A Flexible Bayesian Model for Testing for Transmission Ratio Distortion

Joaquim Casellas,^{*,†} Arianna Manunza,[†] Anna Mercader,^{*} Raquel Quintanilla,[†] and Marcel Amills[†]

^{*}Departament de Ciència Animal i dels Aliments and [†]Department of Animal Genetics, Center for Research in Agricultural Genomics (CSIC-IRTA-UAB-UB), Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain, and [†]IRTA, Genètica i Millora Animal, Torre Marimon, 08140 Caldes de Montbui, Spain

ABSTRACT Current statistical approaches to investigate the nature and magnitude of transmission ratio distortion (TRD) are scarce and restricted to the most common experimental designs such as F_2 populations and backcrosses. In this article, we describe a new Bayesian approach to check TRD within a given biallelic genetic marker in a diploid species, providing a highly flexible framework that can accommodate any kind of population structure. This model relies on the genotype of each offspring and thus integrates all available information from either the parents' genotypes or population-specific allele frequencies and yields TRD estimates that can be corroborated by the calculation of a Bayes factor (BF). This approach has been evaluated on simulated data sets with appealing statistical performance. As a proof of concept, we have also tested TRD in a porcine population with five half-sib families and 352 offspring. All boars and piglets were genotyped with the Porcine SNP60 BeadChip, whereas genotypes from the sows were not available. The SNP-by-SNP screening of the pig genome revealed 84 SNPs with decisive evidences of TRD ($BF > 100$) after accounting for multiple testing. Many of these regions contained genes related to biological processes (e.g., nucleosome assembly and co-organization, DNA conformation and packaging, and DNA complex assembly) that are critically associated with embryonic viability. The implementation of this method, which overcomes many of the limitations of previous approaches, should contribute to fostering research on TRD in both model and nonmodel organisms.

TRANSMISSION ratio distortion (TRD) is the departure from the expected Mendelian inheritance ratio in offspring (Silver 1993; Pardo-Manuel de Villena *et al.* 2000; Casellas *et al.* 2012). Although this phenomenon has been reported in humans (Evans *et al.* 1994; Vorechovsky *et al.* 1999; Pardo-Manuel de Villena and Sapienza 2001) and mice (Wakasugi 1974; Lyon 1991; Wu *et al.* 2005), little is known about its causes and prevalence in livestock species such as pigs. This TRD can originate from diverse biological mechanisms such as germline selection (Huang *et al.* 2013), meiotic drive (Agulnik *et al.* 1990; Lyon 1991; Dyer *et al.* 2007), gametic competition and imprinting errors (Huang *et al.* 2013; Labbe *et al.* 2013), embryo or fetal failure (Wakasugi 1974), and differential postnatal viability (Moore 2006; Casellas 2007).

The incidence, biological nature, and magnitude of TRD in livestock species are mostly unknown. This is a relevant

issue because reproductive performance is one of the traits with a major impact on the economic output of farms. Identifying the genetic factors that modulate differential survival of gametes, zygotes, and embryos may have multiple applications on the improvement of reproductive success. From a scientific viewpoint, TRD can be also a relevant confounding factor in identity-by-descent sharing and transmission disequilibrium test association studies, complicating the identification of the causal mutations underlying the variation of phenotypes under study (Huang *et al.* 2013). Despite its relevance, TRD cannot be easily addressed in domestic species because of incomplete pedigrees and the incidence of ungenotyped individuals in the parental generation. These peculiarities prevented the implementation of classical statistical approaches such as standard χ^2 (Underkoffler *et al.* 2005) or *t*-tests (Shendure *et al.* 1998), and missing genotypes in the parental generation can become a major problem if expected offspring genotypes cannot be anticipated by the nature of the experimental cross itself (e.g., F_2 design).

Herewith, we present a new statistical methodology that represents a generalization of the BF approach developed by Casellas *et al.* (2012) to a much wider range of experimental designs. This method overcomes many of the limitations of

Figure 1 (A–E) Plots of the pairwise distribution of allele frequencies, the transmission ratio distortion (TRD) effect, and the Bayes factor, testing TRD on simulated data sets.

previous approaches, models TRD on the basis of each offspring genotype, and traces back allele inheritance to (i) parents' genotypes, if available, or (ii) population-based allele frequencies, if genotypes are not available. As a proof of concept, we have used it to determine the prevalence and genomic distribution of TRD in the pig genome.

Materials and Methods

Transmission ratio distortion model

General parameterization: Previous studies conducted on the TRD phenomena typically relied on classical experimental designs such as backcrosses (Montagutelli *et al.* 1996; Shendure *et al.* 1998; Vogl and Xu 2000) and F_2 crosses (Casellas *et al.* 2012), where expected frequencies in the last generation were known, even without genotyping parental generations. Such experimental designs allowed researchers to assess the segregation departure by standard χ^2 (Paz-Miguel *et al.* 2001; Underkoffler *et al.* 2005; Casellas *et al.* 2012) or t -tests (Shendure *et al.* 1998), although they cannot be directly extrapolated to more complex scenarios. In addition, missing genotypes in the parental generation can become a major problem if expected offspring genotypes cannot be anticipated by the nature of the experimental cross itself. Within this context, we propose a new statistical approach that expands the original method developed by Casellas *et al.* (2012) to a much wider diversity

of experimental designs within the TRD framework, even accounting for ungenotyped ancestors if necessary.

Assume as a starting point that a given genetic marker has two alleles A_1 and A_2 . Under a flexible inheritance model, the probability of each genotype to be sampled in the offspring generation is generalized to

$$p(A_1A_1) = p_{\text{sire}}(A_1) \times p_{\text{dam}}(A_1)$$

$$p(A_1A_2) = [p_{\text{sire}}(A_1) \times p_{\text{dam}}(A_2)] + [p_{\text{dam}}(A_1) \times p_{\text{sire}}(A_2)]$$

and

$$p(A_2A_2) = p_{\text{sire}}(A_2) \times p_{\text{dam}}(A_2),$$

where p_{sire} and p_{dam} define the probability of inheritance of a given allele from the sire and the dam, respectively. Assuming that all parents came from the same population, probabilities of inheritance related to ungenotyped parents must be assumed as $p_{\text{dam}}(A_1) = p_{\text{sire}}(A_1) = \pi$ and $p_{\text{dam}}(A_2) = p_{\text{sire}}(A_2) = (1 - \pi)$, where π must be viewed as the allelic frequency of the A_1 allele in the population of origin. Note that TRD cannot be independently assessed in these animals and, if any occurs, accumulates into π . For genotyped individuals, the contribution to the offspring generation from A_1A_1 homozygous parents reduces to $p_{\text{sire}}(A_1) = p_{\text{dam}}(A_1) = 1$ and $p_{\text{sire}}(A_2) = p_{\text{dam}}(A_1) = 0$ [alternatively, $p_{\text{sire}}(A_1) = p_{\text{dam}}(A_1) = 0$ and $p_{\text{sire}}(A_2) = p_{\text{dam}}(A_2) = 1$ for A_2A_2 homozygous parents]. Additionally, TRD phenomena

