# Highlights in BioScience

**Research Article**　　　　　　　　　　　　　　　　　　**Open Access**

# Identification of Genes for Wheat Fungal Resistance Using Bioinformatics Techniques

## Ahmed E. Nassar[1], Khaled H. Mousa[1], Ahmed A. Madbouly[2], Shafik D. Ibrahim[3] and Alsamman M. Alsamman[3]*

[1] Genetics Department, Faculty of Agriculture, Mansoura University.
[2] Plant Tissue Culture Department, Date Palm Research Institute, Agricultural Research Institute
[3] MGGM lab., Agricultural Genetic Engineering Research Institute, Agricultural Research Center, Egypt.

* To whom correspondence should be addressed: smahmoud@ageri.sci.eg

**Data Availability Statement:** All relevant data are within the paper and supplementary materials

**Competing interests:** The authors declare that they have no competing interests

## Abstract

For the majority of world populations, wheat (*Triticum aestivum* L.) would be the first essential and economic cereal grain crop. Pests and pathogens in both rich and developing countries are constantly threatening wheat production and sustainable development. Multiple gene pathways were recorded to share an association with fungal pathogens with wheat biological resistance. Our aim to use such tools in order to detect and classify fungal resistance genes in wheat through sequence alignment, protein domain identification and phylogenetic analysis. In addition the introduction for restriction fragment length polymorphism (RFLP) for such genes in the new primer database. Approximately 138 sequences of DNA were recovered from the wheat genome by aligning 3845 anti-fungal amino acids through tblastn tool. The NCBI blastn online tool used to detect sequences with functional genes, where 92 genes have been detected. The total number of nucleotides was 48385, where the smallest DNA sequence have 302 bp and the longest contains 977 bp with an average length of 525.9 bp per sequence. The wheat chromosomes 3D, and 4B have the highest number of sequences (9) followed by chromosomes 3B (7) and 3A(6), where wheat genomes A, B and D have 30, 35 and 27 genes, respectively. Five different amino acids motifs have been revealed among studied wheat amino acid sequences. The gene annotation tools used to infer studied amino acid gene annotation. Amino acid sequences belongs to lectin, kinase, tyrosine-protein kinase (STK), thaumatin, and cysteine-rich repeats representing 2, 9, 8, 19, 23 genes respectively, in addition to 31 hypothetical genes. The proteins chemical content have been assessed through 16 different amino acid chemical and physical characteristics.

Keywords: Wheat, Pathogens, Pathways, Fungal, Phylogenetic, RFLP.

## Introduction

For the majority of world populations, wheat (*Triticum aestivum* L.) would be the first essential and economic cereal grain crop. Its 29 percent share shows the world economy's significance of wheat production from world's 730 million hectares of crops cultivated lands. Depending on FAO figures for 2017, this percentage is comparable to 218 million hectares of wheat region (1). In most of the world, wheat is an key ingredient of human diet. Worldwide, 735 million tons of wheat were produced in 2015/2016, worth around US$ 145 billion (2). Wheat accounts for almost 55% of the carbohydrates and 20% of the world's food calorie intake (3).

In Egypt, the significance of wheat crops returns to its strategic importance in Egyptian dietary commodities as it offers more than one-third of Egyptian consumers ' daily calorie intake and 45% of their daily total protein intake (4). Wheat is Egypt's largest winter grain and is cultivated across the Nile delta and along the Nile Valley length. Wheat is usually planted after summer maize, cotton, or rice crops in early November and collected in late April or early May. Germination happens after planting at 19 to 22 weeks (5). While Egypt's wheat productivity has increased over the past few years, wheat production provides only 45% of its yearly domestic demand. Egypt is still one of the largest importing wheat countries. Wheat imports were approximately 9.8 million tons in 2011, costing approximately US$ 3.2 billion (6).

Pests and pathogens in both rich and developing countries are constantly threatening wheat production and sustainable development (7). A key component for trying to meet this challenge is proper management of fungal diseases that may be accountable for yield losses of 15–20 percent per year. The rusts, blotches and head blight/scab are among the major wheat diseases that make a contribution to these losses(2). The rusts, blotches and head blight/scab are among the major wheat diseases that currently contribute to these losses. There are three diseases of wheat rust, namely stem, stripe and leaf rust, all caused by members of the *Basidiomycete* family, genus *Puccinia*, named *P. graminis* f. SP. tritici, *P. striiformis* f. sp. tritici (PST) and *P. triticina* (Pt), respectively (2).

Multiple gene pathways were recorded to share an association with fungal pathogens with wheat biological resistance. The resistance is attributable to the additive impacts of multiple resistance metabolites and proteins generated through a network of multiple plant R gene structures. Plants recognize pathogenic elicitors or receptors and then stimulate downstream genes to ultimately produce resistance metabolites and proteins that suppress plant pathogen progression (8). Following the view of pathogens and microbes, reactive oxygen species (ROS) were consistently identified to accumulate in the plant and over the years, ROS was postulated as an essential part of the plant's defense response (9).

