

## Comprehensive interactome development technique

Janardan Patil\*, Li Len, Abhinav Bharat, and Xi Li

Department of Computer Science, Nagpur University, India 560987

\*Email: [jpatilnagpuredu@gmail.com](mailto:jpatilnagpuredu@gmail.com)

Phone: +91 9877659876

Interactome depicts the arrangement of all atomic communications in cells, particularly with regards to protein-protein collaborations. We look at different strategies for foreseeing protein-protein collaborations utilizing grouping and structure data. A definitive objective of those methodologies is to introduce the total approach for the programmed choice of communication accomplices utilizing their amino corrosive arrangements as well as three dimensional structures, whenever known. The proposed approval of the hypothetical strategies utilizing test information would be a superior appraisal of their exactness.

Keywords: Protein complexes, Docking.

## INTRODUCTION

The combination of different assets into a solitary, openly accessible metadatabase would consider the simple extraction of significant natural data from the monstrous measure of gathered information. Besides, the ongoing advances in hypothetical approach have been a urgent advance into an orderly comprehension of the sub-atomic apparatus. Most strategies for the expectation of protein-protein communications use grouping data to prepare different regulated AI calculations. Three-dimensional structures gained from the Protein Data Bank (PDB) [1] are

additionally used to improve the precision of forecasts. The forecast of protein-protein communications is a troublesome issue when an investigation of both the protein arrangements and known three-dimensional structures is required. There are in any event two explanations behind that. Initially, the quantity of potential communications that must be considered is incredibly huge. In yeast cells, one would need to dissect a potential 18,000,000 communications between 6,000 proteins encoded in its genome, not including different variations of quality items originating from elective joining or post-translational alterations. Just a little bit of those sets is really present in cells [3, 4].

## MATERIALS AND METHODS

**Information bases** The primary objective for researchers of post-genomic science is understanding the perplexing organization of connecting proteins, DNA, RNA and little synthetic particles in living cells. Proteins are significant pieces of those organizations, and data about their structure, amino corrosive succession and utilitarian setting is significant for increasing a superior understanding into the entire cell. Moreover, information about the communication accomplices of a chose protein assists with getting a more point by point portrayal or even forecast of that protein's capacity [14, 15]. The affirmed protein-protein pairings includes transient connections, and the second spotlights on buildings, for example stable associations. The greater part of the information bases utilize their own personal configuration for

the information, so their present degree of combination is restricted. The hypothetical examination of cooperations relies upon heterogeneous wellsprings of organic data, for example, succession and basic information bases, the writing, and exploratory information. The principle data sets containing trial data about protein-protein collaborations are: the Database of Interacting Proteins (DIP) [17], the Biomolecular Interaction Network Database (BIND) [18], the Molecular Interaction Database (MINT) [19], INTACT [20, 21], and Human Protein Reference Database (HPRD) [22]. The writing information on chose protein successions is accessible from the iHOP [23] and STRING [24] data sets. In the Protein Data Bank (PDB) [1] information base, one can locate the three-dimensional structures of protein edifices, while in SCOP, one can discover protein areas, though protein families in PFAM [25].

Nitty gritty trial information is required for a superior comprehension of the utilitarian principles overseeing cell life on the atomic level. In the course of the most recent decade, we could watch critical advancement in the exploratory methods for the ID of communications between proteins. A few sorts of trial test, for example, the yeast two-crossover measure [37-40] or couple liking cleaning [41], permit the high productivity exploratory examination of protein-protein connections in general proteome scale (for a survey of the trial techniques, if it's not too much trouble allude

to [11, 12, 42-52]). Those advances in test philosophy immediately prompted progress in hypothetical methodologies, for example, those dependent on homology [53], protein pathways investigation [54], multimeric stringing [55], or the expectation of connection locales by docking strategies [56]. The last relies upon recently obtained threedimensional protein structures [57]. New information from high throughput techniques for auxiliary genomics take into account further advances in this field. The connecting locales on protein surfaces are frequently hydrophobic [58, 59], with developmentally preserved polar buildups, alleged "problem areas" [60-62].