can be appropriately modeled on the basis of heterozygous parents only as

$$p_{\text{sire}}(A_1) = p_{\text{dam}}(A_1) = 0.5 + \alpha$$

and

$$p_{\text{sire}}(A_2) = p_{\text{dam}}(A_1) = 0.5 - \alpha,$$

where α is the TRD parameter with values defined within the parametric space between -0.5 and 0.5 . Note that this parameterization assumes the same TRD effect whatever the sex of the parent. At this point, it is important to highlight the limitations of this assumption and a couple of suitable solutions. Some type of TRD can only be observed through one parent. Meiotic drive restricts to the female gamete because males show symmetric meiosis that results in four functional gametes (Agulnik *et al.* 1990), whereas gamete competition restricts to males, at least in mammals (Huang *et al.* 2013). Within these contexts, the previous assumption of sex-independent TRD effects must be seen as wrong and, as a consequence, we should be extremely cautious when trying to implement the TRD model. At least two solutions could be suggested. When genotypes from the parent generation restrict to one sex (see Duroc data below), estimated TRD effects must be free from influences from the ungenotyped sex (*i.e.*, if only sires are genotyped, gamete competition-related TRD effects could be captured, whereas meiotic drive influences could mask within the π parameter, and vice versa). On the other hand, a sex-specific TRD model could be implemented if genotyping data from all parents become available (see the *Appendix*). In this case, two α parameters are required and missing genotypes in the parental generation cannot be ascertained because we lack additional degrees of freedom for the π parameter. Given that these sex-specific TRD patterns have been rarely reported in the scientific literature, we decided to develop the model within the general framework illustrated above, although specific details about the model with sire- and dam-specific α parameters can be found in the *Appendix*.

Following the standard Bayesian implementation proposed by Casellas *et al.* (2012), the conditional posterior distribution of π and α was defined as

$$p(\pi, \alpha | \mathbf{y}) \propto p(\mathbf{y} | \pi, \alpha) p(\pi) p(\alpha),$$

where $\mathbf{y}$ was the column vector of genotypes of the offspring generation. The contribution of each offspring in the likelihood became a straightforward multiplication of the appropriate terms,

$$p(\mathbf{y} | \pi, \alpha) = \prod_n p(A_i A_j),$$

where $A_i A_j$ was the genotype of the offspring and n was the total number of offspring to be evaluated. Given the lack of previous knowledge about π and α , their *a priori* distributions were assumed flat with appropriate boundaries as follows,

Table 1 Distribution of single-nucleotide polymorphisms (SNPs) across chromosomes (Chrom.) and number of heterozygous boars

Chrom.	No. heterozygous boars							SNPs
	0 (%)	1 (%)	2 (%)	3 (%)	4 (%)	5 (%)	6 (%)	
1	8.9	21.6	23.1	23.4	16.2	6.3	0.6	3,421
2	11.9	24.0	33.0	17.3	10.4	3.1	0.2	1,820
3	9.7	23.3	22.9	22.2	17.0	4.5	0.5	1,316
4	8.8	18.0	27.0	24.8	17.7	3.6	0.1	2,442
5	15.2	17.6	30.5	24.3	9.5	2.5	0.5	1,241
6	8.2	23.0	24.3	26.6	13.7	4.0	0.2	1,352
7	8.7	19.4	29.6	25.3	12.3	4.3	0.4	2,441
8	9.3	18.6	26.4	20.5	17.8	7.2	0.2	1,461
9	21.4	21.9	24.7	19.1	10.3	2.6	0.1	1,379
10	7.5	17.9	31.9	26.0	12.7	3.6	0.5	923
11	4.4	22.5	26.6	31.9	10.5	3.7	0.4	1,273
12	5.7	13.5	30.3	22.2	21.1	5.4	1.8	703
13	9.6	17.0	30.8	24.8	11.5	5.3	1.0	2,253
14	11.3	20.5	32.9	23.3	8.6	3.0	0.5	2,666
15	7.1	25.2	29.8	21.3	11.0	4.6	1.0	1,746
16	8.9	17.2	29.6	24.1	13.8	6.0	0.3	1,009
17	4.7	21.8	25.6	23.9	13.5	9.1	1.3	1,117
18	6.6	22.3	28.9	24.5	14.1	2.4	1.2	744
X	9.1	21.2	25.8	27.3	15.2	1.5	0.0	66
Overall	9.6	20.5	28.1	23.5	13.2	4.5	0.5	29,373

$$p(\pi) = 1 \quad \text{if } \pi \in [0, 1] \text{ and } 0 \text{ otherwise}$$

$$p(\alpha) = 1 \quad \text{if } \alpha \in [-1/2, 1/2] \text{ and } 0 \text{ otherwise.}$$

Testing for TRD relevance: The BF is the standard Bayesian tool to compare two competing models and focuses on the ratio between their posterior probabilities (Kass and Raftery 1995). In contrast with likelihood-based approaches for testing significance, it does not require one to define any null or alternative hypothesis model, providing a probability for both candidate models and avoiding the calculation of levels of significance (Kass and Raftery 1995). As noted by Lavine and Schervish (1999), the BF measures the change in the odds in favor of a given model (relative to the competing model) when going from the prior to the posterior probability. Within this context, frequentist-based concepts like bias or type-I error are meaningless, and inference relies on probabilities across the parametric space (Blasco 2001). More specifically, the BF approach applied on our data set derived from Verdinelli and Wasserman (1995) provided an efficient way to compare two nested models that differ in only one parameter (García-Cortés *et al.* 2001; Varona *et al.* 2001). The statistical relevance of parameter α in the TRD multinomial model can be evaluated by computing the BF between the model including α and a virtual model with $\alpha = 0$. This BF requires the analysis of only the most complex model (García-Cortés *et al.* 2001; Varona *et al.* 2001) and reduces to

$$\text{BF} = p(\alpha = 0) / p(\alpha = 0 | \mathbf{y}) = p(\alpha = 0 | \mathbf{y})^{-1}.$$

It is important to highlight that $p(\alpha = 0) = 1$ was previously defined by the *a priori* probability of α , and therefore,

Figure 2 (A and B) Manhattan plots of the absolute value of the transmission ratio distortion (TRD) effects and the associated Bayes factors across pig chromosomes.

$p(\alpha = 0|y)$ sufficed to obtain the BF. For each sampling iteration, $p(\alpha = 0|y)$ can be calculated on the basis of the marginal posterior distribution of α by applying the Rao–Blackwell kernel density estimation (Gelfand and Smith 1990; Wang *et al.* 1994).

