Multi locus sequence typing of pathogenic treponemes isolated from cloven hoofed animals and their comparison to human isolates

Simon R. Clegg[1], Stuart D. Carter[1], Richard J. Birtles[2], Jennifer M. Brown[1], C. Anthony Hart[3*], Nicholas J. Evans[1]

[1]Department of Infection Biology, Institute of Infection and Global Health, School of Veterinary Science, University of Liverpool, Liverpool Science Park IC2, 146 Brownlow Hill, Liverpool. L3 5RF.

[2]School of Environment and Life Sciences, University of Salford, M5 4WT

[3]Department of Medical Microbiology, University of Liverpool, Liverpool L69 3GA, UK

* Deceased

Running title: MLST of digital dermatitis treponemes

Corresponding author email. s.r.clegg@liverpool.ac.uk. [1]Department of Infection Biology, Institute of Infection and Global Health, School of Veterinary Science, University of Liverpool, Liverpool Science Park IC2, 146 Brownlow Hill, Liverpool. L3 5RF.

Abstract word count- 247

Document word count- 4175

Abstract

Treponema species are implicated in many diseases of man and animals. Digital dermatitis (DD) treponemes are reported to cause severe lesions in cattle, sheep, pigs, goats and wild elk, causing substantial global animal welfare issues and economic losses.

The fastidiousness of these spirochetes has previously precluded studies investigating within phylogroup genetic diversity. An archive of treponemes isolated by the authors enabled multi locus sequence typing to quantify diversity and population structure of DD treponemes. Isolates (n=121) were obtained from different animal hosts in nine countries on three continents.

The analyses herein of currently isolated DD treponemes at seven housekeeping gene loci confirms classification into the three previously designated phylogroups, *Treponema medium*, *Treponema phagedenis and Treponema pedis*.

Sequence analysis of seven DD treponemes housekeeping genes revealed a generally low level of diversity among the strains within each phylogroup removing the need for the previously used ‘-like’ suffix. Surprisingly, all isolates within each phylogroup clustered together, regardless of host or geographical origin, suggesting the same sequence types (STs) can infect different animals. Some STs were derived from multiple animals from the same farm, highlighting probable within farm transmission. Several STs infected multiple hosts from similar geographic regions identifying probable frequent between host transmissions.

Interestingly, *T. pedis* appears to be evolving more quickly than the *T. medium* or *T. phagedenis* DD treponeme phylogroups by forming two unique

ST complexes. Lack of phylogenetic discrimination between treponemes isolated from different hosts or geographical regions substantially contrasts with other clinically relevant spirochetes.

Importance

The recent expansion of the host range of digital dermatitis (DD) treponemes, from cattle, to now include sheep, goats, pigs and wild elk, coupled with the high level of 16S rRNA gene sequence similarity across hosts and with human treponemes suggests the same bacterial species can cause disease in multiple different hosts. This multi locus sequence typing (MLST), study further demonstrates that these bacteria isolated from different hosts are indeed very similar, raising potential for cross species transmission between animal species. The study also shows infection spread, both locally and globally, occurs frequently, suggesting transmission by routes other than animal-animal alone. These results indicate that on farm biosecurity is important in controlling disease spread in domesticated species. Continued surveillance and vigilance is important to ascertain the evolution and track any further host range expansion of these important pathogens.

Introduction

Only rarely do we encounter infectious agents rapidly spreading through different animal populations and causing substantial and varied disease manifestations in a wide variety of hosts.

Classically, digital dermatitis (DD) is a disease of dairy cattle, first seen in 1974, and known to cause severe lameness (1). DD is now considered

endemic in dairy cattle in many countries worldwide, and is a serious animal welfare issue on farms. Economic impacts of the disease, due to reduction in milk yields and reproductive performance have been estimated at $190 million per annum in the USA alone (2).

A considerable body of evidence identifies specific *Treponema* species as the aetiological agents of DD. More recently, since it was first reported in 1997, DD has spread through UK sheep farms (3) and in very recent times, has been reported in UK goats (4). In these two host species, the same treponeme phylotypes associated with cattle DD are consistently identified in the foot lesions resulting in severe clinical outcomes which are very difficult to treat. The DD treponemes have been recently associated with foot lesions causing lameness in wild American elk (*Cervus elaphus*) (5). Reports show DD treponemes can be isolated from, and associated with, porcine ear and shoulder skin lesions (6, 7, 8). In man, *Treponema* spp. are considered responsible for periodontal disease and syphilis. Interestingly, whilst oral treponemes are reported as closely related to DD treponemes, the agent of syphilis is substantially different (9, 10).

To date, five major phylotypes of treponeme have been highly associated with DD (9, 10, 11, 12). Three of these DD associated phylogroups have repeatedly been isolated from animals symptomatic for DD and subsequently designated as coherent groups on the basis of genotypic and phenotypic characterisations (10, 13, 14, 15). Previous studies have identified the culturable DD treponemes as highly similar to human periodontal and genito-urinary treponemes, based on their 16S rRNA gene, and due to lack of additional data this led to assignment of a ‘-like’ suffix (10, 13, 14, 15).

Contrastingly, a recent study has suggested the removal of the '-like' suffix for the bovine *T. phagedenis* isolates (16). The three cultivable treponemes have been grouped as *Treponema medium/Treponema vincentii*, *Treponema phagedenis* and *Treponema putidum/Treponema denticola* DD spirochaete phylogroups (10, 13, 14, 15). Subsequently, the latter phylogroup was designated as a novel species, *Treponema pedis* (13).

The fastidious nature of these micro-organisms and the difficulty of obtaining pure treponemes have previously led to a dearth of isolates. However, bacterial culture developments by the authors have enabled accumulation of an archive of treponeme isolates which may now allow comparative analyses to investigate their genetic relationships.

Sequencing 16S rRNA genes has demonstrated clear differences between the three commonly isolated phylogroups of DD treponemes with them sharing only 90.1 % - 92.3% 16S rRNA gene sequence identity and they are therefore considered separate phylogroups/species (10). However, little 16S rRNA gene sequence variation within phylogroups has been identified, with no notable variation between different treponemes within a phylogroup isolated from different hosts (4, 5, 17). Other studies have analysed a number of genetic loci, including intragenic spacer regions (ISR1 and ISR2) and *flab2*, but this did not allow for isolate discrimination, beyond that observed using 16S rRNA gene sequence comparisons (10, 14, 18). To further investigate the DD treponemes, additional genotyping studies are required to allow intra-phylogroup discrimination.

Multi locus sequence typing (MLST) or multi locus sequence analysis (MLSA; which differs in analyses used) schemes, for a range of spirochaetes have now been described (19-23), and serve as key frameworks for the phylogeny and taxonomy of these taxa. Given rapid recent expansions in host range of the DD treponemes, it is particularly timely to determine if the same bacteria are infecting multiple animal species, or whether several different host-specific genotypes exist. The aim of this study was to design an MLST protocol for the three important phylogroups of DD treponemes which can be grown in culture and to investigate cross host species disease transmission. Such comprehensive molecular genetic analyses of the different treponemes isolated from a variety of hosts and geographic regions should ascertain their similarities and identify relevant relationships between human and animal pathogenic treponemes.

This report describes the MLST classification of 121 fastidious treponeme isolates, the vast majority of which have been obtained from animal tissues during the past 10 years. The data reveals interesting insights into the transmission of disease between host species, on varying spatial scales (including within a farm, country, and more globally) and the role of treponeme evolution in such processes.

Methods

Bacterial taxonomy.

In the majority of previous studies, a '-like' suffix has been used for the bovine DD *T. medium* and *T. phagedenis* spirochetes based on their close similarity to human treponeme relatives (using 16S rRNA gene analysis). Towards clarity, this study proposes removal of the '-like' suffix with subsequent reference as the '*T. medium* phylogroup' and the '*T. phagedenis* phylogroup', as previously suggested for bovine DD T. phagedenis (16). Each phylogroup includes isolates which share more than 97%, 16S rRNA gene sequence identity to what are considered the representative strains of the three different phylogroups, namely T19, T320A and T3552B^{T} respectively. The 97% 16S rRNA gene sequence identity criterion has been frequently used in previous taxonomic assignments of bacteria and specifically treponemal species (10, 24, 25, 26).

Treponeme isolates

In this study, 121 isolates were investigated, 116 isolates were regrown, passaged and purified to ensure a pure isolate of each was used for genotyping. Forty eight of these are previously undescribed isolates. All 116 isolates were grown on fastidious anaerobe agar (FAA) plates, supplemented with 5% defibrinated sheep blood and antibiotics (10) from which single colonies were picked into OTEB tubes as previously described (10). The provenances of isolates are presented in Table 1-3. The three USA cattle isolates were a kind gift from Richard Walker (USA) and human isolates were obtained from the American Tissue Culture Collection (ATCC) and the Collection of Institut Pasteur (CIP).

In addition, the DNA sequences of five samples available on Genbank were used in the study; *T. vincentii* OMZ 838 (CP009227; 27) and *T. pedis* (CP004120) isolated from a pig (7), two shotgun sequenced *T. phagedenis* isolates from cattle, one from Iowa, USA (16) and one from Sweden (AQCF00000000 and CDNC01000001-CDNC01000051 respectively), and a human genitourinary *T. phagedenis* isolate were also used (NZ_AEFH00000000).

DNA isolations

For collection of bacterial genomic DNA from OTEB cultures, 2 ml of the culture was centrifuged (5000 X *g*, 10 min, 4°C) in a bench-top centrifuge. DNA was then extracted from the cell pellet using Chelex-100, as previously described (28) and stored at -20°C.

16S rRNA gene PCR.

The 16S rRNA gene was amplified as described previously (10) from the 48 new isolates included in this study (Table 1-3). Isolates were confirmed to contain only a single phylogroup using a nested PCR, specific for the three unique treponeme phylogroups (10).

Multi locus sequence analysis

The genetic loci used within this study were akin to those used in the MLST of another pathogenic spirochete genus, *Brachyspira* (23).

The presence of a single copy of each locus within the genomes of representatives of each of the three DD treponeme phylogroups was confirmed by analysis of almost complete (>93%) genomes available online (*T. medium*, (KE332517.1) *T. phagedenis* F0421 (AEFH01000000), *T. pedis* (CP004120)). Furthermore, the loci were identified as well dispersed around these genomes (>100kb between each loci). Primers were designed to amplify fragments of genes encoding for a heat shock protein (GroEL), recombination protein A (RecA), glycerol kinase (GlpK), adenosine kinase (AdK), glutamate dehydrogenase (GDH), orotidine 5′-phosphate decarboxylase (PyrG) and the large RNA polymerase sub unit (RplB) by reference to the genome sequences described above using Primer3 (29) such that all amplicons were 500-600bp in length (Table 4).

PCR master mixes for each locus were as previously reported (10, 13) but incorporating the new MLST primers (Table 4). All PCRs were carried out using the cycling conditions of: 95°C for 1 minute, followed by 40 cycles of 95°C for 1 minute, 55°C for 1 minute and 72°C for 2 minutes, with a final extension of 72°C for 10 minutes.

