

# L'importance de la variabilité biologique et de la précision des mesures en biologie intégrative

## Importance of biological variability and technical precision in integrative biology

J-F. HOCQUETTE, C. JURIE, A. LISTRAT, I. CASSAR-MALEK

INRA, Unité de Recherches sur les Herbivores, Equipe Croissance et Métabolisme du Muscle, Theix, 63122 Saint-Genès Champanelle, France

### INTRODUCTION

La biologie intégrative a pour but de comprendre le fonctionnement d'un système biologique dans sa globalité en associant des méthodes d'investigation complémentaires. Elle implique un aller-retour entre les différents niveaux d'organisation du vivant en associant des données de génomique et des données biologiques plus classiques. Cette démarche implique toutefois une bonne connaissance de la variabilité des données qui sont associées. En effet, une variabilité d'origine biologique ne doit pas être confondue avec une variabilité due à l'imprécision des mesures sous peine de biaiser les conclusions biologiques. Notre objectif est donc de préciser la variabilité de données généralement considérées comme complémentaires.

### 1. MATERIEL ET METHODES

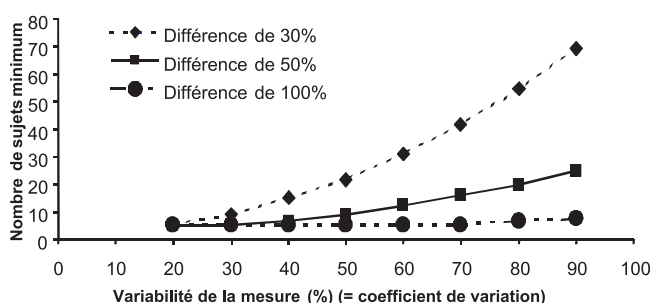
Le nombre de sujets nécessaire pour détecter une différence relative de 30, 50 ou 100 % entre deux groupes d'animaux a été calculé en fonction de l'écart type des données en utilisant la formule du test t de Student (Mera *et al.*, 1998). Les performances de croissance, la composition corporelle et les caractéristiques contractiles et métaboliques des fibres du muscle *Longissimus* ont été déterminées sur 70 taurillons (Hocquette *et al.*, 2001). Les teneurs en collagène de ce muscle ont été déterminées par Listrat et Hocquette (2003). L'activité et la teneur en ARNm de la lipoprotéine-lipase (LPL) ont été déterminées dans le cœur de deux groupes de veaux par Graulet *et al.* (1996). L'étude du transcriptome du muscle de bovin a été réalisée par Sudre *et al.* (2003).

### 2. RESULTATS

#### 2.1. CALCUL DU NOMBRE DE SUJETS NÉCESSAIRE POUR DÉTECTER UNE DIFFÉRENCE DE 30, 50 OU 100 % DE LA MOYENNE

D'après Mera *et al.* (1998), le nombre de sujets ( $n$ ) nécessaire pour détecter une différence entre 2 moyennes  $\mu_i$  et  $\mu_j$  est tel que  $n = 2 \times [t_{5\%} \times \sigma / (\mu_i - \mu_j)]^2$  où  $\sigma$  est l'écart type de la mesure et  $t_{5\%}$  la valeur du  $t$  de Student pour une probabilité de 5 %. Si on exprime  $\sigma$  en fonction du coefficient de variation, la Figure 1 représente  $n$  en fonction de la différence relative que l'on souhaite observer entre les deux moyennes  $[(\mu_i - \mu_j) / \mu_i]$ .

Figure 1 : Nombre d'animaux nécessaire pour détecter une différence significative de 30, 50 ou 100 %



Il faut plus d'animaux par groupe expérimental pour détecter une différence de 30 % lorsque que les mesures sont très variables. Cette différence s'atténue d'autant plus que la différence à observer est grande (figure 1).

### 2.2. DONNÉES BIOLOGIQUES

Les poids des animaux, leur vitesse de croissance et leur note d'état corporel ont un coefficient de variation entre animaux inférieur à 15 % (Hocquette *et al.*, 2001).

Pour les propriétés des fibres musculaires et les teneurs en collagène, ce coefficient varie entre 13 et 31 % (Hocquette *et al.*, 2001 ; Listrat et Hocquette, 2003), à l'exception de l'activité cytochrome-c oxydase (38 %) et de la teneur en collagène insoluble (47 %).

La variabilité entre animaux est de 26 % pour l'activité LPL et de 90 % pour les teneurs en ARNm de la LPL (Graulet *et al.*, 1996). Environ le tiers de cette variabilité est d'ordre technique due à l'imprécision des méthodes de Northern blot (Hocquette et Branstetter, 2002). Cette forte variabilité s'explique aussi par le faible niveau d'expression de ce gène. Enfin, la variabilité technique des mesures du transcriptome est en moyenne de 20 % mais varie de 1 % à plus de 100 % selon les ADNc (Sudre *et al.*, 2003).

### 2.3. INTÉGRATION DES DONNÉES

Pour détecter une différence significative de l'ordre de 30 % avec moins de 10 animaux par groupe expérimental, il est nécessaire que la variabilité de la mesure n'excède pas 30 % (figure 1). Les données de performances zootechniques et la plupart des données biochimiques répondent à ce critère. En revanche, la variabilité de certaines données de biologie moléculaire peut être plus élevée. En génomique, une procédure souvent utilisée consiste à considérer comme non fiables et à éliminer les données avec une variabilité supérieure à 25 % (Sudre *et al.*, 2003). Cette procédure est justifiée dans un objectif de biologie intégrative.

### CONCLUSION

L'association de données issues de plusieurs techniques différentes suppose que leur variabilité soit comparable. Ce n'est pas toujours le cas en fonction de la précision des méthodes utilisées. La biologie intégrative ne trouvera du sens que si les protocoles expérimentaux et les méthodes d'analyse sont adaptés en conséquence.

Graulet, B., Hocquette, J.F., Bauchart, D., 1996. Proc. Nutr. Soc., 55, 18A

Hocquette, J.F., Picard, B., Trillat, G., Normand, J., Boissy, A., Culioli, J., 2001. Renc. Rech. Ruminants, 8, 53-56.

Hocquette, J.F., Branstetter, A., 2002. J. Nutr. Biochem., 13, 370-377.

Listrat, A., Hocquette, J.F., 2003. Renc. Rech. Ruminants, 10.

Mera, R., Thompson, H., Prasad, C. 1998. Nutr. Neurosci., 1, 87-91.

Sudre, K., Leroux, C., Piétu, G., Cassar-Malek, I., Petit, E., Listrat, A., Auffray, C., Picard, B., Martin, P., Hocquette, J.F.

2003. J. Biochem., 133, 745-756.