RESEARCH ARTICLE

Chlamydia trachomatis: when the virulence-associated genome backbone imports a prevalence-associated major antigen signature

Vítor Borges[1], Dora Cordeiro[2], Ana Isabel Salas[2], Zohra Lodhia[2], Cristina Correia[2], Joana Isidro[1], Cândida Fernandes[3], Ana Maria Rodrigues[3], Jacinta Azevedo[4], João Alves[4], João Roxo[5], Miguel Rocha[5], Rita Côrte-Real[3], Luís Vieira[6], Maria José Borrego[2],* and João Paulo Gomes[1],*

Abstract

Chlamydia trachomatis is the most prevalent sexually transmitted bacterium worldwide and the causative agent of trachoma. Its strains are classified according to their *ompA* genotypes, which are strongly linked to differential tissue tropism and disease outcomes [ocular disease, urogenital disease and lymphogranuloma venereum (LGV)]. While the genome-based species phylogenetic tree presents four main clades correlating with tropism/prevalence, namely ocular, LGV, urogenital T1 (more prevalent genotypes) and urogenital T2 (less prevalent genotypes), inter-clade exchange of *ompA* is considered a rare phenomenon probably mediating marked tropism alterations. An LGV epidemic, associated with the clonal expansion of the L2b genotype, has emerged in the last few decades, raising concerns particularly due to its atypical clinical presentation (ulcerative proctitis) and circulation among men who have sex with men (MSM). Here, we report an LGV outbreak, mostly affecting human immunodeficiency virus-positive MSM engaging in high-risk sexual practices, caused by an L2b strain with a rather unique non-LGV *ompA* signature that precluded the laboratory notification of this outbreak as LGV. *C. trachomatis* whole-genome capture and sequencing directly from clinical samples was applied to deeply characterize the genomic backbone of this novel LGV outbreak-causing clone. It revealed a chimeric genome structure due to the genetic transfer of *ompA* and four neighbouring genes from a serovar D/Da strain, likely possessing the genomic backbone associated with the more prevalent urogenital genotypes (T1 clade), to an LGV (L2b) strain. The hybrid L2b/D-Da strain presents the adhesin and immunodominant antigen MOMP (major outer membrane protein) (encoded by *ompA*) with an epitope repertoire typical of non-invasive genital strains, while keeping the genome-dispersed virulence fingerprint of a classical LGV strain. As previously reported for inter-clade *ompA* exchange among non-LGV clades, this novel *C. trachomatis* genomic mosaic involving a contemporary epidemiologically and clinically relevant LGV strain may have implications on its transmission, tissue tropism and pathogenic capabilities. The emergence of variants with epidemic and pathogenic potential highlights the need for more focused surveillance strategies to capture *C. trachomatis* evolution in action.

DATA SUMMARY

Chlamydia trachomatis-specific reads and assembled genome generated in this study were deposited in the European Nucleotide Archive (ENA) (BioProject accession no. PRJEB32243; www.ebi.ac.uk/ena/data/view/PRJEB32243). ENA accession numbers are included in Table S1 (available

with the online version of this article). The L2-L2b/D-Da hybrid *ompA* sequence was also deposited in the National Center for Biotechnology Information under the accession number MN094864.

INTRODUCTION

Sexually transmitted infections (STIs) are a global major public-health concern, resulting in significant morbidity and health-care costs. The World Health Organization (WHO) predicts that more than 1 million curable STIs occur each day, with an estimate of 127 million new cases of *Chlamydia trachomatis* infections in 2016 [1]. Lymphogranuloma venereum (LGV), one of the *C. trachomatis*-associated diseases of high concern, emerged in European countries since 2003, contrasting with the nearly 'inexistent' LGV cases detected until the new millennium [2–7]. A real estimate of LGV infection rates in Europe cannot be robustly performed. In fact, many LGV surveillance systems do not generate data that are considered representative of the national population (although reporting of LGV infection is compulsory in most countries with comprehensive surveillance systems), and there are significant differences in the capacity to identify the LGV-causing *C. trachomatis* lineages [8, 9]. This context represents an enhanced challenge for both health-care providers and public-health authorities, since the current recommended treatment for LGV differs from the one for non-LGV *C. trachomatis* infections [8, 10]. In contrast to the classical LGV presentation (i.e. inguinal syndrome with painful lymphadenopathy), current LGV cases are essentially manifested by proctitis characterized by severe symptoms, such as anorectal pain, rectal tenesmus, haemopurulent discharge and bleeding. This 'anorectal syndrome' mostly afflicts men who have sex with men (MSM), normally co-infected with human immunodeficiency virus (HIV) and other STIs [3, 4, 8]. The proctitis-associated LGV epidemic is being primarily caused by strains from the LGV genotype L2b [11], although some countries have been reporting increasing cases associated with the L2 genotype [12–14]. Genotyping of *C. trachomatis* relies on the polymorphism of the *ompA* gene, which encodes the major outer membrane protein (MOMP), its main antigen. The *ompA*-based phylogeny is incongruent with the genome-based species tree, but *ompA* genotype classification strongly correlates with tissue tropism and disease outcome, as genotypes A–C are strongly associated with the trachoma, genotypes D–K with urogenital infection and genotypes L1–L3 with LGV. This genotype grouping is generally reflected in the genome-based species tree, presenting a well-established topology with four main clades: ocular, LGV, urogenital T1 (enrolling clinically prevalent genotypes – mostly E and F) and urogenital T2 (involving less-prevalent genotypes) [15, 16]. Recombination is a main driver of contemporary *C. trachomatis* evolution and diversification, mainly leading to extensive intra-clade homologous exchange of genomic regions (including *ompA*), since genetic exchange between strains with the same tissue tropism is more likely to occur [15–18]. Inter-clade recombination is less marked. A relevant example enrolling the LGV clade concerns the report

Impact Statement

It is known that subtle evolutionary changes of pathogens may lead to rapid modifications in diagnosis and surveillance strategies. Here, we report the emergence of an outbreak-causing *Chlamydia trachomatis* strain with a novel recombinant genome make-up including a chimeric *ompA* (the classical typing gene and the main antigen/adhesin-encoding gene) genotype. The emergence of such a recombinant variant may have impact mainly at two levels. First, the hybrid L2b/D-Da *ompA* genotype precluded the compulsory national laboratory notification of the outbreak cases as lymphogranuloma venereum (LGV) cases. This challenges the current legal criteria for LGV notification (established by the European Commission and adopted by Portugal), and justifies the usage of broader multi-loci typing techniques in *C. trachomatis* outbreak detection, monitoring and control. Secondly, this genomic variant underlies a singular *C. trachomatis* diversification step, which is marked by the recombination-driven import of a non-LGV *ompA* (and four neighbouring genes) by an LGV strain with high clinical and public-health relevance (i.e. the worldwide disseminated proctitis-associated *C. trachomatis* L2b clone). Considering the strong correlation between *ompA* genotypes (and the mutational signature of the neighbouring genes) and *C. trachomatis* tropism/prevalence, it can be hypothesized that this variant may harbour modified transmission, tissue tropism and pathogenic capabilities.

of a severe clinical case caused by an L2 strain that imported several genome-dispersed regions from a D urogenital strain, although *ompA* exchange was not observed and no further cases were reported [19]. In fact, the exchange of *ompA* between strains from the three disease groups (i.e. ocular, genital and LGV) is considered a rare phenomenon among contemporary clinical strains [16] and is believed to mediate marked tropism alterations [16, 20].

