Chlamydia suis undergoes interclade recombination promoting Tet-island exchange

Helena Seth-Smith[1] , Sankhya Bommana[2] , Deborah Dean[2,3] , Timothy D. Read[4,5] and Hanna Marti[1*]

Abstract

Background The obligate intracellular bacterial family *Chlamydiaceae* comprises a number of different species that cause disease in various vertebrate hosts including humans. *Chlamydia suis*, primarily found in the gastrointestinal tract of pigs, is the only species of the *Chlamydiaceae* family to have naturally gained tetracycline resistance (TetR), through a genomic island (Tet-island), integrated into the middle of chromosomal *invasin*-like gene *inv*. Previous studies have hypothesised that the uptake of the Tet-island from a host outside the *Chlamydiaceae* family was a unique event, followed by spread among *C. suis* through homologous recombination. In vitro recombination studies have shown that Tet-island exchange between *C. suis* strains is possible. Our aim in this study was to gain a deeper understanding of the interclade recombination of the Tet-island, among currently circulating *C. suis* field strains compared to in vitro-generated recombinants, using published whole genome sequences of *C. suis* field strains (*n* = 35) and in vitro-generated recombinants (*n* = 63).

Results We found that the phylogeny of *inv* better reflected the phylogeny of the Tet-island than that of the whole genome, supporting recombination rather than site-specific insertion as the means of transfer. There were considerable differences between the distribution of recombinations within in vitro-generated strains compared to that within the field strains. These differences are likely because in vitro-generated recombinants were selected for a tetracycline and rifamycin/rifampicin resistant background, leading to the largest peak of recombination across the Tet-island. Finally, we found that interclade recombinations across the Tet-island were more variable in length downstream of the Tet-island than upstream.

Conclusions Our study supports the hypothesis that the occurrence of TetR strains in both clades of *C. suis* came about through interclade recombination after a single ancestral horizontal gene transfer event.

Keywords *Chlamydia*, *Chlamydiaceae*, Antibiotic resistance, Horizontal gene transfer, Homologous recombination, Tetracycline, Comparative phylogenetics

*Correspondence:
Hanna Marti
hanna.marti@uzh.ch
[1]Institute of Veterinary Pathology, University of Zurich, Zurich, Switzerland
[2]Division of Infectious Diseases and Global Health, Departments of Medicine and Pediatrics, University of California San Francisco School of Medicine, San Francisco, CA, USA

[3]Joint Graduate Program in Bioengineering, University of California Berkeley, University of California San Francisco, Berkeley and San Francisco, CA, USA
[4]Division of Infectious Diseases, Department of Medicine, Emory University School of Medicine, Atlanta, GA, USA
[5]Department of Human Genetics, Emory University School of Medicine, Atlanta, GA, USA

Background

Chlamydia (*C.*) *suis* belongs to the *Chlamydiaceae* family, Gram-negative, obligate intracellular bacteria that are primarily responsible for respiratory, ocular and urogenital disease in humans and animals [1]. *C. suis* is a pig pathogen and has mostly been associated with mild clinical signs such as conjunctivitis, diarrhoea but also reproductive disorders to a limited degree [2]. Like many of the *Chlamydiaceae*, *C. suis* is a zoonotic pathogen and has been detected in the eyes, pharynges and rectum of people that come in close contact with pigs, namely farmers and slaughterhouse workers [3–5], as well as in the eyes of trachoma-endemic populations that domesticate pigs [6].

While the clinical impact of *C. suis* on porcine health is mild, *C. suis* has received attention in recent years as the only *Chlamydiaceae* species to have naturally obtained an antibiotic resistance gene: the tetracycline resistance-conferring *tet*A(C) gene [7, 8]. This gene has been detected in many different *C. suis* strains from Europe, Israel and the USA [1]. Where present, it is consistently identified to be part of an over 12 kilobase pair (kb) genomic island termed the Tet-island, likely derived from a plasmid originating from Proteobacteria, located within the chromosomal invasin-like gene *inv* [9, 10]. The nearly 4 kb *inv* gene has an unknown function, and is disrupted by the Tet-island insertion [8, 11]. The complete *inv* gene is only found in two chlamydial species, *C. suis* and *C. caviae*, a pathogen found in guinea pigs. In *C. muridarum* and human *C. trachomatis*, the two closest phylogenetic relatives of *C. suis*, the *inv* gene is truncated or entirely absent, respectively [11, 12].

