

## Bioinformatics analysis of the promoter sequence of the *9f-2.8* gene encoding germin

IZABELA SZUĆKO, URSZULA KOWALSKA, LIDIA SKUZA

University of Szczecin, Institute for Research on Biodiversity, Faculty of Biology, Department of Molecular Biology and Cytology, ul. Wąska 13, 71-415 Szczecin, Poland, e-mail: izabela.szucko@usz.edu.pl

**Keywords** *Triticum aestivum* L., bioinformatic tools, germin protein, *in silico* analyse

**Abstract** Bioinformatics is a field of study having an enormous potential, allowing to solve a number of problems arising as a result of dynamic development of natural sciences with the use of computer science methodologies. It is widely used and constitutes a basis for most scientific research conducted in the field of molecular biology. The aim of this study was *in silico* analysis of the promoter sequence of the *9f-2.8* gene encoding isoform of the germin protein considered as a germination marker in common wheat's (*Triticum aestivum* L.). The gene mentioned above has already been characterized, however, with the use of experimental methods instead of bioinformatics. Analysis with the use of TSSP and TSSPlant software identified the promoter region and classified it as the TATA-box containing promoter. For *9f-2.8* gene including 2.8 kbp, the TSSP software indicated that the TATA-box sequence was located in the position 1665 nt, while the TSSPlant tool showed that TSS [+1] was located in the position 1699 nt. At the second stage, transcription factors were analyzed. Four main families of transcription factors were detected within the analyzed region: MADS, AP2, bZIP and NAC. The most common were MADS-box and bZIP motifs. In the final step of analysis the presence of CpG islands have been checked using the PlantPAN software. The region which could be potentially considered as CpG island have been detected and localized. Software used in analysis above is free online tool.

### Analiza bioinformatyczna sekwencji promotora genu *9f-2.8* kodującego germinę

**Słowa kluczowe** *Triticum aestivum* L., narzędzia bioinformatyczne, białko germina, analiza *in silico*

**Streszczenie** Bioinformatyka jest dyscypliną nauki, w której tkwi olbrzymi potencjał. Dyscyplina ta rozwiązuje wiele problemów powstałych w wyniku dynamicznego rozwoju nauk przyrodniczych przy użyciu metodologii nauk informatycznych. Ma szerokie zastosowanie i jest bazą dla prowadzenia większości badań naukowych z dziedziny biologii molekularnej. Celem artykułu jest analiza *in silico* promotora genu *9f-2.8* kodującego izoformę (*9f-2.8*) białka germiny uważaną za marker kiełkowania u pszenicy zwyczajnej (*Triticum aestivum* L.). Gen ten był już wcześniej scharakteryzowany, jednak do jego analizy korzystano z metod eksperymentalnych, nie obliczeniowych. Analiza bioinformatyczna za pomocą programów TSSP i TSSPlant pozwoliła zidentyfikować promotor i potwierdziła jego klasyfikację do grupy promotorów z motywem TATA-box. W genie *9f-2.8* liczącym 2.8 kbp program TSSP wykazał, że sekwencja TATA-box

znajduje się w pozycji 1665 nt, zaś narzędzie TSSPlant wskazało, że TSS [+1] znajduje się w pozycji 1699 nt. Drugim etapem była analiza czynników transkrypcyjnych. W analizowanym obszarze wyróżniono cztery główne rodziny czynników transkrypcyjnych: MADS, AP2, bZIP oraz NAC, z których najliczniejsze były motywy MADS-box oraz bZIP. Ostatnią częścią analiz była kontrola obecności wysp CpG. Zastosowano program PlantPAN, dzięki któremu zlokalizowano region spełniający warunki, pozwalające uznać go za wyspę CpG. Programy, których użyto do scharakteryzowania tych sekwencji są darmowe i ogólnodostępne online.