In the present study, we report an ongoing *C. trachomatis* outbreak caused by a recombinant strain with a rather unique genome make-up, presenting the adhesin and immunodominant antigen MOMP with an epitope repertoire typical of non-invasive genital strains, while keeping the genome-dispersed virulence fingerprint of a classical LGV (L2b) strain. This constitutes what is believed to be the first report of a contemporary inter-clade import of *ompA* to an LGV strain with high clinical and public-health relevance.

METHODS

C. trachomatis molecular diagnosis and surveillance

The National Reference Laboratory (NRL) for Curable STIs of Portugal, hosted at the National Institute of Health (NIH)

Doutor Ricardo Jorge (Lisbon, Portugal), acts as a routine *C. trachomatis* diagnostic laboratory for NIH clients. These are members of the public attending the NIH, and specific risk populations (MSM, migrants, transsexuals and commercial sex workers) attending STIs clinics. The NRL is also responsible for nationwide laboratory surveillance of *C. trachomatis*. In this context, the NRL routinely performs the molecular characterization of all *C. trachomatis* strains, either identified in the NRL itself or by any Portuguese laboratory.

C. trachomatis molecular diagnosis was performed by using a dual-target (chromosome *plus* plasmid) commercial nucleic acid amplification test (NAAT) (Cobas 4800 CT/NG Roche Diagnostics), following the manufacturer's instructions. Anorectal samples tested as *C. trachomatis* positive by the first-line NAAT test were subsequently subjected to a rapid LGV discriminatory NAAT (CLART STIs from Genomica and/or Allplex genital ulcer assay from Seegene, both targeting the *pmpH* gene), following European Guidelines on the Management of LGV and despite the lack of thorough validation of these tests [8]. Also, DNA samples from all *C. trachomatis*-positive specimens (identified either in the NRL or in nationwide laboratories) are routinely subjected to *ompA* sequencing-based genotyping, since LGV laboratory notification is compulsory in Portugal and requires molecular confirmation of *ompA* genotypes L1–L3 [21]. Briefly, PCR and nested PCR were performed using primers NLO and NRO, and primers PCTM3 and SERO2A [22], respectively, as previously described [23], followed by partial nucleotide sequencing of the ~1010 bp PCR product using primer *ompA*-1 (5′-TTA TGA TCG ACG GAA TTC T-3′), BigDye terminator v1.1 and capillary sequencing (3130XL Genetic Analyzer; Applied Biosystems). *ompA* genotypes are currently determined by BLASTN-based comparison (using the ABRIcate tool) [24] with a custom database enrolling reference and variant sequences of all *ompA* genotypes [25]. MEGA software (version 7) [26] is additionally applied for fine-tuned sequence inspection and curation, alignment and phylogenetic reconstructions.

Targeted *C. trachomatis* whole-genome capture and sequencing directly from clinical samples

In order to potentiate the success of the culture-independent whole-genome sequencing (WGS) of *C. trachomatis*, diagnostic DNA samples positive for the novel hybrid L2b/D-Da *ompA* genotype were preliminarily subjected to dsDNA quantification (using Qubit; ThermoFisher Scientific) and subsequent absolute real-time quantification of both the number of *C. trachomatis* (targeting the single-copy *ompA*) and human genome copies (targeting β-actin), as previously described [27]. Samples selected (*n*=12) for targeted WGS had an amount of DNA above the minimum required for library preparation (>10 ng in 7 µl) and a number of *C. trachomatis* copies in the SureSelect XT HS (Agilent Technologies) input volume (7 µl) above 4×10³. Since the success of culture-independent targeted WGS does not seem to depend on the degree of human DNA content [28], this measure did not constitute a rigid criterion (although it was taken into account to discriminate samples with low *C. trachomatis* load) (Table S1).

C. trachomatis whole-genome capture and sequencing directly from selected DNA samples was performed using Agilent Technologies' SureSelect XT HS target enrichment system for Illumina paired-end multiplexed sequencing library protocol (G9702-90000, version C0, September 2018; Agilent Technologies) upon enzymatic fragmentation with SureSelect XT HS and XT low input enzymatic fragmentation kit (G9702-90050, Revision A0, September 2018; Agilent Technologies), according to the manufacturer's instructions. For this, RNA oligonucleotide baits (120 bp in size; a total of 33619 baits) were bioinformatically designed by our group to span the *C. trachomatis* chromosome and plasmid, and were further ordered for production in Agilent Technologies. Our custom bait design accounted for the main genetic variability among the four clades of the *C. trachomatis* species tree [15, 16], where baits with homology to the human genome (after discontiguous MegaBlast searches against the Human Genomic +Transcript database) were excluded. After library quantification and normalization (using the Fragment Analyzer with the HS NGS fragment kit; Advanced Analytical Technologies), *C. trachomatis*-enriched libraries were subjected to cluster generation and paired-end sequencing (2×150 bp) using MiSeq equipment (Illumina), according to the manufacturer's instructions.

Bioinformatics

In order to confirm the outbreak and characterize the genomic backbone and mosaicism of the hybrid L2b/D-Da *C. trachomatis* strain, the following bioinformatics activities were conducted: (i) reads' quality analysis and cleaning/improvement, *de novo* genome assembly and post-assembly optimization using the integrative pipeline INNUca version 4.0.1 [29]; (ii) reference-based mapping and SNP/indel analysis against representative genome sequences of both the worldwide disseminated proctitis-associated *C. trachomatis* L2b strain [L2b/UCH-1/proctitis; National Center for Biotechnology Information (NCBI) accession numbers AM884177.2/NC_010280.2 for the chromosome and AM886279.1 for the plasmid] and the detected L2b/D-Da *C. trachomatis* strain [strain Ct_L2b/D_PT05; European Nucleotide Archive (ENA) accession number CAAKND010000000] using Snippy version 4.1.0 [30]; (iii) whole-genome alignment and inspection using Mauve software version 2.3.1 [31]; (iv) coreSNP-based alignment and recombination inspection/visualization using Parsnp and Gingr tools available at the Harvest suite, respectively [32]; (v) integration of the hybrid L2b/D *C. trachomatis* strain in the *ompA*-based and genome-based species trees [15] by constructing approximately-maximum-likelihood phylogenetic trees using the double-precision mode of Fast-Tree2 under the general time-reversible (GTR) model (1000 bootstraps) [33]; and (vi) locus-based sequence alignment and manipulation using MEGA software (version 7) [26].

