

# Murine Gut Microbiota Is Defined by Host Genetics and Modulates Variation of Metabolic Traits

Autumn M. McKnite<sup>1,9</sup>, Maria Elisa Perez-Munoz<sup>2,9</sup>, Lu Lu<sup>3</sup>, Evan G. Williams<sup>4</sup>, Simon Brewer<sup>5</sup>, Pénélope A. Andreux<sup>4</sup>, John W. M. Bastiaansen<sup>6</sup>, Xusheng Wang<sup>3</sup>, Stephen D. Kachman<sup>7</sup>, Johan Auwerx<sup>4</sup>, Robert W. Williams<sup>3</sup>, Andrew K. Benson<sup>2</sup>, Daniel A. Peterson<sup>2,8</sup>, Daniel C. Ciobanu<sup>1\*</sup>

**1** Animal Science Department, University of Nebraska, Lincoln, Nebraska, United States of America, **2** Department of Food Science and Technology, University of Nebraska, Lincoln, Nebraska, United States of America, **3** Department of Anatomy and Neurobiology, University of Tennessee Health Science Center, Memphis, Tennessee, United States of America, **4** Laboratory for Integrative and Systems Physiology, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, **5** Department of Geography, University of Utah, Salt Lake City, Utah, United States of America, **6** Animal Breeding and Genomics Centre, Wageningen University, Wageningen, The Netherlands, **7** Department of Statistics, University of Nebraska, Lincoln, Nebraska, United States of America, **8** Department of Pathology, John Hopkins University School of Medicine, Baltimore, Maryland, United States of America

## Abstract

The gastrointestinal tract harbors a complex and diverse microbiota that has an important role in host metabolism. Microbial diversity is influenced by a combination of environmental and host genetic factors and is associated with several polygenic diseases. In this study we combined next-generation sequencing, genetic mapping, and a set of physiological traits of the BXD mouse population to explore genetic factors that explain differences in gut microbiota and its impact on metabolic traits. Molecular profiling of the gut microbiota revealed important quantitative differences in microbial composition among BXD strains. These differences in gut microbial composition are influenced by host-genetics, which is complex and involves many loci. Linkage analysis defined Quantitative Trait Loci (QTLs) restricted to a particular taxon, branch or that influenced the variation of taxa across phyla. Gene expression within the gastrointestinal tract and sequence analysis of the parental genomes in the QTL regions uncovered candidate genes with potential to alter gut immunological profiles and impact the balance between gut microbial communities. A QTL region on Chr 4 that overlaps several interferon genes modulates the population of *Bacteroides*, and potentially Bacteroidetes and Firmicutes—the predominant BXD gut phyla. *Irak4*, a signaling molecule in the Toll-like receptor pathways is a candidate for the QTL on Chr15 that modulates Rikenellaceae, whereas *Tgfb3*, a cytokine modulating the barrier function of the intestine and tolerance to commensal bacteria, overlaps a QTL on Chr 12 that influence Prevotellaceae. Relationships between gut microflora, morphological and metabolic traits were uncovered, some potentially a result of common genetic sources of variation.

**Citation:** McKnite AM, Perez-Munoz ME, Lu L, Williams EG, Brewer S, et al. (2012) Murine Gut Microbiota Is Defined by Host Genetics and Modulates Variation of Metabolic Traits. PLoS ONE 7(6): e39191. doi:10.1371/journal.pone.0039191

**Editor:** Bryan A. White, University of Illinois, United States of America

**Received:** September 8, 2011; **Accepted:** May 16, 2012; **Published:** June 18, 2012

**Copyright:** © 2012 McKnite et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by a Layman Award from the Office of Research of the University of Nebraska-Lincoln, the Nebraska Tobacco Settlement Biomedical Research Development Funds, the Ecole Polytechnique Fédérale de Lausanne, the Swiss National Science Foundation, the European Research Council Ideas Programme (Sirtuins; ERC-2008-AdG23118), NIAAA (U01AA013499, U01AA014425), NIDA and NIAAA (P20-DA 21131), and the UT Center for Integrative and Translational Genomics. The Nebraska Tobacco Settlement Biomedical Research Development Funds is not a profit making body and there are no commercial connections. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: dciobanu2@unl.edu

 These authors contributed equally to this work.

## Introduction

The gastrointestinal tract contains a vast community of organisms that collectively comprise the microbiota, which is critical for the development of the intestinal epithelium and mucosal immunity as well as contributing digestive metabolic functionalities [1,2]. The composition of this complex community is established early in life and influenced soon after birth by maternal environment and stochastic exposure to different microbes. Diet, exposure to antibiotics, pathogens, and parasites can also influence compositional features of the microbiota. In a remarkable set of studies, the transition to a low fat diet in overweight humans shifted the gut flora to a composition that resembled that of healthy non-obese matched controls [3,4]. Conversely, the dissemination of a complex gut flora from

overweight animals into the gut of otherwise matched gnotobiotic animals can induce statistically significant weight gains [5,6].

There is increasing evidence that genetics of the host influence and interact with gut microbiota in various mammals [6,7,8,9,10,11]. Early studies focused on enteric pathogens, such as that of Meijerink et al. (2000) who found the adhesion of F18 fimbriated *E. coli* to intestinal mucosa and subsequent susceptibility to swine edema disease to be controlled by fucosyltransferase 1 gene [10]. Several studies in monogenic models have demonstrated the role of innate immune response in altering the composition of mouse gut microbiota and disease susceptibility [6,12,13]. For example, deficiency in T-bet (*Tbx2*) promotes a colitogenic microbial population and ulcerative colitis [12], while deficiency in Toll-like receptor 5 (*TLR5*) alters the abundance of microbiota at species level leading to features characteristic of metabolic

syndrome [6]. Recently, Benson et al. (2010) using Quantitative Trait Locus (QTL) mapping methods detected genome-wide linkages with the relative abundance of several taxa in the gut of a large murine advanced intercross population [7].

The purpose of the present study is to uncover natural genetic variants present in the host that explain variation in mouse gut microbiota and also explore its impact on obesity and other metabolic phenotypes that affect health. We achieved this by combining the power of next-generation sequencing of gut microbiota with genome-wide linkage analysis and a deep multi-scalar analysis of microbiota across an extensive set of physiological phenotypes in the BXD mouse reference population. BXD is mouse genetic resource characterized extensively at molecular and phenotypic level. This population resulted from the combination of C57BL/6J and DBA/2J genomes [14] and displays important differences in susceptibility to obesity and other morphologic, immunologic, behavioral and metabolic traits. While gut microbiota of C57BL/6J-based genetic resources were previously profiled in various environments, here we introduce the gut microbial profile of DBA/2J, a strain known for its high proportion of body fat mass and predisposition to obesity [15,16]. Our analysis of gut microbiota of the BXD strains revealed substantial quantitative differences among strains, which can be explained by complex and polygenic influences of the host.

## Results

### Gut microbial profile of the BXD strains is dominated by Firmicutes and Bacteroidetes and displays substantial variability

Pyrosequencing generated 512,646 sequencing reads of the V1–V2 region of the 16S rRNA gene that passed the filtering criteria. An average of 7,651 sequencing reads was obtained per sample. The reads were assigned to different taxonomic units using three approaches. Using a parallelized version of the CLASSIFIER from Ribosomal Database Project (RDP), 97.12% of the sequences were assigned to five phyla groups. Considerable variability was detected at phyla level, with Firmicutes (79.25%) and Bacteroidetes (15.69%) representing the predominant taxa (Table 1). Bacteroidetes is represented at low levels in strains such BXD77 (4.51–7.60%) where Firmicutes accounted for the majority of microbiota (81.11–97.34%). In contrast, in BXD15 the ratio between Bacteroidetes (34.66–40.36%) and Firmicutes (50.47–56.24%) is well balanced. Evidence of Actinobacteria, Proteobacteria and candidate phylum TM7 was detected at very low levels (<1%). Approximately 62% of the sequences were assigned at genus level. *Lactobacillus* was the predominant group (54.1%) with no other genera contributing more than 3.2% of the microbiota. We detected important differences in the abundance of *Lactobacillus* among BXD strains ranging from an average of approximately 22% in BXD12 to 71% in BXD100. Sex did not have a significant effect on the gut microbiota while age had significant effect on three taxa (Bacillales, Staphylococaceae, *Staphylococcus*) and cage density had effects limited to one taxon (Proteobacteria). The location of the cage (room) had significant effects on 7 taxa that include most of the members of Erysipelotrichi and Clostridia branches. Most of these effects did not reach significance if Bonferroni correction was applied for multiple testing.

