Variation in gut microbiota composition impacts host gene expression

Allison L Richards [1], Amanda L Muehlbauer [3,4], Adnan Alazizi [1], Michael B Burns [3,4], Trevor J Gould [3,4], Camilla Cascardo [1], Roger Pique-Regi [1,2,5], Ran Blekhman [3,4,5], Francesca Luca [1,2,5]

[1]Center for Molecular Medicine and Genetics, and

[2]Department of Obstetrics and Gynecology, Wayne State University, Detroit, Michigan 48201, USA

[3]Department of Genetics, Cell Biology and Development, and

[4]Department of Ecology, Evolution and Behavior, University of Minnesota, Minneapolis, Minnesota, USA

[5]To whom correspondence should be addressed.

Email: fluca@wayne.edu, blekhman@umn.edu, rpique@wayne.edu

October 27, 2017

Abstract

Variation in gut microbiome is associated with wellness and disease in humans, yet the molecular mechanisms by which this variation affects the host are not well understood. A likely mechanism is through changing gene regulation in interfacing host epithelial cells. Here, we treated colonic epithelial cells with live microbiota from five healthy individuals and quantified transcriptional response in host cells. We identified over 5,000 host genes that change expression, including 588 distinct associations between specific taxa and host genes. The taxa with the strongest influence on gene expression, alter the response of genes associated with complex traits. We then created a manipulated microbial community with titrated doses of *Collinsella*, demonstrating that both natural and controlled microbiome composition leads to distinct, and predictable, gene expression profiles in the host. This ability to fine tune the expression of host genes by manipulating the microbiome suggests future therapeutic routes for human wellness.

The microbial community in the human digestive tract, the gut microbiota, is highly complex, and also displays strong variation across individuals [17, 23, 29]. Variability in gut microbiota composition is related to many factors, including medication, diet and genetics [3, 4, 9, 10, 18, 22, 24, 25, 33, 35, 37, 38, 40]. The gut microbiota has a variety of functions within the host, such as metabolism of certain compounds [1, 13, 16, 31], and its composition is correlated with several diseases, such as Crohn's disease and colorectal cancer [5, 14, 21, 28, 32, 34]. In mouse, certain microbial communities can lead to changes in the host's weight and overall health, suggesting that there is a reciprocal effect between the host and the gut microbiota [15]. While recent studies have shown that the gut microbiota can influence host gene expression in mice [6, 8, 11], in humans, our ability to study the effects of the microbiome *in vivo* are severely limited. Recently, we have described an in vitro approach based on human epithelial cells inoculated with live microbial communities [30] that is well suited to study the effects of the microbiome

on human gene regulation. Here we seek to use this in vitro system to determine the extent by which variation in microbiome composition drives differences in gene expression in the host cells. We also seek to determine if specific microbial taxa drive gene expression variation, and if these changes are predictable, i.e. they can be recapitulated by manipulating the composition of the microbiome. These open question are crucial for understanding the causal role of the microbiome in host physiology and designing targeted therapies revolving around interventions on the gut microbiome.

To determine the impact of variation in the gut microbiota on host cells, we treated human colonic epithelial cells with live gut microbiota extracted from 5 healthy, unrelated, human individuals (Figure 1A). These samples are representative of other healthy gut microbiome samples from the Human Microbiome Project (Figure S4) [17, 26]. We then assessed changes in gene expression and microbial composition following 1, 2 and 4 hours of exposure separately. The overall changes in gene expression between each microbiome treatment and control cluster first by time-point (Figure 1B, Table S3) where the strongest response occurs at 2 hours following exposure (3,260 genes across any of the five microbiota samples, BH FDR $< 10\%$, $|\log_2 \mathrm{FC}| > 0.25$). Among these, we identified 669 transcripts (188 genes) that are differentially expressed in all five treatments following 2 hours of treatment (Figure 1C, 1 and 4 hour comparisons in Figure S1A and B). We used meta-analysis to identify genes that change consistently across the treatments and time points (examples in Figure 1D, Figure S2A and B, Table S1) and found that they are enriched for genes that function in protein translation, as well as those on the cell surface, such as in adherens junctions (BH FDR $< 10^{-4}\%$) (Table S2, Figure S3), suggesting a biological function for consistent changes in gene expression that may relate to the host cell's interaction with the microbiota.

Each microbiota sample is derived from a different individual with unique diet and genetic makeup. Therefore, we expect that the microbial composition and diversity of each sample differs. When we considered the uncultured microbiome, we found variability in their microbial composition and diversity (Simpson's index range between 0.94 to 0.98). When we considered how the human colonocytes influenced microbial composition, we found that most taxa were unaffected by the presence of human cells, while 13 taxa showed varying abundance dependent on the presence of host cells (likelihood ratio test, BH FDR $< 10\%$, examples in S5, Table S5). In order to determine how the microbiota composition of each sample influences host gene expression differently, we utilized a likelihood ratio test to compare models including or excluding the individual microbiota effect. We identified 409 genes (1,484 transcripts, BH FDR $< 10\%$, Table S4) with gene expression responses significantly different across colonocytes treated with the five microbiota samples (examples in Figure 1E). These data demonstrate that both the host and the microbiota influence each other and that inter-individual variation in the microbiome can lead to different gene expression responses in interacting host cells.

We hypothesized that the differences in gene expression response to each microbiome could be attributed to specific microbiota features, such as the abundance of specific taxa. For this reason we studied the association between host gene expression (147,555 transcripts) and the abundance of microbial taxa (62 taxa that pass filtering criteria; see Methods and Materials) at the time of treatment. Across all possible associations (9,125,927 tests) we identified 588 transcript-by-taxon pairs with a significant association (BH FDR $= 10\%$, Table S6), corresponding to 121 host genes with changes in expression associated with the abundance of 46 taxa. 35 of these taxa were associated with the expression of more than one host gene (BH FDR $= 10\%$) demonstrating that a single microbe may affect the regulation of many genes, and suggesting that microbes may influence a particular trait in a polygenic manner.

