Longitudinal modeling of serum cytokine levels and gut microbial abundance links IL-17/IL-22 with Clostridia and insulin sensitivity in humans

Xin Zhou\textsuperscript{1,2,6}, Jethro S. Johnson\textsuperscript{1,6}, Daniel Spakowicz\textsuperscript{1,3} Wenyu Zhou\textsuperscript{4}, Yanjiao Zhou\textsuperscript{1,5}, Erica Sodergren\textsuperscript{1}, Michael Snyder\textsuperscript{4} and George M. Weinstock\textsuperscript{1}

\textsuperscript{1}The Jackson Laboratory for Genomic Medicine, Farmington, CT 06032 USA
\textsuperscript{2}Genetics and Genome Sciences, University of Connecticut Health Center, CT 06032 USA
\textsuperscript{3}Ohio State University Comprehensive Cancer Center, Columbus, OH 43210, USA
\textsuperscript{4}Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA
\textsuperscript{5}Department of Medicine, University of Connecticut Health Center, CT 06032 USA
\textsuperscript{6}These authors equally contributed to the paper

Running Title: Clostridia, IL-17/IL-22 and insulin sensitivity

Corresponding Author: Jethro Johnson, The Jackson Laboratory for Genomic Medicine, Farmington, CT, 06032, USA. Telephone: +1-860-873-2021 Email: jethro.johnson@jax.org

Word Count: 3149
Number of Tables: 0
Number of Figures: 4
ABSTRACT

Recent studies using mouse models suggest that interaction between the gut microbiome and IL-17/IL-22 producing cells plays a role in the development of metabolic diseases. We investigated this relationship in humans using data from the prediabetes arm of the Integrated Human Microbiome Project (iHMP). Specifically, we addressed the hypothesis that early in the onset of metabolic diseases there is a decline in serum levels of IL-17/IL-22, with concomitant changes in the gut microbiome. Clustering iHMP study participants on the basis of longitudinal IL-17/IL-22 profiles identified discrete groups. Individuals distinguished by low levels of IL-17/IL-22 were linked to established markers of metabolic disease, including insulin sensitivity. These individuals also displayed gut microbiome dysbiosis, characterized by decreased diversity, and IL-17/IL-22-related declines in the phylum Firmicutes, class Clostridia, and order Clostridiales. The newly released iHMP data therefore supports a link between the gut microbiome to IL-17/IL-22 and the onset of metabolic diseases. This raises the possibility for novel, microbiome-related therapeutic targets that may effectively alleviate metabolic diseases in humans as they do in animal models.
The human gut microbiome consists of trillions of microorganisms that are known to impact host physiology. Variation in composition of the gut microbiome has been linked to multiple aspects of the metabolic syndrome, such as obesity, (1; 2) hypertension, (3; 4) and insulin resistance (5), as well as to Type 2 diabetes (T2D). (6-10)

While much remains to be learned about the functional mechanisms underpinning this relationship, growing evidence in mouse models points to an important role for the microbially-mediated immune system (11; 12). Deficiency in TLR5 (13), RORγt (14), and IL-22 (15) is associated with a variety of metabolic disorders. Additionally, high-fat diet-induced obesity results in a loss of IL-17 producing cells in the small intestine lamina propria (SILP) (14). While, induction of Th17 cells (16), or treatment with gut-homing Th17 cells (17), low dose IL-17 (18), or IL-22 (13; 15) can ameliorate the obesity-associated metabolic phenotype. Collectively, these results suggest a protective role for IL-17/IL-22 producing cells during onset of diet-related metabolic disorders in mice.

In humans, low IL-22 has been associated with impaired fasting glucose and T2D (19). However, the relationship between IL-17/IL-22 production, the microbiome, and metabolic diseases remains controversial (20-22) and comparatively understudied. Large-scale integrated omics studies provide an ideal opportunity to investigate this relationship. Here we present an ancillary study of the recently released prediabetes arm of the iHMP (23), in which our goal was to investigate the relationship between IL-17/IL-22 profiles, the gut microbiome and aspects of the metabolic syndrome. Specifically, we aimed to test hypotheses in humans that have hitherto only been convincingly demonstrated in mice.
RESEARCH DESIGN AND METHODS

iHMP Overview and Data Curation

The iHMP project consists of 105 human participants identified as at risk of developing T2D, who were followed over a four-year period. During this time detailed multi-omic profiling was carried out at quarterly intervals, and more frequently during periods of stress or upper respiratory infection (23). Participants were recruited following Stanford University Institutional Review Board Protocol #23062. Stool samples and blood samples were collected from each subject after an overnight fast (24). Stool samples were processed according to the Human Microbiome Project standard protocol (#07-001. V12.0). Cytokine data were generated by 63-plex Luminex antibody conjugated bead capture assay (Affymetrix, Santa Clara, California). Raw cytokine data were normalized to eliminate batch effects (24). Full details of approaches used for processing sequence and cytokine data can be found in our companion paper introducing this cohort (24).

16S rRNA sequence data and normalized cytokine data in Median Fluorescence Intensity (MFI), were intersected to retain sampling events for which both microbiome data and cytokine data were present. Based on this criterion, there were 92 participants (753 individual samples) whose measurements were retained for further analysis. Age of the participants ranged from 25 to 75 years old, BMI ranged from 19.10 to 40.83 kg/m², steady-state plasma glucose ranged from 40 to 276 mg/dL. According to the manufacturer’s protocol, CHEX1 to CHEX4 are different types of background control for Luminex MFI data. Any data points with substantial background noise (larger than 5 standard derivations plus-minus mean value (mean ± 5 * SD)) were removed. Among the 92 participants, 35 had fewer than five time points, these participants were removed before performing mixture modelling.
Statistical Analysis

A two-sided student’s t-test was used for significance testing when data were normally distributed, otherwise a two-sided Wilcoxon signed-rank test or Mann-Whitney U test was used. A chi-square test was used to determine if the proportion of insulin resistant individuals was different between High-Activity (HA) and Low-Activity (LA) groups. Linear discriminant analysis based on effect size (LEfSe) (25) was performed to determine if the microbial taxon abundances differed between HA and LA groups. All statistical tests were performed using R (version 3.5.0).