The strategies use successions of proteins or their known three-dimensional structures so as to foresee cooperations. Much of the time, the three-dimensional structures of the interfacing atoms are not known. Subsequently, a large portion of the current strategies center around succession based protein-protein association expectations, and use collaborating groupings as their preparation sets. Information about complex structures helps in the determination of positives and negatives for preparing, which is of critical significance for AI strategies. Zhou and Shan [89] anticipated connections between proteins utilizing neural organizations prepared on the grouping profiles of cooperating proteins for 615 sets of non-homologous proteins and anticipated dissolvable availability buildups. In autonomous trial of 129

sets of proteins, their strategy predicts that 70% of 11,004 deposits participate in the arrangement of the complex. The rundown of neighbor deposits on a protein chain and dissolvable openness are not subject to basic changes during complex arrangement. In this manner, the exactness of the technique isn't more awful when single protein gems are utilized rather than three-dimensional complex structures [6].

## RESULTS AND DISCUSSION

Ongoing advances in System Biology, particularly with regards to high-throughput DNA sequencing (genomics), quality articulation (transcriptomics), metabolite and particle examination (metabolomics/ionomics) and protein investigation (proteomics) conveys with it the test of preparing and deciphering the gathering informational collections [37-35]. Openly available information bases and bioinformatic devices are utilized to mine this information so as to channel important connections and make models portraying physiological states [43]. Recreating the organizations of collaborations of the different cell segments as compound exercises and edifices, quality articulation, metabolite pools or pathway transition modes is subsequently conceivable. Notwithstanding, catching the collaborations of organization components requires exploratory arrangements with an assortment of conditions [90, 95-99]. A definitive objective of frameworks science with regards to plant research is to