Of the 121 host genes whose expression is associated with abundance of microbial taxa, 70 genes (219 transcripts) were also differentially expressed when each microbiota treatment was compared to control conditions, and formed two clusters with distinct functions (Figure 2A and D, examples in Figure 2B and C). Genes in the first cluster are positively correlated with genera *Ruminococcus*, *Coprococcus* and *Streptococcus*, and have functions in cell junction assembly (BH FDR $< 10^{-3}\%$, Figure 2E), while the

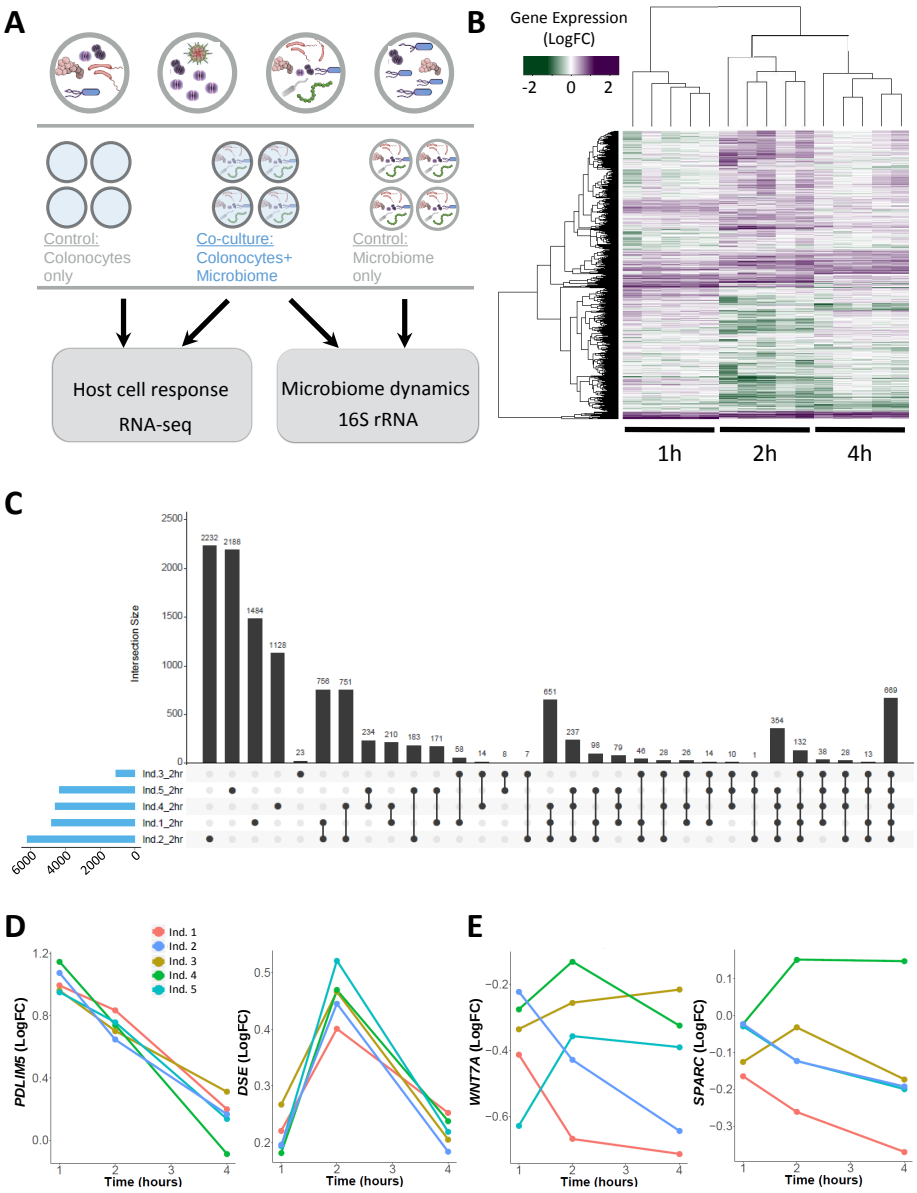

Figure 1: Gene expression changes in Colonocytes Treated with Microbiota from Five Unrelated Individuals. **A** Study design. Human colonocytes were inoculated separately with five microbiota samples from unrelated individuals. **B** Heatmap of gene expression changes induced at each time point by the individual microbiota samples. Purple denotes an increase in gene expression (green shows a reduction) compared to the gene expression in the control (colonocytes cultured alone). Only genes that are differentially expressed in at least one sample are shown. **C** Comparison of transcripts differentially expressed at 2 hours across the five treatments. The blue bars to the left show the total number of differentially expressed transcripts in the given set. The gray vertical bars show the number of transcripts that are in the set denoted below them. Sets with a single dark gray circle show the number of differentially expressed transcripts unique to that sample. **D** Examples of genes (*PDLIM5* and *DSE*) whose changes in expression are consistent across treatments with the five different microbiota. Changes in expression (y-axis) are shown as $\log_2$ fold change as compared to control. **E** Examples of genes (*WNT7A* and *SPARC*) whose changes in expression are significantly different across treatments with the five microbiota samples.

second cluster of genes, positively correlated with microbial genera including *Odoribacter*, *Blautia* and *Collinsella*, function in protein targeting to the endoplasmic reticulum (BH FDR $< 10^{-7}$%, Figure 2F). These results suggest that microbial consortia may work in concert to affect specific functions through changes in host gene regulation.

