Part 2. Biotechnology

UDC 619:616.98:578.828.5:578.2'21:636.22/.28 DOI 10.36016/JVMBBS-2024-10-2-3

IDENTIFICATION OF INTRAMOLECULAR CONSERVED G-QUADRUPLEX MOTIFS IN THE GENOME OF THE BOVINE FOAMY VIRUS

Balak O. K. [1], Balak S. O. [2], Lymanska O. Yu. [3]

[1] Kharkiv National Medical University, Kharkiv, Ukraine
[2] Institute for Problems of Cryobiology and Cryomedicine of the National Academy of Sciences of Ukraine, Kharkiv, Ukraine
[3] National Scientific Center 'Institute of Experimental and Clinical Veterinary Medicine', Kharkiv, Ukraine, e-mail: olgaliman@ukr.net

Summary. G-quadruplexes (G4s) are guanine-rich DNA structures, which play an essential regulatory role in key steps of the viral life cycle (replication, transcription regulation, translation). Currently, there is no relevant information about putative G4s in the bovine foamy virus (BFV) genome. The goal of the present study was the determination of such conservative non-B-DNA structures as conservative G-quadruplexes, which can be formed by two and three G-quartets in the mRNA, sense, and antisense strands of the bovine foamy virus proviral DNA. Bioinformatic analysis was used to search motifs of intramolecular G-quadruplexes in BFV mRNA and proviral DNA and to determine the G-score (a parameter that characterizes the stability of the G-quadruplex in relative units). Based on multiple alignments of 27 BFV isolates 26 putative conservative G-quadruplexes from two G-quartets were found in mRNA and sense strand of BFV proviral DNA with G-score from 30 to 36. 32 G4s formed by two G-quartets with a G-score from 30 to 36 and 2 G4s formed by three G-quartets were found in the antisense strand of BFV proviral DNA with a G-score of 53. These two G4s are direct repeats and are localized in U5 5'LTR and U5 3'-LTR. The density of G4s was 2.1/kbp in the sense strand of BFV proviral DNA and 2.8/kbp in the antisense strand. A localization map of potential stable conserved intramolecular G-quadruplexes formed by two and three G-tetrads on the BFV genome was created. Conservative G4s are unevenly distributed throughout the BFV genome. A distinctive feature of the BFV genomic organization is the fact that the antisense strand of the BFV proviral DNA is characterized by a significantly higher density of G-quadruplexes compared to one of the sense strands. The QGRS Mapper software detects a significantly higher number of potential G4s (34 G4s in the antisense strand of BFV proviral DNA) compared to the G4Hunter software (7 G4s)

Keywords: bovine foamy virus, BFV, G-quadruplex, motif, non-canonical structure, direct repeat, antisense strand

Introduction. Foamy viruses (or spumaviruses) belong to the subfamily Spumaretrovirinae of the family Reoviridae. They were found in many species of animals, including cattle, and they are the most ancient among all known retroviruses (Pinto-Santini, Stenbak and Linial, 2017; Rethwilm and Bodem, 2013). The first foamy virus was described in 1954 (Enders and Peebles, 1954). However, the first bovine foamy virus (BFV), which belongs to the genus *Bovis foamy virus*, was isolated later (Malmquist, Van der Maaten and Boothe, 1969) and since then BFV-seropositive animals have been detected in many countries of the world.

The BFV infection rate for different countries varies considerably. In Canada, it ranged from 40 to 50%. A slightly lower level (39%) was observed in Australia and Great Britain. In Germany, only 7% of tested animals were identified as BFV-seropositive, in Poland among dairy cattle its value was 30%. In Japan, the infection level for different prefectures ranged from 12 to 16%. The highest level of seropositivity was typical for age-related animals (Materniak-Kornas et al., 2019; Okamoto et al., 2020).

Clinical symptoms or diseases associated with BFV have not yet been described, and the virus is considered a non-pathogenic one (Meiering and Maxine, 2001). However, the possibility of mixed infection with bovine leukemia virus, bovine immunodeficiency virus, and foamy virus, considered cofactors, has been established (Materniak-Kornas et al., 2017; Le et al., 2021).

BFV mRNA with a length of about 12.0 kb contains three typical for retroviruses structural *gag*, *pol*, *env* genes and two additional non-structural characteristics for foamy viruses' *bel-1* and *bel-2* genes, which are localized between *env* gene and 3'LTR and encode Bet protein. This protein consists of Tas and Bet2 parts. Unlike the Tas protein, which is a transcriptional activator, the function of the Bet2 protein has not yet been established (Jaguva Vasudevan et al., 2021; Mekata et al., 2021).

