Biomolecular Detection and Quantification 11 (2017) 1–3

Contents lists available at ScienceDirect

Biomolecular Detection and Quantification

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

Editorial

Reproducibility of biomedical research – The importance of editorial vigilance

A B S T R A C T

Many journal editors are a failing to implement their own authors' instructions, resulting in the publication of many articles that do not meet basic standards of transparency, employ unsuitable data analysis methods and report overly optimistic conclusions. This problem is particularly acute where quantitative measurements are made and results in the publication of papers that lack scientific rigor and contributes to the concerns with regard to the reproducibility of biomedical research. This hampers research areas such as biomarker identification, as reproducing all but the most striking changes is challenging and translation to patient care rare.

Concern at the robustness, reliability and relevance of much molecular biomedical research has been mounting for at least 20 years. Sporadic warnings about the lack of reliable quality control procedures in molecular diagnostics [1], reports of cross-contamination of tissue cultures [2] and evidence for between-study heterogeneity of genetic association studies for disease outcome [3] have been reinforced by numerous investigations that expose a lack of reproducibility of much basic and preclinical research and identify the accompanying huge waste of research funding (reviewed in [4]).

There are many reasons for this serious situation [5], but one of the most critical factors is the lack of transparency in the reporting of experimental detail [6,7], which makes it difficult for reviewers and readers alike to judge the validly of the data and conclusions presented in many papers. This problem, which is probably more one of omission that deliberate, has led to solutions being proposed on how to deal with this challenge; for example, there are suggestions for improved training of researchers, enhanced emphasis on appropriate study design, easier access to unpublished primary data and open public discussion of published research [8]. Undoubtedly there is a need to encourage changes in investigator attitude, both in terms of validating reagents and understanding that ambiguity is intrinsic to biological systems [9]. Furthermore, there needs to more of an appreciation that choice of methods for sample preparation and measurement can contribute substantially to technical variance and/or bias. This is rarely a consideration when quantifying nucleic acids, yet can be major reason for poor reproducibility.

There have also been moves towards urging the research community to implement guidelines for the appropriate design, execution and reporting of research methods. There are now a number of guidelines, starting with the MIAME guidelines providing minimal information about microarray experiments [10], minimum information specifications for *in situ* hybridisation and immunohistochemistry experiments (MISFISHIE) [11], through MIQE [6] and digital MIQE [12] guidelines for real time quantitative PCR (qPCR) and digital PCR (dPCR), respectively, to those dealing with standards for reporting *in vivo* animal research [13] and pathology data gathered from animal tissues [14].

Nevertheless, it is also clear that whilst the existence of guidelines is useful, they are only part of the solution [15]. For one, although adherence places additional burdens on the researcher in terms of additional cost and effort, failure to observe basic standards of transparency has few immediate consequences, especially with regards to publication. For example, whilst there have been discussions on the need to improve the publication validation process [16,17] and journal editors have acknowledged that they have failed in their responsibility to monitor the quality of technical information provided and promise to address these issues [16–27] there is considerable evidence that this talk has not necessarily been followed by much action [4]. This is perhaps not surprising in light of their earlier failure to ensure that authors comply with primary data availability required by their own author instructions [28].

This failure is perhaps best demonstrated using an illustrative example that is representative of the types of technical errors and omissions typically found in the biomedical literature. qPCR is probably the most widely used enabling technology in molecular research [29] and results obtained using reverse transcription (RT) of cellular RNA followed by simultaneous amplification and detection of PCR amplicons has been used in tens of thousands of papers to investigate changes in RNA expression levels [30]. It is also a good example of a technology that has been subjected to extensive scrutiny and found to be seriously deficient in design, reporting and analysis [31–41]. A detailed analysis of around 2000 peer-reviewed publications found that the vast majority of papers do not report sufficient technical detail and use inappropriate or misleading data analysis methods [42]. This has been confirmed by follow-up studies [4,15,43–46]. But perhaps a more tangeable

http://dx.doi.org/10.1016/j.bdq.2017.01.002

metric is that, despite the tens of thousands of research papers that have used RT-qPCR to quantify RNA expression over the last 20 years, the number of clinical tests developed measuring cellular RNAs to guide medical treatment can be counted in the ones.

In its 2013 announcement acknowledging that journals such as Nature contribute to "failures in the reliability and reproducibility of published research", the editors admit that journals "compound" them by failing to "exert sufficient scrutiny" and not publishing "enough information for other researchers to assess results properly" [18]. Yet they do not appear to be implementing resolutions to rectify failings. A paper published in a top nature journal in the last two months demonstrates a recent example of this failure. The manuscript, which we will not name, lacks the most basic information required to assess the validity of the experimental protocols, contains critical, yet basic errors that should have been detected during the review process, and makes conclusions based on results that are probably flawed.

