Complete genome sequence of *Trueperella pyogenes* strain Arash114, isolated from the uterus of a water buffalo (*Bubalus bubalis*) in Iran

Iraj Ashrafi Tamai¹, Abdolmajid Mohammadzadeh¹*, Arash Ghalyanchi Langeroudi², Pezhman Mahmoodi¹, Zahra Ziafati Kafi², Babak Pakbin³ and Taghi Zahraei Salehi²

Abstract

**Objective:** *Trueperella pyogenes* has been considered a major causative agent of metritis, abortion, and death in a broad range of domestic and wild animals, including cattle, swine, sheep, goats, camels, buffalo, deer, antelopes, reptiles, and birds.

**Data description:** Here, we report the complete chromosome sequence of *Trueperella pyogenes* strain Arash114, isolated from the uterus of a water buffalo (*Bubalus bubalis*) died due to the infection caused by this pathogen. The genome assembly comprised 2,338,282 bp, with a 59.5% GC content. Annotation of the genome showed 46 tRNA genes, 6 rRNA, 1 CRISPR and 2059 coding sequences. Also, several genes coding for antimicrobial resistance such as *tetW* and virulence factor including *plo*, *nanH*, *nanP*, *cbp* and 4 fimbrial proteins were found. This study will advance our knowledge regarding the metabolism, virulence factors, antibiotic resistance and evolution of Arash114 strain and serve as an appropriate template for future researches.

**Keywords:** *Trueperella pyogenes*, Complete genome sequencing, Uterus infection, Water buffalo

**Objective**

*Trueperella pyogenes*, formerly known as *Actinomyces pyogenes* and *Arcanobacterium pyogenes*, has recently been reclassified based on distinctive 16S rRNA gene sequences. This bacterium is an irregular, nonmotile, non-spore-forming, aerobic, commensal, Gram-positive, short, rod-shaped bacterium which is normally isolated from the upper respiratory, urogenital, and gastrointestinal tracts [1–3]. *T. pyogenes* is proposed as a globally distributed secondary pathogen which may cause diseases such as acute and subclinical endometritis, cutaneous and visceral abscesses, arthritis, pneumonia, endocarditis, osteomyelitis, and several other suppurative infectious diseases in a broad range of domestic and wild animals, including cattle, swine, sheep, goats, camels, buffalo, deer, antelopes, reptiles, and birds. Severe infections caused by *T. pyogenes* have rarely been reported and are usually associated with occupational exposure, due to the bacterium is not the part of the human normal flora bacteria [4–8].

Several pathogenic properties are known in *T. pyogenes* which can increase its pathogenicity. Pyolysin (*plo*), as one of its major virulence factors, causes hemolysis and cytology of leukocytes. On the other hand, bacterial adhesion to epithelial cells and colonization, as well as degradation of DNA and sialic acid, are attributed to H and Pneuraminidases (*nanH* and *nanP*) of this bacterium.
Different types of fimbriae are expressed by *T. pyogenes*, including A, G, E, and C. These types of fimbriae are required for adherence to membranes and epithelial cells. Collagen-binding proteins (*Cbp*) and fibronectin-binding proteins (*Fbp*) are essential for adhesion to collagen-rich tissues (types 1, 2, and 4) and fibronectins. In addition, the protease and DNase of *T. pyogenes* provide nutrients for the bacteria through degradation of proteins and nucleic acids [1, 9, 10]. Despite this bacterium has been known as an opportunistic pathogen for human and animals for many decades, characterization of *T. pyogenes* genomes remains still necessary.

**Data description**

Here, we report the complete chromosome sequence of *T. pyogenes* strain Arash114. The strain was isolated from the uterus of a water buffalo (*Bubalus bubalis*) and is a major causative agent of metritis, abortion, and death. Therefore, we performed whole-genome sequencing. Genomic DNA was extracted using a commercial DNA extraction kit for Gram-positive bacteria according to the manufacturer’s instructions (Bioneer, South Korea). The quantity and quality properties of DNA was measured using Thermo-Fisher Nano-Drop Spectrophotometer model ND1000 (Thermo Fisher Scientific, DE). Sequencing was performed with the Illumina MiSeq platform using paired-end (PE) reads and Nextera library preparation. The sequences were de novo assembled using the CLC Genomics Workbench software (version 8) (Data set 1 and 2) [11, 12]. Genome annotation of the strain was performed using RAST annotation server [13]. Prediction of clustered regularly interspaced short palindromic repeats (CRISPRs) sequences was performed using CRISPRFinder online program [14, 15]. Antimicrobial resistance genes were identified using “ResFinder”, “card” and “NCBI AMRFinderPlus” databases and ABRicate software version 0.8 (https://github.com/tseemann/abricate) [16].