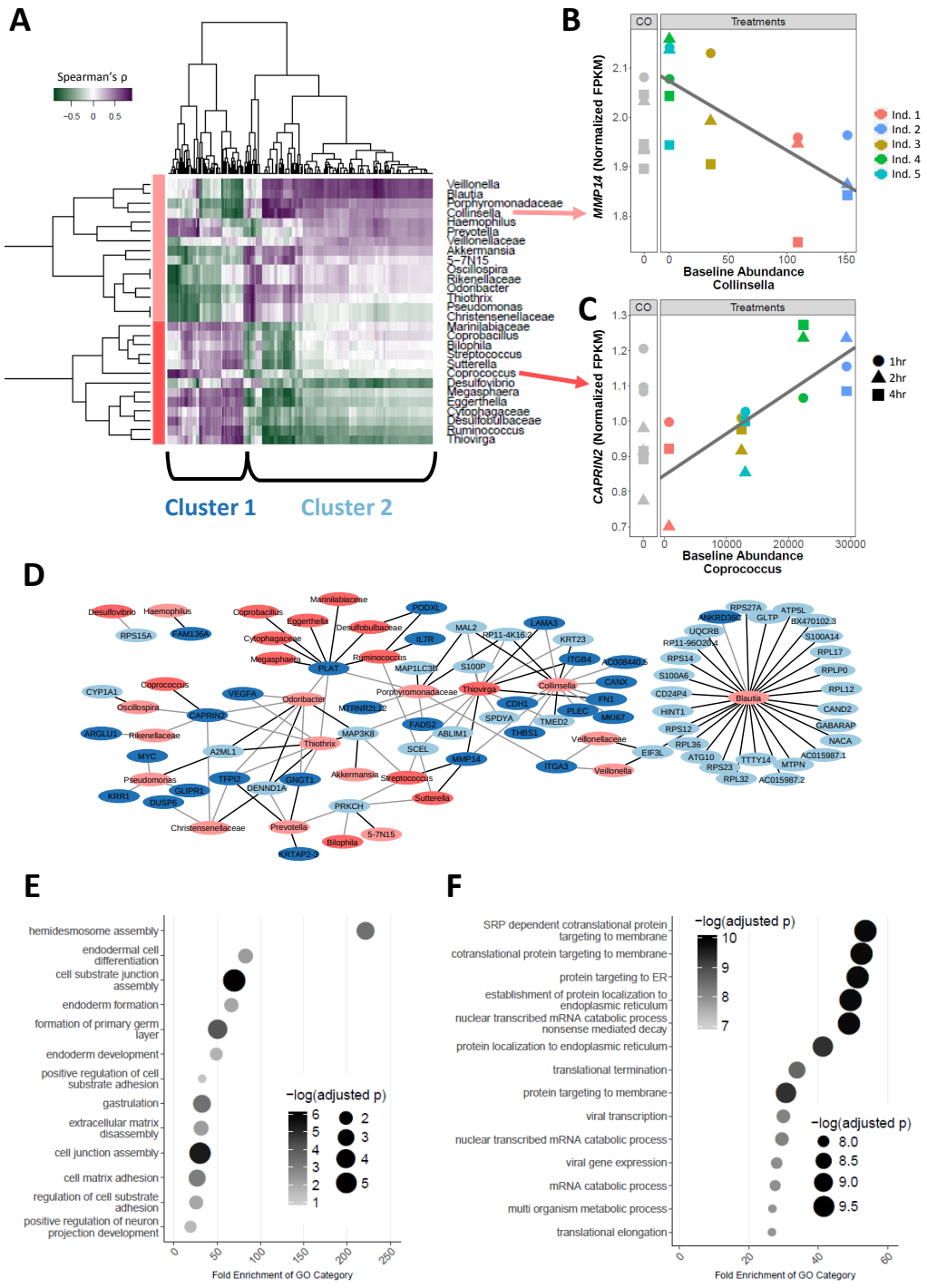

Figure 2: Abundance of microbiome taxa is associated with specific host gene expression changes. A Heatmap of microbiota taxa and colonocyte gene expression correlation (Spearman's ρ). Columns correspond to 28 microbiota taxa and rows correspond to 219 transcripts (70 genes) that had at least one significant correlation in the model. Taxa and transcripts are each clustered via hierarchical clustering showing two major groups indicated by a different shade of red (taxa) / blue (transcripts). **B and C** Examples (*MMP14*, and *CAPRIN2*) of significant association (BH FDR = 7% for both genes) between host gene expression (FPKM quantile normalized) and baseline abundance of specific taxa. **D** Network of associations between taxa and genes from the heatmap in **A**. Nodes in blue denote genes while nodes in red denote microbial taxa. Color shading indicates clusters of genes or taxa from **A**. Black edges indicate a positive correlation while light gray indicates a negative correlation. **E and F** Gene ontology enrichment for cluster 1 and 2 respectively, defined in **A**.

