Mini Review

Open Access

CrossMark

RT PCR data plagiarism their ignorance and impact

Abstract

Polymerase Chain Reaction (PCR) is a technique based on the ability of DNA polymerase to synthesize a new strand of DNA which is complementary to the template strand offered and amplifying this strand to billions of amplicons. There are different types of PCR reactions for different types of experiments but reverse transcription-polymerase chain reaction (RT-PCR) is most commonly used PCR reactions in the field of genomics and proteomics basically at mRNA level. A wide range of procedures have been devised for studying PCR products like gel electrophoresis, cloning and sequencing of PCR products, etc. The PCR products from gel electrophoresis are in the form of images, therefore image processing techniques are commonly used to analyze these images in three main steps i.e. band detection, band matching, and data quantification. Image processing sometimes alters the intensity of image which can at points improve the visualization of the data. But if that image alteration mischaracterizes the data, one has gone too far and this is known as plagiarism. Therefore, it is important to safeguard and protect the original, unaltered image in order to avoid the accusations of misconduct which will stand or fall on the basis of whether or not the original image is available to be compared with its altered copies. For example, investigators whose work fall under the food and drug administration's (FDA's) "Final Rule on Electronic Records and Electronic Signatures" must maintain the integrity of the original image. Similarly, industries whose work products are used in forensic activities or in health insurance portability and accountability act (HIPAA) related aspects of health care might be required to maintain an original image. Therefore, the plagiarism of RT-PCR data has tremendous impact on research if left ignored as seen through the above examples and therefore should be taken care.

Keywords: RT-PCR, image processing, plagiarism, polymerase chain reaction, DNA polymerase

Kanika Gupta, Ashok Kumar
Department of Biology and Bioinformatics, Panjab University, India

Correspondence: Ashok Kumar, Department of Biology and Bioinformatics, Panjab University, India, Tel 9216270057, Email ashokkumar@pu.ac.in

Abbreviations: PCR, polymerase chain reaction; RT-PCR, reverse transcription-polymerase chain reaction; FDA's, food and drug administration's; HIPAA, health insurance portability and accountability act

Introduction

PCR (Polymerase Chain Reaction) is a method developed by Kary Mullis.[1] This technique is based on using the ability of DNA polymerase to synthesize new strand of DNA, complementary to the offered template strand, as DNA polymerase can add a nucleotide only onto a pre-existing 3'-OH group and it needs a primer to which it can add the first nucleotide. This requirement makes it possible to delineate a specific region of template sequence that the researcher wants to amplify.[2] At the end of the PCR reaction, the specific sequence will be accumulated in billions of amplicons.[3] There are different types of PCR reactions for different experiments but RT-PCR is commonly used now a day for the new discoveries in the field of genomics and proteomics at mRNA level.[4] RT-PCR (reverse transcription-polymerase chain reaction) is the most sensitive technique for mRNA detection and quantitation which is currently available.[5] Compared to the two other commonly used techniques for quantifying mRNA levels, Northern blot analysis and RNase protection assay, RT-PCR can be used to quantify mRNA levels from much smaller samples. In fact, this technique is sensitive enough to enable quantitation of RNA from a single cell.[6] When the quantification of RNA in both relative and absolute terms is done then the technique is known as quantitative PCR or qRT-PCR.[7] This technique is considered to be the most powerful, sensitive, and quantitative assay for the detection of RNA levels i.e. more than RT-PCR. It is frequently used in the expression analysis of single or multiple genes, and expression patterns for identifying infections and diseases[8] (42). PCR is often the starting point for a longer series of experiments in which the amplification product is studied in various ways in order to gain information about the DNA molecule that acted as the original template. Many studies of this type could be encountered, where the applications of gene cloning and PCR in research and biotechnology are examined[9] (38). Although a wide range of procedures have been devised for studying PCR products, three techniques are particularly important i.e. Gel electrophoresis of PCR products, Cloning of PCR products, and Sequencing of PCR products, etc[10] (39). PCR experiments are checked by running a portion of the amplified reaction mixture in an agarose gel. A band representing the amplified DNA may be visible after ethidium bromide staining, or if the DNA yield is low the product can be detected by Southern hybridization. Accurate interpretation of DNA banding patterns from electrophoretic images can be laborious and error prone when a large number of bands are interrogated manually.[11]

Data visualization

The amount of substance in each band is estimated by calculating the area of the band, and the molecular weight of each band is estimated by considering the position relative to a predefined reference band.[12] Therefore, obtaining accurate genetic information from gel images depends on several parameters, including the quality of the bands

isolated. Image processing techniques commonly used to analyze gel electrophoresis images require three main steps: band detection, band matching, and quantification.[13]