RESULTS

Detection of a novel hybrid L2b/D-Da *ompA* genotype causing a proctitis-associated LGV outbreak

On behalf of the national laboratory surveillance of *C. trachomatis*, the Portuguese NRL for curable STIs performs *ompA* sequencing-based genotyping, not only to monitor the genetic diversity of circulating strains, but also to comply with the legislation requirements for LGV laboratory notification (i.e. detection of L1–L3 genotypes), which has been compulsory in Portugal since 2017. Since that year, the NRL started detecting clinical specimens tested as LGV positive by a rapid discriminatory commercial NAAT (CLART STIs) that were not confirmed by classic *ompA* genotyping later on. As such, we initially interpreted the incongruent results as a failure of the commercial test. However, given the continuous emergence of these discrepant cases, we retrospectively performed an alternative commercial test (Allplex genital ulcer assay), which confirmed the prior LGV-positive results. We proceeded with the inspection of *ompA* and observed that all cases had an identical *ompA* sequence phylogenetically placed among D/Da strains. Still, a fine scrutiny of the sequence revealed an L2-L2b/D-Da hybrid profile, i.e. while the first 302 bp match to *ompA* L2/L2b reference sequences (L2b/UCH-1/proctitis; GenBank accession number AM884177.1; L2/434/BU, GenBank accession number AM884176.1), the region spanning the 366–1182 bp revealed a D/Da profile matching the reference strain D/UW-3 and a Da variant (CS-908/07) circulating in Portugal [34] (GenBank accession numbers NC_000117.1 and FJ943521, respectively). The conserved region between the 5′–3′ gene position 303–365 was estimated as the

recombination crossover region. The L2-L2b/D-Da hybrid *ompA* sequence was deposited at the NCBI under the accession number MN094864.

Twenty-five cases have been detected (the first case in April 2017). The novel hybrid LGV genotype represents 12.5% (10/80) and 16.5% (15/91) of all LGV cases (i.e. classical LGV genotypes plus the hybrid L2-L2b/D-Da hybrid genotype) detected in the Portuguese NIH during 2017 and 2018, respectively. These represent about 20 % of all *C. trachomatis*-positive specimens already subjected to *ompA* genotyping in the Portuguese NRL in both 2017 (80/405) and 2018 (91/484). Of note, the hybrid L2-L2b/D-Da genotype constitutes about half (in 2017) and one third (in 2018) of the number of cases of the classical L2b genotype associated with worldwide proctitis-related *C. trachomatis* epidemics occurring since 2003. All 25 specimens (most anorectal swabs) were collected from MSM presenting similar symptoms and clinical features (e.g. rectal syndrome, rectal pain, rectal tenesmus, anal discharge, rectal bleeding) (Fig. 1). Sexually transmitted bacterial co-infections were diagnosed for most patients, especially *Neisseria gonorrhoeae* (18 out the 25 patients). HIV status was available for 17 patients, 16 of which being positive. Multiple partners (>10 in last 6 months) and engaging in chemsex, as well as involvement with international sexual networks, were also reported (Fig. 1). Although all the health centres where cases were detected do not have harmonized behavioural-data collection procedures, hampering the identification of common sex venues or social events, known risk factors for rectal LGV, such as practices of fisting, douching and group sex in the last 12 months were reported for many MSM.


Fig. 1. Confirmed cases of the outbreak-causing hybrid L2b/D-Da *C. trachomatis* strain. Schematic summary of the clinical data, co-infections, risk behaviours and laboratory typing results.

The real number of cases cannot be predicted due to lack of geographical coverage of the national surveillance program. Furthermore, there is a delay associated with the genotyping procedure (it is performed in batches and not in a real-time fashion) and the genotyping technique frequently fails (in about 30% of the assays). The last case was detected on 10th November 2018.

Genome-based outbreak confirmation

In order to confirm the outbreak and characterize the genomic backbone of the *C. trachomatis* causative strain, RNA baits-based targeted genome capture and sequencing was performed directly from diagnostic DNA samples (12 samples were selected upon bacterial load quantification through quantitative PCR). Five complete or near-complete genomes were successfully sequenced (a representative assembly was deposited in ENA/NCBI under the accession number CAAKND010000000/CAAKND01) (Table S1). The phylogenetic integration of the hybrid *C. trachomatis* strain in the frame of well-established genome-based species trees [15] places it in the LGV clade, namely within the monophyletic branch reflecting the clonal expansion of the globally

