


# Integrating microbiome, transcriptome and metabolome data to investigate gastric disease pathogenesis: a concise review

Dalla Doohan<sup>1</sup>, Yudith Annisa Ayu Rezkitha<sup>2</sup>, Langgeng Agung Waskito<sup>1</sup>,  
Ratha-korn Vilaichone<sup>3</sup>, Yoshio Yamaoka<sup>4,5</sup> and Muhammad Miftahussurur<sup>5,6</sup> 

## Review

**Cite this article:** Doohan D, Rezkitha YAA, Waskito LA, Vilaichone R-K, Yamaoka Y, Miftahussurur M (2021). Integrating microbiome, transcriptome and metabolome data to investigate gastric disease pathogenesis: a concise review. *Expert Reviews in Molecular Medicine* **23**, e9, 1–8. <https://doi.org/10.1017/erm.2021.8>

Received: 22 February 2021

Revised: 7 June 2021

Accepted: 14 June 2021

### Key words:

Disease burden; gastric disease; metabolome; microbiome; next-generation sequencing; transcriptome

### Authors for correspondence:

Muhammad Miftahussurur,  
E-mail: [muhammad-m@fk.unair.ac.id](mailto:muhammad-m@fk.unair.ac.id);  
Yoshio Yamaoka,  
E-mail: [yamaoka@oita-u.ac.jp](mailto:yamaoka@oita-u.ac.jp)

<sup>1</sup>Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia; <sup>2</sup>Faculty of Medicine, University of Muhammadiyah Surabaya, Surabaya, Indonesia; <sup>3</sup>Gastroenterology Unit, Department of Medicine, Thammasat University Hospital, Pathum Thani, Thailand; <sup>4</sup>Department of Environmental and Preventive Medicine, Faculty of Medicine, Oita University, Yufu, Japan; <sup>5</sup>Division of Gastroentero-Hepatology, Department of Internal Medicine, Faculty of Medicine, Dr Soetomo Teaching Hospital, Universitas Airlangga, Surabaya, Indonesia and <sup>6</sup>Institute of Tropical Disease, Universitas Airlangga, Surabaya, Indonesia

## Abstract

Microbiome, the study of microbial communities in specific environments, has developed significantly since the Human Microbiome Project began. Microbiomes have been associated with changes within environmental niches and the development of various diseases. The development of high-throughput technology such as next-generation sequencing has also allowed us to perform transcriptome studies, which provide accurate functional profiling data. Metabolome studies, which analyse the metabolites found in the environment, are the most direct environmental condition indicator. Although each dataset provides valuable information on its own, the integration of multiple datasets provides a deeper understanding of the relationship between the host, agent and environment. Therefore, network analysis using multiple datasets might give a clearer understanding of disease pathogenesis.

## Introduction

Currently, the human body is viewed not only as a complex organism, but also as an environment within which countless microorganisms live. The number of microorganism genes in the human body is approximately 150 times the number of human genes (Ref. 1). These microorganisms live and interact with the human body in different symbiotic relationships, for example, mutualism, commensalism or even parasitism (Ref. 2). These interactions may be associated with a disturbance in normal physiology. Therefore, microbiome studies may provide a better understanding of the development of human diseases.

The human stomach was once thought to be a 'sterile' environment because of its extremely high acidity (Refs 3, 4). It was then thought to be the most important gastrointestinal barrier to various pathogenic microorganisms. However, this theory was debunked when *Helicobacter pylori* was successfully isolated from the human stomach by Barry Marshal and Robin Warren in the early 1980s (Refs 5, 6). This breakthrough suggested that other microorganisms might be able to survive and colonise the human stomach. However, past efforts to profile stomach microbiota were hampered by the limitations of conventional culture, histology and immunohistochemistry. The culture method required a huge effort to isolate and identify bacteria and could not characterise most of them. Therefore, it most likely underestimated the actual number of microbiota in the stomach.

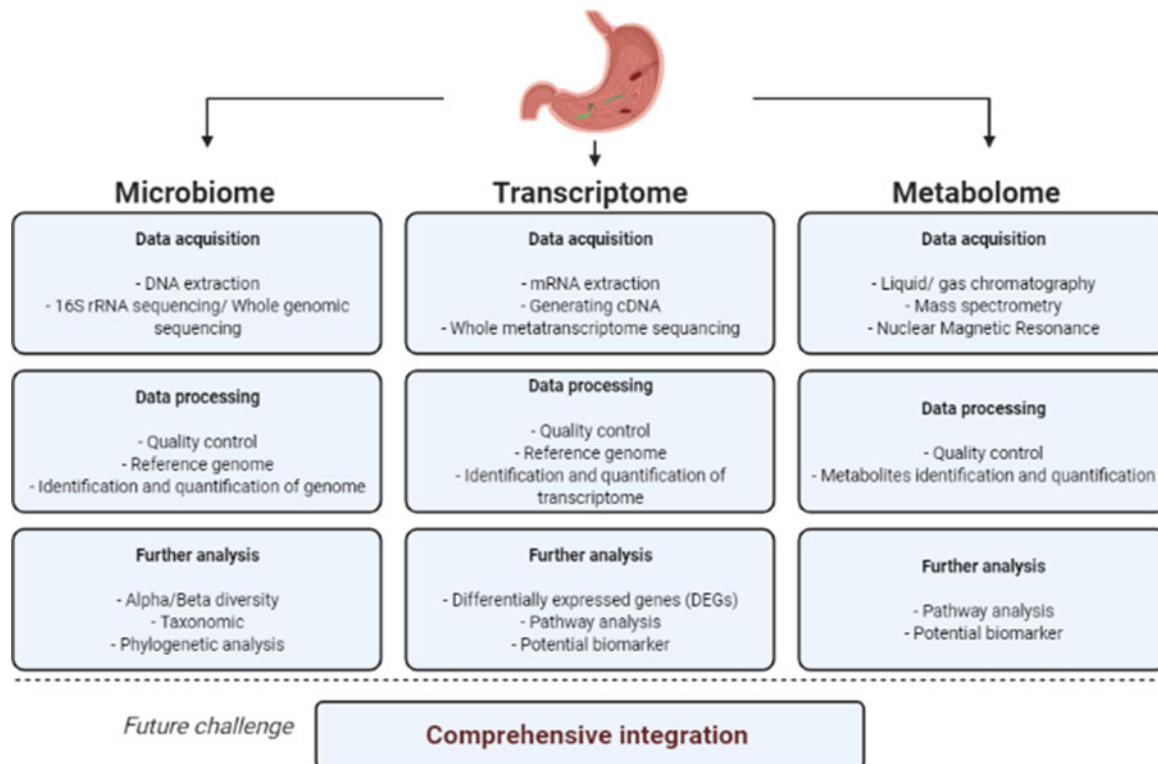
The development of sequencing technology has shed new light on the gastric microbiota. The application of next-generation sequencing (NGS) is a huge leap that has enabled a shift to a culture-independent approach and focused on sequence analysis to identify stomach microorganisms. The microbiome and transcriptome are then used to characterise the microbiota and viable microorganisms in the stomach.

