



OPEN

Genetic sequencing of a 1944 Rocky Mountain spotted fever vaccine

Yongli Xiao^{1✉}, Paul A. Beare², Sonja M. Best³, David M. Morens⁴, Marshall E. Bloom⁵ & Jeffery K. Taubenberger¹

Rocky Mountain spotted fever (RMSF) is a rapidly progressive and often fatal tick-borne disease caused by *Rickettsia rickettsii*. Its discovery and characterization by Howard Ricketts has been hailed as a remarkable historical example of detection and control of an emerging infectious disease, and subsequently led to the establishment of the Rocky Mountain Laboratories (RML). Here, we examined an unopened bottle of a vaccine, labeled as containing RMSF inactivated by phenol-formalin of infected ticks, developed prior to 1944 at RML by DNA analysis using Illumina high throughput sequencing technology. We found that it contains DNA from the Rocky Mountain wood tick (*Dermacentor andersoni*), the vector of RMSF, the complete genome of *Rickettsia rickettsii*, the pathogen of RMSF, as well as the complete genome of *Coxiella burnetii*, the pathogen of Q-fever. In addition to genomic reads of *Rickettsia rickettsii* and *Coxiella burnetii*, smaller percentages of the reads are from *Rickettsia rhipicephali* and *Arsenophonus nasoniae*, suggesting that the infected ticks used to prepare the vaccine carried more than one pathogen. Together, these findings suggest that this early vaccine was likely a bivalent vaccine for RMSF and Q-fever. This study is the among the first molecular level examinations of an historically important vaccine.

Rocky Mountain spotted fever (RMSF), was first recognized as an unknown disease called as “black measles” in the nineteenth century, especially after the 1890s in the Bitterroot Valley of southwest Montana¹. Edward E Maxey^{2,3} provided the first clinical description of the so-called spotted fever of Idaho: “a febrile disease, characterized clinically by a continuous moderately high fever, and a profuse or purpuric eruption in the skin, appearing first on ankles, wrists, and forehead, but rapidly spreading to all parts of body”. The case fatality rate sometimes approached 80%⁴. The 1928 establishment and subsequent development of Rocky Mountain Laboratories (RML), component of the National Institutes of Health (NIH) since 1937, was a direct result of research on Rocky Mountain spotted fever that began around 1900, in the Bitterroot Valley⁵.

The causal agent of RMSF is a tick-borne rickettsial bacterium known as *Rickettsia rickettsii* (*R. rickettsii*), identified by pathologist Howard T. Ricketts⁶ and subsequently named in honor of his discovery⁷. The importance to medical history of this discovery was emphasized by Richard Shope who was presented with the Howard T. Ricketts Prize in 1963⁸. *Rickettsia rickettsii* is a gram-negative, intracellular, coccobacillus bacterium that is around 0.8 to 2.0 μm long⁹. After infection, its initial targets are CD68 + cells (macrophages and/or dendritic cells)¹⁰ which then spread hematogenously throughout the body and infect vascular endothelial cells. The bacteria can proliferate in the nucleus or in the cytoplasm of the infected host cell^{11,12}. *Rickettsia rickettsii* possesses two major immunodominant surface proteins of outer membrane protein A (OmpA, 190 kDa) and outer membrane protein B (OmpB, 135 kDa), which is the most abundant surface protein of *Rickettsia*³. OmpA and OmpB contain species-specific epitopes that provide the basis for rickettsial serotyping in comparative indirect immunofluorescence assays¹³. OmpA is important for *R. rickettsii* adhesion to host cells¹⁴ and interacts with $\alpha\beta 1$ integrin to promote invasion of the bacteria into the host cells¹⁵. OmpB binds to host cell-specific receptor

¹Viral Pathogenesis and Evolution Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 33 North Drive MSC 3203, Bethesda, MD 20892-3203, USA. ²Coxiella Pathogenesis Section, Laboratory of Bacteriology, Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Hamilton, MT, USA. ³Innate Immunity and Pathogenesis Section, Laboratory of Persistent Viral Diseases, Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Hamilton, MT, USA. ⁴Office of the Director, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. ⁵Biology of Vector-Borne Viruses Section, Laboratory of Virology, Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Hamilton, MT, USA. ✉email: yongli.xiao@nih.gov

Ku70 (a subunit of a nuclear DNA-dependent protein kinase with subcellular localization in the cytoplasm and plasma membrane) and contributes to endocytosis and rickettsial internalization¹⁶. Both OmpA and OmpB are conserved throughout the spotted fever group *Rickettsia*, whereas OmpB is conserved in all *Rickettsia* species except for *R. canadensis*^{17–19}. For this reason, some genetic studies have used the OmpA and OmpB gene region to distinguish different *Rickettsia* species^{20–22}.

Dermaecentor andersoni (*D. andersoni*), named in honor of John F. Anderson (1873–1958) in 1908²³, is the primary vector of the RMSF (*R. rickettsii*)²⁴, tularemia²⁵, and Colorado tick fever (CTF) virus²⁶. It also is the vector of Q fever²⁷ and bovine anaplasmosis²⁸. In 1903, John F. Anderson (1873–1958) examined epidemiological data and found that all RMSF cases were associated with tick exposure in the week before the onset of spotted fever²⁹. Adult *D. andersoni* feed on mammals, including humans, dog, horses, cattle, chipmunks, ground squirrels, marmots, and jackrabbits^{30,31}. Its genome was sequenced by Agricultural Research Service, United States Department of Agriculture and submitted to GenBank in 2022 (GenBank Accession: JALBCO000000000).

In 1924, R. R. Spencer and R. R. Parker at RML prepared the first vaccine against RMSF by crushing infected ticks and phenol-inactivating the material³². In 1938, a simpler method of growing *Rickettsia* in the yolk sacs of developing chick embryo was developed by Herald R. Cox (1907–1986), utilizing formaldehyde inactivation, and extraction with ether³³. Both vaccines showed protection in animal studies³⁴ and were used in humans starting in 1927³⁵, but neither vaccine conferred a high level of human immunity³⁶. A subsequent RMSF vaccine developed in duck and chicken embryo culture followed by formalin inactivation had higher immunogenic activity and lower impurity^{37,38}. However, complete protection against RMSF following vaccination with formalin-inactivated vaccines has not been achieved in humans to date, possibly because of alterations in the antigenic determinants due to the fixation method³⁹. Recently, there have been studies of subunit^{40,41} or polypeptide⁴² vaccines of RMSF based on identified immunogenic surface proteins of *R. rickettsii*. Their studies showed that recombinant OmpA, OmpB, and Adr2 protein as antigens can develop antibody and T cell responses and provide protection in guinea-pigs⁴⁰ and mice⁴¹. While Wang, et al.⁴² used immunodominant peptides as antigen in mice, which induce a Th1-type immune response against *R. rickettsii* infection. All these new studies and developments may lead to an effective new RMSF vaccine eventually. However, currently there is no licensed vaccine available for RMSF because of the effective treatment by doxycycline at early stage of infection⁴³, decreased case fatality rate (from 28% in 1944 to 0.1% in 1995)⁴³, and the limited understanding of the protective host response and the *R. rickettsii* antigens involved in stimulating protective immunity⁴⁴.

