The gut microbiome in patients with chronic lymphocytic leukemia

The gut microbiome, an ecosystem formed by commensal, symbiotic, and pathogenic microorganisms colonizing the gastrointestinal tract, may impact both immune function and carcinogenesis. The host immune system plays a vital role in the maintenance of gut microbiome homeostasis by establishing a balance between eliminating invading pathogens and promoting the growth of beneficial microbes. When this balance is disturbed, a state of dysbiosis arises in the microbial ecosystem. Conditions found to be associated with gut dysbiosis include, but are not limited to, inflammatory disorders including inflammatory bowel diseases (IBD),[1] diabetes,[2] as well as obesity,[3] and asthma.[4] The gut microbiome also seems to influence cancer susceptibility, and to correlate with tumorigenesis and progression.[5] Within the area of hematological malignancies, studies focusing on acute leukemias, lymphoproliferative disorders, and multiple myeloma, have found microbiome dysbiosis and decreased microbiome diversity to be related to microenvironmental alterations induced by the disease itself, chemotherapy, and/or antibiotics.[6]

To our knowledge, no previous studies have investigated the gut microbiome in patients with chronic lymphocytic leukemia (CLL). As CLL represents an antigen-driven malignancy with immune dysfunction,[7] the gut microbiome could both be implicated in the pathogenesis of CLL through antigenic drive, and contribute to the distortion of the immune system. However, the CLL microbiome itself may also be impacted by the immune dysfunction as well as reflect the increased prescription of antimicrobials for this patient group. Thus, our study focusing on the fecal microbiome in patients with CLL aimed to describe perturbations in the gut microbial composition, and to characterize potential signature for CLL-related gut dysbiosis.

We included fecal samples from ten patients enrolled during regular out-patient visits. Control cohorts were selected from an array of previously published cohorts[8,9] and matched to the CLL cohort on criteria including mean age, residency, year of sampling, and sampling methods. Using shotgun metagenomic sequencing and taxonomical profiling, we assessed fecal microbiome composition, diversity, and dynamics. Bioinformatic analyses were performed following the state-of-the-art methods as well as using innovative tailor-made approaches.

Ten patients diagnosed with CLL delivered 12 stool samples between January 2016 and October 2018. Sampling was prior to treatment for eight patients and after treatment for four patients, thus, two patients, were sampled both before and after treatment. No patients received antibiotics within 4 months prior to sample collection. For an overview of patient and control baseline characteristics, and patient treatment see the *Online Supplementary Figure S1A* to *C*. Changes in relative abundance of ten major bacterial genera in patients sampled both before and after treatment are demonstrated in the *Online Supplementary Figure S1D*.

We observed high intra-variability among the CLL microbiomes (Figure 1A). At the genus level, *Bacteroides* was the most abundant genus in six of the 12 CLL samples. Additionally, there was a trend of *Bacteroides* acquiring bacterial dominance (>30% relative abundance) in five of 12 samples. Several other bacterial genera such as *Parabacteroides*, *Prevotella*, and *Acinetobacter* also acquired bacterial dominance among CLL patients. We next assessed microbiome diversity in CLL patients and healthy controls. Richness (the observed number of species) was lowest for patients with CLL and was significantly different compared to each of the two control cohorts (Observed: CLL *vs*. C1: median, 53 *vs*. 69; P=0.00057; CLL *vs*. C2: median, 53 *vs*. 73; P=6.8e−05; Figure 1B). Patients with CLL also showed lower α diversity compared to the control groups when assessed by Shannon index (CLL vs. C1: median, 1.90 *vs*. 2.90; P=2.1e−05; CLL *vs*. C2: median, 1.90 *vs*. 2.75; P=0.00057) and *I*nvSimpson index (CLL *vs*. C1: median, 4.18 *vs*. 11.94; P=4.3e−05; CLL *vs*. C2: median, 4.18 *vs*. 10.05; P=0.00057; Figure 1B). The two control cohorts also demonstrated a significant difference in diversity between one another (C1 *vs*. C2: Shannon: median, 2.90 *vs*. 2.75; P=0.00057; InvSimpson: median, 11.94 *vs*. 10.05; P=0.017). In addition to reduced diversity, a difference in the specific microbial composition between the CLL cohort and the two control cohorts was observed: at the phylum level, we focused on the differences in *Bacteroidetes*, *Firmicutes*, *Proteobacteria* and *Actinobacteria* as they comprised 95% of the total bacterial content in CLL patients, on average. The distribution of these four phylotypes across each cohort highlights a significantly higher abundance of *Bacteroidetes* and *Proteobacteria* relative to the controls (Figure 1C). In contrast, both control groups showed greater proportions of *Firmicutes* and *Actinobacteria*.

