



The Promise of Whole Genome Pathogen Sequencing for the Molecular Epidemiology of Emerging Aquaculture Pathogens

Sion C. Bayliss¹, David W. Verner-Jeffreys², Kerry L. Bartie³, David M. Aanensen^{4,5}, Samuel K. Sheppard¹, Alexandra Adams³ and Edward J. Feil^{1*}

¹ The Milner Centre for Evolution, Department of Biology and Biochemistry, University of Bath, Bath, UK, ² Cefas Weymouth Laboratory, Weymouth, UK, ³ Institute of Aquaculture, University of Stirling, Stirling, UK, ⁴ Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, UK, ⁵ The Centre for Genomic Pathogen Surveillance, Wellcome Genome Campus, Cambridge, UK

OPEN ACCESS

Edited by:

Marcelino T. Suzuki,
Sorbonne Universities (UPMC)
and Centre National de la Recherche
Scientifique (CNRS), France

Reviewed by:

Torsten Thomas,
University of New South Wales,
Australia

Andrew C. Barnes,
The University of Queensland,
Australia

*Correspondence:

Edward J. Feil
e.feil@bath.ac.uk

Specialty section:

This article was submitted to
Aquatic Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 19 October 2016

Accepted: 17 January 2017

Published: 03 February 2017

Citation:

Bayliss SC, Verner-Jeffreys DW,
Bartie KL, Aanensen DM,
Sheppard SK, Adams A and Feil EJ
(2017) The Promise of Whole
Genome Pathogen Sequencing
for the Molecular Epidemiology
of Emerging Aquaculture Pathogens.
Front. Microbiol. 8:121.
doi: 10.3389/fmicb.2017.00121

Aquaculture is the fastest growing food-producing sector, and the sustainability of this industry is critical both for global food security and economic welfare. The management of infectious disease represents a key challenge. Here, we discuss the opportunities afforded by whole genome sequencing of bacterial and viral pathogens of aquaculture to mitigate disease emergence and spread. We outline, by way of comparison, how sequencing technology is transforming the molecular epidemiology of pathogens of public health importance, emphasizing the importance of community-oriented databases and analysis tools.

Keywords: aquaculture, genomics, whole genome sequencing, molecular epidemiology, infectious disease, pathogens, bacteria, viruses

THE GLOBAL EXPANSION OF INTENSIVE AQUACULTURE

Aquaculture broadly refers to the farming of aquatic organisms, including finfish, shellfish, crustaceans and plants, primarily for human consumption, and has been practiced since antiquity. Multiple ancient civilisations independently developed aquaculture systems; 4000 year old bas reliefs show Egyptians fishing for tilapia in artificial ponds, a number of freshwater species were farmed and perhaps domesticated in ancient China and other parts of Asia, ancient Greeks and Romans farmed bivalves and fish in freshwater and saltwater ponds and pre-Columbian systems of weirs in the Amazon are thought to have been constructed for fish cultivation (Costa-Pierce, 2008). However, it was not until the mid-20th century that intensive, industrial-scale aquaculture became viable, and recent decades have witnessed a dramatic expansion such that it is currently the fastest growing food producing sector globally. This expansion has been driven by stresses on terrestrial systems owing to population increase and scarcity of resources, combined with a plateauing of production from capture fisheries world-wide (Garcia and Rosenberg, 2010).

In a world where more than 800 million people continue to suffer from chronic malnourishment (FAO, 2015), and where the global population is expected to grow by another 2 billion to reach 9.6 billion people by 2050, with a concentration in coastal urban areas (UN, 2015), aquaculture plays an increasingly critical role in global food security. In 2013, finfish, molluscs, and crustaceans represented about 17% of the animal

protein supply and 6.5% of all protein for human consumption. Finfish alone provided more than 3.1 billion people with almost 20% of their animal protein (FAO, 2014). A milestone was reached in 2014, when for the first time, aquaculture accounted for the majority of fish consumed globally (FAO, 2014). Four decades ago, aquaculture accounted for only 7% of all fish consumed globally, with the vast majority being produced by the capture sector. This increase in aquaculture production has easily outpaced population growth; *per capita* fish supply in 2014 reached a global average of 20 kg, double the global average in the 1960s, with the gap narrowing between developed and developing nations (FAO, 2016). Much of this growth has occurred in Asia, with China alone accounting for 60% of global production, with other major producers being India, Vietnam, Bangladesh, Thailand, and Egypt (FAO, 2016). Ninety-four *per cent* of the 18 million people employed in fish farming in 2014 were Asian, and of the top 10 aquaculture countries, in terms of production, only Norway is classified as high income. Thus the increasing economic reliance on aquaculture is heavily skewed toward developing nations. Total aquaculture animal production in 2014 amounted to 73.8 million tons, worth approximately US\$160.2 billion, of which finfish accounted for around 50 million tons (US\$100 billion) (FAO, 2016).

The intensification of aquaculture production (the so-called “Blue Revolution”) has some parallels to the agricultural “Green Revolution.” However, there is one important difference. Whereas in agriculture there are a limited number of domesticated species, the dramatic increase in scale of aquaculture production has been accompanied by a commensurate diversification of target species. In 2014 a total of 580 species were farmed around the world, including 362 finfishes, 104 molluscs, and 62 crustaceans (FAO, 2016). Moreover, secondary species are now also being cultivated to improve production. For example, cleaner fish, such as wrasse and lumpfish, are now being farmed as biological control agents to manage salmon lice infestation (Leclercq et al., 2014). This diversification in target species has been accompanied by a very wide range of different farming techniques and husbandry regimes. Farms can be based on fresh, brackish or salt water within cages, earth ponds, tanks, open sea, rafts or raceways, in mono- or mixed culture, from tropical climates to the sub-Arctic (Baluyut, 1989).

Maintaining the steep trajectory of global aquaculture production requires addressing a critical trade-off between intensification and sustainability. The challenges to overcome are many and various, and include: (i) increasingly severe competition with other resource (land/water/feed) users, (ii) deteriorating quality of water supplies resulting from aquatic pollution, (iii) successful integration of aquaculture with other farming activities, and (iv) improvements in environmental management including reduction of environmental impacts and avoidance of risks to biodiversity. However, arguably the biggest threat arising as a consequence of the intensification and globalization of the aquaculture is from infectious disease.

EMERGENT INFECTIOUS DISEASE IN AQUACULTURE

Non-communicable disease or sporadic endemic infections have been described in aquatic animals for centuries. However, the development of large-scale intensive aquaculture has been associated with a dramatic increase in severe disease outbreaks caused by a diverse range of pathogens, including parasites, viruses and bacteria (Yin, 2004; Crane and Hyatt, 2011; Woo et al., 2011). These disease outbreaks can be catastrophic for the industry. The Chilean Infectious Salmon Anemia (ISA) outbreak in 2007 bankrupted the industry and left debts of US\$1.8 billion (Alvial et al., 2012). Low income rural communities were particularly badly affected, and an estimated 13,000 jobs were lost. Shellfish can be just as badly affected as finfish. Highly virulent shrimp viral diseases have arisen as a result of selection pressure in shrimp farming, subsequently disseminated via transfer of broodstock and larvae (Flegel, 2012). Bacterial diseases can be just as devastating. Between 2010 and 2014 acute hepatopancreatic necrosis disease (AHPND) was estimated to have caused a US\$1 billion annual loss to the shrimp farming industry (FAO, 2013). The impact of the disease varied between countries, although an extreme example would be the drop in Thailand’s share of the global shrimp market from 40% in 2011 to 10% in 2012. This disease, originally named early mortality syndrome (EMS), was first identified in China in 2009 and spread successively to Vietnam, Malaysia, Thailand and Mexico by 2013 (Tran et al., 2013; Joshi et al., 2014; Nunan et al., 2014; Soto-Rodriguez et al., 2015). In 2015 the putative causative agent of AHPND was identified as a strain, or strains, of the bacterium *Vibrio parahaemolyticus* which had acquired a toxin producing plasmid (Lee et al., 2015). In the intervening period between emergence of the disease and identification of the pathogen AHPND had spread almost globally to nations engaged in shrimp aquaculture (FAO, 2013).

Understanding the conditions underpinning the emergence and spread of infectious disease is thus a major challenge for developing a sustainable global aquaculture industry. Pathogen emergence can result through the evolution of novel strains or the international spread and expansion of previously characterized pathogens. New strains causing a novel pathology can potentially emerge through host-switching as a result of intensive mixed farming (Marcos-López and Gale, 2010; Marcos-López et al., 2010), or alternatively new pathogen strains might evolve from a background of commensal organisms. This can occur as a consequence of mutation or the horizontal acquisition of virulence genes via recombination between previously isolated pathogen populations (Wijegoonawardane et al., 2009). Emergent pathogens that have spread rapidly on international scales, termed ‘transboundary’ pathogens, proliferate either through trade of broodstock, fry, fingerlings or mature animals, or as a response to changing environmental conditions. There also are grounds for arguing that certain husbandry practices, including those designed to mitigate the impact of infectious disease, might inadvertently be driving the evolution of increasingly virulent strains (Kennedy et al., 2016). Key husbandry factors impacting on the spread of existing

disease, and the emergence of new infectious agents, include stocking densities, compression of the rearing cycle, the use of chemotherapy and a reduction in host genetic diversity (Kennedy et al., 2016). Mixed farms, in which more than one species are farmed together, can increase the potential for pathogen host switches, which is often associated with heightened virulence in the new host. A recent study on the bacterial pathogen *Flavobacterium columnare*, a freshwater pathogen of rainbow trout, indicated that a subset of virulent and competitive strains were being selected in the farm environment (Pulkkinen et al., 2010; Sundberg et al., 2016).

Climate change is also likely to play an increasing role in the emergence of infectious disease in the aquatic environment. Vezzulli et al. (2016) have recently shown a concordance between rising sea surface temperature (SST), the abundance of *Vibrio* bacteria associated with plankton samples and the incidence of human infection caused by these bacteria. In aquaculture settings, the outcome of exposure to fish pathogens is often critically dependent upon both water quality, in particular the concentration of dissolved oxygen, and water temperature. Water temperature impacts both on the replication rate of the pathogen and, being poikilothermic animals, on the immune response of the fish. Whilst some bacterial pathogens, such as *Flavobacterium psychrophilum*, are more virulent at cold temperatures (10°C) most, including *Lactococcus garvieae*, are much more virulent above 15°C. Salmonids are particularly vulnerable to relatively narrow change in temperature with respect to disease susceptibility. This temperature sensitivity has potentially grave implications for disease emergence and spread within the context of a changing climate (Marcos-López and Gale, 2010).

Intensive aquaculture also has the potential to increase the infection load in wild stocks through pathogen 'spillover' (Power and Mitchell, 2004). Although it is usually very difficult to gauge the severity of this threat, farming of species outside of their natural geographical range is likely to increase the risks due to the introduction of exotic pathogens into susceptible host populations. A controversial example is that of the impact of sea lice spillover from farmed salmon to wild stocks, with some studies pointing to a correlation between lice abundance on farms and the productivity of wild stocks (Krkosek et al., 2011) whilst other studies suggest that spillover of lice alone does not explain increased mortality of natural populations (Marty et al., 2010). The impact and occurrence of pathogen spillover will vary between pathogens, farming systems and local environment. Whilst disease spillover may not impact directly on aquaculture production the negative impact on public perception and the downstream impact on the valuation of farmed aquatic animals should not be underestimated (Hansen and Onozaka, 2011).

Although the precise drivers for infectious disease emergence in aquaculture settings typically remain unclear, it is evident in many cases that strains vary in their virulence and that specific pathogen variants have adapted to the aquaculture settings. This observation has many parallels in agriculture, where certain strains predominate within intensive farming environments, and also in public health where strains of human pathogens appear adapted to the healthcare setting by virtue

of multidrug resistance. However, the degree to which disease emergence is accompanied by strong selection imposed by the farmed environment, compared to 'opportunistic' infections from environmental reservoirs, depends upon the specific ecology, disease dynamics and etiology.

We discuss some examples of pathogenic Gram-negative (Box 1), Gram-positive (Box 2) bacteria and viruses (Box 3) that are of key importance to aquaculture, with a focus on evidence from molecular typing data. These examples discussed largely affect finfish, as there is typically less detailed evidence available for shellfish pathogens, although we have included a recent study on *Vibrio aestuarianus*, an emerging pathogen of oysters (Goudenège et al., 2015).

THE PROMISE AND CHALLENGES OF WHOLE GENOME SEQUENCE DATA FOR MOLECULAR EPIDEMIOLOGY

Molecular epidemiology encompasses a synthesis between molecular biology, typing, evolutionary biology, population biology and epidemiology (Galvani, 2003). As discussed below, whole genome sequencing represents the 'ultimate' typing methodology in terms of discriminatory power, and the advantages and opportunities for infectious disease management afforded by the advances in sequencing technology are numerous and profound. Moreover, much of the methodology developed for pathogens that infect humans, including bioinformatic pipelines, data archiving and management and statistical approaches to analyzing the data, are directly applicable to aquaculture settings.

Before discussing in more detail the significant benefits from adopting whole genome pathogen sequencing for aquaculture disease management, there are two key caveats to be highlighted. First, it is critical to appreciate that any molecular typing dataset, including whole genome sequencing, is always most productively considered alongside rich metadata pertaining to the epidemiological source and clinical phenotypes of the isolates. Sequence data alone will rarely inform on novel disease management strategies. Second, the contribution sequencing technology will make toward developing a sustainable future for the global aquaculture industry is ultimately dependant on the degree to which these data, and associated metadata, can be freely archived, accessed, and interrogated. As discussed below, the technical ability, and political will, to share resources and combine data from international sources is critical for efficient global surveillance, for providing early warnings of new emerging pathogens, and for fully exploiting the data for novel diagnostics and vaccine development. There are, of course, considerable challenges with respect to agreeing common standards and resolving issues relating to data ownership and commercial sensitivities.

Molecular typing methodologies have long been deployed to understand and monitor the emergence and spread of bacterial pathogens of public health and veterinary relevance. A key advance in the late 1990s was the development of multilocus sequence typing (MLST) which combined advances

BOX 1 | Gram-negative bacteria.

Many different Gram-negative genera are known to cause disease in finfish and shellfish, including *Aeromonas*, *Edwardsiella*, *Flavobacterium*, *Francisella*, *Photobacterium*, *Piscirickettsia*, *Pseudomonas*, *Tenacibaculum*, *Vibrio*, *Weisella*, and *Yersinia* (Austin and Austin, 2012). Disease monitoring and management is often complicated by the ability of bacterial pathogens to persist in the water column, or to asymptomatically colonize farmed species as a constituent of the 'healthy' microbiome. Here, we focus on the molecular epidemiology of four exemplar species: *Y. ruckeri*, *F. psychrophilum*, *A. salmonicida*, and *V. aestuarianus*. The former three species cause disease within stressed, high density populations of farmed finfish, whilst *V. aestuarianus* causes a serious disease of oysters. Management strategies tend to focus on both mitigating the spread of disease through trade and other human activities, and upon understanding and avoiding the conditions that trigger the transition from a commensal lifestyle to serious pathogen.

Yersinia ruckeri is a rod-shaped bacterium, endemic in cool waters, and is the etiological agent of enteric redmouth disease (ERM). This pathogen affects many finfish species, particularly salmonids, and causes considerable economic losses worldwide. It has been recovered from faces, water and sewage sludge (Willumsen, 1989) and is known to readily form biofilms in the aquatic environment, which helps to account for recurrent infections in trout farms (Coquet et al., 2002). The disease may have been exported on multiple occasions from the USA via the movement of rainbow trout eggs and fry (Austin and Austin, 2012), and an isolate from Europe was first described in 1981 (Lesel et al., 1983). Whilst it is thus possible that the spread of the disease was as a direct consequence of trade and the expansion of aquaculture, it is also possible that *Y. ruckeri* is ubiquitous in aquatic environments worldwide and that local outbreaks result from endemic strains adapting the farm environment (Willumsen, 1989).

Some of the earliest fish vaccines targeted *Y. ruckeri*. Disease in rainbow trout is largely caused by serotype O1 (serovar 1) (Kumar et al., 2015). These strains are related to the motile O1 'Hagerman' strains, dating back to the first isolation of *Y. ruckeri* in the 1950s in the Hagerman valley, Idaho, and which have had widespread use in immersion vaccination (Wheeler et al., 2009; Barnes, 2011). The dominance of this clone as a source of infection in intensively farmed rainbow trout is suggestive of either specific adaptations to the host or to the conditions associated with the farm environment.