Another intracellular bacterium, *Coxiella burnetii* (*C. burnetii*), was identified as the causative agent of Q fever in the late 1930s. Q fever was first described by Derrick in abattoir workers in Brisbane, Queensland, Australia⁴⁵. The pathogen of Q fever (*C. burnetii*) was discovered by Burnet (1899–1985) and Mavis Freeman when they studied one of Derrick's patients in 1937⁴⁶ and was near simultaneously isolated from *D. andersoni* at RML by Gordon David (1889–1977) and Herold Cox (1907–1986)²⁷. *Coxiella burnetii*, named to honor both Cox and Burnet, is an obligate intracellular, small gram-negative bacterium (0.2 and 2.0 μm)⁴⁷ with ~ 2 million base pair DNA genome⁴⁸ and its phylogenic neighbors include Legionellae spp, Francisella tularensis, and Rickettsiella spp.⁴⁹. When infected usually by inhalation of infectious aerosols generated by infected domestic livestock reservoirs such as dairy cows, goats, and sheep, Q fever generally presents in humans as an acute influenza-like illness followed by full recovery, particularly after treatment with doxycycline or other antibiotics⁵⁰. However, because its high infectivity through aerosol route^{51,52}, environmental resistance, and ability to cause disease, *C. burnetii* is classified as a Select Agent by the United States Centers for Disease Control and Prevention's Division of Select Agents and Toxins (DSAT) and is a noted bioterrorism "pathogen of interest"⁵³.

Q fever vaccine development started almost immediately following the identification of *C. burnetii* at RML, in infected ticks and later more-efficiently cultured in embryonated chicken eggs. Like the RMSF vaccine, this early Q-fever vaccine was made from formalin-inactivated infected *D. andersoni* tissues, and shown to offer protection in animal models⁵⁴. The first Q fever vaccine introduced for use in humans, comprised of whole-cell, formaldehyde-inactivated, ether-extracted *C. burnetii* with 10% egg yolk sac⁵⁵, and showed protection against high-dose aerosol challenge in US army volunteers⁵⁶. The only currently available Q-fever vaccine, Q-vax, is an iteration of the whole cell vaccine from RML. Despite being different diseases with different epidemiology, bivalent RMSF and Q fever vaccines were made at RML, prepared from formalin-inactivated adult *D. andersoni* that were infected with agents for both Rocky Mountain spotted fever (*R. rickettsii*) and American Q fever (*C. burnetii*), and was protective in guinea pigs following challenge with both organisms⁵⁴.

In the current study, an unopened 1944 vaccine bottle from RML labelled as "Rocky Mountain Spotted Fever Vaccine" was opened, DNA was extracted from the vaccine material and analyzed using Next Generation Sequencing (NGS) technology, which revealed the contents of the first RMSF vaccine produced in last century.

Results

DNA recovery. An unopened, sealed glass bottle of Rocky Mountain Spotted Fever Vaccine was provided by Rocky Mountain Laboratories, National Institute of Health (Fig. 1). The manufacture date of this vaccine is September 1944. In this RMSF vaccine bottle, there were visible dark red small solid pieces floating in the liquid and settling at the bottom (Fig. 1). The pH value of the vaccine solution was 5.5. Test DNA isolations were performed on both solid and liquid portions from 400 μl of the vaccine contents. DNA was only recovered from the solid materials of the vaccine. The solid materials from a 3 ml vaccine suspension were collected from which DNA was isolated, yielding approximately 6 ng of DNA. The DNA size profile was in the range of 50-to-200 bp with a peak at ~ 120 bp (Fig. 2a).

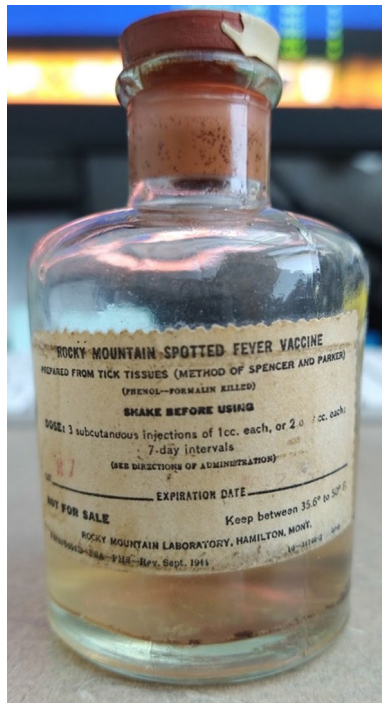


Figure 1. The bottle of Rocky Mountain Spotted Fever Vaccine made in 1944.

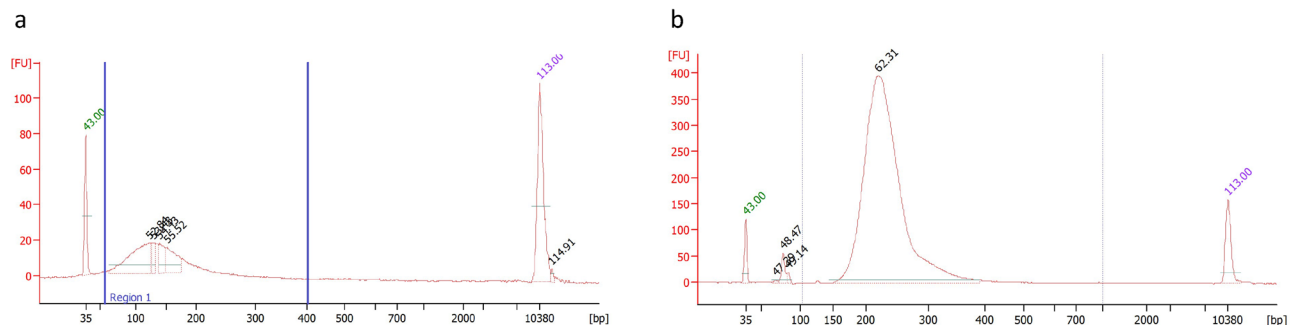


Figure 2. DNA size profile measured using a High Sensitivity DNA chip on an Agilent Bioanalyzer. **(a)** Isolated DNA from RMSF vaccine; **(b)** Final Illumina sequencing library with adaptor-ligated made from DNA isolated from RMSF vaccine. y-axis of the electropherograms represents fluorescent units (FU) and the x-axis represents the nucleotide length in base pair (bp).

Illumina sequencing. Isolated DNA from RMSF vaccine solids was used to make an Illumina library without shearing since the peak DNA length was ~ 120 bp. The final constructed Illumina library profile is shown in Fig. 2b.

The constructed library was diluted and sequenced on a NextSeq 500 sequencer with a 160 bp single read run. At completion, a total of 83.9 Gb raw sequence data (616.45 M raw reads) with 82.3% of total bases \geq Phred score of Q30 was generated. All sequences generated were deposited as a series into the Genbank SRA database (Accession No. PRJNA880551).

Sequence alignment to tick genome. As shown in Fig. 1, this RMSF vaccine was prepared from infected *D. andersoni* tick tissues. All generated Illumina reads were trimmed and aligned to reference genome of *Derma-centor andersoni* (JALBCO000000000.1) using Bowtie2⁵⁷ with default settings. Among the total of 519,012,047 trimmed reads, 396,715,374 reads (76.44%) were aligned to the *D. andersoni* genome (with 203,363,636 (39.18%) aligned exactly 1 time and 193,351,738 (37.25%) aligned $>$ 1 times). The average genome coverage for all 3120 contigs of *D. andersoni* genome was 94.6%. Therefore, DNA sequence analysis confirmed that this vaccine was prepared from infected *D. andersoni* ticks.

Sequence alignment to rickettsial genomes. To confirm the presence of rickettsial species DNA in this vaccine, all sequences were aligned to 14 complete rickettsial genomes that downloaded from NCBI. The

numbers of aligned reads and percentages are shown in Table 1: the species with the most aligned reads was *R. rickettsii*. The average of genome coverage of *R. rickettsii* is 1021.58 times and 99.7% of the genome (1,268,201 bp) is covered by the mapped reads (Table 2).

