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

Emilio Garcia1, Patricia Worsham2, Scott Bearden3, Stephanie Malfatti1, Dorothy Lang1, Frank Larimer4, Luther Lindler5, and Patrick Chain1

1 Biosciences and Biotechnology Division, Chemistry Materials and Life Sciences Directorate, Lawrence Livermore National Laboratory, garcia12@llnl.gov
2 United States Army Medical Research Institute of Infectious Diseases
3 Bacterial Zoonoses Branch, Centers for Disease Control and Prevention
4 Oak Ridge National Laboratory
5 Department of Bacterial Diseases, Walter Reed Army Institute of Research

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 lack of 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.

2.1 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 et al. (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) (Worsham and Roy 2003). Pestoides F, the strain whose genomic characteristics are being presented in this work, is one of a number of “pestoides” strains (known as subspecies *caucasica* in the FSU) originally described by Worsham and colleagues 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, possession of unusual animal host specificity and full virulence by parenteral and aerosol route in the mouse model (Worsham and Hunter 1998; Worsham and Roy 2003).

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.

### 2.2 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).