LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-BOOK-227152

Pestoides F, and Atypical Yersinia pestis Strain from the Former Soviet Union

E. Garcia, P. Worsham, S. Bearden, S. Malfatti, D. Lang, F. Larimer, L. Lindler, P. Chain

January 10, 2007

Proceedings of the 9th International Symposium on Yersinia

Updated October 14, 2003

UCRL-BOOK-227152

Pestoides F, an Atypical *Yersinia pestis* Strain from the Former Soviet Union

Emilio Garcia[1], Patricia Worsham[2], Scott Bearden[3], Stephanie Malfatti[1], Dorothy Lang[1], Frank Larimer[4], Luther Lindler[5] and Patrick Chain[1]

[1] Lawrence Livermore National Laboratory, Biosciences and Biotechnology Division, Chemistry Materials and Life Sciences Directorate, Livermore, CA 94550, garcia12@llnl.gov

[2] United States Army Medical Research Institute of Infectious Diseases, Fort Detrick, MD21702, patricia.worsham@us.army.mil

[3] Centers for Disease Control and Prevention, Bacterial Zoonoses Branch, Fort Collins, CO 80521, zyv3@cdc.gov

[4] Oak Ridge National Laboratory, Oak Ridge, TN 37831, larimerfw@ornl.gov

[5] Department of Bacterial Diseases, Walter Reed Army Institute of Research, Silver Spring, MD 20910, Luther.Lindler@na.amedd.army.mil

1 Abstract

Unlike the classical *Yersinia pestis* strains, members of an atypical group of *Y. pestis* from Central Asia, denominated *Y. pestis* subspecies *caucasica* (also known as one of several pestoides types), are distinguished by a number of characteristics including their ability to ferment rhamnose and melibiose, their lacking the small plasmid encoding the plasminogen activator (*pla*) and pesticin, and their exceptionally large variants of the virulence plasmid pMT (encoding murine toxin and capsular antigen). We have obtained the entire genome sequence of *Y. pestis* Pestoides F, an isolate from the former Soviet Union that has enabled us to carryout a comprehensive genome-wide comparison of this organism's genomic content against the six published sequences of *Y. pestis* and their *Y. pseudotuberculosis* ancestor. Based on classical glycerol fermentation (+ve) and nitrate reduction (+ve) *Y. pestis* Pestoides F is an isolate that belongs to the biovar antiqua. This strain is unusual in other characteristics such as the fact that it carries a non-consensus V antigen (*lcrV*) sequence, and that unlike other Pla$^-$ strains, Pestoides F retains virulence by the parenteral and aerosol routes. The chromosome of Pestoides F is 4,517,345 bp in size comprising some 3,936 predicted coding sequences, while its pCD and pMT plasmids are 71,507 bp and 137,010 bp in size respectively. Comparison of chromosome-associated genes in Pestoides F with those in the other sequenced *Y. pestis*

strains, reveals a series of differences ranging from strain-specific rearrangements, insertions, deletions, single nucleotide polymorphisms, and a unique distribution of insertion sequences. There is a single ~7 kb unique region in the chromosome not found in any of the completed *Y. pestis* strains sequenced to date, but which is present in the *Y. pseudotuberculosis* ancestor. Taken together, these findings are consistent with Pestoides F being derived from the most ancient lineage of *Y. pestis* yet sequenced.

2 Introduction

Most of our knowledge of *Yersinia pestis* at the genomic level has been obtained from studies on strains derived from "classical" isolates of the Americas, Africa and some from Asia that share some broad phenotypic and virulence properties. However, atypical strains from the territories encompassed by the former Soviet Union harbor a number of clearly distinct *Y. pestis* strains described extensively in the recent review by Anisimov and colleagues (Anisimov, Lindler and Pier 2004). Standard biochemical characteristic, their specific animal host range, virulence properties as well as numerical taxonomy (Martinevskii 1969) have helped classify these widely diverse strains of *Y. pestis*. The term "pestoides", in particular, was first coined by Martinevskii to describe *Y. pestis* isolates from natural foci located in the Transcaucasian highland, in the Mountain Altai and Transbikalian regions. The term was later adopted by scientists in the US to designate strains derived from the former Soviet Union (FSU) (P.L. Worsham and C. Roy 2003). Pestoides F, the strain whose genomic characteristics are being presented in this work, is one of a number of "pestoides" strains originally described by Worsham and colleagues (known as subspecies *caucasica* in the FSU) that is characterized by: its atypical biochemical characteristics; its lack of pPCP, one of the unique virulence plasmids of the *Y. pestis* group; the presence of an enlarged pMT plasmid; possessing unusual animal host specificity and displaying full virulence by parenteral and aerosol route in the mouse model.

