Structural and Function Prediction of *Musa acuminata* subsp. *Malaccensis* Protein

Anum Munir[1*], Azhar Mehmood[1], Shumaila Azam[1,2]

[1]*Department of Bioinformatics*
Government Post Graduate College Mandian
Abbottabad 22010, Pakistan
Emails: anummunir786@yahoo.com, azharmehmood35@yahoo.com, shumailaazam@hotmail.com

[2]*Departments of Bioinformatics*
Muhammad Ali Jinnah University
Islamabad 44000, Pakistan

[*]*Corresponding author*

Received: September 03, 2015 **Accepted: March 02, 2016**

Published: March 31, 2016

Abstract: *Hypothetical proteins (HPs) are the proteins whose presence has been anticipated, yet in vivo function has not been built up. Illustrating the structural and functional privileged insights of these HPs might likewise prompt a superior comprehension of the protein-protein associations or networks in diverse types of life. Bananas (Musa acuminata spp.), including sweet and cooking types, are giant perennial monocotyledonous herbs of the order Zingiberales, a sister grouped to the all-around considered Poales, which incorporate oats. Bananas are crucial for nourishment security in numerous tropical and subtropical nations and the most prominent organic product in industrialized nations. In the present study, the hypothetical protein of M. acuminata (Banana) was chosen for analysis and modeling by distinctive bioinformatics apparatuses and databases. As indicated by primary and secondary structure analysis, XP_009393594.1 is a stable hydrophobic protein containing a noteworthy extent of α-helices; Homology modeling was done utilizing SWISS-MODEL server where the templates identity with XP_009393594.1 protein was less which demonstrated novelty of our protein. Ab initio strategy was conducted to produce its 3D structure. A few evaluations of quality assessment and validation parameters determined the generated protein model as stable with genuinely great quality. Functional analysis was completed by ProtFun 2.2, and KEGG (KAAS), recommended that the hypothetical protein is a transcription factor with cytoplasmic domain as zinc finger. The protein was observed to be vital for translation process, involved in metabolism, signaling and cellular processes, genetic information processing and Zinc ion binding. It is suggested that further test approval would help to anticipate the structures and functions of other uncharacterized proteins of different plants and living being.*

Keywords: *Musa acuminata, Homology modeling, Functional annotations, Ab initio.*

Introduction

Hypothetical proteins (HPs) are proteins whose presence has been anticipated yet *in vivo* function has not been built up [11, 25]. HPs generally cover around a large portion of the protein coding regions in many genomes. In spite of the fact that their functions have yet not been very much described, they may have their own significance to complete genomic and proteomic data [17, 23]. Legitimate structural and functional annotations of HPs of specific genome may prompt the locating of new structures and also new functions and help to present extra protein pathways and cascades, in this way finishing our sketchy information on the mosaic of proteins [17]. Illustrating the structural and functional privileged insights of these HPs might likewise prompt a superior comprehension of the protein-protein associations or

networks in diverse types of life, for example, plants, microorganisms, and so forth [14]. Besides, novel HPs might likewise serve as markers and pharmacological targets for drug design, revelation and screening [19, 21].

Bananas (*Musa Acuminata spp*.), including sweet and cooking types, are giant perennial monocotyledonous herbs of the group *Zingiberales*, a sister group to the very much examined Poales, which incorporate grains. Banana is a key for nourishment security in numerous tropical and subtropical nations and the most famous organic product in industrialized nations. Banana cultivars principally include *M. acuminata* (A genome) and *M. balbisiana* (B genome) and are some of the time diploid yet for the most part triploid. *M. acuminata* is diploid wild types of banana. Bananas are perpetual monocotyledonous herbs of the group Zingiberales to which ginger and cardamom additionally have a place. Bananas are critical for nourishment security in numerous tropical and subtropical nations and a vast scope of sorts can be discovered, including those utilized as vegetables and as organic products. While domestication included hybridization between various species and subspecies, a significant part of the present generation depends on clones got from one triploid genotype – Cavendish. *M. acuminata* has a haploid chromosome number of 11, containing 390 Mb with 36,000 protein encoding genes [9].

