products are pathological. Further experimental work is required to confirm the formation of these structures *in vitro* and to determine their precise functional role in the context of prolactin gene regulation in cardiac and pituitary cells.

Conclusions

A total of 24 putative G-quadruplex forming sequences were identified *in silico* within the human prolactin (*PRL*) gene.

The predicted G4–forming sequences are predominantly located on the sense strand (17 sequences), suggesting a potential strand-specific regulatory function.

Five sequences on the sense strand demonstrated exceptionally high predicted stability ("INF" cG/cC score), making them the most promising candidates for future experimental validation.

The presence of these high-confidence PQS in the *PRL* gene opens new avenues for research into the transcriptional and post-transcriptional regulation of prolactin, potentially informing novel therapeutic strategies for peripartum cardiomyopathy.

LITERATURE

1. Recovery from postpartum cardiomyopathy in 2 patients by blocking prolactin release with bromocriptine / D. Hilfiker-Kleiner [et al.]. // Journal of the American College of Cardiology. – 2007. – Vol. 50, № 24. – P. 2354–2355.

2. Arany, Z. Peripartum cardiomyopathy / Z. Arany, U. Elkayam // Circulation. – 2016. – Vol. 133, № 14. – P. 1397–1409.

3. Идентификация и роль гуаниновых квадруплексов в регуляции биоэнергетики, апоптоза и ферроптоза в митохондриальной ДНК Kluyveromyces lactis / А. Н. Коваль, Л. Н. Алексейко, О. С. Логвинович [и др.]. // Актуальные проблемы общей и клинической биохимии – 2025 : сборник материалов республиканской научно-практической конференции с международным участием, посвященной 100-летию со дня рождения академика Ю. М. Островского, Гродно, 27 июня 2025 года. – Гродно: Гродненский государственный медицинский университет, 2025. – С. 223–226.

4. G-quadruplex structures in the human genome as novel therapeutic targets / R. Hänsel-Hertsch [et al.]. // Nature Reviews Molecular Cell Biology. – 2017. – Vol. 18. P. 279–284.

5. National Center for Biotechnology Information // National Library of Medicine (US). – URL: https://www.ncbi.nlm.nih.gov/datasets/ (date of access: 18.03.2026).

6. G4IPDB: A database for G-quadruplex structure forming nucleic acid interacting proteins / S. K. Mishra [et al.]. // Scientific reports. – 2016. – Vol. 6, № 1. – P. 38144.

УДК 577.175.328:[618.33+616.127-053.31]

U. R. S. P. Vintura Silvage

Scientific supervisor: Associate Professor, PhD A. N. Koval

Educational Establishment
"Gomel State Medical University"
Gomel, Republic of Belarus

IDENTIFICATION OF PUTATIVE G-QUADRUPLEX STRUCTURES IN THE PROLACTIN GENE AND THEIR RELEVANCE TO PERIPARTUM CARDIOMYOPATHY

Introduction

Peripartum cardiomyopathy (PPCM) is a potentially life-threatening form of heart failure that occurs in previously healthy women during the late stages of pregnancy or in the months following delivery [1]. While its ethology is complex and multifactorial, accumulating evidence points towards a central role for the 16 kDa prolactin fragment. This N-terminal fragment, generated by proteolytic cleavage of the full-length 23 kDa prolactin hormone, has been shown to induce endothelial damage and myocardial dysfunction, directly implicating the prolactin pathway in the

pathogenesis of PPCM. Consequently, understanding the regulatory mechanisms governing prolactin (PRL) gene expression is of significant clinical interest [2].

G-quadruplexes (G4) are four-stranded secondary structures that can form in guanine-rich regions of DNA or RNA [3]. These non-canonical structures are formed by the stacking of guanine tetrads and are stabilized by monovalent cations, particularly potassium [3]. G4 structures are not merely structural curiosities; they are increasingly recognized as critical regulatory elements involved in key genomic processes, including transcription, replication, and translation [3, 4]. The presence of putative G4–forming sequences (PQS) in gene promoters or within the gene body can act as molecular switches, modulating gene expression in response to cellular conditions. The *in-silico* prediction of such sequences is a crucial first step in understanding their potential biological role.

Goal

The aim of this study was to perform an *in-silico* analysis of the human prolactin (PRL) gene to identify and characterize putative G-quadruplex forming sequences, providing a foundation for future research into their potential role in the regulation of PRL expression in the context of peripartum cardiomyopathy.

