

Common practice in molecular biology may introduce statistical bias and misleading biological interpretation

Jean-François Hocquette^{a,*}, Anna M. Brandstetter^b

^aUnité de Recherches sur les Herbivores, Equipe Croissance et Métabolismes du Muscle, Theix, 63122 Saint-Genès-Champanelle, France

^bKrannert Institute of Cardiology, Indiana University School of Medicine, 1800 North Capitol Avenue, Indianapolis, IN 46202 USA

Received 23 April 2001; received in revised form 5 October 2001; accepted 5 November 2001

Abstract

In studies on enzyme activity or gene expression at the protein level, data are usually analyzed by using a standard curve after subtracting blank values. In most cases and for most techniques (spectrophotometric assays, ELISA), this approach satisfies the basic principles of linearity and specificity. In our experience, this might be also the case for Western-blot analysis. By contrast, mRNA data are usually presented as arbitrary units of the ratio of a target RNA over levels of a control RNA species. We here demonstrate by simple experiments and various examples that this data-normalization procedure may result in misleading conclusions. Common molecular biology techniques have never been carefully tested according to the basic principles of validation of quantitative techniques. We thus prefer a regression-based approach for quantifying mRNA levels relatively to a control RNA species by Northern-blot, semi-quantitative RT-PCR or similar techniques. This type of techniques is also characterized by a lower reproducibility for repeated assays when compared to biochemical analyses. Therefore, we also recommend to design experiments, which allow the detection of a similar range of variance by biochemical and molecular biology techniques. Otherwise, spurious conclusions may be provided regarding the control level of gene expression. © 2002 Elsevier Science Inc. All rights reserved.

Keywords: Statistics; Molecular biology; Reproducibility; Power; Standard curves; Data normalization

1. Introduction

Tissue metabolic and biochemical characteristics have been widely investigated during the last decades [1]. This type of approach has provided significant information about normal or aberrant function of the studied tissues. It is now well known that tissue characteristics depend on the expression of key genes involved in many biological processes such as cell differentiation, tissue metabolism and also on their regulation by hormonal, genetic and nutritional factors [1]. Therefore, gene expression and its regulation are, nowadays, subjects of active investigation. For instance, quantification of mRNA levels has become routinely practice in biological research laboratories. The typical objective is the comparison of mRNA levels of key genes in target tissues between two (or more) groups of subjects (humans, animals) differing in age, nutrition, genetic type or any other physiological factor. Samples of tissues are taken and total RNAs are extracted. Then,

mRNA levels of key genes are quantified and expressed in various units, for instance in molecules or in relative arbitrary units per mg total RNA or per g tissue wet weight taking yields of total extractable RNA into account [2–8].

Scientists frequently face technical difficulties (e.g. a low sensitivity, a poor specificity of the molecular probes, etc) and carefully choose the appropriate molecular techniques depending on the mRNA species they are studying. However, much less effort is usually given for data analysis, especially for the choice of the statistical procedures. Thus, our objective was to compare commonly used procedures of data analysis and their impact on the interpretation of the results. Misleading conclusions can also be drawn from data sets bearing heterogeneous variances within subgroups. We found that the data-normalization procedure was critical to get proper results.

2. Material and methods

2.1. Materials

A polyclonal antibody against rat GLUT4 (glucose transporter 4) was supplied by East-Acres Biological (South-

* Corresponding author. Tel.: +33-4-73-62-42-53; fax: +33-4-73-62-46-39.

E-mail address: hocquet@clermont.inra.fr (J.-F. Hocquette).

bridge, MA). A polyclonal antibody against bovine FABP (fatty acid binding protein) was a gift from Professor J.H. Veerkamp (Nijmegen, The Netherlands). Polystyrene 96-well flat-bottom microtiter enzyme-linked immunosorbent assay (ELISA) plates were obtained from Polylabo (Strasbourg, France). Horseradish peroxidase-conjugated anti rabbit IgG was from Amersham International (Bucks, England). The chemicals for protein electrophoresis and immunoblotting were supplied by Bio-Rad (Munich, Germany), except membranes (Immobilon-P) which were obtained from Millipore (Bedford, MA). The enhanced chemiluminescence (ECL) Western blotting kit and Hyperfilm MP were supplied by Amersham International. The chemicals for RNA extraction and DNA labelling were purchased from Appligène (Illkirch, France) except membranes (Genescreen) which were obtained from Du Pont-NEN (Boston, MA) and radiolabelled nucleotides ($[\alpha\text{-}^{32}\text{P}]\text{dCTP}$ and $[\gamma\text{-}^{32}\text{P}]\text{ATP}$) which were supplied by Amersham International. Other biochemical reagents were from Sigma (St Louis, MO).

2.2. Animals and experimental designs

For the study of lipoprotein lipase (LPL) activity, GLUT4 protein content and mRNA levels of LPL and GLUT4, we used *rectus abdominis* and *longissimus thoracis* muscles of fourteen Montbéliard calves slaughtered at 170 days of age [3–5].

For the study of FABP protein content, we worked on *semimembranosus proprius* muscle samples from 32 rabbits slaughtered at 10 and 20 weeks of age.