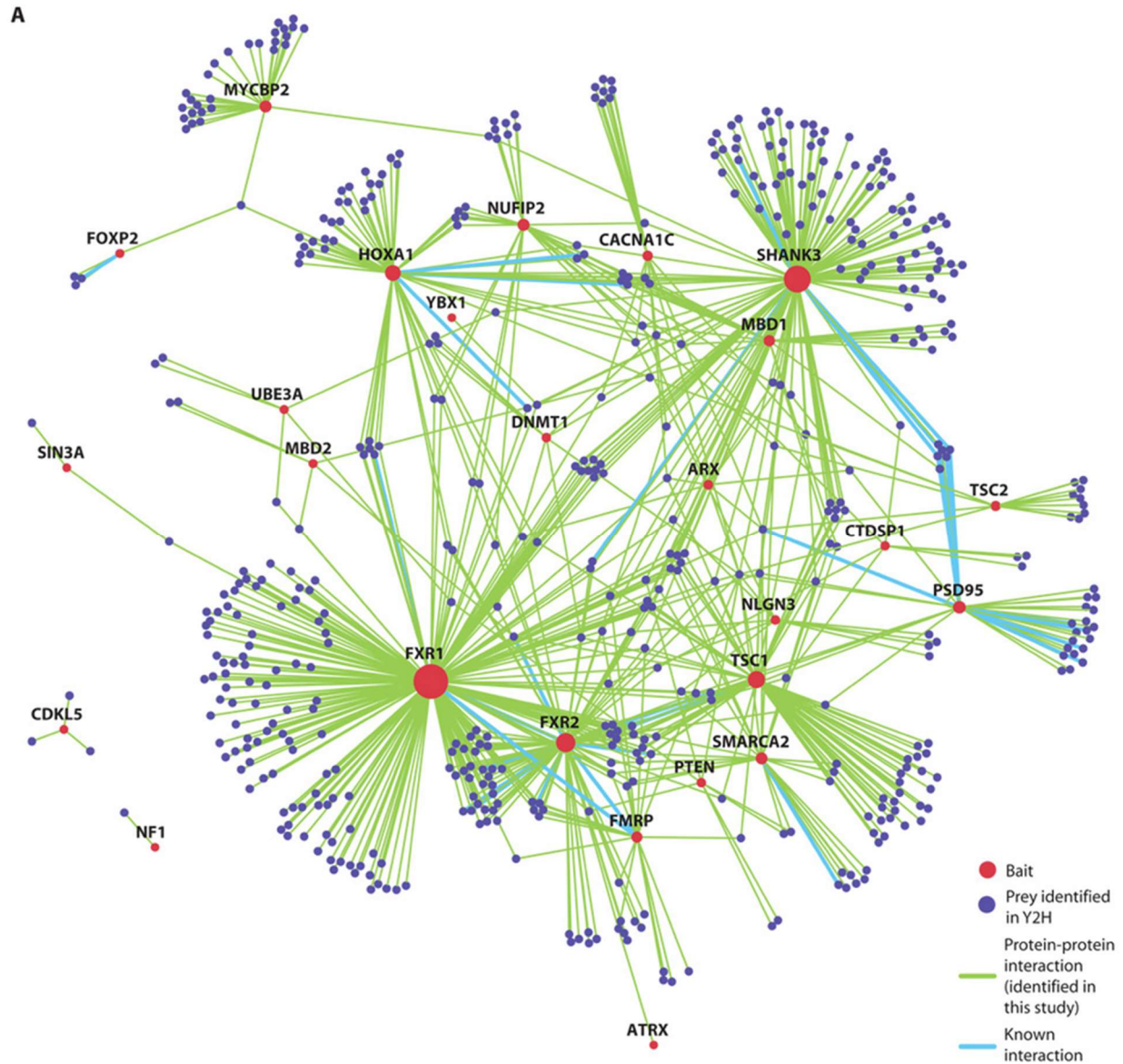


Figure 1: Comprehensive interactome of listed proteins.

comprehend the atomic standards overseeing plant reactions and reliably clarify plant physiology [90, 100, 101]. These methodologies were depicted in detail in this very diary in the 2001-2008 period. Late investigations of protein-protein connections zeroed in on utilizing some worldwide succession or auxiliary highlights. For instance Sprinzak et al. [35] introduced a direct relapse strategy

prepared on nine worldwide protein credits, for example, area signature, overlay type, quality combination, phylogenetic profile, quality setting, protection of neighboring qualities, protein confinement, sort of atomic pathway, mRNA coexpression, or record coregulation. A protein could then be spoken to as focuses in nine-dimensional space utilizing those highlights, and straight relapse was applied. The issue is convoluted and non-unimportant. Appropriate determination of the protein space is essential [46-50]. Notwithstanding unadulterated compound information [109-116] with regards to the Drug Discovery [50-57], there is likewise a requirement for some information on protein-protein associations, the excellent auxiliary expectation of proteins [2, 128-136] and their inhibitors, and a nitty gritty comprehension of how those inhibitors influence the atomic acknowledgment between proteins.

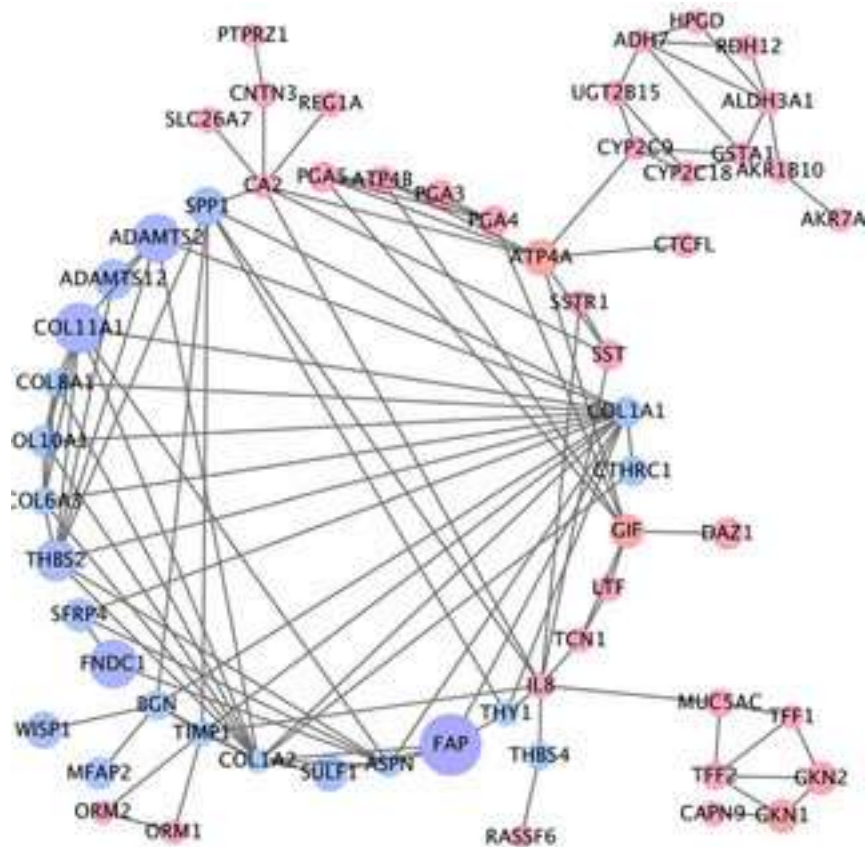


Figure 2: High degree interactome of listed proteins.