## Introduction

Germins are group of proteins present in germinating grains of cereals, among others, in wheat (*Triticum aestivum*) and barley (*Hordeum vulgare*), as well as in some dicotyledonous species, such as *Arabidopsis thaliana* or mustard (*Sinapis arvensis*). However, they mainly occur in the cell wall of monocotyledons. They form an extensive and diverse set of proteins characteristic of plants (Nowakowska, 1998). Germin is characterized by a homopentameric structure and a mass of 125 kDa. Increased concentration of germins, and thus increased expression of genes encoding these proteins is a hallmark of the germination process, hence they are called the germination marker (Nowakowska, 2001). In addition, these proteins are involved in plant defense responses to stress caused by abiotic or biotic factors (Lane, 2000; Davidson et al., 2009). Considering the essence of the functions of this group of proteins, it seems necessary to complement knowledge in this field not only through experimental techniques, but also *in silico* analysis. The rapid development of technology over the past years, which was also associated with enormous progress in bioinformatics, significantly broaden and facilitated the possibilities of conducting analyses. Bioinformatics is based on a variety of mathematical methods, which in a simple and rapid way allow for a detailed analysis of a given sequence. There are many publicly available, free bioinformatic programs that are a great tool for gene analyses (Baxevanisa, Ouellette, 2004, Xiong, 2006, Higgs, Attwood, 2008). Thanks to the use of a number of bioinformatic tools, it is possible to supplement and correct data obtained through the application of experimental methods.

Gene expression is a highly complex and strictly regulated process, because its subsequent stages are closely related and dependent on each other (Szopa et al., 2003). Each stage on the way from the gene to the functional protein can be subject to regulation. In eukaryotes, many cellular processes are regulated at the level of transcription. This process is complicated and we distinguish two main parts in it: transcription initiation and RNA synthesis and processing. The key issue for initiating transcription is to build initiation complexes, whose primary element is the promoter (Molina, Grotewold, 2005). The promoter is a fragment of a sequence lying on a DNA strand, upstream from the transcription start site (TSS) of the gene that has regulatory functions. It is not directly involved in the transcription, but it defines its beginning, direction, time and place (Porto et al., 2014). The identification of promoters and their regulatory elements is one of the main challenges of modern bioinformatics as well as structural and functional genomics, and it allows to predict the expression profiles and location of genes in plants (Rombautus et al., 2003; Porto et al., 2014).

The structure of promoters varies depending on the type of polymerase that will transcribe the gene. However, common elements can be distinguished in all of them, *i.e.*, elements of the basic promoter, *i.e.*, TATA-box, Inr, BRE, DPE, MTE and regulatory elements in its vicinity. They are present upstream the basic sequences. Their presence does not have a decisive influence on the transcription, which is possible even in their absence, but then it occurs with lower

efficiency. There are many factors that determine transcription initiation depending on the type of polymerase involved in this process (Roy, Singer, 2015).

Over 3,000 genes are involved in transcription in plants, and more than half of them encode transcription factors (TFs). The transcription process relies on a number of transcription factors that, through binding to specific DNA sequences, form regulatory regions (Hernandez-Garcia, Finer, 2014). Transcription factors mainly regulate the transcription initiation phase, which is one of the most important points in gene expression regulation. TF regulating effects include interactions with *cis*-acting regulatory elements (CAREs). During this process, TFs function as regulatory *trans*-acting elements that bind to specific *cis*-regulatory elements in the promoters of target genes to activate or repress expression of the target genes. Transcription factors are subject to a complicated classification based on DNA-binding motifs, nevertheless, among plants the most characteristic families include: MADS, AP2, NAC, bZIP, MYB, DATF and WRKY.

In contrast to gene prediction, *in silico* prediction of plant promoters is still underdeveloped; one of the main problems is to define the promoter's location. Although the bioinformatic databases are constantly being updated, there is still a lack of clear and unambiguous descriptions of genomic segments that contain all the elements required for transcription activation. While there are studies on this subject for the model organisms like *Arabidopsis thaliana* or *Medicago truncatula*, it is significantly more difficult to obtain data for plants not belonging to this group.

The following work concerns promoter sequence analysis of the *9f-2.8* gene encoding one of the two germin protein isoforms (9f-2.8) using bioinformatics tools. The previous data on the structure of these genes date back to 1991. Lane and others relied then on tedious methods based on the creation of genomic libraries, restriction enzyme digestions and sequencing.

