Article
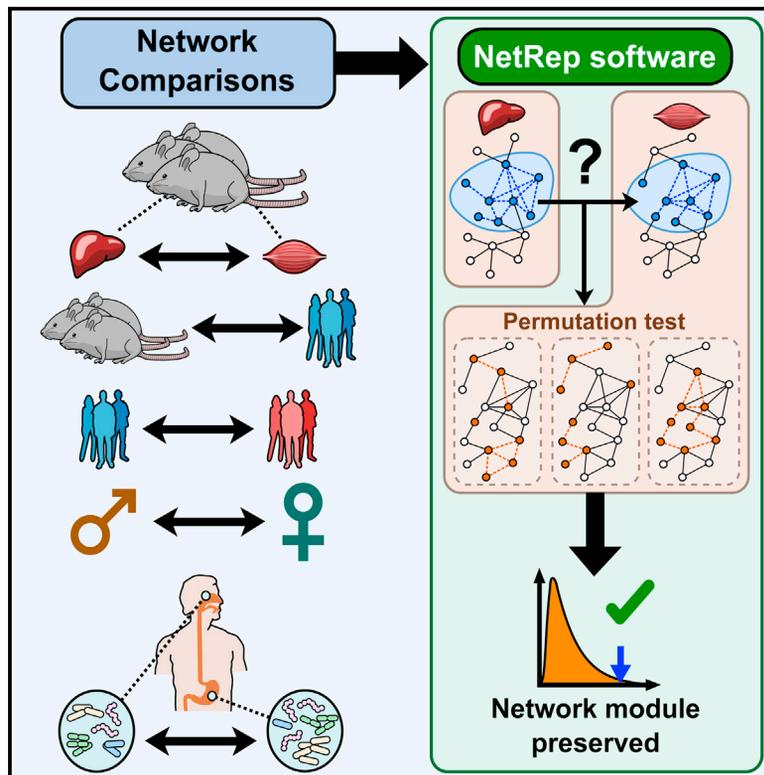
A Scalable Permutation Approach Reveals Replication and Preservation Patterns of Network Modules in Large Datasets

Graphical Abstract

Authors

Scott C. Ritchie, Stephen Watts,
Liam G. Fearnley, Kathryn E. Holt,
Gad Abraham, Michael Inouye

Correspondence

scottr@student.unimelb.edu.au (S.C.R.),
minouye@unimelb.edu.au (M.I.)

In Brief

Ritchie et al. present NetRep, an open-source tool for statistically testing replication and preservation of network modules. NetRep is shown to identify multi-tissue gene networks in mice associated with body weight as well as gut microbial community networks preserved between men and women.

Highlights

- Common network preservation statistics are non-normal

- A fast permutation-based framework and software, NetRep, is presented and tested

- Using NetRep, we identify preserved gene networks across diverse tissues

- We further identify preserved networks of human gut microbiota across genders

CrossMark

CellPress

A Scalable Permutation Approach Reveals Replication and Preservation Patterns of Network Modules in Large Datasets

Scott C. Ritchie,[1,2,*] Stephen Watts,[1,3] Liam G. Fearnley,[1,2,4] Kathryn E. Holt,[1,3] Gad Abraham,[1,2,4] and Michael Inouye[1,2,4,*]

[1]Centre for Systems Genomics, The University of Melbourne, Parkville, VIC 3010, Australia
[2]Department of Pathology, The University of Melbourne, Parkville, VIC 3010, Australia
[3]Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Parkville, VIC 3010, Australia
[4]School of BioSciences, The University of Melbourne, Parkville, VIC 3010, Australia
*Correspondence: scottr@student.unimelb.edu.au (S.C.R.), minouye@unimelb.edu.au (M.I.)
http://dx.doi.org/10.1016/j.cels.2016.06.012

SUMMARY

Network modules—topologically distinct groups of edges and nodes—that are preserved across datasets can reveal common features of organisms, tissues, cell types, and molecules. Many statistics to identify such modules have been developed, but testing their significance requires heuristics. Here, we demonstrate that current methods for assessing module preservation are systematically biased and produce skewed p values. We introduce NetRep, a rapid and computationally efficient method that uses a permutation approach to score module preservation without assuming data are normally distributed. NetRep produces unbiased p values and can distinguish between true and false positives during multiple hypothesis testing. We use NetRep to quantify preservation of gene coexpression modules across murine brain, liver, adipose, and muscle tissues. Complex patterns of multi-tissue preservation were revealed, including a liver-derived housekeeping module that displayed adipose- and muscle-specific association with body weight. Finally, we demonstrate the broader applicability of NetRep by quantifying preservation of bacterial networks in gut microbiota between men and women.

INTRODUCTION

Modern high-throughput technologies generate a large amount of genomic, transcriptomic, metabolomic, and proteomic data. Rather than consider each measurement in isolation, network inference techniques integrate these -omic data to identify meaningful biological relationships between components. In general, these approaches represent each measured variable as a node and the relationships between variables as edges that connect nodes; in aggregate, the connected edges and nodes comprise the network. Statistical analysis of these networks can identify and characterize gene modules, gene regulatory networks, protein-protein interactions, microbial networks and predict diverse molecular interactions (Abraham et al., 2014; Barabási et al., 2011; Dagan, 2011; Faust and Raes, 2012; Lusis and Weiss, 2010; Schadt, 2009).

Typically, a research project investigates one or more subgraphs of these inferred networks, for example, a group of genes associated with disease pathogenesis. These are commonly referred to as network "modules" (Gustafsson et al., 2014; Rotival and Petretto, 2014). The next step for many studies is to assess whether a network module(s) is wholly or partially preserved in an independent dataset(s); preservation is taken as an indication that the module is biologically relevant. Module preservation analysis can be used to quantify the replication of modules (Emilsson et al., 2008; Fuller et al., 2007; Hawrylycz et al., 2012; Miller et al., 2010; Ritchie et al., 2015; Xia et al., 2006), to determine their changes across conditions (Fuller et al., 2007; Keller et al., 2008; van Nas et al., 2009), to examine their tissue specificity (Cai et al., 2010; Keller et al., 2008), and to identify modules conserved across different species (Boyle et al., 2014; Gerstein et al., 2014; Stuart et al., 2003).

Module preservation analyses are both timely and increasingly common, given recent concerns about the reproducibility and generality of research findings (Collins and Tabak, 2014). However, rigorous statistical methodology for assessing module preservation has received little attention. Module preservation is typically assessed via visual inspection and/or tabulation of module composition after application of the same network inference and module detection algorithms in the second (i.e., test) dataset (Boyle et al., 2014; Gerstein et al., 2014; Keller et al., 2008; Miller et al., 2010; van Nas et al., 2009; Xia et al., 2006). A major limitation to these approaches is that they cannot systematically capture information about the network topology, i.e., the relationships between nodes in the module of interest. These relationships encode important biological information. For example, node degree (how many other nodes any given node is connected to) is a common metric analyzed in network studies, as it can indicate relative importance to the network. Genes that are highly connected are often essential to an organism's survival (Carlson et al., 2006; Jeong et al., 2001), and within a module, node degree can be used as a measure of relative

biological importance (Horvath and Dong, 2008; Langfelder et al., 2013).

To address this problem, Langfelder et al. developed a suite of statistics for quantifying the preservation of a module's topology in an independent dataset where the same module or a subset of nodes has been measured (Langfelder et al., 2011). Their module preservation statistics were primarily designed for networks inferred through weighted gene coexpression network analysis (WGCNA) (Zhang and Horvath, 2005). These are weighted, complete networks, which are defined through a power transform on the correlation structure calculated between all pairs of genes. Each gene is connected to every other gene with an edge weight between 0 and 1 denoting connection strength. Modules can either be defined as genes belonging to a pathway of interest or discovered de novo through clustering of the network (Zhang and Horvath, 2005). Seven module preservation statistics are used to quantify module preservation. For convenience, their definitions are given in the Experimental Procedures and their biological interpretation in the Supplemental Experimental Procedures. Broadly speaking, they measure whether the density and connectivity of a module are preserved in a second test dataset. The density statistics assess whether nodes composing a module remain strongly connected in the test dataset and whether measurements in each sample are similar across the module's nodes. The connectivity statistics assess whether the pattern of node-node relationships are similar between the discovery and test datasets (Langfelder et al., 2011). This approach uses Z scores to determine whether any test statistic is significant. The null hypothesis is that the module of interest is not preserved, and thus the value of each statistic should not be higher than expected by chance, assuming each statistic follows a normal distribution. However, Langfelder et al. found that Z scores were typically abnormally large, leading to many false positives in simulation. Consequently, heuristic tests for significance were formulated (Langfelder et al., 2011).

The number of modules and datasets undergoing module preservation analyses is increasing as large multi-omic datasets with dozens of tissues, cell lines, conditions, and corresponding metadata become common and openly available. In particular, studies are now performing unbiased discovery of preserved gene coexpression modules (Cai et al., 2010; Melé et al., 2015) and identifying tissue-specific and multi-tissue modules (Pierson et al., 2015; GTEx Consortium, 2015). Multiple testing becomes a problem for studies assessing preservation of many modules across many datasets; false positives may be detected due to the large number of statistical tests. Typically this is addressed through multiple testing adjustment of p values or thresholds for significance. It is therefore crucial that preservation p values are unbiased and accurately calibrated in order to control type I (false positive) and type II (false negative) errors (Bender and Lange, 2001). Heuristic tests cannot be adjusted for multiple testing, and thus such studies are susceptible to increased type I and II error rates.

