COMMENTARY

A question of faith - or the philosophy of RNA controls

R Ivell

Institute for Hormone and Fertility Research, University of Hamburg, Grandweg 64, 22529 Hamburg, Germany

As an editor, referee, and mostly as an author having to contend with other referees, the problem of controls in determinations of specific mRNA, whether measured by northern-, or *in situ* hybridization, or by some other technique, is perpetual. I know of few other issues in science which appear to encourage more dogmatic outbursts. In what follows, I should like to take a definitely personal viewpoint on this issue of what constitutes a proper control, and in which experimental circumstances each type of control is appropriate.

Let us consider first the situation of northern hybridizations. Here, we must carefully distinguish a variety of different circumstances. First, are we dealing with total RNA or poly(A)-enriched RNA? If the former, then at least 95% of this will comprise the 18S and 28S ribosomal RNA molecules. If all we want to do is to check for even loading of a gel, then a visual estimate of the ribosomal RNA bands by ethidium bromide staining is quite sufficient. Particularly, the 28S RNA is a good visual indicator, since it appears to be more readily degraded than the 18S band. Because all ribosomes have one molecule of each of these structural RNA molecules, under optimal circumstances the 28S RNA band should be present at about double the fluorescence intensity of the 18S band. The main reason why ethidium bromide staining is sufficient is also due to the fact that this reagent is specific for double-stranded nucleic acid, and thus preferentially stains ribosomal RNA which contains a high proportion of double-stranded RNA. When there has been RNase-dependent degradation of RNA, this tends to occur at the single-stranded regions of the ribosomal RNA, leaving the double-stranded regions as discrete and smaller bands often visible between the 28S and 18S bands after ethidium bromide staining. If these smaller bands appear, then there is a strong likelihood that the messenger RNA, which is mostly single-stranded, is also degraded. Therefore, if all I want to do is to control for even loading of intact total RNA onto an agarose gel, then a picture of the 28S and 18S ribosomal RNA bands, and the intermediate region, stained by ethidium bromide, is enough.

This controls for quantity and degradation. However, it does not control for even transfer to the nylon or nitrocellulose membrane. To control for this, one additionally

requires a hybridization step. For the same reasons as outlined above, a probe specific for 28S RNA is probably a good and simple choice. Alternatively, one could choose a messenger RNA for a common householding gene. The best of these are again related to the protein synthetic machinery, because they are very frequent transcripts, and hence represent a relatively large proportion of the total RNA, and tend to be expressed stoichiometrically with the ribosomal RNA. Examples include cDNA probes for ribosomal proteins or for the elongation factor 1 α .

Alternative, so-called householding genes are more problematic. The assumption for such genes is that they are constitutively expressed and relatively common. One of the most frequently used is the gene for glyceraldehyde-3-phosphate dehydrogenase (GAPDH). However, there are a number of difficulties encountered using this mRNA as an internal control. First, it is not very abundant, and hence being a small proportion of the total RNA its levels can change markedly simply due to the relative changes in expression of other more common genes. For this reason, the most commonly encountered problem is that it varies considerably in frequency from one tissue type to another, or even within a single tissue or cell type from an unstimulated to a stimulated condition. This is not due to a direct effect on the GAPDH gene, but is simply a reflection of the amount of its mRNA relative to the expression of other genes. Other householding genes can suffer similar fates, particularly if not expressed at high frequency. Because they are somewhat more common, the cell cycle gene cyclophilin has often been used, as also the gene for ubiquitin, or that for the structural protein, β -actin. These are probably better than GAPDH, because more frequent, but also these can vary markedly in proportion from one tissue to another.