At the family level, *Bacteroidaceae*, *Prevotellaceae*, *Clostidiaceae*, *Lachnospiraceae* and *Ruminococcaceae* were the most abundant five families among both CLL patients and controls (Figure 2A). *Bacteroidaceae* were present in

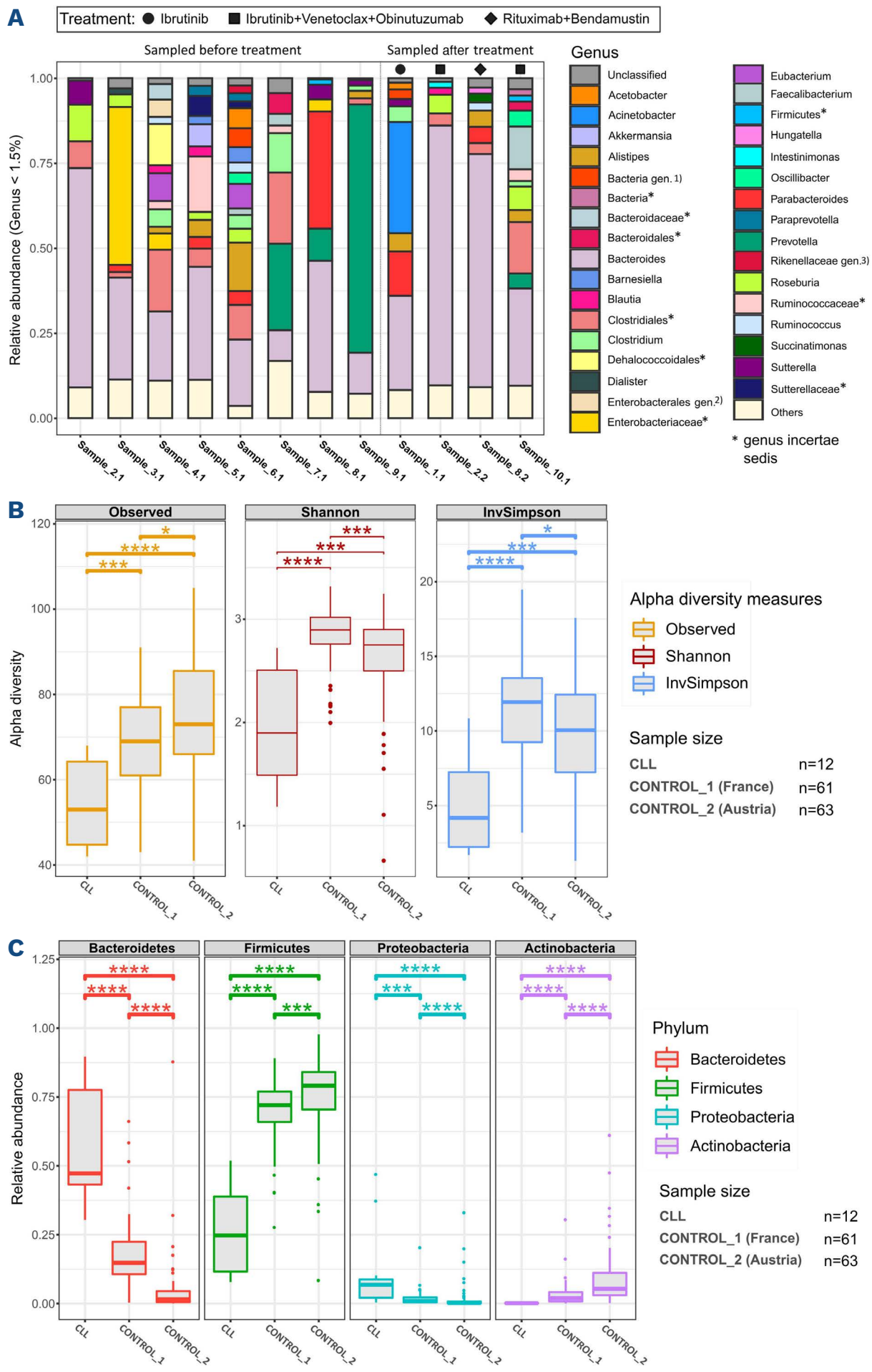

Figure 1. Gut microbiome composition and diversity in chronic lymphocytic leukemia patients and controls. (A) The relative abundance of bacterial genera in all 10 chronic lymphocytic leukemia (CLL) patients (12 samples). Bacterial genera whose abundance was <1.5% in a sample were grouped as 'Others'. Sequences that could not be assigned to a genus were grouped as 'Unclassified'. Taxa having zero counts across all samples were removed prior to all analyses. If the sample was taken after treatment, the treatment regimen is indicated by a corresponding shape on the top of each bar and described in the legend. Bacterial abundance was visualized using stacked barplots from R package *ggplot2*. Unambiguously assigned genera: 1) [Rhodospirillum/Lactobacillus/Azospirillum]; 2) [Enterobacter/Escherichia/Klebsiella/Serratia]; 3) [Tidjanibacter/Alistipes]. (B) Fecal diversity in CLL samples and healthy samples at genus level (α diversity measures: observed number of genera, Shannon and Inverse Simpson indexes). In box plots, box edges represent the 25th and 75th percentiles, the center line shows the median and whiskers extend from the box edges to the most extreme data point. Data beyond the end of the whiskers are plotted individually as