Habitat shapes diversity of gut microbiomes in a wild population of blue tits (*Cyanistes caeruleus*)

Szymon M. Drobniak[1,2,3*#], Mariusz Cichoń[1] & Magdalena Zagalska-Neubauer[4#]

[1] Institute of Environmental Sciences; Jagiellonian University; ul. Gronostajowa 7, 30-387 Kraków, Poland

[2] Ecology & Evolution Research Centre; School of Biological, Environmental and Earth Sciences; University of New South Wales; 2052 Kensington, NSW, Australia

[3] Department of Animal Ecology; Evolutionary Biology Centre; Uppsala University; Norbyvägen 18D, 752 36 Uppsala, Sweden

[4] Department of Behavioural Ecology, University of Wrocław, Sienkiewicza 21, 50-335 Wrocław, Poland

* Corresponsing author; szymek.drobniak@gmail.com

[#] Equal contribution

Introduction

Bacteria are ubiquitous in the environment and constitute the majority of life on Earth. They inhabit not only abiotic environments but also bodies of other living organisms. Being present in virtually all tissues that are in contact with the external environment, such bacteria form what is called a microbiome: a unique, taxon-specific and highly plastic collection of various bacterial, but also fungal, protozoan and viral species that interact synergistically with an organism's physiology and biochemistry (Hird, 2017; Hicks *et al.*, 2018).

There is ample evidence of the impact microbiome diversity has on individual fitness. A huge majority of available published results elaborates on the importance of microorganisms in digestion and nutrients assimilation (Hird, 2017). Symbiotic bacteria can also affect and modulate components of individual immune response, can affect the general metabolic homeostasis of an organism, and finally – they can also modulate individual behaviours through links that exist between organs hosting symbiotic bacteria (e.g., gut) and the brain (Toivanen *et al.*, 2001; Mazmanian *et al.*, 2005; Cryan & O'Mahony, 2011; Davidson *et al.*, 2018; Bergamaschi *et al.*, 2020; Cao *et al.*, 2020). These physiology and condition links often manifest themselves as differences in microbiotic diversity existing between sexes, age classes or developmental stages of individuals within populations, especially if different classes of individuals in populations engage in different types of behaviours or occupy different ecological niches (Spor *et al.*, 2011). For instance, sex differences likely begin to play role already in the prenatal period in mammals, when developing embryos are exposed to compounds secreted by adult female microbiome of their mothers (Jašarević *et al.*, 2016). This sex specificity is amplified later in life, and currently many sex-specific patterns observed in metabolic or psychological disorders prevalence is attributed to, among other things, sex differences in microbiomes and microbial reactivity to sex-specific hormonal profiles (Spor *et al.*, 2011; Jašarević *et al.*, 2016; Beale *et al.*, 2019).

Sex is only one of several individual characteristics influencing symbiotic microbial communities. Unfortunately, majority of evidence about interspecific differences in microbiomes comes from mammals (including humans) and is usually collected in contexts strongly linked to biomedical research. Far less is known about microbiome differences in wild organisms and wild, unmanaged populations and non-mammalian taxa (Benson, 2016; Hird, 2017; Björk *et al.*, 2019). In terms of sexual differentiation, available data suggest the existence of varying, taxon-dependent patterns (e.g., age dependent decreases in microbial diversity in male gorillas, Pafčo et al. 2019; no sexual microbiome dimorphism in baboons, Tung et al. 2015; no significant sexual differentiation in gulls, Noguera et al. 2018; larger diversity of oral and faecal microbiota in male great tits, Kropáčková et al. 2017; markedly larger abundance and diversity of microbes in breeding males in rufous-collared sparrows, Escallón et al. 2019). In most cases the sex more involved in social interactions, or exhibiting more active reproductive behaviour, tended to have richer microbiomes. Similarly, evidence from wild populations suggests that age groups may differ in microbial diversity, with younger individuals often having slightly less diverse microbiomes (Ren *et al.*, 2017; Kohl *et al.*, 2019; Pafčo *et al.*, 2019; but see Noguera *et al.*, 2018).

Relatively the most interesting, but also the scarcest, is evidence for environment-driven modification of host microbiomes. Several factors may contribute to this pattern (e.g., insufficient or not quantified environmental heterogeneity in many wild microbiome projects, insufficient sample sizes collected in wild microbiome studies; Hird 2017). Lack of good estimates of environment-dependent microbiome differences is surprising: the flexibility and

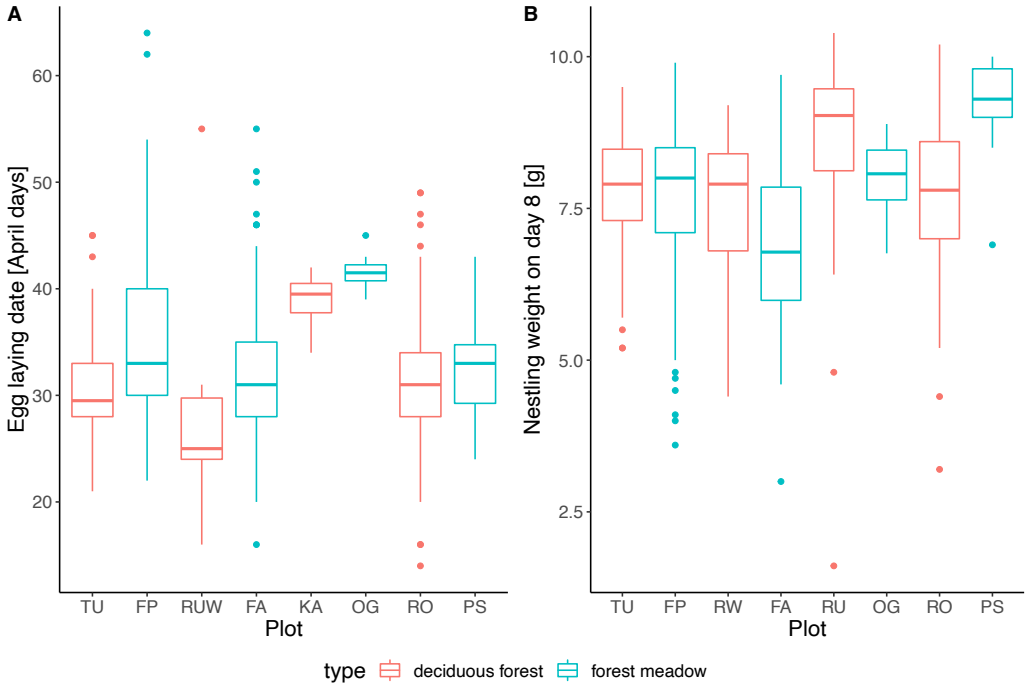
functional diversity of microbiomes may constitute one of important mechanisms conferring phenotypic plasticity and enabling fast, genetically unconstrained, modulation of individual phenotypes (Kolodny & Schulenburg, 2020). Some studies suggest, that in certain systems environmentally-drive microbiome differentiation may play more important role than its modulation by sex or age factors (Ren *et al.*, 2017). Indeed, plasticity and microbiome malleability may be key to adjusting individual physiologies to conditions varying spatially or temporarily (Hicks *et al.*, 2018), and arming individuals with additional physiological pathways necessary while individuals migrate and change habitats (Wu *et al.*, 2018). In systems where multiple related species occur sympatrically, but inhabit varying ecological (e.g., nutritional) niches, microbiomes seem to track this habitat-dependent differentiation (e.g., in Darwin finches microbiome diversity and similarity strongly correlates with phylogenetic differences between related finch species; Loo et al. 2019).

