RECENT DISCOVERIES IN EVOLUTIONARY AND GENOMIC MICROBIOLOGY

EDITED BY : Anton G. Kutikhin and Arseniy E. Yuzhalin
PUBLISHED IN : Frontiers in Microbiology

frontiers Research Topics

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: **researchtopics@frontiersin.org**

RECENT DISCOVERIES IN EVOLUTIONARY AND GENOMIC MICROBIOLOGY

Topic Editors:
Anton G. Kutikhin, Research Institute for Complex Issues of Cardiovascular Diseases, Russian Federation
Arseniy E. Yuzhalin, Cancer Research UK and Medical Research Council Oxford Institute for Radiation Oncology, University of Oxford, UK
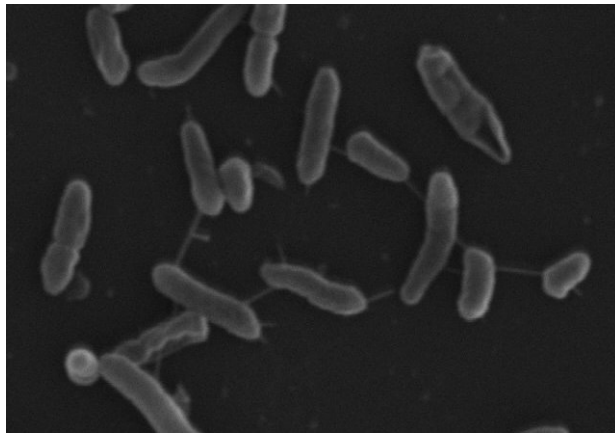
Scanning electron microscopy image of Serratia marsescens, x15,000. Copyright: Anton Kutikhin.

This collection represents certain discoveries that were made in evolutionary and genomic microbiology during the recent ten years. We attempted to shed light on topical issues of microbial evolution and microbiome biology. In our eyes, these articles are of an excellent quality and may be helpful both for casual readers and for specialists in the field.

Table of Contents

Editorial: recent discoveries in evolutionary and genomic microbiology

Anton G. Kutikhin[1]* and Arseniy E. Yuzhalin[2]

[1] Laboratory for Genomic Medicine, Division of Experimental and Clinical Cardiology, Research Institute for Complex Issues of Cardiovascular Diseases, Kemerovo, Russia, [2] Cancer Research UK/MRC Oxford Institute for Radiation Oncology, Department of Oncology, University of Oxford, Oxford, UK

According to the current knowledge, the human body can be viewed as a superorganism comprised of human cells and resident microbial communities called microbiota (Ley et al., 2008). Human microbiota has a wide range of physiological functions, playing an important role in digestion, immunity, and production of certain vitamins; however, most of these microorganisms cannot be cultured in laboratory conditions, and until recently it has been the main challenge for clear determination of the composition of human microbiome (Tlaskalová-Hogenová et al., 2011). New discoveries in methods of molecular biology, including the emergence of omics technologies, have provided a possibility for deciphering the qualitative and quantitative composition of the microbiome (Tlaskalová-Hogenová et al., 2011). This breakthrough has also stimulated in-depth investigation of host-microbiota interactions (Tlaskalová-Hogenová et al., 2011). Basic comparative research revealed that each individual has a unique microbiota, which composition is largely determined during the first years of life but can be altered, reversibly or irreversibly, by a number of factors, such as age, environment, diet, drugs, diseases, and others (Zoetendal et al., 2008).

Therefore, the question of human microbiome biology is extremely intriguing, and large amount of research has been published during the recent years. The investigation of gut, oral, respiratory, skin, vaginal, urinary microbiomes is gaining increasing interest and more attention with time (Sommer and Bäckhed, 2013; Belkaid and Segre, 2014; Xu and Gunsolley, 2014; van de Wijgert et al., 2014; Rogers et al., 2015; Shreiner et al., 2015; Whiteside et al., 2015). We have opened this Research Topic with the idea to provide an outlook for discoveries that have become milestones in the field.

We sincerely thank all researchers who have agreed to contribute to our Research Topic. This collection is divided into three sections. The first one includes three articles in which Eric Bapteste, Christopher House, Matteo Pellegrini, and Sorel Fitz-Gibbon discuss the general issues of microbial evolution whilst Arshan Nasir, Patrick Forterre, Kyung Mo Kim, and Gustavo Caetano-Anolles analyze the distribution of viruses and their impact on evolution of organisms.

The second section is devoted to the biology of human microbiome. We highly recommend an excellent article of Sara Quercia and colleagues who describe the timescales of human gut microbiota adaptation and talk over gut microbiota plasticity that ≪was strategic to face changes in lifestyle and dietary habits along the course of the recent evolutionary history, that has driven the passage from Paleolithic hunter-gathering societies to Neolithic agricultural farmers to modern Westernized societies≫. The article by Noah Voreades, Anne Kozil, and Tiffany Weir extends this topic, focusing primarily on diet as ≪one of the most pivotal factors in the development of the human gut microbiome from infancy to the elderly ≫. Readers interested in biology of aging will appreciate the paper by Sitaraman Saraswati and Ramakrishnan Sitaraman on the causative role of gut microbiota in this process. As immune and neuroendocrine

systems maturate together with gut microbes, we are glad to present a review by Sahar El Aidy, Timothy Dinan, and John Cryan who comprehensively analyze this topic. Gut microbiota also plays a major role in non-communicable diseases, for example, obesity and type 2 diabetes mellitus, and Isabel Moreno-Indias, Fernando Cardona, Francisco Tinahones along with Maria Isabel Queipo-Ortuno present a good story on it. Regarding original research and methods, Eamonn Culligan, Julian Marchesi, Colin Hill, and Roy Sleator demonstrate how the combination of metagenomic and phenomic approaches helps us to identify novel genes within the gut microbiome. Further, Hui Chen and Wen Jiang discuss the role of the oral microbiome in health and diseases. At last, there are several papers on the human mycobiome, and Ablishek Saxena along with Ramakrishnan Sitaraman shed light on the aspects of its osmoregulation.

The last piece of the collection is composed by articles on other topics. Milton Saier and Zhongge Zhang discuss an intriguing principle of directed mutation, Ramakrishnan Sitaraman talks about *Helicobacter pylori,* DNA methyltransferases and the epigenetic field effect in cancerization, and Anton Kutikhin with colleagues summarize recent discoveries on calcifying nanoparticles sometimes referred to as nanobacteria.

We created this Research Topic with the hope that it will be useful for a wide audience, particularly immunologists, microbiologists, graduate, and undergraduate students of biomedical faculties as well as their lecturers.

References

Belkaid, Y., and Segre, J. A. (2014). Dialogue between skin microbiota and immunity. *Science* 346, 954–959. doi: 10.1126/science.1260144

Ley, R. E., Hamady, M., Lozupone, C., Turnbaugh, P. J., Ramey, R. R., Bircher, J. S., et al. (2008). Evolution of mammals and their gut microbes. *Science* 320, 1647–1651. doi: 10.1126/science.1155725

Rogers, G. B., Shaw, D., Marsh, R. L., Carroll, M. P., Serisier, D. J., and Bruce, K. D. (2015). Respiratory microbiota: addressing clinical questions, informing clinical practice. *Thorax* 70, 74–81. doi: 10.1136/thoraxjnl-2014-205826

Shreiner, A. B., Kao, J. Y., and Young, V. B. (2015). The gut microbiome in health and in disease. *Curr. Opin. Gastroenterol.* 31, 69–75. doi: 10.1097/MOG.0000000000000139

Sommer, F., and Bäckhed, F. (2013). The gut microbiota–masters of host development and physiology. *Nat. Rev. Microbiol.* 11, 227–238. doi: 10.1038/nrmicro2974

Tlaskalová-Hogenová, H., Štěpánková, R., Kozáková, H., Hudcovic, T., Vannucci, L., Tučková, L., et al. (2011). The role of gut microbiota (commensal bacteria) and the mucosal barrier in the pathogenesis of inflammatory and autoimmune diseases and cancer: contribution of germ-free and gnotobiotic animal models of human diseases. *Cell. Mol. Immunol.* 8, 110–120. doi: 10.1038/cmi.2010.67

van de Wijgert, J. H., Borgdorff, H., Verhelst, R., Crucitti, T., Francis, S., Verstraelen, H., et al. (2014). The vaginal microbiota: what have we learned after a decade of molecular characterization? *PLoS ONE* 9:e105998. doi: 10.1371/journal.pone.0105998

Whiteside, S. A., Razvi, H., Dave, S., Reid, G., and Burton, J. P. (2015). The microbiome of the urinary tract-a role beyond infection. *Nat. Rev. Urol.* 12, 81–90. doi: 10.1038/nrurol.2014.361

Xu, P., and Gunsolley, J. (2014). Application of metagenomics in understanding oral health and disease. *Virulence* 5, 424–432. doi: 10.4161/viru.28532

Zoetendal, E. G., Rajilic-Stojanovic, M., and de Vos, W. M. (2008). High-throughput diversity and functionality analysis of the gastrointestinal tract microbiota. *Gut* 57, 1605–1615. doi: 10.1136/gut.2007.133603

The origins of microbial adaptations: how introgressive descent, egalitarian evolutionary transitions and expanded kin selection shape the network of life

Eric Bapteste[1,2]*

[1] UPMC, Institut de Biologie Paris Seine, UMR7138 'Evolution Paris Seine', Paris, France
[2] CNRS, Institut de Biologie Paris Seine, UMR7138 'Evolution Paris Seine', Paris, France
*Correspondence: eric.bapteste@snv.jussieu.fr

MICROBIAL EVOLUTION: A GENEALOGICAL PERSPECTIVE

Protists, bacteria and archaea (the prokaryotes), and their mobile genetic elements populate the microbial world. This world is ancient (several billion years old), numerically huge (with 5×10^{30} prokaryotic cells and 10–100 times more viruses!), and genetically extremely diversified. Such a large assemblage cannot be ignored in attempts to understand life's history on Earth, however, how can biologists account for its evolution?

For long, the notion of descent with modification, describing a process of vertical inheritance, defining a tree-like genealogical pattern, when the genetic material, modified by some mutations, is transferred from the genome of a last common ancestor to its direct progeny has offered a promising way to classify organisms and species. Thus, the origin of microbial adaptations can be searched for within lineages, in the changes of genetic material inherited from one ancestor. Yet, such studies are strongly constrained. Non-tree like evolution, generating a reticulate evolutionary pattern, cannot be analyzed with a genealogical tree. Moreover, viruses do not all originate from a single last ancestor (Lima-Mendez et al., 2008), nor do they all display obvious genealogical relationships with cellular organisms, hindering the collective study of mobile genetic elements and cells with a single tree.

Genealogy also plays a central role for explaining the main types of behaviors described in the biological world:

selfishness, mutualism, altruism, and spite (West et al., 2006). The evolution of these interactions can be understood by accounting for kinship between protagonists, under the standard assumption that genealogical proximity between individuals entails their genetic proximity (Huneman, 2013). Thus, knowing the relative kinship, the benefit for the recipient of an interaction and the cost for its actor allows determining when an individual cooperating with a kin, in ways enhancing its reproduction, or preventing distantly related members of a population to reproduce, actually maximizes the reproductive success and the survival of its own genes (van Baalen, 2013). Microbiologists are thus inclined to embrace the conceptual framework of kin selection to analyze many cooperations (Diggle et al., 2007).

However, it is probably not enough to take genealogical relationships into account to explain the diversity, evolution, and interactions in the microbial world. Too systematic a focus on genealogy may even introduce some biases in the explanations of microbial diversity, evolution and interactions, because many crucial biological phenomena result from processes orthogonal to vertical descent.

EXPANDING THE ANALYTICAL FRAMEWORK INSPIRED BY GENEALOGY

The multiplicity of evolutionary processes, their consequences, and the interactions at play within the microbial world are still under-appreciated. The genealogical perspective grounding evolutionary

explanations needs to be completed, because its analytical framework does not accommodate for numerous important biological phenomena, which deeply challenge our background knowledge of the (microbial) world and its evolution.

INTROGRESSION: A CLASS OF NON-TREE-LIKE EVOLUTIONARY PROCESSES

Due to introgressive descent, many adaptations originate from outside rather than from within lineages of vertical descent. While in vertical descent, the genetic material of a particular evolutionary unit is propagated by replication inside its own lineage, in introgressive descent, the genetic material of a particular evolutionary unit propagates into different host structures and is replicated within these host structures (Bapteste et al., 2012). Such host structures are genealogically composite, made of components with distinct genealogical origins. Importantly, introgression is very common in the microbial world, affecting entities from the same or different levels of biological organization. New introgressive mechanisms are constantly discovered (Bapteste, 2013). However, these mechanisms and their actors (viruses, plasmids, conjugative elements, outer membrane vesicles, gene transfer agents, nanotubes, membrane fusion, …) are largely missing from the traditional evolutionary representation. For instance, a gene sequence can propagate into another gene sequence, creating a novel composite gene, whose components come from two different gene lineages. Similarly, a gene sequence

can propagate within a genome, whose ancestor lacked this gene, producing a composite genome with genes originating from different genomes. Likewise, the genome of a mobile element (a virus, a plasmid, etc.) can propagate into a cell born without this element, creating a composite cell with genetic instruction from multiple sources. Or, a part or an entire microbial genome can propagate within a symbiotic association, producing a holobiont with several unrelated genomes. Therefore, the recognition of introgression promotes a substantial expansion of the evolutionary research program: a study of the origins (rather than of the origin) of adaptations and species through the description and analysis of a plurality of processes and objects, some unexpected in the traditional genealogical perspective (Doolittle and Bapteste, 2007; Bapteste, 2013).

GENE–GENE INTROGRESSION: MASSIVE GENE REMODELING

Homology guides comparative analysis in evolutionary biology (Haggerty et al., 2013). Sequences or organs are considered homologous when they evolved in a tree-like fashion from an ancestral form. Thus, gene evolution is often described by a tree with one genealogy per gene family. However, many genes originate from the composition of genetic material from sequences belonging to different gene families. Eukaryotes are the main creators of composite genes (in terms of the proportion of composite genes in their genomes) (Haggerty et al., 2013), yet in terms of absolute numbers, mobile genetic elements operate the most massive gene remodeling on Earth (Jachiet et al., 2013). Therefore, numerous genes display family resemblances (Halary et al., 2013): true similarity caused by introgression between non-homologous sequences. Such family resemblances support the study of the origins of genes and of adaptations at a more global scale than delineated by homology.

GENE–GENOME INTROGRESSION: REMARKABLE PANGENOMES

All conspecific individuals do not own the same gene families. Six percent only of their gene families are distributed in all 60 strains of *Escherichia coli* (Lukjancenko et al., 2010), and experiments showed that

only 61 genes out of 246,065 cannot be transferred into an *E. coli* (Sorek et al., 2007). Members of this species (and many others) exploit a large DNA pool, called pangenome, larger than the size of individual genomes. Therefore, sequencing one individual genome does not always allow describing genetic and functional diversity at the species level. Pangenomes and lateral gene transfer –by small segments or larger chunks- are not restricted to conspecifics (Nelson-Sathi et al., 2012). These observations challenge phylogenetic systematics: they mean that genome evolution is much more than genome genealogy, an increasingly elusive concept, since these objects prove to be ever more composite and their genes do not all coalesce in a single common ancestral genome.

FROM THE MOBILOME NETWORK TO THE SOCIAL NETWORK OF LIFE

Many classes of evolutionary objects (i.e., virus, plasmids, etc.) have fuzzy borders, because many of these objects do not evolve independently at the genetic level. Remarkably, introgression creates novel introgressive mechanisms. Numerous genealogically mosaic mobile elements (autonomous or not: polintons, virophages, R391, phasmides, phage inducible chromosomal islands, transpovirons, etc.) emerge and evolve through the sharing of mobility functions, defining a genetic pool: the pangenome of mobile elements, which unravels a network of shared genes between these elements (Yutin et al., 2013). This network belongs to a larger one: the social network of life, whose edges describe an important biological structure: "what shares genetic material with what," without prejudices about the process involved in these sharings (in part vertical descent, but also introgression since these sequences can be used as common goods by more than one lineage Halary et al., 2010; McInerney et al., 2011). In this latter network, all entities are not genealogically related, but this does not imply their a priori exclusion from the model. Thanks to its diversity of edges and nodes, the social network of life is more inclusive than the tree of life, supposed to be universal but in fact restricted to one type of relationships between one fraction of biological diversity (Halary et al., 2010).

THE CHALLENGING MICROBIAL SOCIAL LIFE

Microbial social life is hard to explain within the framework of kin selection without (at least) deeply expanding this theory. How do bacteria manage to identify their kins and cooperate? In principle, greenbeard genes provide a way to detect other organisms carrying these genes with which an individual can act cooperatively. However, experimental transfers between strains and species of myxobacteria (*M. xanthus* and *M. fulvus*) of the first characterized single greenbeard prokaryotic gene predictably transform their interactions, reprogramming their social interactions. For example, when an isogenic *M. xanthus* strain expresses a *M. fulvus traA* allele, both become efficacious partners. Moreover, strains constructed with two alleles of *traA* cooperate with a broadened range of partners (Pathak et al., 2013). Consequently, the notion of microbial greenbeard gene departs from classical kinship selection: the cooperative behavior targets other individual harboring the same allele, whatever their global genetic proximity. Lateral gene transfer does not only partly uncouple gene and genome evolution, which makes it difficult to conceive of a standard application of kinship within bacterial populations, since bacteria may be similar for some genes without being similar for all, but the transfer of greenbeard genes can also induce cooperation between relatively different microbes. Therefore, cooperation between distantly related individuals must be more largely theorized (Huneman, 2013). The black queen theory provides a good instance of such an explanation in which genealogical relationships between protagonists do not play a role (Morris et al., 2012; Sachs and Hollowell, 2012). This theory would explain why a minority of organisms (*Synechococcus* harboring the *katG* gene) are sufficient to reduce the HOOH in ocean surface waters to a level that allows the dominant types (*Prochlorococcus* and *Candidatus Pelagibacter ubique* who lost this gene) to thrive.

Considerations on the evolution of social life are fundamental: our inability to grow the vast majority of microorganisms in pure cultures (Staley and Konopka, 1985) may largely come from our too limited knowledge on this topic. Furthermore, our general knowledge in

evolutionary microbiology mostly rests on analyses of the rare microbes able to grow in pure cultures. If these organisms are not representative of most of the microbial world, inferences based on a part of this world (e.g., the cultivable microbes) could be mistakenly conflated with general conclusions, which one hopes to be relevant for the whole microbial world. Yet, discoveries such as the Pandoraviruses remind us that much unknown lives outside our Petri dishes (Philippe et al., 2013). Importantly, alleviating some constraints inspired by the genealogical focus is one way to better see the whole rather the parts. Typically, sequences comparison free from the constraints of multiple alignment and a tree-based representation of sequence similarities hints at highly divergent environmental gene forms and lineages, not yet reported in the microbial world (Lynch et al., 2012).

EGALITARIAN EVOLUTIONARY TRANSITIONS AND SYSTEMS WITH MICROBIAL COMPONENTS

The three steps of evolutionary transitions: the association of entities, their stabilization, and their transformation (after which entities originally able to reproduce independently are only able to reproduce as part of a larger whole) result either in fraternal transitions (which can be explained by traditional kin selection), when higher level units emerge from genealogically like components, and in egalitarian transitions, when higher level units emerge from genealogically different components (Huneman, 2013). These latter transitions are common in the microbiological literature, e.g., *Parakaryon myojinensis* (Yamaguchi et al., 2012), the origins of eukaryotes (Alvarez-Ponce et al., 2013), mutualistic viruses and even virophages, seen as components of larger systems (Espagne et al., 2004; Fischer and Suttle, 2011; Roossinck, 2011). The notion of egalitarian evolutionary transition gives credits to the proposal that some elements of microbiomes deserve to be considered as new organs, providing novel physiologies, despite their genealogically different origin from that of the host of these organs (Stahl and Davidson, 2006). The idea that the microbial component influences the physiology and the behavior of the other components of such systems is becoming

increasingly popular (Hoover et al., 2011; McFall-Ngai et al., 2013), i.e., the study of mechanisms of the microbial-brain axis, testifying of the fundamental relevance in terms of adaptations of the interactions between the microbial and macrobial worlds (Collins et al., 2012).

Consequently, many evolutionary objects previously studied as if they were genealogically cohesive organisms or species would rather constitute genealogically heterogeneous systems, composed in parts by microbes or viruses. This microbial contribution to microbe–microbe and microbe–macrobe systems seems a general rule rather than an exception, when one considers the age, abundance, and ubiquity of these minute entities on the planet. This type of discoveries raises a novel fundamental issue: how to model the evolution of systems (and their possible physiological, ecological, and developmental impacts during Earth history), which brings microbial evolutionists very far from the usual reconstruction of a genealogical tree.

CONCLUSION

Genealogical tree-thinking should at least be completed by other perspectives (Doolittle and Bapteste, 2007; Bapteste et al., 2012, 2013). For example networks can be used to adapt current models to the data rather than enforcing the data to fit within pre-existing genealogically constrained models, designed in order to analyze objects and processes far less complex than those affecting the microbial world (Bapteste, 2013).

REFERENCES

Alvarez-Ponce, D., Lopez, P., Bapteste, E., and McInerney, J. O. (2013). Gene similarity networks provide tools for understanding eukaryote origins and evolution. *Proc. Natl. Acad. Sci. U.S.A.* 110, E1594–E1603. doi: 10.1073/pnas.1211371110

Bapteste, E. (2013). *Les Gènes Voyageurs: L'odyssée de L'évolution*. Belin.

Bapteste, E., Lopez, P., Bouchard, F., Baquero, F., McInerney, J. O., and Burian, R. M. (2012).

Evolutionary analyses of non-genealogical bonds produced by introgressive descent. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18266–18272. doi: 10.1073/pnas.1206541109

Bapteste, E., van Iersel, L., Janke, A., Kelchner, S., Kelk, S., McInerney, J. O., et al. (2013). Networks: expanding evolutionary thinking. *Trends Genet.* 29, 439–441. doi: 10.1016/j.tig.2013.05.007

Collins, S. M., Surette, M., and Bercik, P. (2012). The interplay between the intestinal microbiota and the brain. *Nat. Rev. Microbiol.* 10, 735–742. doi: 10.1038/nrmicro2876

Diggle, S. P., Gardner, A., West, S. A., and Griffin, A. S. (2007). Evolutionary theory of bacterial quorum sensing: when is a signal not a signal? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 1241–1249. doi: 10.1098/rstb.2007.2049

Doolittle, W. F., and Bapteste, E. (2007). Pattern pluralism and the tree of life hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 104, 2043–2049. doi: 10.1073/pnas.0610699104

Espagne, E., Dupuy, C., Huguet, E., Cattolico, L., Provost, B., Martins, N., et al. (2004). Genome sequence of a polydnavirus: insights into symbiotic virus evolution. *Science* 306, 286–289. doi: 10.1126/science.1103066

Fischer, M. G., and Suttle, C. A. (2011). A virophage at the origin of large DNA transposons. *Science* 332, 231–234. doi: 10.1126/science.1199412

Haggerty, L. S., Jachiet, P. A., Hanage, W. P., Fitzpatrick, D. A., Lopez, P., O'Connell, M. J., et al. (2013). A Pluralistic account of homology: adapting the models to the data. *Mol. Biol. Evol.* doi: 10.1093/molbev/mst228. [Epub ahead of print].

Halary, S., Leigh, J. W., Cheaib, B., Lopez, P., and Bapteste, E. (2010). Network analyses structure genetic diversity in independent genetic worlds. *Proc. Natl. Acad. Sci. U.S.A.* 107, 127–132. doi: 10.1073/pnas.0908978107

Halary, S., McInerney, J. O., Lopez, P., and Bapteste, E. (2013). EGN: a wizard for construction of gene and genome similarity networks. *BMC Evol. Biol.* 13:146. doi: 10.1186/1471-2148-13-146

Hoover, K., Grove, M., Gardner, M., Hughes, D. P., McNeil, J., and Slavicek, J. (2011). A gene for an extended phenotype. *Science* 333, 1401. doi: 10.1126/science.1209199

Huneman, P. (2013). "Adaptations in transitions: how to make sense of adaptation when beneficiaries emerge simultaneously with benefits?" in *From Groups to Individuals: Evolution and Emerging Individuality*, eds F. Bouchard and P. Huneman (Cambridge: MIT press), 141–172.

Jachiet, P. A., Pogorelcnik, R., Berry, A., Lopez, P., and Bapteste, E. (2013). MosaicFinder: identification of fused gene families in sequence similarity networks. *Bioinformatics* 29, 837–844. doi: 10.1093/bioinformatics/btt049

Lima-Mendez, G., Van Helden, J., Toussaint, A., and Leplae, R. (2008). Reticulate representation of evolutionary and functional relationships between phage genomes. *Mol. Biol. Evol.* 25, 762–777. doi: 10.1093/molbev/msn023

Lukjancenko, O., Wassenaar, T. M., and Ussery, D. W. (2010). Comparison of 61 sequenced Escherichia coli genomes. *Microb. Ecol.* 60, 708–720. doi: 10.1007/s00248-010-9717-3

Lynch, M. D., Bartram, A. K., and Neufeld, J. D. (2012). Targeted recovery of novel phylogenetic diversity from next-generation sequence data. *ISME J.* 6, 2067–2077. doi: 10.1038/ismej.2012.50

McFall-Ngai, M., Hadfield, M. G., Bosch, T. C., Carey, H. V., Domazet-Loso, T., Douglas, A. E., et al. (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proc. Natl. Acad. Sci. U.S.A.* 110, 3229–3236. doi: 10.1073/pnas.1218525110

McInerney, J. O., Pisani, D., Bapteste, E., and O'Connell, M. J. (2011). The public goods hypothesis for the evolution of life on earth. *Biol. Direct.* 6, 41. doi: 10.1186/1745-6150-6-41

Morris, J. J., Lenski, R. E., and Zinser, E. R. (2012). The black queen hypothesis: evolution of dependencies through adaptive gene loss. *MBio.* 3:e00036-12. doi: 10.1128/mBio.00036-12

Nelson-Sathi, S., Dagan, T., Landan, G., Janssen, A., Steel, M., McInerney, J. O., et al. (2012). Acquisition of 1,000 eubacterial genes physiologically transformed a methanogen at the origin of Haloarchaea. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20537–20542. doi: 10.1073/pnas.1209119109

Pathak, D. T., Wei, X., Dey, A., and Wall, D. (2013). Molecular recognition by a polymorphic cell surface receptor governs cooperative behaviors in bacteria. *PLoS Genet.* 9:e1003891. doi: 10.1371/journal.pgen.1003891

Philippe, N., Legendre, M., Doutre, G., Coute, Y., Poirot, O., Lescot, M., et al. (2013).

Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science* 341, 281–286. doi: 10.1126/science.1239181

Roossinck, M. J. (2011). The good viruses: viral mutualistic symbioses. *Nat. Rev. Microbiol.* 9, 99–108. doi: 10.1038/nrmicro2491

Sachs, J. L., and Hollowell, A. C. (2012). The origins of cooperative bacterial communities. *MBio.* 3:e00099-12. doi: 10.1128/mBio.00099-12

Sorek, R., Zhu, Y., Creevey, C. J., Francino, M. P., Bork, P., and Rubin, E. M. (2007). Genome-wide experimental determination of barriers to horizontal gene transfer. *Science* 318, 1449–1452. doi: 10.1126/science.1147112

Stahl, D. A., and Davidson, S. K. (2006). Genomics: blueprints for partnerships. *Nature* 443, 925–927. doi: 10.1038/nature05208

Staley, J. T., and Konopka, A. (1985). Measurement of *in situ* activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annu. Rev. Microbiol.* 39, 321–346. doi: 10.1146/annurev.mi.39.100185.001541

van Baalen, M. (2013). "The units of adaptation, the emergence of individuality, and the loss of evolutionary sovereignty," in *From Groups to Individuals: Evolution and Emerging Individuality*, eds F. Bouchard and P. Huneman (Cambridge: MIT press), 117–140.

West, S. A., Griffin, A. S., Gardner, A., and Diggle, S. P. (2006). Social evolution theory for microorganisms. *Nat. Rev. Microbiol.* 4, 597–607. doi: 10.1038/nrmicro1461

Yamaguchi, M., Mori, Y., Kozuka, Y., Okada, H., Uematsu, K., Tame, A., et al. (2012). Prokaryote or eukaryote? a unique microorganism from the deep sea. *J. Electron Microsc. (Tokyo)* 61, 423–431. doi: 10.1093/jmicro/dfs062

Yutin, N., Raoult, D., and Koonin, E. V. (2013). Virophages, polintons, and transpovirons: a complex evolutionary network of diverse selfish genetic elements with different reproduction strategies. *Virol. J.* 10:158. doi: 10.1186/1743-422X-10-158

Genome-wide gene order distances support clustering the gram-positive bacteria

Christopher H. House[1]*, Matteo Pellegrini[2,3] and Sorel T. Fitz-Gibbon[2,3]

[1] Penn State Astrobiology Research Center and Department of Geosciences, The Pennsylvania State University, University Park, PA, USA
[2] Department of Molecular, Cell, and Developmental Biology, University of California, Los Angeles, Los Angeles, CA, USA
[3] Department of Molecular, Cell, and Developmental Biology, Institute of Genomics and Proteomics, University of California, Los Angeles, Los Angeles, CA, USA

Initially using 143 genomes, we developed a method for calculating the pair-wise distance between prokaryotic genomes using a Monte Carlo method to estimate the conservation of gene order. The method was based on repeatedly selecting five or six non-adjacent random orthologs from each of two genomes and determining if the chosen orthologs were in the same order. The raw distances were then corrected for gene order convergence using an adaptation of the Jukes-Cantor model, as well as using the common distance correction $D' = -\ln(1-D)$. First, we compared the distances found via the order of six orthologs to distances found based on ortholog gene content and small subunit rRNA sequences. The Jukes-Cantor gene order distances are reasonably well correlated with the divergence of rRNA ($R^2 = 0.24$), especially at rRNA Jukes-Cantor distances of less than 0.2 ($R^2 = 0.52$). Gene content is only weakly correlated with rRNA divergence ($R^2 = 0.04$) over all distances, however, it is especially strongly correlated at rRNA Jukes-Cantor distances of less than 0.1 ($R^2 = 0.67$). This initial work suggests that gene order may be useful in conjunction with other methods to help understand the relatedness of genomes. Using the gene order distances in 143 genomes, the relations of prokaryotes were studied using neighbor joining and agreement subtrees. We then repeated our study of the relations of prokaryotes using gene order in 172 complete genomes better representing a wider-diversity of prokaryotes. Consistently, our trees show the Actinobacteria as a sister group to the bulk of the Firmicutes. In fact, the robustness of gene order support was found to be considerably greater for uniting these two phyla than for uniting any of the proteobacterial classes together. The results are supportive of the idea that Actinobacteria and Firmicutes are closely related, which in turn implies a single origin for the gram-positive cell.

Keywords: tree of life, gene order, evolutionary distance, genomics, Actinobacteria, Firmicutes, Archaea

INTRODUCTION

For the past three decades, the comparisons of ribosomal RNA (rRNA) between microorganisms have largely provided the taxonomic and phylogenetic basis for bacteriology (Woese, 1987). During the past 15 years, however, considerable effort has been placed on comparing the similarity of organisms with genome-wide methods or, at least, with methods that use more than a single gene. These methods include the estimation of genomic distances based on the content of genomes, either orthologs, homologs, folds, or protein domains (Gerstein, 1998; Fitz-Gibbon and House, 1999; Snel et al., 1999; Tekaia et al., 1999; Wolf et al., 2002; Deeds et al., 2005; Yang et al., 2005; House, 2009). Genomic distance has also been estimated using direct genome-to-genome sequence comparisons using a variety of approaches like average nucleotide identity (ANI) and the genome-to-genome-distance calculator (GGDC) that can approximate traditional DNA-DNA hybridization results (Konstantinidis and Tiedje, 2005; Goris et al., 2007; Deloger et al., 2009; Richter and Rosselló-Móra, 2009; Auch et al., 2010; Tamura et al., 2012; Meier-Kolthoff et al., 2013).

Also, ever since Nadeau and Taylor (1984) first identified that gene order information was conserved between humans and mice, there has been growing interest in using gene order to estimate the difference between genomes or to solve phylogenetic problems.

Several gene order methods depend on the presence of orthologs adjacent to each other. Watterson et al. (1982) introduced the *breakpoint distance* between genomes, which is the number of orthologs found paired together in one genome but separated in the other Blanchette et al. (1999). Early on, Sankoff et al. (1992) estimated mitochondria gene rearrangements as a means to derive a phylogenetic tree for Eukaryotes. Subsequently, the presence and absence of paired genes has been used to construct trees (Wolf et al., 2001; Korbel et al., 2002) as a gene order method similar in practice to tree building by gene content. A limitation to this approach results from the fact that small groups of laterally transferred genes will be paired after their transfer. Also, a computational method for testing phylogenetic problems using gene order has been presented by Kunisawa (2001). In this method, genomes are searched for cases in which

the arrangement of three genes most parsimoniously suggests that a single transposition has occurred. With the use of an outgroup, the method can be used to test phylogenetic hypotheses, such as the branching order within the Proteobacteria (Kunisawa, 2001) or Gram-positive bacteria (Kunisawa, 2003). The strength of this method is that it can be efficiently applied to a large dataset of genomes and that it reveals (a small number of) interesting cases of transposition. Another gene order approach often implemented is calculating the inversion distance. The *inversion distance* is the minimum possible number of inversions needed to transform one genome into the other (Moret et al., 2001). Recently, Belda et al. (2005) have studied a subset of 244 genes universal to the genomes of 30 γ-Preotobacteria using both the breakpoint distance and the inversion distance. They found the two distances highly correlated suggesting that inversion was the main method of genome rearrangement for these taxa. More recently, models for genome evolution that include rearrangements, duplications, and losses have been developed and tested (Swenson et al., 2008; Zhang et al., 2010; Hu et al., 2011; Lin and Moret, 2011; Shao et al., 2013) have each developed algorithms for using gene order for phylogenetic reconstruction. Furthermore, Lin et al. (2013) and Shifman et al. (2014) have used genome-wide gene order to produce phylogenetic trees. The later work produced a tree of 89 diverse microbial genomes using an algorithm for estimating average genome synteny (Shifman et al., 2014).

In this study, we aimed to develop a simple computational method that could estimate a genome-wide gene order distance between two genomes (even when the genomes were highly diverged). Unlike many previous efforts, our intent was to have the gene order distance not rely on genes that are likely to be in the same operon (such as gene pairs). Here, we present a novel simple Monte Carlo method for estimating *distributed gene order distances* between genomes. In this method, we repeatedly randomly select six non-adjacent orthologs from each of two genomes and determine if the genes are in the same order. The distances are then corrected using an adaptation of the Jukes-Cantor model to account for random gene order convergence.

MATERIALS AND METHODS

Initially, 143 prokaryotic genomes were analyzed (**Table 1**). This represented completed prokaryotic genomes available when the study began in January 2005. All genes from each genome were analyzed as queries using BLAST against each of the other genomes. Ortholog-pairs were identified as cases where two genes from different genomes were each other's BLAST best hit (top hit in both directions). This list of ortholog pairs served as the basis for both calculation of distributed gene order distances and the ortholog gene content distances. As defined by Snel et al. (1999), ortholog gene content similarity (S) was calculated as the number of ortholog pairs found for two genomes divided by the size of the smaller genome. This similarity was then converted to distance as equal to −ln(S), as suggested by Korbel et al. (2002). However, using distance equal to 1-S gives similar correlation results.

Distributed gene order distances were determined using a novel Monte Carlo approach (**Figure 1**). To determine the gene order distance between two genomes, first, six ortholog-pairs

Table 1 | 143 taxa.

	ID
Aeropyrum pernix K1	ap
Agrobacterium tumefaciens C58UW	at
Agrobacterium tumefaciens C58C	atc
Aquifex aeolicus VF5	aa
Archaeoglobus fulgidus DSM4304	af
Bacillus anthracis Ames	baa
Bacillus cereus ATCC 14579	bc
Bacillus halodurans C-125	bh
Bacillus subtilis 168	bs
Bacteroides thetaiotaomicron	bt
Bifidobacterium longum NCC2705	bl
Bordetella bronchiseptica	bbr
Bordetella parapertussis	bpp
Bordetella pertussis	bp
Borrelia burgdorferi B31	bb
Bradyrhizobium japonicum USDA 110	bj
Brucella melitensis	bm
Brucella suis	brs
Buchnera aphidicola Bp	ba
Buchnera aphidicola Sg	bas
Buchnera sp. APS	bu
Campylobacter jejuni NCTC 11168	cj
Candidatus Blochmannia floridanus	cbf
Caulobacter crescentus CB15	cc
Chlamydia trachomatis serovar D	ct
Chlamydia trachomatis MoPn/Nigg	cm
Chlamydophila caviae GPIC	cca
Chlamydophila pneumoniae AR39	cpa
Chlamydophila pneumoniae J138	cpj
Chlamydophila pneumoniae TW183	cpt
Chlamydophila pneumoniae CWL029	cp
Chlorobium tepidum TLS	cte
Chromobacterium violaceum	cv
Clostridium acetobutylicum ATCC 824	ca
Clostridium perfringens	cpe
Clostridium tetani	clt
Corynebacterium diphtheria	cd
Corynebacterium efficiens YS-314	cef
Corynebacterium glutamicum	cg
Coxiella burnetii	cb
Deinococcus radiodurans R1	dr
Enterococcus faecalis V583	ef
Escherichia coli O157:H7 strain EDL933	ece
Escherichia coli K-12 Strain MG1655	ec
Escherichia coli CFT073	ecc
Escherichia coli O157:H7	ech
Fusobacterium nucleatum ATCC 25586	fn
Gloeobacter violaceus	gv
Haemophilus ducreyi	hd
Haemophilus influenzae Rd KW20	hi
Halobacterium sp. NRC-1	hsp
Helicobacter hepaticus ATCC 51449	hh

(Continued)

Table 1 | Continued

	ID
Helicobacter pylori 26695	hp
Helicobacter pylori J99	hpj
Lactobacillus plantarum WCFS1	lp
Lactococcus lactis IL1403	ll
Leptospira interrogans s.l. 56601	li
Listeria innocua clip11262	lin
Listeria monocytogenes EGD-e	lm
Mesorhizobium loti MAFF303099	ml
Methanobacterium thermoautotroph.	mt
Methanococcus jannaschii DSM 2661	mj
Methanopyrus kandleri AV19	mk
Methanosarcina acetivorans C2A	ma
Methanosarcina mazei Goe1	mma
Mycobacterium bovis bovis	mb
Mycobacterium leprae	mle
Mycobacterium tuberculosis H37Rv	mtb
Mycobacterium tuberculosis cdc1551	mtc
Mycoplasma gallisepticum	mga
Mycoplasma genitalium G-37	mg
Mycoplasma penetrans	mpe
Mycoplasma pneumoniae M129	mp
Mycoplasma pulmonis UAB CTIP	mpu
Nanobacterium equitans Kin4-M	neq
Neisseria meningitidis MC58	nmm
Neisseria meningitidis A Z2491	nmz
Nitrosomonas europaea	ne
Nostoc sp. PCC7120	ns
Oceanobacillus iheyensis HTE831	oi
Pasteurella multocida Pm70	pm
Photorhabdus luminescens	pl
Pirellula_sp	pi
Porphyromonas gingivalis	pg
Prochlorococcus marinus CCMP1375	pmc
Prochlorococcus marinus MED4	pmm
Prochlorococcus marinus MIT9313	pma
Pseudomonas aeruginosa PAO1	psa
Pseudomonas putida KT2440	psp
Pseudomonas syringae pv. tomato	pss
Pyrobaculum aerophilum IM2	pa
Pyrococcus abyssi	pab
Pyrococcus furiosus DSM3638	pf
Pyrococcus horikoshii OT3	ph
Ralstonia solanacearum	rs
Rickettsia conorii Malish 7	rc
Rickettsia prowazekii Madrid E	rp
Salmonella enterica Typhi	se
Salmonella enterica Typhi_Ty2	set
Salmonella typhimurium LT2	sty
Shewanella oneidensis	so
Shigella flexneri 2a	sf
Sinorhizobium meliloti 1021	sm
Staphylococcus aureus N315	san

(Continued)

Table 1 | Continued

	ID
Staphylococcus aureus MW2	saw
Staphylococcus aureus Mu50	sam
Staphylococcus epidermidis 12228	sep
Streptococcus agalactiae 2603	sa
Streptococcus agalactiae NEM316	sag
Streptococcus mutans	smu
Streptococcus pneumoniae R6	spn
Streptococcus pneumoniae TIGR4	spt
Streptococcus pyogenes SSI-1	mle
Streptococcus pyogenes MGAS8232	spa
Streptococcus pyogenes MGAS315	spg
Streptococcus pyogenes M1_GAS	spm
Streptomyces avermitilis MA-4680	sav
Streptomyces coelicolor A3(2)	sco
Sulfolobus solfataricusP2	ss
Sulfolobus tokodaii 7	st
Synechococcus sp. WH8102	syo
Synechocystis sp. PCC 6803	sy
Thermoanaerobacter tengcongensis	tt
Thermoplasma acidophilum	ta
Thermoplasma volcanium GSS1	tv
Thermosynechococcus elongatus BP-1	te
Thermotoga maritima MSB8	tm
Treponema pallidum Nichols	tp
Tropheryma whipplei Twist	tw
Tropheryma whipplei TW08_27	twt
Ureaplasma urealyticum serovar 3	uu
Vibrio cholerae serotype O1 (N16961)	vc
Vibrio parahaemolyticus RIMD 2210633	vp
Vibrio vulnificus CMCP6	vv
Vibrio vulnificus YJ016	vvy
Wigglesworthia brevipalpis	wb
Wolinella_succinogenes	ws
Xanthomonas axonopodis pv citri 306	xa
Xanthomonas campestris ATCC 33913	xc
Xylella fastidiosa 9a5c	xf
Xylella fastidiosa Temecula1	xft
Yersinia pestis CO-92 Biovar Orientalis	yp
Yersinia pestis KIM	ypk

were randomly chosen. In order to limit orthologs being chosen from the same operon, the orthologs were required to be at least 5 genes away from each other in either genome. It was then determined if the chosen six ortholog-pairs were in the same order around both circular genomes (irrespective of each genes orientation). For organisms with multiple chromosomes, only the largest chromosome was used in this initial effort. This procedure was repeated for 100,000 iterations to establish one replicate sampling. In the end, 100 replicate samplings were performed for all genome pairs, and these data were either combined to construct one list of distances based on 10 million iterations, or kept separate to make 100 lists of distances for use as bootstrap replicates (nexus files for PAUP are available in Supplementary Material).

FIGURE 1 | Diagram demonstrating the method used to calculate the pair wise distributed gene order distance between genomes.
Repeatedly, six ortholog pairs are chosen randomly (requiring every gene in the six be at least 5 genes away along the genome from each). The six genes are then tested to see if they are in the same order (irrespective of the orientation of the genes). In the case above, the test fails because orthologs C and E are switched. Distributed gene order distance is equal to the fraction of times such a test fails between two genomes. The diagram also works for demonstrating the distributed gene order distance between genomes using five genes (A–E) by ignoring gene F.

Recently diverged genomes begin with close to 100% of their genes arranged in the same order, and with time, the synteny between the genomes decreases. Because there are only 60 different ways to arrange six items on a circle, there is a 1/60 probability of two genomes sharing an arrangement of six orthologs by chance. Therefore, the fraction of six ortholog picks found to be in the same order will ultimately approach 1/60 as divergence time goes to infinity. We, therefore, developed a model of gene order evolution based on the Jukes-Cantor concept that divergence is a logarithmic function with time (Jukes and Cantor, 1969).

The typical Jukes-Cantor correction (Kimura and Ohta, 1972) for nucleotide distance is:

$$D_{JC} = -(3/4) \ln (1 - (4/3)D) \qquad (1)$$

where $D =$ the observed fractional of nucleotides found to be different between two compared genes.

This classical nucleotide Jukes-Cantor correction (Equation 1) accounts for back substitution and is based on a model in which the outcome of any nucleotide substitution can be one of three possibilities. To adopt this logic to gene rearrangements, the Jukes-Cantor equation becomes:

$$D_{JC} = -(59/60) \ln (1 - (60/59)D) \qquad (2)$$

where $D =$ the fraction of iterations in which the six ortholog-pairs chosen are not in the same order.

The classical Jukes-Cantor nucleotide correction (Equation 1) can only be used for raw D up to 0.75. With raw nucleotide distances greater than 0.75, the argument of the logarithm will be zero. To use data in which D is larger than 0.75, Tajima (1993) presented a method using a Taylor series expansion to avoid the logarithm. In our case, Equation (2) fails whenever the raw D is greater than 59/60 (or 0.983). To allow corrections for all of our

genome pair distances, we have adopted the method of Tajima (1993) as follows:

$$D_{JC} = \sum_{i=1}^{k} \frac{k^{(i)}}{i(59/60)^{i-1}n^{(i)}} \qquad (3)$$

where $k^{(i)} = k!/(k-i)!$, $n^{(i)} = n!/(n-i)!$, $k =$ the number of times the six orthologs are not in the same order, and $n =$ the number of iterations used.

Partial reanalysis of the work reported here demonstrated the results are similar when applying $D' = -\ln(1-D)$ as the distance correction rather than the Tajima correction (data not shown), and further future work evaluating this measure of gene order distance is warranted as it is computationally much less intense.

For comparison, Jukes-Cantor corrected rRNA distances were downloaded spring 2006 from the ribosomal database (Cole et al., 2007). The correlations between distributed gene order, gene content, and rRNA distances were performed with SPSS 13 (SPSS, Inc. Chicago, IL) for Mac OS X. Taxonomic assignments for taxa were from the NCBI taxonomic server (Bischoff et al., 2007).

Our follow-up analysis used 172 complete genomes with the aim of being a representative sample of prokaryotes. For this follow-on analysis initiated early in 2014, we used ortholog predictions from the OMA website (Dessimoz et al., 2005). This OMA database is continually updated and includes all chromosomes for each microorganism. The updated analysis here of 172 taxa was done with orthologs downloaded in early 2014. In this case, we also tried searching for five orthologs in the same order rather than six using the same equations, which naturally produces slightly shorter distances overall. In fact, the five gene distances used this last analysis are functionally the same as using the easier to calculate $D' = -\ln(1-D)$. Based on the promising results here, we recommend this simpler distance calculation for future work.

Neighbor Joining (NJ) trees (Saitou and Nei, 1987) were created from data matrices using PAUP 4.0b (Mac and Unix versions; Sinauer Associates, Sunderland, MA). Later, agreement subtrees, which identifies the largest possible pruned tree that is consistent within a set of trees, was used to limit the taxa list in order to minimize possible adverse effects of including genome pairs with very little or no gene order conservation. The agreement subtrees were identified using PAUP 4.0b (Mac) based on a comparison of all of our NJ trees produced from the 100 replicate distances.

We also tried using a hierarchical and iterative approach to produce a series of trees (**Table 2**). This novel method was based on the fact that shorter distances are known with higher confidence than greater distances. The goal of this method of tree building is to provide a systematic and objective way to build a tree that includes as many of the pair wise gene order distances as possible without letting very distant (random) pairs adversely influence the observed phylogenetic positions of the more closely related taxa. We started with a list of genome pairs ranked from shortest to largest gene order distance (available in Supplemental Data). Starting at the top of the list, we moved down the list adding each pair to our working group until enough pair wise distances were included to allow for one or more NJ trees to be built.

Table 2 | Steps used in hierarchical tree building.

1	Construct a ranked list of gene order distances starting with the shortest distances
2	Move down ranked list, forming NJ trees of increasing taxa number estimating reaching a NJ tree of all taxa
3	In turn, evaluate each tree formed starting with the smallest and moving to the largest
4	Keep trees consistent with all previously retained trees, while rejecting any new tree that is incongruent with a previously retained tree
5	Starting with those represented by the smallest gene order pairs, single taxa were added to the largest retained tree if their addition did not disrupt the existing NJ topology (second round of taxa addition)

This process was continued until we had an exhaustive ranked list of possible unrooted NJ trees starting with the top few very closely related taxa and ending ultimately with a NJ tree of all 143 taxa. Moving down the ranked tree list, we evaluated each tree. A tree (unrooted) was rejected if it was found to be incongruent with an earlier unrooted tree. Congruent trees were pared down in number by removing trees that were fully encompassed by another tree and by combining pairs of compatible trees. Trees were combined by building a new NJ tree with the union set of taxa from the two original trees. The trees were only considered compatible for combining if the process did not cause a disruption of either of the original backbone topologies. For each kept tree, we recorded both the rank of the taxa pair that resulted in its initial formation, and the rank of the last taxa pair added. The largest resulting tree (with 37 taxa) was selected for further study. Additional taxa were added using a process of single taxon addition. In this second round of analysis, moving down the ranked list of genome pairs, we attempted to sequentially add additional taxa to the tree. If the addition of the single taxon disrupted the existing NJ topology, then the taxon was not added.

RESULTS AND DISCUSSION

INITIAL TEST OF GENE ORDER AS AN EVOLUTIONARY DISTANCE

The distribution of raw gene order distances for each of the 10,153 genome pairs for our 143 genomes are plotted in **Figure 2A** (and available in Supplementary Material). As expected, with raw gene order distance of 0 (or near 0), the two genomes for *Chlamydia trachomatis*, and separately the four genomes for *Chlamydophila pneumoniae* define the far left of the distribution. The bulk of the genome pairs, however, show raw gene order distances of greater than 0.9 with a peak near, but below, the value expected randomly ($0.98\overline{3}$). 82% of the genome pairs have gene order distances below 0.983. **Figure 2B** shows the same data after an adapted Jukes-Cantor model correction (Equation 2). Using this logarithm–based correction, the gene order distances show a relatively normal distribution with a mean of 7.49 ($SD = 1.68$). This correction, however, is not possible for raw gene order distance larger than $0.98\overline{3}$, and so, such divergent data are missing from **Figure 2B**. **Figure 2C** shows a fuller dataset of gene order distances corrected using the method adopted from Tajima (Equation 3). In this case, a very long tail of very large gene order
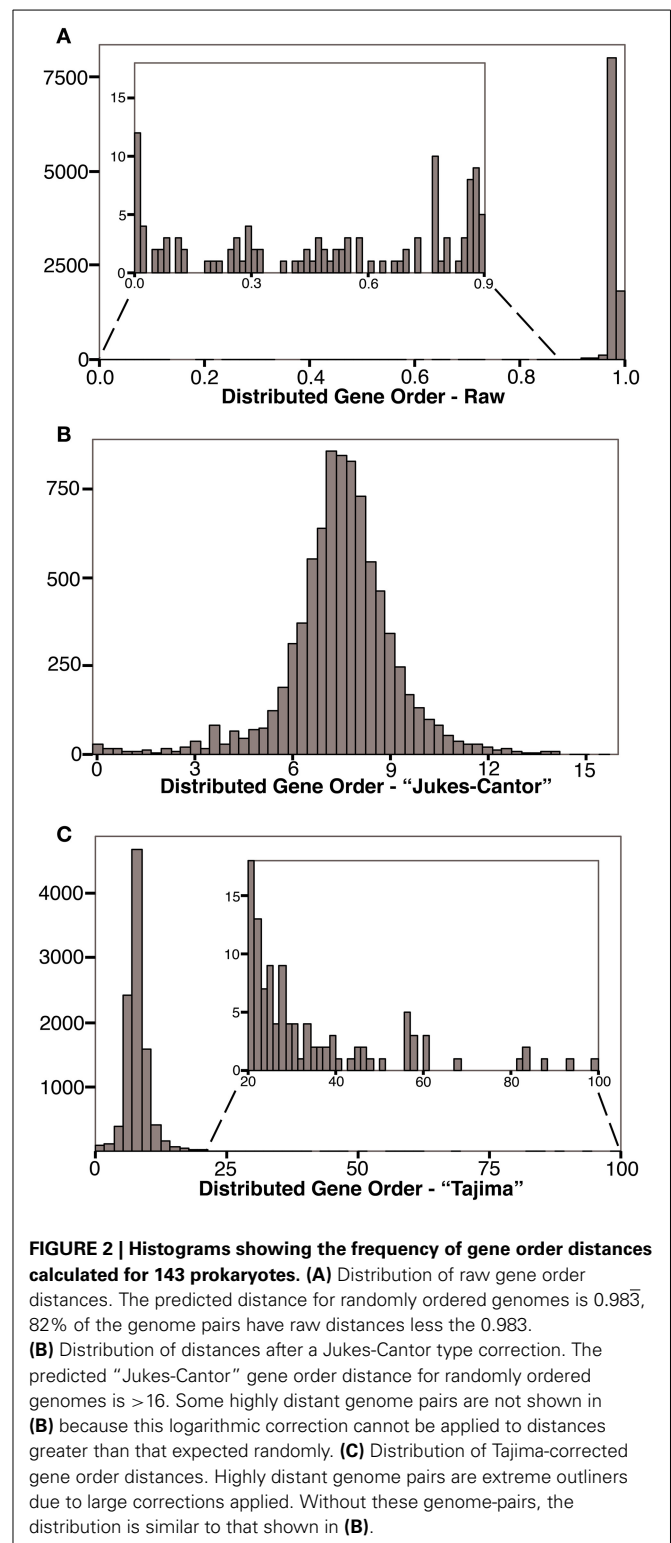
FIGURE 2 | Histograms showing the frequency of gene order distances calculated for 143 prokaryotes. (A) Distribution of raw gene order distances. The predicted distance for randomly ordered genomes is $0.98\overline{3}$, 82% of the genome pairs have raw distances less the 0.983.
(B) Distribution of distances after a Jukes-Cantor type correction. The predicted "Jukes-Cantor" gene order distance for randomly ordered genomes is >16. Some highly distant genome pairs are not shown in **(B)** because this logarithmic correction cannot be applied to distances greater than that expected randomly. **(C)** Distribution of Tajima-corrected gene order distances. Highly distant genome pairs are extreme outliners due to large corrections applied. Without these genome-pairs, the distribution is similar to that shown in **(B)**.

distances is apparent. This tail is caused by large corrections being applied to some dissimilar genome-pairs.

After calculating corrected gene order distances for each genome pair, we compared these values with other measures of genome distance, Jukes-Cantor corrected rRNA distances and

logarithmic gene content distances (data used are available in Supplementary Material). **Figure 3A** shows a strong correlation between the "Jukes-Cantor" corrected gene order distances and the Jukes-Cantor rRNA distances ($R^2 = 0.24$), especially at rRNA distances shorter than 0.2 ($R^2 = 0.52$). Gene content distances show much less significant correlations with rRNA distance (**Figure 3B**; $R^2 = 0.04$), and are actually much more strongly correlated with gene order (**Figure 3C**; $R^2 = 0.22$). However, a very strong correlation between gene content and Jukes-Cantor rRNA distance is apparent at rRNA distances shorter than 0.1

FIGURE 3 | Comparison of "Jukes-Cantor" distributed gene order distances with ortholog gene content and Jukes-Cantor rRNA distances. Select gene pairs have been labeled. **(A)** Gene order distance plotted as a function of rRNA distance. Solid line is linear regression of all data ($R^2 = 0.24$). Dashed line is a linear regression for genome pairs with rRNA distances <0.2 ($R^2 = 0.52$). **(B)** Gene content distance plotted as a function of rRNA distance. Solid line is linear regression of all data ($R^2 = 0.04$). Dashed line is linear regression for genome pairs with rRNA distances <0.1 ($R^2 = 0.67$). **(C)** Gene content distance plotted as a function of gene order distance. Solid line is linear regression of all data ($R^2 = 0.22$).

($R^2 = 0.67$). Apparent in **Figure 2B**, there is a cluster of genome pairs with similar gene content (low gene content distance) but divergent rRNA sequences (rRNA distances between 0.1 and 0.3 and Gene Content distances of 0.0 and 0.2). This population of genome pairs consists of very small genomes in which extreme genome reduction has occurred. In cases of such extensive genome reduction, the proportion of orthologs shared between two genomes is high even when rRNA sequences indicate that the genomes are relatively distant. With an updated list of genomes, we found our Jukes-Cantor gene order distances were correlated with the divergence of conserved protein genes ($R^2 = 0.13$ overall and $R^2 = 0.23$ for cases with less than one amino acid substitution calculated per site). For this comparison, we identified 87 taxa present in both our data (**Table 3**) and (Lang et al., 2013) and we extracted the aligned amino acid sequences from their published alignment of 841 sequences. Using MEGA5 we computed pairwise distances under the Dayhoff matrix based model (Schwartz and Dayhoff, 1978).

In general, these various analyses have shown that the distributed gene order distances are better correlated with sequenced-based distances (both rRNA and conserved proteins) than are gene content distances. Assuming rRNA and protein distances are useful approximations of evolutionary divergence, our results suggest that gene order distance may be useful for taxonomy or phylogenetics in a similar way that genome-to-genome sequence comparison has proven useful as a genomic measure to replace laboratory DNA-DNA hybridization. Both gene content and gene order have their specific issues and so we would not recommend that either genomic method is ever considered the best, but rather that each be used to help inform more traditional molecular sequence analyses when these methods show signal and can be reasonably interpreted. It is also important to note that horizontal gene transfers likely bring in orthologs in the same order, and so gene order does not necessarily avoid complications of integration arising from such transfers, if they are on a large scale.

GENE ORDER TREE BUILDING STARTING FROM 143 TAXA USING NEIGHBOR JOINING

Encouraged by the strong correlation between our distributed gene order distances and those of rRNA, we proceeded to build gene order-based neighbor joining (NJ) trees. **Figure 4** shows a phylogram based on all 143 taxa, although unresolved single taxon are not shown for clarity. Similarly, **Figure 5** shows the bootstrap NJ tree when all taxa are included. The results show that major bacterial taxonomic groups are mostly correctly clustered. With the largest number of resolved taxa, the γ-Proteobacteria shows the best resolution, and has a branching order reasonably consistent with published reports (House et al., 2003; Brown and Volker, 2004; Belda et al., 2005). In contrast, Archaea are almost completely unresolved, with only generic-level similarity yielding clustering (e.g., *Pyrococcus*, *Methanosarcina*, *Sulfolobus*, and *Thermoplasma*). Interestingly, the Actinobacteria are resolved with high confidence as a sister group to the bulk of the Firmicutes (including *Bacillus*, *Clostridium*, *Lactobacillus*, *Listeria*, *Oceanobacillus*, *Staphylococcus*, and *Thermoanaerobacter*). In contrast, our bootstrap tree (**Figure 5**) does not cluster any of
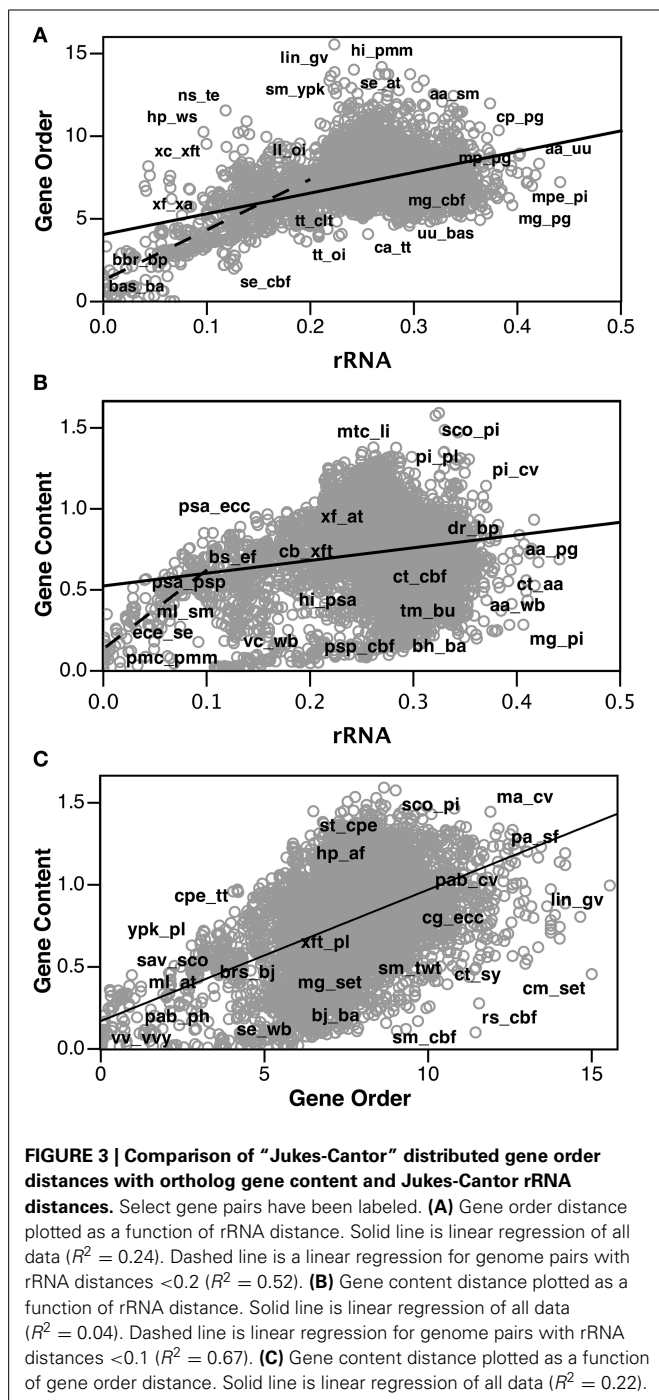
Table 3 | 172 taxa.

	ID
Acidaminococcus fermentans	ACIFV
Acidilobus saccharovorans	ACIS3
Acidimicrobium ferrooxidans	ACIFD
Acidithiobacillus ferrooxidans	ACIF5
Acinetobacter baumannii	ACIBS
Acinetobacter sp.	ACIAD
Aeromonas hydrophila hydrophila	AERHH
Aeromonas salmonicida	AERS4
Alcanivorax borkumensis	ALCBS
Alicyclobacillus acidocaldarius	ALIAD
Alteromonas macleodii	ALTMD
Amycolatopsis mediterranei	AMYMU
Anabaena variabilis	ANAVT
Anoxybacillus flavithermus	ANOFW
Arcanobacterium haemolyticum	ARCHD
Archaeoglobus fulgidus	ARCFU
Archaeoglobus profundus	ARCPA
Archaeoglobus veneficus	ARCVS
Azoarcus sp.	AZOSB
Azotobacter vinelandii	AZOVD
Bacillus amyloliquefaciens	BACA2
Bacillus pumilus	BACP2
Bacillus selenitireducens	BACIE
Beutenbergia cavernae	BEUC1
Bifidobacterium adolescentis	BIFAA
Bifidobacterium animalis animalis	BIFAR
Bifidobacterium animalis lactis	BIFA0
Burkholderia mallei	BURMA
Burkholderia thailandensis	BURTA
Campylobacter jejuni HS:41	CAMJC
Campylobacter lari	CAMLR
Catenulispora acidiphila	CATAD
Caulobacter crescentus	CAUCR
Caulobacter segnis	CAUST
Cellvibrio gilvus	CELGA
Cellvibrio japonicus	CELJU
Cenarchaeum symbiosum	CENSY
Clostridium novyi	CLONN
Clostridium perfringens	CLOPS
Clostridium tetani	CLOTE
Coriobacterium glomerans	CORGP
Corynebacterium jeikeium	CORJK
Corynebacterium kroppenstedtii	CORK4
Corynebacterium urealyticum	CORU7
Dechloromonas aromatic	DECAR
Desulfovibrio vulgaris	DESVV
Desulfurococcus kamchatkensis	DESK1
Desulfurococcus mucosus	DESM0
Dichelobacter nodosus	DICNV
Enterobacter cloacae	ENTCS
Enterobacter sp.	ENT38
Enterococcus faecalis	ENTFA

(Continued)

Table 3 | Continued

	ID
Frankia alni	FRAAA
Frankia sp.	FRASC
Gardnerella vaginalis	GARV4
Geobacillus kaustophilus	GEOKA
Geobacillus sp.	GEOSW
Geobacillus thermodenitrificans	GEOTN
Gloeobacter violaceus	GLOVI
Hahella chejuensis	HAHCH
Halobacterium salinarum	HALSA
Halothermothrix orenii	HALOH
Helicobacter mustelae	HELM1
Helicobacter pylori	HELP5
Hydrogenobacter thermophiles	HYDTT
Kineococcus radiotolerans	KINRD
Korarchaeum cryptofilum	KORCO
Lactobacillus fermentum	LACFC
Lactobacillus helveticus	LACH4
Lactobacillus salivarius	LACSC
Lactococcus lactis cremoris	LACLS
Lactococcus lactis subsp. Lactis	LACLA
Legionella pneumophila	LEGPL
Legionella pneumophila pneumophila	LEGPH
Leuconostoc citreum	LEUCK
Leuconostoc gasicomitatum	LEUGT
Leuconostoc sp.	LEUS2
Listeria monocytogenes serotype 4b	LISMC
Listeria monocytogenes serovar 1/2a	LISMO
Listeria welshimeri serovar 6b	LISW6
Lysinibacillus sphaericus	LYSSC
Magnetococcus sp.	MAGSM
Methanobacterium sp.	METSW
Methanocaldococcus fervens	METFA
Methanocaldococcus infernus	METIM
Methanocaldococcus vulcanius	ETVM
Methanocella conradii	METCZ
Methanococcus aeolicus	META3
Methanococcus vannielii	METVS
Methanococcus voltae	METV3
Methanopyrus kandleri	METKA
Methanosaeta concilii	METCG
Methanosaeta harundinacea	METH6
Methanosaeta thermophile	METTP
Methanosarcina acetivorans	METAC
Methanosarcina barkeri	METBF
Methanosarcina mazei	METMA
Methylobacillus flagellates	METFK
Methylococcus capsulatus	METCA
Microcystis aeruginosa	MICAN
Micromonospora aurantiaca	MICAI
Micromonospora sp.	MICSL
Moraxella catarrhalis	MORCR
Nanoarchaeum equitans	NANEQ

(Continued)

Table 3 | Continued

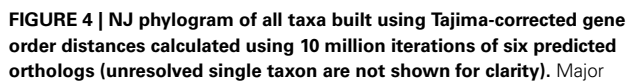|	ID
Natranaerobius thermophiles	NATTJ
Nautilia profundicola	NAUPA
Neisseria meningitides	NEIML
Neisseria meningitidis serogroup B	NEIMG
Nitrosomonas europaea	NITEU
Nitrosomonas eutropha	NITEC
Nitrosopumilus maritimus	NITMS
Nitrososphaera gargensis	NITGG
Nocardia cyriacigeorgica	NOCCG
Nocardia farcinica	NOCFA
Nocardioides sp.	NOCSJ
Nostoc azollae	NOSA0
Nostoc punctiforme	NOSP7
Nostoc sp.	NOSS1
Oceanobacillus iheyensis	OCEIH
Parvularcula bermudensis	PARBH
Pasteurella multocida	PASMU
Prochlorococcus marinus	PROM4
Prochlorococcus marinus pastoris	PROMP
Propionibacterium acnes	PROAC
Propionibacterium propionicum	PROPF
Pseudomonas fulva	PSEF1
Pseudomonas stutzeri	PSEU5
Psychrobacter arcticus	PSYA2
Psychrobacter sp.	PSYWF
Rhizobium etli	RHIEC
Rhizobium meliloti	RHIME
Rhodobacter capsulatus	RHOCB
Rhodobacter sphaeroides	RHOS1
Rhodospirillum centenum	RHOCS
Rhodospirillum rubrum	RHORT
Rickettsia prowazekii	RICPR
Rickettsia typhi	RICTY
Rubrobacter xylanophilus	RUBXD
Saccharomonospora viridis	SACVD
Saccharopolyspora erythraea	SACEN
Sphingomonas wittichii	SPHWW
Staphylococcus carnosus	STACT
Staphylococcus epidermidis	STAES
Staphylococcus lugdunensis	STALH
Streptococcus pyogenes M49	STRPZ
Streptococcus pyogenes M5	STRPG
Streptococcus thermophiles	STRTD
Streptomyces avermitilis	STRAW
Streptomyces coelicolor	STRCO
Streptomyces griseus	STRGG
Streptosporangium roseum	STRRD
Sulfolobus acidocaldarius	SULAC
Sulfolobus islandicus	SULIM
Sulfolobus solfataricus	SULS9
Thermoanaerobacter italicus	THEIA
Thermoanaerobacter mathranii	THEM3

(Continued)

Table 3 | Continued

	ID
Thermoanaerobacter pseudethanolicus	THEP3
Thermobispora bispora	THEBD
Thermococcus onnurineus	THEON
Thermococcus sibiricus	THESM
Thermococcus sp.	THES4
Thermoplasma acidophilum	THEAC
Thermoplasma volcanium	THEVO
Thermoproteus neutrophilus	THENV
Thermoproteus tenax	THETK
Thermoproteus uzoniensis	THEU7
Thiomicrospira crunogena	THICR
Veillonella parvula	VEIPT
Vibrio cholerae serotype O1	VIBCM
Vibrio fischeri	VIBF1
Xanthomonas campestris	XANCP
Xanthomonas oryzae pv. Oryzae	XANOM

the proteobacterial classes together. Finding the Actinobacteria and Firmicutes united is interesting because they are the two phyla that comprise the "gram-positive bacteria." While it has long been considered likely that the gram-positive bacteria are a monophyletic group, it has been to date remarkably hard to find supportive molecular sequence data, genetic or genomic (De Rijk et al., 1995; Olsen, 2001; Fu and Fu-Liu, 2002; Deeds et al., 2005).