Discrepancies between the mutation rate of the Tet-islands and the chromosomes led to the hypothesis that *tet*A(C) acquisition in *C. suis* happened relatively recently, and certainly after the separation of the species into two major ancestral clades [9, 10]. Given current data on sequenced strains, the Tet-island possibly originated in the USA before being later transferred to European *C. suis* strains [9, 10]. Moreover, based on the unique structure of the Tet-island and its invariable position within the *C. suis* chromosome, it has been hypothesised that the original acquisition was a rare and possibly singular horizontal gene transfer (HGT) event followed by spread of the Tet-island among *C. suis* strains through homologous recombination [9, 10]. This hypothesis presupposes the occurrence of interclade recombination events after the Tet-island was integrated into *C. suis*. Intra- and interspecies recombination is well-recorded for many *Chlamydiaceae* species [13], however the extent of interclade recombination has so far not been analysed in detail for *C. suis*.

In this study, we investigated the extent of interclade recombination, in both currently available sequences from field isolates, and in vitro-generated recombinants. We performed phylogenetic comparisons of the *inv* gene to that of the Tet-island and of the whole genome using all available *C. suis* genomes, paying special attention to interclade recombination events. Second, we used an established co-culture model [14] using tetracycline-sensitive (TetS) and tetracycline-resistant (TetR) *C. suis* strains from the two major clades to investigate interclade recombination and Tet-island transfer dynamics in vitro.

Materials and methods

Generation of in vitro-generated recombinants

The detailed protocol has already been published [14]. Briefly, tetracycline-sensitive, rifamycin-resistant strains S45 RIF, 94 Ry and 111 Ry were individually co-cultured with SWA-141 (4-29b), SWA-107 (5-27b) or SWA-110 (1-28b) in LLC-MK2 cells (continuous Rhesus monkey kidney cell line, kindly provided by IZSLER Brescia, Italy) at 37 °C and 5% CO_2. After the first passage, selection was added using inhibitory concentrations of rifampicin (Merck, Darmstadt, Germany) or rifamycin (Merck), and tetracycline (Merck). After the second passage, plaque assays were performed as previously described to obtain individual inclusions [15]. Putative recombinants were initially identified by strain-specific PCR followed and confirmed by stability assays for five to ten passages in the presence and absence of selective antibiotics and PCR re-analysis of passaged cultures. Whole-genome sequencing was performed on a subset of these confirmed recombinants, and the recipient strains, using the Illumina MiSeq platform [14].

Sequence analysis

All read data for field strains [10] and recombinant strains [14] were obtained from SRA or were deposited under PRJNA668469 (Table S1). The complete assembled genome of strain 8-29b (NZ_FTQU01000001) was used as the reference against which to map all reads within CLC Genomics Workbench v20.0.2 and also generate a single nucleotide polymorphism (SNP) phylogeny with parameters that differed from the default as: variant calling with 10x minimum coverage, 10 minimum count and 70% minimum frequency (after which mean coverage was assessed), and SNP tree creation with 10x minimum coverage, 10% minimum coverage, 0 prune distance and including multi-nucleotide variants (MNVs). Genome consensi were extracted from mapped reads using a minimum coverage of 5, a multiple sequence alignment (MSA) was created from the whole genome alignments, exported and run through Gubbins v3.0.00 with 5 iterations. Results were viewed in Artemis [16] and Phandango [17]. FastBaps v 1.0.8 [18] was used to calculate the population structure. BactDating v1.1.1 was run on the

Seth-Smith *et al. BMC Genomics*   (2024) 25:724

Page 3 of 7

gubbins output with up to 5 million iterations and the arc model (Table S6). Traces did not converge, and between three runs, most likely root dates varied from −7433.75 [-20318.78;-109.86] to -5626.65 [-13128.00;24.31], with mu ranging from 2.57e+00 [7.44e-01;7.72e+00] to 2.93e+00 [1.09e+00;8.30e+00].

The *inv* gene was reconstructed from previous assemblies [10] and the Tet-island, where present, was excluded from the resulting MSA, which was generated using MUSCLE [19] in AliView v 1.27 [20]. A phylogeny was created using RAxML (https://raxml-ng.vital-it.ch/) using defaults and 100 bootstraps.