In this study we aimed at supplementing the growing body of evidence about factors driving microbiome differentiation in wild populations of animals. We collected microbiome profiles from over 170 individuals of blue tits inhabiting a wild nest-box population on Gotland (Sweden). Blue tits are an important model species in eco-evolutionary studies, and the data presented in this paper comes from a long-term monitoring project (with over 20 years of continued data collection). In the studied population, we benefited from significant habitat heterogeneity experienced by breeding birds: nest-boxes available to blue tits on Gotland are located either in fertile, rich deciduous forests with dense understory and diverse food base, or in open meadow-like habitats with no understory and sparse one-species tree cover, which translates in observably lower diversity of feeding opportunities. We predicted, that in our study system younger individuals would be characterised by less diverse microbiomes than older (reflecting microbiome development and maturation), that sexes should have similar microbiome profiles (as sexes in blue tits have similar mobility, exploratory behaviour and diet), and finally – that birds from richer forest habitats would significantly differ in terms of microbiome complexity and composition from birds from more open, meadow habitats. We also expected, that blue tit microbiomes will be similar to those of the closely related great tits (which are composed mostly of *Firmicutes* and *Proteobacteria* (Kropáčková *et al.*, 2017; Davidson *et al.*, 2019; Bodawatta *et al.*, 2020), contrary to many scavenging and omnivorous species, where other bacteria phyla, such as *Bacteroidetes*, *Tenericutes* and *Actinobcateria* dominate (Bodawatta *et al.*, 2018; Wu *et al.*, 2018)).

Materials and methods

Field protocol and material

Microbiological material was collected from adult blue tits of both sexes during the 2018 and 2019 fieldwork seasons in the wild population of blue tits inhabiting Gotland – a small Swedish island in the southern part of the Baltic Sea (57°01' N; 18°16' E). In this population blue tits breed in wooden nest-boxes distributed uniformly across 23 study plots of varying size; density of breeding pairs is uniform across plots of different size (unpublished data). Most plots are covered by oak (*Quercus robur*), ash (*Fraxinus excelsior*) and poplar (*Populus* sp.) forests, with dense common hazel undergrowth (*Corylus avellana*). These habitats constitute what we refer to in the following sections as "deciduous forest" habitat. Diversity of plants and very heterogenous environment (with many water-filled ditches, in-forest swamps and small treeless openings) translate into more variable food (winter-moth caterpillars feeding on young oak leaves, large numbers of *Diptera, Coleoptera* and *Hymenoptera* insects; Pitala 2007). Some plots lack the undergrowth and are covered by bright, sparse oak forests with wet, rich hay-meadows abundant in orchids and other perennial plant. We will refer to this type of habitat as "forest meadows" henceforth. The main difference here is in the abundance of different food-sources: in these habitats birds likely feed mostly on caterpillars grazing the oak leaves. Due to these differences, on average birds start breeding later (measured as egg-laying date) in forest-meadow, when compared in pairs of nearby open vs. forest habitats (Supplementary Fig. 1A). In many such pairs of plots, chicks have also lower body mass on day 8 (i.e., in the middle of their growing period, Supplementary Fig. 1B). In plots where both types of habitats can be observed (Öja, Rudvier, Rannarve, Ronnarve) a breeding pair was classified to either of the habitats if all neighbouring nest-boxes (i.e. direct vicinity of the focal nest-box) also were assigned to a give habitat type. This distinction was needed only in case of two largest plots, due to their heterogeneity. Two plots, although entirely classified as deciduous forest (Tuviken and Skoge), have large proportions of conifers (pine *Pinus sylvestris* and spruce *Picea abies*). For this reason in several cases they were excluded in sensitivity analyses to see if this may have had an impact on microbial diversity estimates. In the studied population, tits lay almost exclusively one clutch per year, starting around the 20th of April. Females lay on average 11 eggs (range: 5 – 17) and incubate them for 13 days; chicks fledge at the age of 17 – 20 days.



Supplementary Figure 1. Differences in laying date (A) and day 8 nestling mass between forest-meadow and deciduous forest plots, in pairs of plots located near each other (each two subsequent plots form a pair).

Microbiological samples were collected from adults while they were caught to collect basic morphological measurements and to ring them. Capture was done either by clip-traps mounted inside of a nest-box (i.e., capturing during nestlings' feeding) or by mist-nets setup in the vicinity of a nest-box. All adults were caught at approximately the same stage of nest life, i.e., between 14 and 16 days after hatching. Captured adults were sexed (by the presence of a brood patch) and aged (by the presence of a distinct moult limit in 1-year old individuals), measured for tarsus length, wing length and body weight, and assessed in terms of their aggressiveness while handling (two metrics: bird producing a distress call – yes/no; bird struggling to escape while handling – on scale from 0-3, 0 = no struggling, 3 = very aggressive behaviour, bird difficult to handle).

Faeces were collected using a custom-designed method. Briefly, following measurements, each individual was placed in card-board box (20 x 20 x 30 cm) lined with hot steam-sterilised baking paper. Paper sheets were individual packed in sterile plastic bags to avoid unnecessary contamination. Usually within 2-5 minutes the birds would defecate in the box. After releasing the bird, faeces were collected using a sterile screwcap microtube (Sarstaedt) and transferred to the field laboratory. There, we extracted the bacterial DNA using the PowerSoil Extraction kit (Qiagen) following the manufactures protocol. Extracted material was frozen in -20C and transported frozen (using a portable car freezer) to the Institute of Environmental Sciences in Poland.

At multiple stages of field and laboratory work we collected control samples (e.g., swabs of the clean baking paper, swabs of the field laboratory paper, swabs of the field clothing). We have also performed a couple of extractions without any microbiological material (to establish the "microbiome" of the extraction kit and plastics used in all procedures).

Library preparation and amplicon sequencing

Analysis of microbial DNA was done using a standard protocol designed for the analysis of the 4th variable (V4) region of the bacterial 16S rRNA gene. In the first step we have performed a nested PCR, meant to improve representation of different bacterial lineages (Ganz et al. 2017). The first PCR amplified a long V3-V4 region of rRNA gene using a set of degenerate primers (reverse primer: 1492R, 5'-TACCTTGTTACGACTT; forward primers: a generic primer plus a mixture of primers specific for broad groups of bacteria *Bifidobacteriaceae, Borrelia* oraz *Chlamydiales*, all mixed in proportions 4/1/1/1: 27F-YM 5'-AGAGTTTGATyMTGGCTCAG; 27F-Bif 5'-AGGGTTCGATTCTGGCTCAG; 27F-Bor 5'-AGAGTTTGATCCTGGCTTAG; 27F-Chl 5'-AGAATTTGATCTTGGTTCAG). In the first PCR 12 µl reaction mix were prepared by adding 5 pM of each primer (oligo F mix and 1492R), 2.5 µl of DNA sample, 5 µl of 2xPhanta Max Master Mix (Vazyme), and 3.7 µl pure sterile water. Conditions for this PCR were as follows: 95°C 3 min, (95°C 30 s, 48°C 30s, 72°C 45 s) x 28 cycles. PCR product from the first PCR round was used as template in the second PCR which was performed with a V4-specific pair of primers 515F (5'-TGCCAGCmGCCGCGGTAA) and 806R (5'-GGACTAChvGGGTwTCTAAT). The primers were merged with a portion of the Illumina sequencing adapters, i.e. consisted of a fully complementary primer segment, and partial adapter overhang. In the second PCR 24 µl reaction mix consisted of 1.6 µl of the first PCR product, 1 pM of each primer and 12.5 µl of KAPA HiFi HotStart Ready Mix (KAPA Biosystems). Conditions of the second PCR were as follows: 95°C 3 min, (95°C 30 s, 55°C 30 s, 72°C 30 s) x 22 cycles. Following amplicon generation, the Illumina adapters and library preparation workflow was applied according to the Illumina 16S Metagenomic Sequencing Library Preparation Guide (Part # 15044223 Rev. B). The amplicons were multiplexed with dual-barcode combination for each sample. The samples were sequenced in two 300-bp paired end runs on an Illumina MiSeq platform at the Jagiellonian University, Kraków, Poland. In total we successfully sequenced 175 samples in two MiSeq runs.