The complete genome of *T. pyogenes* was 2,338,282 bp long, with a GC content of 59.5%. The genome of the isolate was composed of a total of 2137 genes, including 2082 coding sequences (CDS) (total) and 2059 coding genes; 23 are pseudogenes, 6 are rRNAs, 46 are tRNAs, 3 are noncoding RNAs (ncRNAs) (Data file 2) [17], 1 is a clustered regularly interspaced short palindromic repeat (CRISPR) (Data file 3) [18] including 33 spacers (Data file 4) [19], and 16 are frameshifted genes. The genome encodes several known and putative virulence factors, including pyolysin, collagen-binding protein, neuraminidases (*nanH* and *nanP*), and 4 fimbrial proteins (Data file 2) [17]. Different antibiotics have often been used for treatment of *T. pyogenes* infections in veterinary practices [6]. In this study, a tetracycline resistance encoding gene (*tetW*) was detected in the genome of *T. pyogenes* strain Arash114 (Data file 5) [20]. Although, resistance to β-lactams, chloramphenicol and macrolides antibiotics have been reported among the *T. pyogenes* strains [21, 22]. No other specific antibiotic resistance genes were identified in strain Arash114 (Data file 6 and Data file 7) [23, 24]. Machado and Bicalho (2014) reported the complete genome sequence of *T. pyogenes* as an important opportunistic pathogen from livestock. They found several virulence factors such as collagen adhesion, fimbrial proteins, pyolysin and cytotoxin in the isolate [25]. Zhang

---

**Table 1** Overview of data files

| Label       | Name of data file/data set                                      | File types (file extension)                        | Data repository and identifier (DOI or accession number) |
|-------------|------------------------------------------------------------------|----------------------------------------------------|----------------------------------------------------------|
| Data set 1  | Complete genome file of *T. pyogenes* Arash114 strain            | Fasta file (.fasta)                                | ENA/EMBL (https://identifiers.org/insdc.gca:GCA_003055835.1) [11] |
| Data set 2  | Raw sequencing data of *T. pyogenes* Arash114 strain             | Fastq file (.fastq.gz)                             | ENA/EMBL (https://identifiers.org/ncbi:CP028833) [12]    |
| Data file 2 | Gene annotation results of *T. pyogenes* Arash114 strain         | Tab separated values (.tsv)                        | HARVARD Dataverse (https://doi.org/10.7910/DVN/HAJYNP) [17] |
| Data file 3 | CRISPR spacers of *T. pyogenes* Arash114 strain                 | Text file (.txt)                                  | HARVARD Dataverse (https://doi.org/10.7910/DVN/HAJYNP) [18] |
| Data file 4 | CRISPR sequences of *T. pyogenes* Arash114 strain               | Text file (.txt)                                  | HARVARD Dataverse (https://doi.org/10.7910/DVN/HAJYNP) [19] |
| Data file 5 | Antimicrobial resistance genes of *T. pyogenes* Arash114 strain based on ResFinder database | Comma-separated values (.csv) | HARVARD Dataverse (https://doi.org/10.7910/DVN/HAJYNP) [20] |
| Data file 6 | Antimicrobial resistance genes of *T. pyogenes* Arash114 strain based on card database | Comma-separated values (.csv) | HARVARD Dataverse (https://doi.org/10.7910/DVN/HAJYNP) [23] |
| Data file 7 | Antimicrobial resistance genes of *T. pyogenes* Arash114 strain based on NCBIAMRFinderPlus database | Comma-separated values (.csv) | HARVARD Dataverse (https://doi.org/10.7910/DVN/HAJYNP) [24] |
et al. [26] also isolated and sequenced the complete genome of T. pyogenes from livestock and they detected different virulence factor encoding genes including pyolysin, cbpA, fimC, nanH, nanP and fimE genes. However, this is the first study reported the complete genome sequence of T. pyogenes isolated from Water Buffalo (Bubalus bubalis) (Table 1).