1 The qPCR protocol published in the supplementary information is limited to six lines and provides no information about RNA integrity, quality or quantity used or the conditions of the reverse transcription or qPCR efficiency. This is despite a wealth of published evidence that highlights the importance of RNA quality [47,48], the variability of the RT step [49–51] and the importance of appropriate PCR conditions [52–55] for achieving reliable results.
2 Data analysis was performed using a less accurate approach, as the chosen method [56] requires an invalid assumption of 100% PCR amplification efficiency across all assays and directly compares the difference in Cqs of different primer sets (a number that will vary with threshold settings if, as is often the case, the assays have different shaped amplification plots). That approach was superseded by a method that only compares the same assay between different samples and factors in the PCR efficiency to provide a far more accurate estimation of the differences [57,58].

The authors use a single, unvalidated reference gene for normalisation, an approach shown to be inappropriate as far back as 2002 [59]. How do the authors know the reference gene is not effected by the study?

The reference gene in question, β-actin, is not an ideal reference gene in the model investigated in this study as it was reported to have "a tendency to be down regulated" in the same types of tumour samples [60].

3 An analysis of the primer sequences published in the supplementary information using PrimerBlast finds no target templates for the one of the primer pairs. A BLAST search finds no 100% sequence identity of another forward primer with any target, whereas the reverse primer has 100% sequence identity with a different, unrelated human mRNA; the study used a mouse model. Another reverse primer has a single mismatch with all published target sequences. Yet another primer pair targeting a specific transcript splice variant has sufficient sequence identity with all other variants to suggest it is not specific at all and could also amplify the other variants.
4 While some of these observations may be less critical if the reported expression changes differed by orders of magnitude, conclusions are drawn on the published qPCR based on 3–5-fold differences that are well within the variability range established for the reverse transcription step [51].

Whilst such differences are measureable using RT-qPCR, with the points highlighted above, doubt is introduced to the conclusions associated with this technique and, arguably far worse, with the lack of detail presented in the manuscript, making it almost impossible to reproduce the study to corroborate the findings. Yet it will not take the reader long, should they be so inclined, to find numerous published studies, where these omissions apply, frequently purporting to find differences of less than two-fold, a feat that, without the most rigorous selection of multiple reference genes, is simply impossible to achieve with any degree of accuracy.

These parameters are elemental for qPCR and it is extraordinary that such lacks of adherence have been allowed to slip through the vetting process. It is truly astonishing that this is but one of hundreds, if not thousands of publications utilising qPCR that are published with similar shortcomings. Biomolecular Detection and Quantification (BDQ) was conceived as a journal committed to publishing molecular studies that adhere to best practice in the areas of study design, measurement, data analysis and reporting. As editors, we handle numerous manuscript and a continuous theme is the initial absence of comprehensive information on aspects of the work being reported. This is corrected in papers published by this journal, but clearly is not done so by the editors of many other journals. It is unfortunate that the editors of Nature, and many other journals, continue to publish papers that have not been adequately reviewed and where the information provided does not follow Nature's own recommendations. Is it a surprise that we continue to face a crisis of reproducibility and that so many papers are published that report results that are meaningless?