Sequencing reads were also subjected to chimeras removal, alignment and subsequently were clustered using the Complete Linkage Clustering tool of the RDP to generate Operational Taxonomic Units (OTU) (Table 1). The OTUs that have an average of at least 30 counts across BXD samples belong to the genus *Lactobacillus* (*L. johnsonii*, *L. murinus* and *L. intestinalis*),

*Staphylococcus* (*S. xylosum* and *S. lentus*) and *Barnesiella*. *Lactobacillus* OTUs are the predominant species accounting for 25.6% of the OTUs. The most abundant OTU had the highest similarity with *L. johnsonii* accounting for 14.9% of the classified sequences. OTU composition varied substantially among BXD strains. For example, *L. murinus* abundance is negligible in several BXD strains such as BXD96, while in others such as BXD66, the contribution is considerable (28.2%).

In the last approach the top most abundant OTU clusters that accounted for 90% of the reads in the dataset were combined into 158 OTUs using a 97% identity cutoff to eliminate overlap between clusters. The entire dataset was compared to these 158 OTUs and the single best BLAST hit identified for all sequences. This allowed us to assign over an average of 90% of the sequence reads to one of these OTUs (s.d. 5.7%, minimum 71% and maximum 98%) and subsequently assign all the reads to five phyla and approximately 73% of the reads to genera (Table S1).

### Host genetics impacts microbial composition of mouse gut

QTL analysis of gut microbiota based on CLASSIFIER output revealed five QTL regions ( $P < 0.05$ ) at the genome-wide level for six taxonomic groups (Table 2, Figure 1). Loci associated with significant effects were concentrated on four chromosomes. The QTLs were restricted to a particular taxon, branch or influenced the variation of taxa across phyla. For example, QTLs mapped on Chr 12 have an effect on Prevotellaceae while a QTL mapped on Chr 17 influenced the variation of Bacillales/Staphylococaceae/*Staphylococcus* branch. In contrast, a QTL located on Chr 4 potentially influenced taxa in different phyla.

Gene expression of the gastrointestinal tract and sequence analysis of parental genomes in the QTL regions were used to uncover potential candidate genes that could explain the variation in gut microbiota. Bacteroidetes displayed the largest number of taxonomic units (*Bacteroides*, Prevotellaceae and Rikenellaceae) influenced by host genetics. A QTL located on Chr 15 (LRS = 19.2, 92.73–97.39 Mb) had a significant effect on Rikenellaceae. *Rapgef2* and *Irak4* are two positional candidate genes that harbor non-synonymous SNPs and display important fold difference in expression between parental alleles (1.59 and 1.21 $\times$ ). A similar example was found for the QTL mapped on Chr 12 for Prevotellaceae. One of the candidates for this QTL was *Tgfb3*, an anti-inflammatory cytokine with a potential role in modulating barrier function of the intestine and tolerance to commensal bacteria [17,18,19]. *Tgfb3* is differentially expressed in parental strains (1.23–1.45 $\times$ ) with the expression increased by *D2* allele in the jejunum, cecum and ileum.

A QTL that has potential effects across phyla was located on Chr 4. This locus (LRS = 17.9, 87.58–95.23 Mb) explains 21% of the observed variation in the abundance of the genus *Bacteroides* (Figure 2). The same locus was also associated with suggestive effects in Firmicutes (LRS = 13.38,  $P = 0.27$ ) and Bacteroidetes (LRS = 12.57,  $P = 0.36$ ) phyla. The C57BL/6J (*B6*) allele (haplotype) from this locus increased the proportion of Firmicutes while the DBA/2J (*D2*) allele increased the proportion of Bacteroidetes. The QTL influenced the variation of the Bacteroidetes/*Bacteroides* branch from phyla to genus, with an increased effect at the tips of the phylogenetic tree (Figure 1). This QTL region is rich in interferon alpha (*Ifna1*, *Ifna2*, *Ifna4* – *Ifna7*, *Ifna9*, *Ifna11*– *Ifna14*, *Ifnab*), beta (*Ifnb1*), zeta (*Ifnz*), and epsilon (*Ifne1*) genes. The expression of this cluster of genes in the gastrointestinal tract was limited to *Ifna1*, *Ifna12*, and *Ifnab*. Considerable fold-difference in expression between parental alleles was detected for *Ifna12* and *Ifnab* with *D2* alleles increasing expression. Additional genes that

**Table 1.** Mean and measures of variability of phyla, genera and Operational Taxonomic Units (OTU) of the BXD strains, their parental lines C57BL/6J and DBA/2J and F1 hybrids (D2B6F1) used in QTL mapping.

Taxa	Mean % (SE)	SD (%)	Min (%)	Max (%)	Median	Samples missing taxa
<i>Phylum</i>						
Actinobacteria	0.92 (0.17)	1.38	0.01	8.57	4.40	0
Bacteroidetes	16.00 (1.43)	11.64	0.45	48.58	13.42	0
Firmicutes	78.84 (1.59)	12.89	46.44	97.31	81.28	0
Proteobacteria	0.37 (0.10)	0.88	0.01	6.96	0.18	0
<i>Genus</i>						
<i>Alistipes</i>	0.30 (0.04)	0.32	0	1.51	0.22	10
<i>Bacteroides</i>	1.78 (0.29)	2.31	0	9.54	0.93	8
<i>Lachnospiraceae</i>	0.34 (0.05)	0.44	0	2.26	0.22	5
<i>Lactobacillus</i>	54.12 (2.34)	18.97	11.80	91.98	54.69	0
<i>Staphylococcus</i>	2.78 (0.71)	5.74	0.01	26.96	0.31	2
<i>OTU</i>						
<i>Barnesiella</i>	0.74 (0.11)	0.91	0	4.48	0.40	4
<i>Lactobacillus johnsonii</i>	11.87 (1.05)	8.57	0	32.93	11.34	3
<i>Lactobacillus murinus</i>	5.65 (0.75)	6.09	0	28.80	3.74	5

doi:10.1371/journal.pone.0039191.t001

displayed important fold-differences in expression include *Ptplad2*, *Cdkn2b*, and *Klhl9*. Variation in DNA sequence between parental genomes is relatively limited to the region spanning 87.0 to 89.0 Mb. Analysis of genes located in this interval and expressed in the gastrointestinal tract revealed non-synonymous SNPs in *BC057079* (6 SNPs), *Ptplad2* (1), *Ifnab* (4) and *Mtap* (1). An early stop codon was introduced by a *D2* SNP variant in one of the *Ptplad2* isoforms. Variants of two of the SNPs located in *Ifnab* (rs13477830 and rs32311578) were shared equally in a diversity panel of laboratory mice lines while others (rs28091853 and rs28091852) had very limited variation among lines, with *B6* being the rare allele.