The whole genome sequence presented in this study serve as a platform for detection of new genes that may contribute to antibiotic resistance and pathogenicity of T. pyogenes strain Arash114. This will advance our knowledge regarding the metabolism, virulence factors, antibiotic resistance and evolution of Arash114 strain and serve as an appropriate template for future researches.

Limitations
Annotations and genomic analysis of T. pyogenes strain Arash114 were performed with validated, novel and robust online and offline bioinformatics tools; consequently, the authors are currently unaware of any drawbacks and limitations of the data.

Abbreviations
Pio: Pyolysin; Cbp: Collagen-binding proteins; Fbp: Fibronectin-binding proteins; DNA: Deoxyribonuclease nucleic acid; RAST: Rapid Annotations using Subsystems Technology; CRISPRs: Clustered Regularly Interspaced Short Palindromic Repeats; CDs: Coding sequences.

Acknowledgements
We thank and appreciate our colleagues and technicians in the Faculty of Veterinary Science, Bu-Ali Sina University of Hamedan.

Authors’ contributions
The project was designed and conceived by IAT, AM and TZS. Data acquisition was carried out by IAT. Genomics analysis and data interpretation was performed by AGL, BP, PM and ZZK. The project was jointly supervised by AM, TZS and AM. AM was the principal investigator. The first draft of the manuscript was written by IAT and revisied by BP, AM, TZS, AGL and ZZK. All authors read and approved the final manuscript.

Funding
This project emanates from Ph.D. thesis of the first author and was financially supported by research grants from Bu-Ali Sina University of Hamedan.

Availability of data and materials
All data files 2–7 described in this data note can be openly and freely accessed on Harvard Dataverse (https://dataverse.harvard.edu [17–20, 23, 24]). Data sets 1 and 2 can be openly and freely accessed on the ENA/EMBL database. Sequence reads have been deposited in the ENA/EMBL sequence read archive under accession number CP028833 (https://identifiers.org/ncbi/CP028833) (Data set 1) [11]. The complete genome of T. pyogenes Arash114 has been deposited in GenBank under accession number CP028833 (https://identifiers.org/ncbi/CP028833) (Data set 1) [11]. The Project accession number for the genome sequencing project of T. pyogenes strain Arash114 is PRJNA449465. See the Table 1 and references for more details and links to all data.

Declarations
Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
1 Department of Pathobiology, Faculty of Veterinary Science, Bu-Ali Sina University, Hamedan, Iran. 2 Department of Microbiology and Immunology, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran. 3 Medical Microbiology Research Center, Qazvin University of Medical Sciences, Qazvin, Iran.