Data Modelling: Mixture model of individuals based on the IL-17/IL-22

Participants sampled more than five times were included in a General Mixture Model (GMM) built using the R package mclust (26). The longitudinal IL-17A, IL-17F and IL-22 MFI data were summarized as mean value and standard deviation (SD) for each individual, and then scaled in R. To determine the optimal number of Gaussian distributed clusters, models with 1~9 clusters were evaluated using the Bayesian Information Criterion (BIC), resulting in three clusters selected for further analyses (Figure S1). 25 individuals were assigned to cluster 1 (Low Activity group), 32 to cluster 2 (Indeterminate Activity group) and 11 to cluster 3 (Hight Activity group), the mixing probability of each cluster was 0.3634941, 0.4749886, and 0.1615172, respectively. Participants assigned to each cluster were associated with a “confidence of assignment” probability (0%~100%); those with < 99% confidence (6 individuals in total) were removed from the subsequent analyses.
Data Modelling: Linear Mixed Model

Linear mixed-models were built using the R package lme4 (27). Models were built separately to test different hypothesis, as described below:

To test if metabolic profiles were different among three groups, the equation we used was:

\[
\text{Target of Interest} \sim \text{Group} + \text{Days} + \text{Gender} + \text{BMI} + \text{Adj.age} + (1|\text{Subject_ID})
\]

Fixed effects included Group as the categorical variable derived from the GMM model, Gender was a binary categorical variable, Body Mass Index (BMI) and adjusted age (Adj.age) were continuous variables. Days was a numerical measurement of how many days after the overall study start date the sample was collected. Adjusted age described the average age of an individual during the study period. Random effects included a random intercept for each participant (1|Subject_ID).

For red cell distribution width (RDW), to adjust the previous known effect of Mean Corpuscular Hemoglobin (MCH) and Mean corpuscular volume (MCV) on the readout of RDW (28), the model was built as

\[
\text{RDW} \sim \text{Group} + \text{Days} + \text{MCH} + \text{MCV} + \text{Gender} + \text{BMI} + \text{Adj.age} + (1|\text{Subject_ID})
\]

For analysis related to microbiome alpha diversity and Firmicutes/Bacteroidetes Ratio, we simply would like to understand the fixed effect in this mixed model, so the model is built as

\[
\text{Microbiome Diversity (or F/B ratio)} \sim \text{Group} + \text{Gender} + \text{BMI} + \text{Adj.age} + (1|\text{Subject_ID})
\]
Bayesian Mixed-Effects Model for Taxa and Cytokine interactions conditional on cluster assignment

A Bayesian negative binomial longitudinal mixed effects model was used to evaluate the relationship between individual microbes and cytokines. To account for the zero-inflated nature of microbiome abundance, we used a Bayesian framework on a sparse matrix with a negative binomial distribution (29). Cytokine-related group (Hight Activity vs. Low Activity) and cytokine (either IL-17A or IL-17F) were scored as the interaction term and fixed effect, respectively, to test the combined effect of cytokine-related group and cytokine on microbe abundance. IL-22 was excluded from this analysis because a large proportion of IL-22 measurements appeared to be lower than the accurate detection threshold of the Luminex assay (Supplementary Fig. 2). We included the cytokine-related group:cytokine interaction term with the aim of testing the hypothesis that significant microbe-cytokine associations may be detected in HA, but not LA individuals, or vice versa.

Each microbe was modeled as the response variable with a random intercept for each participant, and with fixed effects for time, and an interaction term for the cluster identity (defined by the GMM, described above) and the cytokine of interest, thereby evaluating whether the relationship between a microbe and cytokine pair differed depending on the identity of the cluster. This followed standard matrix notation:

\[ M_i \sim X_i \beta + Z_i b_i + \epsilon_i \]

Where \( M_i \) is a vector of microbe relative abundances for each participant \( i \), \( X_i \) is the matrix of fixed effects, \( Z_i \) is the random effects vector of 1’s denoting a random intercept, \( b_i \) is a scalar for each
participant, and \( \varepsilon_i \) is a zero-centered error term. The fixed-effects matrix \( X_i \) is comprised of the days post study start \( D_i \), and an interaction term for cluster \( (C_i) \) and cytokine \( (Y_i) \).

\[
X_i = \begin{bmatrix}
D_{i,1} & C_{i,1} = 3 & Y_{i,1} & Y_{i,1}(C_i = 3) \\
\vdots & \vdots & \vdots & \vdots \\
D_{i,n} & C_{i,n} = 3 & Y_{i,n} & Y_{i,n}(C_i = 3)
\end{bmatrix}
\]

Sampling was performed with four chains with 5000 iterations per sample, and a burn-in of 1000 iterations. Samples were drawn using a No-U-Turn Sampler implemented in the \textit{brms} package (30-32). Chain convergence was confirmed by visual inspection of iteration plots and posterior predictive distributions.

**RESULTS**

**Analysis of the iHMP Prediabetes Cohort**

Here we used the recently-released iHMP dataset as the basis for a detailed longitudinal analysis of the relationship between the gut microbiome and serum cytokine levels. While individuals within this cohort were identified as at risk for T2D, only two individuals showed signs of developing T2D during the course of this study. Therefore, for the purpose of this analysis each individual was assumed to represent a fixed state relating to insulin sensitivity and T2D progression.