Analysis of TRD in the pig genome

High-throughput SNP genotyping of Duroc pigs: Five purebred Duroc boars (Selección Batallé S.A., Riudarenes, Spain) and their 352 offspring were genotyped with the Porcine SNP60 BeadChip (Illumina, San Diego) by applying the Infinium HD Assay Ultra protocol (Illumina). Note that boars were randomly mated with purebred sows from the same Duroc line and each litter contributed a unique male piglet to this study. These piglets were selected at random at an age of 6 weeks, having five families of half-sib piglets with sizes 17, 44, 60, 72, 73, and 86. Although animals were genotyped for 62,163 SNP markers, analyses were performed on 29,373 SNPs (47.3%) after removing all unmapped SNPs (Illumina) and those SNPs with a call rate <95% or a minimum allele frequency <5%.

TRD model for the Duroc data set: Given that boars and offspring were genotyped whereas sows lacked genomic data, the TRD model developed above simplifies to a multinomial expression involving five half-sib families of piglets with a common boar. The contribution of each half-sib family in the likelihood must be viewed as a multinomial process,

$$p(y|\pi, \alpha) \propto \prod_{i=1}^6 \frac{n_i!}{n_{11,i}!n_{12,i}!n_{22,i}!} p_{o,i}(A_1A_1)^{n_{11,i}} \\ \times p_{o,i}(A_1A_2)^{n_{12,i}} p_{o,i}(A_2A_2)^{n_{22,i}},$$

where $p_{o,i}$ is the probability of a given genotype in piglets sired by i ; n_i is the total number of genotyped progeny from boar i ; and $n_{11,i}$, $n_{12,i}$, and $n_{22,i}$ are the within-family numbers of offspring with genotypes A_1A_1 , A_1A_2 , and A_2A_2 , respectively.

Validation of the TRD model through simulation: To evaluate the accuracy and statistical power of the TRD model on Duroc genomic data, additional data sets were generated by assuming the same number of boars and within-boar offspring as in the Duroc data set. A total of 10,000 data sets were generated and genomic data restricted to a unique biallelic genetic marker. The frequency of the A_1 allele (π^*) was randomly sampled from a uniform distribution between 0.05 and 0.95, and the TRD parameter (α^*) was randomly sampled from a uniform distribution between -0.5 and 0.5 . The genotype of each boar was generated by assuming Hardy–Weinberg equilibrium (Hardy 1908; Weinberg 1908), although genotypes were deleted and generated again until at least one boar was heterozygous. For each offspring, the “maternal” allele was inherited with probabilities π^* (allele A_1) and $1 - \pi^*$ (allele A_2), whereas the “paternal” allele from homozygous boars was directly assigned (*i.e.*, A_1 from A_1A_1 boars and A_2 from A_2A_2 boars); focusing on heterozygous boars, alleles A_1 and A_2 were inherited with probabilities $0.5 + \alpha^*$ and $0.5 - \alpha^*$, respectively.

Sampling by Monte Carlo Markov chains: Analyses were performed marker by marker both on simulated and on real genomic data. For each analysis, the Metropolis–Hastings algorithm (Metropolis *et al.* 1953; Hastings 1970) was used to sample from the marginal posterior distribution of π and α . A uniform proposal distribution between 0 and 1 was assumed for π , whereas boundaries moved to -0.5 and 0.5 for the uniform proposal distribution of α . A unique chain of 110,000 iterations was launched for each analysis and the first 10,000 rounds were discarded as burn-in (Raftery and Lewis 1992). Given the high autocorrelation between successive samples related to the Metropolis–Hastings method, a lag interval of 10 iterations was applied and 10,000 samples of model parameters were used to calculate the posterior distribution of each parameter, using the ergodic property of the chain (Gilks *et al.* 1996).

Implementation of the parametric bootstrap to correct for multiple testing: Bootstrap methods, introduced by Efron (1979), have become a routine methodology for approximating the distribution of a statistical quantity of interest, and they have been widely used for multiple-testing correction across multiple genetic markers in genome-wide association studies (Barrett *et al.* 2005; Purcell *et al.* 2007). Within this context, a parametric bootstrap approach was implemented to characterize the expected distribution of the BF specific to each SNP and under the absence of TRD effects. Focusing on a given SNP marker, 10,000 additional data sets were generated by assuming the genotype of each boar as specified in the original data set and characterizing the allelic frequency of the A_1 allele by using one after another, the 10,000 previously collected samples of π during the analysis of the real genomic data. Offspring genotypes were generated as for simulated data sets described above, with the only difference that the paternal allele from heterozygous boars was inherited with probabilities 0.5 (allele A_1) and 0.5 (allele A_2). Each simulated data set was analyzed by the multinomial TRD model and by launching a unique Monte Carlo Markov chain (MCMC) with 10,000 burn-in iterations and 100,000 sampling iterations. The BF was calculated and stored, and the bootstrapping procedure ended with the calculation of the 0.95th percentile of the bootstrap samples of the BF ($BF_{0.95}$). The BF obtained from real genomic data (BF_r) was rescaled by applying $BF_r/BF_{0.95}$ to allow direct comparison among different SNPs with different values of $BF_{0.95}$.

Analysis of gene enrichment in regions displaying TRD signals: We selected regions containing SNPs with $BF > 100$ to carry out gene enrichment analyses. Genes were retrieved by screening the Ensembl database (<http://www.ensembl.org>) with the BioMart data-mining tool. We took into consideration a window of 1 Mb around each selected region. Subsequently, we generated a list of 775 human genes that happened to be orthologous to the porcine loci identified in the previous step, and they were analyzed with two bioinformatic tools, (a) the Cytoscape plugin BiNGO (Maere *et al.* 2005) and (b) DAVID bioinformatic resources (Huang *et al.* 2009). Both software models identify significantly enriched Gene Ontology (GO) terms, among a given number of loci, by using the hypergeometric distribution test (Gold *et al.* 2007) and correcting for multiple testing on the basis of the Benjamini and Hochberg (1995) false discovery rate (FDR) ($\alpha = 0.05$) approach.

Results and Discussion

Evaluating the performance of the multinomial TRD model through simulated data sets

Although a related TRD model approach was successfully implemented and validated on genomic data from F_2 mouse crosses (Casellas *et al.* 2012), that parameterization could not be applied to other experimental designs such as half-sib families, a widely used design for quantitative trait loci detection (Georges *et al.* 1995; Hoeschele *et al.* 1997). Within this

Figure 3 (A and B) Detail of Manhattan plots in chromosome 3, illustrating the absolute value of the transmission ratio distortion (TRD) effects and the associated Bayes factors.

context, a two-parameter TRD model has been developed and adapted to the context of half-sib families when genotyping data become unavailable from one progenitor. Nevertheless, the statistical properties of this model must be validated on simulated data sets before going deep into the analysis of real populations. As shown in Figure 1, both the allele frequency in the ungenotyped ancestor (Figure 1A) and the sire-specific TRD effect (Figure 1B) evidenced a quasi-perfect match between simulated and predicted values. Moreover, both parameters were simulated under independent random processes and, as noted in Figure 1C, the analytical process did not evidence any kind of unexpected connectedness among the unknown parameters of the model during the sampling of their marginal posterior distributions. These plots highlight the suitability of the analytical model for addressing TRD phenomena when model assumptions are satisfied and assuming the same structure as that of the Duroc data set.