The nucleic acid molecules can form non-canonical structures (triplexes, hairpin structures, i-motifs (intercalated

motifs), G-quadruplexes, etc.) that are regulatory elements. Triplexes can regulate gene expression, and participate in chromatin organization and DNA recombination (Jenjaroenpun and Kuznetsov, 2009; Dalla Pozza et al., 2022). RNA hairpin structures called RNA thermometers control and regulate a variety of processes in microorganisms at the translational and transcriptional levels (Abduljalil, 2018). The possibility of the crucial role of i-motif in protein synthesis, chromosome integrity, and mitosis regulation is discussed (Luo et al., 2023).

G-quadruplexes (G4) are formed by planar G-quartets in DNA or RNA guanine-rich sequences under physiological conditions. G-quadruplexes are stabilized by Hoogsteen hydrogen bonds, G-quartets stacking, and monovalent cations K^+, and Na^+. G-quadruplexes are composed of one (intramolecular G4) (Fig. 1), and two or four (intermolecular G4) nucleic acid strands. G-quartets in intra-molecular G4 are connected by different loops with lengths of 1–12 nt and are influenced by G-quadruplexes' stability (Zaccaria and Fonseca Guerra, 2018).

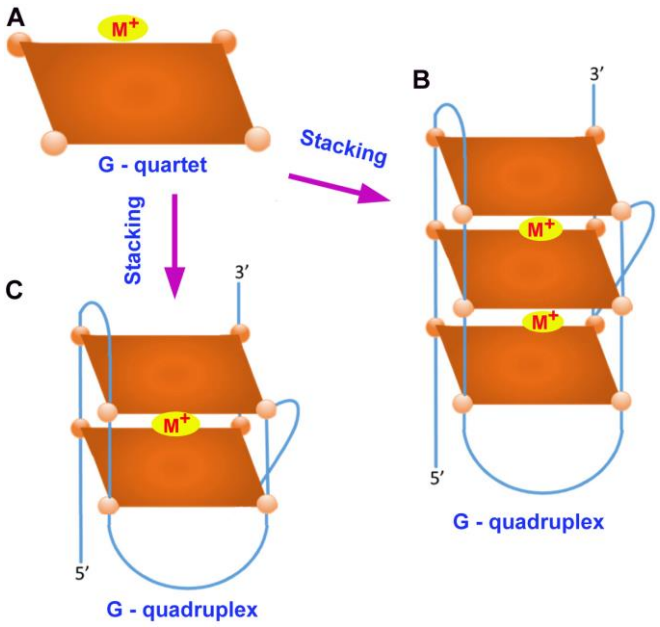

Figure 1. G-quartet involves four coplanar guanines establishing a cyclic array of H-bonds (A). Intramolecular anti-parallel G-quadruplex structure formed by stacking of three (B) and two (C) G-quartets and stabilized by monovalent cations, such as potassium, ammonium, and sodium (Brázda et al., 2020).

G-quadruplexes play significant roles in different biological processes such as translation, transcription, recombination, replication, etc. and they are found in micro- and macroorganisms. G-quadruplexes are associated with some human diseases and are considered prospective targets in molecular medicine (Cagirici, Budak and Sen, 2021; Harris and Merrick, 2015; Perrone et al., 2017). To predict G4 structure formation several methods have been developed (Puig Lombardi and Londoño-Vallejo, 2020).

There are spectroscopic techniques and biochemical methods for G4 identification, visualization, and function determination (Umar et al., 2019).

Regarding the formation of non-canonical structures in the foamy virus genome, information is very limited. The existence of a hairpin structure with the possibility of a stable pseudotriloop structure formation in the genomic RNA of the human foamy virus is shown by NMR spectroscopy and UV melting. This pseudotriloop structure is conserved among all known spumaretroviruses, including BFV, and it may be a recognition sites for proteins (Van Der Werf et al., 2013). As for the existence of motifs with the potential to form G-quadruplexes in the genome of foamy viruses, there is currently no relevant information in the literature.

Aim. In this study, putative conservative G-quadruplexes, which can be formed by two and three G-quartets in the mRNA, sense, and antisense strands of the bovine foamy virus proviral DNA was determined based on the 27 BFV isolates nucleotide sequences with a complete genome.

Materials and methods. Nucleotide sequences of 27 BFV isolates were obtained by searching taxonomic identifier (txid) 207343 in the GenBank database of the National Center for Biotechnology Information (USA).

BioEdit software (version 7.2.5) (Hall, 1999) was used to obtain a consensus sequence of BFV with a complete genome. The search for GGG motifs in the BFV mRNA and proviral DNA, manipulation with nucleotide sequences (reverse, complement), and multiple alignments were performed using the Molecular Evolutionary Genetics Analysis (MEGA) software (version 6.06) (Tamura et al., 2013).

