Interactions between the gut microbiome and mucosal immunoglobulins A, M and G in the developing infant gut

Anders Janzon[1]*, Julia K. Goodrich[1]*, Omry Koren[1,2], the TEDDY Study Group[^], Jillian L. Waters[1] and Ruth E. Ley[1#]

1.Department of Microbiome Science, Max Planck Institute for Developmental Biology, Tübingen 72076, Germany.

2. Azrieli Faculty of Medicine, Bar Ilan University, Safed 1311502, Israel

*co-first authors

[^]Members of the TEDDY Study Group are listed in the Supplemental Online Appendix.

[#]Corresponding author: rley@tuebingen.mpg.de

Keywords: IgA, IgM, IgG, gut microbiome, antibody coating, infant gut development, FACS

Abstract

Objective: Interactions between the gut microbiome and immunoglobulin (Ig) A in infancy are important for future health. IgM and IgG are also present, however, their interactions with the microbiome in the developing infant are less understood.

Design: We employed stool samples sampled 15 times in infancy from 32 healthy subjects at 4 locations in 3 countries (from the TEDDY study). We characterized patterns of microbiome development in relation to levels of IgA, IgG and IgM. For 8 infants from a single location, we FACS-sorted microbial cells from stool by Ig status. We used 16S rRNA gene profiling on full and sorted microbiomes to assess patterns of antibody coating in relation to age and other factors.

Results: All antibodies decreased in concentration with age, but were augmented by breastmilk feeding regardless of infant age. Levels of IgA correlated with the relative abundances of OTUs belonging to the Bifidobacteria and Enterobacteriaceae, which dominated the early microbiome, and IgG levels correlated with *Haemophilus*. The diversity of Ig-coated microbiota was influenced by breastfeeding and age, but birth mode. IgA and IgM coated the same microbiota, while IgG targeted a different subset. *Blautia* generally evaded antibody coating, while members of the *Bifidobacteria* and Enterobacteriaceae were high in IgA/M.

Conclusion: IgA/M have similar dynamics with respect to microbiome development with age, and their interactions with the microbiome are influenced by breastfeeding status. IgG generally does not coat the commensal microbiota.

Summary

What is already known on this subject?

- Secretory IgA coats ~50% of microbiota in the gut
- IgM and IgG are less prevalent and coat a lower fraction in the adult, dynamics in the infant gut are not well characterized.
- Breastmilk is a source of IgA to the infant gut and decreases with time.
- IgA coating of microbial cells in infant gut microbiome decreases over time.

What are the new findings?

- Breastfeeding augments the IgA coating of the microbiome at all ages.
- IgA and IgM coat many of the same cells, whereas few are coated by IgG alone.
- Bifidobacteria, Enterobacteriaceae, *Ruminococcus gnavus* are enriched in IgA/M-coated cell fraction, *Blautia* is enriched in uncoated fraction.
- IgG levels correlated with *Haemophilus*.

How might it impact on clinical practice in the foreseeable future?

- Ig-coated fraction of the gut microbiome could serve as a useful tool for tracking development of the infant gut microbiome, and/or identifying aberrations to immune sensing of the microbiome.

Introduction

The gut microbiota and the immune system and their interactions develop in tandem in infancy [1,2]. The IgA component of breast milk is protective against infection in immunologically immature infants, and may also direct the development of the gut microbiota. Immunoglobulin A (IgA) is secreted into the gut lumen where it binds with antigens from food and microbiota, thereby excluding them from direct contact with the host epithelial cells [3]. At birth, neonates generally have undetectable IgA in meconium [4], and it takes a few weeks for their immune systems to initiate the IgA production and secretion into the gut [5]. Breastmilk is an important early source of IgA and breastfeeding is associated with high levels of fecal IgA in infants [4,6]. Planer *et al* characterized the fraction of IgA-coated fecal microbiota in infants over the first few years of life and reported differences between breastmilk and formula fed infants, which may relate to differences in how the microbiota develop in these two groups [7].

Although IgA is the dominant antibody in the gut, IgM and IgG are also present. The adult gut sees up to 5 grams of secretory immunoglobulin A (IgA) daily, 100-fold less secretory IgM, and 1000-fold less IgG [8]. IgM and IgA are both produced by B cells locally, and the predominant class-switching that occurs in B cells of the gut-associated lymphoid tissue is from IgM to IgA. Both are secreted into the gut via the same mechanism (polymeric Ig receptor); IgM as a pentamer and IgA as a dimer. In contrast, IgG is the most common antibody in circulation, but can also be transported into the gut via a neonatal Fc receptor [9]. Whereas IgA/M are produced in response to luminal microbial epitopes that are sampled by dendritic cells, IgG induction is thought

to require crossing of the barrier by antigens, such that IgG is not produced continuously in response to common gut antigens.

Based on its similarity to IgA, IgM may be expected to follow similar patterns of microbiota-binding, whereas IgG may not. In healthy adults, IgA has been shown to coat a greater proportion of the stool microbiota than IgG or IgM, but whether the diversity of taxa targeted by these antibodies differs has not been reported in controls [10]. To gain a baseline understanding of how IgA, IgG and IgM coat gut microbiota during microbiome development in infancy, here we performed a longitudinal analysis of the fecal microbiome of healthy infants in which we characterized the diversity of microbiota coated with IgA, IgM and IgG as a function of time and with respect to feeding regimen and antibody levels.

