Biomolecular Detection and Quantification 1 (2014) A1–A3

Contents lists available at ScienceDirect

Biomolecular Detection and Quantification

journal homepage: www.elsevier.com/locate/bdq

Editorial

How to make Mathematics Biology's next and better microscope

An essay published in 2004 has the perceptive heading: "Mathematics Is Biology's Next Microscope, Only Better; Biology Is Mathematics' Next Physics, Only Better" [1]. This title neatly summarises the developmental path taken by biology over the past 30 years or so. It emphasises both the exciting role that mathematics and statistics are adopting in the identification and quantification of wholly unimagined and entirely new realms within biology as well as the repercussions biology is precipitating in the prompting of new developments in mathematics and statistics. The title also underlines the intensifying interlinking of numerous disciplines: the complexity of biological and clinical phenomena requires increasingly sophisticated, specialised and discipline-transcending expertise to complete a technical workflow comprising experimental design, generation and analysis of data, interpretation of results and subsequent transparent reporting (Fig. 1).

Natural systems are highly complex and challenging to model, as variables are frequently random, volatile and even contradictory. Hence the drive to constrain parameters by developing simplified model systems such as organotypic or tissue cultures that are less complex but more defined and so can be more easily controlled. Such systems are of course still highly intricate, and so invite further reductionism in the shape of engineered systems that include a few components in a test tube and eventually result in mathematical modelling, which provides ultimate control but may no longer reflect the complexity under initial investigation.

For many years, biological data were descriptive and even the introduction of cell and molecular biology techniques did not have an impact on the publication of data that were small-scale, descriptive and univariate. Results were often referred to as "semiquantitative" and whilst having the appearance of being quantitative, they usually were based on highly variable, subjective interpretations of data images. The analysis of RNA by northern blotting, of proteins by western blotting or the comparison of mRNA levels by reverse transcription-PCR and subsequent gel analysis readily come to mind. However, the development of increasingly sensitive, specific and high throughput techniques such as real-time quantitative PCR (qPCR), microarrays, digital PCR (dPCR), next generation sequencing (NGS) and mass spectrometry has led to their use in a wide variety of applications in a broad range of biological and clinical subjects. This continuous expansion of molecular technologies has swiftly resulted in a demand for more rigorous quantification of nucleic acids, proteins and small molecules in, for example, molecular diagnostics [2]. However, compared with a qualitative yes/no result, reliable, reproducible and biologically relevant quantification poses significant problems in terms of sample preparation and quality control, assay design, optimisation and validation, collection of data, their analysis and, importantly, prudent and transparent reporting [3].

Quantitative experiments investigating physiology frequently involve the investigation of individual cellular processes either by

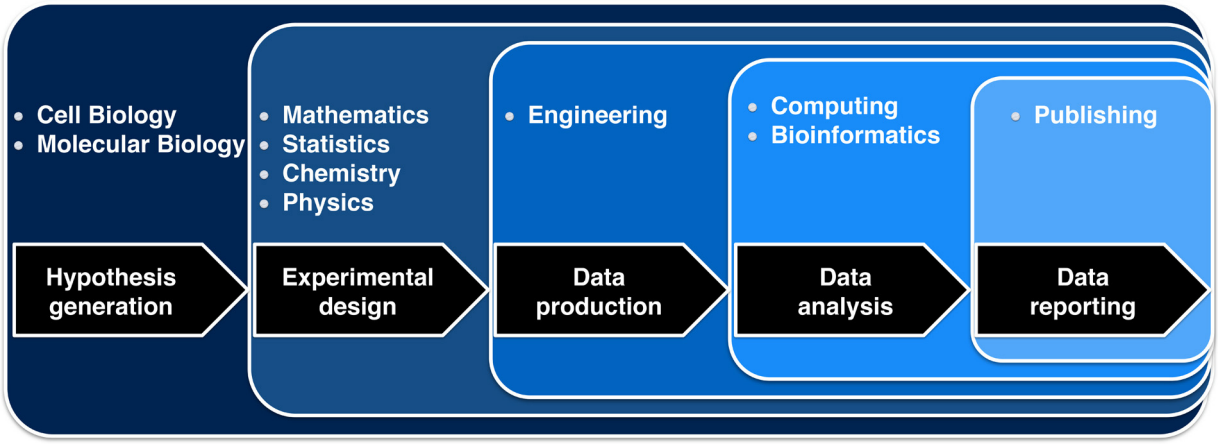

Fig. 1. Interpretation and publication of quantitative results requires interaction between a wide range of scientific disciplines. The diagram does not imply a linear relationship between the disciplines and every stage of the experimental process is benefiting from cross-discipline interaction.

http://dx.doi.org/10.1016/j.bdq.2014.09.001

studying the dynamic interplay of distinct cell types in various tissues or analysing a small number of single cells. In either case, error may be introduced, and this must be both captured and, where possible, reduced using appropriate experimental design and normalisation strategies. Ultimately, the generation of quantitative multivariate measurements enables the application of statistical tests and correlation analyses that allow scientists to report detailed information about their own experiments in a more prescribed and systematised manner.

