

RESEARCH LETTER – Incubator - Scientometrics: Mapping Microbiology

Scientometric analysis of the term 'microbiota' in research publications (1999–2017): a second youth of a century-old concept

Lesya Baudoin^{1,*}, David Sapinho¹, Abdelghani Maddi^{1,2} and Luis Miotti^{1,2}

¹Observatoire des Sciences et Techniques, Hcéres, 2 Rue Albert Einstein, Paris, 75013 France and ²CEPN, UMR-CNRS 723, Université Paris 13, Villetaneuse, 93430 France

*Corresponding author: Observatoire des Sciences et Techniques, Hcéres, 2 Rue Albert Einstein, Paris 75013, France. Tel: +33 0 1 55 55 87 70; E-mail: lesya.baudoin@hceres.fr

One sentence summary: Evolution of the term 'microbiota' in research literature: growth dynamics, field breakdown and topics.

Editor: Manuel Jesús

ABSTRACT

Great progress in microbiota research during last decades resulted in a growing corpus of publications mentioning the term 'microbiota'. Specifically, the human microbiota increasingly recognised nowadays as one of the most important health challenges is becoming an emerging research front. By examining over 28 000 microbiota-related papers from the Web of Science database, our study aims to characterise the evolution of publication patterns in this field between 1999 and 2017. The corpus is first analysed in terms of breakdown by journal subject categories, then an additional insight in the structuring of the microbiota research into different topics is provided by means of topic modelling. Our results demonstrate that over time (i) a substantial increase in the publications number is accompanied by a broad diversification of associated journal subject categories; (ii) the research focus moved outside from its primary research field showing successive shifts from dentistry and ecologically centred areas, through agri-food applied topics, towards the most recent clinical applications. The trends in thematic structure of the field presented from a historical perspective suggest that the current systemic approach to host–microbiota relationship inherited from the ecological background of the concept of microbiota has opened up a number of new research directions and perspectives.

Keywords: microbiota; research trends; scientometrics; scientific publications; growth rate; topic modelling

INTRODUCTION

The term 'microbiota' is commonly used to refer to the association of microorganisms living in and on a multicellular host's body and developing a functionally integrated relationship with a host. Microbial communities inhabiting skin, digestive, urogenital and respiratory tracts of host animals gained increasing research interest over the last decades. Recent advances in understanding their key role in the regulation of the major physiologic processes moved the research focus towards studying the human-dwelling microbial community as a

new therapeutic target for many areas of medicine (Lemon et al. 2012; Kamada et al. 2013; O'Mahony et al. 2015). Though microbiota is often presented as a 'new concept' or even as a 'new organ' in today's science news or in mainstream media, however, a literature search reveals that the term was known for decades, even though used in quite different contexts, and its scope evolved.

The earliest reference to microbiota we have found dates back to 1927 (The Unseen Life of the Soil 1927) and deals with the soil microbiota which comprises 'bacteria, fungi, protozoa [and] lowly plants called algae'. Limnologist W. C. Frohne (Frohne

Received: 15 January 2019; Accepted: 24 June 2019

© The Author(s) 2019. Published by Oxford University Press on behalf of FEMS. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

1939, 1942) characterized the breeding habitats of anopheles mosquitoes as part of preventive actions against the malaria outbreaks and meant by 'microbiota' mainly the unicellular algae and some other microplankton species.

In the 1950s and 1960s, the term entered the regular vocabulary of dental medicine, introducing the notions of oral, gingival and dental plaque microbiota (Crawford and Shankle 1961; Gibbons et al. 1963; Socransky et al. 1963). The easy way to access the human mouth has made the oral microbial community the oldest to be observed (1684, Antoni van Leeuwenhoek) and the earliest studied, with a well-established concept of species diversity. It should be noted that the oral microbiota has been seen mostly as pathogenic.

Research on cellulose digestion by microsymbionts in xylophagous and herbivorous animals also adopted the word 'microbiota' early on. Studies of the rumen microbiota were mainly motivated by the objective to increase the productivity of livestock. As early as the late 19th century, rumen microbiota (formerly called microflora or microfauna) has been described as comprising bacteria, archaea, protozoa, unicellular algae and unicellular fungi. Experimental works in rumen biochemistry carried out in 1940s–1960s led to understanding the key role played by microbiota in plant fibre digestion (Mann, Masson and Oxford 1954; Baldwin, Wood and Emery 1963; Mraz and Cerny 1964). The comprehensive knowledge accumulated by rumen microbiological investigations provided a theoretical basis and enabled the development of anaerobic culture techniques for studying other types of microbial communities (Hungate 1966).