The Tet-island phylogeny was generated from the previous alignment, corrected for the strains used in the whole genome phylogeny (*n*=22), using RAxML as above.

Phylogenies were compared using phylo.io [21].

Analysis of in vitro-generated recombinants
Gubbins alignment of in vitro-generated recombinants and all recipient/donor strains was performed. The alignment was then used for subsequent analysis with the Geneious Prime software (v. 2023.1.1; Biomatters, Auckland, New Zealand). Specifically, we extracted individual recombinants and their associated recipient strains and identified recombinations using the "Find Variation/SNPs" function. All areas with three or more SNPs were identified and marked as motifs. Recombinant regions were then curated following comparison of recipient, donor and recombinant using MAFFT alignment followed by the "Find Variation/SNPs" function and curation of identified motifs. Areas with high variation between all three strains were excluded from downstream recombinant analysis. Genes at the predicted ends of the recombinations were identified following alignment against recombinant strain 217.1, which was annotated using Prokka [22].

Statistical analysis
We used GraphPad Prism (v. 10.0.2, GraphPad Software, Boston, MA, USA) for statistical analyses. Statistical tests included either one-way ANOVA (Kruskal-Wallis test) with Dunn's multiple comparisons test for multiple means, or the Mann-Whitney test if two groups were compared.

Results and discussion
Whole genome phylogeny of field isolates
Available genome data from tetracycline-resistant and -sensitive isolates ("field strains") were collected (*n*=41, Table S1), from previous publications [9, 10]. All genomes with under 15x mean read coverage were excluded (*n*=6) providing *n*=35 genomes to analyse. The phylogeny (Fig. 1) closely agrees with that already published [10], despite the use of alternative analysis tools.

Bayesian clustering was used to confirm the separation into two clades. Recombination analysis of the collection of field strains also identified highly impacted loci (Fig. 1), including the Tet-island, as previously described. Bayesian dating of the ancestor of the species was attempted, giving results unfortunately with high uncertainty (see methods). The most recent common ancestor appears to have emerged several thousand years before the common era; further data is required to improve upon this analysis.

Comparisons of phylogenies of whole genomes, insertion site *inv* and Tet-island of field strains
The complete *inv* gene was reconstructed from genomes containing disrupted versions of the gene, and a phylogeny of *inv* was created to investigate the degree of agreement with a whole genome phylogeny with recombination events removed (Fig. 2A). A high degree of clade mixing is clear in the *inv* phylogeny, which does not reflect the whole genome phylogeny, and suggests recombination of *inv*. A comparison of the phylogeny of *inv* with that of the Tet-islands from Seth-Smith et al. [10], shows a much higher congruence (Fig. 2B), strongly suggesting that the *inv* gene is linked to the Tet-island during Tet-island movement, and not that it serves purely as an integration site. This is particularly clear between clades and speaks for recombination as a means of Tet-island transfer.

