UDC 577.21:632.4

THE HOMOLOGOUS IDENTIFICATION OF THE STEM RUST RESISTANCE GENES *Rpg5*, *Adf3* AND *RGA1* IN THE RELATIVES OF BARLEY

B.V. IVASCHUK [1], D.O. SAMOFALOVA [1], Ya.V. PIRKO [1], G. FEDAK [2], Ya.B. BLUME [1]

[1] Institute of Food Biotechnology and Genomics, NAS of Ukraine, Kyiv
E-mail: antutilo@gmail.com, cellbio@cellbio.freenet.viaduk.net
[2] Eastern Cereal and Oilseed Research Centre, Agriculture and Agrifood Canada, Ottawa

The barley genes Rpg5, RGA1 and Adf3, which provide a strong resistance to many pathotypes of stem rust, were cloned a few years ago, but it was still unclear whether their homologues were represented in wheat and in related species. The paper describes the results of a bioinformatic research to determine the homologues of Rpg5, RGA1 and Adf3 in the genomes of Triticum aestivum and several wild grasses, which breeders usually use as sources of stem rust resistance, and which are available in the genome databases. It was found that the Th. elongatum sequence Q9FEC6 and T. aestivum sequence Q43655 were the high identical homologues of the Adf3 sequence. T. urartu M8A999 sequence and T. aestivum W5FCU1 sequence were found to be the closest homologues of Rpg5 complete protein sequence, but the identity of their kinase domains were not as clear as that of the other domains. The separate Rpg5 kinase part analysis did not provide the strong evidences that its orthologs were presented in our corn species. T. urartu M7ZZX9 sequence and T. aestivum W5FFP0 and W5FI33 sequences were showed to be the homologues of RGA1. The analysis of the predicted active sites allowed finding out the difference between sequences of Rpg5, RGA1, Adf3 protein and their homologues.

Keywords: *stem rust (Puccinia graminis), stem rust resistance genes, barley (Hordeum vulgare), homologous genes.*

Introduction. *Puccinia graminis* is a biotrophic fungus that causes a disease called stem rust and is one of the most devastating parasites of such cultural crops as wheat, barley and rye. Like its related species – *Puccinia stritiformis* and *Puccinia recondite*, it can lead to a significant loss of yield of above mentioned crops during the mass epiphytoties. Its race Ug99, subspecies *Puccinia graminis f. sp. tritici*, has become a major threat to wheat production in Africa and West Asia, and the resistance genes that can protect the vulnerable cultures from the stem rust are the main subjects of the research which many institutions provide in this area [1].

Today we know three genes from *Hordeum vulgare* that have been cloned. Together they provide resistance to Ug99. They are *Adf3*, *Rpg5*, *RGA1* [2]. The description of each gene and their protein products is presented below.

Rpg5 gene has a length of ~ 8.5 kb and consists of seven exons which encode the protein product with a length of ~ 1.38-AA. The protein consists of three domains: NB – nucleotide-binding domain; LRR – domain with leucine rich repeats; PK – protein kinase domain [3]. The kinase part of the Rpg5 protein has a predicted dual serine/threonine (Ser/Thr) motif and possibly a tyrosine activity [2]. Ser/Thr kinases are involved in many processes in the cell, including a participation in the signaling pathways of immunity. More details are required to describe the main features of this domain and its protein family. In the catalytic domain of plant Ser/Thr PK 11 major conservative subdomains are distinguished. The subdomain VI may contain the consensus sequence DLRAAN or DLAARN that indicates tyrosine specificity, whereas consensus DLKPEN is an indicator of Ser/Thr specificity. It was shown that subdomain VIII may contain a conservative tyrosine specific consensus PI/VK/RWT/MAPE and a less conservative Ser/ThrGT/SXXY/FXAPE. The consensus CW (X) 6RPXF in a subdomain XI is a feature of a tyrosine kinase. The investigation of all predicted *Arabidopsis thaliana* kinases has revealed their Ser/Thr and Tyr specificities and has left a question about the existence of the kinases that could fulfill exclusively a phosphorylation of Tyr [4].

Rpg5 amino acid sequence contains a predicted NB (nucleotide binding) domain. The NB domain is a typical constituent of many discovered receptor proteins that provides resistance to different diseases [5]. The type of proteins containing this domain includes a family of NTPases (STAND signal transduction ATPases with numerous domains). The STAND proteins could function like molecular switches in the signal defense pathways. The spatial reconstruction of the R proteins with NB domains

(of TNLs and CNL classes) confirms the suggestion that this domain has three subdomains: NB, ARC1 and ARC2 (the quantity of ARC subdomains in certain cases can vary) [28]. The major part of the conserved patterns lies on a surface of these three subdomains. The motif of P-loop (GVGKTT), the motif of anchor region, the motif MHDV (especially the His residue), the motif GLPL and ATP site are parts of the catalytic center. It is considered that ATP hydrolysis lead to the conformational changes of the R protein and such changes affected the downstream signals [6, 7].

The LRR (leucine rich repeats) domain participates in the protein-ligand interactions. This domain was found in the many types of prokaryotic, eukaryotic and virus proteins, especially in hormone receptors, inhibitors of the ferments and immune receptors of plants and animals. Several articles report that LRR proteins not only take part in the early stages of development of mammals, including the development of the nervous system, in the regulation of gene expression but also play an important role in the apoptosis signaling and cytoskeleton dynamics. Generally this domain consists of a certain number of repeats the quantity of which can vary. The length of one repeat is 20–29 residues, 11 of which are conserved. The consensus sequence LxxLxLxxN/CxL in the non-conserved positions marked as x, can contain any amino acid, while in the conserved positions marked with letter L can contain leucine as well as valine, isoleucine or phenylalanine residues. Also it should be noted that the different organisms have the different consensus sequences of LRR repeats. The plants` LRR proteins is believed to have the consensus sequence LxxLxLxxNxL(t/s)GxIPxxLGxx [8, 9].

RGA1 gene has a length of 4,115 bp; the length of its protein product is 895-AA. The protein has a typical plant immunity receptor proteins structure - it consists of the predicted NB and LRR domains. It seems that these domains are similar to those of this protein family, but analysis of RGA1 NB domain did not reveal its close homologues [10].