Image integrity

Images submitted with a manuscript for review should be minimally processed (for instance, to add arrows to a micrograph). Authors should retain their unprocessed data and metadata files, as editors may request them to aid in manuscript evaluation.[14] If unprocessed data are unavailable, manuscript evaluation may be stalled until the issue is resolved.[15] All digitized images submitted with the final revision of the manuscript must be of high quality and have resolutions of at least 300d.p.i. (dots per inch) for colour, 600d.p.i. for greyscale and 1,200d.p.i. for line art. A certain degree of image processing is acceptable for publication (and for some experiments, fields and techniques is unavoidable), but the final image must correctly represent the original data and conform to community standards.[16] The guidelines will aid in accurate data presentation at the image processing level; Authors must also take care to exercise prudence during data acquisition, where misrepresentation must equally be avoided.[17] Authors should list all image acquisition tools and image processing software packages used. Authors should document key image gathering settings and processing manipulations in the methods.[18] Images gathered at different times or from different locations should not be combined into a single image, unless it is stated that the resultant image is a product of time averaged data or a time lapse sequence. If juxtaposing images is essential, the borders should be clearly demarcated in the figure and described in the legend.[19] The use of touch up tools, such as cloning and healing tools in Photoshop, or any feature that deliberately obscures manipulations, is to be avoided.[20] Processing (such as changing brightness and contrast) is appropriate only when it is applied equally across the entire image and is applied equally to controls. Contrast should not be adjusted so that data disappear. Excessive manipulations, such as processing to emphasize one region in the image at the expense of others (for example, through the use of a biased choice of threshold settings), is inappropriate, as is emphasizing experimental data relative to the control. When submitting revised final figures upon conditional acceptance, authors may be asked to submit original, unprocessed images.[21]

Software for quantification RT PCR

Several software systems have been developed to analyze the electrophoresis gel images automatically like Image J and MATLAB. Some of these systems are semiautomatic and perform band detection by segmenting the image into lanes and locating the peaks of the 1D mean profiles or of the cumulative row difference profile of each lane. However, these methods have major disadvantages because they require the user to select the region of interest and adjust different parameters manually.[22] Other software systems identify bands by extracting the variance and mean variance of the 1D mean profile of the lane and classifying the valleys of the profiles as either noise or bands. Nevertheless, these methods cannot generally locate faint bands, and they sometimes detect false bands due to noise,[13] Proposed system involves four steps:

a. Lane separation: Consists of separating the images into lanes.[13]

b. Lane segmentation: Consists of applying an appropriate automatic thresholding technique in order to separate the bands from the noisy background of the lanes.[13]

c. Band detection: Consists of automatically detecting the location of each band in the lane.[13]

d. Data quantification consists of computing the amount of the substance in each band and its molecular weight.[9]

Obviously, digital images are data upon whose accuracy the scientific community depends. Just as the data that appear in the tables of lab reports can be misrepresented or fabricated, so can digital imagery data.[23] A primary but not always realized source for misrepresenting digital imagery data consists in the fact that each individual element of the image, called a pixel, has a numerical value reflecting RGB (red/green/blue) intensity.[24] Image processing that alters that intensity can improve the visualization of the data. But if that image alteration mischaracterizes the data, one has gone too far.[25] The resolution of dilemmas over whether or not an investigator has gone too far in manipulating an image is very simple. The investigator should make an unaltered, raw image of the data and retain it, preferably in the original file format.[26] This image is never altered or enhanced; only its copies are. If the investigator honestly believes that an altered version of the original image is preferable for publication, he or she should attach both a copy of the original and the altered image to the manuscript being submitted along with a detailed description of why and how the copied image was altered. The journal editors can then decide which image to publish.[27] If the altered image is chosen, the nature of the alteration should be described in the figure legend and explained in the methodology section of the paper.[28] That way, both the investigator and the journal will maintain transparency so that no accusations of deception or misrepresentation will stand.[29] In a nutshell, that's how to resolve this problem, when it is provoked by honest consternation of whether or not some kind of image manipulation is allowable.[30] The remainder of this opinion will discuss various technical details and considerations associated with digital image manipulation. The importance of safeguarding and protecting the unaltered, original image because accusations of misconduct will stand or fall based on whether or not the original is available to compare with its copies.[31] Indeed, investigators whose work falls under the Food and Drug Administration's (FDA's) "Final Rule on Electronic Records and Electronic Signatures" must maintain the integrity of the original image.[32,33] Similarly, industries whose work products are used in forensic activities or in HIPAA (Health Insurance Portability and Accountability Act)-related aspects of health care might be required to maintain an original image.[33–35] Scientists suggest that adjustments to the original image that are usually acceptable are small adjustments in brightness and contrast or reasonable adjustments of the levels and gamma settings. Although cropping an image is usually acceptable, accusations of unethical cropping will occur when the cropping distorts the image, e.g., cropping so as to omit something that contradicts the investigator's hypothesis.[36] They also make the following recommendations:

i. Digital images that are to be compared to one another should be acquired under identical processing considerations. If they are not, the reason should be explained in the publication or in the figure legend.

ii. Enhancing a specific area of an image is extremely questionable. If performed, the selective enhancement(s) must be identified and explained. Similarly, because they can create artifacts in an image, the use of software filters that can lead to misinterpretation are questionable.[37]

iii. Copying objects from one part of an image to another part is

extremely problematic while "the use of cloning techniques to create objects in an image that did not exist there originally (e.g., creating a new gel band) is completely unethical".[38]

iv. JPEG compression reduces the file size but it also changes the XY resolution of the image and the intensity value of any given pixel. This kind of compression should be avoided and replaced by a TIF file format.

v. Care should be exercised when changing the pixel size of a digital image. Decreasing image size will decrease the image's resolution. Increasing image size can cause the software to "guess" at how many pixels need to be created between the existing ones. If the total number of pixels in an image is going to change because of a manipulation, it should be done only once to limit the number of artifacts that might be introduced.[39]

But blot doctoring continues to occur with regularity.[40,41] Photoshop is a powerful image analysis tool, with the help of this tool someone could alter the contrast, rub out extraneous bands or background noise, or present the same bands to represent multiple proteins/mRNAs. There are other tools like GNU Image Manipulation Program (GIMP) which can also be considered for carrying out manipulations in images. GIMP, a free and open source image editing software, is often considered by some to be an alternative for Adobe Photoshop. This is not true. GIMP has never projected itself as a Photoshop alternative, and comparing free software to a mammoth and costly one is, obviously, unfair.[42] These changes are difficult to detect by naked eyes.[43]

Acknowledgements

None.

Conflict of interest

The author declares no conflict of interest.

References

1. Shampo MA, Kyle RA, Kary B. Mullis—Nobel Laureate for procedure to replicate DNA. *Mayo Clin Proc*. 2002;77(7):606.

2. Saiki RK, Gelfand DH, Stoffel S, et al. Primer–directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*. 1988;239(4839):487–491.

3. Cheng S, Fockler C, Barnes WM, et al. Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc Natl Acad Sci U S A*. 1994;91(12):5695–5699.

4. Freeman WM, Walker SJ, Vrana KE. Quantitative RT–PCR: pitfalls and potential. *Biotechniques*. 1999;26(1):112–122.

5. Spackman E, Senne DA, Myers TJ, et al. Development of a real–time reverse transcriptase PCR assay for type A influenza virus and the avian H5 and H7 hemagglutinin subtypes. *J Clin Microbiol*. 2002;40(9):3256–3260.

6. Sharkey FH, Banat IM, Marchant R. Detection and quantification of gene expression in environmental bacteriology. *Appl Environ Microbiol*. 2004;70(7):3795–3806.

7. Varkonyi–Gasic E, Hellens RP. qRT–PCR of Small RNAs. *Methods in Molecular Biology*. 2010;631:109–122.

8. http://groups.molbiosci.northwestern.edu/

9. Brown TA. *Gene Cloning and DNA Analysis: An Introduction*. 7th ed. USA: Wiley–Blackwell; 2016. 374 p.

10. http://www.yourarticlelibrary.com/biology/gene-cloning/4-types-of-polymerised-chain-reaction-pcr-products-used-in-original-dna-molecule/32553/

11. Zimm BH, Levene SD. Problems and prospects in the theory of gel electrophoresis of DNA. *Q Rev Biophys*. 1992;25(2):171–204.

12. Intarapanich A, Kaewkamnerd S, Shaw PJ, et al. Automatic DNA Diagnosis for 1D Gel Electrophoresis Images using Bio–image Processing Technique. *BMC Genomics*. 2015;16(Suppl 12):S15.