Another type of study, the metabolome study, is currently used to support microbiome and transcriptome studies. A metabolome study is a comprehensive analysis of metabolites, where the molecules released by the organisms into the environment are identified and quantified. The metabolome is considered the most direct indicator of alterations in the environment.

In this review, we aim to summarise the current updates on microbiota, microbiome, transcriptome and metabolome studies of the human stomach. We also review the importance of integrating multiple datasets to get a better understanding of the development of gastric diseases.

## Next-generation sequencing

The stomach is unique within the human gastrointestinal tract because of its harsh environment with extremely high acidity, ranging between a pH of 1 and 3 in healthy subjects (Refs 7, 8). Therefore, it is extremely difficult for microorganisms to successfully colonise it.



**Fig. 1.** Schematic pipeline for microbiome, transcriptome and metabolome study.

Most viable microorganisms are eliminated in the stomach, resulting in significantly lower microbial loads in the stomach ( $10^1$ – $10^3$  microbes/g) compared with that in the small intestine ( $10^4$ – $10^7$  microbes/g) or colon ( $10^{11}$ – $10^{12}$  microbes/g) (Refs 9–11). This is one of the causes of the difficulty in profiling the microorganisms within the human stomach. In the past, very few methods could be used for gastric microbiota profiling. A conventional method such as the culture of gastric biopsy and gastric juice can reveal viable microorganisms in the specimens (Refs 12–14). Histology and immunohistochemistry are also frequently used for profiling (Ref. 12). However, these methods are costly, require extensive effort, take a long time to achieve results and have a low sensitivity. Therefore, a more reliable method to provide detailed data about gastric microbiota is needed. NGS has been used extensively for comprehensive profiling of microorganisms ever since the Human Microbiome Project began (Ref. 15). NGS is a ‘culture-independent’ method that can provide a deep and high throughput of DNA sequences. It results in gastric microbiome data, which then can be used to determine the microbiota. Microbiome, in this case, refers to the collection of genomes from all microorganisms in a specific environment, whereas microbiota refers to the specific microorganisms. The pipeline for a microbiome study is shown in Figure 1.

Currently, NGS is widely used for profiling microorganisms in a specific environment, such as the stomach, gut, oral cavity or skin. There are several NGS methods; however, 16S ribosomal RNA (16S rRNA) gene sequencing and shotgun metagenomic sequencing are the two most used. The 16S rRNA sequencing method uses 16S rRNA as a marker gene to be amplified. 16S rRNA is a highly conserved gene with hyper-variable regions within it; it is unique and can be used to distinguish between species and microbiome strains (Refs 16, 17). The hyper-variable regions within the 16S rRNA gene, such as V1–V9, are the amplification target in microbiome studies. A combination of several hyper-variable regions is also commonly used, such as the combination of V1–V3, V2–V3, V3–V4 and V5–V6 (Refs 18–21).

The sequencing reads are then filtered based on the number of expected errors and the length of the reads to obtain high-quality clean tags. Chimeric sequences can be detected using specific software such as UCHIME (Ref. 22). Chimeric sequences are usually removed to obtain the effective tag sequences and to avoid mis-identification. Next, sequences with high similarity are assigned the same cluster name operational taxonomic unit (OTU). Most studies used a 97–99% similarity as the basis of determining the OTUs (Refs 18, 19, 23). However, 16S amplicon sequencing also has limitations. First, it may underestimate the diversity in the community given the biases associated with the polymerase chain reaction amplification method. Second, amplicon sequencing can generate wide estimations of diversity. Third, this method cannot provide data related to biological functions encoded in the genome (Refs 24–26).

On the contrary, shotgun metagenomic sequencing targets all genomic DNA available in the given sample. The library preparation and workflow are similar to other whole-genome sequencing methods. It enables us to provide a better species- (Ref. 27) and strain-level classification of microorganisms (Ref. 28). However, despite the advantages, this method has several challenges, such as relatively complex data leading to complicated bioinformatics analyses, the presence of a lot of host DNA and contaminant genomes.

Various analyses can be conducted from the OTU data. Diversity analysis is one of the basic analyses for microbiome studies. In general, we recognise two types of diversity, alpha-diversity and beta-diversity. Alpha-diversity is used to analyse the diversity of the samples’ richness and evenness through several parameters: observed species, Chao1 index, Shannon index, Simpson index and Good’s coverage (Refs 29, 30). Beta-diversity is used to analyse the species complexity of the samples. Diversity analysis is commonly performed using QIIME software (Ref. 31). Principal coordinate analysis is another method used to obtain the principal coordinates and visualise complex data (Ref. 32).

A study by Johnson *et al.* highlighted the potential of the high-throughput sequencing of the entire 16S rRNA gene to provide better taxonomic resolution than is achieved by only targeting the hyper-variable regions with a short-read sequencing platform. They performed an in-silico comparison of full and partial 16S rRNA sequencing that showed the importance of utilising the full 16S rRNA gene to accurately classify the microorganisms at high taxonomic resolution (Ref. 33).