Bioinformatic processing

Raw sequencing reads were demultiplexed by the MiSeq software and save to paired FASTQ files. Sequencing data was cleaned and processed using the QIIME2 (Bolyen *et al.*, 2019) and DADA2 (Callahan *et al.*, 2016) pipelines. In the first step we trimmed the sequences based on the quality of reads: initial base calls in both forward and reverse reads were of high quality, however we decided to remove the distal 20-30 bases in both reads to keep base calls' quality uniformly above 20, which resulted in truncation of forward reads to 250bp, and reverse reads to 220bp. Then, the reads were cleaned and filtered using the DADA2 pipeline, which clustered sequences into sequence variants, removed chimeres, indel and substitution errors and other artefacts. We used default recommended DADA2 settings.

Final sequence variants (operational taxonomic units, OTUs) were used to construct the full feature table in QIIME2. In order to improve downstream analyses, the feature table was then cleaned using several different levels of filtering. First, following a conservative approach, we have removed from the feature table all variants identified in control samples (samples extracted without the biological material, and PCR reactions using sterile water instead of bacterial DNA; please see the reults section for summary of this filtering). Then, we used a naïve Bayesian taxonomical classifier trained to the rRNA region and read lengths

obtained in our experiment. Training was done using the reference taxonomic set from the Silva database, version 1.38 (Quast *et al.*, 2013). Following taxonomical clustering of OTUs, the feature table was filtered to remove all eukaryotic mitochondrial and chloroplast rRNA genes sequences. Finally, we removed all singletons and doubletons (variants identified only in 1-2 reads).

The final cleaned feature table was used in subsequent analyses to calculate a number of alpha diversity metrics (Shannon index, evenness index, Faith's phylogenetic diversity index) and to represent samples in a low-dimensional space through a number of prevalence and abundance-based metrics used in exploring beta-diversity (Bray-Curtis distances, weighted and unweighted UniFrac distances; for the latter, we have generated a phylogeny of all sequences suing the *phylogeny* plugin in QIIME2). Differential abundance of specific taxons across different subsets of the sample set was explored using the ANCOM algorithm (Mandal *et al.*, 2015); this analysis was performed on the set further filter to only include OTUs representative across multiple individuals in the database (i.e. variants found in at least 10 individuals). The impact of certain factors (individual sex, age, habitat type, experimental plot, year of study) was tested using the *adonis* plugin in QIIME2, utilising a permutation-based ANOVA-type test.

Results

Sequencing performance and overall metrics

Sequencing generated 36659846 raw paired-end reads. 63.4% passed initial quality and error filtering, and 23301944 were merged into complete reads. Two samples dropped-out at this stage as none of their reads passed the filters. 5200156 reads remained after filtering out the chimeras and non-resolvable sequencing errors (range of 42 to 98269 reads per sample, median 27027). Remaining samples were clustered into 3743 unique sequence variants (OTUs).

After subsequent filtering steps (removal of singletons and doubletons, removal of samples with less than 4000 reads) we ended up with 4709554 reads clustered into 2829 OTUs; taxonomic analysis was performed at this stage. Note that all 4 samples removed due to the >4000 reads threshold were animal and not control samples. In order to use a conservative approach, all sequence variants identified in the control samples (i.e. samples without any microbiological material collected at the stage of PCR and DNA extraction, and samples collected from the area of the field lab) were considered as contaminants and were removed from the animal samples – this resulted in removal of 221 OTUs (17 of which were found in all control samples). After this step, and after removal of all sequence variants identified as mitochondrial or chloroplast rRNA genes, we ended up with 170 samples containing 2536 OTUs.

Alpha rarefaction analysis indicated that all individuals were sequenced to near saturation (1 sample with final read count < 2000 dropped out at this stage; Fig. 1). All diversity analyses were done on samples rarefied to the depth of 1500.

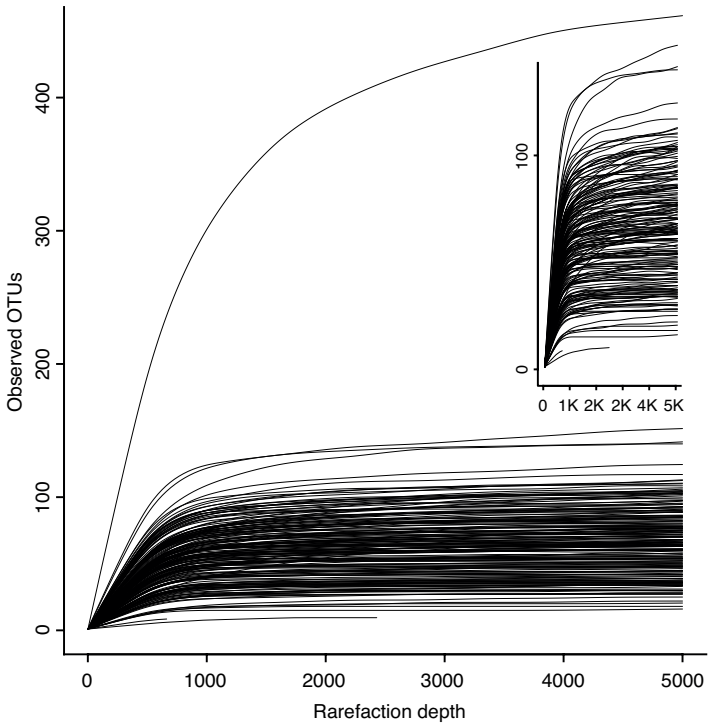


Figure 1. Rarefaction analysis. The inset shows in more details the lower curves. Lines are smoothed traces, each based on 10 rarefactions per depth.

Alpha diversity

Sexes did not differ in terms of alpha diversity measures (Shannon's entropy: Kruskal-Wallis H = 0.41, p = 0.52; evenness index: H = 0.002, p = 0.96; Faith's phylogenetic diversity: H = 1.67, p =0.19; number of observed OTUs: H = 0.09, p = 0.76). Similarly, there were no significant differences between age groups, although older individuals tended to have higher Shannon's diversity index and number of OTUs per individual (1-year old vs. older birds; Shannon's entropy: H = 2.57, p =0.10; evenness index: H = 0.49, p = 0.48; Faith's PD: H = 0.06, p = 0.81; number of observed OTUs: H = 2.94, p = 0.08). Forest-meadow habitats had significantly lower diversity in terms of Shannon's entropy (H = 6.21, p = 0.01). The also tended to have less observed OTUs and lower phylogenetic diversity, as well as higher evenness, although in all cases differences were marginally non-significant (number of observed OTUs: H = 2.21, p = 0.13; Faith's PD: H = 2.02, p = 0.15; evenness: H = 2.19, p = 0.13). Differences between habitats in Shannon's index and numbers of observed taxa were even more pronounced after removing form analyses two plots (Tuviken and Skoge) characterised by significant prevalence of pine and spruce (and hance mixed rather than deciduous forests; Shannon's entropy: H = 7.16, p = 0.007; observed OTUs: H = 3.01, p = 0.08). All these results are summarised in Fig. 2. In spite of now large-scale habitat differences in alpha-diversity, there was large variation in alpha-diversity metrics between the studied plots (Fig. 3A), some of them having significantly different diversity in pairwise plot-plot comparisons (Fig. 3B).

