ORIGINAL ARTICLE

Comparison of genomes of *Coxiella burnetii* strains using formal order analysis

S. N. Shpynov[1], I. V. Tarasevich[1], A. A. Skiba[2], N. N. Pozdnichenko[2] and A. S. Gumenuk[2]

1) *N. F. Gamaleya FRCEM, Moscow, Russia and* 2) *OmSTU, Omsk, Russia*

Abstract

The *Coxiella burnetii* strain NL3262 was isolated during the Q fever outbreak in the Netherlands in 2007–2010. Formal-order analysis (FOA) was used to study the similarity of the genome (chromosome and plasmid) of this strain with the genomes from other strains. Chromosomes from ten *C. burnetii* strains and eight plasmids were studied with FOA tools such as 'Map of genes' and 'Matrix of similarity.' The 'Map of genes' tool showed that the chromosome of strain *C. burnetii* str. NL3262 distanced itself by the index of average remoteness (*g*) of 1.449640 (*x*-axis) from chromosomes of other strains (*g* 1.448295–1.448865). The 'Matrix of similarity' was used for an advanced analysis of the obtained results. The complete similarity of the components of chromosomes and plasmids was determined by pairwise comparison and the identification of nucleotides matching with them. A total of 84.90% of the chromosomal components of *C. burnetii* strain NL3262 coincided completely with the chromosomal components of strain Z3055. For chromosomes of other strains, this percentage varied from 12.06% to 47.14%. The plasmid of strain NL3262 had 50.0% of the components being completely coincident with the components of the plasmid of RSA 331; with RSA 493 it was 29.89%. Thus, *C. burnetii* str. NL3262 is the closest to str. Z3055 by the similarity of the chromosomal components, but on the index of average remoteness of the chromosome and the similarity of the plasmids' QpH1 components, it is the closest to strain RSA 331.

© 2018 The Author(s). Published by Elsevier Ltd.

Keywords: Average remoteness, *Coxiella burnetii*, Epidemiology, Formal order analysis, Genome, Internucleotide distance, 'Map of genes', 'Matrix of similarity', Q fever

Original Submission: 21 November 2017; **Revised Submission:** 27 January 2018; **Accepted:** 6 February 2018

Article published online: 27 February 2018

Corresponding author: S.N. Shpynov, 18, N.F. Gamaleya str., Moscow, 123098 Russia.
E-mail: stanislav63@yahoo.com

Introduction

Q fever is a widespread zoonotic disease caused by *Coxiella burnetii* [1]. Acute Q fever typically arises from inhalation of aerosolized bacteria [2]. Rare but potentially severe chronic disease most commonly manifests as endocarditis [1,3–5]. Goats, sheep and cattle are the primary reservoirs of *C. burnetii,* but strains have been obtained from a large variety of wild vertebrates and arthropods [1,6]. In sheep and goats, massive proliferation of the organism in the female reproductive system can result in late-term abortion. Parturition by infected mammals can consequently spread tremendous numbers of *C. burnetii* into the environment [1]. From 2007 to 2010, the Netherlands has been confronted with the largest global Q fever outbreak ever, involving 4026 human cases. It was shown that the strains responsible for this outbreak, NL3262 and Z3055, are clonal, as they both contain the QpH1 plasmid, have the same MST33 genotype and have the same VNTR profile [7]. The *C. burnetii* strain NL3262 was isolated from an aborted goat placenta during this outbreak [8].

Multiple-locus variable-number tandem repeat analysis revealed the predominant presence of the so-called CbNL01 genotype, and to a lesser extent the CbNL12 genotype, among the strains isolated in the Netherlands [9]. A recent phylogenetic analysis detected four clades (indicated as 1a, 1b, 2 and 3) among *C. burnetii* strains [10]: clade 1, CbNL01 genotypes 1a (including NL3262) and 1b; clade 2, CbNL12 genotype and NM-like