NGS can also provide high-throughput transcriptome sequencing or RNA sequencing data. It enables us to detect, quantify and perform deep analyses of the well-known transcripts and novel, unknown transcripts. Although microarrays are also a powerful method to study the transcriptome, RNA sequencing (RNASeq) using NGS offers several advantages. For example, RNASeq can be performed without prior knowledge of the genome sequences, therefore, without reference genomes. RNASeq can also be used to directly measure all RNA transcripts, whereas microarrays are indirect, measuring fluorescence after hybridisation using a probe sequence. Furthermore, RNASeq can be used to detect novel transcripts, and, finally, RNASeq data provide less noise compared with microarrays, in which cross-hybridisation may increase background signals (Ref. 34).

The types of sequencers used for NGS should also be considered. Because of the huge number of NGS platforms currently available, choosing the most suitable sequencer is important. This will depend mostly on the study design and the project goal. Other considerations are the run time, read length, number of reads per run, maximum output and price.

### Overview of gastric microbiota and factors affecting the composition

The successful identification of *Campylobacter pyloridis* by Barry Marshall and Robin Warren in the early 1980s was a breakthrough in medicine. *Campylobacter pyloridis*, renamed *Helicobacter pylori* in 1984, was shown to successfully colonise the human stomach, which debunked the old 'sterile stomach' theory. This discovery suggested the possibility of human stomach colonisation by other pathogenic microorganisms. However, the comprehensive characterisation of gastric microbiota was still not carried out because of the limitations of conventional methods such as culture, histology or immunohistochemistry. The development of NGS methods has enabled better assessment of the composition of stomach microbiota. In addition, most of the early gastric microbiota studies focused on bacteria rather than other microorganisms such as fungi or parasites.

Because of the anatomical location of the stomach, the microbiota detected in the stomach might include microbiota from the oral cavity or duodenum. For example, there is an abundance of *Lactobacillus* and *Veillonella*, previously described as oral microbiota (Refs 35, 36). However, these are transient bacteria, and their importance in developing gastric diseases is still not clear.

Bik *et al.* characterised the microbial diversity within the human stomach using the 16S rDNA clone library. They identified 128 phylotypes and assigned them into five major phyla: *Proteobacteria*, *Firmicutes*, *Actinobacteria*, *Bacteroidetes* and *Fusobacteria*. Not surprisingly, they found that 67% of phylotypes in the study have been described as phyla commonly found in the oral cavity. They also demonstrated an abundance of transient bacteria (Ref. 37). Importantly, the same study showed that *H. pylori* was the most abundant species in the stomach of *H. pylori*-infected subjects. A study conducted by Nam *et al.* on 20 Korean patients reported that gut microbiota differed between subjects from different geographical regions. They proposed that microbiota might be affected by host and environmental factors, such as genetic variations and diet (Ref. 38). Five dominant

phyla were characterised from the stomachs of Korean subjects: *Actinobacteria*, *Firmicutes*, *Bacteroidetes*, *Fusobacteria* and *Proteobacteria*. Another study by Ferreira *et al.* also supports these results; the gastric microbiota in 135 Portuguese was dominated by five phyla: *Proteobacteria* (69.3%), *Firmicutes* (14.7%), *Bacteroidetes* (9.0%), *Actinobacteria* (4.3%) and *Fusobacteria* (1.3%). Therefore, the studies showed no differences in phyla between subjects from different geographical areas, and the differences might be more pronounced at the level of genus, species or subspecies. Delgado *et al.* examined specimens from 12 healthy subjects from different populations. They found that *Streptococcus*, *Propionibacterium* and *Lactobacillus* were the most abundant genera, and the microbiota was similar at the phylum level (Ref. 39). The gastric microbiota is more dynamic at the genus level, and various factors might affect this level. Gastric microbiota also reportedly changed with the development of gastric diseases, including *H. pylori* infection. Studies have investigated the relationship between microbiota composition and the development of various gastric diseases (Table 1).

The human stomach is a complex organ with various mechanisms to maintain a healthy state. Therefore, the exact mechanisms determining the gastric microbiota composition are not clearly understood. However, factors such as dietary habits, and the use of antibiotics, anti-inflammatory drugs, probiotics and proton-pump inhibitors or H<sub>2</sub>-antagonists, are thought to play important roles. The extensive and long-term use of acid secretion inhibitors such as proton pump inhibitors (PPIs) and H<sub>2</sub>-antagonists is said to be the most important factor influencing the microbiota composition (Ref. 23). Another proposed mechanism influencing the gastric microbial composition is that PPIs may directly target the bacterial and fungi proton-pump mechanisms of organisms such as *H. pylori*, *Clostridium difficile*, *Candida albicans* and *Saccharomyces cerevisiae* (Ref. 40). A high pH level may facilitate bacterial overgrowth in the stomach (Refs 40, 41). Subsequently, the change of gastric microbiota composition will also affect the gut microbiota. A study by Imhann *et al.* revealed that the use of PPI significantly alters gut microbiota, reducing its microbial diversity (Ref. 42). Furthermore, bacterial composition in the PPI user group was clustered differently compared with the non-PPI user group (Ref. 42).