For the study of FABP mRNA levels, we used thirty Montbéliard steers (castrated at 2 months of age) and 30 bulls each allotted to different regimens and slaughter ages as previously described [8]. Four castrated and five intact animals were slaughtered at 4 months of age. Bulls and steers were divided into *ad libitum* (AL) and restricted (R) feeding pens at 9 months of age. The restriction diet contained optimal protein allowances, but energy intake was decreased in order to have 40% inferior average daily gain (ADG) as compared to the AL control group. At the end of this restriction period, four or five animals per treatment were sacrificed at 12 months of age. Then all the remaining animals were given *ad libitum* access to feed and slaughtered at 16 months of age. This allowed the R animals to exhibit compensatory growth. Tissue samples from *semitendinosus* (ST), *triceps brachii* (TB) and *splenius* (SP) muscles were taken immediately after slaughter, quick-frozen in liquid nitrogen and stored at -80°C until subsequent analysis.

2.3. ELISA analysis

To quantitatively measure the heart type FABP (H-FABP) protein, cytosolic protein extracts were prepared as previously described [9,10]. At a concentration of 1 mg/L in

buffer B (50 mmol/L sodium carbonate, pH 9.6), 50 μL aliquots were coated onto the wells of microtiter plates by overnight incubation at 4°C . For calibration, amounts of 0–2 ng of human H-FABP were coated in the same way. The wells were washed and 100 μL of diluted antiserum against H-FABP (1:1000 in buffer C: 5.4 mmol/L sodium phosphate, 1.3 mmol/L potassium phosphate, 150 mmol/L NaCl pH 7.4, 0.5 g/L Tween-20) were added. After incubation for 1 h at 20°C , the wells were washed and 100 μL of diluted horseradish peroxidase-conjugated swine-anti rabbit IgG (1:500 in buffer C) were added and incubated for 1 h at 20°C . The wells were washed and bound peroxidase was assessed with 100 μL of a freshly prepared solution of 2.2 mmol/L *o*-phenylene. The incubation was in the dark for 30 min at 20°C , and was stopped by addition of 50 μL 12.5% H_2SO_4 . The product of the peroxidase reaction was determined at 492 nm using an ELISA plate reader (Lab-systems Multiscan, Helsinki, Finland). The amount of immuno-reactive protein (μg H-FABP/g tissue wet weight) was determined from calibration curves calculated by linear regression analysis [9,10].

2.4. Western-blot analysis

Crude membrane extracts from muscle tissues were prepared by differential centrifugation and analyzed by Western blot as previously described [11]. Membrane proteins were solubilized in Laemmli buffer and size-fractionated on non-denaturing 12% acrylamide gels. After transfer of the proteins to Immobilon-P membranes, they were successively reacted with the polyclonal antibody against rat GLUT4 (final dilution 1:250–1:500) and a horseradish peroxidase-labelled secondary antibody. Detection was with the enhanced chemiluminescence method (ECL Western blotting kit). Membranes were exposed to Hyperfilm MP, and GLUT4 was quantified by scanning densitometry of the autoradiograms (Hoeffer, San Francisco, CA). Results are expressed in arbitrary densitometric units [11].

2.5. Quantification of mRNA levels by Northern-blot analysis or competitive RT-PCR

Extraction of total RNAs and Northern-blot analyses were performed as previously described for the quantification of GLUT4 [2,3], LPL or FABP mRNA levels [4] or for the determination of 18S rRNA levels [8]. Total RNAs were size-fractionated by electrophoresis (1% agarose/2.2 M formaldehyde) and electrophoretically transferred to Genescreen membranes. Prehybridization and hybridizations were performed at 42°C . The membranes were exposed to Hyperfilm MP for several days or to phosphorimager screens overnight, which were scanned (Storm Imager) and analyzed using the Image-Quant software (Molecular Dynamics, Sunnyvale, CA). Results are expressed in arbitrary densitometric units.

Androgen receptor mRNA levels were quantified by

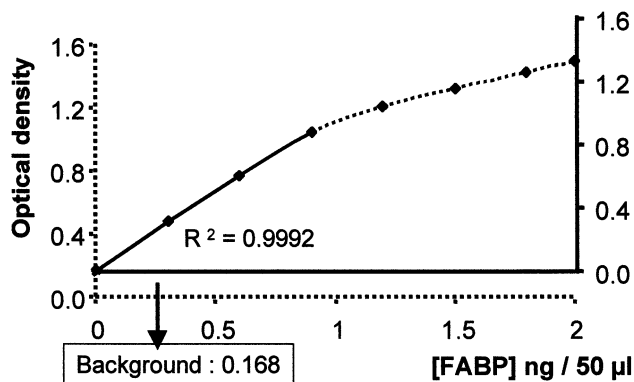


Fig. 1. Standard curve of FABP protein content determined by ELISA.

competitive co-amplification of native target mRNA and an internal androgen receptor standard RNA template using RT-PCR [8]. The PCR products were size-fractionated by electrophoresis and visualized by ethidium bromide staining. Gel images were captured under UV transillumination, and the intensity of the bands was quantified by scanning densitometry using the NIH Image 2.0 program (NIH, Bethesda, MD). Results are expressed in molecules and then in arbitrary units when related to 18S rRNA values [8].

2.6. Statistical analysis

Significant differences for one parameter between two groups of data were evaluated using Student’s *t* tests. ANOVA of the data and regression-based approaches were done using the General Linear Procedure of Statistical Analysis Systems (SAS, 1987). For androgen receptor expression data, the effects tested in the model included muscle type, age, sex, nutrition and all possible two-way interac-

tions between fixed effects. Muscle was considered repeated on the same animal, which was itself nested within age, sex and nutrition. In some analyses, 18S rRNA levels were introduced as a covariable.