The statistical relevance of the TRD parameter was evaluated by adapting a specific BF test (Verdinelli and Wasserman 1995). Allele frequencies of the population of ungenotyped progenitors did not have relevant influences on the BF (Figure 1D). In contrast, the existence of a direct relationship between the magnitude of TRD and the BF was without doubt (Figure 1E). As expected, the BF favoring the model with nonnull TRD increased with the estimated TRD. Taking as reference the percentage of replicates with $BF > 100$, *i.e.*, decisive evidence of TRD according to Jeffreys' (1984) scale, it was almost null (0.5%) for $|\alpha| \leq 0.1$ and quickly increased to 23.0% ($0.1 < |\alpha| \leq 0.2$), 73.1% ($0.2 < |\alpha| \leq 0.3$), and 94.2% ($0.3 < |\alpha| \leq 0.4$). Indeed, 99.4% of replicates had a $BF > 100$ when $|\alpha| > 0.4$, and a lower BF was restricted to replicates with one or two heterozygous boars (see results on

Table 2 Porcine SNPs strongly associated with transmission ratio distortion (TRD) departures (Bayes factor, BF > 100)

SNP	Chr	Location, Mb	Heterozygous boars	Alleles	Bayes Factor	TRD effect ^a
H3GA0040758	14	69, 3	1	A/G	1922.8	0.459
DRGA0001952	1	189, 8	1	A/G	1918.2	0.467
H3GA0006745	2	27, 7	2	A/G	1823.0	0.408
INRA0046035	14	103, 7	1	A/G	1796.5	0.409
ASGA0035226	7	98, 4	2	A/G	1691.1	0.427
ASGA0051058	11	48, 6	1	A/G	1619.3	0.454
ALGA0008856	1	215, 0	3	C/A	1603.4	-0.200
H3GA0049325	17	44, 3	3	A/C	1576.8	0.258
ASGA0099747	6	167, 6	2	A/G	1574.4	0.469
MARC0001159	2	55, 0	1	A/C	1535.0	0.444
M1GA0007166	5	0, 9	4	A/G	1492.6	0.137
MARC0036680	4	78, 2	1	A/C	1481.5	0.463
H3GA0051376	23	1, 4	3	G/A	1478.0	-0.348
ALGA0073287	13	108, 7	4	A/G	1475.0	0.051
DRGA0017069	18	32, 1	4	A/G	1460.9	0.029
ASGA0047391	10	19, 7	1	G/A	1452.9	-0.421
ASGA0063762	14	60, 8	1	A/C	1442.3	0.430
INRA0056144	18	32, 1	4	G/A	1408.3	-0.024
MARC0003819	1	223, 2	4	G/A	1402.5	-0.045
M1GA0005374	4	2, 2	1	A/G	1382.8	0.423
ALGA0039334	7	21, 8	1	C/A	1370.7	-0.427
ALGA0043563	7	97, 6	1	A/G	1357.1	0.428
ASGA0080658	23	1, 7	3	A/G	1352.6	0.348
INRA0005966	1	189, 9	1	A/C	1349.8	0.467
ASGA0078645	18	0, 5	3	A/C	1272.9	0.294
CASI0008963	16	2, 4	1	A/G	1191.6	0.463
DRGA0001304	1	74, 4	2	A/G	1165.5	0.331
ALGA0099095	23	1, 3	3	A/G	1151.9	0.348
ALGA0032618	5	49, 9	2	G/A	1086.7	-0.294
DRGA0014537	14	124, 8	1	C/A	1075.6	-0.454
CASI0006810	17	39, 9	4	C/A	1066.4	-0.119
ALGA0051950	9	12, 6	1	A/G	1060.8	0.419
ALGA0020627	3	49, 0	4	G/A	1037.4	-0.148
ALGA0089288	16	8, 5	3	G/A	1032.6	-0.394
ASGA0044059	9	52, 4	1	A/G	979.0	0.458
ALGA0017095	2	78, 8	2	G/A	872.8	-0.312
DRGA0014376	14	102, 8	1	A/G	871.7	0.460
CASI0009330	6	156, 7	2	A/G	870.3	0.444
ASGA0044057	9	51, 9	1	C/A	856.6	-0.467
INRA0024805	7	34, 5	1	A/G	840.4	0.452
M1GA0008893	6	28, 7	1	G/A	827.3	-0.414
ASGA0076238	17	27, 8	1	A/G	766.0	0.440
MARC0109855	13	206, 0	1	A/G	728.4	0.411
ALGA0054275	9	51, 8	1	A/C	722.7	0.464
H3GA0015497	5	5, 5	1	A/G	689.1	0.449
H3GA0012656	4	39, 9	1	A/G	642.6	0.454
ALGA0008953	1	219, 0	4	G/A	640.2	-0.263
ALGA0099022	23	1, 0	1	A/G	630.0	0.380
MARC0102391	4	41, 0	1	A/G	608.3	0.445
INRA0045792	14	98, 0	1	A/G	565.4	0.440
MARC0013594	16	2, 6	2	A/G	557.2	0.448
H3GA0020344	7	24, 4	1	A/G	544.2	0.424
DRGA0007345	7	23, 5	1	G/A	520.6	-0.428
INRA0016720	4	109, 5	4	A/G	510.2	0.180
ALGA0039405	7	22, 8	1	A/C	509.2	0.394
BGIS0007825	7	26, 2	1	A/G	506.8	0.417
ASGA0060074	13	117, 3	2	G/A	485.7	-0.355
M1GA0020563	15	92, 1	1	A/G	485.6	0.456
MARC0110116	10	24, 3	1	A/G	460.0	0.417
ALGA0014423	2	44, 6	2	A/G	447.8	0.452
ALGA0106178	18	43, 4	3	A/C	417.3	0.288
ASGA0032522	7	33, 9	1	A/G	394.5	0.459

(continued)