© 2018 The Author(s). Published by Elsevier Ltd
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
https://doi.org/10.1016/j.nmni.2018.02.011

genotype (NMI, RSA 331, etc.), representing the genome of strain Z3055, which is equidistant from clusters 1 and 2; and clade 3, scurry genotype (plasmid-less strains: CbuG_Q212), with nonclustered strains for each of the different genotypes (Schperling strain, Dugway 5J108-111, CbuK_Q154, MSU Goat Q177, etc.). It is established that *C. burnetii* strains with different genotypic profiles can infect a variable range of host species with distinct efficiency. CbNL01 genotype strains are predominantly found in goats and in humans, whereas the CbNL12 genotype strains are commonly found in cattle and hardly in goats and humans [9,11,12]. This suggests a higher susceptibility of humans and goats to strains from the CbNL01 genotype than to the CbNL12 genotype.

A new method, formal order analysis (FOA) [13], was recently proposed to classify representatives of the *Rickettsiaceae* family [14]. We believe that FOA may be well suited to study the genomic relationships of the NL3262 strain that has caused the Dutch Q fever outbreak with other *C. burnetii* strains.

Materials and methods

Genome sequences of *Coxiella burnetii* strains

The genomic sequences used in this study are presented in Table 1. All reference genomes were imported from GenBank (https://www.ncbi.nlm.nih.gov/genome). Chromosomes from ten *C. burnetii* strains (Dugway 5J108-111; CbuK_Q154; Z3055; RSA 493; RSA 439; RSA439 clone 4; CbuG_Q212; RSA 331; NL3262; MSU Goat Q177) and eight plasmids (Dugway 5J108-111 pQpDG; CbuK_Q154 pQpRS; RSA 493 pQpH1; RSA 439 pQpH1; RSA439 clone 4 pQpH1; RSA 331 QpH1; NL3262 QpH1; MSU Goat Q177 pQpRS) were studied.

FOA tools

FOA was used to analyse the genomes of *C. burnetii* strains as described previously [13–15]. The foundations of this approach were developed for the study of symbolic sequences and were successfully applied earlier in the research of linguistic and musical texts [16]. The work uses a high-precision and unambiguous numerical representation of the original arrangement of nucleotides in the sequence. In order to do this, numerical characteristics (average remoteness, depth, etc.) based on intersymbol intervals (internucleotide distance [17,18]) were developed [13].

Recently FOA has been improved with the new tools 'Map of genes' (MG) [19] and 'Matrix of similarity' (MS) [20] to permit a more in-depth study of rickettsial genome structure.

Pairs of numeric values of order characteristics from studied genomes and their components $\{<g_i, G_i>\}$ are mapped into pillars of dots on the MG. Components representing individual genomes are placed vertically, and some horizontal lines are formed with similar components in different genomes. The MG tool kit allows one to obtain interactively a detailed description of any component of the genome. Automated identification of similar components is also possible.

The MS represents the similarity values for each pair of analysed genomes. Genomic similarity is determined by comparing the order characteristic values from their components. The MS tool kit enables obtain interactively a list of only similar components of any pair of genomes, and when necessary, sliding window of order characteristics for those components can be obtained. The latter enables visualization of element-by-element similarity of genomic components.

A comparison of coding and noncoding sequences in genomes of different strains of *C. burnetii* sp. is possible using the MG, which can also be useful for differentiation of genes with 100% homology from orthologs. The MG tool identifies inter-strain differences in genomes belonging to *C. burnetii* species. The MG tool describes complete genomes and their components in a graphical form. Genomes are placed (classified) according to the index g on the x-axis, and their components (coding and noncoding sequences) are sited by depth index (G)

TABLE 1. Genome features of sequenced and average remoteness of *Coxiella burnetii* strains