There were a total of 27,909 single nucleotide polymorphisms (SNPs) observed compared to the reference genome (see Materials and Methods) with 11,790 nonsynonymous SNPs (Supplemental Table 1). The consensus *R. rickettsii* vaccine-derived sequence was submitted to GenBank with accession number of CP114277.

As shown in Table 1, *R. rickettsii* has the most aligned reads (2.04%), but *R. slovacica* (2.02%), *R. rhipicephali* (2.0%), and *R. conorii* (1.99%) also had significant aligned reads. All of them belong to the rickettsial spotted fever group and their phylogenetic relationships are close to each other⁵⁸. Previous studies have used OmpA, OmpB, and GltA genes to distinguish different rickettsial species^{20–22,59}. Therefore, the obtained SNPs at OmpA gene region (5,855 bp, from 1,176,559 to 1,182,413 of NC_010263.3), OmpB gene region (4,962 bp, from 1,014,075 to 1,019,036 of NC_010263.3), and GltA gene region (1,305 bp, from 1,212,499 to 1,213,803 of NC_010263.3) from *R. rickettsii* were compared to the differences at these regions of *R. slovacica*, *R. rhipicephali*, and *R. conorii* genomes respectively (Supplemental Table 2).

A total of 221 SNPs were identified in the OmpA region from the vaccine-derived DNA library. Among them, there are 12 called SNPs that have variant base numbers in which > 50% of total aligned read numbers and the remaining SNPs (209) are all minor SNPs with their variant rates ranging from 10.26% to 49.28%. Therefore, there are only 12 SNP differences from the vaccine-derived consensus OmpA gene sequence as compared to the reference *R. rickettsii* (NC_010263.3) sequence, while at all these 221 called SNP sites, 175 of them are different bases between reference genomes of *R. rickettsii* and *R. rhipicephali*; 64 are different between *R. rickettsii* and *R. slovacica*; and 71 are different between *R. rickettsii* and *R. conorii* (Table 3, Supplemental Table 2).

A total of 207 SNPs were identified in the OmpB region from the vaccine-derived DNA library. Among them, only 4 have variant base numbers in which > 50% of total aligned read numbers were observed at these positions and the remaining SNPs (203) are all minor SNPs with their variant rates ranging from 10.64% to 38.05%. Therefore, there are only 4 SNP differences from the vaccine-derived consensus OmpB gene sequence as compared to the reference *R. rickettsii* (NC_010263.3) sequence, while at all these 207 called SNP sites, 200 of them are different between reference genomes of *R. rickettsii* and *R. rhipicephali*; 72 are different between *R. rickettsii* and *R. slovacica*; and 73 are different between *R. rickettsii* and *R. conorii* (Table 3, Supplemental Table 2).

A total of 21 SNPs were identified in the GltA region from the vaccine-derived DNA library and none of them has variant base numbers in which > 50% of total aligned read numbers. So all called SNPs at GltA gene region are all minor SNPs with their variant rates ranging from 14.40% to 31.99%. Therefore, there are no SNP differences from the vaccine-derived consensus GltA gene sequence as compared to the reference *R. rickettsii* (NC_010263.3) sequence, while all these 21 called SNP sites, 15 of them are different between reference genomes of *R. rickettsii* and *R. rhipicephali*; 5 are different between *R. rickettsii* and *R. slovacica*; and 6 are different between *R. rickettsii* and *R. conorii* (Table 3, Supplemental Table 2).

Rickettsial species	GenBank accession	Total reads	Aligned onetime reads	Aligned > 1 time reads	Total aligned reads	Overall alignment rate
<i>R. akari</i>	GCF_000018205.1	519,012,047	5,682,463	6631	5,689,094	1.10%
<i>R. asiatica</i>	GCF_007989425.1	519,012,047	7,149,491	38,500	7,187,991	1.38%
<i>R. australis</i>	GCF_000284155.1	519,012,047	6,295,260	16,953	6,312,213	1.22%
<i>R. bellii</i>	GCF_002078315.1	519,012,047	1,695,792	49,782	1,745,574	0.34%
<i>R. canadensis</i>	GCF_000283915.1	519,012,047	3,892,560	796	3,893,356	0.75%
<i>R. conorii</i>	GCF_000007025.1	519,012,047	10,296,428	54,786	10,351,214	1.99%
<i>R. monacensis</i>	GCF_000499665.2	519,012,047	7,336,041	58,005	7,394,046	1.42%
<i>R. prowazekii</i>	GCF_000277165.1	519,012,047	2,686,599	1268	2,687,867	0.52%
<i>R. rhipicephali</i>	GCF_000284075.1	519,012,047	10,322,074	62,102	10,384,176	2.00%
<i>R. rickettsii</i>	GCF_000017445.4	519,012,047	10,548,147	56,711	10,604,858	2.04%
<i>R. slovacica</i>	GCF_000237845.1	519,012,047	10,408,576	61,641	10,470,217	2.02%
<i>R. sp.MEAM1</i>	GCF_002285905.1	519,012,047	1,216,894	29,885	1,246,779	0.24%
<i>R. tillamookensis</i>	GCF_016743795.1	519,012,047	6,746,940	69,293	6,816,233	1.31%
<i>R. typhi</i>	GCF_000277285.1	519,012,047	2,596,733	1170	2,597,903	0.50%

Table 1. Aligned reads to complete rickettsial genomes.

<i>R. rickettsii</i> genome	Genome length	Mapped reads	Genome covered length	Genome covered rate	Genome average coverage
NC_010263.3	1,268,201	10,604,858	1,264,873	0.9973758	1021.58

Table 2. Alignment results of *R. rickettsii* genome.

	OmpA	OmpB	GltA
Total called SNP number compared to <i>R. rickettsii</i> reference	221	207	21
Bases are different between <i>R. rickettsii</i> and <i>R. rhipicephali</i>	175	200	15
Bases are different between <i>R. rickettsii</i> and <i>R. slovaca</i>	64	72	5
Bases are different between <i>R. rickettsii</i> and <i>R. conorii</i>	71	73	6
SNP matching <i>R. rhipicephali</i>	167	199	15
SNP matching <i>R. rhipicephali</i> only	109	126	10
SNP matching <i>R. slovaca</i>	61	66	5
SNP matching <i>R. slovaca</i> only	9	0	0
SNP matching <i>R. conorii</i>	62	66	6
SNP matching <i>R. conorii</i> only	2	0	1

Table 3. Results of SNP analysis at OmpA, OmpB, and GltA gene regions.

Hence, after analysis of called SNPs at OmpA, OmpB, GltA gene regions, the obtained consensus *Rickettsia* sequence is *R. rickettsii* as expected according to the method labeled on the vaccine bottle and used in RML vaccine preparations³².

Among these compared 175 SNP sites in OmpA region that are different between *R. rickettsii* and *R. rhipicephali*, 167 of called SNP bases based on our obtained sequences are the same as the bases in *R. rhipicephali* and 109 of them are unique to *R. rhipicephali*. Among these compared 64 sites in OmpA region that are different between *R. rickettsii* and *R. slovaca*, 61 of called SNP bases are the same as the bases in *R. slovaca* and only 9 of them are unique to *R. slovaca*. Among these compared 71 sites in OmpA region that are different between *R. rickettsii* and *R. conorii*, 62 of called SNP bases are the same as the bases in *R. conorii*, but only 2 of them are unique to *R. conorii* (Table 3, Supplemental Table 2).