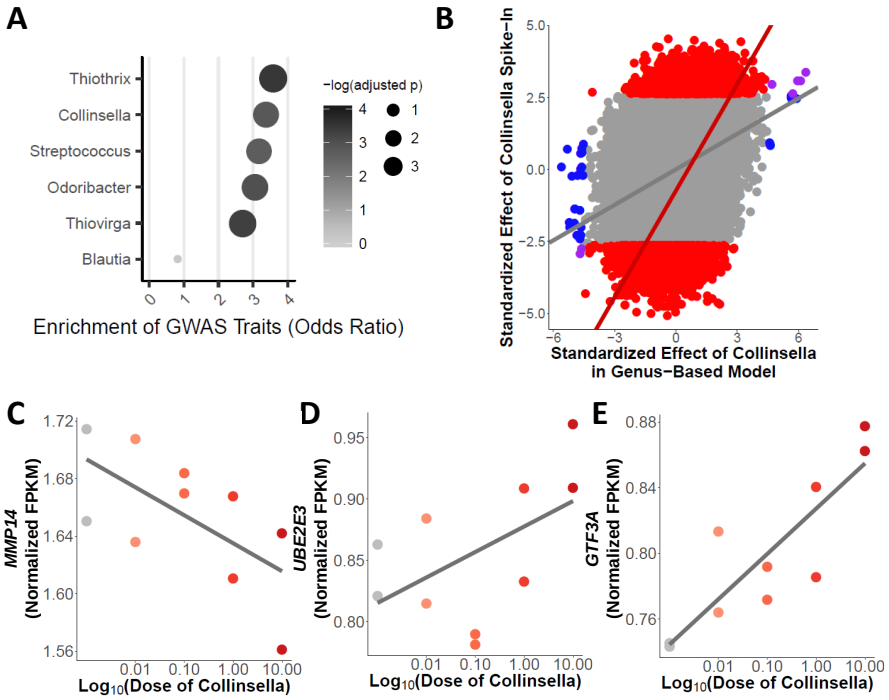

Figure 3: Manipulation of the Microbial Community Induces Predictable Gene Expression Changes in the Host. A Enrichment for complex traits across genes with changes in expression associated with microbial taxa (BH FDR $< 20\%$). **B** Scatterplot of the effect of *Collinsella* abundance from the five microbiome samples (x-axis) and effect of *Collinsella* aerofaciens from the spike-in validation experiment (y-axis). Plotted are $\log_2$ fold changes normalized by the standard error. The red points and line highlight the transcripts that are DE in the spike-in experiment. There is a correlation across all points (p-value $< 10^{-20}$, $\rho = 0.29$) and across transcripts differentially expressed in the spike-in experiment (p-value $< 10^{-20}$, $\rho = 0.46$). **C-E** Examples (*MMP14*, *UBE2E3*, and *GTF3A*) of significant association (BH FDR = 9%, 6% and 7%, respectively) between host gene expression (FPKM quantile normalized) and abundance of *Collinsella* aerofaciens from spike-in validation experiment.

We then focused on microbial taxa associated with changes in expression for a large number of genes, since these microbes are more likely to impact host traits. To test this hypothesis, we focused on the six microbial genera that were associated with the largest number of host genes (at least 30 host genes at p-value $< 3.5 \times 10^{-5}$): *Odoribacter*, *Streptococcus*, *Blautia*, *Thiovirga*, *Thiothrix*, and *Collinsella*. Indeed, all taxa, except for *Blautia*, led to expression changes in genes enriched for complex traits (p-value < 0.005, OR > 2.7, Figure 3A, Table S7) [39]. Moreover, we identified 21 genes that were associated with traits already linked to the gut microbiome, including colorectal cancer [5, 34], obesity [15, 36], and Inflammatory Bowel Disorder (IBD) [2, 12, 20]. Half of the genes (15/30) associated with the genus *Collinsella* are associated with a trait in GWAS (p-value $= 0.001$, OR $= 3.4$, Figure 3A). Previous studies have found that the abundance of *Collinsella* is correlated with several diseases, including colorectal cancer [27], Type 2 Diabetes [7] and irritable bowel syndrome [19]. Interestingly, we identified a gene, *GLTP*, that is involved in glycolipid transfer and has been associated with metabolic syndrome [41], whose expression is influenced by the abundance of the genus *Collinsella* in each of the five microbiota samples (BH FDR = 12.6%). This suggests that microbes of the genus *Collinsella* may influence metabolic syndrome in the host through regulation of genes in the colon, such as *GLTP*. These data also suggest that specific microorganisms, and not simply general exposure to the entire gut microbiota, can lead to changes in many genes' expression. Furthermore, these results support the hypothesis that variation in the abundance of members of the microbiota may influence complex traits.

To validate and further demonstrate the effect of specific microbes on host gene expression, we treated colonocytes with a microbiota sample without any detectable *Collinsella aerofaciens*, and supplemented it with titrated abundances of this bacterium relative to the whole microbiota sample: 0.01%, 0.1%, 1% and 10%. We used RNA-sequencing to study the resulting changes in gene expression, and identified 1,570 genes that change expression (BH FDR = 10%, Table S8) depending on the abundance of *Collinsella aerofaciens*. When we consider the changes in gene expression associated with *Collinsella* abundance in the five microbiota treatments, we found that the effects of *Collinsella* in both experiments are correlated (Figure 3B). We validate 19 out of 29 genes (p-value = 0.0002, OR = 4.1), originally identified (BH FDR = 20%), including *GLTP* and *MMP14* (ENST00000547279, original BH FDR = 7%in Figure 2B, spike-in validation BH FDR = 9% in Figure 3C), demonstrating that *Collinsella* is responsible for changes in the expression of these genes. The large number of genes that change expression in this experiment could be due to several factors, including the increase in power from a larger number of samples. These 1,570 genes are enriched for genes associated with complex traits from GWAS (p-value = 10^{-10}, OR = 1.5, examples in Figure 3D and E) and specifically enriched for genes associated with HDL cholesterol (bonferroni-corrected p-value = 0.018, OR = 2.75) . This spike-in experiment shows that host gene expression can be modulated by changing the abundance of a single bacterial species within the microbiome.

The gut microbiota has recently been associated with several different diseases and disorders [2, 5, 12, 15, 20, 34, 36] but it is still unclear whether inter-individual variation in the microbiome influences host cell gene expression and if this may be a mechanism by which microbiota impact complex human traits. Our results suggest that specific microbes in the microbiome may be important in regulating gene expression in the gut, and that microbes can induce changes to a large number of genes. Furthermore, the host genes that are affected by the microbiome are involved in complex traits, including those that have already been associated with microbiome composition, thereby suggesting a mechanism by which supplementing the microbiome can influence human health. We have shown that by manipulating microbiome composition, we can influence the host cell response. This work and future research will help to determine which microbes may be most beneficial as interventional therapy to improve one's health.

