A unified model for inclusive inheritance in livestock species

Ingrid DAVID*, Anne RICARD†,‡

*GenPhySE, INRA, Université de Toulouse, INPT, ENVT, F-31326 Castanet Tolosan, France
†GABI, INRA, AgroParisTech, Université Paris Saclay, Département Sciences du Vivant,
‡Institut Français du Cheval et de l’Equitation, Département Recherche et Innovation, F-61310 Exmes, France
Article Summary

We propose the “transmissibility” model that takes into account different sources of inheritance (genetic, epigenetic, microbiota, environmental inheritance) to better predict the transmissible potential of each animal in order to improve selection in livestock species. This mixed model differs from the animal model by estimating the path coefficients of inherited information from parent to offspring instead of set values of 0.5. Using simulations, we showed that the transmissibility model outperformed the animal model when the proportion of inheritance of non-genetic origin was high or when the sire and dam path coefficients were very different.
ABSTRACT

For years, animal selection in livestock species has been performed by selecting animals based on genetic inheritance. However, evolutionary studies have reported that non-genetic information that drives natural selection can also be inherited across generations (epigenetic, microbiota, environmental inheritance). In response to this finding, the concept of inclusive heritability, which combines all sources of information inherited across generations, was developed. To better predict the transmissible potential of each animal by taking into account these diverse sources of inheritance and improve selection in livestock species, we propose the “transmissibility model”. Similarly to the animal model, this model uses pedigree and phenotypic information to estimate variance components and predict the transmissible potential of an individual, but differs by estimating the path coefficients of inherited information from parent to offspring instead of using a set value of 0.5 for both the sire and the dam (additive genetic relationship matrix). We demonstrated the structural identifiability of the transmissibility model, and performed a practical identifiability and power study of the model. We also performed simulations to compare the performances of the animal and transmissibility models for estimating the covariances between relatives and predicting the transmissible potential under different combinations of sources of inheritance. The transmissibility model provided similar results to the animal model when inheritance was of genetic origin only, but outperformed the animal model for estimating the covariances between relatives and predicting the transmissible potential when the proportion of inheritance of non-genetic origin was high or when the sire and dam path coefficients were very different.
INTRODUCTION