Next, we tested if the small phylogenetic signal we found with gene order distance was due to the occasional sampling of ribosomal operons, despite the 5 gene exclusion zone. A detailed look at 100,000 randomly sampled gene sets revealed that sets with more than one ribosomal gene do not occur any more frequently for conserved order sets (20%) than non-conserved order sets (21%). Furthermore, there was very little difference in the percentage of each of the following cog-based (Tatusov et al., 2003) protein function categories between the two groups of sets (conserved vs. non-conserved): informational (24 vs. 26%), cellular (17 vs. 17%), metabolism (36 vs. 37%), poorly categorized (15 vs. 14%), no cog match (8 vs. 6%). This suggests that the signal is distributed across many different types of genes and is probably not due to unreliable "jackpot" effects of single operons. We also pruned our data set to remove all ribosomal genes. When this pruned data set was used for building a NJ tree, however, the resolution is reduced resulting in a topology where some well-established microbial phyla are intertwined. This new NJ result does unite the Actinobacteria and Firmicutes, but with very low confidence. Because the dataset with ribosomal genes removed does not fully reproduce the results shown in **Figure 5**, it remains a possibility that a notable proportion of the gene order signal is preserved in ribosomal genes, but that in addition the signal overall appears to be distributed across a variety of other gene functional categories. The most likely way to reconcile these apparently divergent conclusions is that the phylogenetic signal in gene order distance is small, and so, the removal of any class of genes (including

FIGURE 4 | NJ phylogram of all taxa built using Tajima-corrected gene order distances calculated using 10 million iterations of six predicted orthologs (unresolved single taxon are not shown for clarity). Major taxonomic groups are labeled. Actinobacteria are shown in green, and the two clusters of Firmicutes are shown in blue. The Actinobacteria are grouped with the bulk of the Firmicutes.

ribosomal operons) appreciably reduces the robustness of the results.

ADDITIONAL GENE ORDER TREE BUILDING STARTING FROM 143 TAXA

To complement our NJ tree building exercise using all 143 taxa, we aimed to address the fundamental problem that only a portion of our 10,153 pair wise gene order distances were significant and should be useful for tree building. The inclusion of genome pairs that are too diverged with respect to their gene order has the potential to alter the observed position of other taxa on a tree. This concern is not unique to gene order data. It has long been known that with sequence data, the uncertainty on an estimated distance goes up greatly with the magnitude of the divergence (Kimura and Ohta, 1972). Gene order data though provide a dramatic example of how it can be difficult to accurately estimate divergence when organisms are highly diverged. To minimize this problem, we proceeded with two additional tree studies.

We tried developing a novel hierarchical and iterative tree building strategy (see Materials and Methods) based on the principle that our shorter distances are known with a higher degree of confidence than our larger distances. The goal of this approach

is to provide a systematic and objective way to build a tree that includes as many of the pair wise gene order distances as possible without letting very distant (random) pairs adversely influence the observed phylogenetic positions of the more closely related taxa. Detailed results of this work are listed in the Supplemental Material. **Figure 6** shows the largest tree formed starting this process with all 143 taxa. The tree in **Figure 7** has 37 taxa added in the initial clustering process and another 8 taxa added during a second phase (single taxon addition). The result shows reasonable clusters representing the α-Proteobacteria, γ-Proteobacteria, Actinobacteria, and Frimicutes, plus a few other taxa from different poorly represented groups. The midpoint-rooted result again shows the Actinobacteria clustering with the bulk of the Firmicutes in a similar fashion to that shown in **Figures 4**, **5**. The other well-supported trees from this analysis either also show such a clustering or do not contain taxa that can address the relationship between the Actinobacteria and Firmicutes. Also, observed in **Figure 6** is the splitting of the Firmicutes into two groups with the Streptococcaceae (*Streptococcus* and *Lactococcus*) falling away from the bulk of the Firmicutes. A similar result was observed in the **Figure 4**, albeit with a different ultimate affinity for the
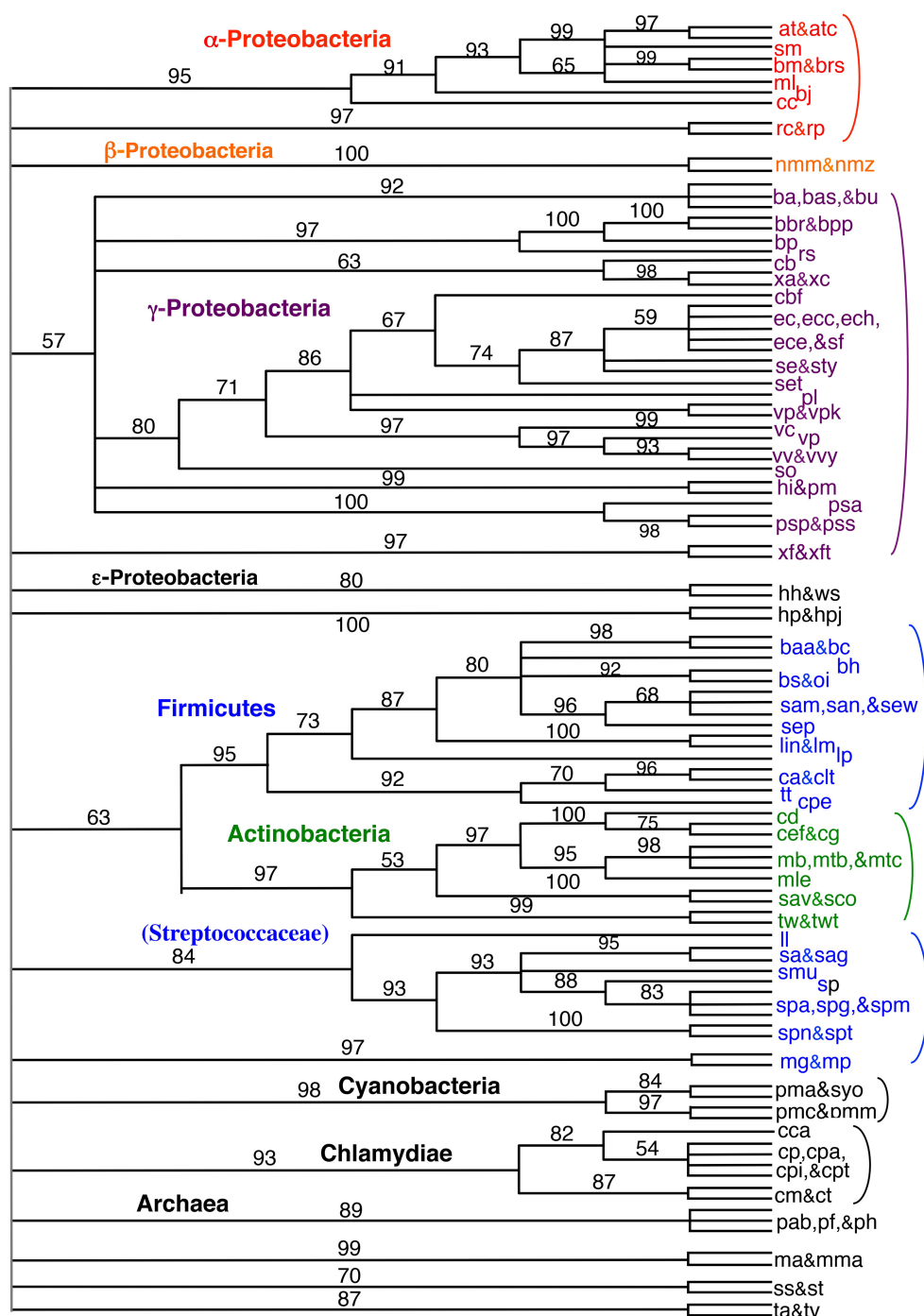
FIGURE 5 | "Bootstrap" NJ cladogram of the gene order distance tree shown in Figure 4. Each node shows the number of times that node appears in 100 replicate trees each using gene order distances based on 100,000 iterations. Select taxonomic groups are labeled with the same color scheme as used in later figures.

Streptococcaceae. The inconsistent placing of this group on the trees found in **Figures 4**, **7**, plus the unresolved placing of this group in **Figure 5** and the exclusion of this group from **Figure 6**, collectively suggests that gene order is unable to confidently place this group on the tree—leaving it inconclusive to the question of whether they belong with the rest of the Firmicutes or even clustered with the gram-positive bacteria, but diverged prior to the Actinobacteria. However, the fact that *Lactobacillus* (labeled lp) is consistently clustered with the bulk of Firmicutes suggests that the Lactobacillales (which includes the Streptococcaceae) do belong
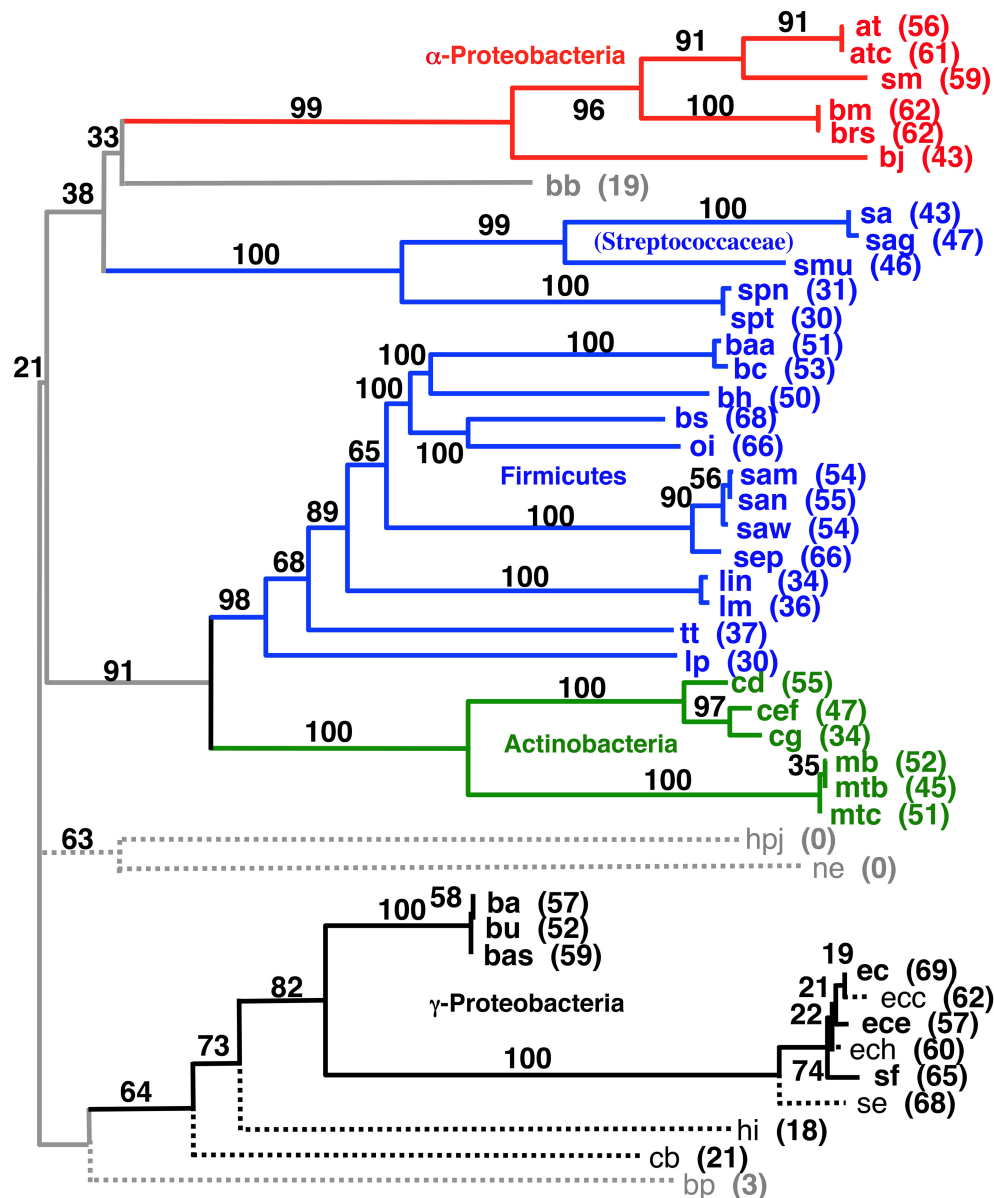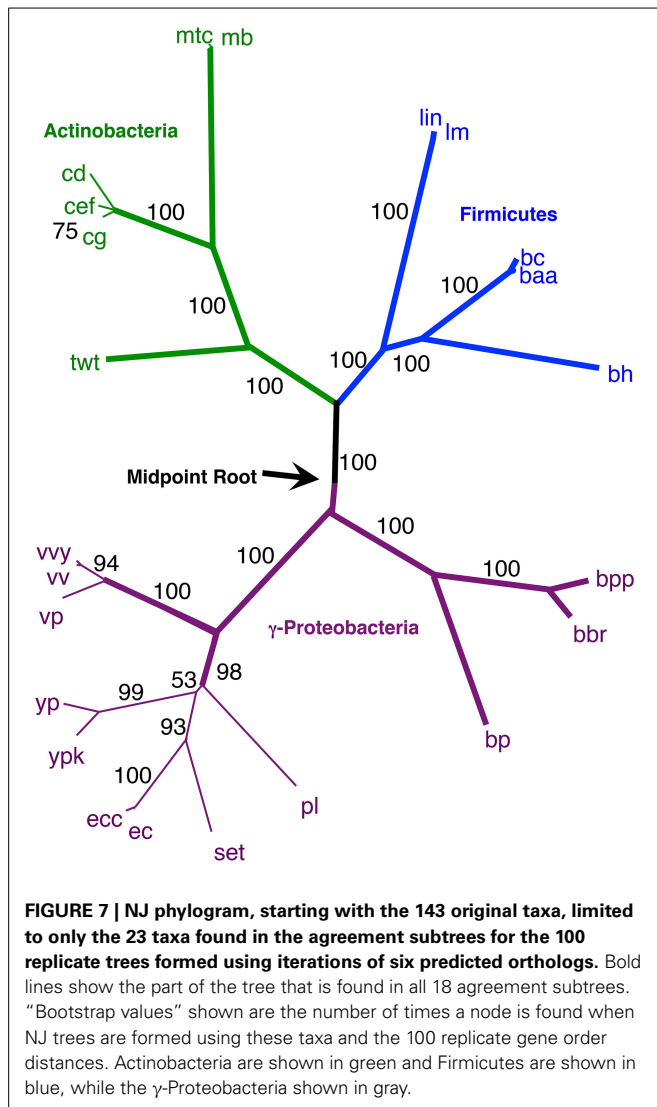
FIGURE 6 | Midpoint-rooted NJ phylogram based on our hierarchical tree building starting with 143 genomes (see Materials and Methods) using the same gene order data distances as the tree shown in Figure 4. This resultant tree includes the most taxa during the initial round of clustering with solid lines and bold font. Taxa connected with dashed lines are those found to be compatible during a second round of single taxon addition. "Bootstrap values" shown are the number of times a node is found when NJ trees are formed using these taxa and the 100 replicate gene order distances. The values listed for individual taxa are the number of times that taxon is found in the biggest tree formed by the initial round of clustering when the 100 replicate gene order distances are used. Taxa not shown that were found 60 or more times in the largest tree after the initial clustering were: vc (70), cbf (67), vv (65), sty (63), set (61), and vp (60).

with the rest of the Firmicutes, and therefore, in this case, the Streptococcaceae appear to be misplaced due to an artifact related to "long branch attraction."

Secondly, using our original NJ trees, we identified the agreement subtrees for the 100 replicate NJ trees that had previously been constructed (and used for bootstrap analysis). Starting with the 100 trees, 18 agreement subtrees (each containing 18 taxa) were found. Together, the agreement subtrees contained a total of 23 different taxa. These 23 taxa were then used to build a

new NJ tree (**Figure 7**) using the dataset constructed from all 10 million iterations. The result shows with high confidence three microbial groups—the Actinobacteria, the Firmicutes, and the γ-Proteobacteria. This pruned tree is the consistent core of the 100 replicate trees, and indicates that there is significant (but small) gene order conservation between these three taxonomic groups. When this tree is midpoint-rooted, the Actinobacteria and Firmicutes are united as sister groups with high confidence, which further suggests that the gram-positive bacteria might be

FIGURE 7 | NJ phylogram, starting with the 143 original taxa, limited to only the 23 taxa found in the agreement subtrees for the 100 replicate trees formed using iterations of six predicted orthologs. Bold lines show the part of the tree that is found in all 18 agreement subtrees. "Bootstrap values" shown are the number of times a node is found when NJ trees are formed using these taxa and the 100 replicate gene order distances. Actinobacteria are shown in green and Firmicutes are shown in blue, while the γ-Proteobacteria shown in gray.

monophyletic (as long as the assumptions inherent to midpoint-rooting are met). Based both of the conservative nature of this agreement substrees approach and the sensible results that it produces, we think that this is our best option for constructing a large gene order-based tree of prokaryotes.

GENE ORDER TREE BUILDING STARTING FROM A MORE REPRESENTATIVE 172 TAXA

Finally, we repeated our agreement subtrees approach for our updated study of the relations of prokaryotes using 172 complete genomes (**Table 3**) better representing a wider-diversity of prokaryotes. With this fuller dataset, starting with 100 replicate NJ trees, the agreement subtree only contained 13 taxa. These 13 taxa were then used to build a NJ tree as before (**Figure 7**). As before, the resultant tree shows with high confidence that Actinobacteria and Firmicutes are sister groups (**Figure 8**). We also repeated this final analysis selecting five orthologs in the same order rather than six. This resulted in a summary agreement subtree with 56 taxa suggesting there is significantly more genomic

gene order signal with five genes than with six. The 56 taxa tree (**Figure 9**), which now includes Archaea and Bacteria, again shows with high confidence that Actinobacteria and Firmicutes are sister groups forming a gram-positive clade (**Figure 9**). The midpoint rooting of this final tree (**Figure 9**) places Archaea as a sister group to the γ-Proteobacteria. At face value, this suggests there is a little more gene order conservation between the Archaea and the γ-Proteobacteria than with any other bacterial group. Gene order conservation between Archaea and the γ-Proteobacteria would argue against the "neomuran origin" for the archaea cell (Cavalier-Smith, 2002). A pairing of Archaea with the γ-Proteobacteria, though, should be taken with significant caution because the result is completely dependent on the midpoint rooting, which may incorrectly represent the history of these evolving groups. Using the Archaea as an outgroup, naturally would place the Proteobacteria with the other bacterial phylum represented. In either case, though, the tree supports the notion that the gram positive bacteria (Actinobacteria and Firmicutes) evolved once from a gram-negative relative. It is also notable that the genome-wide synteny tree of 89 microbes published by Shifman et al. (2014) also shows the Actinobacteria and Firmicutes united as sister groups, even though hat particular work used different genomes and a different approach to estimate gene order similarity across genomes.

IMPLICATIONS OF GENE ORDER CONSERVATION FOUND

At this point, we can conclude that starting from a large number of genomes, we find, perhaps surprising, that there is some gene order conservation between a few major groups, namely Firmicutes, Actinobacteria, and Proteobacteria (**Figures 4–9**) and less robustly the Archaea and Proteobacteria (**Figure 9**). Comparison of genomes from closely related species reveals that inversions are quite common. Large inversions involving up to half of the genome are found frequently between closely related species (e.g., within the *Pyrococcus* genus, Zivanovic et al. (2002), within the *Yersinia* genus, Darling et al., 2008). Given this potentially very rapid rate of divergence in gene order, it is surprising to find residual phylogenetic signal still uniting such distant groups as the Actinobacteria and the Firmicutes. However, while large inversions are common, they are not random in their distribution. For example inversions that disrupt the symmetry of the replicons are frequently not tolerated (Eisen et al., 2000; Zivanovic et al., 2002; Darling et al., 2008). Thus, the rapid changes may be restricted in their range leaving large portions of the genome with potentially conserved gene order over large time scales.

Taken together, our results suggest that the Actinobacteria is a sister group to the Firmicutes, which in turn implies a single origin for the gram-positive cell. Since the first few whole genome sequences were published, some genomic trees have failed to unite these groups (Brown et al., 2001; Fu and Fu-Liu, 2002; Korbel et al., 2002), while others have found weak support for the pairing (House et al., 2003) or have found the pairing under a subset of conditions tried (Deeds et al., 2005). There are three possible disparate causes for these results. First, it is possible that the gram-positive cell has evolved more than once in Earth history. In particular, it has been suggested that *Mycobacterium* may have a close relationship to gram-negative bacteria (Fu and Fu-Liu,
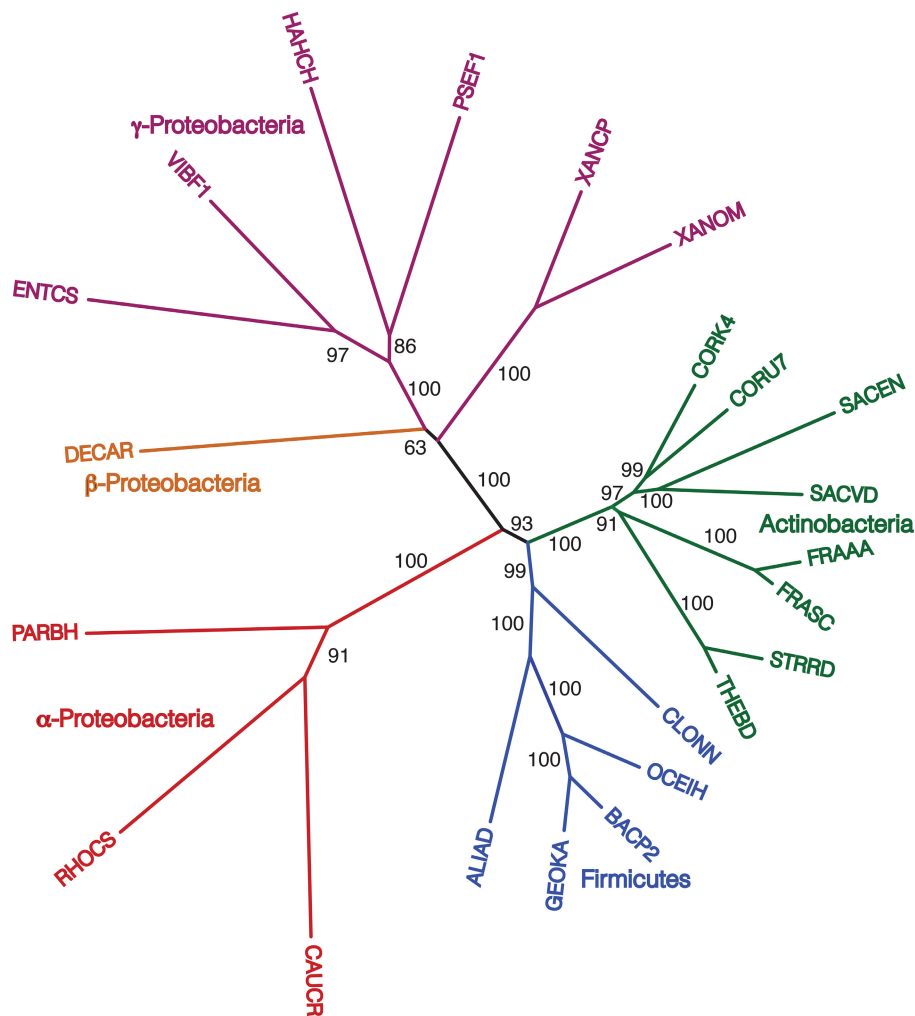
FIGURE 8 | NJ phylogram, starting with 172 representative taxa, limited to only the 23 taxa found in the agreement subtrees for the 100 replicate trees formed using iterations of six predicted orthologs. "Bootstrap values" shown are the number of times a node is found when NJ trees are formed using these taxa and the 100 replicate gene order distances.

2002). Second, it has been hypothesized that gram-positive bacteria are more primitive than gram-negative bacteria (Gupta, 1998; Errington, 2013). Third, some researchers are of the opinion that the failure of genomic methods to unite the gram-positive bacteria together indicates that genomic methods are still inadequate to address this relationship (Olsen, 2001), and that ultimately, we will find that the gram-positive bacteria could be united as a monophyletic group. In particular, the strong similarity in the structure of the cell walls of Firmicutes and Actinobacteria argues for a single origin. The gram-positive cell type, found in both Firmicutes and Actinbacteria, consists of thick layers of peptioglycan with teichoic acids and a single membrane. Gram-negative bacteria have a thin peptidoglycan layer, lack teichoic acid, and have a second outer membrane with lipopolysaccharides.

Considering that our gene order analyses have consistently produced trees with the Actinobacteria united with the bulk of the Firmicutes to the exclusion of other bacterial groups (mostly the Proteobacteria), our results support the uniting of these

groups and argue against multiple origins for the gram-positive cell type. The strongest evidence against a strict monophyletic pairing of the Firmicutes with the Actinobacteria comes from the (unrooted) phylogenetic analysis of 31 concatenated bacterial genes (Wu et al., 2009) and 24 concatenated bacterial genes (Lang et al., 2013), which appear to support a mostly gram-positive clade of Firmicutes, Actinobacteria, Chloroflexi, and Cyanobacteria. Incidentally, Lang et al. (2013) also show the Tenericutes as part of the Firmicutes. At present, we cannot rule out such a larger (primarily) gram-positive clade because it is possible that other phyla (like the Tenericutes) will be included within our Firmicutes/Actinobacteria cluster when taxa sampling increases for gene order studies. Generally, one can argue that because several of our trees (those restricted to agreement subtrees) do not include any taxa from bacterial groups other than the Proteobacteria, we cannot rule out the possibility that one of the other phyla, such as the Cyanobacteria, would break up our Firmicutes/Actinobacteria clade. However, such reasoning does
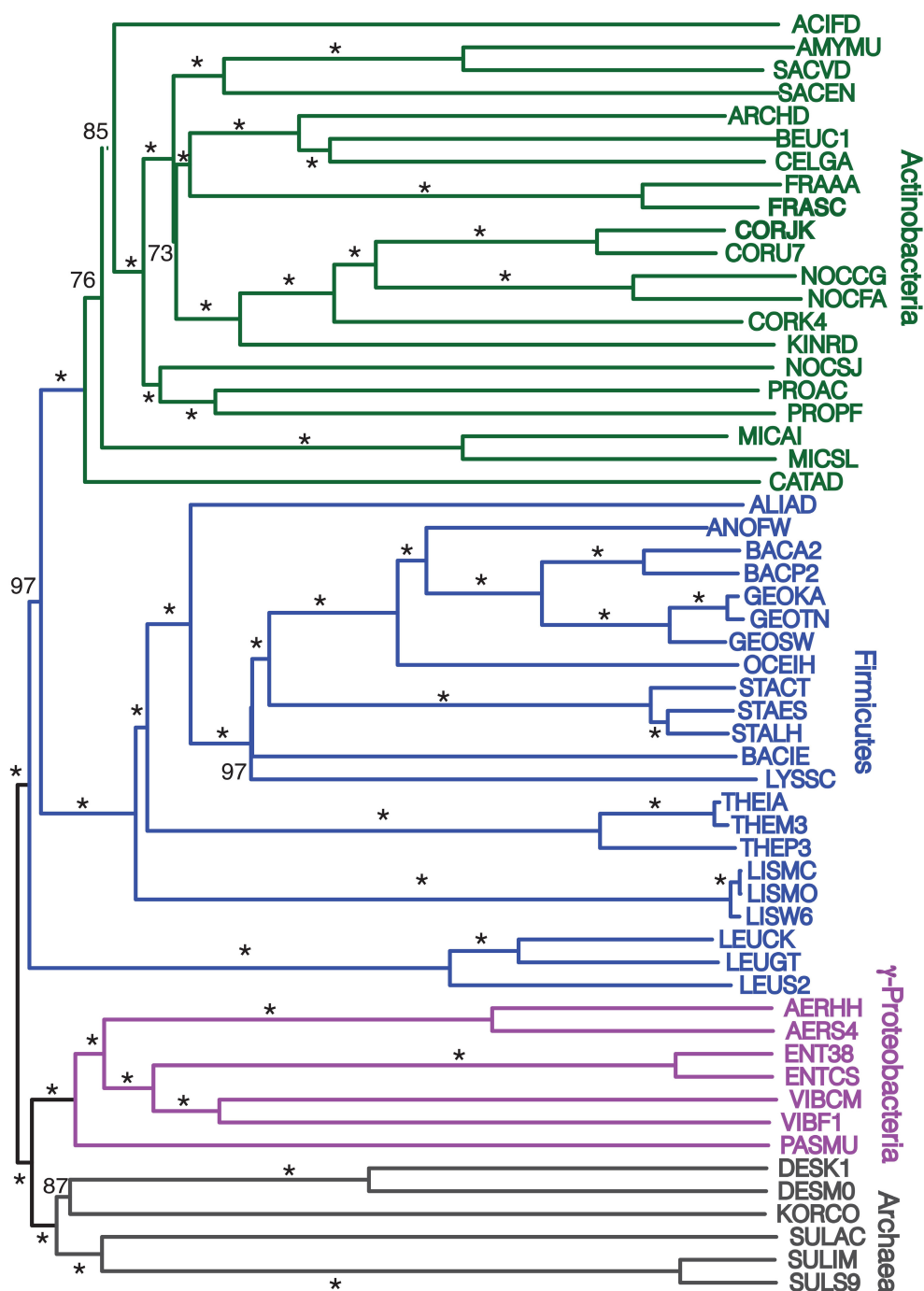
FIGURE 9 | Midpoint-rooted NJ phylogram, starting with 172 representative taxa, limited to only the 56 taxa found in the agreement subtrees for the 100 replicate trees formed using iterations of five predicted orthologs with the same distance equation as before, which ends up functionally equivalent to using $D' = -\ln(1-D)$. "Bootstrap values" shown are the number of times a node is found when NJ trees are formed using these taxa and the 100 replicate gene order distances with * representing a bootstrap value of 100.

requires the taxa within such a phyla to have all scrambled their gene order to the point to which they show no affinity to either the Firmicutes or Actinobacteria in spite of their supposed closer affinity. Our results though still show a uniquely strong conservation of gene order between the Firmicutes and Actinobacteria.

We, therefore, feel that our results are indicative of a tree of life in which most other bacteria phyla diverged prior to the base of a gram-positive cluster (either Firmicutes/Actinobacteria or a larger similar clade). This interpretation in turn implies a single origin for the gram-positive cell. Our results also indicate that

gene order of certain genomes are phylogenetically informative at both low and high taxonomic levels, but that for many other genomes gene order is not conserved for a long time.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fmicb.2014.00785/abstract

REFERENCES

Auch, A. F., von Jan, M., Klenk, H. P., and Göker, M. (2010). Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand. Genomic Sci.* 2, 117. doi: 10.4056/sigs.531120

Belda, E., Moya, A., and Silva, F. J. (2005). Genome rearrangement distances and gene order phylogeny in gamma-Proteobacteria. *Mol. Biol. Evol.* 22, 1456–1467. doi: 10.1093/molbev/msi134

Bischoff, J., Domrachev, M., Federhen, S., Hotton, C., Leipe, D., Soussov, V., et al. (2007). *NCBI Taxonomy Browser*. Available online at: http://www.ncbi.nlm.nih.gov/Taxonomy/

Blanchette, M., Kunisawa, T., and Sankoff, D. (1999). Gene order breakpoint evidence in animal mitochondrial phylogeny. *J. Mol. Evol.* 49, 193–203.

Brown, J. R., Douady, C. J., Italia, M. J., Marshall, W. E., and Stanhope, M. J. (2001). Universal trees based on large combined protein sequence data sets. *Nat. Genet.* 28, 281–285. doi: 10.1038/90129

Brown, J. R., and Volker, C. (2004). Phylogeny of gamma-proteobacteria: resolution of one branch of the universal tree? *Bioessays* 26, 463–468. doi: 10.1002/bies.20030

Cavalier-Smith, T. (2002). The neomuran origin of archaebacteria, the negibacterial root of the universal tree and bacterial megaclassification. *Int. J. Syst. Evol. Microbiol.* 52, 7–76.

Cole, J. R., Chai, B., Farris, R. J., Wang, Q., Kulam-Syed-Mohideen, A. S., and McGarrell, D. M. (2007). The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. *Nucleic Acids Res.* 35, D169–D172. doi: 10.1093/nar/gkl889

Darling, A. E., Miklós, I., and Ragan, M. A. (2008). Dynamics of genome rearrangement in bacterial populations. *PLoS Genet.* 4:e1000128. doi: 10.1371/journal.pgen.1000128

Deeds, E. J., Hennessey, H., and Shakhnovich, E. I. (2005). Prokaryotic phylogenies inferred from protein structural domains. *Genome Res.* 15, 393–402. doi: 10.1101/gr.3033805

Deloger, M., El Karoui, M., and Petit, M. A. (2009). A genomic distance based on MUM indicates discontinuity between most bacterial species and genera. *J. Bacteriol.* 191, 91–99. doi: 10.1128/JB.01202-08

De Rijk, P., van de Peer, Y., van den Broeck, I., and de Wachter, R. (1995). Evolution according to large ribosomal subunit RNA. *J. Mol. Evol.* 41, 366–375.

Dessimoz, C., Cannarozzi, G., Gil, M., Margadant, D., Roth, A., Schneider, A., et al. (2005). "OMA, a comprehensive, automated project for the identification of orthologs from complete genome data: introduction and first achievements," in *Comparative Genomics* (Berlin; Heidelberg: Springer), 61–72.

Eisen, J. A., Heidelberg, J. F., White, O., and Salzberg, S. L. (2000). Evidence for symmetric chromosomal inversions around the replication origin in bacteria. *Genome Biol.* 1, 0011.1–0011.9. doi: 10.1186/gb-2000-1-6-research0011

Errington, J. (2013). L-form bacteria, cell walls and the origins of life. *Open Biol.* 3:120143. doi: 10.1098/rsob.120143

Fitz-Gibbon, S. T., and House, C. H. (1999). Whole genome-based phylogenetic analysis of free-living microorganisms. *Nucleic Acids Res.* 27, 4218–4222. doi: 10.1093/nar/27.21.4218

Fu, L. M., and Fu-Liu, C. S. (2002). Is Mycobacterium tuberculosis a closer relative to Gram-positive or Gram-negative bacterial pathogens? *Tuberculosis* 82, 85–90. doi: 10.1054/tube.2002.0328

Gerstein, M. (1998). Patterns of protein-fold usage in eight microbial genomes: a comprehensive structural census. *Proteins* 33, 518–534.

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., and Tiedje, J. M. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57, 81–91. doi: 10.1099/ijs.0.64483-0

Gupta, R. S. (1998). Protein phylogenies and signature sequences: a reappraisal of evolutionary relationships among archaeabacteria, eubacteria, and eukaryotes. *Microbiol. Mol. Biol. Rev.* 62, 1435–1491.

House, C. H. (2009). "The tree of life viewed through the contents of genomes," in *Horizontal Gene Transfer: Genomes in Flux*, eds M. B. Gogarten, J. P. Gogarten, and L. Olendzenski (New York, NY: Humana Press), 141–161.

House, C. H., Runnegar, B., and Fitz-Gibbon, S. T. (2003). Geobiological analysis using whole genome-based tree building applied to the Bacteria, Archaea, and Eukarya. *Geobiology* 1, 15–26. doi: 10.1046/j.1472-4669.2003.00004.x

Hu, F., Gao, N., Zhang, M., and Tang, J. (2011). "Maximum likelihood phylogenetic reconstruction using gene order encodings," in *Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), 2011 IEEE Symposium* (Paris: IEEE), 1–6.

Jukes, T. H., and Cantor, C. R. (1969). "Evolution of protein Molecules," in *Mammalian Protein Metabolism*, ed H. Munro (New York, NY: Academic Press), 21–132.

Kimura, M., and Ohta, T. (1972). On the stochastic model for estimation of mutational distance between homologous proteins. *J. Mol. Evol.* 2, 87–90. doi: 10.1007/BF01653945

Konstantinidis, K. T., and Tiedje, J. M. (2005). Towards a genome-based taxonomy for prokaryotes. *J. Bacteriol.* 187, 6258–6264. doi: 10.1128/JB.187.18.6258-6264.2005

Korbel, J. O., Snel, B., Huynen, M. A., and Bork, P. (2002). SHOT: a web server for the construction of genome phylogenies. *Trends Genet.* 18, 158–62. doi: 10.1016/S0168-9525(01)02597-5

Kunisawa, T. (2001). Gene arrangements and phylogeny in the class Proteobacteria. *J. Theor. Biol.* 213, 9–19. doi: 10.1006/jtbi.2001.2396

Kunisawa, T. (2003). Gene arrangements and branching orders of gram-positive bacteria. *J. Theor. Biol.* 222, 495–503. doi: 10.1016/S0022-5193(03)00064-X

Lang, J. M., Darling, A. E., and Eisen, J. A. (2013). Phylogeny of bacterial and archaeal genomes using conserved genes: supertrees and supermatrices. *PLoS ONE* 8:e62510. doi: 10.1371/journal.pone.0062510

Lin, Y., Hu, F., Tang, J., and Moret, B. (2013). "Maximum likelihood phylogenetic reconstruction from high-resolution whole-genome data and a tree of 68 eukaryotes," in *Pacific Symposium on Biocomputing* (Waimea, HI).

Lin, Y., and Moret, B. M. (2011). A new genomic evolutionary model for rearrangements, duplications, and losses that applies across eukaryotes and prokaryotes. *J. Comput. Biol.* 18, 1055–1064. doi: 10.1089/cmb.2011.0098

Meier-Kolthoff, J. P., Göker, M., Spröer, C., and Klenk, H. P. (2013). When should a DDH experiment be mandatory in microbial taxonomy? *Arch. Microbiol.* 195, 413–418. doi: 10.1007/s00203-013-0888-4

Moret, B. M. E., Wang, L. S., Warnow, T., and Wyman, S. K. (2001). New approaches for reconstructing phylogenies from gene order data. *Bioinformatics* 17(Suppl. 1), S165–S173. doi: 10.1093/bioinformatics/17.suppl_1.S165

Nadeau, H., and Taylor, B. (1984). Lengths of chromosomal segments conserved since divergence of man and mouse. *Proc. Natl. Acad. Sci. U.S.A.* 81, 814–818. doi: 10.1073/pnas.81.3.814

Olsen, G. J. (2001). The history of life. *Nat. Genet.* 28, 197–198. doi: 10.1038/90014

Richter, M., and Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19126–19131. doi: 10.1073/pnas.0906412106

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Sankoff, D., Leduc, G., Antoine, N., Paquin, B., Lang, B. F., and Cedergren, R. (1992). Gene order comparisons for phylogenetic inference: evolution of the mitochondrial genome. *Proc. Natl. Acad. Sci. U.S.A.* 89, 6575–6579. doi: 10.1073/pnas.89.14.6575

Schwartz, R. M., and Dayhoff, M. O. (1978). "Matrices for detecting distant relationships," in *Atlas of Protein Sequence and Structure*, Vol. 5, ed M. O. Dayhoff (Washington, DC: National Biomedical Research Foundation), 353–358.

Shao, M., Lin, Y., and Moret, B. (2013). Sorting genomes with rearrangements and segmental duplications through trajectory graphs. *BMC Bioinformatics* 14:S9. doi: 10.1186/1471-2105-14-S15-S9

Shifman, A., Ninyo, N., Gophna, U., and Snir, S. (2014). Phylo SI: a new genome-wide approach for prokaryotic phylogeny. *Nucleic Acids Res.* 42, 2391–2404. doi: 10.1093/nar/gkt1138

Snel, B., Bork, P., and Huynen, M. A. (1999). Genome phylogeny based on gene content. *Nat. Genet.* 21, 108–110. doi: 10.1038/5052

Swenson, K. M., Arndt, W., Tang, J., and Moret, B. M. (2008). "Phylogenetic reconstruction from complete gene orders of whole genomes," in *Asia Pacific Bioinformatics Conference Proceedings* (Kyoto), 241–250.

Tajima, F. (1993). Unbiased estimation of evolutionary distance between nucleotide sequences. *Mol. Biol. Evol.* 10, 677–688.

Tamura, T., Matsuzawa, T., Oji, S., Ichikawa, N., Hosoyama, A., Katsumata, H., et al. (2012). A genome sequence-based approach to taxonomy of the genus Nocardia. *Antonie Van Leeuwenhoek* 102, 481–491. doi: 10.1007/s10482-012-9780-5

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., et al. (2003). The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* 4:41. doi: 10.1186/1471-2105-4-41

Tekaia, F., Lazcano, A., and Dujon, B. (1999). The genomic tree as revealed from whole proteome comparisons. *Genome Res.* 9, 550–557.

Watterson, W. A., Ewens, W. J., Hall, T. E., and Morgan, A. (1982). The chromosome inversion problem. *J. Theor. Biol.* 99, 1–7 doi: 10.1016/0022-5193(82)90384-8

Woese, C. R. (1987). Bacterial evolution. *Microbiol. Rev.* 51, 221–271.

Wolf, Y. I., Rogozin, I. B., Grishin, N. V., and Koonin, E. V. (2002). Genome trees and the tree of life. *Trends Genet.* 18, 472–479. doi: 10.1016/S0168-9525(02)02744-0

Wolf, Y. I., Rogozin, I. B., Grishin, N. V., Tatusov, R. L., and Koonin, E. V. (2001). Genome trees constructed using five different approaches suggest new major bacterial clades. *BMC Evol. Biol.* 1:8. doi: 10.1186/1471-2148-1-8

Wu, D., Hugenholtz, P., Mavromatis, K., Pukall, R., Dalin, E., Ivanova, N. N., et al. (2009). A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature* 462, 1056–1060. doi: 10.1038/nature08656

Yang, S., Doolittle, R. F., and Bourne, P. E. (2005). Phylogeny determined by protein domain content. *Proc. Natl. Acad. Sci. U.S.A.* 102, 373–378. doi: 10.1073/pnas.0408810102

Zhang, Y., Hu, F., and Tang, J. (2010). "Phylogenetic reconstruction with gene rearrangements and gene losses," in *Bioinformatics and Biomedicine (BIBM), 2010 IEEE International Conference* (Hong Kong: IEEE), 35–38.

Zivanovic, Y., Lopez, P., Philippe, H., and Forterre, P. (2002). *Pyrococcus* genome comparison evidences chromosome shuffling-driven evolution. *Nucl. Acids Res.* 30, 1902–1910. doi: 10.1093/nar/30.9.1902

The distribution and impact of viral lineages in domains of life

Arshan Nasir[1], Patrick Forterre[2,3], Kyung Mo Kim[4] and Gustavo Caetano-Anollés[1]*

[1] Evolutionary Bioinformatics Laboratory, Department of Crop Sciences, Illinois Informatics Institute, University of Illinois, Urbana-Champaign, Urbana, IL, USA
[2] Unité BMGE, Institute Pasteur, Paris, France
[3] Institut de Génétique and Microbiologie, Université Paris-Sud, CNRS UMR8621, Orsay, France
[4] Microbial Resource Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon, Korea
*Correspondence: gca@illinois.edu

Living organisms can be conveniently classified into three domains, Archaea, Bacteria, and Eukarya (Woese et al., 1990). The three domains are united by several features that support the common origin of life including the presence of ribosomes, double-stranded DNA genomes, a nearly universal genetic code, physical compartments (i.e., membranes), and the ability to carry out metabolism and oxidation-reduction reactions. In comparison, other types of genetic material and particles (e.g., viruses, plasmids, and other selfish genetic elements) are often excluded from the definition of "life" (for opposing views see Raoult and Forterre, 2008; Forterre, 2011, 2012a). However, they can still influence the evolution of cellular organisms, and in conjunction, establish complex life cycles.

Viruses impact our economy, medicine and agriculture due to their infectious nature. Viral infections transform the host cell into a virocell that no longer divides by binary fission but produces more viral particles or a ribovirocell in which the viral and cellular genomes coexist, the cell still dividing while producing virions (Forterre, 2011, 2012a). The virosphere (i.e., collection of all viruses) displays exceptional variability in virion morphologies and replication strategies. Viruses can be classified into DNA or RNA viruses, retroviruses or intermediate forms depending upon the type of replicon present inside the viral particle. Moreover, replicons could be linear, circular, single-stranded, double-stranded, or even segmented. The unprecedented diversity of replicon types has led to the proposal

that viruses first invented DNA as means to trick the host defense systems (Forterre, 2002, 2005). Viruses can also transfer genes between species and enhance biodiversity (Nasir et al., 2012). Even more importantly, viruses appear to create massive amount of new genetic information, part of which can transfer to cells (Abroi and Gough, 2011; Forterre, 2011, 2012b). The discovery of "giant" viruses such as mimiviruses (La Scola et al., 2003), megaviruses (Arslan et al., 2011), pandoraviruses (Philippe et al., 2013), and pithoviruses (Legendre et al., 2014) now creates a continuum in genome size and functional complexity between the virosphere and cells. Still, viruses are neglected in phylogenetic studies because they lack a unifying genetic marker, similar to rRNA for cells, and because many biologists underestimate their genetic creativity. As a consequence, their role in the origin and evolution of modern life, and their impact on the ecology of our biosphere continue to be for the most part unrecognized (Koonin and Wolf, 2012). In this opinion article, we address the impact of viruses on the evolution of cells. We argue that viruses likely initiated major evolutionary shifts. Specifically, we consider that gain and loss of viral lineages often leads to divergent evolutionary trends even in closely related species. We emphasize that no evolutionary theory could be complete without accounting for the viral world and that viruses are responsible for ongoing adaptations in the cellular domains (see also Prangishvili et al., 2006; Forterre and Prangishvili, 2013; Koonin and Dolja, 2013).

The distribution of the association of viral replicon types with cells is extremely biased. For example, RNA viruses are completely absent in Archaea and are rare in Bacteria. In comparison, vertebrates host numerous RNA and retroviruses. Surprisingly, dsDNA viruses are rare in plants while dsRNA viruses are abundant in fungi. Similarly, retroviruses are integrated into the genomes of multicellular eukaryotes but are completely absent in the microbial genomes. In other words, specific relationships exist between the type of viral replicon and the host range. Viruses with a particular replicon may infect one group of organisms but may not replicate in another. Big jumps of viruses from one cellular lineage to another have been observed within the eukaryotic "division" such as animals (opisthokonts) and plants (viridiplantae), when a virus adapts to an established consortium of ecological partners. The same virus can sometimes infect both plant and animal cells when these are linked by their mode of life. One example is the Fiji disease virus (*Reoviridae*) that can replicate in both its insect vector (Delphacidae) and flowering plants (Kings et al., 2012). However, no modern virus is known to cross the barrier between domains. Therefore, while viruses may be able to jump hosts over short evolutionary time spans, crossing domain boundaries is less likely and not expected to compromise our inferences.

To obtain a quantitative view of viral diversity and its distribution among cellular domains, we extracted genome data from the Viral Genomes Resource at NCBI (Bao et al., 2004). This resource

provides accurate, manually curated information about sequenced viral genomes that is minimally redundant. Generally, one sequenced genome portrays many isolates/strains of the same virus. Specifically, we investigated the host preferences for viruses with different replication strategies (**Figure 1A**) and contrasted virion morphologies (borrowed from ViralZone; Hulo et al., 2011) of virus families infecting different domain groups (**Figure 1B**). A birds-eye view of the distribution of viruses among hosts revealed that only 63 were exclusive to the archaeal domain (hereinafter referred to as archaeoviruses) (**Figure 1A**). In comparison, 1251 bacterial (bacterioviruses, formerly bacteriophages) and 2321 eukaryal viruses (eukaryoviruses) were identified. The low number of archaeoviruses is clearly due to a sampling bias (e.g., the low number of archaeal species screened for the presence of viral infection) since it has been shown that four different viruses can infect a single archaeal species (i.e., *Aeropyrum pernix*), each from a different family (Mochizuki et al., 2010, 2011, 2012). Despite their low number, archaeoviruses exhibit greater virion morphotype diversity compared to bacterioviruses [e.g., 4 unique virion morphotypes vs. none (**Figure 1B**); see also Pietilä et al., 2014]. In comparison, bacterial organisms host a vast number of described DNA viruses (1178 out of total 1760) but display very little family and morphotype diversity. In fact, 95% of the dsDNA bacterioviruses belong to just one order (Caudovirales) and only three families (*Myoviridae*, *Podoviridae*, and *Siphoviridae*). Moreover, only 9 virion morphologies have been observed in bacterioviruses (compared to 16 in archaeoviruses) (Pietilä et al., 2014). One explanation for the low diversity of bacterioviruses could be the invention of peptidoglycan-containing cell wall in Bacteria. The inability to traverse this barrier likely resulted in loss of many viral lineages in Bacteria (Forterre and Prangishvili, 2013; Prangishvili, 2013). Taken together, these observations suggest that Archaea are likely infected by a greater number of viral lineages than Bacteria. This is showcased by their virion morphologies diversity **Figure 1**) (Pina et al., 2011; Pietilä et al., 2014),
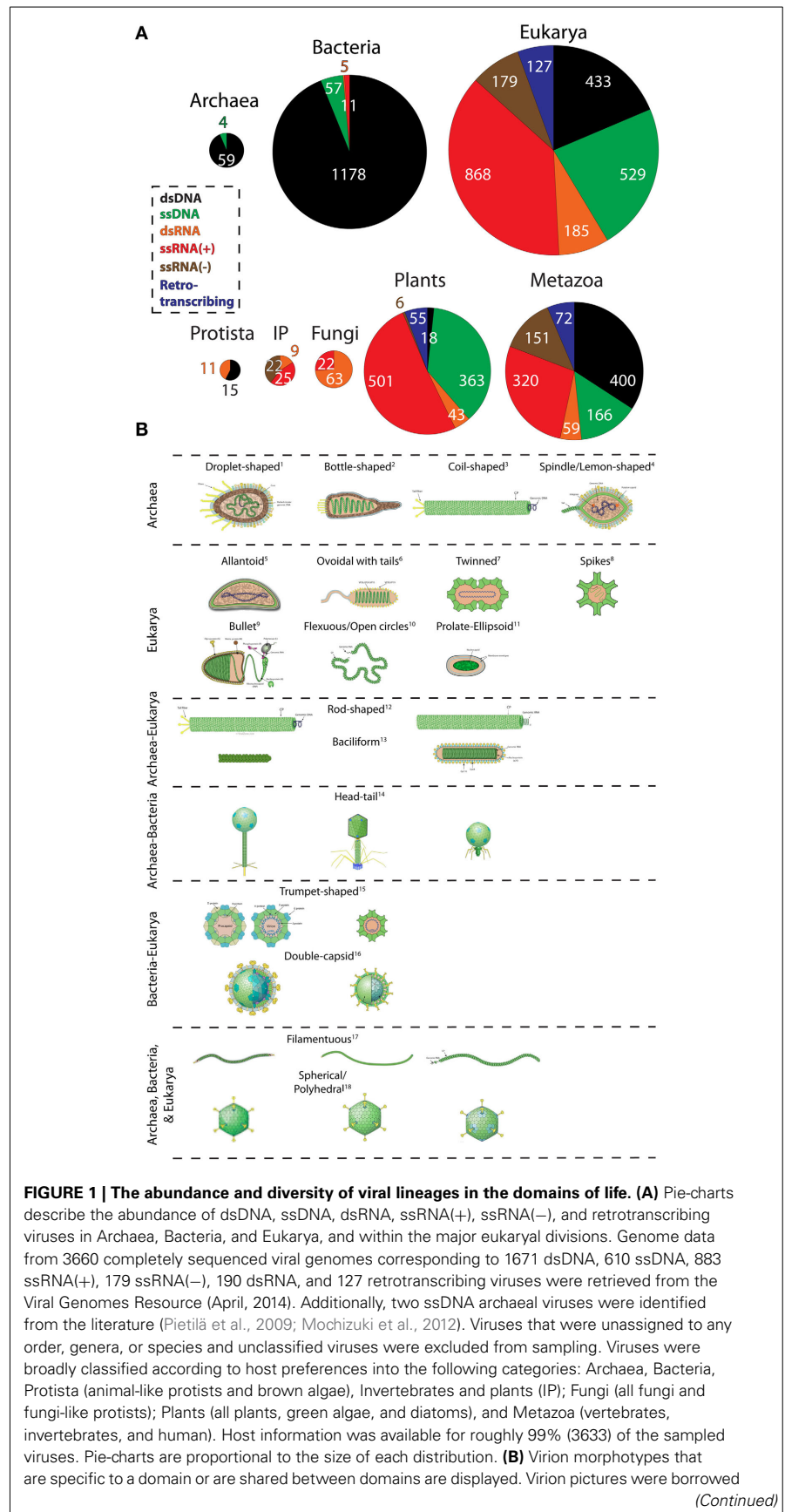

FIGURE 1 | The abundance and diversity of viral lineages in the domains of life. (A) Pie-charts describe the abundance of dsDNA, ssDNA, dsRNA, ssRNA(+), ssRNA(−), and retrotranscribing viruses in Archaea, Bacteria, and Eukarya, and within the major eukaryal divisions. Genome data from 3660 completely sequenced viral genomes corresponding to 1671 dsDNA, 610 ssDNA, 883 ssRNA(+), 179 ssRNA(−), 190 dsRNA, and 127 retrotranscribing viruses were retrieved from the Viral Genomes Resource (April, 2014). Additionally, two ssDNA archaeal viruses were identified from the literature (Pietilä et al., 2009; Mochizuki et al., 2012). Viruses that were unassigned to any order, genera, or species and unclassified viruses were excluded from sampling. Viruses were broadly classified according to host preferences into the following categories: Archaea, Bacteria, Protista (animal-like protists and brown algae), Invertebrates and plants (IP); Fungi (all fungi and fungi-like protists); Plants (all plants, green algae, and diatoms), and Metazoa (vertebrates, invertebrates, and human). Host information was available for roughly 99% (3633) of the sampled viruses. Pie-charts are proportional to the size of each distribution. **(B)** Virion morphotypes that are specific to a domain or are shared between domains are displayed. Virion pictures were borrowed

(Continued)

(which is expected to grow with improvements in our ability to isolate viruses from atypical habitats.

Interestingly, all archaeoviruses possess DNA replicons but no RNA genomes. The complete absence of RNA viruses in Archaea can be linked to high temperature RNA instability (Forterre, 2013). We speculate that escape from RNA viruses could be one major trigger for the evolution of modern Archaea (Forterre, 2013). Thus, loss of RNA viral lineages likely initiated archaeal migration to the harsh environments. One recent study reported the isolation of ssRNA(+) viruses from an archaea-rich community in a hot, acidic spring of Yellowstone National Park (Bolduc et al., 2012). However, their host tropism could not be established with confidence. Finally, four ssDNA viruses were recently isolated from Archaea (Pietilä et al., 2009; Mochizuki et al., 2012; Sencilo et al., 2012). Of these, *Aeropyrum* coil-shaped virus (*Spiraviridae*) is the largest known ssDNA virus and displays unique coil-shaped virion morphology (**Figure 1B**; Mochizuki et al., 2012).

Bacterioviruses are remarkably successful in Bacteria and are highly abundant. Their virions outnumber their bacterial hosts in oceans, balance microbial populations in the marine communities, and regulate biogeochemical cycles (Breitbart and Rohwer, 2005; Suttle, 2007; Rohwer and Thurber, 2009; Zhao et al., 2013). Among the dsDNA bacterioviruses, tailed-bacteriophages exhibit extensive similarities with archaeal caudovirales, suggesting that they form a monophyletic group (Krupovic et al., 2010). Archaeal

and bacterial caudovirales have indeed been grouped in a single major evolutionary lineage, together with *Herpesviridae*. All of these viruses share the same Hong Kong fold (HK97) in their major capsid proteins and homologous packaging ATPases (Baker et al., 2005; Pell et al., 2009; Krupovic et al., 2010; Abrescia et al., 2012). Notably, it has been found recently that the capsid of *Herpesviridae* exhibits a small tail similar to those of *Podoviridae* (Schmid et al., 2012). These data suggest that viruses of the HK97-like lineage are very ancient and originated (most likely) prior to the last common ancestor of cells. Another example of viral lineage shared by the three domains is the so-called "PRD1/Adenovirus lineage" of dsDNA viruses characterized by a major capsid protein containing the double-jelly roll fold and a common packaging ATPase (Abrescia et al., 2012). In comparison, ssDNA bacterioviruses are not as successful in Bacteria and correspond to two major families, *Inoviridae* and *Microviridae* (smallest genomes among DNA viruses; Rosario et al., 2012). Viruses in this group replicate by converting their single-stranded DNA genome into a double-stranded intermediate form engineered by host polymerase. These viruses lack their own polymerase and share this property with the ssDNA viruses of Archaea and Eukarya.

In contrast to DNA viruses, RNA viruses are not as successful in Bacteria. Only, 5 dsRNA, and 11 ssRNA(+) bacterioviruses could be identified. In turn, none of the ssRNA(−) and retrotranscribing viruses associated with bacterial hosts. Among the RNA bacterioviruses,

dsRNA viruses (*Cystoviridae*) encode segmented genomes and infect mostly *Pseudomonas* species (Silander et al., 2005). Interestingly, *Cystoviridae* closely resembles eukaryal dsRNA viruses (i.e., *Reoviridae* and *Totiviridae*) in terms of life cycle and homologous RNA-dependent-RNA-polymerase gene sequences (a virus hallmark) (Butcher et al., 1997). Unlike Archaea, Bacteria are also infected by ssRNA(+) viruses (*Leviviridae*). These viruses are amongst the simplest and smallest known viruses, and historically yielded useful insights into mRNA function (Bollback and Huelsenbeck, 2001). Because RNA viruses (ssRNA and dsRNA) infect both Bacteria and Eukarya, their ancestors likely originated from a putative ancient world of cells with RNA genomes and RNA viruses (Forterre, 2005, 2006a,b). This points to the ancient existence of RNA viruses and suggests their loss from Archaea (since loss in one domain is more likely than the independent gain in two!). The instability of RNA at high temperatures supports this hypothesis, since it is likely that the last common ancestor of Archaea was a hyperthermophile (Brochier-Armanet et al., 2011).

Viruses with all possible types of replicons infect eukaryal organisms. RNA viruses are predominant and cover the entire taxonomic range within Eukarya (**Figure 1A**). Eukaryoviruses also exhibit many unique virion morphotypes not observed in the prokaryotic viruses and are unequally distributed in the major eukaryal groups (**Figure 1**). For example, dsDNA viruses are completely absent in fungi and are rare in plants (i.e., only found in green algae). This suggests that these groups have evolved sophisticated mechanisms to eliminate dsDNA viral infections. A good candidate is the cell wall structure found in plants, fungi, and algae. Differences in cell wall composition and rigidity greatly limit means of viral entry into the cell and serve as barriers to viral infections (Dimmock et al., 2007). However, loss of one viral lineage is apparently offset by the gain of other lineages. This is evident from the high RNA virus distribution among plants and fungi. The origin of the diversity and abundance of RNA viruses in eukaryotes but their near absence in prokaryotes is particularly puzzling (Koonin et al., 2006).

For example, ssRNA(−) and retroviruses are highly successful in vertebrates. At first glance, it seems that organism complexity is proportional to the variety of viral infections. For instance, metazoa are infected by a host of retroviruses. Retroviruses can integrate their genomes into host DNA and thus alter gene expression patterns and trigger genomic rearrangements (Arkhipova et al., 2012). These activities can lead to production of novel genes and advanced machineries (Forterre, 2013). In fact, telomerase enzymes are homologous to retroviral proteins and neocentromeres are formed by epigenetic regulation of transposable elements (Singer, 1995; Chueh et al., 2009), both likely transferred from viruses to host cells much earlier in evolution. This argument is further supported by the absence of RNA and retroviruses from unicellular eukaryotes such as yeast, which resemble a prokaryotic lifestyle (Forterre, 2013). Thus, co-evolution between viruses and their hosts may have led to organism complexity in the eukaryotic domain.

The diversity of eukaryoviruses is intriguing, both in terms of genome structure and virion morphology (see **Figure 1B**). In particular, retrotranscribing, ssRNA(−), and many DNA virus families are only present in eukaryotes. Surprisingly, although Archaea and eukaryotes are very similar in term of their basic molecular biology, there are no viral lineages specific for these two domains (Forterre, 2013). Virions with rod-shaped morphology are up to now specific for Archaea and Eukarya (**Figure 1B**), but they harbor DNA and RNA genomes, respectively, and it is unclear if their major coat proteins are evolutionary related (Goulet et al., 2009). The same is probably also true for bacilliform viruses. Notably, the diversity and specificity of eukaryoviruses is difficult to reconcile with the archaeon-bacterium fusion scenarios for the origin of eukaryotes (e.g., Martin and Müller, 1998), as recently argued (Forterre, 2013).

To conclude, the distribution of viral lineages follows an ancient, highly dynamic and ongoing process that impacts the evolution of organisms. New viral lineages often arise from existing ones and may cross species barriers to infect new hosts (e.g., parvovirus; Shackelton et al., 2005), putting enormous evolutionary pressure on cellular organisms and prompting them to unfold molecular and cellular innovation (Forterre and Prangishvili, 2009) in the search of either simplicity or complexity.

REFERENCES

Abrescia, N. G. L., Bamford, D. H., Grimes, J. M., and Stuart, D. I. (2012). Structure unifies the viral universe. *Annu. Rev. Biochem.* 81, 795–822. doi: 10.1146/annurev-biochem-060910-095130

Abroi, A., and Gough, J. (2011). Are viruses a source of new protein folds for organisms? - Virosphere structure space and evolution. *Bioessays* 33, 626–635. doi: 10.1002/bies.201000126

Arkhipova, I. R., Batzer, M. A., Brosius, J., Feschotte, C., Moran, J. V., Schmitz, J., et al. (2012). Genomic impact of eukaryotic transposable elements. *Mob. DNA* 3, 19. doi: 10.1186/1759-8753-3-19

Arslan, D., Legendre, M., Seltzer, V., Abergel, C., and Claverie, J. M. (2011). Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. *Proc. Natl. Acad. Sci. U.S.A.* 108, 17486–17491. doi: 10.1073/pnas.1110889108

Baker, M. L., Jiang, W., Rixon, F. J., and Chiu, W. (2005). Common ancestry of herpesviruses and tailed DNA bacteriophages. *J. Virol.* 79, 14967–14970. doi: 10.1128/JVI.79.23.14967-14970.2005

Bao, Y., Federhen, S., Leipe, D., Pham, V., Resenchuk, S., Rozanov, M., et al. (2004). National center for biotechnology information viral genomes project. *J. Virol.* 78, 7291–7298. doi: 10.1128/JVI.78.14.7291-7298.2004

Bolduc, B., Shaughnessy, D. P., Wolf, Y. I., Koonin, E. V., Roberto, F. F., and Young, M. (2012). Identification of novel positive-strand RNA viruses by metagenomic analysis of archaea-dominated Yellowstone hot springs. *J. Virol.* 86, 5562–5573. doi: 10.1128/JVI.07196-11

Bollback, J. P., and Huelsenbeck, J. P. (2001). Phylogeny, genome evolution, and host specificity of single-stranded RNA bacteriophage (*family Leviviridae*). *J. Mol. Evol.* 52, 117–128. doi: 10.1007/s002390010140

Breitbart, M., and Rohwer, F. (2005). Here a virus, there a virus, everywhere the same virus? *Trends Microbiol.* 13, 278–284. doi: 10.1016/j.tim.2005.04.003

Brochier-Armanet, C., Forterre, P., and Gribaldo, S. (2011). Phylogeny and evolution of the Archaea: one hundred genomes later. *Curr. Opin. Microbiol.* 14, 274–281. doi: 10.1016/j.mib.2011.04.015

Butcher, S., Dokland, T., Ojala, P., Bamford, D., and Fuller, S. (1997). Intermediates in the assembly pathway of the double-stranded RNA virus Phi6. *EMBO J.* 16, 4477–4487. doi: 10.1093/emboj/16.14.4477

Chueh, A. C., Northrop, E. L., Brettingham-Moore, K. H., Choo, K. A., and Wong, L. H. (2009). LINE retrotransposon RNA is an essential structural and functional epigenetic component of a core neocentromeric chromatin. *PLoS Genet.* 5:e1000354. doi: 10.1371/journal.pgen.1000354

Dimmock, N. J., Easton, A. J., and Leppard, K. N. (2007). *Introduction to Modern Virology, 6th Edn.* Malden, MA: Blackwell Publishing Limited.

Forterre, P. (2002). The origin of DNA genomes and DNA replication proteins. *Curr. Opin. Microbiol.* 5, 525–532. doi: 10.1016/S1369-5274(02)00360-0

Forterre, P. (2005). The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells. *Biochimie* 87, 793–803. doi: 10.1016/j.biochi.2005.03.015

Forterre, P. (2006a). Three RNA cells for ribosomal lineages and three DNA viruses to replicate their genomes: a hypothesis for the origin of cellular domain. *Proc. Natl. Acad. Sci. U.S.A.* 103, 3669–3674. doi: 10.1073/pnas.0510333103

Forterre, P. (2006b). The origin of viruses and their possible roles in major evolutionary transitions. *Virus Res.* 117, 5–16. doi: 10.1016/j.virusres.2006.01.010

Forterre, P. (2011). Manipulation of cellular syntheses and the nature of viruses: the virocell concept. *Comptes Rendus Chimie* 14, 392–399. doi: 10.1016/j.crci.2010.06.007

Forterre, P. (2012a). *Virocell Concept, The. eLS.* Chichester, West Sussex: John Wiley & Sons, Ltd. doi: 10.1002/9780470015902.a0023264

Forterre, P. (2012b). Darwin's goldmine is still open: variation and selection run the world. *Front. Cell. Infect. Microbiol.* 2:106. doi: 10.3389/fcimb.2012.00106

Forterre, P. (2013). The common ancestor of Archaea and Eukarya was not an archaeon. *Archaea* 2013, 372396. doi: 10.1155/2013/372396

Forterre, P., and Prangishvili, D. (2009). The great billion-year war between ribosome- and capsid-encoding organisms (cells and viruses) as the major source of evolutionary novelties. *Ann. N. Y. Acad. Sci.* 1178, 65–77. doi: 10.1111/j.1749-6632.2009.04993.x

Forterre, P., and Prangishvili, D. (2013). The major role of viruses in cellular evolution: facts and hypotheses. *Curr. Opin. Virol.* 3, 558–565. doi: 10.1016/j.coviro.2013.06.013

Goulet, A., Blangy, S., Redder, P., Prangishvili, D., Felisberto-Rodrigues, C., Forterre, P., et al. (2009). Acidianus filamentous virus 1 coat proteins

display a helical fold spanning the filamentous archaeal viruses lineage. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21155–21160. doi: 10.1073/pnas.0909 893106

Hulo, C., de Castro, E., Masson, P., Bougueleret, L., Bairoch, A., Xenarios, I., et al. (2011). ViralZone: a knowledge resource to understand virus diversity. *Nucleic Acids Res.* 39, D576–D582. doi: 10.1093/nar/gkq901

Kings, A. M. Q., Adams, M. J., Carstens, E. B., and Lefkowitz, E. J. (2012). *Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses.* San Diego, CA: Elsevier Academic Press.