The subsequent stage in the improvement of computational techniques depends on a few basic highlights of both collaborating accomplices anticipated from a grouping. The arrangements can be spoken to as sets of short sections with a realized homology profile (work for the entire proteins). Such a methodology was first utilized by Ofran and Rost in 2003 [53], and later by Fernandez-Ballester and Serrano [54] as position-explicit numerous grouping arrangement frameworks. Both wide kinds of communications (steady and transient) are portrayed as collaborations



between some key amino acids of the two accomplices. The best possible determination of significant associating buildups takes into consideration the exhibition of more certain expectations. Along these lines, the investigation of contact maps in protein edifices with anticipated neighborhood auxiliary adaptation of the fundamental chain enhanced by homology profiles will permit the assortment of a more characteristic preparing set for AI techniques. The best possible portrayal of both the succession and structure of the connecting accomplices is of urgent significance for the further improvement of bioinformatics computational strategies for protein-protein cooperation expectation [55, 56]. The arrangement based techniques commonly look for homologues of both communicating proteins with apparatuses, for example, PSI-BLAST [26] or RPS-BLAST. In the event that some communicating accomplices are available in the two sets, the protein pair is probably going to associate. Those strategies are not yet delicate enough on account of removed homology. A few proteins from enormous and assorted superfamilies are disparate as far as succession: just a crease and a couple of significant buildups in a functioning site are protected [69]. The absence of clear succession closeness makes new family ID or capacity expectation more troublesome. Protein-protein collaboration forecast is additionally more unpredictable when no nearby interfacing homologs are known. Structure-based techniques have a restricted zone of handiness because of their reliance on known structures of protein-protein buildings.

## REFERENCES

1. Goh, C.S. and Cohen, F.E. Co-evolutionary analysis reveals insights into protein-protein interactions. *J. Mol. Biol.* 324 (2002) 177-192.
2. Sharan, R., Ideker, T., Kelley, B., Shamir, R. and Karp, R.M. Identification of protein complexes by comparative analysis of yeast and bacterial protein interaction data. *J. Comput. Biol.* 12 (2005) 835-846.
3. Barash, Y., Elidan, G., Kaplan, T. and Friedman, N. CIS: compound importance sampling method for protein-DNA binding site p-value estimation. *Bioinformatics* 21 (2005) 596-600.
4. Sharan, R., Suthram, S., Kelley, R.M., Kuhn, T., McCuine, S., Uetz, P., Sittler, T., Karp, R.M. and Ideker, T. Conserved patterns of protein interaction in multiple species. *Proc. Natl. Acad. Sci. USA* 102 (2005) 1974-1979.
5. Kelley, B.P., Sharan, R., Karp, R.M., Sittler, T., Root, D.E., Stockwell, B.R. and Ideker, T. Conserved pathways within bacteria and yeast as revealed by global protein network alignment. *Proc. Natl. Acad. Sci. USA* 100 (2003) 11394-11399.

6. Salwinski, L., Miller, C.S., Smith, A.J., Pettit, F.K., Bowie, J.U. and Eisenberg, D. The Database of Interacting Proteins: 2004 update. *Nucleic Acids Res.* 32 (2004) D449-D451.
7. Pawar S, Ashraf MI, Mujawar S, Mishra R and Lahiri C (2018) *In silico* Identification of the Indispensable Quorum Sensing Proteins of Multidrug Resistant *Proteus mirabilis*. *Front. Cell. Infect. Microbiol.* 8:269. doi: 10.3389/fcimb.2018.00269
8. Stanam A, Chaudhari M, Pawar S, Rayudu D. Effects of temperature on COVID-19 transmission. medRxiv. doi: <https://doi.org/10.1101/2020.03.29.20044461>
9. Orchard, S., Risse, J., Robbe, K., Roechert, B., Thorneycroft, D., Zhang, Y., Apweiler, R. and Hermjakob, H. IntAct - open source resource for molecular interaction data. *Nucleic Acids Res.* 35 (2007) D561-565.
10. Alfarano, C., Andrade, C.E., Anthony, K., Bahroos, N., Bajec, M., Bantoft, K., Betel, D., Bobechko, B., Boutilier, K., Burgess, E., Buzadzija, K., Cavero, R., D'Abreo, C., Donaldson, I., Dorairajoo, D., Dumontier, M.J., Dumontier, M.R.,

