

Type of the Paper (Article, Review, Communication, etc.)

Francisella frigiditurreis

Julia Khabibulina

¹ ylykhab@gmail.com

Abstract: This review presents several researches on the genome of *Francisella frigiditurreis* including GCcontent analysis, protein length analysis and distribution of protein genes between direct and complimentary chains. According to the results of the researches, some proposals were put forward.

Keywords: *Francisella frigiditurreis*; genome; bioinformatics

1. Introduction

Francisella frigiditurreis sp. nov. was just recently diagnosed in *Francisella* genus due to complexity of delineation of new species when isolates having very close genomic characteristics exhibit different physiological features. *Francisella frigiditurreis* and other isolates' genome was important to sequence to provide a knowledge base for future comparison of *Francisella* organisms identified in clinical samples or environmental surveys. ⁽¹⁾

Francisella was widely used during the last world war as a biological weapon because it causes extremely unpleasant, but mostly not fatal symptoms of a disease called tularemia. The name of the disease was chosen according to the territory it was first isolated from gophers. It was in the province of California- Tulare. It's a gram-negative, small coccobacil which forms capsules, motionless. Obligate aerobic and needs cysteine for growth. The genome is represented by ring DNA.

Known Francisella is pathogenic both for people and rodents, but the natural human immunity is high enough. ⁽²⁾

2. Materials and Methods

To get information on the genome of *Francisella frigiditurreis* you can check on "genetable". Study chromosomes of the bacteria you can on the "feature_table.txt" (you can find them on the section of supplementary materials).

3. Results

1. Genome information

Genome of *Francisella frigiditus* contains one chromosome with a lot of repetitive elements. Chromosome's length is 2,045,775 bp along with a sma;; 3,936 bp plasmid due to research [1]. Also, G+C contain is counted. You can find it using the formula: (GC-nucleotides number)/(general nucleotide number) using a fasta-type file with nucleotide sequence.. In the following table you can see structured information about genome:

	length (bp)	G+C content
chromosome	2,045,775	32%
plasmid	3,936	28,4%

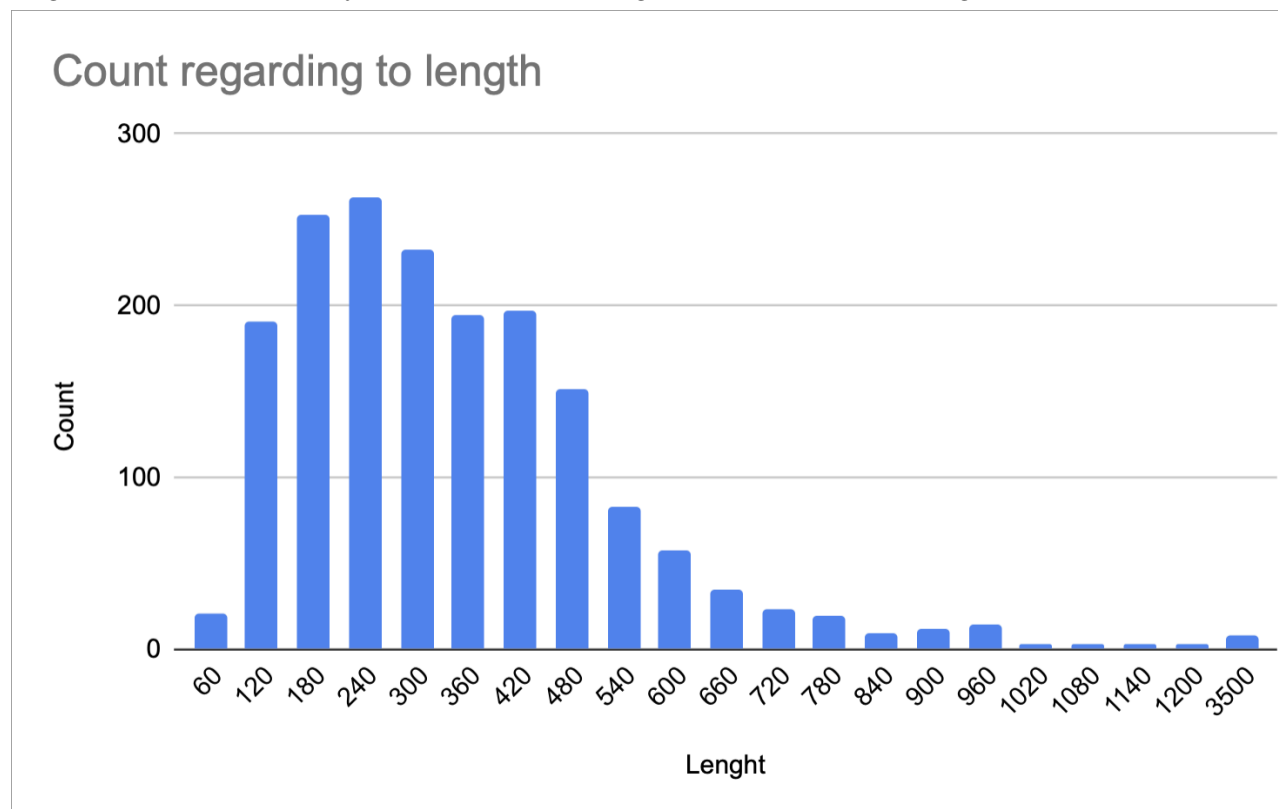
Pic. 1. Length and G+C contain in chromosome and plasmid