The connection between genetic variability identified using distinct molecular marker assays and genes responsible for morphological and physiological characteristics could be identified through sophisticated bioinformatics analysis tools (10–12). These tools could be used to identify and reveal new genes in wheat related to resistance to fungal diseases. Our aim to use such tools in order to detect and classify fungal resistance genes in wheat through sequence alignment, protein domain identification and phylogenetic analysis. In addition the introduction for restriction fragment length polymorphism (RFLP) for such genes in the new primer database.

## Materials and Methods

Anti fungal gene sequences have been downloaded from NCBI database (13). We have downloaded 3845 antifungal protein sequences from NCBI. The draft genome sequence of wheat has been downloaded from Ensemble database (14). The local NCBI blast package (15) was used to build sequence databases using the wheat draft genome and blast all anti-fungal genes with TBLASTN against wheat database. The online NCBI blast tblatn was used to annotate sequences recovered from previous step. MEME suite was used to discover amino acid motifs in the sequence (16). MegaX program was used to construct phylogenetic analysis through maximum likelihood algorithm (17). clustalo (18) tool was used to construct sequence similarity matrix through multiple sequence alignment. Clust-Vis was applied to create Principal Component Analysis (PCA) plots and heatmaps depicting genes similarities. Pepstat (19) was used through in-home per scrips to assess proteins chemical and physical characteristics. The perl script RestrictionDigest was used to detect restriction enzymes recognition sites in DNA sequences (20). BatchPrimer3 online tool was used for designing PCR primers could be used to target studied wheat genes (21).

## Result and Discussion

### Chromosomal distribution of wheat anti-fungal genes

Approximately 138 sequences of DNA were recovered from the wheat genome by aligning 3845 anti-fungal amino acids through tblastn tool. The NCBI blastn

online tool used to detect sequences with functional genes, where 92 genes have been detected (Supplementary file 1). The total number of nucleotides was 48385, where the smallest DNA sequence have 302 bp and the longest contains 977 bp with an average length of 525.9 bp per sequence. The wheat chromosomes 3D, and 4B have the highest number of sequences (9) followed by chromosomes 3B (7) and 3A(6), where wheat genomes A, B and D have 30 , 35 and 27 genes, respectively (Figure 1).
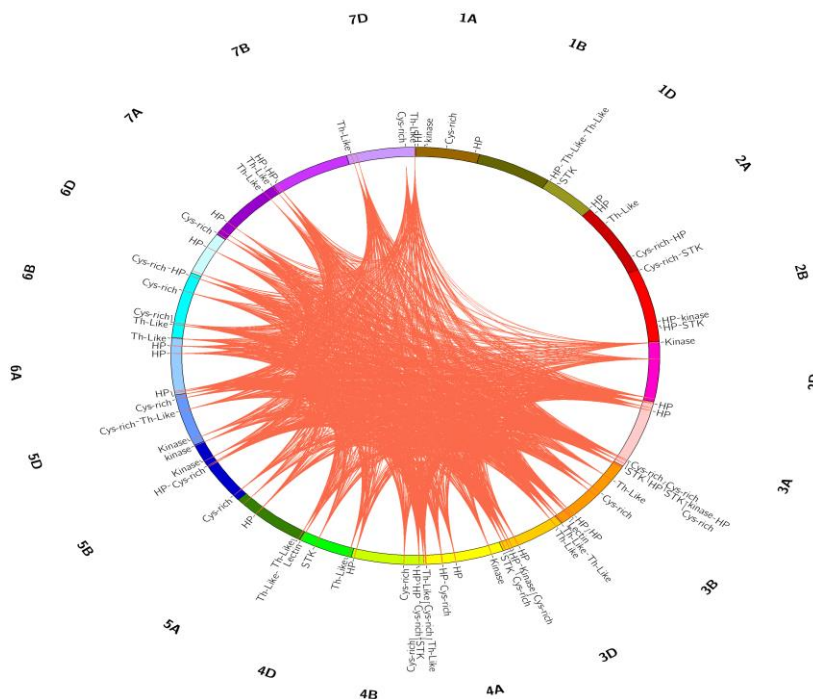


**Figure 1 :** The Circos diagram for the position of predicated anti-fungal genes and their proposed gene description, where the inner connections represent similar genes on separate chromosomes with sequence similarity over 75%.

### Wheat anti-fungal proteins chemical and physical properties

The proteins chemical content have been assessed through 16 different amino acid chemical and physical characteristics (Figure 2 and Table 1). The total molecular weight (MW) of studied fungal amino acid sequences was 1913 KDa with an average of 20 KDa, where 1A:551827789-551828139 has the minimum MW (10.6) and 5D:59043196-59044206 generated the maximum MW (71.911.68) Extinction coefficients are a measure of how much light the protein can be measured at a certain wavelength from the extinction factor. Approximation of this component is needed to be able to track the protein within the spectrophotometer. It is necessary to know the amino acid content to assess the molar extinction coefficient of the protein (22).