Table 2, continued

SNP	Chr	Location, Mb	Heterozygous boars	Alleles	Bayes Factor	TRD effect ^a
ALGA0092079	16	48, 7	2	A/G	383.0	0.424
DIAS0003494	15	75, 9	2	G/A	364.3	-0.436
H3GA0032302	11	54, 4	4	G/A	321.8	-0.103
H3GA0013534	4	91, 7	1	A/C	316.9	0.444
ASGA0064956	14	91, 5	2	A/G	315.0	0.420
ALGA0043884	7	102, 6	4	A/G	304.2	0.040
ALGA0020630	3	49, 2	4	A/G	301.0	0.137
H3GA0056134	2	11, 0	4	A/G	268.4	0.021
MARC0104070	14	74, 5	1	G/A	263.4	-0.437
MARC0030335	1	215, 0	3	C/A	262.4	-0.199
ALGA0099025	23	1, 0	2	A/G	260.9	0.370
ASGA0015742	3	49, 2	4	A/G	229.6	0.149
DRGA0010291	10	8, 3	1	G/A	220.1	-0.421
DIAS0001058	2	151, 3	3	A/G	186.3	0.184
H3GA0052345	15	95, 1	2	G/A	162.3	-0.423
ALGA0020603	3	48, 5	4	A/G	159.7	0.147
ASGA0028023	6	7, 5	4	G/A	155.3	-0.241
ALGA0093687	17	17, 8	3	G/A	145.3	-0.169
ASGA0006696	1	223, 9	4	A/G	143.6	0.013
SIRI0001081	15	12, 5	1	A/G	134.8	0.395
H3GA0050595	18	12, 2	3	G/A	129.2	-0.216
ASGA0079268	18	12, 5	3	A/G	124.2	0.209

^a Positive and negative TRD estimates involve an increased transmission of the first and the second allele, respectively.

multiple-testing correction as an example of the influence of the number of heterozygous boars on the BF). The consistency of this BF approach suggested an appealing statistical behavior and provided new compelling evidence about the usefulness of Verdinelli and Wasserman's (1995) BF. This approach was previously adapted for testing heritability (García-Cortés *et al.* 2001; Casellas and Piedrafita 2006), QTL segregation (Varona *et al.* 2001; Casellas *et al.* 2008), inbreeding (Casellas *et al.* 2009), and gene expression (Casellas and Varona 2008), and its performance was compared with standard frequentist approaches, revealing a stable and consistent behavior (Varona *et al.* 2001; Casellas *et al.* 2008, 2012).

It is important to mention that we have assumed the absence of genotyping errors when evaluating the performance of this BF approach. As previously demonstrated by Meyer *et al.* (2012), transmission disequilibrium tests aimed to detect TRD are very sensitive to genotyping errors, a feature that can increase the rate of false positives. However, in the case of the half-sib design this problem is minimized because genotyping errors in the offspring generation have a low impact on TRD estimates due to the moderate-to-low contribution of each offspring in relation to remaining siblings. In addition, a genotyping error in a given boar must have a null contribution if incorrectly assumed as homozygous.

Implementation of the two-parameter multinomial model to detect paternal TRD in the pig genome

By using the two-parameter multinomial model, we have been able to report the genome-wide distribution of TRD in a livestock species. As a resource population, we have used a Duroc commercial line constituted by five half-sib families and fully described by Gallardo *et al.* (2008). Although 29,373

SNPs in the Duroc data set fulfilled the requirements in terms of call rate, minimum allele frequency, and mapping location (see *Materials and Methods*), 2809 of them (9.56%) were discarded because all Duroc boars displayed homozygous genotypes, thus preventing any further analysis of TRD. It is important to note that these SNPs segregated in the offspring generation (with a minimum allele frequency >0.05). In the sire generation, all chromosomes evidenced a similar pattern of distribution of the SNPs (Table 1), with most of them (>85%) being fixed in at least two boars. Indeed, only 147 SNPs (0.5%) were heterozygous in all five boars. Although the number of SNPs with at least one heterozygous boar clearly differed across chromosomes, they covered the whole genome with an average distance between successive SNPs of 95,297 bp and a maximum distance of 32,312,711 bp.

As shown in Figure 2A, most of the estimates of TRD effects concentrated around 0, without discarding extreme values close to the lower (-0.5; *i.e.*, null inheritance of the A_1 allele) and upper (0.5; *i.e.*, null inheritance of the A_2 allele) biological bounds. Statistical relevance was evaluated by the BF (Figure 2B) that, after appropriate rescaling for multiple-testing correction, revealed decisive evidence ($BF > 100$) of TRD in all chromosomes (84 SNPs; Figure 3, Table 2) with the only exception of chromosome 12 (Figure 2B). Before discussing TRD effects, it is important to highlight the impact of the parametric bootstrap for multiple-testing correction on the final estimates. Average $BF_{0.95}$ increased with the number of heterozygous boars, reaching values of 2.04 (one boar), 3.03 (two boars), 15.37 (three boars), 42.76 (four boars), 135.09 (five boars), and 256.30 (all boars). This procedure prevented a huge incidence of false positives, mainly when the number of heterozygous boars was large.

Table 3 Gene ontology enrichment analysis of the regions with highly significant transmission ratio distortion signals (Bayes factor > 100), obtained with the Cytoscape plugin BiNGO (Maere *et al.* 2005) and DAVID tools (Huang *et al.* 2009)

GO ID	Term	Genes	Fold enrichment	P-value ^a	FDR ^b
BiNGO analysis					
GO:0006334	Nucleosome assembly	18	5.155	9.96×10^{-4}	0.008
GO:0000786	Nucleosome	17	6.196	1.26×10^{-4}	0.011
GO:0031497	Chromatin assembly	18	4.921	5.90×10^{-4}	0.013
GO:0034728	Nucleosome organization	18	4.607	0.002	0.026
GO:0065004	Protein–DNA complex assembly	18	4.656	0.002	0.026
GO:0071103	DNA conformation change	22	3.780	0.003	0.032
GO:0006323	DNA packaging	20	4.043	0.003	0.033
DAVID analysis					
GO:0031497	Chromatin assembly	15	4.319	0.020	0.010
GO:0034728	Nucleosome organization	15	4.041	0.012	0.011
GO:0065004	Protein–DNA complex assembly	15	4.129	0.009	0.011
GO:0006334	Nucleosome assembly	15	4.473	0.013	0.013
GO:0000786	Nucleosome	14	5.420	0.003	0.045

^a Uncorrected P-value.

^b Corrected P-value by Benjamini and Hochberg (1995) false discovery rate (FDR).

The wide distribution of SNPs with significant TRD must be interpreted in the light that pigs were randomly sampled at an age of 6 weeks, and multiple biological mechanisms could contribute to these allelic departures up to this age. Sperm competition [*i.e.*, the preferential transmission of one of the two alleles to the zygote (Agulnik *et al.* 1990; Lyon 1991; Dyer *et al.* 2007; Labbe *et al.* 2013)] could have a role during the early stages of male gametogenesis and fertilization, although embryo or fetal failure (Wakasugi 1974) and differential viability during the first 6 weeks of life (Moore 2006) could also be responsible of TRD effects. Although heritabilities for embryo survival [0.04 (Gama *et al.* 1991)], farrowing survival [0.02–0.10 (Ibáñez-Escriche *et al.* 2009)], and early neonatal life survival [0.03 (Casellas *et al.* 2004)] are low in pigs, they provide compelling evidence that survival traits are regulated by genetic factors. This could unbalance the presence of paternal alleles from heterozygous boars in the offspring generation. It is important to note that all 84 SNPs showing BF_s > 100 clustered (<1 Mb) to other SNPs with BF > 10 (see Figure 3A as an example). This kind of pattern is highly relevant to discard unrealistic TRD due to genotyping errors. Note that, if present, genotyping errors must affect a single SNP, without influencing neighboring SNP. Within this context, our methodological approach must be discouraged if a single gene marker is available. The TRD model must be applied on dense marker data sets and statistically relevant markers should be corroborated by the linked pattern of companion genetic markers.

