



## Genomics Proteomics Bioinformatics

www.elsevier.com/locate/gpb  
www.sciencedirect.com



### PREFACE

# Disease Biomarkers for Precision Medicine: Challenges and Future Opportunities



Edwin Wang<sup>1,\*</sup>, William C.S. Cho<sup>2,\*</sup>, S.C. Cesar Wong<sup>3,\*</sup>, Siqi Liu<sup>4,\*</sup>

<sup>1</sup> Center for Health Genomics and Informatics, University of Calgary Cumming School of Medicine, Calgary, AB T2N 4N1, Canada

<sup>2</sup> Department of Clinical Oncology, Queen Elizabeth Hospital, Hong Kong Special Administrative Region

<sup>3</sup> Department of Health Technology and Informatics, Faculty of Health and Social Sciences, Hong Kong Polytechnic University, Hong Kong Special Administrative Region

<sup>4</sup> CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China

Received 31 March 2017; accepted 3 April 2017

Available online 7 April 2017

Most of the human diseases are complex diseases, which could be caused by many genetic pathways. This means that for a given phenotype (*i.e.*, a complex disease), there are multiple potential genes which could be genomically or epigenetically changed (*i.e.*, mutations, copy number variations, epigenetic modifications, and so on). Therefore, it is understandable that different individuals who share the same phenotype/diseases may have different causal genes and thus, may have different drug targets. For example, mutated genes are rarely common between the cancer patients of the same cancer type [1]; furthermore, for a given drug, only 10%–30% of the patients of the same cancer type respond to that drug [2]. It has been suggested that genomic and other omic features and/or environmental and lifestyle factors, could contribute to these differences such as drug response. It is clear that we should give the ‘right drug’ to the ‘right patient’ at the ‘right time’. One of the missions of the many ongoing precision medicine programs is to reach this goal using omic (*i.e.*, genomic,

proteomic, epigenetic, and so on) and/or environmental and lifestyle factors of the individuals.

High-throughput technologies drive the evolution of biology and medicine. To realize precision medicine, it is essential to identify biomarkers using either omic data alone or in combination with environmental/lifestyle factors. The challenge is how to transform the data into biomarkers that could predict clinical outcomes, drug response or others. In general, it is difficult to identify ‘high-quality’ biomarkers which have high accuracy and robustness [3,4] using omic data such as gene expression data, proteomic data and so on. For example, many omic-based cancer biomarkers are not robust, meaning that a biomarker identified from a patient cohort loses its predictive power in other cohorts of the same cancer type/subtype [3,5]. Efforts have been made to develop new algorithms to overcome this problem. For example, Multiple Survival Screening (MSS) and Significance Analysis of Prognostic Signatures (SAPS) have been developed for identifying robust cancer biomarkers [4,6,7].

In the post-genome sequencing era, genome sequencing gets cheaper and cheaper, which makes genome sequencing become affordable and accessible to the clinic. Therefore, it is very attractive to identify biomarkers using the whole-genome/whole-exome sequencing data. Nonetheless, given the aforementioned features of the complex diseases, it has proven challenging to construct predictive models (*i.e.*, identify biomarkers) using the whole-genome/whole-exome sequencing data [8]. Because multiple gene interactions govern the

\* Corresponding authors.

E-mail: [edwin.wang@ucalgary.ca](mailto:edwin.wang@ucalgary.ca) (Wang E), [chocs@ha.org.hk](mailto:chocs@ha.org.hk) (Cho WCS), [cesar.wong@polyu.edu.hk](mailto:cesar.wong@polyu.edu.hk) (Wong SCC), [siqiliu@big.ac.cn](mailto:siqiliu@big.ac.cn) (Liu S).

<sup>a</sup> ORCID: 0000-0001-8689-5295.

<sup>b</sup> ORCID: 0000-0003-4174-4586.

<sup>c</sup> ORCID: 0000-0003-0058-140X.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

<http://dx.doi.org/10.1016/j.gpb.2017.04.001>

1672-0229 © 2017 The Authors. Production and hosting by Elsevier B.V. on behalf of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

underlying molecular mechanisms of the complex diseases, the linear model approach is not an option for identifying biomarkers using the genome sequencing data. A network-based, non-linear approach could hold promise to solve this problem. For example, the Cancer Hallmark Network Framework (CHNF) [9] provides a solution to the problem. Recently, a CHNF-based algorithm has been developed to successfully predict breast cancer recurrence using the whole-exome sequencing data of the tumors (Milanese et al., unpublished data).