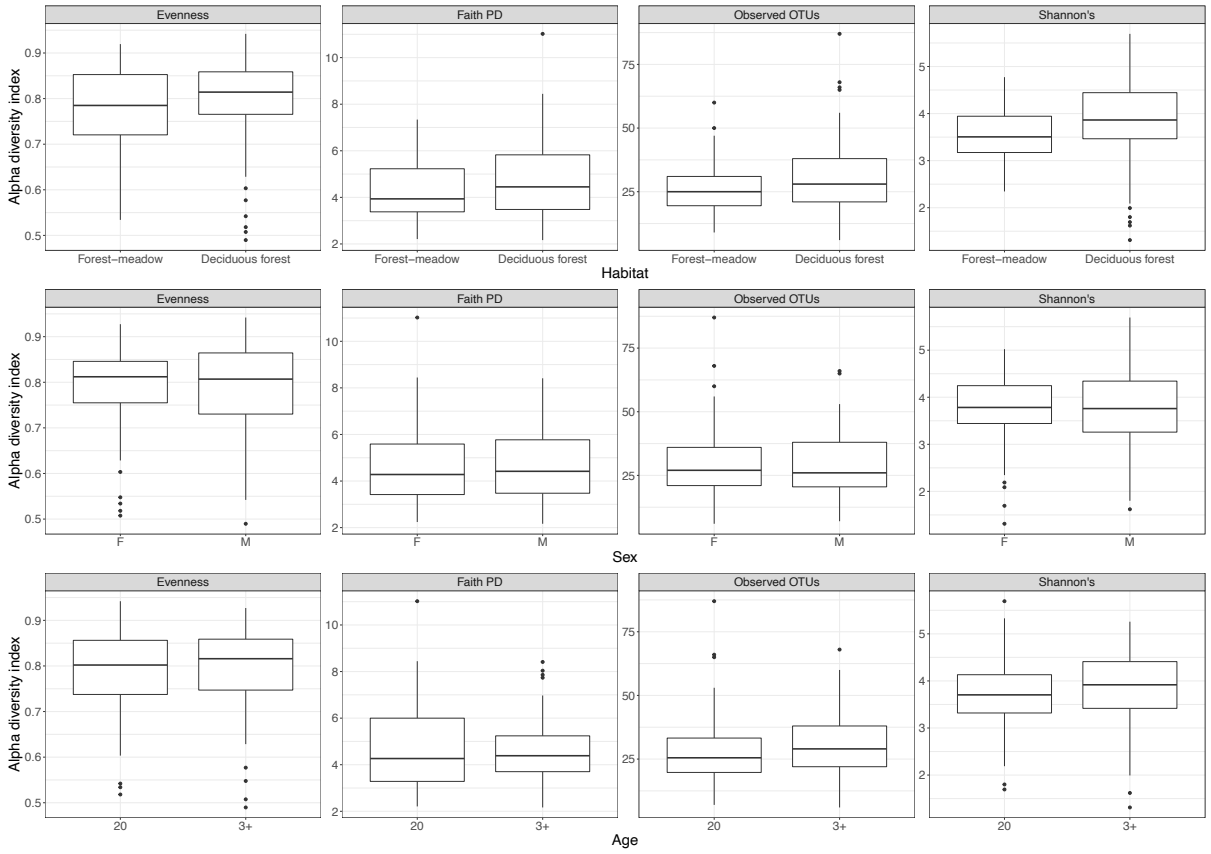


Figure 2. Comparisons of alpha-diversity metrics (columns) between habitats, sexes and age classes. Boxes indicate inter-quantile ranges, horizontal lines are medians, whiskers mark the minimum and maximum ranges, excluding outliers (marked as points).

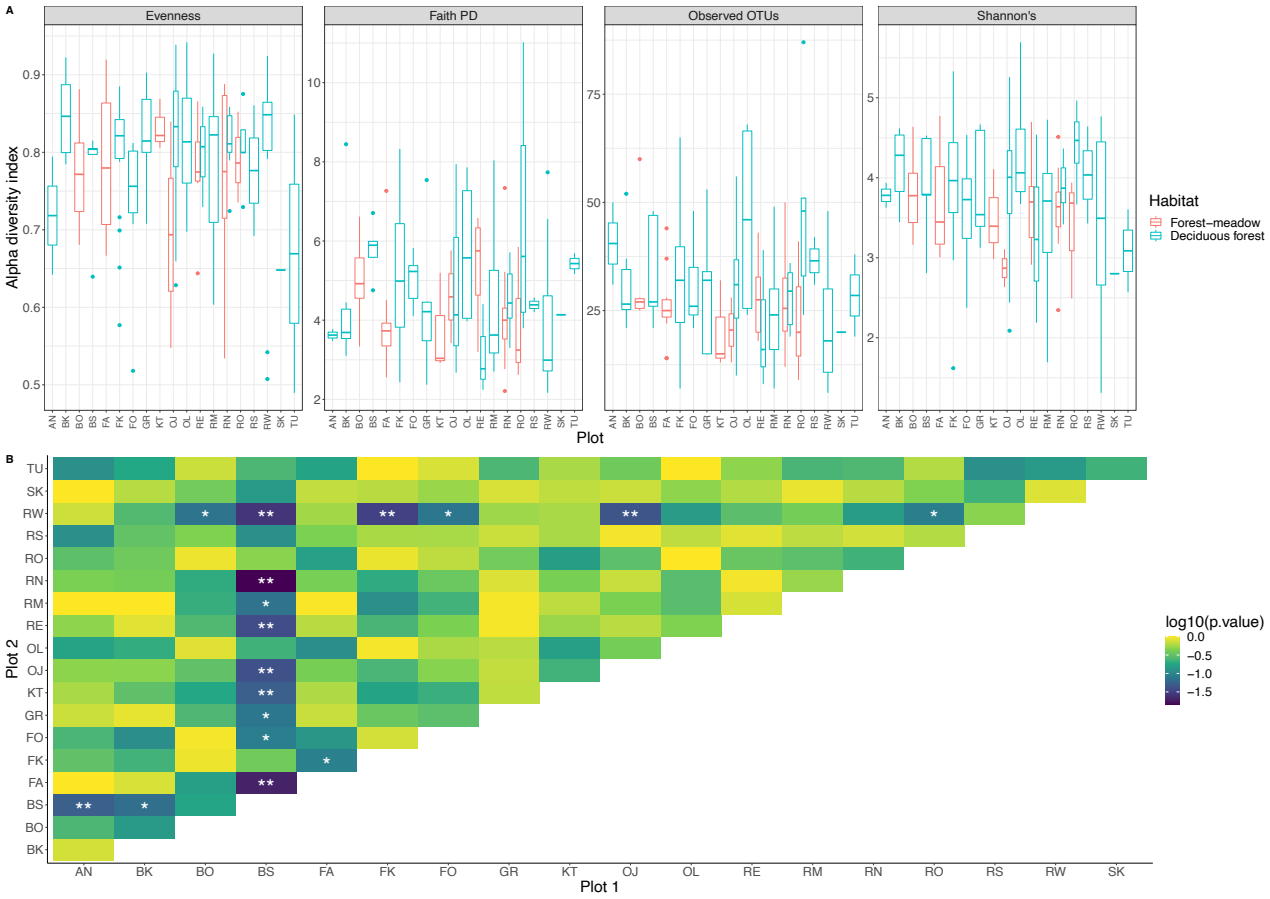


Figure 3. A: comparison of alpha-diversity indices between the studied plots. Plot elements identical with Figure 2. Plots that have both forest-medaow and deciduous forest habitats are split into two narrower box-plots. B: Heatmap of *p*-values from comparisons of Faith's phylogenetic diversity in pairs of plots. Dark values indicate more significant difference. Significant tests marked with asterisk (** <0.05; * <0.1).