We have also investigated whether regions containing the most significant TRD signals are enriched in specific gene ontologies by using BiNGO and DAVID tools. These analyses revealed that the most enriched categories were quite broad (Table 3) and they included subjects such as nucleosome, nucleosome assembly, nucleosome organization, chromatin assembly, protein–DNA complex assembly, DNA conformation change, and DNA packaging. These functional categories included,

fundamentally, genes encoding histones. Paternal histones could play a key role not only during spermatogenesis but also during embryo development. Indeed, sperm cells conserve ~15% of paternal histones during the histone-to-protamine replacement phase that follows meiosis (Oliva 2006). It is well known that post-translational changes of histones have a major impact on chromatin structure and gene expression in the developing embryo (Rugg-Gunn *et al.* 2010). For instance, histone methylation is essential to establish epigenetic asymmetry in the mouse zygote (Arney *et al.* 2002) and regulates its pluripotency (Torres-Padilla *et al.* 2007). Variability near or within the histone genes might lead to the alteration of epigenetic marks or histone expression levels, thus compromising embryo survival. Another relevant category was constituted by genes that encode transcription factors that are deeply involved in embryogenesis such as *PAX5* that is essential for the genesis of the midbrain and cerebellum (Pfeffer *et al.* 2000), *HOXD* genes (*HOXD1*, *HOXD3*, *HOXD4*, *HOXD8*, *HOXD9*, *HOXD12*, and *HOXD13*) that are important for the innervation of the hindlimbs (Tarchini *et al.* 2005) and digit development (Delpretti *et al.* 2012), and *DMRT1* and *DMRT2* that are involved in somitogenesis (*DMRT2*) and sexual determination (Bellefroid *et al.* 2013). Notably, the loss of function of *DMRT1* has been recently reported as a potential cause for human spermatogenic failure (Lopes *et al.* 2013). Interestingly, the categories “mesodermal cell fate determination” and “skeletal system morphogenesis” were also enriched in TRD genomic regions, although at a low significance level.

We also compared our list of genes with those reported in the review of Huang *et al.* (2013) as affected by TRD. We found several coincidences that are worth mentioning: for instance, *DLG5*, which is necessary for the polarization of citron kinase in mitotic neural precursors (Chang *et al.* 2010). This locus is associated with Crohn’s disease and the susceptibility allele is transmission distorted among male progeny (Friedrichs *et al.* 2006). Another relevant gene is *MEGF9* that might have

a tumor suppressor function and showed a TRD pattern in human pedigrees from the Framingham Heart Study (Paterson *et al.* 2009). Finally, *SCA1* is associated with the occurrence of spinocerebellar ataxia type 1 and displays TRD in families affected by this neurodegenerative disease (Riess *et al.* 1997). Other genes of our list for which TRD has not been previously reported might also have a determinant role in gamete or embryo survival. For instance, we have detected 15 *TRIM* genes whose biological functions are related with cell death, tumorigenesis, cell cycle regulation, muscular physiology, and immunity (Napolitano and Meroni 2012). Interestingly, *TRIM32* has been associated with the occurrence of limb girdle muscular dystrophy type 2H and Bardet–Biedl type 11 syndromes (Napolitano and Meroni 2012). Another example is *SIRT1*, whose inactivation provokes genomic instability, impaired DNA damage response, and heterochromatin formation, and, as a consequence of these alterations, embryo lethality (Wang *et al.* 2008).

In summary, we have implemented a new method that allows us to detect TRD at a genome-wide scale and overcomes many of the limitations of preceding methods; *i.e.*, they could be applied only to a very restricted range of population structures and lacked appropriate tools to handle missing genotypes in the parental generation. Moreover, we have used this Bayesian method to screen the pig genome and, in doing so, we have demonstrated that TRD is considerably prevalent, as previously demonstrated in other species. Many of the regions with highly significant TRD signals contained genes displaying functions that are expected to influence offspring viability. These findings and the methodology we have implemented should provide the grounds to investigate in more detail the genetic architecture of gamete, zygote, and embryo survivability not only in model organisms, but also in many other species, such as livestock.

Acknowledgments

The research contract of J. Casellas was partially financed by Spain's Ministerio de Economía y Competitividad (program Ramón y Cajal, reference RYC-2009-04049). This research was funded by grant AGL2010-22208-C02/GAN (Ministerio de Economía y Competitividad).