Although the O1 vaccine was initially very effective, vaccine failures in rainbow trout were reported from the early 1990s (Austin et al., 2003). Welch et al. (2011) characterized a variant serotype O1 strain, biotype 2 (BT2) that was causing outbreaks worldwide even in vaccinated fish. The key distinguishing phenotypes of BT2, which are likely to be linked to the ability to escape the vaccine, are the loss of motility due to the absence of peritrichous flagella and a lack of secreted phospholipase. These phenotypes arise from a very small set of mutations that were found to have occurred independently on at least four occasions.

Early work suggested that the virulence of the *Y. ruckeri* serogroup O1 BT2 clone in trout is linked to an enhanced ability to survive in serum (Davies, 1991b). It has since been shown that the O1 serotype is specifically virulent in rainbow trout, whereas Atlantic salmon are typically more susceptible to infection by *Y. ruckeri*, and a greater diversity of strains are virulent in this host. This was recently confirmed by the phenotypic characterization of 109 *Y. ruckeri* isolates from Atlantic salmon and 26 isolates from rainbow trout (Ormsby et al., 2016). Whereas the majority of the strains isolated from rainbow trout corresponded to the common serogroup O1 BT2 clone, the Atlantic salmon isolates corresponded to 19 different clones. However, all the Atlantic salmon strains in this study corresponded to BT1. A separate study observed that BT2 has arisen independently in Atlantic salmon in Tasmania (Barnes et al., 2016). In this study *Y. ruckeri* isolates were sequenced from a collection that spanned 27 years, covering the entire period of time over which the salmonid farming industry has become established in Tasmania. BT2 was only observed intermittently between 2004 and 2007 and has not been detected since. This may be due in part to the ongoing reduction in motility of BT1 that infected salmon during this time period, which may have allowed BT1 to maintain a selective advantage over BT2. In addition to O-antigen serotyping (Romalde et al., 1993), *Y. ruckeri* populations have been characterized using outer membrane protein (OMP) analysis (Davies, 1991a) and more recently using MLST (Bastardo et al., 2012, 2015). These data shed light on the global biogeography of this species, as well as providing inferences of recent and ancient demographic history. More specifically, the MLST data provided evidence three distinct clades representing isolates from North America, South America, and Europe. The population structure reflects a complex combination of ancient range expansion and more recent migration due to human activity (Bastardo et al., 2015).

Aeromonas salmonicida subsp. *salmonicida* is a non-motile psychrophilic bacteria that causes furunculosis, so called due to the boils (or furuncles) that develop on the skin or musculature typical of chronic disease in older fish. The disease can spread either through direct contact with infected fish or via contaminated water. The acute form of the disease, which more commonly affects younger fish, leads to rapid septicaemia and death within 2–3 days (Dallaire-Dufresne et al., 2014). This is one of the most costly and well-studied diseases of salmonids, particularly Atlantic salmon, and affects both farmed and wild stocks (Lillehaug et al., 2003). Reports of furunculosis epidemics in wild trout in North America and Europe date back to the late 19th century, with major epidemics in wild stocks in the UK and Denmark during the 20th century (Bakke and Harris, 1998). Outbreaks in farmed marine salmon in Scotland and Norway began in the mid-1980s. Although the spread of the disease has been mitigated by the deployment of vaccines since the 1990s, and have been particularly successful for Atlantic salmon, these vaccines are expensive and can cause serious side effects.

A. salmonicida subsp. *salmonicida* is one of five recognized *A. salmonicida* sub-species. However, this scheme is far from comprehensive, and any strains identified as *A. salmonicida* but which do not fall within these sub-species are loosely assigned as "atypical." Atypical strains can also cause ulcerative disease in salmonids and non-salmonids including cod, halibut, wolffish (Dallaire-Dufresne et al., 2014). In particular, atypical *A. salmonicida* can cause furunculosis in wrasse and lumpsuckers (Collins et al., 1991). These 'cleaner fish' are of increasing economic importance due their use as biological control agents for salmon lice infestation and efforts are currently underway both in Scotland and Norway to develop robust farming methods for these species.

A. salmonicida possess an impressive armory of virulence factors, including proteases, lipases and other extracellular molecules as well as a number of toxins (effectors) secreted directly in the host cytoplasm via a plasmid-encoded type three secretion system which is essential for infection in fish. The *A. salmonicida* cell is also covered in a tetragonal protein (encoded by the *vapA* gene) which associates with lipopolysaccharide to form a hydrophobic "A-layer" that protects the cell from macrophages and is responsible for autoagglutination. Gulla et al. (2016) have recently characterized the *vapA* gene from a large collection of typical and atypical *A. salmonicida*. They noted 14 discrete *vapA* sequence clusters that correspond to the known sub-species but also, to a large degree, to different host species, thus pointing to a need for species specific vaccines based on different A-layer types. Although multilocus sequence analysis (MLSA) has been used to study the taxonomy and diversity within the *Aeromonas* genus (Roger et al., 2012; Colston et al., 2014), MLST has not been specifically applied to *A. salmonicida* epidemiology.

Flavobacterium psychrophilum is a psychrotrophic species and the causative agent of rainbow trout fry syndrome (RTFS) and bacterial cold water disease (BCWD) in salmonids (Austin and Austin, 2012). Rainbow trout and coho salmon are particularly susceptible to the disease, and extensive losses have also been reported in farmed and wild ayu (*Plecoglossus altivelis altivelis*) (Wakabayashi et al., 1991). No broad spectrum vaccine is available, and disease management requires close surveillance and the use of antimicrobials. *F. psychrophilum* is endemic in the environment, and can transmit both horizontally and probably also vertically. As with *Y. ruckeri* and other bacterial pathogens, a key question relating to the epidemiology of this species is the contribution of sporadic infection from endemic environmental lineages, compared to the spread via trade of farm-adapted lineages.

Arai et al. (2007) used PFGE to characterize 64 isolates of *F. psychrophilum* from ayu and other fish hosts in Japan, and noted evidence for association between PFGE and host species. Chen et al. (2008) similarly used PFGE to characterise 139 isolates from farmed rainbow trout and wild (spawning) coho salmon mostly from

(Continued)

BOX 1 | Continued

Idaho in the USA. Whilst they noted that samples from both species were characterized by the presence of a small number of predominant PFGE clusters, the isolates from spawning coho salmon were markedly more diverse than those from farmed trout, consistent with transmission within and between farms of specific clones that may have adapted to the farm environment.

An MLST scheme for this pathogen was proposed in 2008 using 11 loci and based on the full genome sequence of the reference strain JIP02-86, which revealed very high rates of recombination (Nicolas et al., 2008). A subsequent MLST study on isolates from Swiss fish farms was published, and these data did not confirm host associated with specific clonal groups (Strepparava et al., 2013). Nilsen et al. (2014) reported a very large MLST study of 560 isolates from four Nordic countries; Norway, Finland, Denmark, and Sweden. This diverse isolate collection was recovered from different hosts and environmental sources between 1983 and 2012. Although the population as a whole was diverse, representing 81 sequence types (STs) and 12 clonal complexes (CC's; clusters of closely related STs), the data confirmed the presence of a dominant clone, ST2 (CC10), which represented 65% of all the isolates. This confirmed other MLST studies using the same scheme that also reported the predominance of ST2 in France (Siekoula-Nguedia et al., 2012) and Switzerland (Strepparava et al., 2013). In Norway this clone is particularly monomorphic, suggesting a recent introduction into this country. ST2 is strongly associated with rainbow trout, and appears to have disseminated rapidly between rainbow trout farms due to some selective advantage, possibly decreased sensitivity to quinolones (Shah et al., 2012). Isolates from other salmonid hosts and from the environment are more diverse and freely recombining. The combination of a freely recombining, diverse 'background' population and a small number of predominant widespread clones conforms to an 'epidemic' population structure as originally proposed for the human pathogen *Neisseria meningitidis* (Smith et al., 1993).

Vibrio aestuarianus has recently been causing commercially damaging disease in farmed oyster stocks, particularly in France. Goudenège et al. (2015) used whole genome sequencing to characterize 14 *V. aestuarianus* strains isolated both before and during the French outbreak, which began in 2012. They noted two distinct clades, but as both clades contained isolates that pre-date the outbreak there was no evidence of the emergence of a single virulent clade that could explain the increase in oyster mortalities. However, this study demonstrated a key role in virulence of a single regulatory gene, *varS*. This gene, a recognized virulence factor in other Gram-negative bacteria, encodes a signal transduction histidine-protein kinase, which regulates the expression of a secreted metalloprotease. The authors noted that a *varS* had undergone a frameshift in one non-virulent strain, leading to reduced expression of the metalloprotease. The authors went on to provide experimental evidence that the *varS* gene is indeed an important virulence factor of *V. aestuarianus* in oysters. This study provides an excellent example of how whole genome sequencing, combined with experimental data, can provide powerful evidence concerning virulence and pathogenicity in shellfish pathogens.

in automated sequencing technology with the rapid expansion of the internet (Maiden et al., 1998; Pérez-Losada et al., 2013). The ability to access data via a website provided the means, for the first time, for near real-time molecular epidemiological surveillance on a global scale (Urwin and Maiden, 2003). Crucially, the establishment of MLST for major pathogenic species was accompanied by the development of freely accessible web-based databases incorporating generic and intuitive software tools¹. These databases grew rapidly as MLST became more widely adopted, thus further increasing the utility of the scheme, and demonstrating the power of community-oriented resources (Brehony et al., 2007). Due to the limitations of sequence technology in the late 1990s, MLST was designed to assay only a tiny proportion of genomic variation (typically, seven genes). This means MLST data can be insufficiently discriminating to distinguish localized disease outbreaks from cases of endemic disease, or to reconstruct the global or national spread of individual clones (Pérez-Losada et al., 2013; Habib et al., 2014).

The development of next-generation sequencing (NGS) platforms, and in particular the Illumina MiSeq, made it possible to generate data for the vast majority of genes within a bacterial genome from large population samples, thus revealing the *minutiae* of genetic change genome-wide. The MLST databases held at pubmlst.org were thus modified and expanded to encompass data for potentially 1000s of genes rather than seven (Sheppard et al., 2012; Maiden et al., 2013). The ability to assay variation in 100s, or even 1000s, of genes revealed layers of hitherto hidden sub-variants within single 'strains' (or MLST genotypes), and often these sub-variants are geographically restricted (Baker et al., 2010; Harris et al., 2010). The unprecedented discriminatory power of whole genome sequencing has been applied to both human and animal pathogens to reconstruct transmission pathways within

outbreaks, track the source of foodborne pathogens (Dearlove et al., 2016; Ronholm et al., 2016) and infer microevolutionary changes within a single host, over time-scales of weeks or months (Falush, 2009; Azarian et al., 2016; Lillebaek et al., 2016). Patterns of pathogen transmission have been inferred across all epidemiological scales; within single hospital outbreaks (Senn et al., 2016; Willems et al., 2016), animal hosts (Sheppard et al., 2013; Mather et al., 2016), hospitals or farms (Chang et al., 2016; Kamath et al., 2016; Reuter et al., 2016), or across countries or continents (Aanensen et al., 2016).

Of course, as with MLST data (Maiden, 2006), a pre-requisite for using genome sequencing for studying the epidemiology and/or evolution of bacterial pathogens is the initial assembly of a representative population sample. Whilst studies on local or geographically restricted outbreaks necessarily require deep sampling from within the locales affected, a broad phylogenetic context can often prove invaluable for inferring when, and from where, any given outbreak may have arisen. Phylogeography is a general term that combines the spatial dynamics of the pathogen with patterns of molecular evolution, and thus describes the distribution of specific variants over space. Bayesian statistical procedures, as implemented in BEAST (Drummond and Rambaut, 2007), can be used for reconstructing the emergence of a new virulent or resistant pathogen, and its subsequent phylogeographic spread (Holden et al., 2013). Similarly, estimates of mutation rate, which can also be made using models implemented BEAST (Drummond and Suchard, 2010), allow the reconstruction of time calibrated phylogenies, and thus the possibility of linking specific evolutionary changes, such as the emergence of new lineages (or the extinction of old ones), with demographic changes resulting from perturbations in the selective landscape. Such perturbations may either result from specific disease management strategies, such as the deployment of vaccines or the administration of a new antibiotic, or broader environmental challenges such as those resulting from climate

¹www.pubmlst.org

BOX 2 | Gram-positive bacteria.

The frequency of reports of aquaculture diseases caused by Gram-positive pathogens has been steadily rising globally over the last 20 years or so. There are three main Gram positive taxa of relevance to aquaculture disease management, the two firmicute genera *Lactococcus* and *Streptococcus*, and *Renibacterium salmoninarum*, a member of the *Micrococcaceae*. Below we will give an overview of two key species, *L. garvieae* and *S. iniae*. These are both very diverse ecological generalists that can cause disease in a wide range of host species including, on occasion, humans. *R. salmoninarum*, the causative agent of bacterial kidney disease in salmonids, was the first fish pathogen to which whole genome sequencing was applied for molecular epidemiology (Brynildsrud et al., 2014), and this study is discussed in the main text.

Lactococcus garvieae is a highly diverse species that causes a fatal hemorrhagic septicaemia called lactococcosis that can result in a loss of 50% of total production. This pathogen causes disease in fish species worldwide, particularly rainbow trout but also a host of other wild and farmed fishes including Japanese yellowtail (Kusuda et al., 1991), gray mullet (Chen et al., 2002), Japanese eel (Kusuda et al., 1991), Red Sea wrasse (Colorni et al., 2003), bottlenose dolphin (Evans et al., 2006), tilapia (Evans et al., 2009), and flounder (Baeck et al., 2006). *L. garvieae* has been also been isolated from farmed giant freshwater prawns (*Macrobrachium rosenbergii*) (Chen et al., 2001). It is also a causative agent of bovine mastitis, and is associated with dairy and meat products, vegetables and cereals, and occasionally poultry (Ferrario et al., 2012). A few cases of human infection have been described, thus it is recognized as an emerging pathogen in public health, and its ability to survive at relatively low temperatures compared to typical human pathogens can help explain the recent isolation of this organism as contaminating agent in blood products prior to transfusion (Kozakai et al., 2016).

Although a “streptococcal” disease was first recorded in rainbow trout in Japan in the 1950s, the causative organism was not described until the 1983, when it was assigned as *Streptococcus garvieae* but reclassified as *L. garvieae* shortly afterward in Schleifer et al. (1985). The type strain of *S. garvieae* (ATCC 43921) was an isolate recovered from a case of bovine mastitis in the UK; it is perhaps unsurprising that no link was made at this time between a pathogen taken from a mastitic udder in the UK and disease outbreaks in rainbow trout in Japan decades earlier. This connection remained unnoticed even when large outbreaks of lactococcosis started to occur in rainbow trout farms in southern Spain from 1988 (Palacios et al., 1993), with the causative organisms being assigned as *Enterococcus* sp. However, the link between fish disease and *L. garvieae* was made in 1991 when archived samples taken from infected fish in Japan dating back to the 1970s were characterized (Kusuda et al., 1991). Although initially erroneously re-assigned as *Enterococcus seriolicida*, subsequent studies finally reclassified *E. seriolicida* as a junior synonym of *L. garvieae* (Doménech et al., 1993).

Subsequent to the initial outbreaks in southern Spain, the disease spread rapidly throughout Europe and elsewhere in the world. Outbreaks associated with high mortality and severe economic losses have been reported in trout farms worldwide, including in Australia, South Africa, Turkey, the USA (Nelson et al., 2016) and Iran (Sharifiyazdi et al., 2010). Within Europe, there was a rapid spread between rainbow trout farms amongst Mediterranean countries particularly in Italy, Portugal and Balkan countries, and further north into France and the UK. The disease was probably transmitted directly through the movement of infected fish, and indirectly (horizontally) through water (Eyngor et al., 2004). *L. garvieae* is now endemic throughout much of continental Europe and the UK.

The appearance of disease symptoms is highly temperature sensitive, with outbreaks on farms most commonly occurring during the summer months when water temperatures exceed 16°C. It is thus likely that increasing summer temperatures resulting from climate change are at least partly responsible for the northward spread of this pathogen within Europe (Marcos-López et al., 2010). The critical role of temperature in clinical outcome was demonstrated by Algöet et al. (2009) who tested the virulence of a *L. garvieae* strains in rainbow trout by intraperitoneal injection and noted substantially lower mortality at 14°C (3%) than at 16°C (28%) or 18°C (67%).