QGRS Mapper software on the web server https://bioinformatics.ramapo.edu/QGRS/index.php (Menendez, Frees and Bagga, 2012; Kikin, D'Antonio and Bagga, 2006) was used for searching motifs of intramolecular G-quadruplexes in mRNA, sense, and antisense strands of the BFV proviral DNA, as well as to determine the G-score (a parameter that characterizes the stability of the G-quadruplex in relative units) with the following parameters: the maximum length of the G-quadruplex is 45 nt, the minimum size of the G-quartet number is 2, the loop — from 0 to 20 nt. Additionally, putative G-quadruplexes were identified by G4Hunter software on the web server https://bioinformatics.ibp.cz (Bedrat, Lacroix and Mergny, 2016; Brázda et al., 2019) with a window of 25, the threshold of 1.2.

Results. G-quadruplexes in DNA and RNA have different topologies. The topology of RNA G4s is limited to the parallel conformation where all four strands are oriented in the same direction. In contrast, DNA G4s can adapt to parallel, antiparallel, or mixed conformations.

Intramolecular interactions within RNA G4s lead to enhanced stability of RNA G4s compared to DNA G4s (Fay, Lyons and Ivanov, 2017).

Sequences forming G-quadruplexes were used as a control of the correctness of the search for G4s motifs. These G4s were experimentally confirmed for the hepatitis B virus (HBV) and found in the genome of tobacco (*Nicotiana tabacum*), which is an ideal plant model in scientific research. The HBV genome is represented by double-stranded and relaxed circular DNA with a length of approximately 3,200 bp. 13 G4s were found in the HBV genome, for the seven of which the determined G-score value was ranged from 19 to 35 (Table 1). For the five G4s most frequently found in the tobacco genome, the G-score varied from 20 to 29 (Table 1).

Table 1 — Conservative G-quadruplex motifs which were determined and used as a control for determining the G-score. G-rich fragments of sequences for putative G-quadruplexes are underlined. Positions of G4s on the HBV genome and the number of G4s in all chromosomes of the tobacco (*Nicotiana tabacum*) nuclear genome are shown in parentheses

Virus/Plant	Length, nt	G-quadruplex motif	G-score
HBV (1207)	19	GGCUGGGGCUUGGUCAUGG (Wang et al, 2023)	35
HBV (1739)	22	GGAGUUGGGGAGGAGAUUAGG (Wang et al, 2023)	34
HBV (1889)	19	GGGUGGCUUUGGGGCAUGG (Wang et al, 2023)	33
HBV (2992)	23	GGAGCUGGAGCAUUCGGGCUGGG (Wang et al, 2023)	33
HBV (3034)	25	GGCCUUUUGGGGUGGAGCCCUCAGG (Wang et al, 2023)	31
HBV (261)	28	GGUGGACUUCUCUCAAUUUUCUAGGGGG (Wang et al, 2023)	19
HBV (1779)	22	GGAGGCUGUAGGCAUAAAUUGG (Wang et al, 2023)	29
Tobacco (1392)	24	GGGGGTGTGTACAGACTCCGGAGG (Song et al, 2024)	23
Tobacco (605)	22	GGGGGCCTCGGGTGTGTTTCGG (Song et al, 2024)	29
Tobacco (587)	23	GGGGTGTGTACAGACTCCGGAGG (Song et al, 2024)	22
Tobacco (599)	26	GGGGGGTTGACTTTTTGATATCGGGGG (Song et al, 2024)	20
Tobacco (575)	22	GGGGGTGTACAGACTCCGGAGG (Song et al, 2024)	25

The sense strand, or coding strand, is a strand within double-stranded DNA that carries information about the translation code in the 5' to 3' direction. Its complementary strand is called the antisense strand, which does not carry a translation code and serves as a matrix during translation. The sense strand of DNA has the same sequence as mRNA, which contains codon sequences for

Taking into account the experimental data on the parameters of the detected G4s in the genome of HBV and tobacco (Table 1), only conservative fragments were analyzed for BFV isolates, which can form G-quadruplexes with G-score ≥ 30. $G_xN_yG_xN_yG_xN_yG_x$ motif with a maximum length of 45 nt and $2 \leq x \leq 4$, $1 \leq y \leq 10$ was used to predict putative G4s (Kikin, D'Antonio and Bagga, 2006).

G4s can affect gene expression also if they are localized on the sense strand concerning the direction of transcription. For this reason, the search of putative G4s motifs was performed both for sense and antisense strands of the BFV proviral DNA.

Based on multiple alignment of 27 BFV isolates 26 putative conservative G-quadruplexes from two G-quartets were found in mRNA and sense strand of BFV proviral DNA with G-scores from 30 to 36 (Table 2). 32 G4s formed by two G-quartets with a G-score from 30 to 36 and 2 G4s formed by three G-quartets were found in the antisense strand of BFV proviral DNA with a G- score of 53 (Table 3). These two G4s are direct repeats and they are localized in 5'- and 3'-LTR. The density of G4s in the sense strand of BFV proviral DNA was 2.1/kbp, and the density of G4s in the antisense strand was 2.8/kbp.