Adf3 gene has a length of 634 bp, and its protein product has 144-AA. The gene belongs to an actin depolymerization (ADF) factor family. It is believed that ADF proteins together with other actin-related proteins regulate actin filament dynamics that lead to cytoskeleton reorganization [10]. Recent researches have confirmed finally the significant role of ADF/cofilin protein family in the immune response and have proposed the possible mechanisms of the resistance of plants to the pathogenic organisms. One of the assumptions is that *A. thaliana* gene A *DF4* is involved in resistance against *Pseudomonas syringae* DC3000, and perhaps the product of this gene is a target of *P. syringae* factor AvrPphB. The investigations prove that there is a significant decrease in the expression of RPS5 (the protein product of another resistance gene) in plants with the mutated gene *ADF4*. Perhaps *ADF4* takes part in the signal delivery to a nucleus that activates the expression of *RPS5*. Importan9 and ADF4 are actively transported into the nucleus, where they are supposed to be involved in expression regulation and localization of actin in the nucleus. Phosphorylation of *ADF4* Ser-6 may regulate its activity and its ability to interact with actin filaments and *RPS5* [11]. Another ADF factor – *Triticum aestivum* TaADF7 accelerates cell death, which is caused by *Nicotiana benthamiana Bax* gene. The *TaADF7* gene which was a knockout in wheat plants resulted in their increased susceptibility to Pst (*P. graminis f. sp. tritici*). In the leaves of TaADF7 knockout plants the fungal hyphae are longer. In these plants, the disease increases the *SOD* and *CAT* genes expression, and vice versa – reducing the *TaNOX* gene expression, which is responsible for the generation of ROS. Thus, the reduced accumulation of reactive oxygen species (ROS) can be observed [12].

The overview of the genes *Adf3*, *Rpg5*, *RGA1* implies that they could not only provide a comprehensive resistance to stem rust due to their complex structure and multi-domain organization (*Rpg5*, *RGA1* protein) but also participate in unknown processes that provide resistance to various pathogens.

The aim of the present study was to determine whether the homologues of these genes were presented in such species as *Aegilops tauschii*, *Triticum urartu, Triticum monococcum* and also in the plants with an available genome data base like *Sorghum bicolor, Setaria italic, Oryza glaberrima* etc. Special attention was also given to the search of the above mentioned gene homologues in *Triticum aestivum* genome because of its exceptional importance as a crop culture. We supposed that *Rpg5* homologues and those of the genes *RGA1* and *Adf3* could be present in wheat genomes. We have been supporting this idea with three facts: three genes are inherited

together; both barley and wheat can be infected by the same races of *P. graminis f. sp. tritici*; the tight evolution of the pathogen and its host lead to the emerging of similar or the same mechanisms of the defense. Our investigation confirms that *Triticum aestivum* has highly identical homologues of the *Adf3* and *RGA1* and that the analog of *Rpg5* is also represented in its genome. We speculate that the wild relatives of wheat *Thinopyrum ponticum* and *Elytrigia repens* or other species, which in the past had been used often as sources of resistance to stem rust, have genes similar to *Rpg5*, *RGA1* and *Adf3* and that these genes could be used to fight new races of stem rust in the future.

Methods. The complete amino acid sequences of the protein products of *Rpg5, RGA1* and *Adf3* genes, and their homologues were obtained using a database of the protein sequences UniPronKB (www.uniprot.org) [13]. BLASTp-scan (BLASTP 2.2.29+) of the database was performed according to the standard parameters [14]. The homologues were selected taking into account the index identity (not less than 30 %) and the completeness of the amino acid sequences. The domain architecture of the proteins were studied with the help of on-line tool SMART [15] (http://smart.embl-heidelberg.de/) and HMMER [16] (http://hmmer.janelia.org/).

The multiple alignments were performed using server T-Coffee (http://www.tcoffee.org/) and applying the two methods: clustalw_msa and t_coffee_msa [17]. The multiple alignments were edited and the gaps were extracted using the Jalview [18]. The phylogenetic analysis was performed using the software package MEGA6.06 [19]. To check the reliability of the OTUs (Operational Taxonomic Units) placement, the bootstrap test was conducted (1000 replica) [20]. A calculation of the evolutionary distances was carried out with the Jones-Taylor-Thornton (JTT) method [21]. The cladistic analysis was performed using Neighbor-Joining method [22]. We modeled differences in the rate of the substitutions among the sites using a gamma distribution.

In addition, we provided the analysis of the substitutions in the animo acid sites of the determined sequences which could be involved in the intermolecular interactions.

At first, the spatial structure of Adf3 protein and those of the separate domains of Rpg5 and RGA1 proteins were reconstructed. For this purpose, an on-line tool Phyre2 was used (http://www.sbg.bio.ic.ac.uk) [23]. After that we found the annotated x-Ray structures in the database of Phyre2 that were the most similar to the reconstructed models that we had obtained for Adf3, Rpg5 and RGA1. We used the x-Ray structures d1ak7a_ та d1f7sa_ to predict the active sites for Adf3 and c2lxxA_ and c2i2qA_ to predict the active sites for Adf2. The x-Ray structure c2a5yB_ and the x-Ray structures c4bstB_ and c3rgxA_ were used to predict the active sites in NB and LRR parts of Rpg5 respectively. The model c3iz8H_ was used to identify the active sites in the NB part of RGA1and the models c4fcgA_ and d1ogqa_ to do the same for RGA1 LRR domain.

Next, the possible protein interaction sites of Adf3, Rpg5 and RGA1 proteins with other molecules were searched using Phyre Investigator (http://www.sbg.bio.ic.ac.uk). We marked the sites of the interactions in the main sequences which were the sequences of Adf3, Rpg5 and RGA1 proteins. After that, we compared the sequences of Adf3, Rpg5, RGA1 and those of their homologues in the predicted active sites using the multiple alignments. Also, we used for analysis the literature data which describe the main motifs of NB-LRR and protein kinase classes of the proteins. Only the full sequences of the selected domains were chosen for the analysis; we removed the gaps from the alignment only in the case if they were in the amino acid sequences of Adf3, Rpg5 and RGA1.

Results and Discussion. ***Search for Adf3, Rpg5 and RGA1 homologues.*** *Adf3 homologues.* We used UniProtKB database to find Adf3 protein homologues. M4THI1 (UniProtKB ID) was chosen as the query sequence. It was found that the sequences Q9FEC6 from *Th. elongatum* (96 % identity) and Q43655 from *T. aestivum* (92 % identity) were the homologues of M4THI1 actin depolimerization factor. In addition, the set of the previously uncharacterized sequences was obtained, which was supposed to be the homologues of Adf3 from other plants. The set contained: I1GLK1 from *Brachypodium distachyon* (69 % identity), K4AM22 from *Setaria italica* (68 % identity), I1PGT7 from *Oryza glaberrima* (66 % identity), C5WV16 from *Sorghum bicolor* (65 % identity), M5XE35 from *Prunus persica* (55 % identity), M0SXV6 from *Musa acuminate subsp. malaccensis* (56 % identity).