In recent years, various hypothetical proteins have been found in the genome of numerous life forms. In any case, because of a few restrictions, for example, the expense and time needed for exploratory methodologies, complete genome annotations have not accomplished yet. In addition, the extensive amount of theoretical proteins in a genome makes their study a troublesome task. Bioinformatics methodologies using distinctive algorithms and databases to estimate protein capacity would be a decent different option for laboratory research facility based techniques. As these algorithms and databases are in light of experimental results, they can be a compelling intends to perform functional and structural annotation of hypothetical proteins [25].

In the present study, the *M. acuminata* hypothetical protein XP_009393594.1, belongs to AAA group was chosen as the primary amino acid sequence of the protein is accessible however structural details are not accessible. The study meant to analyze the physiochemical and secondary structure components, to produce the first three-dimensional (3D) model through *Ab initio* technique, and finally to perform functional annotations. The result of this work will be useful for better comprehension of the mechanism of this protein and discovering other novel proteins and their functions by same method that we have accomplished for *M. acuminata* protein.

Materials and methods

Sequence retrieval

The amino acid sequence of the *M. acuminata* hypothetical protein XP_009393594.1 was retrieved from the Uniprot database (http://www.uniprot.org/).

Physiochemical analysis of the protein

Analysis of the physiochemical characteristics of the studied protein such as molecular weight, theoretical pI, amino acid composition, atomic composition, instability index, and grand average of hydropathicity (GRAVY) were performed using ProtParam tool (http://web.expasy.org/protparam/) [12].

Secondary structure analysis

The server SOPMA was utilized for secondary structure predictions (helix, sheets, and coils) of the hypothetical protein [13]. In addition to that, the PSIPRED [6] and PredictProtein [26] servers were also exploited to validate the results obtained from SOPMA.

Subcellular localization prediction

Subcellular localization of *M. acumianta* was predicted by PSORT [19]. Results were also cross-checked with subcellular localization predictions obtained from SOSUI server and Predict-Protein servers [30].

Homology modeling of the hypothetical protein

The conceivable 3D structure of the protein XP_009393594.1 was constructed through alignment mode in protein structure homology modeling server SWISS-MODEL [2, 20] using the full amino acid sequence of the protein in FASTA format.

Quality assessment of the 3D model and visualization

The initial structural model obtained, was checked for acknowledgement of errors in 3D structure [27] by ERRAT and Verify3D programs included in structural analysis and verification server SAVES (http://nihserver.mbi.ucla.edu/SAVES/) [5, 8]. The Ramachandran plots for the model were constructed using the RAMPAGE server [18], showing the percentage of protein residues in the favored, allowed and outlier regions. The visualization of generated model was performed by Discovery studio 4.1 [23].

Functional annotation of the protein

M. acuminata hypothetical protein XP_009393594.1 was analyzed for the function. Three different bioinformatics tools and databases including ProtFun 2.2 [33], ProFunc [10], and NCBI Conserved Domains Database (NCBI-CDD) [20] were utilized for this reason.

Moreover, KEGG automatic annotation server (KAAS) was used to analyze the involvement of *M. acuminata* hypothetical protein in the metabolic pathways [21].

Submission of the model in protein model database

The model generated for *M. acuminata* hypothetical protein XP_009393594.1 was successfully submitted in Protein Model Database (PMDB) (http://bioinformatics.cineca.it/PMDB/).

Results and discussion

Physiochemical characteristics of XP_009393594.1

The ExPASy's ProtParam server was utilized to examine the theoretical physiochemical characteristics of the amino acid sequence of hypothetical protein XP_009393594.1. The vast majority of the calculations in this server exhibit protein steadiness and stability, in light of the fact that the stability is identified with its appropriate function capacity [1]. The protein was predicted to be comprised 280 amino acids, with a molecular weight of 30302.3 Daltons and an isoelectric point (P^I) of 5.40 demonstrated a negatively charged protein. The instability index of the protein was computed to be 56.52, which demonstrated this protein as stable. The GRAVY index of -0.214 is demonstrative of a hydrophobicity and solubility of protein. The most plenteous amino acid residue was observed to be Serine (30), followed by Glycine (26) and the most minimal amino acid as Tyrosine (3). The sequence had 27 negatively charged residues (Aspartic acid + Glutamic acid) and 20 positively charged residues

(Arginine + Lysine). The molecular formula of the protein was found as $C_{1300}H_{2080}N_{378}O_{414}S_{21}$.