Beta-diversity differences

Principal coordination analysis based on unweighted and weighted (by abundance) UniFrac distances indicated that the forest-meadow and deciduous forest habitats tended to occupy slightly different portions of the community diversity scale, which was especially evident in case of weighted (i.e., taking into account taxa abundance) UniFrac distances (Table 1 & 2, Fig. 4). No sex or age differences were detected, although in case of weighted UniFrac distances there was a close-to-significant age-by-sex interaction (Table 1) indicating, that sexes may differ from each other in terms of microbial community composition in an age-dependent matter. The detected habitat differences were robust to between year variation and the between replicate variation taking into account the two sequencing rounds. Also, removing from the set the two mixed-forest plots did not change the results qualitatively and introduced only slight quantitative differences (not shown).

Table 1. Permutation-based (999 resamplings) ANOVA table from the *adonis* plugin, testing for beta-diversity differences across several grouping factors, based on unweighted UniFrac distances. SS – sum of squared deviations between or within groups. MS – mean SS.

	df	*SS*	*MS*	*F*	*P*
Sex	1	0.161	0.161	0.901	0.557
Age	1	0.199	0.199	1.120	0.270
Habitat	1	0.374	0.374	2.102	0.010
Sequencing round	1	2.498	2.498	14.029	0.001
Year	1	0.432	0.432	2.430	0.003
Sex * Age	1	0.145	0.145	0.815	0.686
Residuals	153	27.245	0.178		
Total	159	31.055			

Table 2. Permutation-based (999 resamplings) ANOVA table from the *adonis* plugin, testing for beta-diversity differences across several grouping factors, based on weighted UniFrac distances. SS – sum of squared deviations between or within groups. MS – mean SS.

	df	*SS*	*MS*	*F*	*P*
Sex	1	0.070	0.070	0.705	0.686
Age	1	0.147	0.147	1.489	0.163
Habitat	1	0.457	0.457	4.627	0.001
Sequencing round	1	0.378	0.378	3.827	0.008
Year	1	0.256	0.256	2.594	0.017
Sex * Age	1	0.185	0.185	1.875	0.069
Residuals	150	15.122	0.098		
Total	156	16.617			

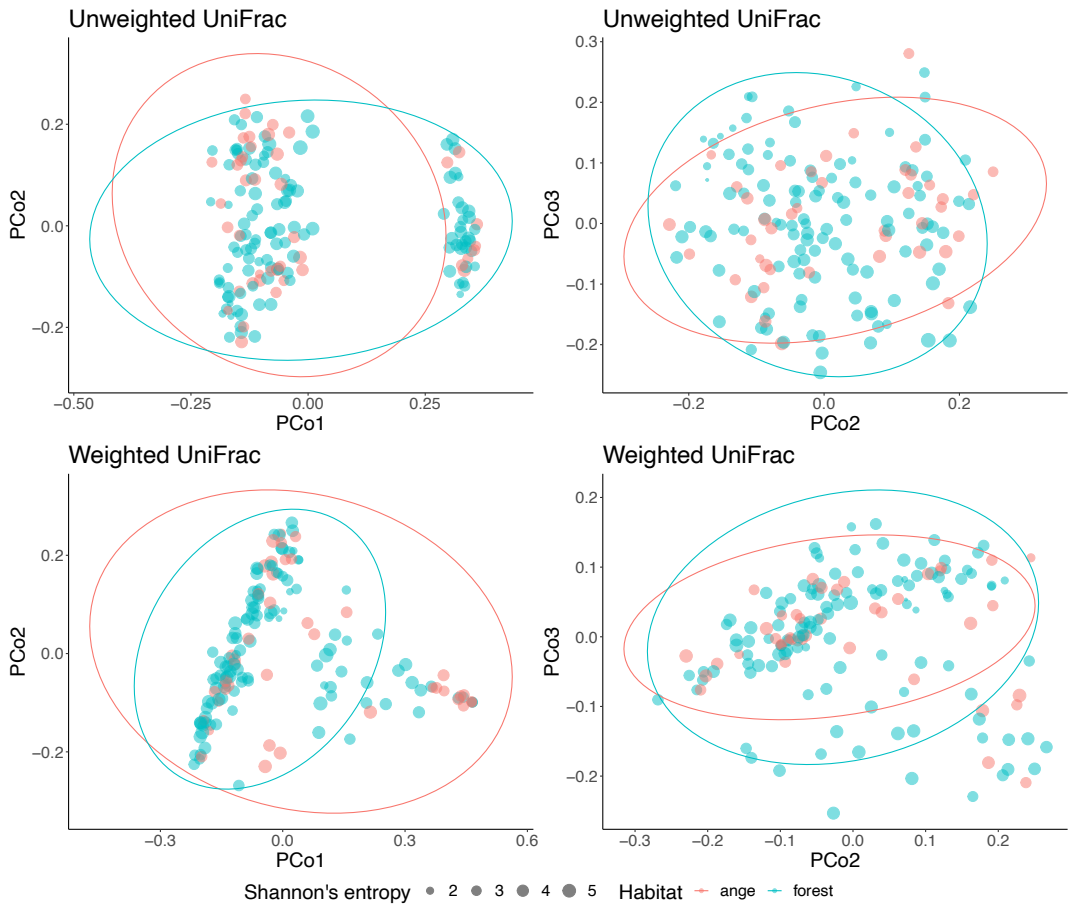


Figure 4. Principal coordination analysis (PCoA) plots based on unweighted (top row) and weighted (bottom row) UniFrac distances. Each point represents one individual sample, habitats are coded by colours. Relative size of points represents the Shannon's entropy index of each sample. Plots present bivariate comparisons of the 1st and 2nd, and the 2nd and 3rd PCo axes. Ellipses mark 95% confidence ellipses around each subgroup.

Taxonomic differences

The most dominant bacterial phyla belonged in both the deciduous forest and forest-meadow classes to: *Proteobacteria* (50.5% vs. 70.1% in forest vs. open habitats, respectively), *Firmicutes* (30.5% vs. 12.6%), *Actinobacteria* (5.5% vs. 2.5%) and *Tenericutes* (4.8% vs. 1.9%). There was also a number of unidentified bacterial taxa (~1% in forest habitats, ~5% in open habitats). Overall, open forest-meadow habitats showed much larger representation of *Proteobacteria* and *Actinobacteria*, whereas deciduous forest habitats showed relatively larger – compared to forest-meadows – prevalence of *Firmicutes* (see Figure 5 for details). At the OUT genus level, differences between the two habitat types were seen in a few taxa. Five most abundant genera were: *Diplorickettsia* (16.3% vs. 43.6% in deciduous forests vs. forest-meadows); *Streptococcus* (18.1% vs. 6.6%); *Acinetobacter* (10.8% vs. 3.6%); an unclassified Bacteria genus (3.1% vs. 11.4%) and *Wolbachia* (2.2% vs. 10.2%). Additionally, 80% of all bacterial genera in deciduous forests comprised also: *Hamiltonella*, *Pseudomonas*, *Carnobacterium*, *Spiroplasma*, *Ureaplasma*, *Anaerococcus*, *Massilia*, *Propionibacterium*, *Janthinobacterium*, *Chryseobacterium*, *Bacillus*, *Lactococcus*, *Lactobacillus*. Forest-

meadows were associated with far less diverse taxa – 80% of its bacterial abundance comprised, beside the abovementioned taxa, also *Pseudomonas*, *Hamiltonella* and *Ureaplasma*.