A comparison of the G4s search effectiveness by two software — QGRS Mapper and G4Hunter — was carried out. 26 putative G4s for mRNA and the sense strand of BFV proviral DNA (Table 2) were identified by QGRS Mapper software, while no G4s were identified by G4Hunter. 34 putative G4s were identified by QGRS Mapper software for antisense strand of BFV proviral DNA, while seven G4s were identified by G4Hunter software (Table 3, motifs are highlighted), six of which are fragments of the G4s identified by QGRS Mapper software.

The formation of non-B-DNA structures requires the unwinding of double-stranded DNA molecules, which occurs during replication and transcription. It is known that G4s in the antisense strand significantly inhibit transcription, unlike G4s in the sense strand, which do not affect transcription (Agarwal et al., 2014). Computer

analysis of 27 BFV isolates with a complete genome showed that G4s are unevenly distributed both among structural genes and in the sense and antisense strands of BFV proviral DNA.

Table 2 — Conservative sequences for putative G-quadruplexes in mRNA and sense strand of bovine foamy virus proviral DNA based on multiple alignments of 27 BFV isolates. G-rich fragments of sequences for putative G4s are underlined. Gene positions are indicated for JX307861 BFV isolate. DR3, DR4, DR5 are the direct repeats

Position	Gene/region	G-quadruplex motif		G-score
56	U3	GGAGGATTGGCTGG	DR3	34
10754	U3	GGAGGATTGGCTGG	DR3	34
486	U3	GGTTCGGAGGATGG	*DR4*	34
11184	U3	GGTTCGGAGGATGG	*DR4*	34
574	U3	GGTTCGAGGTCAGGCGG	DR5	32
11272	U3	GGTTCGAGGTCAGGCGG	DR5	32
1716	*gag*	GGtatctagatatGGtcccttagaagggGGtgattatcagccaGG		34
2183	*gag*	GGCATCCCGGGCTCGCCTTGGAACCCTTGG		35
2914	*gag*	GGaagaagcagccaaGGaacaaacaacacaGGaagttctgcCGG		32
3031	*gag*	GGaaatcaagGGCaatcatcttaaaGGCTACTGGGACTCTGG		31
3616	*pol*	GGCTGATGGGCGCTGGAGGATGG		36
6225	*pol*	GGactcccagtcctGGcctgttggtccaGGAGAGGGTAGCCAGG		36
6582	*pol*	GGCATCGTTGGAAGGAGTGCGGG		31
7309	*env*	GGcactgttacgtGGTACATATTCCGGGGGAAAAGAATGG		36
7688	*env*	GGCACTTCCAGGAACCTGGCACACTCTAGG		32
8770	*env*	GGATCCGAAACTGGCAAGTGGCCAATCTAGG		31
8939	*env*	GGTGGTGGAAGTGGAACCTTGTGGCAACCTGACAGG		34
9332	*env*	GGatctccacgaaGGAagaattactcagactGGatcctgcagcttGG		33
9429	*env*	GGGACCTTACTGGAGAAAGCTGCAGGAACTCTCTTCGG		34
10013	*bel1*	GGAGGAGGAACACCCGG		30
10043	*bel1*	GGTGGAGACCAGGACATGCGGGTCAATACTTCCGG		32
10239	*bel1*	GGTTTTGGACAGGTGATGATAAAGGCCCACAGGAATGG		34
11002	U3	GGCCAAGCAGGCCTCCCGGCGGG		31
11002	U3	GGCCAAGCAGGCCTCCCGGCGGG		31
11272	R	GGTTCGAGGTCAGGCGG		32
11604	U5	GGggaattattgGGaatcccatattgtaGGAAAAGTCAGTTGG		33

We have indicated the sites of putative G4s on the BFV genomic map, which shows the relative positions of structural genes, regulatory elements, genetic markers, and the distance between them (Fig. 2). 4, 3, and 7 G4s were identified in *gag*, *pol* and *env* genes of the sense strand, while 10, 11, and 5 G4s were identified for the above-mentioned genes of the antisense strand, respectively. Represented data indicate that the formation of G4s in sense and antisense strands of BFV proviral DNA occurs asymmetrically, which may result in different effects of G4s, which are identified in sense and antisense strands, on the transcription.