Table 1: The highest and lowest proteins characteristics and their corresponding proteins.

| Protein Characteristic | Total | Mean | Min. | Min. protein | Max. | Max. protein |
|---|---|---|---|---|---|---|
| **A280 Molar Extinction Coefficients cystine bridges** | 2261000 | 24576.09 | 1740 | 3B 716295639-716296412 | 104570 | 5D 59043196-59044206 |
| **A280 Molar Extinction Coefficients reduced** | 2214500 | 24070.65 | 1490 | 3B 716295639-716296412 | 103820 | 5D 59043196-59044206 |
| **Improbability of expression in inclusion bodies** | 65.124 | 0.794 | 0.504 | 7D 45432639-45432947 | 0.977 | 3D 462838207-462838962 |
| **Average Residue Weight** | 9829.348 | 106.841 | 98.498 | 7A 164401571-164402272 | 116.631 | 3D 550382548-550383990 |
| **Isoelectric Point** | 588.9429 | 6.402 | 4.0084 | 1A 20888904-20889590 | 10.3971 | 7B 40654848-40655408 |
| **Molecular weight** | 1913642.24 | 20800.46 | 10658.89 | 1A 551827789-551828139 | 71911.68 | 5D 59043196-59044206 |
| **Non-polar** | 5471.072 | 59.468 | 48.81 | 3A 719941535-719942353 | 69.919 | 5B 461292418-461292831 |
| **Polar** | 3728.357 | 40.526 | 30.081 | 5B 461292418-461292831 | 51.19 | 3A 719941535-719942353 |

| Protein Characteristic | Total | Mean | Min. | Min. protein | Max. | Max. protein |
|---|---|---|---|---|---|---|
| Residues | 17857 | 194.098 | 101 | 7D 45432639-45432947 | 644 | 5D 59043196-59044206 |
| Small | 5629.919 | 61.195 | 40.488 | 6A 610941935-610943086 | 77.027 | 5B 518810105-518810797 |
| Tiny | 3673.776 | 39.932 | 20.27 | 2B 634173118-634173630 | 54.054 | 5B 518810105-518810797 |
| Acidic | 858.259 | 9.329 | 5.172 | 3A 20936815-20937576 | 17.593 | 4B 41828978-41829304 |
| Aliphatic | 2702.584 | 29.376 | 19.048 | 5A 14266438-14266881 | 41.27 | 3D 8508332-8509102 |
| Aromatic | 952.739 | 10.356 | 3.604 | 3B 716295639-716296412 | 16.292 | 4D 29673605-29674213 |
| Basic | 932.965 | 10.141 | 4.082 | 7D 634805112-634805630 | 19.672 | 3B 176321011-176321379 |
| Charge | 24.5 | 0.266 | -9.5 | 3B 690358283-690358846 | 22 | 7B 40654848-40655408 |
| Charged | 1791.221 | 19.47 | 10.204 | 7D 634805112-634805630 | 32.258 | 3D 550382548-550383990 |

The A280 molar extinction coefficients cystine bridges (A280-MECc ) and reduced (A280-MECr) are two different measurements of extinction coefficient, where salt bridges are important motifs of the tertiary protein structure and are mostly correlated with the structural influence force which maintains the stability of the protein. Commonly found on the solvent system and particularly vulnerable to solvent-solute interactions typically with water as well as other cosolvents (23). The minimum scores of A280-MECc and A280-MECr were 1740 and 1490, generated by 3B:716295639-

