Evidence is growing for a role of the intestinal microbiome in the development of type 1 diabetes (T1D) in humans. The composition of the microbiota is heavily influenced by environmental and developmental factors, making the identification of disease-specific microbial signatures difficult. This review summarizes the impact of geographic location, a major confounder of the intestinal microbiota, on the discovery and validation of T1D-microbiota associations as reported in published case-control studies. Few common taxonomic associations were observed across studies and geographic locations, possibly due to the large effect of environmental confounders. In the future, a focus on single geographic regions and integration of multi-omic data will help in identifying disease signatures and potential functional biomarkers of T1D.
Microbiota variation between geographically distinct populations has been consistently observed [14-20]. These findings are not surprising, as several of the studies compared extremely different populations (e.g., Malawians to Amerindians to the USA [14]; Italian to Burkina Faso [15]; USA to Bangladesh [19]), and they primarily attribute ‘geographic’ differences to differences in diet and a western lifestyle. For example, in the study comparing the microbiota of children in Italy and Burkina Faso, the children in Italy were completely weaned from a milk-based diet by 1 year of age, but this occurred closer to 2 years of age in children in Burkina Faso [15]. Thus, the researchers concluded that the microbiota of children in Burkina Faso had a delayed maturation compared to Italian children and attributed this to the difference in age of weaning rather than age itself specifically or even diet or geography [15]. Such studies rely on geography to accumulate the effects of numerous factors on the microbiota in order to capture a maximal amount of variation between different populations. However, this gross simplification overlooks the microbial heterogeneity within a given population.

More constrained studies comparing the microbiota of children from different countries with a western lifestyle also report geographic differences [17,20], drawing attention to possible effects of genetic, cultural, and true geographic influences on the microbiota. Whether microbial variation is attributed directly to surrounding host environment or some other underlying factor, restricting geography in study design is expected to be useful in mitigating the effects of confounding variables in microbiota studies [9].

**Microbiota in T1D Studies and Geography**

There is a lack of consensus on taxonomic associations from studies of IA or T1D and the microbiota. This is not surprising given the extent to which environmental factors can influence microbiota in relation to the small effects of disease. Here results are compared from five longitudinal and three cross-sectional case-control studies describing T1D- or IA-associated microbial signatures and emphasize geographic differences with the expectation that underlying environmental variables may be attributed to differing results and discuss important considerations for future studies.

**Bacteroidetes**

A general trend of increased abundance of members of the phylum Bacteroidetes, specifically *Bacteroides dorei* in Turku, Finland [21], and a decrease in the abundance of Firmicutes prior to seroconversion has been observed in the Type 1 Diabetes Prediction and Prevention (DIPP) study [21,22]. A different cohort of older Finnish children also observed increased *Bacteroides* abundance in children with islet autoantibodies [23]. In the BABYDIET cohort study in Germany, subjects with a high abundance of *Bacteroides* at 6 months of age were at an increased risk of early autoantibody development, but this trend was not particularly prevalent in subjects included in the study [24]. Similar differences were not seen in a longitudinal case-control comparison in DIABIMMUNE, a study of Finnish and Estonian children [25], or in a cross-sectional study conducted in Denver, Colorado [26], possibly due to differences in study protocol, sample size, and geographic location (Table 1). Specifically, Finnish children enrolled in the DIABIMMUNE study were from Espoo, Finland whereas those with increased *B. dorei* enrolled in DIPP were from Turku, Finland. A separate geographic comparison of subjects in DIABIMMUNE reported a higher relative abundance of Bacteroidetes in Finnish samples compared to neighboring Russian Karelia [27]. Kemppainen et al. in 2015 observed an increased abundance of *Bacteroides* in children at risk for T1D in Finland and Colorado compared to Georgia, Washington, Germany and Sweden [20]. This would suggest that while Finnish children tend to harbor a greater relative abundance of *Bacteroides* compared to other nations, the composition of this phyla varies within Finland. Furthermore, this conjecture is supported by a study of a small subset of children enrolled in DIPP from Oulu and Tampere, Finland that did not observe an increase of *B. dorei* but did observe an imbalance of other *Bacteroides* genera compared to controls [28]. This observation was further correlated with a bacteriophage, which was hypothesized to have a modulatory effect on *Bacteroides* populations [28].