Designed genomic map of putative G4s sites has permitted to reveal the features of the structural organization of the sense strand in comparison to the antisense strand of BFV proviral DNA. Three direct repeats (DR3–DR5) formed by two G-tetrads with G-score from 32 to 34 were found in 5'- and 3'-LTR of the sense strand. Two direct repeats (DR1–DR2) were found in 5'- and 3'-LTR of the antisense strand, one of which (DR1) is formed by two G-tetrads (G-score equals 32). Another direct repeat (DR2) is formed by three G tetrads, and the G-score of G4s is 53.

Table 3 — Conservative sequences for putative G-quadruplexes in the antisense strand of bovine foamy virus proviral DNA based on multiple alignments of 27 BFV isolates. G-rich fragments of sequences for putative G4s are underlined. Gene positions are indicated for JX307861 BFV isolate. W=A/T, Y=C/T. DR1 and DR2 are the direct repeats. Identified by G4Hunter software motifs with corresponding G-score values in parenthenses are highlighted

Position	Gene/region	G-quadruplex motif		G-score
11888	U5	GGaaggtaagcccaGGttggagtcaGGttcccactcgcacGG	*DR2*	32
1190	U5	GGaaggtaagcccaGGttggagtcaGGttcccactcgcacGG	*DR2*	32
10715	U3	GGGCAAGTAAAGGGGGGCTGGccyttgctwatttcyGGG	DR1	53
17	U3	GGGCAAGTAAAGGGGGGCTGGccyttgctwatttcyGGG	DR1	53
10671	*bel2*	GGAGGGCATGGCTGG		35
10092	*bel1*	GGTCATCCCTGGCACATTGGCACAAAGAGG		34
9973	*bel1*	GGGGTGGAGG (35)		35
9915	*bel1*	GGTTTTGTCCTAGGATTGCATCGATCGGTCAGATGCCTGG		34
8234	*env*	GGAGGGAGAAGGCTCGG		34
7676	*env*	GGTTCCTGGAAGTGCCTCAGGGTTAAGG		31
7561	*env*	GGAGTATAGGGTAGGGCTACAGAAGG		32
7061	*env*	GGGTTTTGGGTGAGGTTCCAGG		35
6668	*env*	GGTTCATCTGCAGGAGGTTGGTCGTCTCTAGG		31
6222	*pol*	GGCtaccctctcctGGaccaacaggccaGGACTGGGAGTCCAGG		36
6193	*pol*	GGGCAGTGGTCTGTGAAGGAGAAGGG		33
5906	*pol*	GGGGGTGGTAAGG		34
5830	*pol*	GGAGGAGGTGAATGCTGGCCCCTGG		32
5743	*pol*	GGTTGCGTTGGCAGTTTGGGCCTTGGTGGGG (33)		36
5669	*pol*	GGAGGGAGGGAGTGG		35
5165	*pol*	GGTTGGTGGCCTGG		34
4250	*pol*	GGATACATTGGCCTTTGGCTTTGGG		34
3844	*pol*	GGAGAGTTTAGGAATCCCTGGGGGAGG		32
3254	*pol*	GGTCCTGGTTTATACCAGGGAATATCGATAGG		30
2986	*pol*	GGCGGAGTGCCGGCATCTTGTGGGGGTATGGGTGG (34)		36
2874	*gag*	GGGGGCGGGACCTCGG		32
2818	*gag*	GGGGTTGGTTGTCGATAAGGGTTCGG		30
2746	*gag*	GGATTGGGCCTGAGGGGATAACGAGG (31)		33
1987–2006	*gag*	GGCCCTGGAGCAGGTGCAGG		36
1969–1979	*gag*	GGTGGAGGAGG		36
1939–1958	*gag*	GGAGCAGGCAATGGAGCTGG		36
1915–1932	*gag*	GGCTGGAGGTGGTGGAGG (36)		35
1891–1918	*gag*	GGAGGAGCAGGTCCTTGTGGAGCCGCTGG		36
1633	*gag*	GGGTAAGGTTGGGTCCGAGG		34
1573	*gag*	GGTTCCCATTGAGGAATAGGCAAGGGTTGCCCTGCAGG		36

Structural genes of the sense strand are characterized by a different density of G4s compared to one of the antisense strands of BFV proviral DNA. For the *gag* gene of the sense strand, the density of G4s was 2.5/kbp, while one for the antisense strand was 6.3/kbp. For the *pol* gene of the sense strand the density of G4s was 0.8/kbp, while one for the antisense strand was 3.0/kbp. For the *env* gene of the sense strand, the density of G4s was 2.4/kbp, while one for the antisense strand was 1.7/kbp.

The antisense strand of BFV proviral DNA is characterized by an extremely high density of G4s for distinct gene fragments, which is not observed for the

sense strand. In particular, in the fragment of the *pol* gene with a length of 116 nt (positions 1891–2006 nt) 5 G4s were identified and the density of G4s was 43.1/kbp. For the longer *pol* gene fragment (406 nt, positions 1891–2006 nt), taking into account the additional two detected G4s, the density of G4s was 16.1/kbp.

