

11th International Meeting on Microbial Epidemiological Markers (IMMEM XI)

Authors: Vítor Borges*, Alexandra Nunes*, Daniel A Sampaio, Luís Vieira, João P. Gomes.

*Co-first authors

Title

Phylogenomic characterization of the causative strain of one of the largest worldwide outbreaks of Legionnaires' disease occurred in Portugal in 2014

Keywords

Legionnaires' disease; genome; outbreak in Portugal

Background

Legionnaires' disease (LD) is a severe pneumonia with a case fatality rate of 8–12%. Transmission usually occurs by inhalation of aerosols or aspiration of water containing *Legionella* spp, where *L. pneumophila* serogroup 1 accounts for most of the occasional outbreaks. The world's second largest outbreak of LD occurred in 2014 in Vila Franca de Xira, Portugal, yielding more than 400 cases and 14 deaths. The *L. pneumophila* causative strain is from serogroup 1 and displays the novel sequence type (ST) ST1905. Here, we described how whole-genome sequencing (WGS) / Bioinformatics was used for real-time investigation of this LD outbreak as well as for integrating the genetic backbone of the causative strain in the frame of the species phylogeny and diversity.

Materials | Methods

To investigate the genetic relatedness of ST1905 isolates and to attribute the source of LD Portugal outbreak, 10 clinical isolates and one environmental isolate obtained from an industrial cooling tower were selected for WGS on a MiSeq Illumina instrument (depth of coverage >100-fold). Genome assembly was carried out using Velvet. Phylogenetic inferences using alignments of core-genome sequences from the 11 outbreak-associated *L. pneumophila* serogroup 1 isolates and/or multiple publicly available *L. pneumophila* genomes were performed through MEGA5 by using the neighbor-joining method with bootstrapping (1000 replicates).

Results

By using more than 99% of each draft genome sequence, we found no nucleotide differences between each clinical isolate and the environmental isolate, thus suggesting the suspected industrial cooling tower as the source of the LD outbreak. Unexpectedly, the outbreak causative strain phylogenetically clustered apart from the most worldwide studied outbreak-associated *L. pneumophila* serogroup 1 strains (Philadelphia-1, Paris, Lens, Corby, and 2300/99 Alcoy), being more closely related to *L. pneumophila* subsp. *fraseri* strains from serogroups other than serogroup 1 (Figure 1). We found that the novel outbreak strain harbors multiple genomic islands potentially related to virulence and/or adaptation, and carries an exclusive ~38kb genomic region that was only detected so far in one *L. oakridgensis* strain, a species that rarely causes LD. Of note, confirmation of the ST1905 allelic profile highlighted a bias associated with the *in silico* extraction of the allele sequence from WGS data, since the studied strain displayed

non-matching *mompS* copies. This issue would hamper a proper ST attribution if the allelic profile was exclusively determined *in silico*.

Conclusion

WGS enabled an in-depth investigation of the recent LD outbreak that occurred in Portugal, as it allowed attributing, with a greater level of certainty, the source of the epidemic to an industrial cooling tower. Given the unusual genetic features apparently displayed by this outbreak causative strain, it is of the utmost importance to dissect the transmission skills and virulence for this particular strain in a near future, as those traits may justify the magnitude of the outbreak.