Sequencing and sequence analysis

Amplified PCR products were sequenced commercially (Macrogen. Amsterdam. The Netherlands) and the data for each locus were verified and assembled using Chromas Pro 1.7.5 sequence analysis package (Technelysium Pty Ltd). Gene sequences were aligned using CLUSTALW as implemented in MEGA 5.0 (30). Alleles and sequence types were assigned manually and analysed using eBURST (data not shown) (31).

In order to infer a phylogeny from 16S rRNA gene data, an alignment of sequences was subjected to ModelTest, as implemented in Topali (32), which revealed that the best fit model was General Time Reversible (GTR). This was used to produce nucleotide maximum likelihood phylogenetic trees (bootstrap values based on 10,000 iterations). For each isolate, sequence data for the seven MLST loci were concatenated, then concatenated data from different isolates aligned with one another. Phylogenetic inferences from this alignment were made as described above. Concatenated gene trees were drawn using TN93 models (33), and all maximum likelihood trees were produced using 10,000 bootstrap values. Minimum spanning distance trees were drawn using Prim's algorithm (http://pubmlst.org/analysis/). Alignments were screened for evidence of recombination using SplitsTree4 (34) and positive and negative selection using GARD and SLAC available through the Datamonkey web server (35).

Results

16S rRNA gene analysis

Almost complete 16S rRNA gene sequences were obtained for the 48 new DD treponeme isolates obtained in the study. Phylogenetic inference derived from these data and those for the other 73 isolates included in the study, indicated all could be accommodated within one of the three previously described DD treponeme phylogroups (Figure S1).

Thus, the study included 34 isolates belonging to the *T. medium* phylogroup (DD1), 70 isolates belonging to the *T. phagedenis* phylogroup (DD2), and 17 isolates belonging to the *T. pedis* phylogroup (DD3).

MLST data

For all 121 isolates, sequences were obtained for all seven MLST loci. Comparison of the sequence data revealed variation at all loci, with no cases of full gene recombination seen, between each of the three phylogroups. The average dissimilarity between the loci from the three different phylogroups was 28.46%, (range 17.9 (*groEL*) to 39.26 (*adK*)). Furthermore, all loci varied within phylogroups, with dissimilarity ranging from 0.5% (*adK* in the *T. phagedenis* phylogroup) to 17% (*adK* in the *T. medium* phylogroup) (Table 5). Sequence variation at loci was far more pronounced in the *T. medium* phylogroup (mean = 10.9%) than in the *T. phagedenis* (mean = 1.2%) or the *T. pedis* (mean = 2.5%) phylogroups. Even when the outlying *T. vincentii* was excluded from the *T. medium* phylogroup, the mean sequence variation among its remaining members was 4.9% (Table 5).

The number of alleles for each locus ranged from 10 to 18, with between two and nine in each phylogroup (Table 5). Sequence types were assigned based on the MLST allelic profiles. Comparison of allelic profiles revealed a total of 53 STs, 11 within *T. medium* phylogroup (Table 1), 35 within *T. phagedenis* phylogroup (Table 2) and seven within *T. pedis* (Table 3). Unique allelic sequences were obtained from each of the different phylogroups. Of 11 STs within *T. medium* phylogroup, ST1 was encountered most frequently (14/34 (41%) isolates) (Table 1). However, in the other two phylogroups, no ST was

so dominant, suggesting they contain greater sequence variation within each phylogroup. In the *T. phagedenis* phylogroup (Table 2), ST16 was the most common, but only 6 of 67 (8%) isolates possessed this ST. In the *T. pedis* phylogroup, ST5 was the most common, but only 6 of 17 (35%) isolates possessed this ST (Table 3).Of the 53 STs encountered, 29 were represented by only one isolate each.

Minimum spanning trees compare similarities amongst the different ST's isolated, and how closely related they are. Therefore isolates located close on a tree are generally different at one of the MLST loci, whereas the more distant ones have fewer loci in common. The *T. medium* phylogroup minimum spanning tree showed relationships centred around the founder ST, ST1 (Figure 1), which contains both cattle and sheep isolates (Table 1). The *T. vincentii* OMZ 38 sequence type (ST11) and the human *T. medium* ATCC 700293 sequence type (ST2) were outliers in the data, further suggesting that they are profoundly divergent from DD-associated strains. This data also further corroborates that *T. vincentii* is not a member of *T. medium* phylogroup, but a separate species (Figure 1 and 4 and Table 1).

Data for the *T. phagedenis* phylogroup minimum spanning distance tree (Figure 2) suggested that ST2 was the founder ST, with nine other STs as single locus variants (SLV). However, the neighbouring ST, ST9, possessed eight SLVs (Table 2, Figure 2). For both the *T. medium* and *T. phagedenis* phylogroups, human isolates were distant from the animal isolates (Figure 2 and 5 and Table 2).

The *T. pedis* phylogroup minimum spanning tree shows isolates radiating out from ST1 (which contains T3552B^T, the type strain). A larger amount of variation is seen within the *T. pedis* tree, compared to the other two phylogroups (Figures. 3 and 6 and Table 3).

The newer sequences (ST2 and 3), isolated from sheep and goats, form a distinct cluster to the older isolates, which were largely isolated from cattle.

All allelic data was uploaded into pubMLST (36)

Molecular epidemiology

Many STs, in all three phylogroups, were encountered in more than one host species and in multiple geographical locations. Within the *T. medium* phylogroup, four of the five STs (ST 1, 6, 7 and 9) that contained more than one isolate were recovered from different host species In the case of ST1 of the *T. medium* phylogroup, these isolates were from both cattle and sheep. Additionally, three of these *T. medium* phylogroup STs contained isolates which were recovered from animals inhabiting geographically distant countries including ST1 being present in England, Wales and USA. Conversely, we also obtained isolates belonging to different STs of the same phylogroup from the same host species on the same farm (Table 1).

By contrast, the human *T. medium* ATCC 700293*,* and *T. vincentti* OMZ 838 isolates were unique allelic arrangements.

Within the *T. phagedenis* phylogroup, similar patterns were seen, with six of the 15 ST's (ST 1, 2, 3, 9, 17, 27) which contained more than a single isolate being recovered from different host species (Table 2 and figure 5). Twenty of the 35 *T. phagedenis* phylogroup STs were singletons, containing only one isolate. As with the *T. medium* phylogroup, all four human isolates of *T. phagedenis* had unique allelic arrangements (ST7, 18, 19 and 34; Figure 5).

Although *T. pedis* isolate numbers were smaller, two ST's (ST2 and ST5) of three ST's that contained more than a single isolate were recovered from different host species (Table 3).

Of the 19 farms used in this study, 13 of them had isolates circulating on them belonging to more than one ST. (Tables 1, 2, and 3).

Evolutionary features within loci

Nucleotide polymorphisms were seen in all loci tested, from all three of the DD treponeme phylogroups. Within some loci there appeared to be regions of sequence in which single nucleotide polymorphisms (SNPs) were concentrated. For example, among *T. pedis* phylogroup members, 12 of the 13 SNPs in a 421 base pair *adK* PCR product occurred in the final 150 base pairs of the locus. Similarly, among *T. phagedenis* phylogroup members, seven of 10 SNPs in a 560 bp *gdh* PCR product occurred in a 30 bp section between nucleotides 464 - 494.

Analysis of data for each loci did not reveal any evidence of positive selection pressures, although among the *T. medium* phylogroup members, sites within the *adK, pyrG* and *rplB* loci appeared to be under negative or purifying pressure (Table S1).

Splits decomposition analysis suggested that, in general, recombination has had a marked influence on the divergence of STs within all three phylogroups (Figure S2). However, we were unable to find evidence of recombination between different phylogroups (data not shown).

Phylogeny

In concurrence with the phylogeny inferred from alignment of 16S rRNA gene sequence data, the phylogeny inferred from alignment of the concatenated MLST loci sequence data delineated the DD treponemes investigated in this study into three deeply diverging phylogroups. Both the *T. medium* (Figure 5) and *T. phagedenis* (Figure 6) phylogroups formed a single sequence complex. However, the *T. pedis* phylogroup has diverged into two different sequence complexes which are defined as being their similarity to the central allelic profile (Figure 6).

Discussion

The recent expansion in the host range of DD *Treponema* spp., to include a variety of additional food chain animals, has led to a greater number of animal welfare issues and greater substantial economic losses to agricultural industries (4, 5, 7-13, 17, 37-39). Furthermore, the inter- and intra host species spread of these bacteria needs to be given special consideration as

366 isolates from humans and all animal species are considered very similar or
367 identical (4, 5, 7).
368 Therefore, the use of a treponeme isolate archive in this study created a
369 relatively unique opportunity to study bacterial species that can infect and
370 cause disease in multiple animal species. As MLST analyses have previously
371 been used to clarify relationships within a bacterial species, and to
372 differentiate bacteria by host species (23, 40), MLST was used in an attempt
373 to differentiate DD treponemes isolated from different host species.
374 In this study, a collection of 121 DD *Treponema* isolates from nine different
375 countries and three different continents were analysed by MLST to elucidate
376 the relationships between isolates from different host species, but was limited
377 by geographic ranges of species (e.g. elk) and diseases (e.g. CODD). That
378 said, this is the largest and most rigorous molecular genetic analysis of DD
379 treponemes isolated from humans and animals.
380
381 **Cultivable DD treponemes can be classified into three distinct**
382 **phylogroups**
383 All cultivable DD treponeme isolates included within this study fit into the three
384 previously reported phylogroups (10, 14), except for the human periodontal
385 disease associated *T. vincentii* which was unique at each loci tested,
386 suggesting that it is a different phylogroup and unrelated to any farm animal
387 disease associated isolates despite high 16S rRNA gene similarities.
388 The analyses of 16S rRNA and housekeeping gene loci of currently isolated
389 DD treponemes confirms classification into the three previously designated

390 phylogroups, *Treponema medium, Treponema phagedenis* and *Treponema*
391 *pedis* (4, 5, 10, 11, 13, 14, 17, 18).
392 Sequence analysis of seven DD treponemes housekeeping genes revealed a
393 generally low level of diversity among the strains within each phylogroup,
394 removing the need for the previously used '-like' suffix. Taken together, the
395 authors recommend removal of the '-like' suffix, and instead refer to the
396 bacteria as belonging to a phylogroup, such as the *T. medium* phylogroup, in
397 line with similar studies of pathogenic mycobacteria (41). This has also been
398 suggested for *T. phagedenis* isolates recently (16).
399 Although phylogeny and eBURST revealed limited data regarding evolutionary
400 relationships in clonal complexes (32, 42), taken together these approaches
401 show that all isolated treponemes in this study group into three unique
402 phylogroups suggesting that they are of different evolutionary lineages, but
403 from a common ancestor. They also show that the *T. pedis* phylogroup is
404 beginning to form two distinct ST complexes, based on related MLST allelic
405 arrangements, with the newer isolates separating from the older isolates. This
406 raises the importance of continued surveillance and vigilance of DD
407 treponeme infections as emergence of a new species may lead to increased
408 pathogenicity and potentially, host range. Isolation of members of the *T. pedis*
409 phylogroup appears to be less common (or successful) than the other two
410 phylogroups, as only 17 were isolated and analysed in this study, compared to
411 34 *T. medium* and 70 *T. phagedenis* phylogroup treponemes. Isolation of
412 more *T. pedis* phylogroup treponemes in the future will further help to
413 delineate the two ST complexes which this phylogroup appears to be forming.