During the same period, research on gut microbiota of laboratory animals has been driven by practical issues of breeding the healthy stock. The pioneering contribution of René Dubos to the development of modern views on the host–microbiota relationships is highly noteworthy. Trained initially as a soil microbiologist, he may have inherited from the soil ecology his systemic way of thinking. Ahead of his time, Dubos came to perceive the digestive tract as an evolving ecosystem as early as in 1964 (Dubos and Schaedler 1964), and to focus research attention on the non-pathogenic microbiota. His findings led him to advocate the importance of symbiotic gastrointestinal microorganisms for the host's resistance to infectious diseases (Dubos et al. 1965). However, it was not before the 2000s that the cellular basis of the relationship between gut microbiota and immunity found supporting evidence (Gallo and Hooper 2012).

The use of the term 'microbiota' in scientific publications remained quite rare until the mid-1990s; but it is over the past two decades that the publications volume has grown rapidly. Despite an increasing research interest in the field, it has rarely been approached in scientometric studies. It is therefore relevant to provide an overall view on the microbiota-related publication landscape, to give an insight onto the way this subject evolved and to understand the maturation process of the concept and of the related research. To identify the thematic structure of large publications set the citation-based approaches are commonly used (Small 1999; Boyack and Klavans 2010). In parallel, scientometric methods based on content analysis, text-mining and visualisation techniques can provide a synthetic picture revealing an intellectual structure of a field (Song and Kim 2013; Heo et al. 2017), to study an emergent research domain (Raimbault, Cointet and Joly 2016), to map an interdisciplinary research field (Cassi et al. 2017), to explore the synergies between basic research and innovation (Qi et al. 2018). Referring to the approaches used in the above bibliometric studies, we present a first description of topic landscape in the research literature using the term 'microbiota'.

The objectives of this paper are (i) to analyse the changes in the use of the term 'microbiota' across different research fields and (ii) to identify the main research themes in the corpus built with this term and to map their dynamics. We have chosen to focus our study on the period between 1999 and 2017 corresponding both to the strongest increase and to the broadest diversification of the use of the term.

DATA AND METHODS

Data

The data has been extracted from the *Observatoire des sciences et Techniques* (OST) in-house database. It includes five indexes of the Web of Science (WoS) available from Clarivate Analytics (SCIE, SSCI, AHCI, CPCI-SSH and CPCI-S) and corresponds to the WoS content indexed through the week of March 26, 2018.

The query containing the word 'microbiota' in the title, abstract or keywords returned 28 489 publications belonging to 147 WoS Journal Subject Categories (JSCs) for the period 1999–2017. Only publications indexed as 'article' or 'review' have been taken into consideration. The complete frequency distribution of papers by JSCs is provided in Annex 1.

Analysis by subject categories

The classification scheme used in our study was that of WoS JSCs. Full counting was applied: a paper was fully attributed to all JSCs assigned to its journal. Some journals are associated to more than one JSC, therefore paper counts per JSC are not additive. Given the large number of JSCs with very few microbiota papers, the analysis has been restricted to the top 30 categories, covering 95% of the total number of publications (presented in Annex 1 with a grey background).

The evolution of term use has been characterized using two measures.

- (i). The growth rate of microbiota publications within each JSC was estimated from a Poisson regression model especially suitable for analysing count data with skewed distributions (Cameron and Trivedi 2013). This model is described as:

$$E[Y/X] = e^{(b_1 X + b_0)}.$$

The dependent variable Y is the number of papers within the JSC and is supposed to be Poisson-distributed. The independent variable X is the year of publication. The 'microbiota growth rate' (MGR) was derived from the estimated slope coefficient (b_1), by applying formula $MGR = e^{b_1} - 1$. The overall growth rate (OGR) of a JSC may affect the growth of microbiota papers in the JSC: one can expect that in a rapidly-growing field the microbiota growth rate is higher than in a slower field. The growth of microbiota papers has therefore been corrected for the overall growth (OGR) of each JSC—that is estimated with the same Poisson model—to highlight the specific growth of microbiota papers in a JSC. The final 'net microbiota growth rate' (NMGR) was computed as:

$$NMGR_{JSC} = \frac{MGR_{JSC} - OGR_{JSC}}{OGR_{JSC} + 1}.$$

- (ii). The point of emergence of 'microbiota' term use in a JSC defined as the year when the number of publications reached 10% of the cumulated output for the studied period.