## Materials and Methods

The *9f-2.8* gene sequence (accession number M63223) deposited in the NCBI database (National Center for Biotechnology Information) was the base material for the bioinformatic analysis. Publicly available bioinformatic programs were used to determine and analyze promoters, *i.e.*, TSSP – in order to determine the promoter region, TSSPlant – TSS position, CisBP – to analyze transcription factors and their DNA binding motifs and PlantPan – to analyze the occurrence of CpG islands.

## Results

Analysis in the TSSP program showed that the promoter belonged to the group of promoters containing the TATA box. The TAT-box of the *9f-2.8* gene (total sequence length – 2822 nt) is located at position 1695 nt. The TSS position (1699 nt) was determined using the TSSPlant program. Sequences characteristic for the binding of transcription factors belonging to four different families were distinguished in the analyzed gene: MADS, bZIP, NAC and AP2. It was observed that factor motifs belonging to the MADS box were the vast majority, mainly: TaMADS#11 and VRN-B1 (Tables 1 and 3).

The distribution of motifs in the promoter sequences is random, and none of the motif families are present only in one region of the analyzed sequence. There are many motifs belonging to the bZIP family. The motifs of the NAC and AP2 families are significantly less common. The sequences of individual motifs are presented in Table 2 and Figure 1.

Analysis of the gene sequence encoding the germin isoform using PlantPAN showed the presence of the CpG island. PlantPAN has found the CpG island in the *9f-2.8* gene, starting at position 1412, with a length of 1234 nucleotides (Table 2).

Table 1. Families of transcription factors in the promoter of the *9f-2.8* gene

Family of TFs	Name	Number of motifs
MADS box	TaMADS#11	11
	VRN-B1	11
bZIP	HBPIA_WHEAT	3
	HBPIB_WHEAT	2
	EmBP-1	3
NAC/NAM	NAC69-1	3
AP2	A5JTT1_WHEAT	5

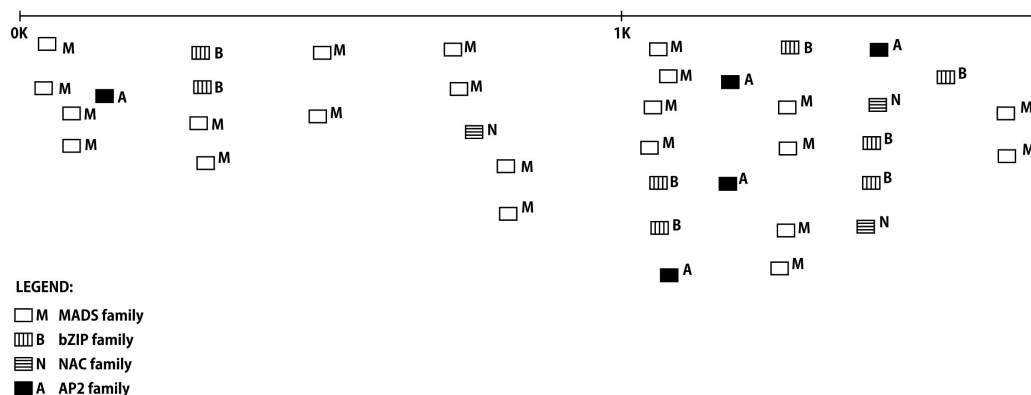


Figure 1. Distribution of motifs in the promoter of the *9f-2.8* gene

Table 2. Motif sequences in the promoter of the *9f-2.8* gene

Name	Sequence
1	2
TaMADS#11	ATATATATGGAAGG CCAAAAAAAAAAT TGCATTTTTGC TATACCGAAAAATGTATT TCTTTTTTTGC ATAAATATCGTACA TTATATATTATTGGCAAC CAGAAACAAAACAAGAAGAGA TTCTCGGCTTTTTCATTTTGC TTTCTCTCTTTTGCT TCCCCTATATAAAGGACT

