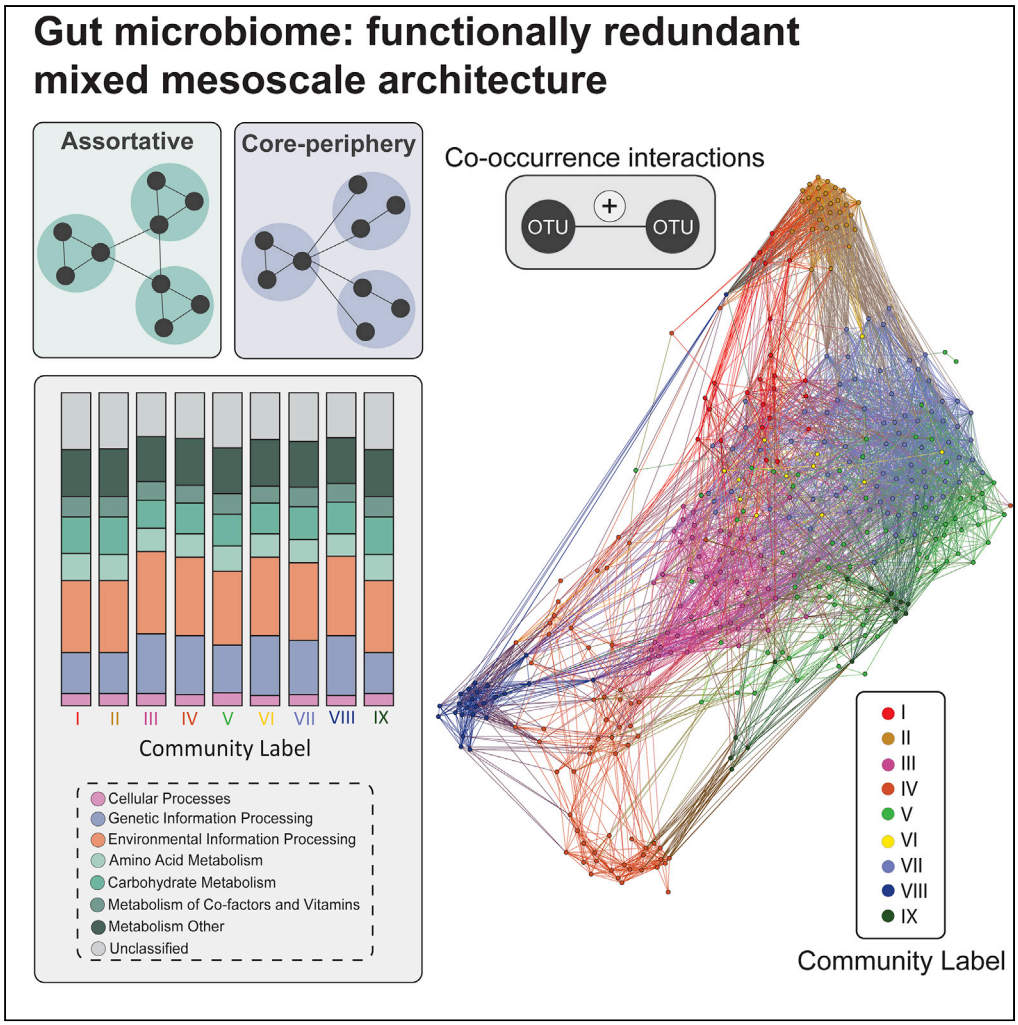
iScience

CellPress

Article

Co-existence of Network Architectures Supporting the Human Gut Microbiome



Caitlin V. Hall, Anton Lord, Richard Betzel, ..., Andrew Zalesky, Graham Radford-Smith, Luca Cocchi

caitlin.hall@qimrberghofer.edu.au

HIGHLIGHTS

The gut microbiome conforms to an assortative and core-periphery mesoscale network

Co-existing mesoscale structures support high functional redundancy

Each community contributes to a diverse number of overlapping metagenomic functions

Hall et al., iScience 22, 380–391
December 20, 2019
https://doi.org/10.1016/j.isci.2019.11.032

iScience

CellPress

Article

Co-existence of Network Architectures Supporting the Human Gut Microbiome

Caitlin V. Hall,[1,2,7,*] Anton Lord,[3] Richard Betzel,[4] Martha Zakrzewski,[5] Lisa A. Simms,[3] Andrew Zalesky,[6] Graham Radford-Smith,[3] and Luca Cocchi[1,2]

SUMMARY

Microbial organisms of the human gut microbiome do not exist in isolation but form complex and diverse interactions to maintain health and reduce risk of disease development. The organization of the gut microbiome is assumed to be a singular assortative network, where interactions between operational taxonomic units (OTUs) can readily be clustered into segregated and distinct communities. Here, we leverage recent methodological advances in network modeling to assess whether communities in the human microbiome exhibit a single network structure or whether co-existing mesoscale network architectures are present. We found evidence for core-periphery structures in the microbiome, supported by strong, assortative community interactions. This complex architecture, coupled with previously reported functional roles of OTUs, provides a nuanced understanding of how the microbiome simultaneously promotes high microbial diversity and maintains functional redundancy.

INTRODUCTION

The human intestinal (gut) microbiome is a complex biological system, whose functions and metabolic processes are the product of multiple interactions between microbial operational taxonomic units (OTUs) (Zelezniak et al., 2015). These diverse interactions can arise from direct or passive mechanisms and may result in beneficial (commensal or mutualistic), neutral, or detrimental (competitive or parasitic) effects to all OTUs involved (Faust and Raes, 2012). Perturbations to microbial interactions may manifest as microbial dysbiosis and have been implicated in a number of pathologies including inflammatory bowel disease (Halfvarson et al., 2017), metabolic dysregulation (Sanz et al., 2014), and neuropsychiatric disorders (Jiang et al., 2015). Understanding the structure, function, and composition of the human microbiome has therefore become an active area of research (The Human Microbiome Project Consortium et al., 2012). Interest in the microbiome has also coincided with the adoption of network science in biology, offering an armory of conceptual and analytical tools to model microbial interactions.

Using co-occurrence and co-exclusion relationships between individual OTUs, network-based approaches allow us to gain insight into the healthy and pathological properties of the microbiome, including organizational features that may contribute to system resilience or vulnerability (Baldassano and Bassett, 2016; Lozupone et al., 2012). Within the microbiome, OTUs are not expected to interact equally but to form smaller communities characterized by dense functional associations. An emerging approach to study the structure and function of the microbiome is therefore to define and characterize co-occurrence interactions at the mesoscale (community level). Mesoscale defines an intermediate level between that of individual OTUs and the whole microbiome. At the mesoscale, networks can exhibit different community structures, including assortative, disassortative, core-periphery, or mixed interactions (Betzel et al., 2018) (Figures 1A–1D). The dominant view emerging from existing work is that the human gut microbiome exhibits segregated and autonomous assortative communities, where OTUs sharing similar phylogenetic or functional properties have a tendency to preferentially cluster together. Evidence for a singular assortative structure in the microbiome has been observed in both empirical (Jackson et al., 2018) and computational modeling work.

Although previous work has provided important insights, the algorithms and heuristic approaches used to detect communities have certain design features such that they can detect only assortative communities. Modularity maximization, for example, is among the more popular community detection methods used in the field. However, this method is only capable of detecting groups of OTUs that are densely intra-connected and sparsely inter-connected: an assortative or modular community structure. Consequently, it is unclear whether the detected assortativity represents a methodological bias or is a reflection of the gut

[1]Clinical Brain Networks Group, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia

[2]School of Biomedical Sciences, University of Queensland, Brisbane, QLD 4006, Australia

[3]Gut Health Lab, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia

[4]Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405, USA

[5]Medical Genomics Group, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia

[6]Melbourne Neuropsychiatry Centre and Department of Biomedical Engineering, The University of Melbourne, Melbourne, VIC 3010, Australia

[7]Lead Contact

*Correspondence: caitlin.hall@qimrberghofer.edu.au

https://doi.org/10.1016/j.isci.2019.11.032

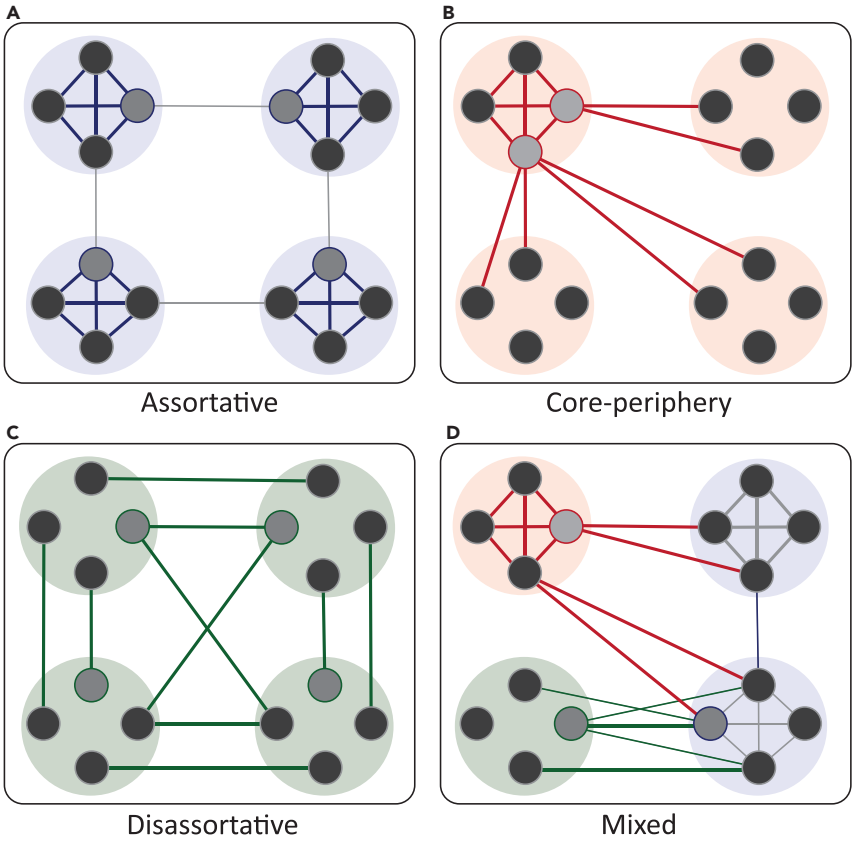


Figure 1. Representation of Possible Community Structures at the Mesoscale Level, Visualized as a Force-Directed Graph Layout

(A–C) Mesoscale network structures can be (A) assortative, where the internal density of interactions (within community) exceeds the external density of interactions (between communities); (B) core-periphery, where there is a central core (connected to the rest of the network) and periphery (with minimal interactions); or (C) disassortative, where the external density of interactions (between communities) exceeds the internal density of interactions (within community).