Subcellular localization of XP_009393594.1

Protein subcellular localization predictions include the computational expectation of where a protein lives in a cell. Predicting subcellular localization of unknown proteins can give information about their cellular functions. This information could be utilized in understanding disease mechanism and developing drugs [32]. The subcellular localization of the query protein was anticipated to be a cytoplasmic protein, analyzed by SOSUI and confirmed by PSORTb v3.2.0 and Predict Protein severs.

Secondary structure of XP_009393594.1

First the secondary structure of the protein was predicted by SOPMA server. The alpha helix was found to be the most predominant (30%), followed by random coil (56%) and extended strand (13%). Likewise, beta turn was found as 0.0%. Second, the similar results were obtained from Predict-Protein and PSIPRED servers. The delegate secondary structure of XP_009393594.1 obtained from the PSIPRED server is shown in Fig. 1.

Homology modeling of XP_009393594.1

We assume these uncharacterized proteins a limitless unexplored field with various opportunities, both as medicinal and industrial tools. *In silico* examination may help with deciding the biological functions of such un-characterized proteins. This can be encouraged by anticipating the 3D structure of the target protein. At the point when the tentatively experimented structure is inaccessible, similar or homology modeling can now and then give a helpful 3D model to the protein of interest that is identified with no less than one known protein structure. Homology modeling predicts the 3D structure of a given protein sequence construct principally with respect to its alignments to one or more proteins of known structure [24].

To perform the homology modeling, the query sequence was given as input in SWISS-MODEL server. The server consequently performed BLASTP search for each protein sequence to identify templates for homology modeling. The highest template identity was 39% which showed that XP_009393594.1 hypothetical protein is novel and no similar template structure is present in any databases. We predicted the 3D structure of XP_009393594.1 hypothetical protein by *Ab initio* method through Phyre 2 server and 3D Jigsaw server which gave 99.3% confidence in model. The 3D model was viewed by Discovery studio 4.1 and shown in Fig. 2.

Quality assessment and visualization

Reliability of the generated model was initially checked by ERRAT that analyzed the statistics of non-bonded interactions between diverse atom types, based on characteristic atomic interactions. The overall quality factor was found as 78.99%, sufficient enough to use this model. As demonstrated by the Verify3D program, the results showed that 8.33% of residues had an average 3D (atomic model) – 1D (amino acid) score ≥ 0.2 meaning that this structures was compatible and genuinely good.