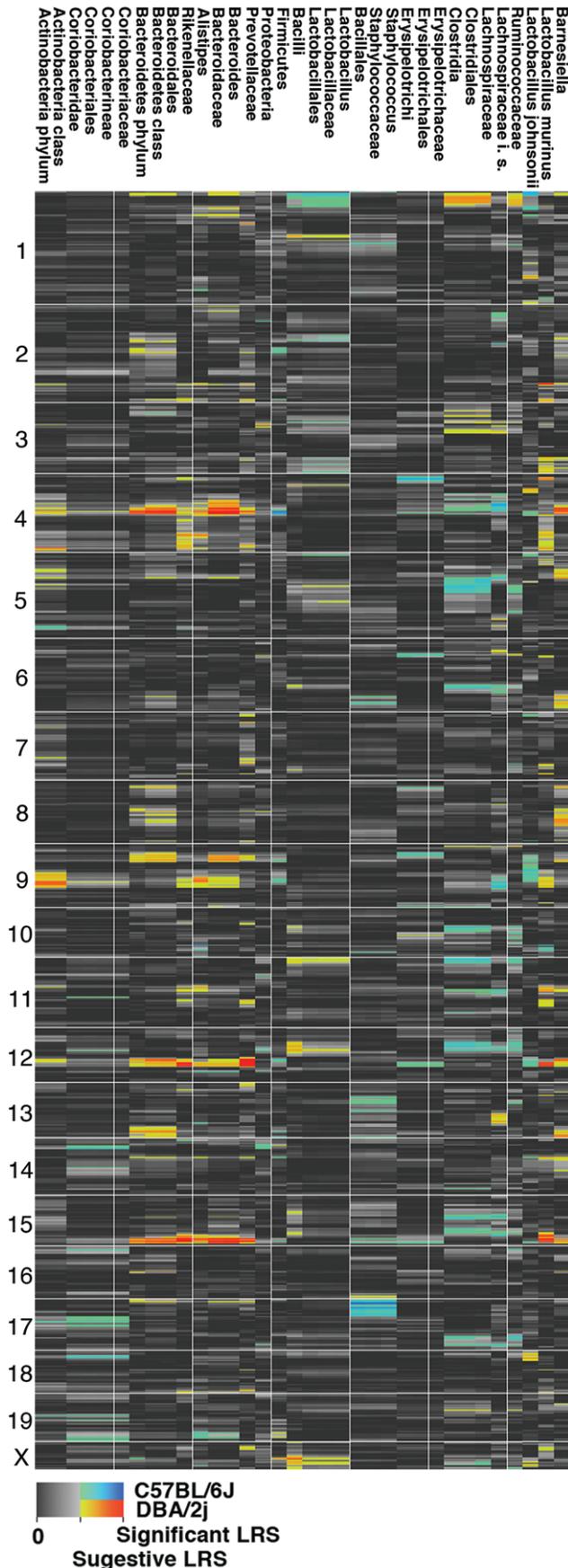
The presence of most of the QTLs was confirmed when sequencing reads were assigned to taxa based on a combination of OTU clustering and CLASSIFIER. The QTL associated with Prevotellaceae was confirmed on the same region of Chr 12 (LRS = 24.09, 88.25–89.93,  $P=0.002$ ). The presence of multiple loci mapped adjacently (31.56, 23.32–26.35 and 7.75–10.17 Mb) on Chr 17 was confirmed for Staphylococcaceae and *Staphylococcus* ( $P<0.02$ ,  $P<0.06$  and  $P<0.09$ ). The QTL located on Chr 4 (87.7–88.04 Mb) was confirmed having effects close to significance in *Bacteroides* (LRS = 17.7, 87.7–88.04 Mb,  $P=0.07$ ) with trends for Bacteroidetes and Firmicutes ( $P<0.42$ ). Suggestive effects were also confirmed for the QTL located on Chr 15 (92.84–94.02 Mb) for Rikenellaceae (LRS = 14.84,  $P=0.20$ ).

A major source of variation of the gut microbiota in BXD is represented by *Lactobacillus*. One of the potential sources of this variation is a suggestive QTL mapped on Chr 10 (LRS = 15.3, 111.13–111.84 Mb,  $P=0.054$ ) for an OTU similar to *L. murinus* with *B6* allele increasing its abundance. The effect of this QTL is not restricted to *L. murinus* but has a potential effect across phyla. A QTL located in the same position potentially influences the variation in abundance of *Alistipes* (LRS = 13.84, 111.13–111.84,  $P=0.23$ ), a member of Bacteroidetes. A strong correlation was detected between *L. murinus* and *Alistipes* ( $r=0.81$ ,  $p<0.0001$ ), potentially as a result of common or linked genetic sources of variation. Important fold-differences in cecum expression between parental alleles of the genes located in this QTL area were found

for *Bbs10*, *Krr1* and *Nap11*. If the sources of these differences in expression are caused by the same regulatory elements, we assumed same modulation mechanisms and direction of effects across BXD tissues [20]. We found consistent evidence of *cis*-regulation of gene expression for *Nap11* in hippocampus [21], whole brain [22] and liver [23] data sets and for *Bbs10* in striatum [24] and retina [25]. The direction of the additive effect was consistent across tissues and in agreement with the expression of parental alleles in the gastrointestinal tract. Presence of non-synonymous DNA sequence variants was limited to *Bbs10*, a chaperonin-like protein and one of the major genetic sources of Bardet-Biedl syndrome in humans with characteristics including postnatal obesity and diabetes among others [26]. The *B6* alleles of the *Bbs10* are the most common across a diverse set of inbred strains.

### Relationships exist between gut microflora, morphologic, and metabolic traits

The BXD strains displays important phenotypic variation for anatomic, developmental and metabolic traits (Figure S1 and S2). Several BXD strains (BXD28, BXD16, and BXD34) are characterized by large weight gains (0.58 to 1.08 g/week of 13 to 20 weeks-old males) while others (BXD69, BXD56, and BXD98) are characterized by lack of or limited growth (−0.10 to 0.08 g/week) (Figure S1). Feed intake and activity level are two of the sources that can explain some of the differences in body weight among BXD strains ( $P<0.005$ ). Feed consumption varies among strains ranging from less than 1.30 g (BXD16, BXD103, and BXD31) to more than 3 g of feed consumed per 24 hr (BXD69, BXD45, BXD31, and BXD68). Activity level during daytime ranges from 1155.8–1304.8 fine movements/hour in extremely active strains (BXD8, BXD75, BXD18, and BXD86) to 424.8–467.1 movements/hour in more sedentary strains (BXD96, BXD71, and BXD69). To examine the sources that could influence susceptibility to metabolic syndrome in BXD, we analyzed the relationship between weight gains and body composition. We found moderate correlations between average weekly gain and fat ( $r=0.48$ ,  $P<0.0001$ ) and lean mass ( $r=-0.45$ ,



**Figure 1. Genome-wide QTL mapping of the gut microbiota in BXD strains.** The heat map is represented in “Grey+Blue+Red” in which more intense colors mark chromosomal regions with statistical values associated with high linkage while the spectrum indicates the allelic effect. The blue-green regions are those in which *B* allele is associated with higher trait values, whereas red-yellow regions are those in which the *D* allele is associated with higher trait values. Grey and black regions correspond to insignificant linkage between traits and DNA markers.

doi:10.1371/journal.pone.0039191.g001

$P < 0.001$ ). Variation in body fat composition ranged from 9.65% (BXD20,  $s.e. = 0.99$ ) to 26.18% (BXD92,  $s.e. = 2.90$ ) with an average of 15.67% ( $s.e. = 0.33$ ) (Figure S2). These differences in body fat composition could explain approximately 13% ( $P < 0.001$ ) of the blood glucose level expressed as area under the curve following a glucose tolerance test.