414 However, the overall variation within the phylogroup is limited, with isolates
415 from pigs, cattle, sheep and goats all being relatively similar.
416 Although variation is seen within each loci of the *T. medium* and *T.*
417 *phagedenis* phylogroups, including the 16S rRNA gene, they all group
418 phylogenetically and both form a single clonal complex.
419
420 **The cultivable DD treponemes show limited genetic variability within**
421 **phylogroups**
422 Identical bacteria were isolated from different host species, and 12 of the 23
423 sequence types with more than a single isolate in them were from different
424 host species, such as with *T. phagedenis* ST1. This contrasts to other
425 clinically significant spirochaetes such as *Brachyspira* spp. where isolates
426 from different hosts generally belong to different bacterial species (23).
427 Furthermore, ST's within several different species of *Leptospira* are generally
428 separated by host and geography (19), whilst there can clearly be identified
429 geographical separation of *Borrelia burgdorferi* between two locations in the
430 USA (43). Therefore, this study demonstrates that MLST may not appear
431 suitable for differentiation of cultivable DD treponemes isolated from different
432 host species or it might be considered that the inability to discriminate
433 identifies frequent transmission events between host species. Alternatively, it
434 may be that the limited geographical sampling and the relatively small isolate
435 numbers included in this study mean that differentiation by MLST is difficult.
436 Indeed, all genes sequenced here, from all three phylogroups, showed
437 relatively little diversity, suggesting that the bacteria potentially have evolved
438 genes which are highly functionally fit, and are under little selection pressure

to evolve further. However, within the three phylogroups, *T. pedis* was the most diverse. Previous studies have suggested that some sections of the *T. pedis* genome has been lost compared to its closest relative, *T. denticola*, which further suggests it is evolving rapidly (7). This increased evolution rate may agree with reports that *T. pedis* is more surface dwelling (44, 45) compared with the other phylogroups, therefore is likely to have to adapt to more rapidly alternating conditions, increasing genetic diversity compared to deeper tissue dwellers (44, 45).

Treponema pallidum, the causative agent of syphilis and yaws, shows a low level of diversity despite multiple isolations over many years, and is highly similar to the related bacterium, *Treponema paraluiscuniculi,* the causative agent of rabbit venereal spirochaetosis (46). These bacteria are similar or identical at the 16S rRNA gene level, but infect two very different hosts. The data presented here shows that the animal and human cultivable DD treponeme phylogroups have an even greater capacity to infect numerous hosts, while undergoing little genetic alteration and evolution.

Treponemes evolve by within phylogroup recombination

This study showed both recombination, and some positive selection, within DD treponeme phylogroups, unlike that seen in *T. denticola* (47). *Treponema denticola* was monophyletic, as were the three DD phylogroups in this study. However, in this study, the *T. pedis* phylogroup was more variable, diverging into two separate ST complexes, suggesting a more rapid evolution than the other phylogroups. Recombination was seen within the DD treponeme phylogroups, but not between phylogroups, as evidenced by the lack of cross

464 reactivity between primers, and this was further confirmed by the splits graphs
465 analysis. The use of different oligonucleotides for amplification of the different
466 phylogroups further supports the continued usage of the three unique
467 groupings of culturable treponemes suggested previously (10). Similar issues
468 were identified in previous studies using *Brachyspira* spp. and *Campylobacter*
469 spp., where it was reported that it is difficult to develop MLST oligonucleotides
470 to amplify genes from an entire genus (23, 40).
471
472 **Bacterial spatial dynamics reveal multiple transmission events, locally**
473 **and globally.**
474 Within each DD treponeme phylogroup, limited evidence of a correlation
475 between genotype and geographical provenance was seen. In some cases,
476 ST's were concentrated on a single farm, or a few localised farms (e.g. ST16
477 in the *T. phagedenis* phylogroup), whereas others were found in different
478 areas of the country (e.g. ST5 of the *T. pedis* phylogroup), and some were
479 more global (e.g. ST7 from the *T. medium* phylogroup). This suggests that
480 STs can spread and circulate worldwide amongst different animals. This is in
481 contrast to *Borrelia* species, which show a geographical delineation with
482 European and American isolates phylogenetically separate (21, 48).
483 Spatial dynamics of the bacterial STs suggest that identical bacteria can
484 circulate on a farm, spreading around a flock or herd, such as was seen in *T.*
485 *medium* phylogroup ST1 and ST7, *T. phagedenis* phylogroup ST16, and *T.*
486 *pedis* phylogroup ST2. On farm spread was more apparent among sheep
487 flocks than cattle herds, possibly due to closer confines and higher stocking
488 density of sheep. Additionally, the clinical manifestation of the disease causes

much greater morbidity in sheep than in cattle, so it may be that cattle appear asymptomatic, whereas sheep show clinical signs quicker, and more noticeably. Many cattle farms also appear to be endemically infected, whereas sheep farms tend to present with episodic epidemics.

Other bacterial species such as ST1 in the *T. medium* phylogroup, ST1 in the *T. phagedenis* phylogroup and ST5 in the *T. pedis* phylogroup can infect multiple species, increasing their transmission. Furthermore the similarities between the isolates from different animal hosts raises the possibility of both inter- and intra host species transmission, but the mechanism for this spread remains unclear.

In future, comparative analyses of full DD treponeme genomes isolated from a range of hosts will further delineate whether the same treponemal strains are indeed responsible for the recent expansion in host range and pathology in line with the results from this current study. Such studies will also increase our knowledge of pathogen evolution and disease transmission, to better inform farm practice, to prevent these severe diseases and enhance global food security.

Funding information

This work was funded by a Biotechnology and Biological Sciences Research Council (BBSRC) CEDFAS Research Grant (BBE0189201), awarded to Stuart D. Carter and Nicolas J. Evans and a BBSRC New Investigator Award

(BB/K009443/1) awarded to Nicholas J. Evans. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Acknowledgements

The authors would like to thank the various veterinary clinicians, foot trimmers and farmers who have been involved in sample collection.

References

1. **Cheli R., Mortellaro C**., 1974. Digital Dermatitis in cattle. In: Proceedings of the 8th 317 International Conference on Diseases of Cattle, Milan, p. 208–213

2. **Losinger WC**. 2006. Economic impacts of reduced milk production associated with 613 papillomatous digital dermatitis in dairy cows in the USA. J Dairy Res.**73**;244-614.

3. **Harwood DH., Cattell JH., Lewis CJ., Naylor RD**. 1997 Virulent footrot in sheep. Vet Rec.**140;**687

4. **Sullivan LE., Evans NJ., Clegg SR., Carter SD, Horsfield JE., Grove-White D., Duncan JS**. 2014. Digital dermatitis treponemes associated with a severe foot disease in dairy goats. Vet Rec.**176;**283

538

539 5. **Clegg SR., Mansfield KG., Newbrook K., Sullivan LE., Blowey R. W.,**

540 **Carter SD., Evans NJ**. 2014. Isolation of digital dermatitis treponemes from

541 hoof lesions in wild North American elk (*Cervus elaphus*) in Washington State,

542 USA. J Clin Micro.**53;**88-94.

543

544 6. **Karlsson F., Svartström O., Belák K., Fellström C., Pringle M**. 2013.

545 Occurrence of *Treponema* spp. in porcine skin ulcers and gingiva. Vet

546 Microbiol.**30**.402- 409.

547

548 7. **Svartström O., Karlsson F., Fellström C., Pringle, M**. 2013.

549 Characterization of *Treponema* spp. isolates from pigs with ear necrosis and

550 shoulder ulcers. Vet Microbiol.**25**;617-23

551

552 **8. Clegg SR., Sullivan LE., Bell J., Blowey RW., Carter SD., Evans NJ.**

553 2016. Detection and isolation of digital dermatitis treponemes from skin and

554 tail lesions in pigs. Res Vet Sci. **104**:64-70.

555

556 9. **Choi BK., Nattermann H., Grund S., Haider W., Gobel, UB**. 1997.

557 Spirochetes from digital dermatitis lesions in cattle are closely related to

558 treponemes associated with human periodontitis. Int. J. Syst.

559 Bacteriol.**47**;175-181.

560 10. **Evans NJ., Brown J.M, Demirkan I, Murray R.D, Vink WD, Blowey RW,**

561 **Hart CA, Carter SD,.** 2008. Three unique groups of spirochetes isolated from

562 digital dermatitis lesions in UK cattle. Vet. Microbiol.**130;**141-150.

11. **Klitgaard K., Boye M., Capion N., Jensen TK**., 2008. Evidence of Multiple *Treponema* Phylotypes Involved in Bovine Digital Dermatitis as Shown by 16S rRNA Gene Analysis and Fluorescence In Situ Hybridization. J Clin Microbiol.**46**;3012-3020.

12. **Nordhoff M., Moter A., Schrank K., Wieler LH**., 2008. High prevalence of treponemes in bovine digital dermatitis-A molecular epidemiology. Vet Microbiol.**131**;293-300.

13. **Evans NJ, Brown JM, Demirkan I., Murray RD, Birtles RJ., Hart CA.,Carter SD**. 2009. *Treponema pedis* sp. nov., a spirochaete isolated from bovine digital dermatitis lesions. Int. J. Syst. Evol. Microbiol.**59;**987–991

14. **Stamm LV., Bergen HL., Walker RL**., 2002. Molecular typing of papillomatous digital dermatitis-associated *Treponema* isolates based on analysis of 16S-23S ribosomal DNA intergenic spacer regions. J Clin Microbiol.**40;**3463-3469

15. **Read DH., Walker, RL**., 1998. Papillomatous digital dermatitis (footwarts) in California dairy cattle: clinical and gross pathologic findings. J. Vet. Diag. Invest.**10;**660 67-76.

16. **Wilson-Welder JH., Elliott MK., Zuerner RL., Bayles DO., Alt DP., Stanton TB**. 2013. Biochemical and molecular characterization of *Treponema phagedenis*-like spirochetes isolated from a bovine digital dermatitis lesion. BMC Microbiol. **13**: 280-289.

17. **Sullivan LE., Clegg SR., Angell JW., Newbrook K., Blowey RW., Carter SD., Bell J., Duncan JS., Grove-White DH., Murray RD., Evans NJ**. 2015. High-level association of bovine digital dermatitis *Treponema* spp. with contagious ovine digital dermatitis lesions and presence of *Fusobacterium necrophorum* and *Dichelobacter nodosus*. J. Clin. Microbiol.**53**:1628-38

18. **Pringle M., Bergsten C., Fernstrom LL., Hook H., Johansson KE**., 2008. Isolation and characterization of *Treponema* phagedenis-like spirochetes from digital dermatitis lesions in Swedish dairy cattle. Acta Vet Scand.**50**; 40-48.