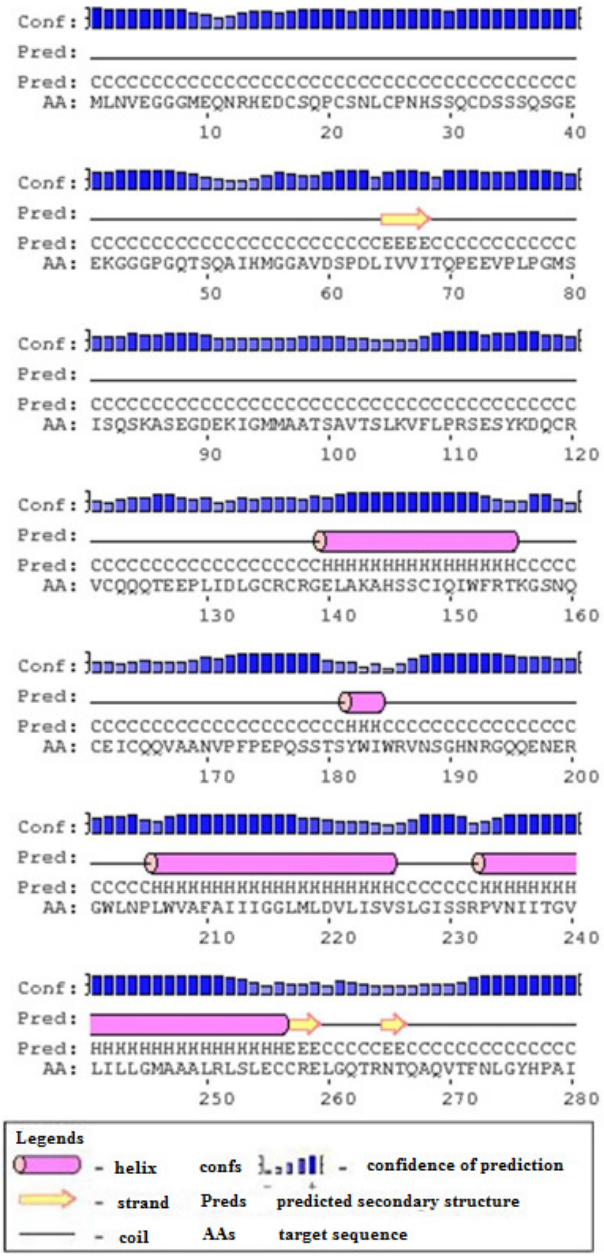

Fig. 1 Predicted secondary structure
of *M. acuminata* hypothetical protein XP_009393594.1 by PSIPRED server

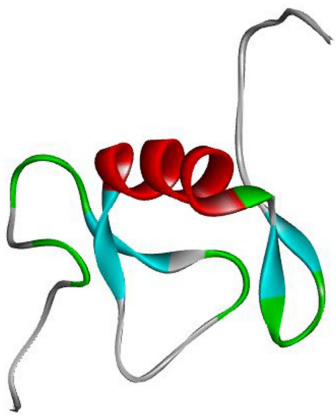

Fig. 2 Structural analysis of *M. acuminata* hypothetical protein XP_009393594.1

Ramachandran plots were resolved. *Z*-score of the query model was checked from PROSAweb. The model's *Z*-score was not shown due to novelty of XP_009393594.1 protein. The *Z*-score is used to estimate the quality of model using structured solved proteins as references [3]. The stereo chemical quality of the model protein was examined using Ramachandran plots through the RAMPAGE server. Ramachandran plot analysis observed 80% of residues of the protein's model structure in the favored region, with 13.8% and 6.2% residues in allowed and outlier regions, individually, indicating that the model was reliable and of good quality shown in Fig. 3. The final protein structure was deposited in PMDB and is available under ID PM0080295.