So far, we have been discussing the issue of controls assuming that constant loading and transfer of total RNA is the ideal we wish to achieve. This is not strictly true. The implications we all want to make in performing northern hybridizations is that a particular gene is being up- or down-regulated under different stimulatory regimes, in different tissues, etc. If this is what we really wanted to assess, then the only correct experiment would be to load the total RNA from a given number of cells, or

cell nuclei, onto a gel. The reference point should therefore be the genomic DNA (i.e. how many copies of the genome have given rise to the RNA I am loading and analysing?). For cell culture, this is feasible, but is virtually never performed. The reason that we can use a constant amount of total RNA for northern hybridization is simply because in general the number of ribosomes (i.e. the protein synthetic machinery) reflects the number of cell nuclei for most cells, and does not vary greatly from one tissue type to another, though it has been noted that brain cells exhibit less protein synthesis than an active gland like the pituitary. Secondly, for most of the 'regulatable' genes we are interested in, the levels of specific mRNA relative to the total RNA are a reasonable, though indirect measure of their transcription. This is not true for so-called constitutive genes, such as GAPDH. There might be true up-regulation of transcription of such genes under particular circumstances, but because a large number of different genes are similarly up-regulated, this would not be evident on a northern blot where a fixed amount of RNA has been loaded. Thus, under certain circumstances, for example, stimulation of cells differentiating in culture, it might be more appropriate not to load constant amounts of total RNA, but to load the RNA equivalent to a fixed number of cell nuclei. The results of such northern blots would then be more comparable to the results obtained from run-on *in vitro* transcription experiments.

For a number of gene products, sufficient sensitivity can be obtained using total RNA. However, for less frequently expressed gene transcripts it may be necessary to enrich the RNA for poly(A)-containing messenger RNA. Depending upon the method used, this may be more or less efficient. So, for example, affinity chromatography using a column containing oligo(dT)-cellulose usually leads to an enrichment of about 10-fold. Given that total RNA comprises between 2 and 5% mRNA, such an enrichment would still leave a large proportion of ribosomal RNA in the enriched fraction. Because of this, some scientists have repeated the enrichment step, to remove more of the ribosomal RNA. More effective is the use of oligo(dT)-conjugated beads (latex or magnetic). The yields with this method appear to be closer to the ideal of about 50-fold enrichment. In terms of controls, there are several problems here. First, yields are often too low to be able to apply reliable UV-spectrophotometric assays. The proportion of ribosomal RNA can be very variable, depending upon the efficiency of the oligo(dT) affinity chromatography and practical issues such as the speed of the column elution. Measuring by using an intercalating dye such as ethidium bromide is for these reasons very unreliable, whereas a non-intercalating dye such as SYBR Green (Molecular Probes, Inc., Eugene, OR, USA) may be better. One of the most reliable quantification procedures is to use a fixed amount of oligo(dT)-conjugated beads with the total RNA in excess, and then to apply all the RNA eluted from the beads to the northern gel. In our

hands this leads to very even loading of mRNA. In terms of controls for loading and transfer, there is only the possibility of hybridizing against a control mRNA which is known to be polyadenylated. Ribosomal protein genes are suitable, but because of the enrichment, other constitutively expressed genes may be more or less appropriate, depending upon the biology of the experiment. An older recommendation suggested to use radiolabelled oligo(dT) or poly(U) itself as a control hybridization probe. However, since stimulation protocols may also influence the poly(A) tail length for a given transcript, this method would appear to be further away from our ideal of assessing the transcription from a given number of cell nuclei.