(D) Mesoscale structures can also occur simultaneously in the network, described as mixed or co-existing architectures (Betzel et al., 2018).

microbiome community structure. The current understanding of the microbiome architecture may therefore be too simplistic. Recent work on microbial interactions support the notion that the microbiome exhibits not only assortative, but also cores, peripheries, and disassortative communities. That is, the microbiome may exhibit densely connected "core" OTUs, whose metabolic and enzymatic processes exert a particularly beneficial role to the rest of the network, including the efficient transfer of nutrients, metabolites, or by-products (Ze et al., 2013). From an ecological perspective, a conserved core-periphery structure is consistent with the concept of a keystone guild. Keystone guilds—groupings of "core" or keystone OTUs—have been described as highly connected structures that exert a considerable influence on the structure and stability of the microbiome (Power et al., 1996; Banerjee et al., 2018). The ability to conform to multiple, co-existing network configurations may therefore reflect an ecological or evolutionary selective advantage to the microbiome. Specifically, this ability may be critical for the human microbiome, where environmental perturbations are frequently introduced, including dietary changes and/or antibiotic administration (Buffie et al., 2012).

Extending upon previous work, we studied the mesoscale architectures underpinning the human gut microbiome by applying the weighted stochastic block model (WSBM), a flexible generative algorithm for detecting community structure (Aicher et al., 2014). Unlike modularity maximization, where a network can only be partitioned into densely connected modular communities (Figure 1A), the WSBM also considers alternative patterns of connectivity that are less spatially compact (Figures 1B–1D). This provides the WSBM the flexibility to uncover multiple network structures beyond assortativity, including disassortative and

core-periphery interactions between communities. The approach has recently been validated to partition and assign community network structures in models representing interactions between remote brain regions (Betzel et al., 2018; Faskowitz et al., 2018). Using the WSBM framework, we sought to (1) establish whether the human microbiome exhibits a unique or heterogeneous mesoscale network architecture; (2) understand the patterns of microbial co-occurrence relationships within and between communities; and (c) identify hub OTUs that may play a key role in supporting the mesoscale network structure or function. Based on emerging research, we hypothesized that the microbiome will exhibit both assortative and non-assortative structures, including core-periphery participation. From a taxonomic perspective, we expected closely related OTUs to form strong, assortative co-occurrence interactions. Non-assortative communities, however, are expected to exhibit greater taxonomic microbial diversity. Information on individual OTUs that exhibit specific network features, including high between- or within-community interactions, may provide additional insights into the key organizational principles supporting the microbiome.

RESULTS

Applying the Weighted Stochastic Block Model to the Human Microbiome

We created a microbial interaction network by fitting the WSBM to a dataset of 58 healthy human intestinal microbiomes, representing 370 nodes (OTUs). We then replicated the analyses in a large (n = 528), independent, and publicly available dataset. Using a Bayesian model optimization method (details in Transparent Methods), the microbiome network was partitioned into $k = 12$ communities, based on where the mean marginal log likelihood (a measure of model fitness) begins to plateau (Figure 2A; Figure S1 for replication). Note that we also performed a control analysis at $k = 13$, which yielded a similar consensus partition and mesoscale structure. At $k = 12$, we repeated the WSBM for 65 fits, each returning an independent partition assigning each OTU to one of 12 communities. The consistency of the community assignment was assessed using normalized mutual information (NMI) from the output of each WSBM partition. To test that the consistency of 65 detected community partitions were above chance level, the results were benchmarked against partitions that were achieved by fitting the WSBM to null networks. Microbiome community partitions were highly consistent with each other (NMI, 0.70 ± 0.02), whereas null networks had low consistency between each WSBM run (NMI, 0.38 ± 0.02). The NMI values obtained from comparing microbiome partitions were significantly higher than those obtained from the null ($P < 0.001$) (Figure S2). To estimate the impact of inter-subject variability in community detection, we performed a cross-validation leave-one-out (LOO) analysis (Figure S3). Results confirm that our group-level co-occurrence network was representative across 58 samples, with low inter-individual variability when compared with the original group-level matrix (as assessed by one minus Mantel's test statistic, $\mu = 0.02 \pm 0.03$) (Figure S3).

We subsequently reduced 65 repetitions from the microbiome dataset to a single consensus partition using a progressive median alignment method (Transparent Methods). The alignment algorithm detects consistent assignment to communities and reduces the complexity of the model if OTU assignment to any given community is low. Across 65 WSBM fits, three communities had low consistency in nodal assignment (i.e., OTUs assigned to these communities had equal preference for assignment with at least one other community). These communities were subsequently removed from the model, and OTUs were assigned to another community that they were more consistently aligned to. The final number of communities was therefore 9, ranging in size from 12 to 80 OTUs (Figure 2B). A complete list of the final consensus partition is presented in Table S1. We visualized the overlap between computational (community assignment) and taxonomic decompositions of the OTUs in Figures 2B–2D. As the WSBM learns from both the presence and weight of edges, thus far we have included both positive and negative interactions. Thus, the mesoscale analysis presents a coarse-graining of the network's overall architecture, accounting for the polarity of microbial interactions. However, given the specificity of our subsequent network-based statistics, we have opted to restrict our analyses to positive correlations only (thresholded at 0.4). This is a common approach in microbiome network studies to reduce statistical noise (Poudel et al., 2016; Jackson et al., 2018), as accurately interpreting the ecological significance of weak positive, and negative correlations is difficult in compositional data (i.e., relative abundance).

Evidence for the Diversity of Mesoscale Architectures in the Microbiome

First, we sought to establish whether communities in the microbiome exhibit a single mesoscale structure or whether we find evidence of many different types. To answer this, we used community motif participation, an approach recently developed and validated to characterize mesoscale interactions between communities of brain regions (Betzel et al., 2018) (Transparent Methods). To estimate an OTU's motif

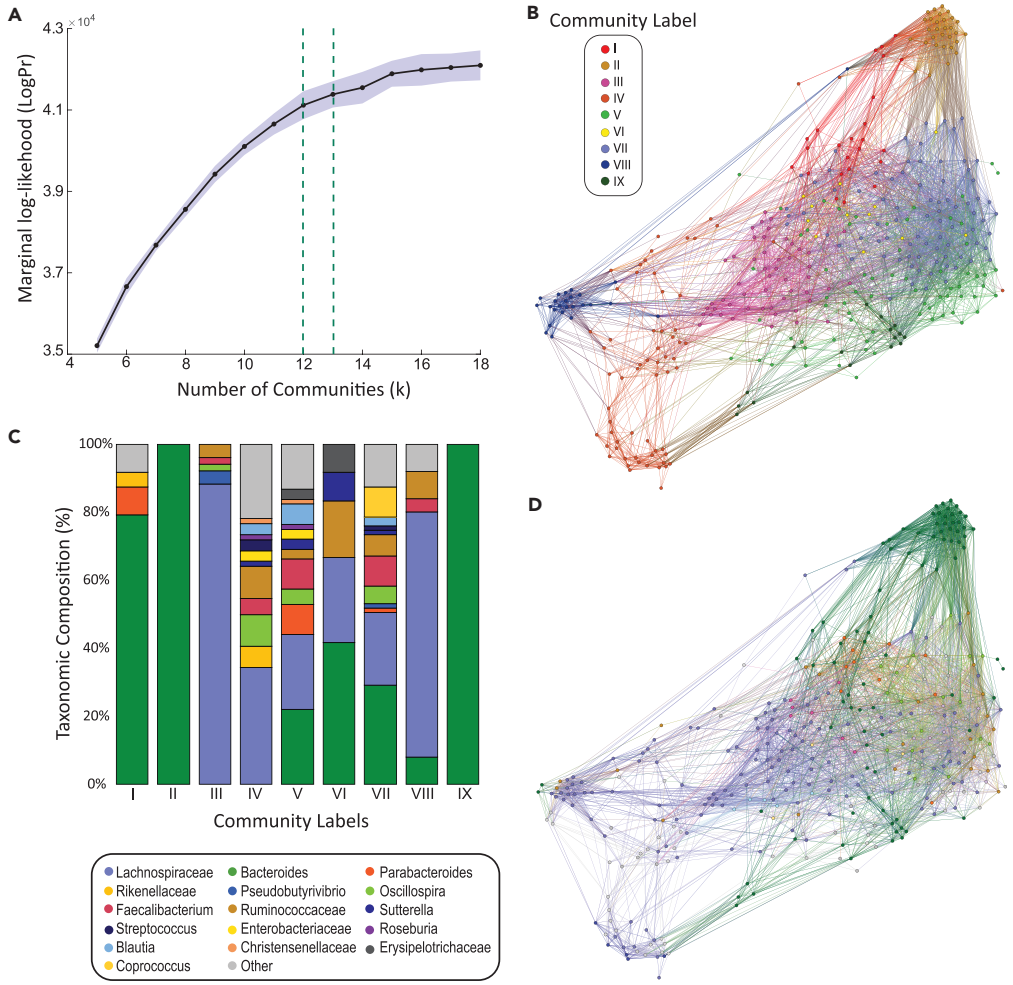


Figure 2. Bayesian Model Selection Using the Weighted Stochastic Block Model to Define the Microbial Interaction Network

(A) We fit the WSBM algorithm across a number of models between $k = 5$ to $k = 18$, repeated 65 times each. The consensus partition, $k = 12$ (green, dashed), was achieved when mean marginal log likelihood, *LogPr* (blue, solid) plateaus.

(B) Co-occurrence microbial interaction network, where each node represents an OTU and edges represent pairwise abundance relationships between them. The co-occurrence network has been reduced for visualization, representing 370 nodes and 4,854 edges at a Pearson correlation cutoff of r = 0.3. OTUs are colored according to community label.

(C) Composition (%) of each community based on taxonomic assignment of the OTUs at the family/genus level.

(D) The same co-occurrence network as seen in (B), but where OTUs have been colored according to taxonomic assignment at the family/genus level.

participation, we first defined a community motif for every pair of communities as the average connection weight within and between those communities. Based on these average connection weights, the communities comprising each motif could be uniquely classified as assortative, disassortative, core, or periphery. Next, we mapped community-level classifications back onto the individual OTUs that comprise each community. Finally, for a given OTU, we calculated motif participation as the proportion of times its community participated in a given motif.

Results showed that the healthy human microbiome exhibited a preference for assortativity, with embedded core-periphery mesoscale architectures. No disassortative interactions were detected across all WSBM fits. We tested this preference for assortativity by calculating a maximum assortativity score, providing an indication of how often each community formed exclusive assortative motifs with all communities (i.e., 100% assortative interactions). Over 65 WSBM fits, maximum assortativity therefore represents the proportion of times the minimum within-community interactions exceeded the maximum

Community Label	Number of Nodes	Core Participation (%)	Peripheral Participation (%)	Community Assortativity	Strength
I	24	0.04 ± 0.02	0.01 ± 0.02	0.08 ± 0.11	32.85 ± 5.30
II	30	0.11 ± 0.00	0.00 ± 0.00	0.56 ± 0.09[f]	43.99 ± 3.99
III	51	0.10 ± 0.04	0.01 ± 0.03	0.14 ± 0.07	37.10 ± 5.80
IV	64	0.02 ± 0.03	0.19 ± 0.16[c]	0.02 ± 0.10	17.50 ± 3.55
V	68	0.00 ± 0.01	0.27 ± 0.13[d]	−0.05 ± 0.05	23.86 ± 5.64
VI	12	0.02 ± 0.03	0.44 ± 0.10[e]	−0.01 ± 0.09	28.43 ± 4.54
VII	80	0.35 ± 0.17[a]	0.13 ± 0.14	0.03 ± 0.07	41.72 ± 7.15
VIII	25	0.31 ± 0.13[b]	0.08 ± 0.13	0.35 ± 0.17[g]	22.21 ± 4.32
IX	16	0.01 ± 0.01	0.03 ± 0.05	0.31 ± 0.10[h]	29.68 ± 3.44

Table 1. Community-Level Network Statistics

Reported as mean ± SD.

