How Causal are Microbiomes? A Comparison with the *Helicobacter pylori* Explanation of Ulcers

Kate E. Lynch [a,b]
[a] Department of Philosophy and Charles Perkins Centre, University of Sydney, Australia
[b] Department of Biological Sciences, Macquarie University, Australia
*Corresponding author: kateemmalynch@gmail.com

Emily C. Parke
Philosophy, School of Humanities, University of Auckland, New Zealand

Maureen A. O'Malley [a,b]
[a] LaBRI, University of Bordeaux
[b] School of History and Philosophy of Science, University of Sydney, Australia

Abstract

Human microbiome research makes causal connections between entire microbial communities and a wide array of traits that range from physiological diseases to psychological states. To evaluate these causal claims, we first examine a well-known single-microbe causal explanation: of *Helicobacter pylori* causing ulcers. This apparently straightforward causal explanation is not so simple, however. It does not achieve a key explanatory standard in microbiology, of Koch's postulates, which rely on manipulations of single-microorganism cultures to infer causal relationships to disease. When Koch's postulates are framed by an interventionist causal framework, it is clearer what the *H. pylori* explanation achieves and where its explanatory strengths lie. After assessing this 'simple', single-microbe case, we apply the interventionist framework to two key areas of microbiome research, in which obesity and mental health states are purportedly explained by microbiomes. Despite the experimental data available, interventionist criteria for explanation show that many of the causal claims generated by microbiome research are weak or misleading. We focus on the stability, specificity and proportionality of proposed microbiome causal explanations, and evaluate how effectively these dimensions of causal explanation are achieved in some promising avenues of research. We suggest some conceptual and explanatory strategies to improve how causal claims about microbiomes are made.

Keywords: microbiome; causal explanation; *Helicobacter pylori*; Koch's postulates, interventionism

Acknowledgements

Introduction

Microbiome research is a new molecular field of study that examines microbial communities ('microbiota', aka 'microbiomes') in host organisms. The field is particularly well-known for its insights into the human gut microbiome and its alleged impact on a huge variety of human health and disease states. As ever more associations are made between microbiomes and the states of their hosts, numerous researchers and commentators have proclaimed that a 'microbiome revolution' is revising everything we know about humans and the biological world around and inside us (e.g., Blaser 2014; Mayer et al. 2014; Rees et al. 2018). Such sentiments have spread far beyond microbiology and other scientific endeavours to popular and social media, where microbiomes are touted as controlling and optimizing all aspects of physical wellbeing (including sporting prowess), mental health, and human evolution.[1] More recently, microbiome research has begun to attract philosophical attention (e.g., Huss 2014; O'Malley and Skillings 2018). A philosophical theme of central importance to this research is how health and disease states are explained by microbiomes, and how those explanatory claims are evaluated.

The main activity in human microbiome research today continues to be identifying bioinformatic associations between microbiome compositions and numerous diseases. To strengthen discoveries of apparently meaningful associations, researchers carry out whole-community experiments. These manipulations are often framed as showing that *entire microbial communities* are a major causal factor in the production as well as treatment of diseases and disorders that range from obesity to anxiety. 'Emergent properties' of whole microbiomes are sometimes argued to explain these negative and positive effects (e.g., Hall et al. 2018). Particularly when whole-microbiome experiments are used to investigate mental health states and general brain function, their apparent success has led to claims about the microbiome being a locus of control for human behaviour and psychology. From mental illness to social preferences, the microbiome is apparently exerting 'microbial mind control' (Sampson and Mazmanian 2015:565). If this were truly the case, it would raise major questions about personal identity and free will. For instance, who or what is in charge of our emotions, behaviours and abilities? Does our microbiome define our 'self' (Rees et al. 2018; Parke et al. 2018)?

Tantalizing as such claims and questions are, we believe that before even considering them, far more fundamental issues need to be addressed. Our aim is to examine whether existing experimental evidence and methodology in microbiome research support basic claims about microbiome causality in human hosts. We argue that in many cases, they do not. Our focus is on the structure and dimensions of causal explanations in microbiome research, and whether they are good causal explanations. We examine candidate explanations by drawing on two traditional methods for establishing causation in microbiology and experimental science more generally: Koch's postulates and interventionist views of causation.

We first discuss a classic example of causal explanation in microbiology: *Helicobacter pylori* and ulcers. A clear contrast between this case and contemporary microbiome explanations is that the latter refer to large communities of hundreds of microbial species, while traditional microbiological explanations focus on single species. A 'simple' textbook case of microbial causation is a good starting point, before taking on the complexities of microbiome explanations.[2] As we will see, even this textbook case is far from straightforward.

---

[1] E.g., https://www.viome.com/services/; https://www.the-scientist.com/notebook/athletes-microbiomes-differ-from-nonathletes-31428; https://kellybroganmd.com/depression-starts-in-your-gut/; https://www.marsvenus.com/blog/john-gray/staying-happy-with-a-healthy-microbiome (accessed October 2018).
[2] We are grateful to John Matthewson for suggesting this comparison.

A textbook case of microbial causation: *Helicobacter pylori* as the cause of peptic ulcers

The discovery of *H. pylori* is a well-known exemplar of how a novel but strong causal explanation can overthrow scientific orthodoxy. This story begins in the 1980s with observations of an unknown microorganism in the stomachs of patients with ulcers. A narrative highpoint occurs when one of the key experimenters drinks a culture of *H. pylori* and rapidly experiences stomach problems. After managing to cure himself, he and his colleague go on to convince the medical community of an indisputable causal link between the bacterium and ulcers, due to the disease being cured by antibiotics. Eventually, both researchers are awarded the Nobel Prize for their findings. Their research is a classic success story for establishing causal explanations in microbiology (Blaser 1996; Malfertheimer et al. 2014).

However, when this narrative is put under the microscope, a much more complicated story of causality emerges. Standard methods used to establish causality in microbiology are not achieved fully. Asymptomatic infections of *H. pylori* are more common than ulcer-producing infections, because background factors play a role in determining whether infections result in ulcers or not. The causal pathways and mechanisms of ulcer production are also contested. Nevertheless, the key intervention of wiping out *H. pylori* with antibiotics is taken by many scientists, clinicians and physicians as the ultimate justification for causal claims about ulcers.

The role of bacteria in ulcer production had been suggested as early as the late nineteenth century but was largely dismissed in the middle of the twentieth century (Kidd and Modlin 1998). Bacteria were believed to be unable to inhabit the harsh acidic environment of the stomach. Instead, excess acid secretion in the stomach was considered the primary cause of ulcer formation. This explanation and associated interventions were not entirely satisfactory, because the primary treatment, acid reduction drugs, at best controlled ulcers rather than healing them.

The dominant explanation was overthrown by the research efforts of a pair of Australian scientists, Robin Warren and Barry Marshall, the latter of whom famously experimented on himself. He drank a culture of *H. pylori* and developed gastritis (Marshall et al. 1985), an inflammatory condition of the stomach lining that can lead to ulcer development. Medical and scientific literature recognize this 'auto-experiment', and the associated body of research results, as discovering the principal cause of peptic ulcers and rejecting acid production as the primary intervention target (e.g., Ford and Talley 2009; Malfertheiner et al. 2014). Yet, even in the heyday of this explanation's acceptance, there was dissent about whether *H. pylori* was conclusively shown to cause peptic ulcers (e.g., van Zanten and Sherman 1994).

Evaluating Warren and Marshall's claims requires some background on explanation in microbiology. From the late nineteenth century until recently, the gold standard for identifying microbial causes required the use of Koch's postulates (named for microbiologist Robert Koch who played a key role in their formulation). These are procedures that establish whether or not a microorganism has caused a disease. Although there are some historical vagaries in their exact wording, commonly they are presented in three steps (see Gillies 2016 for a four-step formulation, and Carter 2003 for five):

1. The microorganism is detected in all disease cases;
2. The microorganism does not occur in other diseases or non-pathogenically;
3. After being fully isolated and repeatedly grown in pure culture, the microorganism can induce the disease by being introduced into a healthy animal host.