3. Results

3.1. Normalization of protein content data determined by ELISA

Classically, data obtained from an ELISA procedure (i.e. optical density values) are normalized as follows: the blank (the optical density value with the antibodies but without any sample) is subtracted from all the other data. Consequently, when increasing amounts of a known standard are quantified, a close relationship exists between concentrations of the standard and optical density values. Furthermore, no sample in the well gives a zero value after correction for the blank. In other words, this relationship is convincing in terms of specificity because it has a y-intercept equal to zero by construction (Fig. 1). This relationship is often linear, because at least the lowest values always fall within the linear range of the calibration curve. Therefore, sample values are required to lie within this linear range of the curve. This premise was met in our experiment, meaning that if for a given sample the optical density is two times higher than for another sample reading, the FABP concentration in this particular sample is also two times higher (Fig. 1).

3.2. Normalization of protein content data determined by Western blot analysis

Theoretically, the determination of the protein content by Western blot should follow the basic procedures of data

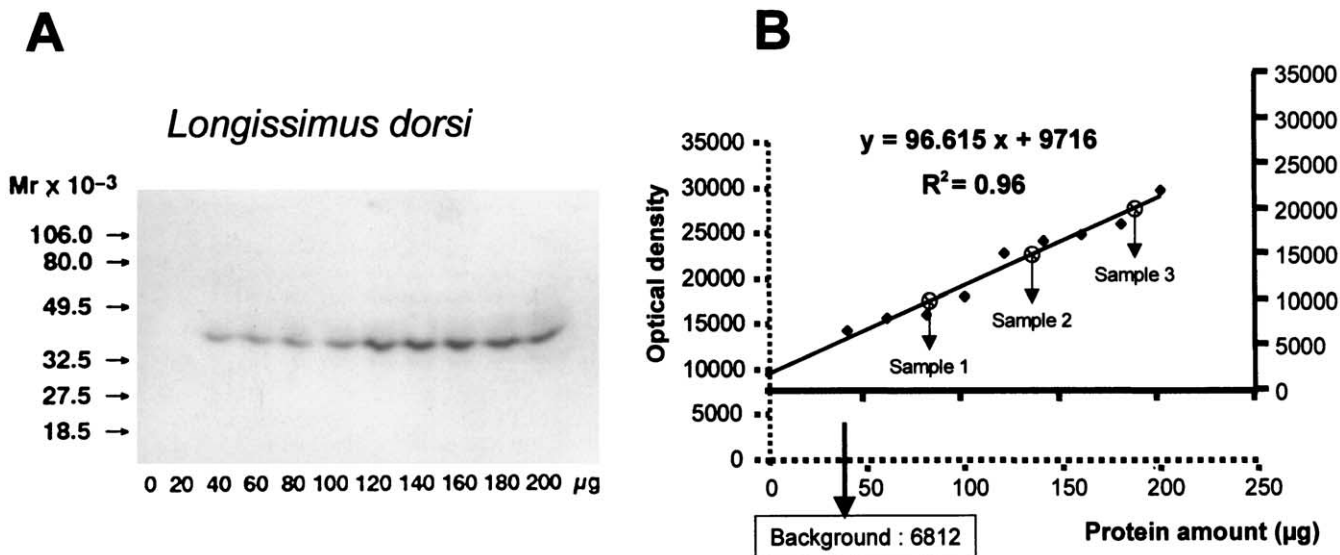


Fig. 2. Standard curve of GLUT4 protein content determined by Western blot. A. Effect of increasing amounts of protein loaded on the gel. B. GLUT4 protein levels plotted against amounts of total protein loaded on the gel. Three theoretical samples not represented on the autoradiograms are indicated.