Differential abundance analysis through the ANCOM pipeline (performed on the core set of taxa, i.e., filtering out OTUs found in less than 10% of all sequenced samples, and collapsing taxa annotations to the level of family) identified several differentially prevalent taxa: *Spiroplasmataceae/Tenericutes* (biased towards deciduous forest); *Diplorickettsiaceae/Proteobacteria* (biased towards deciduous forest); *Burkholderiacaea/Proteobacteria* (biased towards deciduous forest); *Mycoplasmataceae* (biased towards deciduous forest); *Pseudomonadaceae/Proteobacteria* (biased towards deciduous forest); *Moraxellaceae/Proteobacteria* (biased towards deciduous forest); *Pasteurellaceae/Proteobacteria* (biased towards forest-meadows); *Enterobacteriaceae/Proteobacteria* (biased towards deciduous forest); *Micrococcaceae* (biased towards deciduous forest); *Streptococaceae/Firmicutes* (biased towards deciduous forest); *Lactobacillaceae/Firmicutes* (biased towards deciduous forest); *Aerococcaceae/Firmicutes* (biased towards deciduous forest); *Propionibacteriaceae/Actinobacteria* (biased towards deciduous forest).

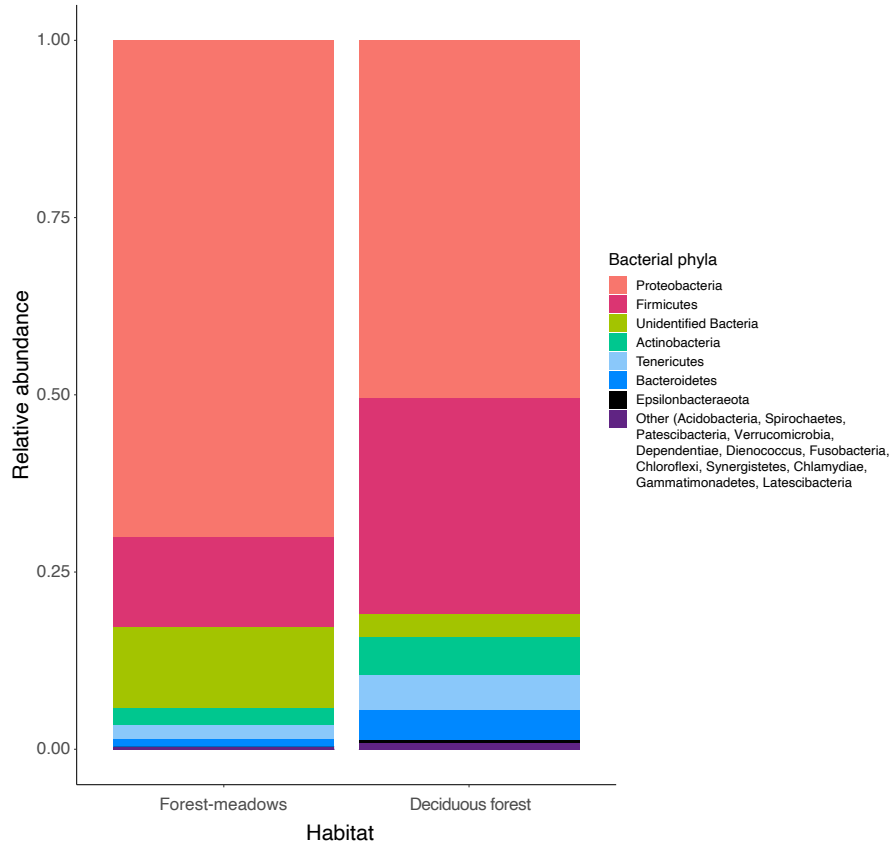


Figure 5. Bacterial phyla identified in the sequenced samples, grouped by the habitat type of each adult bird. The 'Other' group pools together phyla width very small prevalences.

Discussion

In our study, we present the first descriptive account of microbiome diversity and composition in a wild, nest-box population of blue tits from the Swedish island of Gotland. It is one of the few such surveys performed in long-term monitored populations with similar numbers of individuals surveyed. Our study confirms previous results and general patterns seen in wild birds' microbiota. Although we have not observed any significant between-sex or age dependent differences in microbiome composition and diversity, we have detected marked differences between experimental plots that host subsets of nest-boxes in the studied population. Further analysis indicated, that much of the observed differences could be attributed to general habitat differences between the studied plots: individuals breeding in locations covered by dense, rich deciduous forests tended to have visibly more diverse mircobiomes, with higher Shannon's entropy values and more observed taxons. When represented in a reduced-dimensionality space using UniFrac distances, samples coming from the two contrasting habitats occupied disjoint regions of microbial diversity. Differential prevalence analysis indicated several microbial families that were present at contrasting abundances in the two habitat types.

In spite of being an important model species in ecological studies, blue tit microbiomes have not so far been studied using high-throughput sequencing methods. A number of older studies did look into the diversity gut microbes in this species, but in all cases these analyses where done using more traditional DNA fingerprinting methods (Lucas & Heeb, 2005; Benskin *et al.*, 2015) and identified low numbers of actual OTUs. Thus, it is difficult to compare these studies to ours. The closely related great tit (*Parus major*) was already studied using next-generation sequencing microbiome typing and results of those studies largely confirm our general account of microbiome composition in blue tits. Kropáčková *et al.*, (2017) analysed oral and faecal microbiomes from great tits from a Czech population. Bacterial composition of faecal samples was similar to this seen in our study, with domination of *Proteobacteria* and *Firmicutes*, and significant prevalence of *Actinobacteria* and *Tenericutes*. At the genus level, the dominant bacteria noted in this study were also identified as highly prevalent in our population (*Ureaplasma*, *Chryseobacterium*, *Carnobacterium*). Similar microbiome compositions were reported for great tits in several other studies (Davidson *et al.*, 2019, 2020), but population differences may paly role to some extent (e.g., Teyssier *et al.* (2018) reported great tit microbiomes dominated by *Firmicutes* and *Actinobacteria*, with only small prevalence of *Proteobacteria*; notably, their analyses were done on fledglings, contrary to our study and other cited studies which used at least 1-year old adults). Blue tit microbiomes do significantly differ from microbiomes of other bird species (e.g., Darwin's finches – domination of *Firmicutes* and *Actinobacteria* (Loo *et al.*, 2019); swan geese – domination of *Firmicutes* (Wu *et al.*, 2018); white ibises – domination of *Firmicutes* (Murray *et al.*, 2020); great bastards – domination of *Firmicutes* and *Bacteroidetes* (Liu *et al.*, 2020)), which may reflect general differences due to different dietary niches (Waite & Taylor, 2015; Grond *et al.*, 2018).

In our study we did not see any sex or age dependent differences, which agrees with studies looking at similar patterns in great tits (Kropáčková *et al.*, 2017; Davidson *et al.*, 2019). Teyssier *et al.* (2018) did report age-related changes in microbiome composition (reduction in *Proteobacteria* abundance and increase in *Firmicutes* abundance), but that study looked at shifts within a short time window in juveniles (from the age of 8 days to 15 days), i.e., in period when microbiomes may still be in the assembly phase and far from their final composition.

Habitat differences observed in our study indicate, that habitats closer to the typical habitat of blue tits (i.e., a deciduous forest) are linked to larger taxonomic diversity of microbiome communities. They also significantly differ in terms of their beta-diversity metrics (i.e., bacterial communities in each habitat type are on average more similar to each other than to communities in the opposite habitat type). More in-depth analyses are needed to provide better understanding of the actual ecological differences between the two described habitats. Population wide data indicates that they clearly differ in phenological parameters, and also possible in their ability to provide adequate food basis for breeding birds (see Supplementary Figure 1 and the Materials and Methods section). Own, unpublished observations indicate, that deciduous forest habitats are much more heterogenous, less exposed to predators, and support a more diverse array of possible food sources than forest-meadows. Differential abundance of some of the detected microbial clades seems to confirm this. Deciduous forest habitats yielded much larger abundance of *Wolbachia* in the tit microbiome. Unfortunately, taxonomic analysis could not identify specific species/strains of *Wolbachia* – but this result suggest, that deciduous forest can be more abundant in certain *Wolbachia*-carrying insects, and that this translates into dietary differences in the two opposite habitat types. Evidence, that habitat diversity influences blue tit diet was recently presented using taxonomic barcoding of faeces material in this species (Shutt *et al.*, 2020). One of their interesting observations is an increase, with increasing tree diversity, of *Diptera* insects as diet components in blue tits, with only minor changes in abundances of other insect orders. *Diptera*-specific *Wolbachia* could therefore be responsible for the observed microbiome patterns. Increased prevalence of *Spiroplasmataceae* (known insect haemolymph parasites) in deciduous forest microbiomes also aligns with this hypothesis.