1	2
VRN-B1	ATATATATGGAAGG CCAAAAAAAAAAT TGCATTTTTGC TATACCGAAAAATGTATT TCTTTTTTGC ATAAATATCGTACA TTATATATTATTGGCAAC CAGAAACAAAACAAGAAGAGA TTCTCGGCTTTTTCATTTGC TTTCTCTTTTGC TCCCCTATATAAAGGACT
HBP1A_WHEAT	TGACACGTTA CGACGTGT TGCCACGCAA
HBP1B_WHEAT	CATGTCGTCA AGTGACGACA
EmBP-1	TGACACGTTA CGACGTGT TGCCACGCAA
NAC69-1	ATTTGACTA CCTTGACTC CACGCAAG
A5JTT1_WHEAT	CGCAGCCC TGGCGCGACG TACGCCGACG CGGCGGCG CAGCCGACG

 Table 3. CpG islands in the *9f-2.8* gene, PlantPAN program

Start	Finish	Length	Frequency C+G
1412	2663	1234	0,53

## Discussion

Regulation at the level of transcription plays the most important role in the activation or inhibition of expression and is largely controlled by gene promoters and *cis* elements present in their region. The structure of gene promoters transcribed by polymerase II has been most accurately characterized so far. The structure of its promoters has been described in many prokaryotic and eukaryotic organisms, *i.e.*, yeast, fruit fly or human (Morton et al., 2014). The majority of research conducted so far on promoters focused on understanding the structure and functioning of animal promoters, therefore, a clear advantage of information about them is observed when analyzing data available in the scientific literature. However, this information forms a foundation for understanding the structure of plant promoters. Analyses are constantly being conducted to find out their structure and properties depending on their sequence, including curved DNA

(Pandey, Krishnamachari, 2006). Model organisms, which during the studies became the starting point for describing plant promoters, included *Drosophila melanogaster*, *Mus musculus* as well as *Homo sapiens* (Butler, Kadonaga, 2002; Smale, Kadonaga, 2003; Riva, 2012). Research on gene promoters in plants is essential for understanding the global mechanism of gene regulation. Performing a series of bioinformatic analyses on the sequence of a gene encoding germin isoform was designed to analyze the promoter sequences as well as to check the effectiveness and reliability of the programs and to supplement the data published in the 1990s.

Comparing the results with those available in the literature, we could see both a number of similarities, but also differences that may have arisen from the application of different analytical tools (Lane et al., 1991, Nowakowska, 1998, 2001). Differences obtained during the analysis of the data obtained with bioinformatic tools are frequently observed. This is influenced by the quantity of data contained in individual databases as well as model organisms, on the basis of which the programs were designed. The PromPredict computing predicting program is often used to predict promoter regions. It has been used to find promoter regions, e.g., in *A. thaliana* or *O. sativa*. The effectiveness and correctness of promoter region predictions using this program reaches 90%. It analyzes differences in DNA stability of neighboring upstream and downstream regions in relation to TSS (Shahmuradov et al., 2017).

The result for determining the promoter region of the *9f-2.8* gene was consistent with the data available in the literature. The TATA-box position was also similar (1 nt shift) (TSSPlant, 1665-; Lane and others 1991, 1664-). Lane et al. (1991) analyzed the promoter regions of genes coding for two germin isoforms in common wheat. They qualified the analyzed promoter structure to TATA-box promoters, which was confirmed in this work using the TSSPlant program. Promoters containing the TATA box are regulated by the action of biotic and abiotic stimuli. Germins, whose intensive synthesis is observed during germination, are a great example of genes containing the TATA-box motif. The germination process in which they participate is strongly dependent on abiotic factors. Lack of optimal environmental conditions, which include temperature, humidity, soil type, water access or adequate light exposure, may result in plant growth inhibition and development. What is more, pathogen attacks, exemplifying biotic factors, also stimulate the expression of germins, involved in plant defense reactions through the participation in cell wall cross-linking.