13. Neil Yen. *Advances in Computational Biology edited*. Springer; 2010.

14. Freedland KE, Carney RM. Data management and accountability in behavioral and biomedical research. *Am Psychol*. 1992;47(5):640–645.

15. Zlotnik Shaul R, Reid L, Essue B, et al. Dissemination to research subjects: operationalizing investigator accountability. *Account Res*. 2005;12(1):1–16.

16. West R. Data and statistical commands should be routinely disclosed in order to promote greater transparency and accountability in clinical and behavioral research. *J Clin Epidemiol*. 2016;70:254–255.

17. Solberg LI, Mosser G, McDonald S. The three faces of performance measurement: improvement, accountability, and research. *The Joint Commission journal on quality improvement*. 1997;23(3):135–147.

18. Krimsky S. Introduction to special issue of Accountability in Research on conflict of interest in science. *Account Res*. 2004;11(2):79–81.

19. Costa DM, Canady B, Babcock JC. Preliminary report on the Accountability Scale: a change and outcome measure for intimate partner violence research. *Violence Vict*. 2007;22(5):515–531.

20. O'Reilly M, Dixon–Woods M, Angell E, et al. Doing accountability: a discourse analysis of research ethics committee letters. *Sociol Health Illn*. 2009;31(2):246–261.

21. Martin B. Suppressing research data: methods, context, accountability, and responses. *Account Res*. 1999;6(4):333–372.

22. Moreira B, Sousa A, Mendonca AM, et al. Automatic lane segmentation in TLC images using the continuous wavelet transform. *Comput Math Methods Med*. 2013;2013:218415.

23. http://www.atlantactsi.org/documents/ethical-dilemmas/image-manipulation.pdf

24. John Krueger SP, Linda Warren. Image Manipulation. *Science Editor*. 2007;30(6):185.

25. Beautification and fraud. *Nature cell biology*. 2006;8(2):101–102.

26. Image rights and wrongs. *Nature nanotechnology*. 2010;5(9):627.

27. Savulescu J, Chalmers I, Blunt J. Are research ethics committees behaving unethically? Some suggestions for improving performance and accountability. *BMJ*. 1996;313(7069):1390–1393.

28. Joob B, Wiwanitkit V. Research ethics committee: dominance and accountability. *Account Res*. 2015;22(5):301–302.

29. Raven M, Jureidini J, Stuart GW. Transparency and accountability in early psychosis intervention research. *Aust N Z J Psychiatry*. 2013;47(7):680–681.

30. Khorasani R. Too much information! What we need from IT beyond image manipulation tools. *J Am Coll Radiol*. 2007;4(12):925–926.

31. Masic I. Plagiarism in scientific research and publications and how to prevent it. *Mater Sociomed*. 2014;26(2):141–146.

32. Wright DS, Grunbaum JA, Dawkins N, et al. Ensuring accountability in public health prevention research: evaluating the Prevention Research Centers Program at the Centers for Disease Control and Prevention. *J Public Health Manag Pract*. 2011;17(2):E1–E9.

33. McGuire S. International Food Policy Research Institute. Global Nutrition Report 2014: actions and accountability to accelerate the world's progress on nutrition. *Advances in nutrition*. 2014;6(3):278–279.

34. Tereskerz PM. Research accountability and financial conflicts of interest in industry–sponsored clinical research: a review. *Account Res*. 2003;10(3):137–158.

35. Wendler D. One–time general consent for research on biological samples: is it compatible with the health insurance portability and accountability act? *Arch Intern Med*. 2006;166(14):1449–1452.

36. Willerson JT, Kereiakes DJ. Clinical research and future improvement in clinical care: the Health Insurance Portability and Accountability Act (HIPAA) and future difficulties but optimism for the way forward. *Circulation*. 2003;108(8):919–920.

37. Mann H, Shamoo AE. Introduction to special issue of Accountability in Research on the review and approval of biomedical research proposals: a call for a centralized national human research protections system. *Account Res*. 2006;13(1):1–9.

38. Masic I. Twenty years of society of medical informatics of b&h and the journal acta informatica medica. *Acta Inform Med*. 2012;20(1):5–14.

39. Rubenstein LS. Standards of accountability for consent in research. *Account Res*. 1996;4(3–4):197–206.

40. Science under attack. *Nature*. 2006;439(7079):891.

41. Gel slicing and dicing: a recipe for disaster. *Nat Cell Biol*. 2004;6(4):275.

42. https://git.gnome.org//browse

43. White C. Software makes it easier for journals to spot image manipulation. *BMJ*. 2007;334(7594):607.