Initial typing studies based on restriction fragment length polymorphism (RFLP) ribotyping were suggestive of a link between Italian outbreak isolates and the early Japanese isolates, consistent with transmission from Japan into southern Europe (Eldar et al., 1999). However, this link was far from certain as the Japanese isolates were mostly recovered from marine fish and showed serological differences to the isolates from European trout. In fact PFGE data subsequently revealed the Italian-Japanese strains to be quite distinct from the Spanish outbreak isolates and from a French isolate (Vela et al., 2000). The hypothesis that that there was a single introduction into Europe from Japan, followed by regional spread, can therefore be discounted. A more extensive RFLP based ribotyping study revealed clonal regional spread within endemic southern regions, but a more diverse mixture of strains from France where only sporadic infections were reported (Eyngor et al., 2004). The data also revealed that the majority of the isolates within the *L. garvieae* population corresponded to one of two major lineages, with one group corresponding to serotype 1 isolates from Israel and France, and the other serotype 2 isolates from Spain, Greece, and Bulgaria. PFGE data also revealed the presence of two distinct types within Spain, and further confirmed extensive heterogeneity and multiple sources of infection (Vela et al., 2000).

The existence of two distinct and major lineages of *L. garvieae* was later confirmed using multilocus restriction typing (MLRT) (Ferrario et al., 2012) and MLST (Ferrario et al., 2013). These studies characterized isolates from a broad range of ecological sources, including infected fish, raw milk, vegetables, humans, and poultry but, interestingly, found no evidence of concordance between the two major lineages and ecological origin.

Streptococcus iniae is responsible for outbreaks of septicaemia and meningitis (streptococcosis) in at least 27 species of cultured and wild freshwater and marine finfish, including trout, tilapia, bass, yellowtail and flounder (Romalde et al., 1993; Agnew and Barnes, 2007). The name of this species reflects the host from which it was first recovered, *Inia geoffrensis* (a freshwater dolphin). It is associated with warm water, particularly in North America, Asia-Pacific (including Australia) and the Middle East, and it is relatively well-studied partly due to its ability to cause invasive disease in humans through the handling of infected fish. Serotype specific vaccines (including bacterins, attenuated vaccines and DNA vaccines) have generally proved successful, though mutations in the capsular region have been shown underlying vaccine escapes through serotype switching (Zhang et al., 2014; Heath et al., 2016). Molecular typing using ribotyping, random amplification of polymorphic DNA (RAPDs) and rep-PCR have been useful in ruling out putative epidemiological links, for example between North America and Israel (Eldar et al., 1997), and it is likely that local disease is mostly caused by established endemic strains that have gone undetected until relatively recently. Whilst PFGE studies have revealed that isolates recovered from human disease are relatively clonal, isolates recovered from farmed and wild fish are highly diverse, with multiple genotypes being detected within and between different farms. A number of virulence factors in *S. iniae* have been characterized, many of which are similar to those found in group A streptococci (GAS) (Baiano and Barnes, 2009; Zhang et al., 2014).

change. Importantly, just as phylogeographic studies require samples representing, as far as possible, the spatial range of the study population, so the accuracy of mutation rate estimates similarly depends on the temporal spread of the sample. As a rule of thumb, this typically means the inclusion of at least a small number of bacterial strains archived a minimum of 10 years, or ideally 20 years, prior to the contemporary population.

A good example of how these approaches have helped further our understand of the evolution and epidemiology of pathogens important for public health is the study by Holden et al. (2013). The authors sequenced almost 200 genomes of an epidemic clone of Methicillin resistant *Staphylococcus aureus* (EMRSA-15) from multiple countries in Europe and, importantly, included isolates dating back over 20 years. Using

BOX 3 | Viruses.

A number of serious viral aquaculture diseases are notifiable, and are very difficult to treat or vaccinate against (OIE, 2016). However, molecular epidemiological studies have made significant contributions to understanding, and thus managing, the emergence and spread of viral diseases (Snow, 2011). Key aquaculture viruses include infectious salmon anemia virus (ISAV), infectious pancreatic necrosis virus (IPNV), viral hemorrhagic septicemia virus (VHSV), infectious hematopoietic necrosis virus (IHNV), Channel catfish virus (CCV) and Cyprinid herpesvirus 3 (CyHV-3), previously known as Koi herpesvirus (KHV). There are also a range of viral diseases that cause significant problems for farmed shellfish species. These include white spot syndrome virus, yellow head virus genotype 1 and Taura syndrome (OIE, 2016). As has already been discussed with respect to ISA, these viruses can cause devastating outbreaks with severe commercial consequences. Below we discuss CyHV-3 and VHS.

Cyprinid herpesvirus 3, CyHV-3 (previously known as Koi herpesvirus; KHV) is the causative agent of a lethal, highly contagious and notifiable disease of carp and ornamental koi (*Cyprinus carpio*). In addition to being farmed for angling and for exhibitions, carp is also one of most intensively cultivated fish globally for human consumption. As is the case with other aquaculture pathogens, clinical outcome following exposure to CyHV-3 is critically dependent upon water temperature and the disease is highly seasonal. CyHV-3 has spread rapidly to at least 33 countries since it was first observed in Israel, USA and Germany on the late 1990s, probably mostly as a result of global trade and koi exhibitions. As with other members of genus *Cyprinivirus*, it is phylogenetically distinct, and has become a focus of fundamental research into host-virus interactions (Rakus et al., 2013). Full genome sequencing of three isolates from the USA, Israel, and Japan revealed an unusually large genome (295 Kb) consisting of a large central region flanked by inverted 22 Kb inverted repeats each encompassing eight ORFs (Aoki et al., 2007). The presence of loss-of-function mutations in genes predicted to encode membrane glycoproteins has been speculated to be associated with host switching and disease emergence in the carp host from an environmental reservoir population. Although the three genome sequences revealed little genetic diversity, it is clear that variant from Japan is distinct from those isolated from the USA and Israel, which are more similar to each other. Subsequent studies revealed two sub-lineages in Asia, and at least seven within the European lineage. VNTR genotyping has been used successfully to define a new 'intermediate' lineage from Indonesia, and to demonstrate the coexistence of different strains within a single location, which is consistent with multiple introduction events and rapid transmission (Avarre et al., 2011).

Viral Hemorrhagic Septicaemia (VHS) is a highly infectious disease of freshwater and marine fish. In contrast to CyHV-3, where infections are typically asymptomatic at temperatures below 16°C, symptoms of VHS are rarely noted in temperatures above 15°C (Oidtmann et al., 2013). The causative agent, viral hemorrhagic septicemia virus (VHSV) is a World Organisation for Animal Health listed pathogen (OIE, 2016), and management of the disease following detection may involve destruction of fish on affected premises, implementation of movement controls and extensive surveillance prior to resumption of normal trading. The genome is 11-Kb (containing six genes), and it is a single-stranded RNA virus belonging to the genus *Novirhabdovirus*, as does IHNV. Four distinct genotypes have been identified that appear to be associated with different geographical regions, environments (freshwater or marine) and host species (Snow et al., 1999). Genotypes I, II, and III are associated with European waters (with sub-lineages also showing geographical restriction (Abbadi et al., 2016), whereas genogroup IV is typically recovered from North America (Snow et al., 2004). Genotype I has been divided into five subtypes (Einer-Jensen et al., 2004) and genotype IV into three subtypes. VHS has caused substantial die-offs in the Laurentian Great Lakes (Stepien et al., 2015). A recent study of VHSV and IHNV in Italy based on the sequence of the glycoprotein (Abbadi et al., 2016) suggested that there had been multiple introductions of VHSV into Italy.

the Bayesian approaches discussed above, they inferred that this clone emerged in the English West Midlands in the mid-1980s. This inference was strongly supported by a second, independent, line of evidence: an important characteristic of this clone is resistance to flouoroquinolone and, strikingly, the time and date of clonal emergence inferred from the sequence data coincided almost perfectly with the first clinical trials of this drug in the UK. In addition to public health, WGS studies on livestock pathogens have provided insights into the selection pressures resulting from intensive food production processes, and how these processes can increase or limit opportunities for pathogen transmission (Franz et al., 2014). WGS data also underpins an increased understanding of fundamental micro-evolutionary processes in bacterial pathogens, such as homologous recombination, gene gain and loss, genetic drift, and the strength and direction of selection (Castillo-Ramírez et al., 2011; Yahara et al., 2016).

Although the ability to assign an isolate to one type or another using traditional typing methodologies, such as MLST or pulsed-field gel electrophoresis (PFGE), may indirectly provide phenotypic information (e.g., virulence or resistance profile), this is reliant on these traits having already been associated (linked) with the given genotype. In contrast, as WGS data provides a complete catalog of gene content, as well as SNPs (single nucleotide polymorphisms) and indels (insertions and deletions), it is possible to directly mine the data for virulence or resistance genes in order to predict these phenotypes and inform on optimum management options. In particular, mobile genetic elements (MGEs) such as phage, plasmids or genomic

islands commonly carry genes conferring enhanced virulence or resistance to antibiotics (Gordon et al., 2014; Lindsay, 2014; Walker et al., 2015). Alternatively, these MGEs may carry genes underpinning adaptation to specific hosts, ecological niches or environmental pressures (Sheppard et al., 2013; Holt et al., 2015; Hsu et al., 2015). Rapid advances are being made in the development of sophisticated statistical procedures, or genome-wide association studies (GWAS) approaches, to identify genes or single SNPs that are associated with key phenotypic traits such as virulence, resistance, or host adaptation (Farhat et al., 2014; Howell et al., 2014; Skallerup et al., 2015; Falush, 2016; Laabei and Massey, 2016). The broad utility of WGS data thus lies in simultaneously tracking the emergence and spread of a given variant, and also in providing predictive information concerning key phenotypic traits (Feil, 2015). **Figure 1** presents some examples of how whole genome sequencing may be applied to benefit major areas of research that are relevant to pathogen evolution and epidemiology in aquaculture (**Figure 1**).

THE UTILITY OF WHOLE GENOME SEQUENCING WITHIN AQUACULTURE SETTINGS

Many of the questions and examples above have clear and direct relevance for aquaculture settings. It is, in principle, equally feasible and valid to use WGS to track disease emergence and spread within and between individual fish-farms as it is within

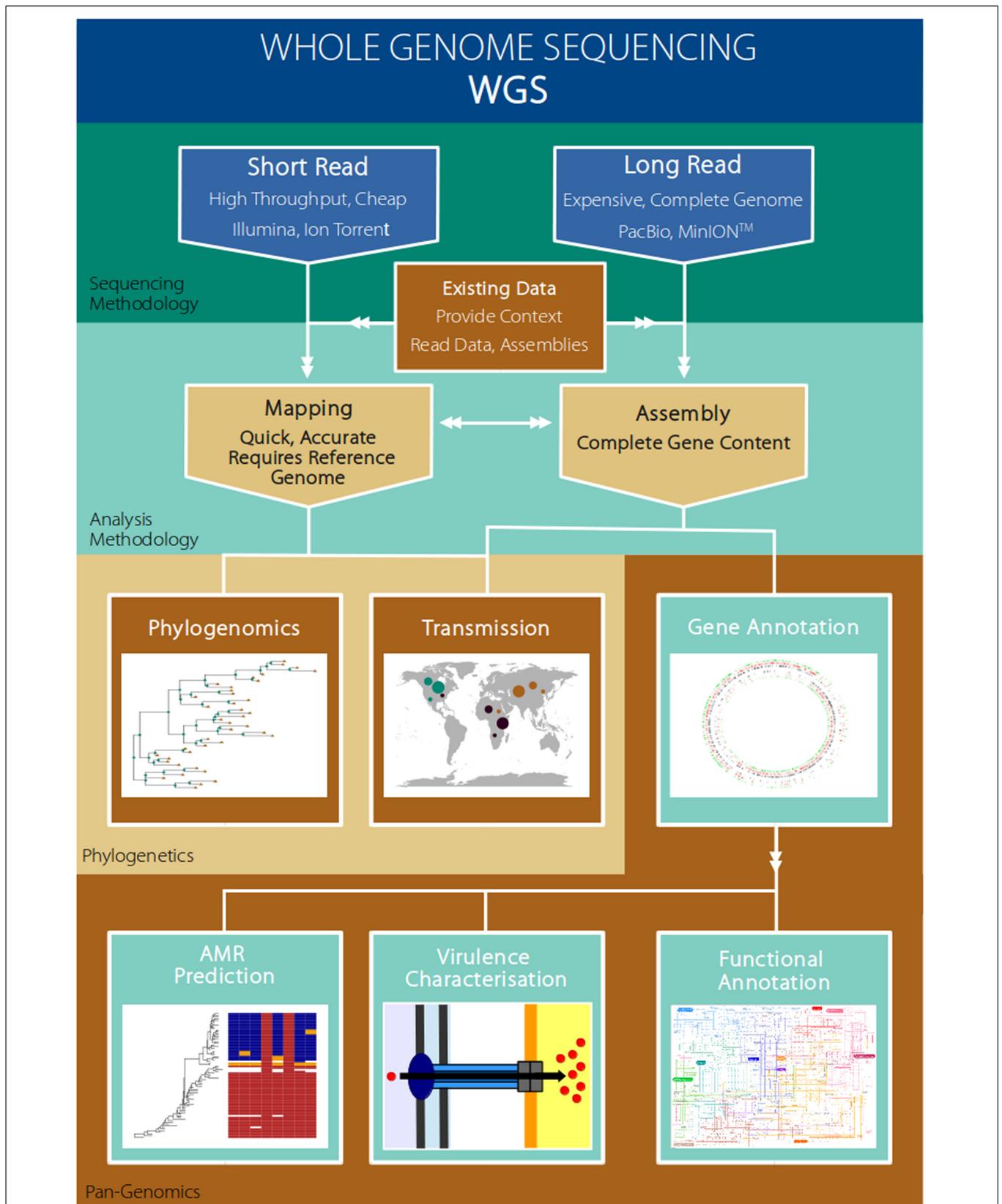


FIGURE 1 | Example workflow of how whole genome sequencing of pathogens maybe used to address research areas of particular relevance to the aquaculture setting.

and between hospitals. Of particular relevance to aquaculture is the ability to identify and track specialized strains that have become adapted to specific hosts, such as *Aeromonas salmonicida* (Box 1) or viral hemorrhagic septicemia virus (Box 3). For example, recent outbreaks of VHSV have occurred in farmed wrasse used to control salmon lice in Scotland and Iceland (Munro et al., 2015). The strains that cause disease in the wrasse were shown to be genogroup III, but these strains do not cause overt disease symptoms in Atlantic salmon (Munro et al., 2015; Matejusova et al., 2016). This is especially relevant for the World Organisation for Animal Health notifiable pathogen species, the presence of which may incur the destruction of stock or a temporary ban on trading between states (OIE, 2016).

Whilst important efforts are being made to incorporate and modify risk-based methods, and other statistical and modeling approaches, from terrestrial to aquaculture systems, these endeavors are often constrained by a lack of basic data concerning pathogen movement, host and environmental adaptations and population dynamics (Murray and Peeler, 2005; Oidtmann et al., 2013). However, given the long and highly successful history of the development and use of robust molecular epidemiological tools in aquaculture (Boxes 1–3) there can be little doubt that whole genome sequencing is set to transform the field (Snow, 2011). There are already a large number of complete reference genomes for a range of aquaculture associated pathogens (Table 1). These

represent a strong foundation to utilize NGS for molecular epidemiology in aquaculture. Phylogeographic analysis, namely the ability to infer geographical structuring and patterns of transmission by distinguishing between almost identical variants of a single pathogenic strain, informs on many of the challenges underpinning the management of aquaculture diseases. These tools are of particular relevance for the identification of transmission routes between aquaculture producing countries or regions, and may help inform novel disease management strategies. More specifically, such approaches can be used to unequivocally demonstrate whether the disease emergence in a specific region has resulted from transmission of the pathogen from elsewhere, or has originated from within the ‘native’ endemic microbial ecosystem. An excellent recent example, and a superb “roadmap” for how WGS might be applied to aquaculture pathogens more broadly, is the recent study by Barnes et al. (2016) (Box 1). These authors sequenced 51 genomes of *Yersinia ruckeri*, an important pathogen of salmonids, from Tasmania ($n = 44$) and New Zealand ($n = 7$) and analyzed these alongside six genomes from other countries already in the public domain. The temporal range of the isolates was 27 years. These data demonstrated that the disease-causing strains in Australia and New Zealand are distinct from those in the Northern Hemisphere, and are instead local endemic strains.

Each specific setting and pathogen will, of course, pose particular questions. We outline below further recent key studies

TABLE 1 | WGS studies of selected bacterial and viral aquaculture pathogens.