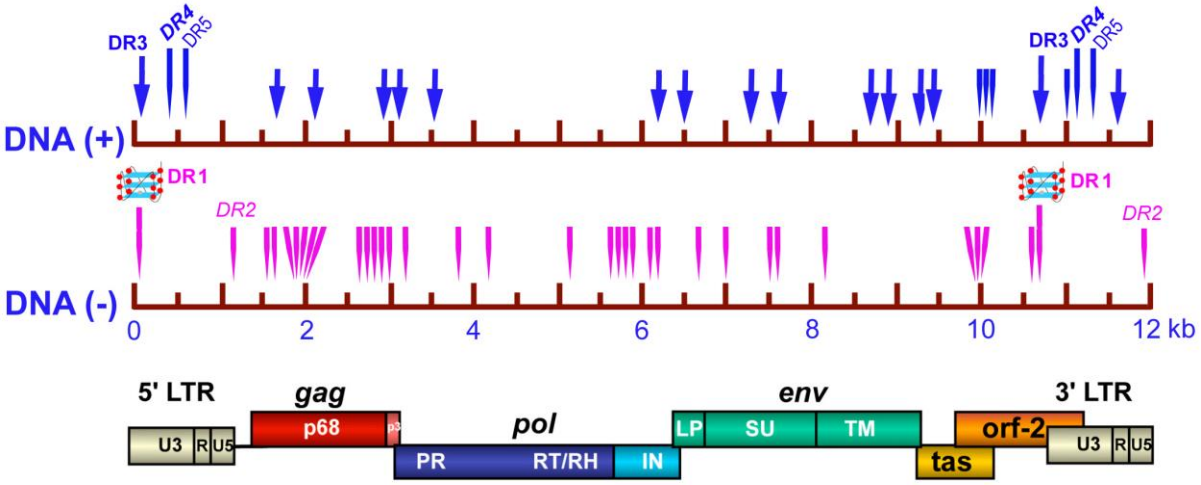

Figure 2. Genomic map of the bovine foamy virus genome with the positions of known genes (Hamann and Lindemann, 2016). Arrows show sites of G-quadruplex motifs in sense and antisense strands of BFV proviral DNA. Small G4 models indicate putative intramolecular G-quadruplexes whose sequences are direct repeats at the 5'- and 3'-ends of the antisense strand of proviral DNA with a G-score of 53.

Conclusions. In the present study, the existence of conservative motifs in mRNA, sense, and antisense strands of BFV proviral DNA, which contain guanine-riched fragments with the potential to form G-quadruplexes was demonstrated. Comparison of the G-score for experimentally determined known from the literature G4s and ones for BFV shows that the G-score for putative G-quadruplexes in the BFV genome (30–53) coincides with the G-score for the hepatitis B virus, *Nicotiana tabacum*.

Conservative G4s are unevenly distributed throughout the BFV genome. A distinctive feature of the BFV genomic organization is the fact that the antisense strand of the BFV proviral DNA is characterized by a significantly higher density of G4-quadruplexes compared to one of the sense strands. The QGRS Mapper program detects a significantly higher number of putative G4s (34 G4s in the antisense strand of BFV proviral DNA) compared to the G4Hunter program (7 G4s).

Represented data indicate that the formation of G4s in sense and antisense strands of BFV proviral DNA occurs asymmetrically, which may result in different effects of G4s on the transcription. Structural genes of the sense strand are characterized by a different density of G4s in comparison to one of the antisense strands of BFV proviral DNA.

A designed genomic map of the distribution of putative G4s raises several questions because the biological function of the vast majority of G4s is uncertain. It is known that G4s, which are localized in the 5'- and 3'-LTR, have different effects on translation (Kikin, D'Antonio and Bagga, 2006). G4s in the antisense strand create blockage for RNAPII unlike G4s in the sense strand and thus reduce transcription (Agarwal et al., 2014).

The identification in the 5'- and 3'-LTR of three direct repeats formed by two G-tetrads for the sense strand and two direct repeats, one of which is formed by three G-tetrads, for the antisense strand indicates that G4s, which are localized in the 5'- and 3'-LTR of the sense and antisense strands, can have different effects on transcription and translation.

References

Abduljalil, J. M. (2018) 'Bacterial riboswitches and RNA thermometers: Nature and contributions to pathogenesis', *Non-coding RNA Research*, 3(2), pp. 54–63. doi: 10.1016/j.ncrna.2018.04.003.

Agarwal, T., Roy, S., Kumar, S., Chakraborty, T. K. and Maiti, S. (2014) 'In the sense of transcription regulation by G-quadruplexes: Asymmetric effects in sense and antisense strands', *Biochemistry*, 53(23), pp. 3711–3718. doi: 10.1021/bi401451q.