Experimentally, the key step is the third one, and as we will see, it is central to understanding causation and what constitutes an adequate experimental intervention. In the *H. pylori* story this postulate is not strictly fulfilled, with Marshall himself acting as the only animal host to which the microorganism was introduced (and he did not develop ulcers). Even Koch, when confronted with difficulties in carrying out a laboratory-based experiment in the case of cholera, drew on a 'natural experiment' of a cholera epidemic to substitute for his third postulate (Ross and Woodward 2016; Gillies 2016). This 'natural experiment' served as an experimental proxy, in which the absence of the microorganisms from the disease context resulted in a lack of observable effect. In short, the reasoning is that 'if the cause is removed, the effect is taken away' (Gillies 2016:12). This interpretation is also invoked in the *H. pylori* case.

What did Warren and Marshall do? They took biopsies from patient stomachs and cultured the only organism apparently growing there. This was a novel organism, a slightly spiral bacterium with a few flagella at one end, first thought to be a *Campylobacter*, then later given a new genus name, *Helicobacter* (Goodwin et al. 1989). Patient symptoms were then correlated with bacterial presence. 100% of duodenal ulcer patients and 75% of gastric ulcer patients had *H. pylori*. More generally, the bacterium was associated with gastritis (Marshall and Warren 1984; Marshall 2005).

Even though the researchers knew 'cause-and-effect cannot be proved in a study of this kind' (Marshall and Warren 1984:1314), Marshall, like Koch, thought he could 'extrapolate backwards'. He postulated that if he could eliminate the bacteria, he would also cure duodenal ulcers and perhaps gastric ones (Marshall 2005:262). Otherwise he feared that *H. pylori* would be seen as only spuriously associated with ulcers and gastritis. Various antibiotics and antibacterials were tried out on culture plates and in patients. The results were good. Cultures on plates were severely inhibited in the treatment zone. Four patients treated with a combination of antibiotics and an antibacterial were cured without recurrence (single antibiotic interventions were and are not successful). Eliminating *H. pylori* in this limited patient sample appeared to be a far more effective long-term treatment for ulcers than conventional stomach acid inhibitors (Marshall 2005).

Early papers on this new bacterium and its role in gastric disease invoked Koch's postulates to make causal claims about the association (e.g., Marshall et al. 1985). But at several points subsequently, even as he accepted the Nobel Prize in 2005 for his role in the research, Marshall admitted that three decades more work on *H. pylori* had not led to Koch's postulates being fulfilled (Marshall 1995; 2005). Part of the problem was the lack of a non-human animal model. Despite colonizing other animal species, *Helicobacter* did not initially produce ulcers or even gastritis in non-humans.[3] For ethical reasons, the infection and ulcers especially cannot be established in healthy humans, which would be required to fulfil Postulate Three without a non-human animal model.

That is why Marshall decided to self-experiment. But crucially, he did not go on to develop an ulcer. He developed gastritis. Gastritis can lead to ulcer formation, but not all gastritis leads to ulcers, nor even most gastritis. On top of that, his gastritis symptoms all abated rapidly, prior to any treatment intervention (Marshall et al. 1985).

Marshall himself noted later that the complexities of ulcer formation might be too hard to capture in any animal model. 'In any case, the pressing need for such an effort may have

---

[3] Although some animals—especially gerbils—now model gastric symptoms concomitant with *H. pylori* infections, they are argued to be 'inadequate' models because *H. pylori* alone cannot produce human-like symptoms (e.g., Chen et al. 2007).

passed', he suggested, once there were straightforward diagnostic tools and *H. pylori* therapies (1995:568). In other words, he concluded that Koch's postulates are useful when there is uncertainty, but once something works as a treatment, the pressure is off to establish causality in the standard way.

Causal explanation and interventionist dimensions

The *H. pylori* case illustrates that there is more to establishing good causal explanations in microbiology than straightforward reliance on Koch's postulates. Although this methodological framework has had many successes in the history of microbiology, it is likely that each success story will have ambiguities and complications that require additional strategies for examining causal claims. Recently, an interventionist account of causation has taken hold in philosophy of science, particularly the philosophy of biology (Woodward 2010). On a broad-stroke view of this account, cause-effect relationships are demonstrated by showing that interventions on the putative cause produce changes in the effect (Woodward 2003). Experimental manipulations are a paradigm case of interventions. The interventionist understanding of causation is precisely what researchers are looking for when they attempt to establish causal explanations in an experimental setting, such as the case of ulcer formation.

Applying this account to the ulcer case illustrates why the *H. pylori* explanation supplanted the received view of acid as a primary cause. Intervening on *H. pylori* using antibiotics and antibacterials successfully cures ulcers. This is in contrast to intervening exclusively on acid production, which at best controls the symptoms of ulcers. Interventionism is therefore a good framework for understanding why *H. pylori* provided a more successful treatment outcome. It is important to note, however, that acid production is causally influenced by *H. pylori,* and so they are not independent causal factors. Moreover, the causal chain linking them both to ulcer formation is complex, and in the case of duodenal ulcers is still contested (see Ford and Talley 2009; Hobsley et al. 2009).

In addition to providing criteria for causation in line with scientific practice (Woodward 2003), interventionism provides tools for evaluating the strength of causal explanations. The framework identifies dimensions that match intuitive scientific ideas about 'good' or 'strong' causal explanations (Box 1). These dimensions can also be applied to what we have learned about *H. pylori* explanations of ulcers.

<div style="border:1px solid">

Box 1: Interventionist dimensions of causal explanation

Specificity (1): The extent to which a cause has fine-grained control over an effect. Variations in the states of a cause will produce corresponding variations in particular states of its effect (Woodward 2010; Griffiths et al. 2015; Calcott 2017).

Specificity (2): The degree to which the causal relationship approximates a one-to-one mapping: i.e., the cause does not result in other outcomes, and the effect is not produced by other causes (Woodward 2010; Ross and Woodward 2016).

Stability: The extent to which a causal relationship is maintained over a large range of background conditions (Woodward 2010

Proportionality (1): The most appropriate cause and effect variables are selected from a causal chain (Kendler 2005; Pocheville et al. 2017).

Proportionality (2): The causal relationship is described neither too broadly (including irrelevant information), nor too narrowly (excluding relevant information) (Yablo 1992; Woodward 2010; Pocheville et al. 2017).

</div>

Recent work has helpfully linked interventionist criteria to Koch's postulates. Ross and Woodward (2016) emphasize that Koch's postulates involve claims about the *specificity* of causal relationships. Specificity is concerned with the tightness of the cause-effect link, and can be understood in two ways. The first is the extent to which a cause has fine-grained control or influence over its effect: variations in a cause's states produce corresponding variations in its effect's states. Explanations that feature causal relationships with higher specificity in this sense are thought to be better explanations than those that feature causal relationships with lower specificity (Box 1).[4]

The second sense of specificity refers to how closely the cause-effect relationship approximates one-to-one mapping (Woodward 2010). It is in this sense that Ross and Woodward argue that Koch's Postulate One is akin to 'specificity of the cause', in which microorganisms of a given type are the only causes of a disease. Postulate Two captures 'specificity of the effect', in which the microorganism does not cause any other kinds of disease. Combined, these two aspects of the cause-effect relationship achieve 'one-to-one' mapping of cause to effect.

The explanation of excess acid causing ulcers does not score well on this dimension of causal specificity. Peptic ulcers encompass two types: duodenal and gastric. While duodenal ulcers are consistently associated with an increase in acid production, the story is more complicated for gastric ulcers, where acid secretion decreases (Kusters et al. 2006; Malfertheimer et al. 2009). The acid explanation thus fails the 'specificity of the cause', which would require excess acid secretion to be present in every case of the disease (gastric and duodenal).

H. pylori, however, is associated with both duodenal and gastric ulcers. Does it meet the specificity criterion? *H. pylori*'s causal role in ulcer production is also limited in specificity but for different reasons. It is well-known that most infections of *H. pylori* are asymptomatic (Graham et al. 2009), which means the bacterial explanation fails the 'specificity of the cause' condition in Postulate One. A further complication is that some ulcers form without any host exposure to *H. pylori* or even to the alternative known cause of ulcers, anti-inflammatory drugs (McColl et al. 1993). *H. pylori* thus fails the 'specificity of the effect' condition of Postulate Two. Combining both failures of specificity means that having an ulcer is not diagnostic of an

---

[4] Typically, more distal causes will have less specificity than more proximate ones, simply because the longer a causal chain, the more associated causes would need to be mapped to unique values of the effect.