Robust and unbiased p values should be determined in the absence of distributional assumptions by permutation testing. When this principle is applied to module preservation analyses, the module preservation statistics are calculated when shuffling node identifiers in the test dataset to determine their distributions under the null hypothesis. The true value of each statistic is then compared to the empirical null distribution to obtain a permutation test p value. However, at least w permutations are required to estimate significance at a threshold of $1/w$ (Phipson and Smyth, 2010). The analysis of large datasets, together with the concomitant multiple testing correction burden, necessitates increasingly stringent significance thresholds, making permutation-based significance testing computationally challenging.

Here, we address this challenge by developing a rapid and efficient approach for assessing module preservation, available as an R package, NetRep. We use NetRep to create and assess the empirical null distributions of Langfelder et al.'s suite of module preservation statistics when inferring weighted gene coexpression networks in a publicly available resource of mouse adipose, brain, liver, and muscle tissue expression (Yang et al., 2006). We show the majority of these statistics have non-normal distributions and are thus in need of a permutation approach. Next, we demonstrate NetRep's scalability to large-scale module preservation analysis by performing permutation tests to quantify cross-tissue gene coexpression module preservation. We identify and characterize multi-tissue modules associated with mouse body weight. Consequently, we uncover a body weight-associated module with differential adipose and muscle tissue expression. Finally, we explore the broader applicability of Langfelder et al.'s suite of module preservation statistics by using NetRep to quantify the preservation of gut microbial community networks between men and women from publicly available 16S rRNA gene sequence data (Human Microbiome Project Consortium, 2012).

RESULTS

Rapid Module Preservation Analysis

We have developed a time- and memory-efficient method for massively parallel calculation of module preservation statistics. The software is available as an R package, NetRep, which can be downloaded from https://github.com/InouyeLab/NetRep. Implementation details are provided in the Supplemental Experimental Procedures.

To examine the null distributions of the module preservation statistics in an empirical setting, we applied NetRep to publicly available gene expression data for brain, adipose, liver, and muscle tissues from a BxH mouse cross (Yang et al., 2006). From 334 total mice, there were 249 brain, 295 adipose, 306 liver, and 319 muscle tissue samples available for analysis (Experimental Procedures). Figure 1 illustrates the workflow of network construction, module detection, and module preservation analysis. We inferred weighted gene coexpression networks (Experimental Procedures; Zhang and Horvath, 2005) for each tissue, identifying 38, 66, 29, and 32 distinct coexpression modules in the brain, liver, adipose, and muscle tissues, respectively (Figure S1). For a module, we refer to the tissue it was initially identified in as its "discovery" tissue, and other tissues where its preservation is being tested as "non-discovery" tissues.

A runtime comparison of NetRep versus WGCNA's modulePreservation function for calculating permutations for these modules is provided in the Supplemental Experimental Procedures (see also Figure S2). Briefly, NetRep was, on
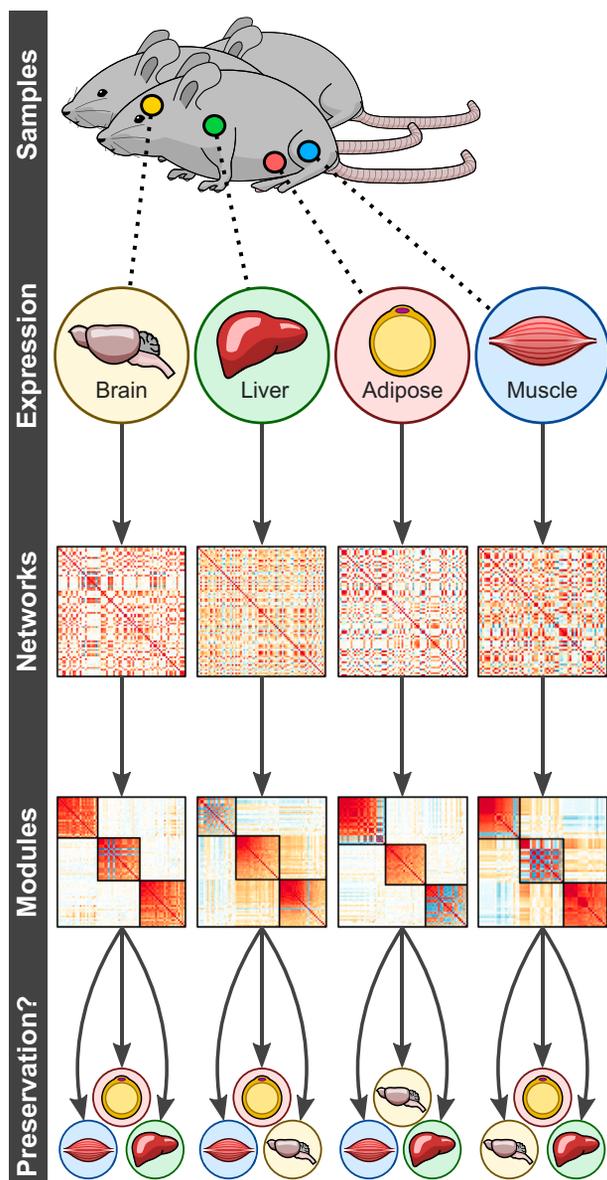
Figure 1. Workflow of Network Construction, Module Detection, and Module Preservation Analysis Workflow on the BxH Mice Tissue Expression

First, the correlation structure (coexpression) between probes was calculated from the gene expression in each tissue. Next the interaction network was inferred and modules were detected using weighted gene coexpression network analysis (WGCNA) (Experimental Procedures). Finally, the topological preservation of each module was assessed in each non-discovery tissue using their respective gene expression, correlation structure, and interaction network matrices.

average, 11 times faster than WGCNA and used less memory when run in parallel. The runtime of NetRep depended on multiple factors, as follows: the number of dataset comparisons, number of permutations for each comparison, sample size, and the sum of the squared sizes of modules included in the analysis. Pairwise comparison of the 165 modules across 4 mouse tissues with 100,000 permutations took approximately 8 days when NetRep was parallelized over 40 cores.

Null Distributions of Module Preservation Statistics

We investigated the normality of the seven module preservation statistics by comparing permutation-based null distribution quantiles to theoretical normal distribution quantiles (Figure S3). For each discovered module, module preservation statistics were calculated on 100,000 random gene sets of identical size in each non-discovery tissue (Experimental Procedures). Across all 38 brain modules, 66 liver modules, 29 adipose modules, and 32 muscle modules, 495 null distributions were generated for each module preservation statistic. We observed strong non-normality of null distributions generated for the average edge weight, density of correlation structure, and concordance of correlation structure statistics (Figure S3). Moderate non-normality was also observed in null distributions for the other four statistics. We also observed increasing non-normality with decreasing module size, particularly for modules of <100 probes (Figure S3). This shows that the assumption of normality required for Z score statistics is often violated.

To determine the consequences of non-normality, we matched Z score p values and permutation p values calculated for each module and preservation test statistic. Substantial inflation of the Z scores and corresponding deflation of p values was observed for average edge weight, density of correlation structure, and concordance of correlation structure (Figure 2). Moderate inflation was also observed for module coherence and average node contribution (Figure 2). These results are consistent with the extremely low p values that motivated heuristic significance thresholds (Langfelder et al., 2011), indicating that a non-parametric approach is necessary to produce unbiased p values.

We further investigated the test statistic with the most extreme deviation from normality, the average edge weight statistic. Edge weights in interaction networks inferred with WGCNA are calculated through a power transform on the absolute correlation coefficient. This power transform acts as a soft-threshold: it penalizes weak correlation coefficients toward zero. The soft-threshold power is chosen under the assumption that the resulting network should be scale free (Experimental Procedures; Zhang and Horvath, 2005). Under this assumption, the distribution of the weighted degree of the network follows an inverse power law where a few hub genes are strongly connected in the network, while most genes are only weakly connected (Barabási and Albert, 1999; Stumpf and Porter, 2012). To test the impact of the scale-free assumption on the null distribution normality, we calculated the average edge weight statistic on 10,000 permutations in the muscle tissue for the 66 liver modules, varying the soft-threshold exponent used to define the edge weights. We observed a trend toward non-normality as the exponent increased, indicating that the scale-free assumption contributes to non-normality for this statistic (Figure S4). We similarly generated null distributions for the concordance of weighted degree as it is also calculated from the edge weights in the interaction network; however, its deviation from normality was only mild and did not change as the exponent increased (Figure S5). These analyses indicate that the non-normality of preservation test statistics can be influenced by the distribution of node degree.