For years, animal selection in livestock species has been performed by selecting animals to be the parents of the next generation based on their breeding values for the trait of interest (FISHER 1918). The theory behind this method is that, in addition to the common environmental effects shared by relatives, phenotypic resemblance between relatives is of genetic origin. Selection is therefore efficient because additive genetically-caused phenotypic variation is intergenerationally stable (MAMELI 2004). Indeed, DNA sequence is transmitted from parent to offspring (i.e. inherited) and, because genes play an important role in the development of all phenotypes, the transmission of this genetic information contributes to the phenotypic similarity between parents and offspring. However, it is very simplistic to consider that only the DNA sequence is transmitted from one generation to the next. Evolutionary studies have reported that non-genetic information that drives evolution and natural selection can also be inherited across generations (see (BONDURIANSKY et al. 2012; JABLONKA and LAMB 2008; O'DEA et al. 2016) for examples describing how non-genetic inheritance could be a key transient mechanism by which organisms adapt to rapid environmental change). In response to this finding, evolutionary biologists have developed the concept of inclusive or general heritability (DANCHIN et al. 2011; MAMELI 2004) that combines all sources of information inherited across generations. The existence of non-genetic inherited effects is one of the many explanations given to the missing heritability problem mainly reported in human (MANOLIO et al. 2009), i.e. genetic variants in Genome-Wide Association Studies cannot completely explain the heritability of traits. In addition to the possible complex genetic architecture of traits (epistasis, imprinting, etc.), it has been postulated that mechanisms unrelated to the DNA sequence may explain the differences between heritability measurements using pedigrees or SNPs (EICHLER et al. 2010; SANDOVAL-MOTTA et al. 2017). Many additional inheritance mechanisms have been
described and sometimes it is difficult to distinguish one from another (Rossiter 1996). For simplicity, we propose to classify these mechanisms into two main categories based on the medium of transmission: 1) the inherited information is transmitted across generations via a physical transmission support as is the case for epigenetic marks (epigenetic inheritance), metabolites and symbionts (microbiota inheritance), and 2) the information is not transmitted via a physical transmission support (environmental inheritance). Numerous studies have emphasized the importance of epigenetics in mammal and plant traits (Charlesworth et al. 2017; Quadrana and Colot 2016). In addition, there is now evidence that some epigenetic marks (histone modifications, mRNA and DNA methylation, etc.) that modulate gene expression are transgenerationally transmitted (Miska and Ferguson-Smith 2016; Van Otterdijk and Michels 2016). Among the other information that is physically transmitted across generations, one of the most important is probably the microbiota. The microbiota consists of the symbiotic microbial cells (bacteria, archaea, viruses and eukaryotic microbes) that reside in and on the bodies of animals and plants (Shreiner et al. 2015; Turner et al. 2013). These microbes have tremendous potential to impact the physiology, both in health and in disease, of their host (Marchesi et al. 2015; Sommer and Backhed 2013). The vertical transmission of the microbiota has been described in different species (Abecia et al. 2007; Cankar et al. 2005; Ley et al. 2006; Sylvain and Derome 2017) confirming that the microbiome is part of the information transmitted across generations (Sandoval-Motta et al. 2017). The term environmental inheritance that we propose here refers to other sources of information transmitted across generations that do not rely on a physical transmission medium, that is information that passes from one generation to another via learning mechanisms (behavioral (Jablonska and Lamb 2014), cultural inheritance (Danchin et al. 2011; Feldman and Cavalli-Sforza 1975)) or transmission of environmental conditions (ecological inheritance). All these phenomena have been explored in evolutionary studies to
understand the mechanisms that drive natural selection (AMELI 2004). Even if discussed (GODDARD and WHITELAW 2014), they are not taken into account as of yet in the context of selection in livestock species. Nonetheless if, as stated in evolution studies, non-genetic inheritance is much more important than genetic inheritance (AMELI 2004), then not considering it for selection could mean that the maximal progress for the selected population may never be reached or, even worse, that the transmissible potential of an individual be misevaluated because the model (animal model) used to study the covariances between relatives is inappropriate.

The aim of this study is to present a model for estimating the inclusive heritability of traits in livestock species and predicting the transmissible potential of individuals based on genealogical and phenotypic information. The first part of the paper gives a brief overview of the mathematical models proposed in the literature for modeling the different sources of non-genetic inheritance and describes our transmissibility model for inclusive heritability. The second part reviews the properties of the transmissibility model: demonstration of the structural identifiability of the model, comparison of the practical identifiability of the transmissibility model with that of a model aiming to disentangle genetic and non-genetic sources of inheritance, and power analysis of the transmissibility model. The third part compares, using simulation studies, the performances of the animal model and the transmissibility model for selection (i.e. for estimating covariances between relatives and predicting the transmissible potential of individuals).

MATERIAL AND METHOD
Overview of the mathematical models for the different sources of inclusive inheritance in a livestock context

The path coefficient diagram describing the transmission of the different inherited factors in livestock species is provided in figure 1. Let’s consider an offspring $i$ born from sire $s$ and dam $d$.

**Genetic inherited effects:** an animal receives half of its DNA from its father and half from its mother. Following the Mendel theory, the additive genetic value $a$ (or breeding value) of animal $i$ is then $a_i = 0.5(a_d + a_s) + \varepsilon_{a,i}$, where $\varepsilon_{a,i}$ is the Mendelian sampling (i.e. the deviation of the breeding value of animal $i$ from the average breeding value for both parents (FALCONER 1974)), considering no inbreeding, $\varepsilon_{a,i} \sim N \left(0, \frac{1}{2} \sigma_a^2\right)$, with $\sigma_a^2$ the additive genetic variance. Thus $a \sim MVN(0, A\sigma_a^2)$ with $A$ the numerator relationship matrix.