It have been believed that gene *Adf2* (*Rpg4*) in the complex with *Rpg5* ensures resistance to Ug99 [24],

but recently it was shown that actually the resistance was provided by *Adf3*, not the *Adf2* [10]. That is why we decided to conduct the search of *Adf2* homologues too, and after that we found out the differences between Adf3 and Adf2. The most identical homologues of Adf2 (UniProtKB: B5LQU6) were the sequences W5FCA4 from *T. aestivum* (98 % identity) and J3LU77 from *Oryza brachyantha* (81 % identity). The remaining sequences were selected using the same pattern as for Adf3. These sequences were I1GLK1 from *B. distachyon* (80 % identity), K4AM22 from *S. italica* (80 % identity), C5WV16 from *S. bicolor* (76 % identity), M5XE35 from *P. persica* (55 % identity), Q9FEC6 (74 % identity) from *T. elongatum*, I1LGG5 from *Glycine max* (74 % identity). It is clearly seen from these data that all found sequences are almost identical for both Adf2 and Adf3, and the reason of this is a high identity between Adf3 i Adf2 (74 % identity). The results of the cladistic analysis are presented at the Fig.1, *A*, *B*. A bootstrap test confirmed the relationship of *Adf3* and *Adf2* with their homologues.

Rpg5 homologues. The next step of our investigation was the search of Rpg5 protein homologues. As a search query the not annotated amino acid sequences M9WLW5 of Rpg5 from *H. vulgare* was chosen. The profile analysis revealed only two sequences with high identity (76 %) to the given protein: potential protein kinase of APK1B type – M8A999 from *Triticum urartu* and not characterized protein W5FCU1 from *T. aestivum*. The analysis of the domain architecture of Rpg5 homologues with the instrument SMART showed the existence of three domains. To expand a sample we used the amino acid sequences of every domain of Rpg5 for a search of the additional sequences that could be the candidates to be the Rpg5 homologues.

Carrying out the search using the sequence of a protein kinase domain, we found several candidates with high identity to its sequences: R7WEV1 from *A.taushii* (94 % identity), Q9LFP7_Y5158 from *A. thaliana* (83 % identity), T1NE20 from *T. urartu* (87 % identity), W5EUI0 from *T. aestivum* (94 % identity), I1IPJ8 from *B. distachyon* (93 % identity). The results of cladistic analysis showed that the kinase domain sequence of Rpg5 protein is not related to those of M8A999 from *T. urartu* and W5FCU1 from *T. aestivum*. This indicates that the divergence between the ancestor of *H. vulgare* and the ancestors of *T. urartu* and *T. aestivum* probably took place earlier than between ancestors of those of *T. urartu* and *T. aesivum*. Although, the PK domain sequence of Rpg5 is located in the cluster together with its highly identical homologues, the bootstrap test rated this OTU (PK Rpg5) as unreliable (Fig. 1, *G*).

Using for the search the sequence of nucleotide-binding (NB) domain of Rpg5 protein we were able to identify in UniprotKB database in addition to the previously mentioned sequences M8A999 (81 % identity of NB domain and 76 % identity of LRR domain with those of Rpg5) and W5FCU1 (77 % identity of NB domain and 74 % identity of LRR domain with those of Rpg5), also such uncharacterized sequences: C5Y1I6 (49 % identity) from *S. bicolor*, I1HE45 (52 % identity) from *B. distachyon*, K3ZL04 (50 % identity) from *S. italica*, M0XND7 (86 % identity) from M0XYV5 (98 % identity) from *H. vulgare*, M8C4A0 (60 % identity) from *A. tauschii*, T1LD43 (82 % identity) from *T. urartu*, W5FPX7 (80 % identity) from *T. aestivum* (Fig. 1, *C*).

Using the sequence of the LRR part of the Rpg5 protein as a query we revealed that the sequence occurred in the following homologues: C5YSD6 from *S. bicolor* (50 % identity), F2E044 from *H. vulgare* (58 % identity), I1IJB3 (49 % identity) and I1IK46 (66 % identity) from *B. distachyon*, J3M864 from *O. brachyantha* (52 % identity), K3ZQ30 from *S. italica* (50 % identity), T1LD43 from *T. urartu* (80 % identity), R7VZB5 from *A. tauschii* (56 % identity), W5D8L8 (59 % identity) and W5FE25 (78 % identity) from *T. aestivum* (Fig. 1, *D*). In addition, we found the sequences that have high identity to the NB-LRR part of Rpg5; they are T1LD43, M0XND7 and M0XYV5. It should be noted that the sequences M8A999, W5FCU1 and Rpg5 have highly identical NB-LRR part of their proteins, but they differ greatly in their protein kinase domains. Also the sequences W5FPX7 and W5FE25 from *T. aestivum* are highly identical to the sequences of Rpg5 NB and LRR domains respectively.

RGA1 homologues. Further we initiated the homologue search to detect *RGA*1 homologues. B5B0D1 protein from *H. vulgare* (unannotated RGA1) was used as a query sequence. Such homologues of RGA1 were revealed: A9UKM2 (51 % identity) from *H. vulgare,* M8BTT3 (51 % identity) from *A. tauschii*, I1I6Q8 (62 % identity) from *B. distachyon*, M7ZZX9 (94 % identity) from *T. ura-*