dots. The *P*-values (adjusted for multiple testing with the Benjamini-Hochberg [BH]) obtained upon Wilcoxon rank-sum tests are indicated, values <0.05 were considered significant. Not significant (Ns) $P>0.05$; *$P<0.05$; **$P< 0.01$; ***$P<0.001$; ****$P<0.0001$. (C) Relative abundance of 4 major bacterial phyla forming the microbiota in the CLL cohort. Box plots are constructed as described in (B).

high proportions in the CLL cohort. Interestingly, a major difference was also observed between the samples from patients with CLL and the healthy controls with lower relative abundance of *Lachnospiraceae* and *Ruminococcaceae* taxa, while the dissimilarities between the control groups were small. We further aimed to identify groups of bacterial taxa differentially abundant between patients with CLL and controls. Analyzing the composition of microbiomes based on the abundance at the genus level, we detected *Bacteroides*, *Sutterella* and *Parabacteroides*, to be overrepresented in CLL relative to the average microbiome (abundance of a taxon across all CLL and healthy samples). In contrast, we identified a group of taxa including *Bifidobacterium*, *Anaerostipes*, and nine other bacterial genera to be underrepresented among patients with CLL as compared to controls (Figure 2B). Bearing in mind the complexity of bacterial communities, we also sought to characterize groups of bacteria co-occurring across all CLL and healthy samples. We calculated proportional co-occurrence coefficient for all pairs of bacter-