*Actually, this information is correct to *Francisella tularensis*, but as *Francisella frigiditus* is one of its closest neighbors we can say the information is very close to truth since there is almost no information on the investigated bacterium.

Both chromosome and plasmid have almost same percentage of G and C nucleotides and it's not very high. As we know pyrimidine provides protein with stability so we can suggest that *Francisella frigiditurris* does not have good temperature resistance.

2. Histogram of proteins lengths

A histogram of protein length was constructed. Due to it the most frequently encountered length is in the range of 120-480 n.a. and very small amount have length lower than 60 n. a. or higher than 540 n. a. [3]



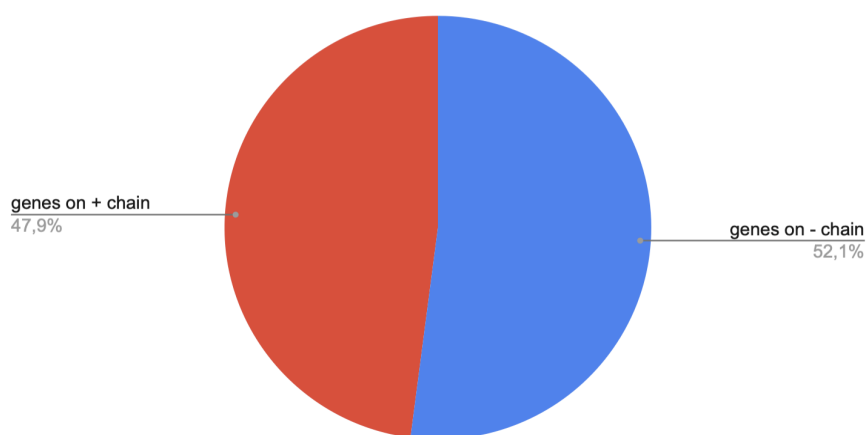
Pic. 2. Histogram of proteins length

Most common length is 240 n.a., but in general length variates between 120-480 n.a. Very small amounts of proteins have length less than 120 n. a. and more than 660 n. a.

3. Distribution of protein genes between direct and complimentary chains

Due to google tablemade diagram based on genome features table it becomes clear that complimentary chain contain a little bit more protein genes than on direct chain. So we can suppose that coding DNA sequence has more non coding genes than non-coding one.

Distribution of protein genes between direct and complimentary chains



Pic. 3. Distribution of protein genes between direct and complimentary chains the text following

4. Gene categories

ribosomal protein	hypothetical protein	transport protein
58	291	3

Amount of ribosomal protein is 5 times less than the hypothetical protein and nuclear transport factor 2 which is needed for efficient protein transport in cell, electron transport complex which controls translocation of ions across the membrane and sodium genes that provide the instructions for making sodium channels. All of them control membrane transport in cell.

5. RNA categories

rRNA	tRNA	tmRNA
35	44	2


There is no transport RNA identified in the cell, but there are two transport-matrical RNAs. Amount of rRNA and tRNA are almost the same.

Supplementary Materials:

To get information on chromosomes you can check feature table: [Genome features](#)

To get more information about genome follow the link: [GCF_000012925.1_ASM1292v1](#)

References

1. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5244304/>
2. https://ru.wikipedia.org/wiki/Francisella_tularensis
3.  Practise 8 - histogram of protein length