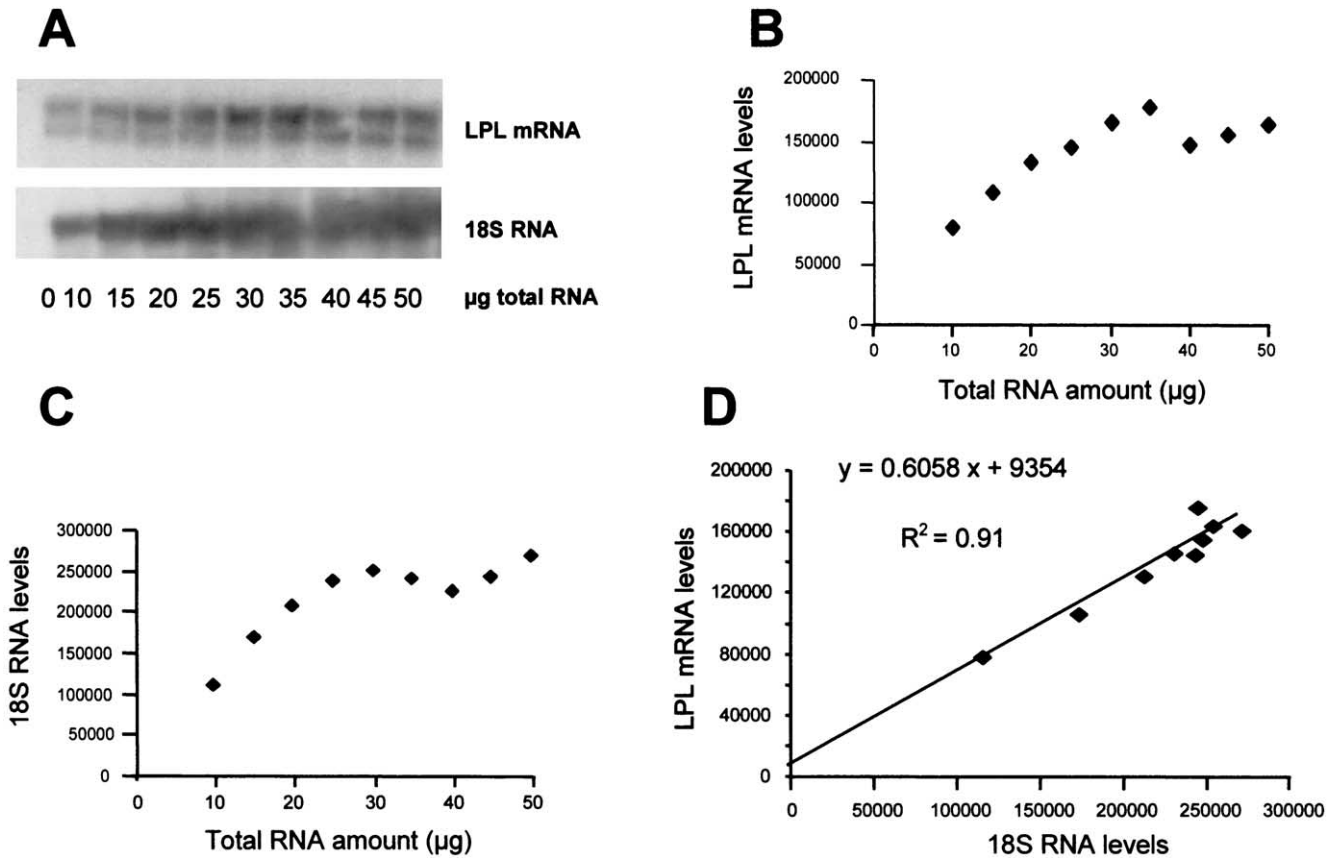


Fig. 3. Standard curve of LPL mRNA levels determined by Northern blot analysis. A. Effect of increasing amounts of total RNA loaded on the gel. B. LPL mRNA levels plotted against amounts of total RNA loaded on the gel. C. 18S RNA levels plotted against amounts of total RNA loaded on the gel. D. LPL mRNA levels (y) plotted against 18S RNA levels (x).

normalization as previously described for ELISA, i.e., the subtraction of the blank, and calculation in the linear range of the standard curve. In our experience, the GLUT4 protein content is linearly related to the amount of total protein loaded on the gel. Consequently, the ordinate at the origin is almost equal to zero (Fig. 2B). This means that the criteria of linearity (highly significant coefficient of correlation) and specificity (blank equal to zero) are achieved. Similar results were previously observed for GLUT4 in goat *masseter* muscles [12].

However, the square of the coefficient of regression was 0.96 in our example (Fig. 2B), which is lower than for ELISA (0.99 as in Fig. 1). In other words, 4% of the variance may be explained by technical problems of reproducibility within the same Western blot analysis (Fig. 2A).

In addition, the blank value directly determined by image analysis (6812 arbitrary densitometric units) is not exactly equal to the value of the intercept of the regression line with the ordinate (9716 arbitrary densitometric units) in our example (Fig. 2B). Therefore, an error of at most 2904 units would be introduced by this approach. This error will have a significant impact on the 1.5-fold difference between two theoretical samples of 10,000 and 15,000 arbitrary densitometric units, because $17,904/12,904 = 1.39$, not 1.5. However, for a similar absolute difference between two theoret-

ical samples of higher absolute values such as 20,000 and 15,000 (Fig. 2B), the impact will not be so important because $20,000/15,000 = 1.33$ and $22,904/17,904 = 1.28$.

In conclusion, the basic criteria of linearity, specificity and reproducibility are almost achieved for Western blot analysis although the results are less satisfactory compared to ELISA analysis. The error will be minimized if relatively high amounts of total protein or if samples that abundantly express the protein of interest are used.

3.3. Normalization of mRNA data determined by Northern blot analysis

When we loaded increasing amounts of total RNA from 10 to 50 µg of the same sample (Fig. 3A), increasing LPL mRNA levels were observed from 10 to 35 µg of total RNA loaded (Fig. 3B). To verify that increasing amounts of total RNA were transferred onto the membrane, we hybridized an oligonucleotide probe against the 18S ribosomal RNA. Similar results were obtained for LPL mRNA and 18S rRNA levels (Fig. 3C). This means that with large amounts of total RNA, an uncontrolled transfer did occur, due perhaps, to saturation of the membrane. The plot of LPL mRNA levels against 18S rRNA levels, which is a common normalization practice in Northern blot analysis [2–7], emphasizes this