References

[1] R. Jung, P. Ahmad-Nejad, M. Wimmer, M. Gerhard, C. Wagener, M. Neumaier, Quality management and influential factors for the detection of single metastatic cancer cells by reverse transcriptase polymerase chain reaction, Eur. J. Clin. Chem. Clin. Biochem. 35 (1997) 3–10.
[2] O. Markovic, N. Markovic, Cell cross-contamination in cell cultures: the silent and neglected danger, In Vitro Cell. Dev. Biol. Anim. 34 (1998) 1–8.
[3] J.P. Ioannidis, E.E. Ntzani, T.A. Trikalinos, Contopoulos-Ioannidis DG. Replication validity of genetic association studies, Nat. Genet. 29 (2001) 306–309.
[4] S.A. Bustin, The reproducibility of biomedical research: sleepers awake!, Biomol. Detect. Quantif. 2 (2014) 35–42.
[5] C.G. Begley, J.P. Ioannidis, Reproducibility in science: improving the standard for basic and preclinical research, Circ. Res. 116 (2015) 116–126.
[6] S.A. Bustin, V. Benes, J.A. Garson, et al., The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments, Clin. Chem. 55 (2009) 611–622.
[7] J. Carp, The secret lives of experiments: methods reporting in the fMRI literature, Neuroimage 63 (2012) 289–300.
[8] F.S. Collins, L.A. Tabak, Policy: NIH plans to enhance reproducibility, Nature 505 (2014) 612–613.
[9] K. Mullane, M. Williams, Unknown unknowns in biomedical research: does an inability to deal with ambiguity contribute to issues of irreproducibility, Biochem. Pharmacol. 97 (2015) 133–136.
[10] A. Brazma, P. Hingamp, J. Quackenbush, et al., Minimum information about a microarray experiment (MIAME)-toward standards for microarray data, Nat. Genet. 29 (2001) 365–371.
[11] E.W. Deutsch, C.A. Ball, J.J. Berman, et al., Minimum information specification for in situ hybridization and immunohistochemistry experiments (MISFISHIE), Nat. Biotechnol. 26 (2008) 305–312.
[12] J.F. Huggett, C.A. Foy, V. Benes, et al., The digital MIQE guidelines: minimum information for publication of quantitative digital PCR experiments, Clin. Chem. 59 (2013) 892–902.
[13] C. Kilkenny, W. Browne, I.C. Cuthill, M. Emerson, D.G. Altman, NC3Rs RGWG. Animal research: reporting in vivo experiments: the ARRIVE guidelines, J. Gene Med. 12 (2010) 561–563.
[14] C.L. Scudamore, E.J. Soilleux, N.A. Karp, et al., Recommendations for minimum information for publication of experimental pathology data: MINPEPA guidelines, J. Pathol. 238 (2016) 359–367.
[15] S.A. Bustin, T. Nolan, Improving the reliability of peer-reviewed publications: we are all in it together, Biomol. Detect. Quantif. 7 (2016) A1–5.
[16] Further confirmation needed, Nat. Biotechnol. 30 (9) (2012) 806 [editorial].
[17] Must try harder, Nature 483 (7391) (2012) 509 [editorial].
[18] Announcement: reducing our irreproducibility, Nature 496 (2013) 398.
[19] Enhancing reproducibility, Nat. Methods 10 (2013) 367.
[20] Raising reporting standards, Nature Cell Biol. 15 (2013) 443.
[21] Raising standards, Nat. Biotechnol. 31 (2013) 366.
[22] Raising standards, Nat. Med. 19 (2013) 508.
[23] Raising standards, Nat. Struct. Mol. Biol. 20 (2013) 533.
[24] Raising standards, Nat. Genet. 45 (2013) 467.
[25] Raising standards, Nat. Neurosci. 16 (2013) 517.