disseminated proctitis-associated *C. trachomatis* L2b lineage (Fig. 2a). This contrasts with its position among serovar D/Da strains in *ompA* phylogeny (Fig. 2b), and strongly supports a scenario of recombination-driven inter-clade *ompA* exchange (detailed in the next section of Results), where the epidemic proctitis-associated L2b lineage was the most likely recipient strain. SNP/indel analysis of the hybrid L2b/D-Da complete (or near-complete) genomes against the classical L2b representative genome (L2b/UCH-1/proctitis strain) confirmed the high genetic relatedness with the parental-like strain and the clonality of the hybrid strains. In fact, apart from a single high SNP density peak observed in the recombinant region (enrolling *ompA*), only 35 variant sites (involving SNPs and small indels spanned throughout the chromosome) were detected among the five hybrid L2b/D-Da strains when compared with L2b/UCH-1/proctitis (Fig. 3a; Table S2). No differences were found in the plasmid. The L2b genomic backbone and recombination event could be also verified for the seven hybrid strains with partial genome data by the detection of other genetic fingerprints beyond the hybrid *ompA* profile detected by classical genotyping. These included the detection of SNP/
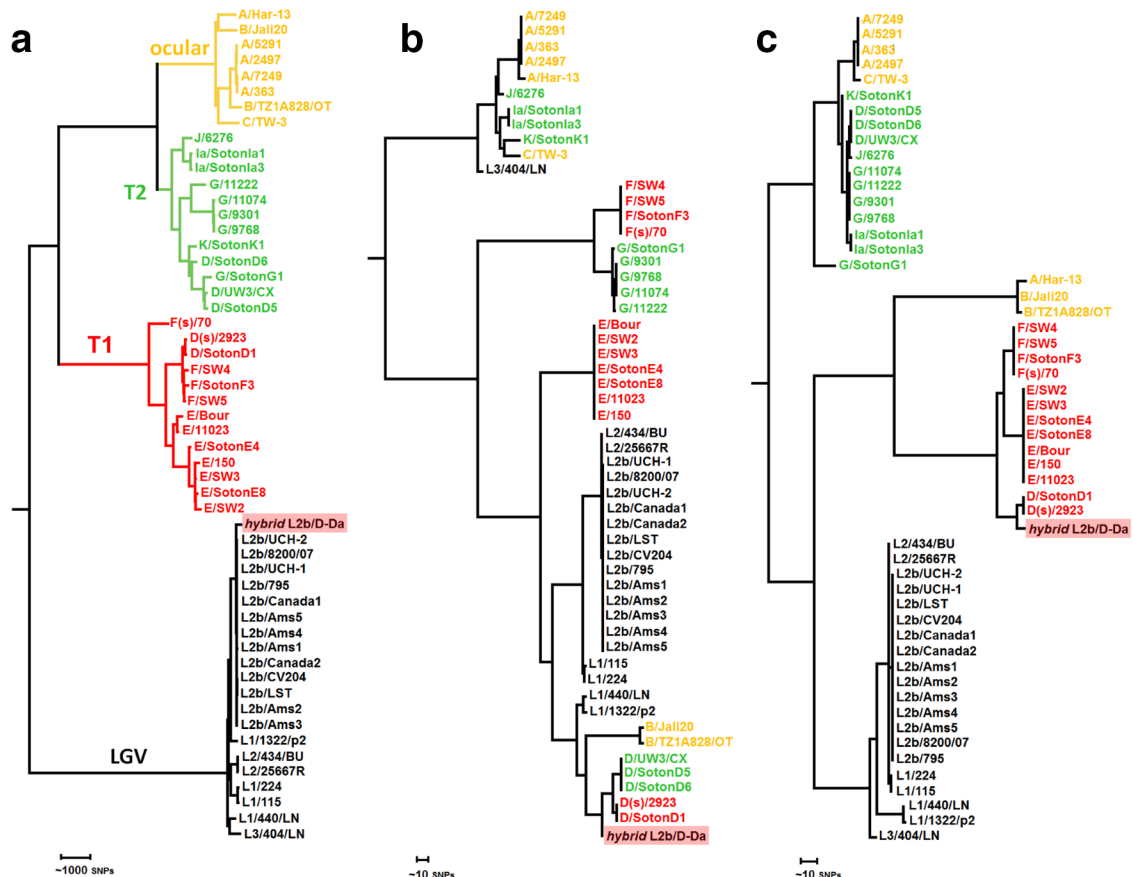
Fig. 2. Phylogenetic integration of the hybrid L2b/D-Da outbreak-causing *C. trachomatis* strain in the genome-based species tree (a), *ompA* tree (b) and concatenated CT677–CT680-based tree (c). Represented strains reflect the species phylogenetic diversity [15] and are coloured by the genome-based phylogenetic clade (ocular, yellow; prevalent genital – T1, red; non-prevalent genital – T2, green; and LGV, black), as marked in (a).
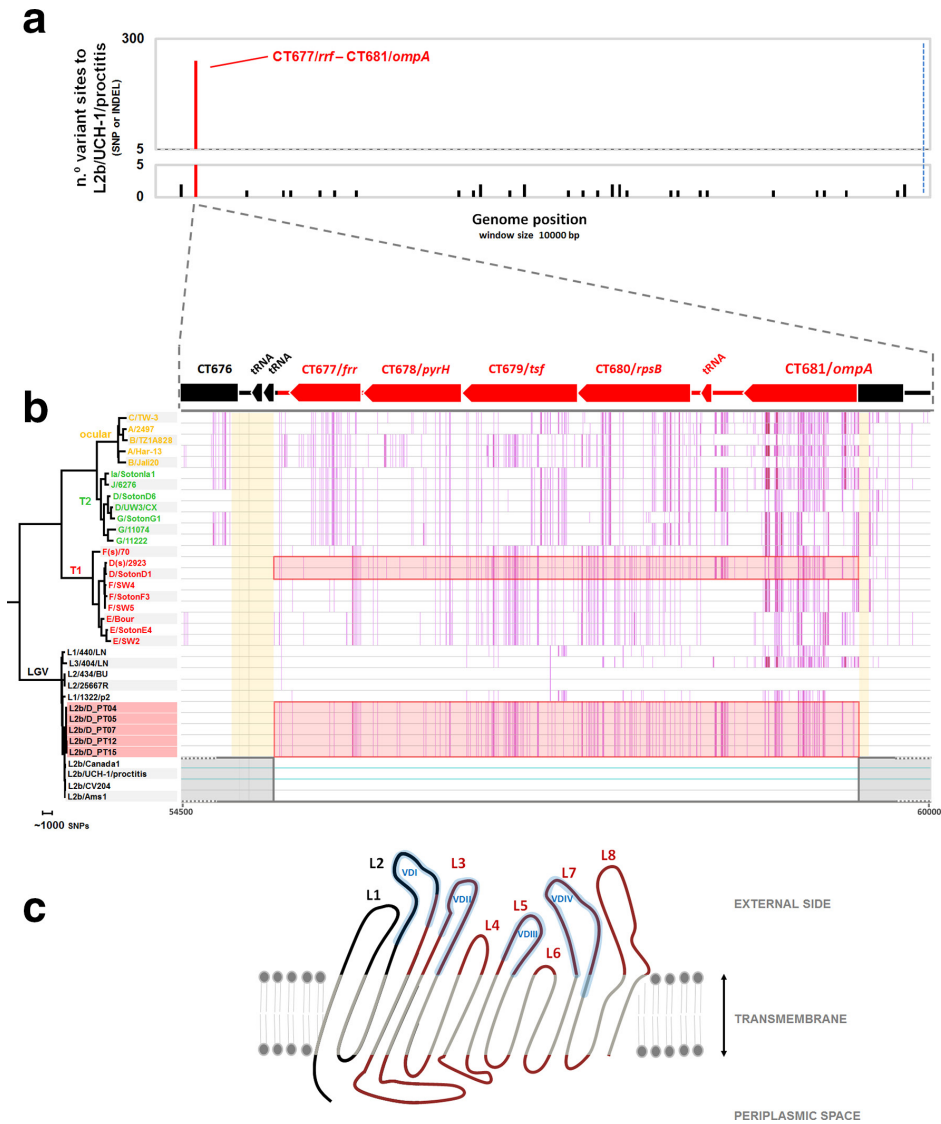
Fig. 3. Genome backbone and mosaic structure of the outbreak-associated *C. trachomatis* L2b/D-Da recombinant strain. (a) Genome backbone and recombination analysis of the L2b/D-Da *C. trachomatis* strain. The graph shows the number of genetic differences (SNPs and indels) detected among five complete or near-complete genomes of outbreak-associated L2b/D-Da clones when comparing with the genome (chromosome plus plasmid) sequence of the L2b/UCH-1/proctitis strain, which is a representative strain of the worldwide disseminated proctitis-associated outbreak of *C. trachomatis* L2b (a probable parental-like strain) occurring since 2003 (GenBank accession numbers: AM884177.2/NC_010280.2 for chromosome and AM886279.1 for the plasmid). Polymorphism is displayed as variant sites (SNPs or indels) in a window size and a step size of 10000 bp each, where variants introduced by the main recombination event are highlighted in red. The 100% conserved plasmid is present at the right side of the blue dashed line. Polymorphisms outside the recombinant fragment are detailed in Table S2. (b) Detail of the recombination event. The L2b/D-Da *C. trachomatis* strain is a recombinant strain resulting from the genetic import of the a ~4.2 kbp genomic fragment from a non-LGV D/Da strain (most likely belonging to T1 clade) by a typical proctitis-associated L2b strain. The recombination fragment (highlighted in red by both boxes and coloured genes) enrols ~75% of CT681/*ompA* and four entire genes (CT677/*frr*, CT678/*pyrH*, CT679/*tsf*, CT680/*rpsB*) (positions 55221–59461 in the L2b/UCH-1 chromosome). Estimated recombination crossovers most likely involved the two following regions (shaded in orange): (i) a region with two contiguous tRNAs located between CT676 and CT677/*rrf*; and (ii) a conserved region (among serogroup B serovars: B/Ba, D/Da, E, L1 and L2) within *ompA* (5'- 3' gene position 303-365). For scheme simplification, the core-genome-based phylogenetic reconstruction enrols sequences from *C. trachomatis* strains representing the main clades of the species tree [15] and from five outbreak-associated L2b/D-Da clones. Vertical pink lines reflect SNPs against the chromosome sequence of the parental-like L2b/UCH-1/proctitis strain (region spanning positions 54500–60000 bp; accession numbers AM884177.2/NC_010280.2) (highlighted by horizontal light blue lines). Gene orientation and annotation refers to the firstly released C. *trachomatis* genome, D/UW-3/CX strain (GenBank accession number NC_000117). (c) Schematic representation of the topology sketch of the recombinant MOMP, where the L2/L2b-like and D/Da-like non-transmembrane parts of the protein are coloured in black and red, respectively. MOMP is oriented from N-terminal to C-terminal (left to right). L1–L8 indicates protein loops (L) facing the outside, and include the four VDs (VDI–VDIV) of MOMP (shaded in blue). Amino acid sequence is detailed in Fig. S2. MOMP structure was drawn based on data from published work [25, 62].