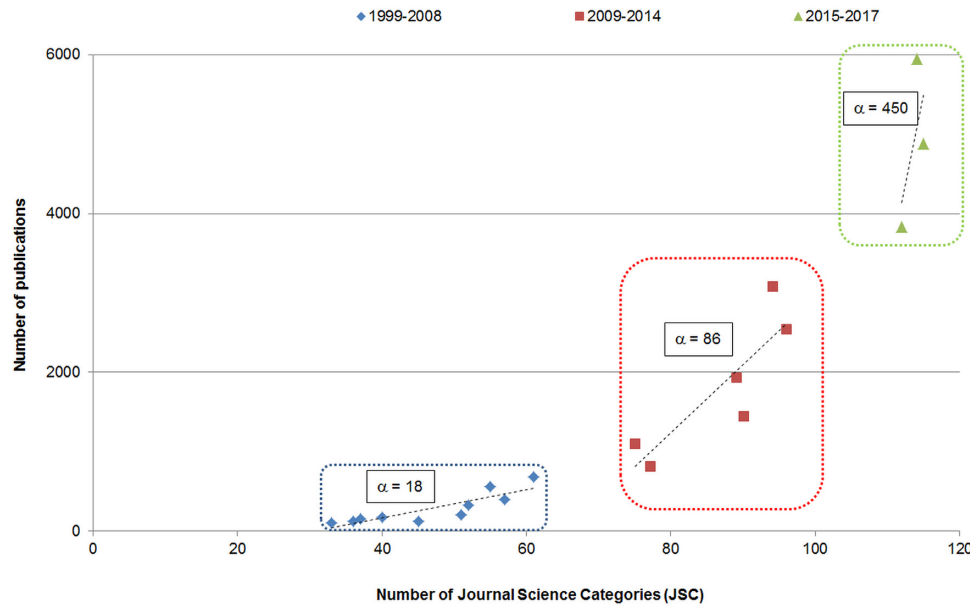


Figure 1. Annual growth of number of papers (y-axis) and associated JSCs (x-axis) of microbiota-related corpus over a 19-year period. Years are represented by points. Successive periods distinguished by different slopes (α) are represented by different colours.

Topic modelling

Topic modelling is as an efficient method to extract a structure of a corpus relying only on textual information. For this analysis, terms were extracted from non-empty Keywords Plus and Author Keywords fields from the overall corpus (28 400 articles–89 dropped out, 60 818 terms). The initial list was subjected to a data preprocessing, in order to reduce the dimensionality of the document–term matrix by conflating the candidate terms, removing infrequent terms and stop-words. The approximate matching was carried out by combining several string matching algorithms: the Levesthein distance, the cosine distance between q and g and the Jaro–Winkler distance. The computations were done with the R ‘stringdist’ package (van der Loo 2014).

To summarize the semantic information, the Latent Dirichlet Allocation (LDA) was applied (Blei, Ng and Jordan 2003). This unsupervised algorithm allows the detection of hidden thematic structures (topics) from a set of selected terms in documents. The method assumes that a probabilistic distribution determines how terms, considered as realisations of latent variables, are randomly attributed to documents (Blei, Ng and Jordan 2003; Blei 2012). Therefore, the observed document–term matrix is approached with the product of document–topic and topic–term matrices, where the first gives the proportion of topics within each document, while the second gives the proportion of terms in each topic.

The LDA model was fitted with the Mallet package (McCallum 2002). R ‘LDavis’ and ‘topicmodels’ packages were used to analyse data, display the results and provide an interactive visualisation of topics (Sievert and Shirley 2014). The proximity in the figure indicates similarity between terms distribution of topics.

The number of topics retained for the analysis was chosen with regard to the perplexity of models provided with different number of topics, a lower perplexity score indicating better generalization performance (Blei, Ng and Jordan 2003). According to this measure, the best compromise was between 20 and 25 topics. The model with 23 topics was finally selected.

The distributions of topics over time were presented in box-plots for each topic, based on the document–topic matrix (proportion of topics within each document) aggregated by year of publication, and weighted by the number of publications in the respective year.

RESULTS

Growth patterns of microbiota publications by journal subject category

We first analyse the distribution of the publications across identified JSCs. Next, we focus on the top 30 selected JSCs.

The annual number of ‘microbiota’ publications increased from 109 in 1999 to 5935 in 2017. The ‘microbiota’ publication set is trans-disciplinary: it spans 147 JSCs in total over the studied period; however, the number of categories concerned varies over time. The number of JSCs dealing with microbiota rises continuously with time ranging from 33 categories in 1999 to 115 in 2017. In order to characterise the use pattern of the term ‘microbiota’ throughout JSCs, the number of microbiota-related papers and the number of corresponding JSCs for each year were plotted in Fig. 1.

A basic trend in the diffusion of the term ‘microbiota’ can be described as broken down into three phases, characterised by different slopes (α), representing the average increase of the number of publications for each new JSC involved. The first phase corresponding to 1999–2008, represents the disciplinary expansion of the term use: the number of JSCs grew while the increase of the number of papers published remained relatively low ($\alpha = 18$). In the second phase, 2009 through 2014, both the number of publications and the number of JSCs increased ($\alpha = 86$). Finally, in the third phase that is still underway, the number of JSCs remained stable while the number of papers continues to rise dramatically.