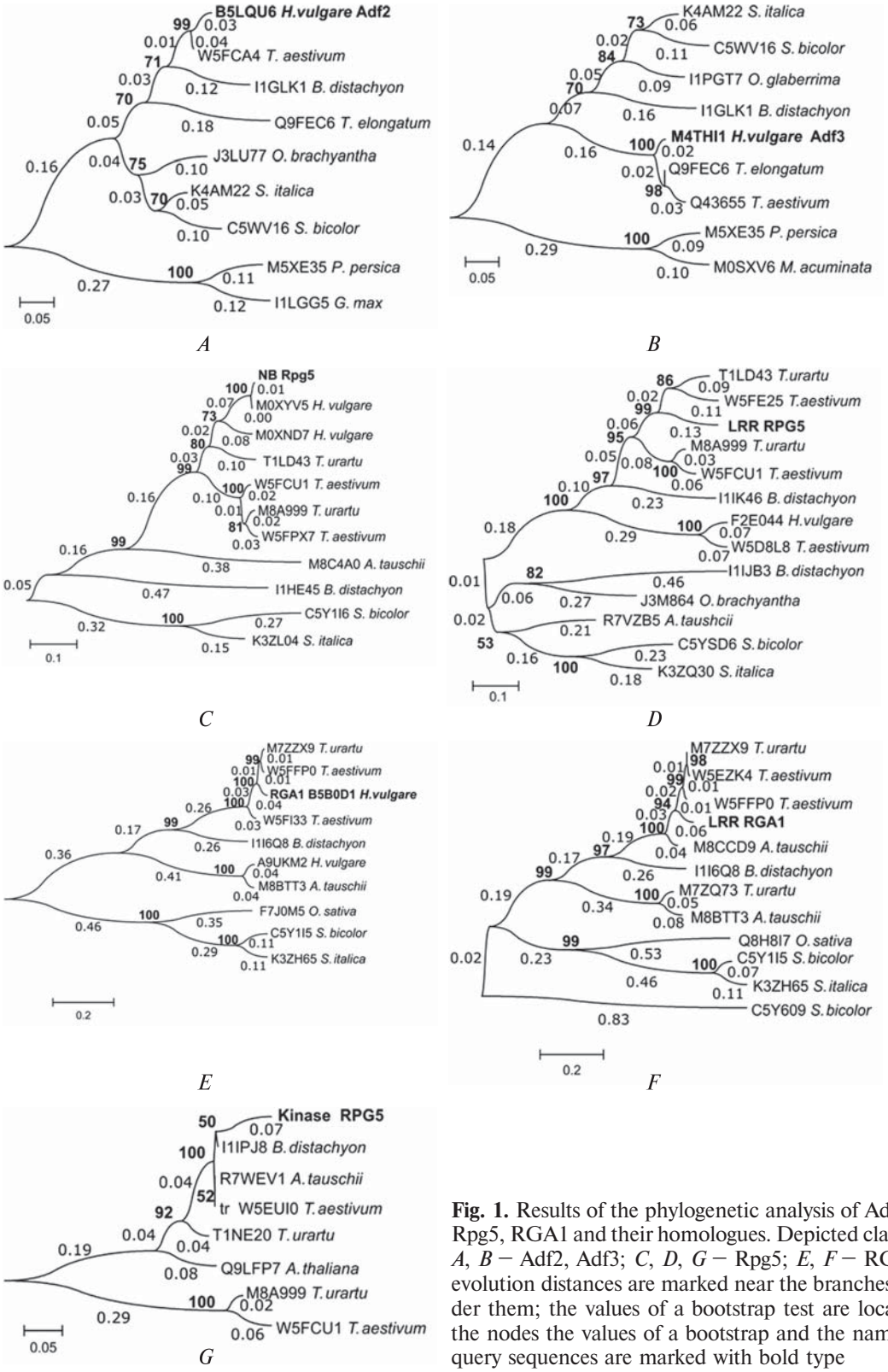

Fig. 1. Results of the phylogenetic analysis of Adf3, Adf2, Rpg5, RGA1 and their homologues. Depicted cladograms: *A*, *B* – Adf2, Adf3; *C*, *D*, *G* – Rpg5; *E*, *F* – RGA1. The evolution distances are marked near the branches and under them; the values of a bootstrap test are located near the nodes the values of a bootstrap and the names of the query sequences are marked with bold type

rtu, W5FFP0 (94 % identity) and W5FI33 (90 % identity) from *T. aestivum*, F7J0M5 (39 % identity) from *Oryza sativa* subsp. *indica*, K3ZH65 (39 % identity) from *S. italica*, C5Y1I5 (38 % identity) from *S. bicolor* (Fig. 1, *E*). In addition, we conducted the separate searches of the homologues for RGA1 using as the queries the sequences of LRR and NB domains, but we could not reveal the complete sequences, which corresponded to NB part of RGA1, but we were able to find them for the LRR part. We discovered that the sequences highly identical to LRR part of RGA1 were: M0XSU6 (100 % identity) from *H. vulgare*, M8CCD9 (87 % identity) from *A. tauschii*, W5FFP0 (92 % identity) та W5EZK4 (91 % identity) from *T. aestivum*. The sequences with lower identity are not described, but one can find them on the cladogram on the Fig. 1, *F*. The brunch topology of the all trees was reliable, so we have used these trees in a farther analysis.

Analysis of the functional sites of RGA1, Rpg5, Adf3 and their homologues. The last stage of our research was the analysis of the amino acid sequences of RGA1, Rpg5, Adf3 and their homologues in the sites that could be involved in the intermolecular interactions.

Analysis of Adf2 and Adf3. Generally, the amino acid differences in sequences of Adf2, Adf3 and their homologues are respectively located in the loops, between the elements of the secondary structures. All predicted F- and G-actin binding sites are conserved and they do not contain substitutions.

It should be noted that the residues Pro and Gly located in the positions 130–131 (on a scale of multiple sequence alignment rather than separate sequences) of Adf2 sequence, are situated between F-actin (129, 132) binding sites, and in Adf3 in these positions Ala and Ser are situated. Perhaps they, especially Pro, cause the structural differences between Adf2 and Adf3 proteins, which in turn affect the interaction with F-actin. Among the set of the predicted functional sites the sequences Adf3 and Adf2 also differ in the positions: E47D, V62T, S63V, R102K, T111N, L114F (Fig. 2, *A*).

Analysis of Rpg5. It was found that a large number of the conserved sites in the Rpg5 protein kinase domain (M9WLW5) sequence differ from that of the homologues. We assume that in this case the protein kinase part of Rpg5 in some sense is a paralog in relation to the other sequences, and fulfils some other functions. As shown at the Fig. 2, *C*, the main difference between protein kinase domain Rpg5 sequence and those of the sequences of its homologues is the presence of Gly in the position 50 in the motif of the catalytic loop (YGDFRTSN).