Bedrat, A., Lacroix, L. and Mergny, J.-L. (2016) 'Re-evaluation of G-quadruplex propensity with G4Hunter', *Nucleic Acids Research*, 44(4), pp. 1746–1759. doi: 10.1093/nar/gkw006.

Brázda, V., Kolomazník, J., Lýsek, J., Bartas, M., Fojta, M., Šťastný, J. and Mergny, J.-L. (2019) 'G4Hunter web application: a web server for G-quadruplex prediction', *Bioinformatics*, 35(18), pp. 3493–3495. doi: 10.1093/bioinformatics/btz087.

Brázda, V., Luo, Y., Bartas, M., Kaura, P., Porubiaková, O., Šťastný, J., Pečinka, P., Verga, D., Da Cunha, V., Takahashi, T. S.,

Forterre, P., Myllykallio, H., Fojta, M. and Mergny, J.-L. (2020) 'G-quadruplexes in the Archaea domain', *Biomolecules*, 10(9), p. 1349. doi: 10.3390/biom10091349.

Cagirici, H. B., Budak, H. and Sen, T. Z. (2021) 'Genome-wide discovery of G-quadruplexes in barley', *Scientific Reports*, 11(1), p. 7876. doi: 10.1038/s41598-021-86838-3.

Dalla Pozza, M., Abdullrahman, A., Cardin, C. J., Gasser, G. and Hall, J. P. (2022) 'Three's a crowd — stabilisation, structure, and applications of DNA triplexes', *Chemical Science*, 13(35), pp. 10193–10215. doi: 10.1039/D2SC01793H.

Enders, J. F. and Peebles, T. C. (1954) 'Propagation in tissue cultures of cytopathogenic agents from patients with Measles', *Experimental Biology and Medicine*, 86(2), pp. 277–286. doi: 10.3181/00379727-86-21073.

Fay, M. M., Lyons, S. M. and Ivanov, P. (2017) 'RNA G-quadruplexes in biology: Principles and molecular mechanisms', *Journal of Molecular Biology*, 429(14), pp. 2127–2147. doi: 10.1016/j.jmb.2017.05.017.

Hall, T. A. (1999) 'BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT', *Nucleic Acids Symposium Series*, 41, pp. 95–98. Available at: https://www.academia.edu/2034992.

Hamann, M. and Lindemann, D. (2016) 'Foamy virus protein–nucleic acid interactions during particle morphogenesis', *Viruses*, 8(9), p. 243. doi: 10.3390/v8090243.

Harris, L. M. and Merrick, C. J. (2015) 'G-quadruplexes in pathogens: A common route to virulence control?', *PLOS Pathogens*, 11(2), p. e1004562. doi: 10.1371/journal.ppat.1004562.

Jaguva Vasudevan, A. A., Becker, D., Luedde, T., Gohlke, H. and Münk, C. (2021) 'Foamy viruses, Bet, and APOBEC3 restriction', *Viruses*, 13(3), p. 504. doi: 10.3390/v13030504.

Jenjaroenpun, P. and Kuznetsov, V. A. (2009) 'TTS Mapping: integrative web tool for analysis of triplex formation target DNA Sequences, G-quadruplets and non-protein coding regulatory DNA elements in the human genome', *BMC Genomics*, 10(Suppl 3), p. S9. doi: 10.1186/1471-2164-10-S3-S9.

Kikin, O., D'Antonio, L. and Bagga, P. S. (2006) 'QGRS Mapper: A web-based server for predicting G-quadruplexes in nucleotide sequences', *Nucleic Acids Research*, 34(Web Server), pp. W676–W682. doi: 10.1093/nar/gkl253.

Le, D. T., Nguyen, S. V., Okamoto, M., Yamashita-Kawanishi, N., Dao, T. D., Bui, V. N., Ogawa, H., Imai, K. and Haga, T. (2021) 'Molecular characterization of bovine foamy virus and its association with bovine leukemia virus infection in Vietnamese cattle', *Journal of Veterinary Medical Science*, 83(8), pp. 1273–1277. doi: 10.1292/jvms.21-0190.

Luo, X., Zhang, J., Gao, Y., Pan, W., Yang, Y., Li, X., Chen, L., Wang, C. and Wang, Y. (2023) 'Emerging roles of i-motif in gene expression and disease treatment', *Frontiers in Pharmacology*, 14, p. 1136251. doi: 10.3389/fphar.2023.1136251.

Malmquist, W. A., Van der Maaten, M. J. and Boothe, A. D. (1969) 'Isolation, immunodiffusion, immunofluorescence, and electron microscopy of a syncytial virus of lymphosarcomatous and apparently normal cattle', *Cancer Research*, 29(1), pp. 188–200. Available at: https://aacrjournals.org/cancerres/article/29/1/188/477078.