To assess the performance of NetRep and the permutation approach for quantifying module preservation, we simulated a
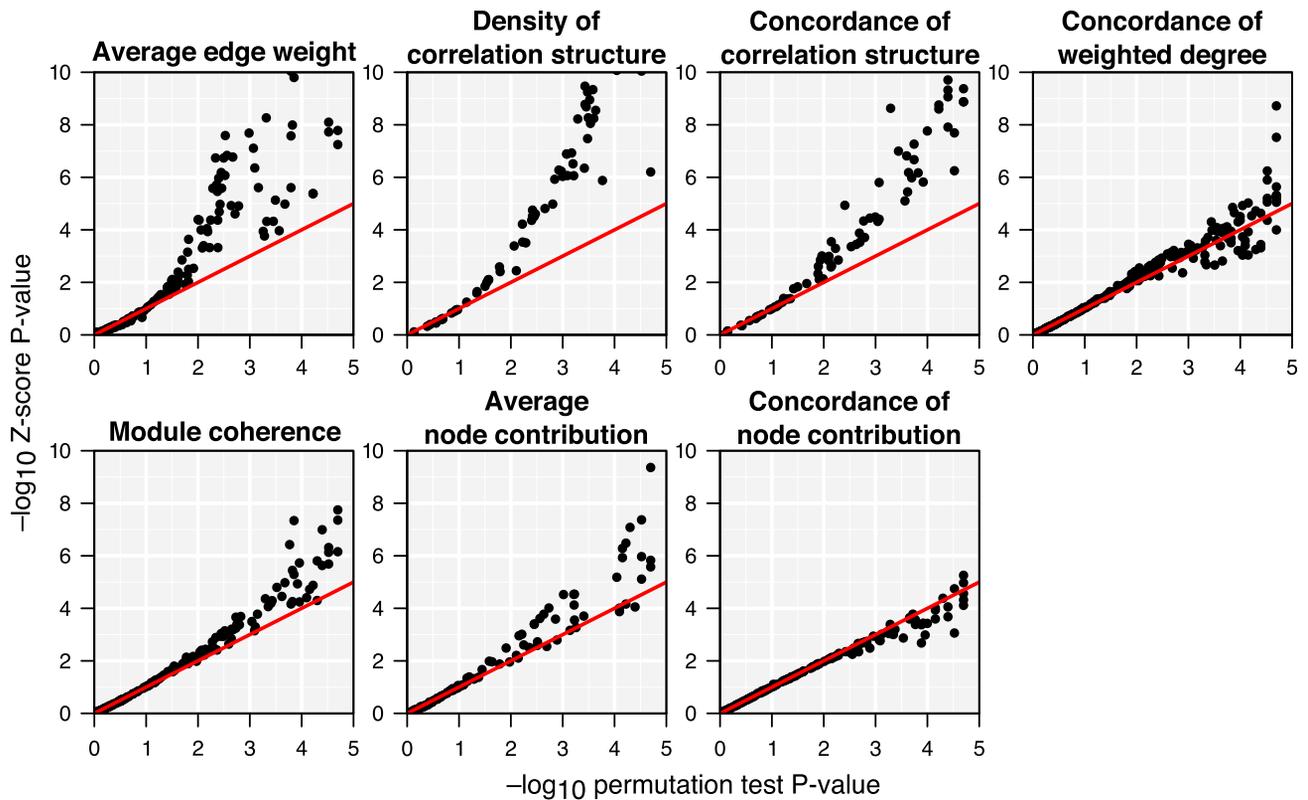
Figure 2. Comparison of Permutation p Values to Z Test p Values for Each of the 165 BxH Mice Modules when Tested in Their Three Non-discovery Tissues

The mean and standard deviation of the 495 null distributions were used to calculate the Z test p values. Null distributions were generated from 100,000 permutations. p values were plotted on a −log10 scale. Tests where the permutation test p value was incomparable to the Z test p values (2,229 of 3,465 tests where permutation test $p \leq 1 \times 10^{-5}$, the smallest permutation p value possible with 100,000 permutations) are not shown. Z tests with $p < 1 \times 10^{-10}$ are not plotted (25 data points for the average edge weight statistic with a minimum $p = 1 \times 10^{-63}$, 11 data points for the concordance of correlation structure statistic with a minimum $p = 1 \times 10^{-19}$, and 21 data points for the density of correlation structure statistic with a minimum $p = 1 \times 10^{-28}$).

discovery gene expression dataset containing negative and positive control modules of varying sizes and three test datasets with varying noise levels (Supplemental Experimental Procedures). Positive control modules were simulated to have identical topology in each test dataset, while negative control modules were simulated as random. We estimated permutation p values for each simulated module in each test dataset using NetRep. In total, we performed 3,000 tests for each module preservation statistic (10 modules × 3 test datasets × 100 simulations), estimating permutation p values with 10,000 permutations per test. At a significance threshold of $p \leq 1 \times 10^{-4}$ (the smallest possible p value that can be obtained from 10,000 permutations; Supplemental Experimental Procedures), NetRep was able to successfully detect preservation of positive control modules while being robust to false positives (Figure S6). Sensitivity varied by statistic but decreased as noise increased or module size decreased (Figure S6). The module preservation statistics were nearly always robust to false positives, with the exception of the module coherence statistic, which falsely detected preservation for large negative control modules (≥500 genes) in the presence of low and medium levels of simulated noise (Figure S6). These results indicate that NetRep is sensitive and can distinguish between true and false positives under most conditions.

Cross-Tissue Module Preservation in Mouse Transcriptomic Data

We next examined the preservation of each discovered module in other tissues by evaluating the permutation p values for each of the 495 null distributions (Experimental Procedures). We defined strong evidence for a module's preservation in another tissue as all test statistics achieving p < 0.0001, weak evidence if one or more, but not all, test statistics were p < 0.0001, and no evidence if no test statistics are p < 0.0001. The significance threshold of 0.0001 was chosen to Bonferroni adjust for the 495 tests performed for each preservation statistic. Figure 3 provides a summary view of the cross-tissue module preservation in the BxH mice, and Figure 4 shows the preservation evidence for each module in each non-discovery tissue.

We observed widespread preservation for modules in all four tissues (Figures 3A and 3B). 85 of 165 modules (52%) had strong evidence of preservation in at least one other non-discovery tissue, and 41 modules (25%) had strong evidence of preservation in all non-discovery tissues (Figure 3A). In contrast, only 21 modules (13%) had no evidence for preservation in any other tissue, suggesting tissue specificity of these modules (Figures 3A and 3C). In comparing NetRep results with those obtained through summary Z score and heuristic thresholds (Supplemental
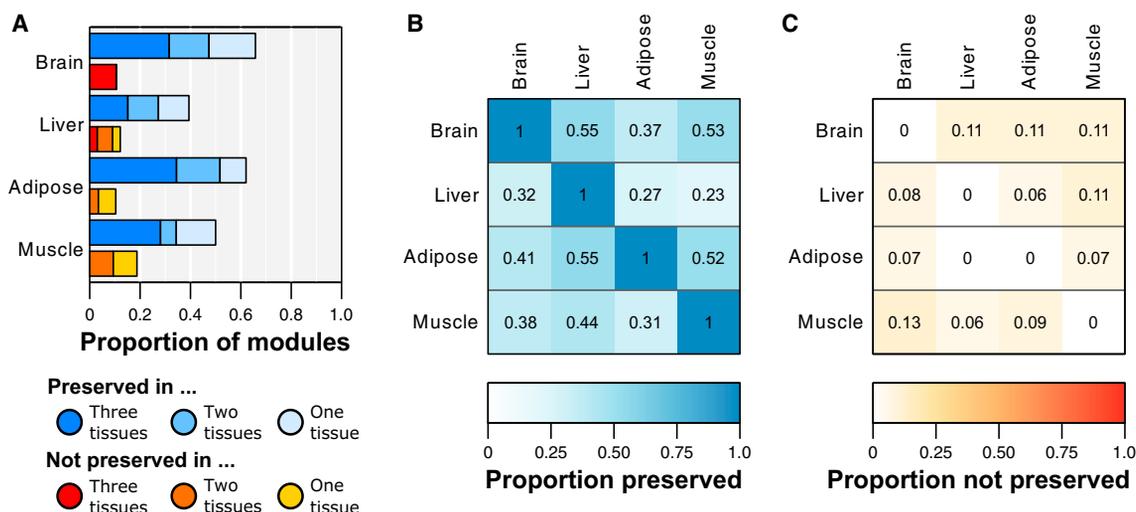
Figure 3. Summary of BxH Mice Cross-Tissue Module Preservation
(A) Summary of preservation evidence for BxH mice modules discovered in each tissue. Here, a module was considered preserved if it had strong evidence of preservation in another tissue, and not preserved if it had no evidence of preservation in another tissue.
(B) Tissue similarity based on the proportion of modules discovered in the row tissue with strong evidence of preservation in the column tissue.
(C) Tissue uniqueness based on the proportion of modules discovered in the row tissue with no evidence of preservation in the column tissue. Note that the heatmap entries in (B) and (C) cannot be read vertically.

Experimental Procedures; Langfelder et al., 2011), the two approaches mostly obtained similar levels of evidence for preservation (Table S1). However, differences in preservation were observed for 120 of the 495 (24%) module preservation tests. In terms of module preservation, 55 (24%) of the modules found to be strongly preserved by heuristics were classified as weakly preserved by NetRep. Similarly, 44 (54%) of the modules found to be not preserved by heuristic were classified as weakly preserved by NetRep (Table S1). Therefore, NetRep was more stringent in the evidence required to call a module either strongly preserved or not preserved.