In recent studies by Achtman et al. (2005), the pestoides isolates have been placed as early offshoots in phylogenetic trees, together with another recently characterized avirulent Y. pestis strain, 91001, which has been given the designation of subspecies Microtus. These studies however, were based on single nucleotide polymorphisms found among the strains available at the time. The result is that isolates closely related to those whose genomes have been sequenced can be firmly placed on the phylogenetic tree, yet distinct isolates that belong to independent non-sequenced lineages can be only loosely placed. We have determined the complete sequence of *Y. pestis* Pestoides F, an isolate we have determined belongs to the FSU *caucasica* subspecies, and have performed preliminary analyses that firmly place this group as the oldest lineage sequenced to date.

3 Materials and Methods

Completed whole genome sequence was derived from a standard genome shotgun approach, deep (~10-fold) sequencing of 3 differently sized libraries (average sizes of 3 kb, 7 kb and 40 kb). Genome closure was performed by a combination of PCR-sequencing and directed primer walking off of clones. Genome assembly was verified with the properly assembled clone end-sequences as well as by PCR where physical gaps remained. All repeat sequences surpassing the length of sequencing reads (~600 bp) were separately assembled to assure correct sequence and assembly.

Whole genome sequence alignments were performed using MUMmer (Kurtz et al. 2004) and/or BLAST (Altschul et al. 1997) and visualized using ACT. Separate gene alignments were conducted using BLAST and/or CLUSTALW (Thompson et al. 1994).

4 Results

The *Y. pestis* Pestoides F genome has been sequenced to completion and is available for download from our Lawrence Livermore Laboratory web site (http://www.llnl.gov/bio/groups/genomics_virulence/Yersinia/YersiniaGenomics.html) as well as from our sister site at the Oak Ridge National Laboratory Computational Biology site, as part of the Department of Energy's Joint Genome Institute (https://maple.lsd.ornl.gov/microbial/ypes_1570/). The genome of Pestoides F consists of a single chromosome of 4,517,345 bp and only two of the three virulence plasmids typical of *Y. pestis* strains, pCD and pMT at 71,507 bp and 137,010 bp in size respectively.

The pCD1 plasmid is essentially identical to that of CO92 but contains a non-consensus *lcrV* (see Figure 1), a gene encoding the V antigen, an important factor involved in immunosuppresion of the host during *Yersinia* infection. In contrast, the pMT plasmid of Pestoides F is substantially larger (137,010 bp in size) than those of the classical *Yersinia* (at roughly 96-100 kb) and is quite similar to that of the pFra plasmid reported by Golubov et al. (2004), although it contains a number of single nucleotide polymorphisms and rearrangements.

As in the other *Y. pestis* and *Y. pseudotuberculosis* genomes (Parkhill et al. 2001; Deng et al. 2002; Song et al. 2004; Chain et al. 2004) the chromosome of Pestoides F displays a large number of rearrangements compared to all other *Y. pestis* chromosomes. There are a total of 106 IS elements (of the four main types: IS100, IS1541, IS 285, IS1661) in the Pestoides F chromosome, which are associated with most of the rearrangement events observed between *Y. pestis* (as well as *Y. pseudotuberculosis*) genomes. In addition to the repeated IS elements, there are a few other observed instances of repeated sequences, most prominently seven copies of the rRNA operon (3 copies of the 16S-tRNA-tRNA-23S-5S, and 4 copies of the 16S-tRNA-23S-5S), as is observed in most *Y. pestis* strains (only 6 are present in the KIM sequence, due to a presumed deletion between IS100 elements).