Gut microbial profiles were subjected to principal component analysis (PCA) to determine if variation in gut microbiota could explain some of the differences in weight, body composition and susceptibility to diabetes. The presence and magnitude of these relationships were influenced by sex. The first eigenvector explained 24.3% of the variation of classified taxa and is associated with weekly body weights of 14 to 20 weeks-old females ( $P < 0.05$ ) but not in males. Phenotypic correlations were also estimated between the abundance of each taxonomic group and a set of 143 metabolic traits available for BXD strains [15]. The number of BXD strains that have both microbiota and metabolic traits profiled varied by trait from 9 to 21 strains. Consistent relationships in both sexes were estimated between *L. johnsonii* and several morphologic and metabolic traits. Increased abundance of *L. johnsonii* in females was associated with increased white adipose ( $r = 0.68$ ,  $P < 0.01$ ) and body fat mass ( $r = 0.57$ ,  $P < 0.05$ ) and suggestive relationships with body weight ( $r = 0.39$ ,  $P < 0.16$ ) and lean mass ( $r = -0.42$ ,  $P < 0.15$ ). Similarly, increased abundance of *L. johnsonii* in males was associated with increased body weight ( $r = 0.52$ ,  $P < 0.05$ ), lean ( $r = -0.49$ ,  $P < 0.05$ ) and fat mass ( $r = 0.45$ ,  $P = 0.05$ ) and suggestive relationships with daytime activity ( $r = -0.46$ ,  $P = 0.07$ ) and white adipose mass ( $r = 0.30$ ,  $P < 0.20$ ). Total body fat ( $r = -0.71$ ,  $p < 0.005$ ), white fat ( $r = -0.57$ ,  $p < 0.05$ ) and body lean mass ( $r = 0.65$ ,  $p < 0.05$ ) of BXD females were associated with the ratio of Firmicutes to Bacteroidetes (Figure S3). Increased level of Bacteroidetes ( $r = 0.68$ ,  $p < 0.01$ ) and reduced level of Firmicutes ( $r = -0.75$ ,  $p < 0.005$ ) indicated high body fat composition. The direction of these relationships was consistent but not significant in males. False discovery rate of the significant relationship presented ( $p < 0.05$ ) vary from 0.09 to 0.47 but the distribution of P values clearly deviates from the null hypothesis.

## Discussion

The random segregation and recombination of the C57BL/6J and DBA/2J genomes into BXD strains generated substantial variability for key physiological phenotypes including feed intake, activity level, energy expenditure, body composition, and susceptibility to metabolic syndrome and other diseases. In this study we have extended the scope of phenotyping by studying host genetics and other sources that underlie variation of gut microbiota in BXD. This is the first study that has carried out a deep multiscalar analysis of the complex gut microbiota in combination with a set of physiological phenotypes.

Murine gut microbiota is influenced by several factors including diet, host-genetics, age, environment, and caging history [7]. Molecular profiling of the gut microbiota revealed important

**Table 2.** Significant QTLs that influence gut microbial composition in the gut of BXD mice. A positive additive effect indicates that *DBA/2J* alleles increase trait values.

Taxa	Chr	Peak position (Mb)	Confidence intervals (95%)	LRS	Additive effect	Variance explained
<i>Order</i>						
Bacillales	17	21.83–23.00	21.71–26.63	19.6	−0.417	0.21
	17	7.75–10.17	7.37–10.44	19.0	−0.412	0.20
<i>Family</i>						
Prevotellaceae	12	90.15	78.36–95.80	24.9	0.556	0.27
Rikenellaceae	15	95.15–95.78	92.73–97.39	19.2	0.552	0.23
Staphylococcaceae	17	7.75–10.17	7.39–10.44	19.0	−0.417	0.21
	17	21.83–23.00	21.71–26.66	18.9	−0.416	0.20
<i>Genus</i>						
<i>Bacteroides</i>	4	87.70–88.04	87.58–95.23	17.9	0.504	0.21
<i>Staphylococcus</i>	17	7.75–10.17	5.26–10.43	19.2	−0.414	0.20
	17	21.83–23.00	21.71–26.70	18.9	−0.412	0.20

In contrast, a negative additive effect indicates that *C57BL/6J* alleles increase the trait value.  
doi:10.1371/journal.pone.0039191.t002

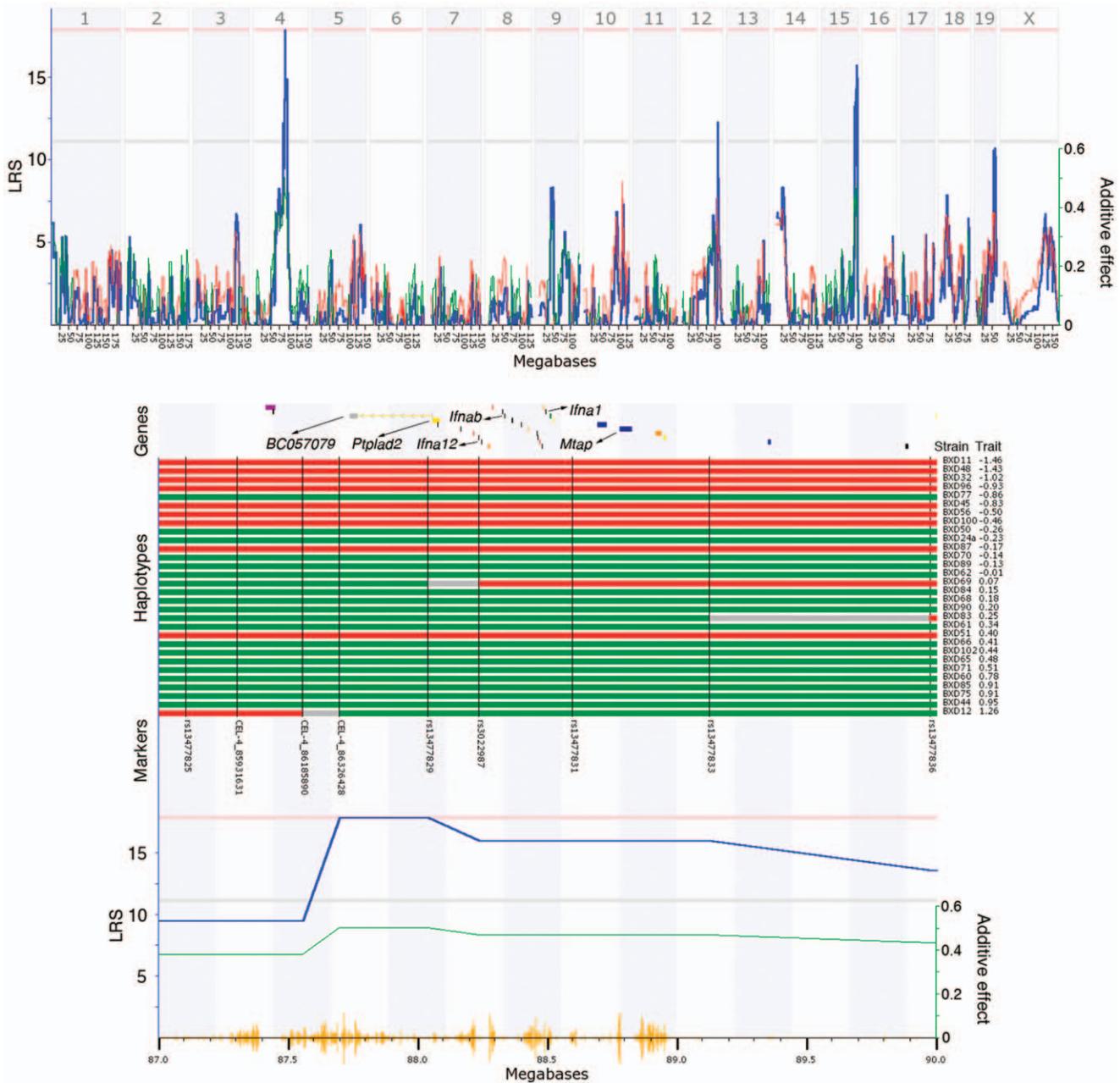
quantitative differences in microbial composition among BXD strains. As in the previous studies, Firmicutes and Bacteroidetes represented the predominant groups of gut microbiota [7,27,28] with the rest of the phyla (Actinobacteria, Proteobacteria and TM7) having a very limited contribution but substantial variation among strains.