The liver had the lowest proportion of modules preserved in at least one other tissue; however, it had twice as many modules in total than any other tissue. Many of these were small (<100 genes) and had only weak evidence for preservation (Figure 4). Only the brain and liver had any modules with no evidence for preservation in all three non-discovery tissues (Figures 3 and 4). These results were broadly consistent with recent results from the GTEx consortium, who observed high similarity between coexpression network modules across nine human tissues (including adipose and muscle tissues) (GTEx Consortium, 2015).

In total, NetRep found that 41 modules (10 discovered in adipose tissue, 12 in brain, 10 in liver, and 9 in muscle) were preserved in all non-discovery tissues. Analysis of Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways for each module (Supplemental Experimental Procedures) showed that these were putative housekeeping modules, which were most frequently enriched for genes involved in translation, and to a lesser extent transcription and basic cellular functions, e.g., cell cycle, apoptosis, and DNA repair (Table S2). The putative housekeeping modules were most frequently enriched for genes coding for ribosomal proteins with 10 of 41 putative housekeeping modules enriched for the Ribosome pathway in KEGG (Table S2).

Multi-tissue Weight-Associated Modules

The BxH mice were bred to enhance differentiation of cardiovascular disease risk traits such as obesity and circulating lipids (Yang et al., 2006). Previous coexpression network analysis of the BxH mice focused on the identification of modules associated with mouse weight (Chen et al., 2008; Ghazalpour et al., 2006). We asked whether modules with strong evidence of preservation in more than one tissue ("multi-tissue" modules) were associated with obesity. We therefore tested each multi-tissue module's summary expression (first principal component; Experimental Procedures) for an association with mouse weight in any tissue for which there was strong evidence for its preservation (Supplemental Experimental Procedures). Significant association with weight was defined as $p < 0.0001$ (Bonferroni correction). Each regression model was adjusted for sex due to its strong effect on both mouse weight and gene expression (Fuller et al., 2007; Ghazalpour et al., 2006; Yang et al., 2006).

Of the 85 multi-tissue modules, 43 modules were significantly associated with mouse body weight in either their discovery tissue or in a tissue where it was strongly preserved, comprising 57 body weight associations in total (Table S3). Twenty-seven (32%) of these multi-tissue modules were also putative housekeeping modules. Weight was most frequently associated with modules in adipose tissue (28 of 57 associations) and liver tissue (24 of 57 associations). Notably, there were many cases where multi-tissue modules were not associated with weight in the tissue they were identified in, but displayed a significant association in one or more of the tissues in which they were preserved (Table S3). In total, 13 multi-tissue modules were associated with mouse weight in multiple tissues (Table S3). Notably, we observed different directions of weight association across tissues for five modules: i.e., in tissue A, an increase in weight was associated with a decrease in module summary expression, but in tissue B an increase in weight was associated with an
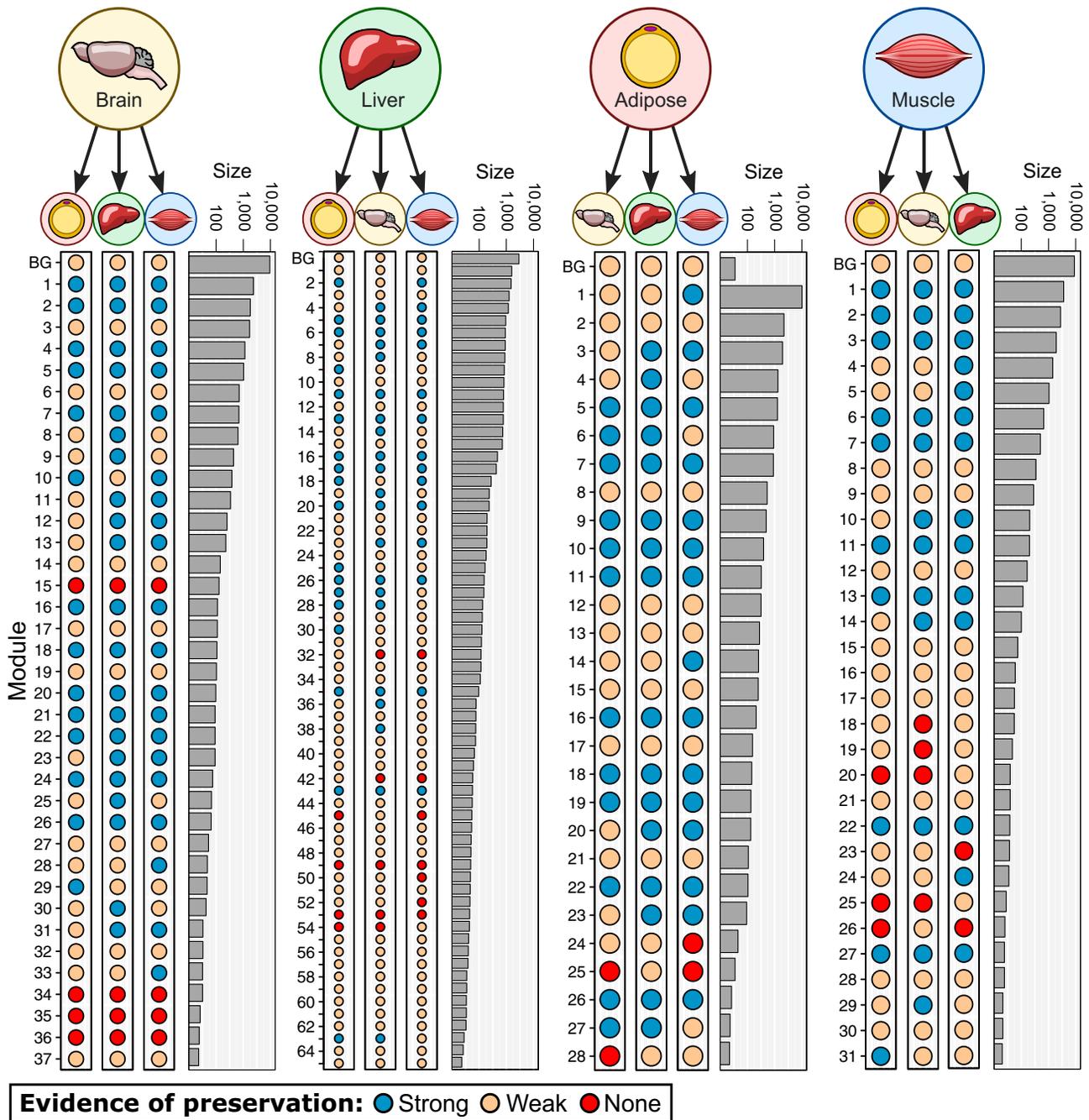
Figure 4. Evidence for Preservation for Each BxH Mice Module in Each Non-discovery Tissue
Module sizes (horizontal bar plots) are shown on a $\log_{10}$ scale. Modules (dots) are sorted according to module size and colored according to preservation evidence ("strong" evidence is blue, "weak" is yellow, "none" is red. "BG": the background module, which contains all nodes that did not cluster into a coexpression module in that tissue.).

increase in module summary expression or vice versa (Table S3). For these five modules, visualization of the network topology indicated the weight-associated differential summary expression reflected differential whole-module expression for two modules: liver module 35 and brain module 20. We focus on liver module 35 (LM35) for subsequent investigation (Figure 5).