[a]p Value = 2.68×10^{-38}.

[b]p Value = 1.80×10^{-08}.

[c]p Value = 8.68×10^{-04}.

[d]p Value = 1.80×10^{-16}.

[e]p Value = 5.66×10^{-12}.

[f]p Value = 5.15×10^{-40}.

[g]p Value = 8.92×10^{-09}.

[h]p Value = 6.33×10^{-05}.

between-community interactions for any pair of communities (Figure 1). On average, each OTU's community exhibited maximum assortativity 36.94 ± 8.45% of the time. We then calculated the proportion of times an OTU's community participated in any other type of structural motif beyond assortativity (e.g., a minimum of one pair of community interactions) and observed that as a whole network, core and peripheral interactions occurred 12.56 ± 16.60% and 13.25 ± 15.87% of the time, respectively. Importantly, comparative observations of assortative, core, and periphery motif interactions were also observed in our replication dataset (Figure S1). Replication of the above-mentioned findings in a distinct dataset (larger sample size, different age range, and broad ethnic diversity) suggests that the co-existence of complex architectures may be a ubiquitous property to the microbiome. The large variances observed between participation in core and periphery structures suggests each OTU's community may exhibit dissimilar patterns of mesoscale proportions, which we investigate below.

Microbial Communities Exhibit Unique Mesoscale Signatures

The above findings corroborate previous analyses of microbiome mesoscale structure, identifying a mostly assortative organization (Baldassano and Bassett, 2016). Using the common community detection algorithm, modularity maximization, we further support the existence of a dominant assortative structure in this dataset (Figure S4). However, we also find evidence of non-assortative (primarily core-periphery) interactions. This suggests that a strictly assortative description may not fully characterize the diversity of mesoscale communities. We next sought to establish whether core-periphery interactions occurred uniformly between all nine communities (with each exhibiting a small degree of core-periphery interactions with at least one other community) or whether these observations are driven by a small subset of non-assortative interactions. We calculated the average number of times (over 65 WSBM fits) each community (Figure 2B) participated with other communities in a core or peripheral motif (Table 1). Results showed that most communities exhibited some degree of participation in non-assortative motifs (as shown by the participation in core and peripheral motifs) but were not uniformly distributed. To establish the significance of core and periphery mesoscale interactions, we used a permutation-based null model to create distributions representing the size and number of each community observed in the empirical dataset (Figure S5). When benchmarked against random community assignments, core motif participation was significantly higher in communities VII and VIII, whereas communities IV, V, and VI exhibited significantly higher participation

in peripheral mesoscale motifs (Table 1, Figure 2B) ($P < 0.05$). These communities are characterized by a range of different OTUs classified at broader (phylum) and finer (family/genus) taxonomic levels, including four dominant bacterial phyla: *Firmicutes, Bacteroidetes, Actinobacteria, and Proteobacteria* (Figure 2C).

To gain a more nuanced understanding of the mesoscale interactions between OTU communities, we additionally calculated nodal assortativity (normalized to the size of each community) (Faskowitz et al., 2018) and nodal strength (Rubinov and Sporns, 2010) (Transparent Methods) and mapped these onto community-level patterns (Table 1). In this context, positive assortativity coefficients (>0) should therefore be interpreted as exhibiting assortative interactions, whereas negative coefficients (<0) suggest this community may participate in disassortative interactions. Communities with assortativity coefficients of ~0 do not exhibit strong assortative interactions as detected using the WSBM. As the WSBM algorithm can detect simultaneous mesoscale structures, it is important to note that a high assortativity coefficient does not preclude the existence of core or periphery participation (Table 1). Communities II, VIII, and IX exhibited the strongest preference for community assortativity, significantly higher than what would be expected based on the size of these communities ($P < 0.05$). These communities are characterized by robust co-occurrence interactions between taxonomically related OTUs: Community II and IX exclusively consisted of the OTUs related to the *Bacteroides* genus (Figure 2C). Although results have thus far suggested that assortative structures coincided with low microbial diversity, there are notable exceptions. Community VIII exhibited both assortative and core motifs and is composed of higher taxonomic microbial diversity at the genus level (Figure. 2C). The simultaneous existence of mesoscale architectures may suggest that assortativity is not a simple derivative of taxonomic relatedness. This hypothesis is further confirmed by an additional analysis assessing community assortativity as a product of taxonomic lineage. When OTUs were forced into communities based on their taxonomic classification at the phylum and family/genus levels (as per Figure 2D), we observed that mean nodal assortativity was lower compared with the WSBM-derived results.

Determining the Functional Contribution of Communities

To gain insight into the functional contribution of each community, we used PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) Langille et al., 2013, a validated computational modeling approach that predicts a microbial community's metagenome from its 16S profile. An estimated functional contribution (i.e., to determine the degree to which OTUs contribute to specific metagenomic processes, %) was attributed at the OTU level. This information was used to estimate the functional contribution of each WSBM community. Results suggest that each community performs a broad repertoire of functional processes (Figure S6). Communities II and IX (composed exclusively of *Bacteroides*) exhibited the largest total functional contribution (%) and were particularly enriched for metabolic and genetic information processing. Results also highlight that some distinct functions were supported by specific communities. For example, processes facilitating cell motility (under *Cellular Processes*) are almost exclusively supported by community V (with a small contribution from community IV).

Identification of Hubs

Insights into the functional role of individual OTUs in the mesoscale network were then assessed by comparing interactions within (within-community *Z* score, Z_i) and outside (participation coefficient, *PC*) WSBM-derived community labels (Figure 2B) (Guimerà and Amaral, 2005) (Transparent Methods). We used previously validated cutoffs as a broad guideline (Guimerà and Amaral, 2005) and subsequently assigned each OTU into one of the two broad classifications: non-hubs nodes and community hubs (Figures 3A, 3B, and 4A). At this stage, the classification of community hubs does not take into account the total abundance of the OTU within the microbiome. For example, a prevalent OTU might share many co-occurrences within the system exclusively by virtue of their sheer abundance (Banerjee et al., 2018). To make the distinction between dominant and true connector/provincial hubs, we assessed each candidate hub in terms of the relative abundance (%) of its corresponding genus (or family, where required). OTUs whose genus/family had a total relative abundance greater than 5% were re-classified as dominant hubs (Figure 3). Candidate OTUs with less than 5% total relative abundance in the microbiome were referred to as "true" provincial or connector hubs. We then visualized the proportion of "true" provincial and connector hubs, dominant hubs, and non-hub OTUs within each community (Figure 4B), as well as the proportion within each genus/family classification (Figure 4C). As the dichotomy between provincial and connector classifications within dominant hubs is informative, we also visualize the decomposition of provincial and connector hubs for the three dominant genera before they were re-classified as dominant hubs (Figures 4D–4F).

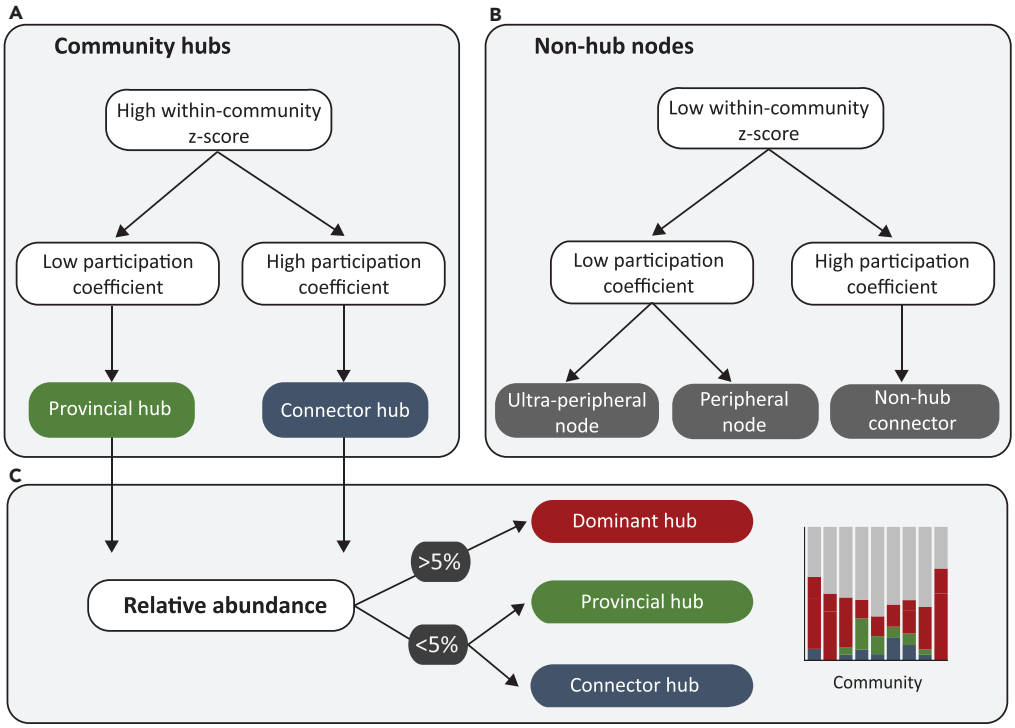


Figure 3. Classification of Individual OTUs Based on Within-and Between-Community Interactions and Relative Abundance

(A) The classification of community hubs (provincial and connector) was based on high within-community *Z* score (>0).

(B) The classification of non-module hubs (ultra-peripheral, peripheral, and non-hub connectors) was based on low within-community *Z* score (<0).

(C) Provincial and connector hubs were further assessed in terms of their relative abundance (%) at the genus level. OTUs whose genus had >5% relative abundance in the microbiome were classified as "Dominant" hubs. The proportion of dominant, connector, and provincial hubs were finally visualized for each mesoscale community (Figure 4).

Dominant hubs including *Faecalibacterium*, *Bacteroides*, and *Lachnospiraceae* were ubiquitous to all communities (red in Figures 4B and 4C) but particularly enriched within communities I, II, and IX (54%, 50%, and 69%, respectively). Specifically, community IX had a high prevalence of dominant hubs that, before re-classification, exhibited characteristics of provincial hubs. True provincial hubs, characterized by high within- and low between-community interactions, were proportionately higher in communities IV and V (23% and 13%, respectively) (green in Figure 4A). For OTUs classified as connector hubs, a high proportion originated from core and peripheral communities VI and VII, exhibiting both strong within- and between-community interactions (Figure 4C) (Rubinov and Sporns, 2010). Community VIII exhibited both core and assortative mesoscale motifs but showed a low prevalence of provincial and connector hubs. This suggests that the ability of this community to participate in core-periphery structures may largely be driven by dominant hubs. Ultra-peripheral (mauve in Figure 4A) and peripheral (red in Figure 4A) OTUs were uniformly present across all communities.

DISCUSSION

This study highlights the co-existence of diverse mesoscale network architectures in the healthy human gut microbiome. Our results provided an alternative interpretation of the microbiome's network organization, in which it exhibited a principal mesoscale structure of assortativity, alongside the existence of clearly delineated cores and peripheries. Our innovative use of advanced network modeling methods provides significant advances to the understanding of the human microbiome, which, until now, has been conceptualized as a singular assortative community structure (Jackson et al., 2018).