716296412, while the highest scores were 104570 and 103820, revealed by 5D:59043196-59044206 with a mean of 24576 and 24070, respectively. Expression improbability in bodies of inclusion (IEIB) is a form of measurement of solubility. For instance, recombinant protein in Escherichia coli could be expressed either as soluble throughout the cytosol or as insoluble in the bodies of inclusion (24). The IEIB of wheat amino acids showed a mean of 0.794, where it ranges from 0.504 (7D:45432639-45432947) to 0.977 (3D: 462838207-462838962).
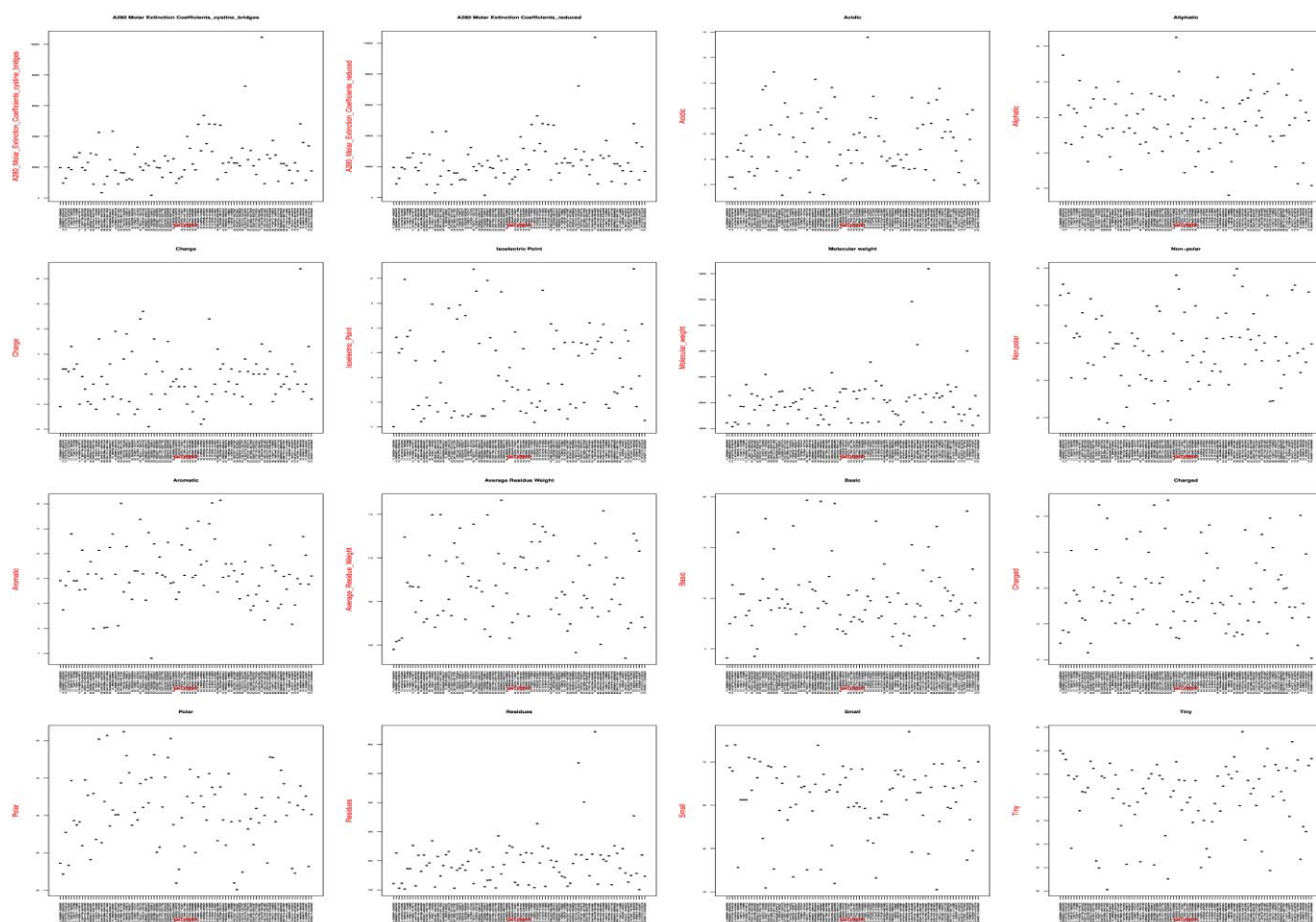


**Figure 2:** The protein chemical composition and physical properties for the predicated anti-fungal genes.

The average residue weight (ARW) stands for the average collective weight for all amino acid sequences according to its length. The total ARW for all amino acid sequences was 9829.3 Da, where 7A:164401571-164402272 revealed the minimum ARW (98.498 Da). 3D:550382548-550383990 generated the maximum (116.631 Da) and the mean was 106.841 Da. The isoelectric point (IP) is the pH level at which protein's net charge is positive and is associated with its production of amino acids and protein conformation (25). The collective wheat fungal genes amino acids was 588.9 with a mean of 6.402, where it ranges from 4 (1A:20888904-20889590) to 10.4 (7B:40654848-40655408).

In addition, a protein's folded structure becomes thermodynamically less desirable because it decreases the protein's disorder or entropy. Nonpolar (water hating) side chains tend to push themselves inside a protein whereas side chains of polar (water loving) prefer to put themselves outside of the molecule (26). The non-polar and polar amino acid scores were ranged from 48.81 and 30.081 (3A:719941535-719942353 and 5B:461292418-461292831) to 69.919 and 51.19 (5B:461292418-461292831 and 3A:719941535-719942353), respectively.

Acidic amino acids have an acidic group in side chain while basic amino acids have some basic group in side chain. Acidic amino acids have low pKa while basic amino acids have high pKa. The score of acidic and basic amino acids range from 5.172 and 4.082 (3A:20936815-20937576 and 7D:634805112-634805630 ) to 17.593 and 19.672 (4B:41828978-41829304 and 3B:176321011-176321379 ). The studied fungal amino acids charges ranged from -9.5 (3B:690358283-690358846) to 22 (7B:40654848-40655408) (Figure 2 and Table 1).