Additional Files

Supplementary File 1 — Supplemental Results.

Supplementary text for materials and methods, additional results, figures and tables.

Acknowledgments

We thank members of the Luca/Pique-Regi group and Blekhman Lab for helpful discussions and comments.

Competing Interests

The authors declare that they have no competing interests.

References

[1] Backhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A., and Gordon, J. I. (2005). Host-bacterial mutualism in the human intestine. *Science*, **307**(5717), 1915–1920.

[2] Bibiloni, R., Mangold, M., Madsen, K. L., Fedorak, R. N., and Tannock, G. W. (2006). The bacteriology of biopsies differs between newly diagnosed, untreated, Crohn's disease and ulcerative colitis patients. *J. Med. Microbiol.*, **55**(Pt 8), 1141–1149.

[3] Blekhman, R., Goodrich, J. K., Huang, K., Sun, Q., Bukowski, R., Bell, J. T., Spector, T. D., Keinan, A., Ley, R. E., Gevers, D., and Clark, A. G. (2015). Host genetic variation impacts microbiome composition across human body sites. *Genome Biol.*, **16**, 191.

[4] Bolnick, D. I., Snowberg, L. K., Hirsch, P. E., Lauber, C. L., Org, E., Parks, B., Lusis, A. J., Knight, R., Caporaso, J. G., and Svanback, R. (2014). Individual diet has sex-dependent effects on vertebrate gut microbiota. *Nat Commun*, **5**, 4500.

[5] Burns, M. B., Lynch, J., Starr, T. K., Knights, D., and Blekhman, R. (2015). Virulence genes are a signature of the microbiome in the colorectal tumor microenvironment. *Genome Med*, **7**(1), 55.

[6] Camp, J. G., Frank, C. L., Lickwar, C. R., Guturu, H., Rube, T., Wenger, A. M., Chen, J., Bejerano, G., Crawford, G. E., and Rawls, J. F. (2014). Microbiota modulate transcription in the intestinal epithelium without remodeling the accessible chromatin landscape. *Genome Res.*, **24**(9), 1504–1516.

[7] Candela, M., Biagi, E., Soverini, M., Consolandi, C., Quercia, S., Severgnini, M., Peano, C., Turroni, S., Rampelli, S., Pozzilli, P., Pianesi, M., Fallucca, F., and Brigidi, P. (2016). Modulation of gut microbiota dysbioses in type 2 diabetic patients by macrobiotic Ma-Pi 2 diet. *Br. J. Nutr.*, **116**(1), 80–93.

[8] Davison, J. M., Lickwar, C. R., Song, L., Breton, G., Crawford, G. E., and Rawls, J. F. (2017). Microbiota regulate intestinal epithelial gene expression by suppressing the transcription factor Hepatocyte nuclear factor 4 alpha. *Genome Res.*, **27**(7), 1195–1206.

[9] Dethlefsen, L. and Relman, D. A. (2011). Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc. Natl. Acad. Sci. U.S.A.*, **108 Suppl 1**, 4554–4561.

[10] Dominianni, C., Sinha, R., Goedert, J. J., Pei, Z., Yang, L., Hayes, R. B., and Ahn, J. (2015). Sex, body mass index, and dietary fiber intake influence the human gut microbiome. *PLoS ONE*, **10**(4), e0124599.

[11] Donohoe, D. R., Garge, N., Zhang, X., Sun, W., O'Connell, T. M., Bunger, M. K., and Bultman, S. J. (2011). The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab.*, **13**(5), 517–526.

[12] Frank, D. N., St Amand, A. L., Feldman, R. A., Boedeker, E. C., Harpaz, N., and Pace, N. R. (2007). Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc. Natl. Acad. Sci. U.S.A.*, **104**(34), 13780–13785.

[13] Fujimura, K. E., Sitarik, A. R., Havstad, S., Lin, D. L., Levan, S., Fadrosh, D., Panzer, A. R., LaMere, B., Rackaityte, E., Lukacs, N. W., Wegienka, G., Boushey, H. A., Ownby, D. R., Zoratti, E. M., Levin, A. M., Johnson, C. C., and Lynch, S. V. (2016). Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nat. Med.*, **22**(10), 1187–1191.

[14] Gevers, D., Kugathasan, S., Denson, L. A., Vazquez-Baeza, Y., Van Treuren, W., Ren, B., Schwager, E., Knights, D., Song, S. J., Yassour, M., Morgan, X. C., Kostic, A. D., Luo, C., Gonzalez, A., McDonald, D., Haberman, Y., Walters, T., Baker, S., Rosh, J., Stephens, M., Heyman, M., Markowitz, J.,

Baldassano, R., Griffiths, A., Sylvester, F., Mack, D., Kim, S., Crandall, W., Hyams, J., Huttenhower, C., Knight, R., and Xavier, R. J. (2014). The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe*, **15**(3), 382–392.

[15] Goodrich, J. K., Waters, J. L., Poole, A. C., Sutter, J. L., Koren, O., Blekhman, R., Beaumont, M., Van Treuren, W., Knight, R., Bell, J. T., Spector, T. D., Clark, A. G., and Ley, R. E. (2014). Human genetics shape the gut microbiome. *Cell*, **159**(4), 789–799.

[16] Hirata, S. I. and Kunisawa, J. (2017). Gut microbiome, metabolome, and allergic diseases. *Allergol Int*.