Linkage analysis of the data generated by deep sequencing of microbial DNA evidenced the influence of host genetics on the composition of gut microbiota in two months-old naive BXD strains. In our previous work using *cis* expression QTL as empirical measures of mapping precision we showed that BXD data sets of a similar size can provide sufficient precision (~1 kb from the source gene to QTL peak) to map QTLs with LODs scores of 3–4 [19]. The QTLs identified in this study were restricted to a particular taxon, branch, or influenced the variation of taxa across phyla. Gene expression of the gastrointestinal tract and sequence analysis of parental genomes in the QTL regions uncovered several candidate genes that have the potential to alter gut immunological profiles and subsequently impact gut microbial composition. A QTL mapped on Chr 4, located in a region rich in interferon genes, influenced the variation of *Bacteroides* and potentially Bacteroidetes and Firmicutes, the predominant BXD phyla. *Tgfb3*, a cytokine with a potential role in modulating the barrier function of the intestine and tolerance to commensal bacteria [17,18,19], represents one of the candidate genes for a QTL mapped on Chr 12 that influenced Prevotellaceae. *Irak4*, a signaling molecule in Toll-like receptor pathways [29], represents a potential source for the QTL mapped on Chr 15 that influenced Rikenellaceae. IRAK4 interacts with MYD88 adapter protein, which is used by several TLR in host defense [30] and control of commensal bacteria [31]. Mutations in *IRAK4* and *MYD88* in humans impair some of the TLR pathways and are associated with predisposition to pyogenic bacterial infections [29]. Loss of *Myd88* in mice impacts the innate immune response to pathogens and influences gut microbiota by increasing the composition of three bacterial families including Rikenellaceae [13]. This microbial profile can reduce the progression of autoimmune diabetes in a *Myd88*-independent manner [13]. We found a strong correlation between the expression of *Irak4* and *Myd88* across 25 different tissues in the parental strains of BXD ( $r = 0.79$ ,  $P < 0.0001$ ) that indicates a tight

functional relationship between these genes. The same or closely linked alleles inherited from DBA/2J are associated with a lower expression of *Irak4* as well as a higher proportion of Rikenellaceae. Recently, Benson et al. found *Irak3*, a modulator of the MYD88-dependent TLR2 pathway, as a potential candidate for a QTL on Chr 10 that influence Coriobacteriaceae and *Lactococcus* [7].

*Lactobacillus* was the predominant taxa at genus level and a substantial contributor to the quantitative differences in gut microbiota among BXD strains. The source of variation could have important implication since *Lactobacillus* is known to have immunomodulating properties [32]. The *B6* allele of the suggestive QTL mapped on Chr 10 increased the abundance of *L. murinus*. A QTL located in the same region was found by Benson et al. to influence *Lactococcus* (order Lactobacillales) [7]. This genus is closely related to *Lactobacillus*, which was associated with low counts in the population used in Benson et al. study.

It is expected that a specific gut immunological profile has the potential to alter gut microbiota [5,6] favoring a microbiome that possesses increased efficiency of extracting energy from food [4,8,33]. Association analyses between gut microbial composition, especially the ratio between Firmicutes to Bacteroidetes, and obesity generated contradictory results [27,33,34,35,36,37]. One of the important roles of intestinal microorganisms is to break down the dietary fiber and starch incompletely hydrolyzed by intestinal enzymes of the host [38,39,40]. Short chain fatty acids (SCFA), specifically acetate, propionate and butyrate, are the main fermentation products of fiber digestion and represent an additional source of energy since they can be used for lipid or glucose *de novo* synthesis [41,42]. Changes in the SCFA profiles were associated with changes in gut microbiota [37,43] as well as with variation in body mass index [37]. Propionate [36,43] and butyrate [43] are favored in overweight and obese subjects, suggesting an important role of SCFA metabolism in obesity. Propionate is absorbed and delivered to hepatocytes where is used as a precursor for gluconeogenesis, lipogenesis and protein synthesis [42,44,45]. The *D2* alleles from two QTLs identified in this study increased the abundance of *Bacteroides* and Prevotellaceae, known propionate producers [37,46]. DBA/2J has a substantial higher proportion of body fat mass than C57BL/6J and is predisposed to obesity [15].



**Figure 2. Genome-wide QTL mapping of *Bacteroides* composition in the gut of BXD strains.** The Left y axis represents the strength of the linkage between *Bacteroides* composition to different DNA marker intervals on A) each chromosome (blue line) and B) detailed region of the QTL located on Chr 4 (LRS = 20.2, 82.12–95.05 Mb) that explains 27.5% of the composition in *Bacteroides*. The Right y axis represents the additive effect and indicates if the *D* (green line) or *B* (red line) allele contributes to an increase in the abundance of *Bacteroides*. Pink and gray horizontal lines indicates the significant (<0.05) and suggestive (<0.67) QTL threshold. b) *B* haplotype (red) is associated with a reduced level of *Bacteroides* while the *D* haplotype (green) is associated with an increased level of *Bacteroides*. Grey regions represent recombination spots. Yellow seismograph represents the SNP density between the sequence of the parental genomes. The QTL region is rich in genes from interferon family such as *Ifna1*, *Ifna12* and *Ifnab*, all expressed in the cecum of BXD. Additional positional candidates include *Ptplad2*, *BC057079* and *Mtap*. doi:10.1371/journal.pone.0039191.g002

The position of the QTL located on Chr 4 that increased the proportion of *Bacteroides*, coincides with a QTL (*Adip11*, peak at 88.710 Mb) associated with inguinal and retroperitoneal fat pads mass in a F2 cross between SM/J x NZB/BINJ (n = 513) mouse strains following 16-weeks of feeding with atherogenic diet [47]. SM/J is characterized by higher adiposity compared to NZB/BINJ. Individual fat pad weights are not available in BXD and there are no QTL located on Chr 4 associated with whole fat body

mass. However, most of the fat in the SM/J x NZB/BINJ F2 cross was represented by mesenteric and gonadal pads (68 to 86% in parental strains of SM/J x NZB/BINJ) which were not influenced by the same QTL located on Chr 4 [47]. Using the genotypes generated by Mouse Diversity Array [48] we compared the parental haplotypes in the peak area of the Chr4 QTL for *Bacteroides* (Figure 2). As expected, we found SM/J haplotype very similar to *DBA/2J* while *C57BL/6J* is more similar to *NZB/BINJ*,

suggesting that a QTL influencing gut microbial profile could impact energy metabolism and affect fat pads weight. This QTL region is rich in genes from the interferon family. Differences in cecum gene expression between parental alleles of the BXD were found for two members of type I interferon (IFN- $\alpha$ ) family of cytokines, *Ifna12* and *Ifnab*, with the later being rich in non-synonymous SNPs. There is limited information about the biological role of *Ifnab* but detailed analysis of *Ifna12* indicated that it has functional characteristics of type I IFN- $\alpha$  proteins [49]. In combination with other mediators such as TNF- $\alpha$  and IL-12, IFN- $\alpha$  is produced in response to microbial products and modulate innate and adaptive immune response [50,51]. In mice, type I IFN consists in multiple IFN- $\alpha$  subtypes and one IFN- $\beta$  and IFN- $\omega$ , all binding a ubiquitously expressed heterodimeric receptor – IFNAR [52]. IFN- $\alpha/\beta$  genes are involved in Toll-like receptor, RIG-I-like receptor, Jak-STAT signaling pathways and in natural killer cell mediated cytotoxicity and as a result could impact the balance between various microbial communities in the gut. The role of IFN- $\alpha/\beta$  in immune function and intestinal homeostasis was underlined by Katakura et al. [53]. Using a murine model they discovered a protective role of IFN- $\alpha/\beta$  in experimental colitis as a result of cross-regulation mechanisms with TNF- $\alpha$ .