Analysis of recombination sites in field and in vitro recombinant isolates
A phylogeny of in vitro recombinant genomes from previous plaque-purified strains generated from Tet-island transfer experiments [14] (*n*=63, all over 25x mean read coverage), including the six parental strains was generated, to compare against that of the field strains (Fig. 3). The recombinant strains cluster phylogenetically around the recipient parental strains, all of which are from Clade 2 (SWA-94 is 10-26b; SWA-111 is 1-28a; and S45), compared to the Clade 1 Tet-donor strains (SWA-141 is 4-29b; SWA-107 is 5-27b; and SWA-110 is 1-28b) (Fig. 1). The distribution of recombinations within the in vitro-derived recombinants is clearly different from that within the field strains. The in vitro-derived recombinants were selected on the basis of having acquired the Tet-island in a rifamycin/rifampicin resistant background. Hence the largest peak of recombination is across the Tet-island with the downstream recombination site extending further relative to the Tet-island compared to the upstream recombination site. The predicted mean recombination size across all recombinations is 10.6 kb, ranging between 4 bp and 135 kb (Table S2), which stands in contrast to comparison of the field strains where the average
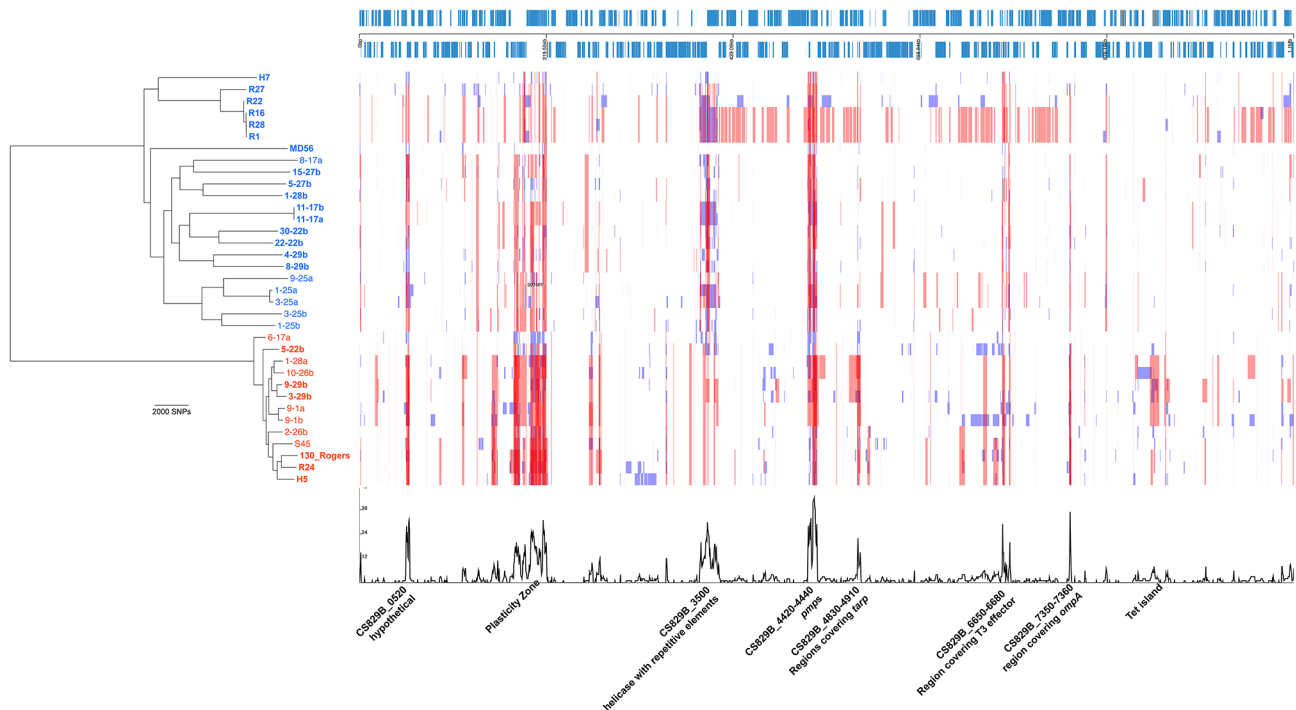
Fig. 1 Phylogeny of *C. suis* field strains (*n*=35) and identified recombinations. The phylogenies with recombinations removed (left) are coloured with Clade 1 in blue and Clade 2 in orange (FastBAPS). Names in bold define strains that carry the Tet-island. The sample names are matched to the genome length tracks (right) where red bars indicate recombinations identified in more than one genome, and blue bars show those only identified in single genomes. The genome with annotated coding sequences (CDSs in blue, both forward and reverse reading frames) of strain 8-29b is shown above these tracks. Below the tracks is a plot of recombination density within the phylogenies, with peaks annotated according to the relevant genome regions, CDSs and predicted encoded proteins. Data around the Tet-island illustrates that not all strains carry the island; *inv* fragments at both sides are commonly predicted to be involved in recombinations. The data was generated using Gubbins [23] and visualised using Phandango [17]

recombination size was predicted to be much shorter with 1.1 kb (2 bp-21.2 kb, Table S3).

We could not identify consistent upstream or downstream recombination junctions involving the Tet-island of the in vitro-derived recombinants, suggesting that a general homologous, rather than site-specific, recombination mechanism is responsible. Further recombinations across the genome are also apparent, despite not having been selected for, showing the promiscuity of recombination under permissive conditions.