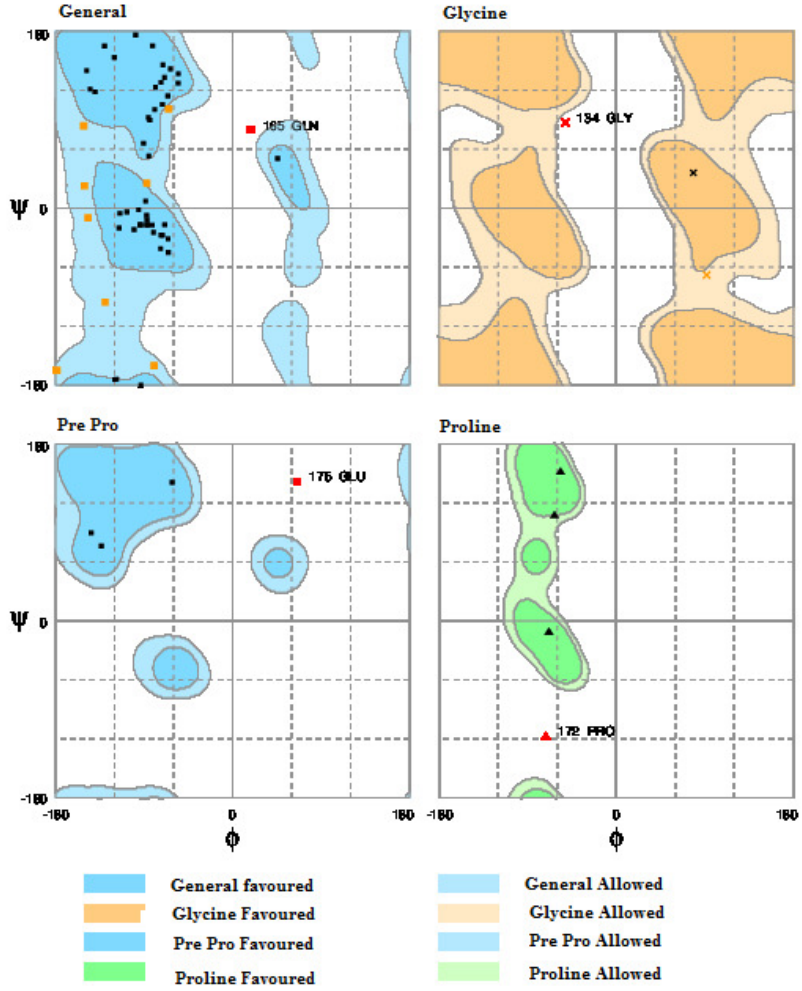

Fig. 3 Ramachandran plot for the 3D model of the studied hypothetical protein XP_009393594.1 by RAMPAGE server

Functional annotation of XP_009393594.1

We utilized three web tools to search the potential functions of XP_009393594.1. In light of predictions made by ProtFun 2.2, and KEGG (KAAS), XP_009393594.1 was suggested as a transcription factor with cytoplasmic domain as a zinc finger. Zinc finger (Znf) domains are moderately small protein motifs which contain multiple finger-like protrusions that make tandem contacts with their target molecule. Some of these domains bind zinc, but numerous do not. The protein was found to be crucial for the translation process, involved in metabolism, signaling and cellular processes and genetic information processing, and Zinc ion binding. The function of XP_009393594.1 is shown in Table 1.

Table 1. Function involvement of XP_009393594.1 in different functional pathways

Metabolism	Genetic information processing	Signaling and cellular processes
Enzymes Protein kinases Protein phosphatases and associated proteins Peptidases Glycosyl transferases Lipopolysaccharide biosynthesis proteins Lipid biosynthesis proteins Poly ketide biosynthesis proteins Prenyl transferases Amino acid related enzymes Cytochrome P450 Photosynthesis proteins	Transcription factors Transcription machinery Messenger RNA biogenesis Spliceosome Ribosome Ribosome biogenesis Transfer RNA biogenesis Translation factors Chaperones and folding catalysts SNAREs Ubiquitin system Proteasome DNA replication proteins Chromosome and associated proteins DNA repair and recombination proteins Mitochondrial biogenesis	Transporters Secretion system proteins Bacterial toxins Two-component system Bacterial motility proteins Cytoskeleton proteins Exosome Prokaryotic defense system G Protein-coupled Receptors Enzyme-linked receptors Cytokine receptors Nuclear receptors Ion Channels GTP-binding proteins Cytokines Cell adhesion molecules and their ligands CD molecules Proteoglycans Glycosaminoglycan binding proteins Lectins Non-coding RNAs

Comparative genome analysis of XP_009393594.1

We utilized NCBI Blast search tool for comparative genome analysis of XP_009393594.1 hypothetical protein of *M. acuminata* with other plant genomes. In the result the XP_009393594.1 showed highest similarities with other uncharacterized hypothetical proteins of several plants.

Conclusion

The present study was directed to create the first 3D structure and propose possible functions of the *M. acuminata* hypothetical protein XP_009393594.1. The 3D model of the protein was constructed using *Ab initio* method as well as refined by few structural assessment methods and the final outcome was genuinely great. We observed that this novel protein is a stable cytoplasmic protein and function as a transcription factor with a zinc finger domain. The protein was observed to be crucial for the translation process, also involved in metabolism, signaling and cellular processes and genetic information processing. The molecular function of protein was found as Zinc ion binding. From genomic similarities we conclude that this hypothetical protein may be checked for same function as of XP_009393594.1. Moreover, this sort of methodology could be helpful in the structure and functions prediction of other uncharacterized proteins.

Acknowledgements

Authors greatly acknowledge the support provided by the Department of Bioinformatics and Bioinformatics Research Club Government Post Graduate College Mandian for conducting

this research work. Authors are thankful to Dr. Sahar Fazal, Professor in Mohammad Ali Jinnah University Islamabad and M.Sc. Zanib Khan, Lecturer of Microbiology at Government Post Graduate College Mandian, Abbottabad, for reviewing the manuscript.