The analyses localized the TSS at the -1699 nt position. The obtained results meet the condition characteristic for promoters with the TATA box, according to which the start site of transcription is located at a distance of about 30–40 base pairs from the TATA-box. Conventional methods used to determine TSS are based on technologies of low and medium throughput, i.e., EST/cDNA or MPSS modification. For *A. thaliana*, the PEAT analysis was applied to determine TSS sites, using transcript digestions with TAP; it allows detecting potential TSS sites and subsequently analyzing them based on the already known TFBS signals (Morton et al., 2014).

The analysis of the sites responsible for transcription factor binding is the key to understanding expression regulation, because they are mainly responsible for the transcription process to occur. In the following tests, the CIS-BP program was used to classify TFs, thanks to which it was possible to determine CAREs characteristic for plants in the analyzed sequence. Four families, characteristic of plant promoters, were localized in the analyzed region: MADS, AP2, bZIP and NAC. Bioinformatic analysis allowed to reveal their sequences, the distribution of individual families against each other and their location, which indicated their presence in both the distal and proximal region of the promoter. Thirty-eight sequence motifs were distinguished.

Complex interactions between protein and DNA lead to activation, enhancement or suppression of transcription. Germins, in addition to their key role during germination, participate in plant reactions related to stress response. The presence of the Dreb1 motif belonging to the AP2 family of genes encoding germin isoforms demonstrated during the analysis confirmed their involvement in stress reactions (Yaish et al., 2010). The Dreb1 motif was detected in *A. thaliana* in the *Rd29A* gene promoter and was associated with the dehydration-triggered plant response (Hernandez-Garcia, Finer, 2014). Motifs belonging to the NAC family have also been localized in *A. thaliana*, and similarly as AP2, they are involved in responses to drought (Tran et al., 2004). The remaining transcription factor motifs, *i.e.*, MADS-box and bZIP are also related to stress response induced by abiotic factors (Jakoby et al., 2002; Schütze et al., 2008). The MADS family additionally participates in processes related to gametophyte development (Heijmans et al., 2012). Literature data indicate that unidirectional AT repeats are present in the promoter region of the *9f-2.8* gene. Nine sequences known as RY, *i.e.*, purine pyrimidine AuxRE-type sequences (auxin-responsive element), conditioning the response to auxin, can also be distinguished. In addition, there were characteristic sequences, similar to the AuxRE element mentioned above, and PS was one of them. It is a homologous fragment to the *PS-IAA4/5* gene. There were also “TG” sequences in this region. The AS motif is another element present in the promoter of this gene, responsible for binding the ASF-1 protein (Nowakowska, 2001). It is possible to regulate transcription factors using compounds, such as auxins, gibberellins, salicylic acid and certain ions or chlorides (Lane et al., 1991; Nowakowska, 1998, 2001).

The use of the PlantPAN program allowed to localize the CpG island, not only in the promoter region, but in the entire sequence of the analyzed gene. Generally, it is assumed that CpG islands are motifs characteristic of animal genomes (Sakowicz, Frasiński, 2014). However, there are references in the literature and the results of analyses demonstrating the presence of these motifs also in plants. Sequence analysis of the *Arabidopsis thaliana* genome, showed the presence of CG-rich segments. Most segments rich in CpG were associated with genes, hence they can be used as landmarks when identifying genes in plants. CpG motifs present in plants meet the criteria applied to identify animal CpG islands (Ashikawa, 2001). PlantPAN indicated the region recognized as the CpG island in the *9f-2.8* gene at 1412-2663. It means that the CpG island is located partly in the promoter region. The presence of CpG islands in plants has been the object of numerous studies for a long time. Epigenetic modifications, *i.e.*, DNA methylation, chromatin remodeling and histone modifications are hereditary changes that affect the expression of genes, and thus the phenotype of organisms. Among them, DNA methylation has the greatest impact on gene expression in plants and animals. DNA methylation involves CpG dinucleotides and CpNpG sites (N = A, C or T). CpG-rich regions are known as CpG islands. They must meet three conditions to classify them as CpG islands: a) they must contain over 50% of CG dinucleotides, b) the length of the CpG/CpNpG region should be more than 200 bp and c) the ratio of observed to expected dinucleotides should be above 0.5. CpG islands present near TSS can regulate tissue-specific gene expression. The plant genome contains more CpG dinucleotides compared to human DNA. Cytosine methylation on CpG islands in the promoter region has been shown to limit the access of transcription factor binding, which inhibits expression. Cytosine methylation patterns are not static, they change with development or under the influence of environmental conditions in the whole plant genome. DNA methylation plays a significant role in plant embryogenesis, seed development, in regulating the immune response to pathogenic infections, environmental adaptations and resistance to stress. Methylation errors cause defects in embryogenesis, *i.e.*, abnormal cell divisions or partial sterility as well as development retardation and plant size reduction.

