

INTRODUCTION

Open Access

InCoB2014: bioinformatics to tackle the data to knowledge challenge

Shoba Ranganathan^{1*}, Tin Wee Tan², Christian Schönbach^{3,4}

From Asia Pacific Bioinformatics Network (APBioNet) Thirteenth International Conference on Bioinformatics (InCoB2014)

Sydney, Australia. 31 July – 2 August 2014

Abstract

Since 2006, the International Conference on Bioinformatics (InCoB) has been publishing selected papers in BMC Bioinformatics. Papers within the scope of the journal from the 13th InCoB July 31-2 August, 2014 in Sydney, Australia have been compiled in this supplement. These span protein and proteome informatics, structural bioinformatics, software development and bioimaging to pharmacoinformatics and disease informatics, representing the breadth of bioinformatics research in the Asia-Pacific.

Introduction

InCoB (the International Conference on Bioinformatics) has served as the annual conference of the Asia-Pacific Bioinformatics Network (APBioNet) [1], since 2002 and Sydney, Australia was the venue for the 13th InCoB, 31 July-2 August, 2014. In order to provide our region with international, peer-reviewed impact factor journal publications rather than printed books of conference proceedings, APBioNet has setup a rigorous peer review protocol and accepted the best InCoB papers in BMC Bioinformatics supplements since 2006, gradually, adding in BMC Genomics and BMC Systems Biology supplements, with BMC Medical Genomics as well this year. We have briefly reviewed the articles in this supplement, providing the 2014 bioinformatics research update from the APBioNet community.

Manuscript submission and review

InCoB2014 provided authors the choice of submitting original research as full manuscripts to either the BMC track (supplement issues of BMC Bioinformatics, BMC Systems Biology or BMC Genomics) or to a special issue of PeerJ. The statistics for paper submission and

acceptance, along with details of the peer review process undertaken as well as the links to the BMC Systems Biology, BMC Medical Genomics and PeerJ supplements are provided in the InCoB2014 BMC Genomics supplement introduction [2], with 16 “bioinformatics” articles briefly overviewed here.

Protein and proteome informatics

Proteins display diverse functionality and these are usually the consequence of specific binding sites or sequence motifs. Dipeptide propensity scores have provided the solution to successfully predicting heme binding proteins [3]. Among biologically important post-translational modifications, O-linked glycosylation of serine and threonine residues is elusive in that there is no clear sequence motif associated with this site. Wu *et al.* [4] have applied support vector machine (SVM) learning to this problem, outperforming three other currently available tools. While mass spectrometry is frequently used to identify phosphorylated proteins, the low abundance of phosphopeptides in a sample is an obstacle to data analysis. iPhos [5] offers an innovative workflow system for streamlining phosphoproteome analysis.

Structural bioinformatics

The biological function of a protein is ascribed to its 3D structure and five papers [6-10] provide updates on

* Correspondence: shoba.ranganathan@mq.edu.au

¹Department of Chemistry and Biomolecular Sciences and ARC Centre of Excellence in Bioinformatics, Macquarie University, Sydney NSW 2109, Australia

Full list of author information is available at the end of the article

structural bioinformatics research. At the outset, for proteins that have neither experimental structural information nor structural homologues, predicting the 3D structure of a protein continues to remain a challenge. Paliwal *et al.* [6] have integrated evolutionary information and secondary structure prediction for protein 3D fold recognition, while Bhageerath-H [7] provides a novel *ab initio* and homology-based hybrid tertiary structure prediction server. For multi-domain proteins, the domain boundaries may be delineated by identifying inter-domain linker regions [8], for subsequent prediction of domain 3D structures using either [6] or [7]. Liu *et al.* [9] have analysed all available protein 3D structures and developed a new approach to discriminate between biologically relevant protein interactions and crystal packing contacts, while IFACEwat by Su *et al.* [10] seeks to predict near native structures of protein-protein complexes with the inclusion of solvent molecules at the interface.

Bioinformatics software

Coevolution is an unusual phenomenon observed among species from different phyla asserting selective pressures on one another, making cophylogenetic analysis of these species very difficult. TreeCollapse addresses the cophylogeny problem indirectly by using common topological patterns [11]. In the area of bioimaging, phenotypic changes associated with development can be quantitatively analysed live imaging of muscle tissue using the software tool, FMAj [12]. These tools enable scientific advances in research areas where high quality software is lacking.

Where a number of bioinformatics software programs exist, quality assurance strategies are required to verify and validate the results generated. Ho and coworkers [13] show that metamorphic testing can be effectively applied to evaluate the results from two programs, BWA and Bowtie.

Pharmacoinformatics and disease informatics

Generating a 3D map of chemical features of known ligands (a 'pharmacophore'), for effective drug design remains a difficult problem which has been addressed by a pharmacophore-assisted iterative closest point (ICP) method [14]. Grover *et al.* [15] have used virtual screening of the glucagon receptor to identify novel natural inhibitors as potential therapeutic candidates for combating type 2 diabetes.

Prevention is better than cure, especially in the case of Alzheimer's disease and Zhang *et al.* [16] propose a genetic algorithm with logistic regression for the early diagnosis of this disease from the information available from non-invasive neuropsychological tests. This could result in treatment options to prevent the onset or slow down the progression of this disease.

On the other hand, serious negative side reactions of drugs leading to organ failure can pose a serious threat to a subset of sensitive patients. To identify such patients prior to drug therapy, nephrotoxicity can be predicted based on only two genes [17] while multi-organ failure can be anticipated, using an integrative prediction score [18] for gene expression profiles.