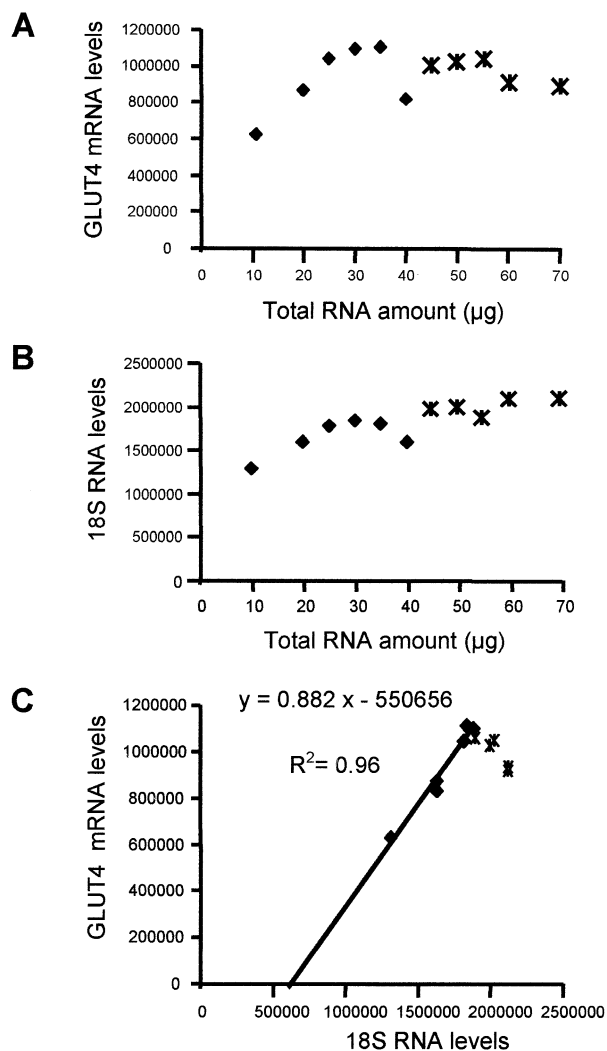


Fig. 4. Standard curve of GLUT4 mRNA levels determined by Northern blot analysis. A. GLUT4 mRNA levels plotted against amounts of total RNA loaded on the gel. B. 18S rRNA levels plotted against amounts of total RNA loaded on the gel. C. GLUT4 mRNA levels (y) plotted against 18S rRNA levels (x).

close relationship between the two RNA species and allows us to correct the results (Fig. 3D). However, the square of the coefficient of regression was 0.91, which means that 9% of the variance may be explained by technical problems of reproducibility within the same membrane (Fig. 3D).

Indeed, the same amount of sample (30 µg of total RNA extracted from adult cow heart) was analyzed in an 8-fold repetition on the same Northern membrane. The coefficient of variation was equal to 34%, 14% and 28% for LPL mRNA levels, 18S rRNA levels and the ratio of LPL mRNA to 18S rRNA levels, respectively (results not shown). This demonstrates the low reproducibility for RNA analysis as for the low abundant LPL mRNA in comparison with 18S rRNA. Expression level and repeatability are directly related: the lower the expression level, the more difficult it is to reproduce the measurement.

In another experiment, we quantified 18S rRNA and

GLUT4 mRNA levels of a given sample in a dilution series from 10 to 70 µg of total RNA loaded (Fig. 4). At low RNA concentrations up to 35 µg, similar results as previously for LPL mRNA levels were obtained (Figs. 4A and 4B). However, beyond 45 µg of total RNA, 18S rRNA values were dissociated from GLUT4 mRNA levels which makes the results invalid (Fig. 4C). Indeed, the square of the coefficient of regression was equal to 0.43 when all the points of Fig. 4C were taken into account, but equal to 0.96 when points between 10 and 40 µg were used. Finally, the relationship between GLUT4 mRNA levels and 18S rRNA levels for the valid points was characterized by a y-intercept that is different from zero (Fig. 4C). This means that for the lowest 18S values (550,656/0.882 arbitrary densitometric units), GLUT4 mRNA was undetectable. This can be explained by the very low expression of GLUT4 mRNA as compared to the 18S rRNA and adds the problem of assay sensitivity.

Potentially, Northern blot analyses are thus prone to methodological errors, which are much greater than those in Western blot analyses: (i) the absence of linearity above 40–45 µg of loaded total RNA even after 18S rRNA correction, (ii) an intercept of the slope not equal to zero when the mRNA of interest is expressed at extremely low levels compared to 18S rRNA. Data normalization by a regression-based approach is a valid option to avoid any statistical bias induced by these technical problems [13].

3.4. Normalization of mRNA data by different statistical approaches

In the line of the previous example for GLUT4, the ratio-method of data normalization leads to improper results compared to a regression-based approach. For instance, FABP mRNA levels were compared by Northern blot analysis in SP muscles of bulls and steers at two ages (12 and 16 months). The quality of the RNA was assessed by 18S rRNA hybridization. The statistical analysis was performed with three different approaches: (1) when crude FABP mRNA levels were analyzed, a significant difference between the two ages was detected (Fig. 5A), (2) when FABP mRNA levels were first divided by 18S rRNA levels, no significant difference was detected between the two ages (Fig. 5B), (3) When 18S rRNA levels were introduced as a co-variable in the statistical model, a significant difference was observed between the two ages (Fig. 5C). In another set of data, androgen receptor mRNA levels were compared by quantification of PCR products (Fig. 6A) in three muscles of the same animals at three ages (4, 12 and 16 months of age). Again, the ratio-method (Fig. 6B) gives different results than the regression method (Fig. 6C). By definition, in a co-variance analysis, a measured value of the co-variable is adjusted to the mean of its subgroup, before its contribution to the overall variance of the mRNA of interest is determined. Assuming a relatively constant expression of the control RNA species, it is adjusted to an average, unlike a