Earles, V., Farrall, R., Feldman, H., Garderman, E., Gong, Y., Gonzaga, R., Grytsan, V., Gryz, E., Gu, V., Haldorsen, E., Halupa, A., Haw, R., Hrvojic, A., Hurrell, L., Isserlin, R., Jack, F., Juma, F., Khan, A., Kon, T., Konopinsky, S., Le, V., Lee, E., Ling, S., Magidin, M., Moniakis, J., Montojo, J., Moore, S., Muskat, B., Ng, I., Paraiso, J.P., Parker, B., Pintilie, G., Pirone, R., Salama, J.J., Sgro, S., Shan, T., Shu, Y., Siew, J., Skinner, D., Snyder, K., Stasiuk, R., Strumpf, D., Tuekam, B., Tao, S., Wang, Z., White, M., Willis, R., Wolting, C., Wong, S., Wrong, A., Xin, C., Yao, R., Yates, B., Zhang, S., Zheng, K., Pawson, T., Ouellette, B.F. and Hogue, C.W. The Biomolecular Interaction Network Database and related tools 2005 update. *Nucleic Acids Res.* 33 (2005) D418-D424.

11. Chatr-Aryamontri, A., Ceol, A., Palazzi, L.M., Nardelli, G., Schneider, M.V., Castagnoli, L. and Cesareni, G. MINT: the Molecular INTeraction database. *Nucleic Acids Res.* 35 (2007) D572-574.

12. Hermjakob, H., Montecchi-Palazzi, L., Lewington, C., Mudali, S., Kerrien, S., Orchard, S., Vingron, M., Roechert, B., Roepstorff, P., Valencia, A., Margalit, H., Armstrong, J., Bairoch, A., Cesareni, G., Sherman, D. and Apweiler, R. IntAct: an open source molecular interaction database. *Nucleic Acids Res.* 32 (2004) D452-D455.

13. Kerrien, S., Alam-Faruque, Y., Aranda, B., Bancarz, I., Bridge, A., Derow, C., Dimmer, E., Feuermann, M., Friedrichsen, A., Huntley, R., Kohler, C., Khadake, J., Leroy, C., Liban, A., Liefertink, C., Montecchi-Palazzi, L., 12 Vol. 14. No. 1. 2009 CELL. MOL. BIOL. LETT.

14. Pawar, S.; Lahiri, C. Quorum sensing: An imperative longevity weapon in Bacteria. Afr J Microbiol Res 2018, 12, 96-104. DOI: 10.5897/AJMR2017.8751.

15. Liu, S.J., Zhang, D.Q., Sui, X.M., Zhang, L., Cai, Z.W., Sun, L.Q., Liu, Y.J., Xue, Y. and Hu, G.F. The inhibition of in vivo tumorigenesis of osteosarcoma (OS)-732 cells by antisense human osteopontin RNA. Cell. Mol. Biol. Lett. 13 (2008) 11-19.

16. Pawar S, Davis CD, Rinehart CA (2011) Statistical analysis of microarray gene expression data from a mouse model of toxoplasmosis. BMC Bioinform 12(Suppl 7):A19

17. Mujawar, S., Mishra, R., Pawar, S., Gatherer, D., Lahiri, C.: Delineating the plausible molecular vaccine candidates and drug targets of multidrug-resistant acinetobacter baumannii. *Front. Cell. Infect. Microbiol.* 9, 203 (2019)
18. Freas, C., Weber, I. T., Pawar, S. D., & Harrison, R. W. (2018). Analysis of drug resistance in HIV protease. *BMC Bioinformatics*, 19, 362. <https://doi.org/10.1186/s12859-018-2331-y>.
19. Peri, S., Navarro, J.D., Amanchy, R., Kristiansen, T.Z., Jonnalagadda, C.K., Surendranath, V., Niranjana, V., Muthusamy, B., Gandhi, T.K., Gronborg, M., Ibarrola, N., Deshpande, N., Shanker, K., Shivashankar, H.N., Rashmi, B.P., Ramya, M.A., Zhao, Z., Chandrika, K.N., Padma, N., Harsha, H.C., Yatish, A.J., Kavitha, M.P., Menezes, M., Choudhury, D.R., Suresh, S., Ghosh, N., Saravana, R., Chandran, S., Krishna, S., Joy, M., Anand, S.K., Madavan, V., Joseph, A., Wong, G.W., Schiemann, W.P., Constantinescu, S.N., Huang, L., Khosravi-Far, R., Steen, H., Tewari, M., Ghaffari, S., Blobe, G.C., Dang, C.V., Garcia, J.G., Pevsner, J., Jensen, O.N., Roepstorff, P., Deshpande, K.S., Chinnaiyan, A.M., Hamosh, A., Chakravarti, A. and Pandey, A. Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome Res.* 13 (2003) 2363-2371.