	Pathogenic species	Reference
Bacterial	<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	Reith et al., 2008; Charette et al., 2012; Valdes et al., 2015; Vincent et al., 2015, 2016
	<i>Aeromonas salmonicida</i> subsp. <i>achromogenes</i>	Han et al., 2013
	<i>Francisella noatunensis</i> subsp. <i>orientalis</i>	Sridhar et al., 2012; Figueiredo et al., 2016; Gonçalves et al., 2016
	<i>Flavobacterium psychrophilum</i>	Duchaud et al., 2007; Wu et al., 2015; Castillo et al., 2016; Neiger et al., 2016
	<i>Lactococcus garvieae</i>	Aguado-Urda et al., 2011; Morita et al., 2011; Ricci et al., 2012
	<i>Piscirickettsia salmonis</i>	Eppinger et al., 2013; Pulgar et al., 2015
	<i>Renibacterium salmoninarum</i>	Wiens et al., 2008; Brynildsrud et al., 2014
	<i>Streptococcus agalactiae</i>	Delannoy et al., 2012; Liu et al., 2013; Pereira Ude et al., 2013; Kayansamruaj et al., 2015
	<i>Streptococcus iniae</i>	El Aamri et al., 2013; Pridgeon et al., 2014; Zhang et al., 2014
	<i>Vibrio anguillarum</i>	Naka et al., 2011; Li et al., 2013; Busschaert et al., 2015; Holm et al., 2015
	<i>Vibrio harveyi</i>	Kondo et al., 2015
	<i>Vibrio parahaemolyticus</i>	Yang et al., 2014; Letchumanan et al., 2016
	<i>Yersinia ruckeri</i>	Nelson et al., 2015; Wang et al., 2015; Barnes et al., 2016; Liu et al., 2016
	<i>Vibrio aestuarianus</i>	Goudenège et al., 2015
	<i>Aeromonas hydrophila</i>	Rasmussen-Ivey et al., 2016
	Viral	Cyprinid herpesvirus (CyHV-3)/Koi herpesvirus (KHV)
Infectious hematopoietic necrosis virus (IHNV)		Jia et al., 2014; Wang et al., 2016
Infectious pancreatic necrosis virus (IPNV)		Dadar et al., 2014
Infectious salmon anemia virus (ISAV)		Markussen et al., 2008; Cottet et al., 2010; Mérour et al., 2011
Taura syndrome virus (TSV)		Mari et al., 2002
Viral hemorrhagic septicemia virus (VHSV)		Schütze et al., 1999; Kim et al., 2013; Ito et al., 2016; Schönherz et al., 2016
White spot syndrome virus (WSSV)/white spot bacilliform virus (WSBV)		Yang et al., 2001; Rodriguez-Anaya et al., 2016
Yellow head nidovirus (YHV)	Sittidilokratna et al., 2008	

that illustrate how whole genome sequencing has already begun to transform the field.

Renibacterium salmoninarum is the causative agent of bacterial kidney disease in salmonids, and a successful Multilocus VNTR (variable number of tandem repeats) Analysis (MLVA) scheme has also been developed (Matejusova et al., 2013). *R. salmoninarum* was also the first aquaculture pathogen for which WGS was used to reconstruct patterns of international transmission and microevolutionary events (Brynildsrud et al., 2014). These data are consistent with the hypothesis that this pathogen was introduced into Europe from North America over the last 50 years, possibly via the trade in eggs. The study also provided proof-of-principle that WGS can be used to detect outbreak variants within a single fish-farm or neighboring farms. The recovery of essentially indistinguishable isolates from the same geographical location, but from different host species, provided strong evidence for frequent host-switching within the farm environment, which has implications for both the management of farmed aquatic animals and also in gaging the risk of spillover into wild stocks. These data also revealed a deep phylogenetic division within the population, possibly reflecting historical allopatry between Pacific and Atlantic wild stocks.

Whole genome sequencing of aquaculture pathogens, including both viruses and bacteria, has also been central to the development of novel diagnostic tools and bespoke typing methodologies. For example, Avarre et al. (2011) noted that a large proportion of the variation present in the genomes of CyHV-3 is restricted to a small number of hyper-variable tandem repeat loci, from which they developed a discriminatory and relatively cheap and easy VNTR scheme (variable number of tandem repeats). A more recent study showed that it is possible to sequence CyHV-3 genomes directly from tissues of infected carp, which revealed high levels of mixed infection (Hammoumi et al., 2016). Whole genome sequencing is also paving the way to the development of new vaccine targets through the identification of novel surface expressed proteins (Barnes and Silayeva, 2016) and reverse vaccinology (Andreoni et al., 2016). Seth-Smith et al. (2016) used whole genome sequencing to characterize two novel intracellular pathogens of gilthead seabream. Many other studies are now using whole genome sequencing to confirm inferences from other typing methods. For example cases of mixed infection of *Streptococcus agalactiae* and *Francisella noatunensis* subsp. *orientalis* in farmed Nile Tilapia (Assis et al., 2016), and comparative genomics were used to supplement a comprehensive MLST study of *Aeromonas hydrophila* infection in farmed catfish in the USA (Hossain et al., 2014).

Recent comparative genomics papers have also focused on understanding pathogenicity and the distribution of virulence factors in bacterial populations (Sudheesh et al., 2012). Busschaert et al. (2015) analyzed 15 draft genomes of *Vibrio anguillarum* and noted very few genetic differences between strains with differing virulence properties, indicating that only very small genomic changes might underlie major clinical differences. Castillo et al. (2016) combined six novel genome sequences of *F. psychrophilum* from fish farms with five publicly available sequences to show very few differences in the distribution of virulence factors or phenotypic traits over large

geographical distances. Finally, a comparative genomics analysis of a range of *A. salmonicida* sub-species provided a detailed picture on the genome dynamics of this species and provided evidence that insertion sequences may be a driving force in the evolution of a psychrophilic lifestyle (Vincent et al., 2016).

Recent advances in third generation sequencing technologies may also impact on the application of WGS to the aquaculture sector. Long read technologies, such as the Pacific Biosciences' PacBio, have dramatically reduced the time and cost of producing high-quality draft and complete genome sequences (Zhang et al., 2012). Another innovation in the field of sequencing technologies, the Oxford Nanopore MinION™, has the potential to reduce the cost and time investment necessary to set-up sequencing centers. The MinION™ has been used for 'in the field' sequencing of the recent Ebola and Zika virus outbreaks (Faria et al., 2016; Quick et al., 2016). Considering that the majority of aquaculture producing regions are in developing nations this may be a revolutionary development for sustainable disease control strategies that use WGS for molecular tracing. Although the cost of the technology is currently prohibitive, future developments may lead to this technology being a cost-effective alternative in regions without existing sequencing infrastructure. In addition, the turnaround time for the MinION™ from sample to sequence has been shown to be short in comparison to current technologies. When used during the Ebola outbreak, sequences of the virus were often available within 24 h of the sample being submitted.

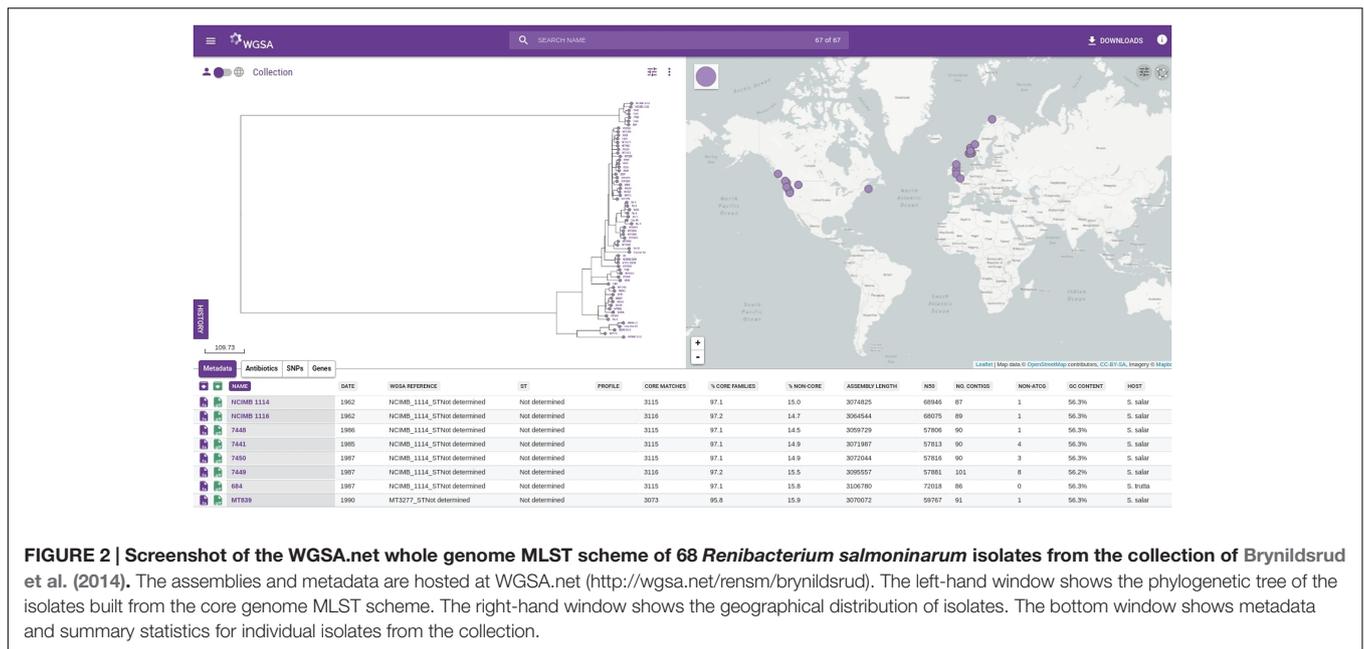
THE MANAGEMENT AND ANALYSIS OF WGS DATA WITHIN A COMMUNITY-ORIENTED FRAMEWORK

The studies cited above illustrate how the generation of WGS data can advance basic bioscience and aquaculture disease management. However, the optimal exploitation of this technology is dependent on a community-oriented infrastructure being in place to facilitate resource sharing. Rapid advances are being made in the development of community-oriented databases and analysis tools for public health pathogens, and these tools are largely transferable to aquaculture pathogens. Freely available databases that store, manipulate and analyze WGS data, include BIGSdb (Maiden et al., 2013), Enterobase², Whole Genome Sequence Analysis³, and commercial systems are also available⁴. Other systems deliver rapid annotation such as the RAST server (Aziz et al., 2008), or can predict antibiotic resistance profiles (Zankari et al., 2012; Bradley et al., 2015), or analyze gene content variation (Page et al., 2015). Bioinformatics tools are rapidly becoming more powerful and more accessible to the community, and the development of cloud-based informatics infrastructure provides the means both to provide easy access to these tools and the necessary computing power to analyze very large datasets (e.g., climb.ac.uk).

²<https://enterobase.warwick.ac.uk/>

³<http://WGSA.net>

⁴<http://www.ridom.com/seqsphere>



In order to illustrate the utility of these databases, we have loaded the *R. salmoninarum* whole genome sequence data of Brynildsrud et al. (2014) on to WGS.net. Additionally, we have used the same dataset to generate a core genome MLST scheme (<http://wgsa.net/rems/brynildsrud>; Figure 2). WGS.net is a web application for uploading, processing, clustering and exploration of microbial genome assemblies in real time. Genome assemblies and metadata files can be added simply through “drag and drop” onto a web browser. The system then provides a suite of analysis options including quality metrics for the data, reconstruction of MLST profiles (if a scheme has been pre-defined), resistance (or virulence) prediction, ‘core’ gene set identification, gene content differences, and phylogenetic analysis. SNPs or genes specific to a given clade (useful for developing novel diagnostics) are provided and data are displayed interactively via trees and maps showing the geographical sources of the isolates. The dataset of Brynildsrud et al. (2014) has been included as a public collection, allowing for other research groups to integrate their own public or private data to provide context to WGS data of local or regional outbreaks of *R. salmoninarum*. Simpler tools have also been developed for visualizing and exploring open-access datasets. Full genome phylogenies of Ebola and Zika viruses, and patterns of genomic variability, are freely available to explore at nextstrain.org/ebola and nextstrain.org/zika respectively. The Microreact web-server (microreact.org) is designed to allow sharing and visualization of metadata and phylogenetic trees and serves as an open platform for linking the outputs of whole genome analysis pipelines to epidemiological and other metadata for collective interpretation (Argimón et al., 2016). We have utilized this tool to visualize the *R. salmoninarum* data of Brynildsrud et al. (2014) (<https://microreact.org/project/N1KdygLt>; Figure 3). This system provides a number of powerful features, such as the ability to select isolates based on geographical location (directly

from the map), phylogenetic position (directly from the tree) or time of isolation.

CONCLUSION: THE FUTURE OF PATHOGEN GENOMICS IN AQUACULTURE

It is clear that whole genome sequencing is set to revolutionize the molecular epidemiology of aquaculture pathogens as it has for those pathogens of relevance to public health. Reconstructing pathogen transmission routes on local and global scales will form an important part of this, but the specific challenges relating to aquaculture disease management are far more multilayered and complex than for public health. These challenges encompass the biology and ecology of both the pathogen and the host (including host adaptation, and different modes of transmission), trading patterns, husbandry and food production, as well as monitoring pathogen emergence and spread due to shifting environmental pressures, in particular climate change. Whole genome sequencing will make important contributions to all these aspects, and others besides. However, in order to capture the full potential of these data, it is critical to develop common standards and pipelines, and open, community-oriented database infrastructures and analysis tools that can be tailored toward these specific questions.

An important first step will be the generation of a set of reference datasets for key pathogens, both bacterial and viral, which represent wide geographical, temporal and phenotypic diversity which can be readily accessed and interrogated. To this end, we are in the process of expanding the dataset of *R. salmoninarum* to include isolates from Chile, and generating datasets for other key aquaculture pathogens including *F. psychrophilum*, *Piscirickettsia salmonis*, *V. anguillarum*, *Y. ruckeri*, *L. garvieae* and CyHV-3. This has been made possible

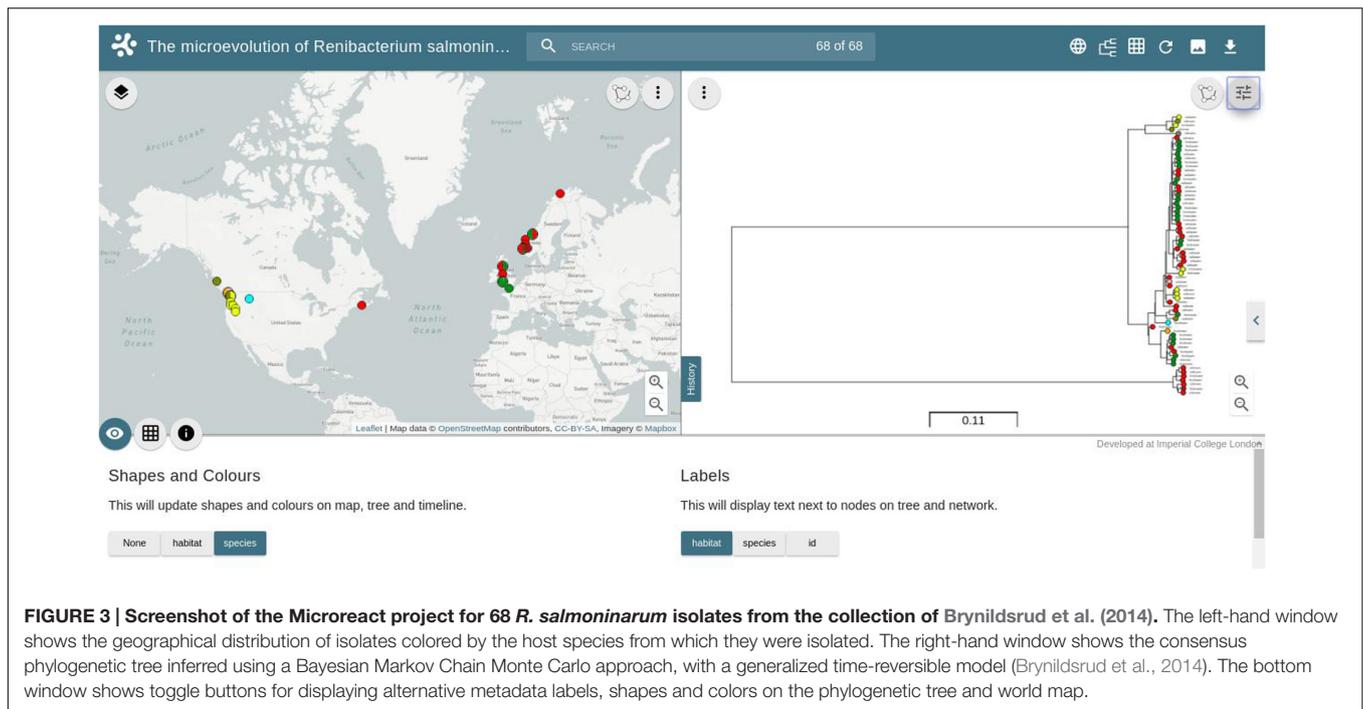


FIGURE 3 | Screenshot of the Microreact project for 68 *R. salmoninarum* isolates from the collection of Brynildsrud et al. (2014). The left-hand window shows the geographical distribution of isolates colored by the host species from which they were isolated. The right-hand window shows the consensus phylogenetic tree inferred using a Bayesian Markov Chain Monte Carlo approach, with a generalized time-reversible model (Brynildsrud et al., 2014). The bottom window shows toggle buttons for displaying alternative metadata labels, shapes and colors on the phylogenetic tree and world map.

in part by the WGS-aqua project (wgs-aqua.net), funded under the UK Research Councils (BBSRC/NERC) sustainable aquaculture call, with important contributions from Cefas. These data will be stored in the standard read and assembly archives but also made available through WGSa.net. These high quality, open-access reference datasets will provide a valuable tool for the aquaculture community and that will showcase the utility of WGS data for understanding the emergence, transmission, and molecular epidemiology of aquaculture pathogens. Furthermore, we hope that they also highlight the importance of community based, community supported and open-access databases for monitoring and management of disease through this and subsequent publications.