H. pylori infection, and having an *H. pylori* infection does not mean you have or will get an ulcer.[5]

Koch's postulates can also be connected to other interventionist dimensions of explanatory strength (Box 1). *Stability* reflects how sensitive a causal relationship is to background factors, and is sometimes referred to as explanatory 'robustness' (e.g., Woodward 2006). Difficulties in reproducing experimental findings may result from the low stability of a focal causal relationship. If repeating an experimental manipulation in a slightly new background does not reliably produce the same outcome, the original causal explanation is deemed unstable.[6] Clinical trials are held to this kind of standard: a causal relationship between a treatment and an outcome is confirmed only if it is shown to hold across a range of different genetic, physiological and socioeconomic backgrounds. Postulate Three, which stipulates that the disease can be induced by the same pathogen in different organisms (including other species), can also be interpreted as requiring stability of the causal relationship in different background conditions.

Stability is problematic in the *H. pylori* case. Infection rates vary across continents and countries, with currently low rates in Oceania (24%) and high rates in Africa (70%) (Hooi et al. 2017). In Western countries, only 20% of infected people will develop an ulcer. Although strain differences in *H. pylori* are often mooted as the difference-maker for pathologic outcomes (some strains being more virulent), this has limited evidential support due to high variation among *H. pylori* strains and at best modest associations between particular strains and disease (Lin and Koskella 2014). Instead, regional diets interact causally with *H. pylori* infections to determine disease outcomes in groups of infected humans (Graham et al. 2009). This suggests a failure of stability. The relationship between the microorganism and disease outcome depends upon particular background factors, which in this case involve diet.

Stable relationships generate better causal explanations because they are generalizable to more situations (Mitchell 2000; Woodward 2010). A focus on stability may explain why some researchers argue that '[c]olonization with *H. pylori* is not a disease in itself but a condition that affects the relative risk of developing various clinical disorders of the upper gastrointestinal tract' (Kusters et al. 2006:458). These authors go on to suggest that the pathology of *H. pylori* infections is not caused by 'direct bacterial activity', but by the general response of the human immune system. This explanation obviously scores very well on stability, because of how general the immune response is. But its very broadness raises another explanatory issue.

Human immune responses do not score well explanatorily in terms of the third interventionist dimension of *proportionality*. Proportionality generally refers to the most appropriate level of description of a cause, and like stability, it can be further broken down into two types (Pocheville et al. 2017). The first focuses on *which* cause is selected along a causal chain. Any causal variable of interest will produce a series of effects, and any given effect will be produced by a chain of causes. Often, more distal causes and effects are less proportionate explanations compared to proximate ones. For example, historical and more recent hypotheses about ulcers argue that stress is at least as important a causal factor as acid secretion and *H. pylori* (e.g., Levenstein et al. 2015). But even if stress is implicated in the causal chain to ulcers, it has a

---

[5] These issues point to a possible interpretation of Koch's postulates as being about necessary and sufficient conditions. While they are sometimes interpreted this way (e.g., Susser 1991), this is not what is claimed under an interventionist account of the postulates (Ross and Woodward 2016).

[6] Stability is also related to proximity. The more proximal a pair of cause and effect variables, the fewer causal links or steps between them. The fewer causal links there are, the fewer background conditions (per causal link) can perturb the total causal relationship. Generally, therefore, distal causal relationships are considered less stable than proximal ones (Woodward, 2010:294).

distal role that initiates a series of more proximate physiological responses. Some of these responses would be relevant to ulcer formation, such as an increase in acid secretion, but stress would also produce other phenotypes (e.g., stress hormones and their immune and nervous system effects). A distal factor, such as stress, often means a more complicated causal pathway to the effect, due to interactions with other contributing causes.[7] Accordingly, acid secretion generates a more proportional causal explanation than stress, and this proportionality is the reason for the former's dominance prior to the *H. pylori* findings.

The second sense of proportionality looks at how narrowly or broadly a cause or an effect variable is described. If a description is too narrow in scope then it excludes other important factors in bringing about the effect; if it is too broad, it will include factors that are irrelevant to the production of the effect. This sense of proportionality can often involve a trade-off with stability. Causes that are stably associated with effects, such as the human immune response discussed above, lack explanatory power due to their disproportionately broad scope and irrelevant information. Immune responses refer to a heterogeneous range of factors that are causally implicated in a large range of different effects. Though less broad than the immune response, acid production is also a disproportionately broad cause compared to *H. pylori* because of the different pathways to gastric and duodenal ulcers, and acid secretion's causal influence on other phenotypes.[8]

A major reason why many researchers see the relationship between *H. pylori* and ulcers as paradigmatically causal is because the explanation generated successful treatment outcomes. Intervening on *H. pylori* is a successful way to get rid of ulcers. Just as for Koch's own analysis of cholera, the elimination of ulcers via the elimination of *H. pylori* seals an inference of causation: 'The ultimate proof of *H. pylori* as the main cause of ulcer disease was the permanent cure of peptic ulcers by eradication of the infection' (Malfertheiner et al. 2009:1450).

In sum, although the *H. pylori* explanation of ulcers is weak on the specificity and stability dimensions, it fits an interventionist account of causation better than alternative explanations, and scores higher on proportionality.

The broader message here is that the causal story linking *H. pylori* to ulcers is much more complicated than the standard textbook version (Thagard 1998).[9] Given that *H. pylori* is merely a single microbe, how well will microbiome explanations of host states fare? Unlike the *H. pylori* case, most microbiome research has not yet resulted in successful treatment regimes for human disease. This gap suggests that causal claims about the microbiome's role in human disease and mental states require careful assessment with causal criteria.

---

[7] Note that this sense of proportionality can apply to both the selection of causes and effects. Whether or not proportionality is used to select causes or effects depends upon the explanatory project: whether it is focused on explaining what causes an outcome, or on explaining what kinds of outcomes are produced by a particular factor. When selecting the most appropriate effect, proximity is also a good indicator of proportionality. This is because proximate effects often encompass the entire causal output, whereas more distal effects are parsed into different variables, some of which have other contributing causes influencing them.

[8] This sense of proportionality can also be applied to both causes and effects, again depending on the explanatory project. When trying to understand the effects of a particular factor, effects may be invoked that are inappropriately broad or narrow, such as claiming that *H. pylori* causes illness, rather than more exactly ulcers.

[9] Philosopher Paul Thagard (1998) has long argued that the interacting causal factors in the *H. pylori*-ulcer explanation make a 'causal network' explanation more effective than a single-cause explanation.

The microbiome 'revolution'

Contemporary microbiome researchers look for microbial causality rather differently from classic microbiology, which focuses on single microbes that can be investigated within the framework provided by Koch's postulates. However, the aim of microbiome research in human hosts is similar: to establish whether microbiomes are causally responsible for particular health and disease states. Standard microbiome investigations start with DNA extraction from a sample of all the microorganisms in a specified environment (including the human gut). The DNA is sequenced and then analysed computationally. Sophisticated bioinformatic methods analyse this mixed-species DNA sequence to find compositional patterns. These microbiome patterns can then be associated with particular host states, such as diseases and disorders. When associations seem strong, experimental efforts are brought to bear on the relationship. We assess two such cases. The first attempts to link microbiomes to obesity and the second to mental health disorders. But first we address an important question: What is the microbiome, anyway? This is crucial to understanding how microbiomes are meant to work causally.

What is the microbiome?

There are at least four different ways to characterize microbiomes, which we conceptualize as 'compositional', 'functional', 'outcome-oriented', and 'causal core' microbiomes. In many cases, including those of obesity and mental health, characterizations vary and the nature of the 'microbiome' being manipulated is not clear.

Generally speaking, any microbiome is a complex entity of multiple interacting components. Human gut microbiomes consist of around 40 trillion organisms that can be sorted into at least 1000 species-like groups (Sender et al. 2016). The microbiome is often first characterized compositionally in attempts to identify these taxa and their relative proportions. When characterized compositionally, host-associated microbiomes differ extensively within human populations, and within individual human hosts over time (Human Microbiome Project Consortium 2012).