Material and Methods

Subjects and selection criteria - This study was conducted using fecal samples collected by the international type-1 diabetes (T1D) epidemiological prospective cohort study “The Environmental Determinants of Diabetes in the Young” (TEDDY) [11]. Subjects were screened for Type 1 Diabetes (T1D) risk through HLA genotyping, and all samples used in this study were collected from children with HLA genotypes that put them at high-risk of developing T1D. None of the subjects included in the current study have developed T1D or seroconverted for any of the 3 measured T1D-associated autoantibodies (glutamic acid decarboxylase autoantibodies, IA-2 autoantibodies, insulin autoantibodies) by 24 months of age and had not developed T1D by December 2016 (approximately 10 years of age). We focused on a set of longitudinal stool

111 samples (9 to 16 timepoints per subject) obtained from 32 age and sex-matched healthy
112 children (468 fecal samples) from the USA (Georgia and Washington), Germany, and
113 Sweden. Subjects were excluded if they provided less than 12 longitudinal samples, if
114 their participation and sample collection was through a long distance protocol, or if they
115 dropped out of the study before or at 24 months of age. Additional data collected as part
116 of the TEDDY study and that were used in the current analysis included physical
117 descriptors (e.g., length, weight), diet, illnesses, hospitalizations, vaccinations and
118 medicines, and social data such as daycare attendance.
119
120 **IgA, IgG, and IgM ELISAs**
121 IgA, IgG, and IgM were quantified in duplicate for all samples. A serial dilution of
122 reference human serum (Bethyl Laboratories, Montgomery, TX) was used for
123 generating a standard curve, and blocking buffer was used as a negative control. The
124 immunoglobulin concentrations were log transformed (with an offset of 0.01 added to
125 IgM and IgG concentrations to handle zero values) before downstream analysis (see
126 supplementary methods).
127
128 **FACS sorting of antibody-coated cells**
129 Microbial FACS was performed on 117 fecal samples derived from 8 subjects
130 from the Georgia study site. Samples were vortexed and centrifuged to separate
131 bacterial cells from large particles and debris, and the resulting supernatant was
132 transferred to a new tube. This supernatant was then resuspended in PBS and
133 centrifuged to remove unbound immunoglobulins. The resulting pellet was then

resuspended in PBS + 0.1% BSA before labeling with the respective IgA, IgG, and IgM fluorophore-labeled antibodies. Samples were incubated for 30 minutes and then washed twice prior to flow cytometry and cell sorting (see supplementary methods).

Microbial diversity analysis

Genomic DNA was extracted from 468 fecal samples (approximately 20 mg per sample) using the PowerSoil - htp DNA isolation kit (MoBio Laboratories Ltd, Carlsbad, CA). The 16S rRNA V4 hypervariable region was then PCR amplified [12], and the resulting amplicons were pooled and sequenced using the Illumina MiSeq 2x250 bp paired end platform at Cornell Biotechnology Resource Center Genomics Facility.

Sequence analysis was performed using the open-source software package QIIME 1.8.0 (Quantitative Insights Into Microbial Ecology) [13]. Briefly, reads were quality filtered before using the open-reference OTU picking at 97% against the Greengenes August 2013 database. The data was rarefied at 18,429 sequences per sample, which was the lowest sample sequencing depth over 1,000, in order to include as many samples and sequences as possible (described in more detail in the online supplementary methods).

Statistical analyses

Linear mixed models - All linear mixed effects models were fit using the lme4 [14] package in R. Significance was assessed using an F-test with a Satterthwaite approximation for denominator degrees of freedom calculated by the R package

lmerTest. Post-hoc pairwise comparisons were performed using Tukey's HSD tests implemented in the lsmeans [15] R package.

Association of subject characteristics with beta diversity and alpha diversity- A linear mixed model was used for associating individual metadata with alpha and beta diversity. The only factors used in the model examining the effect of time since exposure to oral antibiotics on unweighted UniFrac PC4 were: age, the time since oral antibiotic exposure, and the random effect for subject.

Correlations between immunoglobulin levels, age, and microbial diversity - The R package rmcorr [16] was used to calculate pairwise repeated measures correlations (and their significance) between levels of IgA, IgG and IgM, age, the first three unweighted UniFrac PCoA PCs, and Phylogenetic Diversity. A repeated measures correlation was used because it accounts for the multiple sampling of individuals.

For further details, see the online supplementary methods.

Results

The small sample set is representative of the larger TEDDY population

We looked for previously reported patterns of microbiota diversity in relation to metadata in order to establish that the small cohort used here is representative of the larger TEDDY cohort. Two recent papers have reported on the gut microbiome in the TEDDY cohort as well, employing a 16S rRNA survey on 903 subjects, and metagenomics with 783 subjects [17,18]. Here, we used a subset (32 subjects sampled longitudinally, unsorted stool) that were sequenced ~6X more deeply. Our results recapitulate those of Stewart *et al* and Vatanen *et al* in the following ways: (i)

Bifidobacteriaceae and Enterobacteriaceae dominated the infant gut at earlier timepoints, while Firmicutes and Bacteroidetes increased in relative abundance later (see online supplementary figure S1); (ii) breastfeeding status was significantly associated with several PCs from both unweighted and weighted UniFrac PCoA (see online supplementary figures S2, S3; table S1) after correcting for multiple testing; (iii) the only other factor with a significant association to any PC was geographic location (see online supplementary figure S4A); (iv) we observed a weak association between antibiotic exposure and unweighted UniFrac PC4 (see online supplementary figure S4B-C); (v) age had a significant association with alpha-diversity ($p<10^{-10}$, see online supplementary figure S5). Although these findings that age and breastfeeding status are associated with microbiome diversity are not novel, we report them here to underscore the independent repeatability of previous reports, and to establish that the subset of subjects that we used in subsequent analysis are representative of the larger TEDDY cohort.

Stool antibody levels decrease with infant age and are related to breastfeeding status

We observed that levels of IgA, IgG and IgM (measured by ELISA in the same stool samples for which the microbiome was analyzed) were positively correlated to each other across the sample set (see online supplementary figure S6). Age (and therefore also microbiome alpha diversity) was negatively correlated with levels of IgA (repeated measures correlation r_{rm}=-0.45, p=6.14 × 10^{-22}), IgG (r_{rm}=-0.37, p=1.13 × 10^{-14}) and IgM (r_{rm}=-0.23, p=3.29 × 10^{-5}), although the IgG and IgM correlation with age

was not as strong as that for IgA (see online supplementary figure S6). No association of geographic location or HLA genotype was observed with any of the immunoglobulin levels in stool.

We assessed the relationships between stool IgA levels, age and breastfeeding status to ask how IgA levels varied over time. We were particularly curious to see if the cessation of breastfeeding made an impact on stool IgA levels, and if the effect of breastfeeding was age-dependent. We found that IgA levels in stool decreased significantly with age ($p=1.79 \times 10^{-11}$; figure 1). Furthermore, we observed that breastfeeding was significantly associated with higher stool IgA levels ($p = 0.039$), and that this association was not dependent on the age of the infants. These observations corroborate previous reports that IgA levels decrease with age, but that at any age, breastmilk delivers additional IgA to the infant.