No.	*Coxiella burnetii* strains	Clade	Source	GenBank accession no.	Average remoteness (g)	Genome size (bp)	G+C%	Gene count	Plasmid		IS110 copy no.
									Type	Size (bp)	
1	NL3262	1a	Goat		1.4496407330	2 093 477	42.84	2401	QpH1	37 320	106
2	CbRSA331	2	Human	NC_010117.1	1.4488653471	2 016 427	42.74	2272	QpH1	37 317	44
3	CbuG_Q212	3	Human	NC_011527.1	1.4487365584	2 008 870	42.60	2208	integrated		28
4	NMRSA493 (phase I)	2	Tick	NC_002971.4	1.4485974740	1 995 281	42.64	2085	QpH1	37 319	1
5	Z3055	1/2	Sheep	NZ_LK937696.1	1.4484246347	1 995 463	42.60	2197	QpH1		11
6	NMRSA 439 (phase II)	2	Cell culture	NZ_CP020616.1	1.4484206459	1 969 224	42.64	2217	pQpH1	37 319	20
7	NMRSA439 (phase II, clone 4)	2	Human	NZ_CP018005.1	1.4484095625	1 969 245	42.64	2219	pQpH1	37 319	20
8	MSU Goat Q177	a	Goat	NZ_CP018150.1	1.4483845927	2 090 565	42.64	2305	QpRS	39 281	48
9	CbuK_Q154	a	Human	NC_011528.1	1.4483811070	2 063 100	42.64	2302	QpRS	39 280	36
10	Dugway 5J108-111	a	Rodents	NC_009727.1	1.4482959721	2 158 758	42.34	2358	QpDG	54 179	2

[a]Nonclustered strains are different genotypes (clade 4) [10].

© 2018 The Author(s). Published by Elsevier Ltd, *NMNI*, **23**, 86–92
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

on the *y*-axis. Sequences that are completely similar by characteristic are 100% homologous. Strains of the same species show a high degree of homology of their components. It is necessary to introduce a criterion (characteristic) for component identification with a degree of homology <100% in order to compare genomes of different species. Using this approach, it is possible to conduct a selective analysis of all components of genomes individually or by grouping them by feature. Each component can be identified (visualized) among all compared organisms by its name in the annotation, which allows checking its presence in each genome. FOA software is available online (http://foarlab.org/).

Results

The application of the MG tool enabled comparison of the sequence organization of the chromosomes from *C. burnetii* strains according to the index of the average remoteness on the *x*-axis, ranging from 1.4482959721 for the Dugway 5J108-111 strain to 1.4496407330 for the NL3262 strain (Fig. 1). The analysis of annotated components (coding and noncoding) of chromosomes ranked according to depth index *G* (*y*-axis) showed high percentages (more than 80%) of the genomes of the *C. burnetii* Z3055 and NL3262 strains having complete (100%) nucleotide sequence similarity. One hundred six copies of IS110 family transposases were detected in the chromosome of strain NL3262, versus 11 in strain Z3055, and from 0 (RSA 493) to 48 (MSU Goat Q177) in other studied strains (Table 1).

The MS tool was used for extended analysis. The complete similarity of the components of chromosomes and plasmids was determined by their pairwise comparison. Strain NL3262 exhibited the highest percentage of homologous components (84.88%) with strain Z3055, followed by strains RSA 331 (47.13%), RSA 439 (33.54%), RSA439 clone 4 (33.00%) and RSA 493 (32.22%) (Table 2). Strain RSA 493 had a high percentage of homology with its clones RSA439 phase II clone 4 (86.89%) and *C. burnetii* RSA 439 phase II (85.56%). However, the percentage of homologous components between strains RSA 439 phase II clone 4 and RSA 439 was even higher, at 98.16%. In contrast, strain Dugway 5J108-111, isolated from the blood of a chisel-toothed kangaroo rat (*Dipodomys microps*), and strain CbuG_Q212, isolated from a patient with endocarditis and not carrying a plasmid but containing an integrated plasmid in the genome, were the most dissimilar. Strain CbuK_Q154, isolated from a patient with endocarditis, and strain MSU Goat Q177, isolated from a goat, both of which contain a QpRS plasmid, have a high percentage of homologous chromosomic components (76.55%).

