Aquaculture 448 (2015) 464–475

Contents lists available at ScienceDirect

Aquaculture

journal homepage: www.elsevier.com/locate/aqua-online

A new view of the fish gut microbiome: Advances from next-generation sequencing

Mahdi Ghanbari [a,b,*], Wolfgang Kneifel [a], Konrad J. Domig [a]

[a] BOKU — University of Natural Resources and Life Sciences, Department of Food Science and Technology, Institute of Food Science, Muthgasse 18, A-1190 Vienna, Austria
[b] University of Zabol, Faculty of Natural Resources, Department of Fisheries, Zabol, Iran

ARTICLE INFO

Article history:
Received 29 September 2014
Received in revised form 17 June 2015
Accepted 18 June 2015
Available online xxxx

Keywords:
Fish microbiota
Next-generation sequencing
Gut microbiota
Bacterial community membership
Diet

ABSTRACT

The fish gut microbiota contributes to digestion and can affect the nutrition, growth, reproduction, overall population dynamics and vulnerability of the host fish to disease; therefore, this microbial community is highly relevant for aquaculture practice. Recent advances in DNA sequencing technologies and bioinformatic analysis have allowed us to develop a broader understanding of the complex microbial communities associated with various habitats, including the fish gut microbiota. These recent advances have substantially improved our knowledge of bacterial community profiles in the fish intestinal microbiota in response to a variety of factors affecting the host, including variations in temperature, salinity, developmental stage, digestive physiology and feeding strategy. The goal of this review is to highlight the potential of next-generation sequencing platforms for analysing fish gut microbiota. Recent and promising results in this field are presented along with a focus on new perspectives and future research directions of fish gut microbial ecology.

Contents

1. Introduction

In-depth knowledge of community membership as well as the structure and relationships between resident microbes (microbiota) and their hosts can provide insight into both the function and dysfunction of the host organism. Comprising a large diversity of approximately 28,000 species, fish make up nearly half of all vertebrate species and exhibit a wide variety of physiologies, ecologies and natural histories (Wong and Rawls, 2012). Thus, fish represent an important group for understanding the variety and nature of symbioses in vertebrate gut microbial communities (Nayak, 2010). The digestive tract of fish receives water and food that are populated with microorganisms from the surrounding environment; these microbes undoubtedly affect the

* Corresponding author at: BOKU — University of Natural Resources and Life Sciences, Department of Food Science and Technology, Institute of Food Science, Muthgasse 18, A-1190 Vienna, Austria.
E-mail addresses: m.ghanbari@boku.ac.at, ghanbari.msc@gmail.com (M. Ghanbari).

http://dx.doi.org/10.1016/j.aquaculture.2015.06.033
0044-8486/

resident microbiota. Correlations between changes in the composition and activity of the fish gut microbiota with fish physiology and disease have been proposed, increasing the scientific community's interest in this field of research. From this perspective, a comprehensive and detailed view of the fish gut microbiota, including phylogenetic composition as well as the genetic and metabolic potential, is essential to understand the dynamics and possible mechanisms of the cause/effect relationships between gut microbiota and physiology (Austin, 2006, 2011). As numerous studies have indicated that culture-dependent techniques possess dubious sensitivity and often detect only a limited fraction of microbial communities (Austin, 2006; Kim et al., 2007; Namba et al., 2007; Wu et al., 2010, 2012a; Lan and Love, 2012; Larsen et al., 2013), several methods for culture-independent microbial analysis have been developed within the past few decades (Head et al., 1998; Muyzer and Smalla, 1998) and applied in studies of the fish gut. Molecular methods provide faster results and novel, high-resolution insights into the structure and diversity of microbial communities within the digestive tracts of freshwater and marine fish (Austin, 2006; Kim et al., 2007; Namba et al., 2007; Wu et al., 2010, 2012a; Lan and Love, 2012; Larsen et al., 2013).

Despite these recent findings, the cost, technical difficulties and low coverage associated with Sanger sequencing have limited the ability to analyse a large number of samples. Recently, however, rapid and low-cost approaches for next-generation sequencing (NGS) technologies have been introduced to study the composition and genetic potential of densely populated microbial communities such as gut microbiota (Foster et al., 2012). Within the past few years, these techniques have been applied to analyse the composition and functional properties of fish microbial communities. Fish microbiota studies have most frequently utilised the 454/Roche pyrosequencing (e.g., the Roche 454 FLX Titanium and FLX+) and Illumina technologies (e.g., the Illumina MiSeq and HiSeq 2000). Unlike the Sanger sequencing approach, these NGS platforms provide a larger number of reads in a single run, enabling the rapid and cost-effective acquisition of in-depth and accurate sequence data and allowing for the detection of both dominant and low-abundance (rare) microbial community members (Roeselers et al., 2011; Wu et al., 2012b; Star et al., 2013; Wong et al., 2013). With the emergence and availability of NGS techniques for studying complex microbial ecosystems and the growing appreciation of the importance of the indigenous microbiota of fish, this review highlights the potential of NGS platforms for the analysis of fish gut microbial ecology.