However, there is often much more similarity across human microbiomes when understood at a functional level. Functions are usually assessed in terms of biochemical pathways, and it is these pathways that produce phenotypic effects. Metabolic pathways are a common focus here: the functional microbiome is understood on the basis of all of the metabolic pathways and products of the microbiota. But because experimental biochemistry is rarely carried out with microbiome members, functions are typically inferred from genes in the microbiome sequence (Manor and Borenstein 2017). These first two ways of characterizing microbiomes are often closely related: compositional characterizations of microbiomes that focus on genes to identify microbial taxa also allow functional inferences about the metabolic and other biochemical pathways these taxa might possess.

In addition to compositional and functional characterizations of the microbiome, many researchers use what could be called an outcome-oriented characterization. This occurs most obviously in discussions of 'dysbiosis', which is defined as a microbiome state that produces disease states in the host-microbiome system (see Hooks and O'Malley 2017). Because of the interactions between host and microbiome, the state of the latter is frequently characterized in terms of the host state: a dysbiotic microbiome is stipulated to be the microbiome composition found in diseased hosts, such as those with cancer (e.g., Schwabe and Jobin 2013). In these cases, the causal characterization of the microbiome becomes trivially circular in explaining those 'dysbiotic' states. In other words, the dysbiotic nature of the microbiome is inferred from the disease, and the disease is explained by the dysbiotic microbiome. If a mechanism is

speculated on, it is often framed in terms of 'imbalance', with an imbalanced microbiome thought to be causally productive of disease (e.g., Round and Mazmanian 2009). This kind of characterization can also be done the other way, for 'healthy' microbiomes. Terms such as 'homeostasis' are applied to microbiomes associated with healthy host states (e.g., Schwabe and Jobin 2013). This positive outcome-oriented characterization is even looser than dysbiosis claims, because health is simply the absence of any disease when such terms are used.

Increasingly, a fourth characterization of the microbiome is used, which we call the 'causal core' microbiome. Researchers began discussing core microbiomes compositionally, in terms of taxa shared by all healthy humans (e.g., Hamady & Knight 2009; Turnbaugh et al. 2009). It was hoped this would allow diagnosis and treatment to circumvent the high inter-individual variability of microbiomes. However, this taxonomic core proved elusive, unless identified at very coarse levels (e.g., phyla), and even then, the quantities of organisms in such 'core' taxa varied massively between different hosts or across studies (Lloyd-Price et al. 2016). Consequently, some researchers have attempted to discuss the causality of microbiomes in terms of key functions (e.g., metabolic pathways), but whole microbiomes are not well characterized functionally. One solution has been to pick out a few 'core' members that are known or can be shown to be doing something causally (e.g., Surana and Kasper 2017; Neville et al. 2018). Determining the causal-functional contribution of this subgroup means that only some of the microbiome is considered causally efficacious for a particular host state.

These four perspectives show that microbiome descriptions are conceptually heterogeneous. Some researchers assume a compositional account, some a functional account, some an outcome-oriented approach, while yet others focus on a causally relevant 'core'. Single research papers often combine two or more of these concepts when discussing microbiome causality, such as when a 'causal core' microbiome co-explains obesity with a dysbiotic ('outcome-oriented') microbiome (e.g., Truax et al. 2018). Often, even when researchers refer to a whole microbiome (compositional or functional), they really appear to have in mind just a few members with a particular causal role ('causal core'). Furthermore, microbiome composition varies both across individual hosts and within individual hosts over time, so even when a single way of characterizing the microbiome is accepted, it does not refer to a fixed or invariable causal agent. This range of different ways to conceptualise microbiomes is entangled with a range of causal claims about microbiomes. We show how the explanatory problems multiply in the two cases that follow.

Microbiomes and obesity

A microbiome association with obesity was initially discussed compositionally in terms of differences in the ratios of two phyla (a very coarse taxonomic level, indicated by DNA sequence). These relative compositions appeared to predict obesity reliably in hosts (Ley et al. 2006; Turnbaugh et al. 2009). When mice or human gut microbiomes were found to have more Firmicutes than Bacteroidetes (the two main phyla in human and mouse guts), the hosts were very likely to be overweight; conversely, if the microbiomes had more Bacteroidetes, the hosts were likely to be leaner. In a human-based manipulation of these findings, obese humans who lost weight by dieting also lost their over-representation of Firmicutes (Ley et al. 2006).

Findings like these stimulated an experimental regime to test the inferred causality of microbiome-host relationships. The gold standard of experimental techniques quickly became microbiome transplants from conventionally colonized mice with one phenotype (lean or obese) to germ-free mice. Germ-free mice are born in sterile laboratory environments and live their lives without microbiomes until they are experimentally colonized. The absence of a

microbiome is to ensure a microbiome transplant is not confounded with effects from the native microbiome. When germ-free mice receive the transplant, almost always from donor faeces (called a faecal microbiome transplant or FMT), several key studies found the donor phenotype of obesity also transferred (Turnbaugh et al. 2008; 2009). Germ-free mice that received microbiomes from obese mice became obese, and germ-free mice that received microbiomes from lean mice remained lean (their typical condition). One study has explicitly demonstrated the transferability of a lean phenotype from a whole microbiome, to mice in a particular post-surgical condition (Liou et al. 2013).

Transfer-of-phenotype FMT experiments also succeeded from obese humans to mice, with the same effects (e.g., Turnbaugh et al. 2009). Even when human twin samples were used, to remove genetic confounding on the host side, obese phenotypes successfully transferred to the mice via human faeces (Ridaura et al. 2013). One human-mouse study transferred a single bacterial species from a human microbiome to decrease weight in mice (Goodrich et al. 2014a). In all these cases of phenotype transfer, say researchers, 'the microbiome is considered causal' (Goodrich et al. 2014b:250; see also Harley & Karp 2012).

Challenges for microbiome explanations of obesity

However, even in what appears to be a strong causal relationship between microbiomes and obesity, the causal direction cannot be determined due to confounds. For example, Ley et al. (2006) found that obese individuals who lost weight on a diet also lost their over-representation of Firmicutes. In this case, diet may be a common cause of both obesity and microbiome state, rather than the microbiome being the primary cause of obesity via direct mechanisms (Hildebrandt et al. 2009).

The putative causal relationship also appears to be highly unstable. When the background factor of diet is manipulated and germ-free mice are put on low-fat diets, obesity FMTs fail (Ridaura et al. 2013). Additionally, manipulations of diet in germ-free mice have shown that the absence of a microbiome does not prevent weight gain on a high-fat diet (Fleissner et al. 2010). This means that the causal relationship between microbiomes and weight is only maintained under certain background conditions, thus scoring low on the stability dimension. As well, germ-free mice have many abnormalities that make extrapolating effects on them to other mice and humans very problematic, even when the mice receive FMTs from humans (Nguyen et al. 2015). Causal relationships in germ-free mice between microbiomes and phenotypes such as obesity are not, therefore, necessarily indicative of a more general, stable causal relationship between microbiomes and these phenotypes in other organisms, or even within healthy colonized mice.

Further indications of instability are found in the association literature. Systematic reviews and meta-analyses show that changed phyla proportions or decreased diversity in the microbiome are not robust predictors of obesity in humans. Both increases and decreases of the key phylum Firmicutes have now been associated with obesity (e.g., Sze and Schloss 2016; Duvallet et al. 2017), which suggests that as well as the causal structure being unclear, any causal relationship between microbiomes and obesity is sensitive to changes in background conditions.

Specificity in the microbiome-obesity relationship is also undermined by the way microbiome effects are conceptualized. Recall that for the first sense of specificity (see Box 1), different states of a cause permit fine-grained control over its effect's states, so that varying the cause results in corresponding variation in its effect. However, the effects of leanness and obesity are quantitative traits occurring along a continuum. Although cut-off criteria allow individuals to be classified into groups related to these traits, individuals do not exist in the world as either obese or lean. Nevertheless, microbiome research classifies these phenotypes as binary states (obese

versus lean). The clearest experiments on germ-free mice affect only these crude phenotypic classifications, and very rarely induce leanness. This kind of descriptive coarseness also occurs when 'dysbiosis' is postulated. Microbiomes are treated as if they occur in one of two states: healthy or dysbiotic (disease-producing). Binary categorizations such as these suggest that although microbiomes may be causally implicated in a trait like obesity, the known specificity of the microbiome on the categorized trait is low (Duvallet et al. 2017).