CpG/CpNpG analyses indicate the presence of CpG/CpNpG islands in the second half of the promoter region (3' end) in all *OsPRs* genes, except for *OsPR2*, whereas they are absent in *AtPR*, indicating that the CpG islands are present in *Oryza sativa* genome in the *PR* genes (pathogenesis related), whereas they are missing in the *PR* genes in *A. thaliana*. The CpG islands observed only in the *PR* genes of rice indicate that monocotyledonous genomes contain more GC motifs than dicotyledonous ones (Kaur et al., 2017).

Knowledge of the promoters' architecture is crucial for understanding the transcription regulation, which is fundamental to basic life processes. More detailed molecular analyses of these sequences are connected with the isolation of promoter sequences and associated elements, which is a critical point during the regulation of the introduced transgenes, including genes encoding proteins and non-coding sequences involved in gene silencing through the RNAi mechanism.

## References

- Ashikawa, I. (2001). Gene-associated CpG islands in plants as revealed by analyses of genomic sequences. *The Plant Journal*, 26 (6), 617–25.
- Baxevanisa, A.D., Ouellette'a, B.F.F. (2004). *Bioinformatyka. Podręcznik do analizy genów i białek*. Warszawa: Wydawnictwo Naukowe PWN.
- Butler, J.E.F., Kadonaga, J.T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes & Development*, 16, 2583–2592.
- Davidson, R.M., Reeves, P.A., Manosalva, P.M., Leach, J.E. (2009). Germins: A diverse protein family important for crop improvement. *Plant Science*, 177, 499–510.
- Heijmans, K., Morel, P., Vandenbussche, M. (2012). MADS-box genes and floral development: the dark side. *Journal of Experimental Botany*, 63 (15), 5397–5404.
- Hernandez-Garcia, C.M., Finer, J.J. (2014). Identification and validation of promoters and cis-acting regulatory elements. *Plant Science*, 2017–2018, 109–119.
- Higgs, P.G., Attwood, T.K. (2008). *Bioinformatyka i ewolucja molekularna*. Warszawa: Wydawnictwo Naukowe PWN.
- Jakoby, M., Weisshaar, B., Dröge-Laser, W., Vicente-Carbajosa, J., Tiedemann, J., Kroj, T., Parcy, F. (2002). bZIP transcription factors in *Arabidopsis*. *Trends in Plant Science*, 7 (3), 106–111.
- Kaur, A., Pati, P.K., Pati, A.M., Nagpal, A.K. (2017). *In-silico* analysis of cis-acting regulatory elements of pathogenesis-related proteins of *Arabidopsis thaliana* and *Oryza sativa*. *Plos one*, 12 (9), e0184523. <https://doi.org/10.1371/journal.pone.0184523>.
- Lane, B.G., Bernier, F., Dratewka-Kos, E., Shafai, R., Kennedy, T.D., Pyne, C., Munro, J.R., Vaughan, T., Walters, D., Altomare, F. (1991). Homologie between Members of the Germin Gene Family in Hexaploid Wheat and Similarities between These Wheat Germins and Certain Physarum Spherulins. *The American Society for Biochemistry and Molecular Biology*, 266 (16), 10461–104696.
- Lane, B.G. (2000). Oxalate oxidases and differentiating surface structure in wheat: germins. *Biochemical Journal*, 349, 309–321.
- Molina, C., Grotewold, E. (2005). Genome wide analysis of *Arabidopsis* core promoters. *BMC Genomics*, 6, 25.
- Morton, T., Petricka, J., Corcoran, D.L., Li, S., Winter, C.M., Carda, A., Benfey, P.N., Ohler, U., Megraw, M. (2014). Paired-End analysis of Transcription Start Sites in *Arabidopsis* reveals plant-specific promoter signature. *The Plant Cell*, 26, 2746–2760.