Next, we examined the association between stool levels of IgA, IgG, and common OTUs (e.g., those shared by greater than 40% of samples; figure 2). IgA levels were positively associated with several Enterobacteriaceae OTUs and a few Bifidobacteriaceae OTUs. Only one OTU was associated with IgG levels in feces: this OTU belonged to genus *Haemophilus* (BH adjusted $p=1.66 \times 10^{-3}$).

IgA and IgM coat many of the same microbial cells

For the 8 infants from Georgia, for whom 15 timepoints each were available, we sorted cells according to their antibody-coating status using FACS, using their distributions in four quadrants (see supplementary methods). The gatings for these four populations and the resulting quadrants are illustrated in figure 3B. We ensured that

cross-reactivity of the anti-IgA, IgM and IgG antibodies was minimal, thus the patterns observed by flow cytometry indicate that many microbial cells were tagged with multiple antibodies. The gating patterns indicate that IgA and IgM coat many of the same cells, since the patterns of coating overlap. IgG-coated cells, on the other hand, fell into three categories: Q1, those low in IgG and high in IgA/IgM; Q2, those highly coated in IgG, IgA and IgM; Q3, uncoated cells; and Q4, those highly coated in IgG and low in IgA/IgM. We observed that the correlation between IgA and IgM coating was consistently higher than the correlation between IgA and IgG coating ($p<10^{-10}$; figure 3C)

Specific taxa discriminate total and FACS-sorted populations

We compared the diversity of the sorted microbiota (all Qs combined) to the whole microbiota (unsorted) to assess the impact of the FACS process on microbial diversity. We observed that the microbiota composition of the sorted cells (all Qs) is distinct from the total microbiome: a combined PCoA analysis of unweighted UniFrac analysis shows clear separation of unsorted and sorted cells along PC2 (figure 4A). The unsorted population was richer (Chao 1, Phylogenetic Diversity, Observed Species) and exhibited greater evenness (Gini coefficient) than all sorted cells ($p<10^{-10}$ for all metrics; see online supplementary figure S7). We applied linear mixed models restricted to identify taxa that were differentially abundant between the unsorted and sorted populations (all Qs). This analysis identified members of the Bacteroidetes, Verrucomicrobia, gamma-Proteobacteria and most Firmicutes as comparatively enriched in the unsorted fraction, and Actinobacteria and alpha-Proteobacteria as enriched in the sorted fraction (figure 4C, see online supplementary figure S8). The

difference in composition between sorted and unsorted fractions likely stems from the FACS process itself: cells belonging to Actinobacteria and alpha-Proteobacteria may be less likely to clump than others. Thus, the difference between sorted and unsorted microbiomes introduces the caveat that subsequent analyses with the sorted data are performed on a subset of the microbiome.

Antibody-targeted cells exhibit patterns similar to those of the total population

As observed for unsorted cells, the majority of the variation in the sorted cells (PC1 of the unweighted UniFrac PCoA) was explained by age (figure 4B, see online supplementary figure S9). Similarly, as age increased, alpha-diversity also increased (see online supplementary figure S7). Analysis of variance indicated that FACS quadrant was significantly associated with most of the first 10 PCs from the PCoA (restricted to only sorted 16S data) of both unweighted and weighted UniFrac and all four alpha-diversity metrics. Post-hoc pairwise comparisons showed that this association was driven primarily by a difference between Q4 (high IgG-only) and the other three quadrants. This finding might represent an overall diversity difference between the high IgG-only cells compared to others. However, the number of cells sorted into the high-IgG (Q4) quadrant was significantly lower than the other three quadrants (see online supplementary figure S10). Indeed, most of the cells coated in IgG were also coated by IgA and IgM, and are therefore sorted into Q2 rather than Q4. This significant difference in cell number between the quadrants could explain the difference in diversity observed between Q4 and the other three quadrants.

After exclusion of Q4 (high-IgG only), we observed some significant associations between the first 10 PCs of the beta diversity PCoA and the FACS quadrant, as well as several associations with infant age and infant breastfeeding status. The significant associations with unweighted UniFrac were with age (PCs 1, 2, and 3), breastfeeding status (PCs 1, 2, 3, 6, and 7), and FACS quadrant (PCs 7 and 10; see online supplementary table S2). Further analysis of the FACS quadrant associations showed that Q3 (uncoated) was different from both Q1 ($p = 0.030$) and Q2 ($p = 0.001$) along PC7, and that Q2 differed from Q3 ($p = 0.010$) along PC10 (post-hoc pairwise comparisons between the quadrants using Tukey's HSD method to adjust for multiple comparisons). Among the weighted UniFrac PCs, many are associated with FACS quadrant (PCs 1, 5, 7, 8, and 9; post-hoc analysis shows that Q3 is different from Q1 and Q2), PC1 is associated with breastfeeding status, and PC2 with age (see online supplementary table S2). These results indicate that the diversity of cells targeted by antibodies is influenced by breastfeeding status and with infant age. Furthermore, the specific combination of antibodies on the cells is not random.

Variability in the relative abundance of specific taxa between FACS quadrants

We next identified specific OTUs differentially abundant between the quadrants. To identify common OTUs (i.e., OTUs with non-zero sequence counts in greater than 40% of FACS samples examined) with differential relative abundance between quadrants, we performed linear mixed models using each OTU as a response variable. We searched for differences between the coated (Q1 and Q2; excluding Q4 because of low cell population) and uncoated (Q3) populations. When comparing Q1 to Q3, we