indel markers of the hybrid strain, the detection of the L2b plasmid and/or identification of sequencing reads 100% matching to the non-LGV recombinant fragment (Tables S1 and S2). Of note, similarly to the scenario observed during the clonal expansion of the epidemic L2b lineage [35], the SNP/indel-based (micro)evolutionary segregation and discreet diversification of the hybrid strain is marked by non-synonymous mutations targeting several genes potentially linked to pathoadaptation, such as genes encoding inclusion membrane proteins (e.g. CT223 and CT383), type III secretion system effectors (e.g. CT456/Tarp) and surface proteins (e.g. CT242/*ompH*, CT270/*pbp3*) (Table S2). Curiously, some clones harbour a frameshift deletion in CT157 (CTL0413 in the reference strain L2/434/Bu), which encodes a phospholipase D (PLD) protein. This represents the inactivation of an additional PLD protein, besides the other PLD pseudogenes that seem to characterize the LGV strains [11]. It is also noteworthy that several SNP/indels distinguishing the hybrid strain from the L2b/UCH-1/proctitis are homoplasic (Table S2) as they occur in different branches of the *C. trachomatis* species tree.

Characterization of the genome mosaicism of the outbreak-related L2b/D-Da recombinant

Detailed analysis of the detected genetic mosaicism supports that the outbreak-causing L2b/D-Da hybrid *C. trachomatis* strain resulted from the genetic transfer of a ~4.2 kbp genomic fragment (transferring 240 SNPs, 76 of which were within *ompA*) from a D/Da strain to a typical proctitis-associated L2b strain (Figs 3a,b and S1). The non-LGV parental strain was most likely a D/Da strain with E/F-like genomic backbone, i.e. belonging to clade T1 of the species tree (which involves prevalent genotypes) (Figs 2 and 3b). In fact, although NCBI BLAST analyses of the imported fragment found no single genome with 100% similarity, the 'top' hits (query cover=100%; identity >99.75%) are serovar D/Da strains reported to belong to clade T1: D/13–96 [16, 36], D/SotonD1 [15], D/SQ29 and D/SQ32 [37]. The T1 donor backbone is also strongly supported by the existence of multiple SNVs within the transferred region perfectly segregating T1 strains and also SNVs that are exclusive of serovar D/Da strains with a E/F-like genome profile (Figs 2c and 3b). The recombinant fragment (position 55221–59461 in the L2b/UCH-1 chromosome) enrols ~75% of CT681/*ompA*, which encodes the major *C. trachomatis* antigen (MOMP), and four entire genes (CT677/*rrf*, CT678/*pyrH*, CT679/*tsf*, CT680/*rpsB*) located downstream from *ompA*. Estimated recombination crossovers most likely involved the two following regions: (i) a region with two contiguous tRNAs located downstream from CT677/*rrf*; and (ii) a conserved region (among serogroup B serovars: B/Ba, D/Da, E, L1 and L2) within *ompA* (5′–3′ gene position 303–365). Of note, the latter breakpoint region has been also involved in the partial import of ocular *ompA* by strains with urogenital backbone associated with ocular infections [16, 20]. The recombinant L2b/D-Da MOMP displays a total of 19 aa replacements and a 1 aa insertion when compared with the parental-like L2b

MOMP, with 18 of the amino acid variants falling within surface-exposed epitope-enriched highly variable domains (VDs) (Fig. S2). In fact, the recombination event yielded a new combination of epitopes in the *C. trachomatis* immunodominant antigen, MOMP, mixing serovar L2/L2b epitopes (VDI region) and serovar D/Da-like epitopes (from VDII to VDIV) (Figs 3c and S2).

DISCUSSION

WGS-based typing is still not applicable for *C. trachomatis* routine surveillance, despite the need for timely characterization of circulating types and variants to disclose transmission chains, guide therapies and identify emerging public-health harm. However, the recent history of *C. trachomatis* infections illustrates how the dynamic pathogen evolution (rather than the availability of technological advances) prompts the adoption of important modifications in diagnosis, surveillance and therapeutic strategies. The first relevant example of 'evolution in action' with impact on molecular diagnosis/surveillance was the emergence of a *C. trachomatis* variant, detected in 2006 in Sweden, with a deletion in the plasmid region targeted by widespread commercial NAATs [38]. Retrospective studies revealed that this variant had been in circulation at least since 2001 and displayed a rapid spread (reaching in a few years 64% of all diagnosis in Sweden) due to diagnostic failure and a consequent lack of treatment [39, 40]. This situation led to a rapid switch to different molecular diagnostics tests, while leading to the replacement of the single target tests in question by tests that target both the plasmid and the chromosome [41]. More recently, a *C. trachomatis* variant detected in Finland was found to yield false-negative results in over 160 cases, potentially due to a nucleotide substitution in the sequence targeted by the diagnostic test [42].

The evolution of *C. trachomatis* is expectedly potentiated by increasing human population density and mobility, as well as by changes in sexual networks, which enhance the spread and admixture of strains. Accordingly, the global emergence of the successful MSM-associated proctitis-related L2b clone implicated the modification of international guidelines for *C. trachomatis* molecular identification/typing towards the adoption of tests targeting LGV-specific loci and differential treatment of LGV cases [8, 41]. The unique character of the outbreak-causing L2b/D-Da recombinant strain reported in the present study may lead to an 'erroneous' classification and potentially incorrect treatment if only the conventional *ompA* typing is performed, because *ompA* phylogeny places this hybrid strain among serovar D/Da strains (Fig. 2). Although commercial NAATs for rapid LGV/non-LGV differentiation were proficient in detecting these recombinant strains as LGV, they are not widely available and still lack validation [8, 43]. A laboratory sub-notification of LGV cases (compulsory in multiple countries [9]) may also occur, as *ompA* genotyping is still required to comply with the Commission Implementing Decision (EU) 2018/945 of 22 June 2018 [44], which demands L1, L2 or L3 identification as criteria for a confirmed LGV case. For example, in Portugal, where LGV laboratory