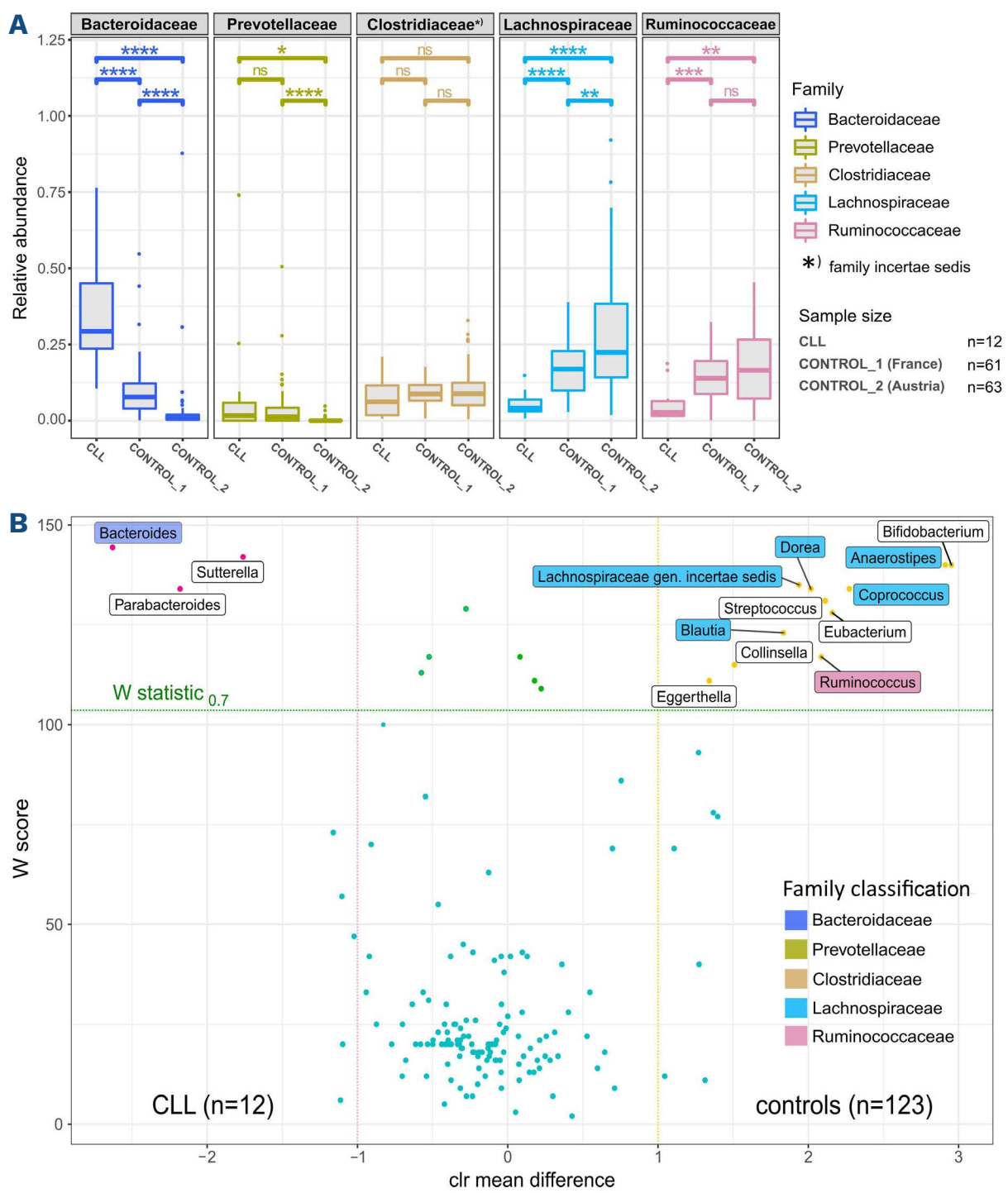

Figure 2. Relative and differential abundance of bacterial families and genera. (A) Relative abundance of five most abundant bacterial families in chronic lymphocytic leukemia (CLL) and healthy feces samples. Box plots are constructed as described in Figure 1(B). Not significant (Ns) $P>0.05$; *$P<0.05$; **$P<0.01$; ***$P<0.001$; ****$P<0.0001$. (B) Differential abundance of bacterial genera between CLL and the merged control cohorts. Bacterial genera are color coded according to its higher taxonomic rank - family. The y-axis W value represents number of times the null hypothesis (H0: the average abundance of a given taxa is equal across cohorts) was rejected for a given taxon. The x-axis value represents the centered log-ratio (clr)-transformed mean difference in abundance of a given taxon between the CLL and healthy groups with respect to average abundance of a given taxon. Positive value at the x-axis indicates bacterial genera being differentially abundant in controls, negative value indicates bacterial genera being differentially abundant in CLL cohort. Only bacterial genera with null hypothesis rejected in >70% of cases and clr mean difference -/+ 1 are labeled. The analysis and volcano plot visualization were done in R by implementation of Analysis of Compositions of Microbiomes (ANCOM).