293 found significant differences in the relative abundances of 80 out of 190 OTUs (figure 5,
294 see online supplementary figure S11 and supplementary table S3). These included
295 mostly OTUs classified as *Blautia*, which had higher relative abundance in Q3
296 (uncoated) compared to other Qs. OTUs that were higher in Q1 (IgM and IgA both high)
297 were mostly classified as *Bifidobacterium*, unclassified Enterobacteriaceae, and
298 *Ruminococcus gnavus*. Similarly, 101 OTUs had differential relative abundances
299 between Q2 (all high) and Q3 (all low), with *Blautia* significantly higher in Q3 (see online
300 supplementary figure S12 and supplementary table S4).
301
302 **Cross-validation for IgA-targeted microbiota**
303 We compared the results of our analysis with those of Planer *et al*, who used
304 FACS followed by 16S rRNA gene sequencing to characterize the IgA coated microbes
305 of 40 twin pairs over the first couple years of life [7]. We used the same reference
306 database as Planer *et al* to classify sequences, therefore we could compare OTUs
307 directly. Only two of the OTUs that the Planer study identified as differentially abundant
308 between the IgA^+ and IgA^- fractions were also among the common OTUs examined in
309 our analysis, but interestingly, both behaved similarly with respect to antibody coating.
310 Greengenes OTU 365385 (genus *Bifidobacterium*) was consistently targeted by IgA in
311 the Planer study and we observed that it was proportionally higher in both Q1 (BH
312 adjusted $p = 3.07 \times 10^{-05}$; figure 5A) and Q2 (BH adjusted $p = 1.88 \times 10^{-05}$) compared to
313 Q3. The other Greengenes OTU detected in both datasets (606927; family
314 Peptostreptococcaceae) was consistently non-targeted by IgA in the Planer study, and
315 similarly, we observed this OTU to have higher relative abundance of this OTU in Q3

316 (uncoated) compared to Q2 (BH adjusted p = 0.013) (see online supplementary figure
317 S12 supplementary table S4). These comparisons indicate that taxon-specific antibody
318 coating profiles can be generalizable across studies.
319
320 **Discussion**
321 This study included infants from four distinct geographic locations, two in the
322 USA and two in Europe. These geographic locations left a discernible imprint on the
323 infant microbiome, which was observed in other studies of the microbiome in the
324 TEDDY group [17,18]. Effects of breastfeeding status on the microbiome and antibody
325 levels, and the decrease in antibody levels with age were also similar across subjects,
326 regardless of the shift in diversity associated with the geographic location of the infants.
327 We also report a strong impact of breastfeeding status and age on gut microbial
328 community structure, which is consistent with the two other TEDDY studies [17,18]. In
329 contrast to these larger studies, we did not observe an effect of birth mode on
330 microbiome diversity in this dataset.
331 Our longitudinal analysis of the Georgia USA infants' gut microbiota targeted by
332 antibodies over time revealed that IgM and IgA target the same microbial populations.
333 IgA and IgM targeting of the same cell population implies they may be induced via the
334 same mechanism, and that like IgA, IgM induction is a local phenomenon. In contrast,
335 we observed patterns of IgG coating were quite distinct from those of IgA and IgM. IgG
336 is induced as a result of the systemic immune system's recognition of a "non-self"
337 antigen and is not, in general, produced locally in the gut. Thus, whereas IgA/M
338 responses mirror the gut microbiome generally, the bacterial targets of IgG are related

to those targeted by the systemic immune system around the time of collection. We noted a fair amount of overlap between the IgA/M and IgG coated fractions, however, suggesting redundancy across all three classes of antibody.

The majority of IgA secreted to the gut is polyclonal and thought to be relatively unselective as it binds with epitopes that are widely shared amongst gut bacteria [19,20]. In addition, secretory IgA affinity maturation by somatic hypermutation is a prominent feature of the human repertoire [21] and a large fraction of secretory IgA may be targeted to specific bacterial species [22]. Hapfelmeier and colleagues showed with a reversible colonization system in germfree mice that a specific strain of *E. coli* induced a sustained IgA response even after the strain was no longer there (>100 days), but that colonization with other bacteria induced a decline in the *E.coli*-specific IgA and an increase in IgA response to the newly introduced bacteria [23]. These results suggested the presence of a long-lived compartment of IgA secreting plasma cells in the intestinal *lamina propria*, but that the compartment has a finite size, so that the IgA secreted into the intestine depends on the dominant luminal microbial species. Based on this model, we would expect the IgA-coated fraction of microbiota to mirror the unsorted microbiota, and allowing for the caveats of FACS sorting (e.g., some phyla aren't well represented in the sorted cells for reasons that are not well understood) this is the general pattern that we observed.

A few OTUs provided exceptions to the general pattern in that some OTUs were more highly represented in the IgA-coated than uncoated cell populations. In particular, OTUs classified as *Bifidobacterium*, unclassified Enterobacteriaceae, and *Ruminococcus gnavus* had higher representation in the IgA-coated cell fraction. In

addition, we observed that levels of IgA antibody in stool was correlated to a few Bifidobacteriaceae OTUs and several Enterobacteriaceae OTUs. Planer *et al* also observed one of the same OTUs as highly coated in IgA in the fecal samples of children from Malawi [7], and Bifidobacteria in particular have been shown to induce high levels of IgA in the gut [24]. Bifidobacteria and Enterobacteriaceae dominated the infant gut microbiomes early on, however the higher-than expected levels of specific OTUs belonging to these taxa in the IgA-coated fraction suggests they may be stimulating IgA production specific to them in excess of what is expected from their relative abundances in the microbiota. Alternatively, given that here we characterize antibody coating of bacteria in the feces, these taxa may be preferentially cleared from the intestine once coated with IgA.

In contrast to the high-IgA-coated microbiota, OTUs classified as *Blautia* showed a lack of antibody coating. In the healthy mouse cecum, specific OTUs have been shown to be less coated than what would be expected by chance [25]. One explanation could be antibody evasion by regulation of epitopes in response to the antibodies. Certain bacterial species, including the gut commensal *Bacteroides thetaiotaomicron*, have been shown to downregulate the expression of epitopes in response to IgA [26–28]. Alternatively, the relative undercoating of specific taxa in stool may reflect patterns of clearance, and here uncoated *Blautia* OTUs are more likely to be shed than their coated counterparts. How IgA coating relates to the microbial ecology of the microbiota, their growth rates, the immune response to specific epitopes, and ultimately shedding of the cells, is complex and not well understood.

Although a large proportion of sorted cells was coated by all three antibodies, we did observe differences between the patterns of IgG coating compared to the patterns of IgA/IgM coating, although the low cell count in the IgG-only coated category stymied the interpretation. However, one OTU was associated with IgG levels in feces: this OTU belonged to the genus *Haemophilus.* IgG is produced systemically in response to infection, and it is also produced in response to immunization. Thus, this pattern may have resulted from the *Haemophilus* vaccine, which is given to infants. It is one of the few vaccines against a bacterium, and it is not a mucosal vaccine. *Haemophilus* colonizes the upper respiratory mucosa and not all are pathogenic. The *Haemophilus* that we observed here may include commensal specie(s) that share capsule epitopes with the vaccine strains.