Regarding the distribution of papers among the top JSCs, the number of publications related to microbiota increased in each of the 30 selected JSCs categories. Net microbiota growth rates

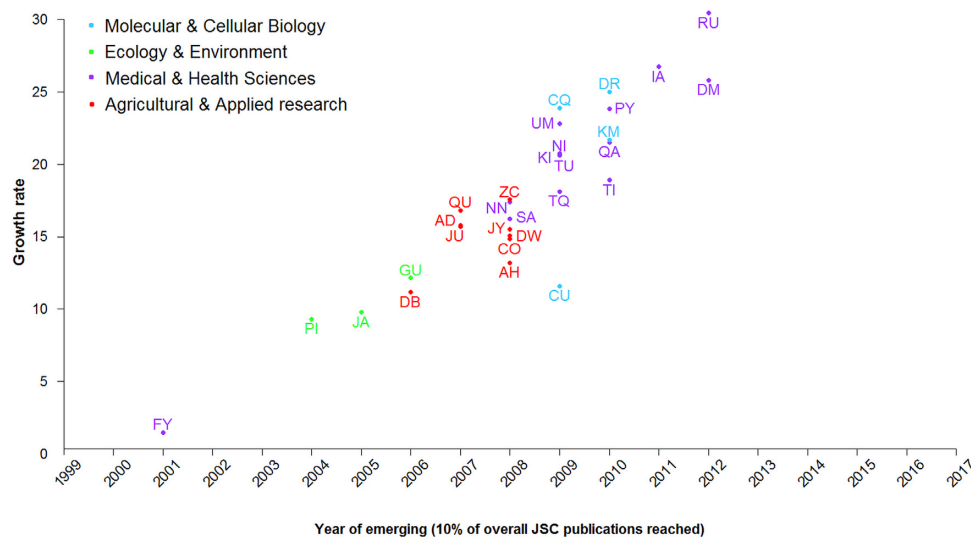


Figure 2. Annual net microbiota growth rates (y-axis) and emergence point (10% of the overall to-date output) for the top 30 JSCs. The large categories are drawn up according to the main study subject, and the colour keys are provided on the left.

(NMGR) were plotted against the emergence points of each JSC (Fig. 2). The scatter plot shows an almost linear relationship between the emergence point and the NMGR: the earlier the subject gets to 10% of its cumulated microbiota output, the slower the number of microbiota-related papers grows within the JSC. The plot presents a variety of subject-specific patterns between two extremities of slowly growing old-timers and newcomers with very high microbiota growth rates. The atypical position of the point Biology (JSC CU) is attributable to re-assigning of JSC for PLoS One (about 1500 microbiota-related papers) from Biology to Multidisciplinary Sciences by Clarivate Analytics in 2012.

From left to right, the succession of JSCs on the plot can be grouped into four larger areas according to their broad disciplinary scope. The left side appears as the old-timers group, with JSCs which were historically among the first 'microbiota' users: Dentistry and three JSCs related to ecology and environment, characterised both by low relative growth rate and reaching the 10% threshold by 2006. This group is followed by the JSCs devoted to agricultural and applied research. Although the scope of use of the term 'microbiota' in this group is close to the old-timers, it shows a relatively higher growth rate. We consider here the Microbiology JSC as belonging to this broad area. The third group encompasses medical JSCs and it spreads till the upper end of the plot. It covers a wide range of growth rates, from medium to the highest, including the latest entrants: Endocrinology & Metabolism, Oncology and Neurosciences. JSCs associated with molecular and cellular biology constitute the fourth group appearing among newcomers, suggesting that microbiota studies enter the technologically driven stage of development.

To sum up, before 2006, the term ‘microbiota’ was mainly used in Dentistry and in JSCs related to ecology and environment. The first health research categories, Nutrition & Dietetics and Infectious Diseases, took on the subject only since 2008, whereas next year several other health categories (Pediatrics, Pharmacology & Pharmacy, Gastroenterology & Hepatology, Immunology, Physiology) were already actively tackling the microbiota-related questions. The shift of research interests from microbiota into clinical fields resulted in important publication growth corresponding to the third phase of growth described above.

Topic analysis

In order to complete the analysis built on the journal-based subject classification, we carried out an analysis based on textual information of publications to identify topics in the corpus.

The model with 23 topics was found to be relevant (i.e. no incoherent topics and all topics are interpretable) and eventually retained. Interpretation and labelling of the topics were done manually, based on the top 30 terms (Annex 2).

Fig. 3 displays the positioning of topics in a 2D space. Regarding the relative position of topics, all clinical topics are situated left or top left on the map. Specifically, the dense aggregation of topics in the left part gathers the pathological conditions associated with the gut microbiota as inflammatory bowel disease, cancer, metabolic syndrome, etc. More loosely situated topics in the top correspond to microbiotas of other locations in the body as vagina, respiratory tract and oral cavity. The group of five topics related to human and animal nutrition is localised at the bottom of the figure. The three topics situated at the right may be defined as unrelated to humans: they comprise food safety, cellulose digestion in ruminants and termites, soil and plants microbiotas. Two topics in intermediate position (3 and 19) focused on laboratory techniques and methods appear as hubs connecting 'human' and 'non-human' topic sets. Similarly, basic research topics (9 and 18) are at the intersection between technology, clinics and ecology.