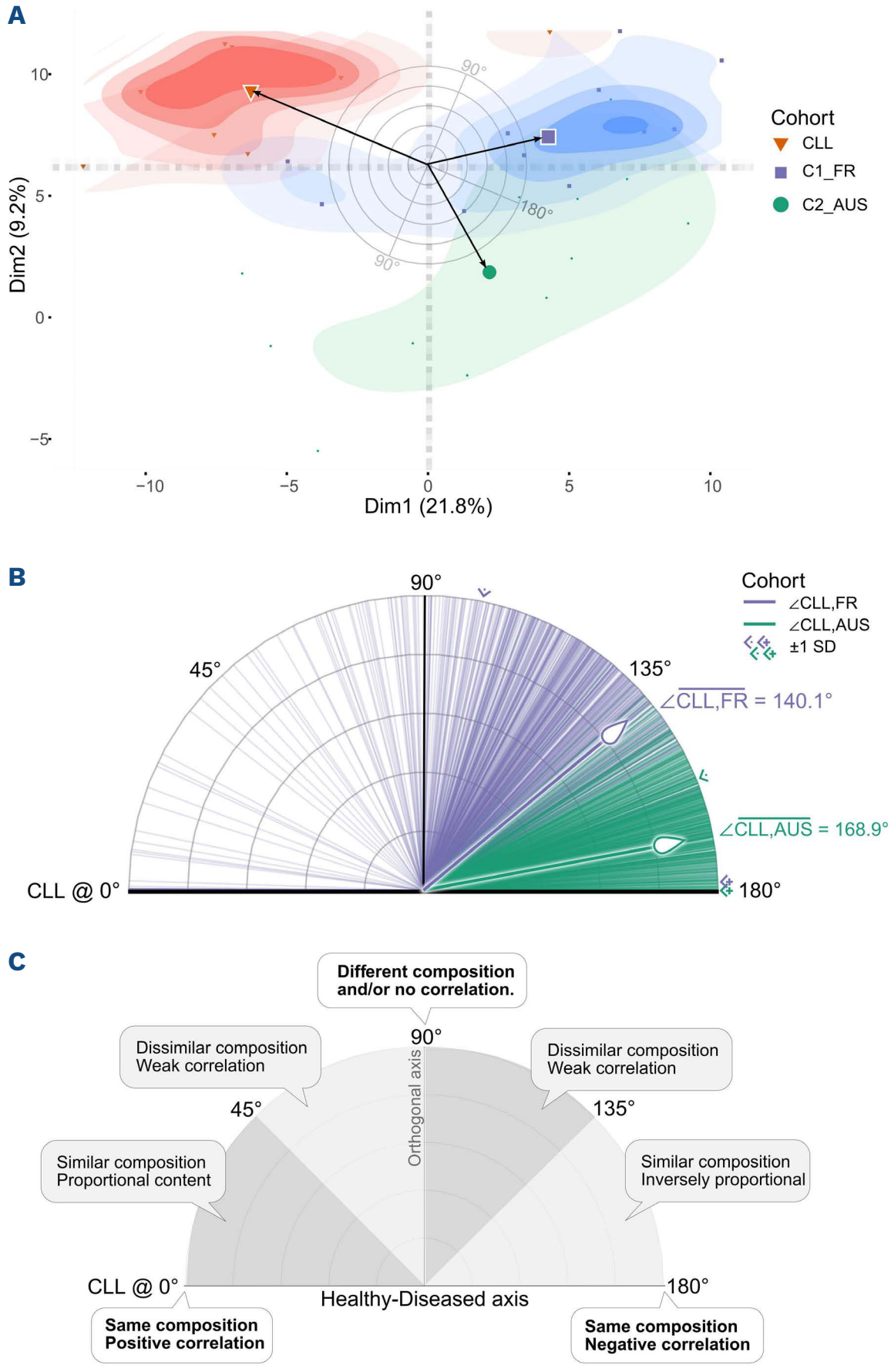

Figure 3. Analysis of covariance by principal component analysis. In order to assess similarity between the chronic lymphocytic leukemia (CLL) microbiome and the control microbiome in a multidimensional space, a principal component analysis (PCA) was performed. (A) The biplot illustrates the distance between the CLL cohort and the 2 control cohorts in terms of a 2-dimensional representative plot of 1,000 iterations run on the original dataset (n=12 per cohort totaling 36 cases per cohort per iteration), delimited by principal component (PC) 1 and PC2. The large symbols (centroids) represent the mean PC score from each cohort. The PC score for each individual is plotted relative to their position along each of the PC. The biplot shows vectors (black) pointing to the centroid of each cohort, as well as the individuals of each cohort (CLL – red triangles, C1_FR – purple squares, C2_AUS – green circles). Colored contour maps represent the density and distribution of individuals grouped by cohorts. The biplot is overlaid with a protractor-like plot displaying degrees from 0 to 180°. The angles between 2 cohorts were calculated as the angles between vectors pointing to centroids of individual cohorts ($\cos\theta = {}^{(a \cdot b)}/_{(||a|| \, ||b||)}$), with CLL being always positioned at 0°. Although C1_FR show certain overlap with CLL, note that the patients with CLL are distinctly clustered from both healthy control cohorts. (B) The protractor-like plot represents all angles identified over the 1,000 iterations between CLL-C1_FR and CLL-C2_AUS. The mean of all centroid vectors per cohort is drawn as a thick line with a white symbol at its end. The standard deviation (SD) is visualized by arrows of a color corresponding to the cohort, on the outside of the plot. (C) The protractor-like plot provides interpretation of the angles on healthy-diseased axis. An angle between vectors is interpreted as an approximation of the correlation and the similarity between the cohorts' variables; i.e., the C1_FR cohort has a dissimilar composition with weaker correlation with the CLL cohort, whereas the C2_AUS control cohort is inversely correlated to the CLL cohort.