The QpH1 plasmid of strain NL3262 exhibited a 50.0% MS value with that of strain RSA 331 and 29.89% with strain RSA 493. (For the RSA 439 and RSA439 clone 4 strains, it was 29.55% and 28.89, respectively.) With the other strains, values

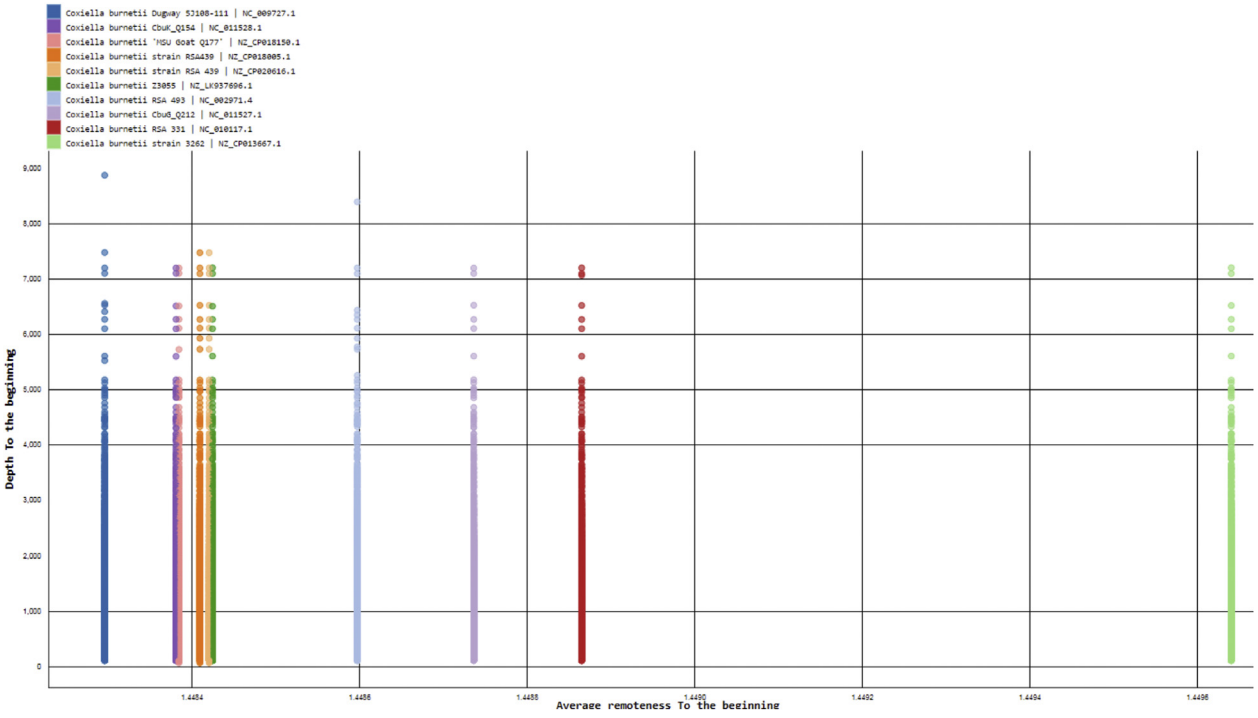

FIG. 1. Respective positions of *Coxiella burnetii* strains using 'Map of genes' tool with application of average remoteness (*x*-axis).

© 2018 The Author(s). Published by Elsevier Ltd, *NMNI*, **23**, 86–92
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

TABLE 2. Homology degree of the chromosomic components of *Coxiella burnetii* strains using 'Matrix of similarity' tool