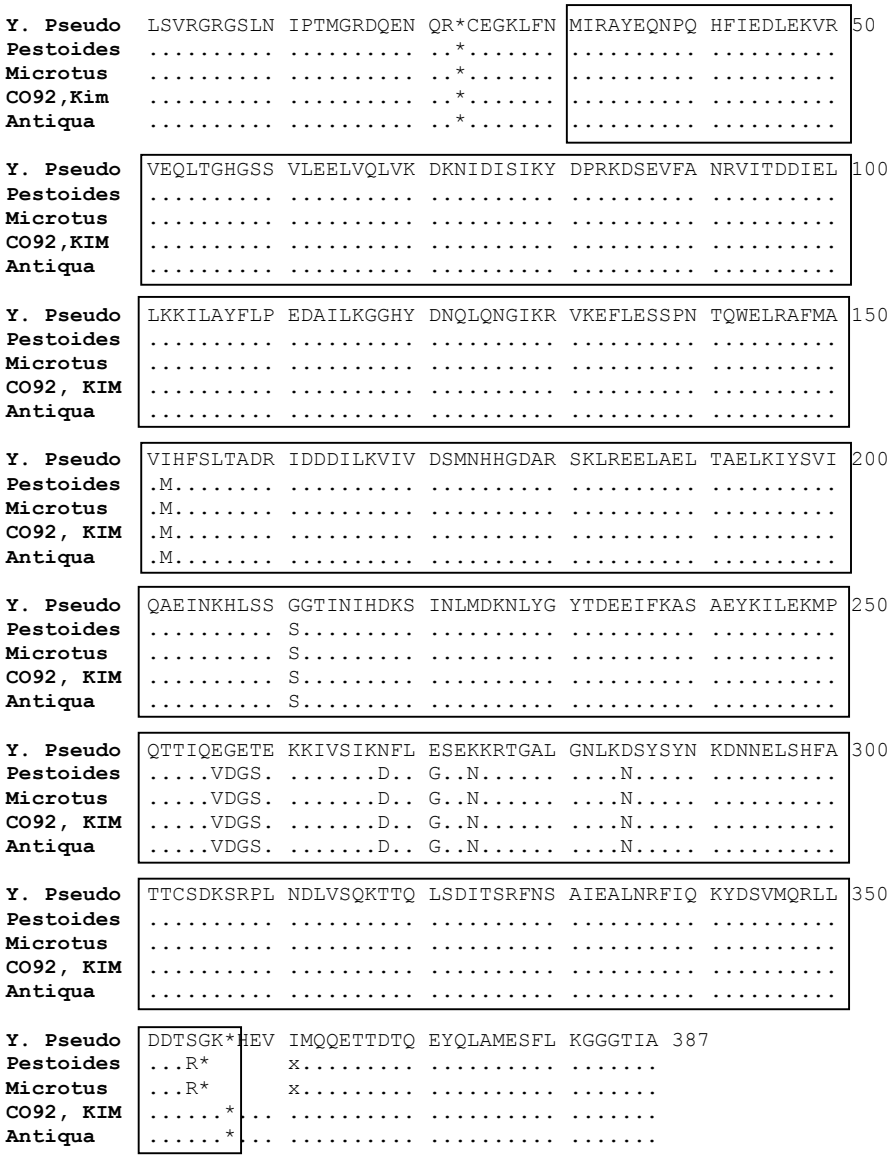

Fig. 1. Alignment of the *lcrV* gene of various *Y. pestis* strains. Dots indicate conserved amino acids; asterisks indicate a stop codon. A 16 nt deletion at the carboxyl end of Pestoides F and Microtus results in a shorter protein and a frameshift.

The Pestoides F strain has been characterized biochemically as an isolate of the Antiqua biovar, despite its classification as a pestoides isolate (in the *caucasica* subspecies). We examined the genes responsible for this classification, as identified by Motin et al. 2001 and Achtman et al. 2005, and have found wild type versions of *glpD*, which in Orientalis strains harbors a deletion responsible for the glycerol fermentation negative phenotype, as well as wild type *napA*, which in Medievalis strains harbors one of two mutations responsible for the nitrate reduction negative phenotype. Of special note is a ~7 kb unique region not found in any other *Y. pestis* strain but shared with the enteropathogen ancestor, *Y. pseudotuberculosis* and several of the non-pathogenic Yersiniae (*Y. intermedia*, *Y. mollaretii* and *Y. frederiksenii*). This region has been found to encode a number of genes involved in various enzymatic activities (apolipoprotein acyl transferase, aminotransferase, aldolase, methyylthioribose kinase).