Omic profiling of cell-free DNA (*i.e.*, liquid biopsy) has opened a new avenue for identifying non-invasive biomarkers, which are extremely useful in clinics. Much more efforts will be made in this direction in the near future. In addition, the recent development of the single-cell omic technology could bring new opportunities for biomarker identification. Finally, almost all of the efforts made have focused on identifying biomarkers using omic data in the past. However, most of the diseases are caused by the interaction of genetics and environmental/lifestyle factors. Therefore, to accurately predict clinical features of the complex diseases, the future work should focus on identifying biomarkers by integrating the data of the omics and environmental/lifestyle factors.

### Competing interests

The authors have declared that there are no competing interests.

### Acknowledgments

We thank all the authors for nice contributions and great input from all the reviewers.

### References

- [1] Cancer Kaiser J. First pass at cancer genome reveals complex landscape. *Science* 2016;313:1370.
- [2] Chan TA, Wolchok JD, Snyder A. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N Engl J Med* 2015;373:1984.
- [3] Diamandis EP. The failure of protein cancer biomarkers to reach the clinic: why, and what can be done to address the problem? *BMC Med* 2012;10:87.
- [4] Li J, Lenferink AE, Deng Y, Collins C, Cui Q, Purisima EO, et al. Identification of high-quality cancer prognostic markers and metastasis network modules. *Nat Commun* 2010;1:34.
- [5] Venet D, Dumont JE, Detours V. Most random gene expression signatures are significantly associated with breast cancer outcome. *PLoS Comput Biol* 2011;7:e1002240.
- [6] Beck AH, Knoblauch NW, Hefti MM, Kaplan J, Schnitt SJ, Culhane AC, et al. Significance analysis of prognostic signatures. *PLoS Comput Biol* 2013;9:e1002875.
- [7] Gao S, Tibiche C, Zou J, Zaman N, Trifiro M, O'Connor-McCourt M, et al. Identification and construction of combinatory cancer hallmark-based gene signature sets to predict recurrence and chemotherapy benefit in stage II colorectal cancer. *JAMA Oncol* 2016;2:37–45.
- [8] Costello JC, Heiser LM, Georgii E, Gönen M, Menden MP, Wang NJ, et al. A community effort to assess and improve drug sensitivity prediction algorithms. *Nat Biotechnol* 2014;32:1202–12.
- [9] Wang E, Zaman N, Mcgee S, Milanese JS, Masoudi-Nejad A, O'Connor-McCourt M. Predictive genomics: a cancer hallmark

network framework for predicting tumor clinical phenotypes using genome sequencing data. *Semin Cancer Biol* 2015;30:4–12.



**Edwin Wang, PhD**, is a Professor and an AIHS Translational Chair in Cancer Genomics/Informatics at the University of Calgary. His primary research interests include computational systems biology and bioinformatics. He is a member of the AACR's cancer systems biology think tank and has served multiple NIH review panels. His pioneering work of the cancer network motifs has been featured in the college textbook, *Genetics* (2014), edited by a Nobel Laureate, Dr. Hartwell. His

pioneering work of the microRNA-human signaling network opens a new research area: network biology of non-coding RNAs. Up to 2016, more than 4000 papers have been published in this area.



**William CS Cho, PhD, MLTB (Part I), RCMP, FHKSMDS, FHKIMLS, FIBMS (UK), Chartered Scientist (UK)**, is a seasoned researcher in proteomics, microRNA and Chinese medicine. His main research interests have been focusing on cancer studies utilizing high-throughput technologies to discover biomarkers for cancer diagnosis, treatment and prognosis. He is also an international grant reviewer of the Hope Funds for Cancer Research (USA),

Cancer Research (UK), MRC Research Grant (UK), Science Foundation (Ireland), The Foundation Fournier-Majoie for Innovation (Brussels) and National Medical Research Council (Singapore), *etc.*



**Cesar Wong, PhD, FF(FSc)RCPA**, is an Associate Professor of Department of Health Technology and Informatics at Hong Kong Polytechnic University. His primary research interests include developing biomarkers with diagnostic and prognostic significance for Asian common cancers such as colorectal cancer, cervical cancer, nasopharyngeal cancer and their pre-malignant states, and developing novel non-invasive molecular and immunocytochemical tests for clinical applications.



**Siqi Liu, PhD**. Professor Liu obtained his PhD degree from University of Texas Medical Branch in 1994 and was appointed as an assistant professor in Department of Medicine, University of Louisville in 1998. In 1999, he worked closely with other three founders to establish Beijing Genomics Institute (BGI), and in 2003 to build up Beijing Institute of Genomics, Chinese Academy of Sciences (BIG, CAS). His main focus is research on pathway regulation in

disease model and technological development in protein science. He had served as the chief scientist for several national “973 Program” projects on tumor and proteome.