notification is compulsory and LGV case definition [21] is in compliance with the European Union legislation, these outbreak-associated cases cannot be formally notified as LGV. This has already triggered a still ongoing revision of the laboratory LGV notification criteria. Altogether, this situation may facilitate the spread of the hybrid outbreak strain especially due to its circulation among MSM (most of them co-infected with HIV and other STI agents), who often report a sexual behaviour marked by engaging in chemsex and, in some cases, involvement in international sexual networks. Considering this methodological and epidemiological context, one can hypothesize that the outbreak has already crossed borders, although remained unnoticed. This whole scenario highlights the need to revise laboratory methods for notification, and supports the international recommendations to target LGV-specific genome loci for detecting, monitoring and controlling LGV outbreaks. Several MLST-like systems capable of LGV-strain identification have been described (e.g. [45–48]) but none of them has been broadly adopted by the scientific community. Curiously, *in silico* extraction of MLST profiles using the three schemes currently available at PubMLST [https://pubmlst.org/ (accessed 15 May 2019)] showed that the recombinant L2b/D-Da strain cannot be distinguished from the parental-like L2b/UCH-1/proctitis lineage by applying these MLST strategies [belonging to ST44 (*Chlamydiales* scheme [46]), ST1 (*C. trachomatis* scheme [47]) and ST58 (*C. trachomatis* Uppsala scheme [45])]. As such, the present outbreak further supports the application of multi-loci typing strategies in addition to *ompA* genotyping. From a clinical point of view, this specific epidemic reinforces the notion that MSM (especially those who are HIV-positive) should be routinely screened for oral, urethral and anal *C. trachomatis* infection. Voluntary partner notification and treatment may also be considered due to the possible sexual network (sex abroad and chemsex). From another perspective, second generation surveillance for LGV (e.g. NRL-centralized integration of both microbiological and demographic/behavioural data, safeguarding the appropriate ethical and social issues) would be advisable in order to track both the 'evolution' and 'context' of *C. trachomatis* (namely LGV) epidemics. For example, in this outbreak, a harmonized behavioural survey among the enrolled STI clinics could have yielded more complete reports about practices of fisting, douching and group sex in the last 12 months, which are known risk factors for rectal LGV [49]. The latter were actually reported by most patients attending one of the clinics where this data are systematically collected.

The inter-strain exchange of the whole or part of the *ompA* gene is a natural phenomenon in *C. trachomatis* [15–18, 20, 50–52]. To the best of our knowledge, in addition to the known ancient *ompA* diversification (leading to the *C. trachomatis ompA*-based phylogenetic grouping being incongruent with the genome-wide phylogenetic topology) [15], the present study provides the first report of an epidemiologically and clinically relevant hybrid strain presenting a classical LGV chromosome backbone, while harbouring most of the major antigen (MOMP) from a non-LGV strain. This

evolutionary step is not surprising in the epidemiological and demographical context it emerged. In fact, the MSM-associated worldwide epidemics of L2b (occurring during the last two decades) potentiate co-infection with highly prevalent urogenital strains (e.g. T1 clade), which are usually associated with asymptomatic infections, favouring multiple rounds of infections and co-infections. Besides the main recombination event, the hybrid strain also revealed discreet genome-dispersed homoplasies that further corroborate a scenario of recurrent exchange of genetic material with other non-LGV strains. This is a microevolutionary mechanism occurring in parallel with the fixation of *de novo* mutations mostly affecting host-interacting proteins with a potential role in pathoadaptation (Table S2). The observed genomic mosaicism results in not only a new combination of MOMP epitopes (mixing D/Da-like and L2/L2b-like epitopes within MOMP) (Figs 3 and S2), but also a novel antigenic composition at the proteome level. In fact, most MOMP epitopes (targeted by both T- and B-cells) belong to strains typically associated with non-invasive genital tropism, whereas the remaining antigen repertoire (e.g. polymorphic membrane proteins – Pmps; chlamydial protease-like activity factor – CPAF; inclusion membrane proteins – Incs; OmcB) [11, 53] is typical of LGV strains causing anorectal and/or inguinal syndromes. We can hypothesize that this novel MOMP chimeric structure influences the strain's ability to deal with the immunological response, as even modest alterations in the immunodominant MOMP may substantially change *C. trachomatis* antigenic presentation [54–56]. Beyond the potential impact on the interaction with the host immune system, it is conceivable that other bacterial functions are also affected, possibly influencing pathogenesis. In fact, the immunodominant MOMP, which constitutes about 60% of the membrane dry weight [57], is also a porin [58] and adhesin [59], which are well established host-interacting bacterial factors underlying differential tissue tropism. Noteworthy is the fact that the expected role of inter-clade *ompA* exchange in mediating alterations in phenotypic and transmission capabilities has been recently corroborated by the independent detection (in Australia and Europe) of ocular *ompA* genotypes introgression into typical urogenital strains (from clades T1 and T2) associated with both ocular and urogenital infections [16, 20]. Although the selective advantage of the recombinant strain most likely relies on the novel signature of the hybrid MOMP, the other exchanged genes neighbouring *ompA* (CT677/*rrf*, CT678/*pyrH*, CT679/*tsf*, CT680/*rpsB*), which play key functional roles (in ribosome recycling, pyrimidine metabolism, translation elongation and structural ribosome constitution, respectively), may also contribute to an enhanced fitness. This may be particularly relevant considering that: (i) their mutational signature in the hybrid L2b/D-Da strain matches the one observed for prevalent clade T1 strains (i.e. strains with E/F-like genomic backbone); and (ii) the CT677–CT680 region shows phylogenetic signals of tropism and prevalence (Fig. 3c) [60, 61]. For example, it has been speculated that increased fitness at or around *ompA* (preventing recombination) may underlie the expansion of genotype E, which is in general the most

successful *C. trachomatis* genotype among STIs as suggested by its high prevalence and global distribution [16, 34, 61]. In conclusion, considering that (i) the hybrid outbreak strain displays a novel antigenic and adhesin fingerprint, while retaining the classical epidemic LGV (L2b) genome-dispersed virulence signature; (ii) L2b strains are capable of dual tropism, i.e. infection of the rectal mucosa (leading to proctitis) and/or dissemination through mononuclear mono-cytes to inguinal nodes (leading to the bubonic disease LGV); and (iii) the presumably parental genital strain belongs to the T1 clade, which includes the ecologically most successful genital strains; one can hypothesize that the observed recombination-driven unique genomic backbone may have implications on the transmission, tissue tropism and pathogenic capabilities of the circulating hybrid outbreak-associated strain. However, a more-definitive information regarding the potentially enhanced phenotypic skills of the hybrid strain must await the detection of the hybrid L2b/D-Da strain from around the world, as well as *in vitro* and *in vivo* assays to assess host-evasion and growth characteristics.

This study reports an ongoing outbreak involving a unique L2b variant (with potential modified fitness) in the context of the parental L2b worldwide epidemics. The risk sexual behaviour (e.g. multiple partners, engaging in chemsex) and sexual network involvement enhances the likelihood of this outbreak to have already spread abroad, which might be soon confirmed if broader *C. trachomatis* genotyping strategies are followed. The non-eligibility of this outbreak as laboratory-confirmed LGV cases additionally highlights the need to revise the European Commission criteria, adopted by several countries, including Portugal, for LGV laboratory notification.

Data bibliography
1. Borges V *et al.* ENA (reads and genome assembly), BioProject acces-sion number PRJEB32243 (2019).

2. Borges V *et al.* NCBI (representative genome assembly), accession number CAAKND01 (2019).

3. Borges V *et al.* NCBI (*ompA*), accession number MN094864 (2019).

References
1. **World Health Organization**. *Report on Global Sexually Transmitted Infection Surveillance*. Geneva: World Health Organization; 2018.