Our results indicate similar dynamics of antibody levels and microbiome development with age and breastfeeding status in infants from different locations. Patterns of IgA/M patterning indicate that sorting for IgM coated cells may not be any more informative than sorting for IgA alone. IgG-coating of the infant gut microbiome may reveal antigen exposure if gut microbes cross-react. These data provide a baseline reference for further investigation of healthy or unhealthy children's gut microbiomes in infancy.

Acknowledgements: We thank William Melvin and Wei Zhang at Cornell University for their assistance. Special thanks to Timothy Bushnell, Matthew Cochran and staff at the Flow Cytometry Core at the University of Rochester Medical Center. This work was

supported by a National Science Foundation Graduate Fellowship (JKG), Swedish Research Council grant 2011-922 (AJ), and the Max Planck Society (REL). The TEDDY study is funded by U01 DK63829, U01 DK63861, U01 DK63821, U01 DK63865, U01 DK63863, U01 DK63836, U01 DK63790, UC4 DK63829, UC4 DK63861, UC4 DK63821, UC4 DK63865, UC4 DK63863, UC4 DK63836, UC4 DK95300, UC4 DK100238, UC4 DK106955, and Contract No. HHSN267200700014C from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Institute of Child Health and Human Development (NICHD), National Institute of Environmental Health Sciences (NIEHS), Juvenile Diabetes Research Foundation (JDRF), and Centers for Disease Control and Prevention (CDC). This work supported in part by the NIH/NCATS Clinical and Translational Science Awards to the University of Florida (UL1 TR000064) and the University of Colorado (UL1 TR001082).

Figure and Table Legends

Table 1. Summary of participant characteristics.

Figure 1. Infant fecal IgA concentrations over the first couple years of life. PC1 from the PCoA of the fecal microbiome unweighted UniFrac distances is plotted against the age of the infant in days at the time of sampling. The points are colored by the log transformed fecal IgA concentrations; lower concentrations are blue and higher concentrations are red.

Figure 2. A heatmap summarizing the association of age, fecal IgA and IgG concentrations, and breastfeeding status with common OTUs in the infant fecal microbiomes. Along the vertical axis are each of the common OTUs (non zero value in >40% of samples tested) and the horizontal axis is each of the fixed factors examined. The heatmap is colored by the –log10 of the p value from a linear mixed model examining the OTU association with the fixed effects; Blue indicates a positive association and red a negative association while white is a p value of 1, and the darker the color the more significant the association. The panel on the left is colored by the family level taxonomic association of the OTU.

Figure 3. IgA and IgM coat many of the same microbial cells. (A) Representative data for FACS sorting of microbial cells in infant fecal samples. Frequencies of IgM versus IgA tagged cells (green and red, respectively). (B) This panel shows how the four Quadrants were gated: Q1-Q4 are indicated on the figure. Q1: IgA, IgM both high; Q2:

All three high; Q3: all three low; Q4: IgG high, others low. (C) Box-plots at each time point of the Pearson correlations (y-axis) between IgA and IgM signals and between IgA and IgG signals from the FACS sorting of each fecal sample.

Figure 4. Fecal microbiome composition is significantly altered by FACS, but remains significantly associated with infant age. (A) PCoA of unweighted UniFrac distances colored by FACS quadrant. Unsorted samples are shown in gray. (B) PCoA of unweighted UniFrac distances colored by participant age. (C) Summary of taxa that are differentially abundant between the total (red) and FACS (green) populations.

Figure 5. Many differentially abundant OTUs between Q1 (IgM and IgA both high) and Q3 (low coating) are classified as *Blautia* and *Bifidobacterium.* (A-B) Plots for a subset of differentially abundant OTUs between Q1 and Q3 illustrating (A) higher relative abundance of *Bifidobacterium* OTUs in Q1 and (B) higher relative abundance of *Blautia* OTUs in Q3 (see online supplementary figure S11 for plots of all OTUs). Plots on the left show the log(Q1 OTU abundance) - log(Q3 OTU abundance) divided by log(Q1 OTU abundance) + log(Q3 OTU abundance) over time where each point represents a sample and is colored by timepoint. Positive values on the y-axis indicate enrichment of the OTU abundance in Q1 and negative values indicate and enrichment of the OTU abundance in Q3; this is similar to the IgA index defined in Planer *et al* 2016. Plots on the right show the average of the log transformed OTU relative abundance in Q1 (Red) and Q3 (Blue) at each time point. OTU Greengenes ID, taxonomic classification, and the p value from the linear mixed model using OTU as a response variable is indicated

above each set of graphs. Only common OTUs (non-zero value in >40% of samples tested) were tested.

476 **References**

477 1 Honda K, Littman DR. The microbiome in infectious disease and inflammation.
478 *Annu Rev Immunol* 2012;**30**:759–95.

479 2 Belkaid Y, Hand TW. Role of the microbiota in immunity and inflammation. *Cell*
480 2014;**157**:121–41.

481 3 Stokes CR, Soothill JF, Turner MW. Immune exclusion is a function of IgA. *Nature*
482 1975;**255**:745–6.

483 4 Maruyama K, Hida M, Kohgo T, *et al.* Changes in salivary and fecal secretory IgA in
484 infants under different feeding regimens. *Pediatr Int* 2009;**51**:342–5.

485 5 Rognum TO, Thrane S, Stoltenberg L, *et al.* Development of intestinal mucosal
486 immunity in fetal life and the first postnatal months. *Pediatr Res* 1992;**32**:145–9.

487 6 Bridgman SL, Konya T, Azad MB, *et al.* Infant gut immunity: a preliminary study of
488 IgA associations with breastfeeding. *J Dev Orig Health Dis* 2016;**7**:68–72.

489 7 Planer JD, Peng Y, Kau AL, *et al.* Development of the gut microbiota and mucosal
490 IgA responses in twins and gnotobiotic mice. *Nature* 2016;**534**:263–6.

491 8 Haneberg B, Aarskog D. Human faecal immunoglobulins in healthy infants and
492 children, and in some with diseases affecting the intestinal tract or the immune
493 system. *Clin Exp Immunol* 1975;**22**:210–22.