Recombinations in in vitro-generated recombinants reveal high variability in selected and non-selected areas

Predicted recombinations within in vitro-generated strains were analysed in detail by finding the variations between individual recombinants with their respective parental strains using Geneious Prime (Table S4). Overall, we identified a total of 237 recombinations with a mean size of 26.6 kb (range: 20 bp to 492 kb). We detected an average of 3.8 recombinations per in vitro-generated strain (*n*=63), all of which included one recombination involving the Tet-island; recombinations across this region were notably longer than regions that were not under specific selective pressure, with a mean

size of 9.5 kb (20 bp to 193 kb) and 73.8 kb (15.2–492 kb) for recombinations without and with the Tet-island, respectively. This confirms the results of our previous study which contained a smaller sample size [14]. These findings are also in line with a co-culture study co-infecting *C. muridarum* and a tetracycline-resistant *C. suis* which identified a 98 kb recombinant region involving the Tet-island [24]. In co-infection experiments that did not include the Tet-island and corresponding tetracycline selection, average recombinant regions were variable. One study concerning *C. muridarum / C. trachomatis*, produced recombinant region sizes similar to those in our study, ranging between 558 bp and 124 kb [25], while other studies investigating *C. trachomatis* yielded longer regions ranging between 200 and 400 kb [26, 27].

We then investigated recombinations involving the Tet-island relative to their distance to *inv* (Table S5). There was no significant difference in the recombination extension length upstream and downstream (*p*=0.7). However, while the upstream recombination site is an average of 17.1 kb upstream of the proximal *inv* fragment (56/63) and never further upstream than 53 kb, the downstream recombination site ranges between 0 kb and 440 kb from the distal fragment of *inv*. The reason for this finding is
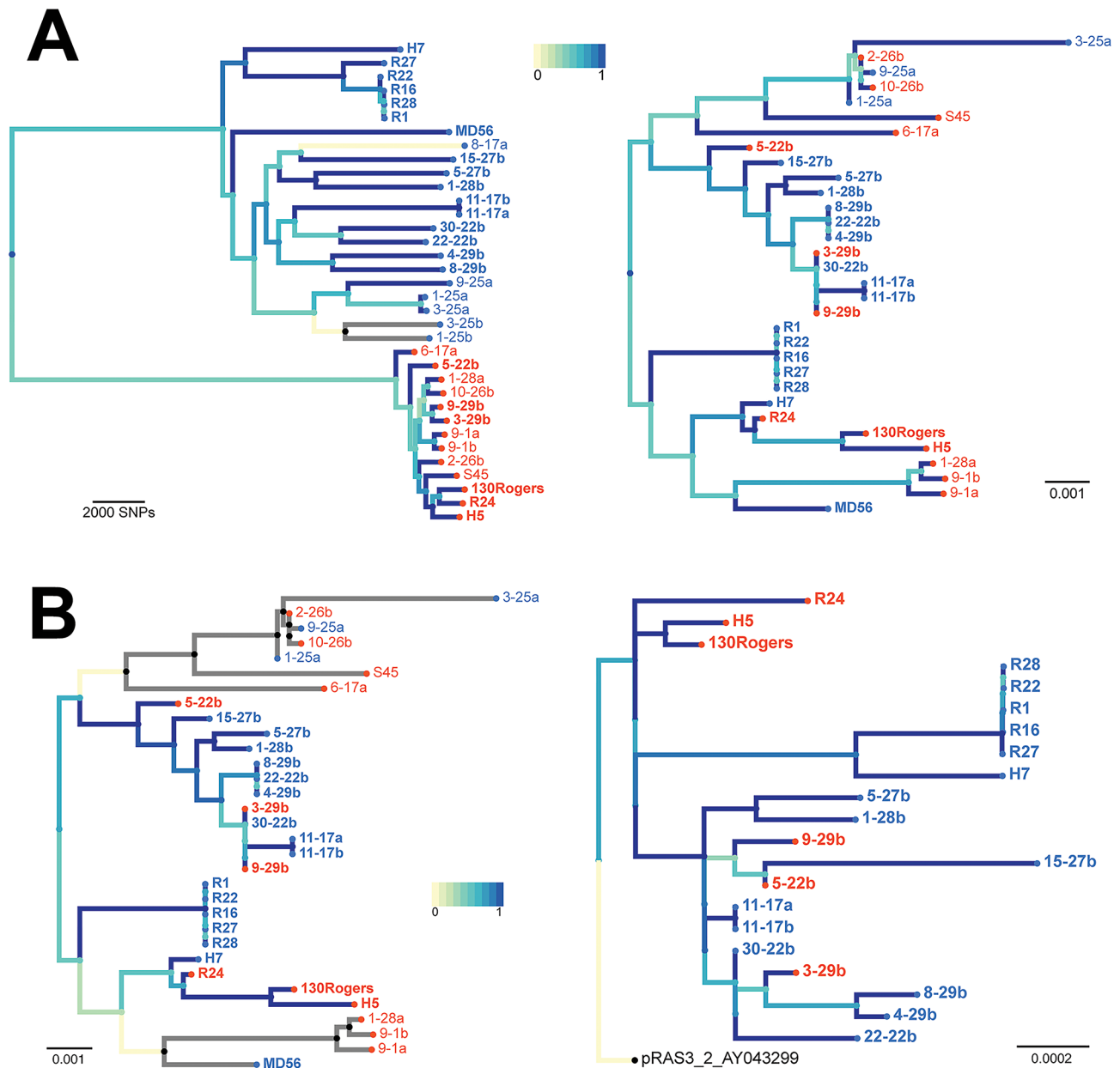
Fig. 2 Comparison of phylogenies. The colour scale of the branches shows the weighting of the similarities. Names in bold define strains that carry the Tet-island. **A**: Whole genome (left) vs. *inv* (right) phylogenies. **B.** *inv* phylogeny (left) vs. Tet-island phylogeny (right; [10]). Leaves are coloured with Clade 1 in blue and Clade 2 in orange