References

1. Oany A. R., T. P. Jyoti, S. A. Ahmad (2014). An *in silico* Approach for Characterization of an Aminoglycoside Antibiotic-resistant Methyltransferase Protein from *Pyrococcus furiosus* (DSM 3638), Bioinformatics and Biology Insights, 8, 65-72.
2. Arnold K., L. Bordoli, J. Kopp, T. Schwede (2006). The SWISS-MODEL Workspace: A Web-based Environment for Protein Structure Homology Modelling, Bioinformatics, 22(2), 195-201.
3. Benkert P., M. Biasini, T. Schwede (2011). Toward the Estimation of the Absolute Quality of Individual Protein Structure Models, Bioinformatics, 27(3), 343-350.
4. Benkert P., S. C. Tosatto, D. Schomburg (2008). QMEAN: A Comprehensive Scoring Function for Model Quality Assessment, Proteins: Structure, Function and Genetics, 71(1), 261-277.
5. Bowie J. U., R. Luethy, D. Eisenberg (1991). A Method to Identify Protein Sequences that Fold into a Known Three-dimensional Structure, Science, 253(5016), 164-170.
6. Buchan D. W., F. Minneci, T. C. Nugent, K. Bryson, D. T. Jones (2013). Scalable Web Services for the PSIPRED Protein Analysis Workbench, Nucleic Acids Research, 41(W1), W349-W357.
7. Butt A. M., M. Batool, Y. Tong (2011). Homology Modeling, Comparative Genomics and Functional Annotation of *Mycoplasma genitalium* Hypothetical Protein MG_237, Bioinformation, 7(6), 299-303.
8. Colovos C., T. O. Yeates (1993). Verification of Protein Structures: Patterns of Nonbonded Atomic Interactions, Protein Science, 2(9), 1511-1519.
9. D'Hont A., F. Denoeud, J.-M. Aury, F.-C. Baurens, F. Carreel, O. Garsmeur, B. Noel, S. Bocs, G. Droc, M. Rouard, C. Da Silva, K. Jabbari, C. Cardi, J. Poulain, M. Souquet, K. Labadie, C. Jourda, J. Lengelle, M. Rodier-Goud, A. Alberti, M. Bernard, M. Correa, S. Ayyampalayam, M. R. Mckain, J. Leebens-Mack, D. Burgess, M. Freeling, D. Mbeguie-A-Mbeguie, M. Chabannes, T. Wicker, O. Panaud, J. Barbosa, E. Hribova, P. Heslop-Harrison, R. Habas, R. Rivallan, P. Francois, C. Poiron, A. Kilian, D. Burthia, C. Jenny, F. Bakry, S. Brown, V. Guignon, G. Kema, M. Dita, C. Waalwijk, S. Joseph, A. Dievart, O. Jaillon, J. Leclercq, X. Argout, E. Lyons, A. Almeida, M. Jeridi, J. Dolezel, N. Roux, A.-M. Risterucci, J. Weissenbach, M. Ruiz, J.-C. Glaszmann, F. Quetier, N. Yahiaoui, P. Wincker (2012). The Banana (*Musa acuminata*) Genome and the Evolution of Monocotyledonous Plants, Nature, 488(7410), 213-219.
10. Finn R. D., J. Mistry, J. Tate, P. Coggill, A. Heger, J. E. Pollington, O. L. Gavin, P. Gunasekaran, G. Ceric, K. Forslund, L. Holm, E. L. L. Sonnhammer, S. R. Eddy, A. Bateman (2010). The Pfam Protein Families' Database, Nucleic Acids Research, 38(Suppl1), D211-D222.
11. Galperin M. Y., E. V. Koonin (2004). 'Conserved Hypothetical' Proteins: Prioritization of Targets for Experimental Study, Nucleic Acids Research, 32(18), 5452-5463.
12. Gasteiger E., C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, A. Bairoch (2005). Protein Identification and Analysis Tools on the ExPASy Server, In: The Proteomics Protocols Handbook, Walker J. M. (Ed.), Humana Press, 571-607.
13. Geourjon C., G. Deléage (1995). SOPMA: Significant Improvements in Protein Secondary Structure Prediction by Consensus Prediction from Multiple Alignments, Computer Applications in Biosciences, 11(6), 681-684.
14. Idrees S., S. Nadeem, S. Kanwal, B. Ehsan, A. Yousaf, S. Nadeen, M. I. Rajoka (2012).