494 9 Yoshida M, Claypool SM, Wagner JS, *et al.* Human neonatal Fc receptor mediates
495 transport of IgG into luminal secretions for delivery of antigens to mucosal dendritic
496 cells. *Immunity* 2004;**20**:769–83.

497 10 van der Waaij LA, Kroese FGM, Visser A, *et al.* Immunoglobulin coating of faecal
498 bacteria in inflammatory bowel disease. *Eur J Gastroenterol Hepatol* 2004;**16**:669–
499 74.

500 11 Hagopian WA, Lernmark A, Rewers MJ, *et al.* TEDDY--The Environmental
501 Determinants of Diabetes in the Young: an observational clinical trial. *Ann N Y Acad*
502 *Sci* 2006;**1079**:320–6.

503 12 Caporaso JG, Lauber CL, Walters WA, *et al.* Global patterns of 16S rRNA diversity
504 at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A* 2011;**108**
505 **Suppl 1**:4516–22.

506 13 Caporaso JG, Kuczynski J, Stombaugh J, *et al.* QIIME allows analysis of high-
507 throughput community sequencing data. *Nat Methods* 2010;**7**:335–6.

508 14 Bates D, Mächler M, Bolker B, *et al.* Fitting Linear Mixed-Effects Models Using
509 lme4. *Journal of Statistical Software, Articles* 2015;**67**:1–48.

15 Lenth R. Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software, Articles* 2016;**69**:1–33.

16 Bakdash JZ, Marusich LR. Repeated Measures Correlation. *Front Psychol* 2017;**8**:456.

17 Stewart CJ, Ajami NJ, O'Brien JL, *et al.* Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* 2018;**562**:583–8.

18 Vatanen T, Franzosa EA, Schwager R, *et al.* The human gut microbiome in early-onset type 1 diabetes from the TEDDY study. *Nature* 2018;**562**:589–94.

19 Stoel M, Jiang H-Q, van Diemen CC, *et al.* Restricted IgA repertoire in both B-1 and B-2 cell-derived gut plasmablasts. *J Immunol* 2005;**174**:1046–54.

20 Wijburg OLC, Uren TK, Simpfendorfer K, *et al.* Innate secretory antibodies protect against natural Salmonella typhimurium infection. *J Exp Med* 2006;**203**:21–6.

21 Benckert J, Schmolka N, Kreschel C, *et al.* The majority of intestinal IgA+ and IgG+ plasmablasts in the human gut are antigen-specific. *J Clin Invest* 2011;**121**:1946–55.

22 Macpherson AJ, Uhr T. Induction of protective IgA by intestinal dendritic cells carrying commensal bacteria. *Science* 2004;**303**:1662–5.

23 Hapfelmeier S, Lawson MAE, Slack E, *et al.* Reversible microbial colonization of germ-free mice reveals the dynamics of IgA immune responses. *Science* 2010;**328**:1705–9.

24 Yasui H, Nagaoka N, Mike A, *et al.* Detection of Bifidobacterium Strains that Induce Large Quantities of IgA. *Microb Ecol Health Dis* 1992;**5**:155–62.

25 Cullender TC, Chassaing B, Janzon A, *et al.* Innate and adaptive immunity interact to quench microbiome flagellar motility in the gut. *Cell Host Microbe* 2013;**14**:571–81.

26 Peterson DA, Planer JD, Guruge JL, *et al.* Characterizing the interactions between a naturally primed immunoglobulin A and its conserved Bacteroides thetaiotaomicron species-specific epitope in gnotobiotic mice. *J Biol Chem* 2015;**290**:12630–49.

27 Lönnermark E, Nowrouzinan F, Adlerberth I, *et al.* Oral and faecal lactobacilli and their expression of mannose-specific adhesins in individuals with and without IgA deficiency. *Int J Med Microbiol* 2012;**302**:53–60.

28 Mantis NJ, Rol N, Corthésy B. Secretory IgA's complex roles in immunity and mucosal homeostasis in the gut. *Mucosal Immunol* 2011;**4**:603–11.