2. **Nieuwenhuis RF, Ossewaarde JM, van der Meijden WI, Neumann HA**. Unusual presentation of early lymphogranuloma venereum in an HIV-1 infected patient: effective treatment with 1 G azithromycin. *Sex Transm Infect* 2003;79:453–455.

3. **Nieuwenhuis RF, Ossewaarde JM, Götz HM, Dees J, Thio HB** *et al.* Resurgence of lymphogranuloma venereum in Western Europe: an outbreak of *Chlamydia trachomatis* serovar L2 proctitis in the Netherlands among men who have sex with men. *Clin Infect Dis* 2004;39:996–1003.

4. **Savage EJ, van de Laar MJ, Gallay A, van der Sande M, Hamouda O** *et al.* Lymphogranuloma venereum in Europe, 2003-2008. *Euro Surveill* 2009;14:19428.

5. **Bébéar C, de Barbeyrac B**. Genital *Chlamydia trachomatis* infec-tions. *Clin Microbiol Infect* 2009;15:4–10.

6. **White JA**. Manifestations and management of lymphogranuloma venereum. *Curr Opin Infect Dis* 2009;22:57–66.

7. **de Vrieze NHN, de Vries HJC**. Lymphogranuloma venereum among men who have sex with men. An epidemiological and clinical review. *Expert Rev Anti Infect Ther* 2014;12:697–704.

8. **de Vries HJC, de Barbeyrac B, de Vrieze NHN, Viset JD, White JA** *et al*. 2019 European guideline on the management of lymphogranu-loma venereum. *J Eur Acad Dermatol Venereol* 2019;33:1821–1828.

9. **European Centre for Disease Prevention and Control**. *Lymphogran-uloma Venereum - Annual Epidemiological Report for 2016.* Stock-holm: ECDC; 2018.

10. **Leeyaphan C, Ong JJ, Chow EPF, Kong FYS, Hocking JS** *et al.* Systematic review and meta-analysis of doxycycline efficacy for rectal lymphogranuloma venereum in men who have sex with men. *Emerg Infect Dis* 2016;22:1778–1784.

11. **Thomson NR, Holden MTG, Carder C, Lennard N, Lockey SJ** *et al*. *Chlamydia trachomatis*: genome sequence analysis of lymphogran-uloma venereum isolates. *Genome Res* 2008;18:161–171.

12. **Rodríguez-Domínguez M, Puerta T, Menéndez B, González-Alba JM, Rodríguez C** *et al*. Clinical and epidemiological char-acterization of a lymphogranuloma venereum outbreak in Madrid, Spain: co-circulation of two variants. *Clin Microbiol Infect* 2014;20:219–225.

13. **Peuchant O, Touati A, Sperandio C, Hénin N, Laurier-Nadalié C** *et al*. Changing pattern of *Chlamydia trachomatis* strains in lymphogran-uloma venereum outbreak, France, 2010–2015. *Emerg Infect Dis* 2016;22:1945–1947.

14. **Isaksson J, Carlsson O, Airell Åsa, Strömdahl S, Bratt G** *et al*. Lymphogranuloma venereum rates increased and *Chlamydia trachomatis* genotypes changed among men who have sex with men in Sweden 2004-2016. *J Med Microbiol* 2017;66:1684–1687.

15. **Harris SR, Clarke IN, Seth-Smith HMB, Solomon AW, Cutcliffe LT** *et al*. Whole-genome analysis of diverse *Chlamydia trachomatis* strains identifies phylogenetic relationships masked by current clinical typing. *Nat Genet* 2012;44:413–419.

16. **Hadfield J, Harris SR, Seth-Smith HMB, Parmar S, Andersson P** *et al*. Comprehensive global genome dynamics of *Chlamydia trachomatis* show ancient diversification followed by contemporary mixing and recent lineage expansion. *Genome Res* 2017;27:1220–1229.

17. **Gomes JP, Bruno WJ, Nunes A, Santos N, Florindo C** *et al*. Evolution of *Chlamydia trachomatis* diversity occurs by widespread interstrain recombination involving hotspots. *Genome Res* 2007;17:50–60.

18. **Rodríguez-Domínguez M, González-Alba JM, Puerta T, Martínez-García L, Menéndez B** *et al*. Spread of a new *Chlamydia trachomatis* variant from men who have sex with men to the heterosexual population after replacement and recombination in *ompA* and *pmpH* genes. *Clin Microbiol Infect* 2017;23:761–766.

19. **Somboonna N, Wan R, Ojcius DM, Pettengill MA, Joseph SJ** *et al*. Hypervirulent *Chlamydia trachomatis* clinical strain is a recombi-nant between lymphogranuloma venereum (L(2)) and D lineages. *mBio* 2011;2:e00045-11.

20. Andersson P, Harris SR, Seth Smith HMB, Hadfield J, O'Neill C *et al*. *Chlamydia trachomatis* from Australian Aboriginal people with trachoma are polyphyletic composed of multiple distinctive lineages. *Nat Commun* 2016;7:10688.

21. Direção-Geral da Saúde. 2016. Despacho no. 15385-A/2016. *Diário da República*. https://dre.pt/application/conteudo/105574339

22. Lan J, Ossewaarde JM, Walboomers JM, Meijer CJ, van den Brule AJ. Improved PCR sensitivity for direct genotyping of *Chlamydia trachomatis* serovars by using a nested PCR. *J Clin Microbiol* 1994;32:528–530.

23. Gomes JP, Bruno WJ, Borrego MJ, Dean D. Recombination in the genome of *Chlamydia trachomatis* involving the polymorphic membrane protein C gene relative to *ompA* and evidence for horizontal gene transfer. *J Bacteriol* 2004;186:4295–4306.

24. Seemann T. ABRIcate. https://github.com/tseemann/abricate

25. Nunes A, Nogueira PJ, Borrego MJ, Gomes JP. Adaptive evolution of the *Chlamydia trachomatis* dominant antigen reveals distinct evolutionary scenarios for B- and T-cell epitopes: worldwide survey. *PLoS One* 2010;5:e13171.

26. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 2016;33:1870–1874.

27. Gomes JP, Borrego MJ, Atik B, Santo I, Azevedo J *et al*. Correlating *Chlamydia trachomatis* infectious load with urogenital ecological success and disease pathogenesis. *Microbes Infect* 2006;8:16–26.

28. Pinto M, Borges V, Antelo M, Pinheiro M, Nunes A *et al*. Genome-scale analysis of the non-cultivable *Treponema pallidum* reveals extensive within-patient genetic variation. *Nat Microbiol* 2016;2:16190.

29. LlarenaA-K, Ribeiro-Gonçalves BF, Nuno Silva D, Halkilahti J, Machado MP. *INNUENDO: a Cross-sectoral Platform for the Integration of Genomics in the Surveillance of Food-borne Pathogens. EFSA Supporting Publication 2018;2018:EN-1498.142*. Parma: EFSA; 2018.

30. SeemannT. Snippy. https://github.com/tseemann/snippy

31. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 2010;5:e11147.

32. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol* 2014;15:524.