The detection of both assortative and core-periphery architectures in the microbiome raises questions about the ecological and evolutionary selection processes that shaped this mesoscale formation. We first

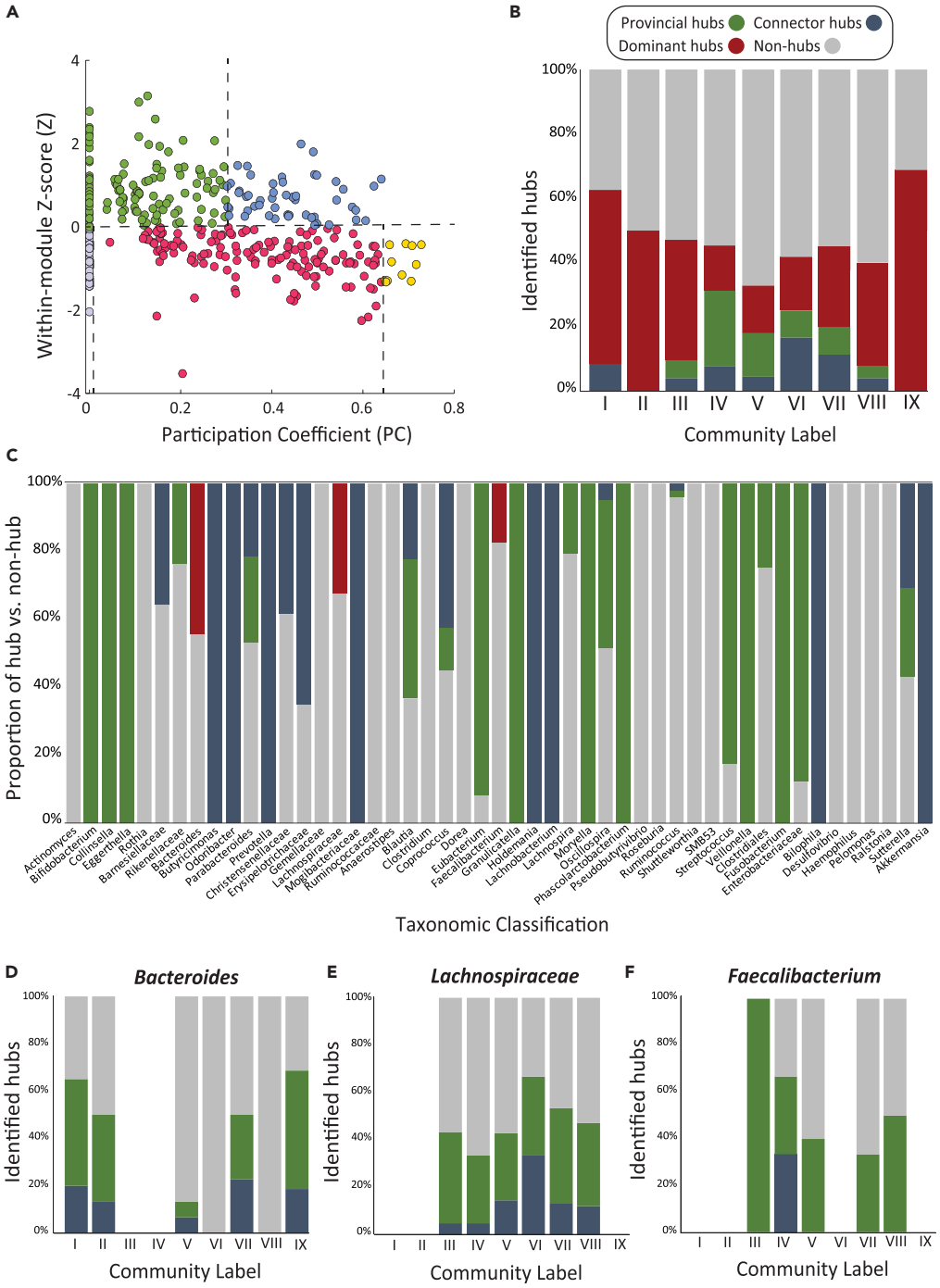


Figure 4. Role of Individual OTUs Based on Within- and Between-Community Co-occurrences

(A) OTUs were partitioned into non-community hubs ($Z < 0$) or community hubs ($Z > 0$) based on their within-community Z score (Guimerà and Amaral, 2005). Non-community hubs were then grouped into ultra-peripheral ($PC \approx 0$, mauve), peripheral ($PC < 0.625$, red), non-hub connectors ($0.625 < PC \angle 0.8$, yellow), and non-hub kinless nodes ($PC > 0.8$, dark purple). Non-community hubs were then grouped into ultra-peripheral (PC ≈ 0, mauve), peripheral (PC < 0.625, red), non-hub connectors (0.625 < PC < 0.8, yellow) and non-hub kinless nodes (PC > 0.8, no detected OTUs). Community hubs were grouped into provincial hubs (0 < PC < 0.3, green), connector hubs (PC < 0.75, blue), or kinless hubs (PC > 0.75, no detected OTUs).

Figure 4. *Continued*

(B–F) (B) Relative proportion (%) of provincial hubs (green), connector hubs (blue), dominant hubs (red), and non-hubs (gray) within each community and (C) when OTUs are assigned to their genus/family level, using the classification pipeline described in Figure 3. Relative proportion of provincial, connector, and non-hub nodes within each community for the dominant hubs (before reclassification), including all OTUs belonging to (D) *Bacteroides*, (E) *Lachnospiraceae*, and (F) *Faecalibacterium*.

consider the notion that assortativity is an important feature for microbiome structure and function. Understanding the advantage of an assortative topology has long been a focus in the study of biological networks, including brain networks (Sporns and Betzel, 2016), protein interactions (Zhang et al., 2010), and complex systems in general (Noldus and Van Mieghem, 2015). Several studies support the notion that assortative communities emerge through environmental filtering (Levy and Borenstein, 2013), whereas others suggest that assortativity represents a selective advantage to increase network robustness and functional redundancy. However, the view that microbial communities exclusively function as independent modules is at odds with findings supporting the existence of integrative core-periphery assemblages in biological systems. This evidence supports the existence of "spatially distinct and highly connected" (core) structures as a means of facilitating efficient interactions between communities (Baldassano and Bassett, 2016; Banerjee et al., 2018; Mouquet et al., 2013). In the context of the human gut microbiome, these interactions may arise from complementary resource acquiring strategies, niche partitioning, or the transfer of resources, including metabolites. Groupings of highly connected hubs, analogous to a keystone guild, may therefore decrease the distance between disparate communities comprising the network and support greater microbial diversity and species survival (Sugihara and Ye, 2009). Although core-periphery structures promote interactions between spatially disparate communities, a network largely dependent on structures increases the risk of ecosystem collapse (Sugihara and Ye, 2009). For example, when a core community or hub is perturbed, significant downstream effects to the wider network may ensue, potentially resulting in microbial dysbiosis. The vulnerability of core-periphery structures is supported by computational work, demonstrating widespread perturbation of the healthy microbiome network following the simulated removal of few hub species (Baldassano and Bassett, 2016). That is, the co-existence of core-periphery and assortative mesoscale structures allow optimal interactions between OTUs while maintaining functional redundancy, respectively (Otokura et al., 2016).

The detection of diverse and co-existing network architectures is critical to understand the broad repertoire of biological phenomena underpinning OTU interactions. A prominent feature of the detected network architecture is the development of strong assortative co-occurrence interactions among closely related OTUs, a finding consistent with previous work (Chaffron et al., 2010; Jackson et al., 2018). This observation is in line with two converging hypotheses regarding the functional organization of the microbiome. The first hypothesis postulates that the assortative grouping of closely related OTUs may be a product of environmental filtering (Stuart, 2018), where challenges including nutrition availability and substrate conditions have favored specific microbial traits. These traits are likely shared among genetically related OTUs, explaining their co-occurrence in the network. It has also been suggested that closely related OTUs have overlapping functional roles, supporting a high degree of compensation or degeneracy (Fornito et al., 2015). However, individual OTUs will still need to maintain minimal niche differentiation, as competitive exclusion remains a dominant driver of co-occurrence patterns (Stuart, 2018). In our WSBM analysis, we observed two assortative communities (II and IX) characterized by strong within-community co-occurrences between OTUs of the abundant genus *Bacteroides*. Specifically, community IX was enriched with provincial hubs above what would be expected in a random null network, suggesting this community exhibits high local integration. *Bacteroides* are known to exhibit a high degree of functional flexibility in response to changing substrate conditions in the microbiome (Rios-Covian et al., 2017). This is supported in our predicted metagenomics assessment at both the OTU and community level, suggesting that *Bacteroides* indeed exhibits a diverse repertoire of functional capacity, alongside the largest overall contribution to microbiome processes (Figure S6).

The above-mentioned results are consistent with previous community detection methods, including those undertaken by modularity maximization algorithms (Jackson et al., 2018). However, our WSBM analyses have also detected coexisting mesoscale structures. Communities IV to VIII exhibited significant assortative, core, and/or periphery interactions and were further characterized by high taxonomic microbial diversity (Figures 2C and 2D). Evidence for "true" (those characterized by low relative abundance in the system) connector and provincial hubs suggest they have a central role in facilitating network integration (Fornito

et al., 2015). This is in line with previous work (Ze et al., 2013) demonstrating that the growth and survival of many OTUs may rely on interactions with a few core OTUs. The existence of connector hubs in a core-periphery architecture also presents some risks, as perturbation to these nodes may result in whole-network destabilization.

Our mesoscale assessment thus far has distilled important organizational principles of the gut microbiome. Just as important, however, is to understand how network topological properties inform function. We performed a metagenomics assessment to discern generalizable functional patterns both within and between our WSBM-derived communities (Figure S6). Although we observed some evidence for a mono-functional system (i.e., distinct functions performed by single or few communities), the microbiome can be better described as a multi-functional system (Moya and Ferrer, 2016). That is, irrespective of taxonomic composition or diversity, each community contributes to a broad number of overlapping functions. Our findings are consistent with recent work (Heintz-Buschart and Wilmes, 2018; Zhu et al., 2015) and supports the gut microbiome's preference to maintain high functional redundancy and stability.

Limitations of the Study

Community detection algorithms, like the WSBM, can offer powerful exploratory tools to elicit insights into the macroscopic (large-scale) and microscopic (e.g., microbe-microbe) patterns underpinning ecological networks. However, it is important to note that no algorithm can capture ground truth mesoscale architecture (Peel et al., 2017). Thus, the insights gained via our approach must be complemented with empirical data to allow a more direct assessment of the posited patterns of interactions and their functional roles. While the community partitions achieved by the WSBM draws parallels with known biological mechanisms, we have balanced the interpretation of the results to reflect these limitations. It is also important to consider the limitations associated with 16S rRNA datasets. Although OTUs were clustered based on the standard 97% similarity thresholds, compared with deeper sequencing techniques (i.e., shotgun metagenomics sequencing), our taxonomic classification lacks the sensitivity to detect all bacterial species. In addition, owing to the bias introduced by the PCR amplification step during the 16S rRNA gene procedure, not all microbes may be represented in the dataset. Therefore, potentially significant co-occurrence interactions between individual species or strains may not have been captured through 16S rRNA gene analysis. Future work adopting metagenomics shotgun sequencing (Almeida et al., 2019) is required to confirm our findings. Such analysis may also achieve a more nuanced level of specificity, as well as facilitate the analysis of directed networks. Directed networks may provide a more complete picture of biological interactions, including the directionality of trophic, metabolic, and signaling pathways of bacteria, fungi, and archaea.