unclear. In *Chlamydia*, both homologous recombination pathways—double-strand break-dependent RecBCD and single-strand break-dependent RecFOR—have been identified [13]. One unknown component is the identity of Chi sites, the site where the double-stranded degradation of RecBCD ends and single-strand degradation begins for the binding of RecA [13]. It is possible that there are one or more Chi sites close to the upstream *rrn* operon resulting in more constrained recombination sites compared to recombination downstream of the

Tet-island. However, we do not have sufficient data to support this hypothesis.

Limitations

Sample and whole genome sequencing of further Tet-island carrying and tetracycline-sensitive isolates of *C. suis* would better inform all aspects of this study.

In the current study, all in vitro recombinations were performed with a Tet donor in Clade 1 and a recipient in Clade 2. To further investigate interclade recombination,
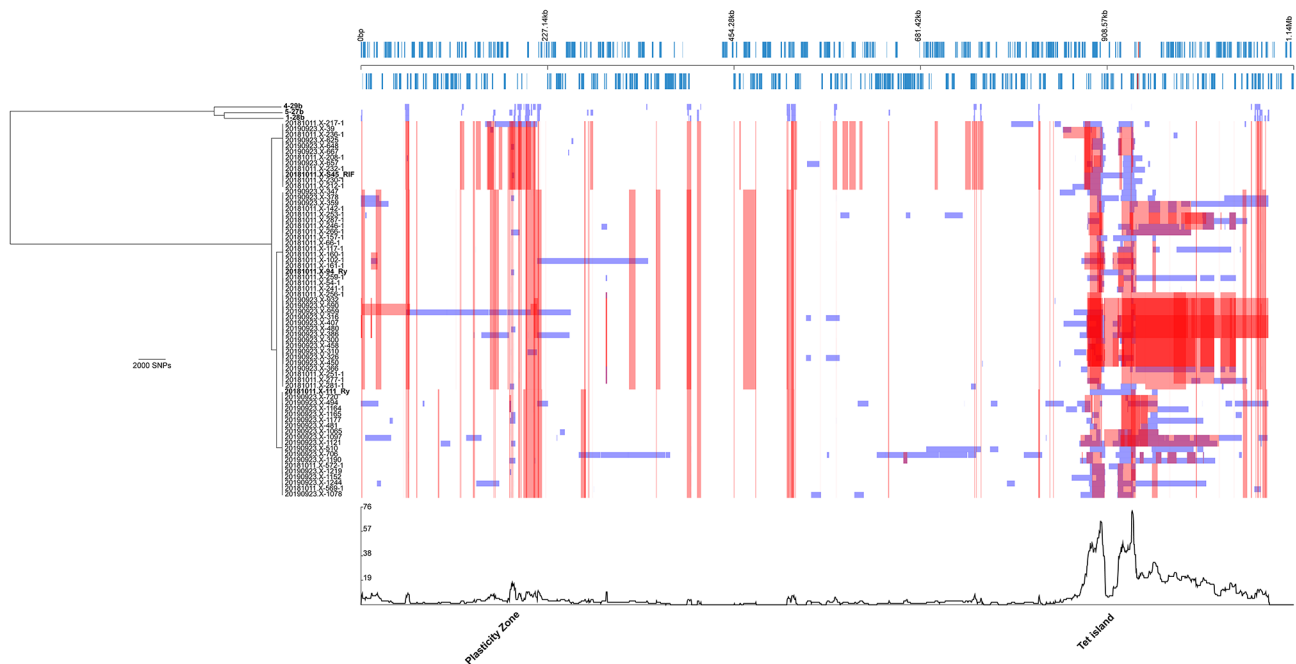
Fig. 3 Phylogeny of *C. suis* in vitro-generated recombinant strains *n* = 63 including Tet-carrying (*n* = 3) and derived resistant parental (*n* = 3) strains (bold) and identified recombinations. The phylogenies with recombinations removed (left) have bold names representing parental strains. The sample names are matched to the genome length tracks (right) where red bars indicate recombinations identified in more than one genome, and blue bars show those only identified in single genomes. Where recombinations are identified in more than one strain (red recombination blocks), this is either due to recombinations identified between the parent strains (Figure S1), or to duplicate picking of plaques from the same parent combination. The genome with annotated coding sequences (CDSs in blue, both forward and reverse reading frames) of strain 8-29b is shown above these tracks. Below the tracks is a plot of recombination density within the phylogenies, with peaks annotated according to the relevant genome regions, CDSs and predicted encoded proteins. The data was generated using Gubbins [23] and visualised using Phandango [17]