**Detection of protein domain and phylogenetic analysis**

A motif for the protein sequence is a brief pattern that nature retains. For proteins, a motif can refer to an enzyme's active site or a structural unit required for proper protein folding. Therefore, sequence motifs are among molecular evolution's basic functional components (16). Five different amino acids motifs have been revealed among studied wheat amino acid sequences (Motif-1: YGLAQCTPDL, Motif-2: CYLRYSNKNF, Motif-3: PNRVWGITWC, Motif-4: YDVS VVDGFN, and Motif-5: NPGDCGQCLQ) (Figure 3 and 4).
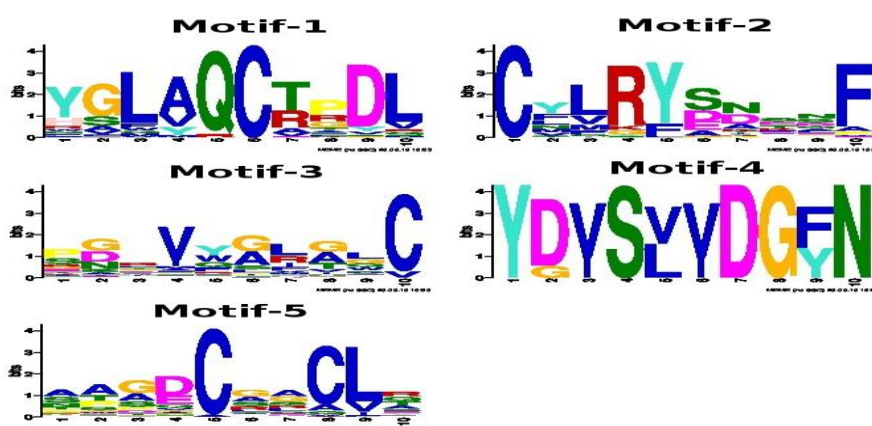


**Figure 3:** The identified protein domains / motifs in predicated wheat genes.

The Motif-1 and Motif-3 has a match to ELME000249 in motif database with p-value of 9.64e-04. This motif is a TRFH domain docking motifs wich are coordinating the telomeres with other proteins. Not only do they form homodimers with their TRFH domains, they also provide specific protein-binding surfaces for interaction (27). TRFHs and other proteins have been proposed to have important implications for the biology and evolution of telomeres. In general, this offers a basis for recognizing and controlling the hierarchical assembly and stoichiometry of telomere subunits during the cell cycle, cell division and senescence (28).

Motif-2 has an ELME000084 match which is ligands of phosphotyrosine bound by domains SH2. Src Homology 2 (SH2) domains identify tiny patterns comprising a residue of

phosphorylated tyrosine. Up to four positions after the pTyr was found mainly to assess additional specificity. In plants SH2 dominates new signaling scenarios and participates in metazoan signal transduction, serving as key mediators of controlled protein-protein interactions with tyrosine-phosphorylated substrates (29, 30).

Moreover, the database motif of ELME000377 was highly similar to Motif-4, this motif codes for Pex14 ligand motif, which belongs to peroxisomes. Peroxisomes are subcellular organelles present in eukaryotes that are single-membrane spherical. Peroxisomes relate together with glyoxysomes found in plants and glycosomes found in trypanosomes to microbody group of organelles.
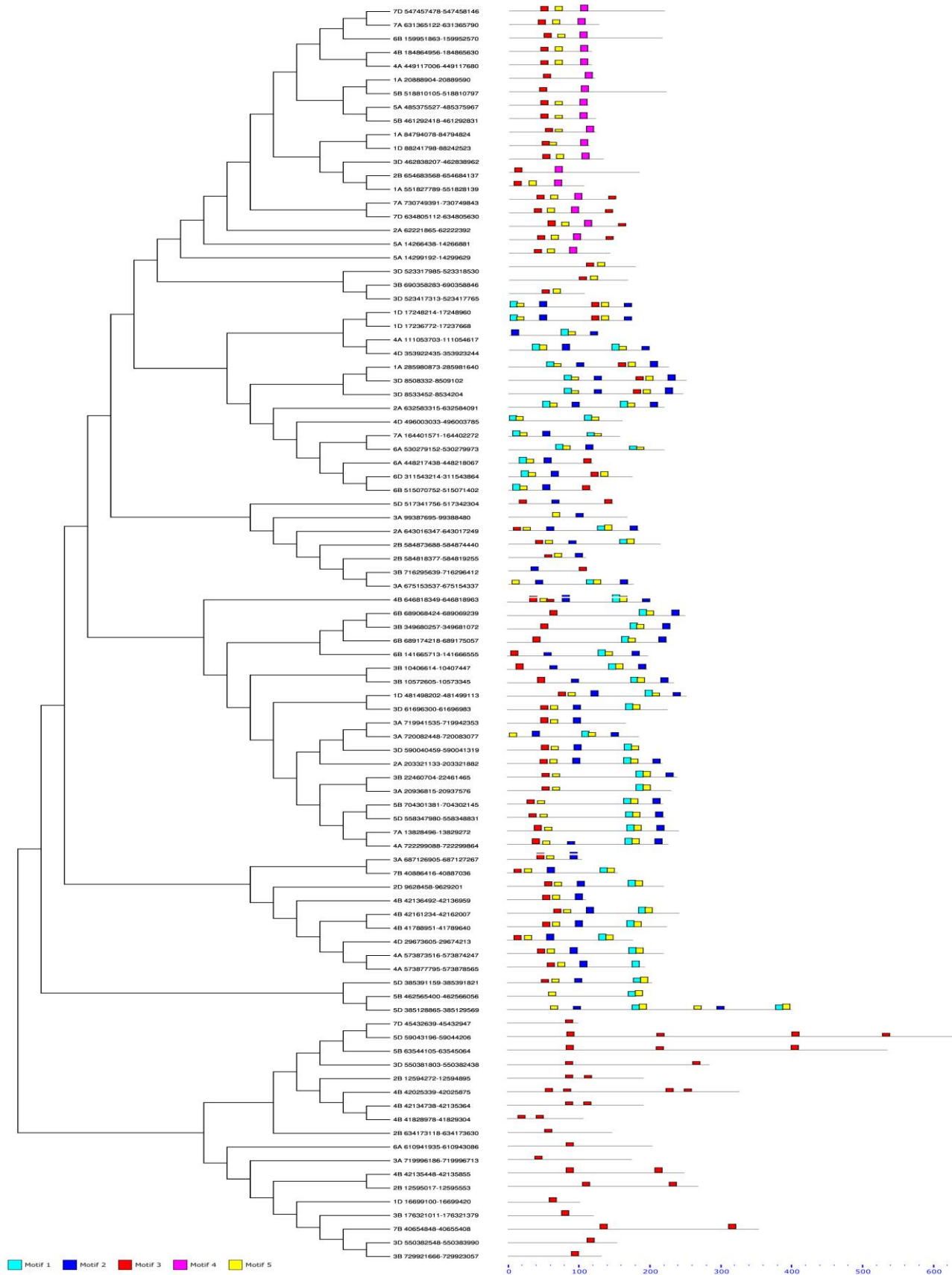
**Figure 4:** The phylogenetic tree developed using multiple sequence alignment of predicted anti-fungal genes as compared to the expected positions of protein domains/motifs.