In summary, our study highlights the importance of characterizing heterogeneous and co-existing mesoscale architectures to understand the ecology and functions of the human gut microbiome. We reported two independent microbial co-occurrence networks exhibiting core-periphery structures, supported by a backbone of strong assortative community interactions. These findings represent an advance over the current view of a modular and segregated microbial environment, presenting opportunities for research and clinical endeavors on the gut microbiome.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

DATA AND CODE AVAILABILITY

Data and code related to this paper may be requested from the authors. Raw sequences are available online (https://www.ncbi.nlm.nih.gov/sra/PRJNA577760). The accession number for the dataset reported in this paper is NCBI: [PRJNA577760].

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2019.11.032.

ACKNOWLEDGMENTS

We thank S. Sonkusare for helpful feedback on the early version of the manuscript. L.C. would like to acknowledge support from the Australian National Health and Medical Research Council (NHMRC), APP1099082. A.Z. was supported by an NHMRC Senior Research Fellowship B, ID:1136649.

AUTHOR CONTRIBUTIONS

Formal Analysis, C.V.H., A.L., R.B., M.Z., A.Z., and L.C.; Data Curation, C.V.H and M.Z.; Review and Editing: C.V.H., M.Z., A.Z., and L.C.; Funding acquisition, A.Z., G.R-S., and L.C.; Supervision, A.Z., and L.C.

DECLARATION OF INTERESTS

The authors declare they have no competing interests.

Received: April 29, 2019
Revised: September 27, 2019
Accepted: November 15, 2019
Published: December 20, 2019

REFERENCES

Aicher, C., Jacobs, A.Z., and Clauset, A. (2014). Learning latent block structure in weighted networks. J. Complex Netw. *3*, 221–248.

Almeida, A., Mitchell, A.L., Boland, M., Forster, S.C., Gloor, G.B., Tarkowska, A., Lawley, T.D., and Finn, R.D. (2019). A new genomic blueprint of the human gut microbiota. Nature *568*, 499–504.

Baldassano, S.N., and Bassett, D.S. (2016). Topological distortion and reorganized modular structure of gut microbial co-occurrence networks in inflammatory bowel disease. Sci. Rep. *6*, 26087.

Banerjee, S., Schlaeppi, K., and Van Der Heijden, M.G.A. (2018). Keystone taxa as drivers of microbiome structure and functioning. Nat. Rev. Microbiol. *16*, 567–576.

Betzel, R.F., Medaglia, J.D., and Bassett, D.S. (2018). Diversity of meso-scale architecture in human and non-human connectomes. Nat. Commun. *9*, 346.

Buffie, C.G., Jarchum, I., Equinda, M., Lipuma, L., Gobourne, A., Viale, A., Ubeda, C., Xavier, J., and Pamer, E.G. (2012). Profound alterations of intestinal microbiota following a single dose of clindamycin results in sustained susceptibility to *Clostridium difficile*-induced colitis. Infect. Immun. *80*, 62–73.

Chaffron, S., Rehrauer, H., Pernthaler, J., and Von Mering, C. (2010). A global network of coexisting microbes from environmental and whole-genome sequence data. Genome Res. *20*, 947–959.

Faskowitz, J., Yan, X., Zuo, X.-N., and Sporns, O. (2018). Weighted stochastic block models of the human connectome across the life span. Sci. Rep. *8*, 12997.

Faust, K., and Raes, J. (2012). Microbial interactions: from networks to models. Nat. Rev. Microbiol. *10*, 538–550.

Fornito, A., Zalesky, A., and Breakspear, M. (2015). The connectomics of brain disorders. Nat. Rev. Neurosci. *16*, 159–172.

Guimerà, R., and Amaral, L. (2005). Cartography of complex networks: modules and universal roles. J. Stat. Mech. *2005*, P02001-1–P02001-13, nihpa35573.

Halfvarson, J., Brislawn, C.J., Lamendella, R., Vazquez-Baeza, Y., Walters, W.A., Bramer, L.M., D'amato, M., Bonfiglio, F., Mcdonald, D., Gonzalez, A., et al. (2017). Dynamics of the human gut microbiome in inflammatory bowel disease. Nat. Microbiol. *2*, 17004.

Heintz-Buschart, A., and Wilmes, P. (2018). Human gut microbiome: function matters. Trends Microbiol. *26*, 563–574.

Jackson, M.A., Bonder, M.J., Kuncheva, Z., Zierer, J., Fu, J., Kurilshikov, A., Wijmenga, C., Zhernakova, A., Bell, J.T., Spector, T.D., and Steves, C.J. (2018). Detection of stable community structures within gut microbiota co-occurrence networks from different human populations. PeerJ *6*, e4303.

Jiang, H., Ling, Z., Zhang, Y., Mao, H., Ma, Z., Yin, Y., Wang, W., Tang, W., Tan, Z., Shi, J., et al. (2015). Altered fecal microbiota composition in patients with major depressive disorder. Brain Behav. Immun. *48*, 186–194.

Langille, M., Zaneveld, J., Caporaso, J., and Mcdonald, D. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat. Biotechnol. *31*, 814–821.

Levy, R., and Borenstein, E. (2013). Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. Proc. Natl. Acad. Sci. U S A *110*, 12804.

Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K., and Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. Nature *489*, 220–230.

Mouquet, N., Gravel, D., Massol, F., and Calcagno, V. (2013). Extending the concept of keystone species to communities and ecosystems. Ecol. Lett. *16*, 1–8.

Moya, A., and Ferrer, M. (2016). Functional redundancy-induced stability of gut microbiota subjected to disturbance. Trends Microbiol. *24*, 402–413.

Noldus, R., and Van Mieghem, P. (2015). Assortativity in complex networks. J. Complex Netw. *3*, 507–542.

Otokura, M., Leibnitz, K., Shimokawa, T., and Murata, M. (2016). Evolutionary core-periphery structure and its application to network function virtualization. Nonlinear Theory Appl. IEICE *7*, 202–216.

Peel, L., Larremore, D.B., and Clauset, A. (2017). The ground truth about metadata and community detection in networks. Sci. Adv. *3*, e1602548.

Poudel, R., Jumpponen, A., Schlatter, D.C., Paulitz, T.C., Gardener, B.B.M., Kinkel, L.L., and Garrett, K.A. (2016). Microbiome networks: a systems framework for identifying candidate microbial assemblages for disease management. Phytopathology *106*, 1083–1096.

Power, M.E., Tilman, D., Estes, J.A., Menge, B.A., Bond, W.J., Mills, L.S., Daily, G., Castilla, J.C., Lubchenco, J., and Paine, R.T. (1996). Challenges in the Quest for Keystones: identifying keystone species is difficult—but essential to understanding how loss of species will affect ecosystems. BioScience *46*, 609–620.

Rios-Covian, D., Salazar, N., Gueimonde, M., and De Los Reyes-Gavilan, C.G. (2017). Shaping the metabolism of intestinal Bacteroides population through diet to improve human health. Front. Microbiol. *8*, 376.

Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. Neuroimage *52*, 1059–1069.

Sanz, Y., Olivares, M., Moya-Pérez, Á., and Agostoni, C. (2014). Understanding the role of gut microbiome in metabolic disease risk. Pediatr. Res. *77*, 236.

Sporns, O., and Betzel, R.F. (2016). Modular brain networks. Annu. Rev. Psychol. *67*, 613–640.

Stuart, S.A. (2018). Phylogenies in ecology: a guide to concepts and methods. In Austral Ecology, *43*, M.W. Cadotte and T.J. Davies, eds. (Princeton University Press), p. e19.

Sugihara, G., and Ye, H. (2009). Cooperative network dynamics. Nature *458*, 979.

The Human Microbiome Project Consortium, Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J.H., Chinwalla, A.T., Creasy, H.H., Earl, A.M., Fitzgerald, M.G., Fulton, R.S., et al. (2012). Structure, function and diversity of the healthy human microbiome. Nature *486*, 207–214.

Ze, X., Le Mougen, F., Duncan, S.H., Louis, P., and Flint, H.J. (2013). Some are more equal than others: the role of "keystone" species in the degradation of recalcitrant substrates. Gut Microbes *4*, 236–240.

Zelezniak, A., Andrejev, S., Ponomarova, O., Mende, D.R., Bork, P., and Patil, K.R. (2015). Metabolic dependencies drive species co-occurrence in diverse microbial communities. Proc. Natl. Acad. Sci. U S A *112*, 6449.

Zhang, S., Ning, X.-M., Ding, C., and Zhang, X.-S. (2010). Determining modular organization of protein interaction networks by maximizing modularity density. BMC Syst. Biol. *4*, S10.

Zhu, A., Sunagawa, S., Mende, D.R., and Bork, P. (2015). Inter-individual differences in the gene content of human gut bacterial species. Genome Biol. *16*, 82.

ISCI, Volume *22*

Supplemental Information

Co-existence of Network Architectures Supporting the Human Gut Microbiome

Caitlin V. Hall, Anton Lord, Richard Betzel, Martha Zakrzewski, Lisa A. Simms, Andrew Zalesky, Graham Radford-Smith, and Luca Cocchi

Supplemental Material

Transparent Methods

Human microbiome datasets

The study was approved by the Human Research Ethics Committees of the Royal Brisbane and Women's Hospital (RBWH), Brisbane, Australia, and QIMR Berghofer Medical Research Institute, Brisbane, Australia. Written informed consent was obtained for all study participants in accordance with the Helsinki Declaration. The first dataset was generated from gut mucosal biopsies from 58 healthy Australian adults (53% female; mean age 52.3 ±12.5 years) who were attending the RBWH Department of Gastroenterology for a routine colonoscopy as part of a colorectal cancer family history screening (Zakrzewski et al., 2019). Gut mucosal biopsies were collected during the endoscope, transported on dry ice to the coordinating site, and stored at -80^{o}C until analyses. Exclusion criteria included previous diagnosis of a gastrointestinal disease or disorder, complex chronic illness, or pregnancy. In the event that a participant presented with abnormal colonoscopy or biopsy results, they were subsequently excluded. For the replication analysis, we accessed an independent and publicly available dataset, generated from 528 human fecal samples (Yatsunenko et al., 2012). The replication dataset included 326 individuals aged 0-17 years, and 202 adults aged 18-70 years, representing healthy Amerindians from the Amazonas of Venezuela, rural Malawians, and residents from USA metropolitan areas.