So far, studies limited to one clinical site tend to report taxonomic differences between cases and controls [21,22,26,29], while studies with subjects from multiple geographically distinct sites do not report such associations [30]. The BABYDIET study, which enrolled children with a first-degree relative with T1D, did not see any taxonomic or functional differences in children prior to seroconversion [30]. The authors hypothesized that this may be due to underlying differences caused by early life environmental exposures [30]. This hypothesis is supported by the fact that these children came from multiple cities across Germany, and thus environmental confounders likely influenced the ability to detect taxa-specific associations with disease. However, they did observe that the microbial interaction networks of children that went on to develop T1D were substantially altered from controls, supporting the notion that a perturbation in microbiota precedes disease [30].

**Diversity**

The microbial diversity of pre-autoimmune children is not well defined, with some studies reporting clear decreases in diversity prior to seroconversion [22,23] while other studies see no difference in diversity measures between those that seroconvert and healthy controls [25,26,30]. Meanwhile, a decrease in microbial diversity after seroconversion but prior to T1D onset was reported [25]. Thus, microbial diversity may be more indicative of events leading to T1D, but not IA, but fewer studies have evaluated the former. It is important to note that all studies reported alpha diversity metrics, general measurements of species richness and/or evenness and are not informative of phylogenetic relationships within a microbial community. These studies used either Chao1 or Shannon diversity metrics (Table 1), which are not comparable. In addition, alpha diversity calculations are highly influenced by how read counts were processed, thus drawing into question the comparability of reported diversity findings across studies. Therefore, differences in diversity measures between T1D-microbiota studies are likely driven by technical variation in addition to geographical location. Of particular importance, the studies that reported an association of diversity with IA or T1D [22,25], respectively, were underpowered (4 case subjects in both studies) compared to other studies considered. The fact that the two larger studies [25,30] did not find a difference in diversity of pre-autoimmune subjects is not surprising as an increase in sample size often leads to an increase in data variance making it more difficult to identify disease-specific differences. Since microbial diversity metrics differ by geographical location [14,15,19,31], it may be helpful to focus on a single location for future autoimmune microbiota studies.

**Promise of Metagenomic and Integrative Studies**

Seroconversion and later T1D onset in Finnish and Estonian children was characterized by metagenomic changes with an increase in the abundance of genes involved in sugar transport and a decrease in genes involved in the biosynthesis of amino acids [25]. Similar
<table>
<thead>
<tr>
<th>Study</th>
<th>Reference</th>
<th>Demographic variables</th>
<th>Technical variables</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Age</td>
<td># Subjects</td>
<td>Geographic location</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Giorgio [22]</td>
<td>4 months - 3 years</td>
<td>8</td>
<td>Turku, Finland</td>
<td>IA</td>
</tr>
<tr>
<td>DIPP</td>
<td>Brown [23]</td>
<td>4 months - 3 years</td>
<td>8</td>
<td>Turku, Finland</td>
</tr>
<tr>
<td>Davis-Richardson [21]</td>
<td>4 months - 2.2 years</td>
<td>76</td>
<td>Turku, Finland</td>
<td>IA</td>
</tr>
<tr>
<td>BABYDIET</td>
<td>Endesfelder [24]</td>
<td>3-36 months</td>
<td>44</td>
<td>Germany</td>
</tr>
<tr>
<td></td>
<td>Endesfelder [24]</td>
<td>6 months</td>
<td>44</td>
<td>Germany</td>
</tr>
<tr>
<td></td>
<td>Kostic [25]</td>
<td>Birth - 3 years</td>
<td>33</td>
<td>Finland, Estonia</td>
</tr>
<tr>
<td></td>
<td>Alkanani [26]</td>
<td>2 - 45 years</td>
<td>111</td>
<td>Denver, Colorado, USA</td>
</tr>
<tr>
<td></td>
<td>de Goffau [27]</td>
<td>4 - 14 years</td>
<td>36</td>
<td>Finland</td>
</tr>
</tbody>
</table>
metagenomic changes were observed in an earlier study even though the two studies did not identify the same taxonomic differences [29]. The metabolomic potential of the microbiota is similar across individuals and geographic locations, even when taxonomic composition differs [25,32,33] and therefore metagenomic studies may be the key to finding common microbial signatures in IA and T1D studies around the world, even if taxonomic differences are not reproducible.

