Testing for allele- versus genotype-specific transmission ratio distortion

J. Casellas[1], S. Id-Lahoucine[2] & A. Cánovas[2]

[1] Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain
joaquim.casellas@uab.cat (Corresponding Author)
[2] Centre for Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, N1G2W1 Guelph, ON, Canada

Summary

The transmission ratio distortion (TRD) is defined as the departure from expected frequencies under Mendelian inheritance. Current statistical approaches to investigate TRD account for two independent scenarios, model parametrizations assuming that TRD relies on allelic or genotypic mechanisms, respectively. If compared, results may provide additional information about the mechanism underlying TRD. Both analytical approaches have been compared by Bayes factors on simulated data sets with 1,000 individuals each. Analyses revealed trustworthy patterns where the model used for simulation was systematically favoured. Moreover, the correlation between simulated and predicted distortion parameters was high (>0.85). The only weakness was suggested for genotypic TRD with small to null dominance distortions. As a whole, the comparison of allelic and genotypic TRD models must be viewed as an appealing tool that should contribute to fostering research on TRD.

Keywords: Bayes factor, genetic marker, transmission ratio distortion

Introduction

The transmission ratio distortion (TRD) is the departure from the expected genotypic frequencies under Mendelian inheritance in offspring (Crow, 1999). The biological background of TRD is complex and can be due to multiple physiological mechanisms during gametogenesis, fertilization, foetal and embryonic development, and early neonatal life (Wakasugi, 1974; Lyon, 1991; Moore, 2006). Within this context, TRD should be considered as a relevant issue with potential outcomes in multiple research fields where its incidence, biological nature and magnitude should be accurately screened across the whole genome of animals and plants.

The analysis of TRD has been systematically addressed with standard statistical procedures such as χ^2- (Paz-Miguel *et al.*, 2001; Underkoffler *et al.*, 2005) or *t*-tests (Shendure *et al.*, 1998), although this masked the genetic mechanism (or inheritance model) involved in the departure. The recent development of TRD-specific analytical approaches (Casellas *et al.,* 2012, 2014, 2017) has contributed useful tools to elucidate the genetic mechanisms underlying TRD, although little is known about biological processes, metabolic pathways and physiological or reproductive stages ruling these TRD. Within this context, additional information can be obtained from current analytical models; the different parametrizations can be evaluated and compared by appropriate methodologies such as Bayes factors (BF; Kass & Raftery, 1995). The main objective of this study focuses on the review of TRD-specific models, their rationale, and interpretation, and the discrimination among

different models by a Bayes factor approach implemented on simulated data.

Material and methods

TRD-specific models

TRD due to single-allele mechanisms

Take as starting point a genetic marker with alleles *A* and *B*. The probability of each genotype to be sampled in the offspring generation (p_{off}) generalizes to:

$$\begin{aligned} p_{off}(AA) &= p_{sire}(A) \times p_{dam}(A) \\ p_{off}(AB) &= [p_{sire}(A) \times p_{dam}(B)] + [p_{sire}(B) \times p_{dam}(A)] \\ p_{off}(BB) &= p_{sire}(B) \times p_{dam}(B), \end{aligned} \quad (1)$$

where p_{sire} and p_{dam} define the probability of inheritance of a given allele from the sire and the dam, respectively. Parent-specific TRD can integrate into previous probabilities for heterozygous parents as follows (Casellas *et al.*, 2014),

$$p_{sire}(A) = 1 - p_{sire}(B) = 0.5 + \alpha_{sire} \quad \text{and} \quad p_{dam}(A) = 1 - p_{dam}(B) = 0.5 + \alpha_{dam}. \quad (2)$$

Note that α_{sire} and α_{dam} are allelic sire- and dam-specific TRD estimates ranging from -0.5 to 0.5. Additionally, probabilities of transmission from homozygous parents reduce to 0 or 1 (e.g., $p(A) = 1 - p(B) = 1$for *AA* sires). This kind of TRD must be linked to haploid reproductive stages (e.g., gametogenesis, transport in the reproductive tract, and fertilization, among others), although some specific genetic phenomena such as imprinting and polar overdominance may also originate them. Within this context, uncertainly arises when crossing heterozygous parents, i.e.,

$$\begin{aligned} p_{off}(AA) &= 0.25 + \alpha_{sire}/2 + \alpha_{dam}/2 + \alpha_{sire}\alpha_{dam} = 1/4(1 + 2\alpha_{sire} + 2\alpha_{dam} + 4\alpha_{sire}\alpha_{dam}) \\ p_{off}(AB) &= 0.5 - 2\alpha_{sire}\alpha_{dam} = 1/2(1 - 4\alpha_{sire}\alpha_{dam}) \\ p_{off}(BB) &= 0.25 - \alpha_{sire}/2 - \alpha_{dam}/2 + \alpha_{sire}\alpha_{dam} = 1/4(1 - 2\alpha_{sire} - 2\alpha_{dam} + 4\alpha_{sire}\alpha_{dam}), \end{aligned} \quad (3)$$

or when crossing heterozygous (e.g., *AB* sire) with homozygous (e.g., *AA* dam) parents, i.e.,

$$p_{off}(AA) = 0.5 + \alpha_{sire} = 1/2(1 + 2\alpha_{sire}) \quad \text{and} \quad p_{off}(AB) = 0.5 - \alpha_{sire} = 1/2(1 - 2\alpha_{sire}). \quad (4)$$