Material and methods of research

The complete genomic sequence of the human prolactin gene (Gene ID: 5617; *PRL*) was retrieved from the National Center for Biotechnology Information (NCBI) Datasets resource (https://www.ncbi.nlm.nih.gov/datasets/gene/) [5]. The corresponding gene sequence file (gene.fna) was downloaded. This FASTA file was then uploaded to the G4IPDB online prediction tool (https://people.iiti.ac.in/~amitk/bsbe/ipdb/pattern2.php) [6]. TheG4IPDB algorithm scans both the sense (5'–>3') and antisense (3'–>5') strands for sequences conforming to the canonical G4 motif: $G_{x-}N_yG_xN_yG_xN_yG_x$, where $x \geq 3$ and y represents a variable loop length. The analysis was performed using the default settings. The resulting output table, containing the position, strand, sequence, length, and confidence scores (cG, cC, and cG/cC), was saved as a text file (PRL_G4.txt). The cG and cC scores represent the propensity of the sequence to form a stable G4 on the G-rich and C-rich strands, respectively. The cG/cC ratio is a key indicator of strand bias and structure stability; a high ratio (>20) or an "INF" value (where cC = 0) indicates a very high probability of stable G4 formation on the G-rich strand.

The results of the research and their discussion

The G4IPDB analysis of the human *PRL* gene identified a total of 24 unique putative G-quadruplex forming sequences. Of these, 17 were predicted on the sense strand, and 7 were predicted on the antisense strand. This uneven distribution suggests a potential strand-specific regulatory bias within the gene.

A detailed examination of the G4IPDB output revealed sequences with a range of predicted stabilities. We selected the top 15 sequences based on their cG/cC scores to highlight the most promising candidates for potential biological activity. These high-confidence sequences, which are most likely to form stable G4 structures under physiological conditions, are presented in Table 1. Notably, several sequences on the sense strand yielded a cG/cC score of "INF", as their cC score was zero. This indicates an extremely high confidence for G4 formation on the G-rich strand, making them primary targets for future *in vitro* validation.

The presence of multiple high-scoring PQS within the PRL gene is a novel finding with potential implications for PPCM. The expression of prolactin is tightly controlled, and its dysregulation is a key event in PPCM pathophysiology. G4 structures, particularly those located in regulatory regions, could act as cis-regulatory elements. For instance, a stable G4 in the promoter region might impede the transcriptional machinery, thereby downregulating PRL expression. Conversely, if located in the transcribed region, it could affect RNA processing or translation efficiency. The

identification of several high-confidence sequences (e.g., those with INF scores at positions 3767, 3704, 4113, 4505, and 1255) suggests that the PRL gene is structurally poised for G4–mediated regulation. These sequences could represent potential therapeutic targets. Small molecules that stabilize or destabilize specific G4 structures could, in theory, be used to modulate PRL expression levels, offering a novel approach to managing conditions like PPCM where prolactin cleavage products are pathological. Further experimental work is required to confirm the formation of these structures *in vitro* and to determine their precise functional role in the context of prolactin gene regulation in cardiac and pituitary cells.

Table 1 – Putative G-quadruplex forming sequences in the human PRL gene with high predicted stability

Strand	Start Position (bp)	Putative G4 Sequence (5'–>3')	Length (bp)	cG Score	cC Score	cG/cC Score
Sense	3767	GGGCTGGGGGATGGGGGCTGGGG	23	136.4	0.0	INF
Sense	3704	GGGAGGAGCAGGGACAGGGAGGG	22	114.9	0.0	INF
Sense	4113	GGGTGGGTGGGATGGG	16	81.4	0.0	INF
Sense	4505	GGGAAGGGAGGGAAGG	16	71.8	0.0	INF
Sense	1255	GGGTGGGAGGGAGGG	15	68.7	0.0	INF
Sense	3043	GGGATGGGAGGGATGGG	17	86.0	2.6	33.08
Sense	4750	GGGCTGGGGCTGGGGG	16	83.8	2.6	32.23
Sense	4890	GGGACGGGCTGGGG	14	60.0	2.6	23.08
Sense	1093	GGGGAGGGGAGGGAGGG	17	92.7	5.2	17.83
Sense	1268	GGGAGGGAGGGAAGG	15	68.7	5.2	13.21
Sense	3763	GGGAGGGCTGGGGGATGGGGGCTGGGG	26	137.3	15.7	8.75
Sense	3765	GGGCTGGGGGATGGGGGCTGGGG	22	132.5	15.7	8.44
Sense	4869	GGGGGACGGGCTGGGG	16	71.5	15.7	4.55
Antisense	966	CCCAGCCCACCCCTCCC	17	0.0	77.6	0.00
Antisense	4279	CCCCAGCCCCAGCCCC	16	0.0	73.3	0.00

Note. "INF" (infinity) indicates a cG/cC ratio where the cC score is zero, signifying an exceptionally high confidence for G4 formation on the G-rich strand.