In terms of the digestive role of bacteria, deciduous forest blue tits had microbiomes with higher abundance of *Streptococcaceae*, *Enterobacteriaceae*, *Propionibacteriaceae* and *Lactobacillaceae* – all containing many fermenting bacteria species. Their presence for sure facilitates break down of many carbohydrate compounds, but without exact knowledge of diets in the two opposing habitats it is difficult to conclude which diet components may be responsible for the observed differences. In a recent dietary manipulation experiment, supplementing great tits with insects did not selectively affect any of these taxa (Davidson *et al.*, 2019), leaving open the question about insect diet composition impact on the fermenting bacteria. In general, the link between within-species/genus diet differences is already well established in the literature (Davidson *et al.*, 2019; Loo *et al.*, 2019). Experimental manipulation of great tit diets by shifting them into more insect-larvae based led to an increase in the proportion of *Firmicutes* and drop in the proportion of *Proteobacteria* in bird microbiomes (Bodawatta *et al.*, 2020) – a trend seen in our study in case of deciduous forest birds. Also, a comparative study of several insectivorous and omnivorous species indicated that insectivorous species tended to have more *Gammaproteobacteria* and *Enterobacteriaceae* than omnivorous taxa. Putting our results in the context of the abovementioned ones will be possible only when more data on habitat-induces diet differences will be available. An important step in determining how much of the observed diversity is fixed at the between individual level, and how much of it depends on year-by-year dietary changes, will require comparing samples from the same individuals breeding in different seasons in different types of habitats.

There were several technical issues that may have affected our results. One of major problems in similar microbiome studies, using low microbial biomass DNA extracts, is dealing with contaminating bacteria that enter samples during fieldwork, and as part of laboratory handling or from laboratory kit microbes (“kitome” and “splashome”) (Edmonds

& Williams, 2017; Eisenhofer *et al.*, 2019; Hornung *et al.*, 2019). Recommendation varies from analysing all identified OTUs together to complete filtering of putative contaminants (Eisenhofer *et al.*, 2019). We have applied a conservative approach of dealing with sample contamination: we removed all OTUs that were present in negative controls (that is, samples extracted without faecal material, samples containing swabs of the field laboratory area and PCRs run without extracted bacterial DNA). Majority of contaminants belonged to the genera *Cutibacter* (mostly *C. acne*) and *Staphylococcus*, i.e., two major bacterial groups commonly seen as opportunistic commensals on human skin and comprised 221 OTUs in total (~8% of the original number of all identified OTUs). It is likely that some of the removed taxa may be genuine components of bird microbiomes and may have cross-contaminated negative controls, or are environmental bacteria genuinely encountered in bird microbiomes (Eisenhofer *et al.*, 2019). Nonetheless, low proportion of removed OTUs ensures that this approach should not affect our results significantly. Also, since bird microbiomes are likely to be different than mammalian ones, our conservative approach should still be robust.

Another technical issue common in faecal microbiome is interference of common faeces' components with downstream molecular techniques. Insectivorous bird faeces contain large amounts of uric acid (general feature of bird faeces), polyphenols (especially important in caterpillar-eating birds, where polyphenols come from large amounts of plant material eaten by insects) and fat which may decrease the efficiency of amplicon-generating PCR (Schrader *et al.*, 2012). The used extraction kit should deal with similar contaminants satisfactorily. Still, some of heterogeneity observed between individuals in prevalence of specific OTUs may result from random amplification failure. Indeed, earlier studies on a closely related species, the great tit, demonstrated that oral microbiome (likely less affected by the PCR inhibitors issue) shows much greater microbial diversity than faecal microbiome (Kropáčková *et al.*, 2017). To certain extent this surely represents genuine body-region variation in microbial communities but establishing real impact of using faecal samples in place of, e.g., swabs requires more large-scale studies similar to ours.

Other confounding factors that might influence our results (such as year and the ID of sequencing round) did not impact the significance of observed community composition differences. We have identified significant divergence between study years, and between sequencing rounds – which could be expected. In the studied population, years typically differ substantially from each other in terms of climatic conditions and consequently food abundance patterns. Sequencing replicates (i.e., two separate kits used in sequencing each half of the included samples) also can be expected to differ as they came from two different production batches. Nonetheless, robustness of microbiome sensitivity to those technical aspects makes the observed pattern even more unequivocal.

To conclude, our study represents one of the first large-scale accounts of the microbial diversity in faecal gut microbiomes from a wild blue tit population. It demonstrates habitat-specific differences in microbiomes that may be attributable to general food-base diversity and habitat richness. Additional studies are needed to elucidate the actual origin of the observed differences, both at the level of the studied species, but also comparatively (i.e., whether other species breeding sympatrically with blue tits in the same population show similar patterns of microbiome differentiation).

Acknowledgments

We thank Lars Gustafsson for access to the Gotland population, and Dorota Lutyk, Katarzyna Janas, Wioleta Oleś, Rita Fragueira, Agnieszka Gudowska, Barbara Regueira, Aneta Arct, Tomasz Kowalczyk and Julia Barczyk for help with fieldwork over the included study years. This project was funded via a Iuventus Plus 5 grant awarded to SMD (no. IP2015 016374).

References

Beale, A.L., Kaye, D.M. & Marques, F.Z. 2019. The role of the gut microbiome in sex differences in arterial pressure. *Biol. Sex Differ.* **10**: 22.

Benskin, C.M.H., Rhodes, G., Pickup, R.W., Mainwaring, M.C., Wilson, K. & Hartley, I.R. 2015. Life history correlates of fecal bacterial species richness in a wild population of the blue tit Cyanistes caeruleus. *Ecol. Evol.* **5**: 821–835.

Benson, A.K. 2016. The gut microbiome - An emerging complex trait. *Nat. Genet.* **48**: 1301–1302. Nature Publishing Group.

Bergamaschi, M., Maltecca, C., Schillebeeckx, C., McNulty, N.P., Schwab, C., Shull, C., *et al.* 2020. Heritability and genome-wide association of swine gut microbiome features with growth and fatness parameters. *Sci. Rep.* **10**: 10134. Springer US.

Björk, J.R., Dasari, M., Grieneisen, L. & Archie, E.A. 2019. Primate microbiomes over time: Longitudinal answers to standing questions in microbiome research. *Am. J. Primatol.* **81**: 1–23.

Bodawatta, K.H., Freiberga, I., Puzejova, K., Sam, K., Poulsen, M. & Jønsson, K.A. 2020. *Flexibility and resilience of Great tit (Parus major) gut microbiomes to changing diets*. In Review.

Bodawatta, K.H., Sam, K., Jønsson, K.A. & Poulsen, M. 2018. Comparative Analyses of the Digestive Tract Microbiota of New Guinean Passerine Birds. *Front. Microbiol.* **9**. Frontiers.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., *et al.* 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**: 852–857.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A. & Holmes, S.P. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**: 581–583.