In silico Sequence Analysis, Homology Modeling and Function Annotation of *Ocimum basilicum* Hypothetical Protein G1CT28_OCIBA, Int J of Bioautomation, 16(2), 111-118.
15. Jones D. T. (1999). Protein Secondary Structure Prediction Based on Position-specific Scoring Matrices, Journal of Molecular Biology, 292, 195-202.
16. Kiefer F., K. Arnold, M. Kunzli, L. Bordoli, T. Schwede (2009). The SWISS-MODEL Repository and Associated Resources, Nucleic Acids Research, 37(1), D387-D392.
17. Loewenstein Y., D. Raimondo, O. C. Redfern, J. Watson, D. Frishman, M. Linial, C. Orengo, J. Thornton, A. Tramontano (2009). Protein Function Annotation by Homology-based Inference, Genome Biology, 10:207.
18. Lovell S. C., I. W. Davis, W. B. Arendall IIIrd, P. I. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, D. C. Richardson (2003). Structure Validation by Cα Geometry: ϕ, ψ and Cβ Deviation, Proteins: Structure, Function, and Genetics, 50, 437-450.
19. Lubec G., L. Afjehi-Sadat, J. W. Yang, J. P. John (2005). Searching for Hypothetical Proteins: Theory and Practice Based upon Original Data and Literature, Progress of Neurobiology, 77(1-2), 90-127.
20. Marchler-Bauer A., S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, D. Jackson, Z. Ke, C. J. Lanczycki, F. Lu, G. H. Marchler, M. Mullokandov, V. Omelchenko, C. L. Robertson, J. S. Song, N. Thanki, R. A. Yamashita, D. Zhang, M. N. Zhang, C. Zheng, S. H. Bryant (2011). CDD: A Conserved Domain Database for the Functional Annotation of Proteins, Nucleic Acids Research, 39(Database Issue), D225-D229.
21. Minion F., E. J. Lefkowitz, M. L. Madsen, B. J. Cleary, S. M. Swartzell, G. G. Mahairas (2004). The Genome Sequence of *Mycoplasma hyopneumoniae* Strain 232, the Agent of Swine Mycoplasmosis, Journal of Bacteriology, 186(21), 7123-7133.
22. Moriya Y., M. Itoh, S. Okuda, A. C. Yoshizawa, M. Kanehisa (2007). KAAS: An Automatic Genome Annotation and Pathway Reconstruction Server, Nucleic Acids Research, 35(Web Server Issue), W182-W185.
23. Nimrod G., M. Schushan, D. M. Steinberg, N. Ben-Tal (2008). Detection of Functionally Important Regions in "Hypothetical Proteins" of Known Structure, Structure, 16(12), 1755-1763.
24. Oany A. R., S. A. I. Ahmad, M. A. A. Siddikey, M. U. Hossain, A. Ferdoushi (2014). Computational Structure Analysis and Function Prediction of an Uncharacterized Protein (I6U7D0) of *Pyrococcus furiosus* COM1, Austin J Comput Biol Bioinform, 1(2):5.
25. Paul S., M. Saha, N. C. Bhoumik, S. N. Talukdar (2015). *In silico* Structural and Functional Annotation of *Mycoplasma genitalium* Hypothetical Protein MG_377, International Journal Bioautmation, 19(1), 15-24.
26. Pettersen E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, J. E. Ferrin (2004). UCSF Chimera: A Visualization System for Exploratory Research and Analysis, Journal of Computational Chemistry, 25(13), 1605-1612.
27. Rost B., G. Yachdav, J. Liu (2004). The PredictProtein Server, Nucleic Acids Research, 32(Web Server Issue), W321-W326.
28. Sippl M. J. (1993). Recognition of Errors in Three-dimensional Structures of Proteins, Proteins, 17, 355-362.
29. Wiederstein M., M. J. Sippl (2007). ProSA-web: Interactive Web Service for the Recognition of Errors in Three-dimensional Structures of Proteins, Nucleic Acids Research, 35(Web Server Issue), W407-W410.
30. Yu C. S., Y. C. Chen, C. H. Lu, J. K. Hwang (2006). Prediction of Protein Subcellular Localization, Proteins: Structure, Function and Bioinformatics, 64, 643-651.