**Epigenetic inherited effects:** Tal et al. (2010) proposed a mathematical model for the transmission of epigenetic marks used later by Varona et al. (2015) and Paiva et al. (2018a) in the context of animal breeding. The model is built on the idea that during the vertical transmission of epigenetic marks from mother to offspring and from father to offspring, an unknown proportion of them is lost (thus not transmitted to the next generation). This unknown probability of changing the epigenetic state during gametogenesis and/or early development is called the reset coefficient $v$ and can be different for the sire and the dam ($v_s$ and $v_d$, respectively). Consequently, the model proposed for the epigenetic value ($epi$) of an individual is: $epi_i = \lambda_s epi_s + \lambda_d epi_d + \varepsilon_{epi,i}$ with $\lambda_s = 0.5(1 - v_s), \lambda_d = 0.5(1 - v_d)$ and (under the assumption that the variance $\sigma_{epi}^2$ of transgenerational epigenetic effects is constant across generations) $\varepsilon_{epi,i} \sim N \left(0, \frac{1}{4}(1 - v_s)^2 - \frac{1}{4}(1 - v_d)^2\right) \sigma_{epi}^2$. Thus $epi \sim MVN(0, T\sigma_{epi}^2)$ with $T$ the matrix of epigenetic relationships between individuals.
Microbiota inherited effects: to our knowledge, no mathematical model for the transmission of the microbiota has been proposed in the literature. However, there is evidence that part of the microbiota received by an offspring is vertically transmitted and another part is horizontally transmitted from environmental sources including comates (Bright and Bulgheresi 2010). The vertical transmission reflects the inheritance of symbionts from the mother and rarely from both parents. This is particularly true in livestock species, where generally no direct contact between sires and offspring is possible. Consequently the model for the transmission of the “microbiota value” ($m_{i}$) of an individual following these rules could be: $m_{i} = \gamma_{m}d + \sum_{j \in I} \gamma_{j}'m_{j} + \varepsilon_{m_{i}}$, where $\gamma$ is the proportion of microbiota inherited from the mother, $\gamma_{j}'$ is the proportion of microbiota transmitted by individual $j$ sharing the same environment $I_{i}$ as the target individual $i$, and $\varepsilon_{m_{i}} \sim N(0, (1 - \gamma^2 - \sum_{j \in I} \gamma_{j}'^2)\sigma^{2}_{m_{i}})$. Then $m_{i} \sim MVN(0, B\sigma^{2}_{m_{i}})$, where $B$ is the microbial relationship matrix.

Cultural and behavioral inherited effects: culture is a system of inheritance via social learning; it has been described to be transmitted through vertical, horizontal and oblique pathways. Nearly 50 years ago, Feldman and Cavalli-Sforza (1975) proposed a model for cultural inheritance. In their model, the cultural value of an individual is defined as the sum of a fraction of the cultural values of its parents and of animals having an influence on the target animal (transmission through nonparental elders and animals of the same age group). Due to the rearing conditions used in livestock species, most often among the two parents only the dam can have a parental cultural influence on her offspring (with several exceptions in meat sheep, meat cattle species). Thus, the model can be written as: $c_{i} = \kappa c_{d} + \sum_{j \in I} \kappa_{j}'c_{j} + \varepsilon_{c_{i}}$ ($0 \leq \kappa \leq 1, 0 \leq \kappa_{j}' \leq 1, \kappa + \sum_{j \in I} \kappa_{j}' = 1$) which is a model equivalent to that proposed for microbiota inheritance. Consequently, it will be difficult to distinguish cultural and microbiota values in practice without a specific experimental design or without additional information on the microbiota and/or the behavior of animals. We thus
propose, in the context of selection in livestock species, to combine them into a single term:
the single-parent transmitted value \( s_{pi} = \delta s_{pd} + \sum_{j \in I} \delta'_j s_{pj} + \epsilon_{sp,i} \). In addition, for
simplicity, we will consider that individuals \( j \) sharing the same environment as the target
individual \( i \) throughout its life are randomly sampled in the population. The model describing
the vertical transmission of the single parent transmitted value can therefore be simplified to:
\( s_{pi} = \delta s_{pd} + \epsilon_{sp,i}, \) \( 0 \leq \delta \leq 1 \) with \( \epsilon_{sp,i} \sim N\left(0, (1 - \delta^2)\sigma^2_{sp}\right) \). It should be noted that this
model is an oversimplification of the transmission of culture/behavior/microbiota, the purpose
of which is solely to underline the difficulty of separating genetic from non-genetic effects,
even in a simple case.