LM35 is a putative housekeeping module consisting of 99 genes (permutation test $p \leq 1 \times 10^{-5}$ for all statistics in the brain, adipose, and muscle tissues). Consistent with previous analysis of the putative housekeeping modules, GO term and KEGG pathway enrichment indicated LM35 was primarily enriched for ribosomal genes involved in translation (Table S4). While a majority of probes in LM35 were specific to the custom microarray design and thus lacking gene annotation, 17 of its 24 annotated genes coded for ribosomal proteins (Table S5). Increased body weight was associated with increased LM35
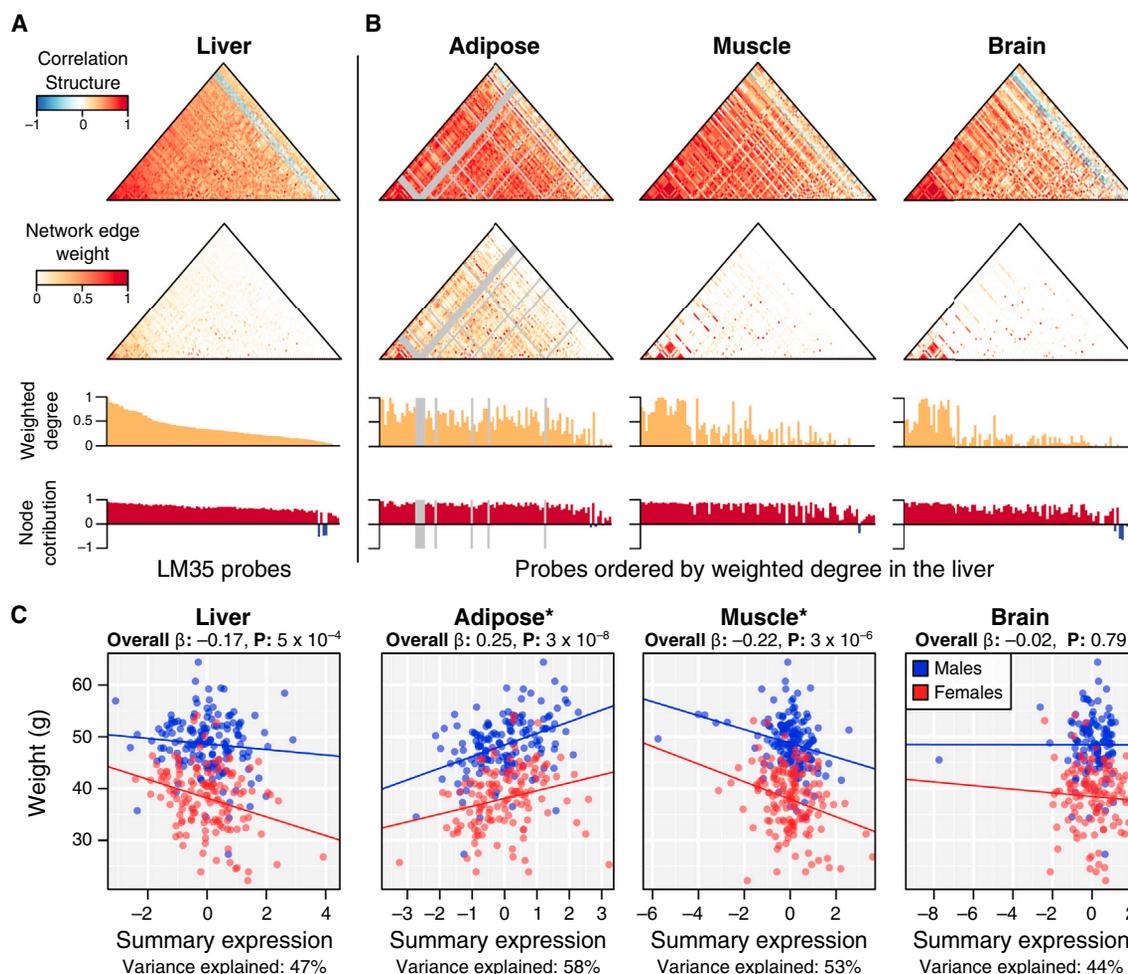
Figure 5. BxH Mice Liver Module 35

(A) Network topology in the liver (discovery tissue). From top to bottom: heatmap of the correlation structure (Pearson correlation), heatmap of the interaction network edge weights, normalized weighted degree (calculated within the module and normalized by the maximum value), and node contribution (Pearson correlation between each probe and the summary expression profile). Probes are ordered by descending order of weighted degree.

(B) Network topology in adipose, muscle, and brain tissues. Probes are ordered as in the liver tissue. Grey bars denote probes either missing, or not passing quality control, in the respective tissue.

(C) Scatter plots of standardized LM35 summary expression versus body weight. Points on the scatter plot are colored by sex (males in blue, females in red), and linear regression models were adjusted for gender (lines shown are fitted within genders). An "*" next to the tissue name indicates significant weight association in Table S3. Models were robust to outliers (mice with summary expression < −3 SD). Variance explained indicates the proportion of variance in liver module 35 (LM35) expression explained by the summary expression vector in each tissue (i.e., the module coherence).

expression in adipose tissue ($p = 3 \times 10^{-8}$) and decreased LM35 expression in muscle tissue ($p = 3 \times 10^{-6}$) (Figure 5). Consistent with this, its summary expression profile was negatively correlated between the adipose and muscle tissues (Pearson's $\rho = -0.13$) and the expression of 64 of its 99 probes were negatively correlated across the two tissues, suggesting that the relationship between body weight and LM35 was tissue specific—genes associated with weight were simultaneously upregulated in the adipose tissue and downregulated in the muscle tissue.

We subsequently tested 20 other cardiometabolic traits for association with LM35 expression in adipose and muscle tissues (Table S6). Consistent with the direction of the weight associations, increased insulin, total cholesterol, and total fat were associated with increased adipose expression and decreased

muscle expression (false discovery rate [FDR] q < 0.025; Table S6). These changes in LM35 expression were also associated with a decrease in the ratio of glucose over insulin (Table S6). Increased LM35 adipose expression was associated with increased glucose, other fat, body length, and monocyte chemotactic protein-1 (MCP-1) (Table S6). On the other hand, decreased LM35 muscle expression was associated with increased abdominal fat, free fatty acids, total cholesterol, and LDL+VLDL, but a decreased ratio of HDL to LDL+VLDL (Table S6). Our findings indicate the tissue specificity of LM35 function and its relationships with phenotypes.

Overall, these analyses highlight that NetRep can be used to determine whether the relationships between genes are preserved, but separate investigation of preserved modules is required to determine whether module function is preserved. In

the case of multi-tissue analyses, this may elucidate differential inter-tissue module regulation.

Preservation of Gut Microbial Community Networks

To demonstrate the broader applicability of NetRep, we inferred microbial community networks in gut samples of 62 healthy adult men and 65 healthy adult women from the Human Microbiome Project (HMP) Consortium (Human Microbiome Project Consortium, 2012). The nodes in these networks corresponded to operational taxonomic units (OTUs), and we generated OTU networks with the commonly utilized SparCC approach (Friedman and Alm, 2012) (Experimental Procedures). From 152 distinct OTUs, we identified 17 and 21 communities of co-occurring OTUs in the male and female gut samples, respectively (Figures 6A and 6B). Using NetRep, we subsequently tested the preservation of the male gut communities in the female network and vice versa. Permutation p values were estimated from null distributions drawn from 10,000 permutations of OTU labeling in the respective test datasets (Figure 6C). We considered each module preservation statistic significant where p < 0.001 to Bonferroni adjust for the 38 tests performed for each statistic.

Unlike weighted gene coexpression networks, where all individuals in a population have more or less the same genes, OTU networks are relatively sparse due to the variable presence/absence of microbial taxa in the gut. Module sizes in OTU networks were also substantially smaller (range: 2–12 nodes). Thus, in applying the module preservation statistics to OTU networks, it was clear that some statistics would be more appropriate than others. We found that concordance of node contribution, concordance of correlation structure, and concordance of weighted degree statistics were not suitable for assessing preservation of these OTU modules. In addition to their small size, OTU modules tended to have uniform structure across nodes in terms of their SparCC correlation coefficient, node contribution, and weighted degree. This led to low values for these statistics in cases where the node contribution, SparCC correlation coefficient, and weighted degree were high across all nodes, due to dramatic changes in node rank caused by tiny variations in these values. Further, these module preservation statistics could not be evaluated where the node contribution, SparCC correlation coefficient, and weighted degree were identical for all nodes in a module. This always occurred for modules composed of two OTUs (Figure 6C), for which the weighted degree was always identical for both nodes and there was only one SparCC correlation coefficient. The sparsity of the network also meant the concordance of weighted degree could often not be calculated. This occurred where a module had no edges between any nodes in the test network (e.g., male module 7 in the female gut samples; Figure 6), which occurred frequently when generating null distributions for all modules, reducing the power of the permutation tests.

Therefore, in applying the module preservation statistics and NetRep to sparse networks and small modules, we recommend assessing module coherence, average node contribution, density of correlation structure, and average edge weight. Using these four statistics, we defined strong evidence for module preservation where all four were significant (p < 0.001), weak evidence if one or more were significant, and no evidence if none of these four were significant. Ignoring modules composed of only two OTUs, for which obvious false negatives were prevalent, we observed widespread preservation of microbial communities found in the men's gut samples, with 50% (4 of 8) of their OTU modules having strong evidence of preservation in the women's gut samples (Figure 6D). The women's gut microbial communities were less preserved, with 20% (2 of 10) of their OTU modules having strong evidence of preservation in the men's gut samples (Figure 6D). However, four of six women's gut OTU modules that had weak evidence of preservation (modules 2, 3, 9, and 10) were almost identical to OTU modules identified in the men's gut samples, suggesting comparative levels of preservation between women's and men's gut microbial communities.

DISCUSSION

Accurate and unbiased assessment of the replication and preservation of network modules requires permutation testing of network feature similarity. However, the current approach employs heuristics to assess significance due to the computational burden of these calculations (Langfelder et al., 2011). While heuristics may be appropriately employed for a small number of modules, the scale of module preservation and replication analyses now requires a rapid and statistically rigorous method to enable adjustment for multiple hypothesis testing, consequently allows confident investigation of the underlying biology. In this study, we have empirically shown that module preservation statistics are typically non-normal under the null hypothesis of non-preservation and thus have developed a rapid and efficient approach for assessing module preservation through permutation testing: NetRep.

In addition to assessment of reproducibility, module preservation analysis can be used to ask questions about conserved biological interactions and functions across spatial locations or species (Langfelder et al., 2011). Application of NetRep to a multi-tissue gene expression dataset showed widespread preservation of gene coexpression network modules across brain, adipose, liver, and muscle tissues in a BxH mouse cross. Housekeeping modules, those preserved in all four tissues, were enriched for genes involved in basic cellular processes, most notably ribosomal genes involved in translation. Subsequent investigation of multi-tissue modules associated with body weight revealed that preserved modules can exhibit differential intramodule expression across tissues, and we have identified a housekeeping module linked to obesity and insulin resistance with increased adipocyte expression and decreased muscle expression in overweight mice.