The homologues in this position have the conserved amino acid Arg. Rpg5 PK sequence has also Glu in the position 73 in the activation loop motif. The homologues have Ala in this position. The PK Rpg5 has Ile and Arg in the positions 82–83 which are placed in the area of the so-called P + 1 loop (not marked), but the other sequences have Ser/Thr and His respectively in these sites. Together these differences may affect the substrate specificity of Rpg5 protein and thus they distinguish it among other relative kinases. In addition, the multiple sequence alignment shows clearly why the PK domain sequences of M8A999 from *T. urartu* and W5FCU1 from *T. aestivum* locate on the tree apart from the other sequences – in many positions, especially in the important motifs; they differ from the rest of the homologues. Overall, PK of Rpg5 has conventional catalytic motifs that are the main features of Ser/Thr-specific kinases, to which the product of this gene belongs.

The sequence of NB domain of Rpg5 does not differ from that of its most identical homologue M0XYV5 from *H. vulgare*. However, the sequences of NB part of Rpg5, M0XND7 and M0XYV5 have many differences in relation to the rest homologues in the positions 69(G), 70(T), 72(H), 82(S), 156(G), 161(V), which are situated near the active sites (Fig. 2, *B*). Rpg5 LRR part differs from those of its homologues in 22 positions without examining each separately (Fig. 3, *A*).

Analysis of RGA1. To analyze the active sites of NB domain, we used the homologues obtained from the full-length sequence of RGA1, and to analyze LRR domain, we chose not only the homologues of a full length RGA1 sequence but also the sequences that were found only for the LRR domain.

We did not find any differences between RGA1 NB domain and those of the homologues M7ZZX9 and W5FFP0. Only the sequence W5FI33 differed from the rest of the homologues in position 113 and 146 (Fig. 3, *B*). The sequence of RGA1 LRR part has 11 positions that differ from the remaining set, except the sequence MOXSU6 which is identical in the active sites to RGA1 (Fig. 3, *C*). The homologues have differences only in two active sites which were predicted equally for both models.

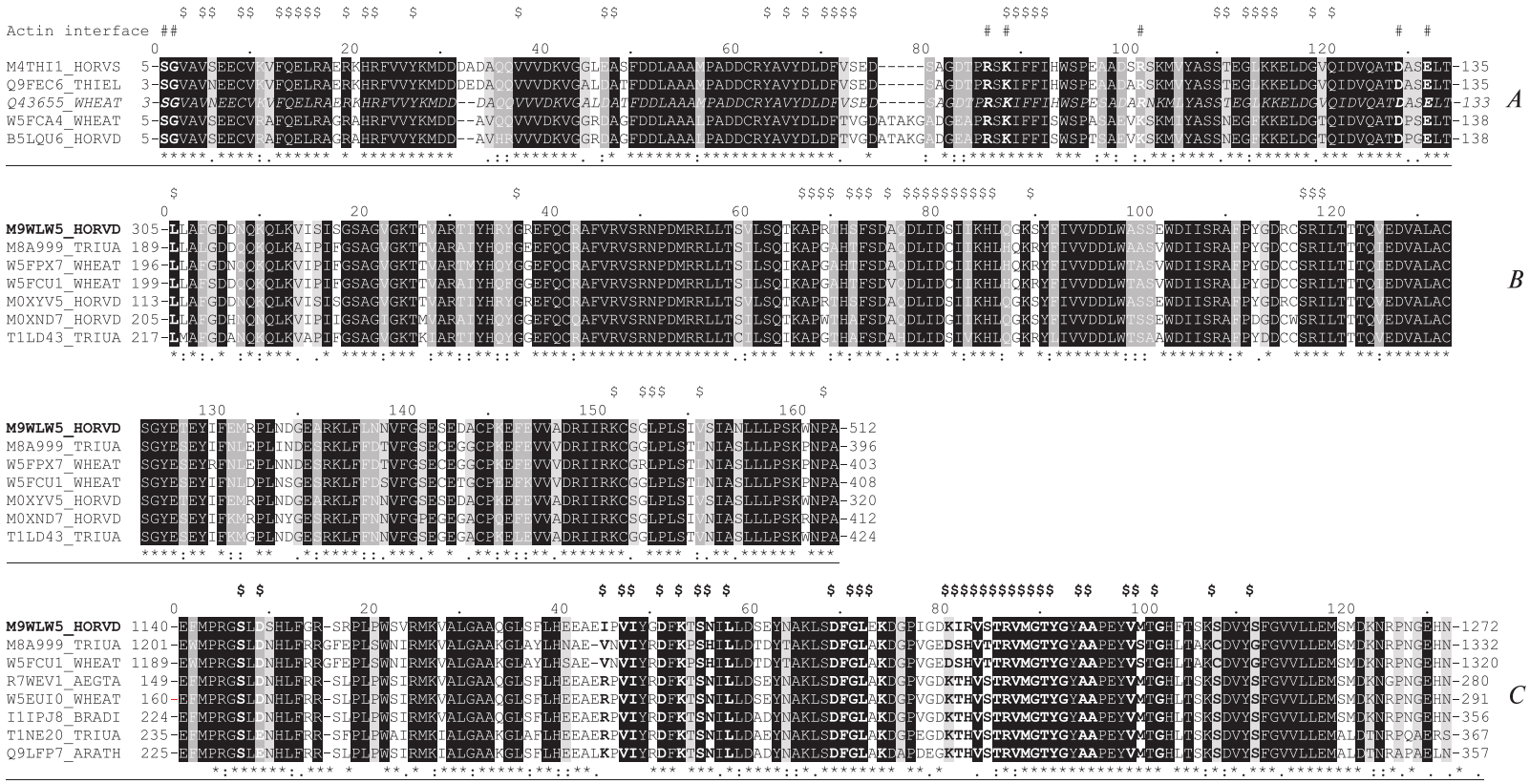

Fig. 2. The fragments of the multiple alignments of Rpg5, Adf3, RGA1 and their homologues: *A* – a multiple alignment of Adf2, Adf3, # - this symbol marks the residues that can interact with actin; *B* – a multiple alignment of Rpg5 NB domain (SMART: PF00931) and those of his homologues; *C* – a multiple alignment of Rpg5 PK domain (SMART: SM000221) and those of his homologues. The notes: the identical residues is marked with black color, the similar residues – dark gray and gray color, the variable residues – white; */20 – the symbols that mark the ordinal numbers of the residues in the multiple alignments; $ – a symbol that marks the predicted residues which could take part in the intermolecular interactions; the numbers near the N- and C-ends of the sequences show their location in the full length amino acid sequence that deposited in the UniProt