Import into the peroxisome of peroxisomal matrix enzyme proteins (PTS1 cargo) includes the identification of the PTS1 cargo by the cytosol Pex5 receptor, the docking of the PTS1-Pex5 complex at the peroxisomal membrane, and the translocation of the PTS1 cargo through the peroxisomal membrane into the matrix. This is preceded by recycling the Pex5 receptor back into the cytosol for another round of export of PTS1 stock. Imported into peroxisome are unfolded, folded oligomeric or cofactor-bound proteins (31).

Additionally, Motif-5 was with high similarity with ELME000159, which is MAPK Phosphorylation Site. Cascades of mitogen-activated protein kinase (MAPK) are highly conserved downstream signaling modules of receptors/sensors that transform extracellular stimuli in eukaryotes into intracellular responses. Plant MAPK cascades play crucial functions against pathogen invasion in signaling plant defense (32). The NCBI-blastx tool used to infer studied amino acid gene annotation. Amino acid sequences belongs to lectin, kinase, tyrosine-protein kinase (STK), thaumatin, and cysteine-rich repeats representing respectively, 2, 9, 8, 19, 23 genes, in addition to 31 hypothetical genes (Figure 1).

Lectins are non-immune proteins which attach carbohydrates directly and reversibly. The lectin's biochemical functions are really diverse. A common theme arising from the identified functions of many plant and animal lectins is their involvement in communicating with other organisms whether symbiosis or defense, as effectors or regulators (33). New types of nucleocytoplasmic plant lectins have been described and defined over the past decade, especially lectins expressed within the nucleus and the cytoplasm of plant cells, much as part of a particular plant response when exposed to different stressors or shifting environmental conditions (34). Lectins include proteins containing at least one non-catalytic domain which helps them to selectively identify and reversibly attach to different glycans that are either freely available or are member of glycoproteins or glycolipids.

Plants release an overwhelming number of highly complex lectins with various molecular structures and attaching specificities to endogenous (plant) glycans as well as exogenous (non-plant) glycans (35). The role of plant lectins in plant defense againt different pathogens including fungi have been reported in previous researches (36,37). There are several receptor-like kinases among the plant proteins proposed to engage in immunity pathways (38). For example, multiple kinase proteins have been recorded for wheat resistance to fungal diseases, which provides temperature-dependent resistance to wheat stripe rust (39), and fungal resistance in Arabidopsis (40).

In this regards , various thaumatin-like proteins have been reported to generate tolerance to fungal pathogens Arabidopsis (41), potato (42) and grape (43). Moreover, cysteine-rich proteins have been informed to confer resistance for fungal diseases such as maize fungal pathogens (44) and legume modules (45-47).

Using maximum likelihood analysis phylogenetic analysis was effective in differentiating between anti-fungal wheat proteins according to their protein domains. Four classes grouped the phylogenetic tree. Group A comprises only Motif-3 domain protein fragments, group B has Motif-3, Motif-4 and Motif-5 domain proteins. In addition some group C contains genes with Motif-1, Motif-2 and Motif-5 domains, and eventually group D contains genes with all protein domains (Figure 4).

**PCR primers and restriction enzymes analysis**

About 85 PCR primers pairs have been designed to target most of the predicted anti-fungal genes (Supplementary file 2). The PCR predicated product size ranged from 300 bp to 708 pb, where the GC content ranges from 36.4 to 66.7% and the primers annealing temperature ranges from 58.05 oC to 62.55 oC. A total of 40 different restriction enzymes (RE) were used to identify potential locations for future RFLP tests for genetic polymorphism inside predicted anti-fungal wheat genes. Figure (5) demonstrates the site redundancy ratio for RE identification within genes. Figure (5) shows that some RE such as BanII, NlaIII and BaeGI have high probability to produce fragment length polymorphism isnide most of predicated anti-fungal genes.