33. Price MN, Dehal PS, Arkin AP. FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One* 2010;5:e9490.

34. Nunes A, Borrego MJ, Nunes B, Florindo C, Gomes JP. Evolutionary dynamics of *ompA*, the gene encoding the *Chlamydia trachomatis* key antigen. *J Bacteriol* 2009;191:7182–7192.

35. Borges V, Gomes JP. Deep comparative genomics among *Chlamydia trachomatis* lymphogranuloma venereum isolates highlights genes potentially involved in pathoadaptation. *Infect Genet Evol* 2015;32:74–88.

36. Putman TE, Suchland RJ, Ivanovitch JD, Rockey DD. Culture-independent sequence analysis of *Chlamydia trachomatis* in urogenital specimens identifies regions of recombination and in-patient sequence mutations. *Microbiology* 2013;159:2109–2117.

37. Suchland RJ, Dimond ZE, Putman TE, Rockey DD. Demonstration of persistent infections and genome stability by whole-genome sequencing of repeat-positive, same-serovar *Chlamydia trachomatis* collected from the female genital tract. *J Infect Dis* 2017;215:1657–1665.

38. Ripa T, Nilsson P. A variant of *Chlamydia trachomatis* with deletion in cryptic plasmid: implications for use of PCR diagnostic tests. *Euro Surveill* 2006;11:E061109.2.

39. Unemo M, Seth-Smith HMB, Cutcliffe LT, Skilton RJ, Barlow D *et al*. The Swedish new variant of *Chlamydia trachomatis*: genome sequence, morphology, cell tropism and phenotypic characterization. *Microbiology* 2010;156:1394–1404.

40. Unemo M, Clarke IN. The Swedish new variant of *Chlamydia trachomatis*. *Curr Opin Infect Dis* 2011;24:62–69.

41. Dahlberg J, Hadad R, Elfving K, Larsson I, Isaksson J *et al*. Ten years transmission of the new variant of *Chlamydia trachomatis* in Sweden: prevalence of infections and associated complications. *Sex Transm Infect* 2018;94:100–104.

42. Rantakokko-Jalava K, Hokynar K, Hieta N, Keskitalo A, Jokela P *et al*. *Chlamydia trachomatis* samples testing falsely negative in the Aptima Combo 2 test in Finland, 2019. *Euro Surveill* 2019;24:1900298.

43. Lanjouw E, Ouburg S, de Vries HJ, Stary A, Radcliffe K *et al*. 2015 European guideline on the management of *Chlamydia trachomatis* infections. *Int J STD AIDS* 2016;27:333–348.

44. European Commission. 2018. Commission implementing decision (EU) 2018/945 of 22 June 2018 on the communicable diseases and related special health issues to be covered by epidemiological surveillance as well as relevant case definitions. Brussels. European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018D0945&from=EN

45. Klint M, Fuxelius H-H, Goldkuhl RR, Skarin H, Rutemark C *et al*. High-resolution genotyping of *Chlamydia trachomatis* strains by multilocus sequence analysis. *J Clin Microbiol* 2007;45:1410–1414.

46. Pannekoek Y, Morelli G, Kusecek B, Morré SA, Ossewaarde JM *et al*. Multi locus sequence typing of *Chlamydiales*: clonal groupings within the obligate intracellular bacteria *Chlamydia trachomatis*. *BMC Microbiol* 2008;8:42.

47. Dean D, Bruno WJ, Wan R, Gomes JP, Devignot S *et al*. Predicting phenotype and emerging strains among *Chlamydia trachomatis* infections. *Emerg Infect Dis* 2009;15:1385–1394.

48. Bom RJM, Christerson L, Schim van der Loeff MF, Coutinho RA, Herrmann B *et al*. Evaluation of high-resolution typing methods for *Chlamydia trachomatis* in samples from heterosexual couples. *J Clin Microbiol* 2011;49:2844–2853.

49. Macdonald N, Sullivan AK, French P, White JA, Dean G *et al*. Risk factors for rectal lymphogranuloma venereum in gay men: results of a multicentre case-control study in the UK. *Sex Transm Infect* 2014;90:262–268.

50. Dean D, Schachter J, Dawson CR, Stephens RS. Comparison of the major outer membrane protein variant sequence regions of B/Ba isolates: a molecular epidemiologic approach to *Chlamydia trachomatis* infections. *J Infect Dis* 1992;166:383–392.

51. Brunham R, Yang C, Maclean I, Kimani J, Maitha G *et al*. *Chlamydia trachomatis* from individuals in a sexually transmitted disease core group exhibit frequent sequence variation in the major outer membrane protein (*omp1*) gene. *J Clin Invest* 1994;94:458–463.

52. Hayes LJ, Yearsley P, Treharne JD, Ballard RA, Fehler GH *et al*. Evidence for naturally occurring recombination in the gene encoding the major outer membrane protein of lymphogranuloma venereum isolates of *Chlamydia trachomatis*. *Infect Immun* 1994;62:5659–5663.

53. de la Maza LM, Zhong G, Brunham RC. Update on *Chlamydia trachomatis* vaccinology. *Clin Vaccine Immunol* 2017;24:e00543-16.

54. Suchland RJ, Eckert LO, Hawes SE, Stamm WE. Longitudinal assessment of infecting serovars of *Chlamydia trachomatis* in Seattle public health clinics: 1988–1996. *Sex Transm Dis* 2003;30:357–361.

55. Kari L, Whitmire WM, Carlson JH, Crane DD, Reveneau N *et al*. Pathogenic diversity among *Chlamydia trachomatis* ocular strains in nonhuman primates is affected by subtle genomic variations. *J Infect Dis* 2008;197:449–456.

56. Kari L, Whitmire WM, Crane DD, Reveneau N, Carlson JH *et al*. *Chlamydia trachomatis* native major outer membrane protein induces partial protection in nonhuman primates: implication for a trachoma transmission-blocking vaccine. *J Immunol* 2009;182:8063–8070.

57. Caldwell HD, Kromhout J, Schachter J. Purification and partial characterization of the major outer membrane protein of *Chlamydia trachomatis*. *Infect Immun* 1981;31:1161–1176.

58. Bavoil P, Ohlin A, Schachter J. Role of disulfide bonding in outer membrane structure and permeability in *Chlamydia trachomatis*. *Infect Immun* 1984;44:479–485.

59. Su H, Watkins NG, Zhang YX, Caldwell HD. *Chlamydia trachomatis*-host cell interactions: role of the chlamydial major outer membrane protein as an adhesin. *Infect Immun* 1990;58:1017–1025.

60. Ferreira R, Antelo M, Nunes A, Borges V, Damião V *et al*. *In Silico* scrutiny of genes revealing phylogenetic congruence with clinical prevalence or tropism properties of *Chlamydia trachomatis* strains. *G3* 2015;5:9–19.

61. Brunelle BW, Sensabaugh GF. Nucleotide and phylogenetic analyses of the *Chlamydia trachomatis ompA* gene indicates it is a hotspot for mutation. *BMC Res Notes* 2012;5:53.

62. Wang Y, Berg EA, Feng X, Shen L, Smith T *et al*. Identification of surface-exposed components of MOMP of *Chlamydia trachomatis* serovar F. *Protein Sci* 2006;15:122–134.