

MUSINGS

Expression profiling: a cost-effective biomarker discovery tool for the personal genome era

David Gurwitz*

Despite the declining cost of 'personal genomes' and the acknowledged usefulness of individual genome sequences to understand human health, for the near future, personal genomes alone are unlikely to become the leading research tool in genomic medicine. I argue that, at present, expression profiling studies are the most promising and cost-effective tool for discovering new disease and drug-response biomarkers.

\$1,000 personal genomes? Not around the corner

The proliferation of studies on personal genomes is fueled by reduced sequencing costs and improved bioinformatics software tools. Substantial advances in understanding human biology in health and disease - and improved diagnostic capacities - are foreseen with the availability of personal genome sequences [1,2]. Yet, expectations from personal genomes should not be overstated: individual genome sequencing *per se* - even when combined with comprehensive individual medical records - will not suffice to decipher the overwhelming complexity of human biology and disease pathology. Our current level of understanding of biology in general, and human biology in particular, is simply too low to interpret the information encoded in individual DNA sequences [3].

Moreover, even with dramatically reduced DNA sequencing costs, the high costs of data interpretation subsequent to the tremendous bioinformatics effort of analyzing 6.2 Gb nucleotides of two copies of one individual genome means that personal genome sequencing will not become clinically routine in the near future. Currently, at the 10th anniversary of the Human Genome Project's conclusion, whole genome sequencing is being put into practice in leading medical academic centers, predominantly for cancer diagnosis and treatment. But it will be many years before it becomes a widespread, routine clinical tool.

The case for expression profiling

Regardless of the pace of application of personal genome sequencing in clinical practice, researchers should bear in mind that additional tools are needed to supplement personal genomes so that insights can be made into the molecular biology of complex diseases. Among such tools, genome-wide transcriptomic profiling seems to be a powerful yet straightforward, affordable and easily validated molecular biology technology for discovering new diagnostic biomarkers and drug targets [4,5]. For example, expression profiling currently costs under US\$400 per sample using commercial microarrays. Comparisons of healthy and diseased tissues from the same individual, or the same tissue (in particular, white blood cells) from an individual over time, such as before and following drug treatment, may yield knowledge not extractable from personal genome sequences. The complexity of the interrelationship between DNA sequences and cell biology is likely to be far higher than is currently understood. For example, a new level of complexity linking the genome to the proteome has recently been introduced: it is well established that gene expression is regulated by short (22 to 23 nucleotides long) non-coding RNA sequences termed microRNAs. Now, an additional level of complexity has been discovered: circular non-coding RNAs, which modulate the action of microRNAs on gene expression [6]. Further surprises are likely to be in store for gene expression regulation by non-coding genome sequences, even though these sequences are part of already published - but little understood - personal genomes.

Another key advantage of expression profiling studies is that they also inform about the consequences of epigenomic modifications, as transcriptomes reflect not merely the output of DNA sequences, but also their interplay with non-genetic modifiers of gene expression. Moreover, transcriptomic studies can inform about alternative splicing events - in particular when RNA sequencing is applied - whereas, at our current level of knowledge, personal genome sequences do not have this capacity.

Expression profiling data can thus be far more informative than personal genomes for deciphering cellular

*Correspondence: gurwitz@post.tau.ac.il
Department of Human Molecular Genetics and Biochemistry, Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv 69978, Israel

networks and disease biomarkers, and indicating drug targets. Certainly, having both personal genomes and longitudinal gene expression profiles from the same study participants has clear advantages. Indeed, an integrative 'personal'omics profile that combines genomic, transcriptomic, proteomic, metabolomic and autoantibody profiles from a single individual over 14 months was recently presented [7]. This is undeniably the best way forward for genomic medicine projects when adequate funding is available. However, in the clinical setting, as well as for most academic research groups, costs for such comprehensive projects - in particular for large cohorts - are prohibitive. In lieu of such funding levels, expression profiling seems to offer the most promising and cost-effective approach for genome-wide searches for disease and drug-response biomarkers.