Strains of *C. burnetii* | GenBank accession no.	Dugway 5J108-111 | NC_009727.1	CbuK_Q154 | NC_011528.1	'MSU goat Q177' | NZ_CP018150.1	RSA439 clone 4 NZ_CP018005.1	RSA 439 | NZ_CP020616.1	Z3055 | NZ_LK937696.1	RSA 493 | NC_002971.4	CbuG_Q212 | NC_011527.1	RSA 331 | NC_010117.1	NL3262 | NZ_CP013667.1
Dugway 5J108-111 | NC_009727.1	100.00%	22.20%	18.73%	19.98%	19.89%	17.60%	20.44%	18.11%	18.93%	17.15%
CbuK_Q154 | NC_011528.1	22.20%	100.00%	76.55%	15.45%	15.36%	14.14%	15.73%	15.01%	14.94%	13.66%
'MSU Goat Q177' | NZ_CP018150.1	18.73%	76.55%	100.00%	18.58%	18.63%	12.12%	17.79%	13.00%	12.91%	12.06%
RSA439 clone 4 NZ_CP018005.1	19.98%	15.45%	18.58%	100.00%	98.16%	33.95%	86.89%	18.71%	37.33%	33.54%
RSA 439 | NZ_CP020616.1	19.89%	15.36%	18.63%	98.16%	100.00%	33.49%	85.56%	18.58%	36.84%	33.00%
Z3055 | NZ_LK937696.1	17.60%	14.14%	12.12%	33.95%	33.49%	100.00%	32.87%	16.43%	46.15%	84.90%
RSA 493 | NC_002971.4	20.44%	15.73%	17.79%	86.89%	85.56%	32.87%	100.00%	19.05%	36.05%	32.24%
CbuG_Q212 | NC_011527.1	18.11%	15.01%	13.00%	18.71%	18.58%	16.43%	19.05%	100.00%	17.40%	16.00%
RSA 331 | NC_010117.1	18.93%	14.94%	12.91%	37.33%	36.84%	46.15%	36.05%	17.40%	100.00%	47.14%
NL3262 | NZ_CP013667.1	17.15%	13.66%	12.06%	33.54%	33.00%	84.90%	32.24%	16.00%	47.14%	100.00%

ranged from 5.56% to 6.74% (Table 3). Plasmids QpRS from strains CbuK_Q154 and MSU Goat Q177 had the highest percentage of homologous components (95.75%). Plasmid QpH1 from strain RSA 493 exhibited MS values of 48.35% and 73.12% with those of its clones RSA 439 and RSA439 clone 4, respectively.

Discussion

The application of the MG analysis showed that the genome from strain NL3262 (*g* 1.4496407330) with genotypic profile CbNL01 significantly differed from those of other studied strains (*g* 1.4482959721–1.4488653471) in structure according to the index of average remoteness. This difference may be caused by the highest number of IS110 elements among all genomes (n = 106; Table 1). The closest genome, that of strain CbRSA 331 (*g* 1.4488653471), had 44 IS110 elements. Sixty-five of the 106 IS110 elements in strain NL3262 strain were 1020 bp in length. Of these, ten had complete homology with 41 IS110 elements from strain RSA 331, and 20 had 100% homology with two IS110 elements from strain Z3055. Sixteen IS110 elements from strain NL3262 with a length of 1095 bp had complete homology with five elements from strain RSA 439 (clone 4) and two from strain RSA 439.

It was confirmed that the Z3055 strain (isolated from a sheep aborted placenta in Germany in 1992), based on the *g* index, actually belongs to the group of clones of strain RSA 493 and was localized between strains RSA 439, RSA439 clone 4 and RSA 493. This was confirmed by the absence of rearrangements in the genomes of strains Z3055 and NMI [10]. Comparative studies of complete genomes of *C. burnetii* showed rearrangements at 0, 21, 6 and 13 chromosomal locations in NMI compared to strains Z3055, Cbuk_Q154, CbuG_Q212 and Dugway 5J108-111, respectively [21]. A recent analysis also showed genomic rearrangements of strain NL3262 compared to strains Z3055 (21 locations), Cbuk_Q154 (31 locations), CbuG_Q212 (21 locations) and Dugway 5J108-111 (23 locations) [10]. It is considered that a large number of transposons in outbreak strains might have resulted in extensive genome rearrangements in response to environmental challenges, including in the host. Such rearrangements can result in DNA insertions or deletions, generation of polymorphisms or pseudogenes (genes disrupted by IS elements) or modulation of the gene expression by gene order rearrangements favouring its growth and survival in different niches [10,21–24].