20. Donthamsetty, S., Pannu, V., Rida, P., Ogden, A., Pawar, S., Bowen, N., Osan, R., Cantuaria, G., and Aneja, R. (2014) KIFCI, a novel putative prognostic biomarker for ovarian adenocarcinomas: delineating protein interaction networks and signaling circuitries. *J. Ovarian Res.* 7, 53

21. Batra, H.; Pawar, S.; Bahl, D. Curcumin in combination with anti-cancer drugs: A nanomedicine review. *Pharm. Res.* 2018, 139, 91–105.

22. Hoffmann, R. and Valencia, A. Implementing the iHOP concept for navigation of biomedical literature. *Bioinformatics* 21 Suppl 2 (2005) ii252-ii258.

23. C.J., Chen, M.J., Liao, Y.L. and Liao, T.N. Polymorphisms of the uridine-diphosphoglucuronosyltransferase 1A1 gene and coronary artery disease. *Cell. Mol. Biol. Lett.* 13 (2008) 1-10. CELLULAR & MOLECULAR BIOLOGY LETTERS

17

24. Huang, B., Chu, C.H., Chen, S.L., Juan, H.F. and Chen, Y.M. A proteomics study of the mung bean epicotyl regulated by brassinosteroids under conditions of chilling stress. *Cell. Mol. Biol. Lett.* 11 (2006) 264-278.

25. Pawar S, Ashraf MI, Mujawar S, Mishra R and Lahiri C (2018) *In silico* Identification of the Indispensable Quorum Sensing Proteins of Multidrug Resistant *Proteus mirabilis*. *Front. Cell. Infect. Microbiol.* 8:269. doi: 10.3389/fcimb.2018.00269
26. Wisniewska, A., Draus, J. and Subczynski, W.K. Is a fluid-mosaic model of biological membranes fully relevant? Studies on lipid organization in model and biological membranes. *Cell. Mol. Biol. Lett.* 8 (2003) 147-159.
27. Pawar, S.; Lahiri, C. Quorum sensing: An imperative longevity weapon in Bacteria. *Afr J Microbiol Res* 2018, 12, 96-104. DOI: 10.5897/AJMR2017.8751.
28. Gronemeyer, H. and Miturski, R. Molecular mechanisms of retinoid action. *Cell. Mol. Biol. Lett.* 6 (2001) 3-52.
29. Pawar, S., Ashraf, M. I., Mujawar, S., Mishra, R., and Lahiri, C. (2018). In silico identification of the indispensable quorum sensing proteins of multidrug resistant *Proteus mirabilis*. *Front. Cell. Infect. Microbiol.* 8:269. doi: 10.3389/fcimb.2018.00269



30. Knizewski, L., Steczkiewicz, K., Kuchta, K., Wyrwicz, L., Plewczynski, D., Kolinski, A., Rychlewski, L. and Ginalski, K. Uncharacterized DUF1574 leptospira proteins are SGNH hydrolases. *Cell Cycle* 7 (2008) 542-544.

31. Korohoda, W. and Wilk, A. Cell electrophoresis - a method for cell separation and research into cell surface properties. *Cell. Mol. Biol. Lett.* 13 (2008) 312-326.

32. Li, J., Ji, C., Zheng, H., Fei, X., Zheng, M., Dai, J., Gu, S., Xie, Y. and Mao, Y. Molecular cloning and characterization of a novel human gene containing 4 ankyrin repeat domains. *Cell. Mol. Biol. Lett.* 10 (2005) 185-193.