While variation in body weight and susceptibility to obesity are known to be strongly influenced by genetic variation ( $h^2 = 0.40–0.80$ ) [54,55,56], recent genome-wide association studies uncovered a relatively modest fraction of genetic variants that could explain variation of body mass index [57,58], predisposition to obesity [57,59,60,61,62] and fat distribution [63,64]. Variation in gut microbiota and complex relationships with host genetics can represent unaccounted sources of differences for physiological phenotypes including susceptibility to obesity.

The BXD reference population has been used since 1973 [65] to characterize genetic factors that control variation of both Mendelian and complex traits [66]. Recently, BXD strains have undergone extensive phenotyping and molecular characterization including deep sequencing of the parental lines and identification of approximately 4.9 million SNPs and hundreds of thousands of insertion/deletions and copy number variants [67]. This deep compendium of sequence polymorphisms represents a genetic resource that models some aspects of sequence variation in human populations. In this study, the BXD population was used to detect and quantify genetic factors that may have a significant influence on the variation of gut microbiota. We have demonstrated that host-genetics is complex and involves many loci. These differences in microbial composition could impact susceptibility to obesity and other metabolic traits. Functional analysis of gut microbiota and characterization of the relationships with host-genotype has important implications to human health and agriculture. The gut microbial composition can be temporarily altered through dietary interventions tailored to host genotype, ultimately mitigating the effects of unfavorable alleles and inducing profiles that promote human health. Genetic variants that influence gut microbiota may also be used in selection programs of livestock to improve feed efficiency, disease resistance, and to reduce dissemination of pathogens associated with zoonotic diseases such as *E.coli* O157:H7 or *Salmonella*.

## Materials and Methods

### Animals, animal care and sample collection

Fecal samples were collected in less than 24 hours period from all 30 BXD strains used in mapping. Fecal samples were also collected from the parental lines of BXD strains, C57BL/6J and DBA/2J and from F1 hybrids (D2B6F1). Fecal pellets were

collected separately from two months-old naive males and female by pooling samples from each cage and storing at  $-80^{\circ}\text{C}$ . The mice were housed at the University of Tennessee Health Science Center (UTHSC) in a specific pathogen-free environment at  $20–24^{\circ}\text{C}$  with a light/dark cycle of 14/10 hr. and *ad libitum* access to food and water. Male and female littermates were housed in separate cages. The cages were located in four different rooms of the same facility with a relatively balanced ratio of cages allocated to each sex per room and with an average of three mice per cage (Table S2). The fat source of the diet was derived from soybean oil providing approximately 17% of energy and containing 0.8% saturated, 1.3% monosaturated, and 2.9% polyunsaturated fatty acids (diet 7019, Harlan Teklad). All animals were kept in accordance with guidelines set by the NIH Guide for the Care and Use of Laboratory Animals and under the prevue of the Institutional Animal Care and Use Committee (IACUC) at the UTHSC. The IACUC at the UTHSC specifically approved the study (Permit No. 680).

### Pyrosequencing of the Gut Microbial DNA

DNA was extracted from a pool of approximately 0.10 mg of fecal pellets collected from each cage using QIAmp DNA stool Mini Kit (Qiagen) following cell lysis as described previously [68]. The 16S rRNA gene was amplified using primers targeting the V1–V2 region and containing bar-coded adapters. The forward primer used was 5'-AGAGTTTGATCMTGGCTCAG-3' and the reverse primer was 5'-CTGCTGCCTYCCGTA-3'. Sequencing of the products was performed from one end of the amplicons using Roche 454 GS FLX Titanium chemistry.

### Pyrosequencing data analysis

Quality control of the raw sequences was based on a filtering protocol that excludes sequences that are short, missing barcodes, have high nucleotide ambiguities ( $>2$ ), or low average quality scores ( $<25$ ). The reads were distributed to each sample based on barcodes. The reads were assigned to different taxonomic units using three approaches:

- A parallelized version of CLASSIFIER from the Ribosomal Database Project (RDP) was used to assign sequences to taxonomic groups as previously described [7]. The reads in each taxonomic bin were normalized to the proportion of the total number of reads per each sample.
- After quality control procedures, sequences were submitted to the open source software package Quantitative Insights into Microbial Ecology (QIIME) [69] to remove chimeric sequences using the Chimera Slayer method. Sequences were subjected to alignment by ALIGNER and subsequently clustered using the Complete Linkage Clustering tool of the RDP to generate Operational Taxonomic Units (OTU). OTUs are defined units that include sequences with similarities that exceed 97%. OTUs were binned by sample using QIIME [69], and those whose average was above 30 sequences per OTU were selected for QTL mapping.
- In the third approach representative sequences were selected from the originally identified OTUs that accounted for 90% of assigned reads in the dataset. These reads were re-clustered into 158 non-overlapping OTUs using a 97% identity cutoff. The entire dataset was then compared to the database of these 158 non-overlapping OTUs and the single best BLAST hit identified for all sequences in the dataset using BioEdit Blastall function. Alignments of 200 bp and greater than 97% identity were tabulated for each OTU. Each OTU representative sequence was assigned a taxonomy using

RDP Seq Match and the counts of those of the same Genus, Family or Phylum were added together for QTL mapping.

### QTL mapping

Proportion values of the microbial profile of the taxa that had at least an average of 20 counts/sample and not more than 10 samples lacking reads were  $\log_{10}$  – transformed. Samples lacking sequencing reads for a given taxon were assigned a value of 0.5 divided by the total number of reads of the sample, which was  $\log_{10}$  – transformed [7]. The influence of sex, age and cage density on gut microbial composition was tested using a linear-model. A linear model that included cage location was employed to generate residuals for all taxa. Cage density and age was incorporated in the model when significant. We computed the regression between the genotypes of 3,785 informative markers and average of the proportion values of males and females samples for each strain using QTL Reaper to map QTLs for gut microbiota as previously described [21]. Genome-wide empirical P values were obtained by permuting data for each phenotype (proportion values) randomly between 1,000 and 1,000,000 times. Confidence intervals for QTLs were obtained by likelihood support interval (1.5-logarithm of the odds, LOD) as described [70].

### Phenotypes

Correlations between relative abundance of gut microbiota and a BXD set of physiological phenotypes stored in GeneNetwork database ([www.genenetwork.org](http://www.genenetwork.org)) were computed using Pearson product-moment correlations. This set includes 143 traits associated with energy metabolism, blood chemistry, morphology and cardiology, that were generated as part of a large and systematic phenotyping study of 17 to 43 BXD strains by Koutnikova et al. (2009, GeneNetwork Record ID: 11017, 11287, 12076, 12821–12960) [16]. False discovery rates of the multiple comparisons were estimated for each taxon based on the p-values resulted from correlation estimates using the R package *qvalue* [71].

### Gene expression

Gene expression profile was employed to uncover genes that are expressed in the gastrointestinal tract and are potential candidates for the QTL. Measurement of mRNA expression was performed by Illumina Mouse WG-6 v2.0 array using tissues collected from

the gastrointestinal tract of C57BL/6J and DBA/2J males including esophagus, stomach, duodenum, jejunum, ileum, cecum, and ascending colon as previously described [19]. Illumina microarray data was normalized using the vendor's rank invariant method ([www.illumina.com](http://www.illumina.com)). Gene expression and expression QTL data [20–24] can be accessed from [www.genenetwork.org](http://www.genenetwork.org).