Koonin, E. V., and Dolja, V. V. (2013). A virocentric perspective on the evolution of life. *Curr. Opin. Virol.* 3, 546–557. doi: 10.1016/j.coviro.2013.06.008

Koonin, E. V., Senkevich, T. G., and Dolja, V. V. (2006). The ancient Virus World and evolution of cells. *Biol. Direct* 1, 29. doi: 10.1186/1745-6150-1-29

Koonin, E. V., and Wolf, Y. I. (2012). Evolution of microbes and viruses: a paradigm shift in evolutionary biology? *Front. Cell. Infect. Microbiol.* 2:119. doi: 10.3389/fcimb.2012.00119

Krupovic, M., Forterre, P., and Bamford, D. H. (2010). Comparative analysis of the mosaic genomes of tailed archaeal viruses and proviruses suggests common themes for virion architecture and assembly with tailed viruses of bacteria. *J. Mol. Biol.* 397, 144–160. doi: 10.1016/j.jmb.2010.01.037

La Scola, B., Audic, S., Robert, C., Jungang, L., de Lamballerie, X., Drancourt, M., et al. (2003). A giant virus in amoebae. *Science* 299, 2033. doi: 10.1126/science.1081867

Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., et al. (2014). Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4274–4279. doi: 10.1073/pnas.1320670111

Martin, M., and Müller, M. (1998). The hydrogen hypothesis for the first eukaryote. *Nature* 392, 37–41. doi: 10.1038/32096

Mochizuki, T., Krupovic, M., Pehau-Arnaudet, G., Sako, Y., Forterre, P., and Prangishvili, D. (2012). Archaeal virus with exceptional virion architecture and the largest single-stranded DNA genome. *Proc. Natl. Acad. Sci. U.S.A.* 109, 13386–13391. doi: 10.1073/pnas.1203668109

Mochizuki, T., Sako, Y., and Prangishvili, D. (2011). Provirus induction in hyperthermophilic archaea: characterization of *Aeropyrum pernix* spindle-shaped Virus 1 and *Aeropyrum pernix* Ovoid Virus 1. *J. Bacteriol.* 193, 5412–5419. doi: 10.1128/JB.05101-11

Mochizuki, T., Yoshida, T., Tanaka, R., Forterre, P., Sako, Y., and Prangishvili, D. (2010). Diversity of viruses of the hyperthermophilic archaeal genus Aeropyrum, and isolation of the *Aeropyrum pernix* bacilliform virus 1, APBV1, the first representative of the family Clavaviridae. *Virology* 402, 347–354. doi: 10.1016/j.virol.2010.03.046

Nasir, A., Kim, K. M., and Caetano-Anolles, G. (2012). Giant viruses coexisted with the cellular ancestors and represent a distinct supergroup along with superkingdoms Archaea, Bacteria and Eukarya. *BMC Evol. Biol.* 12:156. doi: 10.1186/1471-2148-12-156

Pell, L. G., Kanelis, V., Donaldson, L. W., Howell, P. L., and Davidson, A. R. (2009). The phage lambda major tail protein structure reveals a common evolution for long-tailed phages and the type VI bacterial secretion system. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4160–4165. doi: 10.1073/pnas.0900 044106

Philippe, N., Legendre, M., Doutre, G., Couté, Y., Poirot, O., Lescot, M., et al. (2013). Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science* 341, 281–286. doi: 10.1126/science.1239181

Pietilä, M. K., Demina, T. A., Atanasova, N. S., Oksanen, J. M., and Bamford, D. H. (2014). Archaeal viruses and bacteriophages: comparisons and contrasts. *Trends Microbiol.* doi: 10.1016/j.tim.2014.02.007. [Epub ahead of print].

Pietilä, M. K., Roine, E., Paulin, L., Kalkkinen, N., and Bamford, D. H. (2009). An ssDNA virus infecting archaea: a new lineage of viruses with a membrane envelope. *Mol. Microbiol.* 72, 307–319. doi: 10.1111/j.1365-2958.2009.06642.x

Pina, P., Bize, A., Forterre, P., and Prangishvili, D. (2011). The archeoviruses. *FEMS Microbiol. Rev.* 35, 1035–1054. doi: 10.1111/j.1574-6976.2011.00280.x

Prangishvili, D. (2013). The wonderful world of archaeal viruses. *Annu. Rev. Microbiol.* 67, 565–585. doi: 10.1146/annurev-micro-092412-155633

Prangishvili, D., Forterre, P., and Garrett, R. A. (2006). Viruses of the Archaea: a unifying view. *Nat. Rev. Microbiol.* 4, 837–848. doi: 10.1038/nrmicro1527

Raoult, D., and Forterre, P. (2008). Redefining viruses: lessons from mimivirus. *Nat. Rev. Microbiol.* 6, 315–319. doi: 10.1038/nrmicro1858

Rohwer, F., and Thurber, R. V. (2009). Viruses manipulate the marine environment. *Nature* 459, 207–212. doi: 10.1038/nature08060

Rosario, K., Dayaram, A., Marinov, M., Ware, J., Kraberger, S., Stainton, D., et al. (2012). Diverse circular ssDNA viruses discovered in dragonflies (*Odonata: Epiprocta*). *J. Gen. Virol.* 93, 2668–2681. doi: 10.1099/vir.0.045948-0

Schmid, M. F., Hecksel, C. W., Rochat, R. H., Bhella, D., Chiu, W., and Rixon, F. J. (2012). A tail-like assembly at the portal vertex in intact herpes simplex type-1 virions. *PLoS Pathog.* 8:e1002961. doi: 10.1371/journal.ppat.1002961

Sencilo, A., Paulin, L., Kellner, S., Helm, M., and Roine, E. (2012). Related haloarchaeal pleomorphic viruses contain different genome types. *Nucleic Acids Res.* 40, 5523–5534. doi: 10.1093/nar/gks215

Shackelton, L. A., Parrish, C. R., Truyen, U., and Holmes, E. C. (2005). High rate of viral evolution associated with the emergence of carnivore parvovirus. *Proc. Natl. Acad. Sci. U.S.A.* 102, 379–384. doi: 10.1073/pnas.0406765102

Silander, O. K., Weinreich, D. M., Wright, K. M., O'Keefe, K. J., Rang, C. U., Turner, P. E., et al. (2005). Widespread genetic exchange among terrestrial bacteriophages. *Proc. Natl. Acad. Sci. U.S.A.* 102, 19009–19014. doi: 10.1073/pnas.0503074102

Singer, M. F. (1995). Unusual reverse transcriptases. *J. Biol. Chem.* 270, 24623–24626. doi: 10.1074/jbc.270.42.24623

Suttle, C. A. (2007). Marine viruses – major players in the global ecosystem. *Nat. Rev. Microbiol.* 5, 801–812. doi: 10.1038/nrmicro1750

Woese, C. R., Kandler, O., and Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. U.S.A.* 87, 4576–4579. doi: 10.1073/pnas.87.12.4576

Zhao, Y., Temperton, B., Thrash, J. C., Schwalbach, M. S., Vergin, K. L., Landry, Z. C., et al. (2013). Abundant SAR11 viruses in the ocean. *Nature* 494, 357–360. doi: 10.1038/nature11921

frontiers in
MICROBIOLOGY

From lifetime to evolution: timescales of human gut microbiota adaptation

Sara Quercia[1], Marco Candela[1], Cristina Giuliani[2], Silvia Turroni[1], Donata Luiselli[2], Simone Rampelli[1], Patrizia Brigidi[1], Claudio Franceschi[3,4,5,6,7], Maria Giulia Bacalini[3,4], Paolo Garagnani[3,4,8] and Chiara Pirazzini[3,4] *

[1] Department of Pharmacy and Biotechnology, University of Bologna, Bologna, Italy
[2] BiGEA, Department of Biological, Geological and Environmental Sciences, Laboratory of Molecular Anthropology & Centre for Genome Biology, University of Bologna, Bologna, Italy
[3] DIMES, Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Bologna, Italy
[4] CIG, Interdepartmental Centre "L. Galvani" CIG, University of Bologna, Bologna, Italy
[5] IRCSS, Institute of Neurological Sciences of Bologna, Bologna, Italy
[6] IGM-CNR, Institute of Molecular Genetics, Unit of Bologna IOR, Bologna, Italy
[7] CNR, Institute of Organic Synthesis and Photoreactivity (ISOF), Bologna, Italy
[8] CRBA, Center for Applied Biomedical Research, St. Orsola-Malpighi University Hospital, Bologna, Italy

Human beings harbor gut microbial communities that are essential to preserve human health. Molded by the human genome, the gut microbiota (GM) is an adaptive component of the human superorganisms that allows host adaptation at different timescales, optimizing host physiology from daily life to lifespan scales and human evolutionary history. The GM continuously changes from birth up to the most extreme limits of human life, reconfiguring its metagenomic layout in response to daily variations in diet or specific host physiological and immunological needs at different ages. On the other hand, the microbiota plasticity was strategic to face changes in lifestyle and dietary habits along the course of the recent evolutionary history, that has driven the passage from Paleolithic hunter-gathering societies to Neolithic agricultural farmers to modern Westernized societies.

Keywords: gut microbiota, aging, environmental stimuli, co-evolution, biological adaptation

INTESTINAL MICROBIOTA, STRUCTURE, AND ROLE IN HUMAN PHYSIOLOGY

Human beings co-evolved as superorganisms as the result of the mutualistic relationship with the enormous microbial community that resides in human gastrointestinal tract (GIT); this ecosystem is better known as gut microbiota (GM; Turnbaugh et al., 2007). The GM reaches the highest cell concentration in the colon, with a density of 10^{12} CFU/g of intestinal content and represents the most densely populated and biodiverse ecosystem on earth (O'Hara and Shanahan, 2006; Ley et al., 2008a). The GM presents a very particular phylogenetic structure, resulting in a sparsely branched tree, with a high degree of radiation at the ends. Indeed, out of the 100 different bacterial phyla detected on our planet, only seven are found in our gut – Firmicutes, Bacteroidetes, Actinobacteria, Proteobacteria, Verrucomicrobia, Tenericutes, and Fusobacteria – of which Firmicutes and Bacteroidetes represent together up to 90% of the ecosystem (Costello et al., 2009). Conversely, the GM shows an impressive biodiversity at lower phylogenetic levels with up to 1000 different species being detected by next generation sequencing-based approaches (Qin et al., 2010). Interestingly, the species-level GM composition varies dramatically among people, and each subject owns a very unique subset of microorganisms, that consists of hundreds of the more than 1000 species detected in the GM of the entire human population. The total genome of these microorganisms, often referred to as the intestinal microbiome, has been estimated to contain 150 times more genes than the human one, providing the host with essential functional traits that human beings have not evolved on their own (Qin et al., 2010). For instance, the carbohydrate-active enzymes encoded in the microbial glycobiome allow the host to extract energy from otherwise

indigestible polysaccharides (Gill et al., 2006), complementing the poor human glycobiome diversity. Indeed, the functional assignment of the gut microbiome revealed an extraordinary diversity of gene clusters involved in carbohydrate metabolism (Koropatkin et al., 2012). Moreover, the gut microbiome is enriched in genes involved in the production of vitamins, cofactors, and secondary metabolites, further supporting its important role in host nutrition (Bäckhed et al., 2004; Clemente et al., 2012). The GM is also an active component of the human immune system (Round and Mazmanian, 2009; Maynard et al., 2012). The cross-talk with intestinal microbes has been shown to be essential for the maturation of a correct immune function in early life and to preserve a well-balanced immune homeostasis later in life (Kamada et al., 2013). Finally, a new and only sparsely explored role of the GM in human physiology is its potential to modulate the function of the gut-brain axis. Indeed, accumulating data from studies carried out in mouse models suggest how the GM and its metabolites could affect the behavior and pain, in addition to depression, anxiety and other disorders belonging to the central nervous system (Cryan and Dinan, 2012). Gut commensals are capable of activating neural pathways and modulating signaling to the central nervous system through their metabolite production (Foster and McVey Neufeld, 2013). In particular, studies carried out on mice demonstrated a strategic role for commensal bacteria in programming the hypothalamic–pituitary–adrenal (HPA) stress responsiveness at early developmental stages, when brain plasticity is still preserved (Sudo et al., 2004). Indeed, germ-free mice showed an enhanced HPA response to restraint stress, which was reversed by their re-conventionalization at early stages of development. According to the authors, both a cytokine-mediated humoral route and a neural mediated pathway could be involved in the microbiota modulation of the endocrine response in early life. Moreover, as observed by Diaz Heijtz et al. (2011) in mice, the GM can also affect the synaptogenesis during the perinatal period.

The recent use of germ-free and gnotobiotic mice has allowed the field to disentangle the complexity of the GM-host transgenomic metabolism, shedding light on the specific role of GM metabolites in host physiology (Nicholson et al., 2012). In particular, the endpoints of the GM polysaccharide fermentation, short-chain fatty acids (SCFAs) – most abundantly acetate, propionate, and butyrate – are key GM metabolites, with a multifactorial role in human health and homeostasis. These acids have been shown to play a pivotal role in host nutrition and energy homeostasis, controlling energy production, and storage as well as the appetite. Butyrate represents the main energetic substrate for the colonic epithelial cells (Russell et al., 2013) and both butyrate and propionate have been reported to activate intestinal gluconeogenesis (IGN) through two complementary mechanisms: butyrate triggers IGN gene expression in the gut via cAMP-dependent mechanisms, whereas propionate activates IGN by gene expression through gut-brain neural circuits and itself represents a substrate for IGN (De Vadder et al., 2014). According to the authors, this last propionate-dependent mechanism of IGN induction has been defined as strategic to provide the host with several metabolic benefits on body weight and glucose control. Besides the nutritional role, SCFAs have been reported to be involved in the maintenance of immune homeostasis (Arpaia et al., 2013). Through their production the GM controls the epithelium inflammation rate and drives the production and migration of specific immunological cells. Effectively, propionate governs the *de novo* peripheral regulatory T cell (Treg) generation and, together with acetate, drives their homing in the colon. Furthermore, propionate has been involved in the enhancement of hematopoiesis of dendritic cells with an impaired Th2 activation (Trompette et al., 2014). On the other hand, butyrate has the ability to regulate the production of pro-inflammatory cytokines (Segain et al., 2000), exerting a local immunomodulatory activity. Moreover, it is involved in the extrathymic Treg generation. In addition, the protective activity exerted by butyrate on the gut epithelium has been reported, as it stimulates the release of mucins (Petersson et al., 2011).

GENETICS OF HOST AND MICROBIOTA

Host genetics and the GM are linked together by an intense crosstalk and this interaction is dynamic throughout the course of our life. Several studies have been conducted to determine the impact of host genetics on the GM composition providing conflicting results. To address how the host genotype and the environment influence the GM composition, a study on the fecal microbiota of monozygotic and dizygotic twin pairs concordant for leanness or obesity, and their mothers was conducted (Turnbaugh et al., 2009). The authors found that the GM of monozygotic twin pairs had a degree of similarity that was comparable to that of dizygotic twin pairs, highlighting that the environment impacts the GM composition more than the genetics does. It was also reported that family members harbor a similar GM composition and share a "core microbiome" made of several microbial genes. However, a study conducted on related and unrelated children reported that the highest level of similarity was found in genetically identical twins (Stewart et al., 2005).

Several studies reported that single host genes, i.e., MEditerranean FeVer, *APOA1, NOD2*, and *FUT2* affect the GM by altering its composition or by reducing the degree of bacterial diversity (Khachatryan et al., 2008; Petnicki-Ocwieja et al., 2009; Zhang et al., 2010; Frank et al., 2011; Wacklin et al., 2014).

Murine models proved to be very useful to clarify the effect of genetics on the GM. One of the first studies focused on the interaction host genetics-GM was based on observations on the GM reconstruction process occurring after a course of antibiotics (Vaahtovuo et al., 2003). It was observed that the colonization of the GM depends on the genotype of the host, and differences in communities between mouse strains were observed, supporting the idea that the gut community is not established by chance but is influenced by the host genetic background. Kovacs et al. (2011) studied particular recombinant inbred mouse strains to assess the relative role of the host genotype in the GM composition and they reported that the mouse genetic background is a strong determinant in shaping the intestinal microbiota. To address how environmental factors and host genetic factors combine to shape the GM, Benson et al. (2010) explored the associations between host quantitative trait loci (QTL) and the GM composition in mice. Eighteen host's QTL showing a significant association with the relative abundance of specific microbial taxa were identified.

Even if litter and cohort effects accounted for some of the GM variation, according to the authors host genetics had a greater impact on the GM variability.

All these studies provide information that supports the idea that the host genetics and the GM interact with each other deeply, and we can speculate that changes in the GM composition could boost the different genetic make up of every individual.

MICROBIOTA ADAPTATION TO DAILY LIFE: MICROBIOME PLASTICITY IN RESPONSE TO DIFFERENT DIETS AND HOST PHYSIOLOGY

The human GM is a complex dynamic system with the potential for multistability. Indeed, Faith et al. (2013) found that on average 40% of the microbial strains harbored in an adult's intestine was variable in a 5-year sampling period. In a mutualistic context the GM makes sudden jumps from different steady states under the pressure of environmental and endogenous factors, such as diet, age, host genetics, and physiological state (Candela et al., 2012).

The most rapid observable response of the gut microbial community is the reaction to daily dietary changes. Through a study where high-fat low-fiber and low-fat high-fiber diets were compared, changes in the microbiome composition were detected within 24 h of controlled feeding, confirming that the human gut ecosystem plasticity can respond efficiently and rapidly to external variables (Walker et al., 2011). However, as reported by Wu et al. (2011), short- and long-term dietary interventions differently impact the GM composition. Effectively, some bacterial groups are more likely to be influenced only by short-term dietary intervention, while others, namely those referred to as human enterotypes (Arumugam et al., 2011), are affected only by long-term intervention.

The correlation between nutrients and the GM composition was investigated in a caloric restriction study realized in 18 lean subjects over a 4-day period. The outcomes showed that nutritional compounds, like proteins and fibers, affect the phylogenetic and functional structure of the gut microbial community (Muegge et al., 2011). The connection between the GM phylogenetic profile and the ingestion of a specific nutrient, namely fermentable carbohydrate, was also observed in a recent study conducted in 14 overweight men in a 3-week period of intervention (Walker et al., 2011).

It has been recently demonstrated that animal-based and plant-based diets deeply impact on the GM (David et al., 2014). Both diets were administered for 5 days to 10 young US adults and the microbial community composition, metabolic products, and gene expression were analyzed. Interestingly, dietary changes to the plant- or animal-based diet resulted in marked microbiota changes only 1 day after the diet modification. In particular, the plant-based diet was associated with the presence of fibrolytic SCFA producers as *Roseburia*, *Eubacterium rectale*, and *Faecalibacterium prausnitzii*, while the animal-based diet resulted in the increase of potentially putrefactive microorganisms, such as *Bacteroides* and the bile tolerant *Bilophila wadsworthia* and *Alistipes*. It was observed that the animal-based diet had a greater impact on the GM structural and functional layout than the only plant-based diet. Lower levels of metabolic products resulting from the

fermentation of carbohydrates and greater levels of the products resulting from the fermentation of amino acids were reported in individuals with the animal-based diet. Finally, the animal-based diet was associated with an increased expression of genes involved in the biosynthesis of vitamins and genes involved in the metabolism of products resulting from the consumption of meat (David et al., 2014). Interestingly, besides dietary substrates, the GM also relies on host-derived glycans secreted in the mucus as a nutrient source in the gut (Kashyap et al., 2013). Indeed, genetically dictated changes in host mucus glycan composition, such as the presence or absence of terminal fucose residues, have been shown to significantly impact the GM structure and function. This provides a global view where the diet and the host genotype interact to modulate the GM configuration.

The ability of the GM to re-program itself in response to different stimuli is necessary to adapt to the metabolic requirements of the host corresponding to different physiological states. For instance, pregnancy represents a period of deep physiological changes during which the GM composition adjusts according to the growth of the fetus and the lactation period. According to Koren et al. (2012), pregnancy is characterized by a greater inflammation tone, reduced insulin sensitivity, and body fat increase.

These traits are supported by a pregnancy-associated GM profile, whose main features are the expansion of Proteobacteria and Actinobacteria and a decrease in richness. Interestingly, such modifications in the GM composition persist for 1 month after birth and afterward the adult-like microbiota configuration is restored.

While the GM virtually varies in response to any changes in environmental and endogenous factors, the GM adaptation to extreme conditions – as an abnormal dietary sugar and fat intake or chronic inflammation – breaks the microbiota-host mutualistic homeostasis, lowering the ecosystem diversity, and overcoming the resilience of the microbiota-host symbiosis. The microbiota observed in obese people represents an appropriate example of unbalanced GM configuration driven by the Westernized diet and lifestyle (Ridaura et al., 2013). The functional annotation of the obese-type gut microbiome revealed a decreased functional diversity and an enrichment in genes involved in carbohydrate, lipid, and amino acid metabolism, showing an overall increased fermentative capacity with respect to the lean-type microbiome (Turnbaugh et al., 2009). Moreover, very recently, Schulz et al. (2014) demonstrated that a high-fat diet mediates shifts in the GM composition that promote intestinal carcinogenesis by compromising the Paneth-cell-mediated antimicrobial host defenses. On the other hand, inflammatory bowel disease (IBD) is a paradigmatic model to elucidate the self-sustained inflammatory loop that is established in the gut as a result of an inflammation-induced microbiota dysbiosis. In particular, inflammation forces GM to change toward a pro-inflammatory pathobionts-enriched profile, which consolidates the host inflammatory tone (Round and Mazmanian, 2009).

Therefore, the role of the GM as a plastic factor in response to environmental or endogenous stress is essential for the maintenance of the mutualistic relationship with the host. But under some specific circumstances the microbial community can be

forced to shift to a disease-associated configuration with the breaking of the homeostasis balance.

MICROBIOTA ADAPTATION TO DIFFERENT AGES

The human GM describes an evolutionary trajectory along the course of human life. The GM ecosystem changes its structural and functional layout from early infancy to old age, providing the host with ecosystem services finely calibrated for each stage of life (Yatsunenko et al., 2012). For instance, the peculiar GM composition during infancy exerts specific functions for the infant biology, supporting the immune system education, brain development, and host nutrition (Candela et al., 2013). At weaning, the GM gains diversity and develops new physiological functions, in order to fulfill the adult age-related requirements, such as the need to extract energy from the variable array of complex polysaccharides characterizing the adult diet.

The individual microbial layout begins to be formed immediately during delivery (Jost et al., 2012). We are born sterile and environmental microbes immediately colonize us (Palmer et al., 2007). The infant' GIT is firstly colonized, just a few hours after birth, by facultative anaerobic bacteria, i.e., enterobacteria, staphylococci, and streptococci. Over time, the decreased amount of available oxygen allows strictly anaerobic bacteria to settle in the intestine, modifying the intestinal environment (Vael and Desager, 2009). In particular, Jost et al. (2012) analyzed the bacterial composition in feces from seven healthy vaginally delivered, breast-fed neonates at different times after birth. They observed that, during the first days of life, anaerobes, i.e., *Bifidobacterium* and *Bacteroides*, outnumbered facultative anaerobes in all seven neonates, pointing out that anaerobes may become dominant early in life and that the switch from facultative to strict anaerobes may occur at a very early stage. The infant-type microbiota is thus characterized by the dominance of *Bifidobacterium* and the presence of *Staphylococcus*, *Streptococcus,* and *Enterobacteriaceae* as other major components. With a relatively low degree of diversity, the infant-type GM is capable of tremendous fluctuations over time, with an individual-specific temporal pattern of variation in species composition (Palmer et al., 2007; Jost et al., 2012). The delivery mode is one of the factors that most influence early infants' microbiota composition (Dominguez-Bello et al., 2010). Indeed, the authors observed that the vaginally delivered infants acquired bacterial communities resembling their own mother's vaginal microbiota, while the cesarean section infants transiently harbored bacterial communities similar to those found on the mothers' skin surface.

Largely dominated by *Bifidobacterium* and *Enterobacteriaceae* with an extraordinary rate of variation over time, the infant-type microbiota is functionally structured to educate the infant immune system through an intense, yet controlled, immunological dialog. Centanni et al. (2013) demonstrated that in infants the phylogenetic structure of the enterocyte-associated GM fraction was unaffected by the host inflammatory stimulus, probably because the GM of infants is specifically shaped to cope with the dynamic and intense cross-talk with the host immune system that is necessary for immune education. Recently, it has been shown that the diversity in the GM composition in infants is more important than the prevalence of specific bacterial taxa in the determination of

the risk of immunological diseases later in life, i.e., allergic disease and asthma (Bisgaard et al., 2011; Abrahamsson et al., 2013). Furthermore, the infant GM also responds to precise developmental and nutritional needs crucial for the infant, such as the development and functionality of the central nervous system (Sudo et al., 2004; Collins et al., 2012), as well as the specific vitamin requests (Yatsunenko et al., 2012). Recently, fascinating hypotheses extending the GM-dependent immune and metabolic programming to the perinatal period have been advanced (Rautava et al., 2012). However, until confirmed by robust experimental findings, such hypotheses need to be taken with caution (Hanage, 2014).

The infant-type GM is subject to profound fluctuations until weaning when, with the introduction of solid food, it shifts toward the adult-type microbiota, with the progressive acquisition of taxonomic and functional complexity, such as a wide array of carbohydrate-active enzymes. This shift results in a profound change in the GM composition that goes from a bifidobacteria-enriched community to another one dominated by *Firmicutes* and *Bacteroidetes*, resembling more and more the microbiota of an adult, characterized by increased functionality and stability (Koenig et al., 2011). This adult-type microbiota is functionally structured to metabolize the whole complexity of the plant polysaccharides contained in the adult diet and provides mutual benefits to the host (Vanhoutte et al., 2004). Indeed, the microbiota takes advantage of a warm and nutrient-rich environment in which it can settle, while the host can benefit from an easy-fitting metabolic equipment that can provide essential factors and increase the host's digestive capacity (Lozupone et al., 2012). A strong selection toward a readily changeable individual microbiome profile has been shown (Candela et al., 2012). This is the consequence of the inherent degree of plasticity of this bacterial ecosystem in adults, which allows the GM to change in response to environmental/endogenous factors, and the uniqueness of our physiology, lifestyle and history (Costello et al., 2012). These result in a peculiar temporal dynamics of the individual GM, always providing an adaptive response to ensure ecosystem services in the face of personalized physiology, immune system, environmental, or dietary exposure and lifestyle (Candela et al., 2013).

With aging and the onset of pathophysiological conditions (e.g., colon cancer, IBD, obesity, type 2 diabetes, and cardiovascular diseases) the GM-host mutualistic relationship progressively becomes compromised (Biagi et al., 2013). In elderly people diet and lifestyle undergo profound variations that include alterations of taste and smell, of gastrointestinal motility, and mastication, resulting in a nutritionally imbalanced diet (Biagi et al., 2010, 2013; Claesson et al., 2011; Drago et al., 2012). These age-related modifications, together with immunosenescence, affect the phylogenetic and functional structure of the gut ecosystem, leading to a microbial composition that favours the bloom of pathobionts (*Enterobacteriaceae*) to the detriment of immunomodulatory groups (*Clostridium* cluster IV and XIVa, *Bifidobacterium*). This age-associated configuration together with the "inflammaging" process could contribute to the creation of a self-sustained pro-inflammatory loop that is prejudicial for host health (Franceschi et al., 2000a,b; Grignolio et al., 2014). Interestingly, the GM of

the elderly displays a restricted stability and extreme variability. Recently, a functional description of the aged GM was reported (Rampelli et al., 2013). By using Illumina shotgun sequencing, three centenarians' fecal samples were analyzed and a shift from a saccharolytic to a putrefactive metabolism was reported. Indeed, an increase in the proteolytic potential, a reduction of genes involved in the metabolism of carbohydrates and a reduction of genes involved in SCFA production were observed. These modifications are in agreement with the age-related enrichment of genes belonging to pathobionts, and the authors hypothesized the existence of a pro-inflammatory loop in which pathobionts actively promote the worsening of health status with age. They also speculate that in centenarians some readjustments could occur to counteract the detrimental effects of pathobiont accumulation.

MICROBIOTA ADAPTATION DURING DIETARY SHIFT IN HUMAN EVOLUTION

Bacteria are part of the evolutionary history of complex organisms and they occupy every ecological niche of our planet. The human GM is the biggest stable symbiont of our body (Costello et al., 2009) and it is characterized by a long adaptive history.
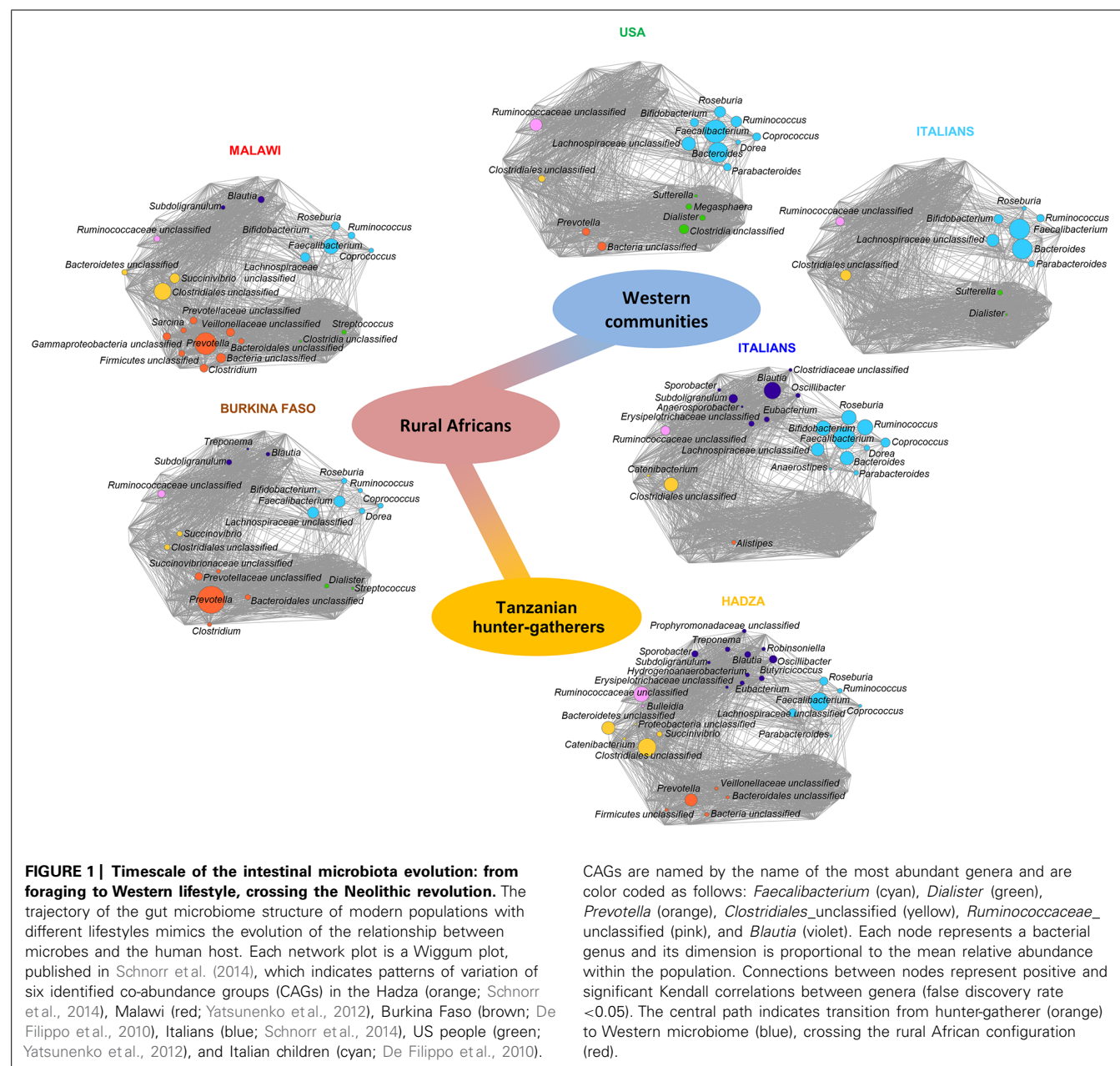
Modern humans, when have moved out of Africa, had to face different environmental challenges (such as food availability, climate changes, and pathogen loads). The main change in the host–microbiota symbiosis likely occurred almost 10,000 years ago, during the *Neolithic revolution*, also called "agricultural revolution" (De Filippo et al., 2010; Ottaviani et al., 2011). This revolution is based on the transition from hunting and gathering to agriculture and permanent settlements. In this period, the agriculture and animal husbandry have led to natural changes of human lifestyle and shaped modern human genomes. Given its high plasticity, the GM is able to change its composition and to adapt itself, according to diet/food availability, and the advent of agricultural societies could have favored microbial communities able to ferment complex substrates like polysaccharides (Hehemann et al., 2010). However, to date, little is known regarding how the GM has changed during human evolution. One of the most constraining aspects in this research field is the impossibility of having suitable fossil record. Indeed, the study of changes in the GM in human history is complicated by the difficulty in finding well-preserved samples of feces or intestinal samples of different periods (Walter and Ley, 2011). Nevertheless, researchers are developing methods to overcome this limitation. In a very recent paper, Sistiaga et al. (2014) applied gas-chromatography-mass spectrometry to Neanderthal's fecal matter to evaluate sterol and stanol level. The authors provide the first evidence that, even if Neanderthals predominantly consumed meat, they also had a remarkable plant intake, and they suggest the presence of a specific GM involved in cholesterol metabolism throughout human evolution. On the other hand, a glimpse of the ancestral human GM configuration could be provided by the GM of close primate relatives (Ley et al., 2008b; Moeller et al., 2012). Interestingly, the GM of modern humans clusters with that of other omnivorous primates, regardless of their affiliation to *Pan* (Ley et al., 2008b). This supports the key role of dietary habits in shaping the composition of the GIT microbial ecosystem.

The dynamics of the GM-host co-evolution and environmental adaptation can be addressed by investigating the GM variability in modern human populations of different culture (Candela et al., 2012). Indeed, the study of the GM from large healthy human populations of different age and socio-economic, geographic, and cultural settings allows researchers to point out the contribution of these environment components to the GM variation. In this context, a very recent paper explored the GM of the Hadza of Tanzania, a modern population of hunter-gatherers that still live as Paleolithic humans (Schnorr et al., 2014). This study elucidated the mechanisms of humans/GM co-evolution and showed a first map of the microbiota composition of the Hadza that reflects the functional adaptation to a foraging lifestyle. For instance, the high bacterial diversity and the enrichment in fibrolytic microorganisms (e.g., xylan-degrading *Prevotella* and *Treponema*) proper of the Hadza GM, represent ecosystem adaptations to provide SCFAs from their heavy plant-based diet. Furthermore, the Hadza show a sex-related divergence in the GM composition reflecting the sexual division of labor and sex differences in diet composition. In particular, the higher relative abundance of *Treponema* found in Hadza women could provide specific functions to deal with their higher intake of tubers and plant foods. In fact, women selectively forage for tubers and plant foods and spend a lot of time in camp, while men are highly mobile foragers and range far from the central camp site to obtain meat and honey. Even if foods are brought back to the camp and shared, men and women tend to consume more of their targeted foods. Finally, the absence of *Bifidobacterium* and a corresponding enrichment of potential opportunists as Proteobacteria and Spirochaetes in the Hadza GM probably correspond to a different tolerogenic layout of their immune system, redefining the notion of what we consider a healthy and an unhealthy GM structure. Indeed, the Hadza have relatively low rates of infectious diseases, metabolic diseases and nutritional deficiencies in comparison with other groups settled in Northern Tanzania (Bennett et al., 1973; Work et al., 1973; Blurton Jones et al., 1992). Moreover, De Filippo et al. (2010) compared the GM of children living in rural Africa and that of European children, and many differences emerged. Children from Boulpon Rural Village in Burkina Faso have a traditional rural African diet that is rich in starch, fibers, and plant polysaccharides and low in fat and animal protein, while European children follow a Western diet. The authors argued that the consumption of sugar, animal fat, and calorie-dense foods in industrialized countries is rapidly limiting the adaptive potential of the microbiota, by reducing microbiota richness and its functionality. Interestingly, the authors reported that only the GM of African children contains *Prevotella*, *Xylanibacter,* and *Treponema* that are involved in cellulose and xylan hydrolysis. It was speculated that the high fiber intake characterizing the African diet could change the GM composition to maximize the metabolic energy extraction from ingested plant polysaccharides (De Filippo et al., 2010). Finally, the same approach was applied to fecal samples from 531 children and adults from the Amazonas of Venezuela, rural Malawi and US metropolitan areas, including parents, siblings, and twins (Yatsunenko et al., 2012). The phylogenetic composition of the GM of these three populations is different, especially for US residents vs. non-US residents (Malawians and Amerindians).

Furthermore, the authors confirmed the importance of *Prevotella* as a discriminatory taxon that distinguishes non-US from US individuals. A meta-analysis of the GM composition in Western populations (i.e., USA and Italian citizens), rural Malawi and Burkina Faso populations, and Hadza hunter-gatherers has been carried out (Schnorr et al., 2014). Data allowed to reconstruct the putative trajectory of GM adaptive evolution that accompanied human beings along the transition from the Paleolithic hunter-gatherer to the Neolithic rural communities until modern Western societies. The diagram reported in **Figure 1** shows the emergence of specific co-abundance groups (CAGs) – groups of microorganisms which correlate and cluster together – along with the most important transition phases in our recent evolutionary history,

such as the higher abundance of *Ruminococcaceae* unclassified CAG distinguishing for the Hadza hunter-gatherers, the emergence of *Clostridiales* unclassified and *Prevotella* CAGs in rural Malawi and Burkina Faso populations, and the dominance of the *Faecalibacterium* CAG in Western populations (Schnorr et al., 2014).

A striking and fascinating example for gene acquisition by a gut microbe as an adaptation to the local diet is described by Hehemann et al. (2010) in a study on Japanese population. By comparing the GM from Japanese and North American populations, it was reported that the GM of Japanese subjects is enriched for genes (probably acquired by contact with marine microbes) that encode enzymes capable of degrading porphyran

FIGURE 1 | Timescale of the intestinal microbiota evolution: from foraging to Western lifestyle, crossing the Neolithic revolution. The trajectory of the gut microbiome structure of modern populations with different lifestyles mimics the evolution of the relationship between microbes and the human host. Each network plot is a Wiggum plot, published in Schnorr et al. (2014), which indicates patterns of variation of six identified co-abundance groups (CAGs) in the Hadza (orange; Schnorr et al., 2014), Malawi (red; Yatsunenko et al., 2012), Burkina Faso (brown; De Filippo et al., 2010), Italians (blue; Schnorr et al., 2014), US people (green; Yatsunenko et al., 2012), and Italian children (cyan; De Filippo et al., 2010).

CAGs are named by the name of the most abundant genera and are color coded as follows: *Faecalibacterium* (cyan), *Dialister* (green), *Prevotella* (orange), *Clostridiales*_unclassified (yellow), *Ruminococcaceae*_ unclassified (pink), and *Blautia* (violet). Each node represents a bacterial genus and its dimension is proportional to the mean relative abundance within the population. Connections between nodes represent positive and significant Kendall correlations between genera (false discovery rate <0.05). The central path indicates transition from hunter-gatherer (orange) to Western microbiome (blue), crossing the rural African configuration (red).

that is contained in seaweeds. In the Japanese culture, the greatest source of porphyran is due to nori (edible seaweed), which is commonly used in the preparation of sushi and it is the most produced and consumed seaweed for centuries in Japan.

CONCLUSION

Human beings faced tremendous changes in lifestyle and dietary habits along the course of their recent evolutionary history. They passed from Paleolithic hunter-gathering societies to Neolithic agricultural farmers until modern Westernized societies in 10,000 years, adapting to dramatic changes in diet and lifestyle in a relatively short evolutionary frame.

Human beings have been recently revised as superorganisms as a result of a close mutualistic relationship with their GM. Recent longitudinal studies highlighted an adaptive role for the GM in human biology, allowing to optimize the superorganism metabolic performances in response to diet, lifestyle, and physiological changes such as aging. This raises the question of whether this adaptive GM potential had a role in our recent evolutionary history, allowing adaptation to the profound lifestyle changes and describing a real GM-host evolutionary trajectory.

The first and very recent description of the Hadza GM structure provided important light in this direction. The peculiar structural and functional configuration of the Hadza gut microbial ecosystem suggests that adaptive functional changes of the GM accompanied the evolutionary trajectory of human beings, allowing the optimization of the superorganism performances in response to the profound changes that characterized our recent evolutionary history.

However, several circumstances characteristic of the Western world are challenging the resilience of the GM-host mutualistic interaction. The high-fat high-sugar energy-dense diet, sanitization, and antibiotic usage – a landmark of Western societies – are forcing GM adaptive changes to deviate from a mutualistic configuration. This raises the need to better comprehend the dynamics involved in this process, controlling any variables to preserve the extraordinary mutualistic relationship we evolved with our microbial counterpart.

ACKNOWLEDGMENTS

REFERENCES

Abrahamsson, T. R., Jakobsson, H. E., Andersson, A. F., Björkstén, B., Engstrand, L., and Jenmalm, M. C. (2013). Low gut microbiota diversity in early infancy precedes asthma at school age. *Clin. Exp. Allergy* 44, 842–850. doi: 10.1111/cea.12253

Arpaia, N., Campbell, C., Fan, X., Dikiy, S., van der Veeken, J., deRoos, P., et al. (2013). Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature* 504, 451–455. doi: 10.1038/nature12726

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., et al. (2011). Enterotypes of the human gut microbiome. *Nature* 473, 174–180. doi: 10.1038/nature09944

Bäckhed, F., Ding, H., Wang, T., Hooper, L. V., Koh, G. Y., Nagy, A. (2004). The gut microbiota as an environmental factor that regulates fat storage. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15718–15723. doi: 10.1073/pnas.0407076101

Bennett, F. J., Barnicot, N. A., Woodburn, J. C., Pereira, M. S., and Henderson, B. E. (1973). Studies on viral, bacterial, rickettsial and treponemal diseases in the hadza of Tanzania and a note on injuries. *Hum. Biol.* 45, 243–272.

Benson, A. K., Kelly, S. A., Legge, R., Ma, F., Low, S. J., Kim, J., 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. doi: 10.1073/pnas.1007028107

Biagi, E., Candela, M., Turroni, S., Garagnani, P., Franceschi, C., and Brigidi, P. (2013). Ageing and gut microbes: perspectives for health maintenance and longevity. *Pharmacol. Res.* 69, 11–20. doi: 10.1016/j.phrs.2012.10.005

Biagi, E., Nylund, L., Candela, M., Ostan, R., Bucci, L., Pini, E., et al. (2010). Through ageing, and beyond: gut microbiota and inflammatory status in seniors and centenarians. *PLoS ONE* 5:e10667. doi: 10.1371/journal.pone.0010667

Bisgaard, H., Li, N., Bonnelykke, K., Chawes, B. L., Skov, T., Paludan-Müller, G., et al. (2011). Reduced diversity of the intestinal microbiota during infancy is associated with increased risk of allergic disease at school age. *J. Allergy Clin. Immunol.* 128, 646–652.e1–5. doi: 10.1016/j.jaci.2011.04.060

Blurton Jones, N. G., Smith, L. C., O'Connell, J. F., Hawkes, K., and Kamuzora, C. L. (1992). Demography of the hadza, an increasing and high density population of savanna foragers. *Am. J. Phys. Anthropol.* 89, 159–181. doi: 10.1002/ajpa.1330890204

Candela, M., Biagi, E., Maccaferri, S., Turroni, S., and Brigidi, P. (2012). Intestinal microbiota is a plastic factor responding to environmental changes. *Trends Microbiol.* 20, 385–391. doi: 10.1016/j.tim.2012.05.003

Candela, M., Biagi, E., Turroni, S., Maccaferri, S., Figini, P., and Brigidi, P. (2013). Dynamic efficiency of the human intestinal microbiota. *Crit. Rev. Microbiol.* 1–7. doi: 10.3109/1040841X.2013.813900

Centanni, M., Turroni, S., Consolandi, C., Rampelli, S., Peano, C., Severgnini, M., et al. (2013). The enterocyte-associated intestinal microbiota of breast-fed infants and adults responds differently to a TNF-a-mediated pro-inflammatory stimulus. *PLoS ONE* 8:e81762. doi: 10.1371/journal.pone.0081762

Claesson, M. J., Cusack, S., O'Sullivan, O., Greene-Diniz, R., de Weerd, H., Flannery, E., et al. (2011). Composition, variability, and temporal stability of the intestinal microbiota of the elderly. *Proc. Natl. Acad. Sci. U.S.A.* 108(Suppl. 1), 4586–4591. doi: 10.1073/pnas.1000097107

Clemente, J. C., Ursell, L. K., Parfrey, L. W., and Knight, R. (2012). The impact of the gut microbiota on human health: an integrative view. *Cell* 148, 1258–1270. doi: 10.1016/j.cell.2012.01.035

Collins, S. M., Surette, M., and Bercik, P. (2012). The interplay between the intestinal microbiota and the brain. *Nat. Rev. Microbiol.* 10, 735–742. doi: 10.1038/nrmicro2876

Costello, E. K., Lauber, C. L., Hamady, M., Fierer, N., Gordon, J. I., and Knight, R. (2009). Bacterial community variation in human body habitats across space and time. *Science* 326, 1694–1697. doi: 10.1126/science.1177486

Costello, E. K., Stagaman, K., Dethlefsen, L., Bohannan, B. J., and Relman, D. A. (2012). The application of ecological theory toward an understanding of the human microbiome. *Science* 336, 1255–1262. doi: 10.1126/science.1224203

Cryan, J. F., and Dinan, T. G. (2012). Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nat. Rev. Neurosci.* 13, 701–712. doi: 10.1038/nrn3346

David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., et al. (2014). Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505, 559–563. doi: 10.1038/nature12820

De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J. B., Massart, S., et al. (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from europe and rural Africa. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14691–14696. doi: 10.1073/pnas.1005963107

De Vadder, F., Kovatcheva-Datchary, P., Goncalves, D., Vinera, J., Zitoun, C., Duchampt, A., et al. (2014). Microbiota-generated metabolites promote metabolic benefits via gut-brain neural circuits. *Cell* 156, 84–96. doi: 10.1016/j.cell.2013.12.016

Diaz Heijtz, R., Wang, S., Anuar, F., Qian, Y., Björkholm, B., Samuelsson, A., et al. (2011). Normal gut microbiota modulates brain development and behavior. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3047–3052. doi: 10.1073/pnas.1010529108

Dominguez-Bello, M. G., Costello, E. K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., et al. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11971–11975. doi: 10.1073/pnas.1002601107

Drago, L., Toscano, M., Rodighiero, V., De Vecchi, E., and Mogna, G. (2012). Cultivable and pyrosequenced fecal microflora in centenarians and young subjects. *J. Clin. Gastroenterol.* 46(Suppl. 1), S81–S84. doi: 10.1097/MCG.0b013e3182693982

Faith, J. J., Guruge, J. L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A. L., et al. (2013). The long-term stability of the human gut microbiota. *Science* 341:1237439. doi: 10.1126/science.1237439

Foster, J. A., and McVey Neufeld, K.-A. (2013). Gut-brain axis: how the microbiome influences anxiety and depression. *Trends Neurosci.* 36, 305–312. doi: 10.1016/j.tins.2013.01.005

Franceschi, C., Bonafè, M., Valensin, S., Olivieri, F., De Luca, M., Ottaviani, E., et al. (2000a). Inflamm-aging. an evolutionary perspective on immunosenescence. *Ann. N. Y. Acad. Sci.* 908, 244–54. doi: 10.1111/j.1749-6632.2000.tb06651.x

Franceschi, C., Valensin, S., Bonafè, M., Paolisso, G., Yashin, A. I., Monti, D., et al. (2000b). The network and the remodeling theories of aging: historical background and new perspectives. *Exp. Gerontol.* 35, 879–896. doi: 10.1016/S0531-5565(00)00172-8

Frank, D. N., Robertson, C. E., Hamm, C. M., Kpadeh, Z., Zhang, T., Chen, H., et al. (2011). Disease phenotype and genotype are associated with shifts in intestinal-associated microbiota in inflammatory bowel diseases. *Inflamm. Bowel Dis.* 17, 179–184. doi: 10.1002/ibd.21339

Gill, S. R., Pop, M., Deboy, R. T., Eckburg, P. B., Turnbaugh, P. J., Samuel, B. S., et al. (2006). Metagenomic analysis of the human distal gut microbiome. *Science* 312, 1355–1359. doi: 10.1126/science.1124234

Grignolio, A., Mishto, M., Faria, A. M., Garagnani, P., Franceschi, C., and Tieri, P. (2014). Towards a liquid self: how time, geography, and life experiences reshape the biological identity. *Front. Immunol.* 5:153. doi: 10.3389/fimmu.2014.00153

Hanage, W. P. (2014). Microbiology: microbiome science needs a healthy dose of scepticism. *Nature* 512, 247–248. doi: 10.1038/512247a

Hehemann, J.-H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., and Michel, G. (2010). Transfer of carbohydrate-active enzymes from marine bacteria to japanese gut microbiota. *Nature* 464, 908–912. doi: 10.1038/nature08937

Jost, T., Lacroix, C., Braegger, C. P., and Chassard, C. (2012). New insights in gut microbiota establishment in healthy breast fed neonates. *PLoS ONE* 7:e44595. doi: 10.1371/journal.pone.0044595

Kamada, N., Seo, S.-U., Chen, G. Y., and Núñez, G. (2013). Role of the gut microbiota in immunity and inflammatory disease. *Nat. Rev. Immunol.* 13, 321–335. doi: 10.1038/nri3430

Kashyap, P. C., Marcobal, A., Ursell, L. K., Smits, S. A., Sonnenburg, E. D., Costello, E. K., et al. (2013). Genetically dictated change in host mucus carbohydrate landscape exerts a diet-dependent effect on the gut microbiota. *Proc. Natl. Acad. Sci. U.S.A.* 110, 17059–17064. doi: 10.1073/pnas.1306070110

Khachatryan, Z. A., Ktsoyan, Z. A., Manukyan, G. P., Kelly, D., Ghazaryan, K. A., and Aminov, R. I. (2008). Predominant role of host genetics in controlling the composition of gut microbiota. *PLoS ONE* 3:e3064. doi: 10.1371/journal.pone.0003064

Koenig, J. E., Spor, A., Scalfone, N., Fricker, A. D., Stombaugh, J., Knight, R., et al. (2011). Succession of microbial consortia in the developing infant gut microbiome. *Proc. Natl. Acad. Sci.* 108(Suppl. 1), 4578–4585. doi: 10.1073/pnas.1000081107

Koren, O., Goodrich, J. K., Cullender, T. C., Spor, A., Laitinen, K., Bäckhed, H. K., et al. (2012). Host remodeling of the gut microbiome and metabolic changes during pregnancy. *Cell* 150, 470–480. doi: 10.1016/j.cell.2012.07.008

Koropatkin, N. M., Cameron, E. A., and Martens, E. C. (2012). How glycan metabolism shapes the human gut microbiota. *Nat. Rev. Microbiol.* 10, 323–335. doi: 10.1038/nrmicro2746

Kovacs, A., Ben-Jacob, N., Tayem, H., Halperin, E., Iraqi, F. A., and Gophna, U. (2011). Genotype is a stronger determinant than sex of the mouse gut microbiota. *Microb. Ecol.* 61, 423–428. doi: 10.1007/s00248-010-9787-2

Ley, R. E., Lozupone, C. A., Hamady, M., Knight, R., and Gordon, J. I. (2008a). Worlds within worlds: evolution of the vertebrate gut microbiota. *Nat. Rev. Microbiol.* 6, 776–788. doi: 10.1038/nrmicro1978

Ley, R. E., Hamady, M., Lozupone, C., Turnbaugh, P. J., Ramey, R. R., Bircher, J. S., et al. (2008b). Evolution of mammals and their gut microbes. *Science* 320, 1647–1651. doi: 10.1126/science.1155725

Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K., and Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature* 489, 220–230. doi: 10.1038/nature11550

Maynard, C. L., Elson, C. O., Hatton, R. D., and Weaver, C. T. (2012). Reciprocal interactions of the intestinal microbiota and immune system. *Nature* 489, 231–241. doi: 10.1038/nature11551

Moeller, A. H., Degnan, P. H., Pusey, A. E., Wilson, M. L., Hahn, B. H., and Ochman, H. (2012). Chimpanzees and humans harbour compositionally similar gut enterotypes. *Nat. Commun.* 3:1179. doi: 10.1038/ncomms2159

Muegge, B. D., Kuczynski, J., Knights, D., Clemente, J. C., González, A., Fontana, L., et al. (2011). Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* 332, 970–974. doi: 10.1126/science.1198719

Nicholson, J. K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., et al. (2012). Host-gut microbiota metabolic interactions. *Science* 336, 1262–1267. doi: 10.1126/science.1223813

O'Hara, A. M., and Shanahan, F. (2006). The gut flora as a forgotten organ. *EMBO Rep.* 7, 688–693. doi: 10.1038/sj.embor.7400731

Ottaviani, E., Ventura, N., Mandrioli, M., Candela, M., Franchini, A., and Franceschi, C. (2011). Gut microbiota as a candidate for lifespan extension: an ecological/evolutionary perspective targeted on living organisms as metaorganisms. *Biogerontology* 12, 599–609. doi: 10.1007/s10522-011-9352-5

Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A., and Brown, P. O. (2007). Development of the human infant intestinal microbiota. *PLoS Biol.* 5:e177. doi: 10.1371/journal.pbio.0050177

Petersson, J., Schreiber, O., Hansson, G. C., Gendler, S. J., Velcich, A., Lundberg, J. O., et al. (2011). Importance and regulation of the colonic mucus barrier in a mouse model of colitis. *Am. J. Physiol. Gastrointest. Liver Physiol.* 300, G327–G333. doi: 10.1152/ajpgi.00422.2010

Petnicki-Ocwieja, T., Hrncir, T., Liu, Y.-J., Biswas, A., Hudcovic, T., Tlaskalova-Hogenova, H., et al. (2009). Nod2 is required for the regulation of commensal microbiota in the intestine. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15813–15818. doi: 10.1073/pnas.0907722106

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65. doi: 10.1038/nature08821

Rampelli, S., Candela, M., Turroni, S., Biagi, E., Collino, S., Franceschi, C., et al. (2013). Functional metagenomic profiling of intestinal microbiome in extreme ageing. *Aging (Albany, NY)* 5, 902–912.

Rautava, S., Luoto, R., Salminen, S., and Isolauri, E. (2012). Microbial contact during pregnancy, intestinal colonization and human disease. *Nat. Rev. Gastroenterol. Hepatol.* 9, 565–576. doi: 10.1038/nrgastro.2012.144

Ridaura, V. K., Faith, J. J., Rey, F. E., Cheng, J., Duncan, A. E., Kau, A. L., et al. (2013). Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* 341:1241214. doi: 10.1126/science.1241214

Round, J. L., and Mazmanian, S. K. (2009). The gut microbiota shapes intestinal immune responses during health and disease. *Nat. Rev. Immunol.* 9, 313–323. doi: 10.1038/nri2515

Russell, W. R., Hoyles, L., Flint, H. J., and Dumas, M.-E. (2013). Colonic bacterial metabolites and human health. *Curr. Opin. Microbiol.* 16, 246–254. doi: 10.1016/j.mib.2013.07.002

Schnorr, S. L., Candela, M., Rampelli, S., Centanni, M., Consolandi, C., Basaglia, G., et al. (2014). Gut microbiome of the hadza hunter-gatherers. *Nat. Commun.* 5:3654. doi: 10.1038/ncomms4654

Schulz, M. D., Atay, C., Heringer, J., Romrig, F. K., Schwitalla, S., Aydin, B., et al. (2014). High-fat-diet-mediated dysbiosis promotes intestinal carcinogenesis independently of obesity. *Nature* 514, 508–512. doi: 10.1038/nature13398

Segain, J. P., Raingeard de la Blétière, D., Bourreille, A., Leray, V., Gervois, N., Rosales, C., et al. (2000). Butyrate inhibits inflammatory responses through NFkappaB inhibition: implications for Crohn's Disease. *Gut* 47, 397–403. doi: 10.1136/gut.47.3.397

Sistiaga, A., Mallol, C., Galván, B., and Summons, R. E. (2014). The eanderthal meal: a new perspective using faecal biomarkers. *PLoS ONE* 9:e101045. doi: 10.1371/journal.pone.0101045

Stewart, J. A., Chadwick, V. S., and Murray, A. (2005). Investigations into the influence of host genetics on the predominant eubacteria in the faecal microflora of children. *J. Med. Microbiol.* 54, 1239–1242. doi: 10.1099/jmm.0.46189-0

Sudo, N., Chida, Y., Aiba, Y., Sonoda, J., Oyama, N., Yu, X.-N., et al. (2004). Postnatal microbial colonization programs the hypothalamic-pituitary-adrenal system for stress response in mice. *J. Physiol.* 558, 263–275. doi: 10.1113/jphysiol.2004.063388

Trompette, A., Gollwitzer, E. S., Yadava, K., Sichelstiel, A. K., Sprenger, N., Ngom-Bru, C., et al. (2014). Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nat. Med.* 20, 159–166. doi: 10.1038/nm.3444

Turnbaugh, P. J., Hamady, M., Yatsunenko, T., Cantarel, B. L., Duncan, A., Ley, R. E., et al. (2009). A core gut microbiome in obese and lean twins. *Nature* 457, 480–484. doi: 10.1038/nature07540

Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., and Gordon, J. I. (2007). The human microbiome project. *Nature* 449, 804–810. doi: 10.1038/nature06244

Vaahtovuo, J., Toivanen, P., and Eerola, E. (2003). Bacterial composition of murine fecal microflora is indigenous and genetically guided. *FEMS Microbiol. Ecol.* 44, 131–136. doi: 10.1016/S0168-6496(02)00460-9

Vael, C., and Desager, K. (2009). The importance of the development of the intestinal microbiota in infancy. [Miscellaneous Article]. *Curr. Opin. Pediatrics* 21, 794–800. doi: 10.1097/MOP.0b013e328332351b

Vanhoutte, T., Huys, G., Brandt, E., and Swings, J. (2004). Temporal stability analysis of the microbiota in human feces by denaturing gradient gel electrophoresis using universal and group-specific 16S rRNA gene primers. *FEMS Microbiol. Ecol.* 48, 437–446. doi: 10.1016/j.femsec.2004.03.001

Wacklin, P., Tuimala, J., Nikkilä, J., Tims, S., Mäkivuokko, H., Alakulppi, N., et al. (2014). Faecal microbiota composition in adults is associated with the FUT2 gene determining the secretor status. *PLoS ONE* 9:e94863. doi: 10.1371/journal.pone.0094863

Walker, A. W., Ince, J., Duncan, S. H., Webster, L. M., Holtrop, G., Ze, X., et al. (2011). Dominant and diet-responsive groups of bacteria within the human colonic microbiota. *ISME J.* 5, 220–230. doi: 10.1038/ismej.2010.118

Walter, J., and Ley, R. (2011). The human gut microbiome: ecology and recent evolutionary changes. *Ann. Rev. Microbiol.* 65, 411–429. doi: 10.1146/annurev-micro-090110-102830

Work, T. H., Ifekwunigwe, A., Jelliffe, D. B., Jelliffe, P., and Neumann, C. G. (1973). Tropical problems in nutrition. *Ann. Intern. Med.* 79, 701–711. doi: 10.7326/0003-4819-79-5-701

Wu, G. D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.-Y., Keilbaugh, S. A., et al. (2011). Linking long-term dietary patterns with gut microbial enterotypes. *Science* 334, 105–108. doi: 10.1126/science.1208344

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., et al. (2012). Human gut microbiome viewed across age and geography. *Nature* 486, 222–227. doi: 10.1038/nature11053

Zhang, C., Zhang, M., Wang, S., Han, R., Cao, Y., Hua, W., et al. (2010). Interactions between gut microbiota, host genetics and diet relevant to development of metabolic syndromes in mice. *ISME J.* 4. 232–241. doi: 10.1038/ismej.2009.112

Diet and the development of the human intestinal microbiome

Noah Voreades, Anne Kozil and Tiffany L. Weir*

Department of Food Science and Human Nutrition, Colorado State University, Fort Collins, CO, USA

The important role of the gut microbiome in maintaining human health has necessitated a better understanding of the temporal dynamics of intestinal microbial communities as well as the host and environmental factors driving these dynamics. Genetics, mode of birth, infant feeding patterns, antibiotic usage, sanitary living conditions and long term dietary habits contribute to shaping the composition of the gut microbiome. This review focuses primarily on diet, as it is one of the most pivotal factors in the development of the human gut microbiome from infancy to the elderly. The infant gut microbiota is characterized by a high degree of instability, only reaching a state similar to that of adults by 2–3 years of age; consistent with the establishment of a varied solid food diet. The diet-related factors influencing the development of the infant gut microbiome include whether the child is breast or formula-fed as well as how and when solid foods are introduced. In contrast to the infant gut, the adult gut microbiome is resilient to large shifts in community structure. Several studies have shown that dietary changes induce transient fluctuations in the adult microbiome, sometimes in as little as 24 h; however, the microbial community rapidly returns to its stable state. Current knowledge of how long-term dietary habits shape the gut microbiome is limited by the lack of long-term feeding studies coupled with temporal gut microbiota characterization. However, long-term weight loss studies have been shown to alter the ratio of the Bacteroidetes and Firmicutes, the two major bacterial phyla residing in the human gastrointestinal tract. With aging, diet-related factors such as malnutrition are associated with microbiome shifts, although the cause and effect relationship between these factors has not been established. Increased pharmaceutical usage is also more prevalent in the elderly and can contribute to reduced gut microbiota stability and diversity. Foods containing prebiotic oligosaccharide components that nurture beneficial commensals in the gut community and probiotic supplements are being explored as interventions to manipulate the gut microbiome, potentially improving health status.

Keywords: enterotype, gut microbiome, aging, dietary patterns, colonization

IMPORTANCE OF THE GUT MICROBIOME

The consortium of single-celled organisms residing in our intestines, the gut microbiome, is rapidly emerging as an important determinant of health. Deterrents to proper bacterial colonization in early life are hypothesized to contribute to food sensitivities, allergic reactions, Type I diabetes, and other autoimmune disorders (Kelly et al., 2007). Association of the microbiome to autoimmune diseases has been explained by the "hygiene hypothesis," which suggests that the absence of a robust microbiome results in defects in development and regulation of the immune system, resulting in a lack of immune tolerance (Okada et al., 2010; Rook, 2012). Later in life, strong evidence supports an important role for intestinal microbiota in weight regulation via contributions to dietary energy harvest and appetite control (Tilg and Kaser, 2011). The gut microbiome has also been implicated in the pathology of several intestinal inflammatory diseases as well as in the development of colorectal, gastric, and prostate cancers and cardiometabolic disorders (Sekirov et al., 2010). Mechanisms giving rise to these conditions include the production of genotoxins by bacterial pathogens, microbial metabolism of dietary components to produce carcinogenic compounds, and inciting local and systemic inflammatory cascades that result in chronic low grade inflammation and damage to affected tissues and organs.

While a dysbiotic microbiota can cause disease, a healthy microbial community is vital to assist the host in maintaining optimal wellness. Thus, there is a need to understand the factors that shape and alter the microbiome throughout the lifespan of an individual. Numerous elements, encompassing environmental exposures, genetics, and other inherent host factors, contribute to the initial colonization of the microbiome in infants and to the subtle shifts that occur in adults, occasionally culminating in microbial decline as observed in frail and unhealthy elderly individuals (Koenig et al., 2011). However, none of these factors may be as important in the development of the microbiome as diet. In this review we will present evidence for the importance of diet in initial colonization events and in determining the composition of a stable adult microbiome. Factors such as malnutrition and pharmaceutical interventions on the aging gut will also be reviewed. Finally, we will discuss potential interventions, including dietary changes that can be used to alter the intestinal microbial community.

EARLY MICROBIAL COLONIZATION AND ESTABLISHMENT

The infant gut is thought to be sterile at birth, although some new research characterizing the placental microbiome challenges that assumption (Aagaard et al., 2014). After birth initial colonization and early establishment of the infant gut is influenced by whether delivery was vaginal or caesarean, feeding patterns, sanitary conditions, and antibiotic administration (Marques et al., 2010). The relative importance of these factors on the long-term structure of the intestinal microbial community and associated health outcomes is still debated. It stands to reason that with constant exposure between the microbiome and food components that diet is one of the primary drivers shaping the changes that occur during infancy and the structure of the adult microbiome that eventually establishes. This section will focus on the diet's role in shaping the infant gut microbiome from birth to ~3 years of age. Specifically, the following topics will be explored in detail: (1) the influence of breast vs. formula-feeding in initial colonization, (2) changes related to beginning of weaning and introduction of solid foods, and (3) factors contributing to a stable gut microbiome profile (**Figure 1**).

BREAST vs. FORMULA FEEDING

Following birth, the infant gut microbiome is characterized by low-species diversity and high rates of bacterial flux until ~2 or 3 years old (Bergström et al., 2014). Facultative anaerobic bacteria including *Staphylococcus, Streptococcus, Escherichia coli* and *Enterobacteria* are thought to be the first colonizers of the gut. Their purpose is to consume oxygen and create an environment for obligate anaerobes to thrive (Palmer et al., 2007; Jost et al., 2012). These are later replaced by facultative anaerobes that dominate the gastrointestinal tract, primarily Actinobacteria and Firmicutes (Turroni et al., 2012). This change in dominant taxa representation can be attributed to the introduction of breast or formula-feeding, signifying the first diet-related colonization event in the infant gut microbiome (Harmsen et al., 2000; Jost et al., 2012). In breast-fed infants, the dominant Actinobacteria are represented by *Bifidobacterium* species, specifically, *B. breve, B. longum, B. dentium, B. infantis*, and *B. pseudocatenulatum* (Harmsen et al., 2000; Jost et al., 2012). The Firmicutes phylum

is represented principally by lactic acid bacteria such as *Lactobacillus* and *Enterococcus* as well as *Clostridium* species (Turroni et al., 2012; Bergström et al., 2014). More than 700 species of bacteria have now been identified in human colostrum and breast milk, including multiple species of lactic acid bacteria as well as species typically colonizing the oral cavity of infants (Cabrera-Rubio et al., 2012). While this may contribute to the intestinal community of breastfed infants, it is still unclear whether the composition of species in breast milk is driven by transfer from infant to mother. The chemical composition of breast milk does influence the gut microbiome through supplying unique oligosaccharides that are selectively utilized by *Bifidobacterium spp.* (Turroni et al., 2012).

There are conflicting reports regarding differences in the relative abundance of these bacteria between breast and formula fed infants. Many studies have reported that formula-fed infants display dominance of *Bifidobacterium spp.* similar to what has been observed in breastfed infants (Harmsen et al., 2000; Fallani et al., 2010, 2011). However, another study reported approximately double the count of *Bifidobacterium* in breast fed infants compared to those fed formula (Bezirtzoglou et al., 2011). Formula feeding was also associated with higher levels of *Atopobium* (Bezirtzoglou et al., 2011); which corroborated reports by Fallani et al. (2010), although they only noted *Atopobium* increases in formula fed infants delivered by Cesearean section or whose mother's had been administered antibiotics. Higher numbers of *Bacteroides spp.* as well as members of the Enterobacteriaceae have also been reported in formula-fed infants (Harmsen et al., 2000; Fallani et al., 2010). Despite significant evidence that *Bifidobacterium* is an important early colonizer in neonates, Palmer et al. (2007) reported that *Bifidobacterium* was not present in significant amounts in the infant gut (Palmer et al., 2007). However, it is important to highlight that within their cohort, there was a mixture of breast and formula-feeding, antibiotics were provided to infants and a small subset required specialized hospitalization.