**Ecological inherited effects:** In livestock species, conversely to wild species, animals do not
choose their environment; it depends entirely on the farming system used. Consequently, we
consider that there is no ecological inheritance in livestock species.

Therefore, the different sources of inheritance in livestock species can be modeled using three
different models: a model in which both parents influence the inherited value with a known
coefficient of transmission (genetic effects), with a unknown coefficient of transmission
(epigenetic effects) and a third model in which, of the two parents, only the dam vertically
transmits the inherited information (microbiota, cultural and behavioral inheritance).

**Model proposed for inclusive inheritance in livestock species**
The phenotype of an individual is the result of the combination of all these inherited factors
and other environmental effects. To our knowledge, no studies have taken into account all of
these sources of variation in a single model in order to explain phenotypic variability.
However, several have investigated two inherited factors at a time (DANCHIN et al. 2013; TAL
et al. 2010; VARONA et al. 2015) by considering additivity of the effects of the different
inherited factors. Under this assumption, the phenotype of an individual considering all the
inherited factors is:
\[ y_i = x_i \beta + a_i + epi_i + sp_i + e_i \]  
(Eq1)

where \( \beta \) is the vector of fixed effects and \( x_i \) the vector that links the fixed effect to the observation of animal \( i \); \( a_i, epi_i \) and \( sp_i \) are the additive genetic, epigenetic and single-parent transmitted values of animal \( i \), respectively, that follow the aforementioned models, \( e_i \) is the residual with \( e \sim \text{MVN}(0, I\sigma^2_e) \). Given the laws of transmission of the different inherited factors; an animal \( i \) will transmit \( \frac{1}{2} a_i + \frac{1}{2} (1 - v_s) epi_i \) to its offspring if \( i \) is a male (=sire transmissible potential of individual \( i \)), while it will transmit \( \frac{1}{2} a_i + \frac{1}{2} (1 - v_d) epi_i + \delta sp_i \) to its offspring if \( i \) is a female (=dam transmissible potential of individual \( i \)). In this model, the dam transmissibility (the dam-offspring regression) is \( b_{OD} = \frac{0.5\sigma_a^2 + 0.5(1 - v_d)\sigma_{epi}^2 + \delta \sigma_{sp}^2}{\sigma_a^2 + \sigma_{epi}^2 + \sigma_{sp}^2 + \sigma_e^2} \) and the sire transmissibility (the sire-offspring regression) is \( b_{OS} = \frac{0.5\sigma_a^2 + 0.5(1 - v_s)\sigma_{epi}^2}{\sigma_a^2 + \sigma_{epi}^2 + \sigma_{sp}^2 + \sigma_e^2} \). These transmissibilities are useful to compute the response to selection: \( R = S_m b_{OS} + S_f b_{OD} \) where \( S_f \) and \( S_m \) are the female and male selection differentials, respectively. Theoretically, the parameters of Eq1 are identifiable (see appendix 1 for the demonstration). However, a very large number of observations and a particular population structure would be necessary to estimate them in practice. For these reasons we propose to combine the different transmissible potentials of an individual \( i \) into a single “transmissible value” \( t_i \) with the following model of transmission:

\[ t_i = \omega_s t_s + \omega_d t_d + \epsilon_{t,i}, \epsilon_{t,i} \sim N(0, (1 - \omega_s^2 - \omega_d^2)\sigma_e^2) \]

where \( \omega_s, \omega_d \) are the path coefficients of transmission from the sire and the dam, respectively, that follow the following constraints: \( 0 \leq \omega_s \leq 1, 0 \leq \omega_d \leq 1, 0 \leq \omega_s + \omega_d \leq 1 \). Thus, in this model, the 2 path coefficients of transmission can be equal to 0.5 as in the traditional animal model, lower than 0.5 in agreement with the epigenetic model or one coefficient can
be higher than 0.5 in agreement with single parent inheritance, which is of particular interest for the dam side. The corresponding model explaining phenotypic variation is:

\[ y_i = x_i \beta + t_i + e_i \]

with \( t \sim MVN(0, M \sigma_t^2) \) and \( e \sim MVN(0, I \sigma_e^2) \) where \( M \) is the matrix of transmission between individuals (transmission relationship matrix) and \( I \) the identity matrix. We call this model the “transmissibility model”. The transmission relationship matrix is a symmetric matrix with 1’s on the diagonal and \( r_{ij} \) as off-diagonal entries. For the case of 2 animals \( i, j \) with \( n \) common ancestor \( l \) \( r_{ij} = \sum_{l=1}^{n} r_{ij,l} \) with \( r_{ij,l} = \omega_s k_{ij,l} + \omega_d q_{ij,l} \), \( k_{ij,l} = k_{il} + k_{jl}, q_{ij,l} = q_{il} + q_{jl} \), where \( k_{il}, q_{il} \) are the number of sire and dam transmissions between ancestor \( l \) and animal \( i \), respectively. The transmission relationship matrix can be computed through its inverse that can be easily obtained by the following decomposition \( M^{-1} = L'D^{-1}L \) where \( D \) is a diagonal matrix with variances of \( \varepsilon_t \) relative to \( \sigma_t^2 \) as components (i.e. \( 1 - \omega_s^2 - \omega_d^2, (1 - \omega_d^2) \) for animals of unknown sire, \( (1 - \omega_s^2) \) for animals of unknown dam and 1 for animals for which both parents are unknown), \( L \) is a lower triangular matrix with 1’s on the diagonal and the negatives of the sire and dam coefficients of transmission as off-diagonal entries.

In the transmissibility model, the proportion of transmitted variance is \( \tau^2 = \frac{\sigma_t^2}{\sigma_t^2 + \sigma_e^2} \), the sire-offspring regression and the dam-offspring regression (sire and dam transmissibilities) are \( b_{OS} = \omega_s \tau^2 \) and \( b_{OD} = \omega_d \tau^2 \), respectively. The response to selection is thus \( R = S_m \omega_s \tau^2 + S_f \omega_d \tau^2 \) where \( S_f \) and \( S_m \) are the female and male selection differentials, respectively.

The parameters of the transmissibility model can be estimated with the restricted maximum likelihood method (REML) using ASReml and the OWN Fortran program that we have developed. This OWN fortran program is freely available on the zenodo website:
10.5281/zenodo.1487869, it can be used to fit a transmissibility model with different or equal path coefficients of transmission for the sire and the dam.

In the transmissibility model, variance $\sigma_I^2$ of the transmissible potential, sire and dam path coefficients of transmission ($\omega_s$, $\omega_d$) estimates depend on the relative proportion of variance and path coefficients of each inherited factor. Let’s describe different schematic cases that help understanding the behavior of the transmissibility model. For simplicity, we consider that phenotypes are derived from Eq1 with sire and dam epigenetic path coefficients of transmission equal to 0.25 and 0.40, respectively, single parent path coefficient of transmission equals to 0.9, total variance equals to 1 and $\sigma_a^2 = \sigma_{epi}^2 + \sigma_{epi}^2 + \sigma_{epi}^2 + \sigma_e^2 = 0.6$ in all the examples. 1/ If the inheritance is mainly of genetic origin ($\sigma_a^2 = 0.5, \sigma_{epi}^2 = 0.05, \sigma_{epi}^2 = 0.05$), the sire and dam path coefficient of transmission estimated with the transmissibility model will be close to 0.5 but slightly higher for the dam side ($\omega_s = 0.46, \omega_d = 0.54$) and $\frac{\sigma_I^2}{\sigma_a^2 + \sigma_{epi}^2 + \sigma_{epi}^2 + \sigma_e^2} = 0.58$ which is close but lower than 0.6 ; 2/ if the three inherited factors equally participate to the inheritance ($\sigma_a^2 = \sigma_{epi}^2 = \sigma_{epi}^2 = 0.2$) the dam path coefficient of transmission obtained with the transmissibility model is higher than the sire one (0.68 versus 0.26) and $\frac{\sigma_I^2}{\sigma_a^2 + \sigma_{epi}^2 + \sigma_{epi}^2 + \sigma_e^2}$ decreases to 0.53. Finally, 3/ if the inheritance is mainly of single-parent origin ($\sigma_a^2 = 0.05, \sigma_{epi}^2 = 0.05, \sigma_{epi}^2 = 0.5$) then the disequilibrium between sire and dam path coefficient of transmission is high ($\omega_s = 0.05, \omega_d = 0.87$) and $\frac{\sigma_I^2}{\sigma_a^2 + \sigma_{epi}^2 + \sigma_{epi}^2 + \sigma_e^2}$ is close to 0.6 (0.57).