Fig. 4 represents the weighted distribution of topics over time. It shows a clear succession of topics entering the scene progressively. We can distinguish among the most precursory topics: Oral microbiota (22), Soil and plants microbial communities (11) and Cellulose digestion in ruminants and termites (2). Most of the to-date production of these topics was published before 2007, implying that they pre-existed the current microbiota boom. The middle period is mainly characterised by setting up topics related to nutrition issues and other applied aspects as food technologies and food safety, aquaculture, animal farming, etc. This period is followed by establishing of a new series of topics related to human health. Gut–liver axis (16) becomes a hot topic as late as in 2009; Obesity (17) is the most recent of the topics. It is notable that almost all recently emerged topics (i.e. 1st quartile of volume attained by 2006) are clinically oriented.

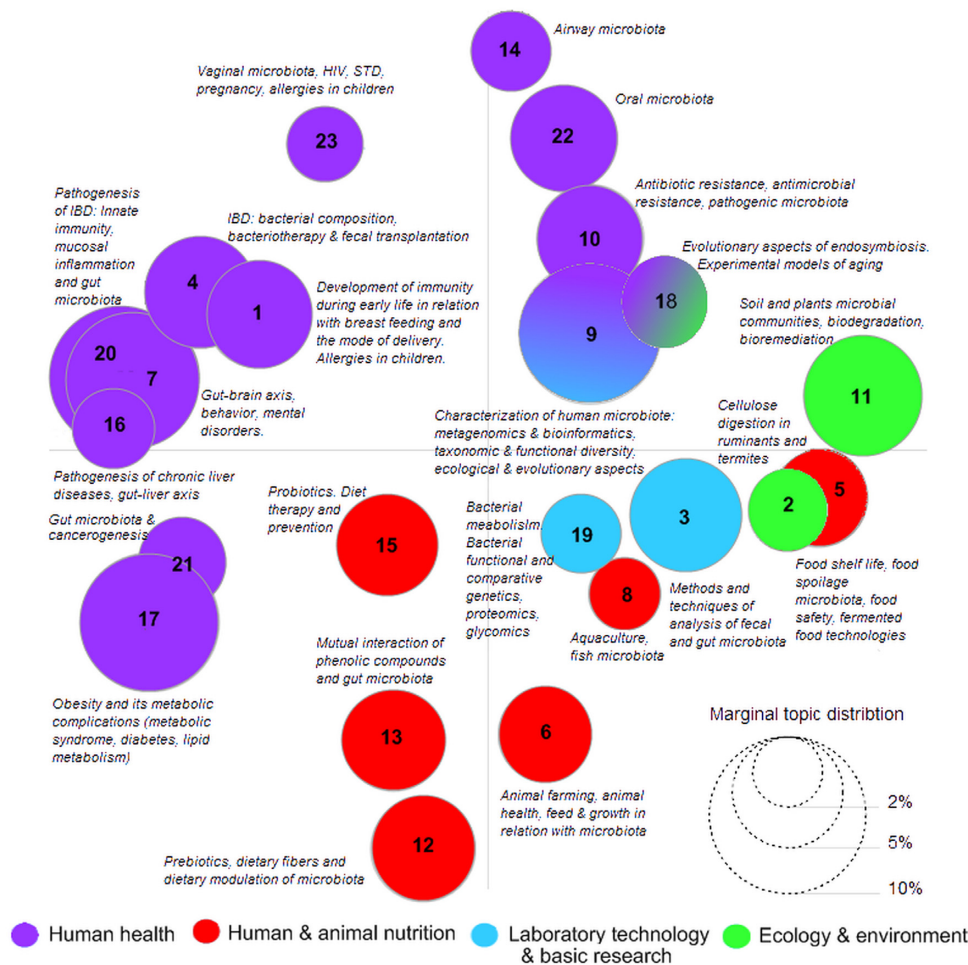


Figure 3. Inter-topic distance map (via multi-dimensional scaling). Topics were interpreted on the basis of top 30 terms (listed in the Annex 2). The large categories are drawn up according to the main study subject. The colour keys are provided below the figure.

It reflects the diversity of new findings showing how microbiota is involved in pathogenesis of different health issues.

These findings are in accordance with the results of the JSC analysis: the topics where the microbiota concept has originated from, as oral microbiota, soil and plants microbiota, ruminants or insects microbiota, progressively lagging behind, while clinical topics take over in the most recent years.