A more specific causal relationship would allow finely tuned manipulations of the microbiome to have similarly precise effects on the host phenotype. For example, more of a certain kind of bacteria would lead to a heavier host. Newer microbiome research is in fact attempting to find very specific groups of bacteria that have causal effects on obesity (e.g., Truax et al. 2018). However, if only some taxa are causally relevant (the 'causal core' concept), this makes the notion of a whole microbiome as a causal entity disproportionately broad. FMTs involve transfers of the entire gut community, and thus invoke the idea of the whole microbiome (conceptualized compositionally, functionally, or both) as the causal agent. But in any situation where the microbiome might cause anything, researchers could be referring to one or more of the four microbiome concepts identified above. Further exacerbating this representation problem, FMT manipulations are usually performed with whatever members of the community emerge in faeces, even though faecal microbiomes are known to be different from gut microbiomes (Momozawa et al. 2011). FMTs also include viruses, Archaea and eukaryotic microbes, but most microbiome studies make causal claims about only bacteria (admittedly the most abundant type of microorganism in the gut, apart from viruses).

These problems indicate a failure of proportionality in characterizing the causes of observed effects in obesity and other FMT studies. When considering the second sense of proportionality (Box 1), an FMT manipulation that changes only a few elements of the microbiome composition, rather than the microbiome as a whole, would be more appropriately described as an intervention on particular bacterial taxa. It is these focal bacteria that cause the disease state observed, rather than the (whole) microbiome. This is the finding of work that implicates a single family of gut bacteria in the prevention of obesity (Truax et al. 2018). Even though the authors claim this focal subgroup ameliorates 'dysbiosis', their intervention manipulates only one particular subset of the microbiome, and even more proximally focuses on a certain anti-inflammatory protein produced by these (and other) bacteria.

In general, the causal components of microbiomes are not known. When causality is attributed to any microbiome, it often serves merely to signal that 'something' in the microbiome is causal. For instance, Grinspan and Kelly (2015:17), when discussing colitis, observe that the results from FMT experiments 'suggest that FMT is a heterogeneous treatment, and the effect may be donor dependent … we still do not understand the active component of FMT. Is it a bacteria or bacterial metabolite, or is it the interaction between microbial and host factors?'

Invoking the whole microbiome (compositionally or functionally) as a causal agent seldom if ever achieves the right level of explanation or the most appropriate causal description for many microbiome findings. As an illustration, the association between microbiome composition and obesity suggests that the relative compositions of Firmicutes and Bacteroidetes are responsible for obesity effects. Explanatorily, phyla ratios would indicate that small-scale characteristics of the microbiome (e.g., smaller taxonomic groups or particular metabolic functions) are causally insufficient for such effects. The unlikeliness of that being true suggests that the Firmicutes-Bacteroidetes causal description is too restricted (to phyla ratios only), and thus disproportional. Conversely, if microbiomes are described too broadly, then they include numerous elements that are irrelevant to the production of the focal effect (Mithieux 2018). To overcome the proportionality problem, a growing number of microbiome studies now home in on much smaller groups of organisms as causal agents in a community context, akin to the 'causal core'

concept. Indeed, advocacy for more 'reductionist' causal ascriptions is ramping up in microbiome research (e.g., Fischbach 2018).

Microbiomes and mental health

The obesity transfer-of-phenotype experiments via FMT became 'proof of principle' for showing causal relationships in microbiome research. This methodological template inspired researchers in other fields to seek causal connections between microbiomes and other bodily states, notably brain functions and behaviours. Using the language of 'gut-brain axis' (a known but poorly understood connection), increasing numbers of studies have tested the involvement of microbiomes in both psychiatric disorders and general brain states. In these experiments, germ-free and other mice are again the main experimental systems. However, the main manipulation in this subfield of microbiome research is not whole microbiome transfers, but instead the administration of probiotics (living microorganisms, usually just a few strains or species) to the mouse gut, to see if this changes the brain or behaviour.

Anxiety and depression are key targets of these probiotic manipulations, with reports of less depression-related and anxiety-like behaviour in mice receiving probiotics (e.g., Bravo et al. 2011). These behaviours are measured by performance on classic mice-testing devices that involve measurable responses to environments. Once behavioural differences are found between probiotic and control groups, mechanisms may be postulated. For example, microbial metabolites are suggested to be 'acting directly on the brain' (Collins et al. 2013:241), or to have indirect effects on host production of neurochemicals.

A modest amount of brain-behaviour microbiome research does manipulate whole microbiomes via FMT instead of using probiotics. One study famously showed that FMTs from a 'bolder' strain of mice into the guts of 'timid' strains made the timid mice bolder (Bercik et al. 2011). Depression-associated symptoms have also been transferred via FMT from depressed humans to antibiotic-treated rats (who thus have a 'depleted' microbiome), and this finding is suggested to provide 'definitive evidence' of a causal relationship between microbiomes and mental health (Kelly et al. 2016:115).

FMTs are believed to change the overall microbiome composition. Probiotics are also suggested to induce positive changes in the community of gut microorganisms, and these broad compositional or functional states are often argued to be the causal agents of phenotypic change (e.g., Arnold et al. 2016). This level of explanation focuses on the whole community, rather than the probiotics themselves or any key taxonomic groups in the FMT. These changes might be described as a simple shift in the relative proportions of taxonomic groups of microorganisms in the community, or as a movement toward a 'normal' community composition (e.g., Ait-Belgnaoui et al. 2012).

Challenges for microbiome explanations of mental health

Just as in the obesity case, the very structure of causality is problematic in this field. In many of the experiments linking microbiomes to mental health, no direct manipulation of the microbiome has been conducted. Probiotics are the intervention. Even if probiotics have effects, their relationship to the microbiome is unclear. Although there is some evidence of positive probiotic effects in mice, there is much less such evidence in humans (Wang et al. 2016). Crucially, probiotic manipulations that appear to produce behavioural outcomes in mice do not produce them in humans (Kelly et al. 2017). Further complications come from research that shows probiotics do not reliably colonize mice intestines (Zmora et al. 2018). Even when

behavioural effects do apparently occur after probiotic treatment of humans, there may be no concomitant changes to their microbiomes after probiotic administration (Kristensen et al. 2016). In other words, the probiotics could be acting as a common cause of changes to the microbiome as well as the measured behavioural outcome, or they may influence the outcome directly without any mediation via the microbiome.

Other studies that manipulate the microbiome more extensively, by carrying out microbiome manipulations with FMTs, also present the common cause problem. FMTs probably cause outcomes via a range of causal pathways. This possibility is addressed by studies that provide placebo interventions alongside FMTs in randomized controlled trials (e.g., Moayyedi et al. 2015), but many FMT experiments are not so stringent. Some studies try to avoid the 'confounding influences' of germ-free mice and their atypical physiology and brains by instead using rodents with microbiomes that have been depleted by antibiotic treatments (e.g., Kelly et al. 2016). However, this strategy introduces a different kind of causal background variable: antibiotic treatment, which is well known to decrease compositional diversity. Many microbiome studies suggest that antibiotic-depleted microbiomes are causally related to disease symptoms (e.g., Cox et al. 2014). Using antibiotic treatments as a proxy for germ-free status thereby introduces another confound.

In addition to an unclear causal structure, there are problems in the proportionality of causal claims made about the microbiome and mental health states. Experiments using probiotics involve at most a few strains of organism, often classic milk-fermenting microorganisms such as lactobacilli. A manipulation involving probiotics might alter a specific subset of the microbiome of an organism (i.e., a 'causal core' microbiome), or it may bring about widespread changes to the community composition (the compositional microbiome). However, probiotics may also cause no changes in the microbiome composition at all (e.g., Kristensen et al. 2016), in which case any putative effect on host brain or behaviour might be a direct effect of the probiotics, or might be due to probiotics changing some other factor. As in the obesity case, this issue points to a problem with the proportionality of the causal explanation. In experiments administering probiotics, it is not appropriate to make causal claims about the entire microbiome, because these statements describe the cause too broadly and include information not related to the intervention.