DNA extraction and 16S rRNA analysis

Following tissue homogenisation using tubes containing 1.4mm ceramic beads (Precellys Lysing Kit), DNA was extracted from samples using DNeasy Blood and Tissue Kit (QIAGEN). DNA was quantitated using Nanodrop 2000 (Thermo Scientific). PCR amplification was performed on the V3-V4 hypervariable region of the 16S rRNA gene, and sequenced on a MiSeq sequencer (Australian Genome Research Facility (AGRF), Brisbane, Australia). Sequence data were processed using Quantitative Insights Into Microbial Ecology (QIIME) software suite v1.9.1 using default settings. Low quality reads were filtered and removed, and remaining sequences were de-multiplexed using a custom script to reduce the possibility of a mismatch. Clustering occurred in a two-step process. Firstly, sequences were clustered into operational taxonomic units (OTUs) based on existing sequences in the Greengenes database v13.5 (97% identity threshold). Secondly, unclustered reads were clustered de-novo (97% identify threshold). The USEARCH package (UCLUST v8.0.1623) was used to assign the representative OTUs to taxonomic lineage, using the Greengenes database as a reference. Chimeric seed sequences as identified by UCHIME were removed, and singleton OTUs were discarded for all downstream analyses. Metagenomics functions were predicted using the PICRUSt algorithm (Langille et al., 2013), following the tutorial steps outlined elsewhere (http://picrust.github.io/picrust/). The functional profiles for each subject were deconvolved into OTU-specific functional profiles, which provides an estimation of how each OTU contributes to the sum of functional processes performed by a given community. The data analysis pipeline for the replication dataset, also using the Greengenes database for OTU clustering, has been described elsewhere [see Methods section of (Yatsunenko et al., 2012)].

Co-occurrence network construction

We prepared our microbial interaction network in accordance with Berry and Widder's (Berry and Widder, 2014) best practices for co-occurrence network construction. As OTU abundance tables are sparse (with values of 0 representing an absence of an OTU in a sample), we applied a threshold to filter out infrequent OTUs and included only those with the largest relative abundances (top 20%) across all 58 samples. For each pair of OTUs, we then computed the Pearson correlation coefficient across individuals to generate an undirected, weighted co-occurrence network between all pairs of OTUs. Accordingly, the matrix represented weighted pairwise interactions for 370 OTUs. The replication dataset (n = 528) was thresholded to include OTUs in the top 10% with the largest absolute abundances and subsequently represented weighted pairwise interactions for 1225 OTUs. The different threshold adopted for the replication dataset was selected to: (i) maintain a comparable number and microbial diversity as seen in the original dataset, and (ii) strike a balance between scientific rigor and computational limitations.

Applying the WSBM to the microbiome

The WSBM is a generative modelling approach used to detect and partition a network's nodes into a number of latent communities, *k*. The WSBM described here extends upon the classical stochastic block model by allowing the inclusion of weighted edges in sparse networks. This model eliminates the need to threshold a

matrix before input to the WSBM and thus, preserves information about the weights of co-occurrence and co-exclusion relationships. MATLAB (The Mathworks, USA) codes for the WSBM were sourced via http://tuvalu.santafe.edu/~aaronc/wsbm/. We applied the WSBM to create a weighted network of the human intestinal microbiome, using steps described in detail elsewhere (Aicher et al., 2014).

To infer community structure from the WSBM, the user is required to input *k,* a free parameter representing the total number of communities. The model selection technique we describe here optimizes *k* based on a value that maximizes the marginal log-likelihood and penalizes model complexity (Aicher et al., 2014). To achieve a representative partition from the WSBM, we varied the number of communities from $k = 5$ to $k = 18$ and repeated the algorithm 65 times. We inferred *k* by determining where the mean and maximum marginal log-likelihood plateaus. For values greater than $k = 12$, we observed a clear plateau, and selected the simplest model (**Figure. 2A**). Given the stochastic nature of the WSBM, we assessed the consistency of the resulting outputs using NMI (Cover, 2012). Specifically, we fit the WSBM to the same dataset for 65 additional independent repetitions at our selected model, $k = 12$, to create a new frequency prior. Our marginal log-likelihood values (50 internal trials over 65 fits) were highly consistent with the original fits, with an NMI of 0.99. These findings were replicated in our supplemental dataset, where we observed a plateau betwen $k = 12$ and $k = 13$ ($\boldsymbol{P} = 0.52$) (Supplemental Material, **Figure. S1**).

Consistency of community assignments

Each stochastic repetition of the WSBM returns a subdivision of non-overlapping groups of OTUs to one of 12 distinct communities. To assess if our detected partitions (n = 65) were significantly more consistent across trials to those obtained from a random network, we employed the *randmio_und* function from the BCT toolbox (Rubinov and Sporns, 2010), permuting the strength of connections while preserving the signed degree distribution. This resulted in 10 independent null networks. For each null, we then applied the WSBM for 65 fits, each with 5 internal trials (established as sufficient to ascertain consistent outputs for the null). The consistency between community assignments within each null was assessed using NMI. We then compared the consistency of the mean NMI scores (n = 650) with our mean MNI observed in our microbiome data.

Creating a consensus partition from WSBM fits

Due to the stochastic nature of the WSBM, the community assignment designated to each OTU may differ between runs. To ensure that each community was represented by a consistent identifier across all runs, we applied a progressive median alignment method using the software language R. This algorithm was developed in house and adapted from previous work (Lord et al., 2012). With our predefined *k*, the alignment script relabels communities to maximize consistency of community labels within each OTU across all runs. As such, the community structure within each subject is preserved, while consistency between labeling between runs is increased. After alignment, a single consensus partition that most accurately represents the 65 WSBM fits can be derived. We observed that, on average, OTUs at the individual level were assigned to the same cluster as the community consensus cluster 57% of the time (chance level 8.3%).

Visualization of the microbial network

The interactive platform Gephi (Version 0.9.2) was used to visualize the co-occurrence network, using the force atlas template (Bastian M., 2009). The network was reduced for visualization, using a cut-off of 0.3. This resulted in a network representing 370 nodes (OTUs) and 4854 edges (co-occurrence interactions). OTUs were colored according to community label, and taxonomic grouping at the family/genus level.

Community Motif Participation

We applied community morphospace analysis (Betzel et al., 2018) to classify community interactions into one of four motifs: assortative, disassortative, core, or periphery interactions. The community-level structures generated from 65 WSBM fits yielded a mesoscale description that mapped onto individual nodes (total n = 370, for the first dataset). Therefore, node *i*'s overall participation in each mesoscale structure was calculated as the number of times that an individual OTU's community interacted with other communities to form an interaction motif. Motif participation was averaged over 65 WSBM fits and expressed at the node (OTU), and community level. The significance of the detected mesoscale community structures was determined using permutation testing. For each community (n = 10), 650 null community partitions were created, comprising the same number and size of communities in the empirical dataset. A two-tailed test of the null hypothesis was performed with the resulting null distributions.

Nodal Assortativity

To determine nodal assortativity (Faskowitz et al., 2018), the weighted connectivity of node i (a_i) based on its assigned community (z) is compared to the maximum weighted connectivity to other communities. For each node, the weighted connection density to community m is defined as:

$$a_i(m) = \frac{1}{n_r} \sum_{j \in r} A_{ij} ,$$

from which the nodal assortativity can be computed, as below:

$$a_i = a_i(z_i) - \max_{m \neq z_i} a_i(m).$$

Participation coefficient

The co-occurrence matrix was thresholded at 0.4, and network measures in the BCT toolbox (Rubinov and Sporns, 2010) were calculated to achieve insights at the global, mesoscale, and local scale. The participation coefficient (*PC*) measures the diversity of inter-community connections of node i, weighted by the importance of connections, and is defined as:

$$PC_i = 1 - \sum_{m \in M} \left(\frac{k_i(m)}{k_i} \right)^2 ,$$

where M is the number of communities, and $k_i\ (m)$ is the total connections between node *i* and all nodes in community *m*. A high value of *P* indicates that node i has strong connections with nodes outside its community, relative to connections within the community.

Within-community degree z-score

The within-community degree z-score, Z_i, measures the degree of intra-community connectivity of node i relative to other nodes in the community, and is defined as:

$$Z_i = \frac{k_i(m_i) - \bar{k}\ (m_i)}{\sigma^{k(m_i)}} ,$$

where m_i is the community containing node i, $k_i(m_i)$ is the within-community degrees between node i and all other nodes, $\bar{k}(m_i)$ and $\sigma^{k(m_i)}$ are the mean and standard deviation of the within-community degrees distribution. Higher values of Z_i (> 0) indicates node i has strong connectivity within its own community.

Nodal Strength

Nodal strength is defined as the sum of weighted edges connected to that node:

$$s_i = \sum_j w_{ij} ,$$

where w_{ij} is the weight of edges between *i* and *j*.

Functional cartography measures

We broadly classified each node (OTU) into one of seven possible roles according to their participation coefficient and within-community Z scores, using previously validated cut-offs as a guideline (Guimerà and Amaral, 2005). This was achieved in a two-step process, involving:

1. *Classification of non-hub nodes (Z_i<0):*
    a) Ultra-peripheral nodes, characterized by exclusive intra-community connections (*P* ≈ 0);
    b) Peripheral nodes, characterized by moderately high intra- and low inter-community connections (*PC* <0.625);
    c) Non-hub connectors, where 62.5 – 80% of connections are within the community (0.625 < *PC* < 0.8);
    d) Non-hub kinless nodes, nodes that cannot be associated with any single community (*PC* >0.8).
2. *Classification of community hubs (Z_i>0)*
    e) Provincial hubs, characterized by high intra- and low inter-community connections (0 < *PC* < 0.30)
    f) Connector hubs, characterized by moderately high inter- and high intra-community connections (*PC* < 0.75)
    g) Kinless hubs, similar to non-hub kinless nodes, are not clearly associated with any single community (*PC* > 0.75).

Supplemental Figure Legends

Figure S1. Bayesian model selection using the WSBM on a replication dataset, Related to Figure 2A. We repeated the WSBM algorithm for our replication dataset (n = 528), across a number of models between k = 10 to k = 16 (repeated 10 times each). Using the Bayesian model optimization method (details in Transparent Methods), the replication dataset was partitioned into k = 12 communities (green, dashed), given that the mean marginal log-likelihood first plateaued at this point (**Figure. S1.**). As observed in the main text, assortativity was the dominant mesoscale community structure detected, alongside observations of core and peripheries. Again, no disassortative mesoscale structures were detected in the replication dataset. In the main text, we demonstrated maximum assortativity was evident 37% of the time across 370 OTUs. We reproduced this analyses and found that the results corroborate those of the main text. We observed maximum assortativity 27% of the time across 1225 OTUs. When calculating the proportion of times a node participated in non-assortative motifs, we observed core and peripheral interactions 3.84 ± 4.0, and 2.00 ± 3.40 percent of the time, respectively. These supplementary findings bolster the findings from the main text, and suggest that our results can be observed in two, independent microbial co-occurrence datasets. The replication dataset demonstrated that the human microbiome can be described as an assortative network with co-existing and nested non-assortative mesoscale architectures, as detected by the WSBM.

Figure S2. Consistency of detected communities as benchmarked against null models, Related to Table 1. The consistency of each community detection partition was assessed by calculating the NMI between each possible combination of partitions (total = 65), and benchmarked against 10 null models (each with 65 null partitions) generated from randomized networks. The community partitions observed in the microbiome were highly consistent (0.70 ± 0.02, red dashed line), and significantly different when compared to 650 random partitions generated from rewired null networks (NMI, 0.38 ± 0.02, blue).