## Conclusion

Predicting using available fungal resistance genes in the public database was very helpful and suggested that several wheat genes could be used to restrict genetic research of genes that hold the key to fungal resistance in wheat. Most of the identified protein domains clarify the genetic structure of anti-fungal genes and suggest a potential role for the MAPK gene family in such pathways. Providing restriction enzyme information and gene-specific PCR primers could be useful to wheat scientists and breeders, saving time and effort.

**Figure 5:** The ratio heat map for redundancy site identification of different restriction enzymes, where red and blue squares represent high and low rate of RE digestion sites, respectively.

## Supplementary Files

The Supplementary Material for this article can be found online at: http://bioscience.highlightsin.org/2018/20185.php

**Supplementary file 1**: The nucleotide sequences of identified fungal resistance genes in wheat genome.
**Supplementary file 2:** The information of PCR primers designed to target identified fungal resistance genes in wheat.

## Referances

1. FAOSTAT. 2017; Available from: http://www.fao.org

2. Figueroa M, Hammond-Kosack KE, Solomon PS. A review of wheat diseases—a field perspective. Mol Plant Pathol. 2018;19(6):1523–36.

3. Kumar P, Yadava RK, Gollen B, Kumar S, Verma RK, Yadav S. Nutritional contents and medicinal properties of wheat: a review. Life Sci Med Res. 2011;22:1–10.

4. Farg E, Arafat SM, El-Wahed MA, El-Gindy AM. Estimation of evapotranspiration ETc and crop coefficient Kc of wheat, in south Nile Delta of Egypt using integrated FAO-56 approach and remote sensing data. Egypt J Remote Sens Sp Sci. 2012;15(1):83–9.

5. McVey D V, Nazim M, Leonard KJ, Long DL. Patterns of virulence diversity in Puccinia triticina on wheat in Egypt and the United States in 1998-2000. Plant Dis. 2004;88(3):271–9.

6. Al-Naggar AMM, Sabry SRS, Atta MMM, El-Aleem OMA. Field Screening of Wheat (Triticum aestivum L.) Genotypes for Salinity Tolerance at Three Locations in Egypt. Ecol Res Intern(JAERI). 2015;4(3):88–104.

7. Skoracka A, Rector BG, Hein GL. The interface between wheat and the wheat curl mite, Aceria tosichella, the primary vector of globally important viral diseases. Front Plant Sci. 2018;9.

8. Kushalappa AC, Yogendra KN, Karre S. Plant innate immune response: qualitative and quantitative resistance. CRC Crit Rev Plant Sci. 2016;35(1):38–55.

9. Lehmann S, Serrano M, L'Haridon F, Tjamos SE, Metraux J-P. Reactive oxygen species and plant resistance to fungal pathogens. Phytochemistry. 2015;112:54–62.

10. Alsamman AM, Ibrahim SD, Hamwieh A. KASPspoon: an in vitro and in silico PCR analysis tool for high-throughput SNP genotyping. Bioinformatics. 2019;

11. Habib PT, Alsamman AM, Hamwieh A. BioAnalyzer: Bioinformatic Software of Routinely Used Tools for Analysis of Genomic Data. Biotechnology. 2019 ;10:33–41.

12. Awan Z. Plant Molecular Biology Databases. Highlights Biosci. 2019;1–7.

13. Maglott D, Ostell J, Pruitt KD, Tatusova T. Entrez Gene: gene-centered information at NCBI. Nucleic Acids Res. 2010;39(suppl_1):D52--D57.

14. Hubbard T, Barker D, Birney E, Cameron G, Chen Y, Clark L, et al. The Ensembl genome database project. Nucleic Acids Res. 2002;30(1):38–41.

15. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25(17):3389–402.

16. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37 :W202--W208.

17. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35(6):1547–9.

18. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22(22):4673–80.

19. Rice P, Longden I, Bleasby A. EMBOSS: the European molecular biology open software suite. Elsevier current trends; 2000.

20. Wang J, Li L, Qi H, Du X, Zhang G. RestrictionDigest: A powerful Perl module for simulating genomic restriction digests. Electron J Biotechnol. 2016; 19(3):36–42.

21. You FM, Huo N, Gu YQ, Luo M, Ma Y, Hane D, et al. BatchPrimer3: a high throughput web application for PCR and sequencing primer design. BMC Bioinformatics. 2008;9(1):253.

22. Gill SC, Von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem. 1989;182(2):319–26.

23. Pylaeva S, Brehm M, Sebastiani D. Salt Bridge in Aqueous Solution: Strong Structural Motifs but Weak Enthalpic Effect. Sci Rep. 2018;8.

24. Davis GD, Elisee C, Newham DM, Harrison RG. New fusion protein systems designed to give soluble expression in Escherichia coli. Biotechnol Bioeng. 1999;65(4):382–8.