As well as the more general worry of which microbiome concept is being invoked (related to issues with Proportionality [2] in Box 1), these studies are subject to problems with proportionality in the first sense, which focuses on how variables along a causal chain are selected as the most relevant causes or effects. We noted above that most causal claims about microbiomes influencing anxiety and depression are not based on observations of depressed or anxious humans. Instead they refer to rodent behaviour on standardized testing devices. For instance, tests of 'anxiety-like' behaviours measure the exploratory tendencies of rodents in different states (e.g., germ-free, FMT recipients, or probiotic recipients), while tests of 'depression-like' behaviours often focus on factors such as the amount of time spent swimming versus floating in a beaker of water. Even if we assume a meaningful mapping for anxiety and depression between rodents and humans (problematic in its own right), these studies do not measure the effect variables 'depression' and 'anxiety'. They measure rough behavioural proxies of these states. Although in animal studies (including those of humans) it is not possible to measure psychological states directly, in human psychological research it is at least possible to collect self-reports of participant mental states. Self-reports provide more proportional evidence for claims about mental states, because they are at least reflective of the states themselves, whereas proxy behaviours may occur downstream from the mental states along a causal chain. In rodents there are also many questions about how adequately such tests measure anything to do with genuine anxiety or depression states. These are standard worries for rodent-

based research, but they are compounded by the problems we have already noted about microbiome variables and causal structure.

Relatedly, depression and anxiety are not binary variables with on-off states, as discussed in relation to the variables of obesity and leanness. Yet what the interventions focus on is anxiety versus non-anxiety-like behaviours, and depression versus non-depression-like behaviours. As we noted for obesity, with only binary effect measures, the specificity of the cause and effect relationship is low.

For the second sense of specificity, approximating one-to-one mapping in mental health and other microbiome research is simply not possible given the breadth of phenotypes believed to be causally influenced by microbiomes. These include diabetes, hypertension, numerous cancers, intestinal diseases, liver diseases, immune disorders, multiple sclerosis, asthma and other respiratory diseases, erectile dysfunction, autism and other neurodevelopmental disorders, neurological conditions, migraines, dermatitis, halitosis, vision loss, and unhealthy ageing (see Gilbert et al. 2016). This proliferation of disease effects leads to a failure of specificity of the cause, in much the same way that 'immune response' fails specificity of the cause in explanations of ulcer formation. Regarding specificity of the effect, traits in which the microbiome is implicated (i.e., obesity or mental health disorders) have for a long time been understood as multi-factorial, with a range of causal contributions (e.g., Reich et al. 1975; Grundy 1998). Microbiome relationships to such disorders will not, therefore, meet specificity of effect criteria—even assuming it is clear what precisely 'the microbiome' is, and whether or not it is being intervened on. As we showed above, often it is not.

Reconstructing microbiome explanations

We have argued that explanations in contemporary microbiome research face a number of problems. Explanations are presented as causally linking the microbiome to particular human health or disease states. But they often fail to establish causality, to intervene on the microbiome itself, or to converge on a consistent understanding of precisely what the microbiome is in the first place. Causal explanations often seem disproportionate, to hold only under certain kinds of background conditions (instability), and to have low specificity. Although some of these problems also beset well-established microbiological explanations (e.g., the case of *H. pylori* and ulcers), they are especially relevant in new fields seeking to establish their causal credentials and treatment potential, such as microbiome research. Issues surrounding the proportionality of causal explanations appear to be particularly problematic in microbiome research, due to the scale of whole microbial communities, and the heterogeneity of ways to conceptualize the microbiome itself.

Microbiome researchers generally agree that better causal explanations are precisely what the field needs now to make progress (e.g., Fischbach 2018). How might clearer and stronger causal explanations be achieved? We answer this question by way of a microbiome success story that not only produces a treatment but also leads back to Koch's postulates.

C. difficile and Koch's postulates

It is well-known that FMTs ameliorate the effects of a particularly nasty pathogen in the gut, *Clostridium difficile*. Administering an FMT, from almost any healthy donor, consistently overcomes the debilitating disease effects induced by *C. difficile* infections (van Nood et al. 2013). Being able to cure *C. difficile* robustly with FMTs indicates the stability of the causal relationship between FMTs and *C. difficile*. In other words, the causal relationship persists

across many different background conditions (individuals and populations). This success has often been interpreted as evidence in favour of the whole microbiome being causally efficacious in curing *C. difficile*. But does this treatment indicate specificity and proportionality?

Recently, a few studies have decomposed FMTs to the members that are most causally potent against *C. difficile*. These studies identify which microorganisms play the strongest causal roles, and the particular pathways involved (e.g., Stein et al. 2013; Buffie et al. 2015). In other words, the earlier 'lumping' of the causal factors into one of the broader microbiome concepts (such as a functional or compositional account) is overcome by fine-tuning a broad cause to a much narrower complement of causal factors. Although some researchers claim this is still a whole microbiome explanation, because the FMT has altered the niche of the microbiome (e.g., Lawley & Walker 2013), an interventionist framework suggests that the most proportionate causal explanation involves a small subgroup of the microbiome. Niche modification refers to the causal background. Factors in the background have no immediate difference-making effects and thus are not part of the central explanatory variable, which is the 'causal core' microbiome. The microbiome captured by this 'causal core' sense is both a specific explanation of *C. difficile* amelioration (in the one-to-one mapping sense of specificity), as well as the most proportional explanation for *C. difficile* cures.

This example of microbiome research provides a good causal explanation because a restricted number of bacterial components explain *C. difficile* cures. The success of this strategy should come as no surprise, because it is a relatively proximal cause. More proximal causal relationships often provide better causal explanations due to proximity increasing both stability and specificity. In contrast, recent findings that dietary manipulations can suppress *C. difficile* (Hryckowian et al. 2018) rely on the notion that diet fosters 'anti-dysbiotic' microbial communities that somehow inhibit *C. difficile*. This account depends on less proximate and less defined causal relationships. Although we do not dispute that coarse-grained interventions can often have useful effects (e.g., diet, antibiotics),[10] there is less explanatory strength according to the interventionist dimensions embedded in many scientific explanations.

One approach to better causal explanations in human microbiome research would, therefore, be to identify more proximal causal relationships between the microbiome and phenotypes such as metabolic pathways. Another would be to decompose the microbiome into component parts that can be understood independently (at least initially) as would occur with Koch's postulates. However, as we have seen in the *Helicobacter* case, Koch's postulates are not always achieved, and yet still a causal explanation is deemed to exist especially if it produces an effective treatment via intervention. The *C. difficile* story, however, shows that invoking the whole microbiome (via FMTs) leads to a failure of proportionality and specificity, and yet a very stable treatment. As well as indicating that successful treatments may not be aligned with the best causal explanation, this case suggests that more philosophical work needs to be done to understand the trade-offs between causal specificity, proportionality, and stability.

An ecological version of Koch's postulates in microbiome research has been touted as an alternative to the traditional single-microbe Koch's postulates (Vonaesch et al. 2018). However, this ecological version simply distributes all causality across the whole microbiome, and thus reinforces the proportionality problem. Should we think of microbiome effects as emergent, and that such emergent features are 'wholly unpredictable from observations of components' of the microbiome (Hall et al. 2018:979-980)? Right now, claims about the causal role of emergent microbiome properties are simply acknowledging that far too little is known about microbiome components and their interactions. Even if such claims eventually find more evidential support,

---

[10] This is generally due to pragmatic factors such as ease of intervention. While diet may be a distal cause, it is often easier to manipulate compared to more proximate physiological factors.

the same interventionist dimensions of explanation should be brought to bear. Such dimensions could be additionally helpful when there appear to be competing levels of microbiome explanation (i.e., whole versus part).

Concluding reflections

Much of the enthusiasm for microbiome research is based on its capacity to provide a 'revolutionary' methodological, conceptual and explanatory alternative to traditional microbiology (e.g., Cho and Blaser 2012). Experimental evidence such as we outline above suggests there is something causal going on in microbiome research. Manipulations are made, and effects are produced often enough to be deemed reliable. But the exact nature of the causal relationships, and the explanatory credence that should be given to causal claims about the microbiome, are not well worked out. There are numerous weaknesses in causal inferences from experimental microbiome research.