19. **Boonsilp S., Thaipadungpanit J., Amornchai P., Wuthiekanun V., Bailey MS., Holden MT., Zhang C., Jiang X., Koizumi N., Taylor K., Galloway R., Hoffmaster AR., Craig S., Smythe LD., Hartskeerl RA., Day NP., Chantratita N., Feil EJ., Aanensen DM., Spratt BG., Peacock SJ.** 2013. A single multilocus sequence typing (MLST) scheme for seven pathogenic *Leptospira* species. PLoS Negl. Trop. Dis. **7;**e1954.

20. **La T., Phillips ND., Harland BL., Wanchanthuek P., Bellgard MI., Hampson DJ**., 2009. Multilocus sequence typing as a tool for studying the molecular epidemiology and population structure of *Brachyspira hyodysenteriae*. Vet Microbiol.**138**;330-338.

21**. Margos G., Gatewood AG., Aanensen D.M., Hanincova K., Terekhova D., Vollmer SA., Cornet M., Piesman J., Donaghy M., Bormane A., Hurn MA., Feil EJ., Fish D., Casjens S., Wormser GP., Schwartz I., Kurtenbach K**. 2008. MLST of housekeeping genes captures geographic population structure and suggests a European origin of *Borrelia burgdorferi*. Proc Natl Acad Sci USA.**105;**8730-8735.

22. **Phillips ND., La T., Amin MM., Hampson DJ.** 2010. *Brachyspira intermedia* strain diversity and relationships to the other indole-positive *Brachyspira species*. Vet Microbiol.**143**;246-254.

23. **Rasback T., Johansson KE., Jansson DS., Fellstrom C., Alikhani MY., La T., Dunn DS., Hampson DJ**. 2007. Development of a multilocus sequence typing scheme for intestinal spirochaetes within the genus *Brachyspira*. Microbiol.**153;**4074-4087.

24. **Moter A, Leist G., Rudolph R., Schrank K., Choi BK., Wagner M., Göbel UB**. 1998. Fluorescence in situ hybridization shows spatial distribution of as yet uncultured treponemes in biopsies from digital dermatitis lesions. Microbiol.**144**:2459-67.

25. **Woese C R. Kandler O., Wheelis M.** 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci USA. **87**;4576–9.

636 26. **Stackebrandt E., Goebel BM**. 1994. Taxonomic Note: A Place for DNA-
637 DNA Reassociation and 16S rRNA Seqence Analysis in the Present Species
638 Definition in Bacteriology. Int J Syst Evol Microbiol.**44**:846-849
639
640 **27. Chan Y, Ma APY, Lacap-Bugler DC, Huo Y-B, Keung Leung W, Leung**
641 **FC, Watt RM**. 2014. Complete genome sequence for *Treponema* sp. OMZ
642 838 (ATCC 700772, DSM 16789), isolated from a necrotizing ulcerative
643 gingivitis lesion. Genome Announc.**2**:e01333-14.
644
645 28. **Chua PK., Corkill JE. Hooi PS. Cheng SC. Winstanley C., Hart CA**.
646 2005 Isolation of *Waddlia malaysiensis*, a novel intracellular bacterium, from
647 fruit bat (*Eonycteris spelaea*). Emerg. Infect. Dis.**11**;271–277
648
649 29. **Untergasser A., Nijveen H., Rao X., Bisseling T., Geurts R.,**
650 **Leunissen JAM**. 2007 Primer3Plus, an enhanced web interface to Primer3.
651 Nucleic Acids Res.**35**:W71-W74
652
653 30. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011.
654 MEGA5: Molecular Evolutionary Genetics Analysis using Maximum
655 Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol
656 Biol Evol.**28**:2731-2739.
657
658 31. **Feil EJ., Li BC., Aanensen DM., Hanage WP., Spratt BG**. 2004.
659 eBURST: Inferring Patterns of Evolutionary Descent among Clusters of

660 Related Bacterial Genotypes from Multilocus Sequence Typing Data. J
661 Bacteriol.**186**;1518–1530.
662
663 32. **Milne I., Lindner D., Bayer M., Husmeier D. McGuire G., Marshall DF.,**
664 **Wright F**. 2009. TOPALi v2: a rich graphical interface for evolutionary
665 analyses of Phylogenetics. Bioinformatics.**25**;126-127
666
667 33. **Tamura K, Nei M**. 1993 Estimation of the number of nucleotide
668 substitutions in the control region of mitochondrial DNA in humans and
669 chimpanzees. Mol Biol Evol.**10**:512-26.
670
671 34. **Huson DH., Bryant D**. 2006 Application of Phylogenetic Networks in
672 Evolutionary Studies. Mol Biol Evol.**23**:254-267.
673
674 35. **Pond SL., Frost SDW**. 2005. Datamonkey: rapid detection of selective
675 pressure on individual sites of codon alignments. Bioinformatics.**21**:2531–
676 2533
677
678 36. **Jolley KA. Maiden MCJ**. 2010. BIGSdb: Scalable analysis of bacterial
679 genome variation at the population level. BMC Bioinformatics.**11**:595-606
680
681 37. **Dhawi A., Hart CA., Demirkan I., Davies IH., Carter SD.** 2005. Bovine
682 digital dermatitis and severe virulent ovine foot rot: a common spirochaetal
683 pathogenesis.Vet J.**169**;232-41.
684

685 38. **Duncan JS., Angell JW., Carter SD., Evans NJ., Sullivan LE., Grove-**
686 **White DH**. 2014. Contagious ovine digital dermatitis: An emerging disease.
687 Vet J.**201**; 265-268
688
689 39. **Sayers G., Marques P., Evans NJ., O'Grady L., Doherty ML., Carter**
690 **SD., Nally JE**. 2009 Identification of spirochetes associated with contagious
691 ovine digital dermatitis. J Clin Microbiol.**47**;1199-201.
692
693 40. **Miller WG., On SLW., Wang G., Fontanoz S., Lastovica AJ., Mandrell**
694 **RE.** 2005. Extended Multilocus Sequence Typing System for *Campylobacter*
695 *coli*, *C. lari*, *C. upsaliensis*, and *C. helveticus*. J Clin Micro.***43***;*2315–2329.
696
697 41. **Huard RC, Fabre M, de Haas P, Lazzarini LC, van Soolingen D,**
698 **Cousins D, Ho JL**. 2006. Novel genetic polymorphisms that further delineate
699 the phylogeny of the *Mycobacterium tuberculosis* complex. J
700 Bacteriol.**188**:4271-87.
701
702 42. **Didelot X.,Falush D**. 2007. Inference of Bacterial Microevolution Using
703 Multilocus Sequence Data. Genetics.**175**.1251-1266
704
705 43**. Hanincova K., Mukherjee P., Ogden NH., Margos G., Wormser GP.,**
706 **Reed KD., Meece JK., Vandermause MF., Schwartz I.** 2013 Multilocus
707 sequence typing of *Borrelia burgdorferi* suggests existence of lineages with
708 differential pathogenic properties in humans. PLoS One.**17**; e73066.
709

710 44. **Ellen, RP., Dawson JR., Yang PF**. 1994. *Treponema denticola* as a
711 model for polar adhesion and cytopathogenicity of spirochetes. TIM.**2**.114-9.
712
713 45. **Briggs B., Colwell F**. 2014 Adapt or Die on the Highway To Hell:
714 Metagenomic Insights into Altered Genomes of Firmicutes from the Deep
715 Biosphere. American Geophysical Union. Fall meeting. San Fransisco. USA.
716
717 46. **Šmajs D., Zobaníková M., Strouhal M., Čejková D., Dugan-Rocha S.,**
718 **Pospíšilová P., Norris SJ., Albert T., Qin X., Hallsworth-Pepin K., Buhay**
719 **C., Muzny DM., Chen L., Gibbs RA., Weinstock GM**. 2011. Complete
720 genome sequence of *Treponema paraluiscuniculi*, strain Cuniculi A: the loss
721 of infectivity to humans is associated with genome decay. PLoS
722 One.**6**:e20415.
723
724 47. **Mo S., You M, Su YC., Lacap-Bugler DC., Huo YB., Smith GJ., Leung**
725 **WK., Watt RM**. 2013 Multilocus sequence analysis of *Treponema denticola*
726 strains of diverse origin. BMC Microbiol.**4**;13-24.
727
728 48. **Richter D., Postic D., Sertour N., Livey I., Matuschka FR., Baranton G**.
729 2006 Delineation of *Borrelia burgdorferi sensu lato* species by multilocus
730 sequence analysis and confirmation of the delineation of *Borrelia spielmanii*
731 *sp. nov*. Int J Syst Evol Microbiol.**56**:873-81.
732

49. **Walker RL., Read DH., Loretz KJ., Nordhausen RW**., 1995. Spirochetes isolated from dairy cattle with papillomatous digital dermatitis and interdigital dermatitis. Vet Microbiol.**47**;343-355.

50. **Umemoto T., Nakazawa F., Hoshino E., Okada K., Fukunaga M., Namikawa I.** 1997. *Treponema medium sp. nov.,* isolated from human subgingival dental plaque. Int J Syst Bacteriol.**47**:67-72.

51. Wallace, A.L., Harris, A., Allen, J.P., 1967. Reiter treponeme. A review of the literature. 394 Bull. World Health Organ. 36, Suppl: 1-103

52. **Smirbert RM**., 1984. Genus III Treponema. p 49–57. In: Krieg, N.R., Holt, J.G. (Eds.), Bergey's Manual of Systematic Bacteriology. Williams and Wilkins, Baltimore/London,

53. **Mushtaq M., Manzoor S., Pringle M., Rosander A., Bongcam-Rudloff E**. 2015. Draft genome sequence of *'Treponema phagedenis'* strain V1, isolated from bovine digital dermatitis. Stand Genomic Sci. **10**:67-74.

54. **Demirkan I., Williams HF., Dhawi A., Carter SD., Winstanley C., Bruce KD., Hart CA**. 2006. Characterization of a spirochaete isolated from a case of bovine digital dermatitis. J Appl Microbiol.**101;**948-955.