[26] Raising standards, Nat. Immunol. 14 (2013) 415.
[27] Journals unite for reproducibility, Nature 515 (2014) 7.
[28] A.A. Alsheikh-Ali, W. Qureshi, M.H. Al-Mallah, J.P. Ioannidis, Public availability of published research data in high-impact journals, PLoS One 6 (2011) e24357.
[29] H.D. VanGuilder, K.E. Vrana, W.M. Freeman, Twenty-five years of quantitative PCR for gene expression analysis, Biotechniques 44 (2008) 619–626.
[30] S.A. Bustin, Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays, J. Mol. Endocrinol. 25 (2000) 169–193.
[31] J. Murphy, S.A. Bustin, Reliability of real-time reverse-transcription PCR in clinical diagnostics: gold standard or substandard, Expert Rev. Mol. Diagn. 9 (2009) 187–197.
[32] S.A. Bustin, Real-time quantitative PCR-opportunities and pitfalls, Eur. Pharm. Rev. 4 (2008) 18–23.
[33] S.A. Bustin, Real-time polymerase chain reaction -towards a more reliable, accurate and relevant assay, Eur. Pharm. Rev. 6 (2008) 19–27.
[34] S. Bustin, Molecular medicine, gene-expression profiling and molecular diagnostics: putting the cart before the horse, Biomark. Med. 2 (2008) 201–207.
[35] T. Nolan, R.E. Hands, S.A. Bustin, Quantification of mRNA using real-time RT-PCR, Nat. Protoc. 1 (2006) 1559–1582.
[36] S.A. Bustin, R. Mueller, Real-time reverse transcription PCR (qRT-PCR) and its potential use in clinical diagnosis, Clin. Sci. (Lond.) 109 (2005) 365–379.
[37] S.A. Bustin, V. Benes, T. Nolan, M.W. Pfaffl, Quantitative real-time RT-PCR–a perspective, J. Mol. Endocrinol. 34 (2005) 597–601.
[38] S.A. Bustin, T. Nolan, Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction, J. Biomol. Tech. 15 (2004) 155–166.
[39] S.A. Bustin, Meaningful quantification of mRNA using real-time PCR, in: T. Griffin, H.G. Weissensteiner, A. Griffin (Eds.), PCR Technology Current Innovations, CRC Press, Boca Raton, FL, 2004, pp. 225–233.
[40] S.A. Bustin, T. Nolan, Basic RT-PCR considerations, in: S.A. Bustin (Ed.), A–Z of quantitative PCR, IUL Press, La Jolla, CA, 2004, pp. 359–396.
[41] S.A. Bustin, Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems, J. Mol. Endocrinol. 29 (2002) 23–39.
[42] S.A. Bustin, V. Benes, J. Garson, et al., The need for transparency and good practices in the qPCR literature, Nat. Methods 10 (2013) 1063–1067.
[43] A.M. Abdel Nour, E. Azhar, G. Damanhouri, S.A. Bustin, Five years MIQE guidelines: the case of the Arabian countries, PLoS One 9 (2014) e88266.
[44] J.R. Dijkstra, L.C. van Kempen, I.D. Nagtegaal, S.A. Bustin, Critical appraisal of quantitative PCR results in colorectal cancer research: can we rely on published qPCR results? Mol. Oncol. 8 (2014) 813–818.
[45] S.C. Taylor, E.M. Mrkusich, The state of RT-quantitative PCR: firsthand observations of implementation of minimum information for the publication of quantitative real-time PCR experiments (MIQE), J. Mol. Microbiol. Biotechnol. 24 (2014) 46–52.
[46] S. Alemayehu, K.C. Feghali, J. Cowden, J. Komisar, C.F. Ockenhouse, E. Kamau, Comparative evaluation of published real-time PCR assays for the detection of malaria following MIQE guidelines, Malar. J. 12 (2013) 277.
[47] J. Vermeulen, K. De Preter, S. Lefever, et al., Measurable impact of RNA quality on gene expression results from quantitative PCR, Nucleic Acids Res. 39 (2011) e63.
[48] S. Fleige, M.W. Pfaffl, RNA integrity and the effect on the real-time qRT-PCR performance, Mol. Aspects Med. 27 (2006) 126–139.
[49] A. Stahlberg, J. Hakansson, X. Xian, H. Semb, M. Kubista, Properties of the reverse transcription reaction in mRNA quantification, Clin. Chem. 50 (2004) 509–515.
[50] A. Stahlberg, M. Kubista, M. Pfaffl, Comparison of reverse transcriptases in gene expression analysis, Clin. Chem. 50 (2004) 1678–1680.
[51] S. Bustin, H.S. Dhillon, S. Kirvell, Variability of the reverse transcription step: practical implications, Clin. Chem. 61 (2015) 202–212.
[52] A. Tichopad, M. Dilger, G. Schwarz, M.W. Pfaffl, Standardized determination of real-time PCR efficiency from a single reaction set-up, Nucleic Acids Res. 31 (2003) e122.
[53] T. Bar, A. Stahlberg, A. Muszta, M. Kubista, Kinetic outlier detection (KOD) in real-time PCR, Nucleic Acids Res. 31 (2003) e105.
[54] C. Ramakers, J.M. Ruijter, R.H. Deprez, A.F. Moorman, Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data, Neurosci. Lett. 339 (2003) 62–66.
[55] W. Liu, D.A. Saint, Validation of a quantitative method for real time PCR kinetics, Biochem. Biophys. Res. Commun. 294 (2002) 347–353.
[56] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method, Methods 25 (2001) 402–408.
[57] M.W. Pfaffl, G.W. Horgan, L. Dempfle, Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR, Nucleic Acids Res. 30 (2002) e36.
[58] X. Rao, X. Huang, Z. Zhou, X. Lin, An improvement of the 2(-delta delta CT) method for quantitative real-time polymerase chain reaction data analysis, Biostat. Bioinf. Biomath. 3 (2013) 71–85.
[59] J. Vandesompele, K. De Preter, F. Pattyn, et al., Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes, Genome Biol. 3 (0034) (2002) (1-0034.11).
[60] S. Dupasquier, A.S. Delmarcelle, E. Marbaix, J.P. Cosyns, P.J. Courtoy, C.E. Pierreux, Validation of housekeeping gene and impact on normalized gene expression in clear cell renal cell carcinoma: critical reassessment of YBX3/ZONAB/CSDA expression, BMC Mol. Biol. 15 (2014) 9.

Stephen A. Bustin*
Faculty of Medical Science , Anglia Ruskin University, Chelmsford, CM1 1SQ, United Kingdom

Jim F. Huggett
Analytical Microbiology, School of Bioscience and Medicine, Faculty of Health and Medical Science, University of Surrey, Guildford, GU2 7XH, United Kingdom

* Corresponding author.
Handled by Jim Huggett
E-mail address: Jim.Huggett@lgcgroup.com
(J.F. Huggett)

Available online 21 February 2017