The way in which the microbiome itself is conceptualized and described differs between (and sometimes within) studies. Compositional and functional accounts of whole microbiomes describe causes too broadly, and target phenotypes such as behavioural measures in mice are often described inexactly, meaning that causal explanations are commonly disproportionate to their effects. The coarse-grained or binary descriptions of outcomes such as obesity, leanness, depression and anxiety limit the degree of causal specificity in microbiome causal explanations. Proposed causal relationships are potentially confounded by common causes, which in many instances could be due to the experimental manipulation itself (i.e., probiotic administration). Well-accepted causal relationships (i.e., microbiome composition and obesity) display relatively low degrees of stability across populations and species, which could account for problems in the replicability of microbiome research.

But despite the recent identification of a narrower causal basis in the FMT treatment of *C. difficile*, this explanatory strategy may not hold for many other outcomes influenced by the microbiome. Explaining obesity, depression, and anxiety by microbiomes is much harder than explaining how a single (albeit dire) infection can be controlled by elements of the microbiome. All of the former are complex multifactorial traits that have heritable components, and are not likely to be captured by the identification of a few proximal pathways. The multi-causal nature of these traits makes it likely that there are many interactions between numerous components of the microbiome, the host, and their environments. It is for this reason that one-to-one specificity is not possible for whole microbiome explanations, and it also suggests that the causal influence of the microbiome as a whole would be extremely unstable, and difficult to isolate experimentally. Current research already indicates this to be the case. Given these limitations, it might be most productive to focus causal explanations on specific microbiome components, and to limit causal explanations to more proportional descriptions of behaviour, rather than complex human traits such as mental health states. But even if complex multi-causal traits challenge an interventionist causal-explanatory framework, its tools can be useful for isolating and evaluating multiple important contributing causes (Ross 2018).

Establishing causation is not easy, and microbiome research is not unique in facing the challenges we outline. To end on a positive note, what kinds of explanatory and other claims can or should be made, given microbiome complexities? Our microbiome cases, and their comparison with the *H. pylori* explanation, show there are two avenues by which explanatory problems can be addressed: more explicit conceptualization of the focal causal entity, and more precise assessment of the dimensions of causal explanation.

Our discussion suggests avenues for clearer communication about the focal causal entity. First, there is the matter of whether the whole microbiome, as a unitary causal entity, is the proper explanatory focus. We have questioned its role as such, at least in many cases of microbiome research to date. First, interventions such as FMTs and probiotics do not convincingly support the inference that the whole microbiome has been intervened on, and second, there are not yet any strong reasons to think that 'emergent' properties of the whole microbiome play a causal role. Even if these properties are eventually identified, specifying what is meant by 'the microbiome' is always a good starting point for better causal explanations. Our four ways to characterize the microbiome could help: if it is clear which aspect of the microbiome is at stake in a given study—its composition, function, outcome, or causal 'core', or some conjunction of these—it will be easier to tell when researchers are talking about the same causal agent.

We realize that at least sometimes researchers target 'the microbiome' without commitment to its status as a unitary causal entity. Rather, this phrase is used as a placeholder for a black-boxed subgroup of human microbiota: some to-be-defined microbial taxa or their biochemical pathways. To the extent that this is the case, explicit articulation of that black-boxing would appropriately shift the focus and facilitate clearer causal explanations. Perhaps in some instances researchers talk about the microbiome in this way because broad microbiome claims are heuristically valuable, especially in exploratory studies of new candidate microbiome effects. Even if this is the intention, clearer articulation of background assumptions would be productive. In particular, conceptual precision requires identifying whether whole microbiomes are causally salient, as opposed to some focal subgroup of the community.

Finally, the second avenue for addressing explanatory complexities in microbiome research is to assess more precisely the dimensions of causal explanation, and communicate more carefully the extent to which a successful causal explanation has been achieved. If microbiome researchers continue to investigate complex traits, whatever their understanding of 'the microbiome', these investigations will benefit from evaluating how prospective explanations rank on the dimensions of stability, specificity, and proportionality. Bearing these dimensions in mind may also help microbiome research to develop additional methods for achieving causal explanations of human health and mental states.

References

Ait-Belgnaoui A, et al. (2012) Prevention of gut leakiness by a probiotic treatment leads to attenuated HPA response to an acute psychological stress in rats. Psychoneuroendocrinology 37:1885–1895

Arnold JW, et al. (2016) Emerging technologies for gut microbiome research. Trends Microbiol 24:887–901

Bercik, P, et al. (2011) The intestinal microbiota affect central levels of brain-derived neurotropic factor and behaviour in mice. Gastroenterology 141:599-609.e3

Blaser MJ (1996) The bacteria behind ulcers. Sci Am 274:104-107

Blaser MJ (2014) The microbiome revolution. J Clin Invest 124:4162-4165

Bravo JA, et al. (2011). Ingestion of *Lactobacillus* strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. Proc Natl Acad Sci 108:16050–16055

Buffie CG, et al. (2015). Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*. Nature 517:205.

Calcott B (2017) Causal specificity and the instructive–permissive distinction. Biol Philos 32:481–505

Carter KC (2003) The rise of causal concepts of disease: case histories. Ashgate

Chen D, et al. (2007) Does *Helicobacter pylori* infection per se cause gastric cancer or duodenal ulcer? Inadequate evidence in Mongolian gerbils and inbred mice. FEMS Immunol Med Microbiol 50:184-189

Cho I, Blaser MJ (2012) The human microbiome: at the interface of health and disease. Nat Rev Genet 13:260–270

Collins SM, et al. (2013). The adoptive transfer of behavioral phenotype via the intestinal microbiota: experimental evidence and clinical implications. Curr Opin Microbiol 16:240–245

Cox LM, et al. (2014) Altering the intestinal microbiota during a critical developmental window has lasting metabolic consequences. Cell 158:705-721

Duvallet C, et al. (2017) Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. Nat Commun 8:1784

Fischbach MA (2018) Microbiome: focus on causation and mechanism. Cell 174:785-790

Fleissner CK, et al. (2010) Absence of intestinal microbiota does not protect mice from diet-induced obesity. Brit J Nutr 104:919–929

Ford AC, Talley NJ (2009) Does *Helicobacter pylori* really cause duodenal ulcers? Yes. BMJ 339:b2784

Gilbert JA, et al. (2016) Microbiome-wide association studies link dynamic microbial consortia to disease. Nature 535:94-103

Gillies DA (2016) Establishing causality in medicine and Koch's postulates. Intl J Hist Philos Med 6:10603

Goodrich JK, et al. (2014a) Human genetics shape the gut microbiome. Cell 159:789–799

Goodrich JK, et al. (2014b) Conducting a microbiome study. Cell 158:250-262

Goodwin CS, et al. (1989) Transfer of *Campylobacter pylori* and *Campylobacter mustelae* to *Helicobacter* gen. nov. as *Helicobacter pylori* comb. nov. and *Helicobacter mustelae* comb. nov. respectively. Intl J Systematic Bacteriol 39:397-405

Graham DY, et al. (2009) African, Asia or Indian enigma, the East Asian *Helicobacter pylori*: facts or medical myths. J Digest Dis 10:77-84

Griffiths PE, et al. (2015). Measuring causal specificity. Philos Sci 82:529–555

Grinspan AM, Kelly CR (2015) Fecal microbiota transplantation for ulcerative colitis: not just yet. Gastroenterology 149:15–18

Grundy SM (1998) Multifactorial causation of obesity: implications for prevention. Am J Clin Nutr 67(Suppl):563S-572S

Hall EK, et al. (2018) Understanding how microbiomes influence the systems they inhabit. Nat Microbiol 3:977-982

Hamady M, Knight R (2009) Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. Genome Research 19:1141–1152

Hanage WP (2014) Microbiome science needs a healthy dose of scepticism. Nature 512:247–249

Harley ITW, Karp CL (2012) Obesity and the gut microbiome: Striving for causality. Molecular Metabolism 1:21–31

Hildebrandt MA, et al. (2009) High-fat diet determines the composition of the murine gut microbiome independently of obesity. Gastroenterology 137:1716–1724

Hobsley M, et al. (2009) Does *Helicobacter pylori* really cause duodenal ulcers? No. BMJ 339:b2788

