Review

CellPress

Discovering new indicators of fecal pollution

Sandra L. McLellan[1] and A. Murat Eren[2]

[1] School of Freshwater Sciences, University of Wisconsin–Milwaukee, Milwaukee, WI, USA
[2] Josephine Bay Paul Center, Marine Biological Laboratory, Woods Hole, MA, USA

Fecal pollution indicators are essential to identify and remediate contamination sources and protect public health. Historically, easily cultured facultative anaerobes such as fecal coliforms, *Escherichia coli*, or enterococci have been used but these indicators generally provide no information as to their source. More recently, molecular methods have targeted fecal anaerobes, which are much more abundant in humans and other mammals, and some strains appear to be associated with particular host sources. Next-generation sequencing and microbiome studies have created an unprecedented inventory of microbial communities associated with fecal sources, allowing reexamination of which taxonomic groups are best suited as informative indicators. The use of new computational methods, such as oligotyping coupled with well-established machine learning approaches, is providing new insights into patterns of host association. In this review we examine the basis for host-specificity and the rationale for using 16S rRNA gene targets for alternative indicators and highlight two taxonomic groups, *Bacteroidales* and *Lachnospiraceae*, which are rich in host-specific bacterial organisms. Finally, we discuss considerations for using alternative indicators for water quality assessments with a particular focus on detecting human sewage sources of contamination.

Sanitation, health, and rationale for alternative indicators

Fecal pollution carries a myriad of pathogens, and contamination of water is a global public health problem [1]. In developing countries, sanitation issues are overt, with 2.4 billion people, approximately 30% of the world population, lacking access to sewage disposal [2,3]. Urban areas can have inadequate sewage treatment infrastructure and in rural areas, residential sewage is routinely handled by piping it directly to rivers and streams that are also impacted by agricultural runoff, resulting in a mixture of human and animal sources [2]. In the USA fecal pollution of waterways is a subtle but persistent problem. More than 44% of the rivers and 30% of the bays and estuaries in the USA are deemed to be impaired, with pathogens often cited as the number one cause of these impairments [4]. Pathogens are not directly measured because there are numerous possible agents and the methods for their detection are time-consuming and expensive. Instead, pathogen presence is assumed based on detection of fecal indicator bacteria. Many rivers run through a combination of

Glossary

Alternative indictor: an organism or non-biological constituent of fecal pollution or sewage that is used to indicate the presence of fecal pollution. Constituents can range from commensal organisms found only in one type of host species to viruses, caffeine, or optical brighteners. These indicators are often used to complement traditional indicators including fecal coliforms, *Escherichia coli*, or enterococci in investigative studies. In this review we examine commensal organisms as indicators that can provide information on the host source.
Core community member: an organism that is found in all individuals of a host species or related group of host species (e.g., ruminants). Core community members are hypothesized to fulfill a crucial role in the host or be specifically adapted to the physiology of the host. A core member is not necessarily exclusive to the host species in which they are found, but may be found in other hosts as either an essential or transient member of the community.
Cosmopolitan: refers to the ubiquitous distribution of an organism among samples of different host species, where they occur in almost all individuals and hosts at varying relative abundances, but with no distinctive pattern.
Entropy: quantification of nucleotide variation at specific positions across sequences. High entropy at specific positions can correspond to a phylogenetic signal.
Genetic marker: a DNA sequence that can be used to track an organism or closely related subpopulation of organisms that share the marker. A genetic marker is the sequence that can be targeted by primers to provide an assay that is specific for a host source.
Host-associated: refers to the abundance pattern of an organism found in one host species, but not every individual of that host, and is absent in other host species.
Host-preferred: describes the abundance pattern of an organism that is dominant in all individuals of a host species and is low or absent in other host species.
Host-specific: may be used as a general reference to organisms that are indicative of host sources. In describing different abundance patterns of host microbial communities, we define host-specific more stringently – to mean strictly host-specific – where organisms are present in all individuals of a host species and absent in all other hosts.
Microbial source tracking: the process of determining the source of fecal bacteria found in contaminated environments. General indicators such as fecal coliforms, *E. coli*, and enterococci in most cases do not provide information as to the host in which they originated. Alternatives to these general indicators (i.e., alternative indicators) are based on commensal bacteria specifically associated with a host.
Oligotyping: a supervised computational method to analyze closely related sequences by only considering high-entropy positions for partitioning. Oligotypes can be formed using any number of nucleotide positions, but generally consist of 20–30 nts chosen from a sequence read of 60–250 nts depending on the sequencing platform. Oligotyping complements OTU clustering methods because it can resolve ecologically distinct organisms that may only vary by a single nucleotide in regions of 16S rRNA genes sequenced for microbial community analyses.
Operational taxonomic units (OTUs): taxonomic units typically defined by a sequence-similarity threshold. The most commonly used similarity threshold to cluster 16S rRNA gene sequences is 97%. This approach minimizes inflated diversity because of sequencing errors while preserving the resolution of closely related organisms.
Taxonomic unit: a general designation for different levels of resolution in classifying organisms (family, genus, species, strain, or operational taxonomic unit defined by sequence analysis). In this review we introduce oligotype as a taxonomic unit. The abbreviated term 'taxa' is often used.

Corresponding author: McLellan, S.L. (mclellan@uwm.edu).
Keywords: sewage; fecal indicators; alternative indicators; water quality; next-generation sequencing; microbial communities; oligotyping.