Figure S3. Individual contribution to the group-level co-occurrence network, Related to Figure 2B&D. To derive information about single-subject community variability and confirm that our WSBM analyses is representative of the group, we performed a cross-validation analysis based on the leave-one-out (LOO) approach. In the LOO strategy, each subject's contribution to the resultant network structure is estimated by leaving that subject out and re-estimating group-level co-occurrence coefficients. To assess the similarity between the original and the n-1 networks, we used Mantel's test statistic, where a value of 1 indicates high similarity between networks (Mantel, 1967). We then assessed the average variability of each subject's contribution (μ) for each pairwise OTU interaction by subtracting the subject with the group-level correlation matrix. To test that our LOO strategy did not result in changes to WSBM-derived mesoscale architectures, we selected subjects with the greatest variability (>1 STD from the mean, subjects 13, 22, 33, 40, and 45) and re-fit these matrices to the WSBM 65 times at k = 12. The resultant consensus partition obtained from each of the networks (mean NMI = 0.90) were highly consistent with the original group-level analyses.

Figure S4. Comparing WSBM with an alternative community detection approach, Related to Figure 2B&D. To assess the validity of our findings, we assessed the putative modular structure using an alternative algorithm: The deterministic spectral modularity maximization algorithm (sites.google.com/site/bctnet/; function modularity_und). Modularity maximization detected five assortative communities that overlapped with the nine detected via the WSBM. Moreover, the mean nodal assortativity was highly consistent between algorithms (**R** = 0.83, **P** <0.001). These results are consistent with previous work comparing the WSBM to modularity maximization (Faskowitz et al., 2018). We here iterate that, contrary to the WSBM, modularity maximization can only detect assortative structures. Our comparative results suggest that the WSBM provides greater sensitivity to the detection of structures that do not conform to the strict assortative description required by other methods.

Figure S5. Proportions of mesoscale motif participation (%) under a permutation-based null model, Related to Table 1. To test that our detected assortative, core, and peripheral mesoscale motif interactions were greater than chance, we created 650 random partitions of community assignment (blue) and re-calculated the percentage of motif participation for a community of the same size and number. We benchmarked our observed community motif interactions (red, dashed) against the null community partitions (blue) and determined significance at $\boldsymbol{P} < 0.05$. (A) Significant core motif interactions were observed within communities VII, and VIII. (B) Significant periphery motif interactions were observed within communities IV, V, and VI. (C) Significant assortative motif interactions were observed within communities II, VIII, and IX.

Figure S6. Predicted functional contributions for each community based on PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) (Langille et al., 2013), Related to Figure 2. Results showed that each community has the ability to perform a broad repertoire of functions.

Supplemental References

Aicher, C., Jacobs, A. Z. & Clauset, A. (2014). Learning latent block structure in weighted networks. *Journal of Complex Networks,* 3**,** 221-248.

Bastian M., H. S., Jacomy M. (2009). Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*.

Berry, D. & Widder, S. (2014). Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Frontiers in microbiology,* 5**,** 219-219.

Betzel, R. F., Medaglia, J. D. & Bassett, D. S. (2018). Diversity of meso-scale architecture in human and non-human connectomes. *Nature Communications,* 9**,** 346.

Cover, T. M., Thomas, J.A. (2012). *Elements of information theory.*, John Wiley & Sons.

Faskowitz, J., Yan, X., Zuo, X.-N. & Sporns, O. (2018). Weighted Stochastic Block Models of the Human Connectome across the Life Span. *Scientific Reports,* 8**,** 12997.

Guimerà, R. & Amaral, L. a. N. (2005). Cartography of complex networks: modules and universal roles. *Journal of statistical mechanics (Online),* 2005**,** nihpa35573-nihpa35573.

Langille, M. G. I., Zaneveld, J., Caporaso, J. G., Mcdonald, D., Knights, D., Reyes, J. A., Clemente, J. C., Burkepile, D. E., Vega Thurber, R. L., Knight, R., Beiko, R. G. & Huttenhower, C. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology,* 31**,** 814.

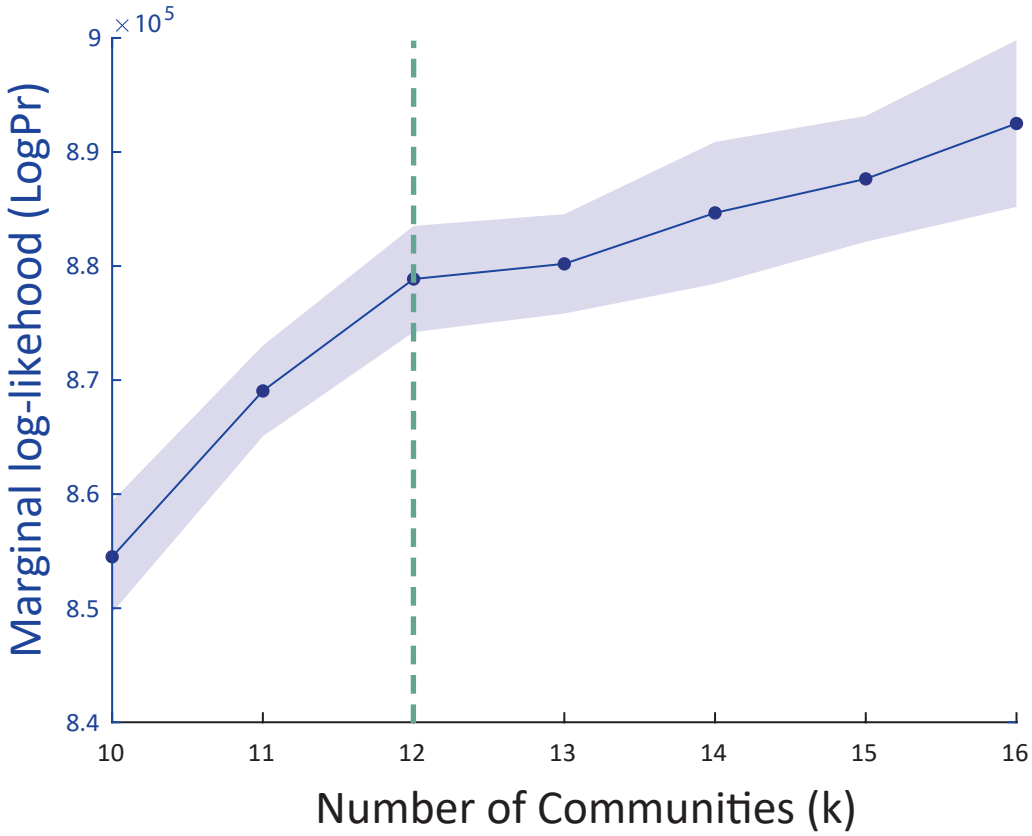
Lord, A., Horn, D., Breakspear, M. & Walter, M. (2012). Changes in Community Structure of Resting State Functional Connectivity in Unipolar Depression. *PLOS ONE,* 7**,** e41282.

Mantel, N. (1967). The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Research,* 27**,** 209.

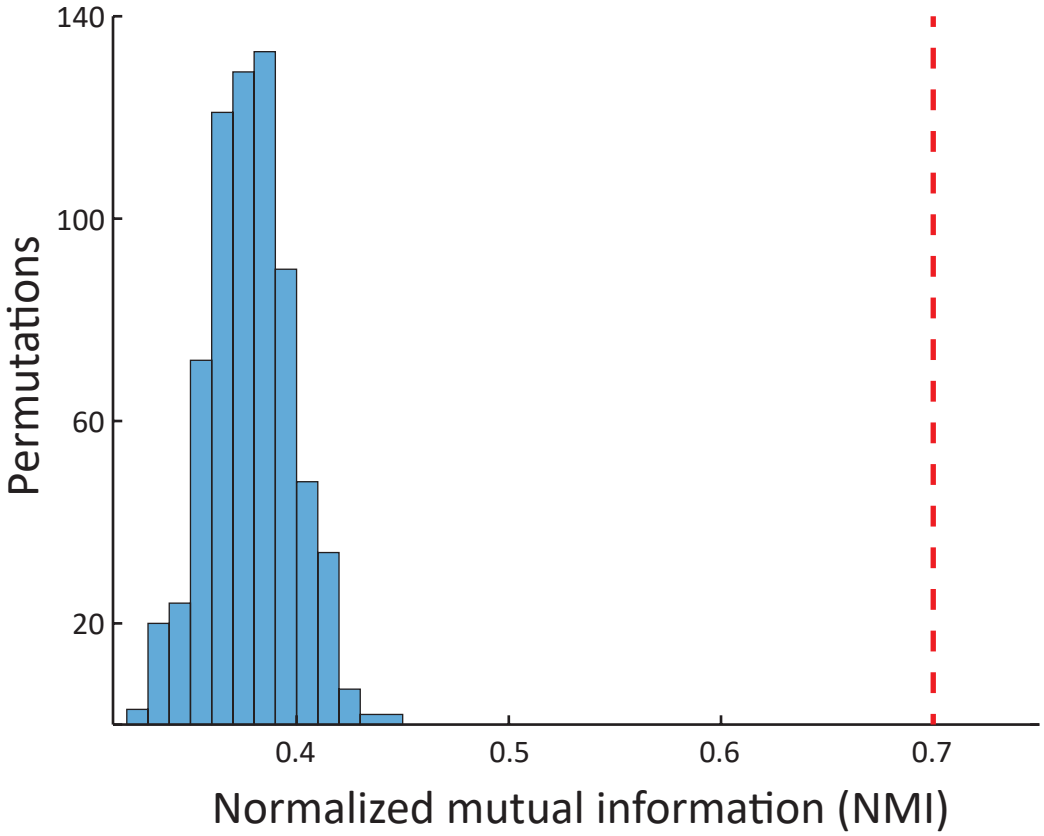
Rubinov, M. & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage,* 52**,** 1059-69.

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R. N., Anokhin, A. P., Heath, A. C., Warner, B., Reeder, J., Kuczynski, J., Caporaso, J. G., Lozupone, C. A., Lauber, C., Clemente, J. C., Knights, D., Knight, R. & Gordon, J. I. (2012). Human gut microbiome viewed across age and geography. *Nature,* 486**,** 222.