### Supporting Information

**Figure S1 Average body weight of BXD strains between 13 to 20 weeks of age.**

(TIF)

**Figure S2 BXD display important difference for morphological and metabolic traits.** A) Proportion of body fat composition, B) Area under the curve for glycemia following a glucose tolerance test, C) Total cholesterol and, D) Triglyceride levels.

(TIF)

**Figure S3 A negative relationship was detected between the ratio of Firmicutes to Bacteroidetes and body fat mass (g) of 19 weeks-old BXD females ( $r = -0.71$ ,  $p < 0.005$ ).**

(TIF)

**Table S1 Mean and measures of variability of the gut microbiota of the BXD strains, their parental lines (C57BL/6J and DBA/2J) and F1 hybrids (D2B6F1) based on the combination of OTU clustering and CLASSIFIER.**

(XLSX)

**Table S2 Cage density and the distribution of BXD mice in experimental rooms.**

(XLSX)

### Acknowledgments

We thank Jaehyoung Kim of the UNL Core for Applied Genomics and Ecology, Pengchong Liu and Sarah Rowe for technical assistance.

### Author Contributions

Conceived and designed the experiments: DC AB DP LL RW. Performed the experiments: DC AM JA LL EW PA. Analyzed the data: AM MPM SB JB XW SK AB DP DC. Contributed reagents/materials/analysis tools: DC DP AB RW JA. Wrote the paper: DC AM MPM SB JA RW AB DP.

### References

- Barbosa T, Rescigno M (2010) Host-bacteria interactions in the intestine: homeostasis to chronic inflammation. *Wiley Interdiscip Rev Syst Biol Med*: 80–97.
- Cerf-Bensussan N, Gaboriau-Routhiau V (2010) The immune system and the gut microbiota: friends or foes? *Nat Rev Immunol* 10: 735–744.
- Scarpellini E, Campanale M, Leone D, Purchiaroni F, Vitale G, et al. (2010) Gut microbiota and obesity. *Intern Emerg Med* 5 Suppl 1: S53–56.
- Turnbaugh PJ, Backhed F, Fulton L, Gordon JI (2008) Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe* 3: 213–223.
- Turnbaugh PJ, Hamady M, Yatsunenkov T, Cantarel BL, Duncan A, et al. (2009) A core gut microbiome in obese and lean twins. *Nature* 457: 480–484.
- Vijay-Kumar M, Aitken JD, Carvalho FA, Cullender TC, Mwangi S, et al. (2010) Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. *Science* 328: 228–231.
- Benson AK, Kelly SA, Legge R, Ma F, Low SJ, et al. (2010) Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. *Proc Natl Acad Sci U S A* 107: 18933–18938.
- Hildebrandt MA, Hoffmann C, Sherrill-Mix SA, Keilbaugh SA, Hamady M, et al. (2009) High-fat diet determines the composition of the murine gut microbiome independently of obesity. *Gastroenterology* 137: 1716–1724 e1711–1712.
- Joller D, Jorgensen CB, Bertschinger HU, Python P, Edfors I, et al. (2009) Refined localization of the Escherichia coli F4ab/F4ac receptor locus on pig chromosome 13. *Anim Genet* 40: 749–752.
- Meijerink E, Neuwenschwander S, Fries R, Dinter A, Bertschinger HU, et al. (2000) A DNA polymorphism influencing alpha(1,2)fucosyltransferase activity of the pig FUT1 enzyme determines susceptibility of small intestinal epithelium to Escherichia coli F18 adhesion. *Immunogenetics* 52: 129–136.
- Niu X, Li Y, Ding X, Zhang Q (2011) Refined mapping of the Escherichia coli F4ab/F4ac receptor gene(s) on pig chromosome 13. *Animal Genetics*: no-no.
- Garrett WS, Lord GM, Punit S, Lugo-Villarino G, Mazmanian SK, et al. (2007) Communicable ulcerative colitis induced by T-bet deficiency in the innate immune system. *Cell* 131: 33–45.
- Wen L, Ley RE, Volchkov PY, Stranges PB, Avanesyan L, et al. (2008) Innate immunity and intestinal microbiota in the development of Type 1 diabetes. *Nature* 455: 1109–1113.
- Williams RW, Gu J, Qi S, Lu L (2001) The genetic structure of recombinant inbred mice: high-resolution consensus maps for complex trait analysis. *Genome Biol* 2: RESEARCH0046.
- Tortorello DV, McMinn J, Chua SC (2004) Dietary-induced obesity and hypothalamic infertility in female DBA/2J mice. *Endocrinology* 145: 1238–1247.
- Koutnikova H, Laakso M, Lu L, Combe R, Paananen J, et al. (2009) Identification of the UBPI locus as a critical blood pressure determinant using a combination of mouse and human genetics. *PLoS Genet* 5: e1000591.