0966-842X/
© 2014 Elsevier Ltd. All rights reserved. http://dx.doi.org/10.1016/j.tim.2014.08.002

CrossMark

agricultural, suburban, and heavily urbanized areas before discharging to bays and estuaries; therefore, sources of fecal pollution are not easily assigned based on land use. Given the regulatory and public health implications of such assignments, empirical measurements of fecal pollution sources are needed.

The use of fecal indicators to detect fecal contamination has evolved over the past 100 years, but has primarily focused on coliforms, fecal coliforms, *Escherichia coli*, or enterococci [5]. These traditional indicators are commonly found in mammals and birds and continue to be used widely because detection methods are relatively fast, easy, and inexpensive. The advent of molecular methods allowed for non-cultured organisms to be used as 'alternative' fecal indicators (see Glossary). Until recently, only a few taxonomic groups such as *Bacteroidales* and *Bifodobacterium* have been explored. Next-generation sequencing technologies have given us an unprecedented inventory of the microbial community in a variety of environments. Before this, clone libraries only captured the most abundant community members, unless a large effort was undertaken [6]. Deep sequencing of the microbiome of humans and animals creates a new opportunity to explore a whole range of bacterial taxonomic groups suited for host-specific indicators. Comparison of microbial communities in humans and animal sources will not only validate the robustness of currently employed indicators but will also allow us to identify new human and animal fecal pollution indictors.

Development of alternative indicators

In an effort to create more informative fecal pollution indicators, several aspects need to be considered. What organisms should be targeted? How are organisms that are uniquely associated with a host source best distinguished and detected? Promising targets for these efforts are organisms that dominate the microbiome but are not easily cultured. Although functional genes may be responsible for the specialized activities of host-specific organisms, universal genes such as the 16S rRNA gene could be used to track these populations. Certain fecal pollution sources are a high priority for development of indicators. Discerning human sources (i.e., sewage) from other animal sources is important because of the implicit health risk posed by human sewage and the very different types of mitigation strategies needed to remediate sewage contamination compared with animal waste that is carried in surface runoff.

Fecal anaerobes as indicators

The intestinal tracts of humans and many animals are dominated by fecal anaerobes [6], making these organisms ideal targets for alternative indicators. By far the most explored taxonomic group is the order *Bacteroidales* [7–10], which are detailed later. Studies have also focused on Firmicutes [11,12], *Bifidobacteria* spp. [11,13–16], and *Methanobrevibacter smithii*, a common archaeon in the intestine [17,18]. Fecal anaerobes may be more indicative of the presence of pathogens because it is unlikely they will grow once released from their host into the environment, in contrast to *E. coli* and enterococci, which have been shown to persist and even grow in beach sand [19–22], algal mats [20,23], and sediment [24–26].

The 16S rRNA gene as a marker

The 16S rRNA gene has several advantages as a genetic marker for host fecal indicators including high-resolution, sensitivity and specificity, particularly when used as a combination of markers (Box 1). The use of multiple indicators and assessment of covariance patterns of host communities could untangle complex fecal pollution signals in watersheds that have multiple sources contributing to poor water quality [27,28]. Employing the 16S rRNA gene as a genetic marker lends itself to community approaches, where fecal pollution sources without a known marker can be characterized at minimum by their unique signature of 16S rRNA gene sequences [29].

Targets other than the 16S rRNA gene have been identified and have utility to detect specific sources. Subtractive hybridization or genomic enrichment have been used to find microbial sequences uniquely associated with a host source [30–32]. Identifying functional genes that are involved in microbial associations with their host is a promising approach, particularly when microbial populations that might differ between host sources cannot be distinguished by the 16S rRNA gene sequences [33]. Although functional gene markers are highly specific, they may occur at lower abundances than 16S rRNA targets, and therefore may not be as sensitive for detection purposes [34]. Nevertheless, they can be employed to detect a specific source of fecal pollution and complement 16S rRNA targets.

Detecting sewage in the environment

Humans and animals each have different pathogens uniquely associated with them; therefore, the relationship between fecal indicators and pathogens is highly dependent on the source of fecal pollution and the disease prevalence in that population [35]. One important division in assessing health risk is distinguishing between human

Box 1. Targets for alternative indicators: a case for 16S rRNA gene

The 16S rRNA gene has long been recognized as a chromatic clock, distinguishing between organisms from different lineages [109]. Recent studies have shown that high-resolution analysis of 16S rRNA gene sequences in closely related organisms reveals host-related patterns (see Oligotyping section). Employing the 16S rRNA gene as a genetic marker for alternative fecal indicators has several advantages:

- The 16S rRNA gene is the gold standard for defining bacterial community structure in the host and physical environment, and is widely used in research studies.
- It is universally present in bacterial genomes and contains hypervariable regions that can distinguish between closely related organisms.
- The 16S rRNA gene usually occurs in multiple copies and therefore is potentially a more a sensitive target than single-copy genes for detecting fecal bacteria in the environment.
- Differential abundance and covariance patterns across hosts can guide accurate classification of host sources when using the 16S rRNA gene to define community structure even when there are no host-specific organisms.
- Multiple levels of resolution, from phylum-level groupings to single-nucleotide variation, can be employed to examine distinguishing patterns among the microbial community in different hosts.
- The presence of highly abundant and diverse targets lends itself to innovative detection strategies.

Download English Version:

https://daneshyari.com/en/article/3422030

Download Persian Version:

https://daneshyari.com/article/3422030

Daneshyari.com