**Individuals show discrete IL-17/IL-22 profiles associated with insulin sensitivity**

Using gaussian mixture modelling based on the mean cytokine level and longitudinal variance in each individual (\textit{Supplementary Fig. 1}), we observed that iHMP study participants could be optimally separated into 3 groups. Individuals at one extreme were characterized by consistently
low cytokine levels and variance while individuals at the other extreme were characterized by high levels and/or variance of at least one cytokine (Fig. 1a, Supplementary Fig. 2). Henceforth we refer to these three groups as Low Activity (LA), Indeterminate Activity (IA), and High Activity (HA), to represent differences in their temporally-integrated levels of serum IL-17 and IL-22 activity (Fig. 1a). Importantly, identifying the three groups in this manner required estimating intra-individual variation (Supplementary Fig. 3), which is not available from cross-sectional data, highlighting the advantage of a longitudinal design.

We next considered the possibility that discrete IL-17/IL-22 cytokine profiles may reflect different stages of metabolic disease progression. Previous studies in mice demonstrated that HFD-induced onset of metabolic disease is associated with loss of CD4+ IL-17 producing cells in the SILP (14), while studies in humans have shown a negative correlation between serum IL-22 levels and physiological indicators of T2D (19). We therefore hypothesized that individuals with a LA profile would show a more severe metabolic phenotype than individuals with a HA profile. Within the two-thirds (n=65) of study participants whose single steady state plasma glucose (SSPG) measurement was available, we observed that SSPG levels were significantly lower in HA individuals compared to LA individuals (two-sided Wilcoxon test, W=64.5, p=0.021, Fig. 1b). Accordingly, LA individuals were more frequently insulin resistant (Supplementary Fig. 4a); however, mean fasting plasma glucose (FPG) levels, age, and body mass index (BMI) did not vary significantly across groups (Supplementary Fig. 4b).

IL-17/IL-22 inactivity is associated with a more severe metabolic phenotype
Longitudinal modelling of clinical data collected across the study period also revealed multiple markers that showed significant differences between LA compared to HA individuals (Fig. 2, Supplementary Fig. 5). Participants classified as HA showed higher plasma high-density lipopolysaccharide (HDL) and lower plasma triglycerides (TGL). The red blood cell distribution width (RDW) was higher in HA individuals, consistent with previous findings of high RDW as being associated with high HDL and low TGL (33). Additionally, serum sodium and chloride levels were lower in the HA group, and high insulin has been associated with increased sodium retention in T2D (34). Finally, we note that established markers of T2D, including A1C and serum glucose, did not vary significantly between groups (Fig. 2, Supplementary Fig. 5). This may be expected, given that the majority of iHMP study participants were identified as at risk for T2D, but had not developed the disease (24).

IL-17/IL-22 activity is associated with variation in the composition of the gut microbiome

The close association between IL-17/IL-22 producing cells and gut microbiota (35-37) prompted us to next ask whether individuals distinguished by cytokine activity levels differed in the composition of their gut microbiome. As an individual’s gut microbiome remained relatively stable throughout the course of this study (24) (Supplementary Fig. 3), we began by comparing mean microbiome abundance for each participant between HA and LA groups. We found LA individuals had a significantly lower alpha diversity (two-sided Wilcoxon test, W=34.0, p=0.003, Fig. 3a; Supplementary Fig. 6) and a lower Firmicutes to Bacteroidetes (F:B) ratio (two-sided Wilcoxon test, W=54.0 p=0.044, Fig. 3b, Supplementary Fig. 7) compared to HA individuals. Linear discriminant analysis based on effect size (LefSE) (25) revealed differences between LA and HA groups at multiple taxonomic levels. Most notably, the classes Bacteroidia and Clostridia were
more abundant in the LA and HA groups, respectively (Fig. 3c), indicating that members of these
taxa were likely responsible for observed differences in the F:B ratio.

We subsequently performed longitudinal modelling to look at pairwise relationships between
individual bacterial genera and cytokine abundances within HA versus LA individuals. Previous
studies identified significant associations between cytokines and gut microbes, but only revealed
inter-individual variation (38; 39) due to their cross-sectional design. The repeated and
longitudinal measurements of the iHMP study allowed us to test the possibility that intra-individual
variation in cytokine/bacteria abundance may be used to identify additional host-microbe
associations of biological interest. To accommodate within-individual correlation, we used a
mixed-effects model with random effects by individual.

We first observed the abundance of *Alistipes* was positively correlated with changes in serum IL-
17F levels (Fig. 4). As models were designed to compare cytokine vs. microbe interactions in the
context of cytokine activity (encoded as HA vs. LA), this result indicates *Alistipes* was
significantly associated with IL-17F in LA individuals. In contrast, 7 bacterial taxa were significant
for the activity group:cytokine interaction term (Fig. 4), indicating their relative abundance was
significantly associated with IL-17F or IL-17A levels in HA individuals. Notably, 6 of these 7
significant relationships involved taxa belonging to the class *Clostridia*, and in all 8 associations,
the cytokine abundance was positively correlated with taxon relative abundance. In conclusion,
analyses of the taxonomic abundance of the gut microbiome between individuals (Fig. 3) and
within individuals (Table 1) both provide evidence that members of the class *Clostridia* are
positively associated with increased levels of IL-17 activity.
DISCUSSION

Using data from the iHMP prediabetes cohort, we present evidence that individuals at risk of developing T2D display distinct, longitudinal IL-17/IL-22 cytokine profiles, which can be associated with altered severity in a number of established markers for metabolic disorders. By subsequently providing evidence for a link between IL-17/IL-22 and the composition of the gut microbiome we validate previous findings in mouse models and thus provide further support for the hypothesis that microbe-immune system interactions are relevant to human metabolic homeostasis.