Previous studies have identified multi-tissue modules driving obesity in mice and humans, with concordant expression across tissues (Chen et al., 2008; Emilsson et al., 2008). Here, we found that multi-tissue modules may be differentially expressed across tissues with corresponding phenotypic differences. The liver module LM35 exhibited negative, positive, and negative associations with body weight in liver, adipose, and muscle tissues, respectively. Perhaps consistent with its tissue-specific directions of body weight association, LM35 was enriched for genes encoding ribosomal proteins, which maintain putative housekeeping functions. However, the gene set comprising LM35
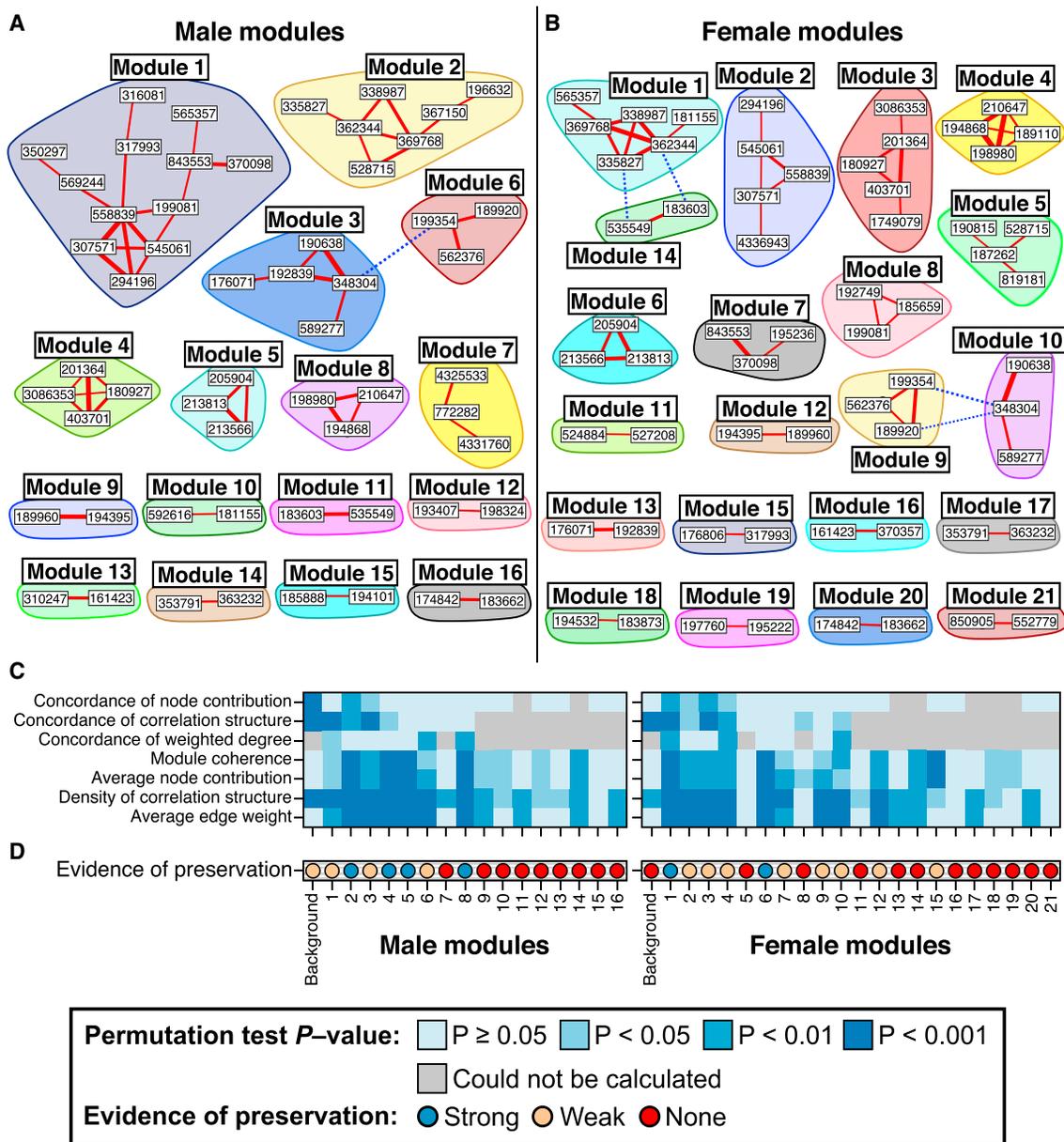
Figure 6. Preservation of Gut Microbial Communities across Males and Females Participating in the Human Microbiome Project

(A and B) Microbial communities inferred from the male (A) and female (B) gut samples. Nodes correspond to operational taxonomic units (OTUs) and numeric labels indicate their unique identifier. The colored shapes drawn around OTUs indicate community (module) assignment. Edge widths indicate strength of the correlation coefficient and color and linetype indicate positive (red, solid line) or negative (blue, dashed line) coefficients. Edge weights were defined as the absolute value of the correlation coefficient for the purpose of module detection (Experimental Procedures). Note that many OTUs present in male modules are also present in female modules and vice versa, and some male and female modules overlap.

(C) Heatmaps of permutation p values when assessing module preservation in the other sex. Permutation p values were estimated from null distributions generated from 10,000 permutations. Grey cells indicate tests where the permutation p value could not be calculated.

(D) Evidence of preservation for each OTU module in the other sex ("strong" evidence is blue, "weak" is yellow, "none" is red). See Table S7 for the taxonomic assignments of OTUs participating in modules with strong evidence of preservation.

was the only multi-tissue housekeeping module that exhibited significant patterns of differential body weight association. Furthermore, LM35 was associated with several obesity related traits, including a decreased ratio of glucose over insulin, suggesting an association with decreased insulin sensitivity. The link between insulin sensitivity, obesity, and adipocytes is well

established (Hotamisligil et al., 1993; Kahn and Flier, 2000; Kahn et al., 2006), and consistent with this link, the adipose expression of LM35 was associated with circulating MCP-1 levels. MCP-1 has been shown to be secreted by adipocyte cells as well as overexpressed in obese mice, and it has been shown to decrease insulin-stimulated glucose uptake in vitro (Kanda

et al., 2006; Sartipy and Loskutoff, 2003). The phenotypic associations of LM35 across tissues may be explained by possible coexpression with obesity-linked genes in the adipose and muscle tissue, which did not coexpress with the module in the liver tissue where the module was identified.

We also showed that NetRep can be successfully applied to OTU networks derived for 16S microbiome data and have offered recommendations for dealing with the relative sparsity of these networks. In doing so, we identified several gut OTU modules that were preserved between men and women in the HMP data. Consistent with expectation, preserved modules largely involved multiple OTUs from the same genus (e.g., *Dialister*, *Bacteroides*, and *Ruminococcus* modules). A more diverse module comprising OTUs from the *Clostridiales* order, particularly *Faecalibacterium prausnitzii*, *Coprococcus*, *Butyrivibrio*, and *Clostridium*, was also preserved (male module 2, female module 5). *F. prausnitzii* has been linked to various human diseases, including inflammatory bowel disease, celiac disease, and obesity, and has been the subject of intense research to understand its specific functions, both individually and as part of communities, in the human gut (Miquel et al., 2013). Our analyses suggest that *F. prausnitzii* is part of a broader preserved *Clostridiales* community that may have functional consequences. Further studies in larger sample sizes may offer more power to detect additional members of this community, its variation across sexes, and its relevance to disease. Identification of gut microbial communities that change in composition between sexes may offer insight into diseases, such as irritable bowel syndrome, which have different prevalences in males and females (Canavan et al., 2014; Kassinen et al., 2007).

In recent years, studies have begun generating and analyzing datasets containing gene expression measured in dozens of tissues and cell types, for example, the Genotype-Tissue Expression (GTEx) Consortium (GTEx Consortium, 2015), the Immunological Genome (ImmGen) (Shay and Kang, 2013), and the Immune Variation (ImmVar) projects (De Jager et al., 2015). Similar scale studies are investigating microbiota spatiotemporally and in conjunction with other -omics data (Alivisatos et al., 2015; Human Microbiome Project Consortium, 2012; Integrative HMP (iHMP) Research Network Consortium, 2014). Already, multiple module preservation analyses have been performed on the GTEx pilot data (Melé et al., 2015; Pierson et al., 2015; GTEx Consortium, 2015), and here we have performed an initial preservation analysis of microbiome network modules between men and women. With large-scale expression studies increasing in scale and complexity, and the emergence of other types of datasets of similar scale, there is an urgent need for powerful and accurate statistical methodologies that quantify module replication and preservation. Here, we have presented an approach for rapid assessment of network module preservation and reproducibility that makes possible unbiased large-scale comparative analysis.

EXPERIMENTAL PROCEDURES

Full experimental procedures and data details can be found in the Supplemental Experimental Procedures. For the Humap Microbiome Project, details of institutional review boards are given in Human Microbiome Project Consortium (2012). For the mouse data, these are given in Yang et al. (2006).