[17] Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J. H., Chinwalla, A. T., Creasy, H. H., Earl, A. M., FitzGerald, M. G., Fulton, R. S., Giglio, M. G., Hallsworth-Pepin, K., Lobos, E. A., Madupu, R., Magrini, V., Martin, J. C., Mitreva, M., Muzny, D. M., Sodergren, E. J., Versalovic, J., Wollam, A. M., Worley, K. C., Wortman, J. R., Young, S. K., Zeng, Q., Aagaard, K. M., Abolude, O. O., Allen-Vercoe, E., Alm, E. J., Alvarado, L., Andersen, G. L., Anderson, S., Appelbaum, E., Arachchi, H. M., Armitage, G., Arze, C. A., Ayvaz, T., Baker, C. C., Begg, L., Belachew, T., Bhonagiri, V., Bihan, M., Blaser, M. J., Bloom, T., Bonazzi, V., Brooks, J., Buck, G. A., Buhay, C. J., Busam, D. A., Campbell, J. L., Canon, S. R., Cantarel, B. L., Chain, P. S., Chen, I. M., Chen, L., Chhibba, S., Chu, K., Ciulla, D. M., Clemente, J. C., Clifton, S. W., Conlan, S., Crabtree, J., Cutting, M. A., Davidovics, N. J., Davis, C. C., DeSantis, T. Z., Deal, C., Delehaunty, K. D., Dewhirst, F. E., Deych, E., Ding, Y., Dooling, D. J., Dugan, S. P., Dunne, W. M., Durkin, A., Edgar, R. C., Erlich, R. L., Farmer, C. N., Farrell, R. M., Faust, K., Feldgarden, M., Felix, V. M., Fisher, S., Fodor, A. A., Forney, L. J., Foster, L., Di Francesco, V., Friedman, J., Friedrich, D. C., Fronick, C. C., Fulton, L. L., Gao, H., Garcia, N., Giannoukos, G., Giblin, C., Giovanni, M. Y., Goldberg, J. M., Goll, J., Gonzalez, A., Griggs, A., Gujja, S., Haake, S. K., Haas, B. J., Hamilton, H. A., Harris, E. L., Hepburn, T. A., Herter, B., Hoffmann, D. E., Holder, M. E., Howarth, C., Huang, K. H., Huse, S. M., Izard, J., Jansson, J. K., Jiang, H., Jordan, C., Joshi, V., Katancik, J. A., Keitel, W. A., Kelley, S. T., Kells, C., King, N. B., Knights, D., Kong, H. H., Koren, O., Koren, S., Kota, K. C., Kovar, C. L., Kyrpides, N. C., La Rosa, P. S., Lee, S. L., Lemon, K. P., Lennon, N., Lewis, C. M., Lewis, L., Ley, R. E., Li, K., Liolios, K., Liu, B., Liu, Y., Lo, C. C., Lozupone, C. A., Lunsford, R., Madden, T., Mahurkar, A. A., Mannon, P. J., Mardis, E. R., Markowitz, V. M., Mavromatis, K., McCorrison, J. M., McDonald, D., McEwen, J., McGuire, A. L., McInnes, P., Mehta, T., Mihindukulasuriya, K. A., Miller, J. R., Minx, P. J., Newsham, I., Nusbaum, C., O'Laughlin, M., Orvis, J., Pagani, I., Palaniappan, K., Patel, S. M., Pearson, M., Peterson, J., Podar, M., Pohl, C., Pollard, K. S., Pop, M., Priest, M. E., Proctor, L. M., Qin, X., Raes, J., Ravel, J., Reid, J. G., Rho, M., Rhodes, R., Riehle, K. P., Rivera, M. C., Rodriguez-Mueller, B., Rogers, Y. H., Ross, M. C., Russ, C., Sanka, R. K., Sankar, P., Sathirapongsasuti, J., Schloss, J. A., Schloss, P. D., Schmidt, T. M., Scholz, M., Schriml, L., Schubert, A. M., Segata, N., Segre, J. A., Shannon, W. D., Sharp, R. R., Sharpton, T. J., Shenoy, N., Sheth, N. U., Simone, G. A., Singh, I., Smillie, C. S., Sobel, J. D., Sommer, D. D., Spicer, P., Sutton, G. G., Sykes, S. M., Tabbaa, D. G., Thiagarajan, M., Tomlinson, C. M., Torralba, M., Treangen, T. J., Truty, R. M., Vishnivetskaya, T. A., Walker, J., Wang, L., Wang, Z., Ward, D. V., Warren, W., Watson, M. A., Wellington, C., Wetterstrand, K. A., White, J. R., Wilczek-Boney, K., Wu, Y., Wylie, K. M., Wylie, T., Yandava, C., Ye, L., Ye, Y., Yooseph, S., Youmans, B. P., Zhang, L., Zhou, Y., Zhu, Y., Zoloth, L., Zucker, J. D., Birren, B. W., Gibbs, R. A., Highlander, S. K., Methe, B. A., Nelson, K. E., Petrosino, J. F., Weinstock, G. M., Wilson, R. K., and White, O. (2012). Structure, function and diversity of the healthy human microbiome. *Nature*, **486**(7402), 207–214.

[18] Jernberg, C., Lofmark, S., Edlund, C., and Jansson, J. K. (2010). Long-term impacts of antibiotic exposure on the human intestinal microbiota. *Microbiology (Reading, Engl.)*, **156**(Pt 11), 3216–3223.

[19] Kassinen, A., Krogius-Kurikka, L., Makivuokko, H., Rinttila, 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**(1), 24–33.