If the current trends in the deployment of WGS for disease management for public health are replicated for aquaculture, then we will witness a transition to the routine use of genomics for molecular epidemiology and the generation of datasets containing 100s, if not 1000s, of isolates. Although technical advances in informatics are already providing the means to openly share, analyze and contribute to these datasets, significant barriers relating to data ownership and commercial sensitivities will remain. The consequences of disclosing the presence of a pathogen on a farm when the farm is not legally obliged to do so may clearly be prohibitive. However, we are optimistic that the significant benefit of access to high quality reference collections with broad geographical and temporal coverage will become readily apparent within the aquaculture community. Just as with public health, both academics and governmental agencies should be encouraged to incorporate WGS into routine investigation and, as far as possible, to make the data available to the community.

Presently the application of WGS to bacterial genomes is more cost-effective and less technically demanding than viral or parasite genomes. However, methodological advances in enrichment for viral sequences may shortly make large scale investigation of the epidemiology of viral genomes feasible (Bonsall et al., 2015). Protists and other eukaryotic parasites represent a further frontier that, due to space constraints, have not been discussed at length in this article and which currently pose more serious methodological challenges. However, the rate at which sequencing technology is developing, with the advent of entirely new platforms capable of generating long reads, such as the Oxford Nanopore Minion™, there are grounds to be optimistic that large genomics datasets, representing 1000s of specimens, will soon be within reach even for these more challenging disease agents.

AUTHOR CONTRIBUTIONS

All authors contributed significantly to the manuscript, with the initial drafts being prepared by SB, DV-J, and EF.

ACKNOWLEDGMENTS

SB, KB, DA, SS, AA, and EF are funded on the wgs-aqua.net project funded by BBSRC/NERC as part of the sustainable aquaculture call (BB/M026388/1). Cefas have also contributed significant funding to this project. Many thanks to Ola Brynildsrud for providing information, advice and the *R. salmoninarum* dataset for Microreact and WGSa.net and to Corin Yeats for setting up the *R. salmoninarum* WGSa scheme and website.

REFERENCES

- Aanensen, D. M., Feil, E. J., Holden, M. T. G., Dordel, J., Yeats, C. A., Fedosejev, A., et al. (2016). Whole-genome sequencing for routine pathogen surveillance in public health: a population snapshot of nvasive *Staphylococcus aureus* in Europe. *MBio* 7:e00444-16. doi: 10.1128/mBio.00444-16
- Abbadì, M., Fusaro, A., Ceolin, C., Casarotto, C., Quartesan, R., Dalla Pozza, M., et al. (2016). Molecular evolution and phylogeography of co-circulating IHNV and VHSV in Italy. *Front. Microbiol.* 7:1306. doi: 10.3389/fmicb.2016.01306
- Agnew, W., and Barnes, A. C. (2007). *Streptococcus iniae*: an aquatic pathogen of global veterinary significance and a challenging candidate for reliable vaccination. *Vet. Microbiol.* 122, 1–15. doi: 10.1016/j.vetmic.2007.03.002
- Aguado-Urda, M., López-Campos, G. H., Gibello, A., Cutuli, M. T., López-Alonso, V., Fernández-Garayzábal, J. F., et al. (2011). Genome sequence of *Lactococcus garvieae* 8831, isolated from rainbow trout lactococcosis outbreaks in Spain. *J. Bacteriol.* 193, 4263–4264. doi: 10.1128/JB.05326-11
- Algöet, M., Bayley, A. E., Roberts, E. G., Feist, S. W., Wheeler, R. W., and Verner-Jeffreys, D. W. (2009). Susceptibility of selected freshwater fish species to a UK *Lactococcus garvieae* isolate. *J. Fish Dis.* 32, 825–834. doi: 10.1111/j.1365-2761.2009.01058.x
- Alvial, A., Kibenge, F., Forster, J., Burgos, J. M., Ibarra, R., and St-Hilaire, S. (2012). *The Recovery of the Chilean Salmon Industry: The ISA Crisis and Its Consequences and Lessons*. Available at: https://gaalliance.org/wp-content/uploads/2015/02/GAA_ISA-Report.pdf
- Andreoni, F., Amagliani, G., and Magnani, M. (2016). Selection of vaccine candidates for fish Pasteurellosis using reverse vaccinology and an in vitro screening approach. *Methods Mol. Biol.* 1404, 181–192. doi: 10.1007/978-1-4939-3389-1_12
- Aoki, T., Hirono, I., Kurokawa, K., Fukuda, H., Nahary, R., Eldar, A., et al. (2007). Genome sequences of three koi herpesvirus isolates representing the expanding distribution of an emerging disease threatening koi and common carp worldwide. *J. Virol.* 81, 5058–5065. doi: 10.1128/JVI.00146-07
- Arai, H., Morita, Y., Izumi, S., Katagiri, T., and Kimura, H. (2007). Molecular typing by pulsed-field gel electrophoresis of *Flavobacterium psychrophilum* isolates derived from Japanese fish. *J. Fish Dis.* 30, 345–355. doi: 10.1111/j.1365-2761.2007.00809.x
- Argimón, S., Abudahab, K., Goater, R. J. E., Fedosejev, A., Bhai, J., Glasner, C., et al. (2016). Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Microb. Genom.* 2. doi: 10.1099/mgen.0.000093
- Assis, G. B. N., Tavares, G. C., Pereira, F. L., Figueiredo, H. C. P., and Leal, C. A. G. (2016). Natural coinfection by *Streptococcus agalactiae* and *Francisella noatunensis* subsp. *orientalis* in farmed Nile tilapia (*Oreochromis niloticus* L.). *J. Fish Dis.* 40, 51–63. doi: 10.1111/jfd.12493
- Austin, B., and Austin, D. A. (2012). *Bacterial Fish Pathogens - Disease of Farmed and Wild Fish*, 5th Edn. Dordrecht: Springer.
- Austin, D. A., Robertson, P. A. W., and Austin, B. (2003). Recovery of a new biogroup of *Yersinia ruckeri* from diseased rainbow trout (*Oncorhynchus mykiss*, Walbaum). *Syst. Appl. Microbiol.* 26, 127–131. doi: 10.1078/072320203322337416
- Avarre, J.-C., Madeira, J.-P., Santika, A., Zainun, Z., Baud, M., Cabon, J., et al. (2011). Investigation of Cyprinid herpesvirus-3 genetic diversity by a multi-locus variable number of tandem repeats analysis. *J. Virol. Methods* 173, 320–327. doi: 10.1016/j.jviromet.2011.03.002
- Azarian, T., Daum, R. S., Petty, L. A., Steinbeck, J. L., Yin, Z., Nolan, D., et al. (2016). Intrahost evolution of methicillin-resistant *Staphylococcus aureus* USA300 among individuals with reoccurring skin and soft-tissue infections. *J. Infect. Dis.* 214, 895–905. doi: 10.1093/infdis/jiw242
- Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi: 10.1186/1471-2164-9-75
- Baeck, G. W., Kim, J. H., Gomez, D. K., and Park, S. C. (2006). Isolation and characterization of *Streptococcus* sp. from diseased flounder (*Paralichthys olivaceus*) in Jeju Island. *J. Vet. Sci.* 7, 53–58. doi: 10.4142/jvs.2006.7.1.53
- Baiano, J. C. F., and Barnes, A. C. (2009). Towards control of *Streptococcus iniae*. *Emerg. Infect. Dis.* 15, 1891–1896. doi: 10.3201/eid1512.090232
- Baker, S., Hanage, W. P., and Holt, K. E. (2010). Navigating the future of bacterial molecular epidemiology. *Curr. Opin. Microbiol.* 13, 640–645. doi: 10.1016/j.mib.2010.08.002
- Bakke, T. A., and Harris, P. D. (1998). Diseases and parasites in wild Atlantic salmon (*Salmo salar*) populations. *Can. J. Fish. Aquat. Sci.* 55, 247–266. doi: 10.1139/d98-021
- Baluyut, E. A. (1989). *Aquaculture Systems and Practises: A Selected Review*. Available at: <http://www.fao.org/docrep/t8598e/t8598e05.htm>
- Barnes, A. C. (2011). “Enteric redmouth disease (*Yersinia ruckeri*),” in *Fish Diseases and Disorders: Viral, Bacterial, and Fungal Infections*, eds P. Woo and D. Bruno (Wallingford: CABI).
- Barnes, A. C., Delamare-Deboutteville, J., Gudkovs, N., Brosnahan, C., Morrison, R., and Carson, J. (2016). Whole genome analysis of *Yersinia ruckeri* isolated over 27 years in Australia and New Zealand reveals geographical endemism over multiple lineages and recent evolution under host selection. *Microb. Genom.* 2. doi: 10.1099/mgen.0.000095
- Barnes, A. C., and Silayeva, O. (2016). Vaccination against streptococcal infections in farmed fish. *Microbiol. Aust.* 37, 118–121. doi: 10.1071/MA16040
- Bastardo, A., Ravelo, C., and Romalde, J. L. (2012). Multilocus sequence typing reveals high genetic diversity and epidemic population structure for the fish pathogen *Yersinia ruckeri*. *Environ. Microbiol.* 14, 1888–1897. doi: 10.1111/j.1462-2920.2012.02735.x
- Bastardo, A., Ravelo, C., and Romalde, J. L. (2015). Phylogeography of *Yersinia ruckeri* reveals effects of past evolutionary events on the current strain distribution and explains variations in the global transmission of enteric redmouth (ERM) disease. *Front. Microbiol.* 6:1198. doi: 10.3389/fmicb.2015.01198
- Bonsall, D., Ansari, M. A., Ip, C., Trebes, A., Brown, A., Klenerman, P., et al. (2015). ve-SEQ: robust, unbiased enrichment for streamlined detection and whole-genome sequencing of HCV and other highly diverse pathogens. *F1000Res.* 4:1062. doi: 10.12688/f1000research.7111.1
- Bradley, P., Gordon, N. C., Walker, T. M., Dunn, L., Heys, S., Huang, B., et al. (2015). Rapid antibiotic-resistance predictions from genome sequence data for *Staphylococcus aureus* and *Mycobacterium tuberculosis*. *Nat. Commun.* 6:10063. doi: 10.1038/ncomms10063
- Brehony, C., Jolley, K. A., and Maiden, M. C. J. (2007). Multilocus sequence typing for global surveillance of meningococcal disease. *FEMS Microbiol. Rev.* 31, 15–26. doi: 10.1111/j.1574-6976.2006.00056.x
- Brynildsrud, O., Feil, E. J., Bohlin, J., Castillo-Ramirez, S., Colquhoun, D., McCarthy, U., et al. (2014). Microevolution of *Renibacterium salmoninarum*: evidence for intercontinental dissemination associated with fish movements. *ISME J.* 8, 746–756. doi: 10.1038/ismej.2013.186
- Busschaert, P., Frans, I., Crauwels, S., Zhu, B., Willems, K., Bossier, P., et al. (2015). Comparative genome sequencing to assess the genetic diversity and virulence attributes of 15 *Vibrio anguillarum* isolates. *J. Fish Dis.* 38, 795–807. doi: 10.1111/jfd.12290
- Castillo, D., Christiansen, R. H., Dalsgaard, I., Madsen, L., Espejo, R., and Middelboe, M. (2016). Comparative genome analysis provides insights into the pathogenicity of *Flavobacterium psychrophilum*. *PLoS ONE* 11:e0152515. doi: 10.1371/journal.pone.0152515
- Castillo-Ramirez, S., Harris, S. R., Holden, M. T. G., He, M., Parkhill, J., Bentley, S. D., et al. (2011). The impact of recombination on dN/dS within recently emerged bacterial clones. *PLoS Pathog.* 7:e1002129. doi: 10.1371/journal.ppat.1002129
- Chang, H.-H., Dordel, J., Donker, T., Worby, C. J., Feil, E. J., Hanage, W. P., et al. (2016). Identifying the effect of patient sharing on between-hospital genetic differentiation of methicillin-resistant *Staphylococcus aureus*. *Genome Med.* 8:18. doi: 10.1186/s13073-016-0274-3
- Charette, S. J., Brochu, F., Boyle, B., Filion, G., Tanaka, K. H., and Derome, N. (2012). Draft genome sequence of the virulent strain 01-B526 of the fish pathogen *Aeromonas salmonicida*. *J. Bacteriol.* 194, 722–723. doi: 10.1128/JB.06276-11
- Chen, S.-C., Liaw, L.-L., Su, H.-Y., Ko, S.-C., Wu, C.-Y., Chung, H.-C., et al. (2002). *Lactococcus garvieae*, a cause of disease in grey mullet, *Mugil cephalus* L., in Taiwan. *J. Fish Dis.* 25, 727–732. doi: 10.1046/j.1365-2761.2002.00415.x
- Chen, S. C., Lin, Y. D., Liaw, L. L., and Wang, P. C. (2001). *Lactococcus garvieae* infection in the giant freshwater prawn *Macrobrachium rosenbergii* confirmed by polymerase chain reaction and 16S rDNA sequencing. *Dis. Aquat. Organ.* 45, 45–52. doi: 10.3354/dao045045
- Chen, Y.-C., Davis, M. A., Lapatra, S. E., Cain, K. D., Snekvik, K. R., and Call, D. R. (2008). Genetic diversity of *Flavobacterium psychrophilum* recovered