In animal models, dietary habit was reported to be associated with the disturbance of normal gastric microbiota (Ref. 43). Moreover, a long-term high-fat diet was associated with gastric dysbiosis. However, the mechanism leading to dysbiosis in the mouse model might also be caused by the metabolic syndrome, which comprises obesity, hyperlipidaemia and insulin resistance. Probiotics contain live microbial components that have a beneficial effect on human health (Ref. 44). They have, therefore, been proposed as an alternative or supplemental therapy for gastrointestinal diseases. For example, Igarashi *et al.* showed the potential of LG21 probiotic to restore the gastric microbiota composition of patients with functional dyspepsia (Ref. 45). The patients had an alteration of gastric microbiota compared with healthy subjects. In the functional dyspepsia group, the level of *Bacteroidetes* was higher than that of *Proteobacteria*, with absent *Acidobacteria*. In contrast, in healthy subjects, the *Proteobacteria* level was higher than that of *Bacteroidetes*, and *Acidobacteria* were present. The LG21 probiotic supplementation was able to return the gastric microbiota of dyspeptic patients to a composition similar to that of healthy subjects, significantly decreasing the *Bacteroidetes/Proteobacteria* ratio and increasing *Acidobacteria* levels. In addition, the abundance of *Lactobacillus*, which is the genus of the LG21 strain was not increased after 12 weeks therapy (Ref. 45), showing that the huge amount of LG21 administered, caused a minimum of side effects and no *Lactobacillus* overgrowth in the stomach.

**Table 1.** Studies investigating the predominant microbiota in gastric diseases

Country	Published year	Total samples	Predominant OTUs in disease			Reference
			<i>H. pylori</i> gastritis	<i>H. pylori</i> atrophic gastritis	Gastric cancer	
China	2017	207	–	–	<i>Peptostreptococcus stomatis</i> , <i>Streptococcus anginosus</i> , <i>Parvimonas micra</i> , <i>Slackia exigua</i> and <i>Dialister pneumosintes</i>	Coker <i>et al.</i> (Ref. 68)
United Kingdom	2017	95	<i>Helicobacteraceae</i>	<i>Helicobacteraceae</i> , <i>Streptococcaceae</i> , <i>Fusobacteriaceae</i> and <i>Prevotellaceae</i>	–	Parsons <i>et al.</i> (Ref. 69)
Malaysia	2017	36	–	–	<i>Lactococcus</i> , <i>Veillonella</i> and <i>Fusobacteriaceae</i>	Castaño-Rodríguez <i>et al.</i> (Ref. 70)
China	2018	276	–	–	<i>Prevotella melaninogenica</i> , <i>Streptococcus anginosus</i> and <i>Propionibacterium acnes</i>	Liu <i>et al.</i> (Ref. 71)
Mongolia	2019	75	<i>Helicobacter</i> genus in <i>H. pylori</i> -positive gastritis	–	–	Gantuya <i>et al.</i> (Ref. 72)
China	2018	62	–	–	<i>Peptostreptococcus</i> , <i>Streptococcus</i> and <i>Fusobacterium</i>	Chen <i>et al.</i> (Ref. 73)
Mongolia	2020	168	<i>H. pylori</i>	<i>H. pylori</i>	<i>Lactobacilli</i> and <i>Enterococci</i>	Gantuya <i>et al.</i> (Ref. 74)
China	2020	132	–	–	Phyla Actinobacteria, Bacteriodes, Firmicutes and Fusobacteria	Wang <i>et al.</i> (Ref. 75)

### Overview of gastric transcriptome studies

The metagenomic approach is a powerful method for profiling the gastric microbiota. It is a culture-independent method for characterising microorganisms within the human stomach. It is considered a breakthrough from conventional culture-based or histological approaches, which could only characterise approximately 30% of gastric microbiota. However, metagenomics can only provide information on the ‘presence’ or ‘absence’ of microorganisms; therefore, another approach is needed to shed light on the functional profile of the gastric microbial community. The value of transcriptome studies in examining viable bacteria and microbial gene expression in the human stomach has been shown by Thorell *et al.* (Ref. 46). Their results showed that *H. pylori* is the predominant microbiota both in infected individuals and in most uninfected individuals. Conventional methods such as urea breath test, serology and culture confirmed this. *H. pylori* abundance is positively correlated with the presence of *Campylobacter*, *Deinococcus* and *Sulphurospirillum*. Importantly, the results also showed that the expression of *H. pylori* genes involved in pH regulation and nickel transport was high (Ref. 46).

Transcriptome studies use a different approach than metagenome studies, focusing on expressed genes in the specific environment (Ref. 47), either from the microbial community or the host. Transcriptome studies analyse gene expression by capturing the mRNA found in a specific environment (Refs 48, 49), thus determining the viable microbiota. The pipeline for transcriptome studies is shown in Figure 1.

As the mRNA is extracted from biopsy or stool specimens, it may contain microbial, and host RNA; differentiating between the two has proven challenging. A suitable transcriptome reference database should be used to provide accurate data on the expressed genes either from the host or the microbes.

Differentiating between microbial and host RNA may help provide sequencing data with less background noise caused by the abundant host RNA. Currently, commercial kits used to separate and isolate bacterial RNA from mixed samples are widely available. Several studies have demonstrated the use of bacterial RNA enhancement kits to clean up samples by removing non-bacterial RNA (Refs 50–52).

It is important to note that transcriptome studies can be performed from either the microbial or host perspective. The biological significance depends on the design and aims of the study. For example, where a study aims to examine the expression of genes responsible for antibiotic resistance mechanisms, the microbial transcriptome may provide more significant data. On the contrary, when the study aims to examine the effect of infection on the host, the host transcriptome may give deeper information on the disease pathogenesis. The researcher should consider the best approach to fulfil the study goal.

### Microbiome, transcriptome and gastric diseases

The success of *H. pylori* in colonising the stomach demonstrates the important role of gastric microbiota in gastroduodenal disease pathogenesis. Other microbiota might be able to survive in the extreme gastric niche environment and be related to disease pathogenesis. Therefore, it is important to examine the association between gastric microbiota and various gastric diseases. The microbiome modification during gastric diseases is shown in Table 2.