Received: 13 May 2021 Accepted: 11 August 2021
Published online: 23 August 2021

References
1. Jost BH, Billington SJ. Arcanobacterium pyogenes: molecular pathogenesis of an animal opportunist. Antonie Van Leeuwenhoek. 2005;88(2):87–102.
2. Quinn PJ, Markey BK, Leonard FC, Hartigan P, Fanning S, Fitzpatrick E. Veterinary microbiology and microbial disease. New York: Wiley; 2011.
3. Yassin A, Hupfer H, Siering C, Schumann P. Comparative chemotaxonomic and phylogenetic studies on the genus Arcanobacterium Collins et al. 1982 emend. Lehnen et al. 2006: proposal for Trueperella gen. nov. and emended description of the genus Arcanobacterium. Int J Syst Evolut Microbiol. 2011;61(6):1265–74.
4. Gahm-Hansen B, Frederiksen W. Human infections with Actinomyces pyogenes (Corynebacterium pyogenes). Diag Microbiol Infect Dis. 1992;15(4):349–54.
5. Narayanan S, Nagaraja T, Wallace N, Staats J, Chengappa M, Oberst R. Biochemical and ribotypic comparison of Actinomyces pyogenes and A pyogenes-like organisms from liver abscesses, ruminal wall, and ruminal contents of cattle. Am J Vet Res. 1998;59(3):271–6.
6. Zhao K-L, Liu Y, Zhang X-Y, Wang H-N, Yue B-S. Detection and characterization of antibiotic-resistance genes in Arcanobacterium pyogenes strains from abscesses of forest musk deer. J Med Microbiol. 2011;60(12):1820–6.
7. Zastempowska E, Lassa H. Genotypic characterization and evaluation of an antibiotic resistance of Trueperella pyogenes (Arcanobacterium pyogenes) isolated from milk of dairy cows with clinical mastitis. Vet Microbiol. 2012;161(1–2):153–8.
8. Ribeiro M, Risetti R, Bolanos C, Caffaro K, De Morais A, Lara G, et al. Trueperella pyogenes multispecies infections in domestic animals: a retrospective study of 144 cases (2002 to 2012). Vet Q. 2015;35(suppl_2):82–7.
9. Silva E, Gaivao M, Leitao S, Jost B, Carneiro C, Vilela C, et al. Genomic characterization of Arcanobacterium pyogenes isolates recovered from the uterus of dairy cows with normal puerperium or clinical metritis. Vet Microbiol. 2008;132(1–2):111–8.
10. Kasimanickam V, Owen K, Kasimanickam R. Detection of genes encoding multidrug resistance and biofilm virulence factor in uterine pathogenic bacteria in postpartum dairy cows. Theriogenology. 2016,85(2):173–9.
11. ENA/EMBL. Assembly. https://identifiers.org/ncbi/CP028833.1 (2020).
12. ENA/EMBL. Sequence read archive. https://identifiers.org/ncbi/insdc:CP028833 (2020).
13. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST server: rapid annotations using subsystems technology. BMC Genom. 2008;9(1):1–15.
14. Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 2007;35(suppl_2):W52–7.
15. Grissa I, Vergnaud G, Pourcel C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinform. 2007;8(1):1–10.
16. Seemann T. ABriCate: mass screening of contigs for antimicrobial and virulence genes. Department of Microbiology and Immunology, The University of Melbourne, Melbourne, Australia. 2018. https://github.com/tseemann/abricate. Accessed 28 Feb 2019.
17. Pakbin B. Data file 2: replication data for *Trueperella pyogenes* Arash114. 2021. Harvard Dataverse. https://doi.org/10.7910/DVN/HAJYNP.
18. Pakbin B. Data file 3: replication data for *Trueperella pyogenes* Arash114. 2021. Harvard Dataverse. https://doi.org/10.7910/DVN/HAJYNP.
19. Pakbin B. Data file 4: replication data for *Trueperella pyogenes* Arash114. 2021. Harvard Dataverse. https://doi.org/10.7910/DVN/HAJYNP.
20. Pakbin B. Data file 5: replication data for *Trueperella pyogenes* Arash114. 2021. Harvard Dataverse. https://doi.org/10.7910/DVN/HAJYNP.
21. Dong W-L, Liu L, O dah KA, Atiah LA, Gao Y-H, Kong L-C, et al. Antimicrobial resistance and presence of virulence factor genes in *Trueperella pyogenes* isolated from pig lungs with pneumonia. Trop Anim Health Prod. 2019;51(7):2099–103.
22. Santos T, Caixeta L, Machado V, Rauf A, Gilbert R, Bicalho R. Antimicrobial resistance and presence of virulence factor genes in *Arcanobacterium pyogenes* isolated from the uterus of postpartum dairy cows. Vet Microbiol. 2010;145(1–2):84–9.
23. Pakbin B. Data file 6: replication data for *Trueperella pyogenes* Arash114. 2021. Harvard Dataverse. https://doi.org/10.7910/DVN/HAJYNP.
24. Pakbin B. Data file 7: replication data for *Trueperella pyogenes* Arash114. 2021. Harvard Dataverse. https://doi.org/10.7910/DVN/HAJYNP.
25. Machado VS, Bicalho RC. Complete genome sequence of *Trueperella pyogenes*, an important opportunistic pathogen of livestock. Genome Announc. 2014;2(2):e00400-14.
26. Zhang SH, Qiu JJ, Yang R, Shen KF, Xu GY, Fu LZ. Complete genome sequence of *Trueperella pyogenes*, isolated from infected farmland goats. Genome Announc. 2016;4(6):e01421-16.

**Publisher’s Note**
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.