2. NGS platforms and technologies

Although capillary sequencers such as the ABI 3730 (Applied Biosystems, USA) continue to be used for small-scale sequencing (Metzker, 2010; Zhou et al., 2010) in fish microbiota studies, there is an increasing trend towards the use of NGS platforms. While there are several different NGS technologies, one benefit common to all of them is the ability to rapidly generate large volumes of sequencing data in parallel. Table 1 compares the main characteristics of NGS sequencers to emphasise the similarities and differences among the platforms. An overview of the NGS technologies that have been used for microbiome studies is given in the following sections. For a more in-depth discussion about the technical aspects of next-generation sequencing platforms and principles, we refer to comprehensive reviews in the literature (Mardis, 2008; Kircher and Kelso, 2010; Metzker, 2010; Foster et al., 2012; Highlander, 2012; Kuczynski et al., 2012; Liu et al., 2012; Di Bella et al., 2013; Hui, 2014; Hodkinson and Grice, 2015).

2.1. The 454 platform family

Developed by 454 Life Sciences, the 454 Roche platforms based on sequencing-by-synthesis with pyrophosphate chemistry (Mardis, 2008; Metzker, 2010) are among the most commonly used NGS techniques for studying fish-associated microbiota. These platforms use "pyrosequencing" chemistry, which involves the incorporation of deoxynucleotide triphosphate (dNTPs) bases into a synthesised DNA chain followed by the release of a pyrophosphate. The pyrophosphate subsequently serves as a substrate for the enzymatic production of ATP. In the presence of luciferase, the ATP produces a quantifiable amount of light, which is then detected by a camera. This reaction is performed on beads containing millions of copies of a single DNA molecule. For DNA amplification, the single DNA molecules are clonally amplified in an oil–water emulsion containing PCR reagents in micelles occupied by only one bead, a process known as emPCR (Dressman et al., 2003). After harvesting, beads carrying single-stranded DNA are placed into the wells of a picotiterplate, which contains millions of wells, in such a way that each well is occupied by a single bead and the enzymatic reagents. Sequencing occurs through the repeated cyclic flow of thin films of dNTPs across the wells. Base incorporation leads to the production of photons that are detected by a camera every 7 s. Each cycle contains a different dNTP, and an image is captured after each cycle to measure the light produced in each well. Thus, sequential collections of images are analysed for light intensity patterns. The amount of light discloses whether a specific dNTP in each flow was incorporated as well as how many dNTPs were incorporated in that flow when homopolymer runs are present. This information is then translated into a DNA sequence for each bead (Ahmadian et al., 2006; Siqueira et al., 2012). The massively parallel 454 platform family has been further improved in recent years to optimise sequence read length and the amount of sequence data obtained (modal read length 750 bp and average read length 700 bp using FLX+ chemistry; see Table 1), resulting in an increase in the performance at different phylogenetic depths. However, the higher costs per base as well as the total output per run of 454 platforms in terms of reads (1 million) and bases (700 Mb) are clearly limiting factors in projects in which coverage is more important than read length, though it is valid for some metagenomic or amplicon sequencing projects.

2.2. Illumina sequencing platforms

The Illumina sequencing platforms were introduced in 2006 and quickly embraced by many researchers because a larger amount of data could be generated in a more cost-effective manner. Major progress in the Illumina platforms has been made in recent years with regard to sequence read length and output (number of reads per run) by technically improving the instruments, the chemistry and the base-calling algorithms. The working principle of this platform is based on sequencing-by-synthesis chemistry and, like the Sanger method, relies on the incorporation of dye terminator nucleotides into the sequence by a DNA polymerase (Siqueira et al., 2012). In this approach, DNA fragments are immobilised on a flow cell surface, which is coated with adapters and complementary adapters. Each of these single DNA fragments then creates a bridge with the complementary adapters. DNA sequencing is initiated by the addition of the reaction mixture containing DNA polymerase, sequencing primers and four reversible terminator nucleotides, each labelled with a different fluorescent dye. After incorporation, the terminator nucleotide and its position on the support surface are detected and recorded by a four-channel fluorescence scanner (Metzker, 2010).

Among the Illumina platforms, HiSeq sequencers produce the largest amount of data per run (up to 1500 Gb) at the lowest cost per base and with the shortest reads (Table 1). The Illumina MiSeq platform (introduced in 2011) generates 44–50 million paired-end reads 2 × 300 bp in length using the new V3 chemistry. The costs per sequenced base are higher compared with the HiSeq instrument; however, the longer read length in combination with the lower read number can be of particular interest for amplicon sequencing projects. The Illumina NextSeq 500 platform is another recently released product that performs at an