#### Gastric cancer

Gastric cancer is the fifth most common malignancy globally (age-standardised rates 11.1 per 100 000 in 2018; Global Cancer Observatory). The incidence of gastric cancer shows wide

**Table 2.** Studies investigating microbiota modification in subjects with diseases

Country	Year	Total samples	Microbiota details	Reference
USA	2005	23 (12 <i>H. pylori</i> -positive and 9 <i>H. pylori</i> -negative)	<i>H. pylori</i> was the most abundant phylotype within the libraries. <i>H. pylori</i> sequences were found in 19 of the 23 gastric samples	Bik <i>et al.</i> (Ref. 37)
China	2009	45 (23 healthy and 22 gastritis)	Significantly higher <i>Firmicutes</i> phylum and <i>Streptococcus</i> genus in antral gastritis patients	Li <i>et al.</i> (Ref. 61)
Sweden	2009	15 (10 gastric cancer and 5 healthy controls)	Gastric cancer patients had a low abundance of <i>H. pylori</i> and the genera <i>Streptococcus</i> , <i>Lactobacillus</i> , <i>Veillonella</i> and <i>Prevotella</i> dominated	Dicksved <i>et al.</i> (Ref. 76)
Venezuela	2010	12 (8 <i>H. pylori</i> -positive and 4 <i>H. pylori</i> -negative)	Higher abundance of Proteobacteria and Acidobacteria in <i>H. pylori</i> -infected and higher abundance of Actinobacteria and <i>Firmicutes</i> in <i>H. pylori</i> -negative	Maldonado-Contreras <i>et al.</i> (Ref. 77)
China	2017	207 (with superficial gastritis, atrophic gastritis, intestinal metaplasia and gastric cancer)	The OTUs of <i>Parvimonas micra</i> , <i>Dialister pneumosintes</i> , <i>Slackia exigua</i> and <i>Peptostreptococcus stomatis</i> were significantly higher in gastric cancer microbiome compared with other precancerous stages	Coker <i>et al.</i> (Ref. 68)
Portugal	2017	135 (54 gastric carcinoma and 81 gastritis)	Reduced diversity in gastric cancer patients. Decreased <i>Helicobacter</i> abundance and enrichment of other intestinal microbials	Ferreira <i>et al.</i> (Ref. 19)
South Korea	2019	556 (268 gastric cancer and 288 controls)	Higher abundance of <i>H. pylori</i> and <i>Propionibacterium acnes</i> in gastric cancer patients. Significantly higher risk for gastric cancer in patients with higher <i>H. pylori</i> and <i>P. acnes</i> abundance	Gunathilake <i>et al.</i> (Ref. 78)
Mongolia	2020	168 (48 gastric cancer, 20 gastritis, 40 atrophy, 40 intestinal metaplasia and 20 healthy controls)	Control group had the highest alpha-diversity compared with cancer and intestinal metaplasia groups. <i>Lactobacilli</i> and <i>Enterococci</i> were predominant in <i>H. pylori</i> -negative gastric cancer patients	Gantuya <i>et al.</i> (Ref. 74)
Indonesia	2020	137 (27 <i>H. pylori</i> -positive and 110 <i>H. pylori</i> -negative)	Significantly lower diversity index in <i>H. pylori</i> -positive patients compared with <i>H. pylori</i> -negative	Miftahussurur <i>et al.</i> (Ref. 18)

variation worldwide, with a notably high prevalence in East Asian countries such as Japan, Korea, Mongolia and China (Refs 53, 54). *H. pylori* is the most successful colonising pathogen in the human stomach. It was classified as a class I carcinogen in 1994 by the International Agency for Research on Cancer. *H. pylori* is one of the most important risk factors for gastric cancer. Several studies have reported a strong association between gastric cancer development and *H. pylori* infection (Refs 55, 56). Gastric cancer develops and progresses slowly over years, starting from gastric inflammation/gastritis, and progressing through gastric atrophy, intestinal metaplasia, finally developing into gastric cancer. *H. pylori* plays an important role in inducing gastritis. Moreover, because of *H. pylori* being able to persistently colonise gastric epithelial cells, chronic inflammation continues in the stomach.

The eradication of *H. pylori* is crucial for reducing gastric cancer risk. Lee *et al.* showed that *H. pylori* eradication prevents gastritis progression into gastric cancer in the INS-GAS mice model (Ref. 57). *H. pylori* infection accelerates the development of pre-malignant lesions, intestinal metaplasia and severe dysplasia, whereas the administration of antibiotics and eradication therapy as early as 8 weeks post-infection prevents dysplasia, atrophy and intestinal metaplasia (Ref. 57). Studies highlighting the significant role of gastric microbiota in gastric disease pathogenesis include another study by Lofgren *et al.* (Ref. 58). The absence of gastric colonisation with microbiota in germ-free INS-GAS mice significantly delayed gastric cancer development. In addition, *H. pylori* alone is capable of promoting gastritis and gastrointestinal intraepithelial neoplasia in INS-GAS mice. Furthermore, *H. pylori*-infected INS-GAS mice with complex microbiota develop neoplasia faster compared with *H. pylori*-mono-association INS-GAS mice with the absence of gastric microbiota.

Transcriptome studies can be used to analyse the host genes that might be responsible for gastric cancer. Zhang *et al.* used transcriptome analysis to identify the genes and pathways involved in gastric adenocarcinoma pathogenesis (Ref. 59). They found that there were 1477 upregulated and 282 downregulated genes in the gastric adenocarcinoma group, compared with the normal controls. Moreover, the functional enrichment analysis and clustering analysis also showed that the upregulated differentially expressed genes (DEGs) were significantly associated with cell adhesion molecule binding, serine hydrolase activity and several inflammation and tumour pathways, such as the p53 pathway, tight junction pathway, apoptosis pathway and tumour necrosis factor signalling pathway. In addition, the relationship of the expression of genes such as *RASGRP3* and *CTHRC1* to prognosis suggests their potential use as prognostic markers.

Transcriptome studies can also be used to investigate novel therapies for gastric cancer by targeting genes associated with cancer development. For example, Ren *et al.* found that the GPNMB gene is highly expressed in gastric cancer patients and indicates a worse prognosis. Knockdown of this gene in the gastric cell lines, such as AGS cell and NCI-N87, inhibits the proliferation and migration of cancer cells. Moreover, GPNMB could affect the expression of coinhibitory molecules in the cancer cell and be involved in the escape of cancer cells from the host immune system. Therefore, the GPNMB gene is a promising target for gastric cancer immunotherapy (Ref. 60).