Figure legends

Figure 1. A minimum spanning distance tree of the isolates of the *T. medium* phylogroup (DD1). Further details of the ST information are shown in Table 1. Each ST is coloured based on the proportion of the sequences within it which were isolated from each host. The numbers, indicated in white correspond to the ST numbers shown in Table 1. Key: Dairy- red. Beef- black, Goat- purple, Elk- orange, Sheep- blue, Human- green

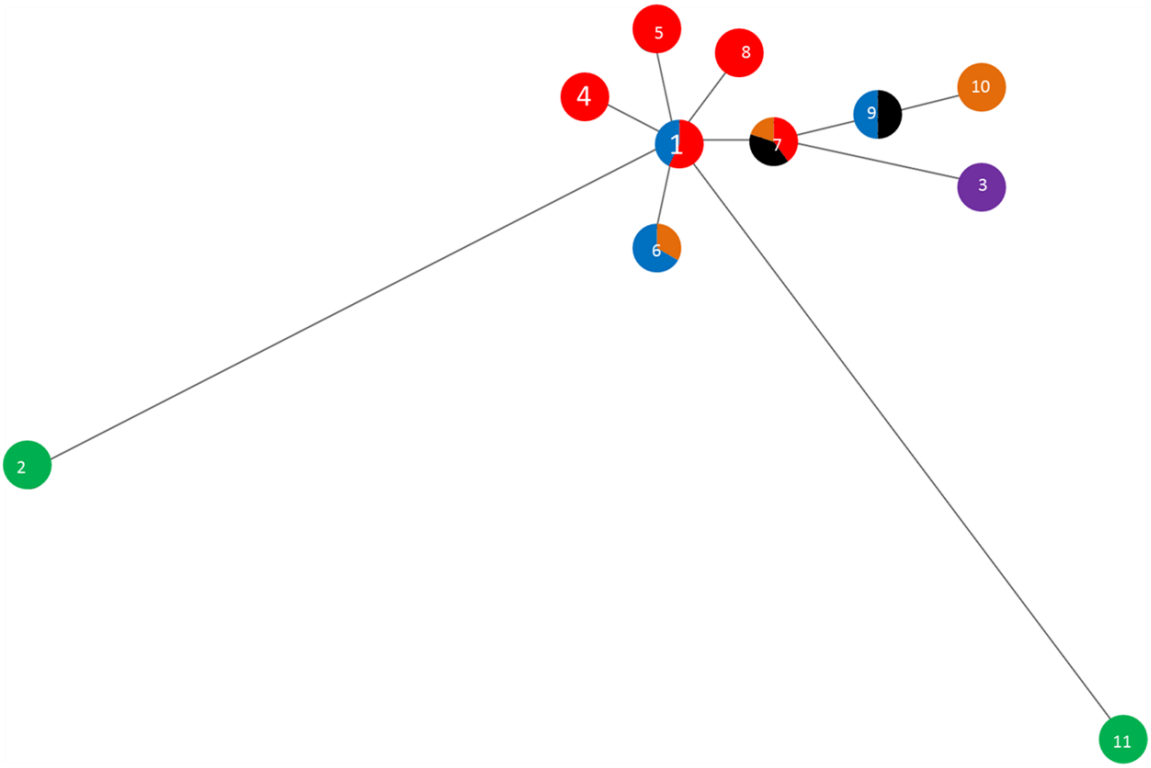
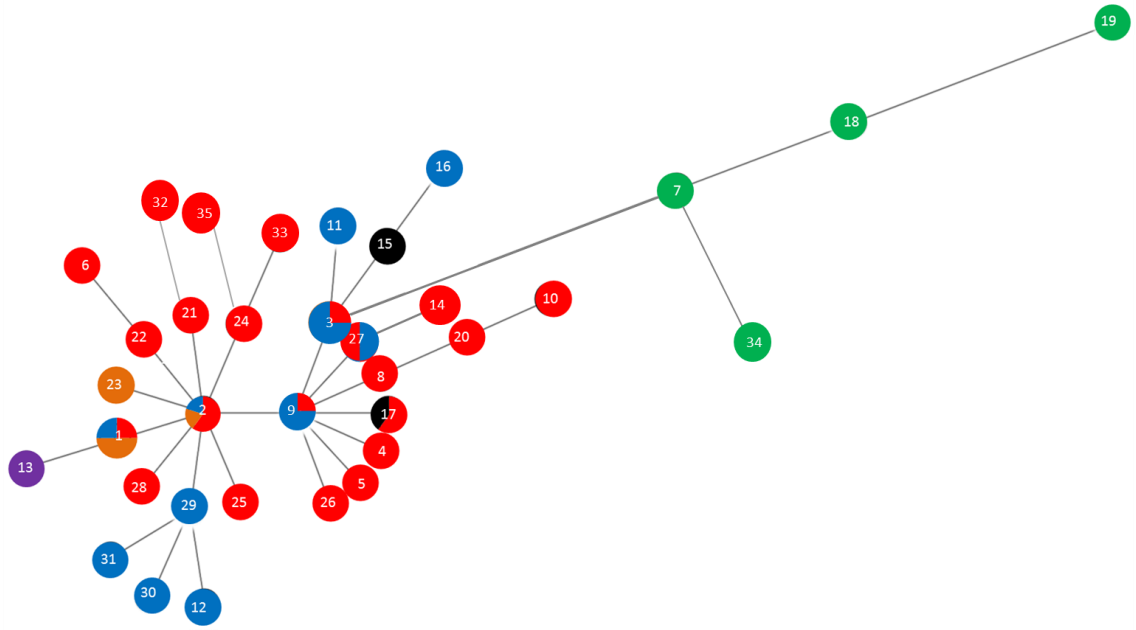

5
8
4
1
7
9
10
3
6
2
11

Figure 2. A minimum spanning distance tree of the isolates of the *T. phagedenis* phylogroup (DD2). Further details of the ST information are shown in Table 2. Each ST is coloured based on the proportion of the sequences within it which were isolated from each host. The numbers, indicated in white correspond to the ST numbers shown in Table 2. Key: Dairy- red. Beef- black, Goat- purple, Elk- orange, Sheep- blue, Human- green


19
18
16
7
32
35
33
11
15
6
21
24
3
14
10
22
27
20
34
8
23
2
9
17
1
4
13
28
29
25
26
5
31
30
12

838 **Figure 3**. A minimum spanning distance tree of the isolates of the *T. pedis*
839 phylogroup (DD3). Further details of the ST information are shown in Table 3.
840 Each ST is coloured based on the proportion of the sequences within it which
841 were isolated from each host. The numbers, indicated in white correspond to
842 the ST numbers shown in Table 3. Key: Dairy- red. Beef- black, Goat-
843 purple, Elk- orange, Sheep- blue, Pig- grey
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863

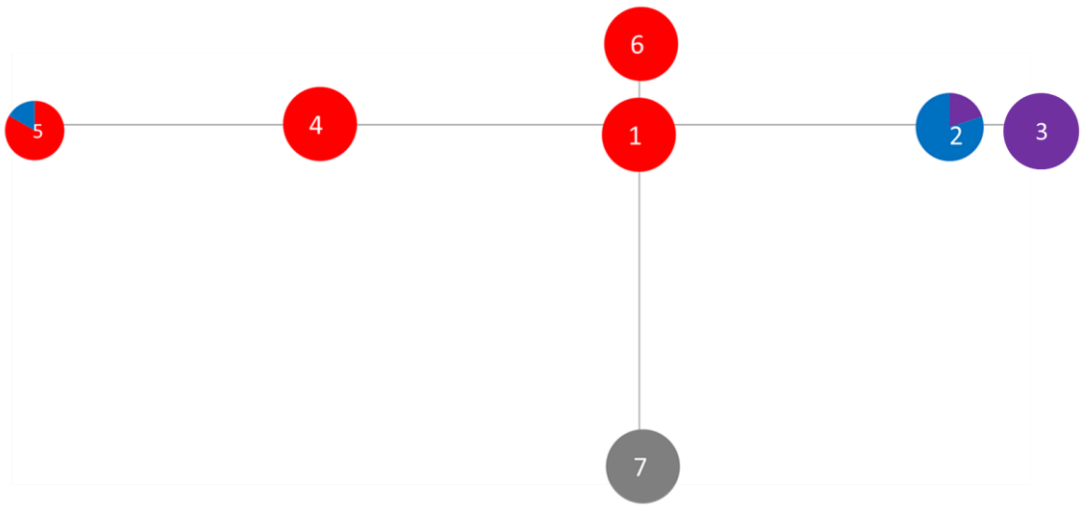
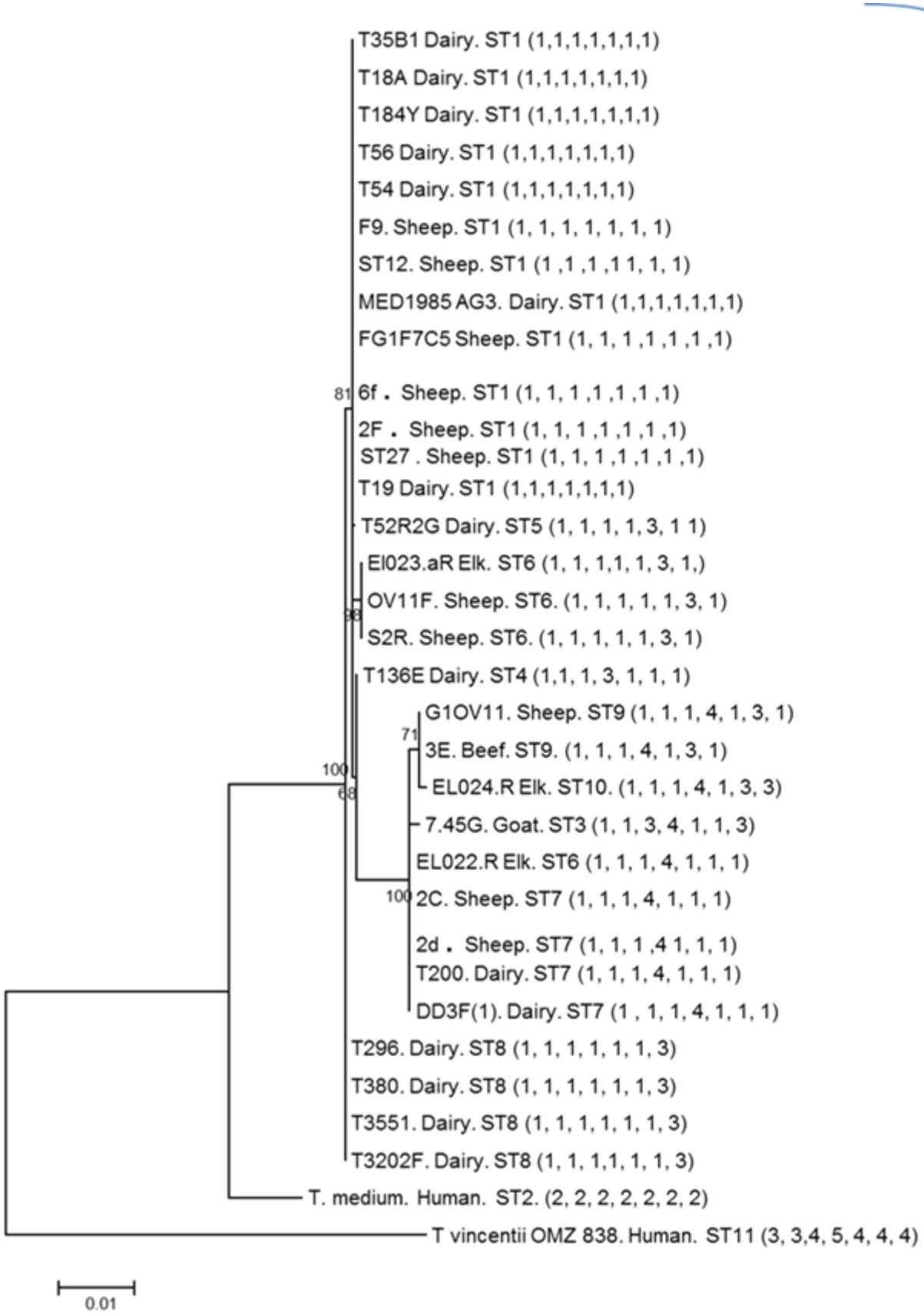