25. Righetti PG. Determination of the isoelectric point of proteins by capillary isoelectric focusing. J Chromatogr A. 2004;1037(1–2):491–9.

26. Broome BM, Hecht MH. Nature disfavors sequences of alternating polar and non-polar amino acids: implications for amyloidogenesis. J Mol Biol. 2000;296(4):961–8.

27. Gouw M, Michael S, Sámano-Sánchez H, Kumar M, Zeke A, Lang B, et al. The eukaryotic linear motif resource--2018 update. Nucleic Acids Res. 2017;46(D1):D428--D434.

28. Giraud-Panis M-J, Ye J, Gilson E. TRFH domain: at the root of telomere protein evolution? Cell Res. 2018;28(1):7–8.

29. Liu BA, Shah E, Jablonowski K, Stergachis A, Engelmann B, Nash PD. The SH2 domain--containing

proteins in 21 species establish the provenance and scope of phosphotyrosine signaling in eukaryotes. Sci Signal. 2011;4(202):ra83--ra83.

30. Williams JG, Zvelebil M. SH2 domains in plants imply new signalling scenarios. Trends Plant Sci. 2004;9(4):161–3.

31. Pires JR, Hong X, Brockmann C, Volkmer-Engert R, Schneider-Mergener J, Oschkinat H, et al. The ScPex13p SH3 domain exposes two distinct binding sites for Pex5p and Pex14p. J Mol Biol. 2003;326(5):1427–35.

32. Meng X, Zhang S. MAPK cascades in plant disease resistance signaling. Annu Rev Phytopathol. 2013;51:245–66.

33. De Hoff PL, Brill LM, Hirsch AM. Plant lectins: the ties that bind in root symbiosis and plant defense. Mol Genet genomics. 2009;282(1):1–15.

34. Lannoo N, Van Damme EJM. Lectin domains at the frontiers of plant defense. Front Plant Sci. 2014;5:397.

35. Van Damme EJM, Lannoo N, Peumans WJ. Plant lectins. In: Advances in botanical research. Elsevier; 2008. p. 107–209.

36. Wang Y, Bouwmeester K. L-type lectin receptor kinases: New forces in plant immunity. PLoS Pathog. 2017;13(8):e1006433.

37. Wawra S, Fesel P, Widmer H, Timm M, Seibel J, Leson L, et al. The fungal-specific $\beta$-glucan-binding lectin FGB1 alters cell-wall composition and suppresses glucan-triggered immunity in plants. Nat Commun. 2016;7:13188.

38. Llorente F, Alonso-Blanco C, Sánchez-Rodriguez C, Jorda L, Molina A. ERECTA receptor-like kinase and heterotrimeric G protein from Arabidopsis are required for resistance to the necrotrophic fungus Plectos-phaerella cucumerina. Plant J. 2005;43(2):165–80.

39. Fu D, Uauy C, Distelfeld A, Blechl A, Epstein L, Chen X, et al. A kinase-START gene confers temperature-dependent resistance to wheat stripe rust. Science (80). 2009;323(5919):1357–60.

40. Wan J, Zhang X-C, Neece D, Ramonell KM, Clough S, Kim S, et al. A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in Arabidopsis. Plant Cell. 2008;20(2):471–81.

41. Misra RC, Kamthan M, Kumar S, Ghosh S, others. A thaumatin-like protein of Ocimum basilicum confers tolerance to fungal pathogen and abiotic stress in transgenic Arabidopsis. Sci Rep. 2016;6:25340.

42. Ali G, Hu X, Reddy ASN. Overexpression of the Arabidopsis thaumatin-like protein 1 in transgenic potato plants enhances resistance against early and late blights. bioRxiv. 2019;621649.

43. Marcato R, Sella L, Lucchetta M, Vincenzi S, Odorizzi S, Curioni A, et al. Necrotrophic fungal plant pathogens display different mechanisms to counteract grape chitinase and thaumatin-like protein. Physiol Mol Plant Pathol. 2017;99:7–15.

44. Patiño B, Vázquez C, Manning JM, Roncero MIG, Córdoba-Cañero D, Di Pietro A, et al. Characterization of a novel cysteine-rich antifungal protein from Fusarium graminearum with activity against maize fungal pathogens. Int J Food Microbiol. 2018;283:45-51.

45. Maróti G, Downie JA, Kondorosi É. Plant cysteine-rich peptides that inhibit pathogen growth and control rhizobial differentiation in legume nodules. Curr Opin Plant Biol. 2015;26:57–63.

46. Mokhtar MM, Adawy SS, El-Assal SE-DS, Hussein EHA. Genic and Intergenic SSR Database Generation, SNPs Determination and Pathway Annotations, in Date Palm (Phoenix dactylifera L.). PLoS One. Public Library of Science; 2016;11(7):e0159268.

47. Adawy SS, Mokhtar MM, Alsamman AM, Sakr MM. Development of annotated EST-SSR database in olive (Olea europaea). Int J Sci Res. 2015;4(9):1063–73.