Epidemics 14 (2016) 26–35

Contents lists available at ScienceDirect

Epidemics

journal homepage: www.elsevier.com/locate/epidemics

Use of bacterial whole-genome sequencing to investigate local persistence and spread in bovine tuberculosis

Hannah Trewby[a], David Wright[b], Eleanor L. Breadon[c], Samantha J. Lycett[d], Tom R. Mallon[c], Carl McCormick[c], Paul Johnson[a], Richard J. Orton[a], Adrian R. Allen[c], Julie Galbraith[e], Pawel Herzyk[e], Robin A. Skuce[b,c], Roman Biek[a], Rowland R. Kao[a,*]

[a] Boyd Orr Centre for Population and Ecosystem Health, Institute for Biodiversity Animal Health and Comparative Medicine, University of Glasgow, Glasgow, UK
[b] School of Biological Sciences, Queen's University Belfast, Belfast, Northern Ireland, UK
[c] Veterinary Sciences Division, Agri-Food and Biosciences Institute, Stormont, Belfast, Northern Ireland, UK
[d] The Roslin Institute, University of Edinburgh, Edinburgh, UK
[e] Glasgow Polyomics, College of Medical Veterinary and Life Sciences, University of Glasgow, Glasgow, UK

ARTICLE INFO

Article history:
Received 16 December 2014
Received in revised form 13 July 2015
Accepted 17 August 2015
Available online 19 October 2015

Keywords:
Bacterial evolution
Livestock disease
Molecular epidemiology
Mycobacterium bovis
Phylogeography

ABSTRACT

Mycobacterium bovis is the causal agent of bovine tuberculosis, one of the most important diseases currently facing the UK cattle industry. Here, we use high-density whole genome sequencing (WGS) in a defined sub-population of *M. bovis* in 145 cattle across 66 herd breakdowns to gain insights into local spread and persistence. We show that despite low divergence among isolates, WGS can in principle expose contributions of under-sampled host populations to *M. bovis* transmission. However, we demonstrate that in our data such a signal is due to molecular type switching, which had been previously undocumented for *M. bovis*. Isolates from farms with a known history of direct cattle movement between them did not show a statistical signal of higher genetic similarity. Despite an overall signal of genetic isolation by distance, genetic distances also showed no apparent relationship with spatial distance among affected farms over distances <5 km. Using simulations, we find that even over the brief evolutionary timescale covered by our data, Bayesian phylogeographic approaches are feasible. Applying such approaches showed that *M. bovis* dispersal in this system is heterogeneous but slow overall, averaging 2 km/year. These results confirm that widespread application of WGS to *M. bovis* will bring novel and important insights into the dynamics of *M. bovis* spread and persistence, but that the current questions most pertinent to control will be best addressed using approaches that more directly integrate WGS with additional epidemiological data.

1. Introduction

The increasing availability of bacterial whole-genome sequence (WGS) data now makes it possible to generate sequence datasets for whole bacterial pathogen populations at high sampling densities. Such comprehensive sequencing has yielded impressive advances in outbreak investigation (Eyre et al., 2013; Harris et al., 2010; Walker et al., 2012), and provided new insights into both spatial dissemination (Gray et al., 2011; Holden et al., 2013) and the complexities of multi-host pathogen systems (Mather et al., 2013; Viana et al., 2014). However, even at the genomic scale the rates of evolutionary change estimated for bacteria can be substantially lower than those commonly seen in rapidly evolving pathogens such as RNA viruses (Biek et al., 2015; Bryant et al., 2013b). The extent to which slow evolution constrains the type and scale of epidemiological processes that can be resolved for bacterial pathogens, and which analytical approaches are most appropriate to deal with this, remains unclear for many systems.

Mycobacterium bovis is one of a group of closely related bacteria which includes the primary cause of human tuberculosis, *M. tuberculosis*, a pathogen estimated to evolve at a rate of around 0.3–0.5 mutations per genome per year over epidemiological timescales (Bryant et al., 2013b; Walker et al., 2012). *M. bovis* is the causative agent of bovine tuberculosis (bTB), an important disease of cattle and other mammals including man. Herd-to-herd movements of infected cows among farms (Gilbert et al., 2005; Green et al.,

* Corresponding author. Tel.: +44 1413305770.
E-mail address: rowland.kao@glasgow.ac.uk (R.R. Kao).

http://dx.doi.org/10.1016/j.epidem.2015.08.003

2008) and infection in Eurasian badger (*Meles meles*) populations (Delahay et al., 2001; Gallagher and Clifton-Hadley, 2000) have both been implicated in the spread of bTB in Britain and Ireland. While much attention has focused on the relative roles of badgers and cattle in the maintenance of bTB, recent studies emphasising the importance of cattle have highlighted the continuing need for a deeper understanding of the role of cattle-based transmission (Brooks-Pollock et al., 2014; Donnelly and Nouvellet, 2013).

Molecular typing of *M. bovis* isolates based on repeated genetic elements has been advocated for some time to aid in the epidemiology and control of bTB (Cousins et al., 1998; Skuce and Neill, 2001), and in Britain and Ireland these typing methods have shown that *M. bovis* molecular types are maintained within well-defined geographic clusters (Skuce et al., 2010; Smith et al., 2006). While such molecular typing has proved useful for identifying local clustering on larger scales, their power to discriminate within-cluster events involved in fine-scale persistence and spread of bTB is limited.