Figure 1

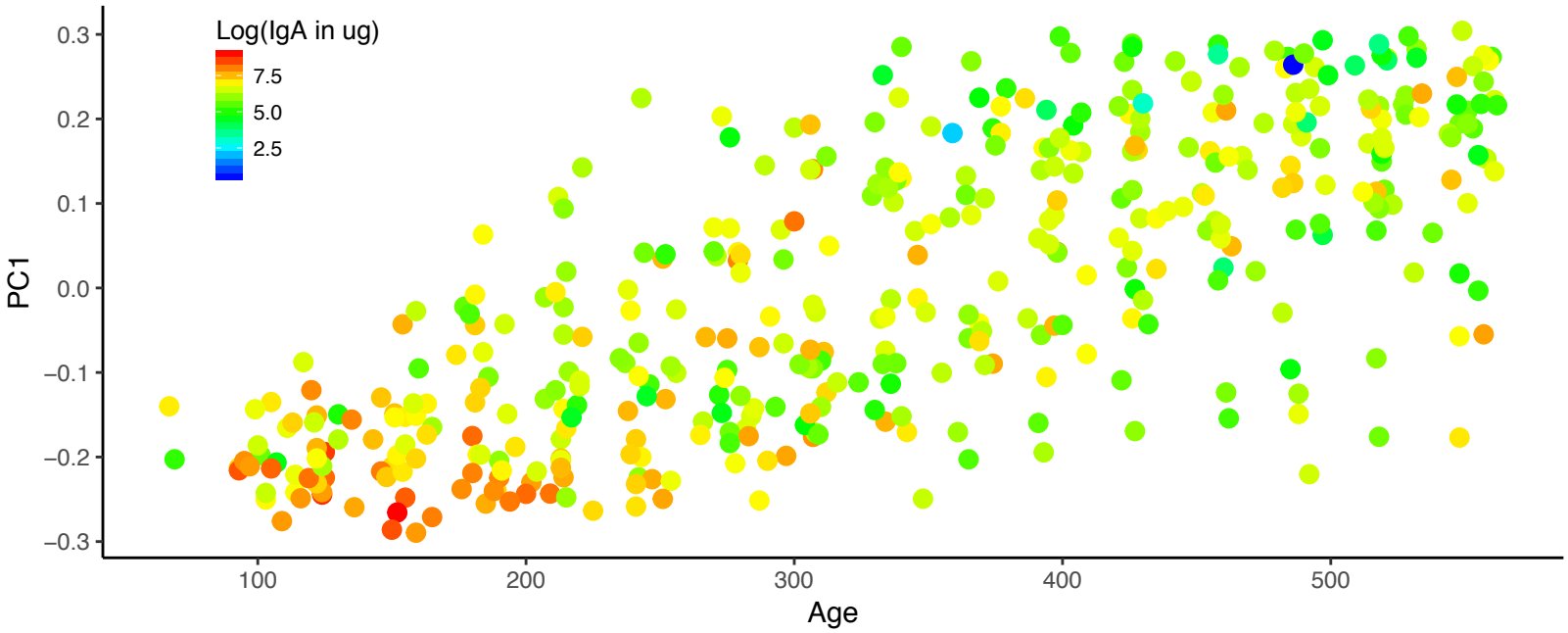

Figure 2

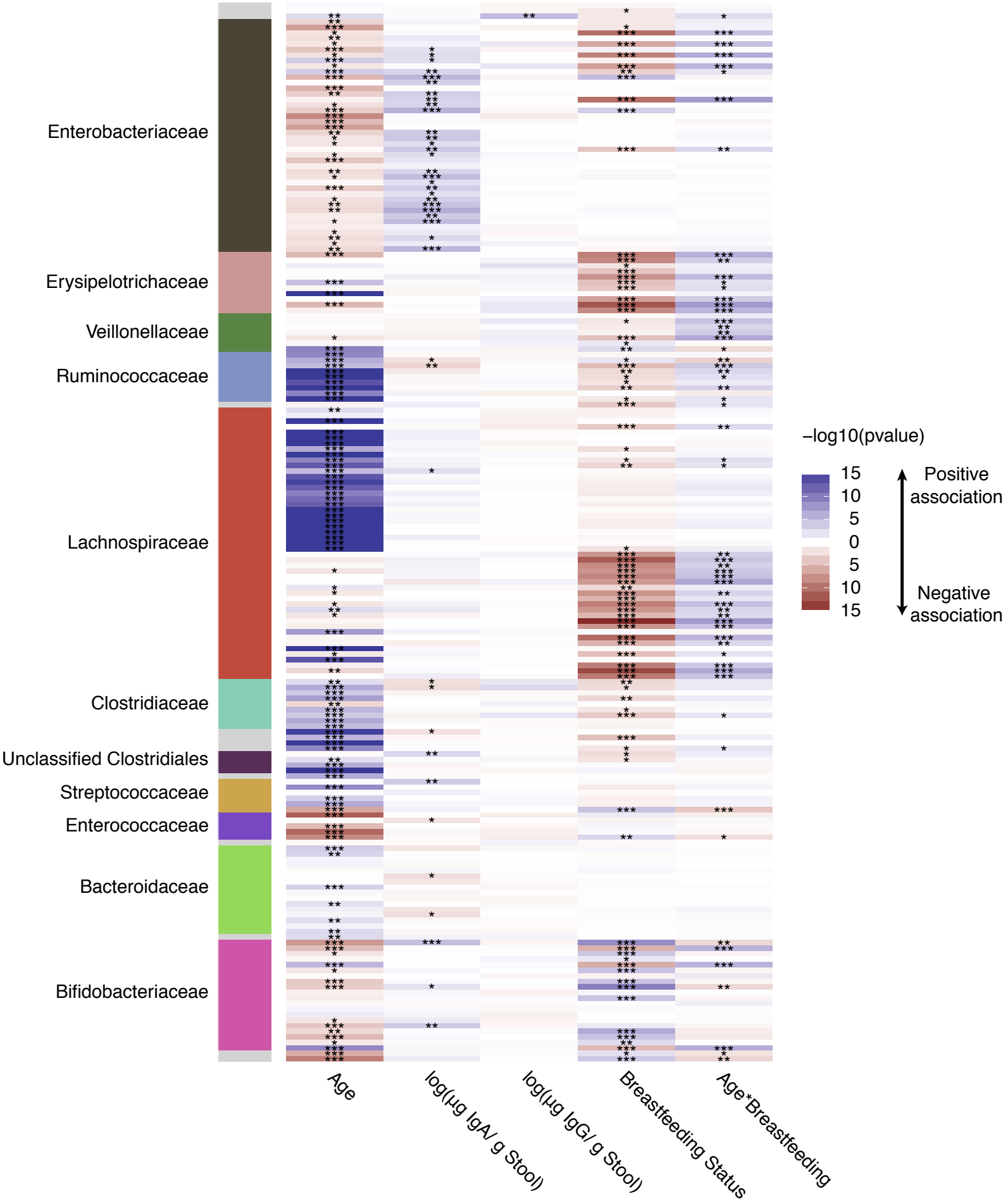