Considerations and limitations

Of course, there are limitations to expression profiling, where each profile represents a single 'snapshot' of a given tissue or cell type at a given time and under distinct physiological conditions. This constraint may be overcome (with extra cost) by performing longitudinal gene expression profiles so that the sequential alterations provide information about molecular events during disease progression, tissue remodeling or drug treatment. This disadvantage of expression profiling can thus be turned into an advantage, in particular for searching drug-response biomarkers, by pinpointing genes or non-coding RNA sequences whose expression levels are modified by a drug of interest. This in turn can be informative for discovery of drug-response biomarkers that can act as companion diagnostics for new drug targets [5].

Personal genome studies entail privacy risks, not only for study participants, but also for their relatives. Researchers are morally obliged to disclose that anonymity promises cannot be made to individuals who consent to genome sequencing [8] and, indeed, there is conclusive evidence that donors of DNA sequences can be identified [9]. By contrast, lesser privacy risks are posed by microarray-based expression profiling studies. This should be considered in particular when studying vulnerable populations such as children [10], where gene expression studies have the additional advantage of affording better privacy protection for research participants.

In conclusion, it seems that considering our limited understanding of biology, and the current research funding situation, expression profiling stands out as the most appropriate and cost-effective methodology to gain new insights into complex disorders and to discover disease and drug-response biomarkers.

Competing interests

The author declares that he has no competing interests.

Acknowledgements

I thank Jeantine Lunshof for critical reading and suggestions.

Published: 14 May 2013

References

1. Berg JS, Khoury MJ, Evans JP: **Deploying whole genome sequencing in clinical practice and public health: meeting the challenge one bin at a time.** *Genet Med* 2011, **13**:499-504.
2. Chen R, Mias GI, Li-Pook-Than J, Jiang L, Lam HY, Chen R, Miriami E, Karczewski KJ, Hariharan M, Dewey FE, Cheng Y, Clark MJ, Im H, Habegger L, Balasubramanian S, O'Huallachain M, Dudley JT, Hillenmeyer S, Haraksingh R, Sharon D, Euskirchen G, Lacroute P, Bettinger K, Boyle AP, Kasowski M, Grubert F, Seki S, Garcia M, Whirl-Carrillo M, Gallardo M, *et al.*: **Personal omics profiling reveals dynamic molecular and medical phenotypes.** *Cell* 2012, **148**:1293-1307.
3. Mardis ER: **The \$1,000 genome, the \$100,000 analysis?** *Genome Med* 2010, **2**:84.
4. Cooper-Knock J, Kirby J, Ferraiuolo L, Heath PR, Rattray M, Shaw PJ: **Gene expression profiling in human neurodegenerative disease.** *Nat Rev Neurol* 2012, **8**:518-530.
5. Boone DN, Lee AV: **Targeting the insulin-like growth factor receptor: developing biomarkers from gene expression profiling.** *Crit Rev Oncol* 2012, **17**:161-173.
6. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J: **Natural RNA circles function as efficient microRNA sponges.** *Nature* 2013, **495**:384-388.
7. Chen R, Mias GI, Li-Pook-Than J, Jiang L, Lam HY, Chen R, Miriami E, Karczewski KJ, Hariharan M, Dewey FE, Cheng Y, Clark MJ, Im H, Habegger L, Balasubramanian S, O'Huallachain M, Dudley JT, Hillenmeyer S, Haraksingh R, Sharon D, Euskirchen G, Lacroute P, Bettinger K, Boyle AP, Kasowski M, Grubert F, Seki S, Garcia M, Whirl-Carrillo M, Gallardo M, *et al.*: **Personal omics profiling reveals dynamic molecular and medical phenotypes.** *Cell* 2012, **148**:1293-1307.
8. Lunshof JE, Chadwick R, Vorhaus DB, Church GM: **From genetic privacy to open consent.** *Nat Rev Genet* 2008, **9**:406-411.
9. Gymrek M, McGuire AL, Golan D, Halperin E, Erlich Y: **Identifying personal genomes by surname inference.** *Science* 2013, **339**:321-324.
10. Gurwitz D, Fortier I, Lunshof JE, Knoppers BM: **Research ethics. Children and population biobanks.** *Science* 2009, **325**:818-819.

doi:10.1186/gm445

Cite this article as: Gurwitz D: Expression profiling: a cost-effective biomarker discovery tool for the personal genome era. *Genome Medicine* 2013, **5**:41.