- from commercially raised rainbow trout, *Oncorhynchus mykiss* (Walbaum), and spawning coho salmon, *O. kisutch* (Walbaum). *J. Fish Dis.* 31, 765–773. doi: 10.1111/j.1365-2761.2008.00950.x
- Collins, R. O., Ferguson, D. A., and Bonniwell, M. A. (1991). Furunculosis in wrasse. *Vet. Rec.* 128:43. doi: 10.1136/vr.128.2.43-a
- Colorni, A., Ravelo, C., Romalde, J. L., Toranzo, A. E., and Diamant, A. (2003). *Lactococcus garvieae* in wild Red Sea wrasse *Coris aygula* (Labridae). *Dis. Aquat. Organ.* 56, 275–278. doi: 10.3354/dao056275
- Colston, S. M., Fullmer, M. S., Beka, L., Lamy, B., Gogarten, J. P., and Graf, J. (2014). Bioinformatic genome comparisons for taxonomic and phylogenetic assignments using *Aeromonas* as a test case. *MBio* 5:e02136. doi: 10.1128/mBio.02136-14
- Coquet, L., Cosette, P., Junter, G.-A., Beucher, E., Saiter, J.-M., and Jouenne, T. (2002). Adhesion of *Yersinia ruckeri* to fish farm materials: influence of cell and material surface properties. *Colloid. Surf. B* 26, 373–378. doi: 10.1016/S0927-7765(02)00023-1
- Costa-Pierce, B. A. (2008). *Ecological Aquaculture: The Evolution of the Blue Revolution*. Chichester: John Wiley & Sons Ltd.
- Cottet, L., Cortez-San Martin, M., Tello, M., Olivares, E., Rivas-Aravena, A., Vallejos, E., et al. (2010). Bioinformatic analysis of the genome of infectious salmon anemia virus associated with outbreaks with high mortality in Chile. *J. Virol.* 84, 11916–11928. doi: 10.1128/JVI.01202-10
- Crane, M., and Hyatt, A. (2011). Viruses of fish: an overview of significant pathogens. *Viruses* 3, 2025–2046. doi: 10.3390/v3112025
- Dadar, M., Peyghan, R., Rajabi-Memari, H., Seifi Abad Shapouri, M. R., Hasanzadeh, R., Moazzami Goudarzi, L., et al. (2014). Phylogenetic relationships of Iranian Infectious Pancreatic Necrosis Virus (IPNV) based on deduced amino acid sequences of genome segment A and B cDNA. *Iran. J. Fish. Sci.* 13, 560–575.
- Dallaire-Dufresne, S., Tanaka, K. H., Trudel, M. V., Lafaille, A., and Charette, S. J. (2014). Virulence, genomic features, and plasticity of *Aeromonas salmonicida* subsp. *salmonicida*, the causative agent of fish furunculosis. *Vet. Microbiol.* 169, 1–7. doi: 10.1016/j.vetmic.2013.06.025
- Davies, R. L. (1991a). Clonal analysis of *Yersinia ruckeri* based on biotypes, serotypes and outer membrane protein-types. *J. Fish Dis.* 14, 221–228. doi: 10.1111/j.1365-2761.1991.tb00591.x
- Davies, R. L. (1991b). Virulence and serum-resistance in different clonal groups and serotypes of *Yersinia ruckeri*. *Vet. Microbiol.* 29, 289–297. doi: 10.1016/0378-1135(91)90136-4
- Dearlove, B. L., Cody, A. J., Pascoe, B., Méric, G., Wilson, D. J., and Sheppard, S. K. (2016). Rapid host switching in generalist *Campylobacter* strains erodes the signal for tracing human infections. *ISME J.* 10, 721–729. doi: 10.1038/ismej.2015.149
- Delannoy, C. M. J., Zadoks, R. N., Lainson, F. A., Ferguson, H. W., Crumlish, M., Turnbull, J. F., et al. (2012). Draft genome sequence of a nonhemolytic fish-pathogenic *Streptococcus agalactiae* strain. *J. Bacteriol.* 194, 6341–6342. doi: 10.1128/JB.01552-12
- Doménech, A., Prieta, J., Fernández-Garayzábal, J. F., Collins, M. D., Jones, D., and Domínguez, L. (1993). Phenotypic and phylogenetic evidence for a close relationship between *Lactococcus garvieae* and *Enterococcus seriolicida*. *Microbiologia* 9, 63–68.
- Drummond, A. J., and Rambaut, A. (2007). BEAST: bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214. doi: 10.1186/1471-2148-7-214
- Drummond, A. J., and Suchard, M. A. (2010). Bayesian random local clocks, or one rate to rule them all. *BMC Biol.* 8:114. doi: 10.1186/1741-7007-8-114
- Duchaud, E., Boussaha, M., Loux, V., Bernardet, J.-F., Michel, C., Kerouault, B., et al. (2007). Complete genome sequence of the fish pathogen *Flavobacterium psychrophilum*. *Nat. Biotechnol.* 25, 763–769. doi: 10.1038/nbt1313
- Einer-Jensen, K., Ahrens, P., Forsberg, R., and Lorenzen, N. (2004). Evolution of the fish rhabdovirus viral haemorrhagic septicaemia virus. *J. Gen. Virol.* 85, 1167–1179. doi: 10.1099/vir.0.79820-0
- El Aamri, F., Acosta, F., Real, F., and Padilla, D. (2013). Whole-genome sequence of the fish virulent strain *Streptococcus iniae* IUSA-1, isolated from gilthead sea bream (*Sparus aurata*) and Red Porgy (*Pagrus pagrus*). *Genome Announc.* 1:e0002513. doi: 10.1128/genomeA.00025-13
- Eldar, A., Goria, M., Ghittino, C., Zlotkin, A., and Bercovier, H. (1999). Biodiversity of *Lactococcus garvieae* strains isolated from fish in Europe, Asia, and Australia. *Appl. Environ. Microbiol.* 65, 1005–1008.
- Eldar, A., Lawhon, S., Frelief, P. F., Assenta, L., Simpson, B. R., Varner, P. W., et al. (1997). Restriction fragment length polymorphisms of 16S rDNA and of whole rRNA genes (ribotyping) of *Streptococcus iniae* strains from the United States and Israel. *FEMS Microbiol. Lett.* 151, 155–162. doi: 10.1111/j.1574-6968.1997.tb12564.x
- Eppinger, M., McNair, K., Zogaj, X., Dinsdale, E. A., Edwards, R. A., and Klose, K. E. (2013). Draft genome sequence of the fish pathogen *Piscirickettsia salmonis*. *Genome Announc.* 1:e00926-13. doi: 10.1128/genomeA.00926-13
- Evans, J. J., Klesius, P. H., and Shoemaker, C. A. (2009). First isolation and characterization of *Lactococcus garvieae* from Brazilian Nile tilapia, *Oreochromis niloticus* (L.), and pintado, *Pseudoplathystoma corruscans* (Spix & Agassiz). *J. Fish Dis.* 32, 943–951. doi: 10.1111/j.1365-2761.2009.01075.x
- Evans, J. J., Pasnik, D. J., Klesius, P. H., and Al-Ablani, S. (2006). First report of *Streptococcus agalactiae* and *Lactococcus garvieae* from a wild bottlenose dolphin (*Tursiops truncatus*). *J. Wildl. Dis.* 42, 561–569. doi: 10.7589/0090-3558-42.3.561
- Eyngor, M., Zlotkin, A., Ghittino, C., Prearo, M., Douet, D.-G., Chilmonczyk, S., et al. (2004). Clonality and diversity of the fish pathogen *Lactococcus garvieae* in Mediterranean countries. *Appl. Environ. Microbiol.* 70, 5132–5137. doi: 10.1128/AEM.70.9.5132-5137.2004
- Falush, D. (2009). Toward the use of genomics to study microevolutionary change in bacteria. *PLoS Genet.* 5:e1000627. doi: 10.1371/journal.pgen.1000627
- Falush, D. (2016). Bacterial genomics: microbial GWAS coming of age. *Nat. Microbiol.* 1:16059. doi: 10.1038/nmicrobiol.2016.59
- FAO (2013). *FAO Fisheries and Aquaculture Report No. 1053*. Available at: <http://www.fao.org/docrep/018/i3422e/i3422e.pdf>
- FAO (2014). *The State of World Fisheries and Aquaculture 2014*. Available at: <http://www.fao.org/3/a-i3720e.pdf>
- FAO (2015). *The State of Food Insecurity in the World 2015*. Available at: <http://www.fao.org/3/a-i4646e.pdf>
- FAO (2016). *The State of World Fisheries and Aquaculture 2016*. Available at: <http://www.fao.org/3/a-i5555e.pdf>
- Farhat, M. R., Shapiro, B. J., Sheppard, S. K., Colijn, C., and Murray, M. (2014). A phylogeny-based sampling strategy and power calculator informs genome-wide associations study design for microbial pathogens. *Genome Med.* 6:101. doi: 10.1186/s13073-014-0101-7
- Faria, N. R., Sabino, E. C., Nunes, M. R. T., Alcantara, L. C. J., Loman, N. J., and Pybus, O. G. (2016). Mobile real-time surveillance of Zika virus in Brazil. *Genome Med.* 8:97. doi: 10.1186/s13073-016-0356-2
- Feil, E. J. (2015). Toward a synthesis of genotypic typing and phenotypic inference in the genomics era. *Future Microbiol.* 10, 1897–1899. doi: 10.2217/fmb.15.113
- Ferrario, C., Ricci, G., Borgo, F., Rollando, A., and Fortina, M. G. (2012). Genetic investigation within *Lactococcus garvieae* revealed two genomic lineages. *FEMS Microbiol. Lett.* 332, 153–161. doi: 10.1111/j.1574-6968.2012.02591.x
- Ferrario, C., Ricci, G., Milani, C., Lugli, G. A., Ventura, M., Eraclio, G., et al. (2013). *Lactococcus garvieae*: where is it from? A first approach to explore the evolutionary history of this emerging pathogen. *PLoS ONE* 8:e84796. doi: 10.1371/journal.pone.0084796
- Figueiredo, H. C. P., Leal, C. A. G., Pereira, F. L., Soares, S. C., Gonçalves, L. A., Dorella, F. A., et al. (2016). Whole-genome sequence of *Francisella noatunensis* subsp. *orientalis* Strain FNO01 isolated from diseased Nile Tilapia in Brazil. *Genome Announc.* 4:e01603-15. doi: 10.1128/genomeA.01603-15
- Flegel, T. W. (2012). Historic emergence, impact and current status of shrimp pathogens in Asia. *J. Invertebr. Pathol.* 110, 166–173. doi: 10.1016/j.jip.2012.03.004
- Franz, E., Delaquis, P., Morabito, S., Beutin, L., Gobius, K., Rasko, D. A., et al. (2014). Exploiting the explosion of information associated with whole genome sequencing to tackle Shiga toxin-producing *Escherichia coli* (STEC) in global food production systems. *Int. J. Food Microbiol.* 187, 57–72. doi: 10.1016/j.ijfoodmicro.2014.07.002
- Galvani, A. P. (2003). Epidemiology meets evolutionary ecology. *Trends Ecol. Evol.* 18, 132–139. doi: 10.1016/S0169-5347(02)00050-2
- Garcia, S. M., and Rosenberg, A. A. (2010). Food security and marine capture fisheries: characteristics, trends, drivers and future perspectives. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 2869–2880. doi: 10.1098/rstb.2010.0171
- Gonçalves, L. A., de Castro Soares, S., Pereira, F. L., Dorella, F. A., de Carvalho, A. F., de Freitas Almeida, G. M., et al. (2016). Complete genome sequences of *Francisella noatunensis* subsp. *orientalis* strains FNO12, FNO24 and FNO190:

- a fish pathogen with genomic clonal behavior. *Stand. Genomic Sci.* 11:30. doi: 10.1186/s40793-016-0151-0
- Gordon, N. C., Price, J. R., Cole, K., Everitt, R., Morgan, M., Finney, J., et al. (2014). Prediction of *Staphylococcus aureus* antimicrobial resistance by whole-genome sequencing. *J. Clin. Microbiol.* 52, 1182–1191. doi: 10.1128/JCM.03117-13
- Goudenège, D., Travers, M. A., Lemire, A., Petton, B., Haffner, P., Labreuche, Y., et al. (2015). A single regulatory gene is sufficient to alter *Vibrio aestuarianus* pathogenicity in oysters. *Environ. Microbiol.* 17, 4189–4199. doi: 10.1111/1462-2920.12699
- Gulla, S., Lund, V., Kristoffersen, A. B., Sørsum, H., and Colquhoun, D. J. (2016). vapA (A-layer) typing differentiates *Aeromonas salmonicida* subspecies and identifies a number of previously undescribed subtypes. *J. Fish Dis.* 39, 329–342. doi: 10.1111/jfd.12367
- Habib, C., Houel, A., Lunazzi, A., Bernardet, J.-F., Olsen, A. B., Nilsen, H., et al. (2014). Multilocus sequence analysis of the marine bacterial genus *Tenacibaculum* suggests parallel evolution of fish pathogenicity and endemic colonization of aquaculture systems. *Appl. Environ. Microbiol.* 80, 5503–5514. doi: 10.1128/AEM.01177-14
- Hammoumi, S., Vallaeys, T., Santika, A., Leleux, P., Borzym, E., Klopp, C., et al. (2016). Targeted genomic enrichment and sequencing of CyHV-3 from carp tissues confirms low nucleotide diversity and mixed genotype infections. *PeerJ* 4, e2516. doi: 10.7717/peerj.2516
- Han, J. E., Kim, J. H., Shin, S. P., Jun, J. W., Chai, J. Y., and Park, S. C. (2013). Draft genome sequence of *Aeromonas salmonicida* subsp. *achromogenes* AS03, an atypical strain isolated from Crucian Carp (*Carassius carassius*) in the Republic of Korea. *Genome Announc.* 1:e00791-13. doi: 10.1128/genomeA.00791-13
- Hansen, H., and Onozaka, Y. (2011). When diseases hit aquaculture: an experimental study of spillover effects from negative publicity. *Mar. Resour. Econ.* 26, 281–291. doi: 10.5950/0738-1360-26.4.281
- Harris, S. R., Feil, E. J., Holden, M. T. G., Quail, M. A., Nickerson, E. K., Chantratita, N., et al. (2010). Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327, 469–474. doi: 10.1126/science.1182395
- Heath, C., Gillen, C. M., Chrysanthopoulos, P., Walker, M. J., and Barnes, A. C. (2016). *Streptococcus iniae* cpsG alters capsular carbohydrate composition and is a cause of serotype switching in vaccinated fish. *Vet. Microbiol.* 193, 116–124. doi: 10.1016/j.vetmic.2016.08.011
- Holden, M. T. G., Hsu, L.-Y., Kurt, K., Weinert, L. A., Mather, A. E., Harris, S. R., et al. (2013). A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant *Staphylococcus aureus* pandemic. *Genome Res.* 23, 653–664. doi: 10.1101/gr.147710.112
- Holm, K. O., Nilsson, K., Hjerde, E., Willassen, N.-P., and Milton, D. L. (2015). Complete genome sequence of *Vibrio anguillarum* strain NB10, a virulent isolate from the Gulf of Bothnia. *Stand. Genomic Sci.* 10:60. doi: 10.1186/s40793-015-0060-7
- Holt, K. E., Wertheim, H., Zadoks, R. N., Baker, S., Whitehouse, C. A., Dance, D., et al. (2015). Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health. *Proc. Natl. Acad. Sci. U.S.A.* 112, E3574–E3581. doi: 10.1073/pnas.1501049112
- Hossain, M. J., Sun, D., McGarey, D. J., Wrenn, S., Alexander, L. M., Martino, M. E., et al. (2014). An Asian origin of virulent *Aeromonas hydrophila* responsible for disease epidemics in United States-farmed catfish. *MBio* 5:e00848-14. doi: 10.1128/mBio.00848-14
- Howell, K. J., Weinert, L. A., Chaudhuri, R. R., Luan, S.-L., Peters, S. E., Corander, J., et al. (2014). The use of genome wide association methods to investigate pathogenicity, population structure and serovar in *Haemophilus parasuis*. *BMC Genomics* 15:1179. doi: 10.1186/1471-2164-15-1179
- Hsu, L.-Y., Harris, S. R., Chlebowicz, M. A., Lindsay, J. A., Koh, T.-H., Krishnan, P., et al. (2015). Evolutionary dynamics of methicillin-resistant *Staphylococcus aureus* within a healthcare system. *Genome Biol.* 16:81. doi: 10.1186/s13059-015-0643-z
- Ito, T., Kurita, J., Mori, K.-I., and Olesen, N. J. (2016). Virulence of viral haemorrhagic septicaemia virus (VHSV) genotype III in rainbow trout. *Vet. Res.* 47:4. doi: 10.1186/s13567-015-0303-z
- Jia, P., Zheng, X.-C., Shi, X.-J., Kan, S.-F., Wang, J.-J., He, J.-Q., et al. (2014). Determination of the complete genome sequence of infectious hematopoietic necrosis virus (IHN) Ch20101008 and viral molecular evolution in China. *Infect. Genet. Evol.* 27, 418–431. doi: 10.1016/j.meegid.2014.08.013
- Joshi, J., Srisala, J., Truong, V. H., Chen, I.-T., Nuangsang, B., Suthienkul, O., et al. (2014). Variation in *Vibrio parahaemolyticus* isolates from a single Thai shrimp farm experiencing an outbreak of acute hepatopancreatic necrosis disease (AHPND). *Aquaculture* 42, 297–302. doi: 10.1016/j.aquaculture.2014.03.030
- Kamath, P. L., Foster, J. T., Drees, K. P., Luikart, G., Quance, C., Anderson, N. J., et al. (2016). Genomics reveals historic and contemporary transmission dynamics of a bacterial disease among wildlife and livestock. *Nat. Commun.* 7:11448. doi: 10.1038/ncomms11448
- Kayansamruaj, P., Pirarat, N., Kondo, H., Hirono, I., and Rodkhum, C. (2015). Genomic comparison between pathogenic *Streptococcus agalactiae* isolated from Nile tilapia in Thailand and fish-derived ST7 strains. *Infect. Genet. Evol.* 36, 307–314. doi: 10.1016/j.meegid.2015.10.009
- Kennedy, D. A., Kurath, G., Brito, I. L., Purcell, M. K., Read, A. F., Winton, J. R., et al. (2016). Potential drivers of virulence evolution in aquaculture. *Evol. Appl.* 9, 344–354. doi: 10.1111/eva.12342
- Kim, J.-O., Kim, W.-S., Nishizawa, T., and Oh, M.-J. (2013). Complete genome sequence of viral hemorrhagic septicemia virus isolated from an olive flounder in South Korea. *Genome Announc.* 1:e00681-13. doi: 10.1128/genomeA.00681-13
- Kondo, H., Van, P. T., Dang, L. T., and Hirono, I. (2015). Draft genome sequence of non-*Vibrio parahaemolyticus* acute hepatopancreatic necrosis disease strain KC13.17.5, isolated from diseased shrimp in Vietnam. *Genome Announc.* 3:e00978-15. doi: 10.1128/genomeA.00978-15
- Kozakai, M., Matsumoto, M., Matsumoto, C., Uchida, S., Nagai, T., Satake, M., et al. (2016). First report of the isolation of *Lactococcus garvieae* from a platelet concentrate in Japan. *Transfusion* 56, 2602–2606. doi: 10.1111/trf.13752
- Krkosek, M., Connors, B. M., Morton, A., Lewis, M. A., Dill, L. M., and Hilborn, R. (2011). Effects of parasites from salmon farms on productivity of wild salmon. *Proc. Natl. Acad. Sci. U.S.A.* 108, 14700–14704. doi: 10.1073/pnas.1101845108
- Kumar, G., Menanteau-Ledouble, S., Saleh, M., and El-Matbouli, M. (2015). Yersinia agent of the causative agent of enteric redmouth disease in fish. *Vet. Res.* 46:103. doi: 10.1186/s13567-015-0238-4
- Kusuda, R., Kawai, K., Salati, F., Banner, C. R., and Fryer, J. L. (1991). *Enterococcus seriolicida* sp. nov., a fish pathogen. *Int. J. Syst. Bacteriol.* 41, 406–409. doi: 10.1099/00207713-41-3-406
- Laabei, M., and Massey, R. (2016). Using functional genomics to decipher the complexity of microbial pathogenicity. *Curr. Genet.* 62, 523–525. doi: 10.1007/s00294-016-0576-4
- Leclercq, E., Davie, A., and Migaud, H. (2014). Delousing efficiency of farmed ballan wrasse (*Labrus bergylta*) against *Lepeophtheirus salmonis* infecting Atlantic salmon (*Salmo salar*) post-smolts. *Pest Manag. Sci.* 70, 1274–1282. doi: 10.1002/ps.3692
- Lee, C.-T., Chen, I.-T., Yang, Y.-T., Ko, T.-P., Huang, Y.-T., Huang, J.-Y., et al. (2015). The opportunistic marine pathogen *Vibrio parahaemolyticus* becomes virulent by acquiring a plasmid that expresses a deadly toxin. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10798–10803. doi: 10.1073/pnas.1503129112
- Lesel, R., Lesel, M., Gavini, F., and Vuillaume, A. (1983). Outbreak of enteric redmouth disease in rainbow trout, *Salmo gairdneri* Richardson, in France. *J. Fish Dis.* 6, 385–387. doi: 10.1111/j.1365-2761.1983.tb00091.x
- Letchumanan, V., Ser, H.-L., Chan, K.-G., Goh, B.-H., and Lee, L.-H. (2016). Genome sequence of *Vibrio parahaemolyticus* VP103 strain isolated from shrimp in Malaysia. *Front. Microbiol.* 7:1496. doi: 10.3389/fmicb.2016.01496
- Li, G., Mo, Z., Li, J., Xiao, P., and Hao, B. (2013). Complete genome sequence of *Vibrio anguillarum* M3, a serotype O1 strain isolated from Japanese Flounder in China. *Genome Announc.* 1:e00769-13. doi: 10.1128/genomeA.00769-13
- Lillebaek, T., Norman, A., Rasmussen, E. M., Marvig, R. L., Folkvardsen, D. B., Andersen, ÅB., et al. (2016). Substantial molecular evolution and mutation rates in prolonged latent *Mycobacterium tuberculosis* infection in humans. *Int. J. Med. Microbiol.* 306, 580–585. doi: 10.1016/j.ijmm.2016.05.017
- Lillehaug, A., Lunestad, B. T., and Grave, K. (2003). Epidemiology of bacterial diseases in Norwegian aquaculture—a description based on antibiotic prescription data for the ten-year period 1991 to 2000. *Dis. Aquat. Organ.* 53, 115–125. doi: 10.3354/dao053115
- Lindsay, J. A. (2014). *Staphylococcus aureus* genomics and the impact of horizontal gene transfer. *Int. J. Med. Microbiol.* 304, 103–109. doi: 10.1016/j.ijmm.2013.11.010