Literature Cited

- Agulnik, S. I., A. I. Agulnik, and A. O. Ruvinsky, 1990 Meiotic drive in female mice heterozygous for the HSR insert on chromosome 1. *Genet. Res.* 55: 97–100.
- Arney, K. L., S. Bao, A. J. Bannister, T. Kouzarides, and M. A. Surani, 2002 Histone methylation defines epigenetic asymmetry in the mouse zygote. *Int. J. Dev. Biol.* 46: 317–320.
- Barrett, J., B. Fry, J. Maller, and M. Daly, 2005 Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21: 263–265.
- Bellefroid, E. J., L. Leclère, A. Saulnier, M. Keruzore, M. Sirakov *et al.*, 2013 Expanding roles for the evolutionary conserved Dmrt sex transcriptional regulators during embryogenesis. *Cell. Mol. Life Sci.* 70: 3829–3845.
- Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57: 289–300.
- Blasco, A., 2001 The Bayesian controversy in animal breeding. *J. Anim. Sci.* 79: 2023–2046.
- Casellas, J., 2007 Survival quantitative trait locus fine mapping by measuring and testing for Hardy-Weinberg and linkage disequilibrium. *Genetics* 176: 721–724.
- Casellas, J., and J. Piedrafit, 2006 Bayes factor for testing the genetic background of quantitative threshold traits. *J. Anim. Breed. Genet.* 123: 301–306.
- Casellas, J., and L. Varona, 2008 Between-groups within-gene heterogeneity of residual variances in microarray gene expression data. *BMC Genomics* 9: 319.
- Casellas, J., J. L. Noguera, L. Varona, A. Sánchez, M. Arqué *et al.*, 2004 Viability of Iberian $\times$ Meishan F₂ newborn pigs. II. Survival analysis up to weaning. *J. Anim. Sci.* 82: 1925–1930.
- Casellas, J., L. Varona, G. Muñoz, O. Ramírez, C. Barragán *et al.*, 2008 Empirical Bayes factor analyses of quantitative trait loci for gestation length in Iberian $\times$ Meishan F₂ sows. *Animal* 2: 177–183.
- Casellas, J., J. Piedrafit, G. Caja, and L. Varona, 2009 Analysis of founder-specific inbreeding depression on birth weight in Ripollés lambs. *J. Anim. Sci.* 87: 72–79.
- Casellas, J., R. J. Gualarte, C. R. Farber, L. Varona, M. Mehrabian *et al.*, 2012 Genome scans for transmission ratio distortion regions in mice. *Genetics* 191: 247–259.
- Chang, Y., O. Klezovitch, R. S. Walikonis, V. Vasioukhin, and J. J. LoTurco, 2010 Discs large 5 is required for polarization of citron kinase in mitotic neural precursors. *Cell Cycle* 9: 1990–1997.
- Delpretti, S., J. Zakany, and D. Duboule, 2012 A function for all posterior Hoxd genes during digit development? *Dev. Dyn.* 241: 792–802.
- Dyer, K. A., B. Charlesworth, and J. Jaenike, 2007 Chromosome-wide linkage disequilibrium as a consequence of meiotic drive. *Proc. Natl. Acad. Sci. USA* 104: 1587–1592.
- Efron, N., 1979 Bootstrap methods: another look at the jackknife. *Ann. Stat.* 7: 1–26.
- Evans, K., A. Fryer, C. Inglehearn, J. Duvall-Young, J. L. Whittaker *et al.*, 1994 Genetic linkage of cone-rod retinal dystrophy to chromosome 19q and evidence for segregation distortion. *Nat. Genet.* 6: 210–213.
- Friedrichs, F., S. Brescianini, V. Annese, A. Latiano, K. Berger *et al.*, 2006 Evidence of transmission ratio distortion of DLG5 R30Q variant in general and implication of an association with Crohn disease in men. *Hum. Genet.* 119: 305–311.
- Gallardo, D., R. N. Pena, M. Amills, L. Varona, O. Ramírez *et al.*, 2008 Mapping of quantitative trait loci for cholesterol, LDL, HDL, and triglyceride serum concentrations in pigs. *Physiol. Genomics* 35: 199–209.
- Gama, L. T., K. G. Boldman, and R. K. Johnson, 1991 Estimates of genetic parameters for direct and maternal effects on embryonic survival in swine. *J. Anim. Sci.* 69: 4801–4809.
- García-Cortés, L. A., C. Cabrillo, C. Moreno, and L. Varona, 2001 Hypothesis testing for the genetic background of quantitative traits. *Genet. Sel. Evol.* 33: 3–16.
- Gelfand, A., and A. Smith, 1990 Sampling based approaches to calculating marginal densities. *J. Am. Stat. Assoc.* 85: 398–409.
- Georges, M. D., M. Nielsen, A. Mackinnon, R. Mishra, R. Okimoto *et al.*, 1995 Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. *Genetics* 139: 907–920.
- Gilks, W. R., S. Richardson, and D. J. Spiegelhalter, 1996 *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London.
- Gold, D. L., K. R. Coombes, J. Wang, and B. Mallick, 2007 Enrichment analysis in high-throughput genomics—accounting for dependency in the NULL. *Brief. Bioinform.* 8: 71–77.