The variability reported with regard to *Bifidobacterium* abundance could be driven by differences in infant formula composition. Formulas supplemented with the prebiotics

0-9 Months (Newborn)		9-18 Months (Infant-Pre-Toddler)	18-36 Months (Toddler)
Breast-Fed Characteristics (BF) • Low Species Diversity • Bacterial Composition Flux • Major Phyla: *Actinobacteria & Firmicutes*	**Formula-Fed Characteristics (FF)** • Low Species Diversity • Bacterial Composition Flux • Major Phyla: *Actinobacteria & Bacteriodetes*	**Introduction of Weaning & Solid Food** • Increased Species Diversity • Bacterial Composition Flux Persists • Increasing Butyrate Producing Bacteria • Major Phyla: *Bacteriodetes & Firmicutes*	**Diet-Influenced Microbiome Profile** • Stable Gut Microbiome Formation • Increased Species Diversity • Breast-Feeding History Ceases To Impact Gut Microbiome Profile • Increasing Butyrate Producing Bacteria Abundance • Dietary Intake Strongly Influences Abundances (*Prevotella* vs *Firmicutes*) • Major Phyla: *Bacteriodetes & Firmicutes*

FIGURE 1 | Representation of the infant gut microbiome development from birth to 3 years of age. By 3 years old, toddler's microbiomes are similar to that in adults and long-term dietary patterns are beginning to establish.

galacto-oligosaccharide (GOS) and fructo-oligosaccharide (FOS) may account for high levels of *Bifidobacterium* found in many formula-fed infants (Marques et al., 2010; Oozeer et al., 2013). A recent review discusses evidence supporting GOS and FOS supplementation effects on the gut (Oozeer et al., 2013). Infant gut microbial populations provided with either human breast milk or prebiotic supplemented infant formula had similar levels of *Bifidobacterium;* whereas gut microbial populations of infants given traditional formula was reported to have about 20% fewer *Bifidobacterium* (Knol et al., 2005). Additionally, the species composition of *Bifidobacterium* was similar between infants given human breast milk and those on prebiotic supplemented formula. However, traditional formula fed infants had markedly different gut microbial communities and even the specific *Bifidobacterium* species differed with higher relative abundances of *B. cantenulatum* and *B. adolescentis,* which are typically represented in adult populations. Another potential explanation for the variation in studies reporting bacterial abundances, particularly with regard to breast-feeding could be due to differences in the maternal-diet (Cabrera-Rubio et al., 2012). Characterization of the placental microbiome suggests that it is colonized by the mother's oral microbiome (Aagaard et al., 2014). Another recent study showing that pre and post-natal maternal consumption of a high fat diet, independent of obesity in the mother, resulted in dysbiosis of the infant gut in a primate model (Ma et al., 2014). Together, these studies suggest that maternal diet may play a significant but previously unrecognized role in determining early colonization and establishment of the infant microbiome. Conduct of randomized trials in which the maternal diet is controlled or large-scale cross-sectional studies of pregnant mothers adhering to different diets (Western, vegetarian, gluten-free, etc) are necessary to further develop this hypothesis.

WEANING AND THE SHIFT TOWARD AN ADULT MICROBIOME

Around the age of 1–2 years old, the infant gut microbiome undergoes its second shift and the stable adult microbiome begins to emerge, further supporting the significant role of the diet in influencing the microbial community (De Filippo et al., 2010; Bergström et al., 2014). One study reported that although there were differences in the microbiome pre- and post-weaning, the impacts of earlier colonization events (delivery mode, formula or breastfed, etc.) were still apparent (Fallani et al., 2011). Another study comparing Italian vs. African children's gut microbiomes showed that after weaning and solid foods were introduced there was a significant diet-related shift in the gut microbiome profiles. Prior to the introduction of their respective Western or African diets, the children across both populations that were still breast-feeding clustered together and had similar *Bifidobacterium* species dominance. Only children who were already weaned reliably clustered together into distinct geographic groupings. This study reinforced two important points related to dietary drivers of the gut microbiome development in children. First, breast-feeding, regardless of duration supports a specific bacterial state that is unique and markedly different from that observed in individuals consuming solid foods. Second, once solid foods are

introduced, its role in shaping long-term gut microbiome profiles is so strong that individual's cluster based on diet type over other environmental and physiological factors (De Filippo et al., 2010).

A similarly significant shift was reported by Bergström et al. (2014) in a 3 year Danish study with a cohort of 330 infants. They reported that between 9 and 18 months, the infant gut bacterial abundances changed drastically with the introduction of solid foods. Specifically, *Bacteroidetes*-related species increased. Whereas *Bifidobacterium* and *Lactobacillus* species and Enterobacteriaceae declined, various species within Firmicutes phylum were also reported to increase. This bacterial taxa shift is logical given that breast and/or formula-feeding has ceased, depleting the primary fuel source for these bacteria. In addition, butyrate producing bacteria such as *Clostridium leptum* group, *E. halli*, and *Roseburia* species increased. Typically, butyrate producing bacteria are responsible for the breakdown of otherwise indigestible complex plant polysaccharides and resistant starches. Anecdotally, this study found that the longer infants were breast and/or formula-fed, the lower their levels of butyrate producing bacteria. Additionally, more and different species begin to appear with introduction of solid foods (Koenig et al., 2011; Bergström et al., 2014).

EMERGENCE OF A STABLE GUT PROFILE

From 18 to 36 months, the infant gut microbiome undergoes its final significant shift to a more stable microbial profile composed primarily of the bacterial phyla Bacteriodetes and Firmicutes. This shift represents a temporal change that can be attributed to the continued influence of a varied solid food diet (De Filippo et al., 2010; Koenig et al., 2011; Bergström et al., 2014). The earlier that solid food is introduced into the diet, the more quickly the gut microbiome begins to resemble a stable adult-like microbiome (Bergström et al., 2014). The specific proportion of Firmicutes and Bacteroidetes is strongly influenced by diet. This was best demonstrated in the previously discussed work by De Filippo et al. (2010) where the distinct microbial signatures of the two groups of children were indicative of their respective dietary habits. The most compelling evidence for this was the dominance of *Prevotella*, capable of digesting complex plant polysaccharides, in African children and its absence in Italian children. Similar diet-driven influences were reported in a detailed temporal study of a single infant. This study demonstrated that introduction of peas, formula, and other solid foods led to an emerging co-dominance between Firmicutes and Bacteroidetes, with the increase in Bacteroidetes potentially resulting from requirements for the breakdown of newly introduced plant polysaccharides (Koenig et al., 2011). The previously mentioned emergence of a stable gut microbiome can be substantially derailed if the infant experiences either severe acute malnutrition or moderate acute malnutrition. Emerging research is demonstrating that either of these malnutrition states has the potential to significantly alter the development of a healthy gut microbiome profile, regardless of diet-based interventions (Subramanian et al., 2014). These recent findings not only support a link between diet and the development of a particular gut microbiota and microbiome, but illustrate that nutrient quantity can impact development too.

THE ADULT MICROBIOME

The typical adult intestinal microbiome is primarily comprised of approximately six or seven different bacterial phyla, of which Bacteroidetes and Firmicutes dominate (Eckburg et al., 2005). Less abundant phyla can include Proteobacteria, Verrucomicrobiota, Actinobacteria, and *Euryarchaeota*. A recent study followed changes in the microbiome of 37 adults for up to 5 years and reported that ∼60–70% of the bacterial strains present remained unchanged over the course of the study and that the most stable members of the microbiome tended to be the most abundant (Faith et al., 2013). They also observed that at the phyla level, Bacteroidetes and Actinobacteria populations were less susceptible to perturbations whereas Firmicutes and Proteobacteria were significantly less stable. These results are fairly consistent with findings from an earlier study utilizing a microarray-based approach to determine molecular taxonomy and which followed a smaller cohort over a longer period of time (Rajilic-Stojanovic et al., 2013). Both studies reported that the taxa present in an individual remain fairly consistent over time, although the relative abundances of these taxa were subject to change. However, data from Rajilic-Stojanovic et al. (2013) suggests that larger fluctuations occur between samples taken at longer intervals while Faith et al. (2013) report the opposite trend, with larger fluctuations occurring in samples taken over shorter periods of time compared to those that are temporally farther apart. Despite this resilience, there is evidence that the diet shapes the relative abundance of dominant phyla and populations of specific bacterial groups are influenced by the composition of macronutrients consumed.
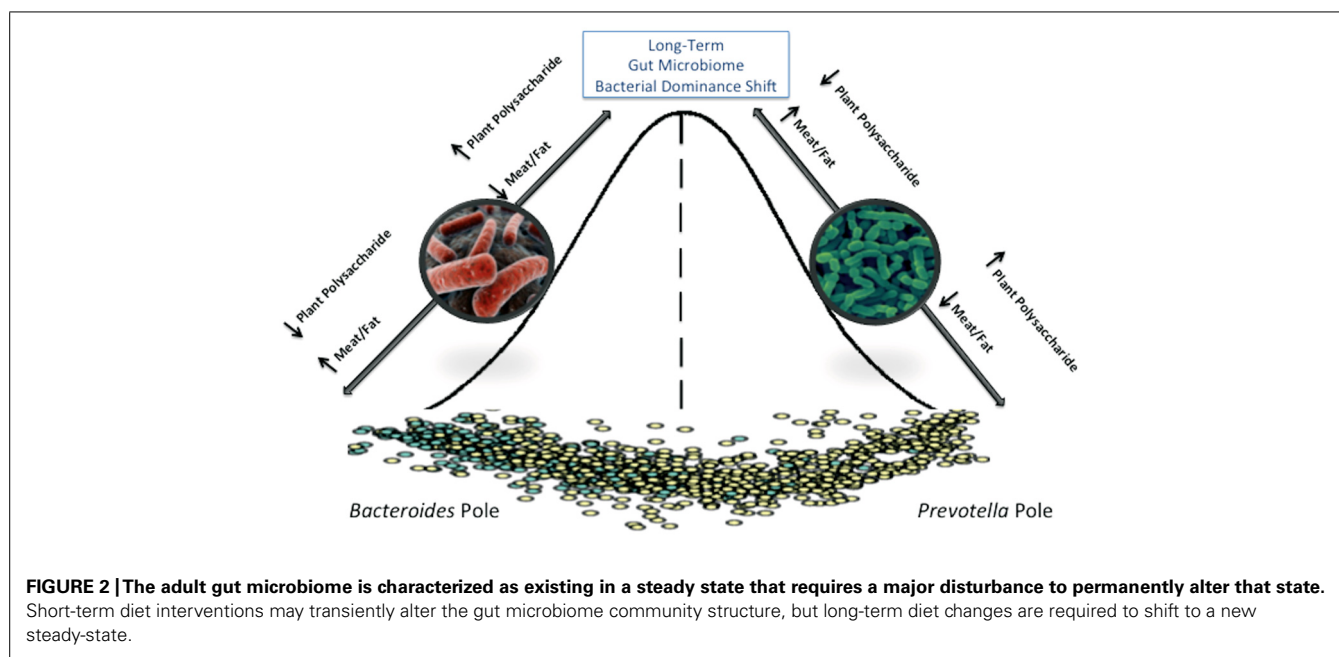
DIET-DRIVEN ENTEROTYPES

There have been numerous attempts to identify a "core" microbiota, usually defined as bacterial taxa that are shared between 95% of individuals tested (Huse et al., 2012). Identification of a core microbiome is important for defining a "normal" healthy state from which major variations may indicate a dysbiotic system that can result from or contribute to disease development. One barrier to defining an intestinal core microbiome has been the vast degree of variation between individuals. The microbial communities identified in samples collected from an individual over time are more similar to each other than microbial communities between two individuals, although related persons share more bacterial strains than unrelated individuals (Palmer et al., 2007; Yatsunenko et al., 2012; Faith et al., 2013). Although a consensus for what constitutes a core gut microbiome has been elusive, one report suggested that an international cohort of 39 individuals could be assigned to one of three distinct clusters or "enterotypes" based on metagenomic sequences (Arumugam et al., 2011). They found that each cluster was dominated by a particular bacterial genus (*Bacteroides*, *Prevotella*, and *Ruminococcus*) with positive or negative associations with a number of other genera in the community. They also reported that each cluster was enriched for specific gene functions that reflected different microbial trophic chains. Two of the three original enterotypes, *Bacteroides*, and *Prevotella*, were later confirmed and long term dietary patterns were identified as the primary predictor of an individual's enterotype (Wu et al., 2011). The *Bacteroides* enterotype was associated with a Western-type diet high in proteins and fat, while the *Prevotella*

enterotype was associated with plant fiber consumption. These enterotypes appear to be extremely stable, and several studies utilizing short-term interventions failed to result in a change in the assigned enterotype of participants (David et al., 2014; Roager et al., 2014).

The existence of enterotypes provided a convenient way of classifying individuals based on their fecal microbiota (although some argue a more appropriate term would be "faecotype") and speculation has begun as to whether enterotypes can be used as a predictor of long term health risks. However, a microbial survey of several body sites, including stool, conducted with more than 200 individuals showed only minimal segregation into the *Bacteroides* and *Prevotella* enterotypes rather than the distinct and well separated clusters previously reported (Huse et al., 2012). These discrepancies could be due to the fact that the method for assigning enterotypes is not consistent across studies. An analysis of archived 16S sequences also showed that enterotype determination is sensitive to clustering methods and distance metrics used and that there is a continuum of *Bacteroides* abundances across samples rather than a bimodal distribution (Koren et al., 2013). These studies suggest that the enterotype concept is not be as clear cut as previously believed, and that standard methods for defining enterotypes should be developed and employed before they can be meaningfully tied with clinical outcomes.

LONG TERM DIETARY PATTERNS AND THE MICROBIOME

Whether enterotypes truly exist or not, it is clear that diet is an important factor in shaping the microbiome (**Figure 2**). In addition to the divergence in microbial composition of Italian children and those from Burkina Faso shortly after weaning (De Filippo et al., 2010); other studies have shown microbiota segregation of individuals from Malawi, Venezuela, and the United States (Yatsunenko et al., 2012); children from Bangladesh and the United States (Lin et al., 2013), and between rural Africans and African Americans (Ou et al., 2013) that are at least partially diet-driven. In the Yatsunenko et al. (2012) study, metagenomic sequences revealed that enzyme classifications associated with protein degradation and bile salt metabolism were enriched in samples from the U.S. population where protein and fat consumption is high. Conversely, glutamate synthase and starch degrading enzymes were more abundant in the Amerindian and Malawian samples; consistent with protein poor diets of corn and cassava. This has been further demonstrated in a recent study of the diversity and metabolism of the microbiome of a Tanzanian hunter gatherer tribe, the Hadza. This study identified differences in the microbiome between the sexes which were consistent with their division of labor with regard to foraging (Schnorr et al., 2014). They also have many bacterial species associated with fermentation of plant-based fibers and are completely deficient in *Bifidobacterium*, which was hypothesized to result from the lack of meat and dairy in the diet; substrates that allow these bacteria to continue to colonize Westerners into adulthood. Although comparative studies between populations with different diets has been useful in identifying how dietary patterns shape the microbiome, these studies have utilized international cohorts that introduce confounding factors such as extreme differences in culture and

FIGURE 2 | The adult gut microbiome is characterized as existing in a steady state that requires a major disturbance to permanently alter that state. Short-term diet interventions may transiently alter the gut microbiome community structure, but long-term diet changes are required to shift to a new steady-state.

environment. Relatively few studies have been conducted that examine the effects of diet on homogenous populations. One study looked at correlations between specific dietary components and microbial function and structure in the intestines of a human cohort known for keeping meticulous diet logs (Muegge et al., 2011). They found that there were significant correlations between microbial gene function (Kegg orthologs) and protein intake, confirming the difference that was seen across multiple *mammalian* species between carnivores and herbivores. They also reported a correlation between insoluble fiber consumption and bacterial community membership. A large-scale microbiome sequencing effort called the American Gut Project is currently underway and is attempting to address the effects of diet on the adult microbiome capturing extremes within the American diet (i.e., vegan, paleo, etc) where cultural and environmental factors will be minimized.

DIETARY INTERVENTIONS INTRODUCE TRANSIENT AND SUBTLE CHANGES IN THE MICROBIOME

Short-term dietary interventions that include introducing novel food components or altering macronutrient levels have also been examined for their effects on intestinal microbial populations. The first of these studies followed obese individuals partitioned to restricted calorie diet groups that controlled for either fat or carbohydrate intake (Ley et al., 2006). Regardless of the macronutrient composition of the diet, individuals that lost a significant amount of body weight had a change in their ratio of Bacteroidetes to Firmicutes, driven by increases in the Bacteroidetes. Weight-loss driven changes in the microbiome was recently confirmed in individuals consuming a calorie restricted liquid diet where it was demonstrated that weight stability of an individual was a better predictor of fecal microbiome stability than time between sample collections (Faith et al., 2013). However, this and another study (Duncan et al., 2008) noted changes in members of the

Firmicutes rather than an increase in Bacteroidetes when corresponding weight loss occurred. Calorie restriction in obese and overweight individuals has also been shown to increase microbial gene richness, a parameter that was correlated to improved metabolic parameters (Cotillard et al., 2013; Le Chatelier et al., 2013).

Several studies have noted rapid but transient changes in fecal microbial composition immediately following the start of a dietary intervention study. Wu et al. (2011) conducted a controlled feeding experiment in ten individuals randomized to high fat/low fiber or high fiber/low fat diets and found that although there was no increase in community similarity between individuals on the same diet over a period of 10 days, the first 24 h period was considered an outlier because transient dramatic shifts occurred in the fecal communities of all individuals. Similarly, switching between animal and plant-based diets produces similar results (David et al., 2014). Another interesting finding of the David et al. (2014) study was that foodborne microbes transiently colonized the gut, introducing the idea that food may not only select for commensal bacterial species, but serve as a reservoir for new microbial introductions. Intentional introduction of food-borne microorganisms (probiotics) as well as prebiotic food ingredients and foods high in fiber can also be a means of subtly changing the relative abundance of bacterial species in the gut (Preidis and Versalovic, 2009). Thus, despite the inherent stability of the microbiome over time, changes related to weight loss and diet composition continue to subtly alter the composition and relative abundance of our commensal organisms, driving the development of our gut microbiome throughout adulthood.

THE AGING GUT

As a person ages, the stability and diversity of their gut microbiota declines with the state of their health. If health remains intact however, microbiota composition often retains the stability

and compositional make-up of a healthy younger adult (Claesson et al., 2012). The most prevalent age-related factors influencing the microbial population of the gut are: (1) physiological changes, (2) dietary choices and malnutrition, (3) living situation (community-dwelling, hospitalized, or long-term care), and (4) use of antibiotics (Bartosch et al., 2004; Woodmansey, 2007; Claesson et al., 2012) and other prescription drugs (Qato et al., 2008). This section will explore dietary alterations and antibiotic usage as drivers of change in the elderly gut microbiome and discuss the use of probiotics and prebiotics as potential solutions for the restoration of a healthy gut.

Diet is a major influence on the bacterial makeup of the aging gut. Physiological changes, such as loss of taste and smell, difficulty chewing or swallowing, impaired digestive function, and lack of physical mobility can leave elderly individuals consuming a narrow and nutritionally imbalanced diet, setting the stage for malnutrition (Bartosch et al., 2004; Claesson et al., 2012). Relocation from an in-home community setting to a long-term care facility can change dietary intake as well. The move often contributes to a greater consumption of fat and a decreased intake of fiber, fruits, vegetables, and meat. These dietary alterations are associated with a decrease in microbial diversity and increased frailty (Claesson et al., 2012).

The use of antibiotics in elderly populations is especially prevalent in hospital and long-term care facilities. Antibiotics create an environment of instability by diminishing the population of total and commensal bacteria and opening the door for pathogenic bacteria to overpopulate (Claesson et al., 2011). The use of broad-spectrum antibiotics is associated with the overgrowth of *Clostridium difficile* which flourishes in the antibiotic-weakened gut, often resulting in a life threatening infection (Macfarlane, 2014). As health issues compound and antibiotic use increases, elderly often see a decline in commensal anaerobes (*Bacteroides, Lactobacillus* and *Bifidobacterium)* accompanied by a rise in proteolytic and pathogenic bacteria (*Fusobacteria, Propionibacteria, Clostridia,* and *E. coli*; Wu et al., 2011). Studies indicate that probiotics may have potential as a therapeutic tool to replenish and recolonize beneficial bacterial species like *Bifidobacterium* and *Lactobacillus,* bringing the elderly gut back into balance (Likotrafiti et al., 2014).

EFFECTS OF DIET AND MALNUTRITION ON THE ELDERLY MICROBIOME
A number of proposed factors contribute to alterations in the elderly gut ecosystem and diet is a significant driver of change (Claesson et al., 2012). Dietary intake can change for a number of reasons with advanced age. Decline in physical mobility may limit access to the grocery store or inhibit the ability to cook. Some elderly lose the desire to eat due to loss of smell and taste or due to slow digestion and prolonged satiety (Britton and McLaughlin, 2013). Malnutrition is often an unintended consequence of age-related physiological changes that can lead to changes in the elderly gut microbiome. Furthermore, studies have shown that compositional dietary changes can result in almost immediate alterations in microbial populations. Wu et al. (2011) found that changes in microbiome composition were detectable within 24 h of dietary alteration and occurred even faster than transit time of food through the gut. In an infant population, malnutrition

was shown to delay the maturation of the intestinal microbiota (Subramanian et al., 2014), and it is likely to have consequences of a similar magnitude in the elderly gut.

Dietary changes that come with age are also impacted by living situation. Claesson et al. (2012) found distinct dietary differences between elderly individuals living in a traditional community setting compared to those in long-term care facilities. Community-dwellers may be healthier than their institutionalized counterparts for a number of reasons, but they broadly stated that community-dwellers eat a healthier and more diverse diet and have a distinct microbiota from those in long-term care facilities (Claesson et al., 2012). The largest dietary differences were seen in consumption of fruits, vegetables, and meat. Community-dwellers correlated 98% with a moderate fat/high fiber diet and long-term care dwellers correlated 83% with high fat/low fiber diet (Claesson et al., 2012). The gut microbiota of community-dwellers was more diverse than long-stay subjects and grouped more closely with healthy young adults, indicating that age itself is not the driving factor of microbial change. Similar to young adults, community-dwellers had a higher proportion of phylum Firmicutes and unclassified bacteria, and abundant populations of genera *Coprococcus, Roseburia, Ruminococcus,* and *Butyricoccus* when compared to long-term stay individuals. Long-stay subjects had a higher incidence of frailty accompanied by a proportional increase in Bacteroidetes and an increased abundance of *Alistipes* and *Oscillibacter* when compared to healthier community-dwelling elderly (Claesson et al., 2012). Increasingly frail individuals showed a significant 26-fold reduction in the number of *Lactobacillus* and a significant sevenfold increase in the number of *Enterobacteriaceae* compared to less frail subjects (van Tongeren et al., 2005).

ANTIBIOTICS
The compounded effects of poor diet, ailing health, and prolonged stays in a hospital or long-term care facility reduce the prevalence of protective gut microbiota and give way to detrimental populations (Bartosch et al., 2004; Wu et al., 2011). This leaves the elderly individual vulnerable to infection and disease and a prime candidate for antibiotic usage. Unfortunately, antibiotic therapies only exacerbate the flux and instability of the already fragile gut microbiome in unhealthy elderly. The use of antibiotics in elderly populations is especially prevalent in hospital and long-term care facilities and it is estimated that nearly 20% of elderly patients in hospitals are receiving antibiotic treatment at any given time (Bartosch et al., 2004).

Antibiotics cause significant disturbances in gut microbiota resulting in the suppression of both beneficial and pathogenic species, allowing the overgrowth of antibiotic-resistant strains. In young, healthy volunteers administered two separate courses of the antibiotic ciprofloxacin, a dramatic change in the microbiota was noted, followed by the return to an alternative stable state of undetermined consequences (Dethlefsen and Relman, 2011). Use of broad-spectrum antibiotics is associated with the opportunistic bacterium *Clostridium difficile* which flourishes in the antibiotic-weakened gut and results in severe diarrhea (Macfarlane, 2014). Elderly hospital patients and others with fragile immune systems are especially susceptible to this life-threatening infection.

Most often, elderly individuals exposed to antibiotics see an increased relative abundance of Bacteroidetes and a significant increase in Bacteroidetes:Firmicutes ratio (Claesson et al., 2011). Beneficial anaerobic species in the colon such as *Bifidobacterium*, *Lactobacillus*, and *Bacteroides* can be drastically reduced or even eradicated with the use of antibiotics (Bartosch et al., 2004). *Bifidobacterium* and *Lactobacillus* are producers of short chain fatty acids (SCFA's), a nutrient vital to the proper function of intestinal cells; the loss of these bacteria can be especially detrimental. A study examining the differences in bacterial colonies between healthy elderly, hospitalized patients, and hospitalized patients receiving antibiotics, found that the hospitalized patients receiving antibiotics saw a significant reduction in the numbers of *Bifidobacterium* spp. and an increased relative abundance of *Enterococcus faecalis* compared to the other two groups. In some patients, the antibiotic treatment eliminated certain bacterial communities altogether (Bartosch et al., 2004).

Effects of antibiotic treatment on gut microbiota can differ significantly with the type and dose of antibiotic administered. A study by Bartosch et al. (2004) following elderly patients receiving antibiotics, found that the same antibiotic, clarithromycin, had different effects on gut microbiota at different doses. A low dose of the antibiotic decreased the proportion of Bacteroidetes (*Bacteroides* and *Parabacteroides*) and increased Firmicutes (*Alistipes*) and a high dose increased the proportion of Bacteroidetes (*Parabacteroides*) and decreased the proportion of Firmicutes (*Alistipes*; Claesson et al., 2011). Countless variables must be considered with the use of antibiotics in elderly individuals. What seems like a lifesaving drug may have detrimental effects on the aging microbiome and the health of the individual. Additional research is needed to inform practitioners on the safest ways to use antibiotics on the elderly while supporting their potentially fragile gut microbiota.

PROBIOTICS AND PREBIOTICS

Probiotics and prebiotics, when taken together or individually, may be particularly beneficial in restoring the proper microbial balance to the elderly gut microbiota, helping to mitigate the detrimental effects of antibiotic usage and under nutrition. Probiotics are live microbes that when administered in sufficient quantities are beneficial to the host. Prebiotics are non-digestible food ingredients such as inulin or various oligosaccharides, which have been show to selectively stimulate growth of beneficial bacterial populations in the large intestine. Probiotic foods and supplements often contain *Bifidobacterium* and/or *Lactobacillus* organisms, both of which are extremely important to proper function of the intestine (Duncan and Flint, 2013). *Bifidobacterium* and *Lactobacillus* are often depleted in elderly individuals as health deteriorates. Research shows that consumption of probiotics containing these strains can result in a notable rise in their abundance along with a reduction of more pathogenic microorganisms in the gut (Toward et al., 2012). Prebiotics may support the *Bifidobacterium* and *Lactobacillus* species delivered via probiotic supplementation by providing a fermentable food source for these bacteria, allowing them to flourish. More specifically, it has been reported that prebiotics have the ability to exert a bifidogenic effect on human subjects (O'Connor et al., 2014).

A recent *in vitro* study showed promise that the elderly gut microbiota can in fact be modulated with appropriate probiotics. Species of *Bifidobacterium* and *Lactobacillus* along with two prebiotics were added to the fecal batch culture of elderly participants. The addition of the beneficial bacteria significantly increased the *Bifidobacterium* and decreased the *Bacteroides* count after fermentation (Likotrafiti et al., 2014). Both probiotic/prebiotic combinations added to the culture increased the *Bifidobacterium* and *Lactobacillus* count in the vessel representing the distal colon. These results represented a major shift in the gut microbiota toward a healthier colon (Likotrafiti et al., 2014). However, prebiotics alone have also been shown to improve the health and alter the gut microbial composition of elderly populations. A study providing inulin supplementation to an elderly cohort increased *Bifidobacterium* levels (Guigoz et al., 2002). Multiple studies using either fructo or GOSs demonstrated both bifidogenic effects and beneficial immune-related effects. Specific immune related effects included reduction in pro-inflammatory cytokines and an increase in the anti-inflammatory cytokine, IL-10.

While probiotic supplementation has become a widely utilized tool to positively impact health by assisting with digestion, bolstering intestinal barrier function and coordinating with the body to regulate both the innate and specific immune responses, the mechanisms by which they exert these beneficial effects is poorly understood (Siciliano and Mazzeo, 2012). Proteomic-based probiotic research is beginning to inform both researchers and industry that adaptation and adherence properties specific to probiotic strains influence their ability to colonize the host (Siciliano and Mazzeo, 2012; van de Guchte et al., 2012). Additionally, these adaptation and adherence mechanisms have been reported to potentially be strain specific, making it difficult to globally apply these mechanisms to all probiotic bacterial strains (Siciliano and Mazzeo, 2012).

Experimentation on the effects of probiotics and prebiotics of the elderly gut microbiome is still limited, but results of the available research lends merit to the notion that beneficial bacteria in the form of probiotics and the indigestible fibers of prebiotics has potential to help restore stability, increase diversity and beneficially alter the immune system in the aging gut (Vulevic et al., 2008). However, these beneficial effects must be placed in perspective given the lack of a mutually agreed upon selection criteria, evaluation methodologies and a clear mechanistic model. With the reduced cost of sequencing and continued proteomic research, hopefully researchers will be able to speak with increased certainty as to the reasons probiotics can be beneficial to human host.

CONCLUSION

The microbes that reside in our gastrointestinal tract comprise a dynamic community that changes throughout the lifespan of an individual. The early years of infancy and childhood are characterized by a microbial state that has been described as chaotic because of the rapid and dramatic fluctuations observed. While the microbiota of small children begins to resemble that of adults at a very early age, there is a paucity of studies examining temporal microbial community shifts in children beyond infancy, so the stability

of their microbiota is not known. Once stable dietary patterns are established, the microbiota of adults remains relatively unaltered; however, significant weight changes have been associated with a higher amount of microbial instability. Finally, factors related to aging, including increased use of pharmaceuticals and changes in diet likely play an important role in shaping the microbial communities residing in the elderly. Changes in physical activity and hormone levels may also be important determinants of the elderly microbiome, but they have not yet been investigated with sufficient depth. Some evidence suggests that the microbial communities of healthy elderly individuals are similar to that of younger adults, but whether the health of the individual contributes to microbial stability or vice versa is not known. Current data suggest that diet is an important driver in the development of the gut microbiome and could serve as a means of therapeutic intervention for prevention of diseases. Studies linking the composition and function of the gut microbiome and disease development certainly highlight the need for a better understanding of temporal microbiome dynamics and their predictors.

ACKNOWLEDGMENTS

REFERENCES

Aagaard, K., Ma, J., Antony, K. M., Ganu, R., Petrosino, J., and Versalovic, J. (2014). The placenta harbors a unique microbiome. *Sci. Transl. Med.* 6, 237ra265. doi: 10.1126/scitranslmed.3008599

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., et al. (2011). Enterotypes of the human gut microbiome. *Nature* 473, 174–180. doi: 10.1038/nature09944

Bartosch, S., Fite, A., Macfarlane, G. T., and McMurdo, M. E. (2004). Characterization of bacterial communities in feces from healthy elderly volunteers and hospitalized elderly patients by using real-time PCR and effects of antibiotic treatment on the fecal microbiota. *Appl. Environ. Microbiol.* 70, 3575–3581. doi: 10.1128/AEM.70.6.3575-3581.2004

Bergström, A., Skov, T. H., Bahl, M. I., Roager, H. M., Christensen, L. B., Ejlerskov, K. T., et al. (2014). Establishment of intestinal microbiota during early life: a longitudinal, explorative study of a large cohort of Danish infants. *Appl. Environ. Microbiol.* 80, 2889–2900. doi: 10.1128/AEM.00342-14

Bezirtzoglou, E., Tsiotsias, A., and Welling, G. W. (2011). Microbiota profile in feces of breast-and formula-fed newborns by using fluorescence in situ hybridization (FISH). *Anaerobe* 17, 478–482. doi: 10.1016/j.anaerobe.2011.03.009

Britton, E., and McLaughlin, J. T. (2013). Ageing and the gut. *Proc. Nutr. Soc.* 72, 173–177. doi: 10.1017/S0029665112002807

Cabrera-Rubio, R., Collado, M. C., Laitinen, K., Salminen, S., Isolauri, E., and Mira, A. (2012). The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery. *Am. J. Clin. Nutr.* 96, 544–551. doi: 10.3945/ajcn.112.037382

Claesson, M. J., Cusack, S., O'Sullivan, O., Greene-Diniz, R., de Weerd, H., Flannery, E., et al. (2011). Composition, variability, and temporal stability of the intestinal microbiota of the elderly. *Proc. Natl. Acad. Sci. U.S.A.* 108(Suppl. 1), 4586–4591. doi: 10.1073/pnas.1000097107

Claesson, M. J., Jeffery, I. B., Conde, S., Power, S. E., O'Connor, E. M., Cusack, S., et al. (2012). Gut microbiota composition correlates with diet and health in the elderly. *Nature* 488, 178–184. doi: 10.1038/nature11319

Cotillard, A., Kennedy, S. P., Kong, L. C., Prifti, E., Pons, N., Le Chatelier, E., et al. (2013). Dietary intervention impact on gut microbial gene richness. *Nature* 500, 585–588. doi: 10.1038/nature12480

David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., et al. (2014). Diet rapidly and reproducibly alters

the human gut microbiome. *Nature* 505, 559–563. doi: 10.1038/nature12820

De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J. B., Massart, S., et al. (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14691–14696. doi: 10.1073/pnas.1005963107

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, 4554–4561. doi: 10.1073/pnas.1000087107

Duncan, S. H., and Flint, H. J. (2013). Probiotics and prebiotics and health in ageing populations. *Maturitas* 75, 44–50. doi: 10.1016/j.maturitas.2013.02.004

Duncan, S. H., Lobley, G. E., Holtrop, G., Ince, J., Johnstone, A. M., Louis, P., et al. (2008). Human colonic microbiota associated with diet, obesity and weight loss. *Int. J. Obes.* 32, 1720–1724. doi: 10.1038/ijo.2008.155

Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M., et al. (2005). Diversity of the human intestinal microbial flora. *Science* 308, 1635–1638. doi: 10.1126/science.1110591

Faith, J. J., Guruge, J. L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A. L., et al. (2013). The long-term stability of the human gut microbiota. *Science* 341, 44–52. doi: 10.1126/science.1237439

Fallani, M., Amarri, S., Uusijarvi, A., Adam, R., Khanna, S., Aguilera, M., et al. (2011). Determinants of the human infant intestinal microbiota after the introduction of first complementary foods in infant samples from five European centres. *Microbiology* 157, 1385–1392. doi: 10.1099/mic.0.042143-0

Fallani, M., Young, D., Scott, J., Norin, E., Amarri, S., Adam, R., et al. (2010). Intestinal microbiota of 6-week-old infants across Europe: geographic influence beyond delivery mode, breast-feeding, and antibiotics. *J. Pediatr. Gastroenterol. Nutr.* 51, 77–84. doi: 10.1097/MPG.0b013e3181d1b11e

Guigoz, Y., Rochat, F., Perruisseau-Carrier, G., Rochat, I., and Schiffrin, E. J. (2002). Effects of oligosaccharide on the faecal flora and non-specific immune system in elderly people. *Nutr. Res.* 22, 13–25. doi: 10.1016/S0271-5317(01)00354-2

Harmsen, H. J., Wildeboer-Veloo, A. C., Raangs, G. C., Wagendorp, A. A., Klijn, N., Bindels, J. G., et al. (2000). Analysis of intestinal flora development in breast-fed and formula-fed infants by using molecular identification and detection methods. *J. Pediatr. Gastroenterol. Nutr.* 30, 61–67. doi: 10.1097/00005176-200001000-00019

Huse, S. M., Ye, Y., Zhou, Y., and Fodor, A. A. (2012). A core human microbiome as viewed through 16S rRNA sequence clusters. *PLoS ONE* 7:e34242. doi: 10.1371/journal.pone.0034242

Jost, T., Lacroix, C., Braegger, C. P., and Chassard, C. (2012). New insights in gut microbiota establishment in healthy breast fed neonates. *PLoS ONE* 7:e44595. doi: 10.1371/journal.pone.0044595

Kelly, D., King, T., and Aminov, R. (2007). Importance of microbial colonization of the gut in early life to the development of immunity. *Mutat. Res.* 622, 58–69. doi: 10.1016/j.mrfmmm.2007.03.011

Knol, J., Boehm, G., Lidestri, M., Negretti, F., Jelinek, J., Agosti, M., et al. (2005). Increase of faecal bifidobacteria due to dietary oligosaccharides induces a reduction of clinically relevant pathogen germs in the faeces of formula-fed preterm infants. *Acta Paediatr. Suppl.* 94, 31–33. doi: 10.1080/08035320510043529

Koenig, J. E., Spor, A., Scalfone, N., Fricker, A. D., Stombaugh, J., Knight, R., et al. (2011). Succession of microbial consortia in the developing infant gut microbiome. *Proc. Natl. Acad. Sci. U.S.A.* 108(Suppl. 1), 4578–4585. doi: 10.1073/pnas.1000081107

Koren, O., Knights, D., Gonzalez, A., Waldron, L., Segata, N., Knight, R., et al. (2013). A guide to enterotypes across the human body: meta-analysis of microbial community structures in human microbiome datasets. *PLoS Comput. Biol.* 9:e1002863. doi: 10.1371/journal.pcbi.1002863

Le Chatelier, E., Nielsen, T., Qin, J., Prifti, E., Hildebrand, F., Falony, G., et al. (2013). Richness of human gut microbiome correlates with metabolic markers. *Nature* 500, 541–546. doi: 10.1038/nature12506

Ley, R. E., Turnbaugh, P. J., Klein, S., and Gordon, J. I. (2006). Microbial ecology: human gut microbes associated with obesity. *Nature* 444, 1022–1023. doi: 10.1038/4441022a

Likotrafiti, E., Tuohy, K. M., Gibson, G. R., and Rastall, R. A. (2014). An *in vitro* study of the effect of probiotics, prebiotics and synbiotics on the elderly faecal microbiota. *Anaerobe* 27, 50–55. doi: 10.1016/j.anaerobe.2014.03.009

Lin, A., Bik, E. M., Costello, E. K., Dethlefsen, L., Haque, R., Relman, D. A., et al. (2013). Distinct distal gut microbiome diversity and composition in healthy

children from Bangladesh and the United States. *PLoS ONE* 8:e53838. doi: 10.1371/journal.pone.0053838

Ma, J., Prince, A. L., Bader, D., Hu, M., Ganu, R., Baquero, K., et al. (2014). High-fat maternal diet during pregnancy persistently alters the offspring microbiome in a primate model. *Nat. Commun.* 5, 3889. doi: 10.1038/ncomms4889

Macfarlane, S. (2014). Antibiotic treatments and microbes in the gut. *Environ. Microbiol.* 16, 919–924. doi: 10.1111/1462-2920.12399

Marques, T. M., Wall, R., Ross, R. P., Fitzgerald, G. F., Ryan, C. A., and Stanton, C. (2010). Programming infant gut microbiota: influence of dietary and environmental factors. *Curr. Opin. Biotechnol.* 21, 149–156. doi: 10.1016/j.copbio.2010.03.020

Muegge, B. D., Kuczynski, J., Knights, D., Clemente, J. C., González, A., Fontana, L., et al. (2011). Diet drives convergence in gut microbiome functions across *Mammalian* phylogeny and within humans. *Science* 332, 970–974. doi: 10.1126/science.1198719

O'Connor, E. M., O'Herlihy, E. A., and O'Toole, P. W. (2014). Gut microbiota in older subjects: variation, health consequences and dietary intervention prospects. *Proc. Nutr. Soc.* 13, 1–11. doi: 10.1017/S0029665114000597

Okada, H., Kuhn, C., Feillet, H., and Bach, J. F. (2010). The "hygiene hypothesis" for autoimmune and allergic diseases: an update. *Clin. Exp. Immunol.* 160, 1–9. doi: 10.1111/j.1365-2249.2010.04139.x

Oozeer, R., van Limpt, K., Ludwig, T., Ben Amor, K., Martin, R., Wind, R. D., et al. (2013). Intestinal microbiology in early life: specific prebiotics can have similar functionalities as human-milk oligosaccharides. *Am. J. Clin. Nutr.* 98, 561S–571S. doi: 10.3945/ajcn.112.038893

Ou, J., Carbonero, F., Zoetendal, E. G., DeLany, J. P., Wang, M., Newton, K., et al. (2013). Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans. *Am. J. Clin. Nutr.* 98, 111–120. doi: 10.3945/ajcn.112.056689

Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A., and Brown, P. O. (2007). Development of the human infant intestinal microbiota. *PLoS Biol.* 5:e177. doi: 10.1371/journal.pbio.0050177

Preidis, G. A., and Versalovic, J. (2009). Targeting the human microbiome with antibiotics, probiotics, and prebiotics: gastroenterology enters the metagenomics era. *Gastroenterology* 136, 2015–2031. doi: 10.1053/j.gastro.2009.01.072

Qato, D. M., Alexander, G., Conti, R. M., Johnson, M., Schumm, P., and Lindau, S. (2008). Use of prescription and over-the-counter medications and dietary supplements among older adults in the united states. *J. Am. Med. Assoc.* 300, 2867–2878. doi: 10.1001/jama.2008.892

Rajilic-Stojanovic, M., Heilig, H. G., Tims, S., Zoetendal, E. G., and de Vos, W. M. (2013). Long-term monitoring of the human intestinal microbiota composition. *Environ. Microbiol.* 15, 1146–1159. doi: 10.1111/1462-2920.12023

Roager, H. M., Licht, T. R., Poulsen, S. K., Larsen, T. M., and Bahl, M. I. (2014). Microbial enterotypes, inferred by the prevotella-to-bacteroides ratio, remained stable during a 6-month randomized controlled diet intervention with the new nordic diet. *Appl. Environ. Microbiol.* 80, 1142–1149. doi: 10.1128/AEM.03549-13

Rook, G. W. (2012). Hygiene hypothesis and autoimmune diseases. *Clin. Rev. Allergy Immunol.* 42, 5–15. doi: 10.1007/s12016-011-8285-8

Schnorr, S. L., Candela, M., Rampelli, S., Centanni, M., Consolandi, C., Basaglia, G., et al. (2014). Gut microbiome of the Hadza hunter-gatherers. *Nat. Commun.* 5, 3654. doi: 10.1038/ncomms4654

Sekirov, I., Russell, S. L., Antunes, L. C. M., and Finlay, B. B. (2010). Gut microbiota in health and disease. *Physiol. Rev.* 90, 859–904 doi: 10.1152/physrev.00045.2009

Siciliano, R. A., and Mazzeo, M. F. (2012). Molecular mechanisms of probiotic action: a proteomic perspective. *Curr. Opin. Microbiol.* 15, 390–396. doi: 10.1016/j.mib.2012.03.006

Subramanian, S., Huq, S., Yatsunenko, T., Haque, R., Mahfuz, M., Alam, M. A., et al. (2014). Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* 510, 417–421. doi: 10.1038/nature13421

Tilg, H., and Kaser, A. (2011). Gut microbiome, obesity, and metabolic dysfunction. *J. Clin. Invest.* 121, 2126–2132. doi: 10.1172/JCI58109

Toward, R., Montandon, S., Walton, G., and Gibson, G. R. (2012). Effect of prebiotics on the human gut microbiota of elderly persons. *Gut Microbes* 3, 57–60. doi: 10.4161/gmic.19411

Turroni, F., Peano, C., Pass, D. A., Foroni, E., Severgnini, M., Claesson, M. J., et al. (2012). Diversity of bifidobacteria within the infant gut microbiota. *PLoS ONE* 7:e36957. doi: 10.1371/journal.pone.0036957

van de Guchte, M., Chaze, T., Jan, G., and Mistou, M.-Y. (2012). Properties of probiotic bacteria explored by proteomic approaches. *Curr. Opin. Microbiol.* 15, 381–389. doi: 10.1016/j.mib.2012.04.003

van Tongeren, S. P., Slaets, J. P., Harmsen, H. J., and Welling, G. W. (2005). Fecal microbiota composition and frailty. *Appl. Environ. Microbiol.* 71, 6438–6442. doi: 10.1128/AEM.71.10.6438-6442.2005

Vulevic, J., Drakoularakou, A., Yaqoob, P., Tzortzis, G., and Gibson, G. R. (2008). Modulation of the fecal microflora profile and immune function by a novel trans-galactooligosaccharide mixture (B-GOS) in healthy elderly volunteers. *Am. J. Clin. Nutr.* 88, 1438–1446.

Woodmansey, E. J. (2007). Intestinal bacteria and ageing. *J. Appl. Microbiol.* 102, 1178–1186. doi: 10.1111/j.1365-2672.2007.03400.x

Wu, G. D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y. Y., Keilbaugh, S. A., et al. (2011). Linking long-term dietary patterns with gut microbial enterotypes. *Science* 334, 105–108. doi: 10.1126/science.1208344

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., et al. (2012). Human gut microbiome viewed across age and geography. *Nature* 486, 222–227. doi: 10.1038/nature11053

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Aging and the human gut microbiota—from correlation to causality

Sitaraman Saraswati[1] and Ramakrishnan Sitaraman[2]*

[1] Department of Biochemistry, Dayananda Sagar College of Arts, Science, and Commerce, Bangalore, India
[2] Department of Biotechnology, TERI University, New Delhi, India
*Correspondence: minraj@gmail.com

INTRODUCTION

The human gastrointestinal (GI) tract harbors the largest number and concentration of microbes found in the human body. Perturbations in the gut microbial ecosystem have also been associated with conditions as diverse as chronic GI diseases (e.g., Crohn's disease, ulcerative colitis), metabolic disorders (e.g., diabetes types 1 and 2, obesity) and antibiotic use (for a review see Sekirov et al., 2010). Metagenomic culture-independent methods have enabled the unraveling of the complexity of the gut microbiota (Rajilić-Stojanović et al., 2009). Given the considerable inter-individual diversity in the actual composition of the microbiota, significant collaborative attempts have been made to systematize the available knowledge (Arumugam et al., 2011; Human Microbiome Project Consortium, 2012) and identify "core" microbiota that are conserved among humans to facilitate meaningful comparisons (Huse et al., 2012). Changes in the microbial composition also take place with age, with a high degree of variability at the two extremes of infancy and old age, punctuated by comparative stability during adulthood (for reviews, see Woodmansey, 2007; O'Toole and Claesson, 2010). Given that increases in life expectancy will likely result in an increase in the elderly population worldwide, analysis of the contribution of the microbiota to healthy aging assumes greater significance (for a recent review, see Tiihonen et al., 2010).

Age-related spatio-temporal variations in the microbiota are best viewed within an ecological- evolutionary framework (see review by Costello et al., 2012). Diet is a major, controllable environmental factor influencing the composition of the host microbiome, with the high-fat, sugar-rich Western diet contributing to a *Bacteroides*-dominant microbiome and high-fiber diet to one dominated by Firmicutes with a strong correlation between long-term diet and enterotypes (Wu et al., 2011). In terms of ecological succession, the Bifidobacterium-dominated microbiota of the infant changes over time into the Bacteroidetes- and Firmicutes-dominated microbiota of the adult (Ottman et al., 2012), remaining fairly stable through adulthood in the absence of perturbations like long-term dietary changes or repeated antibiotic intervention. Pathogens may then be viewed as invasive species in the ecological sense, constantly testing the resilience of the native ecosystem, resulting in their elimination, low-level persistence (enabling future opportunism), or establishment causing disease.

AGE- AND ENVIRONMENT-RELATED CHANGES IN THE GUT MICROBIOTA

The most noticeable feature in the microbiota of elderly individuals is an alteration in the relative proportions of the Firmicutes and the Bacteroidetes, with the elderly having a higher proportion of Bacteroidetes while young adults have higher proportions of Firmicutes (Mariat et al., 2009). Significant decreases in Bifidobacteria, *Bacteriodes,* and *Clostridium* cluster IV have also been reported (Zwielehner et al., 2009). Variability among individuals ranges from 3% to 92% for Bacteroidetes and between 7% and 94% for Firmicutes. The microbiota of individual subjects however exhibit less temporal variability (Claesson et al., 2011).

Changes occurring in the microbiota during aging can have an impact on host health. van Tongeren et al. (2005) studied the relationship between microbial diversity and frailty scores in elderly individuals. A significant reduction in the proportion of lactobacilli, *Bacteroides/Prevotella* and *Faecalibacterium prausnitzii*, and an increase in the proportion of *Ruminococcus*, *Atopobium,* and Enterobacteriaceae was seen in individuals with high frailty scores. Recently, Claesson et al. (2012) studied the relationship between diet, host health, environment, and the gut microbiota. Specifically, association was observed between microbial diversity and the functional independence measure (FIM), the Barthel index (used to evaluate performance in daily routine activities) and nutrition. Decreased microbial diversity correlated with increased frailty, decreased diet diversity and health parameters, and with increased levels of inflammatory markers. Individuals living in a community had the most diverse microbiota and were healthier as compared to those in short- or long-term residential care. Bartosch et al.

(2004) reported that hospitalization itself appeared to result in a decreased abundance of the *Bacteroides-Prevotella* group. Later studies by Claesson et al. (2012) further detailed the effects of residence location on gut microbial diversity. Residence location also affects the microbiota of patients on antibiotic treatment, with highest levels of bifidobacteria in the community-dwelling group and lowest in those in long-term residential care. The levels of *Lactobacillus* in the antibiotic-untreated group were higher in rehabilitation (hospital stay < 6 weeks) as compared to long-stay or community-dwellers.

Predictably, antibiotic treatment has been reported to affect both richness and diversity of the microbiota and is associated with decreases in bifidobacteria (Bartosch et al., 2004; Woodmansey et al., 2004; O'Sullivan et al., 2013) as well as the *Bacteroides-Prevotella* group (Bartosch et al., 2004; Woodmansey et al., 2004). Lactobacilli, however, are observed to have increased in antibiotic-treated elderly subjects (Woodmansey et al., 2004; O'Sullivan et al., 2013); similarly, an increase in clostridial diversity has also been reported (Woodmansey et al., 2004). The changes taking place in response to antibiotics are more apparent at the genus rather than the family or phylum levels (O'Sullivan et al., 2013). Treatment with antibiotics can result in *Clostridium difficile* infection in the elderly, manifesting as *C. difficile*-associated diarrhea (CDAD). Reduced species diversity in CDAD patients compared to healthy elderly and young adults accompanied by a large reduction in bifidobacteria, *Bacteroides,* and *Prevotella* has been reported (Hopkins and MacFarlane, 2002). However, an increase in facultative bacteria along with an increase in diversity of clostridial and lactobacilli species in CDAD patients was reported in the same study. A recent study also detected differences at the genus level between *C. difficile* -negative and -positive subjects, and patients with CDAD (Rea et al., 2012). Incidentally, the isolation of the hypervirulent *C. difficile* R027 ribotype from one asymptomatic individual in this study who exhibited greater microbial diversity compared to CDAD patients, serves to highlight the importance of an intact and unperturbed gut ecosystem in resisting colonization by pathogens. Restoration

of the microbiota and curing CDAD by fecal microbiota transplantation (FMT) in recent years presents a novel therapeutic strategy that is under intense scrutiny (for a discussion see Vrieze et al., 2013). In contrast to antibiotics, the common usage of non-steroidal anti-inflammatory drugs (NSAIDs) does not appear to significantly perturb the microbiota (Tiihonen et al., 2008; Mäkivuokko et al., 2010).

Interestingly, centenarians harbor less diverse microbiota, though Bacteroidetes and Firmicutes still constitute the dominant phyla (Biagi et al., 2010), with enrichment for potentially pathogenic Proteobacteria. Biagi et al. (2010) reported higher levels of *Akkermansia* in the elderly, compared to young adults, in contrast to an earlier study (Collado et al., 2007) that reported a decrease in this genus with age. Subsequent functional microbiome profiling of selected, well-characterized samples from this cohort indicated increased abundance of genes involved in aromatic amino acid metabolism, decreased abundance of those involved in short-chain (≤ 6) fatty acid production and an enrichment of "pathobionts"—low-abundance microbiota that promote and sustain pro-inflammatory conditions (Rampelli et al., 2013). This supports an earlier finding by Collino et al. (2013) that increased levels of phenylacetylglutamine (PAG) and *p*-cresol-sulfate (PCS), derived from the catabolism of aromatic amino acids, were excreted in the urine of centenarians. Thus, the changes in the gut microbiota of the elderly are reflected in the changes in the microbial metabolism.

FROM CORRELATION TO CAUSALITY—SOME GENERAL CONSIDERATIONS

High-throughput analytical tools and meta-"omics" enable probing of the host-microbiota relationship at high resolution, helping correlate healthy or diseased states with the detailed composition of the microbiota, and informing the use of well-characterized (e.g., probiotic) or largely unknown (e.g., stool transplants) mixtures of microorganisms for restorative or maintenance purposes. However, complicating matters further is the existence of distinct ecological niches all along the

alimentary canal, indicating that the common (and convenient) method of fecal sampling for microbiota studies may not adequately reflect the situation *in vivo* (Li et al., 2011). Ideally, we would like to determine the identity of the molecules that mediate host-microbiota interactions, and how their deployment is regulated. Here, information about host-pathogen interactions and general microbiology offers insights into the range of intra- and inter-species interactions, and even inter-kingdom ones (**Table 1**). However, given our current inability to convincingly delineate the contextually most significant effector mechanisms involved in the host-microbiota interaction over a lifetime, it is difficult to tease apart causality from correlation. Moreover, the host and the microbiota impact each other reciprocally, and the microbiota themselves interact in many modes among themselves. While current host signals may modulate the microbiota, it is an open question whether these signals themselves were induced, at least in some measure by components of the microbiota themselves. Theoretically, the host could also modulate the microbiota so that microbial responses are, in turn, beneficial to itself. The landmark study of Claesson et al. (2012) points to the possibility of such a reciprocal (and more confusingly, recursive) relationship between host health and microbial diversity.

From an evolutionary standpoint, we would also like to know how much these interactions and associations are modulated over the host lifetime and during co-evolution in order to benefit both partners. Their persistence is also dependent on the forces of selection operative at a given time (Sancar, 2008; Lukeš et al., 2011), such as bacteriophage infection (Reyes et al., 2012; Koskella, 2013). The recent discovery that an unknown secreted protein from human intestinal cells decreases conjugation efficiency in *E. coli* indicates that the host can potentially influence the composition, the rate of evolution and lateral gene transfer among its microbiota (Machado and Sommer, 2014). An unexplored consideration is the potential influence of host hormones and their changing levels over age on the microbiota. Additionally, the relative abundance of a given enterotype or species may also not be an unambiguous

Table 1 | Examples of effector mechanisms involved in inter-species and inter-kingdom interactions.

Effector molecule(s)	Relevant aspects for microbiota studies	Reference(s)
Secreted phosopholipases (PldA and PldB) of *Pseudomonas aeruginosa*.	Killing of competing bacteria (PldA, PldB), promoting *P. aeruginosa* internalization by host cells (PldB).	Russell et al., 2013; Jiang et al., 2014
Autoinducer-2 (AI-2)	Mediates inter-species signaling among the microbiota. Production of AI-2 by *Bifidobacterium breve* UCC2003 (a probiotic strain) is required for gut colonization and protection against *Salmonella* infection in *C. elegans*.	Christiaen et al., 2014
Autoinducer-3 (AI-3), epinephrine and norepinephrine	Autoinducer 3 (AI-3) produced by the intestinal microbiota can cross-signal with host hormones epinephrine and norepinephrine.	Sperandio et al., 2003
	Epinephrine/Norepinephrine can be sensed by enterohemorrhagic *E. coli*, inducing the expression of the LEE pathogenicity island and the flagella regulon. The AI-3 signaling cascade is conserved across several bacterial species.	Reading and Sperandio, 2006

predictor of relative importance in the ecological sense. Therefore, identifying keystone species that might have an effect on the ecosystem disproportionate to their abundance would be very valuable for focused studies of the microbiota. We surmise that the benefits arising from probiotic administration could be due to their temporary assumption of such a "keystone" role. Notably, current studies of microbiota concentrate solely on the doubtlessly important bacteria, omitting archaea and clinically relevant eukaryotes (fungi), an approach that could potentially miss less abundant but nevertheless important species. A recent finding on the important role of Dectin-1, a C-type lectin receptor and an innate immune sensor of fungi, in preventing intestinal colitis-associated pathology underlines the importance of interactions between the human host and the numerically less abundant intestinal fungi (Iliev et al., 2012).

Natural selection operates simultaneously at multiple ecological levels, ranging from single unicellular organisms to entire communities and ecosystems. The magnitude and relative importance of multiple, and often stochastic, selection pressures acting over a human lifetime, therefore need careful consideration. Thus, it could be inaccurate to ascribe a given microbiota profile solely to a single factor (e.g., diet) even though there may be some correlation between the said profile and a single contributing factor (see review

by Yeoman et al., 2011). Additionally, non-equilibrium (co-) evolutionary processes may not necessarily result in optimality. Rather, they are governed by the actual functionality of the given arrangement ("phenotype") and its ability to propagate itself (fitness), not on the details of the arrangement itself (components, genotypes etc.). Microbiota research, in the context of aging or otherwise, will greatly benefit from the integration of several disparate pieces of mechanistic information within an evolutionary-ecological framework in order to determine the causes underlying our observations, and the formulation of plausible mechanistic models describing how these causes result in the observed effects.

ACKNOWLEDGMENTS

REFERENCES

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., et al. (2011). Enterotypes of the human gut microbiome. *Nature* 473, 174–180. doi: 10.1038/nature09944

Bartosch, S., Fite, A., MacFarlane, G. T., and McMurdo, M. E. T. (2004). Characterization of bacterial communities in feces from healthy elderly volunteers and hospitalized elderly patients by using real-time PCR and effects of antibiotic treatment on the fecal microbiota. *Appl. Environ. Microbiol.* 70, 3575–3581. doi: 10.1128/AEM.70.6.3575-3581.2004

Biagi, E., Nylund, L., Candela, M., Ostan, R., Bucci, L., Pini, E., et al. (2010). Through ageing, and

beyond: gut microbiota and inflammatory status in seniors and centenarians. *PLoS ONE* 5:e10667. doi: 10.1371/journal.pone.0010667

Christiaen, S. E. A., O'Connell Motherway, M., Bottacini, F., Lanigan, N., Casey, P. G., Huys, G., et al. (2014). Autoinducer-2 plays a crucial role in gut colonization and probiotic functionality of Bifidobacterium breve UCC2003. *PLoS ONE* 9:e98111. doi: 10.1371/journal.pone.0098111

Claesson, M. J., Cusack, S., O'Sullivan, O., Greene-Diniz, R., de Weerd, H., Flannery, E., et al. (2011). Composition, variability, and temporal stability of the intestinal microbiota of the elderly. *Proc. Natl. Acad. Sci. U.S.A.* 108 Suppl. 1, 4586–4591. doi: 10.1073/pnas.1000097107

Claesson, M. J., Jeffery, I. B., Conde, S., Power, S. E., O'Connor, E. M., Cusack, S. et al. (2012). Gut microbiota composition correlates with diet and health in the elderly. *Nature* 488, 178–184. doi: 10.1038/nature11319

Collado, M. C., Derrien, M., Isolauri, E., de Vos, W. M., and Salminen, S. (2007). Intestinal integrity and *Akkermansia muciniphila*, a mucin-degrading member of the intestinal microbiota present in infants, adults, and the elderly. *Appl. Environ. Microbiol.* 73, 7767–7770. doi: 10.1128/AEM.01477-07

Collino, S., Montoliu, I., Martin, F.-P. J., Scherer, M., Mari, D., Salvioli, S., et al. (2013). Metabolic signatures of extreme longevity in northern Italian centenarians reveal a complex remodeling of lipids, amino acids, and gut microbiota metabolism. *PLoS ONE* 8:e56564. doi:10.1371/journal.pone.0056564

Costello, E. K., Stagaman, K., Dethlefsen, L., Bohannan, B. J. M., and Relman, D. A. (2012). The application of ecological theory toward an understanding of the human microbiome. *Science* 336, 1255–1262. doi: 10.1126/science.1224203

Hopkins, M. J., and MacFarlane, G. T. (2002). Changes in predominant bacterial populations in human faeces with age and with *Clostridium difficile* infection. *J. Med. Microbiol.* 51, 448–454. Available online at: http://jmm.sgmjournals.org/content/51/5/448.full.pdf+html

Huse, S. M., Ye, Y., Zhou, Y., and Fodor, A. A. (2012). A core human microbiome as viewed through 16S rRNA sequence clusters. *PLoS ONE* 7:e34242. doi: 10.1371/journal.pone.0034242

Iliev, I. D., Funari, V. A., Taylor, K. D., Nguyen, Q., Reyes, C. N., Strom, S. P., et al. (2012). Interactions between commensal fungi and the C-type lectin receptor Dectin-1 influence colitis. *Science* 336, 1314–1317. doi: 10.1126/science.1221789

Jiang, F., Waterfield, N. R., Yang, J., Yang, G., and Jin, Q. (2014). A Pseudomonas aeruginosa type VI secretion phospholipase D effector targets both prokaryotic and eukaryotic cells. *Cell Host Microbe* 15, 600–610. doi: 10.1016/j.chom.2014.04.010

Koskella, B. (2013). Phage-mediated selection on microbiota of a long-lived host. *Curr. Biol.* 23, 1256–1260. doi: 10.1016/j.cub.2013.05.038

Li, X., LeBlanc, J., Truong, A., Vuthoori, R., Chen, S. S., Lustgarten, J. L., et al. (2011). A metaproteomic approach to study human-microbial ecosystems at the mucosal luminal interface. *PLoS ONE* 6:e26542. doi: 10.1371/journal.pone.0026542

Lukeš, J., Archibald, J. M., Keeling, P. J., Doolittle, W. F., and Gray, M. W. (2011). How a neutral evolutionary ratchet can build cellular complexity. *IUBMB Life* 63, 528–537. doi: 10.1002/iub.489

Machado, A. M. D., and Sommer, M. O. A. (2014). Human intestinal cells modulate conjugational transfer of multidrug resistance plasmids between clinical Escherichia coli isolates. *PLoS ONE* 9:e100739. doi: 10.1371/journal.pone.0100739

Mäkivuokko, 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, 227–234. doi: 10.1017/S0007114509991553

Mariat, D., Firmesse, O., Levenez, F., Guimarães, V., Sokol, H., Doré, J., et al. (2009). The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age. *BMC Microbiol.* 9:123. doi: 10.1186/1471-2180-9-123

O'Sullivan, O., Coakley, M., Lakshminarayanan, B., Conde, S., Claesson, M. J., Cusack, S., et al. (2013). Alterations in intestinal microbiota of elderly Irish subjects post-antibiotic therapy. *J. Antimicrob. Chemother.* 68, 214–221. doi: 10.1093/jac/dks348

O'Toole, P. W., and Claesson, M. J. (2010). Gut microbiota: changes throughout the lifespan from infancy to elderly. *Int. Dairy J.* 20, 281–291. doi: 10.1016/j.idairyj.2009.11.010

Ottman, N., Smidt, H., de Vos, W. M., and Belzer, C. (2012). The function of our microbiota: who is out there and what do they do? *Front. Cell. Infect. Microbiol.* 2:104. doi: 10.3389/fcimb.2012.00104

Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214. doi: 10.1038/nature11234

Rajilić-Stojanović, M., Heilig, H. G. H. J., Molenaar, D., Kajander, K., Surakka, A., Smidt, H., et al. (2009). Development and application of the human intestinal tract chip, a phylogenetic microarray: analysis of universally conserved phylotypes in the abundant microbiota of young and elderly adults. *Environ. Microbiol.* 11, 1736–1751. doi: 10.1111/j.1462-2920.2009.01900.x

Rampelli, S., Candela, M., Turroni, S., Biagi, E., Collino, S., Franceschi, C., et al. (2013). Functional metagenomic profiling of intestinal microbiome in extreme ageing. *Aging (Albany. NY).* 5, 902–912. Available online at: http://www.impactaging.com/papers/v5/n12/full/100623.html

Rea, M. C., O'Sullivan, O., Shanahan, F., O'Toole, P. W., Stanton, C., Ross, R. P., et al. (2012). *Clostridium difficile* carriage in elderly subjects and associated changes in the intestinal microbiota. *J. Clin. Microbiol.* 50, 867–875. doi: 10.1128/JCM.05176-11

Reading, N. C., and Sperandio, V. (2006). Quorum sensing: the many languages of bacteria. *FEMS Microbiol. Lett.* 254, 1–11. doi: 10.1111/j.1574-6968.2005.00001.x

Reyes, A., Semenkovich, N. P., Whiteson, K., Rohwer, F., and Gordon, J. I. (2012). Going viral: next generation sequencing applied to phage populations in the human gut. *Nat. Rev. Microbiol.* 10, 607–617. doi: 10.1038/nrmicro2853

Russell, A. B., LeRoux, M., Hathazi, K., Agnello, D. M., Ishikawa, T., Wiggins, P. A., et al. (2013). Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. *Nature* 496, 508–512. doi: 10.1038/nature12074

Sancar, A. (2008). The intelligent clock and the Rube Goldberg clock. *Nat. Struct. Mol. Biol.* 15, 23–24. doi: 10.1038/nsmb0108-23

Sekirov, I., Russell, S. L., Antunes, L. C. M., and Finlay, B. B. (2010). Gut microbiota in health and disease. *Physiol. Rev.* 90, 859–904. doi: 10.1152/physrev.00045.2009

Sperandio, V., Torres, A. G., Jarvis, B., Nataro, J. P., and Kaper, J. B. (2003). Bacteria-host communication: the language of hormones. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8951–8956. doi: 10.1073/pnas.1537100100

Tiihonen, K., Ouwehand, A. C., and Rautonen, N. (2010). Human intestinal microbiota and healthy ageing. *Ageing Res. Rev.* 9, 107–116. doi: 10.1016/j.arr.2009.10.004

Tiihonen, K., Tynkkynen, S., Ouwehand, A., Ahlroos, T., and Rautonen, N. (2008). The effect of ageing with and without non-steroidal anti-inflammatory drugs on gastrointestinal microbiology and immunology. *Br. J. Nutr.* 100, 130–137. doi: 10.1017/S000711450888871X

van Tongeren, S. P., Slaets, J. P. J., Harmsen, H. J. M., and Welling, G. W. (2005). Fecal microbiota composition and frailty. *Appl. Environ. Microbiol.* 71, 6438–6442. doi: 10.1128/AEM.71.10.6438-6442.2005

Vrieze, A., de Groot, P. F., Kootte, R. S., Knaapen, M., van Nood, E., and Nieuwdorp, M. (2013). Fecal transplant: a safe and sustainable clinical therapy for restoring intestinal microbial balance in human disease? *Best Pract. Res. Clin. Gastroenterol.* 27, 127–137. doi: 10.1016/j.bpg.2013.03.003

Woodmansey, E. J. (2007). Intestinal bacteria and ageing. *J. Appl. Microbiol.* 102, 1178–1186. doi: 10.1111/j.1365-2672.2007.03400.x

Woodmansey, E. J., McMurdo, M. E. T., MacFarlane, G. T., and MacFarlane, S. (2004). Comparison of compositions and metabolic activities of fecal microbiotas in young adults and in antibiotic-treated and non-antibiotic-treated elderly subjects. *Appl. Environ. Microbiol.* 70, 6113–6122. doi: 10.1128/AEM.70.10.6113-6122.2004

Wu, G. D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.-Y., Keilbaugh, S. A., et al. (2011). Linking long-term dietary patterns with gut microbial enterotypes. *Science* 334, 105–108. doi: 10.1126/science.1208344

Yeoman, C. J., Chia, N., Yildirim, S., Miller, M. E. B., Kent, A., Stumpf, R., et al. (2011). Towards an evolutionary model of animal-associated microbiomes. *Entropy* 13, 570–594. doi: 10.3390/e13030570

Zwielehner, J., Liszt, K., Handschur, M., Lassl, C., Lapin, A., and Haslberger, A. G. (2009). Combined PCR-DGGE fingerprinting and quantitative-PCR indicates shifts in fecal population sizes and diversity of Bacteroides, bifidobacteria and Clostridium cluster IV in institutionalized elderly. *Exp. Gerontol.* 44, 440–446. doi: 10.1016/j.exger.2009.04.002

Immune modulation of the brain-gut-microbe axis

Sahar El Aidy [1,2]*, Timothy G. Dinan [1,3] and John F. Cryan [1,4]

[1] Laboratory of Neurogastroenterology, Alimentary Pharmabiotic Centre, University College Cork, Cork, Ireland
[2] Department of Industrial Biotechnology, Genetic Engineering and Biotechnology Research Institute, Sadat City University, Sadat City, Egypt
[3] Department of Psychiatry, University College Cork, Cork, Ireland
[4] Department of Anatomy and Neuroscience, University College Cork, Cork, Ireland
*Correspondence: s.elaidy@ucc.ie

Only recently have we fully appreciated that the classically separated domains of neurology, endocrinology, immunology and microbiology, with their various organs- the brain, glands, gut, immune cells and microbiota, could actually be joined to each other in a multidirectional network of communication, in order to maintain homeostasis. For example, local and systemic immune activation have profound neural and behavioral effects (Campos-Rodríguez et al., 2013), neuroendocrine hormones regulate immune cytokines, and together, the nervous system and immune system work together in synergy to protect the body from infection (Steinman, 2004). Analogously, the gut microbes greatly impact the host immunological, psychological, and overall well-being of the host (Collins and Bercik, 2013; El Aidy and Kleerebezem, 2013a; Wang and Kasper, 2013; Moloney et al., 2014). However, definitive mechanisms that orchestrate a functionally relevant communication within this network, in particular, during the early life development, are yet to be elucidated. A potential unifying mechanism very likely involves multiple-functioning molecules and their receptors, as they are produced by, act upon and move from one system to another linking the brain, gut, immune system, and microbiota. These messenger molecules include (among others) neurotransmitters, neuropeptides, endocrine hormones, and cytokines.