Systems biology generates prodigious amounts of complex, integrated data from the fields of functional genomics, proteomics and metabolomics and has started to emerge as a driving force at the interface of biology, biochemistry, mathematics, computing engineering, and physics [4]. It has become accepted, albeit also widely ignored, that the complexity of a biological organism cannot be described by a static listing of even the most well-characterised components such as DNA sequence, genetic and epigenetic modifications, levels of coding and noncoding RNAs, proteins and their many isoforms or the range of metabolites being synthesised or entering individual cells. Instead, it is the temporal, special and dynamic arrays of interactions of each one of these constituents that determines cell responses and, ultimately, an organism's behaviour. Consequently, an understanding of complex biological systems requires the extensive support of computational tools [5] together with agreed standards to enhance information exchange, results compatibility and mutual understanding [6–8].

There are numerous issues relating to the entire quantification workflow that are difficult to standardise and automate. These include variables such as the wide variety of sample types that can be tested, the complex steps involved in sample preparation and, not least, a frequent switch to novel technical platforms leading to inconsistent results as well as complex post-testing analysis methods [9]. Sample preparation affects the reliability of PCR [10], the sensitivity and robustness of mass spectrometric analysis [11] and the accuracy of protein-DNA interactions mapped using chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) [12]. Similarly, ELISA measurements can be compromised by technical issues [13], interfering endogenous and exogenous effects [14] or measurement methodology [15], with significant consequences for diagnosis and or clinical decision-making. One consequence of these challenges is many omics assays are insufficiently sensitive or specific to meet clinical needs [16]. Worse, most biomedical research cannot be independently replicated [17], a finding consistent with an estimate that around 85% of research funding is wasted [18].

The need for more rigour has been recognised for a long time, with detailed discussions published regarding the reliability of results obtained from qPCR [19,20], dPCR [21–23] or microarray analyses [24,25]. These have resulted in a range of proposed solutions aimed at improving the quantitative nature of such technologies, for example recommendations such as the MIAME [26], MIQE [27,28] and dPCR MIQE [29] guidelines. Unfortunately there is a wide disregard for these recommendations, an attitude characterised by one "scientist's" comment that compliance with PCR guidelines was being "too dogmatic". It has also become apparent that unless proper guidance is accorded to NGS [30–33], this powerful technique will join all the others and also go powerfully wrong when not appropriately used or reported.

Discussion of the many methods in use today, as well as their advantages and disadvantages needs to be coupled to improvements and innovations that can only arise from the interplay between different expertise, ideas and communication. These must thrive in a milieu that adapts the scientific rigour of mathematics, statistics, chemistry and physics whilst retaining the versatile and dynamic characteristics of biological scientists.

Biomolecular Detection and Quantification (BDQ) is an open access, peer-reviewed international journal dedicated to championing excellence in molecular study design, measurement, data analysis and reporting. Its focus is on the application of qualitative and quantitative molecular methodologies to all areas of clinical and life sciences. The journal has two main aims:

- to provide a forum for discussion and recommendation of guidelines designed to improve the accuracy of molecular measurement, its data analysis and the transparency of its subsequent reporting;
- to publish molecular biology based studies that adhere to best practice guidelines, both current and future.

The deliberately broad scope of the journal covers clinical areas such as cancer, epigenetics, metagenomics, and infectious diseases as well non clinical subjects including environmental, microbiology and food science. BDQ revolves around the common theme of promoting excellence in molecular measurement and its data analysis. It will serve as a repository for sharing key findings across what may otherwise be disparate specialties.