6
5
4
1
2
3
7

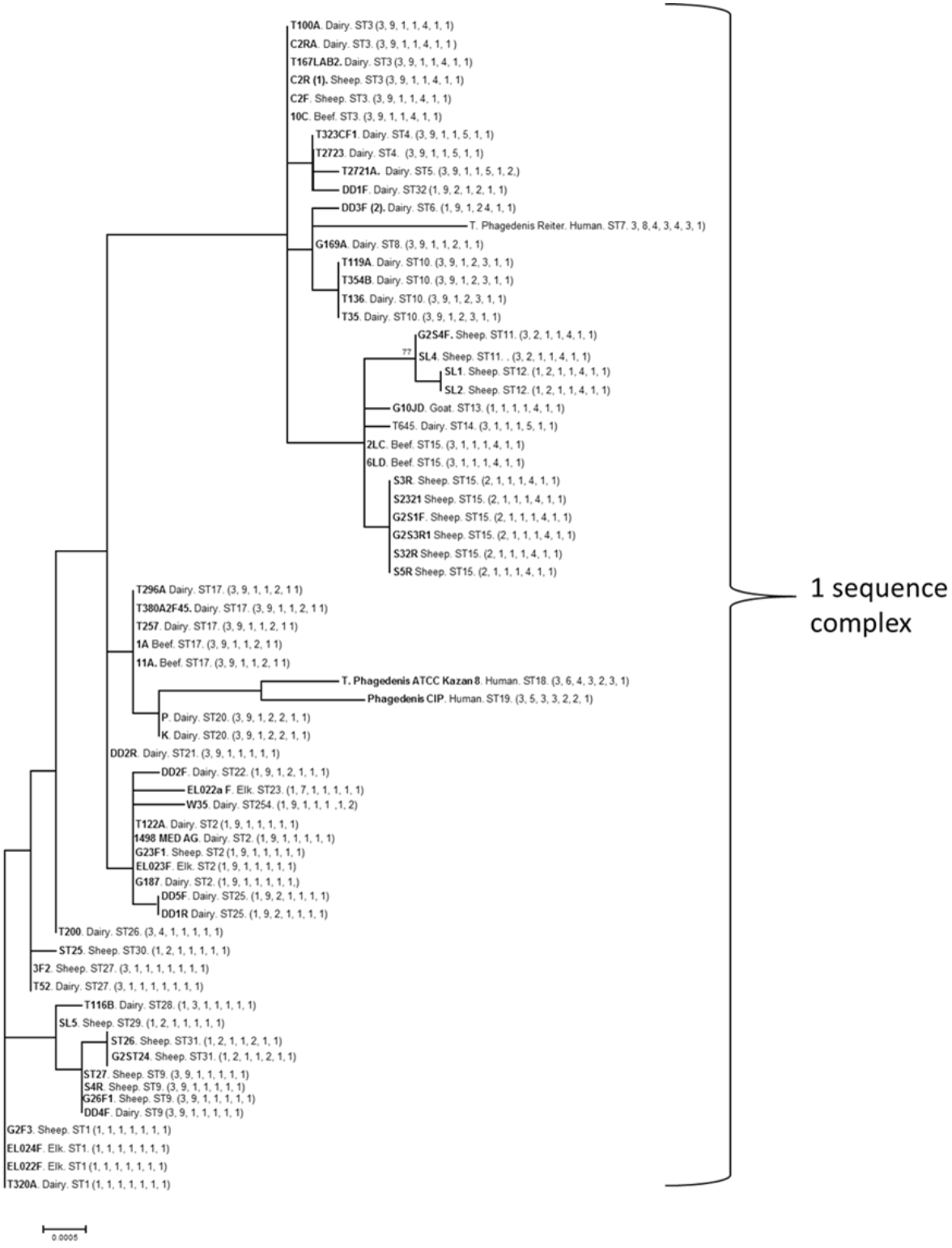
Figure 4. A concatenated gene DNA phylogenetic tree of the seven housekeeping genes for the *T. medium* phylogroup (DD1). Each bacteria is labelled with the isolate name (in bold), host it came from (dairy or beef cattle, sheep, goat, elk, human), the ST it belongs to (Table 1) and the allelic arrangement for that isolate in parenthesis.


T35B1 Dairy. ST1 (1,1,1,1,1,1,1)
T18A Dairy. ST1 (1,1,1,1,1,1,1)
T184Y Dairy. ST1 (1,1,1,1,1,1,1)
T56 Dairy. ST1 (1,1,1,1,1,1,1)
T54 Dairy. ST1 (1,1,1,1,1,1,1)
F9. Sheep. ST1 (1, 1, 1, 1, 1, 1, 1)
ST12. Sheep. ST1 (1 ,1 ,1 ,1 1, 1, 1)
MED1985 AG3. Dairy. ST1 (1,1,1,1,1,1,1)
FG1F7C5 Sheep. ST1 (1, 1, 1 ,1 ,1 ,1 ,1)
81
6f . Sheep. ST1 (1, 1, 1 ,1 ,1 ,1 ,1)
2F . Sheep. ST1 (1, 1, 1 ,1 ,1 ,1 ,1)
ST27 . Sheep. ST1 (1, 1, 1 ,1 ,1 ,1 ,1)
T19 Dairy. ST1 (1,1,1,1,1,1,1)
T52R2G Dairy. ST5 (1, 1, 1, 1, 3, 1 1)
El023.aR Elk. ST6 (1, 1, 1,1, 1, 3, 1,)
OV11F. Sheep. ST6. (1, 1, 1, 1, 1, 3, 1)
S2R. Sheep. ST6. (1, 1, 1, 1, 1, 3, 1)
T136E Dairy. ST4 (1,1, 1, 3, 1, 1, 1)
G1OV11. Sheep. ST9 (1, 1, 1, 4, 1, 3, 1)
71
3E. Beef. ST9. (1, 1, 1, 4, 1, 3, 1)
100
68
EL024.R Elk. ST10. (1, 1, 1, 4, 1, 3, 3)
7.45G. Goat. ST3 (1, 1, 3, 4, 1, 1, 3)
EL022.R Elk. ST6 (1, 1, 1, 4, 1, 1, 1)
100
2C. Sheep. ST7 (1, 1, 1, 4, 1, 1, 1)
2d . Sheep. ST7 (1, 1, 1 ,4 1, 1, 1)
T200. Dairy. ST7 (1, 1, 1, 4, 1, 1, 1)
DD3F(1). Dairy. ST7 (1 , 1, 1, 4, 1, 1, 1)
T296. Dairy. ST8 (1, 1, 1, 1, 1, 1, 3)
T380. Dairy. ST8 (1, 1, 1, 1, 1, 1, 3)
T3551. Dairy. ST8 (1, 1, 1, 1, 1, 1, 3)
T3202F. Dairy. ST8 (1, 1, 1,1, 1, 1, 3)
T. medium. Human. ST2. (2, 2, 2, 2, 2, 2, 2)
T vincentii OMZ 838. Human. ST11 (3, 3,4, 5, 4, 4, 4)
0.01

908

909

Figure 5. A concatenated gene DNA phylogenetic tree of the seven housekeeping genes for the *T. phagedenis* phylogroup (DD2). Each bacteria is labelled with the isolate name (in bold), host it came from (dairy or beef cattle, sheep, goat, elk, human), the ST it belongs to (Table 2) and the allelic arrangement for that isolate in parenthesis.



935

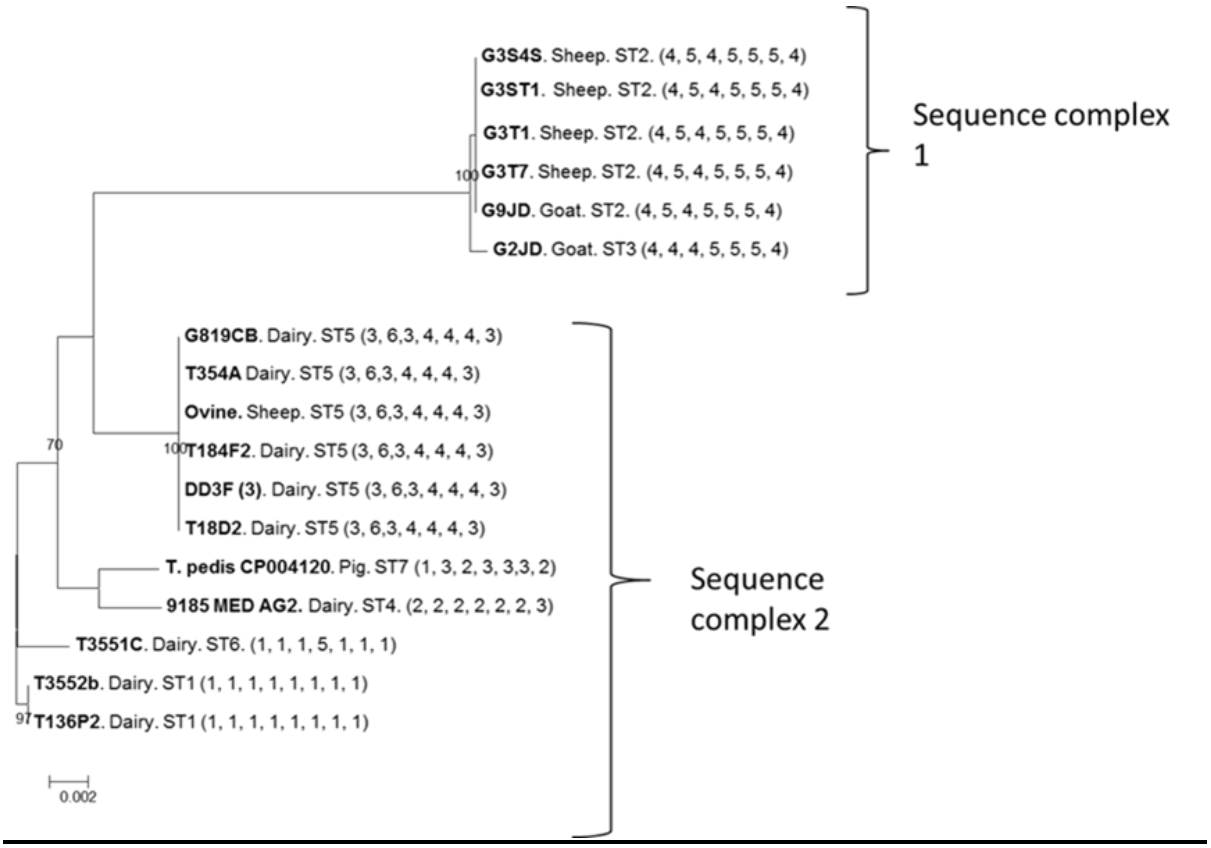
936

937

938

939

Figure 6. A concatenated gene DNA phylogenetic tree of the seven housekeeping genes for the *T. pedis* phylogroup (DD3). Each bacteria is labelled with the isolate name (in bold), host it came from (dairy or beef cattle, sheep, goat, or pig), the ST it belongs to (Table 3) and the allelic arrangement for that isolate in parenthesis.