31. Yu N. Y., J. R. Wagner, M. R. Laird, G. Melli, S. Rey, R. Lo, P. Dao, S. C. Sahinalp, M. Ester, L. J. Foster, F. S. Brinkman (2010). PSORTb 3.0: Improved Protein Subcellular Localization Prediction with Refined Localization Subcategories and Predictive Capabilities for All Prokaryotes, Bioinformatics, 26(13), 1608-1615.
32. Zdobnov E. M., R. Apweiler (2001). InterProScan: An Integration Platform for the Signature-recognition Methods in InterPro, Bioinformatics, 17(9), 847-848.
33. Zhang R., H. Y. Ou, C. T. Zhang (2004). DEG: A Database of Essential Genes, Nucleic Acids Research, 32(Database Issue), D271-D272.

Anum Munir, B.S. (hons) Student
E-mail: anummunir786@yahoo.com

Anum Munir is a student of B.S. (hons) bioinformatics at Government Post Graduate College Mandian (GPGCM), Abbottabad, Pakistan. She works as a bioinformatics researcher and is an active participant in certain ongoing research projects. She is a dedicated member of Bioinformatics Research Club GPGCM. She has command over many bioinformatics tools and software used in drug design and in other activities related to bioinformatics. Her major interests and expertise are in drug discovery, drug design, phylogenetic analysis, protein structure prediction and molecular biology. Anum Munir also works on various programming languages such as R, MATLAB, Java, C++ and Python. Her recent publications are "Computational Drug Zinpip-Analog as an Ultimate Solution to Cure Conserved Domains of Mutant EGFR, ALK, and BRAF Proteins in NSCLC" and "Repositioning of Methotrexate to Cure Mutations of DYR, TYMS and PURA Gene, Bleomycin in the Treatment of Mutant DNA, TOPa, POLa, and RIR Genes". Another two publications of Anum Munir are under processing.

Assoc. Prof. Azhar Mehmood, Ph.D.
E-mail: azharmehmood35@yahoo.com

Azhar Mehmood has done his Ph.D. in Phyto-Sociology. He has a M. Phil. in Tissue Culture. He is working as an Associate Professor and he is a head of Department of Bioinformatics at GPGCM, Abbottabad, Pakistan. Azhar Mehmood is also a team member of Bioinformatics Research Club GPGCM. He has a teaching experience of 18 years and his major interests are in phyto-sociology, genetics and tissue culture.

Shumaila Azam, M. Phil. (hons) Student
E-mail shumailaazam@hotmail.com

Shumaila Azam is a M. Phil. student in Bioinformatics at Mohammad Ali Jinnah University (MAJU), Islamabad, Pakistan. She completed her M.Sc. degree from the same university. She is also working as a lecturer of Bioinformatics at GPGCM, Abbottabad, Pakistan and Shaheed Benazir University Swabi, Pakistan. She is a dedicated member of research group of MAJU and also a team leader of Bioinformatics Research Club GPGCM. She has command over many bioinformatics tools and software. Her major interest is in drug discovery, drug design, phylogenetic analysis and metabolic pathway determination. She is also participating in many ongoing research projects. Shumaila Azam also works on various programming languages such as R, MATLAB, Java, C++, Perl, Python and Oracle. Her recent publications are "Computational Drug Zinpip-Analog as an Ultimate Solution to Cure Conserved Domains of Mutant EGFR, ALK, and BRAF Proteins in NSCLC" and "Repositioning of Methotrexate to Cure Mutations of DYR, TYMS and PURA Gene, Bleomycin in the Treatment of Mutant DNA, TOPa, POLa, and RIR Genes". Another two research manuscripts – P53 signaling pathway and *Musa* genome resistant cultivar – are under processing.