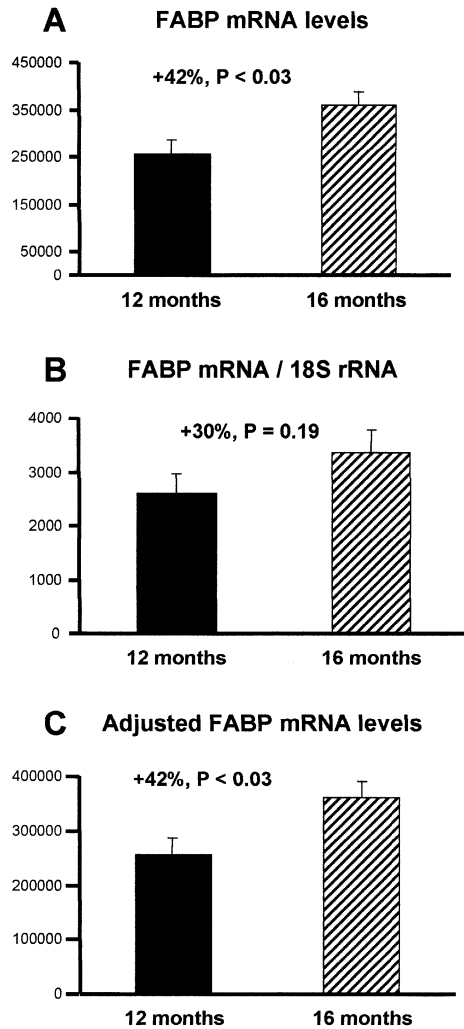


Fig. 5. LS means for FABP mRNA levels in *splenius* (SP) muscle from bulls and steers at 12 and 16 months of age. Comparison of three approaches to analyze the data: A. Crude data expressed in arbitrary densitometric units per μg of total RNA loaded on the gel. B. FABP mRNA levels divided by 18S rRNA levels. C. FABP mRNA levels adjusted for comparable 18S rRNA levels (the statistical model contained 18S rRNA levels as a covariable).

simple ratio, where the control RNA as the divisor has a proportional impact on the value. Co-variance contributes to explain a part of variance of the mRNA of interest and thus results in diminishing the error term and strengthening the testing sensitivity.

Our example points out the misleading conclusions that can be drawn from inappropriate data normalization and show the increase in testing sensitivity that can be gained by introducing a constant reference RNA species as a co-variable.

3.5. Comparison of differences between two groups of samples

A frequent task in molecular biology is to determine the level of regulation of a known gene of interest. For example,

both the lipoprotein lipase (LPL) activity and LPL mRNA levels are measured in the same samples. If a difference in LPL activity between two groups of subjects is detected with a concomitant difference in LPL mRNA levels, a regulation at the pre-translational level (i.e. a regulation at the level of gene expression) will be concluded. By contrast, a regulation at the level of the protein translation or maturation will be postulated, if LPL activity and LPL mRNA levels in two groups of subjects do not vary concomitantly. However, the statistical approach may be essential to the interpretation of these data, as described in the following example.

In a preliminary experiment, hearts of calves exhibited an average LPL activity of 1,026 mU/g wet weight (SD 264) and an average LPL mRNA levels of 1,000 arbitrary units (SD 899) [14]. The level of regulation of LPL expression is determined in an experimental design which aims to compare two groups of ten calves differing in their nutrition. In this theoretical example, the difference in LPL activity is 500 mU/g (1,500 vs 1,000 mU/g), the difference in mRNA levels is 500 arbitrary units (1,500 vs 1,000 units) and the standard deviations are similar for the two groups of ten calves and equal to the above values (264 and 899 units for LPL activity and mRNA levels respectively). Intuitively, the investigator would make the hypothesis that the difference in LPL activity is explained by the difference in mRNA levels due to a 1.5-fold difference for both parameters. But, the investigator applies the Student's t test and calculates the t value: $t = [(\text{difference between the two means}) / \text{standard deviation}] \times \sqrt{(\text{number of animals per group} / 2)}$. For LPL activity, $t = 500 / 264 \times \sqrt{10 / 2} = 4.2$ ($P < 0.001$ according to the t table for 18 degree of freedom), but for LPL mRNA levels, $t = 500 / 899 \times \sqrt{10 / 2} = 1.24$ (difference not significant even for $P < 0.20$ according to the t table). In other words, the high standard deviation for LPL mRNA levels requires at least 28 animals to detect any significant 1.5-fold difference in LPL mRNA levels whereas more than 3 animals may be enough to detect any 1.5-fold difference in LPL activity [15].

4. Discussion

The validation of any laboratory procedure corresponds to the elaboration of proofs that the method is well adapted to its objectives. Three major criteria are often sought to be achieved, i.e. the specificity, the linearity and the fidelity of a laboratory method. For specificity, the possible influence of buffers and other matrix or of similar molecules is tested. For linearity, standard curves are analyzed to calculate regression equations. For fidelity, limits and variations of repeatability and reproducibility are calculated [16]. In protein studies and molecular biology, a great deal of efforts is generally undertaken to verify the specificity of the antibodies or the molecular probes. Therefore, the specificity of a method is assumed given in the present context. By contrast, the linearity and the fidelity