[20] Knights, D., Silverberg, M. S., Weersma, R. K., Gevers, D., Dijkstra, G., Huang, H., Tyler, A. D., van Sommeren, S., Imhann, F., Stempak, J. M., Huang, H., Vangay, P., Al-Ghalith, G. A., Russell, C., Sauk, J., Knight, J., Daly, M. J., Huttenhower, C., and Xavier, R. J. (2014). Complex host genetics influence the microbiome in inflammatory bowel disease. *Genome Med*, **6**(12), 107.

[21] Larsen, N., Vogensen, F. K., van den Berg, F. W., Nielsen, D. S., Andreasen, A. S., Pedersen, B. K., Al-Soud, W. A., Sørensen, S. J., Hansen, L. H., and Jakobsen, M. (2010). Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PLoS ONE*, **5**(2), e9085.

[22] Lee, S., Sung, J., Lee, J., and Ko, G. (2011). Comparison of the gut microbiotas of healthy adult twins living in South Korea and the United States. *Appl. Environ. Microbiol.*, **77**(20), 7433–7437.

[23] Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K., and Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature*, **489**(7415), 220–230.

[24] Makivuokko, H., Tiihonen, K., Tynkkynen, S., Paulin, L., and Rautonen, N. (2010). The effect of age and non-steroidal anti-inflammatory drugs on human intestinal microbiota composition. *Br. J. Nutr.*, **103**(2), 227–234.

[25] Maurice, C. F., Haiser, H. J., and Turnbaugh, P. J. (2013). Xenobiotics shape the physiology and gene expression of the active human gut microbiome. *Cell*, **152**(1-2), 39–50.

[26] Methe, B. A., Nelson, K. E., Pop, M., Creasy, H. H., Giglio, M. G., Huttenhower, C., Gevers, D., Petrosino, J. F., Abubucker, S., Badger, J. H., Chinwalla, A. T., Earl, A. M., FitzGerald, M. G., Fulton, R. S., Hallsworth-Pepin, K., Lobos, E. A., Madupu, R., Magrini, V., Martin, J. C., Mitreva, M., Muzny, D. M., Sodergren, E. J., Versalovic, J., Wollam, A. M., Worley, K. C., Wortman, J. R., Young, S. K., Zeng, Q., Aagaard, K. M., Abolude, O. O., Allen-Vercoe, E., Alm, E. J., Alvarado, L., Andersen, G. L., Anderson, S., Appelbaum, E., Arachchi, H. M., Armitage, G., Arze, C. A., Ayvaz, T., Baker, C. C., Begg, L., Belachew, T., Bhonagiri, V., Bihan, M., Blaser, M. J., Bloom, T., Bonazzi, V. R., Brooks, P., Buck, G. A., Buhay, C. J., Busam, D. A., Campbell, J. L., Canon, S. R., Cantarel, B. L., Chain, P. S., Chen, I. M., Chen, L., Chhibba, S., Chu, K., Ciulla, D. M., Clemente, J. C., Clifton, S. W., Conlan, S., Crabtree, J., Cutting, M. A., Davidovics, N. J., Davis, C. C., DeSantis, T. Z., Deal, C., Delehaunty, K. D., Dewhirst, F. E., Deych, E., Ding, Y., Dooling, D. J., Dugan, S. P., Dunne, W., Durkin, A., Edgar, R. C., Erlich, R. L., Farmer, C. N., Farrell, R. M., Faust, K., Feldgarden, M., Felix, V. M., Fisher, S., Fodor, A. A., Forney, L., Foster, L., Di Francesco, V., Friedman, J., Friedrich, D. C., Fronick, C. C., Fulton, L. L., Gao, H., Garcia, N., Giannoukos, G., Giblin, C., Giovanni, M. Y., Goldberg, J. M., Goll, J., Gonzalez, A., Griggs, A., Gujja, S., Haas, B. J., Hamilton, H. A., Harris, E. L., Hepburn, T. A., Herter, B., Hoffmann, D. E., Holder, M. E., Howarth, C., Huang, K. H., Huse, S. M., Izard, J., Jansson, J. K., Jiang, H., Jordan, C., Joshi, V., Katancik, J. A., Keitel, W. A., Kelley, S. T., Kells, C., Kinder-Haake, S., King, N. B., Knight, R., Knights, D., Kong, H. H., Koren, O., Koren, S., Kota, K. C., Kovar, C. L., Kyrpides, N. C., La Rosa, P. S., Lee, S. L., Lemon, K. P., Lennon, N., Lewis, C. M., Lewis, L., Ley, R. E., Li, K., Liolios, K., Liu, B., Liu, Y., Lo, C. C., Lozupone, C. A., Lunsford, R., Madden, T., Mahurkar, A. A., Mannon, P. J., Mardis, E. R., Markowitz, V. M., Mavrommatis, K., McCorrison, J. M., McDonald, D., McEwen, J., McGuire, A. L., McInnes, P., Mehta, T., Mihindukulasuriya, K. A., Miller, J. R., Minx, P. J., Newsham, I., Nusbaum, C., O'Laughlin, M.,

Orvis, J., Pagani, I., Palaniappan, K., Patel, S. M., Pearson, M., Peterson, J., Podar, M., Pohl, C., Pollard, K. S., Priest, M. E., Proctor, L. M., Qin, X., Raes, J., Ravel, J., Reid, J. G., Rho, M., Rhodes, R., Riehle, K. P., Rivera, M. C., Rodriguez-Mueller, B., Rogers, Y. H., Ross, M. C., Russ, C., Sanka, R. K., Sankar, P., Sathirapongsasuti, J., Schloss, J. A., Schloss, P. D., Schmidt, T. M., Scholz, M., Schriml, L., Schubert, A. M., Segata, N., Segre, J. A., Shannon, W. D., Sharp, R. R., Sharpton, T. J., Shenoy, N., Sheth, N. U., Simone, G. A., Singh, I., Smillie, C. S., Sobel, J. D., Sommer, D. D., Spicer, P., Sutton, G. G., Sykes, S. M., Tabbaa, D. G., Thiagarajan, M., Tomlinson, C. M., Torralba, M., Treangen, T. J., Truty, R. M., Vishnivetskaya, T. A., Walker, J., Wang, L., Wang, Z., Ward, D. V., Warren, W., Watson, M. A., Wellington, C., Wetterstrand, K. A., White, J. R., Wilczek-Boney, K., Wu, Y. Q., Wylie, K. M., Wylie, T., Yandava, C., Ye, L., Ye, Y., Yooseph, S., Youmans, B. P., Zhang, L., Zhou, Y., Zhu, Y., Zoloth, L., Zucker, J. D., Birren, B. W., Gibbs, R. A., Highlander, S. K., Weinstock, G. M., Wilson, R. K., and White, O. (2012). A framework for human microbiome research. *Nature*, **486**(7402), 215–221.