Among these compared 200 SNP sites in OmpB region that are different between *R. rickettsii* and *R. rhipicephali*, 199 of called SNP bases based on our obtained sequences are the same as the bases in *R. rhipicephali* and 126 of them are unique to *R. rhipicephali*. Among these compared 72 sites in OmpB region that are different between *R. rickettsii* and *R. slovaca*, 66 of called SNP bases are the same as the bases in *R. slovaca*, but none of them are unique to *R. slovaca*. Among these compared 73 sites in OmpB region that are different between *R. rickettsii* and *R. conorii*, 66 of called SNP bases are the same as the bases in *R. conorii*, none of them are unique to *R. conorii* either (Table 3, Supplemental Table 2).

Among these compared 15 SNP sites in GltA region that are different between *R. rickettsii* and *R. rhipicephali*, all 15 of called SNP bases based on our obtained sequences are the same as the bases in *R. rhipicephali* and 10 of them are unique to *R. rhipicephali*. Among these compared 5 sites in GltA region that are different between *R. rickettsii* and *R. slovaca*, 5 of called SNP bases are the same as the bases in *R. slovaca*, but none of them are unique to *R. slovaca*. Among these compared 6 sites in GltA region that are different between *R. rickettsii* and *R. conorii*, 6 of called SNP bases are the same as the bases in *R. conorii*, but only 1 of them are unique to *R. conorii* (Table 3, Supplemental Table 2).

Therefore, after analysis of called SNPs at OmpA, OmpB, GltA gene regions with base differences among *R. rickettsii*, *R. rhipicephali*, *R. slovaca*, and *R. conorii*, this RMSF vaccine likely contains at least 10% of *R. rhipicephali*.

Bacterial metagenomic analysis. Bacterial metagenomic analysis was performed on trimmed, filtered and collapsed reads using Kraken2⁶⁰ and the result is shown in Fig. 3. It not only confirms that about 24% of the reads align to rickettsial genomes, but 65% of reads aligned to the *C. burnetii* genome, and 7% of the reads aligned to the *Arsenophonus nasoniae* (*A. nasoniae*) genome that is also carried by ticks^{61,62}. *Arsenophonus nasoniae* is a gram-negative gammaproteobacterial, secondary-endosymbiont that infects a wide range of insects and arachnids⁶³.

Sequence alignment to *Coxiella burnetii*. Because metagenomic analysis showed large portion of the bacterial reads belong to *C. burnetii*, all generated Illumina reads were trimmed and aligned to reference genome of *C. burnetii* RSA 493 (PRJNA57631) using Bowtie2⁵⁷. Among the total of 519,012,047 trimmed reads, 35,411,467 reads (6.86%) were aligned to the *C. burnetii* genome (34,704,131 (6.69%) aligned exactly 1 time and 707,336 (0.14%) aligned >1 times), which covered the *C. burnetii* genome and its plasmid with 100% with an average genome coverage of 1954 times (Table 4). Therefore, this RMSF vaccine also contains whole genomes of *C. burnetii*. Compared to the reference *C. burnetii* genome RSA 493 (PRJNA57631), there are a total of 273 SNPs (268 in the chromosome and 5 in the plasmid) identified SNPs (Materials and methods) with only 7 nonsynonymous changes (Supplemental Table 3). The consensus sequence of this obtained *C. burnetii* was submitted to GenBank with accession number of CP115461.

Sequence alignment to *Arsenophonus nasoniae*. Another bacterial group with significant match identified by Kracken2 was *Arsenophonus nasoniae* (*A. nasoniae*). All generated Illumina reads were trimmed and aligned to reference genome of *A. nasoniae* (PRJNA529362). Among the total of 519,012,047 trimmed reads,

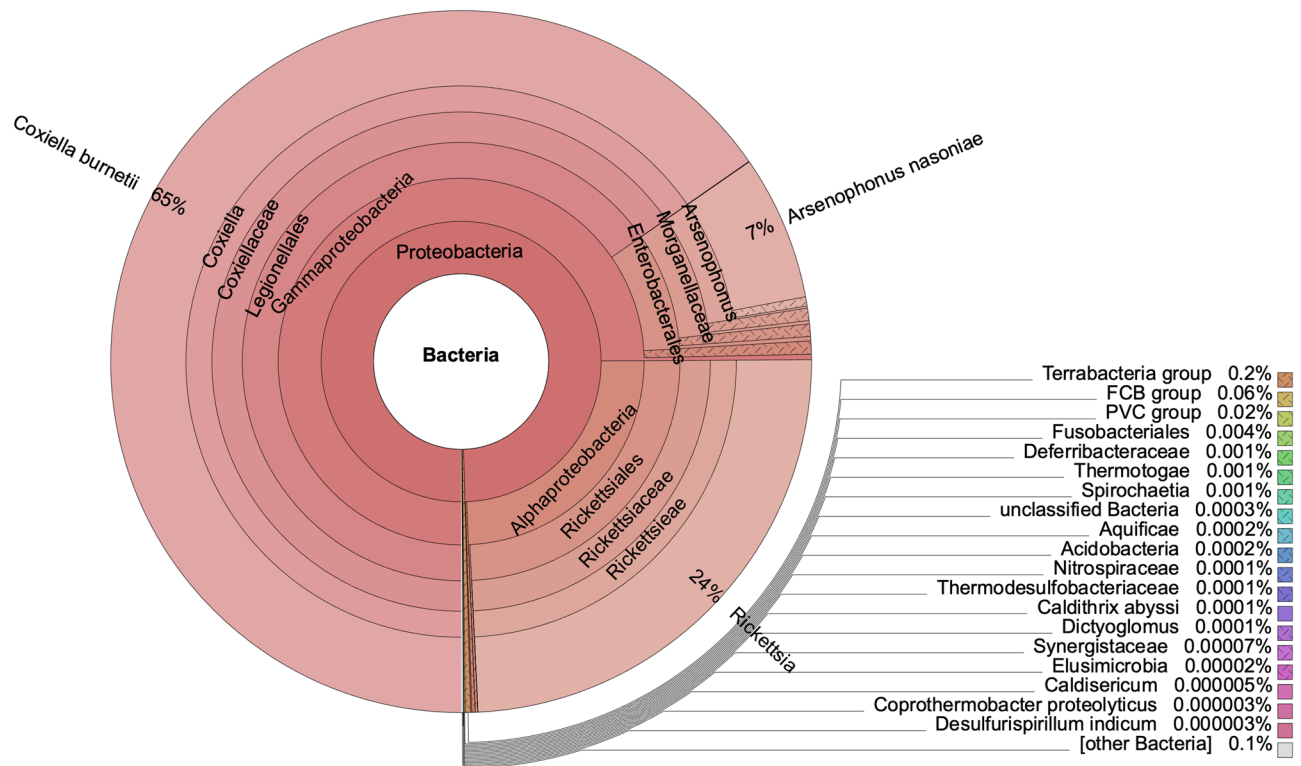


Figure 3. Metagenomics analysis of bacterial group in sequenced RMSF vaccine.

<i>C. burnetii</i> genome	Type	Genome length	Mapped reads	Genome covered length	Genome covered rate	Genome average coverage
NC_002971.4	Chromosome	1,995,488	34,776,864	1,995,488	1.0000000	1954
NC_004704.2	Plasmid	37,319	634,603	37,319	1.0000000	1954

Table 4. Alignment results of *C. burnetii* genome.

only 3,590,436 reads (0.69%) were aligned to *A. nasoniae* genome (3,297,383 (0.64%) aligned exactly 1 time and 293,053 (0.14%) aligned > 1 times), with a 32.9% coverage of the *A. nasoniae* genome with chromosome coverage of 40.8% (Table 5). This confirms that this RMSF vaccine contained a low level of *A. nasoniae* DNA.