Cao, J., Hu, Y., Liu, F., Wang, Y., Bi, Y., Lv, N., *et al.* 2020. Metagenomic analysis reveals the microbiome and resistome in migratory birds. *Microbiome* **8**: 1–18.

Cryan, J.F. & O'Mahony, S.M. 2011. The microbiome-gut-brain axis: From bowel to behavior. *Neurogastroenterol. Motil.* **23**: 187–192.

Davidson, G., Wiley, N., Cooke, A., Johnson, C., Fouhy, F., Reichert, M., *et al.* 2019. Diet induces parallel changes to the gut microbiota and problem solving performance in a wild bird. *Sci. Rep.* 1–13. Nature Publishing Group UK.

Davidson, G.L., Cooke, A.C., Johnson, C.N. & Quinn, J.L. 2018. The gut microbiome as a driver of individual variation in cognition and functional behaviour. *Philos. Trans. R. Soc. B Biol. Sci.* **373**.

Davidson, G.L., Somers, S.E., Wiley, N., Johnson, C.N., Reichert, M.S., Ross, P.R., *et al.* 2020. A time-lagged association between the gut microbiome, nestling growth and nestling survival in wild great tits. *biorXiv*, doi: doi.org/10.1101/2020.09.30.320804.

Edmonds, K. & Williams, L. 2017. The Role of the Negative Control in Microbiome Analyses. *FASEB J.* **31**: 940.3-940.3.

Eisenhofer, R., Minich, J.J., Marotz, C., Cooper, A., Knight, R. & Weyrich, L.S. 2019. Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. *Trends Microbiol.* **27**: 105–117. Elsevier Ltd.

Escallón, C., Belden, L.K. & Moore, I.T. 2019. The Cloacal Microbiome Changes with the Breeding Season in a Wild Bird. *Integr. Org. Biol.* **1**.

Grond, K., Sandercock, B.K., Jumpponen, A. & Zeglin, L.H. 2018. The avian gut microbiota: community, physiology and function in wild birds. *J. Avian Biol.* **49**: 1–19.

Hicks, A.L., Lee, K.J., Couto-Rodriguez, M., Patel, J., Sinha, R., Guo, C., *et al.* 2018. Gut microbiomes of wild great apes fluctuate seasonally in response to diet. *Nat. Commun.* **9**. Springer US.

Hird, S.M. 2017. Evolutionary biology needs wild microbiomes. *Front. Microbiol.* **8**: 1–10.

Hornung, B.V.H., Zwittink, R.D. & Kuijper, E.J. 2019. Issues and current standards of controls in microbiome research. *FEMS Microbiol. Ecol.* **95**.

Jašarević, E., Morrison, K.E. & Bale, T.L. 2016. Sex differences in the gut microbiome–brain axis across the lifespan. *Philos. Trans. R. Soc. B Biol. Sci.* **371**: 20150122. Royal Society.

Kohl, K.D., Brun, A., Caviedes-Vidal, E. & Karasov, W.H. 2019. Age-related changes in the gut microbiota of wild House Sparrow nestlings. *Ibis* **161**: 184–191.

Kolodny, O. & Schulenburg, H. 2020. Microbiome-mediated plasticity directs host evolution along several distinct time scales. *Philos. Trans. R. Soc. B Biol. Sci.* **375**: 20190589. Royal Society.

Kropáčková, L., Pechmanová, H., Vinkler, M., Svobodová, J., Velová, H., Těšičký, M., *et al.* 2017. Variation between the oral and faecal microbiota in a free-living passerine bird, the great tit (Parus major). *PLoS ONE* **12**: 1–18.

Liu, G., Meng, D., Gong, M., Li, H., Wen, W., Wang, Y., *et al.* 2020. Effects of Sex and Diet on Gut Microbiota of Farmland-Dependent Wintering Birds. *Front. Microbiol.* **11**. Frontiers.

Loo, W.T., García-Loor, J., Dudaniec, R.Y., Kleindorfer, S. & Cavanaugh, C.M. 2019. Host phylogeny, diet, and habitat differentiate the gut microbiomes of Darwin's finches on Santa Cruz Island. *Sci. Rep.* **9**: 1–12.

Lucas, F.S. & Heeb, P. 2005. Environmental factors shape cloacal bacterial assemblages in great tit Parus major and blue tit P. caeruleus nestlings. *J. Avian Biol.* **36**: 510–516.

Mandal, S., Van Treuren, W., White, R.A., Eggesbø, M., Knight, R. & Peddada, S.D. 2015. Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb. Ecol. Health Dis.* **26**: 1–7.

Mazmanian, S.K., Cui, H.L., Tzianabos, A.O. & Kasper, D.L. 2005. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* **122**: 107–118.

Murray, M.H., Lankau, E.W., Kidd, A.D., Welch, C.N., Ellison, T., Adams, H.C., *et al.* 2020. Gut microbiome shifts with urbanization and potentially facilitates a zoonotic pathogen in a wading bird. *PLOS ONE* **15**: e0220926. Public Library of Science.

Noguera, J.C., Aira, M., Pérez-Losada, M., Domínguez, J. & Velando, A. 2018. Glucocorticoids modulate gastrointestinal microbiome in a wild bird. *R. Soc. Open Sci.* **5**: 1–8.

Pafčo, B., Sharma, A.K., Petrželková, K.J., Vlčková, K., Todd, A., Yeoman, C.J., *et al.* 2019. Gut microbiome composition of wild western lowland gorillas is associated with individual age and sex factors. *Am. J. Phys. Anthropol.* **169**: 575–585.

Pitala, N. 2007. Causes and consequences of variation in nestling immune function. University of Helsinki.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., *et al.* 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**: D590–D596.

Ren, T., Boutin, S., Humphries, M.M., Dantzer, B., Gorrell, J.C., Coltman, D.W., *et al.* 2017. Seasonal, spatial, and maternal effects on gut microbiome in wild red squirrels. *Microbiome* **5**: 163.

Schrader, C., Schielke, A., Ellerbroek, L. & Johne, R. 2012. PCR inhibitors – occurrence, properties and removal. *J. Appl. Microbiol.* **113**: 1014–1026.

Shutt, J.D., Nicholls, J.A., Trivedi, U.H., Burgess, M.D., Stone, G.N., Hadfield, J.D., *et al.* 2020. Gradients in richness and turnover of a forest passerine's diet prior to breeding: A mixed model approach applied to faecal metabarcoding data. *Mol. Ecol.* 0–2.

Spor, A., Koren, O. & Ley, R. 2011. Unravelling the effects of the environment and host genotype on the gut microbiome. *Nat. Rev. Microbiol.* **9**: 279–290. Nature Publishing Group.

Teyssier, A., Lens, L., Matthysen, E. & White, J. 2018. Dynamics of gut microbiota diversity during the early development of an avian host: Evidence from a cross-foster experiment. *Front. Microbiol.* **9**: 1–12.

Toivanen, P., Vaahtovuo, J. & Eerola, E. 2001. Influence of Major Histocompatibility Complex on Bacterial Composition of Fecal Flora. *Infect. Immun.* **69**: 2372–2377.

Tung, J., Barreiro, L.B., Burns, M.B., Grenier, J.-C., Lynch, J., Grieneisen, L.E., *et al.* 2015. Social networks predict gut microbiome composition in wild baboons. *eLife* **4**: e05224. eLife Sciences Publications, Ltd.

Waite, D.W. & Taylor, M.W. 2015. Exploring the avian gut microbiota: Current trends and future directions. *Front. Microbiol.* **6**: 1–12.

Wu, Y., Yang, Y., Cao, L., Yin, H., Xu, M., Wang, Z., *et al.* 2018. Habitat environments impacted the gut microbiome of long-distance migratory swan geese but central species conserved. *Sci. Rep.* **8**: 1–11. Springer US.