Sustained loss of IL-17/IL-22 activity in iHMP study participants was associated with increased insulin resistance, as well as variation in metabolic markers that included lower HDL, and increased triglycerides. A trend for lower A1C was not statistically significant, which may support the observation that only 2 of 105 individuals were classified as diabetic at any point during the course of this four-year study. Nonetheless, our observations are consistent with evidence that RORrt-/- and IL-22-/- mouse models show reduced insulin sensitivity on a chow diet. Furthermore, low dose administration of IL-17 (18) or IL-22 (15) suppresses the metabolic phenotype induced by a high-fat diet. Taken together, such murine studies suggest that circulating levels of IL-17/IL-22 can be protective. Our work suggests that similar mechanisms may also apply to humans.

One explanation for the protective effects of IL-17/IL-22 is that these cytokines are directly, or indirectly involved in regulating the composition of the gut microbiome. For example, by
regulating the production of antimicrobial peptides that limit the abundance of potentially pathogenic taxa (40-42). Another possibility is that IL-17/IL-22 may influence tight junction function, meaning their deficiency could result in a leaky gut. (41; 43-46) This could in turn contribute to translocation of gut bacteria to the blood, which has been associated with T2D (47).

Alternatively, IL-22 may directly influence beta cell function, (48) in which case changes in the gut microbiome may be correlative, rather than directly contributing to the phenotype reported here.

While IL-17/IL-22 may influence the gut microbiome, the composition of the gut microbiome may reciprocally affect metabolic diseases via the ability of certain taxa to directly, or indirectly influence IL-17/IL-22 production. In a cross-sectional study of the influence of the human gut microbiome on cytokine production Schirmer et al. (49) previously demonstrated that IL-17 production from PBMCs exposed to *Staphylococcus aureus* was positively correlated with the relative abundance of *Clostridium* in the host gut microbiome. This is consistent with our observations that i) IL-17/IL-22 HA individuals had greater mean relative abundance of *Clostridium* in their gut microbiome across the course of this study, and ii) that in longitudinal analysis, three members of the class *Clostridia* (*Clostridium IV, Clostridium XI, Clostridium XIVa*) were positively correlated with IL-17F production and one (*Clostridium XIVb*) with IL-17A production. Human isolates from three of these *Clostridium* clusters (*XIVa*, *XIVb*, *IV*) were previously found to induce Th17 cells (a major producer of IL-17 and IL-22 (50; 51)) in mice. (52) Notably, germ-free mice do not carry Th17 cells in the SILP, (53) but inoculation with certain bacteria, including segmented filamentous bacteria (SFB), can induce Th17 cell development (54; 55). While SFB remain an ambiguous clade, they may be related to the family *Clostridiales*. (56)
In conclusion, our study supports growing evidence that members of the class *Clostridia* can induce development of IL-17/IL-22 producing cells. This in turn represents a plausible way in which the gut microbiome could influence human cytokine profiles and thereby influence metabolic diseases.

Establishing the causative relationship between IL-17/IL-22, the gut microbiome, and metabolic diseases will benefit from a deeper understanding of the mechanisms that underpin this interaction. One possibility is that gut microbiome-derived Aryl hydrocarbon receptor (AhR) signaling is critical for maintenance of Th17 cells (57; 58). Bacterial-produced AhR ligand is reduced after mice are switched to a high fat diet (59). Alternatively, Qin and colleagues (6) established a correlational association between the loss of members of the Clostridia in T2D patients and reduced short-chain fatty acid (SCFA) production. A later study confirmed this observation, and demonstrated that supplementing SCFA-producing bacteria strains to T2D patients can improve their clinical outcomes (60), possibly via the ability of SCFA to influence Th17 production (61).

A close link between IL-17/IL-22 and members of the *Clostridia* supports previous assertions that manipulating the relationship between cytokines and the gut microbiome presents novel therapeutic opportunities. In humans transferring the gut microbiome from lean donors to patients with metabolic syndrome has been shown to increase the insulin sensitivity of the recipients. Notably, among 16 bacteria strains that increased in gut microbiome of recipients post procedure, 12 strains belonged to the class *Clostridia* (62). In mice, metabolic disorders accompanied by a lack of Th17 cells and IL-22 could be rescued, either by induction or adaptive transfer of Th17 cells (16; 17), low dose IL-17 (18), and IL-22 (15), or by supplementing the gut microbiome with
symbiotics that increase Th17 cell abundance (14). Our demonstration of a close relationship between IL-17/IL-22 and members of the Clostridia therefore provides valuable insight into the biological processes that underpin the efficacy of these approaches.

In conclusion, our analysis of the newly released iHMP data set suggests novel avenues of research and raises the possibility that therapeutic targets related to IL-17/IL-22 that may effectively alleviate metabolic diseases in humans as they do in animal models.

ACKNOWLEDGEMENTS

Funding: This work was supported by the NIH Common Fund Human Microbiome Project (HMP) (1U54DE02378901). We thank Monika Avina and The Human Immune Monitoring Center (HIMC) for performing cytokine assays. Author Contributions: GMW and MS designed study. WZ, YZ, ES, GMW, and MS oversaw data collection, curation and storage. XZ, JSJ, and DS analyzed data. XZ, JSJ, DS, and GMW wrote manuscript. JSJ is the guarantor of this work and takes responsibility for the integrity of the data and accuracy of the data analysis. Data Availability: Raw data included in this study can be downloaded from the HMP data depository website https://www.hmpdacc.org/ihmp/.
REFERENCES

1. Yan Q, Gu Y, Li X, Yang W, Jia L, Chen C, Han X, Huang Y, Zhao L, Li P, Fang Z, Zhou J, Guan X, Ding Y, Wang S, Khan M, Xin Y, Li S, Ma Y: Alterations of the Gut Microbiome in Hypertension. Front Cell Infect Microbiol 2017;7:381

2. Al Khodor S, Reichert B, Shatat IF: The Microbiome and Blood Pressure: Can Microbes Regulate Our Blood Pressure? Front Pediatr 2017;5:138