Discussion

As shown on the label of the bottle, the 1944 vaccine was prepared from phenol-formalin inactivated infected tick tissues using the method of Spencer and Parker³² (Fig. 1). In the vaccine bottle, pieces of dark red material, as large as 2 mm squared, were visible floating in the solution and settling at the bottom. According to the original method, the precipitate that formed after adding phenol was separated by slow centrifugation and only the supernatant was used as a vaccine, which had a moderate turbidity³². Therefore, the dark red material that is visible in the vaccine currently probably represents tick tissue re-precipitation after 78 years. In our study, DNA was recovered only from the solid materials in the vaccine, not the solute, which also may help explain that the potency of the original vaccine was destroyed if passed through a Berkefeld filter³². In addition, because the final vaccine process is adding phenol to kill extraneous organisms, the pH value of the vaccine solution of 5.5 is explicable.

The first generation RMSF vaccine was made from crushed Rocky Mountain wood ticks (*D. andersoni*) that fed on *R. rickettsii*-infected guinea pigs³². Therefore, sequences from *D. andersoni* and *R. rickettsii* were expected and our sequence results confirm this. However, the unexpected result is that, within the bacterial metagenomics analysis, the bacterial genome with the most reads was not *Rickettsia*, but *C. burnetii*, the pathogen causing Q fever²⁷. In 1940 at RML, Cox and John E. Bell prepared 10 vaccines from adult *D. andersoni* that were infected with both Rocky Mountain spotted fever and Q fever pathogens. They showed complete protection of guinea pigs against at least 1000 infectious doses of spotted fever and against at least 10,000 doses of Q fever rickettsiae (the genus name *Coxiella* was not established at that time)⁵⁴. Although the label of this bottle of vaccine sequenced here does not mention Q fever, this vaccine is thus possibly one of these bivalent vaccines prepared from ticks harboring the two infectious agents simultaneously⁵⁴. It seems reasonable that a vaccine against both agents

<i>A. nasoniae</i>	Type	Genome length	Mapped reads	Genome covered length	Genome covered rate
NZ_CP038613.1	Chromosome	3,871,978	3,389,947	1,578,353	0.4076348
NZ_CP038614.1	Plasmid	222,851	85,105	19,898	0.0892884
NZ_CP038615.1	Plasmid	85,190	3341	2591	0.0304144
NZ_CP038616.1	Plasmid	85,274	6085	2486	0.0291531
NZ_CP038617.1	Plasmid	133,306	14,249	6201	0.046517
NZ_CP038618.1	Plasmid	61,872	14,176	4934	0.0797453
NZ_CP038619.1	Plasmid	121,472	4817	4282	0.0352509
NZ_CP038620.1	Plasmid	120,926	23,469	6750	0.0558193
NZ_CP038621.1	Plasmid	51,789	1602	1311	0.0253143
NZ_CP038622.1	Plasmid	50,328	1590	1261	0.0250556
NZ_CP038623.1	Plasmid	46,250	2742	1526	0.0329946
NZ_CP038624.1	Plasmid	34,725	34,938	5886	0.1695032
NZ_CP038625.1	Plasmid	33,626	8132	6648	0.1977042
NZ_CP038626.1	Plasmid	32,417	2	105	0.003239
NZ_CP038627.1	Plasmid	15,977	0	0	0
NZ_CP038628.1	Plasmid	8312	3	115	0.0138354
NZ_CP038629.1	Plasmid	3641	62	76	0.0208734
NZ_CP038630.1	Plasmid	7173	176	503	0.0701241

Table 5. Alignment results of *A. nasoniae* genome.

would be useful to laboratory personnel working with these bacteria at that time since RMSF was often a fatal disease and Q fever was highly infectious in laboratory settings.

Low levels of *R. rhipicephali* DNA were also identified in this vaccine. Since *R. rhipicephali* was first isolated from the brown dog tick (*Rhipicephalus sanguineus*) in Mississippi⁶⁴, it has been isolated in diverse tick genera (*Haemaphysalis juxtakochi*, *Ixodes ricinus*, *D. occidentalis*, *D. andersoni*, and *D. variabilis*) with wide geographic distribution^{65–67}. Although *R. rhipicephali* belongs to the spotted fever group *Rickettsia* as does *R. rickettsii*, it has not been identified as a human pathogen, only showing moderately severe disease in meadow voles inoculated with it⁶⁸. However, at least six of subspecies of spotted fever rickettsiae have been isolated from ticks, and later found to be pathogenic to humans⁶⁹. The first mixed infection of *R. belli*, *R. montanensis*, and *R. rickettsii* in one tick was reported in 2006⁷⁰. However, no ticks were found to be co-infected with *R. rhipicephali* and another spotted fever group *Rickettsia*⁷¹. In the current study, DNA sequence analysis suggests a mixed infection of *R. rickettsii* and *R. rhipicephali* in the infected ticks from which the vaccine was prepared. However, the possibility that different ticks infected with *R. rickettsii* and *R. rhipicephali* respectively were mixed in the vaccine preparation cannot be ruled out. In addition, the SNPs reported in Supplemental Table 1, especially the ones at low levels, probably reflect the mixture of *R. rickettsii* and *R. rhipicephali* in the RMSF vaccine.

We identified the bacterium *A. nasoniae* in this vaccine, which was first isolated from the parasitic wasp *Nasonia vitripennis* and is the causative agent of the son-killer trait in that species⁷². *Arsenophonus nasoniae* is a maternally-inherited parasitic bacterium that can cause lethality in approximately 80% of male embryos produced by infected female wasps⁷³. *Arsenophonus nasoniae* has been found in wide range of insects and arachnids⁷⁴, including ticks (*Dermacentor andersoni*, *Dermacentor variabilis*, and *Ixodes ricinus*)^{62,75,76}. Therefore, its presence in our study is not unexpected.

Lastly, avoiding DNA contamination during the whole process of the experiments was critical. Handling of the vaccine vial and processing the samples were always performed in biological safety cabinets. In addition, the laboratory where the extraction and sequencing occurred and laboratories in the same building have never conducted prior research on these bacteria. The lowest obtained reads that we reported in the RMSF vaccine has more than 3.3 million from *A. nasoniae*, which is highly unlikely from the material introduced during our experimental handling of the vaccine.

Conclusions

Recent studies of an historical 1902 vaccine used against smallpox have likewise revealed unexpected DNA results, e.g., that it did not contain the vaccinia virus, as expected, but the horsepox virus instead^{77,78}. Here, we have documented the history of another important vaccine by revealing that a 1944 RMSF vaccine prepared at RML contained *D. andersoni* and *R. rickettsii* DNA, which confirms its preparation from *R. rickettsii*-infected Rocky Mountain Wood ticks (*D. andersoni*). However, our work also reveals that it included significant amounts of *C. burnetii* DNA and low amount of *R. rhipicephali* DNA, which suggests it was a bivalent RMSF and Q-fever vaccine produced from ticks with a mixed infection of both *R. rickettsii* and *R. rhipicephali*. This is the first study applying modern high throughput sequencing technology to investigate an early RMSF vaccine made at RML and is thus represents an invaluable piece of history for the Rocky Mountain Laboratories and the National Institutes of Health.