The application of the MS tool enables establishing that strains NL3262 and Z3055 have the highest percentage (84.8%) of homologous components (genes and noncoding sequences). At the same time, a large proportion of the remaining

© 2018 The Author(s). Published by Elsevier Ltd, *NMNI*, **23**, 86–92
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

TABLE 3. Homology degree of the plasmidic components of *Coxiella burnetii* strains using 'Matrix of similarity' tool

Strains/plasmid *C. burnetii* | GenBank accession no.	RSA 493 plasmid pQpH1 | NC_004704.2	RSA 439 plasmid QpH1 | NZ_CP020617.1	RSA439 clone 4 plasmid QpH1 | NZ_CP018005.1	NL3262 plasmid QpH1 | NZ_CP013668.1	RSA 331 plasmid QpH1 | NC_010115.1	CbuK_Q154 plasmid pQpRS_K_Q154 | NC_011526.1	MSU goat Q177 plasmid QpRS | NC_010258.1	Dugway 5J108-111 plasmid pQpDG | NC_009726.1
RSA 493 plasmid pQpH1 | NC_004704.2	100.00%	48.35%	73.12%	29.89%	34.48%	10.87%	10.87%	7.21%
RSA 439 plasmid QpH1 | NZ_CP020617.1	48.35%	100.00%	76.60%	29.55%	43.18%	8.60%	8.60%	8.93%
RSA439 clone 4 plasmid QpH1 | NZ_CP018005.1	73.12%	76.60%	100.00%	28.89%	35.56%	10.53%	10.53%	8.77%
NL3262 plasmid QpH1 | NZ_CP013668.1	29.89%	29.55%	28.89%	100.00%	50.00%	6.74%	6.74%	5.56%
RSA 331 plasmid QpH1 | NC_010115.1	34.48%	43.18%	35.56%	50.00%	100.00%	8.99%	8.99%	7.41%
CbuK_Q154 plasmid pQpRS_K_Q154 | NC_011526.1	10.87%	8.60%	10.53%	6.74%	8.99%	100.00%	95.75%	8.85%
MSU Goat Q177 plasmid QpRS | NC_010258.1	10.87%	8.60%	10.53%	6.74%	8.99%	95.75%	100.00%	8.85%
Dugway 5J108-111 plasmid pQpDG | NC_009726.1	7.21%	8.93%	8.77%	5.56%	7.41%	8.85%	8.85%	100.00%

nonhomologous genomic fragments is explained by the difference in 95 IS110 elements (Table 1). All genomes from other strains differ significantly from those of strains NL3262 (genotype CbNL01: clade 1a) and Z3055 (genotype CbNL01: clade 1b) due to genes containing single nucleotide polymorphisms. At the same time, using the MS tool, it was found that strains Z3055 and RSA 493 have only 32.89% homologous genes (Table 2). The results obtained in this study are confirmed by recently published research data showing that differences between strains Z3055 and the reference NMI (RSA 493) genome are mostly point mutations and insertions or deletions (indels). Nonsynonymous mutations were increased significantly in coding genes for membrane proteins (16/156 vs. 103/1757), ankyrin repeat domains containing proteins (2/9 vs. 117/1904), transcription factors (7/53 vs. 112/1860) and translation proteins (15/144 vs. 109/1655) [7]. The evolution of this strain may have been driven by mutations in critical genes. Possibly the MS index of complete homology is as informative as the index of orthologous genes when comparing chromosomes. This index takes into account point mutations, which are important differences between Z3055 and the reference NMI which consisted of point mutations and insertions or deletions [7].

Thus, it is more probable that strain NL3262 (and others) which caused the Q fever outbreak in the Netherlands in 2007–2010 belongs to the same clone as strain Z3055 because they have a large number of homologous sequences in their genomes. Additionally, a significant reorganization of the genome took place in strain NL3262 (and other strains from the Dutch outbreak) due to a large increase in the number of IS110 elements, which could be the reason for an increase of virulence and/or spread.

At the same time, the percentage of genome homology between clones of one strain NMI phase I (RSA 493) (tick) and NM phase II (RSA439 from human) was 86.89%. The percentage of homologous components between the genomes of the clones of NMII RSA 439 and NMII RSA 439 (cell culture) themselves was 98.16%. NMII has the ~26 kb chromosomal deletion that eliminates several LPS biosynthetic genes and is associated with the production of a severely truncated LPS [25]. This indicates on the one hand the close relationships (high homology of the genomes) between the strains isolated from the goat and the sheep, and on the other hand, this indicates a significant reorganization of the genome during the transition of the strain to phase II.