Additional expected estimates in the transmissibility model compared to parameters of the true model of inheritance are provided in figures S1, S2 and S3.

Structural, practical identifiability and power analysis

The parameters of the transmissibility model are structurally identifiable (see appendix 2 for the demonstration). Their practical identifiability was assessed by using methods proposed by
Visscher and Goddard (2015), Raffa and Thompson (2016) and Raue et al. (2009). Detailed description of this practical identifiability study is provided in appendix 3. Briefly, the practical identifiability study is as follows: the expected likelihood ratio test (ELRT) comparing the true to the estimation model is literally expressed as a function of the parameters and the structure of the population (i.e. number and size of families with determined relationships). To do so, a simplified population structure with homogeneous families is considered. Then, the profile likelihood (PL) which reflects the minimal value of ELRT is computed for each parameter to estimate. A close inspection of changes in the PL with the value of the parameter to estimate enables the detection of directions where the likelihood flattens out. In such situations, the parameter is considered as practically non-identifiable. In addition, the power for testing a given hypothesis can be derived from the ELRT as described by Raffa and Thompson (2016).

We investigated the practical identifiability of the transmissibility model and of a model that aims at distinguishing genetic from non-genetic transmitted effects. For simplicity, we considered for this practical identifiability study that non-genetic effects were of epigenetic origin only, i.e. the true model is $y_i = a_i + epi_i + e_i$ and the two models of estimation are $y_i = a_i + epi_i + e_i$ (mod1) and $y_i = t_i + e_i$. To ensure structural identifiability in mod1, six different types of family relationship are sufficient: sire – offspring, dam – offspring, paternal half-sibs, maternal half-sibs, uncle-nephew and aunt-nephew. Those relationships are also sufficient to ensure the structural identifiability of the transmissibility model. We thus investigated the practical identifiability of both mod1 and the transmissibility model for three different pedigree sizes (small, medium and large) that differed according to the number of family relationships as described in table 1. Practical identifiability can be evaluated for an infinite number of combinations of the true genetic, epigenetic variances and reset coefficients. We considered, as an illustration, the case of a trait of moderate heritability (0.2),
equal genetic and epigenetic variances and six different values of the sire and dam epigenetic
coefficient of transmission combination ($\lambda_s, \lambda_d$: 0.2 0.2, 0.2 0.3, 0.2 0.4, 0.3 0.3, 0.3 0.4 and
0.4 0.4).

We then computed the power of the transmissibility model to detect non-genetic inheritance
(i.e. at least one of the two path coefficients of transmission differs from 0.5) while the true
model of transmission was $y_i = a_i + epi_i + e_i$ using the three aforementioned pedigree
structures and considering different combinations of the genetic, epigenetic variances (i.e.
\[
\frac{\sigma^2_a + \sigma^2_{epi}}{\sigma^2_a + \sigma^2_{epi} + \sigma^2_e} = 0.2, 0.4 \text{ or } 0.6; \quad \frac{\sigma^2_a + \sigma^2_{epi}}{\sigma^2_a + \sigma^2_{epi}} = 0.3, 0.5 \text{ or } 0.7
\]
in the case of equal sire and dam path coefficients of transmission. Finally, we also computed, for medium pedigree sizes, the power
to detect non-genetic inheritance when the sire and dam have different epigenetic path
coefficients of transmission in the true model and different path coefficients of transmission in
the transmissibility model used for the estimation.