Network Inference and Module Detection
Network inference and module detection were performed on a per-tissue basis for the BxH mouse cross using WGCNA v.1.43.1 with the default parameters (Langfelder and Horvath, 2008). First, the correlation structure (coexpression) for each tissue was calculated as the Pearson correlation coefficient between all probes passing quality control (Supplemental Experimental Procedures). Next, the network of interactions between probes was constructed by taking the element-wise absolute value of the correlation coefficient and exponentiating it to the smallest power such that the distribution of the weighted node degree (i.e., the sum of all edge weights for each node) of the resulting network was approximately scale free (scale-free topology criterion $R^2 > 0.85$) (Zhang and Horvath, 2005). This results in a dense, complete network where edge weights can take values between 0 and 1, most pairs of probes are connected with extremely small edge weights, and the comparatively few strongly correlated probes are connected with strong edge weights. The automated selection procedure selected the exponents of 12, 5, 4, and 12 for the brain, liver, adipose, and muscle tissues, respectively (Figure S7). Subsequently, the topological overlap dissimilarity (Zhang and Horvath, 2005) between probes was calculated and hierarchically clustered using the average linkage method. Hierarchically nested modules were identified from the results dendrogram using the dynamic tree cut algorithm with default parameters (Langfelder et al., 2008). Similar modules were merged together using an iterative process in which modules whose summary expression profile (first principal component, see below) clustered together (hierarchical average linkage) below a height of 0.2 were joined.

Network inference and module detection were performed separately for the HMP male and female gastrointestinal samples. First, 16S rDNA reads were clustered by sequence similarity ($\geq 97\%$) to representative sequences with known taxonomic assignments (Supplemental Experimental Procedures). Subsequent OTU tables were filtered to gastrointestinal samples collected on the first visit for 127 individuals. Next, the correlation structure between the 152 non-rare OTUs (Supplemental Experimental Procedures) was calculated using SparCC, a method for calculating unbiased correlation coefficients in sparse, compositional data (Friedman and Alm, 2012). The interaction network between OTUs was defined as the magnitude of the correlation where the SparCC correlation coefficient was significant at $p < 0.005$ in a bootstrap test. Modules were subsequently defined as groups of OTUs connected with significant positive SparCC correlation coefficients. The bootstrap p values were calculated using the estimator described by Phipson and Smyth (2010) (Supplemental Experimental Procedures), and the threshold $p < 0.005$ was chosen as it provided the best separation of OTUs into distinct modules for testing with NetRep.

Module Preservation
Seven statistics were used to quantify whether the relationships and correlation structure between nodes composing each module were replicated or preserved when measured in a different dataset (Langfelder et al., 2011). Here, we have renamed the statistics so that they are accessible to a wider audience and meaningful when applied to networks inferred from data sources (Table 1). Each module preservation statistic—their biological interpretation, application to different data sources, and network inference methods—are discussed in the Supplemental Experimental Procedures.

A permutation procedure was employed to characterize the distribution of each statistical test under the null hypothesis of non-replication and non-preservation. Specifically, each module preservation statistic was re-calculated when shuffling the node labels in the test dataset. The node labels in the discovery dataset were left unchanged. Nodes that were not present in both the discovery and test dataset were ignored both when calculating the module preservation statistics and when shuffling the node labels in the test dataset. Under the alternate hypothesis of replication/preservation, the test statistics calculated on the non-permuted dataset were expected to be higher than when calculated on random sub-graphs in the test dataset. Permutation p values were then calculated from these null distributions using the estimator described by Phipson and Smyth (2010), which provides a conservative estimate of the p value appropriate for multiple testing adjustment (Supplemental Experimental Procedures).

Module Summary Profiles
The summary profile for each module was calculated as the first principal component of module *w*. Specifically, each summary profile was calculated

Table 1. Definitions of the Module Preservation Statistics

NetRep Name	WGCNA Name	Definition
Module coherence	proportion of variance explained	$mean((cor(g_i^{[t](w)}, Eig_1^{[t](w)}))^2)$
Average node contribution	mean sign-aware module membership	$mean(sign(cor(g_i^{[d](w)}, Eig_1^{[d](w)})) \cdot cor(g_i^{[t](w)}, Eig_1^{[t](w)}))$
Concordance of node contributions	correlation of module membership	$cor(cor(g_i^{[d](w)}, Eig_1^{[d](w)}) , cor(g_i^{[t](w)}, Eig_1^{[t](w)}))$
Density of correlation structure	mean sign-aware coexpression	$mean(sign(C^{[d](w)}) \cdot C^{[t](w)})$
Concordance of correlation structure	correlation of coexpression	$cor_{i \neq j}(C^{[d](w)}, C^{[t](w)})$
Average edge weight	mean adjacency	$mean_{i \neq j}(a_{ij}^{[t](w)})$
Concordance of weighted degree	correlation of intramodular connectivities	$cor((\sum_{i \neq j}^{j} a_i)^{[d](w)}, (\sum_{i \neq j}^{j} a_i)^{[t](w)})$

The NetRep name indicates the name of the statistic in the main text, while the WGCNA name indicates the name given to the statistics by Langfelder et al. (2011). Mathematical symbols are as follows: for n variables measured across m samples, G refers to the $m \times n$ matrix of observations, C refers to the $n \times n$ square matrix containing the pairwise correlation coefficients between variables, and A refers to the $n \times n$ square adjacency matrix denoting the connection strength (edge weight) between each pair of variables (nodes). Lowercase g, c, and a refer to individual elements of the matrices denoted by their respective uppercase letter. The superscripts [d] and [t] indicate whether the respective entity, formula, or network is obtained/calculated from the discovery or test dataset, respectively. The subscript letters i and j denote individual variables/nodes in module w. The superscript (w) indicates that the entity/formula that it is attached to is obtained/calculated on all nodes j (or all pairs of nodes i,j) in module w. For example, $g_i^{[t](w)}$ denotes a vector of observations for node i (which belongs to module w) in the test dataset, and $a_{ij}^{[d](w)}$ indicates the edge weight between nodes i and j (both of which belong to module w) in the discovery dataset. $Eig_1^{(w)}$ refers to the summary profile of the module w (first principal component; Experimental Procedures). The $sign$ function evaluates to 1 if its argument is a positive value or -1 if its argument is a negative value. The cor function calculates Pearson's correlation coefficient between two vectors.

as the first eigenvector of a singular value decomposition of $G^{(w)}$. Two solutions exist for every eigenvector; both contain the same values, but with opposite signs. Thus, the summary profile is chosen as the eigenvector that is positively correlated with the average of $G^{(w)}$ across samples. This ensures that the eigenvector is oriented in the same direction as the data.

For interpretability, we refer to the summary profile as the "summary expression" profile when calculated on the BxH mice gene coexpression network modules, although this vector is commonly referred to as the "module eigengene" in the weighted gene coexpression network literature (Langfelder and Horvath, 2008). For the HMP gut community modules, we refer to the summary profiles as the community "summary abundance."

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.cels.2016.06.012.

AUTHOR CONTRIBUTIONS

M.I. and S.C.R. conceived and designed the study. S.C.R., S.W., G.A., L.G.F., K.E.H., and M.I. performed analyses. S.W. and K.E.H. contributed data and methods. S.C.R. and M.I. wrote the paper with input from all authors.

REFERENCES

Abraham, G., Bhalala, O.G., de Bakker, P.I.W., Ripatti, S., and Inouye, M. (2014). Towards a molecular systems model of coronary artery disease. Curr. Cardiol. Rep. 16, 488.

Alivisatos, A.P., Blaser, M.J., Brodie, E.L., Chun, M., Dangl, J.L., Donohue, T.J., Dorrestein, P.C., Gilbert, J.A., Green, J.L., Jansson, J.K., et al.; Unified Microbiome Initiative Consortium (2015). MICROBIOME. A unified initiative to harness Earth's microbiomes. Science 350, 507–508.

Barabási, A.-L., and Albert, R. (1999). Emergence of scaling in random networks. Science 286, 509–512.

Barabási, A.-L., Gulbahce, N., and Loscalzo, J. (2011). Network medicine: a network-based approach to human disease. Nat. Rev. Genet. 12, 56–68.

Bender, R., and Lange, S. (2001). Adjusting for multiple testing–when and how? J. Clin. Epidemiol. 54, 343–349.

Boyle, A.P., Araya, C.L., Brdlik, C., Cayting, P., Cheng, C., Cheng, Y., Gardner, K., Hillier, L.W., Janette, J., Jiang, L., et al. (2014). Comparative analysis of regulatory information and circuits across distant species. Nature 512, 453–456.

Cai, C., Langfelder, P., Fuller, T.F., Oldham, M.C., Luo, R., van den Berg, L.H., Ophoff, R.A., and Horvath, S. (2010). Is human blood a good surrogate for brain tissue in transcriptional studies? BMC Genomics 11, 589.

Canavan, C., West, J., and Card, T. (2014). The epidemiology of irritable bowel syndrome. Clin. Epidemiol. 6, 71–80.

Carlson, M.R.J., Zhang, B., Fang, Z., Mischel, P.S., Horvath, S., and Nelson, S.F. (2006). Gene connectivity, function, and sequence conservation: predictions from modular yeast co-expression networks. BMC Genomics 7, 40.

Chen, Y., Zhu, J., Lum, P.Y., Yang, X., Pinto, S., MacNeil, D.J., Zhang, C., Lamb, J., Edwards, S., Sieberts, S.K., et al. (2008). Variations in DNA elucidate molecular networks that cause disease. Nature 452, 429–435.