- Liu, G., Zhang, W., and Lu, C. (2013). Comparative genomics analysis of *Streptococcus agalactiae* reveals that isolates from cultured tilapia in China are closely related to the human strain A909. *BMC Genomics* 14:775. doi: 10.1186/1471-2164-14-775
- Liu, T., Wang, K.-Y., Wang, J., Chen, D.-F., Huang, X.-L., Ouyang, P., et al. (2016). Genome sequence of the fish pathogen *Yersinia ruckeri* SC09 provides insights into niche adaptation and pathogenic mechanism. *Int. J. Mol. Sci.* 17:557. doi: 10.3390/ijms17040557
- Maiden, M. C. J. (2006). Multilocus sequence typing of bacteria. *Annu. Rev. Microbiol.* 60, 561–588. doi: 10.1146/annurev.micro.59.030804.121325
- Maiden, M. C. J., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., et al. (1998). Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. U.S.A.* 95, 3140–3145. doi: 10.1073/pnas.95.6.3140
- Maiden, M. C. J., Jansen van Rensburg, M. J., Bray, J. E., Earle, S. G., Ford, S. A., Jolley, K. A., et al. (2013). MLST revisited: the gene-by-gene approach to bacterial genomics. *Nat. Rev. Microbiol.* 11, 728–736. doi: 10.1038/nrmicro3093
- Marcos-López, M., Gale, P., Oidtmann, B. C., and Peeler, E. J. (2010). Assessing the impact of climate change on disease emergence in freshwater fish in the United Kingdom. *Transbound. Emerg. Dis.* 57, 293–304. doi: 10.1111/j.1865-1682.2010.01150.x
- Marcos-López, M., and Gale, P. (2010). Assessing the impact of climate change on disease emergence in freshwater fish in the United Kingdom. *Transbound. Emerg. Dis.* 57, 293–304. doi: 10.1111/j.1865-1682.2010.01150.x
- Mari, J., Poulos, B. T., Lightner, D. V., and Bonami, J.-R. (2002). Shrimp Taura syndrome virus: genomic characterization and similarity with members of the genus cricket paralysis-like viruses. *J. Gen. Virol.* 83, 915–926. doi: 10.1099/0022-1317-83-4-915
- Markussen, T., Jonassen, C. M., Numanovic, S., Braeen, S., Hjortaa, M., Nilsen, H., et al. (2008). Evolutionary mechanisms involved in the virulence of infectious salmon anaemia virus (ISAV), a piscine orthomyxovirus. *Virology* 374, 515–527. doi: 10.1016/j.virol.2008.01.019
- Marty, G. D., Saksida, S. M., and Quinn, T. J. II (2010). Relationship of farm salmon, sea lice, and wild salmon populations. *Proc. Natl. Acad. Sci. U.S.A.* 107, 22599–22604. doi: 10.1073/pnas.1009573108
- Matejusova, I., Bain, N., Colquhoun, D. J., Feil, E. J., McCarthy, U., McLennan, D., et al. (2013). Multilocus variable-number tandem-repeat genotyping of *Renibacterium salmoninarum*, a bacterium causing bacterial kidney disease in salmonid fish. *BMC Microbiol.* 13:285. doi: 10.1186/1471-2180-13-285
- Matejusova, I., Noguera, P. A., Hall, M., McBeath, A. J. A., Urquhart, K., Simons, J., et al. (2016). Susceptibility of goldsinny wrasse, *Ctenolabrus rupestris* L. (Labridae), to viral haemorrhagic septicaemia virus (VHSV) genotype III: experimental challenge and pathology. *Vet. Microbiol.* 186, 164–173. doi: 10.1016/j.vetmic.2016.02.022
- Mather, A. E., Lawson, B., de Pinna, E., Wigley, P., Parkhill, J., Thomson, N. R., et al. (2016). Genomic analysis of *Salmonella enterica* serovar *Typhimurium* from wild passerines in England and Wales. *Appl. Environ. Microbiol.* 82, 6728–6735. doi: 10.1128/AEM.01660-16
- Mérour, E., LeBerre, M., Lamoureux, A., Bernard, J., Brémont, M., and Biacchesi, S. (2011). Completion of the full-length genome sequence of the infectious salmon anaemia virus, an aquatic orthomyxovirus-like, and characterization of mAbs. *J. Gen. Virol.* 92, 528–533. doi: 10.1099/vir.0.027417-0
- Morita, H., Toh, H., Oshima, K., Yoshizaki, M., Kawanishi, M., Nakaya, K., et al. (2011). Complete genome sequence and comparative analysis of the fish pathogen *Lactococcus garvieae*. *PLoS ONE* 6:e23184. doi: 10.1371/journal.pone.0023184
- Munro, E. S., McIntosh, R. E., Weir, S. J., Noguera, P. A., Sandilands, J. M., Matejusova, I., et al. (2015). A mortality event in wrasse species (Labridae) associated with the presence of viral haemorrhagic septicaemia virus. *J. Fish Dis.* 38, 335–341. doi: 10.1111/jfd.12237
- Murray, A. G., and Peeler, E. J. (2005). A framework for understanding the potential for emerging diseases in aquaculture. *Prev. Vet. Med.* 67, 223–235. doi: 10.1016/j.prevetmed.2004.10.012
- Naka, H., Dias, G. M., Thompson, C. C., Dubay, C., Thompson, F. L., and Crosa, J. H. (2011). Complete genome sequence of the marine fish pathogen *Vibrio anguillarum* harboring the pJM1 virulence plasmid and genomic comparison with other virulent strains of *V. anguillarum* and *V. ordalii*. *Infect. Immun.* 79, 2889–2900. doi: 10.1128/IAI.05138-11
- Neiger, R., Thomas, M., Das, S., Barnes, M., Fletcher, B., Snekvik, K., et al. (2016). Draft genome sequences of three *Flavobacterium psychrophilum* strains isolated from coldwater disease outbreaks at three production hatcheries. *Genome Announc.* 4:e00035-16. doi: 10.1128/genomeA.00035-16
- Nelson, M. C., LaPatra, S. E., Welch, T. J., and Graf, J. (2015). Complete genome sequence of *Yersinia ruckeri* strain CSF007-82, etiologic agent of Red Mouth Disease in Salmonid Fish. *Genome Announc.* 3:e01491-14. doi: 10.1128/genomeA.01491-14
- Nelson, M. C., Varney, J. S., Welch, T. J., and Graf, J. (2016). Draft genome sequence of *Lactococcus garvieae* strain PAQ102015-99, an outbreak strain isolated from a commercial trout farm in the Northwestern United States. *Genome Announc.* 4:e00781-16. doi: 10.1128/genomeA.00781-16
- Nicolas, P., Mondot, S., Achaz, G., Bouchenet, C., Bernardet, J.-F., and Duchaud, E. (2008). Population structure of the fish-pathogenic bacterium *Flavobacterium psychrophilum*. *Appl. Environ. Microbiol.* 74, 3702–3709. doi: 10.1128/AEM.00244-08
- Nilsen, H., Sundell, K., Duchaud, E., Nicolas, P., Dalsgaard, I., Madsen, L., et al. (2014). Multilocus sequence typing identifies epidemic clones of *Flavobacterium psychrophilum* in Nordic countries. *Appl. Environ. Microbiol.* 80, 2728–2736. doi: 10.1128/AEM.04233-13
- Nunan, L., Lightner, D., Pantoja, C., and Gomez-Jimenez, S. (2014). Detection of acute hepatopancreatic necrosis disease (AHPND) in Mexico. *Dis. Aquat. Organ.* 111, 81–86. doi: 10.3354/dao02776
- Oidtmann, B., Peeler, E., Lyngstad, T., Brun, E., Bang Jensen, B., and Stärk, K. D. C. (2013). Risk-based methods for fish and terrestrial animal disease surveillance. *Prev. Vet. Med.* 112, 13–26. doi: 10.1016/j.prevetmed.2013.07.008
- OIE (2016). *Aquatic Animal Health Code*. Available at: <http://www.oie.int/international-standard-setting/aquatic-code/>
- Ormsby, M. J., Caws, T., Burchmore, R., Wallis, T., Verner-Jeffreys, D. W., and Davies, R. L. (2016). *Yersinia ruckeri* isolates recovered from diseased Atlantic salmon (*Salmo salar*) in Scotland are more diverse than those from rainbow trout (*Oncorhynchus mykiss*) and represent distinct sub-populations. *Appl. Environ. Microbiol.* 82, 5785–5794. doi: 10.1128/AEM.01173-16
- Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., et al. (2015). Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31, 3691–3693. doi: 10.1093/bioinformatics/btv421
- Palacios, M. A., Zamora, M. J., Velazquez, J., and Zamora, E. (1993). *Streptococcus* in rainbow trout (*Oncorhynchus mykiss*) in Spain. *Boll. Soc. It. Patol. Ittica* 13, 11–16.
- Pereira Ude, P., Rodrigues Dos Santos, A., Hassan, S. S., Aburjaile, F. F., Soares Sde, C., Ramos, R. T., et al. (2013). Complete genome sequence of *Streptococcus agalactiae* strain SA20-06, a fish pathogen associated to meningoencephalitis outbreaks. *Stand. Genomic Sci.* 8, 188–197. doi: 10.4056/sigs.3687314
- Pérez-Losada, M., Cabezas, P., Castro-Nallar, E., and Crandall, K. A. (2013). Pathogen typing in the genomics era: MLST and the future of molecular epidemiology. *Infect. Genet. Evol.* 16, 38–53. doi: 10.1016/j.meegid.2013.01.009
- Power, A. G., and Mitchell, C. E. (2004). Pathogen spillover in disease epidemics. *Am. Nat.* 164(Suppl. 5), S79–S89. doi: 10.1086/424610
- Pridgeon, J. W., Zhang, D., and Zhang, L. (2014). Complete genome sequence of a virulent strain, *Streptococcus iniae* ISET0901, isolated from diseased Tilapia. *Genome Announc.* 2:e00553-14. doi: 10.1128/genomeA.00553-14
- Pulgar, R., Travisany, D., Zuñiga, A., Maass, A., and Cambiazo, V. (2015). Complete genome sequence of *Piscirickettsia salmonis* LF-89 (ATCC VR-1361) a major pathogen of farmed salmonid fish. *J. Biotechnol.* 212, 30–31. doi: 10.1016/j.jbiotec.2015.07.017
- Pulkkinen, K., Suomalainen, L.-R., Read, A. F., Ebert, D., Rintamäki, P., and Valtonen, E. T. (2010). Intensive fish farming and the evolution of pathogen virulence: the case of columnaris disease in Finland. *Proc. Biol. Sci.* 277, 593–600. doi: 10.1098/rspb.2009.1659
- Quick, J., Loman, N. J., Duraffour, S., Simpson, J. T., Severi, E., Cowley, L., et al. (2016). Real-time, portable genome sequencing for Ebola surveillance. *Nature* 530, 228–232. doi: 10.1038/nature16996
- Rakus, K., Ouyang, P., Boutier, M., Ronsmans, M., Reschner, A., Vancsok, C., et al. (2013). Cyprinid herpesvirus 3: an interesting virus for applied and fundamental research. *Vet. Res.* 44:85. doi: 10.1186/1297-9716-44-85
- Rasmussen-Ivey, C. R., Hossain, M. J., Odom, S. E., Terhune, J. S., Hemstreet, W. G., Shoemaker, C. A., et al. (2016). Classification of a hypervirulent *Aeromonas*