DISCUSSION

In this paper, we explored the field evolution and the topic dynamics of microbiota research using a scientometric approach. As our aim was to provide an overview of the publications on microbiota, we did not compare our findings with those of previous bibliometric studies dealing with specific issues as microbial communities of wastewater (Garrido-Cardenas, Polo-López and Oller-Alberola 2017), gut microbiota under the diabetic condition (Tian et al. 2017), relationship between obesity and microbiota (Yao et al. 2018). These analyses describe the basic bibliographic distributions (countries, institutions, authors, journals, frequent terms) or focus on highly cited papers. Whereas these works contribute to understanding microbiota research, these issues are out of the scope of this paper. It could be a challenge for future studies to further

investigate citation impact, to identify major geographical and institutional players in the field, collaboration patterns, etc.

Our study does not claim to provide a comprehensive overview of research dealing with microbiota. It only examines the changes occurred over time in the use of the term 'microbiota' as well as the semantic shift observed during the recent years. If we were to better cover the larger domain of microbiota-related research, we should have widened the corpus by adding to the query more keywords as 'microflora', 'bacterial flora', 'microfauna', 'microbiome', etc. frequently used in the literature as synonyms for 'microbiota'.

The term 'microbiome' deserves particular attention since nowadays it is often used interchangeably with 'microbiota', though some authors consider that microbiome refers rather to the collective genome of the microbial community. This terminological ambiguity has been pointed out (Huss 2014; Eisen 2015); however, it still persists. The term 'microbiome' is more recent than 'microbiota'. Some uncertainty in its initial definitions (Lederberg 2000, 2001a,b) suggests that at the time, the concept was still in maturation. We add some comments about the origin and the evolution of both terms to shed light on the observed dynamics of categories and topics.

Basically, biota could be interpreted as related to species richness, in terms of the flora and fauna composition, as it was first

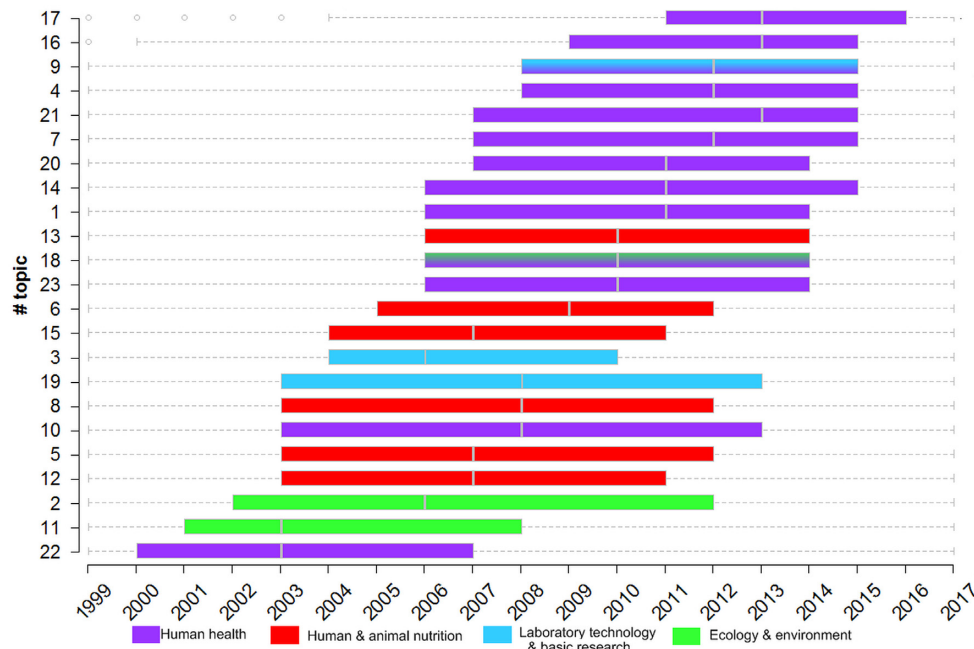


Figure 4. Distribution of topics over the time. Each topic is represented by a box-plot showing quartiles distribution of publications aggregated by year and weighted by the number of publications in the respective year. The large categories (identical to those of the Fig. 3) are drawn up according to the main study subject. The colour keys are provided below the figure.

defined by zoologist Leonhard Stejneger (Stejneger 1901). Initially, microbiota-related studies evolved in line with this primary definition of biota. Whereas in macro-communities the species determination is a rather trivial task built on the extensive past experience, this is not so easy when dealing with the inhabitants of anaerobic environments. The culture techniques developed in 60s has reached its limits since a majority of microbial species turned out to be non-culturable. In 1990s, the breakthrough has been made possible by applying the metagenomic approaches for studying environmental samples mostly from aquatic and soil ecosystems allowing analysis of complex mixtures of species (Riesenfeld, Schloss and Handelsman 2004). Subsequently, the deeper analysis of the human-dwelling microbial communities was prompted not only by the availability of new techniques, but also by the change of paradigm adopting the ecological concepts of biodiversity, equilibrium and host-microbiome–environment interactions. The concept of biome refers rather to ‘an association of diverse, mutually dependent organisms in a natural ecological unit’ (Odum 1945: 194) and implies a particular structure, dynamics and functional relationship. Hence, the functional metagenomics marked a turning point towards the concept of ‘biome’ promoting holistic, integrative approaches to study complex interaction between host, its environment and its symbionts. This conceptual shift is observed in our study through the field and topic evolution. This important shift allowed microbiota research to extend its scope by crossing disciplinary boundaries and to give rise to innovative clinical applications.