Conclusion

The articles in this supplement cover protein, proteome and structural bioinformatics, software packages as well as bioinformatics applications for drug development, early diagnosis of diseases and possible prevention of drug toxicity issues. We believe the Asia-Pacific is on track to participate in the ongoing NIH Big Data to Knowledge (DB2K) [19] and other similar global initiatives. We welcome you to attend our 2015 InCoB meeting to be held jointly with the Genome Informatics Workshop (GIW) in Tokyo, Japan [20], to contribute to this regional bioinformatics effort.

Competing interests

None declared.

Authors' contributions

SR wrote the introduction. CS and SR (Program Committee Co-chairs) managed the review and editorial processes, respectively. TWT supported the post-acceptance manuscript processing.

Acknowledgements

We are indebted to all members of the Program Committee and additional reviewers for their efforts and time. We are grateful to the NSW T&I Conference grant awarded to SR for hosting InCoB2014, our sponsors: the Australian Bioinformatics Network, the International Society for Computational Biology, Qiagen, Millennium Science and ABSciEx, and for material support from Macquarie University and Sydney Business Events. Special thanks go to ASN Events for running the conference smoothly. Finally, we are deeply grateful to Isobel Peters and Jennifer Egar of BMC who have supported us through the supplement publication process.

Declarations

This article has been published as part of *BMC Bioinformatics* Volume 15 Supplement 16, 2014: Thirteenth International Conference on Bioinformatics (InCoB2014): Bioinformatics. The full contents of the supplement are available online at <http://www.biomedcentral.com/bmcbioinformatics/supplements/15/S16>.

Authors' details

¹Department of Chemistry and Biomolecular Sciences and ARC Centre of Excellence in Bioinformatics, Macquarie University, Sydney NSW 2109, Australia. ²Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117599. ³Department of Biology, School of Science and Technology, Nazarbayev University, Astana 010000, Republic of Kazakhstan. ⁴Center for AIDS Research, Kumamoto University, Kumamoto 860-0811, Japan.

Published: 8 December 2014

References

1. The Asia-Pacific Bioinformatics Network. [<http://www.apbionet.org>].
2. Schönbach C, Tan TW, Ranganathan S: InCoB2014: mining biological data from genomics for transforming industry and health. *BMC Genomics* 2014, **15**(Suppl 9):11.

3. Liou YF, Charoenkwan P, Srinivasulu YS, Vasylenko T, Lai SC, Lee HC, Chen YH, Huang HL, Ho SY: **SCMHBP: Prediction and analysis of heme binding proteins using propensity scores of dipeptides.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S4.
4. Wu HY, Lu CT, Kao HJ, Chen YJ, Chen YJ, Lee TY: **Characterization and identification of protein O-GlcNAcylation sites with substrate specificity.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S1.
5. Yang TH, Chang HT, Hsiao ESL, Sun JL, Wang CC, Wu HY, Liao PC, Wu WS: **iPhos: a toolkit to streamline the alkaline phosphatase-assisted comprehensive LC-MS phosphoproteome investigation.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S10.
6. Paliwal KK, Sharma A, Lyons J, Dehzangi A: **Improving protein fold recognition using the amalgamation of evolutionary-based and structural based information.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S12.
7. Jayaram B, Dhingra P, Mishra A, Kaushik R, Mukherjee G, Singh A, Shekhar S, Bhageerath-H: **A homology/ab initio hybrid server for predicting tertiary structures of monomeric soluble proteins.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S7.
8. Shatnawi M, Zaki N, Yoo PD: **Protein inter-domain linker prediction using random forest and amino acid physiochemical properties.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S8.
9. Liu Q, Li Z, Li J: **Use B-factor related features for accurate classification between protein binding interfaces and crystal packing contacts.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S3.
10. Su CTT, Nguyen TD, Zheng J, Kwok CK: **IFACEwat: the interfacial water-implemented reranking algorithm to improve the discrimination of near native structures for protein rigid docking.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S9.
11. Drinkwater B, Charleston MA: **Introducing TreeCollapse: a novel greedy algorithm to solve the cophylogeny reconstruction problem.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S14.
12. Kuleesha Y, Choo PW, Feng L, Wasser M: **FMAJ: a tool for high content analysis of muscle dynamics in Drosophila metamorphosis.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S6.
13. Giannoulatou E, Park SH, Humphreys DT, Ho JWK: **Verification and validation of bioinformatics software without a gold standard: a case study of BWA and Bowtie.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S15.
14. Zhou L, Griffith R, Gaeta BA: **Combining spatial and chemical information for clustering pharmacophores.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S5.
15. Grover S, Dhanjal JK, Goyal S, Grover A, Sundar D: **Computational identification of novel natural inhibitors of glucagon receptor for checking type II diabetes mellitus.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S13.
16. Johnson P, Vandewater L, Wilson W, Maruff P, Savage G, Graham P, Macaulay SL, Ellis KA, Szoeki C, Martins RN, Rowe CC, Masters CL, Ames D, Zhang P: **Genetic algorithm with logistic regression for prediction of progression to Alzheimer's disease.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S11.
17. Su R, Li Y, Zink D, Loo LH: **Supervised prediction of drug-induced nephrotoxicity based on interleukin-6 and -8 expression levels.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S16.
18. Kim J, Shin M: **An integrative model of multi-organ drug-induced toxicity prediction using gene-expression data.** *BMC Bioinformatics* 2014, **15**(Suppl 16):S2.
19. **Big Data to Knowledge (BD2K) initiative.** [<http://bd2k.nih.gov/>].
20. **GIW-InCoB.** 2015 [<http://incob.apbionet.org/incob15>].

doi:10.1186/1471-2105-15-S16-11

Cite this article as: Ranganathan *et al.*: InCoB2014: bioinformatics to tackle the data to knowledge challenge. *BMC Bioinformatics* 2014 **15** (Suppl 16):11.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