We should not forget the issue of the nature of the probes themselves, and the abomination known as 'normalization'. In spite of the many ways that probes can be prepared and labelled, either by nick translation, random priming, or by 'tailing' using terminal transferase, either radioactive or non-radioactive, either as DNA or RNA, or oligonucleotides, they can all be separated into either double-stranded or single-stranded probes. Double-stranded probes rely upon the creation of multiple networks of complementary single-stranded molecules, such that for a given target RNA, there may be many probe molecules from mutually complementary strands linked to it. This amplifies the signal. Although the method is quantitative, larger amounts of target RNA yielding stronger signals, the method is not constantly stoichiometric, and the signal-to-target ratio is non-linear. This is not true for single-stranded probes (cRNA or oligonucleotides) complementary to the target RNA. Here there is a simple stoichiometry and a linear signal-to-target relationship. However, this relationship is disrupted if radioactive probes are visualized by autoradiography, and subsequently quantified by densitometry. Autoradiography film has a limited capacity to convert emulsion into black silver grains; strong signals lead to saturation of the photographic emulsion and to a non-linear signal-to-target relationship. Linearity only occurs for weak signals (i.e. relatively short autoradiographic exposure times). Only when single-stranded probes are used together with direct signal detection (e.g. by phosphorimager) is true stoichiometric quantification possible. For non-radioactive single-stranded probes, quantification is possible if all steps towards visualization are stoichiometric and reagents are in excess and not exhausted. In general, this issue is not important if loading and transfer of RNA samples has been even, and we are only interested in a relative up- or down-regulation of a given gene transcript. What is not legitimate is to 'normalize' the signals from the desired transcript where loading has not been even, by dividing signal strength from a densitometrically scanned autoradiogram by the signal from a 'control' gene product. As outlined above, except under closely defined circumstances, there is only a non-linear signal-to-target relationship, which prohibits calculations of this sort. For

publication purposes, where loading is unfortunately uneven for whatever reason, one has to show the actual results and let the reader (or the referee) draw his or her own conclusions. It is far better, however, to repeat the experiment and ensure even loading. The purpose of the control is to check for even loading and transfer, not to correct for uneven loading and transfer.

Finally, while on the subject of northern hybridization and the use of double-stranded versus single-stranded probes, it is important to be sure that the specific bands that are visualized on the northern hybridization are really what you think they are. Besides the more obvious issues of cross-hybridization of probes to related or homologous gene products, or even to ribosomal RNA (because of insufficiently stringent washing), there are a couple of situations which warrant especial care. The first relates to the observation that in some tissues a genomic sequence may be transcribed in a different way from another tissue where the gene has been fully characterized. We encountered this situation for the vasopressin gene transcripts in the rat testis which yielded a band on northern hybridization of approximately the same size as in the principal site of hormone synthesis in the hypothalamus. More detailed analysis, however, showed that all these transcripts were missing the normal first exon; instead a novel sequence from further upstream was present which yielded a transcript of the same apparent size. This novel RNA did not give rise to an open reading frame, and was non-functional (Foo *et al.* 1991). Many testicular transcripts of genes characterized from other tissues appear to represent similarly apparently functionless transcripts (reviewed in Ivell 1992). Thus, an important caveat to performing northern hybridizations, particularly in cells and tissues where the gene has not been fully characterized, is to be sure that the transcripts visualized really do represent the gene product you expect.

A second related phenomenon, also reported initially from the testis as well as other tissues, is that some genes may be transcribed in both directions, creating both sense and anti-sense RNA. The best example described to date is the gene for basic fibroblast growth factor (FGF), particularly in the testis (Knee *et al.* 1994). Whereas the testis expresses a moderate amount of normal sense transcripts encoding the bFGF peptide, there is also transcription from a downstream promoter on the opposite strand, to yield an antisense RNA. This RNA is polyadenylated and is expressed in relatively large amounts at a size approximately similar to the conventional bFGF transcript (Knee *et al.* 1994). It also appears to have an open-reading frame, though it is not known whether any protein product is made *in vivo*, or what function this anti-sense transcript may have. The important point to note here is that this product is not detectable using single-strand probes for the conventional bFGF transcript, but both sense and anti-sense products are detected and confused using double-stranded probes.