References

[1] Cohen JE. Mathematics is biology's next microscope, only better; biology is mathematics' next physics, only better. PLoS Biol 2004;2:e439.
[2] Wallace PS, MacKay WG. Quality in the molecular microbiology laboratory. Methods Mol Biol 2013;943:49–79.
[3] Bustin S. Transparency of reporting in molecular diagnostics. Int J Mol Sci 2013;14:15878–84.
[4] Li P, Dada JO, Jameson D, Spasic I, Swainston N, et al. Systematic integration of experimental data and models in systems biology. BMC Bioinformatics 2010;11:582.
[5] Ghosh S, Matsuoka Y, Asai Y, Hsin KY, Kitano H. Software for systems biology: from tools to integrated platforms. Nat Rev Genet 2011;12:821–32.
[6] Brazma A, Krestyaninova M, Sarkans U. Standards for systems biology. Nat Rev Genet 2006;7:593–605.
[7] Jones AR, Miller M, Aebersold R, Apweiler R, Ball CA, et al. The Functional Genomics Experiment model (FuGE): an extensible framework for standards in functional genomics. Nat Biotechnol 2007;25:1127–33.
[8] Taylor CF, Field D, Sansone SA, Aerts J, Apweiler R, et al. Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. Nat Biotechnol 2008;26:889–96.
[9] Gomah ME, Turley JP, Lu H, Jones D. Modeling complex workflow in molecular diagnostics: design specifications of laboratory software for support of personalized medicine. J Mol Diagn 2010;12:51–7.
[10] Hedman J, Radstrom P. Overcoming inhibition in real-time diagnostic PCR. Methods Mol Biol 2013;943:17–48.
[11] Bylda C, Thiele R, Kobold U, Volmer DA. Recent advances in sample preparation techniques to overcome difficulties encountered during quantitative analysis of small molecules from biofluids using LC-MS/MS. Analyst 2014;139:2265–76.
[12] Bailey T, Krajewski P, Ladunga I, Lefebvre C, Li Q, et al. Practical guidelines for the comprehensive analysis of ChIP-seq data. PLoS Comput Biol 2013;9:e1003326.
[13] Noble JE, Wang L, Cerasoli E, Knight AE, Porter RA, et al. An international comparability study to determine the sources of uncertainty associated with a non-competitive sandwich fluorescent ELISA. Clin Chem Lab Med 2008;46:1033–45.
[14] Tate J, Ward G. Interferences in immunoassay. Clin Biochem Rev 2004;25:105–20.
[15] Klauenberg K, Ebert B, Voigt J, Walzel M, Noble JE, et al. Bayesian analysis of an international ELISA comparability study. Clin Chem Lab Med 2011;49:1459–68.
[16] Sung J, Wang Y, Chandrasekaran S, Witten DM, Price ND. Molecular signatures from omics data: from chaos to consensus. Biotechnol J 2012;7:946–57.
[17] Prinz F, Schlange T, Asadullah K. Believe it or not: how much can we rely on published data on potential drug targets? Nat Rev Drug Discov 2011;10:712.
[18] Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. Lancet 2009;374:86–9.
[19] Bustin SA, Benes V, Garson J, Hellemans J, Huggett J, et al. The need for transparency and good practices in the qPCR literature. Nat Methods 2013;10:1063–7.
[20] Bustin SA. Why the need for qPCR publication guidelines? The case for MIQE. Methods 2010;50:217–26.
[21] Sanders R, Mason DJ, Foy CA, Huggett JF. Considerations for accurate gene expression measurement by reverse transcription quantitative PCR when analysing clinical samples. Anal Bioanal Chem 2014.
[22] Nixon G, Garson JA, Grant P, Nastouli E, Foy CA, et al. Comparative study of sensitivity, linearity, and resistance to inhibition of digital and nondigital

polymerase chain reaction and loop mediated isothermal amplification assays for quantification of human cytomegalovirus. Anal Chem 2014;86:4387–94.
[23] Huggett JF, Whale A. Digital PCR as a novel technology and its potential implications for molecular diagnostics. Clin Chem 2013;59:1691–3.
[24] Miklos GL, Maleszka R. Microarray reality checks in the context of a complex disease. Nat Biotechnol 2004;22:615–21.
[25] Chagovetz A, Blair S. Real-time DNA microarrays: reality check. Biochem Soc Trans 2009;37:471–5.
[26] Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, et al. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet 2001;29:365–71.
[27] Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 2009;55:611–22.
[28] Bustin SA, Beaulieu JF, Huggett J, Jaggi R, Kibenge FS, et al. MIQE precis: practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments. BMC Mol Biol 2010;11:74.
[29] Huggett JF, Foy CA, Benes V, Emslie K, Garson JA, et al. The digital MIQE guidelines: minimum information for publication of quantitative digital PCR experiments. Clin Chem 2013;59:892–902.
[30] Bhargava V, Head SR, Ordoukhanian P, Mercola M, Subramaniam S. Technical variations in low-input RNA-seq methodologies. Sci Rep 2014;4:3678.
[31] O'Brien CP, Taylor SE, O'Leary JJ, Finn SP. Molecular testing in oncology: problems, pitfalls and progress. Lung Cancer 2014.
[32] DeWoody JA, Abts KC, Fahey AL, Ji Y, Kimble SJ, et al. Of contigs and quagmires: next-generation sequencing pitfalls associated with transcriptomic studies. Mol Ecol Resour 2013;13:551–8.
[33] Xuan J, Yu Y, Qing T, Guo L, Shi L. Next-generation sequencing in the clinic: promises and challenges. Cancer Lett 2013;340:284–95.

Jim Huggett*
Jim Huggett, LGC, Queens Road, Teddington, Middlesex TW11 0LY, UK

Justin O'Grady
Norwich Medical School, University of East Anglia, Colney Lane, Norwich NR4 7GJ, UK

Stephen Bustin
Postgraduate Medical Institute, Faculty of Medical Science, Anglia Ruskin University, Bishop Hall Lane, Chelmsford, Essex, UK

* Corresponding author.
E-mail addresses: jim.huggett@lgcgroup.com (J. Huggett), justin.OGrady@uea.ac.uk (J. O'Grady), Stephen.bustin@anglia.ac.uk (S. Bustin).

Available online 10 September 2014