The physiological phenomenon of maturation of the immune and neurological systems, as well as the microbial colonization, initiated within the fetal period, are dynamic in their character and are expanding in time through the first months and even years of human's life. The mode of delivery, be it vaginal birth or caesarean-section, has recently been shown to be critical in determining the pioneer microbial composition of neonates (Dominguez-Bello et al., 2010). Particularly over the first few years of life, as the microbiota develops, there is a greater potential for disruption of the long-term microbial state upon the repeated use of antibiotics (Lemon et al., 2012). Not only does the antibiotics use disrupt the microbial community in healthy infants but also it amplifies the microbial dysbiosis in pediatric patients with Crohn's disease (Gevers et al., 2014). Several reports have shown that dysbiosis alone as may result from antibiotic treatment is sufficient to drive intestinal inflammation (Hooper et al., 2012). Additionally, alterations of the microbial composition are often associated with changes in brain development and plasticity and alterations in motor, anxiety and social behavior (Sudo et al., 2004; Diaz-Heijtz et al., 2011; Neufeld et al., 2011; Clarke et al., 2013; Desbonnet et al., 2014). Consequently, abrupt shifts during the infant's unique developmental path through this early unstable phase may have longer term health implications (Costello et al., 2012). Activation of the (innate) immune response during the primary colonization involves the induction of toll-like receptors (TLRs) (Carvalho et al., 2012; El Aidy et al., 2012a, 2013b) and is linked to the stress induced by colonization, which increases gut permeability (Dinan and Cryan, 2012) to the colonizing microbes and their metabolites. Although very crucial in the recognition of the gut microbes, the scenario of immune activation via TLRs appear not to respond promptly at the initial stage of colonization (El Aidy et al., 2013c), suggesting that the very rapid actions of neurotransmitters and hormones might be important to control the priming and migration of cells of the first line of defense. Several members of pioneer gut colonizers are able to produce neurotransmitters. *Escherichia* and *Streptococcus*, for instance, can produce norepinephrine and serotonin (5-HT), whereas, *Lactobacillus* and *Bifidobacterium* can produce GABA and acetylcholine (Roshchina, 2010; Lyte, 2011). Mobile cells of the immune system express receptors for neurotransmitters (Pert et al., 1985). For example, migration of immature dendritic cells (DCs) to lymph nodes is mediated via $\alpha 1$ adrenergic receptors (Maestroni, 2000), emphasizing the early effects of the sympathetic nervous system (SNS) at the start of a local immune response. Norepinephrine, a neurotransmitter of the SNS (among other molecules stored in sympathetic vesicles), has pro-inflammatory effects at low concentration mediated through binding to $\alpha 2$ adrenoceptors and reduction of cAMP levels (Spengler et al., 1990). On the other hand, acetylcholine, the principle vagal neurotransmitter, attenuates the release of cytokines [tumor necrosis factor-alpha (TNF- α), interleukin (Il)-1β, Il-6 and Il-18, but not the anti-inflammatory cytokine Il-10, in lipopolysaccharides (LPS) stimulated human macrophage cultures (Borovikova et al., 2000)]. Moreover, choline acetyltransferase (ChAT), the key enzyme in the synthesis of acetylcholine, is expressed by B cells, DCs, and macrophages in the mucosal-associated lymphoid tissue (MALT). Reardon et al.

reported that ChAT expression begins after microbial colonization, following birth, and requires MyD88-dependent signaling derived from the intestinal microbiota (Reardon et al., 2013). Monocytes such as macrophages and other leucocytes travel through the blood and when they come within scenting distance of a given neurotransmitter, they begin to chemotaxically orient toward it, and then communicate with other immune cells in the adaptive arm such as B and T lymphocytes to ensure well-coordinated immune response.

In parallel, activation of the nervous system, such as afferent vagus nerve fibers by cytokines stimulates neuronal anti-inflammatory responses (Sternberg, 1997). Immune cells can produce various neurotransmitters and other factors, which alert the brain to the changes that occur in the body and affect the plasticity of the local and central nervous systems, thereby can regulate mood and behavior. Leukocytes, for example, synthesize and release corticotropin (ACTH) and endorphins in response to bacterial LPS (Harbour-McMenamin et al., 1985). Mounting an innate immune response during pathogenesis and transiently, during primary gut colonization results in the production of pro-inflammatory cytokines such as Il-1β, TNF- α, and IFN-γ (El Aidy et al., 2013c). The pro-inflammatory cytokine- Il-1β, for instance, is able to inhibit the release of norepinephine from noradrenoceptor axon terminals, in the intestine via the induction of nitric oxide (Rühl et al., 1994; Rühl and Collins, 1997). To note, norepinephrine via β adrenergic signaling, inhibits many aspects of the innate and Th1 mediated immune response (Straub et al., 2006), whereas activated macrophages and other innate immune cells produce nerve repellent factors directed toward sympathetic nerve to counteract the sympathetic inhibitory effect (Miller et al., 2004). Importantly, another possible mechanism underlying the reduced levels of norepinephrine during mounting innate and Th1 mediated immune responses is the decreased L-DOPA decarboxylase activity-the enzyme that converts L-DOPA to norepinephrine, as observed in both inflamed and non-inflamed colonic mucosa of Crohn's patients (Magro et al., 2002).

One mechanism through which immune activation or immunomodulation may affect physiology and behavior is via actions on serotonergic systems (Lowry et al., 2007). Analogous to the action of gut microbiota during primary colonization, Lowry et al. found that the non-pathogenic *Mycobacterium vaccae* led to stimulation of the peripheral immune system. The Th1 and Treg but not Th2 mediated immune activation stimulated a specific subset of serotonergic neurons in the dorsal raphe nucleus of mice and increased serotonin metabolism within the ventro-medial prefrontal cortex (Lowry et al., 2007). This finding demonstrates that the type of peripheral immune response is important in determining the effects on serotonergic neurons. Additionally, the study of Lowry and co-workers suggests that the immune-responsive subpopulation of serotonergic neurons in the dorsal raphe appears to play an important role in the regulation of mood during the health state. Afferent fibers within the vagus nerve could be involved in transferring signals of peripheral immune activation to the CNS (Maier et al., 1998). The regulating mechanism appears to involve enhancement of c-Fos expression in dorsal raphe nucleus serotonergic neurons (Hollis et al., 2006). It is therefore tempting to speculate that immune activation under physiological conditions stimulates a subset of serotonergic neurons, distinct from those activated under pathological conditions or by uncontrollable stressors. Nevertheless it is unclear if this association reflects a causal or reactionary response. Immune activation induces symptoms of depression and anxiety in human patients including disorders such as irritable bowel syndrome (IBS), a disorder of the gut-brain axis (Clarke et al., 2009) and in patients receiving treatment with interferon (Felger et al., 2013). Treatment with serotonergic antidepressant drugs prevents the onset of depressive symptoms in such situations (Capuron and Miller, 2004; Capuron et al., 2004). Moreover, recent data has shown that the TNF antagonist infliximab reduces depression symptoms in a subset of patients with high baseline inflammatory biomarkers (Raison et al., 2013). Of note, local and systemic depletion of the 5-HT precursor- Tryptophan (Trp) is associated

with elevation of the immunomodulatory enzyme Indoleamine-pyrrole 2,3-dioxygenase (IDO), which occurs during immune activation (Moffett and Namboodiri, 2003) and transiently during primary gut colonization (El Aidy et al., 2012b, 2013b). Moreover, 5-HT is altered by the gut microbiota, being elevated in conventionally raised mice, only when gut colonization occurs at birth (Clarke et al., 2013). Collectively, these findings illustrate evidence that serotonergic systems may play an important role in the relationship between the immune function, gut microbiota, and psychological state.

Local and systemic elevation of pro-inflammatory cytokines, in particular, Il-1β and Il-6 and altered 5-HT levels cause activation of the hypothalamic-pituitary-adrenal (HPA) axis and production of corticotrophin releasing factor (CRF). CRF, which is also elevated in stress response, causes disturbance to other neuropeptides, leading to changes in mood and behavior (Dinan and Cryan, 2012). CRF stimulates the anterior pituitary gland to release the stress hormone; ACTH; which in turn stimulates release of cortisol from the adrenal gland. Cortisol, through a feedback loop, regulates the levels of CRF, and ACTH. Importantly, immune cells, through the COX-2 pathway and production of PEG-2 during (pro-) inflammation, stimulate the adrenal gland to produce corticosterone, which exclusively supports the β adrenergic pathway through which the SNS perform its anti-inflammatory effect (Straub et al., 2006). This indicates that under normal conditions the neuroendocrine and immune systems coordinate to ensure maintenance of homeostasis. Indeed, patients with IBS show inadequately low concentration of the anti-inflammatory steroid hormone, cortisol (Straub et al., 2002).

In conclusion, we assume that the intimate communication between the microbiota-immune-neuroendocrine systems involves multiple-functioning molecule(s). Those same molecule(s) appear to be produced by and signal member(s) of the gut microbial community, the neuroendocrine and immune systems to mount or avoid rising an attack response against the commensals. This would be of crucial importance early in life. During

this vulnerable period, it is believed to exist a narrow window during which colonization with a "healthy" microbiota exerts effects that may decrease susceptibility to diseases and ensure normal development of the mucosal and systemic immunity and metabolism as well as the development of HPA axis,which impacts on the gut through its action on the enteric nervous system, immune system and the CNS (Sudo et al., 2004; Shreiner et al., 2008). The most important question needed to be answered at this point, what would be the molecular mechanisms underlying the intimate cross-talk between the immune system and the microbiota-gut-brain axis at its various nodes of interaction?

REFERENCES

Borovikova, L. V., Ivanova, S., Zhang, M., Yang, H., Botchkina, G. I., Watkins, L. R., et al. (2000). Vagus nerve stimulation attenuates the systemic inflammatory response to endotoxin. *Nature* 405, 458–462. doi: 10.1038/35013070

Campos-Rodríguez, R., Godínez-Victoria, M., Abarca-Rojano, E., Pacheco-Yépez, J., Reyna-Garfias, H., Barbosa-Cabrera, R. E., et al. (2013). Stress modulates intestinal secretory immunoglobulin A. *Front. Integr. Neurosci.* 7:86. doi: 10.3389/fnint.2013.00086

Capuron, L., and Miller, A. H. (2004). Cytokines and psychopathology: lessons from interferon-alpha. *Biol. Psychiatry* 56, 819–824. doi: 10.1016/j.biopsych.2004.02.009

Capuron, L., Ravaud, A., Miller, A. H., and Dantzer, R. (2004). Baseline mood and psychosocial characteristics of patients developing depressive sy.mptoms during interleukin-2 and/or interferon-alpha cancer therapy. *Brain Behav. Immun.* 18, 205–213. doi: 10.1016/j.bbi.2003.11.004

Carvalho, F. A., Aitken, J. D., Vijay-Kumar, M., and Gewirtz, A. T. (2012). Toll-like receptor-gut microbiota interactions: perturb at your own risk! *Annu. Rev. Physiol.* 74, 177–198. doi: 10.1146/annurev-physiol-020911-153330

Clarke, G., Grenham, S., Scully, P., Fitzgerald, P., Moloney, R. D., Shanahan, F., et al. (2013). The microbiome-gut-brain axis during early life regulates the hippocampal serotonergic system in a sex-dependent manner. *Mol. Psychiatry* 18, 666–673. doi: 10.1038/mp.2012.77

Clarke, G., Quigley, E. M., Cryan, J. F., and Dinan, T. G. (2009). Irritable bowel syndrome: towards biomarker identification. *Trends Mol. Med.* 15, 478–489. doi: 10.1016/j.molmed.2009.08.001

Collins, S. M., and Bercik, P. (2013). Gut microbiota: intestinal bacteria influence brain activity in healthy humans. *Nat. Rev. Gastroenterol. Hepatol.* 10, 326–327. doi: 10.1038/nrgastro.2013.76

Costello, E. K., Stagaman, K., Dethlefsen, L., Bohannanm, B. J., and Relman, D. A. (2012). The application of ecological theory toward an understanding of the human microbiome. *Science* 336, 1255–1262. doi: 10.1126/science.1224203

Desbonnet, L., Clarke, G., Shanahan, F., Dinan, T. G., and Cryan, J. F. (2014). Microbiota is essential for social development in the mouse. *Mol. Psychiatry* 19, 146–148. doi: 10.1038/mp.2013.65

Diaz-Heijtz, R., Wang, S., Anuar, F., Qian, Y., Björkholm, B., Samuelsson, A., et al. (2011). Normal gut microbiota modulates brain development and behavior. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3047–3052. doi: 10.1073/pnas.1010529108

Dinan, T. G., and Cryan, J. F. (2012). Regulation of the stress response by the gut microbiota: implications for psychoneuroendocrinology. *Psychoneuroendocrinology* 37, 1369–1378. doi: 10.1016/j.psyneuen.2012.03.007

Dominguez-Bello, M. G., Costello, E. K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., et al. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11971–11975. doi: 10.1073/pnas.1002601107

El Aidy, S., Derrien, M., Aardema, R., Hooiveld, G., Richards, S. E., Dane, A., et al. (2013b). Transient inflammatory-like state and microbial dysbiosis are pivotal in establishment of mucosal homeostasis during colonisation of germ-free mice. *Benef. Microbes* 5, 67–77. doi: 10.3920/BM2013.0018

El Aidy, S., and Kleerebezem, M. (2013a). Molecular signatures for the dynamic process of establishing intestinal host-microbial homeostasis: potential for disease diagnostics? *Curr. Opin. Gastroenterol.* 29, 621–627. doi: 10.1097/MOG.0b013e328365d365

El Aidy, S., Kunze, W., Bienenstock, J., and Kleerebezem, M. (2012b). The microbiota and the gut-brain axis: insights from the temporal and spatial mucosal alterations during colonisation of the germfree mouse intestine. *Benef. Microbes.* 3, 251–259. doi: 10.3920/BM2012.0042

El Aidy, S., Merrifield, C. A., Derrien, M., van Baarlen, P., Hooiveld, G., Levenez, F., et al. (2013c). The gut microbiota elicits a profound metabolic reorientation in the mouse jejunal mucosa during conventionalisation. *Gut* 62, 1306–1314. doi: 10.1136/gutjnl-2011-301955

El Aidy, S., van Baarlen, P., Derrien, M., Lindenbergh-Kortleve, D. J., Hooiveld, G., Levenez, F., et al. (2012a). Temporal and spatial interplay of microbiota and intestinal mucosa drive establishment of immune homeostasis in conventionalized mice. *Mucosal Immunol.* 5, 567–579. doi: 10.1038/mi.2012.32

Felger, J. C., Li, L., Marvar, P. J., Woolwine, B. J., Harrison, D. G., Raison, C. L., et al. (2013). Tyrosine metabolism during interferon-alpha administration: association with fatigue and CSF dopamine concentrations. *Brain Behav. Immun.* 31, 153–160. doi: 10.1016/j.bbi.2012.10.010

Gevers, D., Kugathasan, S., Lee, D. A., Va'zquez-Baeza, Y., Van Treuren, W., Ren, B., et al. (2014). The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* 15, 382–392. doi: 10.1016/j.chom.2014.02.005

Harbour-McMenamin, D., Smith, E. M., and Blalock, J. E. (1985). Bacterial lipopolysaccharide induction of leukocyte-derived corticotropin and endorphins. *Infect. Immun.* 48, 813–817.

Hollis, J. H., Evans, A. K., Bruce, K. P., Lightman, S. L., and Lowry, C. A. (2006). Lipopolysaccharide has indomethacin-sensitive actions on Fos expression in topographically organized subpopulations of serotonergic neurons. *Brain Behav. Immun.* 20, 569–577. doi: 10.1016/j.bbi.2006.01.006

Hooper, L. V., Littman, D. R., and Macpherson, A. J. (2012). Interactions between the microbiota and the immune system. *Science* 336, 1268–1273. doi: 10.1126/science.1223490

Lemon, K. P., Armitage, G. C., Relman, D. A., and Fischbach, M. A. (2012). Microbiota-targeted therapies: an ecological perspective. *Sci. Transl. Med.* 4, 137rv5. doi: 10.1126/scitranslmed.3004183

Lowry, C.A., Hollis, J. H., de Vries, A., Pan, B., Brunet, L. R., Hunt, J. R., et al. (2007). Identification of an immune-responsive mesolimbocortical serotonergic system: potential role in regulation of emotional behavior. *Neuroscience* 146, 756–772. doi: 10.1016/j.neuroscience.2007.01.067

Lyte, M. (2011). Probiotics function mechanistically as delivery vehicles for neuroactive compounds: microbial endocrinology in the design and use of probiotics. *Bioessays* 33, 574–581. doi: 10.1002/bies.201100024

Maestroni, G. J. (2000). Dendritic cell migration controlled by alpha 1b-adrenergic receptors. *J. Immunol.* 165, 6743–6747.

Magro, F., Vieira-Coelho, M. A., Fraga, S., Serrão, M. P., Veloso, F. T., Ribeiro, T., et al. (2002). Impaired synthesis or cellular storage of norepinephrine, dopamine, and 5-hydroxytryptamine in human inflammatory bowel disease. *Dig. Dis. Sci.* 47, 216–224. doi: 10.1023/A:1013256629600

Maier, S. F., Goehler, L. E., Fleshner, M., and Watkins, L. R. (1998). The role of the vagus nerve in cytokine-to-brain communication. *Ann. N.Y. Acad. Sci.* 840, 289–300. doi: 10.1111/j.1749-6632.1998.tb09569.x

Miller, L. E., Weidler, C., Falk, W., Angele, P., Schaumburger, J., Schölmerich, J., et al. (2004). Increased prevalence of semaphorin 3C, a repellent of sympathetic nerve fibers, in the synovial tissue of patients with rheumatoid arthritis. *Arthritis Rheum.* 50, 1156–1163. doi: 10.1002/art.20110

Moffett, J. R., and Namboodiri, M. A. (2003). Tryptophan and the immune response. *Immunol. Cell Biol.* 81, 247–265. doi: 10.1046/j.1440-1711.2003.t01-1-01177.x

Moloney, R. D., Desbonnet, L., Clarke, G., Dinan, T. G., and Cryan, J. F. (2014). The microbiome: stress, health and disease. *Mamm. Genome* 25, 49–74. doi: 10.1007/s00335-013-9488-5

Neufeld, K. M., Kang, N., Bienenstock, J., and Foster, J. A. (2011). Reduced anxiety-like behavior and central neurochemical change in germ-free mice. *Neurogastroenterol. Motil.* 23, 255–264, e119. doi: 10.1111/j.1365-2982.2010.01620.x

Pert, C. B., Ruff, M. R., Weber, R. J., and Herkenham, M. (1985). Neuropeptides and their receptors: a psychosomatic network. *J. Immunol.* 135, 820s–826s.

Raison, C. L., Rutherford, R. E., Woolwine, B. J., Shuo, C., Schettler, P., Drake, D. F., et al. (2013). A randomized controlled trial of the tumor necrosis factor antagonist infliximab for treatment-resistant depression: the role of baseline inflammatory biomarkers. *JAMA Psychiatry* 70, 31–41. doi: 10.1001/2013.jamapsychiatry.4

Reardon, C., Duncan, G. S., Brüstle, A., Brenner, D., Tusche, M. W., Olofsson, P. S., et al. (2013). Lymphocyte-derived ACh regulates local innate but not adaptive immunity. *Proc. Natl. Acad. Sci. U.S.A.* 110, 1410–1415. doi: 10.1073/pnas.1221655110

Roshchina, V. V. (2010). "Evolutionary considerations of neurotransmitters in microbial, plant, and animal cells," in *Microbial Endocrinology: Interkingdom Signaling in Infectious Diseases and Health*, eds M. Lyte and P. P. E. Freestone (New York, NY: Springer), 17–52. doi: 10.1007/978-1-4419-5576-0_2

Rühl, A., and Collins, S. M. (1997). Role of nitric oxide in norepinephrine release from myenteric plexus *in vitro* and in Trichinella spiralis-infected rats. *Neurogastroenterol. Motil.* 9, 33–39. doi: 10.1046/j.1365-2982.1997.d01-5.x

Rühl, A., Hurst, S., and Collins, S. M. (1994). Synergism between interleukins 1 beta and 6 on noradrenergic nerves in rat myenteric plexus. *Gastroenterology* 107, 993–1001.

Shreiner, A., Huffnagle, G. B., and Noverr, M. C. (2008). "The "microflora hypothesis" of allergic disease," in *GI Microbiota and Regulation of the Immune System*, eds G. B. Huffnagle and M. C. Noverr (Landes Bioscience and Springer Science + Business Media), 113–34. doi: 10.1007/978-0-387-09550-9_10

Spengler, R. N., Allen, R. M., Remick, D. G., Strieter, R. M., and Kunkel, S. L. (1990). Stimulation of alpha-adrenergic receptor augments the production of macrophage-derived tumor necrosis factor. *J. Immunol.* 145, 1430–1434.

Steinman, L. (2004). Elaborate interactions between the immune and nervous systems. *Nat. Immunol.* 5, 575–581. doi: 10.1038/ni1078

Sternberg, E. M. (1997). Neural-immune interactions in health and disease. *J. Clin. Invest.* 100, 2641–2647. doi: 10.1172/JCI119807

Straub, R. H., Günzler, C., Miller, L. E., Cutolo, M., Schölmerich, J., and Schill, S. (2002). Anti-inflammatory cooperativity of corticosteroids and norepinephrine in rheumatoid arthritis synovial tissue *in vivo* and *in vitro*. *FASEB J.* 16, 993–1000. doi: 10.1096/fj.02-0085com

Straub, R. H., Wiest, R., Strauch, U. G., Härle, P., and Schölmerich, J. (2006). The role of the sympathetic nervous system in intestinal inflammation. *Gut* 55, 1640–1649. doi: 10.1136/gut.2006.091322

Sudo, N., Chida, Y., Aiba, Y., Sonoda, J., Oyama, N., Yu, X. N., et al. (2004). Postnatal microbial colonization programs the hypothalamic-pituitary-adrenal system for stress response in mice. *J. Physiol.* 558, 263–275. doi: 10.1113/jphysiol.2004.063388

Wang, Y., and Kasper, L. H. (2013). The role of microbiome in central nervous system disorders. *Brain. Behav. Immun.* doi: 10.1016/j.bbi.2013.12.015. [Epub ahead of print].

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Impact of the gut microbiota on the development of obesity and type 2 diabetes mellitus

Isabel Moreno-Indias[1,2], Fernando Cardona[1,2] *, Francisco J. Tinahones[1,2] and María Isabel Queipo-Ortuño[1,2] *

[1] Unidad de Gestion Clínica de Endocrinología y Nutrición, Laboratorio del Instituto de Investigación Biomédica de Málaga (IBIMA), Hospital Universitario de Málaga (Virgen de la Victoria), Málaga, Spain
[2] Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y la Nutrición, Madrid, Spain

Obesity and its associated disorders are a major public health concern. Although obesity has been mainly related with perturbations of the balance between food intake and energy expenditure, other factors must nevertheless be considered. Recent insight suggests that an altered composition and diversity of gut microbiota could play an important role in the development of metabolic disorders. This review discusses research aimed at understanding the role of gut microbiota in the pathogenesis of obesity and type 2 diabetes mellitus (TDM2). The establishment of gut microbiota is dependent on the type of birth. With effect from this point, gut microbiota remain quite stable, although changes take place between birth and adulthood due to external influences, such as diet, disease and environment. Understand these changes is important to predict diseases and develop therapies. A new theory suggests that gut microbiota contribute to the regulation of energy homeostasis, provoking the development of an impairment in energy homeostasis and causing metabolic diseases, such as insulin resistance or TDM2. The metabolic endotoxemia, modifications in the secretion of incretins and butyrate production might explain the influence of the microbiota in these diseases.

Keywords: gut microbiota, obesity, type 2 diabetes mellitus, inflammation, LPS, SCFA

INTRODUCTION

The prevalence of obesity and its associated disorders, such as type 2 diabetes mellitus (TDM2), has increased substantially worldwide over the last decades. Recent insight suggests that an altered composition and diversity of gut microbiota could play an important role in the development of metabolic disorders. Most of the gut microorganisms reside in the large intestine, which contains an estimated 10^{11-12} bacterial concentrations per gram of content (Leser and Molbak, 2009). These gut microbiota play a number of physiological roles involving digestion, metabolism, extraction of nutrients, synthesis of vitamins, prevention against colonization by pathogens, and immunomodulation (Jumpertz et al., 2011; Purchiaroni et al., 2013). In addition to an increased energy harvest from the diet, several mechanisms, including chronic low-grade endotoxemia, regulation of biologically active fatty acid tissue composition, and the modulation of gut-derived peptide secretion, have been proposed as links between gut microbiota and obesity (Musso et al., 2010). However, the contribution of gut microbiota to obesity and diabetes in humans is unclear. This is probably for various reasons, such as the fact that the heterogeneous etiology of obesity and diabetes can be associated with different microbes, studies have involved participants of diverse ethnic origin and food habits, the large inter-individual variation in the composition of gut microbiota, and in particular the different methods that have been used to profile the microbiota in these studies (Tremaroli et al., 2012). On the other hand, the differences between gut microbiota in lean and obese individuals as well as the impact of diet in the composition of the gut microbiome are still not wholly understood. Thus, manipulation of the gut microbiome represents a novel approach to treating obesity although it is in no way a substitute for diet and exercise. This review discusses the research conducted in understanding the role of gut microbiota in the pathogenesis of obesity and TDM2.

GUT MICROBIOTA COMPOSITION

Microorganisms colonize all surfaces of the human body that are exposed to the environment, with most residing in the intestinal tract. Bacterial communities at a particular body site have more similarity among different subjects than in the same subject but at different body sites; i.e., there is more similarity between oral bacterial communities of different individuals than between the bacterial communities of the skin and the mouth in a single individual (Costello et al., 2009), although there is also considerable inter-individual variability (Costello et al., 2009; Robinson et al., 2010). The bacterial component of the microbiota has hardly been studied in recent years, driven by large-scale projects such as

the Human Microbiome Project (Turnbaugh et al., 2007; Peterson et al., 2009) and MetaHIT (Qin et al., 2010). Research about gut microbiota, mainly using ribosomal 16S RNA and whole-genome sequencing (WGS – whole-genome shotgun; Turnbaugh et al., 2009b), has provided a general view of the commensal microbial communities and their functional capacity. For instance, in 2010, a catalog was established of 3.3 million gut microbial genes (Qin et al., 2010), with another wider catalog published soon after (Human Microbiome Project Consortium, 2012a,b). These studies have shown the great variability in microbiota composition among healthy subjects, even between twins sharing less than 50% of their bacterial taxons at the species level (Turnbaugh et al., 2010). However, this does not mean genetics has no role in the establishment and conformation of the gut microbiota, and it has been demonstrated that bacterial community composition is influenced by host-specific genomic locus (Benson et al., 2010; Koenig et al., 2011). Metagenomic studies have established that in spite of the high interpersonal variability, some bacterial groups share functionalities (Turnbaugh et al., 2009b; Burke et al., 2011). The main bacterial phyla are: Firmicutes (Gram-positive), Bacteroidetes (Gram-negative), and Actinobacteria (Gram-positive). Firmicutes is found in the highest proportion (60%), with more than 200 genera, the most important of which are: *Mycoplasma*, *Bacillus*, and *Clostridium*; Bacteroidetes and Actinobacteria each comprise about 10% of the gut microbiota, with the rest belonging to over 10 minority families. In total there are more than 1000 different species in the gut. It has also been suggested that the microbiota of most individuals can be categorized into three predominant enterotypes dominated by three different genera: *Bacteroides*, *Prevotella*, and *Ruminococcus*, which are independent of age, gender, ethnicity, or body mass index (BMI; Benson et al., 2010; Arumugam et al., 2011). Nevertheless an important debate has recently started about the concept of enterotypes (Jeffery et al., 2012; Yong, 2012), with a number of studies failing to identify the three distinct categories described by Arumugam et al., 2011 (Claesson et al., 2012; Huse et al., 2012).

MICROBIOTA ESTABLISHMENT

Changes are produced in our microbiota from birth to adulthood. The fetal intestinal tract is sterile until birth, after which the newborn tract begins to be colonized. Infants are exposed to a great variety of microorganisms from different environments during and immediately after birth, either in their encounter with their maternal vagina or by the cutaneous microorganisms depending on the type of delivery (Adlerberth and Wold, 2009; Dominguez-Bello et al., 2010). Infants born vaginally have similar communities to those found in the vaginal microbiota of their mothers. In contrast, those born by Caesarian section have the characteristic microbiota of the skin, with taxons like *Staphylococcus* and *Propionibacterium* spp. (Dominguez-Bello et al., 2010). Moreover, these infants have lower intestinal bacteria counts with less diversity in the early weeks of life (Grölund et al., 1999; Axad et al., 2013). Another factor influencing the microbiota concerns the method of feeding. The microbiota of breast-fed infants is dominated by *Bifidobacterium* (Turroni et al., 2012; Yatsunenko et al., 2012) and *Ruminococcus*

(Morelli, 2008), with significantly lower rates of colonization by *Escherichia coli*, *C. difficile*, *Bacteroides fragilis*, and *Lactobacillus* than those observed in exclusively formula-fed infants (Penders et al., 2006). The microbiota of formula-fed infants is more complex and includes enterobacterial genera, *Streptococcus*, *Bacteroides*, and *Clostridium*, as well as *Bifidobacterium* and *Atopobium* (Bezirtzoglou et al., 2011). But, the composition of the microbiota changes with the introduction of solid foods and a more complex and stable community similar to the adult microbiota becomes established at 2–3 years of age (Palmer et al., 2007; Koenig et al., 2011; Ravel et al., 2011; Yatsunenko et al., 2012), with Firmicutes and Bacteroidetes predominating. During adulthood the microbiota is relatively stable until old age, when this stability is reduced (McCartney et al., 1996). The ELDERMET consortium studied the microbiota of elderly Irish subjects, finding a different characteristic microbiota composition to that of young persons, particularly in the proportions of *Bacteroides* spp. and *Clostridium* groups (Claesson et al., 2011).

EFFECT OF DIET ON THE TEMPORAL DYNAMICS OF MICROBIOTA

Human-related microorganisms have been enumerated and categorized (Costello et al., 2009) and their temporal dynamics have been described (Caporaso et al., 2011). To understand the stability of microbiota within an individual over time is an important step to predict diseases and develop therapies to correct dysbiosis (microbial community mismatches). Data from longitudinal studies show the microbiota composition is relatively stable in healthy adults over time and is only transiently altered by external disturbances such as diet, disease, and environment (Delgado et al., 2006). Particularly, changes in diet have shown important effects on the composition of the intestinal microbiota. Indeed, dietary changes could explain 57% of the total structural variation in gut microbiota whereas changes in genetics accounted for no more than 12% (Zhang et al., 2010). Diet provides nutrients for both the host and the bacteria of the gastrointestinal tract. Changes in the composition of the gut microbiota in response to dietary intake take place because different bacterial species are better equipped genetically to utilize different substrates (Scott et al., 2008). Many studies have demonstrated that an increase in fat intake produces an increase in the Gram-negative/Gram-positive index of our microbiota. Recent studies have found that mice [humanized germ-free (GF)] changed from a diet low in fat and rich in vegetable polysaccharides to a diet rich in fat and sugar and low in plant polysaccharides (western diet) changed their microbiota in just 1 day. Mice on the "western diet" experienced an increase in the abundance of bacteria of the phylum Firmicutes and a decrease in the abundance of those of the phylum Bacteroidetes (Turnbaugh et al., 2009a,b). Hildebrandt et al. (2009) also found important changes in the abundance of the gut microbiota of mice after changing from a standard chow to a high-fat diet, which was associated with a decrease in the abundance of bacteria of the phylum Bacteroidetes and an increase in that of both Firmicutes and Proteobacteria phyla. Moreover, murine studies have shown that carbohydrate-reduced diets result in enriched populations of bacteria from the Bacteroidetes phyla (Walker et al., 2011) while

calorie-restricted diets prevent the growth of *C. coccoides*, *Lactobacillus* spp., and *Bifidobacteria* spp., which are all major butyrate producers required for colonocyte homeostasis (Santacruz et al., 2009). Only a limited number of human clinical trials have assessed the effects of changes in dietary patterns on the intestinal microbiota (De Palma et al., 2009; Muegge et al., 2011; Walker et al., 2011). In a controlled-feeding study with humans consuming a high-fat/low-fiber or low-fat/high-fiber diet, notable changes were found in gut microbiota in just 24 h, highlighting the rapid effect that diet can have on the intestinal microbiota (Wu et al., 2011). Interestingly, De Filippo et al. (2010) found that European children have a microbiota depleted of Bacteroidetes and enriched in Enterobacteriaceae compared to rural African children, which the authors attributed to low dietary fiber intake by Europeans (Wu et al., 2011). These authors postulated that gut microbiota co-evolved with the plant-rich diet of the African children, allowing them to maximize energy extraction from dietary fiber while also protecting them from inflammation and non-infectious intestinal diseases (De Filippo et al., 2010). Another study demonstrated that subjects consuming a vegan or vegetarian diet had a lower stool pH and significantly lower total counts of culturable *Bacteroides* spp., *Bifidobacterium* spp., *E. coli*, and Enterobacteriaceae spp. than controls (Zimmer et al., 2011). A vegetarian diet has also been shown to decrease the amount and change the diversity of *Clostridium* cluster IV and *Clostridium* clusters XIV and XVII (Liszt et al., 2009). However, large well-controlled trials are needed to elucidate the mechanisms that link dietary changes to alterations in microbial composition as well as the implications of key population changes for health and disease.

MODULATION OF GUT MICROBIOTA DIVERSITY BY ANTIBIOTICS

Much evidence now exists concerning an important change in our microbiota over recent decades, with some species increasing and others decreasing, though one of the most striking findings is that in developed countries there is a loss in the diversity of our microbiota. One of the most important factors that can disturb microbiota composition is the increased use of antibiotic treatment. There is evidence of important alterations in microbiota after antibiotic treatment (Sullivan et al., 2001; Jernberg et al., 2007; Dethlefsen et al., 2008). Although affected taxons vary among subjects, some taxons are not recovered even several months after treatment, and in general, there is a long-term reduction in bacterial diversity after the use of antibiotics (Jernberg et al., 2010; Dethlefsen and Relman, 2011). A correlation has recently been proposed between the increasing global use of antibiotics and weight gain or obesity in humans (Thuny et al., 2010). Several studies have indicated that some antibiotics are associated with weight gain in malnourished children, neonates, and adults (Ajslev et al., 2011; Trehan et al., 2013), but the precise mechanisms by which antibiotics improve weight are not well characterized. It has been suggested that antibiotics, such as avoparcin (a glycopeptide structurally related to vancomycin), exert selective pressure on Gram-positive bacteria and that gut colonization by *Lactobacillus* spp., which are known to be resistant to glycopeptides, used as a growth promoter in animals and found at a high concentration in the feces of obese patients, could be responsible for the weight gain observed in patients who had been treated with vancomycin. These data suggest that nutritional programs and follow-up of weight should be undertaken in patients under such treatment (Thuny et al., 2010). Other recent studies have also demonstrated the beneficial effects of antibiotics on metabolic abnormalities in obese mice, giving rise to reduced glucose intolerance, body weight gain, metabolic endotoxemia, and markers of inflammation and oxidative stress (Bech-Nielsen et al., 2012). Moreover, these effects were associated with a reduced diversity of gut microbiota (Murphy et al., 2013). Antibiotic treatment combined with a protective hydrolyzed casein diet has been found to decrease the incidence and delay the onset of diabetes in a rat model (Brugman et al., 2006). A recent study also reported that antibiotic-treated humans showed greater and less balanced sugar anabolic capabilities than non-treated individuals (Hernandez et al., 2013). However, the majority of clinical studies are focused primarily on the characterization of the composition and diversity of gut microbes, it remaining uncertain whether antibiotic-induced gut microbiota alteration in human subjects with metabolic disorders is associated with improvements in metabolic derangements as observed in animal studies.

ROLE OF GUT MICROBIOTA IN METABOLIC DISEASES

Recent decades have seen an increase in the prevalence of metabolic diseases in developed countries. Environmental factors, such as the increase in energy intake and the decrease in physical activity, have been considered causes of this spectacular increase in the prevalence of metabolic diseases. However, even when the energy intake does not increase and physical activity does not decrease, the prevalence continues growing exponentially, so other environmental factors must be taken into account, including changes in gut microbiota. One of the challenges is to elucidate the molecular origin of metabolic diseases, though the great diversity and social differences among humans make this difficult. During the last half century, with the advances in molecular biology, researchers have been investigating the genetics of metabolic diseases. In spite of the great efforts and the identification of some mutations in the genome, no global view has yet been established. The discovery of candidate genes in studies of pangenomic associations (GWAS – genome-wide association studies) has helped to identify new genes associated with sensitivity/resistance to diabetes and extreme metabolic phenotypes (Jacquemont et al., 2011). However, the global diversity of metabolic diseases cannot be explained, especially given the studies in monozygotic twins, discordant for TDM2 and obesity (Medici et al., 1999; Beck-Nielsen et al., 2003).

A second step toward the comprehension of the origin of metabolic diseases involves epigenetic and environmental factors. A drastic change in feeding habits in which dietary fiber has been replaced by a high fat diet contributes to the origin of metabolic diseases. However, this simple concept cannot explain why some people are sensitive and others are resistant to the development of these metabolic diseases. In mice, a metabolic adaptation is frequently observed (Burcelin et al., 2002). Genetically identical mice in the same box and with a fat-rich diet for 6–9 months can develop both obesity and diabetes, or only one of the diseases. There is a need to find a new paradigm that takes into account

the genetic diversity, the environmental factor impact, the rapid development of metabolic diseases, and the individual behavior to develop diabetes and obesity. The conclusion reached concerns the concept of personalized medicine in which the individual characteristics should be identified in order to adapt a suitable therapeutic strategy for small patient groups.

INFLUENCE OF GUT MICROBIOTA COMPOSITION IN THE DEVELOPMENT OF OBESITY

Studies during the last decade have associated the gut microbiota with the development of metabolic disorders, especially diabetes and obesity. Although incompletely understood, the gut microbiota is implicated in the programing and control of many physiological functions, including gut epithelial development, blood circulation, innate and adaptive mechanisms (Mackie et al., 1999; Dethlefsen et al., 2006). A new theory shows microbiota as a contributor to the regulation of energy homeostasis. Thus, with the environmental vulnerabilities, gut microbiota could provoke the development of impairment in energy homeostasis, causing metabolic diseases.

The first discovery was related to the fact that mice with a mutation in the leptin gene (metabolically obese mice) have different microbiota as compared with other mice without the mutation (Ley et al., 2005). In this obese animal model, the proportion of the dominant gut phyla, Bacteroidetes and Firmicutes, is modified with a significant reduction in Bacteroidetes and a corresponding increase in Firmicutes (Ley, 2010). Ley et al. (2006) were the first to report an altered gut microbiota similar to that found in obese mice (a larger proportion of Firmicutes and relatively fewer Bacteroidetes) in 12 obese subjects compared with 2 lean controls. Later, Armougom et al. (2009) confirmed a reduction in Bacteroidetes accompanied by a rise in *Lactobacillus* species belonging to the Firmicutes phylum. Turnbaugh et al. (2009b) and Furet et al. (2010) showed a different pattern based on a lower representation of Bacteroidetes (*Bacteroides/Prevotella*) in obese individuals with no differences in Firmicutes phylum. Collado et al. (2008) reported increases in species belonging to both Firmicutes (*Staphylococcus aureus*) and Bacteroidetes (*Bacteroides/Prevotella*) in overweight women. Million et al. (2012) described changes in the composition of Firmicutes based on an increase in *Lactobacillus reuteri* coupled with a reduction in *L. paracasei* and *L. plantarum*. Finally, other studies have found no differences between Firmicutes and Bacteroidetes at the phylum level (Duncan et al., 2008; Mai et al., 2009; Jumpertz et al., 2011).

The shift in the relative abundance observed in these phyla is associated with the increased capacity to harvest energy from food and with increased low-grade inflammation. The increase in Firmicutes and the decrease in the proportion of Bacteroidetes observed in obese mice could be related with the presence of genes encoding enzymes that break down polysaccharides that cannot be digested by the host, increasing the production of monosaccharides and short-chain fatty acids (SCFA) and the conversion of these SCFA to triglycerides in the liver (**Figure 1**). These SCFAs are able to bind and activate two G-protein-coupled receptors (GPR41 and GPR43) of the gut epithelial cells. The activation of these receptors induces peptide YY secretion, which suppresses gut

motility and retards intestinal transit. By this mechanism of SCFA-linked G-protein-coupled receptor activation, the gut microbiota may contribute markedly to increased nutrient uptake and deposition, contributing to the development of metabolic disorders (Erejuwa et al., 2014). Moreover, gut microbiota have also been shown to decrease the production of the fasting-induced adipose factor [FIAF; a secreted lipoprotein lipase (LPL)] by the intestinal cells, which inhibits LPL activity, increasing the storage of liver-derived triglycerides (Backhed et al., 2007).

Turnbaugh et al. (2006), in a study using ob/ob mice, found a reduced calorie content in the feces of obese mice as compared with lean mice. Other studies have suggested that obese subjects might be able to extract more energy from nutrients due to hydrogen transfer between taxa. In fact, a simultaneous increase in both hydrogen-producing Prevotellaceae and hydrogen-utilizing methanogenic Archaea has been previously associated with obesity by Zhang et al. (2009), suggesting a higher energy harvest in obese patients. For instance, intestinal starch digestion produces hydrogen, the increase of which inhibits digestion and methanogenic Archaea are able to transform this hydrogen into methane (**Figure 1**). Thus, there is a specific microbiota that obtains more energy from the same energy intake (Turnbaugh et al., 2009a). These findings agree with the observation in which GF mice fed with a fat-rich diet gained less weight than conventional mice (Backhed et al., 2004).

The most relevant experiment dealing with the causality between microbiota and obesity was done by Turnbaugh et al. (2006). In this study, they demonstrated that microbiota transplantation from genetically obese mice to axenic mice provokes a very significant weight increase compared with the axenic mice transplanted with the microbiota from lean mice.

Surprisingly, the phenotype with increase capacity for energy harvest is simply transmitted by transplantation of the obesity-associated gut microbiota in to healthy and lean donors (Turnbaugh et al., 2006, 2008). But within a phylum, not all the genera have the same role, so that bacterial genera have been related with either beneficial or harmful characteristics associated within the same phylum. Kalliomäki et al. (2008) undertook a prospective study in which they followed 49 children from birth to 7 years of age. Stool was collected at 6 and 12 months of life and it was found that the children who were 7 years old with a normal weight had a higher number of *Bifidobacterium* spp. and a smaller number of *Staphylococcus aureus* than the children who became overweight several years later. The authors concluded that the alteration in the microbiota precedes the alteration in weight, an explanation that is relevant for obesity prevention. The authors also proposed that *Staphylococcus aureus* may act as a trigger of low-grade inflammation, contributing to the development of obesity (Kalliomäki et al., 2008).

On the other hand, *Lactobacillus* spp. and bifidobacteria represent a major bacterial population of the small intestine where lipids and simple carbohydrates are absorbed, especially in the duodenum and jejunum. Recent publications reveal that the *Bifidobacteria* and *Lactobacillus* are not all the same and they may have different characteristics according to the species. For example, within the genus *Lactobacillus*, *L. plantarum*, and *L. paracasei* have been associated with leanness whereas *L. reuteri*
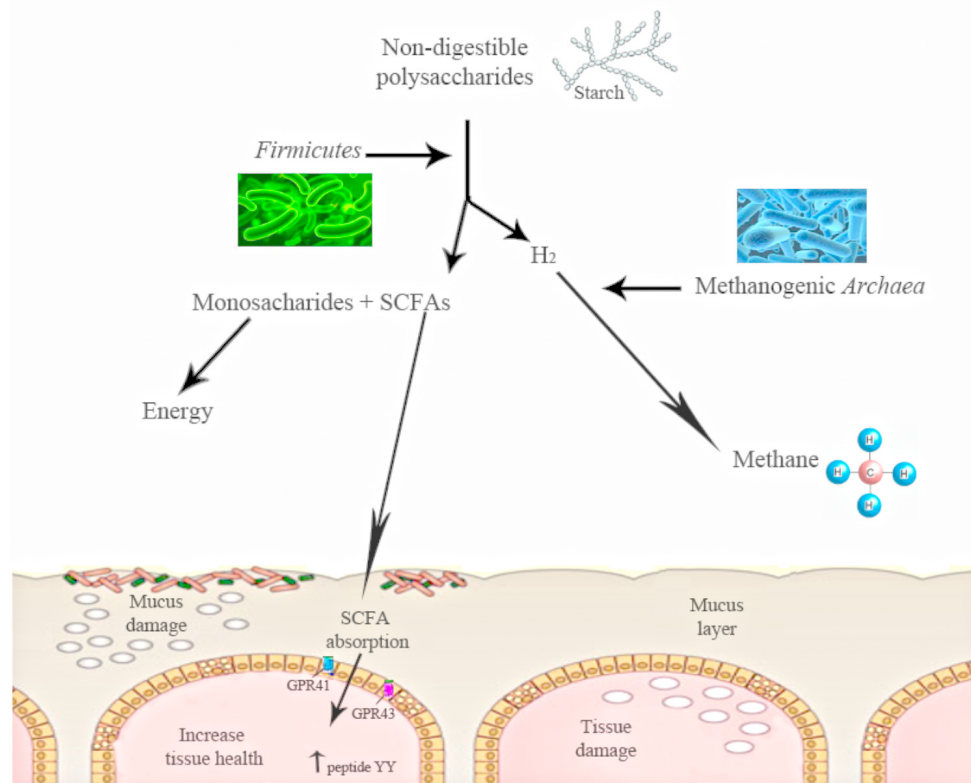
FIGURE 1 | The action of gut microbiota is needed to digest some polysaccharides. Gut microbiota converts polysaccharides into monosaccharides and short-chain fatty acids (SCFA). These SCFAs are able to bind and activate two G-protein-coupled receptors (GPR41 and GPR43) of the gut epithelial cells. The activation of these receptors induces peptide YY. Starch digestion is an example of this process: H_2 is produced, and its increase inhibits starch digestion, moment at which other bacterial groups work and transform the H_2 into methane.

is associated with obesity (Million et al., 2012). Moreover, Drissi et al. (2014) have shown that weight gain-associated *Lactobacillus* spp. encode more bacteriocins and appear to lack enzymes involved in the catabolism of fructose, a defense against oxidative stress and the synthesis of dextrin, L-rhamnose and acetate than weight protection-associated *Lactobacillus* spp., which encodes for a significant gene amount of glucose permease. Regarding lipid metabolism, thiolases were only encoded in the genome of weight gain-associated *Lactobacillus* spp. The results of this study revealed that weight protection-associated *Lactobacillus* spp. have developed defense mechanisms for enhanced glycolysis and defense against oxidative stress while weight gain-associated *Lactobacillus* spp. possess a limited ability to break down fructose or glucose and might reduce ileal brake effects (Drissi et al., 2014).

MICROBIOTA AND ITS RELATIONSHIP WITH TYPE 2 DIABETES MELLITUS

Type 2 diabetes mellitus is the consequence of an increase in the production of glucose in the liver and a deficit in the secretion and action of insulin. Other physiological functions are altered, such as the central and autonomous nervous systems, leading to an impaired secretion of hormones like glucagon and incretins. However, a common feature of obesity and TDM2 is the presence
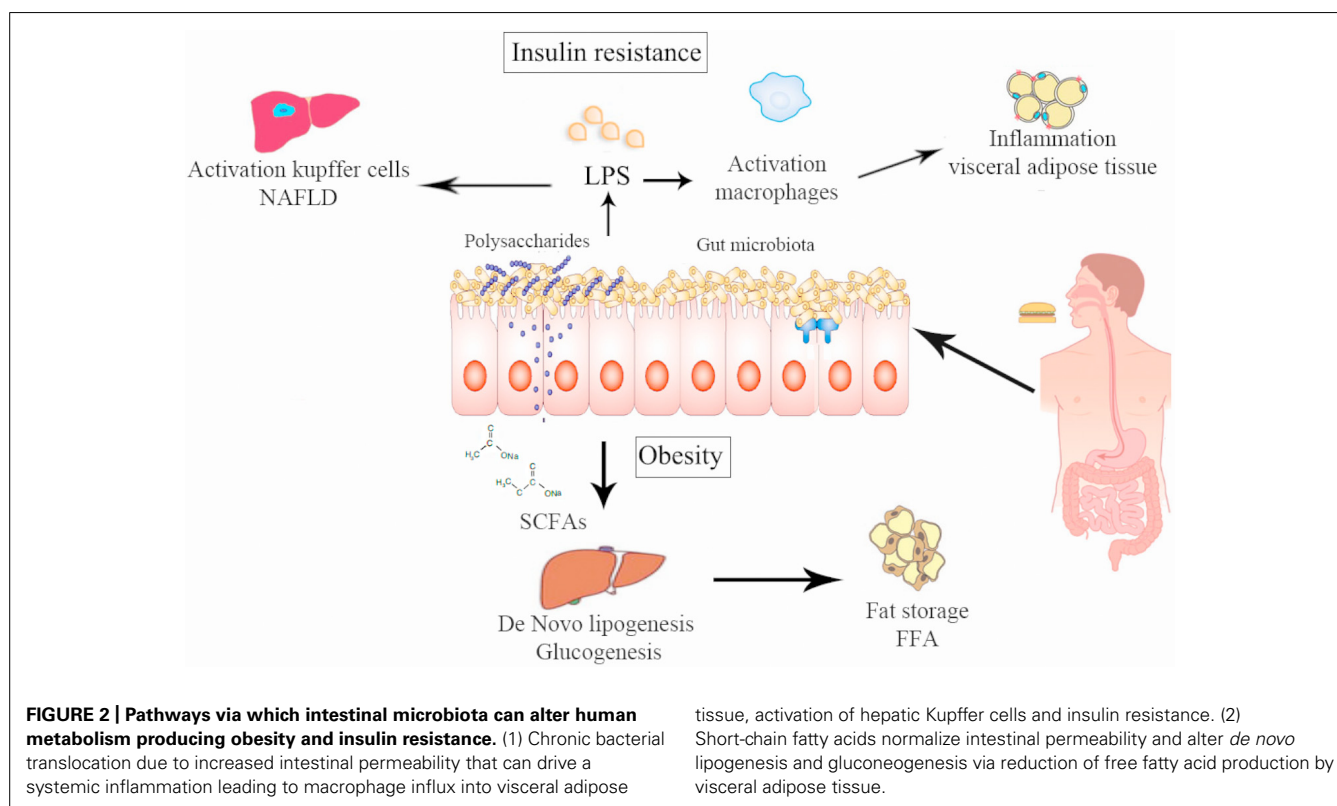
of a low-grade inflammatory component described in tissues involved in metabolism regulation, such as the liver, adipose tissue, and muscles (Pickup and Crook, 1998). This metabolic inflammation is characterized by a moderate excess in cytokine production, including interleukin (IL)-6, IL-1, or tumor necrosis factor alpha (TNF-α), that injures cellular insulin signals and contributes to insulin resistance and diabetes (Hotamisligil, 2006; Shoelson et al., 2006). Weight increase would be an initiating factor of low-grade inflammation. When adipocyte hypertrophy is produced as a response to excess energy intake, an increase in TNF-α production in the adipose tissue is also produced and this stimulates the production of chemotactic factors resulting in adipose tissue being infiltrated by proinflammatory macrophages that produce an increase in the production of IL-6 and IL-1. Recently, two studies have shown that the intestinal microbiome might be an important contributor to the development of TDM2. Both studies also showed that TDM2 subjects were characterized by a reduction in the number of Clostridiales bacteria (*Roseburia* species and *Faecalibacterium prausnitzii*), which produce the SCFA butyrate (Qin et al., 2012; Karlsson et al., 2013). Also, another study found microbiota changes in patients with diabetes or insulin resistance as compared with subjects without alterations in carbohydrate metabolism (Serino et al., 2013). In addition, changes in the amount of *Bifidobacterium*, *Lactobacillus*,

and *Clostridium* as well as a reduced Firmicutes to Bacteroidetes ratio in gut microbiota have also been recently reported in type 1 diabetic children. This study also showed that bacteria involved in the maintenance of gut integrity were significantly lower in diabetic patients than in healthy controls (Murri et al., 2013). Similar changes in the composition of intestinal microbiota have also been reported in TDM2 patients (Larsen et al., 2010; Qin et al., 2012). Several other studies linking the gut microbiota to metabolic disorders, such as obesity, insulin resistance and diabetes mellitus, have been reviewed by other authors (Caricilli and Saad, 2013; Stachowicz and Kiersztan, 2013; Tagliabue and Elli, 2013). Moreover, probiotic (Amar et al., 2011) and prebiotic treatments (Cani et al., 2007b) control gut microbiota and metabolic diseases.

Various mechanisms have been proposed to explain the influence of the microbiota on insulin resistance and TDM2, such as metabolic endotoxemia, modifications in the secretion of the incretins and butyrate production.

The lipopolysaccharides (LPS) are endotoxins commonly found in the outer membrane of Gram-negative bacteria that cause metabolic endotoxemia, which is characterized by the release of proinflammatory molecules (Manco et al., 2010). A rise in LPS levels has been observed in subjects who increased their fat intake (Amar et al., 2008). Similar results were found in mice (Cani et al., 2007b) and in mutant mice (like the leptin-deficient mice) even feeding with a normal diet (Cani et al., 2008), which suggests that a change in the proportion of Gram-negative bacteria in the gut or a change in the gut permeability were produced by the LPS rise in serum (Cani et al., 2008, 2009b) and this increase is

directly related with the degree of insulin resistance. Cani et al. (2007a,b) reported that modulation of the intestinal microbiota by using prebiotics in obese mice acts favorably on the intestinal barrier, lowering the high-fat diet-induced LPS endotoxemia and systemic and liver inflammation (**Figure 2**). LPS are absorbed by enterocytes and they are conveyed into plasma coupled to chylomicrons (Clemente Postigo et al., 2012). In this way, dietary fats can be associated with increased absorption of LPS which in turn can be related with changes in the gut microbiota distinguished by a decrease in the *Eubacterium rectale–C. coccoides* group, Gram-negative *Bacteroides* and in *Bifidobacterium* (Caricilli and Saad, 2013). This causal role of LPS was demonstrated by infusing LPS in mice with a normal diet inducing hepatic insulin resistance, glucose intolerance, and an increase in the weight of adipose tissue (Cani et al., 2007a). It has been recently shown that the LPS-induced signaling cascade via Toll-like receptor 4 (TLR4) impairs pancreatic β-cell function via suppressed glucose-induced insulin secretion and decreased mRNA expression of pancreas-duodenum homebox-1 (PDX-1; Rodes et al., 2013). LPS binds to the CD14/TLR4 receptor present on macrophages and produces an increase in the production of proinflammatory molecules. When LPS injections were administrated to mice with a genetic absence of the CD14/TLR4 receptor they did not develop these metabolic characteristics and there was no start of TDM2 or obesity, showing the important role of LPS in the mechanism of CD14/TLR4. Moreover, knockout CD14/TLR4 mice were even more sensitive to insulin than wild type controls (Cani et al., 2007a; Poggi et al., 2007). LPS can also promote the expression of NF-κB (nuclear factor kappa-light-chain-enhancer of activated

FIGURE 2 | Pathways via which intestinal microbiota can alter human metabolism producing obesity and insulin resistance. (1) Chronic bacterial translocation due to increased intestinal permeability that can drive a systemic inflammation leading to macrophage influx into visceral adipose tissue, activation of hepatic Kupffer cells and insulin resistance. (2) Short-chain fatty acids normalize intestinal permeability and alter *de novo* lipogenesis and gluconeogenesis via reduction of free fatty acid production by visceral adipose tissue.

B cells) and activation of the MAPK (mitogen-activated protein kinase) pathway in adipocytes with several target genes (Chung et al., 2006).

An increase of *Bifidobacterium* spp. modulates inflammation in obese mice by an increase in the production of incretins like the glucagon-like peptide (GLP), also reducing intestinal permeability (Cani et al., 2009b). There is evidence that the rise in *Bifidobacterium* spp. produced by some prebiotics is accompanied by an increase in GLP1 and YY peptide secretions by the intestine. These two molecules have favorable effects, decreasing insulin resistance and the functionality of beta cells (Cani and Delzenne, 2009). In addition, modulation of the gut flora with prebiotics increases GLP2 production in the colon and this increase in GLP2 production is associated with higher expression of zonula occludens-1 (ZO-1), which improves the mucosal barrier function leading to a decrease in plasma LPS (Cani et al., 2009a,b). The study by Qin et al. (2012) showed that subjects with TDM2 suffered from a moderate intestinal dysbiosis and an increase in the number of various opportunistic gut pathogens, more than a change in a specific microbial species, having a direct association with the pathophysiology of TDM2. Specifically, they experienced a decrease in their butyrate-producing bacteria (Qin et al., 2012). This is significant because butyrate is the preferred source of energy, repair and maintaining cell health in the human digestive system. In the colon, the predominant butyrate-producing bacteria are the *C. coccoides* and the *Eubacterium rectale* groups. These changes in intestinal bacteria have recently been reported in patients with colorectal cancer (Wang et al., 2012) and in elderly people (Biagi et al., 2010). Thus, butyrate-producing bacteria could have a protecting role against a functional dysbiosis. Moreover, as other intestinal diseases show a loss of butyrate-producing bacteria with a commensurate increase in opportunistic pathogens, a possible hypothesis is that this change in the microbiota can cause an increase in susceptibility to a wide variety of diseases. The analysis of genetic bacterial functions shows an increase in functions related to the response to intestinal oxidative stress. This is of interest, because previous studies have shown that a high oxidative stress level is related to a predisposition to diabetic complications (Kashyap and Farrugia, 2011).

CONCLUSION

Metabolic diseases are caused by many factors, including a higher consumption of energy-rich diets, reduced physical activity, and a hereditary disposition. In the past 6 years, much evidence suggests that gut microbiota may play an important role in the regulation of energy balance and weight in animal models and in humans. However, although metagenomic tools have provided an important amount of data concerning the characterization and the potential role of this gut microbiota in the development of human obesity and TDM2, the causal relationship between this microbiota and obesity still needs to be confirmed in humans. In the future, larger human studies conducted at the species level and taking into account all of the possible confounding variables (such as age, gender, ethnicity, diet, and genetic factors) are needed to allow us to use the gut microbiota composition and modulation as novel diagnostic or therapeutic strategies to treat obesity and TDM2.

REFERENCES

Adlerberth, I., and Wold, A. E. (2009). Establishment of the gut microbiota in Western infants. *Acta Paediatr.* 98, 229–238. doi: 10.1111/j.1651-2227.2008.01060.x

Ajslev, T. A., Andersen, C. S., Gamborg, M., Sorensen, T. I., and Jess, T. (2011). Childhood overweight after establishment of the gut microbiota: the role of delivery mode, pre-pregnancy weight and early administration of antibiotics. *Int. J. Obes. (Lond.)* 35, 522–529. doi: 10.1038/ijo.2011.27

Amar, J., Burcelin, R., Ruidavets, J., Cani, P., Fauvel, J., Alessi, M., et al. (2008). Energy intake is associated with endotoxemia in apparently healthy men. *Am. J. Clin. Nutr.* 87, 1219–1223.

Amar, J., Chabo, C., Waget, A., Klopp, P., Vachoux, C., Bermudez-Humaran, L. G., et al. (2011). Intestinal mucosal adherence and translocation of commensal bacteria at the early onset of type 2 diabetes: molecular mechanisms and probiotic treatment. *EMBO Mol. Med.* 3, 559–572. doi: 10.1002/emmm.201100159

Armougom, F., Henry, M., Vialettes, B., Raccah, D., and Raoult, D. (2009). Monitoring bacterial community of human gut microbiota reveals an increase in *Lactobacillus* in obese patients and methanogens in anorexic patients. *PLoS ONE* 4:e7125. doi: 10.1371/journal.pone.0007125

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., et al. (2011). Enterotypes of the human gut microbiome. *Nature* 473, 174–180. doi: 10.1038/nature09944

Axad, M. B., Konya, T., Maughan, H., Guttman, D. S., Field, C. J., Chari, R. S., et al. (2013). Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months. *Can. Med. Assoc. J.* 185, 385–394. doi: 10.1503/cmaj.121189

Backhed, F., Ding, H., Wang, T., Hooper, L. V., Koh, G. Y., Nagy, A., et al. (2004). The gut microbiota as an environmental factor that regulates fat storage. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15718–15723. doi: 10.1073/pnas.0407076101

Backhed, F., Manchester, J. K., Semenkovich, C. F., and Gordon, J. I. (2007). Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *Proc. Natl. Acad. Sci. U.S.A.* 104, 979–984. doi: 10.1073/pnas.0605374104

Bech-Nielsen, G. V., Hansen, C. H., Hufeldt, M. R., Nielsen, D. S., Aasted, B., Vogensen, F. K., et al. (2012). Manipulation of the gut microbiota in C57BL/6 mice changes glucose tolerance without affecting weight development and gut mucosal immunity. *Res. Vet. Sci.* 92, 501–508. doi: 10.1016/j.rvsc.2011.04.005

Beck-Nielsen, H., Vaag, A., Poulsen, P., and Gaster, M. (2003). Metabolic and genetic influence on glucose metabolism in type 2 diabetic subjects – experiences from relatives and twin studies. *Best. Pract. Res. Clin. Endocrinol. Metab.* 17, 445–467. doi: 10.1016/S1521-690X(03)00041-1

Benson, A. K., Kelly, S. A., Legge, R., Ma, F., Low, S. J., Kim, J., 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. doi: 10.1073/pnas.1007028107

Bezirtzoglou, E., Tsiotsias, A., and Welling, G. W. (2011). Microbiota profile in feces of breast-and formula-fed newborns by using fluorescence in situ hybridization (FISH). *Anaerobe* 17, 478–482. doi: 10.1016/j.anaerobe.2011.03.009

Biagi, E., Nylund, L., Candela, M., Ostan, R., Bucci, L., Pini, E., et al. (2010). Through ageing, and beyond: gut microbiota and inflammatory status in seniors and centenarians. *PLoS ONE* 5:e10667. Erratum in: *PLoS ONE* 2010;5(6). doi: 10.1371/journal.pone.0010667

Brugman, S., Klatter, F. A., Visser, J. T., Wildeboer-Veloo, A. C., Harmsen, H. J., Rozing, J., et al. (2006). Antibiotic treatment partially protects against type 1 diabetes in the Bio-Breeding diabetes-prone rat. Is the gut flora involved

in the development of type 1 diabetes? *Diabetologia* 49, 2105–2108. doi: 10.1007/s00125-006-0334-0

Burcelin, R., Crivelli, V., Dacosta, A., Roy-Tirelli, A., and Thorens, B. (2002). Heterogeneous metabolic adaptation of C57BL/6J mice to high-fat diet. *Am. J. Physiol. Endocrinol. Metab.* 282, E834–E842.

Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., and Thomas, T. (2011). Bacterial community assembly based on functional genes rather than species. *Proc. Natl. Acad. Sci. U.S.A.* 108, 14288–14293. doi: 10.1073/pnas.1101591108

Cani, P. D., Amar, J., Iglesias, M. A., Poggi, M., Knauf, C., Bastelica, D., et al. (2007a). Metabolic endotoxemia initiates obesity and insulin resistance. *Diabetes* 56, 1761–1772. doi: 10.2337/db06-1491

Cani, P. D., Neyrinck, A. M., Fava, F., Knauf, C., Burcelin, R. G., Tuohy, K. M., et al. (2007b). Selective increases of bifidobacteria in gut microflora improve high-fat-diet-induced diabetes in mice through a mechanism associated with endotoxaemia. *Diabetologia* 50, 2374–2383. doi: 10.1007/s00125-007-0791-0

Cani, P. D., Bibiloni, R., Knauf, C., Waget, A., Neyrinck, A. M., Delzenne, N. M., et al. (2008). Changes in gut microbiota control metabolic endotoxemia-induced inflammation in high-fat diet-induced obesity and diabetes in mice. *Diabetes* 57, 1470–1481. doi: 10.2337/db07-1403

Cani, P. D., and Delzenne, N. M. (2009). The role of the gut microbiota in energy metabolism and metabolic disease. *Curr. Pharm. Des.* 15, 1546–1558. doi: 10.2174/138161209788168164

Cani, P. D., Lecourt, E., Dewulf, E. M., Sohet, F. M., Pachikian, B. D., Naslain, D., et al. (2009a). Gut microbiota fermentation of prebiotics increases satietogenic and incretin gut peptide production with consequences for appetite sensation and glucose response after a meal. *Am. J. Clin. Nutr.* 90, 1236–1243. doi: 10.3945/ajcn.2009.28095

Cani, P. D., Possemiers, S., Van de Wiele, T., Guiot, Y., Everard, A., Rottier, O., et al. (2009b). Changes in gut microbiota control inflammation in obese mice through a mechanism involving GLP-2-driven improvement of gut permeability. *Gut* 58, 1091–1103. doi: 10.1136/gut.2008.165886

Caporaso, J. G., Lauber, C. L., Costello, E. K., Berg-Lyons, D., Gonzalez, A., Stombaugh, J., et al. (2011). Moving pictures of the human microbiome. *Genome Biol.* 12, 50. doi: 10.1186/gb-2011-12-5-r50

Caricilli, A. M., and Saad, M. J. (2013). The role of gut microbiota on insulin resistance. *Nutrients* 5, 829–851. doi: 10.3390/nu5030829

Chung, S., Lapoint, K., Martinez, K., Kennedy, A., Boysen Sandberg, M., and McIntosh, M. K. (2006). Preadipocytes mediate lipopolysaccharide-induced inflammation and insulin resistance in primary cultures of newly differentiated human adipocytes. *Endocrinology* 147, 5340–5351. doi: 10.1210/en.2006-0536

Claesson, M. J., Cusack, S., O'Sullivan, O., Greene-Diniz, R., de Weerd, H., Flannery, E., et al. (2011). Composition, variability, and temporal stability of the intestinal microbiota of the elderly. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4586–4591. doi: 10.1073/pnas.1000097107

Claesson, M. J., Jeffery, I. B., Conde, S., Power, S. E., O'Connor, E. M., Cusack, S., et al. (2012). Gut microbiota composition correlates with diet and health in the elderly. *Nature* 488, 178–184. doi: 10.1038/nature11319

Clemente Postigo, M., Queipo Ortuño, M. I., Murri, M., Boto Ordoñez, M., Pérez Martínez, P., Andres Lacueva, C., et al. (2012). Endotoxin increase after fat overload is related to postprandial hypertriglyceridemia in morbidly obese patients. *J. Lipid Res.* 53, 973–978. doi: 10.1194/jlr.P020909

Collado, M. C., Isolauri, E., Laitinen, K., and Salminen, S. (2008). Distinct composition of gut microbiota during pregnancy in overweight and normal-weight women. *Am. J. Clin. Nutr.* 88, 894e9.