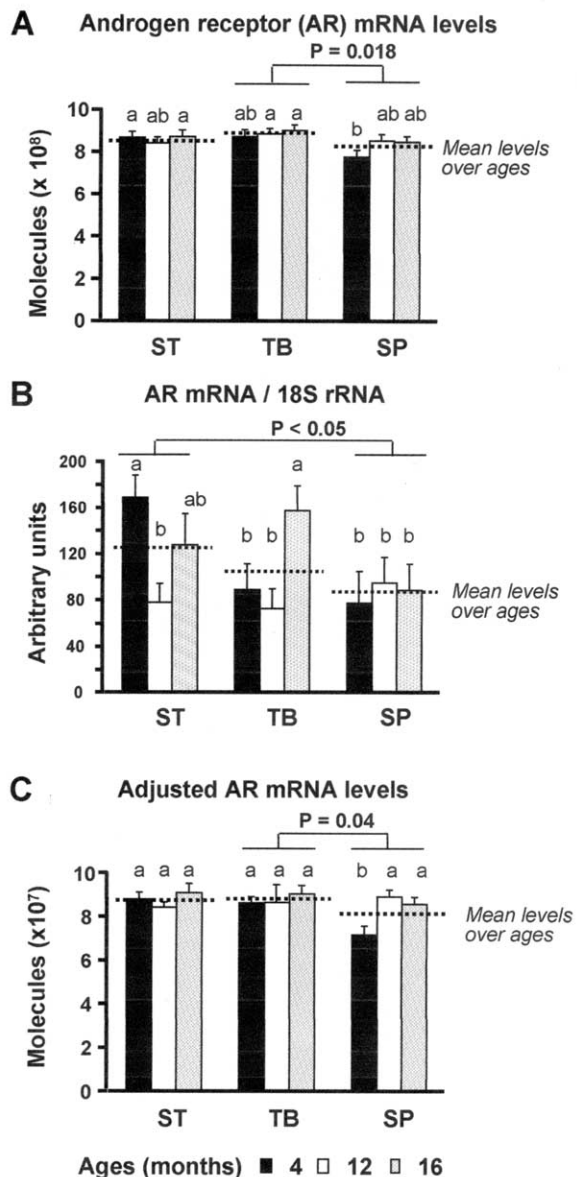


Fig. 6. LS means for androgen receptor mRNA levels in three skeletal muscles: *semitendinosus* (ST), *triceps brachii* (TB) and *splenius* (SP) of bulls and steers at 4, 12 and 16 months of age. Comparison of three approaches to analyze the data: A. Crude data expressed in number of androgen receptor mRNA molecules per 200 ng of total RNA. B. Androgen receptor mRNA levels divided by 18S rRNA levels. C. Androgen receptor mRNA levels adjusted for comparable 18S rRNA levels (the statistical model contained 18S rRNA levels as a covariable).

of the methods depend on normalization procedures and statistical analysis of resulting data.

We examined the effect of two data-normalization methods on the interpretation of molecular data (referred to as the ratio method, in Fig. 5B and 6B, and the regression method, in Fig. 5C and 6C). The normalization procedure using the ratio method is based on the assumption that the signal values are proportional to the concentrations of the analyzed samples. This is true only within the linear range of the standard curve, and when no sample in the test tube

gives a zero signal value (i.e. when the intercept of the standard curve with the signal scale is equal to zero). This occurs rarely in practice. Most researchers intuitively use a regression-based approach, when classical laboratory techniques such as spectrophotometric methods (for instance ELISA) are used. The blank is always subtracted from experimental data and calculation is always made in the linear range of the standard curve. Under these circumstances and assuming no technical errors, the ratio normalization of raw data is appropriate, because the relationship between the variable of interest (y) and the control variable (x) is $y = ax$, where a is a constant term.

However, without any subtraction of the blank, data could be analyzed with an appropriate statistical model that applies the regression method [13]. In this case, the relationships between the variable of interest (y) and the control variable (x) is: $y = ax + b$, where b is a constant term equal to the y -intercept of the standard curve which is different from zero. In other words, the ratio-method, when technical errors are negligible, is the specific case with $b =$ zero of a more general picture. This is why the regression-based approach will avoid the statistical bias induced by the ratio method when b is different from zero [13].

Although molecular biology data are very costly to obtain, involving the time-demanding optimization of the appropriate techniques, little effort is taken to verify the applicability of the most commonly used ratio normalization of resulting data. As discussed above, this bears *a priori* the implication that the intercept of the mRNA standard curve is equal to zero. However, the relationships between target RNA levels and RNA levels of a reference (for instance the 18S rRNA) may not regress to a zero intercept (Fig. 4C). This is the reason why the ratio method may introduce a major bias, which has been previously outlined in other biological fields [13]. Despite this observation, the use of the ratio method in molecular biology is still common practice, probably because of its simplicity.

In an example from our previous study on FABP and androgen receptor expression in bovine muscles [8], we show that the two data-normalization methods (the regression-based calculation and the ratio method) provide opposite results (Figs. 5 and 6). We demonstrated a major limitation of Northern blot analyses by using increasing amounts of a known RNA sample (Fig. 4). This technical problem can be explained by the different conditions of image capture of the electrophoresis or autoradiograms and the quantification of the specific bands. For 18S rRNA, the time of exposure of the Northern blot is short and background low. By contrast, for low abundance mRNA species (for instance, GLUT4 mRNA), time of exposure is much longer, which considerably increases the sensitivity of the procedure, even if background values are also higher. Under these conditions, it is likely to introduce a technical bias, which may result in misleading interpretation of the data.

We also examined the effect of data variance on the general physiological conclusions, when biochemical and

molecular data are analyzed concomitantly and set in a context to each other. We demonstrate in examples from our previous work, that coefficients of variation of means for mRNA levels are usually much higher than coefficients of variation of means for biochemical or enzymatic data. This observation was confirmed when we quantified increasing amounts of standards or representative samples by any of these methods in the same assay: the linearity of the standard curve is always less for Northern blot analysis ($r^2 = 0.91$ to 0.96 maximum after normalization by using 18S values and removal of invalid points) than for Western blot analysis ($r^2 = 0.96$ without any normalization or validation steps) or ELISA analysis ($r^2 = 0.99$). In our understanding, it is indeed logical that Northern blot analysis is less reproducible than biochemical techniques, since Northern blot involves many steps: quantification, loading, electrophoresis and transfer of RNA to membranes, then hybridization and image capture for quantification. On the contrary, enzymatic assays or ELISA are simpler and more quantitative, since the careful use of pipettes checks for sample or reagent volume. Because of this higher risk of technical errors in Northern blot analysis than in an enzyme assay or ELISA, increased sample sizes are required to balance the source of error.