Conclusions

A total of 24 putative G-quadruplex forming sequences were identified *in silico* within the human prolactin (*PRL*) gene.

The predicted G4–forming sequences are predominantly located on the sense strand (17 sequences), suggesting a potential strand-specific regulatory function.

Five sequences on the sense strand demonstrated exceptionally high predicted stability ("INF" cG/cC score), making them the most promising candidates for future experimental validation.

The presence of these high-confidence PQS in the *PRL* gene opens new avenues for research into the transcriptional and post-transcriptional regulation of prolactin, potentially informing novel therapeutic strategies for peripartum cardiomyopathy.

LITERATURE

1. Arany, Z. Peripartum cardiomyopathy / Z. Arany, U. Elkayam // Circulation. – 2016. – Vol. 133, № 14. – P. 1397–1409.

2. Recovery from postpartum cardiomyopathy in 2 patients by blocking prolactin release with bromocriptine / D. Hilfiker-Kleiner [et al.]. // Journal of the American College of Cardiology. – 2007. – Vol. 50, № 24. – P. 2354–2355.

3. Идентификация и роль гуаниновых квадруплексов в регуляции биоэнергетики, апоптоза и ферроптоза в митохондриальной ДНК Kluyveromyces lactis / А. Н. Коваль, Л. Н. Алексейко, О. С. Логвинович [и др.]. //

Актуальные проблемы общей и клинической биохимии – 2025 : сборник материалов республиканской научно-практической конференции с международным участием, посвящённой 100-летию со дня рождения академика Ю. М. Островского, Гродно, 27 июня 2025 года. – Гродно: Гродненский государственный медицинский университет, 2025. – С. 223–226.

4. G-quadruplex structures in the human genome as novel therapeutic targets / J. Bidzinska [et al.]. // Molecules. – 2013. – Vol. 18, № 10. – P. 12368–12395.

5. National Center for Biotechnology Information // National Library of Medicine (US). – URL: https://www.ncbi.nlm.nih.gov/datasets/ (date of access: 19.03.2026).

6. G4IPDB: A database for G-quadruplex structure forming nucleic acid interacting proteins / S. K. Mishra [et al.]. // Scientific reports. – 2016. – Vol. 6, № 1. – P. 38144.

УДК [577.213.3:577.117.2]:618.19-006.6

R. Vuggu

Scientific supervisor: Associate Professor, PhD, A. N. Koval

Educational Establishment
"Gomel State Medical University"
Gomel, Republic of Belarus

INTEGRATED IN SILICO ANALYSIS OF G-QUADRUPLEX FORMING SEQUENCES AND EXPRESSION PATTERNS OF THE *ESR1* GENE IN BREAST CANCER

Introduction

Breast Cancer (BC) is a highly heterogeneous disease, remaining a primary cause of cancer-related mortality worldwide. Its clinical management is dictated by the presence or absence of specific receptors, leading to the classification into four primary molecular subtypes: Luminal A (ER+, PR+, HER2–), Luminal B (ER+, PR+/–, HER2+/–), HER2–Enriched, and Basal-like (Triple-Negative Breast Cancer, TNBC) [1, 2].

The *ESR1* gene encodes the Estrogen Receptor alpha (ERα), a ligand-activated transcription factor. Approximately 80% of breast cancers are hormone receptor-dependent, meaning they rely on estrogen signaling to fuel tumor growth and survival. When estrogen binds to the receptor encoded by *ESR1*, it translocates to the nucleus, binding to Estrogen Response Elements (EREs) to trigger the transcription of genes involved in cell proliferation and anti-apoptosis. *ESR1* expression is inextricably linked to Progesterone Receptor (PR) status, as PR is a downstream target of ER signaling. The co-expression of *ESR1* and PR serves as a functional indicator of an intact estrogen signaling pathway and a critical biomarker for predicting the success of endocrine therapies like Tamoxifen or Aromatase Inhibitors [3].

G-quadruplexes (G4) are four-stranded DNA structures forming in guanine-rich regions, recognized as epigenetic regulators that influence gene expression by modulating transcription factor binding and RNA polymerase processivity [4]. Given the central role of *ESR1* in driving luminal BCs, identifying G4–forming sequences within this gene could uncover novel regulatory mechanisms explaining inter-individual variations in gene expression contributing to disease susceptibility or treatment response.

Goal

To perform an integrated in silico analysis of the human *ESR1* gene by: (1) identifying and characterizing putative G-quadruplex forming sequences (PQS) within the gene, and [2] analyzing its expression patterns across different breast cancer subtypes using public genomic data, to evaluate the potential regulatory significance of G4 structures in the context of gynecological oncology.