In a previous study, Biek et al. (2012) established the potential of bacterial WGS in investigating the epidemiology of bTB at a local (i.e., farm-to-farm) scale. By sequencing 30 bacterial isolates from a spatially dense cluster of bTB cases within one recently emerged *M. bovis* molecular type (VNTR-10) in Northern Ireland (NI), the study demonstrated (i) close relatedness of bacteria isolated from cattle and badgers, (ii) persistence of bacterial lineages on the same farm, and (iii) that genetic similarity between isolates correlated with geographic distance between sampling locations. The study also showed that, due to slow evolution, even WGS is unlikely to provide sufficient resolution to resolve transmission at the animal-to-animal scale for *M. bovis*, similar to findings in human tuberculosis (Bryant et al., 2013b; Roetzer et al., 2013; Walker et al., 2012), and is more suited to do so at the between-farm scale.

While providing a proof of concept, this previous study was targeted towards a subsample of VNTR-10 infected cattle within a small (approx. 5 km) spatial radius. This spatially restricted sampling precluded a more systematic investigation of processes occurring on a wider, population level scale within the bacterial strain, including the potential identification of under-sampled reservoirs, the rate and mode of spatial spread, and transmission links between bTB breakdowns. A herd breakdown is defined as the period during which movements of cattle out of a herd are restricted due to the detection of bTB in the herd, starting at the detection of one or more infected animals (either through the tuberculin skin test or through abattoir surveillance for bTB lesions), and ending when the herd has undergone two consecutive negative whole-herd tests at least 60 days apart, or a single negative test where the breakdown was not laboratory confirmed.

Here, we extend the analysis of the Biek et al. (2012) study by examining WGS data from all 145 available VNTR-10 isolates in NI since 2003. In NI all cattle herds are tested for bTB on an annual basis, and for over a decade *M. bovis* isolates cultured from test-positive cattle have been extensively typed and stored. These archived samples therefore gave us the opportunity to target a genetically defined sub-population of *M. bovis* (VNTR-10) for high-density sampling with respect to cattle infections, although VNTR-10 infections in any other population would not be accessible through this sampling strategy. To gain insights into the mode and rates of transmission, we used intensive sampling and WGS of *M. bovis* isolates from cattle to address the following questions:

(1) Does WGS of VNTR-10 isolates from cattle indicate contributions from another host population which is under-sampled under the above sequencing strategy?
(2) Does the genetic relatedness between sequenced isolates correlate with recorded movements and/or with spatial distance between premises?
(3) What are the rate and mode of *M. bovis* dispersal across the landscape at the between-breakdown scale as determined by WGS?

2. Materials and methods

2.1. Molecular-typing of M. bovis in NI

In NI, *M. bovis* isolates have been stored and typed since the early 2000s using spoligotyping, more recently combined with Variable Number Tandem Repeat (VNTR) typing, to differentiate molecular types (Skuce et al., 2005). Spoligotyping gives a relatively coarse-grained discrimination of the *M. bovis* population, and is based on the presence or absence of multiple spacer oligonucleotides within the direct repeat region of the genome (Kamerbeek et al., 1997). VNTR-typing indexes the number of short nucleotide repeats present at several VNTR loci identified within the mycobacterial genome (Mazars et al., 2001), and provides greater discriminatory power than spoligotyping alone, although with a relatively higher chance of homoplasy (i.e., separate lineages converging on the same molecular type). In NI between 2003 and 2008 one *M. bovis* isolate was VNTR-typed and stored from each herd breakdown for which *M. bovis* was isolated using a panel of 7 VNTR loci optimised for this population of *M. bovis* (Skuce et al., 2010), while from 2009 onwards bacteria were VNTR-typed and stored from all culturable cattle cases, therefore resulting in more intensive sampling in recent years. In addition to these cattle isolates, *M. bovis* has also been typed by spoligotyping and VNTR-typing and archived when isolated from a survey of badgers killed on the roads in NI (Abernethy et al., 2011).

2.2. Bacterial samples and sequencing

Cultures of *M. bovis* were isolated from bovine granulomatous tissue using conventional methods (Skuce et al., 2010). Confirmed isolates were grown on LJ slopes to single colonies, following which single colonies were grown up and DNA was extracted using standard CTAB and solvent extraction protocols (Van Soolingen et al., 2001). A total of 144 VNTR-10 *M. bovis* isolates were included in this study, from 66 herd breakdowns (see earlier definition) occurring in 51 herds between 1996 and 2011. The WGS dataset consisted of the raw reads from 31 VNTR-10 samples originally sequenced in the preceding study (26 cattle and 5 badger isolates; Biek et al., 2012), in addition to 114 VNTR-10 samples (113 cattle and 1 badger isolate) sequenced for the first time in this study. VNTR-10 is located predominantly within the Newtownards area of NI (178/195 VNTR-10 infections recorded between 1996 and 2011 were from the Newtownards district veterinary office), and statistical comparison to other VNTR types confirms that VNTR-10 is generally representative of strains circulating in this area (see supplemental information). Details of the accession numbers for the raw sequencing reads are given in Table S1.

To provide broader evolutionary context we additionally sequenced five samples from VNTR-types thought to be ancestral to VNTR-10, namely four isolates of VNTR-1 (1 cattle isolate (sample A), and 3 badger isolates (samples B–D, Figs. 1 and 2)), and one cattle isolate of VNTR-4. VNTR-1 is thought to be the direct ancestor of VNTR-10 based on the following observations: VNTR-1 and VNTR-10 are separated by a single tandem repeat difference; VNTR-1 has been recorded at a high and approximately stable prevalence in NI since routine VNTR-typing commenced (Skuce et al., 2010, 2005), whereas VNTR-10 has been found in low but increasing numbers suggestive of a newly emerged strain; VNTR-1 is found across a wider spatial range than VNTR-10 (Fig. 2); and a minimum spanning tree of all NI VNTR-types within spoligotype SB0140 shows