3. Parekh PJ, Balart LA, Johnson DA: The Influence of the Gut Microbiome on Obesity, Metabolic Syndrome and Gastrointestinal Disease. Clin Transl Gastroenterol 2015;6:e91

4. Jiao N, Baker SS, Nugent CA, Tsompana M, Cai L, Wang Y, Buck MJ, Genco RJ, Baker RD, Zhu R, Zhu L: Gut microbiome may contribute to insulin resistance and systemic inflammation in obese rodents: a meta-analysis. Physiol Genomics 2018;50:244-254

5. Caricilli AM, Saad MJ: The role of gut microbiota on insulin resistance. Nutrients 2013;5:829-851

6. Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, Liang S, Zhang W, Guan Y, Shen D, Peng Y, Zhang D, Jie Z, Wu W, Qin Y, Xue W, Li J, Han L, Lu D, Wu P, Dai Y, Sun X, Li Z, Tang A, Zhong S, Li X, Chen W, Xu R, Wang M, Feng Q, Gong M, Yu J, Zhang Y, Zhang M, Hansen T, Sanchez G, Raes J, Falony G, Okuda S, Almeida M, LeChatelier E, Renault P, Pons N, Batto JM, Zhang Z, Chen H, Yang R, Zheng L, Li S, Yang H, Wang J, Ehrlich SD, Nielsen R, Pedersen O, Kristiansen K, Wang J: A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature 2012;490:55-60

7. Larsen N, Vogensen FK, van den Berg FW, Nielsen DS, Andreasen AS, Pedersen BK, Al-Soud WA, Sorensen SJ, Hansen LH, Jakobsen M: Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. PloS one 2010;5:e9085

8. Karlsson FH, Tremaroli V, Nookaew I, Bergstrom G, Behre CJ, Fagerberg B, Nielsen J, Backhed F: Gut microbiota in European women with normal, impaired and diabetic glucose control. Nature 2013;498:99-103

9. Forslund K, Hildebrand F, Nielsen T, Falony G, Le Chatelier E, Sunagawa S, Prifti E, Vieira-Silva S, Gudmundsdottir V, Pedersen HK, Arumugam M, Kristiansen K, Voigt AY, Vestergaard H, Hergoc R, Costea PI, Kultima JR, Levenez F, Dore J, Meta HITc, Nielsen HB, Brunak S, Raes J, Hansen T, Wang J, Ehrlich SD, Bork P, Pedersen O: Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. Nature 2015;528:262-266

10. Wu H, Esteve E, Tremaroli V, Khan MT, Caesar R, Manneras-Holm L, Stahlman M, Olsson LM, Serino M, Planas-Felix M, Xifra G, Mercader JM, Torrents D, Burcelin R, Ricart W, Perkins R, Fernandez-Real JM, Backhed F: Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug. Nat Med 2017;23:850-858

11. Grossmann V, Schmitt VH, Zeller T, Panova-Noeva M, Schulz A, Laubert-Reh D, Juenger C, Schnabel RB, Abt TG, Laskowski R, Wiltink J, Schulz E, Blankenberg S, Lackner KJ, Munzel T, Wild PS: Profile of the Immune and Inflammatory Response in Individuals With Prediabetes and Type 2 Diabetes. Diabetes Care 2015;38:1356-1364

12. Shu CJ, Benoist C, Mathis D: The immune system's involvement in obesity-driven type 2 diabetes. Semin Immunol 2012;24:436-442

13. Vijay-Kumar M, Aiiken JD, Carvalho FA, Cullender TC, Mwangi S, Srinivasan S, Sitaraman SV, Knight R, Ley RE, Gewirtz AT: Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. Science 2010;328:228-231

14. Garidou L, Pomie C, Kloppe P, Waget A, Charpentier J, Aloulou M, Giry A, Serino M, Stenman L, Lahtinen S, Dray C, Iacovoni JS, Courteminy M, Collet X, Amar J, Servant F, Lelouvier B, Valet P, EberL G, Fazilleau N, Douin-Echinard V, Heymes C, Burcelin R: The Gut Microbiota Regulates Intestinal CD4 T Cells Expressing RORgammat and Controls Metabolic Disease. Cell Metab 2015;22:100-112

15. Wang X, Ota N, Manzanillo P, Kates L, Zavala-Solorio J, Eidsenschink C, Zhang J, Lesch J, Lee WP, Ross J, Diehl L, van Bruggen N, Kolumam G, Ouyang W: Interleukin-22 alleviates metabolic disorders and restores mucosal immunity in diabetes. Nature 2014;514:237-241

For Peer Review Only
16. Martins LMS, Perez MM, Pereira CA, Costa FRC, Dias MS, Tostes RC, Ramos SG, de Zoete MR, Ryffel B, Silva JS, Carlos D: Interleukin-23 promotes intestinal T helper type17 immunity and ameliorates obesity-associated metabolic syndrome in a murine high-fat diet model. Immunology 2018;

17. Hong CP, Park A, Yang BG, Yun CH, Kwak MJ, Lee GW, Kim JH, Jang MS, Lee EJ, Jeun EJ, You G, Kim KS, Choi Y, Park JH, Hwang D, Im SH, Kim JF, Kim YK, Seoh JY, Surh CD, Kim YM, Jang MH: Gut-Specific Delivery of T-Helper 17 Cells Reduces Obesity and Insulin Resistance in Mice. Gastroenterology 2017;152:1998-2010

18. Mohamed R, Jayakumar C, Chen F, Fulton D, Stepp D, Gansevoort RT, Ramesh G: Low-Dose IL-17 Therapy Prevents and Reverses Diabetic Nephropathy, Metabolic Syndrome, and Associated Organ Fibrosis. J Am Soc Nephrol 2016;27:745-765

19. Shen J, Fang Y, Zhu H, Ge W: Plasma interleukin-22 levels are associated with prediabetes and type 2 diabetes in the Han Chinese population. J Diabetes Investig 2018;9:33-38