TRD due to genotype-related mechanisms

As developed by Casellas *et al.* (2012), TRD can be modeled on a genotype basis by assuming additive (α_G) and dominance (δ_G) parameters, regardless of the origin of each allele. In contrast to the previous model, genotypic TRD must be linked to diploid stages such as zygote, embryo, foetus or neonate survival, embryo implantation, … The original development (Casellas *et al.*, 2012) restricted to an F_2 design with heterozygote parents, where the probability of each genotype in the offspring generation can generalize to

$p_{off}(AA) = 1/4(1 + \alpha_G + \delta_G)$,
$p_{off}(AB) = 1/2(1 - \delta_G)$
(5)
$p_{off}(BB) = 1/4(1 - \alpha_G + \delta_G)$.

Given that distortion phenomena depends on the genotype itself, probabilities can be straightforwardly deduced when crossing heterozygous (e.g. *AB* sire) with homozygous (e.g., *AA* dam) parents, i.e.,

$p_{off}(AA) = 1/2(1 + \alpha_G/2 + \delta_G)$ and $p_{off}(AB) = 1/2(1 - \alpha_G/2 - \delta_G)$.
(6)

Additive TRD from single-allele and genotype-related TRD become equivalent when $\alpha_{sire} = \alpha_{dam} = \alpha_G/4$. As anticipable, genotype-specific TRD captures all possible dominance deviations (Casellas *et al*., 2012), whereas allele-specific TRD forces a dominance-like term when crossing two heterozygous parents (i.e., $4\alpha_{sire}\alpha_{dam}$), this unnoticely escaping from a pure additive model (Casellas *et al*., 2014, 2017).

Model comparison on simulated data sets

Simulation of TRD on genetic markers

A biallelic marker (alleles *A* and *B*) was assumed with allelic frequencies $f(A) = p = 0.5$ and $f(B) = 1 - p = q$. For each offspring, parents' genotypes were randomly sampled with probability $p(AA) = p^2$, $p(AB) = 2pq$ and $p(BB) = q^2$, and a total of 1,000 offspring were originated for each population. Single-allele- or genotype-specific TRD was generated in the offspring generation on the basis of TRD models described above. For each population, TRD parameters were stated (α_{sire} and α_{dam} for allelic TRD, and α_G and δ_G for genotypic TRD).

TRD analysis and model comparison

Each data set was analysed twice by implementing both TRD-specific models under a Bayesian framework. The conditional probability of each offspring genotype was assumed as described above, and priors for each TRD-specific parameters were defined as follows:

$p(\alpha_{sire}) = p(\alpha_{dam}) = 1$ $[-0.5 \leq \alpha_{sire}, \alpha_{dam} \leq 0.5]$
(7)
$p(\alpha_G) = p(\delta_G) = 0.5$ $[-1 \leq \alpha_G, \delta_G \leq 1]$

Each model was analyzed by launching a unique Monte Carlo Markov chain with 10,000 sampling iterations after a burn-in period of 1,000 iterations. Both models were compared by a BF (Newton & Raftery, 1994).

Results and discussion

Results from simulated data sets including 1,000 individuals each evidenced trustworthy patterns. When sire-specific allelic TRD was assumed ($\alpha_{dam} = 0$) during the simulation

process, this model was clearly favoured against genotypic TRD (Figure lA). The correlation between simulated and predicted (posterior mean) α_{sire} was 0.998, whereas predicted α_{dam} (posterior mean) ranged between 0.023 and -0.064. None out of 1,000 analyzed populations revealed lower-than-1 BF (the minimum BF was 1.201), although this lack of false negatives was also influenced by population size (i.e., 1,000 individuals per population).

A similar pattern was obtained for data sets simulated under genotypic TRD. The BF favored the genotypic TRD model in 914 out of 1,000 populations, and little impact was revealed for simulated α_G (results not shown). However, small δ_G increased the chance of ambiguous BF (~1; Figure 1B, suggesting that a certain degree of dominance was required for the genotypic TRD model. Average correlation between simulated and predicted (posterior mean) α_G was 0.877, whereas this increased up to 0.985 for δ_G.