ial genera across the CLL and healthy cohorts. When we visualized the relationships between the 23 most strongly proportional taxa, three clusters were formed (*Online Supplementary Figure S2A*). Abundance/depletion of the proportional bacterial taxa across all samples revealed clear discrimination between CLL patients and healthy controls (*Online Supplementary Figure S2B; Online Supplementary Table S1*). An overlap of the two analytical methodologies indicated lack of bacteria mainly from *Lachnospiraceae* and *Ruminococcaceae* families among the CLL samples. In particular, lower abundances of *Anaerostipes*, *Bifidobacterium*, *Blautia*, *Coprococcus, Dorea*, *Eubacterium*, *Ruminococcus*, and *Streptococcus* were observed in CLL samples when compared to controls. The two approaches also indicated higher abundances of *Bacteroides* and *Parabacteroides* in CLL compared to controls.

While canonical metrics have revealed statistically significant differences between CLL and healthy cohorts, we note that controls are also significantly different from one another across several metrics. Thus, we set out to answer whether there exists a healthy-diseased axis such that healthy patients from both control cohorts co-vary and are distinct from CLL patients. In order to do this we characterized the structure of covariance in each cohort by their centroids in PCA-space (Figure 3A). We assessed the similarity between cohorts by calculating the angles between their centroids within this reduced multidimensional space (interpretation of angles in Figure 3C). On average, CLL gut profiles were inversely proportional, i.e., dissimilar, to C1 & C2 cohorts (∠CLL,C1 =140.1±38.5°, purple; ∠CLL,C2 =168.9±11°, green; Figure 3B). We also observed an acute relationship, i.e., similarity, between healthy cohorts (∠C1,C2 =51±43.9°). These results highlight the presence of a strong inverse axis between diseased and healthy states that is defined by the angles between centroids of the CLL and the two control groups.