Tet donors and recipients in reverse clades would prove interesting.

Conclusion

In this study, by comparing phylogenies of genomes and genomic elements, we show that the *inv* gene, as the location of the important Tet-island in *C. suis*, appears to transport with the Tet-island during recombination events, rather than acting as an integration site. Additionally, the distribution of the Tet-island across the phylogeny does not lend itself to the inference of a simple vertical inheritance, and recombination must be invoked to explain the pattern most parsimoniously. Some recombination events appear to have occurred been between Clade 1 and Clade 2 (interclade). A high proportion of sequenced *C. suis* field strains carry the Tet-island (Fig. 1), which may represent bias in the strains selected for sequencing. Sadly, historical isolates were not collected, and the history of the insertion of the Tet-island remains unknown. We add further evidence of the high rates of recombination in *C. suis*, occurring with high frequency between the two main clades, and readily transferring the Tet-island.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10606-6.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Seth-Smith *et al. BMC Genomics*   (2024) 25:724

Page 7 of 7

Data availability

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References

1. Sachse K, Borel N. Recent Advances in Epidemiology, Pathology and Immunology of Veterinary Chlamydiae. In: Tan M, Hegeman JH, Sütterlin C, editors. Chlamydia Biology: From Genome to Disease [Internet]. Norfolk: Caister Academic Press; 2020 [cited 2020 Mar 6]. pp. 403–28. https://www.caister.com/chlamydia2.
2. Schautteet K, Vanrompay D. *Chlamydiaceae* infections in pig. Vet Res. 2011;42(1):29.
3. De Puysseleyr K, De Puysseleyr L, Dhondt H, Geens T, Braeckman L, Morré SA, et al. Evaluation of the presence and zoonotic transmission of *Chlamydia suis* in a pig slaughterhouse. BMC Infect Dis. 2014;14(1):560.
4. De Puysseleyr L, De Puysseleyr K, Braeckman L, Morré SA, Cox E, Vanrompay D. Assessment of *Chlamydia suis* infection in Pig Farmers. Transbound Emerg Dis. 2017;64(3):826–33.
5. Kieckens E, Van den Broeck L, Van Gils M, Morré S, Vanrompay D. Co-occurrence of *Chlamydia suis* DNA and *Chlamydia suis*-specific antibodies in the Human Eye. Vector-Borne Zoonotic Dis. 2018;18(12):677–82.
6. Dean D, Rothschild J, Ruettger A, Kandel RP, Sachse K. Zoonotic *Chlamydiaceae* species Associated with Trachoma, Nepal. Emerg Infect Dis. 2013;19(12):1948–55.
7. Lenart J, Andersen AA, Rockey DD. Growth and development of tetracycline-resistant *Chlamydia suis*. Antimicrob Agents Chemother. 2001;45(8):2198–203.
8. Dugan J, Rockey DD, Jones L, Andersen AA. Tetracycline Resistance in *Chlamydia suis* mediated by genomic islands inserted into the Chlamydial *inv*-like Gene. Antimicrob Agents Chemother. 2004;48(10):3989–95.
9. Joseph SJ, Marti H, Didelot X, Read TD, Dean D. Tetracycline selective pressure and homologous recombination shape the evolution of *Chlamydia suis*: a recently identified Zoonotic Pathogen. Genome Biol Evol. 2016;8(8):2613–23.