Whether or not the publication corpus supplemented with papers using the term ‘microbiome’ shows the different patterns and topics remains to be explored. Representing the microbiota research in a more comprehensive way, using more exhaustive search query, could be a challenge for future studies.

CONCLUSION

Our work provides a first description of the topic landscape of scholarly publications using the term ‘microbiota’ and demonstrates a rising and broadening usage of the term between 1999 and 2017.

The disciplinary structure of microbiota-related publications approached through JSCs as proxies of research fields shows the spread of the term to different fields.

The dynamics of the term use is characterised by three phases. During the emergence phase the microbiota research is still concentrated in several fields where it was historically developed; few papers appear in many new fields indicating a nascent interest. In the course of the expansion phase, the researchers in many fields get hold of the subject and find new applications owing to technological advances. The increased awareness of the microbiota effect on human health attracts numerous funding, in turn promoting the publication of new results. During the third phase, microbiota research bridges technological advances and accumulated experimental evidence, initiating the involvement of numerous new research teams and generating an important research activity.

The disciplinary composition of the corpus changed over time. The fields where the concept of microbiota emerged earlier (‘old-timers’) displaying currently slower growth of microbiota-related papers than the fields which adopted the term more recently (‘newcomers’). These findings suggest that in the oldest fields, the concept of microbiota did not go through any ground-breaking discoveries, and that these fields are currently not in the epicentre of the recent developments. On the contrary, some other fields where the term ‘microbiota’ is relatively new are involved in front-line research and command the major attention of the research community.

The themes identified through topic modelling reveal a dynamic pattern consistent with that of the JSCs. The focus shifted from studying the microbial communities of natural ecosystems to those of human body viewed as a host macro-organism with its associated microbiota. Simultaneously, the term meaning shifted towards an understanding of microbiota as a symbiotic community inhabiting human body and playing an important role in host's metabolism. Our analysis suggests that the ongoing development of the microbiota research arises from the conceptual background built by the community ecology which shaped novel dimensions in understanding the dialogue between the host and its microbiota and opened the way for new interesting basic findings, diagnostics, preventive and therapeutic approaches.

SUPPLEMENTARY DATA

Supplementary data are available at [FEMSLE](https://femsle.onlinelibrary.com/) online.

ACKNOWLEDGEMENTS

The authors would like to thank Frédérique Sachwald for her comments on an earlier version of the paper and Elisabeth de Turckheim for her guidance in improving the manuscript.

Conflict of Interest. None declared.