### Gastritis

Li *et al.* examined the microbiota composition of healthy subjects and non-*H. pylori* gastritis patients (Ref. 61). By using 16S rRNA sequencing, they found that *Firmicutes* was the most abundant phylum and *Streptococcus* the most abundant genus in patients with antral gastritis, compared with healthy controls. In addition, there was an under-representation of Proteobacteria in antral

gastritis patients. *Streptococcus*, *Prevotella*, *Porphyromonas*, *Neisseria* and *Haemophilus* were the most common genera in antral gastritis patients (Ref. 61).

Nookaew *et al.*, using the microarray method, examined the transcriptomic changes in corpus atrophic gastritis patients facilitated by *H. pylori* infection (Ref. 62). They found a significant depletion of DEGs in both antrum and corpus of corpus atrophic gastritis patients. Interestingly, they observed the antralisation process in the corpus atrophy patients, characterised by increased gastrin expression and downregulation of several corpus-specific genes. They also revealed that in corpus atrophy patients, the acidic mammalian chitinase (*AMCase*) gene had the most significant reduction of expression because of the high expression of this gene in normal corpus mucosa (Ref. 62).

### Metabolome studies complement metagenome and transcriptome studies

Metabolomics is the comprehensive analysis of metabolites from the specimen. The small molecules released by the organisms into the environment are identified and quantified to describe alterations in the environment. The metabolome is considered the most direct indicator of a healthy or altered environment's condition. Variations in the production of signature metabolites are related to changes in the activity of metabolic routes; therefore, metabolomics is useful for pathway analysis (Ref. 63). In addition, metabolomics shows promise for use in drug discovery and pharmacogenomics (Refs 64, 65).

Metabolome data are obtained by a very different method compared with microbiome and transcriptome data. The latter data are obtained through nucleotide sequencing, whereas metabolome data are obtained by mass spectrometry, chromatography and nuclear magnetic resonance. Mass spectrometry has been widely used to measure metabolites with high sensitivity. Chromatography is used to separate more complex mixtures of metabolites (Ref. 66). These methods can be combined as liquid chromatography-mass spectrometry, gas spectrometry-mass spectrometry or capillary electrophoresis-mass spectrometry (CE-MS) to provide reliable data with high precision (Ref. 66).

A systematic review by Huang *et al.* analysed 52 molecular epidemiologic metabolomics studies of human upper gastrointestinal cancers (Ref. 67). Metabolomic studies have facilitated effective biomarker detection in gastric cancer, supporting the potential of applying metabolomic profiling in cancer prevention and disease management. Although several metabolites have been identified for gastric cancer, the identification of putative metabolomic biomarkers has remained inadequate. Application of metabolomic profiling to molecular epidemiologic studies on gastric cancers may provide insights into the biological significance of crucial metabolites and metabolic pathways, but there is no information on the underlying mechanisms. Given the multi-stage progression of gastric carcinogenesis, metabolic biomarkers associated with precancerous and early gastric cancers must be identified to improve screening and early diagnosis in high-risk populations.

Considering the importance of metabolome studies, it is logical to integrate them with microbiome and transcriptome studies. This addition could provide insights into the outcome of changes in gene expressions, which may lead to differential expression of specific metabolites that impact the health of the host environment. Understanding the whole ecosystem will open exciting approaches for generating new knowledge.

### Conclusion and future insights

In future, the integration of multiple datasets obtained from microbiome, transcriptome and metabolome studies will provide

solutions to some current challenges. The application of individual methods might not be enough to provide accurate data to support hypotheses because of their limitations. The integration of different studies will allow the researchers to build and test models of microbial activity and inter-microbe or microbe–host interactions. This will enable a better understanding of the association between the environment and the microbial community. For example, the combination of metagenomics and metatranscriptomics may reveal overexpression or underexpression of particular functions and, in some cases, the activities of specific organisms. However, network analysis is crucial for analysing the combination of microbiota, expressed genes and metabolites data obtained with each method to draw accurate conclusions. Only by performing complex network analysis between microbiota, transcriptome and metabolome data we can shed new light on the association between the host, agent and environment.

**Acknowledgements.** This study was funded by the Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan to YY. LAW and DD were PhD students supported by the Japanese government (MEXT) scholarship program for 2015 and 2016, respectively.

**Conflict of interest.** The authors declare no competing interests.

### References

1. Qin J *et al.* (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65.
2. Eloe-Fadrosh EA and Rasko DA (2013) The human microbiome: from symbiosis to pathogenesis. *Annual Review of Medicine* **64**, 145–163.
3. Nardone G and Compare D (2015) The human gastric microbiota: is it time to rethink the pathogenesis of stomach diseases? *United European Gastroenterology Journal* **3**, 255–260.
4. Petra CV, Rus A and Dumitraşcu DL (2017) Gastric microbiota: tracing the culprit. *Clujul Medical (1957)* **90**, 369–376.
5. Warren J and Marshall B (1983) Unidentified curved bacilli on gastric epithelium in active chronic gastritis. *Lancet (London, England)* **1**(8336), 1273–1275.
6. Marshall B and Warren J (1984) Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet (London, England)* **1**(8390), 1311–1315.
7. Ayazi S *et al.* (2009) Measurement of gastric pH in ambulatory esophageal pH monitoring. *Surgical Endoscopy* **23**, 1968–1973.
8. McLauchlan G *et al.* (1989) Comparison of gastric body and antral pH: a 24 hour ambulatory study in healthy volunteers. *Gut* **30**, 573–578.
9. Berg RD (1996) The indigenous gastrointestinal microflora. *Trends in Microbiology* **4**, 430–435.
10. Sekirov I *et al.* (2010) Gut microbiota in health and disease. *Physiological Reviews* **90**, 859–904.
11. O'Hara AM and Shanahan F (2006) The gut flora as a forgotten organ. *EMBO Reports* **7**, 688–693.
12. Sanduleanu S *et al.* (2001) Non-*Helicobacter pylori* bacterial flora during acid-suppressive therapy: differential findings in gastric juice and gastric mucosa. *Alimentary Pharmacology & Therapeutics* **15**, 379–388.
13. Ryan KA *et al.* (2008) Isolation of lactobacilli with probiotic properties from the human stomach. *Letters in Applied Microbiology* **47**, 269–274.
14. Khosravi Y *et al.* (2014) Culturable bacterial microbiota of the stomach of *Helicobacter pylori* positive and negative gastric disease patients. *The Scientific World Journal* **2014**, 610421.
15. Peterson J *et al.* (2009) The NIH human microbiome project. *Genome Research* **19**, 2317–2323.
16. Osman MA *et al.* (2018) 16S rRNA gene sequencing for deciphering the colorectal cancer gut microbiome: current protocols and workflows. *Frontiers in Microbiology* **9**, 767.
17. Bharti R and Grimm DG (2021) Current challenges and best-practice protocols for microbiome analysis. *Briefings in Bioinformatics* **22**, 178–193.
18. Miftahussurur M *et al.* (2020) Gastric microbiota and *Helicobacter pylori* in Indonesian population. *Helicobacter* **25**, e12695.
19. Ferreira RM *et al.* (2018) Gastric microbial community profiling reveals a dysbiotic cancer-associated microbiota. *Gut* **67**, 226–236.