The lower fecal diversity observed in the CLL cohort is in line with previous findings of reduced bacterial diversity in other inflammatory conditions including IBD, type 1 and 2 diabetes, and obesity,[10,11] as well as hematologic malignancies.[6,12] We suggest that the decreased diversity/increase in *Bacteroidetes* and *Proteobacteria* observed in the CLL cohort could imply a general loss of bacteria, but may also indicate a loss of complexity for the remaining microbiome.

Most of the bacterial genera in the microbial signature depleted in patients with CLL belonged to *Lachnospiraceae* and *Ruminococcaceae* family. Members of the *Lachnospiraceae* and *Ruminococcaceae* families are among the main producers of short chain fatty acids, which are known to modulate the surrounding microbial environment and to directly interact with the host immune system.[13] Interestingly, *Blautia* (*Lachnospiraceae* family) and *Ruminococcus* (*Ruminococcaceae* family) were also identified with higher abundance among controls than among patients with CLL. High abundance of intestinal *Blautia* has been associated with improved survival upon graft-*versus*-host disease,[14] and is together with *Ruminococcus* often underrepresented in feces samples from patients with colorectal cancer.[15]

In conclusion, despite low sample size, the CLL microbiome demonstrated lower microbiome diversity and lower enrichment of short chain fatty acid-producing bacterial taxa when compared to healthy controls. We hypothesize that the overabundance of bacteria from the *Bacteroidetes* phylum together with depletion of *Lachnospiraceae* and *Ruminococcaceae* family bacteria might play a role in – or is observed due to – immune dysfunction among CLL patients. This microbiome signature is warranting validation and refinement in larger CLL cohorts including patients needing treatment, patients assigned to watch and wait, patients with other hematological malignancies, as well as other healthy cohorts. This as a focus of ours in a follow-up study will hopefully lay foundation for defining microbiome characteristics of hematological malignancies and microbiome signatures discriminating subgroups of patients with CLL.

Authors

Tereza Faitová,[1] Rebecka Svanberg,[1] Caspar da Cunha-Bang,[1] Emma E. Ilett,[2] Mette Jørgensen,[2] Marc Noguera-Julian,[3,4] Roger Paredes,[3,5] Cameron R. MacPherson[2] and Carsten U. Niemann[1,6]

[1]Department of Hematology, Rigshospitalet, Copenhagen, Denmark; [2]PERSIMUNE Center of Excellence, Rigshospitalet, Copenhagen, Denmark; [3]Institut de Recerca de la Sida-IrsiCaixa, Hospital Universitari Germans Trias i Pujol, Badalona, Catalonia, Spain; [4]University of Vic-Central University of Catalonia, Barcelona, Spain; [5]Infectious Diseases Department, Hospital Universitari Germans Trias i Pujol, Badalona, Catalonia, Spain and [6]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

Correspondence:
C. U. NIEMANN - carsten.utoft.niemann@regionh.dk

https://doi.org/10.3324/haematol.2021.280455

Received: December 1, 2021.
Accepted: May 5, 2022.
Prepublished: May 12, 2022.

Disclosures

CUN received research funding and/or consultancy fees from

Abbvie, AstraZeneca, CSL Behring, Janssen, Octapharma, Roche and Takeda outside this work. MNJ is founder and shareholder of Nano1Health S.L.

Contributions

CUN, RS, and TF designed the study; EI, CC and CUN were responsible for sample collection and inclusion of patients; MNJ and RP performed the sequencing; TF, MJ, and CRM performed the bioinformatic and statistical analyses; TF and CUN wrote the first version of the manuscript. All authors read, contributed and approved the final version of the manuscript.