965

966

967

968

969

970

971

972

973

974

975

976

977

978

Tables

980

Isolate name	Host from which isolate was obtained	Year of isolation	Farm and Geographical provenance of isolate	ST	*groEL*	*recA*	*glpK*	*adK*	*gdh*	*pyrG*	*rplB*	16S rRNA gene Genbank number	Reference
T19	Dairy	2003	Farm A, Merseyside, England	1	1	1	1	1	1	1	1	EF061249	10
2F	Sheep	2013	Farm B, Conwy, Wales	1	1	1	1	1	1	1	1	KP063172	17
ST27	Sheep	2013	Farm C, Conwy, Wales	1	1	1	1	1	1	1	1	KR025808	This study
G1F7C5	Sheep	2013	Farm C, Conwy, Wales	1	1	1	1	1	1	1	1	KP063152	17
G1F9	Sheep	2013	Farm C, Conwy, Wales	1	1	1	1	1	1	1	1	KP063153	17
6F	Sheep	2013	Farm B Conwy, Wales	1	1	1	1	1	1	1	1	KP063174	17
T56	Dairy	2003	Farm A, Merseyside, England	1	1	1	1	1	1	1	1	EF061251	10
T54	Dairy	2003	Farm A, Merseyside, England	1	1	1	1	1	1	1	1	EF061250	10
T184Y	Dairy	2003	Farm A, Merseyside, England	1	1	1	1	1	1	1	1	AY387410	10
T18A	Dairy	2003	Farm A, Merseyside, England	1	1	1	1	1	1	1	1	EF061252	10
T35B1	Dairy	2003	Farm A, Merseyside, England	1	1	1	1	1	1	1	1	KR025809	This study
ST12	Sheep	2013	Farm B. Conwy, Wales	1	1	1	1	1	1	1	1	KR025810	This study
MED1985 AG 3	Dairy	1994	Farm D. California USA	1	1	1	1	1	1	1	1	KR025853	49
T200BA2	Dairy	2004	Farm E, Shropshire,	1	1	1	1	1	1	1	1	KR025811	This study

			England										
T medium ATCC 700293	Human	1972	Japan	2	2	2	2	2	2	2	2	D85437	51
7.45 G	Goat	2013	Farm F, Lancashire, England	3	1	1	3	4	1	1	3	KR025812	This study
T136E	Dairy	2004	Farm G, Shropshire, England	4	1	1	1	3	1	1	1	FJ204242	13
T52R2G	Dairy	2004	Farm G, Shropshire, England	5	1	1	1	1	3	1	1	FJ204241	13
OV11F	Sheep	2009	Farm H, Gloucestershire, England	6	1	1	1	1	1	3	1	KR025813	This study
EL023 aR	Elk	2013	Washington State, USA	6	1	1	1	1	1	3	1	KM586669	5
S2R	Sheep	2009	Farm I, Cheshire, England	6	1	1	1	1	1	3	1	KP063164	17
T200BA1	Dairy	2004	Farm G, Shropshire, England	7	1	1	1	4	1	1	1	KR025814	This study
EL022R	Elk	2013	Washington State, USA	7	1	1	1	4	1	1	1	KM586668	5
DD3F (1)	Dairy	2009	Farm J, Merseyside, England	7	1	1	1	4	1	1	1	KR025815	This study
2c	Beef	2012	Farm K, Gloucestershire, England	7	1	1	1	4	1	1	1	KP859546	This study
2D	Beef	2012	Farm K, Gloucestershire, England	7	1	1	1	4	1	1	1	KP859544	This study
T296	Dairy	2004	Farm L, Cheshire, England	8	1	1	1	1	1	1	3	KR025816	This study
T380	Dairy	2004	Farm J, Merseyside, England	8	1	1	1	1	1	1	3	KR025817	This study
T3551	Dairy	2004	Farm J, Merseyside, England	8	1	1	1	1	1	1	3	KR025818	This study

T3202F	Dairy	2004	Farm J, Merseyside, England	8	1	1	1	1	1	1	3	KR025819	This study
3E	Beef	2012	Farm K, Gloucestershire, England	9	1	1	1	4	1	3	1	KP859545	This study
G1OV11	Sheep	2009	Farm H, Gloucestershire, England	9	1	1	1	4	1	3	1	KP063157	17
EL024 R	Elk	2013	Washington State, USA	10	1	1	1	4	1	3	3	KM586673	5
T vincentii OMZ 838 (CP009227)	Human	1998	China	11	3	3	4	5	4	4	4	CP009227	Unpublished

981
982

983

984 **Table 1**. *T. medium* phylogroup (DD1) isolate details. This includes allelic arrangements (DNA) for the 34 isolates from the *T.*

985 *medium* phylogroup analysed. In addition, the Genbank accession numbers for the 16S rRNA-encoding gene, and papers which it

986 is previously referenced are also shown.

987

988

989

Sample	Origin	Year of isolation	Farm and Geographical provenance of isolate	ST	*groEL*	*recA*	*glpK*	*adK*	*gdh*	*pyrG*	*rplB*	16SrRNA gene genbank number	Reference
T320A	Dairy	2004	Farm J, Merseyside, England	1	1	1	1	1	1	1	1	EF061261	10
G2F3	Sheep	2013	Farm B, Conwy, Wales	1	1	1	1	1	1	1	1	KP063156	17
EL024 F	Elk	2013	Washington State, USA	1	1	1	1	1	1	1	1	KM586672	5
EL022 F	Elk	2013	Washington State, USA	1	1	1	1	1	1	1	1	KM586667	5
EL023 F	Elk	2013	Washington State, USA	2	1	9	1	1	1	1	1	KM586670	5
G187	Dairy	2004	Farm M, Gloucestershire, England	2	1	9	1	1	1	1	1	EF061266	10
G23F1	Sheep	2013	Farm N, Anglesey, Wales	2	1	9	1	1	1	1	1	KP063175	17
1498 MED AG	Dairy	1994	Farm D, California, USA	2	1	9	1	1	1	1	1	KR025851	49
T122A	Dairy	2005	Farm L, Cheshire, England	2	1	9	1	1	1	1	1	FJ204239	13
C2R (1)	Sheep	2009	Farm I, Cheshire, England	3	3	9	1	1	4	1	1	KR025821	This study
C2F	Sheep	2009	Farm I, Cheshire, England	3	3	9	1	1	4	1	1	KR025822	This study
10C	Beef	2012	Farm K, Gloucestershire, England	3	3	9	1	1	4	1	1	KP859543	This study
C2RA	Dairy	2009	Farm L, Cheshire, England	3	3	9	1	1	4	1	1	KR025820	This study
T167LAB2	Dairy	2003	Farm L, Cheshire, England	3	3	9	1	1	4	1	1	EF061253	10
T100A	Dairy	2005	Farm L, Cheshire, England	3	3	9	1	1	4	1	1	FJ204239	13
T323CF1	Dairy	2004	Farm A, Merseyside, England	4	3	9	1	1	5	1	1	EF061263	10
T2723	Dairy	2004	Farm A, Merseyside, England	4	3	9	1	1	5	1	1	FJ204237	13
T2721A	Dairy	2004	Farm A, Merseyside, England	5	3	9	1	1	5	1	2	EF061260	10
DD3F (2)	Dairy	2009	Farm J, Merseyside, England	6	1	9	1	2	4	1	1	KR025823	This study
T phagedenis Reiter	Human	1926	Germany	7	3	8	4	3	4	3	1	KR025824	51
G169A	Dairy	2004	Farm M, Gloucestershire, England	8	3	9	1	1	2	1	1	EF061265	10

ST27	Sheep	2013	Farm B, Conwy, Wales	9	3	9	1	1	1	1	1	KR025825	This study
G26F1	Sheep	2013	Farm O, Denbighshire, Wales	9	3	9	1	1	1	1	1	KP063173	17
DD4F	Dairy	2009	Farm J, Merseyside, England	9	3	9	1	1	1	1	1	KR025826	This study
S4R	Sheep	2009	Farm I, Cheshire, England	9	3	9	1	1	1	1	1	KR025827	This study
T136	Dairy	2004	Farm G, Shropshire, England	10	3	9	1	2	3	1	1	EF061255	10
T119A	Dairy	2004	Farm G, Shropshire, England	10	3	9	1	2	3	1	1	EF061256	10
T354B	Dairy	2004	Farm L, Cheshire, England	10	3	9	1	2	3	1	1	EF061259	10
T35	Dairy	2004	Farm J, Merseyside, England	10	3	9	1	2	3	1	1		This study
SL4	Sheep	2013	Farm N, Anglesey, Wales	11	3	2	1	1	4	1	1	KR025828	This study
G2S4F	Sheep	2009	Farm I, Cheshire, England	11	3	2	1	1	4	1	1	KP063166	17
SL2	Sheep	2013	Farm N, Anglesey, Wales	12	1	2	1	1	4	1	1	KR025829	This study
G2SL1	Sheep	2013	Farm N, Anglesey, Wales	12	1	2	1	1	4	1	1	KP063167	17
G10JD	Goat	2013	Farm F, Lancashire, England	13	1	1	1	1	4	1	1	KJ206532	4
T645C3	Dairy	2004	Farm A, Merseyside, England	14	3	1	1	1	5	1	1	FJ204236	13
6LD	Beef	2013	Farm P, Anglesey, Wales	15	3	1	1	1	4	1	1	KP859539	This study
2LC	Beef	2013	Farm P, Anglesey, Wales	15	3	1	1	1	4	1	1	KP859540	This study
G2S1F	Sheep	2009	Farm Q, Cheshire, England	16	2	1	1	1	4	1	1	KP063163	17
S2321	Sheep	2009	Farm Q, Cheshire, England	16	2	1	1	1	4	1	1	KR025830	This study
S5R	Sheep	2009	Farm Q, Cheshire, England	16	2	1	1	1	4	1	1	KR025831	This study
G2S3R1	Sheep	2009	Farm Q Cheshire, England	16	2	1	1	1	4	1	1	KP063165	17
S32R	Sheep	2009	Farm I, Cheshire, England	16	2	1	1	1	4	1	1	KR025832	This study
S3R	Sheep	2009	Farm I, Cheshire, England	16	2	1	1	1	4	1	1	KR025833	This study
11A	Beef	2012	Farm K, Gloucestershire, England	17	3	9	1	1	2	1	1	KP859541	This study
1A	Beef	2012	Farm K, Gloucestershire, England	17	3	9	1	1	2	1	1	KC907379	This study