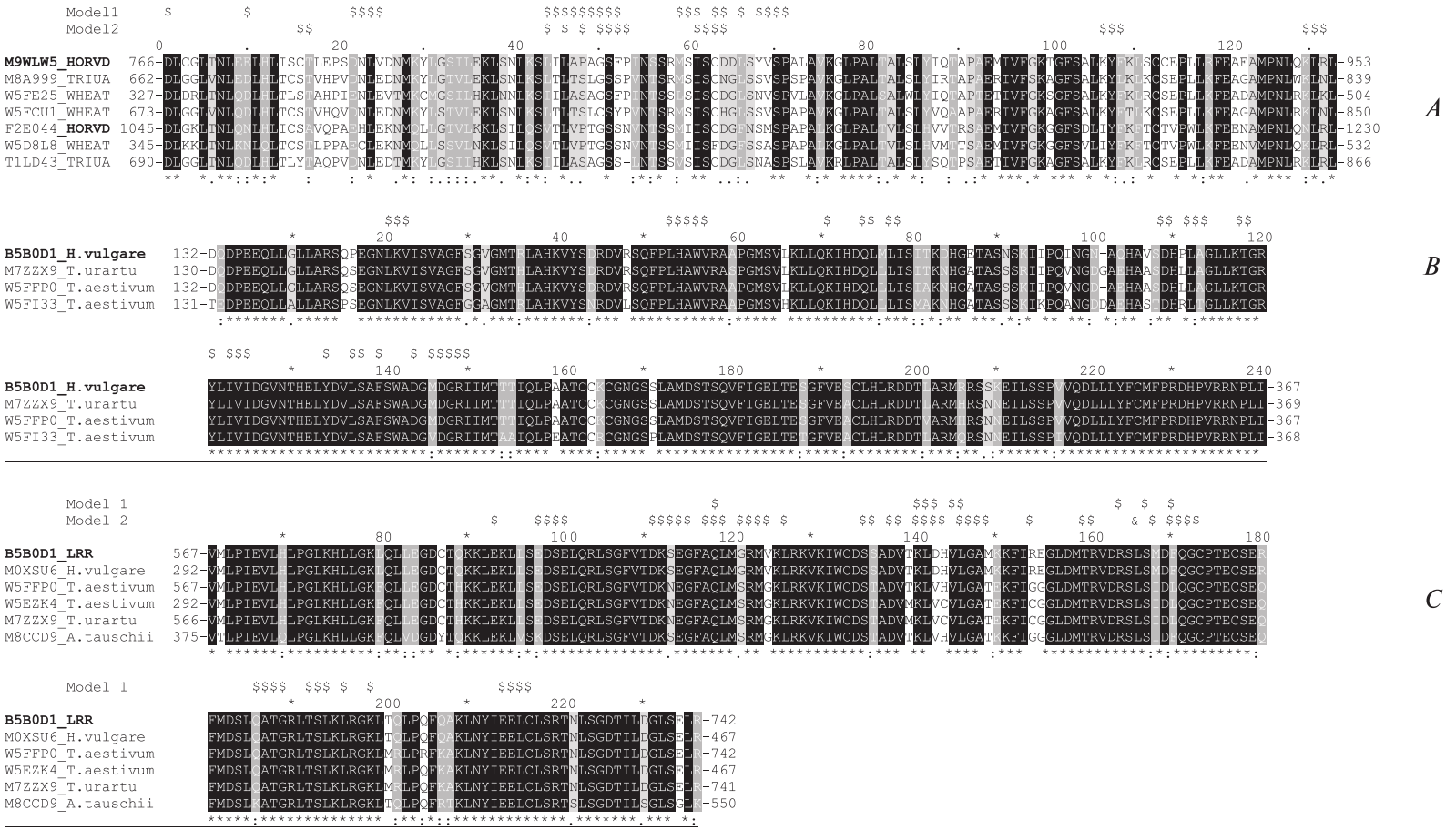

Fig. 3. The fragments of the multiple alignments of Rpg5, RGA1 and their homologues: *A* – a multiple alignment of Rpg5LRR domain (SMART: PF12799) and those of his homologues; *B* – a multiple alignment of RGA1NB domain (SMART: PF00931) and those of his homologues; *C* – a multiple alignment of RGA1 LRR domain (SMART: PF12799) and those of his homologues. The notes: the identical residues is marked with black color, the similar residues – dark gray and gray color, the variable residues - white; */20 – the symbols that mark the ordinal numbers of the residues in the multiple alignments; $ – a symbol that marks the predicted residues which could take part in the intermolecular interactions; the numbers near the N- and C-ends of the sequences show their location in the full length amino acid sequence that deposited in the UniProt

If we take into account this fact, then we can say that RGA1 has a much more conserved sequence of LRR domain than that of RPG5. We suggest that *RGA1* could be a trap for the pathogen effector molecules, and should be sufficiently conservative to fulfill this role.

The Q9LFP7_Y5158 sequence from *A. thaliana* is the only homologue among all which were found for Rpg5, RGA and Adf3, that has the clear data about its intermolecular interactions. It is known that this protein interacts with EXO70E2 (exocyst subunit exo70, family protein e2) and AT4G24690 (ubiquitin-associated (UBA) zinc-finger and PB1 domain-containing protein) [25]. EXO70E2 is required for the formation of exocyst-positive organelle (EXPO) [26], and AT4G24690, due to PB1 domain, could participate in many processes like activation of NADPH oxidase, which is very important for the immune defense of the mammals. And since this domain is conserved in plants, animals, fungi and even amoebas, it could be assumed that AT4G24690 homologues take part in the similar processes in plants [27].

Based on existing knowledge, it is possible to divide the different resistance genes into two groups. In the first class R genes are, most of which belong to the NB-LRR class. The second is the adult plant resistance genes (APR) which function only in adult plants. There are also no guarantees that some APR genes do not indeed include race specific, weak R genes which may be of the NB-LRR class [28]. It is not clear what position the genes like *Adf3* occupy in this classification. But it was shown previously that ADF4 from *A. thaliana* during plant innate immune signaling regulates actin dynamics in order to execute key events associated with PTI (pattern triggered immunity), such as cell wall fortification and transcriptional activation of defense gene markers [29]. In addition to this, recent research connects a basis of *Sr2* APR resistance with germine-like proteins, which in turn are linked with ROS [30]. As we mentioned above, TaADF7 also is tightly connected with reactive oxygen species that makes possible a participation of ADF in the different mechanisms of non-host resistance.