REFERENCES

- Baldwin R, Wood W, Emery R. Conversion of Glucose-C14 to propionate by rumen microbiota. *J Bacteriol* 1963;**85**:1346–9.
- Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. *J Mach Learn Res* 2003;**3**:993–1022.
- Blei DM. Probabilistic topic models. *Commun ACM* 2012;**55**:77–84.
- Boyack KW, Klavans R. Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *J Am Soc Inf Sci Technol* 2010;**61**:2389–404.
- Cameron AC, Trivedi PK. *Regression Analysis of Count Data*. Cambridge: Cambridge University Press, 2013.
- Cassi L, Lahatte A, Rafols I et al. Improving fitness: Mapping research priorities against societal needs on obesity. *J Informetr* 2017;**11**:1095–113.
- Crawford JJ, Shankle RJ. Application of newer methods to study the importance of root canal and oral microbiota in endodontics. *Oral Surg Oral Med Oral Pathol* 1961;**14**:1109–23.
- Dubos R, Schaedler R. The digestive tract as an ecosystem. *Trans Assoc Am Physicians* 1964;**77**:110–7.
- Dubos R, Schaedler RW, Costello R et al. Indigenous, normal, and autochthonous flora of the gastrointestinal tract. *J Exp Med* 1965;**122**:67–76.
- Eisen J. What does the term microbiome mean? And where did it come from? A bit of a surprise. *The Winnower* 2015, doi: 10.15200/winn.142971.16196.
- Frohne WC. Anopheline breeding: Suggested classification of ponds based on characteristic desmids. *Public Health Rep* 1896–1970 1939;**54**:1363–87.
- Frohne WC. Reconnaissance of anopheline larval habitats and characteristic desmids of the okefenokee swamp, Georgia. *Public Health Rep* 1896–1970 1942;**57**:1209–17.
- Gallo RL, Hooper LV. Epithelial antimicrobial defence of the skin and intestine. *Nat Rev Immunol* 2012;**12**:503–16.
- Garrido-Cardenas JA, Polo-López MI, Oller-Alberola I. Advanced microbial analysis for wastewater quality monitoring: metagenomics trend. *Appl Microbiol Biotechnol* 2017;**101**:7445–58.
- Gibbons RJ, Socransky SS, Sawyer S et al. The microbiota of the gingival crevice area of man. II. The predominant cultivable organisms. *Arch Oral Biol* 1963;**8**:281–9.
- Heo GE, Kang KY, Song M et al. Analyzing the field of bioinformatics with the multi-faceted topic modeling technique. *BMC Bioinform* 2017;**18**:251.
- Hungate RE. *Rumen and Its Microbes*. Academic Press, 1966.
- Huss J. Methodology and ontology in microbiome research. *Biol Theory* 2014;**9**:392–400.
- Kamada N, Chen GY, Inohara N et al. Control of pathogens and pathobionts by the gut microbiota. *Nat Immunol* 2013;**14**:685–90.
- Lederberg J. [Notes from talk at centennial celebration of the Nobel Prize for Medicine and Physiology]. T168 2001a. <https://profiles.nlm.nih.gov/ps/retrieve/ResourceMetadata/BBBBLY> (27 December 2018, date last accessed).
- Lederberg J. [Address at the bloomberg school of public health convocation, class of 2001]. T176 2001b. <https://profiles.nlm.nih.gov/ps/retrieve/ResourceMetadata/BBBBDLZ> (27 December 2018, date last accessed).
- Lederberg J. [Postscript to "Infectious History"]. 2000. <https://profiles.nlm.nih.gov/ps/retrieve/ResourceMetadata/BBBCVQ> (27 December 2018, date last accessed).
- Lemon KP, Armitage GC, Relman DA et al. Microbiota-Targeted therapies: An ecological perspective. *Sci Transl Med* 2012;**4**:137rv5.
- Mann SO, Masson FM, Oxford AE. Effect of feeding aureomycin to calves upon the establishment of their normal rumen microflora and microfauna. *Br J Nutr* 1954;**8**:246–52.
- McCallum AK. . *MAchine Learning for LanguagE Toolkit*. 2002, <http://mallet.cs.umass.edu/>.
- Mraz O, Cerny L. Functional composition + main metabolites of rumen microflora in cattle with definite food intake. *Folia Microbiol (Praha)* 1964;**9**:194–.
- O'Mahony SM, Clarke G, Borre YE et al. Serotonin, tryptophan metabolism and the brain-gut-microbiome axis. *Behav Brain Res* 2015;**277**:32–48.
- Odum EP. The concept of the biome as applied to the distribution of north american birds. *Wilson Bull* 1945;**57**:191–201.
- Qi Y, Zhu N, Zhai Y et al. The mutually beneficial relationship of patents and scientific literature: topic evolution in nanoscience. *Scientometrics* 2018;**115**:893–911.
- Raimbault B, Cointet J-P, Joly P-B. Mapping the emergence of synthetic biology. *PLOS ONE* 2016;**11**:e0161522. <http://mallet.cs.umass.edu/>
- Riesenfeld CS, Schloss PD, Handelsman J. Metagenomics: Genomic analysis of microbial communities.. *Annu Rev Genet* 2004;**38**:525–52.
- Sievert C, Shirley K. LDavis: A method for visualizing and interpreting topics. *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*. Baltimore, Maryland, USA: Association for Computational Linguistics, 2014, 63–70.
- Small H. Visualizing science by citation mapping. *J Assoc Inf Sci Technol* 1999;**50**:799.
- Socransky SS, Gibbons RJ, Dale AC et al. The microbiota of the gingival crevice area of man. I. Total microscopic and viable counts and counts of specific organisms. *Arch Oral Biol* 1963;**8**:275–80.
- Song M, Kim SY. Detecting the knowledge structure of bioinformatics by mining full-text collections. *Scientometrics* 2013;**96**:183–201.

- Stejneger L. Scharff's history of the european fauna. *Am Nat* 1901;35:87–116.
- The unseen life of the soil. *Science* 1927;65:x–xiv.
- Tian J, Li M, Lian F et al. The hundred most-cited publications in microbiota of diabetes research: A bibliometric analysis. *Medicine (Baltimore)* 2017;96:e7338.
- van der Loo MPJ. The stringdist package for approximate string matching. *R J* 2014;6:111–22.
- Yao H, Wan J-Y, Wang C-Z et al. Bibliometric analysis of research on the role of intestinal microbiota in obesity. *PeerJ* 2018;6:e5091.