10. Seth-Smith HMB, Wanninger S, Bachmann N, Marti H, Qi W, Donati M, et al. The *Chlamydia suis* genome exhibits high levels of diversity, plasticity, and Mobile Antibiotic Resistance: Comparative Genomics of a recent livestock cohort shows influence of treatment regimes. Genome Biol Evol. 2017;9(3):750–60.
11. Burall LS, Liu Z, Rank R, Bavoil PM. The chlamydial *invasin*-like protein gene conundrum. Microbes Infect. 2007;9(7):873–80.
12. Jordan S, Nelson D, Geisler W. *Chlamydia trachomatis* Infections. In: Tan M, Hegeman JH, Sütterlin C, editors. *Chlamydia* Biology: From Genome to Disease [Internet]. Norfolk: Caister Academic Press; 2020. pp. 1–30. https://www.caister.com/chlamydia2.
13. Marti H, Suchland RJ, Rockey DD. The impact of lateral gene transfer in *Chlamydia*. Front Cell Infect Microbiol. 2022;12:184.
14. Marti H, Bommana S, Read TD, Pesch T, Prähauser B, Dean D, et al. Generation of tetracycline and rifamycin resistant *Chlamydia suis* recombinants. Front Microbiol. 2021;12:630293.
15. Marti H, Kim H, Joseph SJ, Dojiri S, Read TD, Dean D. *Tet*(C) gene transfer between *Chlamydia suis* strains occurs by homologous recombination after co-infection: implications for spread of tetracycline-resistance among *Chlamydiaceae*. Front Microbiol. 2017;8(FEB):156.
16. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. Bioinformatics. 2000;16(10):944–5.
17. Hadfield J, Croucher NJ, Goater RJ, Abudahab K, Aanensen DM, Harris SR. Phandango: an interactive viewer for bacterial population genomics. Bioinformatics. 2018;34(2):292–3.
18. Tonkin-Hill G, Lees JA, Bentley SD, Frost SDW, Corander J. Fast hierarchical bayesian analysis of population structure. Nucleic Acids Res. 2019;47(11):5539–49.
19. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792–7.
20. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics. 2014;30(22):3276–8.
21. Robinson O, Dylus D, Dessimoz C. Phylo.io: interactive viewing and comparison of large phylogenetic trees on the web. Mol Biol Evol. 2016;33(8):2163–6.
22. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068–9.
23. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res. 2015;43(3):e15.
24. Suchland RJ, Sandoz KM, Jeffrey BM, Stamm WE, Rockey DD. Horizontal transfer of tetracycline resistance among *Chlamydia spp.* in vitro. Antimicrob Agents Chemother. 2009;53(11):4604–11.
25. Suchland RJ, Carrell SJ, Wang Y, Hybiske K, Kim DB, Dimond ZE et al. Chromosomal recombination targets in *Chlamydia* interspecies lateral gene transfer. DiRita VJ, editor. Journal of Bacteriology [Internet]. 2019 Dec 9 [cited 2020 Feb 29];201(23). http://jb.asm.org/lookup/doi/https://doi.org/10.1128/JB.00365-19.
26. DeMars R, Weinfurter J. Interstrain gene transfer in *Chlamydia trachomatis* in vitro: mechanism and significance. J Bacteriol. 2008;190(5):1605–14.
27. Jeffrey BM, Suchland RJ, Eriksen SG, Sandoz KM, Rockey DD. Genomic and phenotypic characterization of in vitro-generated *Chlamydia trachomatis* recombinants. BMC Microbiol. 2013;13(1):142.

Publisher's Note