So, it is very important to determine the link between Adf3 protein function and the signaling that accompanies its activation and how this protein works under a pathogen attack together with RGA1 and Rpg5.

Conclusion. The results of this study confirm that barley close relatives such as *A. tauschii*, *B. distachyon*, *T. urartu*, *T. aestivum*, *O. sativa*, *S. italica*, *S. bicolor* have the homologues of *Adf3, Rpg5* and *RGA1*. We found the highly identical *Adf3* homologues in *T. elongatum* and *T. aestivum* genomes, but we could not find them in the genomes of *T. urartu* and *A. tauschii*. We found only two sequences with the same length as Rpg5 in *T. urartu* and *T. aestivum* genomes, but it was discovered that their kinase domains differed significantly from that of the Rpg5 protein. The search for every Rpg5 domain separately revealed the additional homologues in the genomes of some relatives of barley, which were available in the databases. We identified highly identical sequences for the full length protein product of RGA1 and for its LRR domain too, but we could not detect the sequences that were similar to NB part of RGA1 in the genomes of the other plants. Our results suggest that *T. aestivum* has similar sequences to *Adf3, Rpg5* and *RGA1* complex of genes and the further investigation should elucidate the function of this unknown defense mechanism.

The authors declare that there is no conflict of interests regarding the publication of this paper.

ИДЕНТИФИКАЦИЯ ГОМОЛОГОВ ГЕНОВ *Rpg5*, *Adf3* И *RGA1* УСТОЙЧИВОСТИ К БУРОЙ РЖАВЧИНЕ СРЕДИ РОДСТВЕННИКОВ ЯЧМЕНЯ

Б.В. Иващук, Д.А. Самофалова, Я.В. Пирко, Дж. Федак, Я.Б. Блюм

Гены ячменя Rpg5, RGA1 и Adf3, обеспечивающие устойчивость ко многим патотипам стеблевой ржавчины, клонированы несколько лет назад, но неизвестно имеются ли их гомологи в пшенице и других родственных видах. В настоящей работе представлены результаты биоинформационного поиска по определению гомологов Rpg5, RGA1 и Adf3 в геноме *Triticum aestivum*, а также геномах нескольких диких злаков, которые селекционеры обычно используют как источники устойчивости к бурой ржавчине и которые сейчас доступны в базах данных. Определено, что последовательность Q9FEC6 с *Th. elongatum* и последовательность Q43655 с *T. aestivum* являются высокоидентичными гомологами белковой последовательности Adf3. Последовательность M8A999 с *T. urartu* и последовательность W5FCU1 с *T. aestivum* оказались гомологами полной белковой последовательности Rpg5, но идентичность киназных доменов этих последо-

вательностей ниже, чем для других доменов. Отдельный анализ киназной части Rpg5 не предоставил четких доказательств наличия ортологов Rpg5 среди других видов злаков. Последовательность M7ZZX9 с *T. urartu*, а также последовательности W5FFP0 и W5FI33 с *T. aestivum* оказались гомологами RGA1. Анализ активных сайтов позволил определить разницу между белковыми последовательностями Rpg5, RGA1, Adf3 и последовательностями их гомологов.

ІДЕНТИФІКАЦІЯ ГОМОЛОГІВ ГЕНІВ *Rpg5*, *Adf3* І *RGA1* СТІЙКОСТІ ДО СТЕБЛОВОЇ ІРЖІ СЕРЕД РОДИЧІВ ЯЧМЕНЮ

Б.В. Іващук, Д.О. Самофалова, Я.В. Пірко, Дж. Федак, Я.Б. Блюм

Гени ячменю *Rpg5*, *RGA1* та *Adf3*, що забезпечують стійкість до багатьох патотипів стеблової іржі, були клоновані кілька років тому, але було невідомо чи наявні їхні гомологи в пшениці та інших споріднених видах. Дана робота представляє результати біоінформаційного пошуку по визначенню гомологів *Rpg5*, *RGA1* та *Adf3* в геномах *Triticum aestivum* та кількох диких злаків, які селекціонери зазвичай використовують як джерело стійкості до стеблової іржі і які наразі є доступними в базах даних. Визначено, що послідовність Q9FEC6 з *Th. elongatum* і послідовність Q43655 з *T. aestivum* є високоідентичними гомологами білкової послідовності Adf3. Послідовність M8A999 з *T. urartu* і послідовність W5FCU1 з *T. aestivum* виявились гомологами повної білкової послідовності Rpg5, але ідентичність кіназних доменів цих послідовностей є нижчою, ніж для інших доменів. Окремий аналіз кіназної частини Rpg5 не надав чітких доказів наявності ортологів Rpg5 серед інших видів злаків. Послідовність M7ZZX9 з *T. urartu* і послідовності W5FFP0 та W5FI33 з *T. aestivum* виявились гомологами RGA1. Аналіз активних сайтів дозволив визначити різницю між білковими послідовностями Rpg5, RGA1, Adf3 та послідовностями їх гомологів.