- hydrophila* pathotype responsible for epidemic outbreaks in warm-water fishes. *Front. Microbiol.* 7:e1615. doi: 10.3389/fmicb.2016.01615
- Reith, M. E., Singh, R. K., Curtis, B., Boyd, J. M., Bouevitch, A., Kimball, J., et al. (2008). The genome of *Aeromonas salmonicida* subsp. *salmonicida* A449: insights into the evolution of a fish pathogen. *BMC Genomics* 9:427. doi: 10.1186/1471-2164-9-427
- Reuter, S., Török, M. E., Holden, M. T. G., Reynolds, R., Raven, K. E., Blane, B., et al. (2016). Building a genomic framework for prospective MRSA surveillance in the United Kingdom and the Republic of Ireland. *Genome Res.* 26, 263–270. doi: 10.1101/gr.196709.115
- Ricci, G., Ferrario, C., Borgo, F., Rollando, A., and Fortina, M. G. (2012). Genome sequences of *Lactococcus garvieae* TB25, isolated from Italian cheese, and *Lactococcus garvieae* LG9, isolated from Italian rainbow trout. *J. Bacteriol.* 194, 1249–1250. doi: 10.1128/JB.06655-11
- Rodriguez-Anaya, L. Z., Gonzalez-Galaviz, J. R., Casillas-Hernandez, R., Lares-Villa, F., Estrada, K., Ibarra-Gamez, J. C., et al. (2016). Draft Genome sequence of white spot syndrome virus isolated from cultured *Litopenaeus vannamei* in Mexico. *Genome Announc.* 4:e01674-15. doi: 10.1128/genomeA.01674-15
- Roger, F., Marchandin, H., Jumas-Bilak, E., Kodjo, A., colBVH study group, and Lamy, B. (2012). Multilocus genetics to reconstruct aeromonad evolution. *BMC Microbiol.* 12:62. doi: 10.1186/1471-2180-12-62
- Romalde, J. L., Magariños, B., Barja, J. L., and Toranzo, A. E. (1993). Antigenic and molecular characterization of *Yersinia ruckeri* proposal for a new intraspecies classification. *Syst. Appl. Microbiol.* 16, 411–419. doi: 10.1016/S0723-2020(11)80274-2
- Ronholm, J., Naseri, N., Petronella, N., and Pagotto, F. (2016). Navigating microbiological food safety in the era of whole-genome sequencing. *Clin. Microbiol. Rev.* 29, 837–857. doi: 10.1128/CMR.00056-16
- Schleifer, K. H., Kraus, J., Dvorak, C., Kilpper-Bälz, R., Collins, M. D., and Fischer, W. (1985). Transfer of *Streptococcus lactis* and related streptococci to the genus *Lactococcus* gen. nov. *Syst. Appl. Microbiol.* 6, 183–195. doi: 10.1016/S0723-2020(85)80052-7
- Schönherz, A. A., Lorenzen, N., Guldbandsen, B., Buitenhuis, B., and Einer-Jensen, K. (2016). Ultra-deep sequencing of VHSV isolates contributes to understanding the role of viral quasispecies. *Vet. Res.* 47:10. doi: 10.1186/s13567-015-0298-5
- Schütze, H., Mundt, E., and Mettenleiter, T. C. (1999). Complete genomic sequence of viral hemorrhagic septicemia virus, a fish rhabdovirus. *Virus Genes* 19, 59–65. doi: 10.1023/A:1008140707132
- Senn, L., Clerc, O., Zanetti, G., Basset, P., Prod'hom, G., Gordon, N. C., et al. (2016). The stealthy superbug: the role of asymptomatic enteric carriage in maintaining a long-term hospital outbreak of ST228 methicillin-resistant *Staphylococcus aureus*. *MBio* 7:e02039-15. doi: 10.1128/mBio.02039-15
- Seth-Smith, H. M. B., Dourala, N., Fehr, A., Qi, W., Katharios, P., Ruetten, M., et al. (2016). Emerging pathogens of gilthead seabream: characterisation and genomic analysis of novel intracellular β -proteobacteria. *ISME J.* 10, 1791–1803. doi: 10.1038/ismej.2015.223
- Shah, S. Q. A., Nilsen, H., Bottolfsen, K., Colquhoun, D. J., and Sørum, H. (2012). DNA gyrase and topoisomerase IV mutations in quinolone-resistant *Flavobacterium psychrophilum* isolated from diseased salmonids in Norway. *Microb. Drug Resist.* 18, 207–214. doi: 10.1089/mdr.2011.0142
- Sharifyazdi, H., Akhlaghi, M., and Tabatabaei, M. (2010). Isolation and characterization of *Lactococcus garvieae* from diseased rainbow trout (*Oncorhynchus mykiss*, Walbaum) cultured in Iran. *Iran. J. Vet. Res.* 11, 342–350.
- Sheppard, S. K., Didelot, X., Meric, G., Torralbo, A., Jolley, K. A., Kelly, D. J., et al. (2013). Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in *Campylobacter*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11923–11927. doi: 10.1073/pnas.1305591110
- Sheppard, S. K., Jolley, K. A., and Maiden, M. C. J. (2012). A gene-by-gene approach to bacterial population genomics: whole genome MLST of *Campylobacter*. *Genes* 3, 261–277. doi: 10.3390/genes3020261
- Siekoula-Nguedia, C., Blanc, G., Duchaud, E., and Calvez, S. (2012). Genetic diversity of *Flavobacterium psychrophilum* isolated from rainbow trout in France: predominance of a clonal complex. *Vet. Microbiol.* 161, 169–178. doi: 10.1016/j.vetmic.2012.07.022
- Sittidilokratna, N., Dangtip, S., Cowley, J. A., and Walker, P. J. (2008). RNA transcription analysis and completion of the genome sequence of yellow head nidovirus. *Virus Res.* 136, 157–165. doi: 10.1016/j.virusres.2008.05.008
- Skallerup, P., Espinosa-Gongora, C., Jørgensen, C. B., Guardabassi, L., and Fredholm, M. (2015). Genome-wide association study reveals a locus for nasal carriage of *Staphylococcus aureus* in Danish crossbred pigs. *BMC Vet. Res.* 11:290. doi: 10.1186/s12917-015-0599-y
- Smith, J. M., Smith, N. H., O'Rourke, M., and Spratt, B. G. (1993). How clonal are bacteria? *Proc. Natl. Acad. Sci. U.S.A.* 90, 4384–4388. doi: 10.1073/pnas.90.10.4384
- Snow, M. (2011). The contribution of molecular epidemiology to the understanding and control of viral diseases of salmonid aquaculture. *Vet. Res.* 42:56. doi: 10.1186/1297-9716-42-56
- Snow, M., Bain, N., Black, J., Taupin, V., Cunningham, C. O., King, J. A., et al. (2004). Genetic population structure of marine viral haemorrhagic septicaemia virus (VHSV). *Dis. Aquat. Organ.* 61, 11–21. doi: 10.3354/dao061011
- Snow, M., Cunningham, C. O., Melvin, W. T., and Kurath, G. (1999). Analysis of the nucleoprotein gene identifies distinct lineages of viral haemorrhagic septicaemia virus within the European marine environment. *Virus Res.* 63, 35–44. doi: 10.1016/S0168-1702(99)00056-8
- Soto-Rodriguez, S. A., Gomez-Gil, B., Lozano-Olvera, R., Betancourt-Lozano, M., and Morales-Covarrubias, M. S. (2015). Field and experimental evidence of *Vibrio parahaemolyticus* as the causative agent of acute hepatopancreatic necrosis disease of cultured shrimp (*Litopenaeus vannamei*) in Northwestern Mexico. *Appl. Environ. Microbiol.* 81, 1689–1699. doi: 10.1128/AEM.03610-14
- Sridhar, S., Sharma, A., Kongshaug, H., Nilsen, F., and Jonassen, I. (2012). Whole genome sequencing of the fish pathogen *Francisella noatunensis* subsp. *orientalis* Toba04 gives novel insights into *Francisella* evolution and pathogenicity. *BMC Genomics* 13:598. doi: 10.1186/1471-2164-13-598
- Stepien, C. A., Pierce, L. R., Leaman, D. W., Niner, M. D., and Shepherd, B. S. (2015). Gene diversification of an emerging pathogen: a decade of mutation in a novel fish viral hemorrhagic septicemia (VHS) substrain since its first appearance in the Laurentian Great Lakes. *PLoS ONE* 10:e0135146. doi: 10.1371/journal.pone.0135146
- Strepparava, N., Nicolas, P., Wahli, T., Segner, H., and Petrini, O. (2013). Molecular epidemiology of *Flavobacterium psychrophilum* from Swiss fish farms. *Dis. Aquat. Organ.* 105, 203–210. doi: 10.3354/dao02609
- Sudheesh, P. S., Al-Ghabshi, A., Al-Mazrooei, N., and Al-Habsi, S. (2012). Comparative pathogenomics of bacteria causing infectious diseases in fish. *Int. J. Evol. Biol.* 2012:457264. doi: 10.1155/2012/457264
- Sundberg, L.-R., Ketola, T., Laanto, E., Kinnula, H., Bamford, J. K. H., Penttinen, R., et al. (2016). Intensive aquaculture selects for increased virulence and interference competition in bacteria. *Proc. Biol. Sci.* 283:20153069. doi: 10.1098/rspb.2015.3069
- Tran, L., Nunan, L., Redman, R. M., Mohney, L. L., Pantoja, C. R., Fitzsimmons, K., et al. (2013). Determination of the infectious nature of the agent of acute hepatopancreatic necrosis syndrome affecting penaeid shrimp. *Dis. Aquat. Organ.* 105, 45–55. doi: 10.3354/dao02621
- UN (2015). *The World Population Prospects: 2015 Revision*. Available at: https://esa.un.org/unpd/wpp/publications/files/key_findings_wpp_2015.pdf
- Urwin, R., and Maiden, M. C. J. (2003). Multi-locus sequence typing: a tool for global epidemiology. *Trends Microbiol.* 11, 479–487. doi: 10.1016/j.tim.2003.08.006
- Valdes, N., Espinoza, C., Sanhueza, L., Gonzalez, A., Corsini, G., and Tello, M. (2015). Draft genome sequence of the Chilean isolate *Aeromonas salmonicida* strain CBA100. *FEMS Microbiol. Lett.* 362:fnu062. doi: 10.1093/femsle/fnu062
- Vela, A. I., Vázquez, J., Gibello, A., Blanco, M. M., Moreno, M. A., Liébana, P., et al. (2000). Phenotypic and genetic characterization of *Lactococcus garvieae* isolated in Spain from lactococcosis outbreaks and comparison with isolates of other countries and sources. *J. Clin. Microbiol.* 38, 3791–3795.
- Vezzulli, L., Grande, C., Reid, P. C., Hélaouët, P., Edwards, M., Höfle, M. G., et al. (2016). Climate influence on *Vibrio* and associated human diseases during the past half-century in the coastal North Atlantic. *Proc. Natl. Acad. Sci. U.S.A.* 113, E5062–E5071. doi: 10.1073/pnas.1609157113
- Vincent, A. T., Tanaka, K. H., Trudel, M. V., Frenette, M., Derome, N., and Charette, S. J. (2015). Draft genome sequences of two *Aeromonas salmonicida* subsp. *salmonicida* isolates harboring plasmids conferring antibiotic resistance. *FEMS Microbiol. Lett.* 362, 1–4. doi: 10.1093/femsle/fnv002

- Vincent, A. T., Trudel, M. V., Freschi, L., Nagar, V., Gagné-Thivierge, C., Levesque, R. C., et al. (2016). Increasing genomic diversity and evidence of constrained lifestyle evolution due to insertion sequences in *Aeromonas salmonicida*. *BMC Genomics* 17:44. doi: 10.1186/s12864-016-2381-3
- Wakabayashi, H., Horiuchi, M., Bunya, T., and Hoshiai, G. (1991). Outbreaks of cold-water disease in coho salmon in Japan. *Fish Pathol.* 26, 211–212. doi: 10.3147/jsfp.26.211
- Walker, T. M., Kohl, T. A., Omar, S. V., Hedge, J., Del Ojo Elias, C., Bradley, P., et al. (2015). Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: a retrospective cohort study. *Lancet Infect. Dis.* 15, 1193–1202. doi: 10.1016/S1473-3099(15)00062-6
- Wang, C., Zhao, L. L., Li, Y. J., Tang, L. J., Qiao, X. Y., Jiang, Y. P., et al. (2016). Analysis of the genome sequence of infectious hematopoietic necrosis virus HLJ-09 in China. *Virus Genes* 52, 29–37. doi: 10.1007/s11262-015-1263-0
- Wang, K.-Y., Liu, T., Wang, J., Chen, D.-F., Wu, X.-J., Jiang, J., et al. (2015). Complete genome sequence of the fish pathogen *Yersinia ruckeri* strain SC09, isolated from diseased *Ictalurus punctatus* in China. *Genome Announc* 3:e01327-14. doi: 10.1128/genomeA.01327-14
- Welch, T. J., Verner-Jeffreys, D. W., Dalsgaard, I., Wiklund, T., Evenhuis, J. P., Cabrera, J. A. G., et al. (2011). Independent emergence of *Yersinia ruckeri* biotype 2 in the United States and Europe. *Appl. Environ. Microbiol.* 77, 3493–3499. doi: 10.1128/AEM.02997-10
- Wheeler, R. W., Davies, R. L., Dalsgaard, I., Garcia, J., Welch, T. J., Wagley, S., et al. (2009). *Yersinia ruckeri* biotype 2 isolates from mainland Europe and the UK likely represent different clonal groups. *Dis. Aquat. Organ.* 84, 25–33. doi: 10.3354/dao02039
- Wiens, G. D., Rockey, D. D., Wu, Z., Chang, J., Levy, R., Crane, S., et al. (2008). Genome sequence of the fish pathogen *Renibacterium salmoninarum* suggests reductive evolution away from an environmental *Arthrobacter* ancestor. *J. Bacteriol.* 190, 6970–6982. doi: 10.1128/JB.00721-08
- Wijegoonawardane, P. K. M., Sittidilokratna, N., Petchampai, N., Cowley, J. A., Gudkovs, N., and Walker, P. J. (2009). Homologous genetic recombination in the yellow head complex of nidoviruses infecting *Penaeus monodon* shrimp. *Virology* 390, 79–88. doi: 10.1016/j.virol.2009.04.015
- Willems, S., Kampmeier, S., Bletz, S., Kossow, A., Köck, R., Kipp, F., et al. (2016). Whole-genome sequencing elucidates epidemiology of nosocomial clusters of *Acinetobacter baumannii*. *J. Clin. Microbiol.* 54, 2391–2394. doi: 10.1128/JCM.00721-16
- Willumsen, B. (1989). Birds and wild fish as potential vectors of *Yersinia ruckeri*. *J. Fish Dis.* 12, 275–277. doi: 10.1111/j.1365-2761.1989.tb00313.x
- Woo, P. T. K., Leatherland, J. F., and Bruno, D. W. (2011). *Fish Diseases and Disorders*. Available at: https://books.google.co.uk/books?id=S0w3fB_PDNIC
- Wu, A. K., Kropinski, A. M., Lumsden, J. S., Dixon, B., and MacInnes, J. I. (2015). Complete genome sequence of the fish pathogen *Flavobacterium psychrophilum* ATCC 49418(T.). *Stand. Genomic Sci.* 10:3. doi: 10.1186/1944-3277-10-3
- Yahara, K., Didelot, X., Jolley, K. A., Kobayashi, I., Maiden, M. C. J., Sheppard, S. K., et al. (2016). The Landscape of realized homologous recombination in pathogenic bacteria. *Mol. Biol. Evol.* 33, 456–471. doi: 10.1093/molbev/msv237
- Yang, F., He, J., Lin, X., Li, Q., Pan, D., Zhang, X., et al. (2001). Complete genome sequence of the shrimp white spot bacilliform virus. *J. Virol.* 75, 11811–11820. doi: 10.1128/JVI.75.23.11811-11820.2001
- Yang, Y.-T., Chen, I.-T., Lee, C.-T., Chen, C.-Y., Lin, S.-S., Hor, L.-I., et al. (2014). Draft genome sequences of four strains of *Vibrio parahaemolyticus*, three of which cause early mortality syndrome/acute hepatopancreatic necrosis disease in Shrimp in China and Thailand. *Genome Announc* 2:e00816-14. doi: 10.1128/genomeA.00816-14
- Yin, L. K. (ed.) (2004). *Current Trends in the Study of Bacterial and Viral Fish and Shrimp Diseases*. Available at: <http://www.worldscientific.com/worldscibooks/10.1142/5465>
- Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., et al. (2012). Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* 67, 2640–2644. doi: 10.1093/jac/dks261
- Zhang, B. C., Zhang, J., and Sun, L. (2014). *Streptococcus iniae* SF1: complete genome sequence, proteomic profile, and immunoprotective antigens. *PLoS ONE* 9:e91324. doi: 10.1371/journal.pone.0091324
- Zhang, X., Davenport, K. W., Gu, W., Daligault, H. E., Munk, A. C., Tashima, H., et al. (2012). Improving genome assemblies by sequencing PCR products with PacBio. *Biotechniques* 53, 61–62. doi: 10.2144/0000113891

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Bayliss, Verner-Jeffreys, Bartie, Aanensen, Sheppard, Adams and Feil. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.