20. Abdel-Moneim A, Bakery HH, Allam G: The potential pathogenic role of IL-17/Th17 cells in both type 1 and type 2 diabetes mellitus. Biomed Pharmacother 2018;101:287-292

21. Galvan DL, Danesh FR: Paradoxical Role of IL-17 in Progression of Diabetic Nephropathy. J Am Soc Nephrol 2016;27:657-658

22. Yousefidaredor H, Zare-Bidaki M, Hakimi H, Assar S, Bagheri V, Arababadi MK: IL-17A plays an important role in induction of type 2 diabetes and its complications. Asian Pacific Journal of Tropical Disease 2014;4:412-415

23. Integrative HMP RNC: The Integrative Human Microbiome Project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease. Cell Host Microbe 2014;16:276-289

24. Wenyu Zhou MRS, Kévin Contrepois, Yanjiao Zhou, Sara Ahadi, Shana R. Leopold, Martin J. Zhang, Varsha Rao, Monika Avina, Tejaswini Mishra, Jethro Johnson, Brittany Lee, Songjie Chen, Dong-Binh Tran, Hoan Nguyen, Xin Zhou, Brandon Albright, Bo-Young Hong, Lauren Petersen, Eddy Bautista, Blake Hanson, Lei Chen, Daniel Spakowicz, Denis Salins, Benjamin Leopold, Melanie Ashland, Orit Shannon Rego, Pats Limcaoco, Elizabeth Colbert, Candice Allister, Dalia Perelman, Colleen Craig, Eric Wei, Hassan Chaib, Daniel Hornburg, Jessilyn Dunn, Liang Liang, Sophia Miryam Schüssler-Fiorenza Rose, Brian Piening, Hannes Rost, David Tse, Tracey McLaughlin, Erica Sodergren, George M. Weinstock, Michael Snyder: Complex host-microbial dynamics in prediabetes revealed through longitudinal multi-omics profiling. Nature 2019;569:663-671

25. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garret WS, Huttenhower C: Metagenomic biomarker discovery and explanation. Genome biology 2011;12:R60

26. Scrucca L, Fop M, Murphy TB, Raftery AE: mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. R J 2016;8:289-317

27. Bates D, Mächler M, Bolker B, Walker S: Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:14065823 2014;

28. Pilling LC, Atkins JL, Kuchel GA, Ferrucci L, Melzer D: Red cell distribution width and common disease onsets in 240,477 healthy volunteers followed for up to 9 years. PLoS One 2018;13:e0203504

29. Anders S, Huber W: Differential expression analysis for sequence count data. Genome biology 2010;11:R106

30. Bürkner P-C: Advanced Bayesian Multilevel Modeling with the R Package brms. arXiv preprint arXiv:170511123 2017;

31. Bürkner P-C: brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software 2017;80:1-28

32. Hoffman MD, Gelman A: The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research 2014;15:1593-1623

33. Pilling LC, Atkins JL, Duff MO, Beaumont RN, Jones SE, Tyrrell J, Kuo CL, Ruth KS, Tuie MA, Yaghootkar H, Wood AR, Murray A, Weedon MN, Harries LW, Kuchel GA, Ferrucci L, Frayling TM, Melzer D: Red blood cell distribution width: Genetic evidence for aging pathways in 116,666 volunteers. PloS one 2017;12:e0185083
34. Brands MW, Manhiani MM: Sodium-retaining effect of insulin in diabetes. Am J Physiol Regul Integr
Comp Physiol 2012;303:R1101-1109

35. Cavallari JF, Denou E, Foley KP, Khan WI, Schertzer JD: Different Th17 immunity in gut, liver, and
adipose tissues during obesity: the role of diet, genetics, and microbes. Gut Microbes 2016;7:82-89

36. Chewning JH, Weaver CT: Development and Survival of Th17 Cells within the Intestines: The
Influence of Microbiome- and Diet-Derived Signals. The Journal of Immunology 2014;193:4769-4777

37. Jin W, Dong C: IL-17 cytokines in immunity and inflammation. Emerg Microbes Infect 2013;2:e60

38. Li Y, Oosting M, Deelen P, Ricaño-Ponce I, Smeekens S, Jaeger M, Matzaraki V, Swertz MA, Xavier
RJ, Franke L, Wijmenga C, Joosten LAB, Kumar V, Netea MG: Inter-individual variability and genetic
influences on cytokine responses to bacteria and fungi. Nature Medicine 2016;22:952

39. Schirmer M, Smeekens SP, Vlamakis H, Jaeger M, Oosting M, Franzosa EA, ter Horst R, Jansen T,
Jacobs L, Bonder MJ, Kurilshikov A, Fu J, Joosten LAB, Zhernakova A, Huttenhower C, Wijmenga C,
Netea MG, Xavier RJ: Linking the Human Gut Microbiome to Inflammatory Cytokine Production
Capacity. Cell 2016;167:1125-1136.e1128

40. Abusleme L, Moutsopoulos NM: IL-17: overview and role in oral immunity and microbiome. Oral Dis
2017;23:854-865

41. Zenewicz LA, Yin X, Wang G, Elinav E, Hao L, Zhao L, Flavell RA: IL-22 deficiency alters colonic
microbiota to be transmissible and colitogenic. J Immunol 2013;190:5306-5312

42. Dixon BR, Radin JN, Piazuelo MB, Contreras DC, Algood HM: IL-17a and IL-22 Induce Expression
of Antimicrobials in Gastrointestinal Epithelial Cells and May Contribute to Epithelial Cell Defense against
Helicobacter pylori. PLoS One 2016;11:e0148514

43. Dudakov JA, Hanash AM, van den Brink MR: Interleukin-22: immunobiology and pathology. Annu
Rev Immunol 2015;33:747-785