An important aspect of the study of the MS index is the significant difference between NL3262 and another goat strain, MSU Goat Q177, with the percentage of homologous sequences being as low as 12.06%. This indicates the difference in the origin of the two isolates, in contrast to the similarity of the animal and human strains isolated during the outbreak in the Netherlands and strain Z3055 isolated from the sheep in Germany [7,8].

© 2018 The Author(s). Published by Elsevier Ltd, *NMNI*, **23**, 86–92
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

The increase in synonymic recombinations (compared to strain Z3055) associated with the increase in IS110 may have contributed to the increase in virulence of *C. burnetii* strains that caused the outbreak in the Netherlands. The increase in the number of IS110 leads to a reorganization of the genome (causing a significant increase in the collinear blocks), which we detected using a gene map with an average remoteness measure. At the same time, the application of the similarity matrix makes it possible to show that the number of homologous sequences in NL3262 and Z3055 significantly exceeds (by 1.8 times or more) those of other of *C. burnetii* strains.

Our classification completely supports the phylogenetic relationships among sequenced *C. burnetii* genomes based on single nucleotide polymorphism analysis [10] except for the position of a plasmid-less strain (CbuG_Q212), in which the plasmid is integrated in the chromosome.

We propose a comprehensive approach based on the application of two FOA tools (MG and MS) developed by our group as an additional tool for analysis of *C. burnetii* genomes for the purpose of studying the origin of outbreak strains and the epidemiologic significance of different strains.

Conclusions

The *C. burnetii* strain NL3262 isolated from a goat during the outbreak in the Netherlands is the closest to strain Z3055 isolated from a sheep in Germany according to the similarity of the chromosomal components (MS index). But by the index of average remoteness (g) of the chromosome, as well as the MS index of the QpH1 plasmid, strain NL3262 is closer to strain RSA 331. However, the genetic differences between the chromosomes of the two strains are minimal in comparison with eight other *C. burnetii* strains.

It will be necessary to sequence more strains isolated from different sources, and primarily from humans, in order to elucidate the origin of the strains that caused the massive recent outbreak in the Netherlands.

Conflict of interest

None declared.