Acknowledgments

The authors would like to express sincere thanks to the patients providing samples for the study and making this kind of research possible. This research would not have been possible without our collaborator, PERSIMUNE Center of Excellence, which provided the infrastructure as well as financial support and exceptional expertise. The staff at the Hematology Department and at the PERSIMUNE biobank at Rigshospitalet were essential for this study as they organized and collected the samples for our research. Computerome team provided computational platform and technical support.

Funding

The study was funded by the Danish Cancer Society (grant R269-A15924) and the Novo Nordisk Foundation (grant NNF16OC0019302). The infrastructure for sampling and analyses was within the Danish National Research Foundation funded PERSIMUNE project (grant 126). The work was also supported by the CLL-CLUE project funded by the European Union.

Data-sharing statement

The datasets generated and analyzed during this study are derived from patients treated in Denmark. The datasets contain sensitive patient data governed by GDPR and Danish law. Due to Danish legislation (Act No. 502 of 23 May 2018) and approvals granted by the Danish Data Protection Agency, it is not possible to upload raw data to a publicly available database. However, access to these data can be made available from the corresponding author on reasonable request, provided a data transfer agreement is entered into according to current regulations.

References

1. Alam MT, Amos GCA, Murphy ARJ, Murch S, Wellington EMH, Arasaradnam RP. Microbial imbalance in inflammatory bowel disease patients at different taxonomic levels. Gut Pathog. 2020;12:1.
2. Larsen N, Vogensen FK, van den Berg FWJ, et al. Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. PloS One. 2010;5(2):e9085.
3. Turnbaugh PJ, Hamady M, Yatsunenko T, et al. A core gut microbiome in obese and lean twins. Nature. 2009;457(7228):480-484.
4. Abrahamsson TR, Jakobsson HE, Andersson AF, Björkstén B, Engstrand L, Jenmalm MC. Low gut microbiota diversity in early infancy precedes asthma at school age. Clin Exp Allergy. 2014;44(6):842-850.
5. Zitvogel L, Galluzzi L, Viaud S, et al. Cancer and the gut microbiota: an unexpected link. Sci Transl Med. 2015;7(271):271ps1.
6. Uribe-Herranz M, Klein-González N, Rodríguez-Lobato LG, Juan M, de Larrea CF. Gut Microbiota influence in hematological malignancies: from genesis to cure. Int J M Sci. 2021;22(3):1026.
7. Svanberg R, Janum S, Patten PEM, Ramsay AG, Niemann CU. Targeting the tumor microenvironment in chronic lymphocytic leukemia. Haematologica. 2021;106(9):2312-2324.
8. Zeller G, Tap J, Voigt AY, et al. Potential of fecal microbiota for early-stage detection of colorectal cancer. Mol Syst Biol. 2014;10(11):766.
9. Feng Q, Liang S, Jia H, et al. Gut microbiome development along the colorectal adenoma–carcinoma sequence. Nat Commun. 2015;6(1):6528.
10. Valdes AM, Walter J, Segal E, Spector TD. Role of the gut microbiota in nutrition and health. BMJ. 2018;361:k2179.
11. Wilkins LJ, Monga M, Miller AW. Defining dysbiosis for a cluster of chronic diseases. Sci Rep. 2019;9(1):12918.
12. Cozen W, Yu G, Gail MH, et al. Fecal microbiota diversity in survivors of adolescent/young adult Hodgkin lymphoma: a study of twins. Br J Cancer. 2013;108(5):1163-1167.
13. Yoo JY, Groer M, Dutra SVO, Sarkar A, McSkimming DI. Gut Microbiota and immune system interactions. Microorganisms. 2020;8(10):1587.
14. Ilett EE, Jørgensen M, Noguera-Julian M, et al. Associations of the gut microbiome and clinical factors with acute GVHD in allogeneic HSCT recipients. Blood Adv. 2020;4(22):5797-5809.
15. Weir TL, Manter DK, Sheflin AM, Barnett BA, Heuberger AL, Ryan EP. Stool microbiome and metabolome differences between colorectal cancer patients and healthy adults. PloS One. 2013;8(8):e70803.