Figure 3

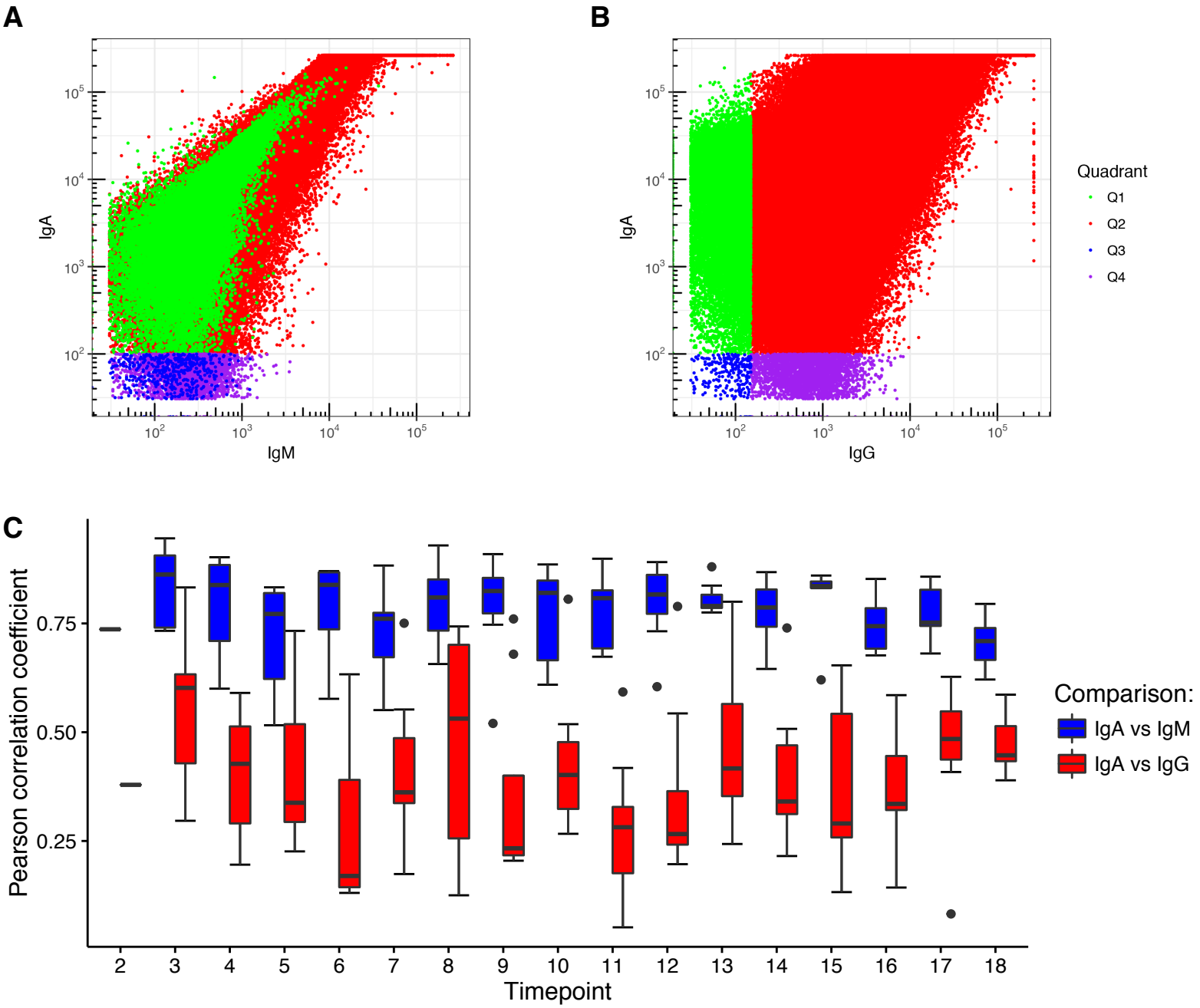

Figure 4

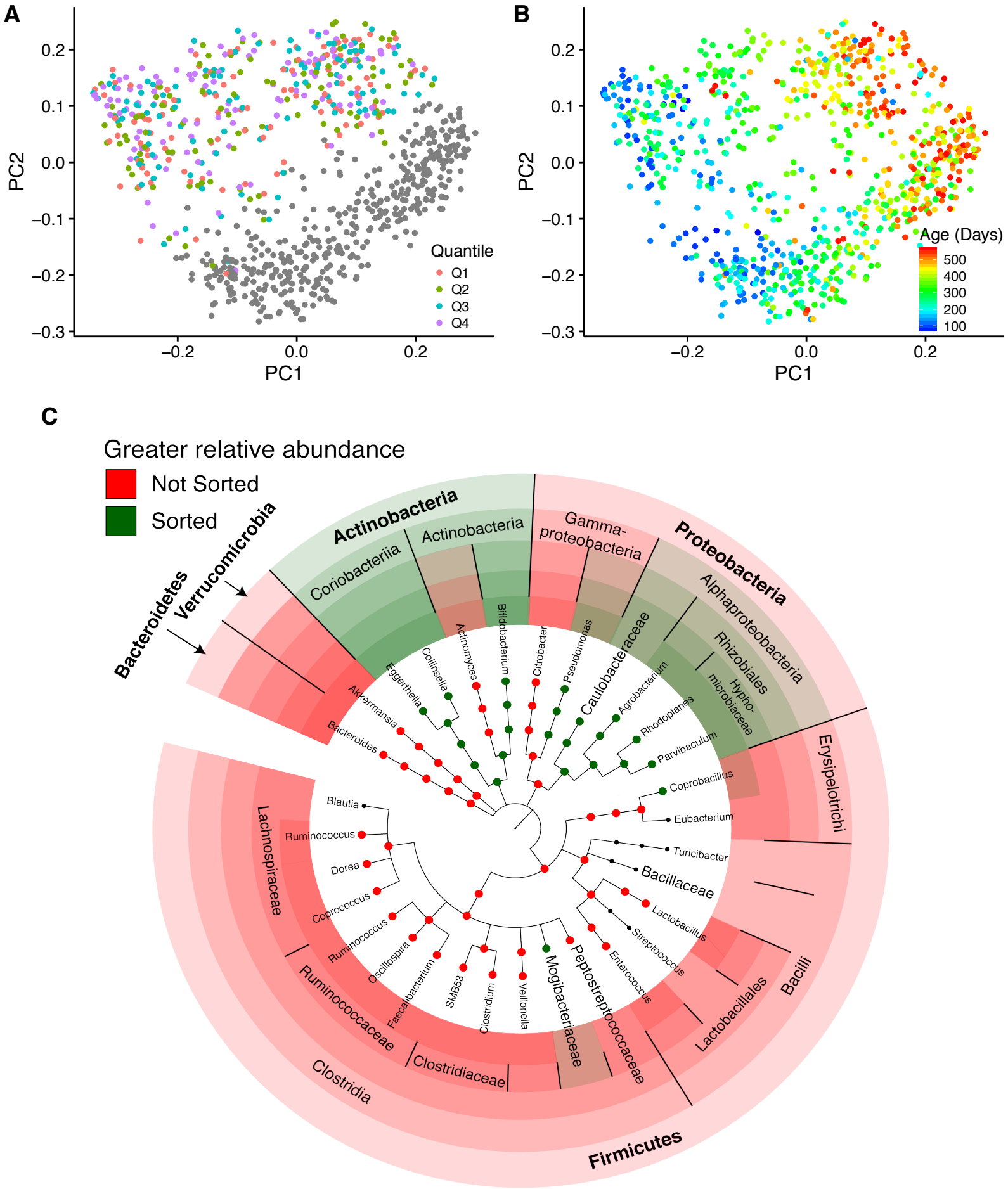

Figure 5