[27] Moore, W. E. and Moore, L. H. (1995). Intestinal floras of populations that have a high risk of colon cancer. *Appl. Environ. Microbiol.*, **61**(9), 3202–3207.

[28] Ott, S. J., Musfeldt, M., Wenderoth, D. F., Hampe, J., Brant, O., Folsch, U. R., Timmis, K. N., and Schreiber, S. (2004). Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. *Gut*, **53**(5), 685–693.

[29] Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D. R., Li, J., Xu, J., Li, S., Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Xie, Y., Tap, J., Lepage, P., Bertalan, M., Batto, J. M., Hansen, T., Le Paslier, D., Linneberg, A., Nielsen, H. B., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H., Yu, C., Li, S., Jian, M., Zhou, Y., Li, Y., Zhang, X., Li, S., Qin, N., Yang, H., Wang, J., Brunak, S., Dore, J., Guarner, F., Kristiansen, K., Pedersen, O., Parkhill, J., Weissenbach, J., Bork, P., Ehrlich, S. D., Wang, J., Antolin, M., Artiguenave, F., Blottiere, H., Borruel, N., Bruls, T., Casellas, F., Chervaux, C., Cultrone, A., Delorme, C., Denariaz, G., Dervyn, R., Forte, M., Friss, C., van de Guchte, M., Guedon, E., Haimet, F., Jamet, A., Juste, C., Kaci, G., Kleerebezem, M., Knol, J., Kristensen, M., Layec, S., Le Roux, K., Leclerc, M., Maguin, E., Minardi, R. M., Oozeer, R., Rescigno, M., Sanchez, N., Tims, S., Torrejon, T., Varela, E., de Vos, W., Winogradsky, Y., and Zoetendal, E. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*, **464**(7285), 59–65.

[30] Richards, A. L., Burns, M. B., Alazizi, A., Barreiro, L. B., Pique-Regi, R., Blekhman, R., and Luca, F. (2016). Genetic and transcriptional analysis of human host response to healthy gut microbiota. *mSystems*, **1**(4).

[31] Rowland, I., Gibson, G., Heinken, A., Scott, K., Swann, J., Thiele, I., and Tuohy, K. (2017). Gut microbiota functions: metabolism of nutrients and other food components. *Eur J Nutr*.

[32] Schwiertz, A., Taras, D., Schafer, K., Beijer, S., Bos, N. A., Donus, C., and Hardt, P. D. (2010). Microbiota and SCFA in lean and overweight healthy subjects. *Obesity (Silver Spring)*, **18**(1), 190–195.

[33] Seto, C. T., Jeraldo, P., Orenstein, R., Chia, N., and DiBaise, J. K. (2014). Prolonged use of a proton pump inhibitor reduces microbial diversity: implications for Clostridium difficile susceptibility. *Microbiome*, **2**, 42.

[34] Sinha, R., Ahn, J., Sampson, J. N., Shi, J., Yu, G., Xiong, X., Hayes, R. B., and Goedert, J. J. (2016). Fecal Microbiota, Fecal Metabolome, and Colorectal Cancer Interrelations. *PLoS ONE*, **11**(3), e0152126.

[35] Tims, S., Derom, C., Jonkers, D. M., Vlietinck, R., Saris, W. H., Kleerebezem, M., de Vos, W. M., and Zoetendal, E. G. (2013). Microbiota conservation and BMI signatures in adult monozygotic twins. *ISME J*, **7**(4), 707–717.

[36] Turnbaugh, P. J., Backhed, F., Fulton, L., and Gordon, J. I. (2008). Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe*, **3**(4), 213–223.

[37] Turnbaugh, P. J., Hamady, M., Yatsunenko, T., Cantarel, B. L., Duncan, A., Ley, R. E., Sogin, M. L., Jones, W. J., Roe, B. A., Affourtit, J. P., Egholm, M., Henrissat, B., Heath, A. C., Knight, R., and Gordon, J. I. (2009a). A core gut microbiome in obese and lean twins. *Nature*, **457**(7228), 480–484.

[38] Turnbaugh, P. J., Ridaura, V. K., Faith, J. J., Rey, F. E., Knight, R., and Gordon, J. I. (2009b). The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci Transl Med*, **1**(6), 6ra14.

[39] Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L., and Parkinson, H. (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.*, **42**(Database issue), D1001–1006.

[40] Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R. N., Anokhin, A. P., Heath, A. C., Warner, B., Reeder, J., Kuczynski, J., Caporaso, J. G., Lozupone, C. A., Lauber, C., Clemente, J. C., Knights, D., Knight, R., and Gordon, J. I. (2012). Human gut microbiome viewed across age and geography. *Nature*, **486**(7402), 222–227.

[41] Zabaneh, D. and Balding, D. J. (2010). A genome-wide association study of the metabolic syndrome in Indian Asian men. *PLoS ONE*, **5**(8), e11961.