- Hardy, G. H., 1908 Mendelian proportions in a mixed population. *Science* 28: 49–50.
- Hastings, W. K., 1970 Monte Carlo sampling methods using Markov chains and their application. *Biometrika* 57: 97–109.
- Hoeschele, I., P. Uimari, F. E. Grignola, Q. Zhang, and K. M. Gage, 1997 Advances in statistical methods to map quantitative trait loci in outbred populations. *Genetics* 147: 1445–1457.
- Huang, D. W., B. Sherman, and R. A. Lempicki, 2009 Systematic and integrative analysis of large gene lists using DAVID bioinformatic resources. *Nat. Protoc.* 4: 44–57.
- Huang, L. O., A. Labbe, and C. Infante-Rivard, 2013 Transmission ratio distortion: a review of concept and implications for genetic association studies. *Hum. Genet.* 132: 245–263.
- Ibáñez-Escribá, N., L. Varona, J. Casellas, R. Quintanilla, and J. L. Noguera, 2009 Bayesian threshold analysis of direct and maternal genetic parameters for piglet mortality at farrowing in Large White, Landrace, and Pietrain populations. *J. Anim. Sci.* 87: 80–87.
- Jeffreys, H., 1984 *Theory of Probability*. Clarendon Press, Oxford.
- Kass, R. E., and A. E. Raftery, 1995 Bayes factors. *J. Am. Stat. Assoc.* 90: 773–795.
- Labbe, A., L. O. Huang, and C. Infante-Rivard, 2013 Transmission ratio distortion: a neglected phenomenon with many consequences in genetic analysis and population genetics, pp. 265–285 in *Epigenetics and Complex traits*, edited by A. K. Naumova, and C. M. T. Greenwood. Springer-Verlag, New York.
- Lavine, M., and M. J. Schervish, 1999 Bayes factors: what they are and what they are not. *Am. Stat.* 53: 119–122.
- Lopes, A. M., K. I. Aston, E. Thompson, F. Carvalho, J. Gonçalves *et al.*, 2013 Human spermatogenic failure purges deleterious mutation load from the autosomes and both sex chromosomes, including the gene DMRT1. *PLoS Genet.* 9: e1003349.
- Lyon, M. F., 1991 The genetic basis of transmission-ratio distortion and male sterility due to the t-complex. *Am. Nat.* 137: 349–358.
- Maere, S., K. Heymans, and M. Huiper, 2005 BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in biological networks. *Bioinformatics* 21: 3448–3449.
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, 1953 Equation of state calculations by fast computing machines. *J. Chem. Phys.* 21: 1087–1092.
- Meyer, W. K., B. Arberthuber, C. Ober, T. Ebner, I. Tiemann-Boege *et al.*, 2012 Evaluating the evidence for transmission distortion in human pedigrees. *Genetics* 191: 215–232.
- Montagutelli, X., R. Turner, and J. H. Nadeau, 1996 Epistatic control of non-Mendelian inheritance in mouse specific crosses. *Genetics* 143: 1739–1752.
- Moore, C. S., 2006 Postnatal lethality and cardiac anomalies in the Ts65Dn Down syndrome mouse model. *Mamm. Genome* 17: 1005–1012.
- Napolitano, L. M., and G. Meroni, 2012 TRIM family: pleiotropy and diversification through homomultimer and heteromultimer formation. *IUBMB Life* 64: 64–71.
- Oliva, R., 2006 Protamines and male infertility. *Hum. Reprod. Update* 12: 417–435.
- Pardo-Manuel de Villena, F., and C. Sapienza, 2001 Transmission ratio distortion in offspring of heterozygous female carriers of Robertsonian translocations. *Hum. Genet.* 108: 31–36.
- Pardo-Manuel de Villena, F., E. de la Casa-Esperón, T. L. Briscoe, and C. Sapienza, 2000 A genetic test to determine the origin of maternal transmission ratio distortion: meiotic drive at the mouse Om locus. *Genetics* 154: 333–342.
- Paterson, A. D., D. Waggott, A. Schillert, C. Infante-Rivard, S. B. Bull *et al.*, 2009 Transmission-ratio distortion in the Framingham Heart Study. *BMC Proc.* 3(Suppl. 7): S51.
- Paz-Miguel, J., F. Pardo-Manuel de Villena, P. Sánchez-Velasco, and F. Leyva-Cobián, 2001 H2-haplotype-dependent unequal transmission of the 17¹⁶ translocation chromosome from Ts65Dn females. *Mamm. Genome* 12: 83–85.
- Pfeffer, P. L., M. Bouchard, and M. Busslinger, 2000 Pax2 and homeodomain proteins cooperatively regulate a 435 bp enhancer of the mouse Pax5 gene at the midbrain boundary. *Development* 127: 1017–1028.
- Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira *et al.*, 2007 PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81: 559–575.
- Raftery, A. E., and S. M. Lewis, 1992 How many iterations in the Gibbs sampler? pp. 763–773 in *Bayesian Statistics IV*, edited by J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith. Oxford University Press, Oxford.
- Riess, O., J. T. Epplen, G. Amoiridis, H. Przuntek, and L. Schöls, 1997 Transmission ratio distortion of the mutant alleles in spinocerebellar ataxia. *Hum. Genet.* 99: 282–284.
- Rugg-Gunn, P. J., B. J. Cox, A. Ralston, and J. Rossant, 2010 Distinct histone modifications in stem cell lines and tissue lineages from the early mouse embryo. *Proc. Natl. Acad. Sci. USA* 107: 10783–10790.
- Shendure, J., J. A. Melo, K. Pociask, R. Derr, and 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) F₁ mice. *Mamm. Genome* 9: 812–815.
- Silver, L. M., 1993 The peculiar journey of a selfish chromosome: mouse t haplotypes and meiotic drive. *Trends Genet.* 9: 250–254.
- Tarchini, B., T. H. Huynh, G. A. Cox, and D. Duboule, 2005 HoxD cluster scanning deletions identify multiple defects leading to paralysis in the mouse mutant Ironside. *Genes Dev.* 19: 2862–2876.
- Torres-Padilla, M. E., D. E. Parfitt, T. Kouzarides, and M. Zernicka-Goetz, 2007 Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* 445: 214–218.
- Underkoffler, L. A., L. E. Mitchell, Z. S. Abdulali, J. N. Collins, and R. J. Oakey, 2005 Transmission ratio distortion in offspring of mouse heterozygous carriers of a 7.18 Robertsonian translocation. *Genetics* 169: 843–848.
- Varona, L., L. A. García-Cortés, and M. Pérez-Enciso, 2001 Bayes factors for detection of quantitative trait loci. *Genet. Sel. Evol.* 33: 133–152.
- Verdinelli, I., and L. Wasserman, 1995 Computing Bayes factors using a generalization of the Savage-Dickey density ratio. *J. Am. Stat. Assoc.* 90: 614–618.
- Vogl, C., and S. Xu, 2000 Multipoint mapping of viability and segregation distorting loci using molecular markers. *Genetics* 155: 1439–1447.
- Vorechovsky, I., A. D. Webster, A. Plebani, and L. Hammarstrom, 1999 Genetic linkage of IgA deficiency to the major histocompatibility complex: evidence for allele segregation distortion, parent-of-origin penetrance differences, and the role of anti-IgA antibodies in disease predisposition. *Am. J. Hum. Genet.* 64: 1096–1109.
- Wakasugi, N., 1974 A genetically determined incompatibility system between spermatozoa and eggs leading to embryonic death in mice. *J. Reprod. Fertil.* 41: 85–96.
- Wang, C. S., J. J. Rutledge, and D. Gianola, 1994 Bayesian analysis of mixed linear models via Gibbs sampling with an application to litter size in Iberian pigs. *Genet. Sel. Evol.* 26: 91–115.
- Wang, R. H., K. Sengupta, C. Li, H. S. Kim, L. Cao *et al.*, 2008 Impaired DNA damage response, genome instability, and tumorigenesis in SIRT1 mutant mice. *Cancer Cell* 14: 312–323.
- Weinberg, W., 1908 Über den Nachweis der Vererbung beim Menschen. *Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg* 64: 368–382.
- Wu, G., L. Hao, Z. Han, S. Gao, K. E. Latham *et al.*, 2005 Maternal transmission ratio distortion at the mouse Om locus results from meiotic drive at the second meiotic division. *Genetics* 170: 327–334.

Communicating editor: C. Sabatti

Appendix

When genotypes from sires and dams are fully available, sex-specific TRD effects can differentiate by appropriately modifying the model outlined above. As in previous parameterization, homozygous parents do not contribute relevant information to the analysis, where $p_{\text{sire}}(A_1) = p_{\text{dam}}(A_1) = 1$ and $p_{\text{sire}}(A_2) = p_{\text{dam}}(A_1) = 0$ for A_1A_1 homozygous parents [alternatively, $p_{\text{sire}}(A_1) = p_{\text{dam}}(A_1) = 0$ and $p_{\text{sire}}(A_2) = p_{\text{dam}}(A_2) = 1$ for A_2A_2 homozygous parents]. When one or both parents are heterozygous, TRD can be modeled by assuming

$$p_{\text{sire}}(A_1) = 0.5 + \alpha_s$$

$$p_{\text{dam}}(A_1) = 0.5 + \alpha_d$$

and

$$p_{\text{sire}}(A_2) = 0.5 - \alpha_s$$

$$p_{\text{dam}}(A_2) = 0.5 - \alpha_d,$$

where α_s and α_d are sire- and dam-specific TRD parameters with values defined within the parametric space between -0.5 and 0.5 . Under a Bayesian framework, this model generalizes to

$$p(\pi, \alpha_s, \alpha_d | \mathbf{y}) \propto p(\mathbf{y} | \pi, \alpha_s, \alpha_d) p(\alpha_s) p(\alpha_d),$$

where $\mathbf{y}$ is the column vector of genotypes of the offspring generation. In this case, *a priori* distributions for TRD parameters can be stated as

$$p(\alpha_s) = p(\alpha_d) = 1 \quad \text{if } \alpha \in [-1/2, 1/2] \text{ and } 0 \text{ otherwise,}$$

given our *a priori* lack of knowledge about the expected pattern of TRD phenomena.