A recently acquired panel of *Y. pestis* isolates have been obtained from the National Center for Disease Control of the Republic of Georgia and, based on both phenotypic properties and an IS100-based genotyping (Motin et al. 2002), have been placed as being closest in properties to Pestoides F. We therefore compared plasmid composition, unique regions and pattern of gene inactivation between Pestoides F and the Georgian isolates. It was determined that in each case, the Pestoides F characteristics closely resemble those of the Georgian strains. This is in agreement with Pestoides F belonging to the *Y. pestis* subspecies *caucasica* in the classification used for strains from the former Soviet Union (Anisimov et al. 2004), and together with the above results, suggests that this group may belong to one of the oldest lineages of *Y. pestis*.

5 Conclusion

Detailed genome comparisons among closely related bacterial strains offer an incomparable opportunity to ascertain important evolutionary relationships that are more difficult, or impossible, to obtain by other means. The genome sequence of Pestoides F has been determined and found to be similar in content to other *Y. pestis* strains with a number of important exceptions. The comparison between this newly-sequenced atypical *Y. pestis* strain and those from previously sequenced members of this group has led us to conclude that Pestoides F likely belongs to the oldest linage of *Y. pestis* thus far sequenced. Facts that support this conclusion include: the lack of pPCP plasmid that encodes the plasminogen activator and pesticin (presumably not acquired by this isolate); the “unique” region of the genome of Pestoides F compared with all other completed *Y. pestis* is shared with *Y. pseudotuberculosis* (the *Y. pestis* progenitor). Together with the observed genome rearrangement and IS element abundance/patterns these findings strongly suggest that Pestoides F derives from the most ancient lineage of *Y. pestis* yet studied.

6 Acknowledgments

This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48. This work was financially supported by the Department of Homeland Security, Office of Science, Chemical and Biological Division.

References

Achtman, M., Morelli, G., Zhu, P., Wirth, T., Diehl, I., Kusecedk, B., Vogler, A.J., Wagner, D.M., Allende, C.J, Easterday, W.R., et al. (2005) Microevolution and history of the plague bacillus, *Yersinia pestis*. Proc. Ntl. Acad. Sci, USA. 101, 17837-17842.

Altschul S.F., Madden, T.L, Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389-3402.

Anisimov, A.P., Lindler, L.E. and Pier, G.B. (1996) Intraspecific Diversity of *Yersinia pestis*. Clin. Microbiol. Rev. 17, 434-464.

Chain, P.S.G., Carniel, E., Larimer, F.W., Lamerdin, J., Stoutland, P.O., Regala, W.M., Georgescu, A.M. Vergez, L.M., Land, L., Motin, V.L., et al. (2004) Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. Proc. Ntl.Acad. Sci. USA, 101,13826-13831.

Deng, W., Burland, V., Plunkett III, G., Boutin, A., Mayhew, G.F., Liss, P. Perna, N.T., Rose, D.J., Mau, B., Zhou, S., Schwartz, D.C, Fetherston, J., et al. (2002) Genome Sequence of Yersinia pestis KIM. J. Bacteriol. 184, 4601-4611.

Golubov, A., Neubauer, H., Nolting, C. Heesemann, J. and Rakin, A. (2004) Structural Organization of the pFra Virulence-Associated Plasmid of Rhamnose-Positive *Yersinia pestis*. Infect. & Immun. 72, 5613-5621.

Kurtz, S., A. Phillippy, A. L. Delcher, M. Smoot, M. Shumway, C. Antonescu, and Salzberg, S. L. (2004) Versatile and open software for comparing large genomes. Genome Biol. 5:R12.

Martinevskii, I.L. (1969) Biology and genetic features of plague and plague-related microbes. Meditsina Press, Moscow, USSR.

Song, Y., Tong, Z., Wang, J., Wang, L., Guo, Z., Han, Y., Zhang, J., Pei, D., Zhou, D., Qin, H., Pang, X., et al. (2004) Complete Genome Sequence of *Yersinia pestis* Strain 91001. DNA Res, 11, 179-197.

Thompson J.D., Higgins,D.G. and Gibson,T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680.

Worsham, P.L. Hunter, M. (1998) Characterization of Pestoides F, an atypical strain of *Yersinia pestis*. Med Microbiol. 6(Suppl.II):34-35.