Costello, E. K., Lauber, C. L., Hamady, M., Fierer, N., Gordon, J. I., and Knight, R. (2009). Bacterial community variation in human body habitats across space and time. *Science* 326, 1694–1697. doi: 10.1126/science.1177486

De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J. B., Massart, S., et al. (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14691–14696. doi: 10.1073/pnas.1005963107

Delgado, S., Suárez, A., and Mayo, B. (2006). Identification of dominant bacteria in feces and colonic mucosa from healthy Spanish adults by culturing and by 16S rDNA sequence analysis. *Dig. Dis. Sci.* 51, 744–751. doi: 10.1007/s10620-006-3201-4

De Palma, G., Nadal, I., Collado, M. C., and Sanz, Y. (2009). Effects of a gluten-free diet on gut microbiota and immune function in healthy adult human subjects. *Br. J. Nutr.* 102, 1154–1160. doi: 10.1017/S0007114509371767

Dethlefsen, L., Eckburg, P. B., Bik, E. M., and Relman, D. A. (2006). Assembly of the human intestinal microbiota. *Trends. Ecol. Evol.* 21, 517–523. doi: 10.1016/j.tree.2006.06.013

Dethlefsen, L., Huse, S., Sogin, M. L., and Relman, D. A. (2008). The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol.* 6:e280. doi: 10.1371/journal.pbio.0060280

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, 4554–4561. doi: 10.1073/pnas.1000087107

Dominguez-Bello, M. G., Costello, E. K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., et al. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11971–11975. doi: 10.1073/pnas.1002601107

Drissi, F., Merhej, V., Angelakis, E., El Kaoutari, A., Carrière, F., Henrissat, B., et al. (2014). Comparative genomics analysis of *Lactobacillus* species associated with weight gain or weight protection. *Nutr. Diabetes* 4, e109. doi: 10.1038/nutd.2014.6

Duncan, S. H., Lobley, G. E., Holtrop, G., Ince, J., Johnstone, A. M., Louis, P., et al. (2008). Human colonic microbiota associated with diet, obesity and weight loss. *Int. J. Obes.* 32, 1720–1724. doi: 10.1038/ijo.2008.155

Erejuwa, O. O., Sulaiman, S. A., and Wahab, M. S. W. (2014). Modulation of gut microbiota in the management of metabolic disorders: the prospects and challenges. *Int. J. Mol. Sci.* 15, 4158–4188. doi: 10.3390/ijms15034158

Furet, J. P., Kong, L. C., Tap, J., Poitou, C., Basdevant, A., Bouillot, J. L., et al. (2010). Differential adaptation of human gut microbiota to bariatric surgery-induced weight loss: links with metabolic and low-grade inflammation markers. *Diabetes* 59, 3049–3057. doi: 10.2337/db10-0253

Grölund, M. M., Lehtonen, O. P., Eerola, E., and Kero, P. (1999). Fecal microflora in healthy infants born by different methods of delivery: permanent changes in intestinal flora after cesarean delivery. *J. Pediatr. Gastroenterol. Nutr.* 28, 19–25. doi: 10.1097/00005176-199901000-00007

Hernandez, E., Bargiela, R., Diez, M. S., Friedrichs, A., Perez-Cobas, A. E., Gosalbes, M. J., et al. (2013). Functional consequences of microbial shifts in the human gastrointestinal tract linked to antibiotic treatment and obesity. *Gut Microbes* 4, 306–315. doi: 10.4161/gmic.25321

Hildebrandt, M. A., Hoffmann, C., Sherrill-Mix, S. A., Keilbaugh, S. A., Hamady, M., Chen, Y. Y., et al. (2009). High-fat diet determines the composition of the murine gut microbiome independently of obesity. *Gastroenterology* 137, 1716–1724. doi: 10.1053/j.gastro.2009.08.042

Hotamisligil, G. S. (2006). Inflammation and metabolic disorders. *Nature* 444, 860–867. doi: 10.1038/nature05485

Human Microbiome Project Consortium. (2012a). Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214. doi: 10.1038/nature11234

Human Microbiome Project Consortium. (2012b). A framework for human microbiome research. *Nature* 486, 215–221. doi: 10.1038/nature11209

Huse, S. M., Ye, Y., Zhou, Y., and Fodor, A. A. (2012). A core human microbiome as viewed through 16S rRNA sequence clusters. *PLoS ONE* 7:e3424225. doi: 10.1371/journal.pone.0034242

Jacquemont, S., Reymond, A., Zufferey, F., Harewood, L., Walters, R. G., Kutalik, Z., et al. (2011). Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. *Nature* 478, 97–102. doi: 10.1038/nature10406

Jeffery, I. B., Claesson, M. J., O'Toole, P. W., and Shanahan, F. (2012). Categorization of the gut microbiota: enterotypes or gradients? *Nat. Rev. Microbiol.* 10, 591–592. doi: 10.1038/nrmicro2859

Jernberg, C., Löfmark, S., Edlund, C., and Jansson, J. K. (2007). Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. *ISME J.* 1, 56–66. doi: 10.1038/ismej.2007.3

Jernberg, C., Löfmark, S., Edlund, C., and Jansson, J. K. (2010). Long-term impacts of antibiotic exposure on the human intestinal microbiota. *Microbiology* 156, 3216–3223. doi: 10.1099/mic.0.040618-0

Jumpertz, R., Le, D. S., Turnbaugh, P. J., Trinidad, C., Bogardus, C., Gordon, J. I., et al. (2011). Energy-balance studies reveal associations between gut microbes, caloric load, and nutrient absorption in humans. *Am. J. Clin. Nutr.* 94, 58–65. doi: 10.3945/ajcn.110.010132

Kalliomäki, M., Collado, M. C., Salminen, S., and Isolauri, E. (2008). Early differences in fecal microbiota composition in children may predict overweight. *Am. J. Clin. Nutr.* 87, 534–538.

Karlsson, F. H., Tremaroli, V., Nookaew, I., Bergström, G., Behre, C. J., Fagerberg, B., et al. (2013). Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* 498, 99–103. doi: 10.1038/nature 12198

Kashyap, P., and Farrugia, G. (2011). Oxidative stress: key player in gastrointestinal complications of diabetes. *Neurogastroenterol. Motil.* 23, 111–114. doi: 10.1111/j.1365-2982.2010.01659.x

Koenig, J. E., Spor, A., Scalfone, N., Fricker, A. D., Stombaugh, J., Knight, R., et al. (2011). Succession of microbial consortia in the developing infant gut microbiome. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4578–4585. doi: 10.1073/pnas.1000081107

Larsen, N., Vogensen, F. K., van den Berg, F. W., Nielsen, D. S., Andreasen, A. S., Pedersen, B. K., et al. (2010). Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PLoS ONE* 5:e9085. doi: 10.1371/journal.pone.0009085

Leser, T. D., and Molbak, L. (2009). Better living through microbial action: the benefits of the mammalian gastrointestinal microbiota on the host. *Environ. Microbiol.* 11, 2194–2206. doi: 10.1111/j.1462-2920.2009.01941.x

Ley, R. E. (2010). Obesity and the human microbiome. *Curr. Opin. Gastroenterol.* 26, 5–11. doi: 10.1097/MOG.0b013e328333d751

Ley, R. E., Backhed, F., Turnbaugh, P., Lozupone, C. A., Knight, R. D., and Gordon, J. I. (2005). Obesity alters gut microbial ecology. *Proc. Natl. Acad. Sci. U.S.A.* 102, 11070–11075. doi: 10.1073/pnas.0504978102

Ley, R. E., Turnbaugh, P. J., Klein, S., and Gordon, J. I. (2006). Microbial ecology: human gut microbes associated with obesity. *Nature* 444, 1022–1023. doi: 10.1038/4441022a

Liszt, K., Zwielehner, J., Handschur, M., Hippe, B., Thaler, R., and Haslberger, A. G. (2009). Characterization of bacteria, clostridia and bacteroides in faeces of vegetarians using qPCR and PCR-DGGE fingerprinting. *Ann. Nutr. Metab.* 54, 253–257. doi: 10.1159/000229505

Mackie, R. I., Sghir, A., and Gaskins, H. R. (1999). Developmental microbial ecology of the neonatal gastrointestinal tract. *Am. J. Clin. Nutr.* 69, 1035S–1045S.

Mai, V., McCrary, Q. M., Sinha, R., and Glei, M. (2009). Associations between dietary habits and body mass index with gut microbiota composition and fecal water genotoxicity: an observational study in African American and Caucasian American volunteers. *Nutr. J.* 8, 49. doi: 10.1186/1475-2891-8-49

Manco, M., Putignani, L., and Bottazzo, G. F. (2010). Gut microbiota, lipopolysaccharides, and innate immunity in the pathogenesis of obesity and cardiovascular risk. *Endocr. Rev.* 31, 817–844. doi: 10.1210/er.2009-0030

McCartney, A., Wenzhi, W., and Tannock, G. (1996). Molecular analysis of the composition of the bifidobacterial and lactobacillus microflora of humans. *Appl. Environ. Microbiol.* 62, 46080–44613.

Medici, F., Hawa, M., Ianari, A., Pyke, D. A., and Leslie, R. D. (1999). Concordance rate for type II diabetes mellitus in monozygotic twins: actuarial analysis. *Diabetologia* 42, 146–150. doi: 10.1007/s001250051132

Million, M., Maraninchi, M., Henry, M., Armougom, F., and Raoult, D. (2012). Obesity-associated gut microbiota is enriched in *Lactobacillus reuteri* and depleted in *Bifidobacterium animalis* and *Methanobrevibacter smithii*. *Int. J. Obes.* 36, 817–825. doi: 10.1038/ijo.2011.153

Morelli, L. (2008). Postnatal development of intestinal microflora as influenced by infant nutrition. *J. Nutr.* 138, 1791S–1795S.

Muegge, B. D., Kuczynski, J., Knights, D., Clemente, J. C., González, A., Fontana, L., et al. (2011). Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* 332, 970–974. doi: 10.1126/science.1198719

Murphy, E. F., Cotter, P. D., Hogan, A., O'Sullivan, O., Joyce, A., Fouhy, F., et al. (2013). Divergent metabolic outcomes arising from targeted manipulation of the gut microbiota in diet-induced obesity. *Gut* 62, 220–226. doi: 10.1136/gutjnl-2011-300705

Murri, M., Leiva, I., Gomez-Zumaquero, J. M., Tinahones, F. J., Cardona, F., Soriguer, F., et al. (2013). Gut microbiota in children with type 1 diabetes differs from that in healthy children: a case–control study. *BMC Med.* 11:46. doi: 10.1186/1741-7015-11-46

Musso, G., Gambino, R., and Cassader, M. (2010). Obesity, diabetes, and gut microbiota: the hygiene hypothesis expanded? *Diabetes Care* 33, 2277–2284. doi: 10.2337/dc10-0556

Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A., and Brown, P. O. (2007). Development of the human infant intestinal microbiota. *PLoS Biol.* 5:e177. doi: 10.1371/journal.pbio.0050177

Penders, J., Thijs, C., Vink, C., Stelma, F. F., Snijders, B., Kummeling, I., et al. (2006). Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* 118, 511–521. doi: 10.1542/peds.2005-2824

Peterson, J., Garges, S., Giovanni, M., McInnes, P., Wang, L., Schloss, J. A., et al. (2009). The NIH Human Microbiome Project. *Genome. Res.* 19, 2317–2323. doi: 10.1101/gr.096651.109

Pickup, J. C., and Crook, M. A. (1998). Is type II diabetes mellitus a disease of the innate immune system? *Diabetologia* 41, 1241–1248. doi: 10.1007/s0012 50051058

Poggi, M., Bastelica, D., Gual, P., Iglesias, M. A., Gremeaux, T., Knauf, C., et al. (2007). C3H/HeJ mice carrying a Toll-like receptor 4 mutation are protected against the development of insulin resistance in white adipose tissue in response to a high-fat diet. *Diabetologia* 50, 1267–1276. doi: 10.1007/s00125-007-0654-8

Purchiaroni, F., Tortora, A., Gabrielli, M., Bertucci, F., Gigante, G., Ianiro, G., et al. (2013). The role of intestinal microbiota and the immune system. *Eur. Rev. Med. Pharmacol. Sci.* 17, 323–333.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65. doi: 10.1038/nature08821

Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., et al. (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490, 55–60. doi: 10.1038/nature11450

Ravel, J., Gajer, P., Abdo, Z., Schneider, G. M., Koenig, S. S., McCulle, S. L., et al. (2011). Vaginal microbiome of reproductive-age women. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4680–4687. doi: 10.1073/pnas.1002611107

Robinson, C. J., Bohannan, B. J., and Young, V. B. (2010). From structure to function: the ecology of host-associated microbial communities. *Microbiol. Mol. Biol. Rev.* 74, 453–476. doi: 10.1128/MMBR.00014-10

Rodes, L., Khan, A., Paul, A., Coussa-Charley, M., Marinescu, D., Tomaro-Duchesneau, C., et al. (2013). Effect of probiotics *Lactobacillus* and *Bifidobacterium* on gut-derived lipopolysaccharides and inflammatory cytokines: an in vitro study using a human colonic microbiota model. *J. Microbiol. Biotechnol.* 23, 518–526. doi: 10.4014/jmb.1205.05018

Santacruz, A., Marcos, A., Warnberg, J., Marti, A., Martin-Matillas, M., Campoy, C., et al. (2009). Interplay between weight loss and gut microbiota composition in overweight adolescents. *Obesity (Silver Spring)* 17, 1906–1915. doi: 10.1038/oby.2009.112

Scott, K. P., Duncan, S. H., and Flint, H. J. (2008). Dietary fibre and the gut microbiota. *Nutr. Bull.* 33, 201–211. doi: 10.1111/j.1467-3010.2008.00706.x

Serino, M., Fernández-Real, J. M., Fuentes, E. G., Queipo-Ortuño, M., Moreno-Navarrete, J. M., Sánchez, A., et al. (2013). The gut microbiota profile is associated with insulin action in humans. *Acta Diabetol.* 50, 753–761. doi: 10.1007/s00592-012-0410-5

Shoelson, S., Lee, J., and Goldfine, A. (2006). Inflammation and insulin resistance. *J. Clin. Invest.* 116, 1793–1801. doi: 10.1172/JCI29069

Stachowicz, N., and Kiersztan, A. (2013). The role of gut microbiota in the pathogenesis of obesity and diabetes. *Postepy Hig. Med. Dosw.* 67, 288–303. doi: 10.5604/17322693.1044746

Sullivan, A., Edlund, C., and Nord, C. E. (2001). Effect of antimicrobial agents on the ecological balance of human microflora. *Lancet Infect. Dis.* 1, 101–114. doi: 10.1016/S1473-3099(01)00066-4

Tagliabue, A., and Elli, M. (2013). The role of gut microbiota in human obesity: recent findings and future perspectives. *Nutr. Metab. Cardiovasc. Dis.* 23, 160–168. doi: 10.1016/j.numecd.2012.09.002

Thuny, F., Richet, H., Casalta, J. P., Angelakis, E., Habib, G., and Raoult, D. (2010). Vancomycin treatment of infective endocarditis is linked with recently acquired obesity. *PLoS ONE* 5:e9074. doi: 10.1371/journal.pone.0009074

Trehan, I., Goldbach, H. S., LaGrone, L. N., Meuli, G. J., Wang, R. J., Maleta, K. M., et al. (2013). Antibiotics as part of the management of severe acute malnutrition. *N. Engl. J. Med.* 368, 425–435. doi: 10.1056/NEJMoa1202851

Tremaroli, V., and Bäckhed, F. (2012). Functional interactions between the gut microbiota and host metabolism. *Nature* 489, 242–249. doi: 10.1038/nature11552

Turnbaugh, P., Ridaura, V., Faith, J., Rey, F., Knight, R., and Gordon, J. (2009a). The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci. Transl. Med.* 1, 6–14. doi: 10.1126/scitranslmed.3000322

Turnbaugh, P. J., Hamady, M., Yatsunenko, T., Cantarel, B. L., Duncan, A., Ley, R. E., et al. (2009b). A core gut microbiome in obese and lean twins. *Nature* 457, 480–484. doi: 10.1038/nature07540

Turnbaugh, P. J., Bäckhed, 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, 213–223. doi: 10.1016/j.chom.2008.02.015

Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., and Gordon, J. I. (2007). The human microbiome project. *Nature* 449, 804–810. doi: 10.1038/nature06244

Turnbaugh, P. J., Ley, R. E., Mahowald, M. A., Magrini, V., Mardis, E. R., and Gordon, J. I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444, 1027–1031. doi: 10.1038/nature05414

Turnbaugh, P. J., Quince, C., Faith, J. J., McHardy, A. C., Yatsunenko, T., Niazi, F., et al. (2010). Organismal, genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins. *Proc. Natl. Acad. Sci. U.S.A.* 107, 7503–7508. doi: 10.1073/pnas.1002355107

Turroni, F., Peano, C., Pass, D. A., Foroni, E., Severgnini, M., Claesson, M J., et al. (2012). Diversity of *Bifidobacteria* within the infant gut microbiota. *PLoS ONE* 7:e36957. doi: 10.1371/journal.pone.0036957

Walker, A. W., Ince, J., Duncan, S. H., Webster, L. M., Holtrop, G., Ze, X., et al. (2011). Dominant and diet-responsive groups of bacteria within the human colonic microbiota. *ISME J.* 5, 220–230. doi: 10.1038/ismej.2010.118

Wang, T., Cai, G., Qiu, Y., Fei, N., Zhang, M., Pang, X., et al. (2012). Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *ISME J.* 6, 320–329. doi: 10.1038/ismej.2011.109

Wu, G. D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y. Y., Keilbaugh, S. A., et al. (2011). Linking long-term dietary patterns with gut microbial enterotypes. *Science* 334, 105–108. doi: 10.1126/science.1208344

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., et al. (2012). Human gut microbiome viewed across age and geography. *Nature* 486, 222–227. doi: 10.1038/nature11053

Yong, E. (2012). Gut microbial 'enterotypes' become less clear-cut. *Nature News.* Available at: http://www.nature.com/news/gut-microbial-enterotypes-become-less-clear-cut-1.10276

Zhang, C., Zhang, M., Wang, S., Han, R., Cao, Y., Hua, W., et al. (2010). Interactions between gut microbiota, host genetics and diet relevant to development of metabolic syndromes in mice. *ISME J.* 4, 232–241. doi: 10.1038/ismej.2009.112

Zhang, H., DiBaise, J. K., Zuccolo, A., Kudrna, D., Braidotti, M., Yu, Y., et al. (2009). Human gut microbiota in obesity and after gastric bypass. *Proc. Natl. Acad. Sci. U.S.A.* 106, 2365–2370. doi: 10.1073/pnas.0812600106

Zimmer, J., Lange, B., Frick, J., Sauer, H., Zimmermann, K., Schwiertz, A., et al. (2011). A vegan or vegetarian diet substantially alters the human colonic faecal microbiota. *Eur. J. Clin. Nutr.* 66, 53–60. doi: 10.1038/ejcn.2011.141

Combined metagenomic and phenomic approaches identify a novel salt tolerance gene from the human gut microbiome

Eamonn P. Culligan[1,2], Julian R. Marchesi[1,3,4]*, Colin Hill[1,2]* and Roy D. Sleator[1,5]*

[1] Alimentary Pharmabiotic Centre, Biosciences Institute, University College Cork, Cork, Ireland
[2] School of Microbiology, University College Cork, Cork, Ireland
[3] Cardiff School of Biosciences, Cardiff University, Cardiff, UK
[4] Department of Surgery and Cancer, Centre for Digestive and Gut Health, Imperial College London, London, UK
[5] Department of Biological Sciences, Cork Institute of Technology, Cork, Ireland

In the current study, a number of salt-tolerant clones previously isolated from a human gut metagenomic library were screened using Phenotype MicroArray (PM) technology to assess their functional capacity. PM's can be used to study gene function, pathogenicity, metabolic capacity and identify drug targets using a series of specialized microtitre plate assays, where each well of the microtitre plate contains a different set of conditions and tests a different phenotype. Cellular respiration is monitored colorimetrically by the reduction of a tetrazolium dye. One clone, SMG 9, was found to be positive for utilization/transport of L-carnitine (a well-characterized osmoprotectant) in the presence of 6% w/v sodium chloride (NaCl). Subsequent experiments revealed a significant growth advantage in minimal media containing NaCl and L-carnitine. Fosmid sequencing revealed putative candidate genes responsible for the phenotype. Subsequent cloning of two genes did not replicate the L-carnitine-associated phenotype, although one of the genes, a σ^{54}-dependent transcriptional regulator, did confer salt tolerance to *Escherichia coli* when expressed in isolation. The original clone, SMG 9, was subsequently found to have lost the original observed phenotype upon further investigation. Nevertheless, this study demonstrates the usefulness of a phenomic approach to assign a functional role to metagenome-derived clones.

Keywords: metagenomics, functional metagenomics, gut microbiome, microbiota, salt tolerance, BIOLOG, phenotype microarray, transcriptional regulator

INTRODUCTION

The ability to adapt to and tolerate increases in extracellular osmolarity is an important characteristic that enables bacteria to survive in stressful environments. Increased osmolarity, caused by sodium chloride (NaCl) for example, initiates a phased response in bacteria. Firstly during the primary response, potassium ions are rapidly accumulated within the bacterial cell to offset the detrimental effects of water loss and influx of toxic sodium and chloride ions (Sleator and Hill, 2002; Epstein, 2003). Once the cell has been stabilized, the secondary response begins and involves the synthesis or, more often the more energetically favorable, uptake of osmoprotectant compounds (also termed compatible solutes or osmolytes). Osmoprotectants are compatible with cellular functions and can accumulate to very high concentrations within the cell and function to protect proteins and to restore cell volume and thus, turgor pressure (Kempf and Bremer, 1998; Sleator and Hill, 2002; Kunte, 2006). Osmoprotectants can be grouped broadly as amino acids, polyols, sugars, trimethyl- and quaternary-ammonium compounds and their derivatives (Kempf and Bremer, 1998). The most widely utilized and best characterized osmoprotectants are glycine betaine, carnitine, proline, and ectoine. Numerous studies have shown carnitine to be important not only for salt tolerance, but also for survival *in vivo*

and pathogenesis of infection (Sleator et al., 2001; Wemekamp-Kamphuis et al., 2004). Carnitine is also found abundantly in animal tissues and red meat and is an important compound in the host environment; for the human pathogen *Listeria monocytogenes*, carnitine and its uptake system OpuC are critical for infection in mice (Sleator et al., 2003). In addition to its osmoprotective properties, carnitine may also be catabolised as a carbon or nitrogen source to generate energy (Wargo and Hogan, 2009).

The emergence of "omics" technologies, an umbrella term to include analyses such as genomics, metagenomics, transcriptomics, proteomics, metabolomics, and phenomics to name a few, have been used to gain valuable information about the functions and interactions of various biological systems as a whole and can provide more information than more traditional and reductive approaches to biological problems. Sequence-based and functional metagenomic approaches have led to the discovery of many novel and diverse genes (Beja et al., 2000; Gillespie et al., 2002; Banik and Brady, 2008; Culligan et al., 2013; Yoon et al., 2013). While identifying clones which display a specific phenotype through functional metagenomic screening yields worthwhile results, characterizing the functional mechanisms responsible for the observed phenotype can be sometimes difficult owing to the fact that a large proportion of metagenome derived

genes will be annotated as hypothetical proteins or have no known function or homology to existing proteins (Bork, 2000; Qin et al., 2010). Phenomic approaches can be used to study hundreds of phenotypic profiles of different bacterial strains concurrently. Comparing phenomic profiles of wild-type and mutant derivatives or host strains and clones identified through metagenomic screening can reveal differences between strains relating to gene function, pathogenicity and metabolism for example (Bochner et al., 2001; Bochner, 2009).

With this in mind, we have utilized combined metagenomic and phenomic approaches in this study to characterize a salt tolerant clone identified from a human gut microbiome metagenomic fosmid library. An overview of the study design using this combined approach can be seen in **Figure 1**.The BIOLOG phenotype microarray (PM) system was used to compare phenotypes between metagenomic clones and cloning host (carrying empty fosmid vector). PM plates measure cellular respiration colorimetrically via reduction of a tetrazolium dye with electrons from NADH generated during the process of respiration. Strongly metabolized substrates generate a more intense purple color which is recorded with a camera on the Omnilog instrument. If desired, thousands of phenotypes may be monitored simultaneously using the different available PM plates which can be grouped as those that measure carbon, nitrogen, phosphorous and sulfur metabolism, response to different pH conditions, osmolytes and chemicals such as anitbiotics. The full range of PM plates and their constituents, for investigation of bacterial phenotypes can be found at: http://www.biolog.com/products-static/phenotype_microbial_cells_use.php

As our primary interest is salt tolerance we used the PM osmolyte plate (PM9) for analysis. The PM screen indicated that clone SMG 9 was positive for L-carnitine utilization in the presence of 6% w/v NaCl. Sequencing of the fosmid insert and cloning of two genes identified a novel salt tolerance gene, but did not replicate the carnitine-associated phenotype originally observed. Subsequent investigation revealed SMG 9 had lost this phenotype. Notwithstanding this phenomenon, this study demonstrates the usefulness of a phenomic approach to assign a functional role to metagenomic library-derived clones.

MATERIALS AND METHODS

BACTERIAL STRAINS AND GROWTH CONDITIONS
Bacterial strains, plasmids, and oligonucleotide primers (Eurofins, MWG Operon, Germany) used in this study are listed in **Table 1**. *Escherichia coli* EPI300::pCC1FOS (Epicentre Biotechnologies, Madison, WI, USA) was grown in Luria-Bertani (LB) medium containing 12.5 μg/ml chloramphenicol (Cm). *E. coli* MKH13 was grown in LB medium and in LB medium supplemented with 20 μg/ml Cm for strains transformed with the plasmid pCI372. LB media was supplemented with 1.5% w/v agar when required. All overnight cultures were grown at 37°C with shaking.

METAGENOMIC LIBRARY CONSTRUCTION AND SCREENING
A previously constructed fosmid clone library (Jones and Marchesi, 2007; Jones et al., 2008), created from metagenomic DNA isolated from a fecal sample from a healthy 26 year old

Caucasian male was used to screen for salt-tolerant clones. The library was screened as outlined previously (Culligan et al., 2012). Briefly, a total of 23,040 clones from the library were screened on LB agar supplemented with 6.5% (w/v) NaCl (a concentration which inhibits the growth of the cloning host, *E. coli* EPI300) and 12.5 μg/ml Cm using a Genetix QPix 2 XT colony picking/gridding robotics platform to identify clones with an increased salt tolerance phenotype compared to the cloning host (*E. coli* EPI300) carrying an empty fosmid vector (pCC1FOS). Identification of any salt tolerant clones will therefore most likely be due to a gene (or genes) present on the metagenomic insert from the human gut microbiota. Clones were gridded onto Q-trays (Genetix) using the robotics platform. Q-trays were incubated at 37°C for 3 days and checked twice daily for growth of likely salt-tolerant clones. Salt-tolerant clones were subsequently replica plated onto LB agar containing 12.5 μg/ml Cm and 6.5% NaCl and onto LB containing 12.5 μg/ml Cm, but without 6.5% NaCl (which represented a positive control plate). Each salt tolerant clone identified was streaked on LB agar + 12.5 μg/ml Cm to ensure a pure culture and all clones were maintained as glycerol stocks at −80°C.

PHENOTYPE MICROARRAY (PM) ASSAY
The PM9 osmolytes microplate was used to compare the cellular phenotypes of SMG 9 and the cloning host, *E. coli* EPI300::pCC1FOS (containing an empty fosmid vector) under 96 different conditions. Preparation of the different IF (Inoculating Fluids; proprietary formulation supplied by BIOLOG) solutions and inoculation of the PM plates was performed according to the BIOLOG PM protocol for *E. coli* and other Gram negative bacteria. Briefly, IF-0 solution was prepared by adding 25 ml of sterile water to 125 ml of 1.2× IF-0. IF-0 + dye mix A solution was prepared by adding 1.8 ml of dye mix A and 23.2 ml of sterile water to another bottle containing 125 ml of 1.2× IF-0. IF-10 solution was prepared by adding 1.5 ml of dye mix A and 23.5 ml of sterile water to a bottle containing 125 ml of 1.2× IF-10. *E. coli* strains were grown overnight on LB agar at 37°C by streaking from a frozen stock. Cells were sub-cultured by streaking again on LB agar and grown overnight again. Isolated colonies were removed from the agar plate using a sterile swab and added to a tube containing 16 ml of IF-0 solution until a cell suspension of 42% T (transmittance) was achieved using the BIOLOG Turbidimeter. 15 ml of this 42% T solution was diluted in 75 ml of IF-0 + dye mix A to achieve 85% T. 600 μl of the 85% T cell suspension solution was added to 120 ml of IF-10 + dye mix A. Finally, 100 μl of the final cell suspension was inoculated to each well of the PM 9 microplate. Plates were incubated at 37°C for 24 h in the Omnilog plate reader (BIOLOG).

FOSMID SEQUENCING AND ANALYSIS
Fosmid DNA was isolated from SMG 9 as described above to a concentration >200 ng/μl (approximately 5 μg in total). Sequencing of the full fosmid insert of SMG 9 was performed by GATC Biotech (Konstanz, Germany) using 454-pyrosequencing on a titanium mini-run of the Roche GS-FLX platform, achieving approximately 65-fold coverage. Sequencing
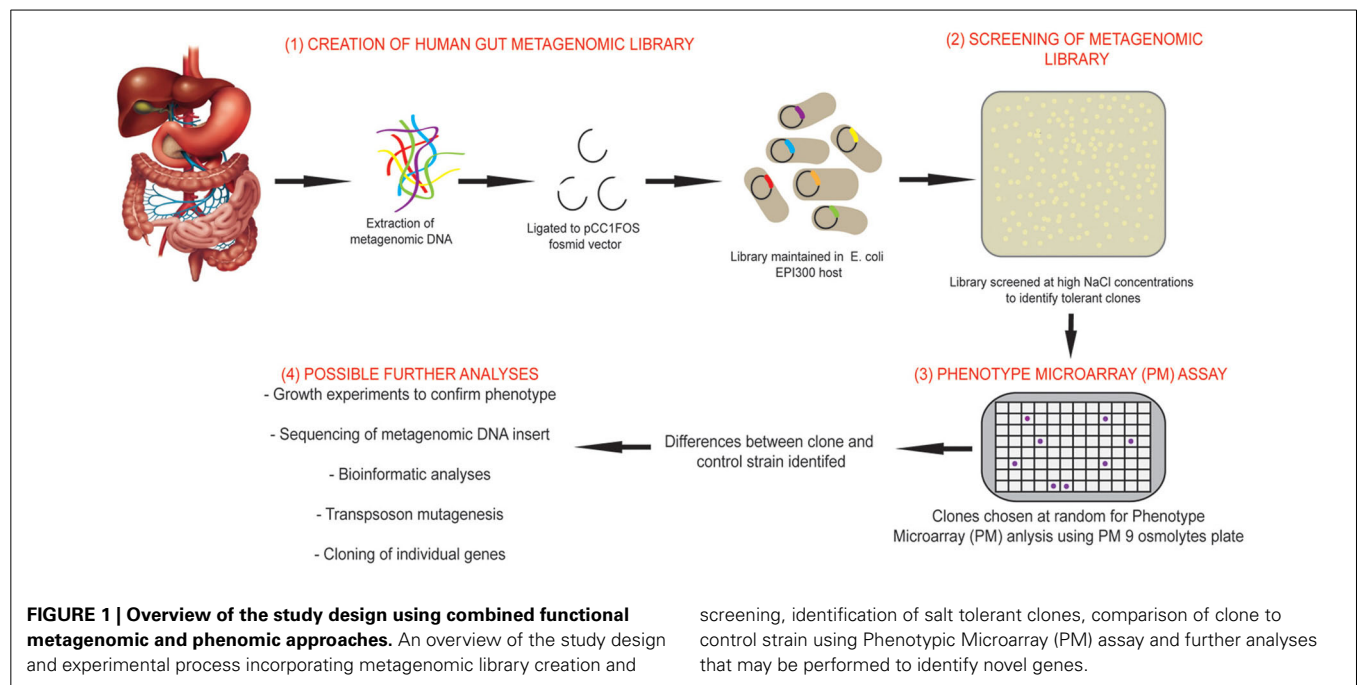
FIGURE 1 | Overview of the study design using combined functional metagenomic and phenomic approaches. An overview of the study design and experimental process incorporating metagenomic library creation and screening, identification of salt tolerant clones, comparison of clone to control strain using Phenotypic Microarray (PM) assay and further analyses that may be performed to identify novel genes.

Table 1 | Bacterial strains, plasmids and oligonucleotide primers used in this study.

Strain or plasmid	Genotype or characteristic(s)	Source or references
STRAINS		
E. coli EPI300	F⁻ *mcrA* Δ(*mrr-hsd*RMS-*mcrBC*) Φ80d*lacZ*ΔM15 Δ*lac*X74 *recA*1 *endA*1 *araD*139 Δ(*ara, leu*)7697 *galU galK* λ⁻ *rpsL nupG trfA dhfr*; high-transformation efficiency of large DNA	Epicentre Biotechnologies, Madison, WI, USA
E.coli MKH13	MC4100Δ(*putPA*)101D(*proP*)2D(*proU*)	Haardt et al., 1995
E. coli MKH13::pCl372-*mfsT*	MKH13 containing pCl372 with *mfsT* gene from SMG 9 (similar to *Bacteroides* sp. CAG:545)	This study
E. coli MKH13::pCl372-*sdtR*	MKH13 containing pCl372 with *sdtR* gene from SMG 9 (similar to *Bacteroides* sp. CAG:545)	This study
E. coli EPI300::pCl372-*mfsT*	EPI300 containing pCl372 with *mfsT* gene from SMG 9 (similar to *Bacteroides* sp. CAG:545)	This study
E. coli EPI300::pCl372-*sdtR*	EPI300 containing pCl372 with *sdtR* gene from SMG 9 (similar to *Bacteroides* sp. CAG:545)	This study
PLASMIDS		
pCl372	Shuttle vector between *E. coli* and *L. lactis*, Cm^R	Hayes et al., 1990
pCC1FOS	Fosmid cloning vector, Cm^R	Epicentre Biotechnologies, Madison, WI, USA
PRIMERS		
pCl372 FP	CGGGAAGCTAGAGTAAGTAG	This study
pCl372 RP	CCTCTCGGTTATGAGTTAG	This study
mfsT FP	AAAA<u>CTGCAG</u>CGGACTCGTGGTGGATGAC	This study
mfsT RP	GC<u>TCTAGA</u>CAAACACCTTGGTGTCATTAGC	This study
sdtR FP	AAAA<u>GTCGAC</u>CGGCTGTATGGAAGTTCCTG	This study
sdtR RP	GC<u>TCTAGA</u>GCCTCCAAATTTTGCACCAAG	This study

FP, forward primer; RP, reverse primer; CmR, chlorpamphenicol resistance; Restriction enzyme cut sites are underlined; PstI, CTGCAG; XbaI, TCTAGA; SalI, GTCGAC.

reads were assembled into a single contig by GATC Biotech. The retrieved sequence was analyzed using the FGENESB software program (Softberry) to identify putative open reading frames and translated nucleotide sequences were subjected to BLASTP analysis to assign putative functions to the encoded proteins and identify homologous sequences from the National Centre for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih. gov/blast/Blast.cgi). The full fosmid insert sequence of SMG 9 was submitted to GenBank and assigned the following accession number; KJ524644.

DNA MANIPULATIONS AND CLONING OF *mfsT* AND *sdtR* GENES

Extraction of fosmids containing metagenomic DNA: 5 ml of bacterial culture was grown overnight with 12.5 μg/ml Cm. One millilitre of culture was used to inoculate 4 ml of fresh LB broth. To this, 5 μl of 1000× Copy Control Induction solution (Epicentre Biotechnologies) and 12.5 μg/ml Cm were added. The mixture was incubated at 37°C for 5 h with vigorous shaking (200–250 rpm) to ensure maximum aeration. Cells were harvested from the whole 5 ml of induced culture by centrifuging at $2100 \times g$ for 12 min. Qiagen QIAprep Spin mini-prep kit was used to extract fosmids as per manufacturer's instructions. PCR products were purified with a Qiagen PCR purification kit and digested with *XbaI* and *PstI* (Roche Applied Science) for *mfsT* and with *SalI* and *XbaI* for *sdtR*, followed by ligation using the FastLink DNA ligase kit (Epicentre Biotechnologies) to similarly digested plasmid pCI372. Electrocompetent *E. coli* MKH13 and *E. coli* EPI300 were transformed with the ligation mixture and plated on LB agar plates containing 20 μg/ml Cm for selection. Colony PCR was performed on all resistant transformants using primers across the multiple cloning site (MCS) of pCI372 to confirm the presence and size of the insert.

CONFIRMATION TESTS FOR OBSERVED PHENOTYPE

Growth experiments were performed in defined M9 minimal media (M9MM) (Fluka) to confirm the observed phenotype. Single isolated colonies of SMG 9 and EPI300::pCC1FOS were grown overnight in M9MM (containing final concentrations of; D-glucose (0.4%), Bacto™ casamino acids (w/v 0.2%) (Becton, Dickinson and Co, Sparks, MD, USA), magnesium sulfate ($MgSO_4$) (2 mM), calcium chloride ($CaCl_2$) (0.1 mM) and 12.5 μg/ml Cm). Reagents were purchased from Sigma Aldrich (St. Louis, MO, USA) unless otherwise stated. Cells were harvested by centrifugation, washed in ¼ strength Ringers solution and resuspended in fresh M9MM. A 2% v/v inoculum was sub-cultured in fresh M9MM containing various concentrations of sodium chloride (0–8% w/v NaCl) and 1 mM of L-carnitine when required. Triplicate wells of a 96-well microtitre plate were inoculated with 200 μl of the appropriate cell suspension. Plates were incubated at 37°C for 24–48 h in an automated spectrophotometer (Tecan Genios) which recorded the optical density at 595 nm ($OD_{595\,nm}$) every hour. After 48 h the data was retrieved and analyzed using the Magellan 3 software program and graphs were created with Sigma Plot 10.0 (Systat Software Inc, London, UK). Results are presented as the average of triplicate experiments, with error bars being representative of the standard error of the mean (SEM).

RESULTS

SCREENING OF METAGENOMIC LIBRARY

Approximately 23,000 clones from a metagenomic fosmid library from the human gut microbiome were screened previously and resulted in the identification of 53 salt tolerant clones which could grow on LB agar supplemented with 6.5% NaCl (a concentration which inhibits growth of the cloning host, *E. coli* EPI300::pCC1FOS) (Culligan et al., 2012). The salt tolerant clones identified were annotated as SMG 1-53 (Salt MetaGenome 1–53).

PM OSMOLYTE PLATE ASSAY

A number of clones initially identified from the metagenomic library were chosen at random and phenotypically screened using PM 9 osmolytes plate from BIOLOG. The layout and contents of PM9 can be viewed at http://www.biolog.com/pdf/pm_lit/PM1-PM10.pdf. One clone, SMG 9, gave a positive result under one of the 96 different conditions tested. It was found that SMG 9 had an increased metabolic response (causing a reduction of the tetrazolium dye to generate a purple color, **Figure 2A**) compared to the cloning host containing an empty fosmid vector, *E. coli* EPI300::pCC1FOS (**Figure 2B**) in 6% w/v NaCl supplemented with L-carnitine. The color formation within each well was measured by BIOLOG's Omnilog machine, which produces a color-coded graph. A comparison of the kinetic data output from SMG 9 and EPI300::pCC1FOS can be seen in **Figure 3**.

CONFIRMATORY EXPERIMENTATION OF PM ANALYSIS

In an attempt to confirm and replicate the result from the PM analysis, SMG 9 and EPI300::pCC1FOS were grown in M9MM containing various concentrations of NaCl (0–6% w/v) and supplemented with 1 mM L-carnitine when appropriate. **Figure 4A** shows growth of both clones in M9MM in the presence and absence of 1 mM L-carnitine. In the presence of L-carnitine, SMG
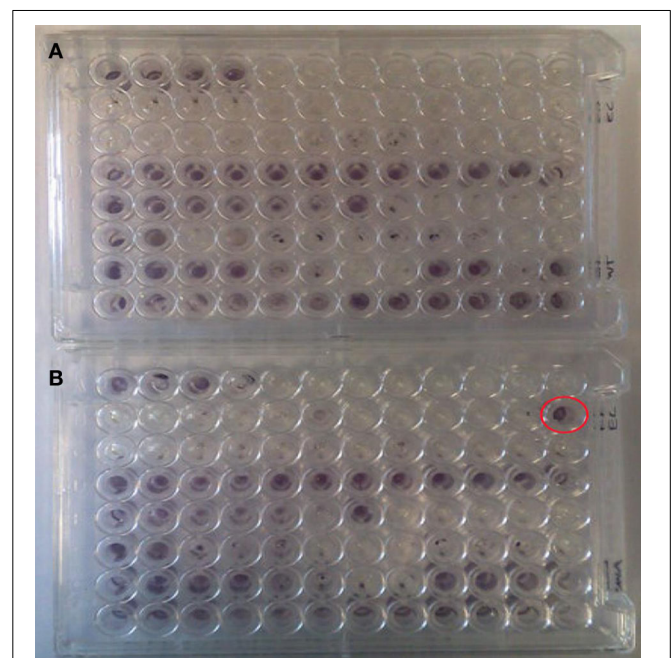

FIGURE 2 | Appearance of PM 9 plates after incubation for 24 hours at 37°C. (A) Control EPI300::pCC1FOS and **(B)** SMG 9. PM plates measure cellular respiration colorimetrically via reduction of a tetrazolium dye with electrons from NADH generated during the process of respiration. Strongly metabolized substrates generate a more intense purple color. Development of a strong purple color can be seen in well B12 in **Figure 2B** (circled in red), which was inoculated with SMG 9, while no color development is visible in B12 of the control plate. This indicates SMG 9 has a greater ability to transport and utilise L-carnitine compared to the EPI300::pCC1FOS host strain.

9 displays a growth defect, while growth is similar under all other conditions. The growth defect is alleviated when grown at 4% w/v NaCl + 1 mM L-carnitine and there is no difference in growth between clones either in the presence or absence of L-carnitine (**Figure 4B**). At 5% w/v NaCl however, SMG 9 has a significant growth advantage compared to EPI300::pCC1FOS both in the presence and absence of 1 mM L-carnitine (**Figure 4C**). The positive effect of L-carnitine on the growth of SMG 9 is evident with cells entering logarithmic phase growth sooner and reaching a

much higher final optical density ($OD_{595\,nm}$). A similar, positive effect for L-carnitine on growth of SMG 9 is also seen at 6% w/v NaCl (**Figure 4D**).

SEQUENCING OF SMG 9 FOSMID INSERT AND ANALYSIS

Fosmid SMG 9 was fully sequenced (454-pyrosequencing) and assembled by GATC Biotech. This generated a total of 2.3×10^6 base pairs of sequence data in 6939 sequencing reads. The average read length was 334 base pairs and coverage of $64.5\times$ was achieved. Following vector trimming the length of the insert was approximately 36.5 kb and the %G+C content was 49.43%. Twenty-four putative open reading frames were predicted using Softberry's FGENESB, bacterial operon and gene prediction software (www.softberry.com) (Mavromatis et al., 2007). Translated nucleotide sequences were functionally annotated by homology searches using the BLASTP program to identify homologous sequences and determine their taxonomic origin. All but two of the genes encoded proteins with high identity (98–100% at the amino acid level) to *Bacteroides* sp. CAG:545. A list of the genes on SMG 9, their encoded functions and putative domains are presented in **Table 2**. The full fosmid insert sequence of SMG 9 has been submitted to GenBank and assigned the accession number, KJ524644.

CLONING OF *mfsT* AND *sdtR* GENES

Following initial inspection of the encoded proteins on SMG 9, the presence of an L-carnitine or general osmoprotectant transporter, nor indeed any protein with a functional link to carnitine metabolism was not immediately obvious. Transposon mutagenesis was attempted in order to create a knock-out mutant; this
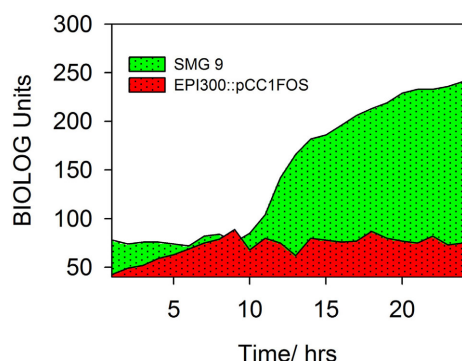

FIGURE 3 | Kinetic data measured by BIOLOG Omnilog system. Color formation within each well was measured by BIOLOG's Omnilog machine, which produces a color coded graph. Kinetic data from two clones can be compared. EPI300::pCC1FOS is shown in red and that from SMG 9 is shown in green. The green color indicates more rapid metabolism by SMG 9 under the conditions in the well (6% NaCl + L-carnitine).
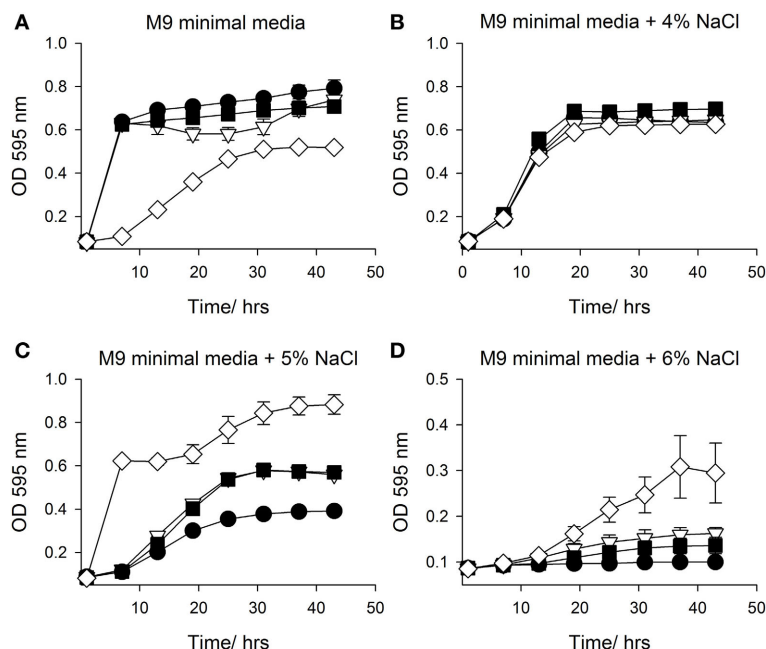

FIGURE 4 | Growth in M9 minimal media with NaCl +/− 1mM L-carnitine. Growth of *E. coli* EPI300::pCC1FOS and SMG 9 in **(A)** M9 minimal media, **(B)** M9 minimal media + 4% NaCl, **(C)** M9 minimal media + 5% NaCl and **(D)**

M9 minimal media + 6% NaCl. **Legend:** *E. coli* EPI300::pCC1FOS (● closed circle); SMG 9 (▽ open triangle); *E. coli* EPI300::pCC1FOS + 1mM L-carnitine (■ closed square); SMG 9 + 1 mM L-carnitine, (◇ open diamond).

Table 2 | Proteins predicted to be encoded on SMG 9 fosmid insert.

Gene #	Putative encoded function	# a.a	Closest hit organism	% Coverage	e-value	% ID	Domains
1	ATP-dependent chaperone ClpB	554	*Bacteroides* sp. CAG:545	97	0.00E + 00	99	COG0714; AAA ATPase
2	Preprotein translocase SecG subunit	121	*Bacteroides* sp. CAG:545	100	3.00E-77	100	SecG
3	Putative uncharacterized protein	187	*Bacteroides* sp. CAG:545	100	3.00E-134	100	None detected
4	Putative uncharacterized protein	177	*Bacteroides* sp. CAG:545	86	3.00E-108	99	LptE
5	Transcriptional regulator	432	*Bacteroides* sp. CAG:545	100	0.00E + 00	99	AAA ATPase; σ^{54} interaction domain; HTH_8 bacterial reg protein, Fis family domain
6	Putative uncharacterized protein	545	*Bacteroides* sp. CAG:545	100	0.00E + 00	99	TadD
7	Putative uncharacterized protein	1015	*Bacteroides* sp. CAG:545	100	0.00E + 00	99	SecD, SecF
8	OmpA/MotB domain protein	618	*Bacteroides* sp. CAG:545	71	0.00E + 00	99	PD40, similar to WD40 domain
9	Putative uncharacterized protein	155	*Bacteroides* sp. CAG:545	100	4.00E-106	100	NfeD
10	uPF0365 protein AL1_06760	317	*Bacteroides* sp. CAG:545	100	0.00E + 00	99	YdfA_immunity superfamily
11	Putative uncharacterized protein	142	*Bacteroides* sp. CAG:545	100	1.00E-97	98	Lipocalin_4
12	Subtilisin-like serine protease	678	*Bacteroides* sp. CAG:545	100	0.00E + 00	99	Peptidase_S8_S53 superfamily
13	Uncharacterized protein	812	*Bacteroides* sp. CAG:545	99	0.00E + 00	99	None detected
14	RagB/SusD family protein	547	*Bacteroides* sp. CAG:545	100	0.00E + 00	99	Two SusD superfamily
15	Outer membrane receptor for ferrienterochelin and colicins	1068	*Bacteroides* sp. CAG:545	100	0.00E + 00	100	Can_B2; Plug; OM channel; OMP_RagA_SusC
16	Alpha-L-fucosidase-like	513	*Bacteroides* sp. CAG:545	100	0.00E + 00	98	COG3669; Alpha_L_fucos; F5_F8_Type_C
17	Putative uncharacterized protein	446	*Bacteroides* sp. CAG:545	100	0.00E + 00	99	DHQ_FE-ADH (Dehydroquinate iron aldehyde dehydrogenase)
18	Major facilitator superfamily MFS_1	487	*Bacteroides* sp. CAG:545	100	0.00E + 00	100	MFS; UhpC, sugar phosphate permease
19	ThiF family protein	242	*Bacteroides* sp. CAG:545	100	4.00E-174	98	YgdL_like
20	Putative uncharacterized protein	649	*Bacteroides* sp. CAG:545	100	0.00E + 00	98	Glyco_hydro_97
21	DNA mismatch repair protein MutS	894	*Bacteroides* sp. CAG:545	98	0.00E + 00	99	PRK05399; MutS-I; MutS_II; MutS_III; ABC_MutS_1
22	Glycosyl transferase group 1	378	*Bacteroides* sp. CAG:545	100	0.00E + 00	99	RfaG
23	No significant similarity found	68	N/A	N/A	N/A	N/A	N/A
24	Hypothetical protein Fjoh_3657	162	*Flavobacterium johnsoniae* UW101	80	4.00E-25	40	AdkA, archaeal adenylate kinase

Functional assignment is based on BLASTP of amino acid sequences predicted by Softberry's FGENESB. Abbreviations: # a.a, number of predicted amino acids; %ID, % identity at the amino acid level; N/A, not applicable.

however, proved unsuccessful. Two genes (gene 5 and gene 18), which we felt may be likely to have a possible role in L-carnitine utilization based on bioinformatic analysis were cloned in isolation to further examine the phenotype. Genes 5 and 18 were annotated *sdtR* for *s*igma-*d*ependent *t*ranscriptional *r*egulator and *mfsT* for *m*ajor *f*acilitator *s*uperfamily *t*ransporter, respectively.

Both *mfsT* and *sdtR* were cloned in the vector pCI372 and expressed in both *E. coli* EPI300 and the osmosensitive strain *E. coli* MKH13 (Haardt et al., 1995). The effect of each gene on the growth of each strain under salt stress in the presence and absence of L-carnitine was assessed. The *mfsT* gene had no effect on growth under any of the conditions tested (data not shown). The *sdtR* gene on the other hand conferred a significant salt tolerance phenotype to *E. coli* MKH13 when grown in media supplemented with both 3 and 4% w/v NaCl (**Figures 5C,D**, respectively), while growth was similar in media lacking NaCl and in media supplemented with 2% w/v NaCl (**Figures 5A,B**, respectively). Cloning and expression of *sdtR* in EPI300 resulted in an increase in salt tolerance compare to wild-type EPI300 carrying an empty copy of the plasmid pCI372. Addition of 1 mM L-carnitine increased the growth rate and final optical density of both strains, but its effect on the *sdtR*+ strain was not significant relative to the EPI300::pCI372 control (**Figures 6A,B**).

Although *sdtR* did confer a salt tolerance phenotype, neither of the cloned genes replicated the original phenotype related to L-carnitine. We re-examined clone SMG 9 and carried out further studies, however these revealed that SMG 9 had lost the carnitine-associated phenotype seen originally and unfortunately, it was therefore not possible to identify the gene(s) responsible.

DISCUSSION

In the present study we have identified a novel salt tolerance gene from a metagenomic library clone from the human gut microbiome using a combined functional metagenomic and PM approach. The clone, SMG 9, was identified from a previous library screen to identify salt tolerant clones (Culligan et al., 2012) and was further characterized in this study using PM osmolyte plates. From the PM screen, SMG 9 showed an increased metabolic profile in the presence of 6% NaCl + 1 mM L-carnitine compared to the control strain carrying an empty fosmid vector (EPI300::pCC1FOS), indicating this clone could utilize or transport L-carnitine. Experiments to confirm the findings of the PM assay showed that SMG 9 displayed an increased growth profile at 5% and 6% NaCl in the presence of 1 mM L-carnitine compared to controls, similar to observations in the PM assay. Transposon mutagenesis was attempted to create phenotypic knock out mutants, using the EZTn5 system (Epicentre Biotechnologies; Goryshin and Reznikoff, 1998) but this proved unsuccessful. This may be due to the presence of a gene encoding a DNA repair protein MutS on the fosmid insert, which has been associated with transposon excision (specifically Tn5 and Tn10) (Lundblad and Kleckner, 1985).

Next generation sequencing of the full fosmid insert of SMG 9 and functional assignment of the encoded proteins using BLASTP revealed sequences shared highest genetic identity to *Bacteroides* sp. CAG:545. Species of *Bacteroides* are commonly found in the human gut, where the resident microbiota is largely composed of species from two dominant phyla, the *Bacteroidetes* and *Firmicutes* (Qin et al., 2010). The %G+C content of the SMG 9 insert was 49.44% which close to the reported range of 40–48% for genomes
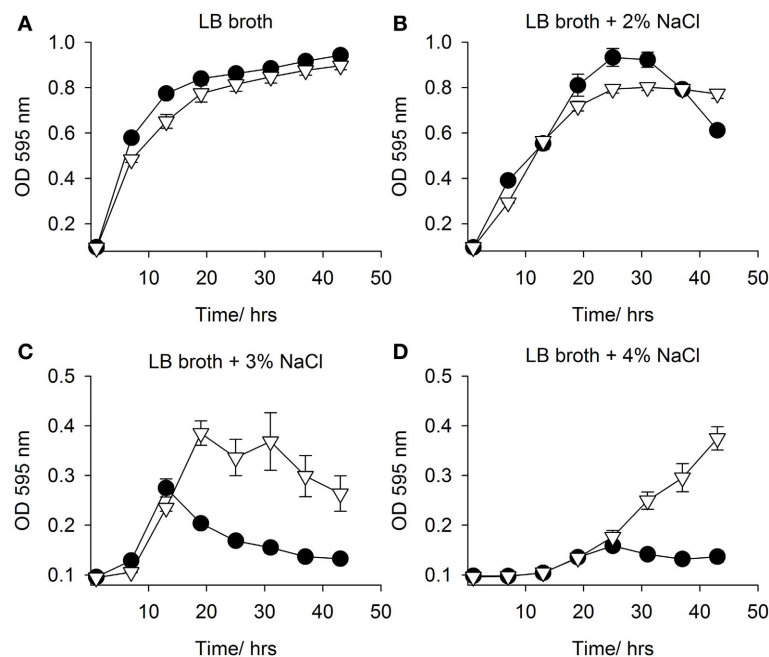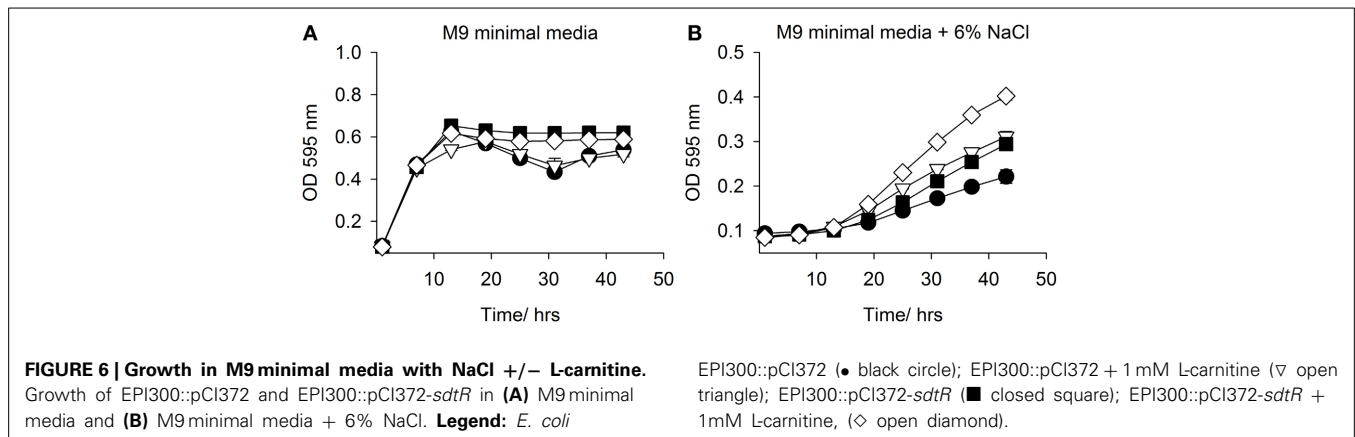

FIGURE 5 | Growth in LB broth with NaCl. Growth in of *E. coli* MKH13::pCI372 (● black circle) and MKH13::pCI372-*sdtR* (▽ open triangle) in **(A)** LB broth and LB broth supplemented with **(B)** 2% NaCl, **(C)** 3% NaCl and **(D)** 4% NaCl.

FIGURE 6 | Growth in M9 minimal media with NaCl +/− L-carnitine. Growth of EPI300::pCI372 and EPI300::pCI372-*sdtR* in **(A)** M9 minimal media and **(B)** M9 minimal media + 6% NaCl. **Legend:** *E. coli* EPI300::pCI372 (● black circle); EPI300::pCI372 + 1 mM L-carnitine (▽ open triangle); EPI300::pCI372-*sdtR* (■ closed square); EPI300::pCI372-*sdtR* + 1mM L-carnitine, (◇ open diamond).

of species of *Bacteroides* (Shah, 1992). Sequencing and subsequent functional annotation did not reveal any obvious genes related to known L-carnitine transport or utilization systems, suggesting a novel mechanism may be involved. We conducted further bioinformatic analysis of the encoded proteins in an attempt to identify any link to salt tolerance or osmoprotectant transport. Gene 5 (*sdtR*) is predicted to encode a σ[54]-dependent transcriptional regulator, which contains a number of domains (see **Table 2**), including a helix-turn-helix 8 (HTH_8), Fis-family protein domain, while gene 18 (*mfsT*) is predicted to encode a major facilitator superfamily (MFS) protein with an UhpC sugar phosphate permease domain. Both genes were chosen for further study and cloning as Fis is a regulatory protein involved the regulation of proline (another important osmoprotectant) uptake (Xu and Johnson, 1995, 1997; Typas et al., 2007), while we further reasoned that *sdtR* could be regulating host EPI300 genes, contributing the L-carnitine-associated phenotype, while MFS transporters also play a role in osmoprotectant uptake (Culham et al., 1993; Haardt et al., 1995; Pao et al., 1998). Despite the presence of a sugar phosphate permease domain, indicating sugar transport, MFS transporters are known to have a diverse substrate range (Pao et al., 1998; Saier, 2000; Law et al., 2008).

The *mfsT* gene did not confer a salt tolerance or the L-carnitine associated phenotype to transformed cells (data not shown). The *sdtR* gene also did not confer the L-carnitine associated phenotype to *E. coli*, but did however confer an increased salt tolerance phenotype. *sdtR* therefore represents a novel salt tolerance gene and most likely functions by influencing expression (either positively or negatively) of host *E. coli* genes, although further work, comprising expression studies and microarray analysis, will be required to elucidate the genes involved as transcriptional regulators can influence a wide variety of genes. Transcriptional regulators are commonly involved in the response different stresses in bacteria (Hengge-Aronis et al., 1991; Cheville et al., 1996; Battesti et al., 2011; Hoffmann et al., 2013), while σ[54] (RpoN) has been shown to play a role in osmotolerance in *Listeria monocytogenes* (Okada et al., 2006).

The loss of the carnitine-associated phenotype of SMG 9 prevented further characterization of this clone and ultimately the identification of the gene(s) responsible. It is difficult to pinpoint the cause of this phenotypic reversion, but a clue may be

evident from **Figure 4A**, where a growth defect for SMG 9 is apparent when grown in M9MM + 1 mM L-carnitine. This indicates L-carnitine may be increasing the metabolic load on the cell and this metabolic stress is only relieved in the presence of NaCl, when L-carnitine may be utilized efficiently in an osmoprotective capacity. If the gene is constitutively expressed, a mutation may have occurred to counteract this phenomenon. The presence of a gene encoding MutS may also be relevant as mutations to MutS can result in a mutator phenotype in *E. coli* cells (Wu and Marinus, 1994). Furthermore, it is possible the original SMG 9 clone acquired a mutation on the fosmid insert that conferred the carnitine-associated phenotype and a subsequent suppressor mutation occurred to silence this mutation, returning the clone to its original phenotype.

In conclusion, we have identified a novel salt tolerance gene from the human gut microbiome using a combined functional metagenomic and PM approach. The gene originates from a species of *Bacteroides* and encodes a putative transcriptional regulator. Overall this study demonstrates the utility of functional metagenomics and phenomics for novel gene discovery and functional characterization of metagenome-derived clones.

REFERENCES

Banik, J. J., and Brady, S. F. (2008). Cloning and characterization of new glycopeptide gene clusters found in an environmental DNA megalibrary. *Proc. Natl. Acad. Sci. U.S.A.* 105, 17273–17277. doi: 10.1073/pnas.080 7564105.

Battesti, A., Majdalani, N., and Gottesman, S. (2011). The RpoS-mediated general stress response in *Escherichia coli*. *Annu. Rev. Microbiol.* 65, 189–213. doi: 10.1146/annurev-micro-090110-102946

Beja, O., Aravind, L., Koonin, E. V., Suzuki, M. T., Hadd, A., Nguyen, L. P., et al. (2000). Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* 289, 1902–1906. doi: 10.1126/science.289.5486.1902

Bochner, B. R. (2009). Global phenotypic characterization of bacteria. *FEMS Microbiol. Rev.* 33, 191–205. doi: 10.1111/j.1574-6976.2008.00149.x

Bochner, B. R., Gadzinski, P., and Panomitros, E. (2001). Phenotype microarrays for high-throughput phenotypic testing and assay of gene function. *Genome Res.* 11, 1246–1255. doi: 10.1101/gr.186501

Bork, P. (2000). Powers and pitfalls in sequence analysis: the 70% hurdle. *Genome Res.* 10, 398–400. doi: 10.1101/gr.10.4.398

Cheville, A. M., Arnold, K. W., Buchrieser, C., Cheng, C. M., and Kaspar, C. W. (1996). rpoS regulation of acid, heat, and salt tolerance in *Escherichia coli* O157:H7. *Appl. Environ. Microbiol.* 62, 1822–1824.

Culham, D. E., Lasby, B., Marangoni, A. G., Milner, J. L., Steer, B. A., Van Nues, R. W., et al. (1993). Isolation and sequencing of Escherichia coli gene proP reveals unusual structural features of the osmoregulatory proline/betaine transporter, ProP. *J. Mol. Biol.* 229, 268–276. doi: 10.1006/jmbi.1993.1030

Culligan, E. P., Sleator, R. D., Marchesi, J. R., and Hill, C. (2012). Functional metagenomics reveals novel salt tolerance loci from the human gut microbiome. *ISME J.* 6, 1916–1925. doi: 10.1038/ismej.2012.38

Culligan, E. P., Sleator, R. D., Marchesi, J. R., and Hill, C. (2013). Functional environmental screening of a metagenomic library identifies *stlA*; a unique salt tolerance locus from the human gut microbiome. *PLoS ONE* 8:e82985. doi: 10.1371/journal.pone.0082985

Epstein, W. (2003). The roles and regulation of potassium in bacteria. *Prog. Nucleic Acid Res. Mol. Biol.* 75, 293–320. doi: 10.1016/S0079-6603(03)75008-9

Gillespie, D. E., Brady, S. F., Bettermann, A. D., Cianciotto, N. P., Liles, M. R., Rondon, M. R., et al. (2002). Isolation of antibiotics turbomycin a and B from a metagenomic library of soil microbial DNA. *Appl. Environ. Microbiol.* 68, 4301–4306. doi: 10.1128/AEM.68.9.4301-4306.2002

Goryshin, I. Y., and Reznikoff, W. S. (1998). Tn5 *in vitro* transposition. *J. Biol. Chem.* 273, 7367–7374.

Haardt, M., Kempf, B., Faatz, E., and Bremer, E. (1995). The osmoprotectant proline betaine is a major substrate for the binding-protein-dependent transport system ProU of *Escherichia coli* K-12. *Mol. Gen. Genet.* 246, 783–786.

Hayes, F., Daly, C., and Fitzgerald, G. F. (1990). Identification of the minimal replicon of *Lactococcus lactis* subsp. lactis UC317 Plasmid pCI305. *Appl. Environ. Microbiol.* 56, 202–209.

Hengge-Aronis, R., Klein, W., Lange, R., Rimmele, M., and Boos, W. (1991). Trehalose synthesis genes are controlled by the putative sigma factor encoded by rpoS and are involved in stationary-phase thermotolerance in *Escherichia coli*. *J. Bacteriol.* 173, 7918–7924.

Hoffmann, R. F., Mclernon, S., Feeney, A., Hill, C., and Sleator, R. D. (2013). A single point mutation in the listerial betL sigma (A) -dependent promoter leads to improved osmo- and chill-tolerance and a morphological shift at elevated osmolarity. *Bioengineered* 4, 401–407. doi: 10.4161/bioe.24094

Jones, B. V., Begley, M., Hill, C., Gahan, C. G., and Marchesi, J. R. (2008). Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13580–13585. doi: 10.1073/pnas.0804437105

Jones, B. V., and Marchesi, J. R. (2007). Transposon-aided capture (TRACA) of plasmids resident in the human gut mobile metagenome. *Nat. Methods* 4, 55–61. doi: 10.1038/nmeth964

Kempf, B., and Bremer, E. (1998). Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Arch. Microbiol.* 170, 319–330.