Despite not having a similar taxonomic composition, children from Germany and Finland developing islet autoantibodies have a reduction in the number of butyrate-producing bacteria before seroconversion, and after the appearance of autoantibodies [23,24,29]. They also have an increased abundance of Bacteroides and a decrease in the abundance of mucin degrading bacteria, such as Akkermansia [21-24,29]. Increased risk of early islet autoantibody development was found in German children with decreased abundance of genes involved with butyrate production via acetate-co-fermentation pathway, and early introduction to solid foods, suggesting that early feeding practices may play a role in the establishment of a pre-autoimmune microbiota [24]. This, coupled with an increase in the abundance of genes involved in sulfur metabolism seen in Finnish children [29] has helped put forth an improved hypothesis for the role of Bacteroides and Akkermansia in butyrate production and gut health [24]. These findings show that through incorporation of multiple data types within microbiota analyses, it is possible to find common disease-specific trends when combining geographically distinct populations.

**Comparability of Studies and Considerations for Future Studies**

Most publications that examined microbiota differences prior to IA or T1D have stemmed from three major prospective cohort studies: DIPP, BABYDIET, and DIABIMMUNE. These cohorts are well suited to address microbiota confounders, as several demographic variables were recorded and samples were subjected and processed under the same protocols, limiting the likelihood of reporting associations due to technical variation. This consistency also makes such cohort studies valuable as additional samples may serve in future validation experiments. Integration of omics data with such carefully collected data on environmental exposures have produced hypotheses about the role of the microbiota in T1D [23-25,29] that can be further validated with intervention studies and/or studies in rodent models. Such studies will be important for validating current findings and developing potential therapeutics or intervention strategies for T1D.

Results from upcoming cohort studies, including All Babies in SouthEast Sweden (ABIS) [34], the Environmental Determinants of Islet Autoimmunity (ENDIA) [35], and The Environmental Determinants of Diabetes in the Young (TEDDY) study [36] (Table 1) will be valuable for further elucidating the role of microbiota in IA and T1D, as well as conducting meta-analyses to identify any potential global disease-microbiota associations. The ongoing TEDDY study, with subjects enrolled from three countries in the EU and three states in the USA, will be especially suited for providing a direct look at the impact of geographic location in the context of IA and T1D.

**Discussion and Conclusion**

Disease-specific microbiota signatures are dwarfed by the variability derived from other environmental factors. Therefore, it’s important to account, or control, for these confounders in studies of disease-microbiota interactions. As a variable, geographic location can be used to capture a multitude of less well-defined environmental factors and social practices that more directly influence microbial composition. Metagenomic studies may be better suited to identify common disease-specific signatures across geographically distinct populations as microbial functions are shared among different taxa. Also, a proper meta-analysis of available T1D-microbiota datasets would reduce technical variation arising from different methods of processing sequencing reads, and help comparability of studies. A more thorough approach is required to fully understand the complex interactions between host, microbiota, and the environment as they relate to autoimmunity for T1D and its different etiologies.

**Conflict of Interest**

The authors declare no conflict of interest in the submission of this manuscript.

**References**


---

**Table 1:** Human studies investigating role of the gut microbiota in development of IA or T1D. Comparison of demographic and technical variables, and findings of microbial diversity and taxonomic differences.

<table>
<thead>
<tr>
<th>Study</th>
<th>Duration</th>
<th>Participants</th>
<th>Geographic Location</th>
<th>Disease Status</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEDDY [25]</td>
<td>2 - 75 months</td>
<td>836</td>
<td>Sweden, Finland, Germany, US (Colorado, Washington, Georgia/ Florida)</td>
<td>IA; T1D</td>
<td>Ongoing study</td>
</tr>
<tr>
<td>ABIS[25]</td>
<td>Birth onward</td>
<td>17,055 children enrolled in full cohort study</td>
<td>Southeast Sweden</td>
<td>IA; T1D</td>
<td>Ongoing study</td>
</tr>
<tr>
<td>ENDIA [26]</td>
<td>Early pregnancy - childhood</td>
<td>Still enrolling subjects</td>
<td>Australia</td>
<td>IA; T1D</td>
<td>Ongoing study</td>
</tr>
</tbody>
</table>

*Please refer to main reference section of manuscript for cited references.

**Abbreviations:** IA: Islet Autoimmunity; T1D: Type 1 Diabetes; SDI: Shannon Diversity Index; FDR: First Degree Relative; SCFA: Short Chain Fatty Acid; BF: Breastfeeding.