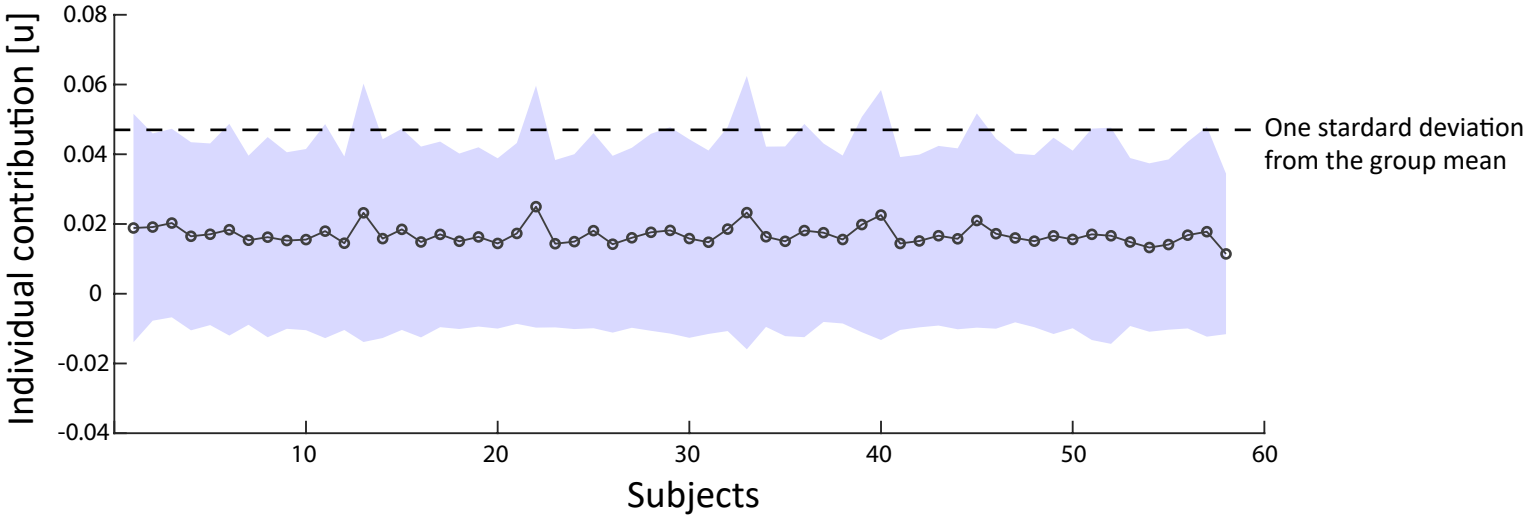
Zakrzewski, M., Simms, L. A., Brown, A., Appleyard, M., Irwin, J., Waddell, N. & Radford-Smith, G. L. (2019). IL23R-Protective Coding Variant Promotes Beneficial Bacteria and Diversity in the Ileal Microbiome in Healthy Individuals Without Inflammatory Bowel Disease. *J Crohns Colitis,* 13**,** 451-461.



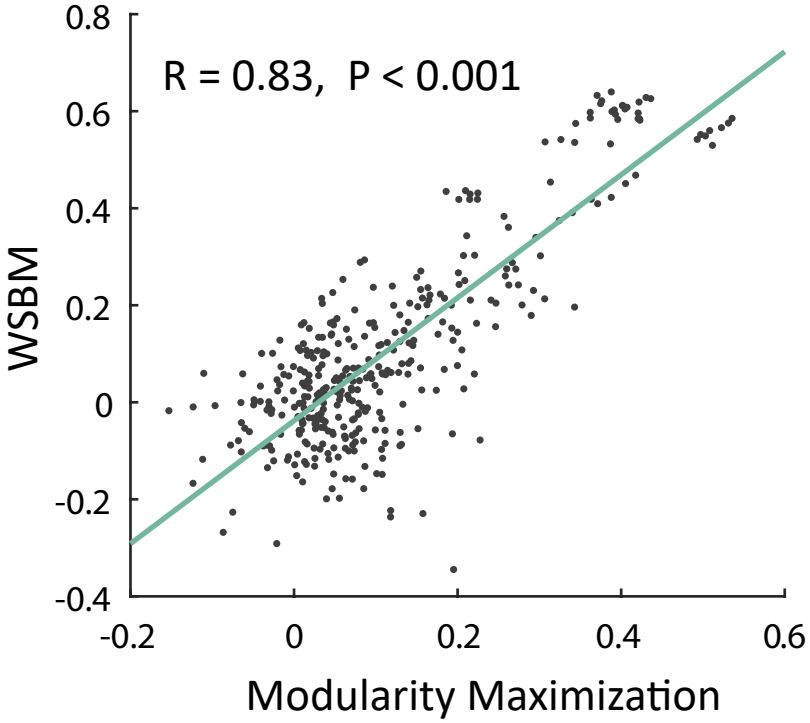
Supplemental Figure 1



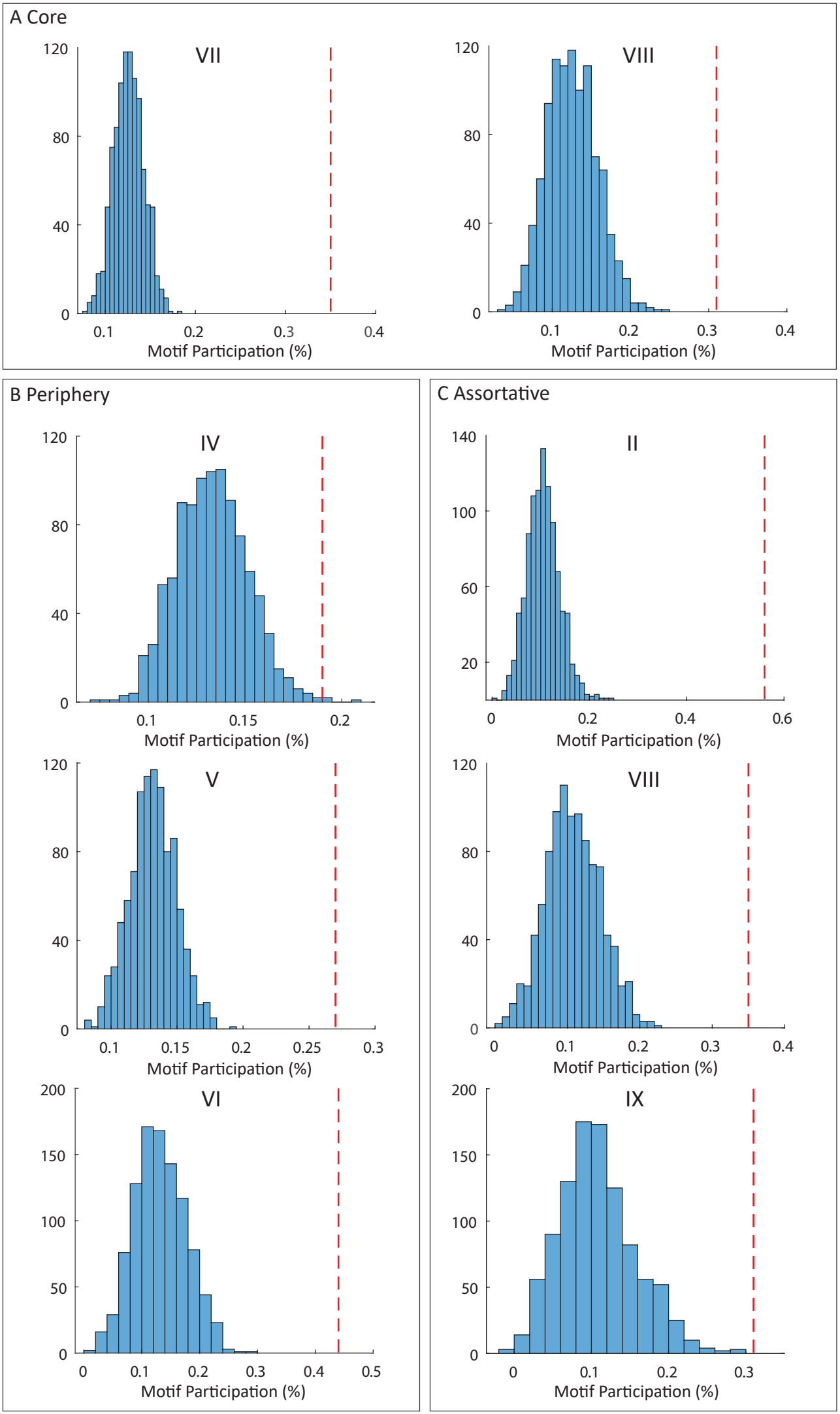
Supplemental Figure 2



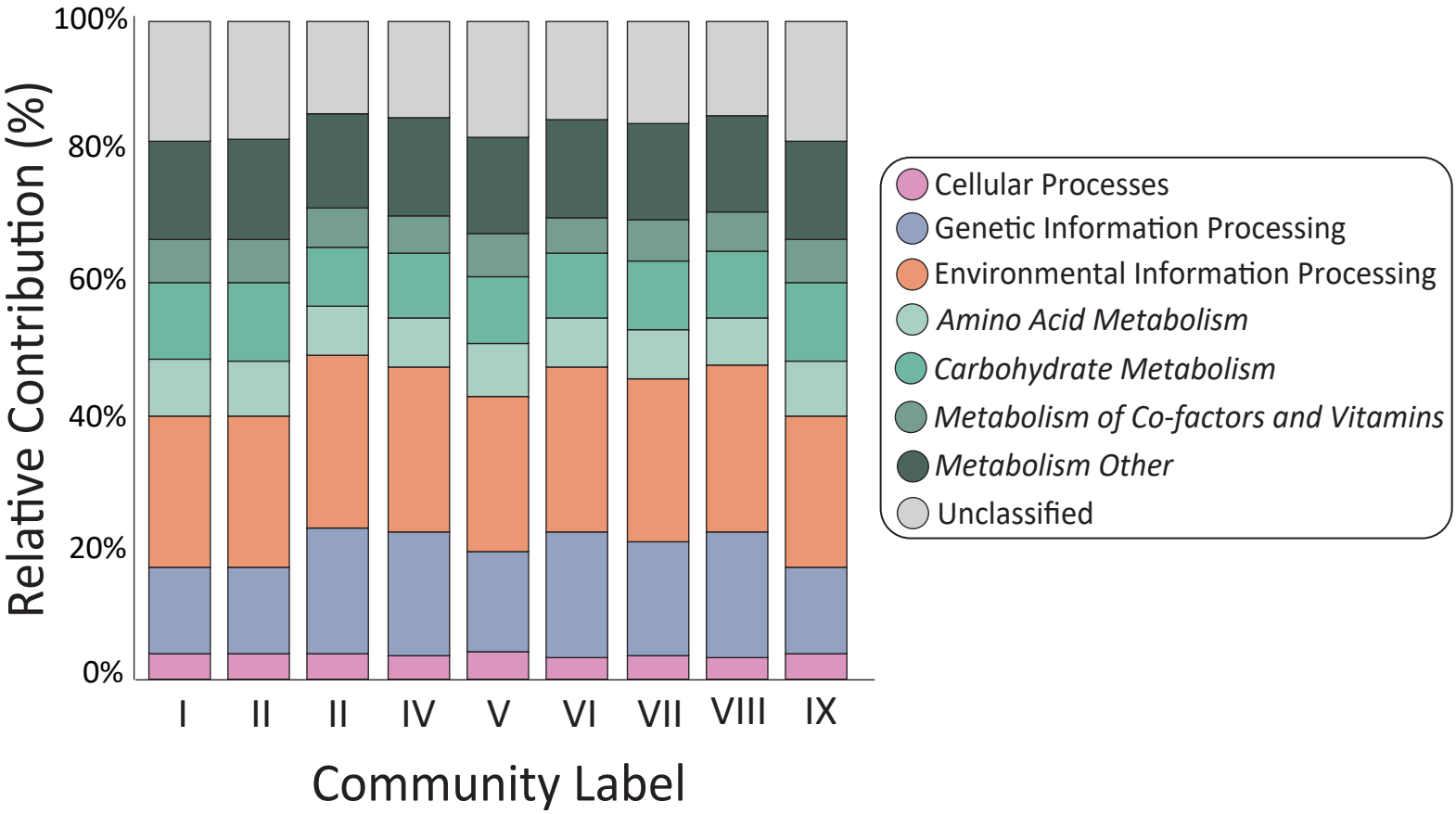
Supplemental Figure 3



Supplemental Figure 4



Supplemental Figure 5



Supplemental Figure 6