T296A	Dairy	2004	Farm L, Cheshire, England	17	3	9	1	1	2	1	1	EF061258	10
T257	Dairy	2004	Farm L, Cheshire, England	17	3	9	1	1	2	1	1	EF061257	10
T380A2F45	Dairy	2004	Farm A, Merseyside, England	17	3	9	1	1	2	1	1	EF061262	10
T phagedenis ATCC Kazan 8	Human	1984	Russia	18	3	6	4	3	2	3	1	KR025835	52
T phagedenis CIP	Human	1962	France	19	3	5	3	3	2	2	1	KR025834	10
P	Dairy	2000	Farm A, Cheshire, England	20	3	9	1	2	2	1	1	KR025836	This study
K	Dairy	2000	Farm A, Cheshire, England	20	3	9	1	2	2	1	1	KR025837	This study
DD2R	Dairy	2009	Farm J, Merseyside, England	21	3	9	1	1	1	1	1	KR025838	This study
DD2F	Dairy	2009	Farm J, Merseyside, England	22	1	9	1	2	1	1	1	KR025839	This study
EL022a F	Elk	2013	Washington State, USA	23	1	7	1	1	1	1	1	KM586666	5
W35	Dairy	2004	Farm L, Cheshire, England	24	1	9	1	1	1	1	2	EF061264	10
DD1R	Dairy	2009	Farm J, Merseyside, England	25	1	9	2	1	1	1	1	KR025840	This study
DD5F	Dairy	2009	Farm J, Merseyside, England	25	1	9	2	1	1	1	1	KR025841	This study
T200	Dairy	2004	Farm G, Shropshire, England	26	3	4	1	1	1	1	1	FJ204240	13
T52	Dairy	2004	Farm G, Shropshire, England	27	3	1	1	1	1	1	1	EF061254	13
3F2	Sheep	2014	Farm N, Anglesey, Wales	27	3	1	1	1	1	1	1	KR025842	This study
T116B	Dairy	2005	Farm A, Merseyside, England	28	1	3	1	1	1	1	1	FJ204238	13
G2SL5	Sheep	2013	Farm N, Anglesey, Wales	29	1	2	1	1	1	1	1	KP063168	17
ST25	Sheep	2013	Farm B, Conwy, Wales	30	1	2	1	1	2	1	1	KR025843	This study
ST26	Sheep	2013	Farm B, Conwy, Wales	31	1	2	1	1	2	1	1	KR025844	This study
G2ST24	Sheep	2013	Farm B, Conwy, Wales	31	1	2	1	1	2	1	1	KP063168	17
DD1F	Dairy	2009	Farm J, Merseyside, England	32	1	9	2	1	2	1	1	KR025845	This study
T. phagedenis 4A	Dairy	unknown	Iowa, USA	33	3	9	1	1	4	1	3	AQCF00000000	16
T. phagedenis F0421	Human	unknown	USA	34	3	7	5	3	4	3	1	NZ_AEFH00000000	16

T. phagedenis V1	Dairy	Unknown	Sweden	35	1	9	1	1	2	1	1	CDNC01000001-CDNC01000051	53

Table 2. Isolation details, with allelic arrangements (DNA) for the 70 isolates from the *T. phagedenis* phylogroup (DD2) analysed as part of this study. In addition, the Genbank accession number for the 16S rRNA gene, and papers which it is previously referenced in are also shown.

Sample	Origin	Year of isolation	Farm and Geographical provenance of isolate	ST	*groEL*	*recA*	*glpK*	*adK*	*gdh*	*pyrG*	*rplB*	16S rRNA gene genbank number	Reference
T3552B^{T}	Dairy	2004	Merseyside, England	1	1	1	1	1	1	1	1	EF061268	10
T136P2	Dairy	2004	Farm E, Shropshire, England	1	1	1	1	1	1	1	1	FJ204243	13
G3ST1	Sheep	2014	Farm R, Shropshire, England	2	4	5	4	5	5	5	4	KP063171	17
G3S4S	Sheep	2014	Farm R, Shropshire, England	2	4	5	4	5	5	5	4	KP063170	17
G3T1	Sheep	2014	Farm R, Shropshire, England	2	4	5	4	5	5	5	4	KR025846	This study
G3T7	Sheep	2014	Farm R, Shropshire, England[1]	2	4	5	4	5	5	5	4	KR025847	This study
G9JD	Goat	2013	Farm F, Lancashire, England	2	4	5	4	5	5	5	4	KJ206531	4
G2JD	Goat	2013	Farm F, Lancashire, England	3	4	4	4	5	5	5	4	KJ206528	4
9185 Med Ag 2	Dairy	1994	Farm D, California, USA	4	2	2	2	2	2	2	3	KR025852	49
T184F2	Dairy	2003	Farm A, Merseyside, England	5	3	6	3	4	4	4	3	KR025848	This study
T18D2 (T18B)	Dairy	2003	Farm A, Merseyside, England	5	3	6	3	4	4	4	3	EF061270	10
DD3F (3)	Dairy	2009	Farm J, Merseyside, England	5	3	6	3	4	4	4	3	KR025849	This study
T354A	Dairy	2004	Farm L, Cheshire,	5	3	6	3	4	4	4	3	EF06126	10

			England									7	
G819CB	Dairy	2004	Farm M, Gloucestershire, England	5	3	6	3	4	4	4	3	EF061269	10
Ovine	Sheep	2006	Farm S, Northern Ireland	5	3	6	3	4	4	4	3	AF363634	54
T3551C	Dairy	2004	Farm A, Merseyside, England	6	1	1	1	5	1	1	1	KR025850	This study
T. pedis CP004120	Pig	2013	Sweden	7	1	3	2	3	3	3	2	CP004120	7

Table 3. Isolation details, with allelic arrangements (DNA) for the 17 isolates from the *T. pedis* phylogroup (DD3) analysed as part of this study. In addition, the Genbank accession numbers for the 16S rRNA gene, and papers in which it is previously referenced are also shown.

Locus	Treponeme group	Putative gene protein	Predicted Product size	Position[1]	Forward primer (5'-3')	Reverse primer (5'-3')
groEL	DD1	Heat shock protein	545 bp	768883-769428	CTTGAATTAAAGCGCGGTATG	AAAATAGCGATATCTTCGAGCATT
	DD2		549 bp		CTTGAGCTGAAACGAGGAATG	GGTAAGAATAGCAATATCTTCAAGCA
	DD3		542 bp		GCTTGAATTAAAACGCGGAAT	CTGCAATATCTTCAAGCATTTCTTT
recA	DD1	Recombination protein A	571 bp	2449887 - 2450338	CTACAAATCGAAAAGGAGTTTGGA	CGTACGCAATACCGATTTTCAT
	DD2		572 bp		GCCTTCAAATCGAAAAACAATTC	GAACATAACGCCGATTTTCAT
	DD3		560 bp		AAATTGAAAAACAATTCGGACAG	AACACCGATTTTCATTCTTATTTGA
glpK	DD1	Glycerol kinase	613 bp	1797272 - 1797770	TATTTTATCATTCGATCAGGGAACA	AATATTCAGTTCCGTCAGAATTTCA
	DD2		610 bp		ATATTTTAGCACTTGATCAGGGAAC	CCGAGTTCTTGTAAAATCTCATCAT
	DD3		589 bp		ATCTTTTGACCAAGGAACTACAAGT	TAACTCATTATCCCATTCCAAAGTC
adK	DD1	Adenosine kinase	517 bp	2265510 - 2265903	CTGCAAAATATTATGGTATCCCTCA	GCATCCAAAGTTATGAGCAGTTTT
	DD2		499 bp		GCTATCAAATCCCGCATATTTC	TTTGCGAGTACATTTTTCTTTTCAT
	DD3		526 bp		TCAAAGTTGTACAAGATACCGCATA	ATGAGGGACGTGCGTCAATA
gdh	DD1	glutamate dehydrogenase	647 bp	275169 - 275682	CGTCAATACTAACGGACAGATTATG	GGTTCTGTACCCATTCAAAGTAAGA
	DD2		643 bp		GTCAACACAAACGGGCAAATAAT	TCTGAACCCATTCAAAGTAAGAAAC
	DD3		623 bp		GTGGGTACAAATGCGAAAATTATG	CATTCAAAATACGAAACAATTACCC
pyrG	DD1	orotidine 5' phosphate dehydrogenase	601 bp	2320945 - 2321441	CAGGTTATCCCGCATGTTACC	ACGCTTCGCTTACGCTTAAATAC
	DD2		611 bp		GTACAAGTTGTCCCGCATGTAAC	GCAGTCAGCGCTTCACTCAC
	DD3		596 bp		GTACCCCATGTAACCGATGAA	AGGGCTTCCACTACGCTTAAATA
rplB	DD1	large polymerase sub unit	565 bp	953257 - 953715	ATATAAGCCTATAACACCGGGTATG	ACCGATTGTTGCATAGCATTTT
	DD2		575 bp		ATAAGCCTATAACACCGGGACTAAG	ATTTCCAACTTCACCGATTGTC
	DD3		575 bp		TCTAAAAGAATATAAGCCGATGACG	CGCCTATGGTAGCATAACATTTTT

1012

Table 4. Treponeme MLST primer design; PCR primers used to generate amplicons of housekeeping genes for the three treponeme phylogroups. One primer set each was developed for the *T. medium* phylogroup (DD1), *T. phagedenis* phylogroup (DD2) and the *T. pedis* phylogroup (DD3).

1. Position of the genes corresponds to *T. vincentii* OMZ 383 (Genbank accession code- CP009227)

Locus	*T. medium* (DD1) phylogroup (n= 33)					*T. phagedenis (DD2)* phylogroup (n= 71)					*T. pedis* (DD3) phylogroup (n=17)				
	Amplicon Size (bp)	Number (and percentage) of variable sites (DNA): with [and without] inclusion of *T. vincentii*	Number (and percentage) of variable sites (AA) with [and without] inclusion of *T. vincentii*	DNA Alleles	AA alleles	Gene Size (bp)	Number (and percentage) of variable sites (DNA)	Number (and percentage) of variable sites AA	AA alleles	DNA Alleles	Gene Size (bp)	Number (and percentage) of variable sites (DNA)	Number (and percentage) of variable sites AA)	DNA Alleles	AA Alleles
groEL	448	40 (9%); [15, 3%]	0; 0	3	1	456	6 (1.3%)	4 (3%)	2	3	441	13 (3%)	0	4	1
recA	475	64, (13%); [12,2%]	59, (37%); [11, 7%]	3	3	472	12 (2.5%)	4 (3%)	3	9	477	10 (2%)	0	6	3
glpK	507	34 (7%), [20, 11%]	31 (7%) ; [18, 10%]	4	4	521	4 (0.7)	3 (1.7%)	4	5	508	5 (1%)	5 (3%)	4	4
adK	416	69 (17%); [27, 6%]	57 (41%); [23,17%]	5	6	394	2 (0.5)	1 (0.7%)	2	3	421	13 (3%)	3 (2)	4	4
gdh	514	47 (9%); [11, 6%]	7 (1%) ; [2, 1%]	4	4	560	10 (1.8%)	9 (5%)	2	5	520	22 (4%)	16 (11%)	5	4
pyrG	501	52, (10%);[21,4%]	47, 28%; [18, 11%]	4	4	527	5 (0.9%)	0	1	3	507	21 (4%)	0	5	2
rplB	469	54 (11 %); [8, 2%]	47 (30%); [8, 5%]	4	4	475	3 (0.65)	2 (1%)	2	2	502	10 (2%)	0	4	4

Table 5. Analysis of individual genes. Gene sizes and allelic arrangement, both at nucleotide and amino acid level are shown. As *T. vincentii* appears to form a separate species, it was analysed in conjunction with, and separately from the *T. medium* phylogroup.