Materials and methods

DNA isolation. Test DNA isolation was performed on 400 µl of well-shaken RMSF vaccine solution. Liquid and solid portions of the vaccine solution were obtained by centrifuge at 13,000 rpm at 4 °C for 10 min. DNA isolation from the liquid portion was performed using NucleoSpin cfDNA isolation kit from Takara Bio (San Jose, CA) following manufacturer's instructions. DNA isolation from the solid portion was performed using Zymo Quick-DNA™ Miniprep Plus Kit from Zymo Research (Irvine, CA) following manufacturer's instructions. Final volumes of isolated DNA were 10 µl and 1 µl from each isolation, and they were measured using the Agilent High Sensitivity DNA Kit (Agilent Technologies, Santa Clara, CA) and Qubit (Thermo Fisher Scientific, Waltham, MA). The DNA concentration from solid portion was 0.05–0.1 ng/µl and no DNA could be detected from the isolation of liquid portion by either method. Therefore, after centrifuging 2 × 1.5 ml vaccine solution at 13,000 rpm at 4 °C for 10 min, supernatants were put back into original RMSF vaccine bottle and precipitations were used to isolated DNA using Zymo Quick-DNA™ Miniprep Plus Kit. From this larger preparation, about 6 ng DNA was obtained for Illumina library construction.

Library construction and sequencing. Illumina sequencing library was made from isolated DNA by using NEBNext® Ultra™ II DNA Library Prep Kit from New England Biolabs (Ipswich, MA) following manufacture instructions. The final sequencing library was analyzed with the Agilent 2100 Bioanalyzer using the Agilent High Sensitivity DNA Kit (Agilent Technologies, Santa Clara, CA) (Fig. 2). The constructed library was diluted following Illumina sequencing standard protocols and sequenced on a NextSeq 500 sequencer with 160 bp single read run using NextSeq 500 High Output Kit v2.5 (Illumina, San Diego, CA).

Data analysis. Illumina reads were trimmed using Trimmomatic⁷⁹ and mapped to reference genomes including *Dermaecentor andersoni* (JALBCO000000000.1 and the access date is 07/28/2022), *Coxiella burnetii* (PRJNA57631 and the access date is 05/09/2022), *Arsenophonus nasoniae* (PRJNA529362 and the access date is 05/16/2022), and 14 complete *Rickettsia* genomes (GenBank accession numbers are in Table 1 and access date is 08/05/2022) that downloaded from NCBI using Bowtie2 (version 2.3.4.1) with default settings⁵⁷. *Rickettsia rickettsii* consensus sequence was generated based on reference *R. rickettsii* (CP000848.1). SAMtools mpileup (version 2.1.0)⁸⁰ was used to generate mpileup files that were subsequently used to make SNP calls by VarScan2⁸¹. Bacterial genome classification in all the samples was performed using Kraken2⁶⁰, utilizing its standard database on the trimmed using Trimmomatic with default settings⁷⁹, filtered using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) with -q 20 -p 80, and collapsed to remove duplicates using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) reads. A reported SNP call using VarScan2⁸¹ was the one that satisfied the following criteria at the SNP position: (1) more than 100 reads at that position, (2) minimum base Phred quality score as 25, (3) the different bases were more than 10% of the aligned reads, (4) pass VarScan2 Strand Filter, (5) VarScan2 SNP call p-value less than 0.05. Consensus nucleotide sequence was generated by VarScan2 with minimum coverage of 100 reads, minimum average quality value of 25, minimum variant frequency 0.5, and SNP call p-value less than 0.05.

Data availability

The datasets generated and/or analyzed during the current study are available in the NCBI SRA database with accession numbers: PRJNA880551, CP114277, and CP115461.

Received: 20 October 2022; Accepted: 20 March 2023

Published online: 22 March 2023

References

- Harden, V. A. *Rocky mountain spotted fever: History of a twentieth-century disease* Vol. 16 (Johns Hopkins University Press, 1990).
- Maxey, E. E. M., G T; Leary, J W. (1899) Some observations on the so-called spotted fever of Idaho. *Med Sentinel*, 7: 433–438.
- Dantas-Torres, F. Rocky mountain spotted fever. *Lancet Infect. Dis.* 7, 724–732. [https://doi.org/10.1016/S1473-3099\(07\)70261-X](https://doi.org/10.1016/S1473-3099(07)70261-X) (2007).
- Wilson, L. B. & Chowning, W. M. Studies in pyroplasmiasis hominis. ("Spotted fever" or "tick fever" of the rocky mountains.) (with map, charts, and plates I, li). *J. Infect. Dis.* 1, 31–57. <https://doi.org/10.1093/infdis/1.1.31> (1904).
- NIAID. *History of Rocky mountain labs (RML)*, <<https://www.niaid.nih.gov/about/rocky-mountain-history>> (2022).
- Ricketts, H. T. the study of "rocky mountain spotted fever" (tick fever?) By means of animal inoculations. A preliminary communication. *J. Am. Med. Assoc.* XLVII, 33–36. <https://doi.org/10.1001/jama.1906.25210010033001j> (1906).
- Knoop, F. C. *Reference module in biomedical sciences* (Elsevier, 2014).
- Shope, R. E. The epidemiology of the origin and perpetuation of a new disease. *Perspect. Biol. Med.* 7, 263–278. <https://doi.org/10.1353/pbm.1964.0039> (1964).
- Perlman, S. J., Hunter, M. S. & Zchori-Fein, E. The emerging diversity of *Rickettsia*. *Proc. Biol. Sci.* 273, 2097–2106. <https://doi.org/10.1098/rspb.2006.3541> (2006).
- Walker, D. H., Hudnall, S. D., Szaniawski, W. K. & Feng, H.-M. Monoclonal antibody-based immunohistochemical diagnosis of rickettsialpox: The macrophage is the principal target. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol.* 12, 529–533 (1999).
- Harrell, G. T. & Aikawa, J. K. Pathogenesis of circulatory failure in rocky mountain spotted fever: Alterations in the blood volume and the thiocyanate space at various stages of the disease. *Arch. Intern. Med.* 83, 331–347. <https://doi.org/10.1001/archinte.1949.00220320085007> (1949).
- Lacz, N. L., Schwartz, R. & Kapila, R. Rocky Mountain spotted fever. *J. Eur. Acad. Dermatol. Venereol.* 20, 411–417 (2006).
- Parola, P., Paddock, C. D. & Raoult, D. Tick-borne rickettsioses around the world: Emerging diseases challenging old concepts. *Clin. Microbiol. Rev.* 18, 719–756 (2005).
- Li, H. & Walker, D. H. rOmpA is a critical protein for the adhesion of *Rickettsia rickettsii* to host cells. *Microb. Pathog.* 24, 289–298 (1998).
- Hillman, R. D. Jr., Baktash, Y. M. & Martinez, J. J. OmpA-mediated rickettsial adherence to and invasion of human endothelial cells is dependent upon interaction with $\alpha 2\beta 1$ integrin. *Cell. Microbiol.* 15, 727–741 (2013).