44. Maxwell JR, Zhang Y, Brown WA, Smith CL, Byrne FR, Fiorino M, Stevens E, Bigler J, Davis JA,
Rottman JB, Budelsky AL, Symons A, Towne JE: Differential Roles for Interleukin-23 and Interleukin-17
in Intestinal Immune regulation. Immunity 2015;43:739-750

45. Lee JS, Tato CM, Joyce-Shaikh B, Gulen MF, Cayatte C, Chen Y, Blumenschein WM, Judo M,
Ayanoglu G, McClanahan TK, Li X, Cua DJ: Interleukin-23-Independent IL-17 Production Regulates
Intestinal Epithelial Permeability. Immunity 2015;43:727-738

46. O'Connor W, Jr., Kamanaka M, Booth CJ, Town T, Nakae S, Iwakura Y, Kolls JK, Flavell RA: A
protective function for interleukin 17A in T cell-mediated intestinal inflammation. Nat Immunol
2009;10:603-609

47. Sato J, Kanazawa A, Ikeda F, Yoshihara T, Goto H, Abe H, Komiya K, Kawaguchi M, Shimizu T,
Ogihara T, Tamura Y, Sakurai Y, Yamamoto R, Mita T, Fujiyama Y, Fukuda H, Nomoto K, Takahashi T,
Asahara T, Hirose Y, Nagata S, Yamashiro Y, Watada H: Gut dysbiosis and detection of "live gut bacteria"
in blood of Japanese patients with type 2 diabetes. Diabetes Care 2014;37:2343-2350

48. Hasnain SZ, Borg DJ, Harcourt BE, Tong H, Sheng YH, Ng CP, Das I, Wang R, Chen AC, Loudovaris
T, Kay TW, Thomas HE, Whitehead JP, Forbes JM, Prins JB, McGuickin MA: Glycemic control in diabetes
is restored by therapeutic manipulation of cytokines that regulate beta cell stress. Nat Med 2014;20:1417-
1426

49. Schirmer M, Smeekens SP, Vlamakis H, Jaeger M, Oosting M, Franzosa EA, Ter Horst R, Jansen T,
Jacobs L, Bonder MJ, Kurilshikov A, Fu J, Joosten LAB, Zhernakova A, Huttenhower C, Wijmenga C,
Netea MG, Xavier RJ: Linking the Human Gut Microbiome to Inflammatory Cytokine Production
Capacity. Cell 2016;167:1125-1136.e1128
53. Ivanov, II, Frutos Rde L, Manel N, Yoshinaga K, Rifkin DB, Sartor RB, Finlay BB, Littman DR: Specific microbiota direct the differentiation of IL-17-producing T-helper cells in the mucosa of the small intestine. Cell Host Microbe 2008;4:337-349
54. Ivanov, II, Atarashi K, Manel N, Brodie EL, Shima T, Karaoz U, Wei D, Goldfarb KC, Santee CA, Lynch SV, Tanoue T, Imaoka A, Itoh K, Takeda K, Umesaki Y, Honda K, Littman DR: Induction of intestinal Th17 cells by segmented filamentous bacteria. Cell 2009;139:485-498
55. Tan TG, Sefik E, Geva-Zatorsky N, Kua L, Naskar D, Teng F, Pasman L, Ortiz-Lopez A, Jupp R, Wu HJ, Kasper DL, Benoist C, Mathis D: Identifying species of symbiont bacteria from the human gut that, alone, can induce intestinal Th17 cells in mice. Proc Natl Acad Sci U S A 2016;113:E8141-E8150
56. Yin Y, Wang Y, Zhu L, Liu W, Liao N, Jiang M, Zhu B, Yu HD, Xiang C, Wang X: Comparative analysis of the distribution of segmented filamentous bacteria in humans, mice and chickens. The ISME journal 2013;7:615-621
57. Chewning JH, Weaver CT: Development and survival of Th17 cells within the intestines: the influence of microbiome- and diet-derived signals. Journal of immunology 2014;193:4769-4777
58. Xue J, Nguyen DT, Habtezion A: Aryl hydrocarbon receptor regulates pancreatic IL-22 production and protects mice from acute pancreatitis. Gastroenterology 2012;143:1670-1680
59. Natividad JM, Agus A, Planchais J, Lamas B, Jarry AC, Martin R, Michel ML, Chong-Nguyen C, Roussel R, Straube M, Jegou S, McQuitty C, Le Gall M, da Costa G, Lecornet E, Michaudel C, Modoux M, Glodt J, Brionneau C, Sovran B, Dupraz L, Bado A, Richard ML, Langella P, Hansel B, Launay JM, Xavier RJ, Duboc H, Sokol H: Impaired Aryl Hydrocarbon Receptor Ligand Production by the Gut Microbiota Is a Key Factor in Metabolic Syndrome. Cell Metab 2018;28:737-749 e734
60. Zhao L, Zhang F, Ding X, Wu G, Lam YY, Wang X, Fu H, Xue X, Lu C, Ma J, Yu L, Xu C, Ren Z, Xu Y, Xu S, Shen H, Zhu X, Shi Y, Shen Q, Dong W, Liu R, Ling Y, Zeng Y, Wang X, Zhang Q, Wang J, Wang L, Wu Y, Zeng B, Wei H, Zhang M, Peng Y, Zhang C: Gut bacteria selectively promoted by dietary fibers alleviate type 2 diabetes. Science 2018;359:1151-1156
61. Asarat M, Apostolopoulos V, Vasiljevic T, Donkor O: Short-Chain Fatty Acids Regulate Cytokines and Th17/Treg Cells in Human Peripheral Blood Mononuclear Cells in vitro. Immunol Invest 2016;45:205-222
62. Vrieze A, Van Noord E, Holleman F, Salojarvi J, Kootse RS, Bartelsman JF, Dallinga-Thie GM, Ackermans MT, Serlie MJ, Oozer R, Derrien M, Druesne A, Van Hylckama Vlieg JE, Bloks VW, Groen AK, Heilig HG, Zoetendal EG, Stroes ES, de Vos WM, Hoekstra JB, Nieuwdorp M: Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. Gastroenterology 2012;143:913-916 e917
FIGURE LEGENDS