33. Wang, Y. F., Wong-Sam, A., Agniswamy, J., Pawar, S., Ghosh, A. K., Harrison, R. W., & Weber, I. T. (2019). Structural studies of antiviral inhibitor with HIV-1 protease bearing drug resistant substitutions of V32I, I47V and V82I. *Biochemical and Biophysical Research Communications*, 514(3), 974–978.  
doi:10.1016/j.bbrc.2019.05.064

34. Miyamoto, T., Sato, H., Yogev, L., Kleiman, S., Namiki, M., Koh, E., Sakugawa, N., Hayashi, H., Ishikawa, M., Lamb, D.J. and Sengoku, K. Is a genetic defect in

Fkbp6 a common cause of azoospermia in humans? *Cell. Mol. Biol. Lett.* 11 (2006) 557-569.

35. Mittal, K., Choi, D.H., Klimov, S. *et al.* A centrosome clustering protein, KIFC1, predicts aggressive disease course in serous ovarian adenocarcinomas. *J Ovarian Res* 9, 17 (2016). <https://doi.org/10.1186/s13048-016-0224-0>

36. Ashraf, M. I., Ong, S. K., Mujawar, S., Pawar, S., More, P., Paul, S., et al. (2018). A side-effect free method for identifying cancer drug targets. *Sci. Rep.* 8:6669. doi: 10.1038/s41598-018-25042-2

37. Cottage, A., Mullan, L., Portela, M.B., Hellen, E., Carver, T., Patel, S., Vavouri, T., Elgar, G. and Edwards, Y.J. Molecular characterisation of the SAND protein family: a study based on comparative genomics, structural bioinformatics and phylogeny. *Cell. Mol. Biol. Lett.* 9 (2004) 739-753

38. Pawar S, Stanam A, Chaudhari M, Rayudu D. Effects of temperature on COVID-19 transmission. medRxiv. doi: <https://doi.org/10.1101/2020.03.29.20044461>

39. Wladyka, B. and Pustelny, K. Regulation of bacterial protease activity. *Cell. Mol. Biol. Lett.* 13 (2008) 212-229.
40. Lahiri, C., Pawar, S., Sabarinathan, R., Ashraf, M. I., Chand, Y., and Chakravorty, D. (2014). Interactome analyses of Salmonella pathogenicity islands reveal SicA indispensable for virulence. *J. Theor. Biol.* 363, 188–197. doi: 10.1016/j.jtbi.2014.08.013
41. Pawar, S., Ashraf, M. I., Mehata, K. M., and Lahiri, C. (2017). “Computational identification of indispensable virulent proteins of Salmonella Typhi CT18,” in *Current Topics in Salmonella and Salmonellosis*, ed M. Mares (InTech Publishers), 21–39.
42. Agoston, V., Cemazar, M., Kajan, L. and Pongor, S. Graph-representation of oxidative folding pathways. *BMC Bioinformatics* 6 (2005) 19.
43. Ashraf MI, Mujawar S, Mishra R and Lahiri C (2018) *In silico* Identification of the Indispensable Quorum Sensing Proteins of Multidrug Resistant *Proteus mirabilis*. *Front. Cell. Infect. Microbiol.* 8:269. doi: 10.3389/fcimb.2018.00269

44. Pawar S, Stanam A, Chaudhari M, Rayudu D. Effects of temperature on COVID-19 transmission. medRxiv. doi: <https://doi.org/10.1101/2020.03.29.20044461>
45. DeLano, W.L. Unraveling hot spots in binding interfaces: progress and challenges. *Curr. Opin. Struct. Biol.* 12 (2002) 14-20.
46. Kajan, L., Kertesz-Farkas, A., Franklin, D., Ivanova, N., Kocsor, A. and Pongor, S. Application of a simple likelihood ratio approximant to protein sequence classification. *Bioinformatics* 22 (2006) 2865-2869.
47. Kocsor, A., Kertesz-Farkas, A., Kajan, L. and Pongor, S. Application of compression-based distance measures to protein sequence classification: a methodological study. *Bioinformatics* 22 (2006) 407-412.
48. Vlahovicek, K., Kajan, L., Agoston, V. and Pongor, S. The SBASE domain sequence resource, release 12: prediction of protein domain-architecture using support vector machines. *Nucleic Acids Res.* 33 (2005) D223-D225.
49. Pawar, S.; Lahiri, C. Quorum sensing: An imperative longevity weapon in Bacteria. *Afr J Microbiol Res* 2018, 12, 96-104. DOI: 10.5897/AJMR2017.8751.