Kunte, H. J. (2006). Osmoregulation in bacteria: compatible solute accumulation and osmosensing. *Environ. Chem.* 3, 94–99. doi: 10.1071/EN06016

Law, C. J., Maloney, P. C., and Wang, D. N. (2008). Ins and outs of major facilitator superfamily antiporters. *Annu. Rev. Microbiol.* 62, 289–305. doi: 10.1146/annurev.micro.61.080706.093329

Lundblad, V., and Kleckner, N. (1985). Mismatch repair mutations of *Escherichia coli* K12 enhance transposon excision. *Genetics* 109, 3–19.

Mavromatis, K., Ivanova, N., Barry, K., Shapiro, H., Goltsman, E., Mchardy, A. C., et al. (2007). Use of simulated data sets to evaluate the fidelity of metagenomic processing methods. *Nat. Methods* 4, 495–500. doi: 10.1038/nmeth1043

Okada, Y., Okada, N., Makino, S., Asakura, H., Yamamoto, S., and Igimi, S. (2006). The sigma factor RpoN (sigma54) is involved in osmotolerance in Listeria monocytogenes. *FEMS Microbiol. Lett.* 263, 54–60. doi: 10.1111/j.1574-6968.2006.00405.x

Pao, S. S., Paulsen, I. T., and Saier, M. H. Jr. (1998). Major facilitator superfamily. *Microbiol. Mol. Biol. Rev.* 62, 1–34.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65. doi: 10.1038/nature08821

Saier, M. H. Jr. (2000). Families of transmembrane sugar transport proteins. *Mol. Microbiol.* 35, 699–710. doi: 10.1046/j.1365-2958.2000.01759.x

Shah, H. (1992). "The genus bacteroides and related taxa," in *The Prokaryotes*, eds A. Balows, H. Trüper, M. Dworkin, W. Harder, and K.-H. Schleifer (New York, NY: Springer), 3593–3607.

Sleator, R. D., Francis, G. A., O'Beirne, D., Gahan, C. G., and Hill, C. (2003). Betaine and carnitine uptake systems in Listeria monocytogenes affect growth and survival in foods and during infection. *J. Appl. Microbiol.* 95, 839–846. doi: 10.1046/j.1365-2672.2003.02056.x

Sleator, R. D., and Hill, C. (2002). Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. *FEMS Microbiol. Rev.* 26, 49–71. doi: 10.1111/j.1574-6976.2002.tb00598.x

Sleator, R. D., Wouters, J., Gahan, C. G., Abee, T., and Hill, C. (2001). Analysis of the role of OpuC, an osmolyte transport system, in salt tolerance and virulence potential of Listeria monocytogenes. *Appl. Environ. Microbiol.* 67, 2692–2698. doi: 10.1128/AEM.67.6.2692-2698.2001

Typas, A., Stella, S., Johnson, R. C., and Hengge, R. (2007). The -35 sequence location and the Fis-sigma factor interface determine sigmas selectivity of the proP (P2) promoter in Escherichia coli. *Mol. Microbiol.* 63, 780–796. doi: 10.1111/j.1365-2958.2006.05560.x

Wargo, M. J., and Hogan, D. A. (2009). Identification of genes required for *Pseudomonas aeruginosa* carnitine catabolism. *Microbiology* 155, 2411–2419. doi: 10.1099/mic.0.028787-0

Wemekamp-Kamphuis, H. H., Sleator, R. D., Wouters, J. A., Hill, C., and Abee, T. (2004). Molecular and physiological analysis of the role of osmolyte transporters BetL, Gbu, and OpuC in growth of Listeria monocytogenes at low temperatures. *Appl. Environ. Microbiol.* 70, 2912–2918. doi: 10.1128/AEM.70.5.2912-2918.2004

Wu, T. H., and Marinus, M. G. (1994). Dominant negative mutator mutations in the mutS gene of *Escherichia coli*. *J. Bacteriol.* 176, 5393–5400.

Xu, J., and Johnson, R. C. (1995). Fis activates the RpoS-dependent stationary-phase expression of proP in *Escherichia coli*. *J. Bacteriol.* 177, 5222–5231.

Xu, J., and Johnson, R. C. (1997). Activation of RpoS-dependent proP P2 transcription by the Fis protein *in vitro*. *J. Mol. Biol.* 270, 346–359. doi: 10.1006/jmbi.1997.1133

Yoon, M. Y., Lee, K. M., Yoon, Y., Go, J., Park, Y., Cho, Y. J., et al. (2013). Functional screening of a metagenomic library reveals operons responsible for enhanced intestinal colonization by gut commensal microbes. *Appl. Environ. Microbiol.* 79, 3829–3838. doi: 10.1128/aem.00581-13

Application of high-throughput sequencing in understanding human oral microbiome related with health and disease

Hui Chen and Wen Jiang*

Department of Conservative Dentistry and Periodontics, Affiliated Hospital of Stomatology, College of Medicine, Zhejiang University, Hangzhou, China

*Correspondence:
Hui Chen, Department of Conservative Dentistry and Periodontics, Affiliated Hospital of Stomatology, College of Medicine, Zhejiang University, No. 395 Yanan Road, Hangzhou, Zhejiang 310006, China
e-mail: huic66@hotmail.com

The oral microbiome is one of most diversity habitat in the human body and they are closely related with oral health and disease. As the technique developing, high-throughput sequencing has become a popular approach applied for oral microbial analysis. Oral bacterial profiles have been studied to explore the relationship between microbial diversity and oral diseases such as caries and periodontal disease. This review describes the application of high-throughput sequencing for characterization of oral microbiota and analyzing the changes of the microbiome in the states of health or disease. Deep understanding the knowledge of microbiota will pave the way for more effective prevent dentistry and contribute to the development of personalized dental medicine.

Keywords: oral microbiome, high-throughput sequencing, dental caries, periodontitis, apical periodontitis

INTRODUCTION

Improvements in bio-technologies have spurred a large number of studies aimed at obtaining a better understanding of the composition and effect in microbiota and its associations with various human diseases. Early studies of the human microbiome have revealed most of the microbes that occupy in different habitats of human body and are ~10 times more numerous than our own cells, more attention has paid to viewing ourselves as a supraorganism (Gill et al., 2006).

Traditional culture independent methods, such as DNA–DNA hybridization or cloning sequencing of DNA is widely used to identify oral organisms and found over 700 bacterial species in oral cavity (Aas et al., 2005). However, they still have significant biases that do not allow microbial diversity to be fully studied, as many low richness species cannot be detected.

Recently a major advance over conventional sequencing techniques is the development of high-throughput sequencing methods, which is a part of the next-generation sequencing (NGS) techniques (Ronaghi, 2001). These methods can largely be grouped into three main types: sequencing by synthesis (Roche 454 pyrosequencing, Illumina, The Ion Torrent system), sequencing by ligation (SOLiD, Polonator G.007 system), and single-molecule sequencing (Helicos, Pacific BioSciences). This has been proving

amenable for use in massively parallel signature sequencing technologies (Rothberg and Leamon, 2008), including Roche 454 genome sequencers, Illumina sequencers, Applied Biosystems SOLiD sequencer, Life Technologies Ion Torrent, Helicos biosciences HeliScope, Pacific Biosciences SMRT DNA sequencer. The most frequently used methods are the 454 pyrosequencing (Roche, Bradford, CT, USA), Illumina (Illumina, San Diego, CA, USA), and SOLiD (Applied Biosystems, Foster City, CA, USA). Each of them had their own character. Nucleotide detection in Illumina and SOLiD systems is performed one at a time. As a result, homopolymer regions can be accurately sequenced. Second advantage of them is their high output per run compared to 454 pyrosequencing, which lead it soon became a workhorse for whole-genome resequencing applications and for exploring metabolic processing potential and pathway representation in health and disease. Owning to its optical signal decay and dephasing, these systems have relative short-read length (Zhou et al., 2010). On the other hand, the 454 pyrosequencing platform had long read length and relatively short run time. Furthermore it does not need to carry out an extra chemical deblocking step, which would reduce the chances of premature chain termination and non-simultaneous extension (Metzker, 2010; Zhou et al., 2010). Although the drawback of relatively high cost per

megabase sequencing output, the 454 pyrosequencing has still become one of the most prevalent used worldwide (Rogers and Venter, 2005).

Oral cavity is one of the indispensable parts of the human microbiome habitat. The oral bacterial community dynamics is complicated and still far way to full understand. In this paper, we reviewed the human oral microbiome composite and its shifting which related with health and disease by application of high-throughput sequencing.

THE CONCEPT OF HUMAN ORAL MICROBIOME

Oral microbiome, which is referred to as the oral microflora or oral microbiota, is defined as all the microorganisms residing in the human oral cavity and their collective genome. It was firstly coined by Lederberg and Mccray (2001) "to signify the ecological community of commensal, symbiotic, and pathogenic microorganisms that literally share our body space and have been all but ignored as determinants of health and disease." The oral cavity harbors one of the most diverse microbiomes in the human body. And bacteria predominated in the oral cavity, while others are in relatively low proportions in the most circumstance.

Oral microbiome harbors on teeth, gingival sulcus, tongue, cheeks, hard and soft palates, and tonsils and it is a critical component of oral health and disease.

THE DIVERSITY AND COMPOSITION OF ORAL MICROBIOME

Oral cavity, which is one of the largest and most complex human-associated microbial habitats, harbors large numbers of bacteria that can have important effects on health. During the past 40 years, a wealth of knowledge has been gathered about these bacteria: over 250 oral species have been isolated and characterized by cultivation, and over 450 species have been identified by culture-independent molecular approaches (Paster et al., 2006). By 454 pyrosequencing, we studied in 120 children at age 6 and found that there were about 2,000 phylotypes by clustering at 3% dissimilarity level in each sample of caries-active and caries-free children; and there were more phylotypes in saliva than in dental plaque samples from the same groups. At the genus level, sequences from saliva and plaque represented 203 different genera, while 153 different genera were found in dental plaque (120 genera in caries-active samples and 116 genera in caries-free samples) and 156 different genera were found in saliva (132 genera in caries-active samples and 115 genera in caries-free samples; Ling et al., 2010). No gender differences in oral microbiome diversity were detected. By using whole-metagenome sequencing approaches, Firmicutes, Actinobacteria, Bacteroidetes, Fusobacteria, and Proteobacteria were found account for 80–95% of the entire oral microbiome. At the genus level, a total of 58 distinct genera are present at an 0.1% abundance. The most abundant genera comprise previously characterized oral bacteria: *Actinomyces, Prevotella, Streptococcus, Fusobacterium, Leptotrichia, Corynebacterium, Veillonella, Rothia, Capnocytophaga, Selenomonas, Treponema*, and TM7 genera 1 and 5 (Liu et al., 2012).

THE COMPOSITION OF ORAL MICROBIOTA IN HEALTH

Oral microbiome in 60 caries free children were analyzed with 454 pryosequencing in one of our studies (Ling et al., 2010). The result agreed with the other studies more and less, found that more than fourteen phyla, including Firmicutes, Bacteroidetes, Proteobacteria, Actinobacteria, Spirochaetes, and Fusobacteria, Euryarchaeota, Chlamydia, Chloroflexi, SR1, Synergistetes, Tenericutes, Cyanobacteria, OD2, and TM7 in healthy subjects (Zaura et al., 2009; Bik et al., 2010; Griffen et al., 2011). Among them, the vast majority (containing more than 80% of the taxa) of oral bacteria belong to Firmicutes, Bacteroidetes, Proteobacteria, Actinobacteria, Spirochaetes, and Fusobacteria (Ling et al., 2010). At genus level, over 200 genera were found in the oral microbiota. The most abundant genera include *Streptococcus, Prevotella, Neisseria, Haemophilus, Porphyromonas, Gemella, Rothia, Granulicatella, Fusobacterium, Actinomyces*, and *Veillonella*. At species level, it had been estimated that the number of species–level phylotypes were between 500 and 10000 (Keijser et al., 2008; Lazarevic et al., 2009), and each oral niches harbored 266 "species-level" phylotypes on average (Zaura et al., 2009). It is higher than the previously reported 10–81 species per site by using a 16S rRNA gene-based microarray (Preza et al., 2009). In Hasan et al.'s (2014) study, the microbiota of human saliva was analyzed by using the Illumina GAIIx and HiSeq 2000 instrument, and more than 175 bacterial species were found at >90% accuracy, including bacteria *Haemophilus influenzae, Neisseria meningitidis, Streptococcus pneumoniae*, and *Gammaproteobacteria*.

THE CORE MICROBIOME IN HEALTH

In our study, there were 3530 OTUs (operational taxonomic unit) shared with the oral microbiota in each of the intact enamel surfaces of 60 children. Five phyla including Proteobacteria, Firmicutes, Actinobacteria, Fusobacteria, and Bacteroidetes were found in healthy individuals, while nine genera were common in all subjects, including *Actinomyces, Capnocytophaga, Corynebacterium, Derxia, Leptotrichia, Neisseria, Prevotella, Streptococcaceae Streptococcus*, and *Veillonella*. Thus, in agreement with other studies it was proposed that there might be a core microbiome in the oral environment (Shade and Handelsman, 2012; Jiang et al., 2014; Xu et al., 2014). The core microbiome is shared with most of individuals and comprised of the predominant species in healthy conditions of oral cavity (Zarco et al., 2012).

In Zaura's study, 26% of the unique sequences, 47% of the OTUs were shared with oral microbiome in each sample of three healthy subjects. At the higher taxonomic levels, 72% of all taxa (genus level or above) were common to the oral microbiome of three adults, contributing to 99.8% of all reads. The result also suggested the existence of a core microbiome (Zaura et al., 2009). Moreover, Lazarevic et al. (2010) found salivary microbial community appeared to be stable at different time points (from 5 to 29 days), supported the concept of a core microbiome in health state. However, It is just the beginning of the understanding whether there is the core microbiome in different oral niches and needed more researches to keep verifying this concept.

THE SITE-SPECIFICITY MICROBIOME

According to the Huse' research, oral bacterial microbiota might be site-specificity and showed the different richness. Hard palate

showed the lowest estimate of total richness, while the gingival plaque showed the highest estimate of total richness. The genus *Corynebacterium* had at least eight OTUs with five different profiles and colonized in different niches. For instance, *Corynebacterium matruchotii* was present almost exclusively in the supragingival plaque, while *Corynebacterium argentoratense* mostly in saliva and to a lesser extent on the hard palate (Huse et al., 2012). Further study also found similar result. It may due to the shedding of the epithelial cells and the shear forces from chewing in the buccal fold and the hard palate. Genera *Eubacterium* and *Prevotella* showed a significant association with the tongue dorsum. The papillary structure and the low redox potential of its surface might explain its significant site-specific bacterial association. *Lautropia mirabilis* was the only species significantly associated with the supragingival plaque, while *Treponema socranskii* was found only in the subgingival plaque (Preza et al., 2009). The anaerobic environment of subgingival plaque may explain their significant site-specific association. In the oropharynx, the distribution of Firmicutes, Proteobacteria, and Bacteroidetes was similar to that of saliva, and more Proteobacteria than that in the mouth (Lemon et al., 2010).

THE MICROBIOME VARYING DURING DIFFERENT PERIODS OF AGE
The oral microbiota is various during different periods of age, which colonizes in the oral cavity and changes with the developmental status such as primary and permanent tooth. *Veillonella, Neisseria, Rothia, Haemophilus, Gemella, Granulicatella, Leptotrichia,* and *Fusobacterium* were predominant genera in infant samples, while *Haemophilus, Neisseria, Veillonella, Fusobacterium, Oribacterium, Rothia, Treponema,* and *Actinomyces* were present at higher levels in their parents. Saliva bacterial microbiome in adults had greater bacteria diversity than that in infants (Cephas et al., 2011). Crielaard et al. (2011) found a higher proportion of Proteobacteria (*Gammaproteobacteria, Moraxellaceae*) than that of Bacteroidetes in the deciduous dentition and Bacteroidetes (mainly genus *Prevotella*), *Veillonellaceae* family, Spirochaetes, and candidate division TM7 increased with increasing age. It may reflect variation of oral microbiome driven by biological changes with age (Crielaard et al., 2011). And by comparing salivary microbiota from healthy children and adults, it was found that the mean level of seven genera including *Moraxella, Leptotrichia, Peptostreptococcus, Eubacterium, Neisseriaceae, Flavobacteriaceae,* and SR1 were significant differences between children and adults in another research, implying the microbiome shifts during different ages (Ling et al., 2013).

ORAL MICROBIOTA IN DENTAL CARIES
THE ORAL MICROBIOTA IN CARIES
Dental caries is one of the most prevalent worldwide chronic infectious diseases (Petersen et al., 2005). The desire of a core theme in studying the characterization of the oral microbiota has being pursued to understanding the particular organisms with tooth decay in a way that implies causation. Most researches have suggested that *Streptococcus mutans* is the major pathogen of dental caries, for it is the most frequently detected bacteria in the caries lesions (Hamada and Slade, 1980; Loesche, 1986; Matee et al., 1992).

However, some recent studies indicate that the relationship between MS and caries is not absolute: high proportions of MS may persist on tooth surfaces without lesion development, and caries can develop in the absence of *Streptococcus mutans* (Bowden, 1997). And very recently, researches found that acidogenic and aciduric bacteria other than MS, are responsible for the initiation of caries (Kashket et al., 1996). As new pyosequencing technique applied in the oral microbiology, ever greater numbers of bacteria have been identified as being associated with caries. In our recent study, applying high-throughput barcoded pyrosequencing combined with PCR-denaturing gradient gel electrophoresis, around 120 genera were found in the oral microbiota of saliva and supragingival plaques from children aged 3–6 years old with and without dental caries. Our study showed that oral microbiota in children was far more diverse than previous studies reported and more than 200 genera belonging to 10 phyla were found in the oral cavity. Six genera (*Streptococcus, Veillonella, Actinomyces, Granulicatella, Leptotrichia,* and *Thiomonas*) were significantly different between caries-active and caries-free samples in plaque (Ling et al., 2010). Further research also found that three genera including *Streptococcus, Granulicatella,* and *Actinomyces* exhibited a relative higher abundance in severe early children caries subjects, whereas caries free subjects exhibited a relative higher abundance of *Aestuariimicrobium,* indicating that there might be no specific pathogens but rather pathogenic populations structure shafting would lead to the occurrence of dental caries (Jiang et al., 2013). Yang et al. (2012) found that caries microbiomes were significantly more variable. And 147 OTUs were associated with adults dental caries (Yang et al., 2012).

THE MICROBIOTA SHIFTING IN THE DIFFERENT STAGE OF CARIES
Our further research also found that oral microbiota was specific at different stages of caries progression. Gomar-Vercher et al. (2014) collected 110 saliva samples from 12-year-old children and divided into six groups according to the International Caries Detection and Assessment System II criteria. They found that *Porphyromonas* and *Prevotella* showed an increasing percentage compared to healthy individuals and bacterial diversity diminished as the severity of the disease increased (Gomar-Vercher et al., 2014). We also studied microbiome of plaque from caries-active subjects in different caries stages including intact enamel, white spot lesions and carious dentin lesions by pyrosequencing technique. And the result showed that the diversity of the total plaque bacterial community in the health subjects were more complex than caries subjects, which is in accordance with Gomar-Vercher's study. Moreover thirteen genera (including *Capnocytophaga, Fusobacterium, Porphyromonas, Abiotrophia, Comamonas, Tannerella, Eikenella, Paludibacter, Treponema, Actinobaculum, Stenotrophomonas, Aestuariimicrobium,* and *Peptococcus*) were associated with dental health, eight genera (including *Cryptobacterium, Lactobacillus, Megasphaera, Olsenella, Scardovia, Shuttleworthia, Cryptobacterium,* and *Streptococcus*) increased significantly in cavitated dentin lesions, and *Actinomyces* and *Corynebacterium* were present at significant high levels in white spot lesions, while *Flavobacterium, Neisseria, Bergeyella,* and *Derxia* were enriched in the intact surfaces of caries individuals (Jiang et al., 2014). Relatively high proportions of *Atopobium,*

Prevotella, or *Propionibacterium* with *Streptococcus* or *Actinomyces* dominated in carious dentin lesions in Obata's study (Obata et al., 2014).

ORAL MICROBIOTA OF APICAL PERIODONTITIS

Apical periodontitis develops around the apex of the dental root and is caused primarily by root canal infection (Siqueira, 2001). Bacterial biofilm communities established in the apical part of infected root canals are conceivably of the most importance in the pathogenesis of apical periodontitis (Siqueira, 2002). For there was no strong evidence of the specific involvement of a single species with any particular sign or symptom of apical periodontitis been found with advancing technique. By using massive parallel pyrosequencing analysis, 187 bacterial species-level phylotypes, 84 genera and 10 phyla were found in the apical part of root canals of teeth with apical periodontitis. The most abundant and prevalent phyla were Proteobacteria, Firmicutes, Bacteroidetes, Fusobacteria, and Actinobacteria. And the mean number of species-level phylotypes per sample was 37. These results indicate that bacterial communities in apical periodontitis are more diverse than previously demonstrated (Siqueira et al., 2011). Another study investigated the microbial diversity in symptomatic and asymptomatic canals with primary endodontic infections by using GS FLX Titanium pyrosequencing. The result showed that the vast majority of sequences belonged to seven phyla including *Actinobacteria*, *Bacteroidetes*, *Firmicutes*, *Fusobacteria*, *Proteobacteria*, *Spirochetes*, and *Synergistetes*. And *Pyramidobacter*, *Streptococcus*, *Leptotrichia* constituted nearly 50% of microbial community in asymptomatic teeth, whereas *Neisseria*, *Propionibacterium*, and *Tessaracoccus* were frequently found in symptomatic teeth (Lim et al., 2011). Santos et al. (2011) performed barcoded multiplex pyrosequencing to compare the microbiota of dental root canal infections associated with acute or chronic apical periodontitis. They found that the most abundant phyla in acute infections were *Firmicutes*, *Fusobacteria*, and *Bacteroidetes*, while in chronic infections, the dominants were *Firmicutes*, *Bacteroidetes*, and *Actinobacteria*. And the most prevalent genera in acute infections were *Fusobacterium* and *Parvimonas* (Santos et al., 2011). In Hong's report, the diversity of bacterial community profile of intracanal microbiota in primary and persistent endodontic infections associated with asymptomatic chronic apical periodontitis showed no significantly different. And Bacteroidetes was the most abundant phylum in both primary and persistent infections. Other reports also found Bacteroidetes was the most abundant phylum in both primary and persistent infections by using pyrosequencing (Hong et al., 2013). And in symptomatic periapical lesions, the most abundant phyla were Proteobacteria and Firmicutes, while the predominated genera were *Fusobacterium*, *Streptococcus*, *Prevotella*, *Corynebacterium*, *Porphyromonas*, and *Actinomyces* (Saber et al., 2012). Another research analyzed endodontic infections by deep coverage pyrosequencing and found that 179 bacterial genera in 13 phyla. Among them, Bacteroidetes was the most prevalent bacterial phylum (Li et al., 2010).

ORAL MICROBIOTA IN PERIODONTITIS

Periodontitis is an inflammatory disease in which oral bacteria play an important role in the progress of disease. It is thought to be concerned to a polymicrobial etiology, and comprehensive studies were performed to elucidate differences of the complex communities between health and disease (Ashimoto et al., 1996). By comparing the periodontally healthy controls and subjects with chronic periodontitis, Griffen found that community diversity was higher than that in disease. 123 species were identified which were significantly more abundant in individuals with chronic periodontitis and 53 species were identified in health controls. Among them, Spirochaetes, Synergistetes, and Bacteroidetes were health-associated, whereas Proteobacteria, Clostridia, *Negativicutes* and *Erysipelotrichia* were associated with disease (Griffen et al., 2012). There are significantly different in abundance comparing the oral microbiome in deep (diseased) and shallow (healthy) sites by sequencing 16S rRNA genes. In the deep sites, 14 genus-level OTUs, including *Streptococcus*, *Actinomyces*, and *Veillonella*, were decreased, whereas 37 genus-level OTUs were present in increased abundance compared to shallow sites such as *Prevotella*, *Porphyromonas*, *Treponema*, and *Fusobacterium* (Ge et al., 2013). By utilizing pyrosequencing technique, the gram-negative genera *Selenomonas*, *Prevotella*, *Treponema*, *Tannerella*, *Haemophilus*, and *Catonella* are significantly enriched in periodontal disease, whereas a set of gram-positive genera are significantly enriched in healthy samples (*Streptococcus*, *Actinomyces*, and *Granulicatella*; Liu et al., 2012). Bacteroidetes was the most abundant phylum in samples of periodontal disease, whereas Actinobacteria and Proteobacteria were significantly increased in plaque of periodontal health in another metagenomic sequencing analysis. At genus level, microbial community of periodontal health were dominated by *Streptococcus*, *Haemophilus*, *Rothia*, and *Capnocytophaga*, while microbiota in periodontal disease exhibited high level of *Prevotella* (Wang et al., 2013). Another 16S rRNA gene sequencing analysis found that *Fusobacterium*, *Porphyromonas*, *Treponema*, *Filifactor*, *Eubacterium*, *Tannerella*, *Hallella*, *Parvimonas*, *Peptostreptococcus*, and *Catonella* showed higher relative abundances in the periodontitis group (Li et al., 2014).

By using an Ion Torrent Personal Genome Machine, the diversity of bacterial community increased after scaling and root planning therapy (SPR). The most striking difference was that periodontal pathogenic species including the genera *Porphyromonas*, *Tannerella*, *Treponema*, and *Filifactor* were removed only in the group treated with SPR and antibiotics (Jünemann et al., 2012). And the post-treatment plaque samples retained the highest similarity to pre-treatment samples of the same individual (Schwarzberg et al., 2014).

CONCLUSION

The high through-put technique has largely expended our knowledge regarding the composition of the bacterial communities associated with healthy and disease. The oral microbiota is far more diverse than previous thought. And as a number of uncultivated organisms discovery, it has being shed light on the relationship between oral microbiota and the caries process and periodontitis developing and other oral disease. The new high-throughput methodologies are likely to approach our understanding of bacterial ecology in oral disease.

REFERENCES

Aas, J. A., Paster, B. J., Stokes, L. N., Olsen, I., and Dewhirst, F. E. (2005). Defining the normal bacterial flora of the oral cavity. *J. Clin. Microbiol.* 43, 5721–5732. doi: 10.1128/JCM.43.11.5721-5732.2005

Ashimoto, A., Chen, C., Bakker, I., and Slots, J. (1996). Polymerase chain reaction detection of 8 putative periodontal pathogens in subgingival plaque of gingivitis and advanced periodontitis lesions. *Oral Microbiol. Immunol.* 11, 266–273. doi: 10.1111/j.1399-302X.1996.tb00180.x

Bik, E. M., Long, C. D., Armitage, G. C., Loomer, P., Emerson, J., Mongodin, E. F., et al. (2010). Bacterial diversity in the oral cavity of 10 healthy individuals. *ISME J.* 4, 962–974. doi: 10.1038/ismej.2010.30

Bowden, G. (1997). Does assessment of microbial composition of plaque/saliva allow for diagnosis of disease activity of individuals? *Community Dent. Oral Epidemiol.* 25, 76–81. doi: 10.1111/j.1600-0528.1997.tb00902.x

Cephas, K. D., Kim, J., Mathai, R. A., Barry, K. A., Dowd, S. E., Meline, B. S., et al. (2011). Comparative analysis of salivary bacterial microbiome diversity in edentulous infants and their mothers or primary care givers using pyrosequencing. *PLoS ONE* 6:e23503. doi: 10.1371/journal.pone.0023503

Crielaard, W., Zaura, E., Schuller, A. A., Huse, S. M., Montijn, R. C., and Keijser, B. J. (2011). Exploring the oral microbiota of children at various developmental stages of their dentition in the relation to their oral health. *BMC Med. Genomics* 4:22. doi: 10.1186/1755-8794-4-22

Ge, X., Rodriguez, R., Gunsolley, J., and Xu, P. (2013). Oral microbiome of deep and shallow dental pockets in chronic periodontitis. *PLoS ONE* 8:e65520. doi: 10.1371/journal.pone.0065520

Gill, S. R., Pop, M., DeBoy, R. T., Eckburg, P. B., Turnbaugh, P. J., Samuel, B. S., et al. (2006). Metagenomic analysis of the human distal gut microbiome. *Science* 312, 1355–1359. doi: 10.1126/science.1124234

Gomar-Vercher, S., Cabrera-Rubio, R., Mira, A., and Almerich-Silla, J. (2014). Relationship of children's salivary microbiota with their caries status: a pyrosequencing study. *Clin. Oral Investig.* doi: 10.1007/s00784-014-1200-y [Epub ahead of print].

Griffen, A. L., Beall, C. J., Campbell, J. H., Firestone, N. D., Kumar, P. S., Yang, Z. K., et al. (2012). Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. *ISME J.* 6, 1176–1185. doi: 10.1038/ismej.2011.191

Griffen, A. L., Beall, C. J., Firestone, N. D., Gross, E. L., DiFranco, J. M., Hardman, J. H., et al. (2011). CORE: a phylogenetically-curated 16S rDNA database of the core oral microbiome. *PLoS ONE* 6:e19051. doi: 10.1371/journal.pone.0019051

Hamada, S., and Slade, H. D. (1980). Biology, immunology, and cariogenicity of *Streptococcus* mutans. *Microbiol. Rev.* 44, 331–384.

Hasan, N. A., Young, B. A., Minard-Smith, A. T., Saeed, K., Li, H., Heizer, E. M., et al. (2014). Microbial community profiling of human saliva using shotgun metagenomic sequencing. *PLoS ONE* 9:e97699. doi: 10.1371/journal.pone.0097699

Hong, B. Y., Lee, T. K., Lim, S. M., Chang, S. W., Park, J., Han, S. H., et al. (2013). Microbial analysis in primary and persistent endodontic infections by using pyrosequencing. *J. Endod.* 39, 1136–1140. doi: 10.1016/j.joen.2013.05.001

Huse, S. M., Ye, Y., Zhou, Y., and Fodor, A. A. (2012). A core human microbiome as viewed through 16S rRNA sequence clusters. *PLoS ONE* 7:e34242. doi: 10.1371/journal.pone.0034242

Jiang, W., Ling, Z., Lin, X., Chen, Y., Zhang, J., Yu, J., et al. (2014). Pyrosequencing analysis of oral microbiota shifting in various caries states in childhood. *Microb. Ecol.* 67, 962–969. doi: 10.1007/s00248-014-0372-y

Jiang, W., Zhang, J., and Chen, H. (2013). Pyrosequencing analysis of oral microbiota in children with severe early childhood dental caries. *Curr. Microbiol.* 67, 537–542. doi: 10.1007/s00284-013-0393-7

Jünemann, S., Prior, K., Szczepanowski, R., Harks, I., Ehmke, B., Goesmann, A., et al. (2012). Bacterial community shift in treated periodontitis patients revealed by ion torrent 16S rRNA gene amplicon sequencing. *PLoS ONE* 7:e41606. doi: 10.1371/journal.pone.0041606

Kashket, S., Zhang, J., and Van Houte, J. (1996). Accumulation of fermentable sugars and metabolic acids in food particles that become entrapped on the dentition. *J. Dent. Res.* 75, 1885–1891. doi: 10.1177/00220345960750111101

Keijser, B., Zaura, E., Huse, S., Van Der Vossen, J., Schuren, F., Montijn, R., et al. (2008). Pyrosequencing analysis of the oral microflora of healthy adults. *J. Dent. Res.* 87, 1016–1020. doi: 10.1177/154405910808701104

Lazarevic, V., Whiteson, K., Hernandez, D., François, P., and Schrenzel, J. (2010). Study of inter-and intra-individual variations in the salivary microbiota. *BMC Genomics* 11:523. doi: 10.1186/1471-2164-11-523

Lazarevic, V., Whiteson, K., Huse, S., Hernandez, D., Farinelli, L., Østerås, M., et al. (2009). Metagenomic study of the oral microbiota by Illumina high-throughput sequencing. *J. Microbiol. Methods* 79, 266–271. doi: 10.1016/j.mimet.2009.09.012

Lederberg, J., and Mccray, A. (2001). 'Ome sweet 'omics—a genealogical treasury of words. *Scientist* 15, 8.

Lemon, K. P., Klepac-Ceraj, V., Schiffer, H. K., Brodie, E. L., Lynch, S. V., and Kolter, R. (2010). Comparative analyses of the bacterial microbiota of the human nostril and oropharynx. *MBio* 1:e00129-00110. doi: 10.1128/mBio.00129-10

Li, L., Hsiao, W., Nandakumar, R., Barbuto, S., Mongodin, E., Paster, B., et al. (2010). Analyzing endodontic infections by deep coverage pyrosequencing. *J. Dent. Res.* 89, 980–984. doi: 10.1177/0022034510370026

Li, Y., He J., Zhou, Y., Yuan, M., Xu, X., Sun, F., et al. (2014). Phylogenetic and functional gene structure shifts of the oral microbiomes in periodontitis patients. *ISME J.* 8, 1879–1891. doi: 10.1038/ismej.2014.28

Lim, S. M., Lee, T. K., Kim, E. J., Park, J. H., Lee, Y., Bae, K. S., et al. (2011). Microbial profile of asymptomatic and symptomatic teeth with primary endodontic infections by pyrosequencing. *J. Korean Acad. Conserv. Dent.* 36, 498–505. doi: 10.5395/JKACD.2011.36.6.498

Ling, Z., Kong, J., Jia, P., Wei, C., Wang, Y., Pan, Z., et al. (2010). Analysis of oral microbiota in children with dental caries by PCR-DGGE and barcoded pyrosequencing. *Microb. Ecol.* 60, 677–690. doi: 10.1007/s00248-010-9712-8

Ling, Z., Liu, X., Wang, Y., Li, L., and Xiang, C. (2013). Pyrosequencing analysis of the salivary microbiota of healthy Chinese children and adults. *Microb. Ecol.* 65, 487–495. doi: 10.1007/s00248-012-0123-x

Liu, B., Faller, L. L., Klitgord, N., Mazumdar, V., Ghodsi, M., Sommer, D. D., et al. (2012). Deep sequencing of the oral microbiome reveals signatures of periodontal disease. *PLoS ONE* 7:e37919. doi: 10.1371/journal.pone.0037919

Loesche, W. J. (1986). Role of *Streptococcus* mutans in human dental decay. *Microbiol. Rev.* 50, 353–380.

Matee, M., Mikx, F., Maselle, S., and Van Palenstein Helderman, W. (1992). Mutans streptococci and lactobacilli in breast-fed children with rampant caries. *Caries Res.* 26, 183–187. doi: 10.1159/000261440

Metzker, M. L. (2010). Sequencing technologies—the next generation. *Nat. Rev. Genet.* 11, 31–46. doi: 10.1038/nrg2626

Obata, J., Takeshita, T., Shibata, Y., Yamanaka, W., Unemori, M., Akamine, A., et al. (2014). Identification of the microbiota in carious dentin lesions using 16S rRNA gene sequencing. *PLoS ONE* 9:e103712. doi: 10.1371/journal.pone.0103712

Paster, B. J., Olsen, I., Aas, J. A., and Dewhirst, F. E. (2006). The breadth of bacterial diversity in the human periodontal pocket and other oral sites. *Periodontol. 2000* 42, 80–87. doi: 10.1111/j.1600-0757.2006.00174.x

Petersen, P. E., Bourgeois, D., Ogawa, H., Estupinan-Day, S., and Ndiaye, C. (2005). The global burden of oral diseases and risks to oral health. *Bull. World Health Organ.* 83, 661–669.

Preza, D., Olsen, I., Willumsen, T., Grinde, B., and Paster, B. J. (2009). Diversity and site-specificity of the oral microflora in the elderly. *Eur. J. Clin. Microbiol. Infect. Dis.* 28, 1033–1040. doi: 10.1007/s10096-009-0743-3

Rogers, Y. H., and Venter, J. C. (2005). Genomics: massively parallel sequencing. *Nature* 437, 326–327. doi: 10.1038/437326a

Ronaghi, M. (2001). Pyrosequencing sheds light on DNA sequencing. *Genome Res.* 11, 3–11. doi: 10.1101/gr.11.1.3

Rothberg, J. M., and Leamon, J. H. (2008). The development and impact of 454 sequencing. *Nat. Biotechnol.* 26, 1117–1124. doi: 10.1038/nbt1485

Saber, M. H., Schwarzberg, K., Alonaizan, F. A., Kelley, S. T., Sedghizadeh, P. P., Furlan, M., et al. (2012). Bacterial flora of dental periradicular lesions analyzed by the 454-pyrosequencing technology. *J. Endod.* 38, 1484–1488. doi: 10.1016/j.joen.2012.06.037

Santos, A. L., Siqueira, J. F. Jr., Rôças, I. N., Jesus, E. C., Rosado, A. S., and Tiedje, J. M. (2011). Comparing the bacterial diversity of acute and chronic dental root canal infections. *PLoS ONE* 6:e28088. doi: 10.1371/journal.pone.0028088

Schwarzberg, K., Le, R., Bharti, B., Lindsay, S., Casaburi, G., Salvatore, F., et al. (2014). The personal human oral microbiome obscures the effects of treatment on periodontal disease. *PLoS ONE* 9:e86708. doi: 10.1371/journal.pone.0086708

Shade, A., and Handelsman, J. (2012). Beyond the Venn diagram: the hunt for a core microbiome. *Environ. Microbiol.* 14, 4–12. doi: 10.1111/j.1462-2920.2011.02585.x

Siqueira, J. (2001). Aetiology of root canal treatment failure: why well-treated teeth can fail. *Int. Endod. J.* 34, 1–10. doi: 10.1046/j.1365-2591.2001.00396.x

Siqueira, J. F. (2002). Endodontic infections: concepts, paradigms, and perspectives. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endodontol.* 94, 281–293. doi: 10.1067/moe.2002.126163

Siqueira, J. F. Jr., Alves, F. R., and Rôças, I. N. (2011). Pyrosequencing analysis of the apical root canal microbiota. *J. Endod.* 37, 1499–1503. doi: 10.1016/j.joen.2011.08.012

Wang, J., Qi, J., Zhao, H., He, S., Zhang, Y., Wei, S., et al. (2013). Metagenomic sequencing reveals microbiota and its functional potential associated with periodontal disease. *Sci. Rep.* 3, 1843. doi: 10.1038/srep01843

Xu, H., Hao, W., Zhou, Q., Wang, W., Xia, Z., Liu, C., et al. (2014). Plaque bacterial microbiome diversity in children younger than 30 months with or without caries prior to eruption of second primary molars. *PLoS ONE* 9:e89269. doi: 10.1371/journal.pone.0089269

Yang, F., Zeng, X., Ning, K., Liu, K. L., Lo, C. C., Wang, W., et al. (2012). Saliva microbiomes distinguish caries-active from healthy human populations. *ISME J.* 6, 1–10. doi: 10.1038/ismej.2011.71

Zarco, M., Vess, T., and Ginsburg, G. (2012). The oral microbiome in health and disease and the potential impact on personalized dental medicine. *Oral Dis.* 18, 109–120. doi: 10.1111/j.1601-0825.2011.01851.x

Zaura, E., Keijser, B., Huse, S., and Crielaard, W. (2009). Defining the healthy. *BMC Microbiol.* 9:259. doi: 10.1186/1471-2180-9-259

Zhou, X., Ren, L., Li, Y., Zhang, M., Yu, Y., and Yu, J. (2010). The next-generation sequencing technology: a technology review and future perspective. *Sci. China Life Sci.* 53, 44–57. doi: 10.1007/s11427-010-0023-6

Osmoregulation and the human mycobiome

Abhishek Saxena and Ramakrishnan Sitaraman*

Department of Biotechnology, TERI University, New Delhi, India
*Correspondence: minraj@gmail.com

INTRODUCTION

The last one-and-a half decades have made it amply clear that the human microbiota have a very significant role to play in health and disease. The human body can (or should) be better viewed as a complex ecosystem inhabited by micro-organisms that outnumber human cells 10 to 1 (Ley et al., 2006). However, most research in this field has been focused on the prokaryotic (specifically, bacterial) component of the microbiota. Sampling, in turn, is carried out mostly from sites that are readily accessible (Human Microbiome Project Consortium, 2012). Such sampling might not be genuinely representative of the situation *in vivo*, even for the bacteria under study. The human alimentary tract is a complex entity that exhibits extensive variation in biological characteristics such as tissue/cell types and secretions, as well as parameters such as temperature, pH, oxygen levels and osmolarity along its entire length. A proteomic study sampling mucosal lavages at multiple colonic sites indicated significant differences in protein profiles between the proximal and distal colon, which was supportive of the concept of their functional and developmental distinctness (Li et al., 2011). The colonization of the human infant by microbes, initially during the process of birth, exhibits an ecological succession of microbial species over time, and plays a prominent role in the maturation of the immune system as well (reviewed in Costello et al., 2012).

The fungal members of the microbiota are not very numerous compared to bacteria. Large-scale metagenomic sequencing of fecal samples of 124 individuals found that only about 0.1% of genes detected were of eukaryotic origin (Qin et al., 2010). The most commonly encountered genera constituting the fungal microbiota

or "mycobiome" (Huffnagle and Noverr, 2013) are *Candida*, *Saccharomyces* and *Cladosporium* (Hoffmann et al., 2013). The bacterial microbiota and a functional immune system are thought to keep the numbers of opportunistic fungal pathogens, such as *Candida* spp., under check in the absence of any perturbations. However, information from studies of polymicrobial diseases points to subtler adjustments, dependent on environmental conditions and cross-kingdom signals that eventually influence (positively and negatively) modes and rates of growth (reviewed in Peleg et al., 2010). The sensing of, and responses to, biotic and abiotic stimuli by fungi (as in other organisms) involves multiple signaling pathways that can interact to either augment or attenuate one another, as will be discussed below.

OSMOREGULATION AND STRESS RESPONSES IN *SACCHAROMYCES CEREVISIAE*

S. cerevisiae, a well-studied fungus, employs two strategies for responding to stress, both involving extensive signaling by MAP kinases (MAPK, alternatively termed SAPK-stress activated protein kinase). The first is stress-specific, e.g., involved in the response to pheromone, spore wall formation and adapting to hyperosmotic conditions etc. (Gustin et al., 1998). The HOG (high osmolarity glycerol) pathway that is activated in response to hyperosmotic conditions leads to the accumulation of compatible solutes (glycerol being the most important) and also results in the closure of the aquaglyceroporin Fps1p, enabling retention of glycerol. The HOG pathway functions through two signaling branches. The SLN1 branch involves the two-component membrane sensor protein Sln1p complexed

with Ypd1p and Ssk1p that, under hyperosmotic conditions, is unable to inactivate downstream MAPKKKs (functionally redundant Ssk2p and Ssk22p) by phosphorylation. This results in the dephosphorylation of these kinases that phosphorylate the MAPKK Pbs2p which, in turn, phsophorylates the MAPK Hog1p. Phosphorylated Hog1p moves into the nucleus and interacts with transcription factors such as Hot1p, Msn1p, Msn2p etc. activating the transcription of, among other genes, including those encoding phosphatases (e.g., Ptp2p, Ptp3p) that dephosphophorylate Hog1p, causing feedback inhibition, limiting the duration of Hog1p activity. The SHO1 branch involves two functionally redundant mucin-like transmembrane osmosensors, Msb2p and Hkr1p, that recruit the Pbs2p MAPKK directly to the cytoplasmic face of the cell membrane as part of a macromolecular complex. Notably, SHO1 branch proteins are shared with other signaling pathways, and it is activated when hyperosmolarity occurs as a result of other cellular responses (reviewed in Hohmann, 2002; Hohmann et al., 2007).

The second strategy for coping with stress is the environmental stress response (ESR) that enables adaptation to the long-term effects of various stresses, in contrast to the more specific and short-term response of other MAPK pathways. The ESR was first described as an increased expression of ~300 genes and repression of ~600 genes in response to diverse environmental conditions to which *S. cerevisiae* was subjected (Gasch et al., 2000; Causton et al., 2001). In both these studies, various stress conditions were tested including temperature, hyper- and hypo-osmotic shock, oxidative (H_2O_2) stress etc. Induced genes included those involved

in a wide variety of processes, including carbohydrate metabolism, detoxification of reactive oxygen species, cellular redox reactions, cell wall modification, protein folding and degradation, DNA damage repair, fatty acid metabolism, metabolite transport, vacuolar and mitochondrial functions, autophagy, and intracellular signaling (Gasch et al., 2000). Genes encoding cytoplasmic ribosomal proteins, tRNA synthases, proteins required for processing rRNAs, and a subset of translation initiation factors were repressed (Causton et al., 2001).

The ESR provides a "cross-protective effect" wherein *S. cerevisiae* subjected to mild heat stress as the primary stress becomes adapted to higher levels of heat as well as oxidative (H_2O_2) stresses (Berry et al., 2011). This is especially relevant as stresses under natural conditions don't occur singly or sequentially, but simultaneously. It may seem at first sight that the mild dosage of primary/initial stress is irrelevant, as the ultimate adaptation to the secondary stress(es) is achieved anyway. However, Berry et al. (2011) demonstrated that distinct subsets of genes were activated due to primary and secondary stresses. Earlier work indicated that the cross-protective effect is not universal, but specific to the primary/secondary stress combination (Berry and Gasch, 2008). The major transcription factors mediating the ESR are Msn2p and Msn4p (Berry et al.,

2011). Msn2p and Msn4p (see below) play specific roles depending on the stress combination and are even regulated in a condition-specific manner. Besides, there are other transcription factors activated during the ESR and subsequent "acquired stress resistance," like Hsf1p (heat stress), Yap1p (oxidative stress), and Hot1p & Sko1p (hyperosmotic stress), that can also activate Msn2p/4p target genes (Berry and Gasch, 2008).

Adding to the mechanistic complexity of stress responses is recent evidence that Hog1p induces transcription of a long non-coding RNA whose presence is required for chromosome remodeling around the *CDC28* gene encoding a cyclin-dependent kinase and its subsequent induction. This is accompanied by cell cycle delay, and increased Cdc28p levels ensure faster recovery following the stress application (Nadal-Ribelles et al., 2014). Osmoadapted *S. cerevisiae* exhibit HOG activation upon shmooing in response to pheromone (Baltanás et al., 2013). Hog1p also imposes checkpoints on the mating pathway if pheromone is sensed during a period of hyperosmotic stress. It phosphorylates the protein kinase Rck2p that inhibits translation of Fus3p (the MAPK of mating pathway) by phosphorylating EF2 (elongation factor 2). Ste50p, a shared component of both the HOG and mating pathways is phosphorylated (and inhibited) by Hog1p (Nagiec and Dohlman,

2012). Thus, the osmoregulation response is not elicited solely by hyperosmolarity *per se*, but is also influenced by the spatio-temporal modulation of cell-cycle events involving stimuli impinging on other signaling pathways.

THE STRESS RESPONSES AND ADAPTIVE POTENTIAL OF *CANDIDA* SPP.

Candida spp. are dimorphic yeasts that occur practically throughout the alimentary tract, and *C. albicans* and *C. glabrata* well-known opportunistic pathogens. *C. albicans* can grow as a planktonic unicellular organism (yeast) or as a filamentous organism (hypha). The yeast form is suitable for dissemination, while the hyphal form is more adapted to colonization and biofilm formation. *Candida albicans* also exhibits parasexuality, in which mating-competent (so-called "opaque") diploid strains mate to form tetraploids, whose progeny later undergo chromosome losses to regenerate the diploid state (Hull et al., 2000; Magee and Magee, 2000). More recent work has demonstrated the existence of a viable and stable haploid form generated by chromosome losses implying that *C. albicans* may not be an "obligate diploid" as originally thought (Hickman et al., 2013). This increases the overall repertoire of genetic diversity in the *Candida* population, that could confer an adaptive advantage on the organism. Incidentally, these

Table 1 | Proteins involved in osmoadaptation/stress responses as well as other phenotypes in pathogenic fungi.

Gene/Protein	Function summary	References
Hog1p	Enhances virulence of Ca	Alonso-Monge et al., 1999; Zhang et al., 2009
Msn2p	ESR-related transcription factor required for full virulence in entomopathogenic fungi	Liu et al., 2013
Ste11p	Shared MAPKKK in mating response and hyperosmotic stress response, implicated in Cg virulence	Calcagno et al., 2005
Hog1p	In Cn, it antagonizes the effect of Mpk1p (a MAPK) and calcineurin, which promote cell wall integrity	Kojima et al., 2006;
	Overactivation of the HOG pathway by fludioxonil (antifungal) results in fludioxinol hypersensitivity	
	Inhibition of Hog1p in Cn is associated with increased ergosterol in the cell membrane, promoting sensitivity to ergosterol-binding antifungals	Young-Sun Bahn et al., 2012
Cta1p	Catalase gene of Cg regulated by multiple stress-related transcription factors like Msn2p, Msn4p, Skn7p, Yap1p	Cuéllar-Cruz et al., 2008
Ste12p	Transcription factor induced by pheromone and nitrogen starvation in Sc, promotes virulence in Cg	Calcagno et al., 2003
Slt2p	The cell wall integrity pathway MAPK in Cg; downstream target Rlm1p (a transcription factor) is involved in micafungin tolerance	Miyazaki et al., 2010
Chk1p	Two component signal transduction protein in Ca; ortholog of oxidative stress sensors (Mak2p and Mak3p) in *S. pombe*. Involved in quorum sensing (farnesol response)	Kruppa et al., 2004
Mln1p	Msn2p-like protein enabling long-term resistance to weak acid in Ca	Ramsdale et al., 2008

Sc, Saccharomyces cerevisiae; Ca, Candida albicans; Cg, Candida glabrata; Cn, Cryptococcus neoformans.

polymorphic transitions caution us that reliance on metagenomic and quantitative approaches to study the gut microbiota may not be adequately reflective of significant, populations and sub-populations that arise transiently by random and/or adaptive mechanisms.

Orthologs of the various MAPK pathway genes and also of the factors involved in the ESR have been discovered in *Candida* spp. *C. albicans* is thought to have diverged from *S. cerevisiae* more than 200 million years ago (Kurtzman and Piškur, 2006) The *C. albicans* ESR (CaESR) is not as extensive in genetic terms as in *S. cerevisiae*. Only a small number of genes are involved in CaESR (~24 upregulated genes and ~37 downregulated genes) (Enjalbert et al., 2006; Gasch, 2007). This suggests the absence of a core environmental response/general stress response in *C. albicans* (Enjalbert et al., 2003). Later studies confirmed that CaMsn4p only weakly complements the inability of an *msn2△msn4△* double mutant in *S. cerevisiae* to activate a STRE-*lacZ* reporter (STRE-Stress response element) while CaMln1p (*Candida albicans* Msn2p/Msn4p-like protein) does not complement the defect at all (Nicholls et al., 2004). The transcription factors finally activated in the CaESR have not been conclusively identified. Therefore a complete picture of activation and regulation of ESR in *C. albicans* is as yet unavailable. However, the number of genes activated/repressed during the CaESR were more in response to oxidative stress (5mM H_2O_2) than in response to osmotic stress (0.3 M NaCl) and heavy metal stress (0.5 mM $CdSO_4$) (Enjalbert et al., 2006). Other studies have reported that components of stress signaling pathways may be important in virulence, drug tolerance, or quorum sensing, among other phenotypes. Interestingly, Msn2p homologs in entomopathogenic fungi *Beauveria bassiana* and *Metarhizium robertsii* augment virulence (Zhang et al., 2009; Liu et al., 2013). **Table 1** lists some components of stress signaling known to influence other phenotypes in pathogenic fungi.

CONCLUSIONS

In contrast to bacteria, a beneficial role for fungi in microbiota-human interactions has not emerged, though their role as opportunistic pathogens (cf. ecologically invasive species) has been extensively studied. The major fungal probiotic in use today, *Saccharomyces cerevisiae* var. *boulardii* (Sb), that is not indigenous to the human gut, provides some examples of potential benefits of fungal proteins for the host (Czerucka et al., 1994; Dahan et al., 2003), and can help in the maintenance and/or restoration of host-microbiota homeostasis.

The strategies employed by the members of the mycobiome to adapt to changing conditions along the length of the gut and host immune responses may depend significantly on adapting to continuously changing environmental parameters that could also serve as indicators for spatial location, microbiota composition and the physiological state of the host. Signaling pathways that respond to different stimuli are not watertight modules, but can interact in unforeseen ways to produce an integrated behavioral response (Baltanás et al., 2013). Thus, components of the osmoregulatory pathway may also participate in the process of mounting a coordinated response to environmental stiumuli.

REFERENCES

Alonso-Monge, R., Navarro-García, F., Molero, G., Diez-Orejas, R., Gustin, M., Pla, J., et al. (1999). Role of the mitogen-activated protein kinase Hog1p in morphogenesis and virulence of *Candida albicans*. *J. Bacteriol.* 181, 3058–3068.

Baltanás, R., Bush, A., Couto, A., Durrieu, L., Hohmann, S., and Colman-Lerner, A. (2013). Pheromone-induced morphogenesis improves osmoadaptation capacity by activating the HOG MAPK pathway. *Sci. Signal.* 6:ra26. doi: 10.1126/scisignal.2003312

Berry, D. B., and Gasch, A. P. (2008). Stress-activated genomic expression changes serve a preparative role for impending stress in yeast. *Mol. Biol. Cell* 19, 4580–4587. doi: 10.1091/mbc.E07-07-0680

Berry, D. B., Guan, Q., Hose, J., Haroon, S., Gebbia, M., Heisler, L. E., et al. (2011). Multiple means to the same end: the genetic basis of acquired stress resistance in yeast. *PLoS Genet.* 7:e1002353. doi: 10.1371/journal.pgen.1002353

Calcagno, A.-M., Bignell, E., Rogers, T. R., Jones, M. D., Mühlschlegel, F. A., and Haynes, K. (2005). *Candida glabrata* Ste11 is involved in adaptation to hypertonic stress, maintenance of wild-type levels of filamentation and plays a role in virulence. *Med. Mycol.* 43, 355–364. doi: 10.1080/13693780400006088

Calcagno, A.-M., Bignell, E., Warn, P., Jones, M. D., Denning, D. W., Mühlschlegel, F. A., et al. (2003). *Candida glabrata* STE12 is required for wild-type levels of virulence and nitrogen starvation induced filamentation. *Mol. Microbiol.* 50, 1309–1318. doi: 10.1046/j.1365-2958.2003.03755.x

Causton, H. C., Ren, B., Koh, S. S., Harbison, C. T., Kanin, E., Jennings, E. G., et al. (2001). Remodeling of yeast genome expression in response to environmental changes. *Mol. Biol. Cell* 12, 323–337. doi: 10.1091/mbc.12.2.323

Costello, E. K., Stagaman, K., Dethlefsen, L., Bohannan, B. J. M., and Relman, D. A. (2012). The application of ecological theory toward an understanding of the human microbiome. *Science* 336, 1255–1262. doi: 10.1126/science.1224203

Cuéllar-Cruz, M., Briones-Martin-del-Campo, M., Cañas-Villamar, I., Montalvo-Arredondo, J., Riego-Ruiz, L., Castaño, I., et al. (2008). High resistance to oxidative stress in the fungal pathogen *Candida glabrata* is mediated by a single catalase, Cta1p, and is controlled by the transcription factors Yap1p, Skn7p, Msn2p, and Msn4p. *Eukaryot. Cell* 7, 814–825. doi: 10.1128/EC.00011-08

Czerucka, D., Roux, I., and Rampal, P. (1994). *Saccharomyces boulardii* inhibits secretagogue-mediated adenosine 3',5'-cyclic monophosphate induction in intestinal cells. *Gastroenterology* 106, 65–72.

Dahan, S., Dalmasso, G., Imbert, V., Peyron, J.-F., Rampal, P., and Czerucka, D. (2003). *Saccharomyces boulardii* interferes with enterohemorrhagic escherichia coli-induced signaling pathways in T84 cells. *Infect. Immun.* 71, 766–773. doi: 10.1128/IAI.71.2.766-773.2003

Enjalbert, B., Nantel, A., and Whiteway, M. (2003). Stress-induced gene expression in *Candida albicans*: absence of a general stress response. *Mol. Biol. Cell* 14, 1460–1467. doi: 10.1091/mbc.E02-08-0546

Enjalbert, B., Smith, D. A., Cornell, M. J., Alam, I., Nicholls, S., Brown, A. J. P., et al. (2006). Role of the Hog1 stress-activated protein kinase in the global transcriptional response to stress in the fungal pathogen *Candida albicans*. *Mol. Biol. Cell* 17, 1018–1032. doi: 10.1091/mbc.E05-06-0501

Gasch, A. P. (2007). Comparative genomics of the environmental stress response in ascomycete fungi. *Yeast* 24, 961–976. doi: 10.1002/yea.1512

Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., et al. (2000). Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* 11, 4241–4257. doi: 10.1091/mbc.11.12.4241

Gustin, M. C., Albertyn, J., Alexander, M., and Davenport, K. (1998). MAP kinase pathways in the yeast *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* 62, 1264–1300.

Hickman, M. A., Zeng, G., Forche, A., Hirakawa, M. P., Abbey, D., Harrison, B. D., et al. (2013). The "obligate diploid" *Candida albicans* forms mating-competent haploids. *Nature* 494, 55–59. doi: 10.1038/nature11865

Hoffmann, C., Dollive, S., Grunberg, S., Chen, J., Li, H., Wu, G. D., et al. (2013). Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. *PLoS ONE* 8:e66019. doi: 10.1371/journal.pone.0066019

Hohmann, S. (2002). Osmotic stress signaling and osmoadaptation in yeasts. *Microbiol. Mol. Biol. Rev.* 66, 300–372. doi: 10.1128/MMBR.66.2.300-372.2002

Hohmann, S., Krantz, M., and Nordlander, B. (2007). Yeast osmoregulation. *Meth. Enzymol.* 428, 29–45. doi: 10.1016/S0076-6879(07)28002-4

Huffnagle, G. B., and Noverr, M. C. (2013). The emerging world of the fungal microbiome. *Trends Microbiol.* 21, 334–341. doi: 10.1016/j.tim.2013.04.002

Hull, C. M., Raisner, R. M., and Johnson, A. D. (2000). Evidence for mating of the "asexual" yeast *Candida albicans* in a mammalian host. *Science* 289, 307–310.

Human Microbiome Project Consortium (2012). A framework for human microbiome research. *Nature* 486, 215–221. doi: 10.1038/nature11209

Kojima, K., Bahn, Y.-S., and Heitman, J. (2006). Calcineurin, Mpk1 and Hog1 MAPK pathways independently control fludioxonil antifungal sensitivity in *Cryptococcus neoformans*. *Microbiology* 152, 591–604. doi: 10.1099/mic.0.28571-0

Kruppa, M., Krom, B. P., Chauhan, N., Bambach, A. V., Cihlar, R. L., and Calderone, R. A. (2004). The two-component signal transduction protein Chk1p regulates quorum sensing in *Candida albicans*. *Eukaryot. Cell* 3, 1062–1065. doi: 10.1128/EC.3.4.1062-1065.2004

Kurtzman, C., and Piškur, J. (2006). "Taxonomy and phylogenetic diversity among the yeasts," in *Comparative Genomics Topics in Current Genetics*, eds. P. Sunnerhagen and J. Piskur (Berlin, Heidelberg: Springer), 29–46.

Ley, R. E., Peterson, D. A., and Gordon, J. I. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* 124, 837–848. doi: 10.1016/j.cell.2006.02.017

Li, X., LeBlanc, J., Truong, A., Vuthoori, R., Chen, S. S., Lustgarten, J. L., et al. (2011). A metaproteomic approach to study human-microbial ecosystems at the mucosal luminal interface. *PLoS ONE* 6:e26542. doi: 10.1371/journal.pone.0026542

Liu, Q., Ying, S.-H., Li, J.-G., Tian, C.-G., and Feng, M.-G. (2013). Insight into the transcriptional regulation of Msn2 required for conidiation, multi-stress responses and virulence of two entomopathogenic fungi. *Fungal Genet. Biol.* 54, 42–51. doi: 10.1016/j.fgb.2013.02.008

Magee, B. B., and Magee, P. T. (2000). Induction of mating in *Candida albicans* by construction of MTLa and MTLalpha strains. *Science* 289, 310–313. doi: 10.1126/science.289.5477.310

Miyazaki, T., Inamine, T., Yamauchi, S., Nagayoshi, Y., Saijo, T., Izumikawa, K., et al. (2010). Role of the Slt2 mitogen-activated protein kinase pathway in cell wall integrity and virulence in *Candida glabrata*. *FEMS Yeast Res.* 10, 343–352. doi: 10.1111/j.1567-1364.2010.00611.x

Nadal-Ribelles, M., Solé, C., Xu, Z., Steinmetz, L. M., de Nadal, E., and Posas, F. (2014). Control of Cdc28 CDK1 by a Stress-Induced lncRNA. *Mol. Cell* 53, 549–561. doi: 10.1016/j.molcel.2014.01.006

Nagiec, M. J., and Dohlman, H. G. (2012). Checkpoints in a yeast differentiation pathway coordinate signaling during hyperosmotic stress. *PLoS Genet.* 8:e1002437. doi: 10.1371/journal.pgen.1002437

Nicholls, S., Straffon, M., Enjalbert, B., Nantel, A., Macaskill, S., Whiteway, M., et al. (2004). Msn2- and Msn4-like transcription factors play no obvious roles in the stress responses of the fungal pathogen *Candida albicans*. *Eukaryot. Cell* 3, 1111–1123. doi: 10.1128/EC.3.5.1111-1123.2004

Peleg, A. Y., Hogan, D. A., and Mylonakis, E. (2010). Medically important bacterial-fungal interactions. *Nat. Rev. Microbiol.* 8, 340–349. doi: 10.1038/nrmicro2313

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65. doi: 10.1038/nature08821

Ramsdale, M., Selway, L., Stead, D., Walker, J., Yin, Z., Nicholls, S. M., et al. (2008). MNL1 regulates weak acid-induced stress responses of the fungal pathogen *Candida albicans*. *Mol. Biol. Cell* 19, 4393–4403. doi: 10.1091/mbc.E07-09-0946

Young-Sun Bahn, Ko, Y.-J., Maeng, S.-A., Jung, K. W., and Kim, G. B. (2012). *Use of the Genes in the HOG, Ras and cAMP Pathway for Treatment of Fungal Infection. US Patent Application Number US* 2012/0093817.

Zhang, Y., Zhao, J., Fang, W., Zhang, J., Luo, Z., Zhang, M., et al. (2009). Mitogen-activated protein kinase hog1 in the entomopathogenic fungus *Beauveria bassiana* regulates environmental stress responses and virulence to insects. *Appl. Environ. Microbiol.* 75, 3787–3795. doi: 10.1128/AEM.01913-08

Transposon-mediated directed mutation controlled by DNA binding proteins in *Escherichia coli*

Milton H. Saier Jr * and Zhongge Zhang

Division of Biological Sciences, Department of Molecular Biology, University of California at San Diego, La Jolla, CA, USA
*Correspondence: msaier@ucsd.edu

INTRODUCTION

It is a basic principle of genetics that the likelihood of a particular mutation occurring is independent of its phenotypic consequences. The concept of directed mutation, defined as genetic change that is specifically induced by the stress conditions that the mutation relieves (Cairns et al., 1988), challenges this principle (Foster, 1999; Rosenberg, 2001; Wright, 2004). The topic of directed mutation is controversial, and its existence, even its *potential* existence, as defined above, has been altogether questioned (Roth et al., 2006).

Part of the justifiable skepticism concerning directed mutation resulted from experiments that were purported to demonstrate this phenomenon, but were subsequently shown to be explainable by classical genetics (Roth et al., 2006). Mutation rates vary with environmental conditions (e.g., growth state) and genetic background (e. g., mutator genes), a phenomenon known as "adaptive" mutation (Wright, 2004; Foster, 2005), but this does not render the mutation "directed." To establish the principle of directed mutation, it is necessary to show that the adaptive mutation is "directed" to a specific site, characterize the mechanism responsible, identify the proteins involved, and provide the evolutionary basis for its appearance.

One frequently encountered type of mutation results from the hopping of transposable genetic elements, transposons, which can activate or inactivate critical genes (Mahillon and Chandler, 1998; Chandler and Mahillon, 2002). For example, activation of the normally cryptic β-glucoside (*bgl*) catabolic operon in *E. coli* can be accomplished by insertion of either of two insertion sequences, IS1 or IS5, upstream of the *bgl* promoter (Schnetz and Rak, 1992).

The *E. coli* glycerol (*glp*) regulon consists of five operons, two of which (*glpFK* and *glpD*) are required for aerobic growth on glycerol (Lin, 1976). Both operons are subject to negative control by the DNA-binding *glp* regulon repressor, GlpR (Zeng et al., 1996), which also binds glycerol-3-phosphate, the inducer of the *glp* regulon. The *glpFK* operon is additionally subject to positive regulation by the cyclic AMP receptor protein (CRP) complexed with cyclic adenosine monophosphate (cAMP; Freedberg and Lin, 1973; Campos et al., 2013), although *glpD* is not appreciably subject to regulation by this mediator of catabolite repression (Weissenborn et al., 1992). The *glpFK* regulatory region contains four GlpR binding sites, O1–O4, and two CRP binding sites, CrpI and CrpII, which overlap O2 and O3, respectively, (**Figure 1A**). The strong CRP dependency of *glpFK* transcription is reflected by the fact that *crp* and *cya* (adenylate cyclase) mutant cells are unable to utilize glycerol (Zhang and Saier, 2009a). We have found that binding of GlpR and the cAMP-CRP complex to the *glpFK* upstream control region negatively influences IS5 hopping specifically into the single site that strongly activates *glpFK* promoter activity (Zhang and Saier, 2009a, unpublished observations).

Glp+ MUTATIONS IN A *crp* GENETIC BACKGROUND

When *crp* cells were incubated on solid glycerol minimal medium, Glp+ colonies appeared. We tested the growth of a *crp* Glp+ strain on glycerol in defined liquid medium. The growth rate was greater than that of wild type (*wt*) *E. coli* (Zhang and Saier, 2009b).