20. **Leiva-Gea I et al.** (2018) Gut microbiota differs in composition and functionality between children with type 1 diabetes and MODY2 and healthy control subjects: a case-control study. *Diabetes Care* **41**, 2385–2395.
21. **Jo HJ et al.** (2016) Analysis of gastric microbiota by pyrosequencing: minor role of bacteria other than *Helicobacter pylori* in the gastric carcinogenesis. *Helicobacter* **21**, 364–374.
22. **Edgar RC et al.** (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics (Oxford, England)* **27**, 2194–2200.
23. **Jackson MA et al.** (2016) Proton pump inhibitors alter the composition of the gut microbiota. *Gut* **65**, 749–756.
24. **Sharpton TJ** (2014) An introduction to the analysis of shotgun metagenomic data. *Frontiers in Plant Science* **5**, 209.
25. **Hong S et al.** (2009) Polymerase chain reaction primers miss half of rRNA microbial diversity. *The ISME Journal* **3**, 1365–1373.
26. **Schloss PD** (2010) The effects of alignment quality, distance calculation method, sequence filtering, and region on the analysis of 16S rRNA gene-based studies. *PLoS Computational Biology* **6**, e1000844.
27. **Jovel J et al.** (2016) Characterization of the gut microbiome using 16S or shotgun metagenomics. *Frontiers in Microbiology* **7**, 459.
28. **Rausch P et al.** (2019) Comparative analysis of amplicon and metagenomic sequencing methods reveals key features in the evolution of animal metaorganisms. *Microbiome* **7**, 133.
29. **Willis AD** (2019) Rarefaction, alpha diversity, and statistics. *Frontiers in Microbiology* **10**, 2407.
30. **Prehn-Kristensen A et al.** (2018) Reduced microbiome alpha diversity in young patients with ADHD. *PLoS One* **13**, e0200728.
31. **Navas-Molina JA et al.** (2013) Advancing our understanding of the human microbiome using QIIME. *Methods in Enzymology* **531**, 371–444.
32. **Goodrich JK et al.** (2014) Conducting a microbiome study. *Cell* **158**, 250–262.
33. **Johnson JS et al.** (2019) Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications* **10**, 5029.
34. **Rao MS et al.** (2018) Comparison of RNA-Seq and microarray gene expression platforms for the toxicogenomic evaluation of liver from short-term Rat toxicity studies. *Frontiers in Genetics* **9**, 636.
35. **Burcham ZM et al.** (2020) Patterns of oral microbiota diversity in adults and children: a crowdsourced population study. *Scientific Reports* **10**, 2133.
36. **Caufield PW et al.** (2015) Oral lactobacilli and dental caries: a model for niche adaptation in humans. *Journal of Dental Research* **94**, 110s–118s.
37. **Bik EM et al.** (2006) Molecular analysis of the bacterial microbiota in the human stomach. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 732–737.
38. **Nam YD et al.** (2011) Comparative analysis of Korean human gut microbiota by barcoded pyrosequencing. *PLoS One* **6**, e22109.
39. **Delgado S et al.** (2013) Microbiological survey of the human gastric ecosystem using culturing and pyrosequencing methods. *Microbial Ecology* **65**, 763–772.
40. **Vesper BJ et al.** (2009) The effect of proton pump inhibitors on the human microbiota. *Current Drug Metabolism* **10**, 84–89.
41. **Theisen J et al.** (2000) Suppression of gastric acid secretion in patients with gastroesophageal reflux disease results in gastric bacterial overgrowth and deconjugation of bile acids. *Journal of Gastrointestinal Surgery* **4**, 50–54.
42. **Imhann F et al.** (2016) Proton pump inhibitors affect the gut microbiome. *Gut* **65**, 740–748.
43. **He C et al.** (2018) High-fat diet induces dysbiosis of gastric microbiota prior to gut microbiota in association with metabolic disorders in mice. *Frontiers in Microbiology* **9**, 639.
44. **Kechagia M et al.** (2013) Health benefits of probiotics: a review. *ISRN Nutrition* **2013**, 481651.
45. **Igarashi M et al.** (2017) Alteration in the gastric microbiota and its restoration by probiotics in patients with functional dyspepsia. *BMJ Open Gastroenterology* **4**, e000144.
46. **Thorell K et al.** (2017) In vivo analysis of the viable microbiota and *Helicobacter pylori* transcriptome in gastric infection and early stages of carcinogenesis. *Infection and Immunity* **85**(10), 1–15.
47. **Aguiar-Pulido V et al.** (2016) Metagenomics, metatranscriptomics, and metabolomics approaches for microbiome analysis. *Evolutionary Bioinformatics Online* **12**, 5–16.
48. **Wolf JB** (2013) Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial. *Molecular Ecology Resources* **13**, 559–572.
49. **Wang Z, Gerstein M and Snyder M** (2009) RNA-seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* **10**, 57–63.
50. **Kamminga T et al.** (2020) Combined transcriptome sequencing of mycoplasma hypopneumoniae and infected pig lung tissue reveals up-regulation of bacterial F1-like ATPase and down-regulation of the P102 cilium adhesin in vivo. *Frontiers in Microbiology* **11**, 1679.