50. Pellegrini, M., Marcotte, E.M. and Yeates, T.O. A fast algorithm for genome-wide analysis of proteins with repeated sequences. *Proteins* 35 (1999) 440-446.
51. Eisen, M.B., Spellman, P.T., Brown, P.O. and Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* 95 (1998) 14863-14868.
52. Sprinzak, E. and Margalit, H. Correlated sequence-signatures as markers of protein-protein interaction. *J. Mol. Biol.* 311 (2001) 681-692.
53. Bock, J.R. and Gough, D.A. Predicting protein--protein interactions from primary structure. *Bioinformatics* 17 (2001) 455-460.
54. Gallet, X., Charlotiaux, B., Thomas, A. and Brasseur, R. A fast method to predict protein interaction sites from sequences. *J. Mol. Biol.* 302 (2000) 917-926.
55. Ofra, Y. and Rost, B. Predicted protein-protein interaction sites from local sequence information. *FEBS Lett.* 544 (2003) 236-239.

56. Jones, S. and Thornton, J.M. Principles of protein-protein interactions. Proc. Natl. Acad. Sci. USA 93 (1996) 13-20.

57. Nooren, I.M. and Thornton, J.M. Diversity of protein-protein interactions. Embo J. 22 (2003) 3486-3492.

58. Pawar S. (2019) Web-Based Application for Accurately Classifying Cancer Type from Microarray Gene Expression Data Using a Support Vector Machine (SVM) Learning Algorithm. In: Rojas I., Valenzuela O., Rojas F., Ortuño F. (eds) Bioinformatics and Biomedical Engineering. IWBBIO 2019. Lecture Notes in Computer Science, vol 11466. Springer, Cham. [https://doi.org/10.1007/978-3-030-17935-9\\_14](https://doi.org/10.1007/978-3-030-17935-9_14)

59. S. Pawar, A. Stanam and Y. Zhu, "Evaluating the computing efficiencies (specificity and sensitivity) of graphics processing unit (GPU)-accelerated DNA sequence alignment tools against central processing unit (CPU) alignment tool", *J. Bioinf. Sequence Anal.*, vol. 9, no. 2, pp. 10-14, 2018.

60. Tuck Onn Liew, Aditya Stanam, Chandrajit Lahiri. Common cancer biomarkers of breast and ovarian types identified through artificial intelligence. *Chemical Biology and Drug Design*, 995-1004, <https://doi.org/10.1111/cbdd.13672>

61. Lahiri, Chandrajit, Pawar, Shrikant and Mishra, Rohit (2019) *Precision medicine and future of cancer treatment*. Precision Cancer Medicine. ISSN 2617-2216

62. Pawar, S., Stanam, A. A Six-Gene-Based Prognostic Model Predicts Survival in Head and Neck Squamous Cell Carcinoma Patients. *J. Maxillofac. Oral Surg.* 18, 320–327 (2019). <https://doi.org/10.1007/s12663-019-01187-z>

63. Pawar, S., Yao, X. & Lu, C. Spermine and oxacillin stress response on the cell wall synthesis and the global gene expression analysis in Methicillin-resistance *Staphylococcus aureus*. *Genes Genom* 41, 43–59 (2019). <https://doi.org/10.1007/s13258-018-0735-8>

64. Nooren, I.M. and Thornton, J.M. Structural characterisation and functional significance of transient protein-protein interactions. *J. Mol. Biol.* 325 (2003) 991-1018.

65. Bahadur, R.P., Chakrabarti, P., Rodier, F. and Janin, J. A dissection of specific and non-specific protein-protein interfaces. *J. Mol. Biol.* 336 (2004) 943-955.

66. Ofraim, Y. and Rost, B. Analysing six types of protein-protein interfaces. *J. Mol. Biol.* 325 (2003) 377-387.

67. Saha, R.P., Bahadur, R.P. and Chakrabarti, P. Interresidue contacts in proteins and protein-protein interfaces and their use in characterizing the homodimeric interface. *J. Proteome Res.* 4 (2005) 1600-1609.

68. Bordner, A.J. and Abagyan, R. Statistical analysis and prediction of protein-protein interfaces. *Proteins* 60 (2005) 353-366.

69. Neuvirth, H., Raz, R. and Schreiber, G. ProMate: a structure based prediction program to identify the location of protein-protein binding sites. *J. Mol. Biol.* 338 (2004) 181-199.