Materniak-Kornas, M., Osiński, Z., Rudzki, M. and Kuźmak, J. (2017) 'Development of a recombinant protein-based ELISA for detection of antibodies against bovine foamy virus', *Journal of Veterinary Research*, 61(3), pp. 247–252. doi: 10.1515/jvetres-2017-0034.

Materniak-Kornas, M., Tan, J., Heit-Mondrzyk, A., Hotz-Wagenblatt, A. and Löchelt, M. (2019) 'Bovine foamy virus: Shared and unique molecular features *in vitro* and *in vivo*', *Viruses*, 11(12), p. 1084. doi: 10.3390/v11121084.

Meiering, C. D. and Maxine, L. L. (2001) 'Historical perspective of foamy virus epidemiology and infection', *Clinical Microbiology Reviews*, 14(1), pp. 165–176. doi: 10.1128/CMR.14.1.165-176.2001.

Mekata, H., Okagawa, T., Konnai, S. and Miyazawa, T. (2021) 'Molecular epidemiology and whole-genome analysis of bovine foamy virus in Japan', *Viruses*, 13(6), p. 1017. doi: 10.3390/v13061017.

Menendez, C., Frees, S. and Bagga, P. S. (2012) 'QGRS-H Predictor: A web server for predicting homologous quadruplex forming G-rich sequence motifs in nucleotide sequences', *Nucleic Acids Research*, 40(W1), pp. W96–W103. doi: 10.1093/nar/gks422.

Okamoto, M., Oguma, K., Yamashita-Kawanishi, N., Ichijo, T., Hatama, S., Endo, M., Ishikawa, M. and Haga, T. (2020) 'Genomic characterization and distribution of bovine foamy virus in Japan', *Journal of Veterinary Medical Science*, 82(11), pp. 1607–1613. doi: 10.1292/jvms.20-0429.

Perrone, R., Lavezzo, E., Palù, G. and Richter, S. N. (2017) 'Conserved presence of G-quadruplex forming sequences in the Long Terminal Repeat Promoter of Lentiviruses', *Scientific Reports*, 7(1), p. 2018. doi: 10.1038/s41598-017-02291-1.

Pinto-Santini, D. M., Stenbak, C. R. and Linial, M. L. (2017) 'Foamy virus zoonotic infections', *Retrovirology*, 14(1), p. 55. doi: 10.1186/s12977-017-0379-9.

Puig Lombardi, E. and Londoño-Vallejo, A. (2020) 'A guide to computational methods for G-quadruplex prediction', *Nucleic Acids Research*, 48(1), pp. 1–15. doi: 10.1093/nar/gkz1097.

Rethwilm, A. and Bodem, J. (2013) 'Evolution of foamy viruses: The most ancient of all retroviruses', *Viruses*, 5(10), pp. 2349–2374. doi: 10.3390/v5102349.

Song, K., Li, B., Li, H., Zhang, R., Zhang, X., Luan, R., Liu, Y. and Yang, L. (2024) 'The characterization of G-quadruplexes in tobacco genome and their function under abiotic stress', *International Journal of Molecular Sciences*, 25(8), p. 4331. doi: 10.3390/ijms25084331.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) 'MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0', *Molecular Biology and Evolution*, 30(12), pp. 2725–2729. doi: 10.1093/molbev/mst197.

Umar, M. I., Ji, D., Chan, C.-Y. and Kwok, C. K. (2019) 'G-quadruplex-based fluorescent turn-on ligands and aptamers: From development to applications', *Molecules*, 24(13), p. 2416. doi: 10.3390/molecules24132416.

Van Der Werf, R., Wijmenga, S. S., Heus, H. A. and Olsthoorn, R. C. L. (2013) 'Structural and thermodynamic signatures that define pseudotriloop RNA hairpins', *RNA*, 19(12), pp. 1833–1839. doi: 10.1261/rna.039636.113.

Wang, J., Huang, H., Zhao, K., Teng, Y., Zhao, L., Xu, Z., Zheng, Y., Zhang, L., Li, C., Duan, Y., Liang, K., Zhou, X., Cheng, X. and Xia, Y. (2023) 'G-quadruplex in Hepatitis B virus pregenomic RNA promotes its translation', *Journal of Biological Chemistry*, 299(9), p. 105151. doi: 10.1016/j.jbc.2023.105151.

Zaccaria, F. and Fonseca Guerra, C. (2018) 'RNA versus DNA G-quadruplex: The origin of increased stability', *Chemistry — A European Journal*, 24(61), pp. 16315–16322. doi: 10.1002/chem.201803530.