REFERENCES

1. Schumann, G.L., and Leonard, K.J., Stem rust of wheat (black rust), *Plant Health Instr.*, DOI: 10.1094/PHI-I-2000-0721-01, Updated 2011.
2. Arora, D., Gross, T., and Brueggeman, R., Allele characterization of genes required for *rpg4*-mediated wheat stem rust resistance identifies *Rpg5* as the *R* gene, *Phytopathology*, 2013, vol. 103, no. 11, pp. 1153–1161.
3. Brueggeman, R., Druka, A., Nirmala, J., Cavileer, T., Drader, T., Rostoks, N., Mirlohi, A., Bennypaul, H., Gill, U., Kudrna, D., Whitelaw, C., Kilian, A., Han, F., Sun, Y., Gill, K., Steffenson, B., and Kleinhofs, A., The stem rust resistance gene Rpg5 encodes a protein with nucleotide-binding-site, leucine-rich, and protein kinase domains, *Proc. Natl Acad. Sci. USA*, 2008, vol. 105, no. 39, pp. 14970–14975.
4. Rudrabhatla, P., Reddy M., and Rajasekharan, R. Genome-wide analysis and experimentation of plant serine/threonine/tyrosine-specific protein kinases, *Plant Mol. Biol.*, 2006, vol. 60, no. 2, pp. 293–319.
5. van der Linden, C.G., Wouters, D.C., Mihalka, V., Kochieva, E.Z., Smulders, M.J., and Vosman, B., Efficient targeting of plant disease resistance loci using NBS profiling, *Theor. Appl. Genet.*, 2004, vol. 109, no. 2, pp. 384–393.
6. McHale, L., Tan, X., Koehl, P., and Michelmore, R.W., Plant NBS-LRR proteins: adaptable guards, *Genome Biol.*, 2006, vol. 7, no. 4, pp. 212.
7. Lukasik, E., and Takken, F.L., STANDing strong, resistance proteins instigators of plant defense, *Curr. Opin. Plant Biol.*, 2009, vol. 12, no. 4, pp. 427–436.
8. Kobe B., and Kajava A.V., The leucine-rich repeat as a protein recognition motif, *Curr. Opin. Struct. Biol.*, 2001, vol. 11, no. 6, pp. 725–732.
9. Helft, L., Reddy, V., Chen, X., Koller, T., Federici, L., Fernбndez-Recio, J., Gupta, R., and Bent, A., LRR conservation mapping to predict functional sites within protein leucine-rich repeat domains, *PLoS One*, 2011, vol. 6, no. 7, e21614.
10. Wang, X., Richards, J., Gross, T., Druka, A., Kleinhofs, A., Steffenson, B., Acevedo, M., and Brueggeman, R., The *rpg4*-mediated resistance to wheat stem rust (*Puccinia graminis*) in barley (*Hordeum vulgare*) requires *Rpg5*, a second NBS-LRR gene, and an actin depolymerization factor, *Mol. Plant Microbe Interact.*, 2013, vol. 26, no. 4, pp. 407–418.
11. Porter, K., and Day, B., Actin branches out to link pathogen perception and host gene regulation, *Plant Signal. Behav.*, 2013, vol. 8, no. 3, e23468.1.
12. Fu, Y., Duan, X., Tang, C., Li, X., Voegele, R.T., Wang, X., Wei, G., Kang, Z., TaADF7, an actin-depolymerizing factor, contributes to wheat resistance against *Puccinia striiformis* f. sp. *tritici*, *Plant J.*, 2014, vol. 78, no. 1, pp. 16–30.
13. The Universal Protein Resource (UniProt), The UniProt Consortium, *Nucl. Acids Res.*, 2008, vol. 36, iss. suppl. 1, pp. D190–D195.
14. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L., BLAST+: architecture and applications, *BMC Bioinformatic*, 2009, vol. 10, pp. 421.
15. Letunic, I., Doerks, T., and Bork, P., Smart 7: recent updates to the protein domain annotation resource, *Nucl. Acids Res.*, 2012, vol. 40, iss. D1, pp. D302–D305.
16. Finn, R.D., Clements, J., and Eddy, S.R., HMMER web server: interactive sequence similarity searching, *Nucl. Acids Res.*, 2011, vol. 139, iss. suppl. 2, pp. W29–W37.

17. Notredame, C., Higgins, D. G., and Heringa, J., T-Coffee: A novel method for fast and accurate multiple sequence alignment, *J. Mol. Biol.*, 2000, vol. 302, no. 1, pp. 205–217.
18. Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., and Barton, G.J., Jalview Version 2 – a multiple sequence alignment editor and analysis workbench, *Bioinformatics*, 2009, vol. 25, no. 9, pp. 1189–1191.
19. Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S., MEGA6: molecular evolutionary genetics analysis version 6.0, *Mol. Biol. Evol.*, 2013, vol. 30, no. 12, pp. 2725–2729.
20. Felsenstein, J., Confidence limits on phylogenies: An approach using the bootstrap, *Evolution*, 1985, vol. 39, no. 4, pp. 783–791.
21. Jones, D.T., Taylor, W.R., and Thornton, J.M., The rapid generation of mutation data matrices from protein sequences, *Comput. Appl. Biosci.*, 1992, vol. 8, no. 3, pp. 275–282.
22. Saitou, N., and Nei, M., The neighbor-joining method: A new method for reconstructing phylogenetic trees, *Mol. Biol. Evol.*, 1987, vol. 4, no. 4, pp. 406–425.
23. Kelley, L.A., and Stenberg, M.J., Protein structure preduction on the Web: a case study using the Phyre server, *Nat. Protoc.*, 2009, vol. 4, no. 3, pp. 363–371.
24. Brueggeman, R., Steffenson, B. J., and Kleinhofs, A., The *rpg4/Rpg5* stem rust resistance locus in barley: resistance genes and cytoskeleton dynamics, *Cell Cycle*, 2009, vol. 8, no. 7, pp. 977–981.
25. Braun, P., Carvunis, A.R., Charloteau, B., Dreze, M., Ecker, J.R., Hill, D.E., et al., Evidence for network evolution in an Arabidopsis interactome map, *Science*, 2011, vol. 333, no. 6042, pp. 601–607.
26. Ding, Y., Wang, J., Chun, Lai, J.H., Ling Chan, V.H., Wang, X., Cai, Y., Tan, X., Bao, Y., Xia, J., Robinson, D.G., and Jiang, L., Exo70E2 is essential for exocyst subunit recruitment and EXPO formation in both plants and animals, *Mol. Biol. Cell*, 2014, vol. 25, no. 3, pp. 412–426.
27. Sumimoto, H., Kamakura, S., and Ito, T., Structure and function of the PB1 domain, a protein interaction module conserved in animals, fungi, amoebas, and plants, *Sci. STKE*, 2007, vol. 2007, no. 401, re6.
28. Ellis, J.G., Lagudah, E.S., Spielmeyer, W., and Dodds, P.N., The past, present and future of breeding rust resistant wheat, *Front. Plant Sci.*, 2014, vol. 5, pp. 641.
29. Henty-Ridilla, J.L., Li, J., Day, B., and Staigera, C.J., Actin depolymerizing factor4 regulates actin dynamics during innate immune signaling in Arabidopsis, *Plant Cell*, 2014, vol. 26, no. 1, pp. 340–352.
30. Mago, R., Tabe, L., Vautrin, S., Simkova, H., Kubalakova, M., Upadhyaya, N., Berges, H., Kong, X., Breen, J., Dolezel, J., Appels, R., Ellis, J.G., and Spielmeyer, W., Major haplotype divergence including multiple germin-like protein genes, at the wheat *Sr2* adult plant stem rust resistance locus, *BMC Plant Biol.*, 2014, vol. 14, pp. 379.

Received 18.08.15