16. Martinez, J. J., Seveau, S., Veiga, E., Matsuyama, S. & Cossart, P. Ku70, a component of DNA-dependent protein kinase, is a mammalian receptor for *Rickettsia conorii*. *Cell* **123**, 1013–1023 (2005).
17. Blanc, G. *et al.* Molecular evolution of rickettsia surface antigens: Evidence of positive selection. *Mol. Biol. Evol.* **22**, 2073–2083 (2005).
18. Sears, K. T. *et al.* Surface proteome analysis and characterization of surface cell antigen (SCA) or autotransporter family of *Rickettsia typhi*. (2012).
19. Ngwamidiba, M., Blanc, G., Raoult, D. & Fournier, P.-E. Sca 1, a previously undescribed paralog from autotransporter protein-encoding genes in *Rickettsia* species. *BMC Microbiol.* **6**, 1–11 (2006).
20. Pan, L., Zhang, L., Wang, G. & Liu, Q. Rapid, simple, and sensitive detection of the ompB gene of spotted fever group rickettsiae by loop-mediated isothermal amplification. *BMC Infect. Dis.* **12**, 254. <https://doi.org/10.1186/1471-2334-12-254> (2012).
21. Peniche-Lara, G., Zavala-Velazquez, J., Dzul-Rosado, K., Walker, D. H. & Zavala-Castro, J. Simple method to differentiate among *Rickettsia* species. *J. Mol. Microbiol. Biotechnol.* **23**, 203–208. <https://doi.org/10.1159/000348298> (2013).
22. Blanda, V. *et al.* New real-time PCRs to differentiate *Rickettsia* spp. and *Rickettsia conorii*. *Molecules* **25**(789), 963 (2020).
23. Barry, J. C. *Notable Contributions to Medical Research by Public Health Service Scientists: A Bibliography to 1940*. (U.S. Department of Health, Education, and Welfare, U.S. Public Health Service, 1960).
24. Burgdorfer, W. Ecology of tick vectors of American spotted fever. *Bull. World Health Organ* **40**, 375–381 (1969).
25. Parker, R. R., Spencer, R. R. & Francis, E. Tularaemia: XI tularaemia infection in ticks of the species dermacentor andersoni stiles in the bitterroot valley Mont. *Public Health Rep.* **39**, 1057–1073. <https://doi.org/10.2307/4577151> (1924).
26. Emmons, R. W. Ecology of Colorado tick fever. *Annu. Rev. Microbiol.* **42**, 49–64. <https://doi.org/10.1146/annurev.mi.42.100188.000405> (1988).
27. Davis, G. E., Cox, H. R., Parker, R. & Dyer, R. A filter-passing infectious agent isolated from ticks. *Public Health Rep.* **53**, 2259–2311 (1938).
28. Yunker, C., Keirans, J., Clifford, C. & Easton, E. Dermacentor ticks (Acari: Ixodidae: Ixodidae) of the new world: A scanning electron microscope atlas. *Proc. Entomol. Soc. Wash.* **88**, 609–627 (1986).
29. Anderson, J. F. & Laboratory, U. S. D. o. t. U. S. M.-H. S. H. *Spotted fever (tick fever) of the rocky mountains: A new disease*. (US Government Printing Office, 1903).
30. Bishopp, F. & Trembley, H. L. Distribution and hosts of certain North American ticks. *J. Parasitol.* **31**, 1–54 (1945).
31. James, A. M. *et al.* Distribution, seasonality, and hosts of the Rocky Mountain wood tick in the United States. *J. Med. Entomol.* **43**, 17–24 (2006).
32. Spencer, R. R. & Parker, R. R. Rocky mountain spotted fever: Vaccination of monkeys and man. *Public Health Rep.* **1896–1970**(40), 2159–2167. <https://doi.org/10.2307/4577679> (1925).
33. Cox, H. R. Rocky mountain spotted fever: Protective value for guinea pigs of vaccine prepared from *Rickettsiae* cultivated in embryonic chick tissues. *Public Health Rep.* **1896–1970**(54), 1070–1077. <https://doi.org/10.2307/4582917> (1939).
34. Lackman, D. & Parker, R. R. Comparison of the immunogenic and anaphylactogenic properties of rocky mountain spotted fever vaccines prepared from infected yolk sacs and from infected tick tissue. *Am. J. Public Health Nations Health* **38**, 1402–1404. <https://doi.org/10.2105/ajph.38.10.1402> (1948).
35. Parker, R. R. Rocky mountain spotted fever: Results of ten years' prophylactic vaccination. *J. Infect. Dis.* **57**, 78–93. <https://doi.org/10.1093/infdis/57.1.78> (1935).
36. DuPont, H. L. *et al.* Rocky Mountain spotted fever: A comparative study of the active immunity induced by inactivated and viable pathogenic *Rickettsia rickettsii*. *J. Infect. Dis.* **128**, 340–344. <https://doi.org/10.1093/infdis/128.3.340> (1973).
37. Kenyon, R. H., Acree, W. M., Wright, G. G. & Melchior, F. W. Jr. Preparation of vaccines for rocky mountain spotted fever from rickettsiae propagated in cell culture. *J. Infect. Dis.* **125**, 146–152. <https://doi.org/10.1093/infdis/125.2.146> (1972).
38. Kenyon, R. H. & Pedersen, C. E. Jr. Preparation of Rocky Mountain spotted fever vaccine suitable for human immunization. *J. Clin. Microbiol.* **1**, 500–503. <https://doi.org/10.1128/jcm.1.6.500-503.1975> (1975).
39. Osterloh, A. The neglected challenge: Vaccination against *Rickettsiae*. *PLoS Negl. Trop. Dis.* **14**, e0008704. <https://doi.org/10.1371/journal.pntd.0008704> (2020).
40. Sumner, J. W., Sims, K. G., Jones, D. C. & Anderson, B. E. Protection of guinea-pigs from experimental Rocky mountain spotted fever by immunization with baculovirus-expressed *Rickettsia rickettsii* rOmpA protein. *Vaccine* **13**, 29–35. [https://doi.org/10.1016/0264-410x\(95\)80007-z](https://doi.org/10.1016/0264-410x(95)80007-z) (1995).
41. Gong, W. *et al.* Enhanced protection against *Rickettsia rickettsii* infection in C3H/HeN mice by immunization with a combination of a recombinant adhesin rAdr2 and a protein fragment rOmpB-4 derived from outer membrane protein B. *Vaccine* **33**, 985–992. <https://doi.org/10.1016/j.vaccine.2015.01.017> (2015).
42. Wang, P. *et al.* Th1 epitope peptides induce protective immunity against *Rickettsia rickettsii* infection in C3H/HeN mice. *Vaccine* **35**, 7204–7212. <https://doi.org/10.1016/j.vaccine.2017.09.068> (2017).
43. CDC. *Rocky mountain spotted fever (RMSF)*, <<https://www.cdc.gov/rmsf/prevention/index.html>> (2022).
44. Alhassan, A. *et al.* *Rickettsia Rickettsii* whole-cell antigens offer protection against Rocky mountain spotted fever in the canine host. *Infect. Immun.* <https://doi.org/10.1128/IAI.00628-18> (2019).
45. Derrick, E. “Q” fever, a new fever entity: Clinical features, diagnosis and laboratory investigation. *Med J Aust* **2**, 281–299 (1937).
46. Burnet, F. M. & Freeman, M. Experimental studies on the virus of “Q” fever. *Med. J. Aust.* **2**, 299–305 (1937).
47. Shaw, E. I. & Voth, D. E. *Coxiella burnetii*: A Pathogenic Intracellular Acidophile. *Microbiology* **165**, 1–3. <https://doi.org/10.1099/mic.0.000707> (2019).
48. Seshadri, R. *et al.* Complete genome sequence of the Q-fever pathogen *Coxiella burnetii*. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 5455–5460. <https://doi.org/10.1073/pnas.0931379100> (2003).
49. Raoult, D., Marrie, T. & Mege, J. Natural history and pathophysiology of Q fever. *Lancet Infect. Dis.* **5**, 219–226. [https://doi.org/10.1016/S1473-3099\(05\)70052-9](https://doi.org/10.1016/S1473-3099(05)70052-9) (2005).
50. Eldin, C. *et al.* From Q fever to *Coxiella burnetii* infection: A paradigm change. *Clin. Microbiol. Rev.* **30**, 115–190 (2017).
51. Ormsbee, R., Peacock, M., Gerloff, R., Tallent, G. & Wike, D. Limits of rickettsial infectivity. *Infect. Immun.* **19**, 239–245. <https://doi.org/10.1128/iai.19.1.239-245.1978> (1978).
52. Scott, G. H., Williams, J. C. & Stephenson, E. H. Animal models in Q fever: Pathological responses of inbred mice to phase I *Coxiella burnetii*. *J. Gen. Microbiol.* **133**, 691–700. <https://doi.org/10.1099/00221287-133-3-691> (1987).
53. Madariaga, M. G., Rezai, K., Trenholme, G. M. & Weinstein, R. A. Q fever: A biological weapon in your backyard. *Lancet Infect. Dis.* **3**, 709–721. [https://doi.org/10.1016/s1473-3099\(03\)00804-1](https://doi.org/10.1016/s1473-3099(03)00804-1) (2003).
54. Cox, H. R. *Rickettsia diaporica* and American Q fever. *Am. J. Trop. Med. Hyg.* **20**, 463–469 (1940).
55. Smadel, J. E., Snyder, M. J. & Robbins, F. C. Vaccination against Q fever. *Am. J. Epidemiol.* **47**, 71–81. <https://doi.org/10.1093/oxfordjournals.aje.a119187> (1948).
56. Benenson, A. S. & Tigert, W. D. Studies on Q fever in man. *Trans. Assoc. Am. Phys.* **69**, 98–104 (1956).
57. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359. <https://doi.org/10.1038/nmeth.1923> (2012).
58. Murray, G. G., Weinert, L. A., Rhule, E. L. & Welch, J. J. The phylogeny of rickettsia using different evolutionary signatures: How tree-like is bacterial evolution?. *Syst. Biol.* **65**, 265–279. <https://doi.org/10.1093/sysbio/syv084> (2016).