Hooi JKY, et al. (2017) Global prevalence of *Helicobacter pylori* infection: systematic review and meta-analysis. Gastroenterology 153f:420-429

Hooks KB, O'Malley MA (2017) Dysbiosis and its discontents. mBio 8(5): e01492-17

Hryckowian AJ, et al. (2018) Microbiota-accessible carbohydrates suppress *Clostridium difficile* infection in a murine model. Nat Microbiol 3:662–669

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

Huss J (2014) Methodology and ontology in microbiome research. Biol Theor 9: 392–400

Kelly JR, et al. (2017) Lost in translation? The potential psychobiotic *Lactobacillus rhamnosus* (JB-1) fails to modulate stress or cognitive performance in healthy male subjects. Brain Behav Immun 61:50–59

Kelly JR, et al. (2016) Transferring the blues: depression-associated gut microbiota induces neurobehavioural changes in the rat. J Psychiatr Res 82:109–118

Kendler KS (2005) "A gene for...": the nature of gene action in psychiatric disorders. Am J Psychiatr 162:1243–1252

Kidd M, Modlin IM (1998) A century of *Helicobacter pylori*. Digestion 59:1–15

Kristensen NB, et al. (2016) Alterations in fecal microbiota composition by probiotic supplementation in healthy adults: a systematic review of randomized controlled trials. Genome Med 8:52

Kusters JG, et al. (2006) Pathogenesis of *Helicobacter pylori* infection. Clin Microbiol Rev 19:449-490

Lawley TD, Walker AW (2013) Intestinal colonization resistance. Immunology 138:1–11

Levenstein S, et al. (2015) Psychological stress increases risk for peptic ulcer, regardless of *Helicobacter pylori* infection or use of nonsteroidal anti-inflammatory drugs. Clin Gastroenterol Hepatol 13:498-506

Ley RE, et al. (2006) Microbial ecology: human gut microbes associated with obesity. Nature 444:1022–1023

Lin D, Koskella B (2014) Friend and foe: factors influencing the movement of the bacterium *Helicobacter pylori* along the parasitism-mutualism continuum. Evol Applications 8:9-22

Liou AP, et al. (2013) Conserved shifts in the gut microbiota due to gastric bypass reduce host weight and adiposity. Sci Transl Med 5:doi:10.1126/scitranslmed.3005687

Lloyd-Price J, et al. (2016) The healthy human microbiome. Genome Med 8:51

Malfertheiner P, et al. (2009) Peptic ulcer disease. Lancet 374:1449-1461

Malfertheiner P, et al. (2014) *Helicobacter pylori*: perspectives and time trends. Nat Rev Gastroenterol Hepatol 11:628-638

Manor O, Borenstein E (2017) Systematic characterization and analysis of the taxonomic drivers of functional shifts in the human microbiome. Cell 21:254-267

Marshall BJ (1995) *Helicobacter pylori* in peptic ulcer: have Koch's postulates been fulfilled? Ann Med 27:565-568

Marshall BJ (2005) *Helicobacter* connections (Nobel lecture, December 8, 2005). Reprinted in ChemMedChem (2006) 1:783-802

Marshall BJ, et al. (1985) Attempt to fulfil Koch's postulates for pyloric *Campylobacter*. Med J Australia 142: 436-439

Marshall BJ, Warren JR (1984) Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. Lancet June 16:1311-1315

McColl KEL, et al. (1993) A study of the pathogenesis of *Helicobacter pylori* negative chronic duodenal ulceration. Gut 34: 762-768

Mayer EA, et al. (2014) Gut microbes and the brain: paradigm shift in neuroscience. J Neurosci 34:15490-15496

Mitchell SD (2000) Dimensions of scientific law. Philos Sci 67:242–265

Mithieux G (2017) Gut microbiota and host metabolism: what relationship? Neuroendocrinology 106:352-356

Moayyedi P, et al. (2015). Fecal microbiota transplantation induces remission in patients with active ulcerative colitis in a randomized controlled trial. Gastroenterology 149:102-109.e6

Momozawa Y, et al. (2011) Characterization of bacteria in biopsies of colon and stools by high throughput sequencing of the V2 region of bacterial 16S rRNA gene in human. PLoS One 6:e16952

Neville BA, et al. (2018) Commensal Koch's postulates: establishing causation in human microbiota research. Curr Opin Microbiol 42:47-52

Nguyen TLA, et al. (2015). How informative is the mouse for human gut microbiota research? Dis Model Mech 8:1–16

O'Malley MA, Skillings DJ (2018) Methodological strategies in microbiome research and their explanatory implications. Persp Sci 26:239-265

Parke EC, et al. (2018) A cautionary note for claims about the microbiome's impact on the "self." PLoS Biol 16(9):e2006654

Pocheville A, et al. (2017) Comparing causes – an information-theoretic approach to specificity, proportionality and stability. In Leitgeb H, et al. (eds), Proceedings of the 15th Congress of Logic, Methodology and Philosophy of Science (pp. 250–275). College Publications

Rees T, et al. (2018) How the microbiome challenges our concept of self. PLoS Biol 16(2):e2005358

Reich T, et al. (1975) The multifactorial model of disease transmission: I. Description of the model and its use in psychiatry. Brit J Psychiat 127:1-10

Ridaura VK, et al. (2013) Gut microbiota from twins discordant for obesity modulate metabolism in mice. Science 341:1241214

Ross LN, Woodward JF (2016) Koch's postulates: An interventionist perspective. Stud Hist Philos Biol Biomed Sci 59:35–46

Ross LN (2018) Causal selection and the pathway concept. Phil Sci 85: 551-572.

Round JL, Mazmanian SK (2009) The gut microbiota shapes intestinal immune responses during health and disease. Nat Rev Immunol 9:313–323

Sampson TR, Mazmanian SK (2015) Control of brain development, function, and behavior by the microbiome. Cell Host Microbe 17:565-576

Schwabe RF, Jobin C (2013) The microbiome and cancer. Nat Rev Genet 13:800-812

Sender R, et al. (2016) Are we really vastly outnumbered? Revisiting the ratio of bacterial to host cells in humans. Cell 164:337–340

Stein RR, et al. (2013). Ecological modeling from time-series inference: insight into dynamics and stability of intestinal microbiota. PLOS Comput Biol 9:e1003388.

Susser M (1991) What is a cause and how do we know one? A grammar for pragmatic epidemiology. Am J Epid 133.7:635-648

Sze MA, Schloss PD (2016) Looking for a signal in the noise: revisiting obesity and the microbiome. mBio 7:e01018-16

Thagard P (1998) Explaining disease: correlations, causes, and mechanisms. Minds Mach 8:61-78

Truax AD, et al. (2018) The inhibitory innate immune sensor NLRP12 maintains a threshold against obesity by regulating gut microbiota homeostasis. Cell Host Microbe 24:364-378

Turnbaugh PJ, et al. (2008) Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. Cell Host Microbe 3:213–223

Turnbaugh PJ, et al. (2009) A core gut microbiome in obese and lean twins. Nature 457:480–484

van Nood E, et al. (2013) Duodenal infusion of donor feces for recurrent *Clostridium difficile*. NEJM 368:407–415

van Zanten SJOV, Sherman PM (1994) *Helicobacter pylori* infection as a cause of gastritis, duodenal ulcer, gastric cancer and nonulcer dyspepsia: a systematic overview. Can Med Assoc J 150:177-185

Vonaesch P, et al. (2018) Pathogens, microbiome and the host: emergence of the ecological Koch's postulates. FEMS Microbiol Rev 42:273-292

Wang H, et al. (2016) Effect of probiotics on central nervous system functions in animals and humans: a systematic review. J Neurogastroenterol Motil 22:589–605

Woodward J (2003) Making things happen: A theory of causal explanation. OUP

Woodward J (2006) Some varieties of robustness. J Econ Methodol 13:219-240.

Woodward J (2010) Causation in biology: stability, specificity, and the choice of levels of explanation. Biol Philos 25:287–318

Yablo S (1992) Mental causation. Philos Rev 101:245–280

Zmora N, et al. (2018) Personalized gut mucosal colonization resistance to empiric probiotics is associated with unique host and microbiome features. Cell 174:1388-1405.e21