Much of what has been discussed above is equally important for the detection of RNA transcripts by *in situ* hybridization. Here, the issue of controls and quantification is also problematic. Aside from the central issue of being sure that what one is 'seeing' with the probe truly is the functional transcript one imagines, what controls are possible and desirable? Conventionally, most scientists use single-stranded anti-sense probes (either cRNA or oligo-nucleotide), and it is usual to apply a similar probe but from the sense strand as control. The assumption is that the sense strand will have the same GC ratio and GC distribution as the probe, not encode any sensible product, and thus check for non-specific binding of the nucleic acid. This is only true, however, if the linearization of the plasmid used for *in vitro* transcription is truly complementary. Some probes are prepared by cleavage at an internal restriction site such that the sense and anti-sense transcription products are not complementary. In such cases, the so-called sense control could be effectively replaced by some random sequence. Also critical here, is that the probes are similar in length and are labelled to identical specific activity. This is often difficult to check, especially for non-radioactive probes, and may not be true if using only a single modified nucleotide and this nucleotide is not equally distributed on both strands. A more significant problem is reported for *in situ* hybridization in the testis. It has been shown that a specific GC-rich sequence within the multiple cloning site of the pBluescript plasmid often used for the preparation of cRNA is able to bind with high quasi-specificity to the residual bodies within the seminiferous tubules (Millar *et al.* 1994). This effect is dependent upon the promoter used by the RNA polymerase, and hence the cloning orientation of the insert into the plasmid. If this GC-rich sequence is colinear with the anti-sense probe, then the consequence is to yield a false result which conforms to all the criteria of a true positive, with a clearly negative sense-strand control.

Another issue relates to quantitation. Whereas radio-labelling of single-strand cRNA probes clearly provides a linear and stoichiometric signal-to-target relationship, the morphology provided by this technique is not ideal. Better microscopic visualization can be obtained using non-radioactively labelled probes. However, we and others have observed a clear 'threshold phenomenon' with this method. It would appear that in general non-radioactive probes are less sensitive than radiolabelled probes, and when a positive signal is present, this is amplified by the subsequent immunohistochemical steps to such a degree that a quantitative signal-to-target relationship is not observed. The result is that if a signal can be detected by this method, it is represented already as a relatively strong signal, implying the presence of more target RNA than in fact may be the case.

Ultimately, with *in situ* hybridization as with conventional northern hybridization, controls are only as good as the biological knowledge we have about the system in the

first place. Ideally, all such techniques should be confirmed by some quite independent methodology. For example, the structure of transcripts visualized in northern blots could be confirmed by RT-PCR and sequencing from the same tissue. *In situ* hybridizations could be performed using probes derived from different exons of the same gene, or be confirmed by immunohistochemistry using a specific antibody. At the end of the day, we need to have a description of the biological situation which involves several different techniques, which are mutually confirmatory in what they say, so that from all the experiments and controls the experimenter and the reader of an article can be convinced, and believe what is being discussed. No single control is perfect, and the issue becomes very much one of 'belief' only, if we cannot provide that independent supporting evidence. The problem now is that all the readers of this commentary will start checking back on my own papers and realise that I have myself not always adhered to these recommendations. But then science is very much a learning process. At the same time it is a question of believing, until alternative evidence or controls destroy that conviction. In discussing here the issue of what is a suitable control, I hope to have reduced that zealously dogmatic approach which many of us encounter on trying to publish our results (?beliefs). There is no perfect control, but together several pieces of independent evidence can convince.

Acknowledgements

I am very grateful to all my colleagues and friends with whom I have often discussed these issues at length, but particularly to Dr Philippa Saunders who has shared with me much of the same exasperation, and who critically read this manuscript.

References

- Foo NC, Carter D, Murphy D & Ivell R 1991 Vasopressin and oxytocin gene expression in rat testis. *Endocrinology* **128** 2118–2128.
- Ivell R 1992 All that glitters is not gold – common testis gene transcripts are not always what they seem. *International Journal of Andrology* **15** 85–92.
- Knee RS, Pitcher SE & Murphy PR 1994 Basic fibroblast growth factor sense (FGF) and antisense (GFG) RNA transcripts are expressed in unfertilized human oocytes and in differentiated adult tissues. *Biochemical and Biophysical Research Communications* **205** 577–583.
- Millar MR, Sharpe RM, Maguire SM, Gaughan J, West AP & Saunders PTK 1994 Localization of mRNAs by *in situ* hybridization to the residual body at stages IX–X of the cycle of the rat seminiferous epithelium: fact or artefact? *International Journal of Andrology* **17** 149–160.

Received 6 July 1998

Accepted 13 July 1998