51. **Robbe-Saule M et al.** (2017) An optimized method for extracting bacterial RNA from mouse skin tissue colonized by *Mycobacterium ulcerans*. *Frontiers in Microbiology* **8**, 512.
52. **Jorth P et al.** (2013) Probing bacterial metabolism during infection using high-resolution transcriptomics. *Journal of Bacteriology* **195**, 4991–4998.
53. **Rawla P and Barsouk A** (2019) Epidemiology of gastric cancer: global trends, risk factors and prevention. *Przegląd Gastroenterologiczny* **14**, 26–38.
54. **GBD 2017 Stomach Cancer Collaborators** (2020) The global, regional, and national burden of stomach cancer in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease study 2017. *The Lancet Gastroenterology and Hepatology* **5**, 42–54.
55. **Uemura N et al.** (2001) *Helicobacter pylori* infection and the development of gastric cancer. *New England Journal of Medicine* **345**, 784–789.
56. **Kumar S et al.** (2020) Risk factors and incidence of gastric cancer after detection of *Helicobacter pylori* infection: a large cohort study. *Gastroenterology* **158**, 527–536.e527.
57. **Lee CW et al.** (2008) *Helicobacter pylori* eradication prevents progression of gastric cancer in hypergastrinemic INS-GAS mice. *Cancer Research* **68**, 3540–3548.
58. **Lofgren JL et al.** (2011) Lack of commensal flora in *Helicobacter pylori*-infected INS-GAS mice reduces gastritis and delays intraepithelial neoplasia. *Gastroenterology* **140**, 210–220.
59. **Zhang W et al.** (2018) Transcriptome sequencing identifies key pathways and genes involved in gastric adenocarcinoma. *Molecular Medicine Reports* **18**, 3673–3682.
60. **Ren F et al.** (2020) Transcriptome analysis reveals GPNMB as a potential therapeutic target for gastric cancer. *Journal of Cellular Physiology* **235**, 2738–2752.
61. **Li XX et al.** (2009) Bacterial microbiota profiling in gastritis without *Helicobacter pylori* infection or non-steroidal anti-inflammatory drug use. *PLoS One* **4**, e7985.
62. **Nookaew I et al.** (2013) Transcriptome signatures in *Helicobacter pylori*-infected mucosa identifies acidic mammalian chitinase loss as a corpus atrophy marker. *BMC Medical Genomics* **6**, 41.
63. **Gu J et al.** (2016) Metabolomic analysis reveals altered metabolic pathways in a rat model of gastric carcinogenesis. *Oncotarget* **7**, 60053–60073.
64. **Wishart DS** (2016) Emerging applications of metabolomics in drug discovery and precision medicine. *Nature Reviews. Drug Discovery* **15**, 473–484.
65. **Tuyiringire N et al.** (2018) Application of metabolomics to drug discovery and understanding the mechanisms of action of medicinal plants with anti-tuberculosis activity. *Clinical and Translational Medicine* **7**, 29.
66. **Pan Z and Raftery D** (2007) Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics. *Analytical and Bioanalytical Chemistry* **387**, 525–527.
67. **Huang S et al.** (2020) A systematic review of metabolomic profiling of gastric cancer and esophageal cancer. *Cancer Biology & Medicine* **17**, 181–198.
68. **Coker OO et al.** (2018) Mucosal microbiome dysbiosis in gastric carcinogenesis. *Gut* **67**, 1024–1032.
69. **Parsons BN et al.** (2017) Comparison of the human gastric microbiota in hypochlorhydric states arising as a result of *Helicobacter pylori*-induced atrophic gastritis, autoimmune atrophic gastritis and proton pump inhibitor use. *PLoS Pathogens* **13**, e1006653.
70. **Castaño-Rodríguez N et al.** (2017) Dysbiosis of the microbiome in gastric carcinogenesis. *Scientific Reports* **7**, 15957.
71. **Liu X et al.** (2019) Alterations of gastric mucosal microbiota across different stomach microhabitats in a cohort of 276 patients with gastric cancer. *EBioMedicine* **40**, 336–348.
72. **Gantuya B et al.** (2019) Gastric microbiota in *Helicobacter pylori*-negative and -positive gastritis among high incidence of gastric cancer area. *Cancers (Basel)* **11**(4), 1–12.
73. **Chen XH et al.** (2019) Mucosa-associated microbiota in gastric cancer tissues compared with non-cancer tissues. *Frontiers in Microbiology* **10**, 1261.
74. **Gantuya B et al.** (2020) Gastric mucosal microbiota in a Mongolian population with gastric cancer and precursor conditions. *Alimentary Pharmacology & Therapeutics* **51**, 770–780.

75. Wang Z *et al.* (2020) Changes of the gastric mucosal microbiome associated with histological stages of gastric carcinogenesis. *Frontiers in Microbiology* **11**, 997.
76. Dicksved J *et al.* (2009) Molecular characterization of the stomach microbiota in patients with gastric cancer and in controls. *Journal of Medical Microbiology* **58**, 509–516.
77. Maldonado-Contreras A *et al.* (2011) Structure of the human gastric bacterial community in relation to *Helicobacter pylori* status. *The ISME Journal* **5**, 574–579.
78. Gunathilake MN *et al.* (2019) Association between the relative abundance of gastric microbiota and the risk of gastric cancer: a case-control study. *Scientific Reports* **9**, 13589.