- Nowakowska, J. (1998). Gene expression and oxalate oxidase activity of two germin isoforms induced by stress. *Acta Physiologiae Plantarum*, 20 (01), 10–33.
- Nowakowska, J. (2001). Germiny u roślin wyższych. *Biuletyn Instytutu Hodowli i Aklimatyzacji Roślin*, 217, 19–27.
- Pandey, S.P., Krishnamachari, A. (2006). Computational analysis of plant RNA Pol-II promoters. *Biosystems*, 83, 38–50.
- Porto, M.S., Pinheiro, M.P.N.P., Batista, V.G.L.B., dos Santos, C., Albuquerque Melo Filko, P., de Lima, L.M. (2014). Plant Promoters: an approach of structure and function. *Molecular Biotechnology*, 56, 38–49.
- Riva, A. (2012). The MAPPER2 Database: a multi-genome catalog of putative transcription factor binding sites. *Nucleic Acids Research*, 40, 155–161.
- Rombauts, S., Florquin, K., Lescot, M., Marchal, K., Rouze, P., Van de Peer, Y. (2003). Computational Approaches to Identify Promoters and cis Regulatory Elements in Plant Genomes. *American Society of Plant Biologists*, 132, 1162–1176.
- Roy, A.L., Singer, D. (2015). Core promoters in transcription: old problem, new insights. *Trends in Biochemical Sciences*, 40 (3), 165–171.
- Sakowicz, T., Frasiński, S. (2014). Roślinne promotory polimerazy RNA II – struktura i identyfikacja. *Postępy Biologii Komórki*, 21 (2), 347–360.
- Schütze, K., Harter, K., Chaban, C. (2008). Post-translational regulation of plant bZIP factors. *Trends in Plant Science*, 13 (5), 247–255.
- Shahmuradov, I.A., Gammerman, A.J., Hancock, J.M., Bramley, P.M., Solovyev, V.V. (2003). PlantProm: a database of plant promoter sequences. *Nucleic Acids Research*, 31 (1), 114–117.
- Smale, S.T., Kadonaga, J.T. (2003). The RNA Polymerase II core promoter. *Annual Review of Biochemistry*, 72, 449–479.
- Szopa, J., Czuj, T., Łukaszewicz, M. (2003). Analiza genów rodziny 14-3-3. *Biotechnologia*, 3 (62), 95–106.
- Tran, L.P., Nakashima, K., Sakuma, Y., Simpson, S.D., Fujita, Y., Maruyama, K., Fujita, M., Seki, M., Shinozaki, K., Yamaguchi-Shinozaki, K. (2004). Isolation and Functional Analysis of Arabidopsis Stress-Inducible NAC Transcription Factors That Bind to a Drought-Responsive cis-Element in the early responsive to dehydration stress 1 Promoter. *The Plant Cell*, 16, 2481–2498.
- Xiong, J. (2006). *Podstawy bioinformatyki*. Warszawa: UW.
- Yaish, M.W., El-kereamy, A., Zhu, T., Beatty, P.H., Good, A.G., Bi, Y., Rothstein, S.J. (2010). The APETALA-2-Like Transcription Factor OsAP2-39 Controls Key Interactions between Abscisic Acid and Gibberellin in Rice. *PLoS Genetics*, 6 (9).

**Cite as:** Szućko, I., Kowalska, U., Skuza, L. (2018). Bioinformatics analysis of the promoter sequence of the 9f-2.8 gene encoding germin. *Acta Biologica*, 25, 131–139. DOI: 10.18276/ab.2018.25-11.