Collins, F.S., and Tabak, L.A. (2014). Policy: NIH plans to enhance reproducibility. Nature 505, 612–613.

Dagan, T. (2011). Phylogenomic networks. Trends Microbiol. 19, 483–491.

De Jager, P.L., Hacohen, N., Mathis, D., Regev, A., Stranger, B.E., and Benoist, C. (2015). ImmVar project: Insights and design considerations for future studies of "healthy" immune variation. Semin. Immunol. 27, 51–57.

Emilsson, V., Thorleifsson, G., Zhang, B., Leonardson, A.S., Zink, F., Zhu, J., Carlson, S., Helgason, A., Walters, G.B., Gunnarsdottir, S., et al. (2008). Genetics of gene expression and its effect on disease. Nature 452, 423–428.

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

Friedman, J., and Alm, E.J. (2012). Inferring correlation networks from genomic survey data. PLoS Comput. Biol. *8*, e1002687.

Fuller, T.F., Ghazalpour, A., Aten, J.E., Drake, T.A., Lusis, A.J., and Horvath, S. (2007). Weighted gene coexpression network analysis strategies applied to mouse weight. Mamm. Genome *18*, 463–472.

Gerstein, M.B., Rozowsky, J., Yan, K.-K., Wang, D., Cheng, C., Brown, J.B., Davis, C.A., Hillier, L., Sisu, C., Li, J.J., et al. (2014). Comparative analysis of the transcriptome across distant species. Nature *512*, 445–448.

Ghazalpour, A., Doss, S., Zhang, B., Wang, S., Plaisier, C., Castellanos, R., Brozell, A., Schadt, E.E., Drake, T.A., Lusis, A.J., and Horvath, S. (2006). Integrating genetic and network analysis to characterize genes related to mouse weight. PLoS Genet. *2*, e130.

GTEx Consortium (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science *348*, 648–660.

Gustafsson, M., Nestor, C.E., Zhang, H., Barabási, A.-L., Baranzini, S., Brunak, S., Chung, K.F., Federoff, H.J., Gavin, A.-C., Meehan, R.R., et al. (2014). Modules, networks and systems medicine for understanding disease and aiding diagnosis. Genome Med. *6*, 82.

Hawrylycz, M.J., Lein, E.S., Guillozet-Bongaarts, A.L., Shen, E.H., Ng, L., Miller, J.A., van de Lagemaat, L.N., Smith, K.A., Ebbert, A., Riley, Z.L., et al. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. Nature *489*, 391–399.

Horvath, S., and Dong, J. (2008). Geometric interpretation of gene coexpression network analysis. PLoS Comput. Biol. *4*, e1000117.

Hotamisligil, G.S., Shargill, N.S., and Spiegelman, B.M. (1993). Adipose expression of tumor necrosis factor-alpha: direct role in obesity-linked insulin resistance. Science *259*, 87–91.

Human Microbiome Project Consortium (2012). A framework for human microbiome research. Nature *486*, 215–221.

Integrative HMP (iHMP) Research Network Consortium (2014). The Integrative Human Microbiome Project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease. Cell Host Microbe *16*, 276–289.

Jeong, H., Mason, S.P., Barabási, A.L., and Oltvai, Z.N. (2001). Lethality and centrality in protein networks. Nature *411*, 41–42.

Kahn, B.B., and Flier, J.S. (2000). Obesity and insulin resistance. J. Clin. Invest. *106*, 473–481.

Kahn, S.E., Hull, R.L., and Utzschneider, K.M. (2006). Mechanisms linking obesity to insulin resistance and type 2 diabetes. Nature *444*, 840–846.

Kanda, H., Tateya, S., Tamori, Y., Kotani, K., Hiasa, K., Kitazawa, R., Kitazawa, S., Miyachi, H., Maeda, S., Egashira, K., and Kasuga, M. (2006). MCP-1 contributes to macrophage infiltration into adipose tissue, insulin resistance, and hepatic steatosis in obesity. J. Clin. Invest. *116*, 1494–1505.

Kassinen, A., Krogius-Kurikka, L., Mäkivuokko, H., Rinttilä, T., Paulin, L., Corander, J., Malinen, E., Apajalahti, J., and Palva, A. (2007). The fecal microbiota of irritable bowel syndrome patients differs significantly from that of healthy subjects. Gastroenterology *133*, 24–33.

Keller, M.P., Choi, Y., Wang, P., Davis, D.B., Rabaglia, M.E., Oler, A.T., Stapleton, D.S., Argmann, C., Schueler, K.L., Edwards, S., et al. (2008). A gene expression network model of type 2 diabetes links cell cycle regulation in islets with diabetes susceptibility. Genome Res. *18*, 706–716.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics *9*, 559.

Langfelder, P., Zhang, B., and Horvath, S. (2008). Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. Bioinformatics *24*, 719–720.

Langfelder, P., Luo, R., Oldham, M.C., and Horvath, S. (2011). Is my network module preserved and reproducible? PLoS Comput. Biol. *7*, e1001057.

Langfelder, P., Mischel, P.S., and Horvath, S. (2013). When is hub gene selection better than standard meta-analysis? PLoS ONE *8*, e61505.

Lusis, A.J., and Weiss, J.N. (2010). Cardiovascular networks: systems-based approaches to cardiovascular disease. Circulation *121*, 157–170.

Melé, M., Ferreira, P.G., Reverter, F., DeLuca, D.S., Monlong, J., Sammeth, M., Young, T.R., Goldmann, J.M., Pervouchine, D.D., Sullivan, T.J., et al.; GTEx Consortium (2015). Human genomics. The human transcriptome across tissues and individuals. Science *348*, 660–665.

Miller, J.A., Horvath, S., and Geschwind, D.H. (2010). Divergence of human and mouse brain transcriptome highlights Alzheimer disease pathways. Proc. Natl. Acad. Sci. USA *107*, 12698–12703.

Miquel, S., Martín, R., Rossi, O., Bermúdez-Humarán, L.G., Chatel, J.M., Sokol, H., Thomas, M., Wells, J.M., and Langella, P. (2013). Faecalibacterium prausnitzii and human intestinal health. Curr. Opin. Microbiol. *16*, 255–261.

Phipson, B., and Smyth, G.K. (2010). Permutation p-values should never be zero: calculating exact p-values when permutations are randomly drawn. Stat. Appl. Genet. Mol. Biol. *9*, Article39.

Pierson, E., Koller, D., Battle, A., Mostafavi, S., Ardlie, K.G., Getz, G., Wright, F.A., Kellis, M., Volpi, S., and Dermitzakis, E.T.; GTEx Consortium (2015). Sharing and specificity of co-expression networks across 35 human tissues. PLoS Comput. Biol. *11*, e1004220.

Ritchie, S.C., Würtz, P., Nath, A.P., Abraham, G., Havulinna, A.S., Fearnley, L.G., Sarin, A.-P., Kangas, A.J., Soininen, P., Aalto, K., et al. (2015). The biomarker GlycA is associated with chronic inflammation and predicts longterm risk of severe infection. Cell Syst. *1*, 293–301.

Rotival, M., and Petretto, E. (2014). Leveraging gene co-expression networks to pinpoint the regulation of complex traits and disease, with a focus on cardiovascular traits. Brief. Funct. Genomics *13*, 66–78.

Sartipy, P., and Loskutoff, D.J. (2003). Monocyte chemoattractant protein 1 in obesity and insulin resistance. Proc. Natl. Acad. Sci. USA *100*, 7265–7270.

Schadt, E.E. (2009). Molecular networks as sensors and drivers of common human diseases. Nature *461*, 218–223.

Shay, T., and Kang, J. (2013). Immunological Genome Project and systems immunology. Trends Immunol. *34*, 602–609.

Stuart, J.M., Segal, E., Koller, D., and Kim, S.K. (2003). A gene-coexpression network for global discovery of conserved genetic modules. Science *302*, 249–255.

Stumpf, M.P.H., and Porter, M.A. (2012). Mathematics. Critical truths about power laws. Science *335*, 665–666.

van Nas, A., Guhathakurta, D., Wang, S.S., Yehya, N., Horvath, S., Zhang, B., Ingram-Drake, L., Chaudhuri, G., Schadt, E.E., Drake, T.A., et al. (2009). Elucidating the role of gonadal hormones in sexually dimorphic gene coexpression networks. Endocrinology *150*, 1235–1249.

Xia, K., Xue, H., Dong, D., Zhu, S., Wang, J., Zhang, Q., Hou, L., Chen, H., Tao, R., Huang, Z., et al. (2006). Identification of the proliferation/differentiation switch in the cellular network of multicellular organisms. PLoS Comput. Biol. *2*, e145.

Yang, X., Schadt, E.E., Wang, S., Wang, H., Arnold, A.P., Ingram-Drake, L., Drake, T.A., and Lusis, A.J. (2006). Tissue-specific expression and regulation of sexually dimorphic genes in mice. Genome Res. *16*, 995–1004.

Zhang, B., and Horvath, S. (2005). A general framework for weighted coexpression network analysis. Stat. Appl. Genet. Mol. Biol. *4*, e17.