59. Paris, D. H. *et al.* Real-time multiplex PCR assay for detection and differentiation of rickettsiae and orientiae. *Trans. R. Soc. Trop. Med. Hyg.* **102**, 186–193. <https://doi.org/10.1016/j.trstmh.2007.11.001> (2008).
60. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257. <https://doi.org/10.1186/s13059-019-1891-0> (2019).
61. Clayton, K. A., Gall, C. A., Mason, K. L., Scoles, G. A. & Brayton, K. A. The characterization and manipulation of the bacterial microbiome of the Rocky Mountain wood tick *Dermacentor andersoni*. *Parasit. Vectors* **8**, 632. <https://doi.org/10.1186/s13071-015-1245-z> (2015).
62. Bohacsova, M., Mediannikov, O., Kazimirova, M., Raoult, D. & Sekeyova, Z. Arsenophonus nasoniae and Rickettsiae Infection of Ixodes ricinus due to parasitic wasp Ixodiphagus hookeri. *PLoS ONE* **11**, e0149950. <https://doi.org/10.1371/journal.pone.0149950> (2016).
63. Trowbridge, R. E., Dittmar, K. & Whiting, M. F. Identification and phylogenetic analysis of Arsenophonus- and Photorhabdus-type bacteria from adult Hippoboscidae and Streblidae (Hippoboscoidea). *J. Invertebr. Pathol.* **91**, 64–68. <https://doi.org/10.1016/j.jip.2005.08.009> (2006).
64. Burgdorfer, W. *et al.* Rhipicephalus sanguineus: Vector of a new spotted fever group rickettsia in the United States. *Infect. Immun.* **12**, 205–210. <https://doi.org/10.1128/iai.12.1.205-210.1975> (1975).
65. Parola, P. *et al.* Update on tick-borne rickettsioses around the world: A geographic approach. *Clin. Microbiol. Rev.* **26**, 657–702. <https://doi.org/10.1128/cmr.00032-13> (2013).
66. Pascucci, I. *et al.* One health approach to rickettsiosis: A five-year study on spotted fever group rickettsiae in ticks collected from humans animals and environment. *Microorganisms* <https://doi.org/10.3390/microorganisms10010035> (2021).
67. Wikswo, M. E. *et al.* Detection and identification of spotted fever group rickettsiae in Dermacentor species from southern California. *J. Med. Entomol.* **45**, 509–516. [https://doi.org/10.1603/0022-2585\(2008\)45\[509:daiosf\]2.0.co;2](https://doi.org/10.1603/0022-2585(2008)45[509:daiosf]2.0.co;2) (2008).
68. Paddock, C. D. Rickettsia parkeri as a paradigm for multiple causes of tick-borne spotted fever in the western hemisphere. *Ann. N. Y. Acad. Sci.* **1063**, 315–326. <https://doi.org/10.1196/annals.1355.051> (2005).
69. Parola, P., Davoust, B. & Raoult, D. Tick- and flea-borne rickettsial emerging zoonoses. *Vet. Res.* **36**, 469–492. <https://doi.org/10.1051/vetres:2005004> (2005).
70. Carmichael, J. R. & Fuerst, P. A. A rickettsial mixed infection in a Dermacentor variabilis tick from Ohio. *Ann. N. Y. Acad. Sci.* **1078**, 334–337. <https://doi.org/10.1196/annals.1374.064> (2006).
71. Padgett, K. A. *et al.* The eco-epidemiology of Pacific coast tick fever in California. *PLoS Negl. Trop. Dis.* **10**, e0005020 (2016).
72. Gherna, R. L. *et al.* Arsenophonus nasoniae gen. nov., sp. nov., the causative agent of the son-killer trait in the parasitic Wasp Nasonia vitripennis. *Int. J. Syst. Evol. Microbiol.* **41**, 563–565. <https://doi.org/10.1099/00207713-41-4-563> (1991).
73. Balas, M. T., Lee, M. H. & Werren, J. H. Distribution and fitness effects of the son-killer bacterium in Nasonia. *Evol. Ecol.* **10**, 593–607 (2005).
74. Nováková, E., Hypsa, V. & Moran, N. A. Arsenophonus, an emerging clade of intracellular symbionts with a broad host distribution. *BMC Microbiol.* **9**, 143. <https://doi.org/10.1186/1471-2180-9-143> (2009).
75. Grindle, N., Tyner, J. J., Clay, K. & Fuqua, C. Identification of Arsenophonus-type bacteria from the dog tick Dermacentor variabilis. *J. Invertebr. Pathol.* **83**, 264–266 (2003).
76. Dergousoff, S. & Chilton, N. Detection of a new arsenophonus-type bacterium in Canadian populations of the rocky mountain wood tick Dermacentor Andersoni. *Exp. Appl. Acarol.* **52**, 85–91. <https://doi.org/10.1007/s10493-010-9340-5> (2010).
77. Schrick, L. *et al.* An early American smallpox vaccine based on horsepox. *N. Engl. J. Med.* **377**, 1491–1492. <https://doi.org/10.1056/NEJMc1707600> (2017).
78. Brinkmann, A., Souza, A. R. V., Esparza, J., Nitsche, A. & Damaso, C. R. Re-assembly of nineteenth-century smallpox vaccine genomes reveals the contemporaneous use of horsepox and horsepox-related viruses in the USA. *Genome Biol.* **21**, 286. <https://doi.org/10.1186/s13059-020-02202-0> (2020).
79. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170> (2014).
80. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352> (2009).
81. Koboldt, D. C. *et al.* VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576. <https://doi.org/10.1101/gr.129684.111> (2012).

Acknowledgements

This work was supported by the Intramural Research Programs of the NIH and NIAID.

Author contributions

Y.X. performed molecular experiments and drafted the article. Y.X. and P.A.B. analyzed and interpreted data. S.M.B. and M.E.B. initiated the project and provided the vaccine. Y.X., S.M.B., M.E.B., and J.K.T. conceived and designed experiments. P.A.B., S.M.B., D.M.M., M.E.B., and J.K.T. revised the article.

Funding

Open Access funding provided by the National Institutes of Health (NIH).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-31894-0>.

Correspondence and requests for materials should be addressed to Y.X.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2023