Figure 1. Participants grouped according to IL-17/IL-22 cytokines
A) Gaussian Mixture Modelling of cytokine mean abundance and variance separates study participants into three discrete groups (columns). Lines within each panel represent repeated measurements of serum cytokine abundance for one individual over the study period. Rows represent serum cytokines associated with cytokine activity (IL-17A, IL-17F, IL-22). CHEX4 is a measurement of background fluorescent intensity and can be treated as a negative control. (Note: different scales on Y axis for each row). B) Steady State Plasma Glucose (SSPG, mg/dL) measurement by group. P-values for pairwise Wilcoxon test are labelled above the bar plot, the p-value for a one-way ANOVA test is labelled under the bar plot.

LA: Low-Activity group, IA: Indeterminate-Activity group, HA: High-Activity group

Figure 2. Linear Mixed Model estimates on fixed effects introduced by Low-Activity (LA) and High-Activity (HA) group
Results for full LMMs are shown in Supplementary Fig. 4. The comparisons of Active vs Inactive group are presented here. Dashed lines represent the LA group, while regression estimates for the HA group are displayed as horizontal lines. The center of each horizontal line is the beta coefficient of regression, while thick lines represent 50% Credible Intervals (CI) or ± 1 SD, and thin lines represent 95% CI or ± 2SD.

A1C: Hemoglobin A1c
HDL: High density lipoprotein
RDW: Red blood cell distribution width
NaCl: sodium chloride
Figure 3. Differences in the gut microbiome of IL-17/IL-22 Low-Activity (LA) and High-Activity (HA) individuals

A) Shannon diversity estimates for the HA and LA. Mean value of diversity for each participant across the study period is used to generate this plot. The p-value from a Wilcoxon test is labelled above the plot. B) Firmicutes to Bacteroidetes ratio of HA and LA. Mean value of $F/B$ ratio for each participant across the study period is used to generate this plot. The p-value from a Wilcoxon test is labelled above the plot. C) Cladogram representing the LEfSe results for comparing taxa abundance between HA and LA groups. Circles on the cladogram represent the phylogenetic relationship of taxa that are tested, with phylum at the center and OTU on the edges. Each point represents a taxonomic unit. Red color covering a dot/region indicates the taxa that are more abundant in the HA group, blue color covering a dot/area indicates the taxa are more abundant in the LA group.

Figure 4. Bacterial genera whose abundance correlates with serum IL-17

Significant correlations between serum IL-17 and bacterial genus abundance are shown for HA individuals (red panels) and LA individuals (blue panel). Distributions show estimated effect sizes from Bayesian Markov chain Monte Carlo draws after parameter convergence. Panels show bacteria for which the estimated effect is significantly greater or less than 0 (95% credible interval does not include 0).
FIGURE 1

A

Low-Activity Group  Indeterminate-Activity Group  High-Activity Group

Median Fluorescent Intensity

Date of Collection

B

Anova, p = 0.018

SSPG

LA  IA  HA
FIGURE 2

**Glucose Metabolism**

- A1C

**Lipid Metabolism**

- HDL
- Triglycerides

**NaCl Metabolism**

- Sodium
- Chloride

**Red Blood Cell Volume**

- RDW

Lower in HA Group  Higher in HA Group

For Peer Review Only
Diabetes

**FIGURE 3**

**A**

\[
p = 0.0015
\]

**B**

\[
p = 0.045
\]

**C**

Differently Abundant Taxa:
- a: Bacteroidales
- b: Granulicatella
- c: Carnobacteriaceae
- d: Oscillibacter
- e: unclassified_Ruminococcaceae
- f: Ruminococcaceae
- g: Clostridiales

- Higher Abundant in LA Group
- Higher Abundant in HA Group
SUPPLEMENTARY FIGURE 1

A) Modelling input variables and result

B) Model selection based on Bayesian Information Criterion (BIC)

Generalized Mixture Model of individuals based on levels of IL-17 and IL-222

A) Pairwise correlation of three cytokines mean value and SD. Color represents the group which the sample is assigned to (blue: LA group, red: IA group, green: HA group). B) The Bayesian information criterion (BIC) result for the potential number of groups (components). The highest three BIC indicate the most likely number of groups.
Density distributions of cytokines for each group and the lower limit of detection

Density distributions show the abundance of each cytokine within each group. Groups are marked with colors. Dashed lines represent the lower limits of detection (LLOD) for the Luminex assay used to calculate cytokine abundance. LLODs are calculated as 3 standard deviations above the mean absorbance for the negative control associated with each measurement. They can be interpreted as conservative thresholds below which the accuracy of measurements may become limited. (IL-17A: 64.4; IL-17F: 140.5, IL-22: 199.9)
Non-metric multidimensional scaling (nMDS) plots of study participants based on serum cytokine and gut microbiome compositions

nMDS plots depicting variation between repeated measurements of A) microbiome and B) cytokine profiles. Repeated measurements from different individuals are colored separately to highlight the relationship between intra- and inter-individual variation. For clarity, only data for individuals with >15 repeated measurements are shown.
SUPPLEMENTARY FIGURE 4

A) Percentage of insulin resistant individuals in each of the group. Total number of individuals are labelled on top of each column. B) Comparisons of age (years), Body Mass Index (BMI, kg/m2), and Fasting Plasma Glucose (FPG, mmol/L) across groups. Pairwise t test was performed (* 0.01<P<0.05; ns: p>0.05). Group 1: Low Activity (LA), Group 2: Indeterminate Activity (IA), Group 3: High Activity (HA)