In conclusion, this paper provides evidence that a regression-based approach to normalize mRNA data is suitable in any case [8], whereas the ratio method can be used in specific cases only. In addition, using the ratio method may even introduce a significant statistical bias and misleading biological interpretations. Furthermore, the higher variance for mRNA level data compared to biochemical or enzymatic data implies that differences brought about by any experimental treatment need to be relatively large in size before they can be detected by molecular approaches. Consequently, the detection limit of a significant difference in a given experimental setting must be assessed prior to the experiment, which allows deducing the required sample sizes for each technical approach in the experimental design. Otherwise, misleading physiological conclusions can be made.

Acknowledgments

We thank Drs P. Gasqui (INRA, Unité d'Epidémiologie Animale, Theix), F. Bornes, B. Graulet and C. Piot (INRA, Unité de Recherches sur les Herbivores) for helpful discussions. We also thank Nicole Guivier, and Geneviève Gentes for their skilled technical assistance.

References

- [1] J.F. Hocquette, I. Ortigues-Marty, D.W. Pethick, P. Herpin, X. Fernandez, Nutritional and hormonal regulation of energy metabolism in

- skeletal muscles of meat-producing animals, *Livest Prod Sci* 56 (1998) 115–143.
- [2] J.F. Hocquette, B. Graulet, C. Castiglia-Delavaud, F. Bornes, N. Lepetit, P. Ferré, Insulin-sensitive glucose transporter transcript levels in calf muscles assessed with a bovine GLUT4 cDNA fragment, *Int. J Biochem Cell Biol* 28 (1996) 795–806.
- [3] J.F. Hocquette, C. Castiglia-Delavaud, B. Graulet, P. Ferré, B. Picard, M. Vermorel, Weaning marginally affects glucose transporter (GLUT4) expression in muscles and adipose tissues in the calf, *Brit J Nutr* 78 (1997) 251–271.
- [4] J.F. Hocquette, B. Graulet, T. Olivecrona, Lipoprotein lipase activity and mRNA levels in bovine tissues, *Comp Biochem Phys* 121 (B) (1998) 85–96.
- [5] J.F. Hocquette, B. Graulet, M. Vermorel, D. Bauchart, Weaning affects lipoprotein lipase activity and gene expression only in adipose tissues and in masseter but not in other muscles of the calf, *Brit J Nutr* 86 (2001) 433–441.
- [6] T. Raclot, R. Groscolas, D. Langin, P. Ferré, Site-specific regulation of gene expression by n-3 polyunsaturated fatty acids in rat white adipose tissues, *J Lipid Res* 38 (1997) 1963–1972.
- [7] J.J. Lee, P.J. Smith, S.K. Fried, Mechanisms of decreased lipoprotein lipase activity in adipocytes of starved rats depend on duration of starvation, *J Nutr* 128 (1998) 940–946.
- [8] A.M. Brandstetter, M.W. Pfaffl, J.F. Hocquette, D.E. Gerrard, B. Picard, Y. Geay, H. Sauerwein, Effects of muscle type, castration, age and compensatory growth rate on androgen receptor mRNA expression in bovine skeletal muscle, *J Anim Sci* 78 (2000) 629–637.
- [9] R.J.A. Paulussen, M.J.H. Geelen, A.C. Beynen, J.H. Veerkamp, Immunochemical quantitation of fatty-acid-binding proteins. I. Tissue and intracellular distribution, postnatal development and influence of physiological conditions on rat heart and liver FABP, *Biochim Biophys Acta* 1001 (1989) 201–209.
- [10] C. Piot, J.F. Hocquette, P. Herpin, J.H. Veerkamp, D. Bauchart, Dietary coconut oil affects more lipoprotein lipase activity than the mitochondria oxidative capacities in muscles of preruminant calves, *J Nutr Biochem* 11 (2000) 231–238.
- [11] J.F. Hocquette, F. Bornes, M. Balage, P. Ferré, J. Grizard, M. Vermorel, Glucose-transporter (GLUT4) protein content in oxidative and glycolytic skeletal muscles from calf and goat, *Biochem J* 305 (1995) 465–470.
- [12] M. Balage, J.F. Hocquette, B. Graulet, P. Ferré, J. Grizard, Skeletal muscle glucose transporter (GLUT4) protein is decreased in lactating goats, *Anim Sci* 65 (1997) 257–265.
- [13] E.T. Poehlman, M.J. Toth, Mathematical ratios lead to spurious conclusions regarding age- and sex-related differences in resting metabolic, *Am J Clin Nutr* 61 (1995) 482–485.
- [14] B. Graulet, J.F. Hocquette, D. Bauchart, Lipoprotein lipase activity and gene expression in heart and skeletal muscles in the preruminant calf, *Proc Nutr Soc* 55 (1996) 18A.
- [15] R. Mera, H. Thompson, C. Prasad, How to calculate sample size for an experiment: a case-based description, *Nutr Neurosci* 1 (1998) 87–91.
- [16] I. Ortigues-Marty, J.F. Hocquette, M. Vermorel, Methodologies and techniques used in the study of energy metabolism: Elements of discussion and perspectives, 15th Symposium on Energy Metabolism in Animals, 11–16 September 2000, Snekkersten (Denmark), in: *Energy Metabolism in Animals*, A. Chwalibog, K. Jakobsen (Ed.), EAAP Publication N° 103, Wageningen Pers. (The Netherlands), 2001, pp. 81–89.