References

[1] Maurin M, Raoult D. Q fever. Clin Microbiol Rev 1999;12:518–53.
[2] Tissot-Dupont H, Raoult D. Q fever. Infect Dis Clin N Am 2008;22:505–14.
[3] Raoult D, Etienne J, Massip P, Iaocono S, Prince MA, Beaurain P, et al. Q fever endocarditis in the south of France. J Infect Dis 1987;155:570–3.
[4] Raoult D, Marrie T, Mege J. Natural history and pathophysiology of Q fever. Lancet Infect Dis 2005;5:219–26.
[5] Parker NR, Barralet JH, Bell AM. Q fever. Lancet 2006;367(9511):679–88.
[6] Arricau Bouvery N, Souriau A, Lechopier P, Rodolakis A. Experimental *Coxiella burnetii* infection in pregnant goats: excretion routes. Vet Res 2003;3:423–33.
[7] D'Amato F, Rouli L, Edouard S, Tyczka J, Million M, Robert C, et al. The genome of *Coxiella burnetii* Z3055, a clone linked to the Netherlands Q fever outbreaks, provides evidence for the role of drift in the emergence of epidemic clones. Comp Immunol Microbiol Infect Dis 2014;37:281–8.
[8] Kuley R, Smith HE, Janse I, Harders FL, Baas F, Schijlen E, et al. First complete genome sequence of the Dutch veterinary *Coxiella burnetii* strain NL3262, originating from the largest global Q fever outbreak, and draft genome sequence of its epidemiologically linked chronic human isolate NLhu3345937. Genome Announc 2016;4. e00245–16.
[9] Roest HI, Ruuls RC, Tilburg JJ, Nabuurs-Franssen MH, Klaassen CH, Vellema P, et al. Molecular epidemiology of *Coxiella burnetii* from ruminants in Q fever outbreak, the Netherlands. Emerg Infect Dis 2011;17:668–75.
[10] Kuley R, Kuijt E, Smits MA, Roest HIJ, Smith HE, Bossers A. Genome plasticity and polymorphisms in critical genes correlate with increased virulence of Dutch outbreak-related *Coxiella burnetii* strains. Front Microbiol 2017;8:1526.
[11] Tilburg JJ, Roest HJ, Buffet S, Nabuurs-Franssen MH, Horrevorts AM, Raoult D, et al. Epidemic genotype of *Coxiella burnetii* among goats, sheep, and humans in the Netherlands. Emerg Infect Dis 2012;18:887–9.
[12] Mori M, Boarbi S, Michel P, Bakinahe R, Rits K, Wattiau P, et al. *In vitro* and *in vivo* infectious potential of *Coxiella burnetii*: a study on Belgian livestock isolates. PLoS One 2013;8:e67622.
[13] Gumenyuk AS, Pozdnichenko N, Rodionov I, Shpynov SN. About the means of formal analysis system of nucleotide chains [in Russian]. Mat Bio Bioinform 2013;8:373–97.
[14] Shpynov S, Pozdnichenko N, Gumenuk A. Approach for classification and taxonomy within family *Rickettsiaceae* based on the formal order analysis. Microbe Infect 2015;17:839–44.
[15] Shpynov SN, Gumenuk AS, Pozdnichenko NN. Application of the numerical characteristic of formal order analysis of the prokaryotic genomes for reclassification within the genus *Rickettsia* [in Russian]. Mat Bio Bioinform 2016;11:336–50.
[16] Gumenyuk A, Kostyshin A, Simonova S. An approach to the research of the structure of linguistic and musical texts. Glottometrics 2002;3:61–9.
[17] Afreixo V, Bastos CA, Pinho AJ, Garcia SP, Ferreira PJ. Genome analysis with inter-nucleotide distances. Bioinformatics 2009;25:3064–70.
[18] Achuth S, Nair Achuthsankar S, Mahalakshmi T. Visualization of genomic data using inter-nucleotide distance signals. In: Proceedings of IEEE genomic signal processing. Bucharest: Romania; 2005.
[19] Gumenuk AS, Pozdnichenko NN, Skiba AA, Shpynov SN. Computer program 'Map of genes' [in Russian]. Certificate of State Registration of the Computer Program in the Register of Computer Programs 2017616730, 13.06. 2017.
[20] Gumenuk AS, Pozdnichenko NN, Skiba AA, Shpynov SN. Computer program 'Matrix of similarity of nucleotide sequences by their components' [in Russian]. Certificate of State Registration of the Computer Program in the Register of Computer Programs 2017616679, 09.06. 2017.
[21] Beare PA, Unsworth N, Andoh M, Voth DE, Omsland A, Gilk SD, et al. Comparative genomics reveal extensive transposon-mediated genomic

© 2018 The Author(s). Published by Elsevier Ltd, *NMNI*, **23**, 86–92
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

plasticity and diversity among potential effector proteins within the genus *Coxiella*. Infect Immun 2009;77:642–56.

[22] Hall G. Is the occurrence of some spontaneous mutations directed by environmental challenges? New Biol 1991;3:729–33.

[23] Seshadri R, Paulsen IT, Eisen JA, Read TD, Nelson KE, Nelson WC, et al. Complete genome sequence of the Q-fever pathogen *Coxiella burnetii*. Proc Natl Acad Sci U S A 2003;100:5455–60.

[24] Pallen MJ, Wren BW. Bacterial pathogenomics. Nature 2007;449:835–42.

[25] Millar JA, Beare PA, Moses AS, Martens CA, Heinzen RA, Raghavan R. Whole-genome sequence of *Coxiella burnetii* Nine Mile RSA439 (phase II, clone 4), a laboratory workhorse strain. Genome Announc 2017;5.e00471–17.

© 2018 The Author(s). Published by Elsevier Ltd, *NMNI*, **23**, 86–92
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).