CrossMark
click for updates

Complete Genome Sequence of *Yersinia ruckeri* Strain CSF007-82, Etiologic Agent of Red Mouth Disease in Salmonid Fish

Michael C. Nelson,[a] Scott E. LaPatra,[b] Timothy J. Welch,[c] Joerg Graf[a]

Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut, USA[a]; Clear Spring Foods, Inc., Research Division, Buhl, Idaho, USA[b]; National Center for Cool and Cold Water Aquaculture, Agricultural Research Service, U.S. Department of Agriculture, Kearneysville, West Virginia, USA[c]

We present the complete, closed, and finished chromosomal and extrachromosomal genome sequences of *Yersinia ruckeri* strain CSF007-82, the etiologic agent of enteric red mouth disease in salmonid fish. The chromosome is 3,799,036 bp with a G+C content of 47.5% and encodes 3,530 predicted coding sequences (CDS), 7 ribosomal operons, and 80 tRNAs.

Address correspondence to Timothy J. Welch, tim.welch@ars.usda.gov, or Joerg Graf, joerg.graf@uconn.edu.

Yersinia ruckeri, a Gram-negative gammaproteobacterium, is the etiologic agent of enteric red mouth disease (ERM) a hemorrhagic septicemia of farmed salmonid fish species worldwide ([1], [2]). *Y. ruckeri* strains include several O-serotypes, yet nearly all outbreaks are caused by a genetically homogenous group of serotype 01 strains. Successful bacterin vaccines for ERM were developed in the 1970s; however, outbreaks in vaccinated fish have recently occurred and have been associated with the emergence of novel variants of this pathogen ([3], [4]). The reemergence of *Y. ruckeri* mediated ERM has renewed interest in the virulence and evolution of this pathogen.

Presently, only two draft *Y. ruckeri* genomes are available; the genome of the type strain, ATCC 29473[T], consists of 174 contigs, while the genome of a strain from Chile consists of 75 contigs ([5]). An Illumina-based assembly of the pathogenic *Y. ruckeri* strain CSF007-82 ([6]) yielded 63 contigs with N_{50} of 160,889 bp. Because of this, we resequenced strain CSF007-82 using PacBio in order to obtain a complete, circularized genomic reference for use in future studies examining determinants of *Y. ruckeri* pathogenicity.

Large insert libraries were prepared and sequenced on a Pacific Biosciences RS II instrument using the P5-C3 sequencing chemistry at the Yale Center for Genomic Analysis, generating over 1.64 Gbp of sequencing data from ~200,000 raw reads. The resulting data was assembled using the HGAP version 2 assembler ([7]) with standard parameters, yielding two unique contigs ~100 kbp and ~3.8 Mbp in size. Overlapping nucleotides were manually edited and removed from the contig ends to create circularized replicons, which were then used as reference sequences for remapping of the raw sequencing reads to identify and correct any remaining errors. Interestingly, a contig corresponding to the plasmid pYR2 identified in the Illumina assembly was not generated, however, sequencing reads did map to pYR2 during the reference assembly. The largest contig, representing the chromosome, is 3,799,036 bp with a G+C content of 47.5% while the second contig is 103,917 bp with a G+C content of 48.4%. pYR2 is 16,923 bp with a G+C content of 47.2%.

The three finished contigs were functionally annotated by RAST ([8]). The chromosome encodes 3,517 predicted coding sequences (CDS), 22 rRNAs organized into 7 ribosomal operons, and 80 tRNAs covering all standard amino acids. Numerous phage and transposable element proteins were identified. The ~100 kbp contig encoded 99 predicted CDS, 16 of which belonged to the IncI1 class of replication proteins. No RNAs were predicted for this contig, indicating that it represents a previously unknown plasmid we named pYR3.

The availability of a complete, finished *Y. ruckeri* genome sequence, in combination with additional genome sequences should allow for more thorough analysis and identification of novel virulence factors associated with ERM. Additionally, the availability of a closed genome sequence should enable further studies of the evolution of the *Yersinia* genus as well as an examination of the transmission and acquisition of mobile genetic elements.

Nucleotide sequence accession numbers. The sequencing data is available under the GenBank/ENA/DDBJ accession no. PRJEB6967. The annotated genome is available under accession numbers LN681229 to LN681231.

REFERENCES

1. **Busch RA.** 1978. Enteric redmouth disease (Hagerman strain). Mar Fish Rev **40:**467–472.
2. **Ewing WH, Ross AJ, Brenner DJ, Fanning GR.** 1978. *Yersinia ruckeri* sp. nov., the redmouth (RM) bacterium. Int J Syst Bacteriol **28:**37–44. http://dx.doi.org/10.1099/00207713-28-1-37.
3. **Welch TJ, Verner-Jeffreys DW, Dalsgaard I, Wiklund T, Evenhuis JP, Cabrera JAG, Hinshaw JM, Drennan JD, Lapatra SE.** 2011. Independent emergence of *Yersinia ruckeri* biotype 2 in the United States and Europe. Appl Environ Microbiol **77:**3493–3499. http://dx.doi.org/10.1128/AEM.02997-10.
4. **Wheeler RW, Davies RL, Dalsgaard I, Garcia J, Welch TJ, Wagley S, Bateman KS, Verner-Jeffreys DW.** 2009. *Yersinia ruckeri* biotype 2 isolates from mainland Europe and the UK likely represent different clonal groups. Dis Aquat Org **84:**25–33. http://dx.doi.org/10.3354/dao02039.

5. **Navas E, Bohle H, Henríquez P, Grothusen H, Bustamante F, Bustos P, Mancilla M.** 2014. Draft genome sequence of the fish pathogen *Yersinia ruckeri* strain 37551, serotype O1b, isolated from diseased, vaccinated Atlantic salmon (*Salmo salar*) in Chile. Genome Announc **2**(4):e00858-14. http://dx.doi.org/10.1128/genomeA.00858-14.

6. **Evenhuis JP, Lapatra SE, Verner-Jeffreys DW, Dalsgaard I, Welch TJ.** 2009. Identification of flagellar motility genes in *Yersinia ruckeri* by transposon mutagenesis. Appl Environ Microbiol **75**:6630–6633. http://dx.doi.org/10.1128/AEM.01415-09.

7. **Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J.** 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods **10**:563–569. http://dx.doi.org/10.1038/nmeth.2474.

8. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O.** 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics **9**:75. http://dx.doi.org/10.1186/1471-2164-9-75.