The relative rates of Glp+ mutation were determined in minimal and complex media. On glycerol plates, colonies first appeared after about 3 days (**Figure 1B**) although wild type (*wt*) and *crp* Glp+ *E. coli* cells formed visible colonies in <2 days (**Figure 1C**). When the same cells were plated as before, but small numbers of *crp* Glp+ mutant cells were included with the *crp* cells before plating, colonies appeared from the *crp* Glp+ cells within 2 days, and new Glp+ mutants arose at the same rate as before (**Figure 1C**). This experiment showed that the mutants that arose on the plates had not existed in the cell population when initially plated. Thus, the Glp+ mutants arising from *crp* cells on these plates arose during incubation on the plate, and no growth inhibitor was present. The rate of mutation on glycerol plates proved to be 10 times higher than in minimal sorbitol or complex LB medium, and it was over 100 times higher than in
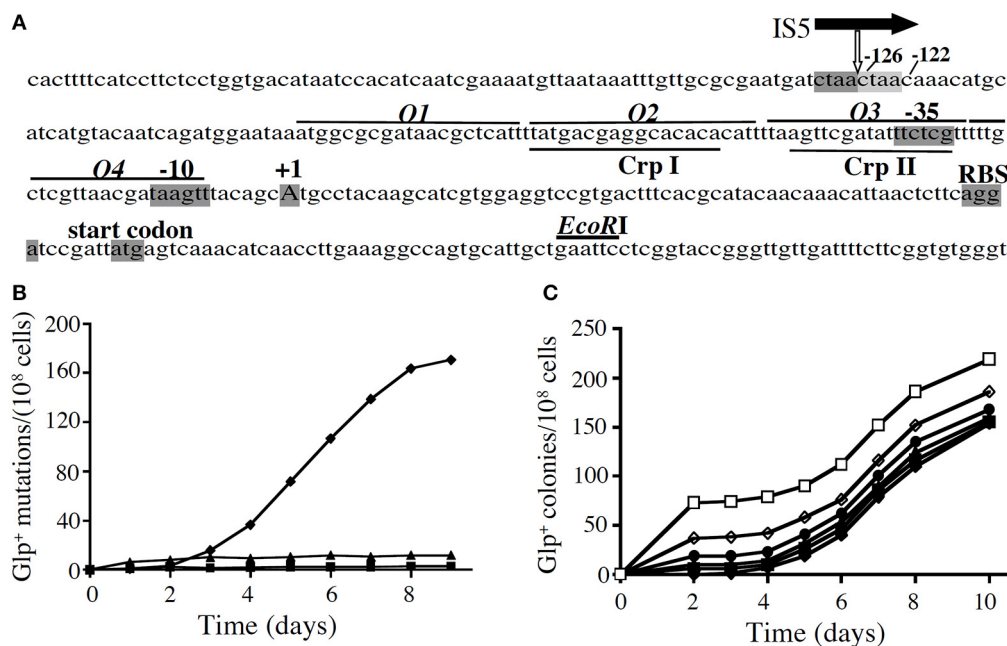
FIGURE 1 | The appearance of Glp⁺ mutations in a *crp* genetic background. (A) The *glpFK* promoter region. The transcriptional initiation site (+1), the −10 and −35 hexamers of the promoter, the ribosome binding site (RBS) and the start codon for *glpF* are shaded. The GlpR binding sites (O1–O4; lines above the sequence) and CRP binding sites (CrpI and CrpII; lines below the sequence) are also shown. The location of the IS5 element upstream of the promoter is indicated by the vertical white arrow below the horizontal black arrow representing IS5. The two 4-nucleotide direct repeats (ctaa), caused by IS5 insertion, are shaded dark and light gray, respectively. **(B)** Glp⁺ mutations in a *crp* genetic background in various media: solid M9 minimal media + 1% glycerol (◆), 0.01% glucose (■), or 1% sorbitol (▲). *crp* cells from an overnight culture (from a single colony) were applied onto agar plates, and the plates were incubated at 30°C. On glycerol minimal plates, mutations were determined by the presence of colonies on the plate. On sorbitol and glucose plates, mutations were determined by washing all cells off the plates and measuring colony formation both on LB plates (total cells) and on minimal glycerol plates (Glp⁺ cells) after 36–48 h of growth. **(C)** The same minimal glycerol agar plates, where the *crp* mutant cells were plated together with various numbers of *crp* Glp⁺ cells. Numbers of *crp* Glp⁺ cells were: (□, 72; ◇, 38; ●, 19; ▲, 10; ■, 5; and ◆, 0). These cells were mixed with *crp* cells and then applied onto M9 glycerol agar plates.

glucose-containing medium (Zhang and Saier, 2009a).

Mutation proved to be due to IS5 hopping to a discrete site, between base pairs 126 and 127 upstream of the transcriptional start site, and always in the same orientation. The increase in mutation rate was specific to the *glpFK* operon and did not occur in three other operons examined (Zhang and Saier, 2009a). Only the downstream 177 bp region of IS5 was required for activation of the *glpFK* operon, and this proved to be due to the presence of a permanent bend and an overlapping IHF binding site, each of which was responsible for half of the activation (Zhang and Saier, 2009b). This mechanism of activation (but not the actual transposon hopping event), presumably involving DNA looping, could also be demonstrated for the lactose (*lac*) operon in a *crp* genetic background of *E. coli* (Zhang and Saier, 2009b).

DEPENDENCY OF THE Glp⁺ MUTATION RATE ON GlpR

Glycerol is phosphorylated by glycerol kinase (GlpK) to glycerol-3-phosphate which binds to and releases GlpR from its operators (Lin, 1976). When GlpR dissociates from its operators, due to glycerol-3-phosphate binding, a conformational change could be transmitted through the DNA, promoting insertion of IS5 at the target CTAA site upstream of the *glpFK* promoter. In other words, GlpR binding might have two functions: repression of gene expression and suppression of IS5 transposition to the upstream activating site.

To test this possibility, the *glpR* gene was deleted, and the rates of appearance of Glp⁺ mutations in the *crp glpR* double mutant background were measured in the absence and presence of glycerol. The numbers of Glp⁺ cells arising was 10-fold higher in the *crp glpR*

double mutant than in the *crp* mutant when glycerol was absent. In the presence of glycerol, the loss of GlpR was without effect. Thus, deletion of *glpR* is equivalent to the inclusion of excess glycerol in the growth medium. High level overexpression of *glpR* decreased mutation rate to background levels (Zhang and Saier, 2009a). Clearly, GlpR regulates the increased rate of IS5-mediated insertional activation of the *glpFK* promoter by glycerol.

GlpR OPERATORS DIFFERENTIALLY CONTROL *glpFK* EXPRESSION AND Glp⁺ MUTATION RATE

There are four GlpR binding sites, O1–O4, in the upstream *glpFK* operon regulatory region (see **Figure 1A**), identified by DNA footprinting (Freedberg and Lin, 1973; Zeng et al., 1996). We mutated the far upstream site (O1) and the far downstream site (O4) so they no longer could

bind GlpR, and compared the effects on *glpFK* expression using a *lacZ* reporter gene fusion construct vs. mutation rate to Glp$^+$ during growth in LB medium. Mutation of *O4* increased *glpFK* operon expression about 5-fold although mutation of *O1* was almost without effect (Zhang and Saier, 2009a). By contrast, loss of *O1* yielded a 7-fold increase in mutation rate although loss of *O4* had only a 2-fold effect on mutation rate. We confirmed that IS5 was always in the same position and orientation (Zhang and Saier, 2009a). Thus, *O1* primarily controls the rate of IS5 insertion into the activating site, while *O4* primarily controls *glpFK* expression.

CONTROL OF IS5-MEDIATED *glpFK* OPERON ACTIVATION BY THE cAMP-CRP COMPLEX

As noted above, IS5-mediated activation of the *glpFK* promoter was observed in a *crp* genetic background. Initial attempts in our laboratory and elsewhere to isolate such mutants in a wild type genetic background proved unsuccessful (Ibarra et al., 2002; Honisch et al., 2004; Fong et al., 2005; Zhang and Saier, 2011, unpublished observations; Cheng et al., 2014). Since *crp* mutants are not found in nature, this brought into question the suggestion that our discovery of directed mutation in a *crp* mutant of *E. coli* was relevant to the wild type situation, and hence whether it

had actually evolved under the pressures of natural selection.

We consequently undertook a systematic analysis of the cAMP dependency of IS5-mediated *glpFK* activation to understand why *crp* mutants but not *wt* cells gave rise to IS5-activated mutants. The *cya* gene, encoding the cAMP biosynthetic enzyme, adenylate cyclase (Cya), was deleted, and as expected, IS5-mediated *glpFK* activation was observed at high frequency. When sub-mM concentrations of cAMP were added to the growth medium, the frequency of these mutations was drastically reduced; 1 mM exogenous cAMP essentially prevented the appearance of these mutants. When *glpR* was deleted in
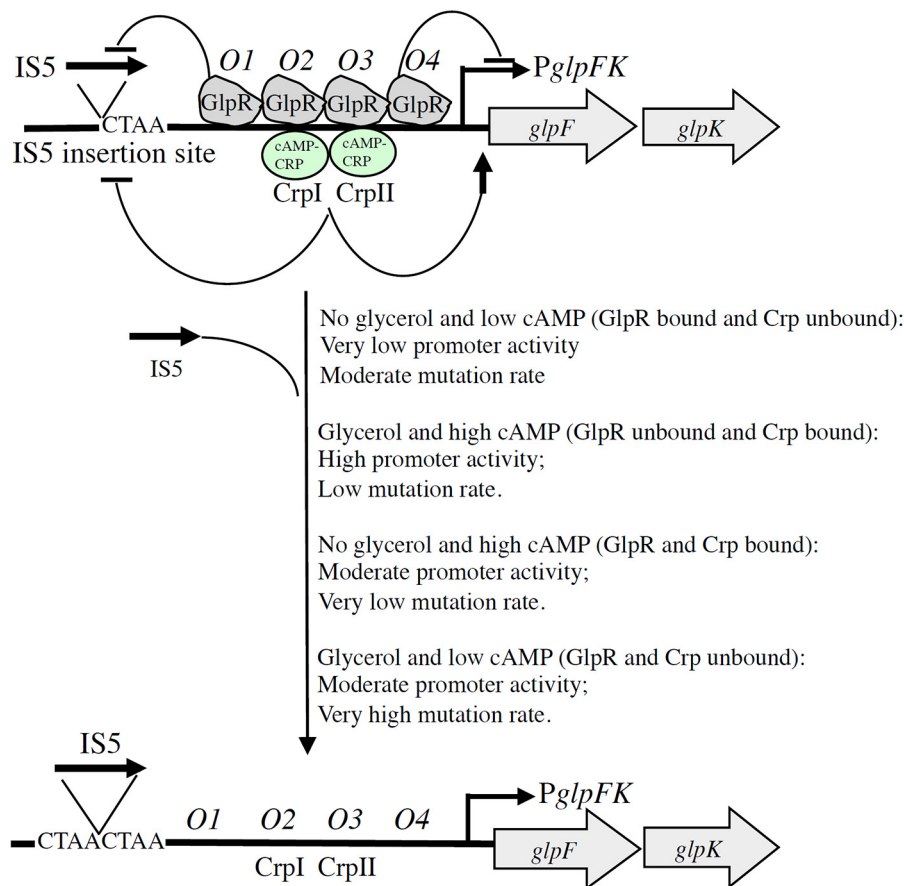

FIGURE 2 | Schematic diagram illustrating dual GlpR-mediated/cAMP-CRP-mediated control of (right) the level of *glpFK* transcription and (left) the rate of IS5 hopping into the activating site upstream of the *glpFK* promoter (directed mutation). With GlpR bound to its operators (*O1–O4*) (in the presence of GlpR and the absence of glycerol), transcription and IS5 hopping both occur at low rates. When GlpR is not bound to its operators (in the absence of GlpR or in the presence of glycerol), both transcriptional initiation

and IS5 hopping increase about 10×. Binding of GlpR to operator *O1* preferentially blocks IS5 insertion, while binding of GlpR to operator *O4* preferentially blocks transcription as indicated. Binding of cAMP-CRP to its transcriptional activating sites, CrpI and CrpII, similarly inhibits IS5 hopping even though binding of this complex promotes *glpFK* transcription. Glucose inhibits IS5 insertion by a mechanism independent of glycerol, GlpR, cAMP, and CRP. (⟶, activation; ⊢, inhibition or repression).

the *cya* genetic background, the frequencies of IS5 insertion increased by about 20-fold (Zhang and Saier, unpublished observations).

These observations led us to experiment with *wt E. coli* cells. Glycerol utilization in these cells is strongly inhibited by the presence of the non-metabolizable glucose analogs, 2-deoxyglucose and α-methyl glucoside, which also lower cytoplasmic cAMP levels by inhibiting adenylate cyclase activity (Saier and Feucht, 1975; Castro et al., 1976; Saier et al., 1976; Feucht and Saier, 1980; Saier, 1989). Wild type cells were therefore plated on minimal salts medium containing 0.2% glycerol and 0.1% 2-deoxyglucose or 0.5% α-methyl glucoside. Not surprisingly, IS5 insertional directed mutants could be isolated under these conditions (Zhang and Saier, unpublished results). Abolition of the CRP binding sites in the *glpFK* upstream regulatory region greatly enhanced the frequency of these mutants in an otherwise *wt* genetic background, even in the absence of a non-metabolizable glucose analog, showing that binding of the cAMP-CRP complex to the *glpFK* control region negatively regulates IS5 insertion. These experiments showed that directed mutation is negatively regulated by binding of both the glycerol repressor, GlpR, and the cAMP-CRP complex to the *glpFK* control region. This explains why IS5 insertion into the activating site of the *glpFK* regulatory region in wild type cells depends on both high glycerol and low cyclic AMP concentrations.

DIRECTED MUTATIONS PROMOTING EXPRESSION OF OTHER OPERONS IN *E. COLI*

Following the reports of Zhang and Saier (2009a,b) cited above, Wang and Wood (2011) demonstrated directed mutation of the operon (*flhDC*) encoding the flagellar transcriptional master switch in *E. coli*, FlhDC. IS5 insertion into the upstream control region of the *flhDC* operon was responsible. Although the mechanism was not established, the frequency of IS5 insertion clearly increased under swarming conditions in soft agar compared to growth in liquid medium or on solid agar plates where swarming does not occur (Wang and Wood, 2011). We

have confirmed and extended their results (Zhang et al., 2013). Moreover, preliminary results suggest that the *E. coli fuc* (fucose; propanediol) operon may also be subject to IS5-mediated directed mutation (Zhang and Saier, 2011; Zhang et al., 2013). It seems that transposon-mediated directed mutation will prove to be important to microbial evolution, and possibly to that in other organisms, partly accounting for the prevalence of these genetic elements in virtually all living organisms.

CONCLUSIONS AND PERSPECTIVE

Directed mutation has been defined as genetic change that is specifically induced by the stress conditions that the mutation relieves, but until recently, in no case had such a mechanism been established. We have demonstrated that mutations in the *glpFK* control region, allowing growth of *E. coli crp* or *cya* mutants on glycerol plates, or *wt* cells on glycerol plus 2-deoxyglucose or α-methyl glucoside plates, are specifically induced by the presence of these compounds. The glycerol regulon repressor, GlpR, which binds to its four operators (*O1–O4*) in front of the *glpFK* operon (**Figure 1A**) and is displaced from these sites when α-glycerol phosphate binds allosterically to GlpR (Zeng et al., 1996), not only controls gene expression, but also controls mutation rate. In this case, GlpR binding negatively influences both *glpFK* operon expression and operon activation by IS5. Our results established that binding of GlpR to *O4*, which overlaps the −10 promoter region, primarily controls gene expression, while binding to *O1* primarily controls IS5 hopping into the specific CTAA site, between 127 and 122 base pairs upstream of the *glpFK* transcriptional start site, that activates the *glpFK* promoter. Binding of the cAMP-CRP complex to its two binding sites overlapping *O2* and *O3* also blocks IS5 hopping to the activating site. This dual mechanism may involve changes in DNA conformation or supercoiling. The results serve to dissociate two functions of both GlpR and CRP.

The mechanism of IS5-mediated *glpFK* promoter activation in wild type cells provides relief from starvation when glycerol is present and a cAMP-depressing toxic substance, such as 2-deoxyglucose, is simultaneously present. Such non-metabolizable sugar derivatives

are synthesized by many organisms and therefore are present in nature. This example of directed mutation could therefore have been selected for during evolution. It appears to be a genuine example of directed mutation, with mutations arising at a greater rate under conditions that allow benefit to the organism. The fact that mutation rate is influenced by the presence of glycerol in a process mediated by the glycerol repressor, and by cAMP in a process mediated by CRP, provides mechanistic explanations for IS5-mediated directed mutational control. This mechanism, illustrated in **Figure 2**, allows rationalization of the presence of four GlpR binding sites and two CRP binding sites in the control region of the *glpFK* operon. Our studies also provide the rationale for the evolution of this elaborate mechanism of gene activation.

REFERENCES

Cairns, J., Overbaugh, J., and Miller, S. (1988). The origin of mutants. *Nature* 335, 142–145. doi: 10.1038/335142a0

Campos, D., Matos, S., and Oliveira, J. L. (2013). A modular framework for biomedical concept recognition. *BMC Bioinformatics* 14:281. doi: 10.1186/1471-2105-14-281

Castro, L., Feucht, B. U., Morse, M. L., and Saier, M. H. Jr. (1976). Regulation of carbohydrate permeases and adenylate cyclase in *Escherichia coli*. Studies with mutant strains in which enzyme I of the phosphoenolpyruvate:sugar phosphotransferase system is thermolabile. *J. Biol. Chem.* 251, 5522–5527.

Chandler, M., and Mahillon, J. (2002). "Insertion sequences revised," in *Mobile DNA II*, eds N. L. Craig, R. Craigie, M. Gellert, and A. M. Lambowitz (Washington, DC: ASM Press), 305–306.

Cheng, K. K., Lee, B. S., Masuda, T., Ito, T., Ikeda, K., Hirayama, A., et al. (2014). Global metabolic network reorganization by adaptive mutations allows fast growth of *Escherichia coli* on glycerol. *Nat. Commun.* 5, 3233. doi: 10.1038/ncomms4233

Feucht, B. U., and Saier, M. H. Jr. (1980). Fine control of adenylate cyclase by the phosphoenolpyruvate:sugar phosphotransferase systems in *Escherichia coli* and Salmonella typhimurium. *J. Bacteriol.* 141, 603–610.

Fong, S. S., Joyce, A. R., and Palsson, B. O. (2005). Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene expression states. *Genome Res.* 15, 1365–1372. doi: 10.1101/gr.3832305

Foster, P. L. (1999). Mechanisms of stationary phase mutation: a decade of adaptive

mutation. *Annu. Rev. Genet.* 33, 57–88. doi: 10.1146/annurev.genet.33.1.57

Foster, P. L. (2005). Stress responses and genetic variation in bacteria. *Mutat. Res.* 569, 3–11. doi: 10.1016/j.mrfmmm.2004.07.017

Freedberg, W. B., and Lin, E. C. (1973). Three kinds of controls affecting the expression of the *glp* regulon in *Escherichia coli*. *J. Bacteriol.* 115, 816–823.

Honisch, C., Raghunathan, A., Cantor, C. R., Palsson, B. O., and Van Den Boom, D. (2004). High-throughput mutation detection underlying adaptive evolution of *Escherichia coli*-K12. *Genome Res.* 14, 2495–2502. doi: 10.1101/gr.2977704

Ibarra, R. U., Edwards, J. S., and Palsson, B. O. (2002). *Escherichia coli* K-12 undergoes adaptive evolution to achieve *in silico* predicted optimal growth. *Nature* 420, 186–189. doi: 10.1038/nature 01149

Lin, E. C. (1976). Glycerol dissimilation and its regulation in bacteria. *Annu. Rev. Microbiol.* 30, 535–578. doi: 10.1146/annurev.mi.30.100176. 002535

Mahillon, J., and Chandler, M. (1998). Insertion sequences. *Microbiol. Mol. Biol. Rev.* 62, 725–774.

Rosenberg, S. M. (2001). Evolving responsively: adaptive mutation. *Nat. Rev. Genet.* 2, 504–515. doi: 10.1038/35080556

Roth, J. R., Kugelberg, E., Reams, A. B., Kofoid, E., and Andersson, D. I. (2006). Origin of mutations under selection: the adaptive mutation controversy. *Annu. Rev. Microbiol.* 60, 477–501. doi: 10.1146/annurev.micro.60.080805. 142045

Saier, M. H. Jr. (1989). Protein phosphorylation and allosteric control of inducer exclusion and catabolite repression by the bacterial phosphoenolpyruvate: sugar phosphotransferase system. *Microbiol. Rev.* 53, 109–120.

Saier, M. H. Jr., and Feucht, B. U. (1975). Coordinate regulation of adenylate cyclase and carbohydrate permeases by the phosphoenolpyruvate:sugar phosphotransferase system in Salmonella typhimurium. *J. Biol. Chem.* 250, 7078–7080.

Saier, M. H. Jr., Feucht, B. U., and Hofstadter, L. J. (1976). Regulation of carbohydrate uptake and adenylate cyclase activity mediated by the enzymes II of the phosphoenolpyruvate: sugar phosphotransferase system in *Escherichia coli. J. Biol. Chem.* 251, 883–892.

Schnetz, K., and Rak, B. (1992). IS5: a mobile enhancer of transcription in *Escherichia coli. Proc. Natl. Acad. Sci. U.S.A.* 89, 1244–1248. doi: 10.1073/pnas.89.4.1244

Wang, X., and Wood, T. K. (2011). IS5 inserts upstream of the master motility operon *flhDC* in a quasi-Lamarckian way. *ISME J.* 5, 1517–1525. doi: 10.1038/ismej. 2011.27

Weissenborn, D. L., Wittekindt, N., and Larson, T. J. (1992). Structure and regulation of the glpFK operon encoding glycerol diffusion facilitator and glycerol kinase of *Escherichia coli* K-12. *J. Biol. Chem.* 267, 6122–6131.

Wright, B. E. (2004). Stress-directed adaptive mutations and evolution. *Mol. Microbiol.* 52, 643–650. doi: 10.1111/j.1365-2958.2004. 04012.x

Zeng, G., Ye, S., and Larson, T. J. (1996). Repressor for the sn-glycerol 3-phosphate regulon of *Escherichia coli* K-12: primary structure and identification of the DNA-binding domain. *J. Bacteriol.* 178, 7080–7089.

Zhang, Z., and Saier, M. H. Jr. (2009a). A mechanism of transposon-mediated directed mutation. *Mol. Microbiol.* 74, 29–43. doi: 10.1111/j.1365-2958.2009.06831.x

Zhang, Z., and Saier, M. H. Jr. (2009b). A novel mechanism of transposon-mediated gene activation. *PLoS Genet.* 5:e1000689. doi: 10.1371/journal.pgen.1000689

Zhang, Z., and Saier, M. H. Jr. (2011). Transposon-mediated adaptive and directed mutations and their potential evolutionary benefits. *J. Mol. Microbiol. Biotechnol.* 21, 59–70. doi: 10.1159/000333108

Zhang, Z., Wang, J., Shlykov, M. A., and Saier, M. H. Jr. (2013). "Transposon mutagenesis in disease, drug discovery, and bacterial evolution," in *Stress-Induced Mutagenesis*, ed D. Mittelman (New York, NY: Springer), 59–77. doi: 10.1007/978-1-4614-6280-4_4

frontiers in MICROBIOLOGY

Helicobacter pylori DNA methyltransferases and the epigenetic field effect in cancerization

Ramakrishnan Sitaraman *

Department of Biotechnology, TERI University, New Delhi, India
*Correspondence: minraj@gmail.com

INTRODUCTION

Helicobacter pylori, a Gram-negative, microaerophilic bacterium, has co-existed with humans beings as a prominent member of their gastric microbiota for approximately 10^5 years (Moodley et al., 2012). It infects approximately half the world's population, and most infected individuals are asymptomatic, but histologically exhibit superficial gastritis (The EUROGAST Study Group, 1993). Only a minority of infected individuals develop gastric or duodenal ulcers that necessitate treatment. Prolonged inflammation caused by chronic (often lifelong) infection predisposes a small fraction of infected individuals to develop gastric adenocarcinoma or lymphoma of the mucosa-associated lymphoid tissue (MALT lymphoma) (Passaro et al., 2002). Unfortunately, the prognosis for cases of gastric cancer is very poor, with 5-year survival rates being lower than 15% (Peek and Blaser, 2002).

A mechanism for carcinogenesis ensuing from *H. pylori*-triggered inflammation was first proposed by Pelayo Correa (Correa, 1992; Correa and Piazuelo, 2012). Briefly, chronic inflammation causes superficial gastritis that progresses over time to multifocal atrophic gastritis (MAG), characterized by the destruction of gastric glands. This is followed by intestinal metaplasia, wherein gastric epithelium undergoes an "epithelial-mesenchymal transition" and begins to exhibit an intestinal phenotype. The subsequent stage consists of dysplasia culminating in invasive carcinoma, which completes the "pre-cancerous cascade." The final outcome is also dependent on host and pathogen genotypes, as well as

environmental factors such as socioeconomic indicators, a high-salt diet, low fruit/vegetable intake and smoking (Khalifa et al., 2010). Most notably, *H. pylori* is the sole bacterium to be classified by the WHO as a class I carcinogen (IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, 1994).

A MECHANISM FOR EPIGENETIC FIELD CANCERIZATION BY *H. pylori* DNA MTases

Field cancerization is a concept first proposed in 1953 in the context of oral stratified squamous epithelium (Slaughter et al., 1953), and subsequently extended to other tissues. Briefly, upon exposure to a carcinogen at sufficient intensity for a significant duration, grossly normal-looking tissue near tumor sites suffers microscopic (essentially, molecular) changes that eventually result in carcinogenesis. Aberrant methylation of cytosine residues within CpG islands (CGI) in genomic DNA has been reported in a variety of cancers, including gastric cancer (Laird and Jaenisch, 1994; Laird, 2005). This is an epigenetic change that could contribute to cancer development by the process of field cancerization (Ramachandran and Singal, 2012). Chronic *H. pylori* infection in humans is associated with hypermethylation of promoter sequences of different categories of genes, resulting in downregulation of transcription. Some of these are: *CDH1* that codes for E-cadherin, a transmembrane glycoprotein involved in maintaining epithelial integrity (Chan et al., 2003); *GATA4* and *GATA5* encoding transcription factors (Wen et al., 2010); and *TFF2* (encoding trefoil factor 2) (Peterson et al., 2010) and *FOXD3* (encoding a

forkhead box transcriptional regulator) that are tumor suppressors (Cheng et al., 2013). However, one unexplored possibility is that one or more of the several functional DNA methyltransferases (MTases) of *H. pylori* could enter host cells and methylate their recognition sequences in chromosomal DNA in an unregulated manner. The result would be the creation of an epigenetic field of cancerization.

THE NUMEROUS DNA MTases OF *H. pylori*

DNA MTases are sequence-specific DNA-binding enzymes that methylate adenine or cytosine residues in the context of their cognate recognition sequences using S-adenosylmethionine as the methyl donor, and are widespread among prokaryotes (Roberts et al., 2010). Depending on the enzyme in question, a methyl group may be added at the N^6 position in adenine forming N^6-methyladenine (m6A) or the N^4 or C^5 positions in cytosine forming N^4-methylcytosine (m4C) or C^5-methylcytosine (m5C) respectively. DNA methylation of regulatory sequences is known to result in changes in gene expression in a wide variety of organisms, both prokaryotes and eukaryotes. Therefore, should DNA MTases encoded by pathogens gain entry into host cells by specific or non-specific means, there is a strong possibility that they could modify host regulatory DNA sequences (nuclear, and perhaps even organellar), making the process at least partially inflammation-independent.

Several pathogens, including *H. pylori*, are known to introduce virulence factors into eukaryotic host cell by a variety of mechanisms. Recently, a type I DNA methyltransferase subunit (HsdM)

of *Klebsiella pneumoniae* was found to have a nuclear localization signal (NLS). When expressed in in the COS-1 (African green monkey kidney) cell line, HsdM localized to the nucleus. Surprisingly, HsdM was capable of methylating DNA even in the absence of the specificity subunit (HsdS), albeit at much lower levels (Lee et al., 2009). A recent study demonstrated that a transposase (Tnp) of *Acinetobacter baumanii* was targeted to the nuclei of A549 (a human lung carcinoma) and COS-7 (African green monkey kidney) cell lines, and that this resulted in specific CpG methylation of the *CDH1* (E-cadherin) promoter (Moon et al., 2012).

A survey of the database of restriction enzymes (REBASE; http://rebase.neb.com) indicates that *H. pylori* encodes a noticeably large number of DNA MTases, known or putative—ranging from 25 in the strain SouthAfrica7 to 37 in strain Puno135. Very few prokaryotes encode such a large number of DNA MTases or restriction-modification (R-M) systems. A majority of the predicted/known DNA MTases encoded by *H. pylori*, both adenine- and cytosine-specific, are type II enzymes (http://tools.neb.com/∼vincze/genomes/index.php?page=H). In this class of DNA MTases, the functions of sequence-specific DNA binding and DNA methylation are carried out by the same protein, and do not require any accessory protein factors for full activity. DNA transfer experiments between *H. pylori* strains clearly demonstrated sequence-specific DNA methylation in cell extracts (Donahue et al., 2000). Several studies have indicated that many of the DNA MTases encoded by the *H. pylori* genome are expressed and functional (Vitkute et al., 2001; Takata et al., 2002; Vale and Vítor, 2007; Kumar et al., 2012a), and can affect *H. pylori* protein expression in a strain-specific manner (Donahue et al., 2002; Takata et al., 2002; Kumar et al., 2012a; Vitoriano et al., 2013).

ENTRY OF *H. pylori* DNA MTases INTO HOST CELLS

There are at least three mutually non-exclusive routes by which DNA MTases could gain entry into host cells. Firstly, while *H. pylori* is predominantly extracellular, studies have indicated that it may be a facultatively intracellular as well (Kwok

et al., 2002; Necchi et al., 2007; Liu et al., 2012). As a chronic pathogen, its intracellular persistence could conceivably result in the DNA MTases gaining access into the host cell cytoplasm. Secondly, *H. pylori* is also known to release membrane vesicles containing cellular proteins, and it is possible that DNA MTases could be transported into to the host cytoplasm in these vesicles. However, a recent proteomic study of *H. pylori* vesicles failed to detect any DNA MTases in them, indicating that this is unlikely, but it is possible that growth of *H. pylori* on plates or in broth might not correspond to the situation *in vivo* (Olofsson et al., 2010). Thirdly, many *H. pylori* strains encode components of a type IV secretion system (termed the *cag* pathogenicity island, *cag* PAI) that is capable of translocating a protein, CagA (Odenbreit et al., 2000), and peptidoglycan (Viala et al., 2004) into host cells. Presently, there is no conclusive data on whether or not DNA MTases or other cell components could be similarly translocated, though a recent computational prediction using indicates that 1–3 DNA MTases could translocated by the *cag*PAI (Wang et al., 2014), based on the model underlying the prediction.

Regardless of the mechanism by which bacterial proteins might enter the host cells, the fact remains that none of the known DNA MTases (or restriction endonucleases) of *H. pylori* possess any recognizable nuclear localization signals, so that the actual mechanism of nuclear translocation required for DNA methylation, if it happens, is still open to question.

INFERENCES FROM STUDIES IN THE MONGOLIAN GERBIL (*Meriones unguiculatus*) MODEL

Humans are the only known natural host for *H. pylori*. However, *H. pylori*-infected Mongolian gerbils reproducibly develop gastric adenocarcinoma upon oral *N*-methyl-*N*-nitosourea administration, and have therefore been used in animal studies for more than 15 years now (Watanabe et al., 1998). Niwa *et al.* have used this model system to examine DNA methylation in gastric cancer in detail over a duration of up to a year (Niwa et al., 2013, 2010). In their earlier study, they first demonstrated that carcinogenesis is accompanied by hypermethylation of

promoters, and that cyclosporin A (CsA), an anti-inflammatory agent, does not interfere with bacterial colonization of the animals, but abrogates DNA hypermethylation significantly. Their studies demonstrated that the inflammatory response to *H. pylori* infection in Mongolian gerbils is associated with an increase in DNA methylation in gastric epithelial cells (GECs). This was taken by them to imply that methylation is not directly attributable to any bacterial effectors such as CagA or DNA MTases (Niwa et al., 2010). However, the same study also observed an unexpected decrease in the transcription levels of the host DNA MTases (Dnmts) in the GECs of infected gerbils compared to uninfected controls. Is it possible that cellular Dnmts are down-regulated in response to the chronic burden of a large number of bacterial DNA MTases, and the significant association of *H. pylori* with cancer development is due, in some part, to its large complement of DNA MTases? An additional fact to consider is that bacterial DNA adenine MTases, depending on their specificity, could modify adenine residues in regulatory regions in the DNA. More importantly, the specificity of some adenine MTases may also be relaxed, resulting in cytosine methylation at the N^4 position (Jeltsch et al., 1999). A cytosine DNA MTase of *H. pylori* (M.HpyAVIB) was found to exhibit relaxed specificity upon mutation (Kumar et al., 2012b). Lastly, given that adenine methylation is not routinely examined in studies targeting promoter hypermethylation in humans on the basis of the very low incidence of m6A in mammalian DNA, it may well have been missed in studies concentrating on CGI methylation.

CONCLUSIONS

While the link between viruses and cancer has been extensively researched, it is notable that *Helicobacter pylori* has remained the best-studied and, for nearly two decades, the sole bacterial pathogen systematically linked with any type of cancer in clinical practice. Some bacterial effector molecules associated with carcinogenesis, such as CagA and VacA, have been studied in great detail. Owing to a combination of unique characteristics—encoding a large number of functional DNA MTases, lifelong persistence in the

host and facultative intracellularity—*H. pylori* may well be a unique member of the stomach microbiota that affects its host in unforeseen ways. The investigation of the effects of the entry of DNA MTases (and restriction endonucleases, including methylation-dependent restriction enzymes, that can cause DNA breaks) and other proteins of the microbiome into host cells has the potential to uncover novel interactions between evolutionarily disparate species. More generally, it is possible that these proteins are effectors of inter-specific epigenetic signals, that perhaps enable commensals, symbionts and pathogens to adapt to their ecological niches by modulating host gene expression. While housekeeping DNA MTases (e.g., the Dam methylase of *E. coli*) of bacteria, pathogenic or non-pathogenic, are known to be important for bacterial viability (Marinus and Casadesus, 2009), the role of bacterial DNA MTases in infectious diseases and importantly, in the evolution and maintenance of host-microbiome interactions remains unclear, and perhaps merits fresh consideration in terms of the epigenetic modulation of host physiology.

ACKNOWLEDGMENTS

REFERENCES

(1994). Infection with *Helicobacter pylori*. *IARC Monogr. Eval. Carcinog. Risks Hum.* 61, 177–240.

Chan, A. O.-O., Lam, S.-K., Wong, B. C.-Y., Wong, W.-M., Yuen, M.-F., Yeung, Y.-H., et al. (2003). Promoter methylation of E-cadherin gene in gastric mucosa associated with *Helicobacter pylori* infection and in gastric cancer. *Gut* 52, 502–506. doi: 10.1136/gut.52.4.502

Cheng, A. S. L., Li, M. S., Kang, W., Cheng, V. Y., Chou, J.-L., Lau, S. S., et al. (2013). *Helicobacter pylori* causes epigenetic dysregulation of FOXD3 to promote gastric carcinogenesis. *Gastroenterology* 144, 122.e9–133.e9. doi: 10.1053/j.gastro.2012.10.002

Correa, P. (1992). Human gastric carcinogenesis: a multistep and multifactorial process–first american cancer society award lecture on cancer epidemiology and prevention. *Cancer Res.* 52, 6735–6740.

Correa, P., and Piazuelo, M. B. (2012). The gastric precancerous cascade. *J. Dig. Dis.* 13, 2–9. doi: 10.1111/j.1751-2980.2011.00550.x

Donahue, J. P., Israel, D. A., Peek, R. M., Blaser, M. J., and Miller, G. G. (2000). Overcoming the restriction barrier to plasmid transformation of *Helicobacter pylori*. *Mol. Microbiol.* 37, 1066–1074. doi: 10.1046/j.1365-2958.2000.02036.x

Donahue, J. P., Israel, D. A., Torres, V. J., Necheva, A. S., and Miller, G. G. (2002). Inactivation of a *Helicobacter pylori* DNA methyltransferase alters dnaK operon expression following host-cell adherence. *FEMS Microbiol. Lett.* 208, 295–301. doi: 10.1016/S0378-1097(02)00479-2

Jeltsch, A., Christ, F., Fatemi, M., and Roth, M. (1999). On the substrate specificity of DNA methyltransferases. adenine-N6 DNA methyltransferases also modify cytosine residues at position N4. J. Biol. Chem. 274, 19538–19544. doi: 10.1074/jbc.274.28.19538

Khalifa, M. M., Sharaf, R. R., and Aziz, R. K. (2010). *Helicobacter pylori*: a poor man's gut pathogen? *Gut Pathog.* 2, 2. doi: 10.1186/1757-4749-2-2

Kumar, R., Mukhopadhyay, A. K., Ghosh, P., and Rao, D. N. (2012a). Comparative transcriptomics of *H. pylori* strains AM5, SS1 and their hpyAVIBM deletion mutants: possible roles of cytosine methylation. *PLOS ONE* 7:e42303. doi: 10.1371/journal.pone.0042303

Kumar, R., Sabareesh, V., Mukhopadhyay, A. K., and Rao, D. N. (2012b). Mutations in hpyAV-IBM, C^5 cytosine DNA methyltransferase from *Helicobacter pylori* result in relaxed specificity. *FEBS J.* 279, 1080–1092. doi: 10.1111/j.1742-4658.2012.08502.x

Kwok, T., Backert, S., Schwarz, H., Berger, J., and Meyer, T. F. (2002). Specific entry of *Helicobacter pylori* into cultured gastric epithelial cells via a zipper-like mechanism. *Infect. Immun.* 70, 2108–2120. doi: 10.1128/IAI.70.4.2108-2120.2002

Laird, P. W. (2005). Cancer epigenetics. *Hum. Mol. Genet.* 1, R65–R76. doi: 10.1093/hmg/ddi113

Laird, P. W., and Jaenisch, R. (1994). DNA methylation and cancer. *Hum. Mol. Genet.* 1487–1495.

Lee, J. C., Kim, D. S., Moon, D. C., Lee, J.-H., Kim, M. J., Lee, S. M., et al. (2009). Prediction of bacterial proteins carrying a nuclear localization signal and nuclear targeting of HsdM from *Klebsiella pneumoniae*. *J. Microbiol.* 47, 641–645. doi: 10.1007/s12275-009-0217-4

Liu, H., Semino-Mora, C., and Dubois, A. (2012). Mechanism of *H. pylori* intracellular entry: an *in vitro study*. *Front. Cell. Infect. Microbiol.* 2:13. doi: 10.3389/fcimb.2012.00013

Marinus, M. G., and Casadesus, J. (2009). Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more. *FEMS Microbiol. Rev.* 33, 488–503. doi: 10.1111/j.1574-6976.2008.00159.x

Moodley, Y., Linz, B., Bond, R. P., Nieuwoudt, M., Soodyall, H., Schlebusch, C. M., et al. (2012). Age of the association between *Helicobacter pylori* and man. *PLoS Pathog.* 8:e1002693. doi: 10.1371/journal.ppat.1002693

Moon, D. C., Choi, C. H., Lee, S. M., Lee, J. H., Kim, S. I., Kim, D. S., et al. (2012). Nuclear translocation

of *Acinetobacter baumannii* transposase induces DNA methylation of CpG regions in the promoters of E-cadherin gene. *PLoS ONE* 7:e38974. doi: 10.1371/journal.pone.0038974

Necchi, V., Candusso, M. E., Tava, F., Luinetti, O., Ventura, U., Fiocca, R., et al. (2007). Intracellular, intercellular, and stromal invasion of gastric mucosa, preneoplastic lesions, and cancer by *Helicobacter pylori*. *Gastroenterology* 132, 1009–1023. doi: 10.1053/j.gastro.2007.01.049

Niwa, T., Toyoda, T., Tsukamoto, T., Mori, A., Tatematsu, M., and Ushijima, T. (2013). Prevention of *Helicobacter pylori*-induced gastric cancers in gerbils by a DNA demethylating agent. *Cancer Prev. Res. (Phila).* 6, 263–270. doi: 10.1158/1940-6207.CAPR-12-0369

Niwa, T., Tsukamoto, T., Toyoda, T., Mori, A., Tanaka, H., Maekita, T., et al. (2010). Inflammatory processes triggered by *Helicobacter pylori* infection cause aberrant DNA methylation in gastric epithelial cells. *Cancer Res.* 70, 1430–1440. doi: 10.1158/0008-5472.CAN-09-2755

Odenbreit, S., Püls, J., Sedlmaier, B., Gerland, E., Fischer, W., and Haas, R. (2000). Translocation of *Helicobacter pylori* CagA into gastric epithelial cells by type IV secretion. *Science* 287, 1497–1500. doi: 10.1126/science.287.5457.1497

Olofsson, A., Vallström, A., Petzold, K., Tegtmeyer, N., Schleucher, J., Carlsson, S., et al. (2010). Biochemical and functional characterization of *Helicobacter pylori* vesicles. *Mol. Microbiol.* 77, 1539–1555. doi: 10.1111/j.1365-2958.2010.07307.x

Passaro, D. J., Chosy, E. J., and Parsonnet, J. (2002). *Helicobacter pylori*: consensus and controversy. *Clin. Infect. Dis.* 35, 298–304. doi: 10.1086/341245

Peek, R. M. Jr., and Blaser, M. J. (2002). *Helicobacter pylori* and gastrointestinal tract adenocarcinomas. *Nat. Rev. Cancer* 2, 28–37. doi: 10.1038/nrc703

Peterson, A. J., Menheniott, T. R., O'Connor, L., Walduck, A. K., Fox, J. G., Kawakami, K., et al. (2010). *Helicobacter pylori* infection promotes methylation and silencing of trefoil factor 2, leading to gastric tumor development in mice and humans. *Gastroenterology* 139, 2005–2017. doi: 10.1053/j.gastro.2010.08.043

Ramachandran, K., and Singal, R. (2012). DNA methylation and field cancerization. *Epigenomics* 4, 243–245. doi: 10.2217/epi.12.12

Roberts, R. J., Vincze, T., Posfai, J., and Macelis, D. (2010). REBASE–a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.* 38, D234–D236. doi: 10.1093/nar/gkp874

Slaughter, D. P., Southwick, H. W., and Smejkal, W. (1953). Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer* 6, 963–968.

Takata, T., Aras, R., Tavakoli, D., Ando, T., Olivares, A. Z., and Blaser, M. J. (2002). Phenotypic and genotypic variation in methylases involved in type II restriction-modification systems in *Helicobacter pylori*. *Nucleic Acids Res.* 30, 2444–2452. doi: 10.1093/nar/30.11.2444

The EUROGAST Study Group (1993). Epidemiology of, and risk factors for, *Helicobacter pylori* infection among 3194 asymptomatic subjects in 17 populations. Gut 34, 1672–1676.

Vale, F. F., and Vítor, J. M. B. (2007). Genomic methylation: a tool for typing *Helicobacter pylori* isolates. *Appl. Environ. Microbiol.* 73, 4243–4249. doi: 10.1128/AEM.00199-07

Viala, J., Chaput, C., Boneca, I. G., Cardona, A., Girardin, S. E., Moran, A. P., et al. (2004). Nod1 responds to peptidoglycan delivered by the *Helicobacter pylori* cag pathogenicity island. *Nat. Immunol.* 5, 1166–1174. doi: 10.1038/ni1131

Vitkute, J., Stankevicius, K., Tamulaitiene, G., Maneliene, Z., Timinskas, A., Berg, D. E., et al. (2001). Specificities of eleven different DNA methyltransferases of *Helicobacter pylori* strain 26695. *J. Bacteriol.* 183, 443–450. doi: 10.1128/JB.183.2.443-450.2001

Vitoriano, I., Vítor, J. M. B., Oleastro, M., Roxo-Rosa, M., and Vale, F. F. (2013). Proteome variability among *Helicobacter pylori* isolates clustered according to genomic methylation. *J. Appl. Microbiol.* 114, 1817–1832. doi: 10.1111/jam.12187

Wang, Y., Wei, X., Bao, H., and Liu, S.-L. (2014). Prediction of bacterial type IV secreted effectors by C-terminal features. *BMC Genomics* 15:50. doi:10.1186/1471-2164-15-50

Watanabe, T., Tada, M., Nagai, H., Sasaki, S., and Nakao, M. (1998). *Helicobacter pylori* infection induces gastric cancer in mongolian gerbils. *Gastroenterology* 115, 642–648.

Wen, X.-Z., Akiyama, Y., Pan, K.-F., Liu, Z.-J., Lu, Z.-M., Zhou, J., et al. (2010). Methylation of GATA-4 and GATA-5 and development of sporadic gastric carcinomas. *World J. Gastroenterol.* 16, 1201–1208. doi: 10.3748/wjg.v16.i10.1201

Conflict of Interest Statement: Work on *Helicobacter pylori* phosopholipases in my laboratory is supported by a grant from the Department of Biotechnology,

Calcifying nanoparticles: one face of distinct entities?

Anton G. Kutikhin[1,2,3]*, Arseniy E. Yuzhalin[4], Vadim V. Borisov[1], Elena A. Velikanova[1], Alexey V. Frolov[1], Vera M. Sakharova[1], Elena B. Brusina[1,2] and Alexey S. Golovkin[1]

[1] Division of Experimental and Clinical Cardiology, Research Institute for Complex Issues of Cardiovascular Diseases under the Siberian Branch of the Russian Academy of Medical Sciences, Kemerovo, Russia
[2] Department of Epidemiology, Kemerovo State Medical Academy, Kemerovo, Russia
[3] Central Research Laboratory, Kemerovo State Medical Academy, Kemerovo, Russia
[4] Department of Oncology, Cancer Research UK and Medical Research Council Oxford Institute for Radiation Oncology, University of Oxford, Oxford, UK
*Correspondence: antonkutikhin@gmail.com

Calcifying nanoparticles (CNPs, nanobacteria, nanobacteria-like particles) were discovered as cell culture contaminants by Kajander et al. more than 25 years ago, and the first results of their work were published some years later (Kajander et al., 1997; Kajander and Ciftçioglu, 1998). The nature of CNPs has been obscured so far. Possibly representing a new proposed class of living organisms or inorganic nanostructures, CNPs exist as coccoid, coccobacillar, or bacillary particles of 80–500 nm in diameter, consisting of a central cavity surrounded by the hydroxyapatite shell, and possessing an ability to grow and divide in culture medium forming biofilms. The phenomenon of CNPs raised intensive discussions amongst the scientific community; there are active debates on their nature and potential role in clinical medicine. According to a number of fundamental and clinical studies, CNPs are suspected to cause ectopic calcification-related diseases such as atherosclerosis, heart valve calcification, placental calcification, nephrolithiasis, cholecystolithiasis, type III chronic prostatitis/chronic pelvic pain syndrome, and testicular microlithiasis. Electron microscopy is considered to be a gold standard for the visualization of CNPs; researchers usually observe the colonies of CNPs after the culturing in DMEM or RPMI-1640 for 4–8 weeks. Other means used by multiple groups for the detection of CNPs are various serological methods, including ELISA, immunohistochemistry, immunofluorescence reaction, immunoblotting, and Ouchterlony immunodiffusion. Although

PCR has been previously used to identify possible genome of CNPs, there are doubts on credibility of this method since primers could have been designed based not on putative genome sequences of CNPs but on the genome sequences of contaminating bacteria. The general features of CNPs are indicated in **Table 1**.

Since the beginning of 2012, when our group published a comprehensive review on the role of CNPs in biology and medicine (Kutikhin et al., 2012), a number of newer findings have been reported. Currently, it is possible to divide all researchers studying CNPs into three groups: (1) those that presume that CNPs are living organisms (nanobacteria), (2) those that consider them as culprits of pathological calcification, and (3) those that completely reject their pathological role and assume them as the result of physiological calcification.

A group from China investigated the role of CNPs in placental calcification. The authors successfully isolated and cultured CNPs from the majority of calcified placental samples but not from normal placental tissues, concluding that CNPs are associated with this pathology (Guo et al., 2012; Lu et al., 2012a,b). Further, the authors also detected ≪nucleic acid-like materials≫ within individual CNPs in consistency with previous results by Agababov et al. (2007). Despite the fact that sequencing of putative nanobacterial 16s rRNA was successful, the obtained sequences were 99% similar to *Agrobacterium tumefaciens* and *Rheinheimera* spp., which suggests sample contamination (Guo et al., 2012; Lu

et al., 2012a,b). In addition, the researchers speculated that genome of CNPs isolated from calcified placental samples may differ from the putative genome of CNPs isolated from human blood and other sources (Guo et al., 2012); nevertheless, it is arguable that this assumption can be proved by *in silico* analysis whether CNPs have nucleic acids or not.

The same research group recently carried out a first investigation of the impact of nanohydroxyapatites (nHAPs) and CNPs on growth and viability of cancer cells (Zhang et al., 2014). Both nHAPs and CNPs had considerable cytotoxic effects on MDA-MB-231 breast cancer cell line, including altered size and morphology, formation of large cytoplasmic vacuoles, inhibition of proliferation, and induction of apoptosis (Zhang et al., 2014). Importantly, CNPs caused more pronounced cytotoxic effects as compared to nHAPs, as they induced both early and late apoptosis and necrosis, whereas nHAPs triggered early apoptosis only (Zhang et al., 2014). After the internalization of CNPs by cancer cells, cell shrinkage, chromatin condensation, nuclear fragmentation, nuclear dissolution, and the formation of apoptotic bodies were detected, although nHAPs did not exhibit such an effect (Zhang et al., 2014). No cytotoxic effects were observed in the untreated control group (Zhang et al., 2014).

Another investigation on the role of CNPs in calcification-related diseases was performed by a group from Spain concerning aortic valve calcification (Barba et al., 2012). The authors successfully

Table 1 | Properties of calcifying nanoparticles (CNPs).

Properties of calcifying nanoparticles	
Morphological properties	80–500 nm in diameter (can pass through 100-nm filters)
	Usually have coccoid, coccobacillar, or bacillar form, often remind mineralized «igloo»
	Have hydroxyapatite shell, cellular-membranous structure, and central cavity
	Can form microscopic colonies (>1 mm in diameter) under the low concentration of nutrients in the environment
	Divide by binary fission, fragmentation, and gemmation
	Can form thermoresistant biofilms
Tinctorial properties	Gram-negative, can be stained by DNA-specific dyes (Hoechst 33258 at a concentration of 5 µg/ml during 50 min, propidium iodide, PicoGreen—they are specific after the filtration through 0.1–0.22 µm pores, staining after a demineralization is optimal), can be revealed by von Kossa staining, staining by 2% uranyl acetate (possibly with lead citrate) can detect specific mucus on the hydroxyapatite shell, in the mineralized state can be stained by alizarin red S, can be stained by phosphotungstic acid
Resistance and sensitivity	Resistant to 90°C heating during 1 h
	Resistant to γ-irradiation up to 30 kGray
	Resistant to 5% NaCl solution
	In the mineralized state are resistant to lyzozyme, proteinase K, certain other proteinases, lipases, amylases, alkali, ultrasound, X-ray, detergents, and solvents
	Temperature under 37°C suppresses replication and prevents biofilms formation
	Resistant to wide spectrum of antimicrobial therapeutics: aminoglycosides (in pharmacological concentrations), chloramphenicol, lincosamides, cephalosporins, macrolides, fluoroquinolones, glycopeptides (in pharmacological concentrations), polymyxins, antituberculous agents, aminocyclitol, spectinomycin
	Sensitive to tetracyclines, ampicillin, trimethoprim, trimethoprim-sulfamethoxazole, nitrofurantoin, 5-fluorouracil, cytosine arabinoside, antimycin A, sodium azide, potassium cyanide, bisphosphonates (etidronate, clodronate), 6-aminocaproic acid, *in vitro* growth can be inhibited by EDTA, EGTA, and citrate
Culture properties	Doubling time is 3 days, in media without serum—6 days
	Passage can be performed in DMEM (Dulbecco's modified Eagle's medium) or RPMI-1640 independently of the presence of serum
	Optimal atmosphere for growth should contain 5% of CO_2 and 95% of air
	Sensitive to β-mercaptoethanol, which stimulates growth of anaerobes, but cannot be cultivated under the strictly anaerobic conditions
	Calcify when serum concentration in the culture media is lower than 5%
	Cytotoxic for fibroblasts and lymphocytes
Biochemical properties	Metabolism is 10,000 times slower than in *E. coli*
	Incorporate uridine (into the expected nucleic acid), methionine, and aspartic acid (into the expected system of protein biosynthesis)
	Calcify under the physiological pH (7.4)
	Urease negative
Detection methods	Bacterioscopic (DNA-specific dyes Hoechst 33258, propidium iodide, PicoGreen, staining after a demineralization is optimal) using scanning and transmission electron microscopy, von Kossa staining which is specific for calcium compounds, staining by 2% uranyl acetate (possibly with lead citrate) to detect specific mucus on the hydroxyapatite shell, staining by alizarin red S in the mineralized state, staining by phosphotungstic acid, after the long-term cultivation light microscopy with von Kossa staining is possible to be used for detection
	Bacteriological (cultivation in DMEM or RPMI-1640 without serum under 37°C during 4–6 weeks after the filtration through 0.1–0.22 µm pores), replication can be assessed by spectrophotometry (650 nm wavelength)
	Serological CNP antigens and anti-CNP monoclonal antibodies 8/0 (to porin), 5/2 (to peptidoglycan) and 8D10 (to porin) of NanoBiotech Pharma, Tampa, FL, USA—ELISA, immunohistochemistry, immunofluorescence reaction, immunoblotting, Ouchterlony immunodiffusion
	Genomic (PCR), but there is doubt that existing primers are obtained on the basis of CNP nucleic acids, and not on the basis of nucleic acids of contaminating bacteria
	Proteomic (sodium dodecyl sulfate polyacrylamide gel electrophoresis with further identification of protein bands by mass spectrometry)

(Continued)

Table 1 | Continued

Properties of calcifying nanoparticles	
Methods of treatment of CNP-associated diseases (atherosclerosis, heart valve calcification, placental calcification, nephrolithiasis, cholecystolithiasis, type III chronic prostatitis/chronic pelvic pain syndrome, and testicular microlithiasis)	comET-therapy (tetracycline, EDTA, and mix of nutrients)

Abbreviations: CNPs, calcifying nanoparticles; EDTA, ethylenediaminetetraacetic acid; EGTA, ethyleneglycoltetraacetic acid; DMEM, Dulbecco's modified Eagle's medium; RPMI-1640, Roswell Park Memorial Institute 1640 medium; ELISA, enzyme-linked immunosorbent assay; PCR, polymerase chain reaction. Cited from Kutikhin et al. (2012).

cultured and isolated CNPs from calcified aortic valves but not from uncalcified control valves, showing a feasible pathological role of CNPs; however, no metabolic activity was observed in the samples, and authors failed to detect CNPs' presupposed nucleic acids by real-time PCR (Barba et al., 2012).

Other researchers did not attempt to extract CNPs from clinical specimens; instead, they investigated the inorganic properties of CNPs and conditions of their formation. Wu et al. (2013a) demonstrated that various charged elements and ions may form mineralo-organic nanoparticles (so-called bions) with bacteria-like morphology and similar properties in biological fluids. Upon formation, bions precipitated with phosphate, accumulated carbonate apatite, incorporated additional elements and thus reflected the ionic milieu of the biological fluid in which they formed (Wu et al., 2013a). Bions were able to increase in size and number and to be sub-cultured in fresh culture medium (Wu et al., 2013a). So, many morphological and cultural features of bions are similar to ones typical for CNPs. The authors suggested that bions may represent a part of a physiological cycle that regulates the function, transport, and disposal of mineral ions in the body, and the accumulation of bions may cause pathological processes in the human body under the conditions of altered calcium homeostasis and disturbed clearance mechanisms (Wu et al., 2013a). So, a proposed hypothesis was that bions may form under both physiological and pathological conditions (Wu et al., 2013a). In addition, the same research group suggested that membrane vesicles secreted by various

cells may contain a number of serum proteins and can induce hydroxyapatite precipitation during incubation in cell culture medium forming nanostructures which morphologically resemble CNPs (Wu et al., 2013b). Treatment of these membrane vesicles with anti-phosphatidylserine antibodies resulted in decrease of their mineral seeding activity suggesting that phosphatidylserine may provide nucleating sites for calcium phosphate deposition on the vesicles (Wu et al., 2013b).

Using dynamic light scattering, Peng et al. (2013) showed that serum and ion concentrations within the physiological range may form nanoparticles below 100 nm in diameter which can be phagocytosed by macrophages in a size-independent manner. However, only large nanoparticles or their aggregates were able to induce the production of mitochondrial reactive oxygen species, caspase-1 activation, and secretion of interleukin-1β and therefore cause inflammation (Peng et al., 2013). In addition, the authors found that the set of particle-bound proteins does not depend on particle size and curvature (Peng et al., 2013). According to the work of Baum et al. (2012), aggregates of hemoglobin and various salts from physiological environment can also produce morphological structures resembling CNPs. Finally, Kumon et al. investigated one of the original CNP isolates from urinary stones (P-17) (Kumon et al., 2014). The authors developed anti-P-17 IgM monoclonal antibodies, CL-15, which were specific for oxidized lipids, and combined immunoelectron microscopy with ultrastructural and elemental analysis (Kumon

et al., 2014). They suggested that lamellar structures consisted of acidic/oxidized lipids provided structural scaffolds for carbonate apatite and that lipid peroxidation induced by γ-irradiation of fetal bovine serum (FBS) was a major cause of CNP propagation (Kumon et al., 2014). Moreover, it was proposed that oxidized lipids may be a common platform for ectopic calcification in atherosclerosis-prone (ApoE$^{-/-}$) mice, thus CNPs were suggested to be by-products rather than etiological agents of chronic inflammation (Kumon et al., 2014). However, it was also noted that propagation of CNPs largely depended on the amount of oxidized lipids available and therefore could play a role in disease progression (Kumon et al., 2014).

However, it should be clearly stated that distinct processes may lead to the formation of nanostructures which are morphologically similar. Bions, hemoglobin-salt aggregates and oxidized lipids with acidified functional groups may lead to the occurrence of nanomorphological phenomenon called CNPs; however, there is no reason to discount many studies where CNPs were significantly associated with ectopic calcification-related diseases. In patients with these diseases, CNPs were detected significantly more frequently by electron microscopy and cultural features in comparison with control samples. Therefore, despite the fact that the emergence of CNPs can be due to physiological processes such as hemoglobin-salt aggregation, they can also form in greater numbers under pathological conditions such as alteration of metabolic and mineral ion homeostasis or lipid peroxidation regardless of their nature, in this case being the culprits of the ectopic

calcification-related diseases. In addition, as demonstrated by Peng et al. (2013), large CNPs or their aggregates may cause chronic inflammation which plays a major role in the development of all these diseases. They should not be by-products of inflammation since: (a) injection of CNPs caused artery calcification (Schwartz et al., 2008), nephrolithiasis (Hu et al., 2010), cholecystolithiasis (Wang et al., 2006), and type III chronic prostatitis (Shen et al., 2010) in animal models; (b) treatment of CNPs by comET-therapy (tetracycline HCl, EDTA, and mixture of nutrients) led to significant decline of CNP detection rate, decreased calcification, and substantial therapeutic improvement in patients with coronary artery disease (CAD) and type III chronic prostatitis/chronic pelvic pain syndrome (Maniscalco and Taylor, 2004; Shoskes et al., 2005; Zhou et al., 2008). So, the phenomenon of CNPs definitely should be taken into account when we talk about the etiology and pathogenesis of ectopic calcification-related diseases.

Regarding the hypothesis of CNPs as of the smallest self-replicating life form on Earth (nanobacteria), the absence of a fairly accurately sequenced genome aborts all discussions about their putative living nature. Nevertheless, the diverse nature of CNPs (bions, hemoglobin-salt aggregates, products of lipid peroxidation, calcified membrane vesicles, other inorganic entities) still leaves a place for speculations that certain nanoscale organisms may also be one of the entities of CNPs. However, this suggestion should be interpreted with caution.

From our point of view, immunological and cytotoxic properties of CNPs are still underinvestigated in the light of their potential pathogenic role. The investigations of Zhang et al. (2014) and Peng et al. (2013) are good examples of this kind; however, new studies are clearly needed, particularly due to increasing role of nanomedicine in clinical practice (nanovesicles for drug delivery, nanobiosensors for disease diagnosis, therapeutic nanoparticles which possibly may act as nucleation agents, etc.). Possibly, B cells may produce antibodies to CNPs, at least to large CNPs and their aggregates, and it is necessary to clarify this issue due to its feasible importance for understanding of the immune response against

CNPs and for their serological detection. Moreover, anti-CNP antibodies may be of distinct structure due to feasibly different entities of CNPs, their different size and different proteins coating them. In addition, the significance of CNPs as etiological agents of ectopic calcification-related diseases may be also tested in animal models and clinical trials using anti-CNP treatment. Notwithstanding, current standards of CAD and peripheral artery disease (PAD) therapy do not include drugs specifically directed against the calcification. Unfortunately, comET-therapy which demonstrated certain clinical efficacy in treatment of CAD and type III chronic prostatitis/chronic pelvic pain syndrome should not be used for a prolonged treatment due to the increasing hazard of antibiotic resistance. Possibly, larger clinical trials of comET-therapy analogues that do not contain antibiotics but include other calcium-chelating agents can be worthwhile.

To conclude, these two years have revealed some new facts about CNPs:

- CNPs can be formed by distinct entities such as bions, hemoglobin-salt aggregates, products of lipid peroxidation, calcified membrane vesicles, other inorganic entities, and possibly by living organisms; all these entities may possess similar looking nanomorphological structure;
- CNPs can be formed both under the both physiological and pathological conditions depending on their entity; impairment of metabolic and mineral ion homeostasis and lipid peroxidation may cause increased propagation of CNPs in the human body that may be a direct cause of ectopic calcification-related diseases;
- Large CNPs or their aggregates may cause chronic inflammation by the enhanced production of mitochondrial reactive oxygen species, cytokines, and by activation of pro-apoptotic enzymes;
- CNPs and nHAPs possess a number of cytotoxic effects on cancer cell lines such as alteration of cell size and morphology, inhibition of proliferation, and induction of apoptosis;
- Placental calcification may be definitely added to the list of CNP-caused pathologies.

What is clear is that at the present time there is no universal theory that can entirely characterize CNPs, their characteristics, and biological/medical significance. We should not disclaim the fact that CNPs may form under the physiological conditions; however, there is irrefutable evidence that an appearance of CNPs in the living organism may cause ectopic calcification-related diseases, and CNPs definitely should not be considered as just physiological phenomenon. In addition, the fact that CNPs are more likely to be inorganic structures than life forms definitely should not cause underestimation of their potential pathogenic role. No doubt, further investigations will shed light on the nature of CNPs, their biological properties, and their role in clinical medicine.

REFERENCES

Agababov, R. M., Abashina, T. N., Suzina, N. E., Vainshtein, M. B., and Schwartsburd, P. M. (2007). Link between the early calcium deposition in placenta and nanobacterial-like infection. *J. Biosci.* 32, 1163–1168. doi: 10.1007/s12038-007-0118-9

Barba, I., Villacorta, E., Bratos-Perez, M. A., Antolín, M., Varela, E., Sanchez, P. L., et al. (2012). Aortic valve-derived calcifyng nanoparticles: no evidence of life. *Rev. Esp. Cardiol. (Engl. Ed.).* 65, 813–818. doi: 10.1016/j.recesp.2012.03.018

Baum, J. L., Jones, R. L., Manning, T. J., Nienow, J., and Phillips, D. (2012). Hemoglobin aggregates studied under static and dynamic conditions involving the formation of nanobacteria-like structures. *Acta. Pharm.* 62, 201–209. doi: 10.2478/v10007-012-0011-4

Guo, Y., Zhang, D., Lu, H., Luo, S., and Shen, X. (2012). Association between calcifying nanoparticles and placental calcification. *Int. J. Nanomedicine* 7, 1679–1686. doi: 10.2147/IJN.S29786

Hu, W. G., Wang, X. F., Xu, T., Li, J. X., Chen, L., Yu, C. F., et al. (2010). Establishment nephrolithiasis rat model induced by nanobacteria and analysis of stone formation. Beijing Da Xue Xue Bao 42, 433–435. doi: 10.3969/j.issn.1671-167X.2010.04.016

Kajander, E. O., and Ciftçioglu, N. (1998). Nanobacteria: an alternative mechanism for pathogenic intra- and extracellular calcification and stone formation. *Proc. Natl. Acad. Sci. U.S.A.* 95, 8274–8279. doi: 10.1073/pnas.95.14.8274

Kajander, E. O., Kuronen, I., Akerman, K., Pelttari, A., and Ciftcioglu, N. (1997). Nanobacteria from blood, the smallest culturable autonomously replicating agent on earth. *Proc. Soc. Photo Opt. Instrum. Eng.* 3111, 420–428. doi: 10.1117/12.278796

Kumon, H., Matsuura, E., Nagaoka, N., Yamamoto, T., Uehara, S., Araki, M., et al. (2014). Ectopic calcification: importance of common nanoparticle scaffolds containing oxidized acidic lipids. Nanomedicine 10, 441–450. doi: 10.1016/j.nano.2013.08.010

Kutikhin, A. G., Brusina, E. B., and Yuzhalin, A. E. (2012). The role of calcifying nanoparticles in biology and medicine. *Int. J. Nanomedicine.* 7, 339–350. doi: 10.2147/IJN.S28069

Lu, H., Guo, Y. N., Liu, S. N., and Zhang, D. C. (2012b). Nanobacteria may be linked to calcification in placenta. *Ultrastruct. Pathol.* 36, 160–165. doi: 10.3109/01913123.2011.653710

Lu, H., Guo, Y. N., Liu, S. N., Zhu, H., and Zhang, D. C. (2012a). Isolation, cultivation and identification of nanobacteria from placental calcification. J. Matern. Fetal. Neonatal. Med. 25, 2182–2185. doi: 10.3109/14767058.2012.684106

Maniscalco, B. S., and Taylor, K. A. (2004). Calcification in coronary artery disease can be reversed by EDTA-tetracycline long-term chemotherapy. *Pathophysiology* 11, 95–101. doi: 10.1016/j.pathophys.2004.06.001

Peng, H. H., Wu, C. Y., Young, D., Martel, J., Young, A., Ojcius, D. M., et al. (2013). Physicochemical and biological properties of biomimetic mineralo-protein nanoparticles formed spontaneously in biological fluids. Small 9, 2297–2307. doi: 10.1002/smll.201202270

Schwartz, M. A., Lieske, J. C., Kumar, V., Farell-Baril, G., and Miller, V. M. (2008). Human-derived nanoparticles and vascular response to injury in rabbit carotid arteries: proof of principle. *Int. J. Nanomedicine* 3, 243–248. doi: 10.2147/IJN.S2473

Shen, X., Ming, A., Li, X., Zhou, Z., and Song, B. (2010). Nanobacteria: a possible etiology for type III prostatitis. J. Urol. 184, 364–369. doi: 10.1016/j.juro.2010.03.003

Shoskes, D. A., Thomas, K. D., and Gomez, E. (2005). Anti-nanobacterial therapy for men with chronic prostatitis/chronic pelvic pain syndrome and prostatic stones: preliminary experience. J. Urol. 173, 474–477. doi: 10.1097/01.ju.0000150062.60633.b2

Wang, L., Shen, W., Wen, J., An, X., Cao, L., and Wang, B. (2006). An animal model of black pigment gallstones caused by nanobacteria. *Dig. Dis. Sci.* 51, 1126–1132. doi: 10.1007/s10620-006-8019-6

Wu, C. Y., Martel, J., Cheng, W. Y., He, C. C., Ojcius, D. M., and Young, J. D. (2013b). Membrane vesicles nucleate mineralo-organic nanoparticles and induce carbonate apatite precipitation in human body fluids. *J. Biol. Chem.* 288, 30571–30584. doi: 10.1074/jbc.M113.492157

Wu, C. Y., Young, L., Young, D., Martel, J., and Young, J. D. (2013a). Bions: a family of biomimetic mineralo-organic complexes derived from biological fluids. PLoS ONE 8:e75501. doi: 10.1371/journal.pone.0075501

Zhang, M. J., Liu, S. N., Xu, G., Guo, Y. N., Fu, J. N., and Zhang, D. C. (2014). Cytotoxicity and apoptosis induced by nanobacteria in human breast cancer cells. *Int. J. Nanomedicine* 9, 265–271. doi: 10.2147/IJN.S54906

Zhou, Z., Hong, L., Shen, X., Rao, X., Jin, X., Lu, G., et al. (2008). Detection of nanobacteria infection in type III prostatitis. *Urology* 71, 1091–1095. doi: 10.1016/j.urology.2008.02.041

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ADVANTAGES OF PUBLISHING IN FRONTIERS

FAST PUBLICATION

Average 90 days
from submission
to publication

**COLLABORATIVE
PEER-REVIEW**

Designed to be rigorous –
yet also collaborative, fair and
constructive

RESEARCH NETWORK

Our network
increases readership
for your article

OPEN ACCESS

Articles are free to read,
for greatest visibility

TRANSPARENT

Editors and reviewers
acknowledged by name
on published articles

GLOBAL SPREAD

Six million monthly
page views worldwide

COPYRIGHT TO AUTHORS

No limit to
article distribution
and re-use

IMPACT METRICS

Advanced metrics
track your
article's impact

SUPPORT

By our Swiss-based
editorial team