As a whole, both allelic and genotypic TRD models can be easily implemented and compared, they providing relevant information about the origin of each TRD phenomenon.

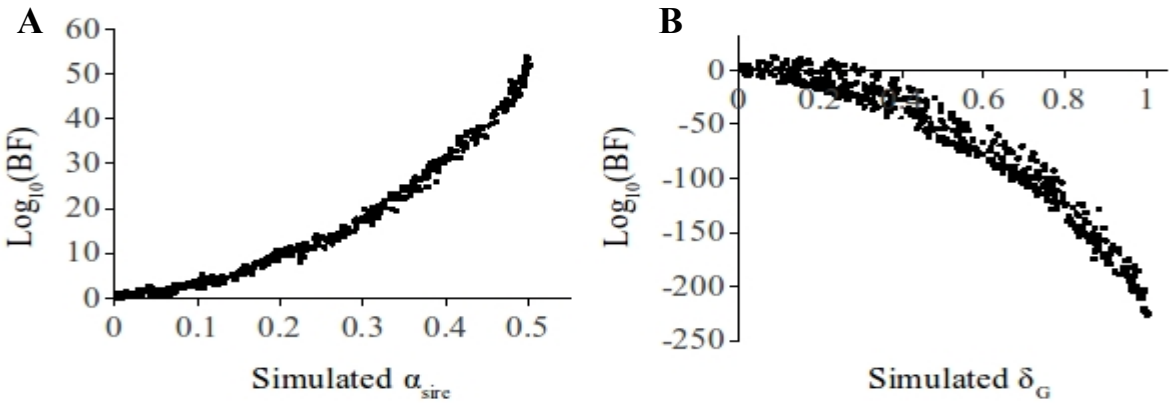

Figure 1. Bayes factor (BF) of allelic versus genotypic transmission ratio distortion (TRD) models for data sets simulated under allelic (A) and genotypic (B) TRD.

Acknowledgements

The authors acknowledge the finical support from Beef Farmers Ontario (BFO-16-04, Canada) and Ministerio de Ciencia e Innovación (CGL2016--80155-R, Spain).

List of References

Casellas, J., J.J. Cañas-Álvarez, A. González-Rodríguez, A. Puig-Oliveras, M. Fina, J. Piedrafita, A. Molina, C. Díaz, J.A. Baró & L. Varona, 2017. Bayesian analysis of parent-specific transmission ratio distortion in Spanish beef cattle breeds. Anim. Genet. 48: 93-96.

Casellas, J., R.J. Gularte, C.R. Farber, L. Varona, M. Mehrabian, E.E. Schadt, A.J. Lusis, A.D. Attie, B.S. Yandell & J.F. Medrano, 2012. Genome scans for transmission ratio distortion in mice. Genetics 191: 247-259.

Casellas, J., A. Manunza, A. Mercader, R. Quintanilla & M. Amills, 2014. A flexible Bayesian model for testing for transmission ratio distortion. Genetics 198: 1357-1367.

Crow, J.F., 1999. Unmasking a cheating gene. Science 283: 1651.

Kass, R.E., & A.E. Raftery, 1995. Bayes factors. J. Am. Stat. Assoc. 90: 773-795.

Lyon, M.F., 1991. The genetic basis of transmission-ratio distortion and male sterility due to the t-complex. Am. Nat. 137: 349-358.

Moore, C.S., 2006. Postnatal lethality and cardiac anomalies in the Ts65Dn Down syndrome mouse model. Mamm. Genome 17: 1005-1012.

Newton, M.A., & A.E. Raftery, 1994. Approximate Bayesian inference by the weighted likelihood bootstrap (with discussion). J. Royal Stat. Soc. B 56: 1-48.

Paz-Miguel, J., F. Pardo-Manuel de Villena, P. Sánchez-Velasco & F. Leyva-Cobián, 2001. H2-haplotype-dependent unequal transmission of the 17^{16} translocation chromosome from Ts65Dn females. Mamm. Genome 12: 83-85.

Shendure, J., J.A. Melo, K. Pociask, R. Derr & L.M. Silver, 1998. Sex-restricted non-mendelian inheritance of mouse chromosome 11 in the offspring of crosses between C57BL/6J and (C57BL/6J × DBA/2J) F_1 mice. Mamm. Genome 9: 812-815.

Underkoffler, L.A., L.E. Mitchell, Z.S. Abdulali, J.N. Collins & R.J. Oakey, 2005. Transmission ratio distortion in offspring of mouse heterozygous carriers of a (7.18) Robertsonian translocation. Genetics 169: 843-848.

Wakasugi, N., 1974. A genetically determined incompatibility system between spermatozoa and eggs leading to embryonic death in mice. J. Reprod. Fertil. 41: 85-96.