17. Bahrami B, Macfarlane S, Macfarlane GT (2011) Induction of cytokine formation by human intestinal bacteria in gut epithelial cell lines. *J Appl Microbiol* 110: 353–363.
18. Dignass AU, Podolsky DK (1993) Cytokine modulation of intestinal epithelial cell restitution: central role of transforming growth factor beta. *Gastroenterology* 105: 1323–1332.
19. Planchon SM, Martins CA, Guerrant RL, Roche JK (1994) Regulation of intestinal epithelial barrier function by TGF-beta 1. Evidence for its role in abrogating the effect of a T cell cytokine. *J Immunol* 153: 5730–5739.
20. Mozhui K, Ciobanu DC, Schikorski T, Wang X, Lu L, et al. (2008) Dissection of a QTL hotspot on mouse distal chromosome 1 that modulates neurobehavioral phenotypes and gene expression. *PLoS Genet* 4: e1000260.
21. Overall RW, Kempermann G, Peirce J, Lu L, Goldowitz D, et al. (2009) Genetics of the hippocampal transcriptome in mouse: a systematic survey and online neurogenomics resource. *Front Neurosci* 3: 55.
22. Peirce JL, Li H, Wang J, Manly KF, Hitzemann RJ, et al. (2006) How replicable are mRNA expression QTL? *Mamm Genome* 17: 643–656.
23. Gatti D, Maki A, Chesler EJ, Kirova R, Kosyk O, et al. (2007) Genome-level analysis of genetic regulation of liver gene expression networks. *Hepatology* 46: 548–557.
24. Rosen GD, Pung CJ, Owens CB, Caplow J, Kim H, et al. (2009) Genetic modulation of striatal volume by loci on Chrs 6 and 17 in BXD recombinant inbred mice. *Genes Brain Behav* 8: 296–308.
25. Geisert EE, Lu L, Freeman-Anderson NE, Templeton JP, Nassr M, et al. (2009) Gene expression in the mouse eye: an online resource for genetics using 103 strains of mice. *Mol Vis* 15: 1730–1763.
26. Stoetzel C, Laurier V, Davis EE, Muller J, Rix S, et al. (2006) BBS10 encodes a vertebrate-specific chaperonin-like protein and is a major BBS locus. *Nat Genet* 38: 521–524.
27. Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD, et al. (2005) Obesity alters gut microbial ecology. *Proc Natl Acad Sci U S A* 102: 11070–11075.
28. Murphy EF, Cotter PD, Healy S, Marques TM, O'Sullivan O, et al. (2010) Composition and energy harvesting capacity of the gut microbiota: relationship to diet, obesity and time in mouse models. *Gut* 59: 1635–1642.
29. Suhir H, Etzioni A (2010) The role of Toll-like receptor signaling in human immunodeficiencies. *Clin Rev Allergy Immunol* 38: 11–19.
30. Akira S, Uematsu S, Takeuchi O (2006) Pathogen recognition and innate immunity. *Cell* 124: 783–801.
31. Rakoff-Nahoum S, Paglino J, Eslami-Varzaneh F, Edberg S, Medzhitov R (2004) Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell* 118: 229–241.
32. Zeuthen LH, Christensen HR, Frokier H (2006) Lactic acid bacteria inducing a weak interleukin-12 and tumor necrosis factor alpha response in human dendritic cells inhibit strongly stimulating lactic acid bacteria but act synergistically with gram-negative bacteria. *Clin Vaccine Immunol* 13: 365–375.
33. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, et al. (2006) An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444: 1027–1031.
34. Collado MC, Isolauri E, Laitinen K, Salminen S (2008) Distinct composition of gut microbiota during pregnancy in overweight and normal-weight women. *Am J Clin Nutr* 88: 894–899.
35. Duncan SH, Lohley GE, Holtrop G, Ince J, Johnstone AM, et al. (2008) Human colonic microbiota associated with diet, obesity and weight loss. *Int J Obes (Lond)* 32: 1720–1724.
36. Ley RE, Turnbaugh PJ, Klein S, Gordon JI (2006) Microbial ecology: human gut microbes associated with obesity. *Nature* 444: 1022–1023.
37. Schwirtz A, Taras D, Schafer K, Beijer S, Bos NA, et al. (2010) Microbiota and SCFA in lean and overweight healthy subjects. *Obesity (Silver Spring)* 18: 190–195.
38. Bird AR, Hayakawa T, Marsono Y, Gooden JM, Record IR, et al. (2000) Coarse brown rice increases fecal and large bowel short-chain fatty acids and starch but lowers calcium in the large bowel of pigs. *J Nutr* 130: 1780–1787.
39. Louis P, Scott KP, Duncan SH, Flint HJ (2007) Understanding the effects of diet on bacterial metabolism in the large intestine. *J Appl Microbiol*: 1197–1208.
40. Topping DL, Clifton PM (2001) Short-chain fatty acids and human colonic function: roles of resistant starch and nonstarch polysaccharides. *Physiol Rev* 81: 1031–1064.
41. Macfarlane GT, Gibson GR (1997) Carbohydrate fermentation, energy transduction and gas metabolism in the human large intestine.; Mackie RI, White BA, editors. New York, USA: Chapman & Hall.
42. Wolever TM, Brighenti F, Royall D, Jenkins AL, Jenkins DJ (1989) Effect of rectal infusion of short chain fatty acids in human subjects. *Am J Gastroenterol* 84: 1027–1033.
43. Duncan SH, Louis P, Flint HJ (2007) Cultivable bacterial diversity from the human colon. *Lett Appl Microbiol* 44: 343–350.
44. Musso G, Gambino R, Cassader M (2011) Interactions between gut microbiota and host metabolism predisposing to obesity and diabetes. *Annu Rev Med* 62: 361–380.
45. Vernay M (1987) Origin and utilization of volatile fatty acids and lactate in the rabbit: influence of the faecal excretion pattern. *Br J Nutr* 57: 371–381.
46. Mahowald MA, Rey FE, Seedorf H, Turnbaugh PJ, Fulton RS, et al. (2009) Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. *Proc Natl Acad Sci U S A* 106: 5859–5864.
47. Stylianou IM, Korstanje R, Li R, Sheehan S, Paigen B, et al. (2006) Quantitative trait locus analysis for obesity reveals multiple networks of interacting loci. *Mamm Genome* 17: 22–36.
48. Yang H, Wang JR, Didion JP, Buus RJ, Bell TA, et al. (2011) Subspecific origin and haplotype diversity in the laboratory mouse. *Nat Genet*.
49. Tsang SL, Leung PC, Leung KK, Yau WL, Hardy MP, et al. (2007) Characterization of murine interferon-alpha 12 (MuIFN-alpha12): biological activities and gene expression. *Cytokine* 37: 138–149.
50. Beutler B (2004) Inferences, questions and possibilities in Toll-like receptor signalling. *Nature* 430: 257–263.
51. Theofilopoulos AN, Baccala R, Beutler B, Kono DH (2005) Type I interferons (alpha/beta) in immunity and autoimmunity. *Annu Rev Immunol* 23: 307–336.
52. Mancuso G, Midiri A, Biondo C, Beninati C, Zummo S, et al. (2007) Type I IFN signaling is crucial for host resistance against different species of pathogenic bacteria. *J Immunol* 178: 3126–3133.
53. Katakura K, Lee J, Rachmilewitz D, Li G, Eckmann L, et al. (2005) Toll-like receptor 9-induced type I IFN protects mice from experimental colitis. *J Clin Invest* 115: 695–702.
54. Maes HH, Neale MC, Eaves IJ (1997) Genetic and environmental factors in relative body weight and human adiposity. *Behav Genet* 27: 325–351.
55. Malis C, Rasmussen EL, Poulsen P, Petersen I, Christensen K, et al. (2005) Total and regional fat distribution is strongly influenced by genetic factors in young and elderly twins. *Obes Res* 13: 2139–2145.
56. Stunkard AJ, Foch TT, Hrubec Z (1986) A twin study of human obesity. *JAMA* 256: 51–54.
57. Loos RJ, Lindgren CM, Li S, Wheeler E, Zhao JH, et al. (2008) Common variants near MC4R are associated with fat mass, weight and risk of obesity. *Nat Genet* 40: 768–775.
58. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, et al. (2010) Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 42: 937–948.
59. Frayling TM (2007) Genome-wide association studies provide new insights into type 2 diabetes aetiology. *Nat Rev Genet* 8: 657–662.
60. Lowe JK, Maller JB, Pe'er I, Neale BM, Salit J, et al. (2009) Genome-wide association studies in an isolated founder population from the Pacific Island of Kosrae. *PLoS Genet* 5: e1000365.
61. Scuteri A, Sanna S, Chen WM, Uda M, Albai G, et al. (2007) Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet* 3: e115.
62. Wang K, Li WD, Zhang CK, Wang Z, Glessner JT, et al. (2011) A genome-wide association study on obesity and obesity-related traits. *PLoS One* 6: e18939.
63. Heid IM, Jackson AU, Randall JC, Winkler TW, Qi L, et al. (2010) Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat Genet* 42: 949–960.
64. Lindgren CM, Heid IM, Randall JC, Lamina C, Steinthorsdottir V, et al. (2009) Genome-wide association scan meta-analysis identifies three Loci influencing adiposity and fat distribution. *PLoS Genet* 5: e1000508.
65. Taylor BA, Wnek C, Kotus BS, Roemer N, MacTaggart T, et al. (1999) Genotyping new BXD recombinant inbred mouse strains and comparison of BXD and consensus maps. *Mamm Genome* 10: 335–348.
66. Belknap JK (1998) Effect of within-strain sample size on QTL detection and mapping using recombinant inbred mouse strains. *Behav Genet* 28: 29–38.
67. Wang X, Agarwala R, Capra JA, Chen Z, Church DM, et al. (2010) High-throughput sequencing of the DBA/2J mouse genome. *BMC Bioinformatics* 11: O7.
68. Martínez I, Wallace G, Zhang C, Legge R, Benson AK, et al. (2009) Diet-induced metabolic improvements in a hamster model of hypercholesterolemia are strongly linked to alterations of the gut microbiota. *Appl Environ Microbiol* 75: 4175–4184.
69. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7: 335–336.
70. Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121: 185–199.
71. Storey JD (2002) A direct approach to false discovery rates. *Journal of the Royal Statistical Society Series B*, 64, 479–498.