Validation of the transmissibility model using simulations

Phenotype simulation: We performed a simulation study to evaluate the ability of the
transmissibility model in complex pedigrees to: 1) detect non-genetic inheritance, 2) estimate
regression coefficients for different types of relatives, and 3) predict the transmissible values
of individuals. In the two latter cases, the performances of the transmissibility model were
compared with those of the animal model.

Datasets consisted of a four-generation population. The first generation comprised 25 sires
and 100 dams that were randomly mated to give birth to 800 offspring (8 offspring per dam,
sex ratio = 1/2). Among the progeny, 25 males and 100 females were sampled randomly to be
the parents of the next generation. Same process was repeated for each generation. The final
population comprised 2,525 individuals.
The genetic, epigenetic and single-parent transmitted values for each animal $i$ in the first generation were sampled from the following distributions:

$$a_i \sim N(0, \sigma_{a}^2), \ epi_i \sim N(0, \sigma_{epi}^2), \ sp_i \sim N(0, \sigma_{sp}^2).$$

Then the different inherited values of animal $j$ born to sire $s$ and dam $d$ in the subsequent generations were obtained by sampling from:

$$a_j \sim N\left(\frac{1}{2} a_s + \frac{1}{2} a_d, \frac{\sigma_{a}^2}{2}\right),$$

$$epi_j \sim N\left(\frac{1}{2} (1 - v_s) epi_s + \frac{1}{2} (1 - v_d) epi_d, \left[1 - \frac{1}{4} (1 - v_s)^2 - \frac{1}{4} (1 - v_d)^2\right] \sigma_{epi}^2\right),$$

$$sp_j \sim N\left(\delta sp_d, (1 - \delta^2) \sigma_{sp}^2\right)$$

Phenotypes were then generated for all the individuals of the first three generations and for $\frac{1}{4}$ of the animals of the last generation (randomly sampled) by:

$$y \sim \mathcal{MVN}(X \beta + a + epi + sp, I \sigma_{e}^2)$$

where $\beta = [10 \ 20 \ 30]^T$ is the vector of the effect of one factor with 3 levels. The level of the factor is randomly sampled for each individual from a uniform distribution. For the remaining $\frac{3}{4}$ of the animals of the last generation (600), no phenotypes were generated. These individuals will be used to evaluate the ability of the models to predict the transmissible potential of an individual based on information from relatives only (no phenotypic information). In addition, among these individuals, 400 were considered as having only one known parent; only the sire was known for 200 and only the dam was known for 200, randomly sampled. To do so, we modified the pedigree to change the ID of the sire or the dam to unknown.

We simulated the four datasets described in Table 2. The sum of the sire and dam transmissibility was the same in all datasets but datasets differed in the ratio of heritable variance to phenotypic variance $\frac{\sigma_{a}^2 + \sigma_{epi}^2 + \sigma_{sp}^2}{\sigma_{a}^2 + \sigma_{epi}^2 + \sigma_{sp}^2 + \sigma_{e}^2}$, in the proportion of genetic origin for the heritable variance and in the disequilibrium between sire and dam transmissibility. In dataset 1, the heritable variance was of genetic origin only: sire and dam transmissibilities were equal ($b_{OS}/b_{OD} = 1$). In datasets 2 and 3, sire and dam transmissibilities were different.
The ratio of heritable variance to phenotypic variance and the ratio of sire transmissibility to dam transmissibility was the same for the 2 datasets 

\[
\frac{\sigma_a^2 + \sigma_{ep}^2 + \sigma_{ep}^2}{\sigma_a^2 + \sigma_{ep}^2 + \sigma_{ep}^2 + \sigma^2_e} = 0.409, \frac{b_{OS}}{b_{OD}} = 0.78
\]

but the proportion of genetic origin for the heritable variance differed (63% in dataset 2 and 10% in dataset 3). Finally, the last dataset (dataset 4) had, as in dataset 3, a low proportion of genetic origin for the transmitted variance (8%) associated with a high disequilibrium between sire and dam transmissibility 

\(\frac{b_{OS}}{b_{OD}} = 0.37\).

Parameter estimation: Two models were used to analyze the phenotypes: the animal model 

\(y_i = x_i\beta + a_i + e_i\) and the transmissibility model \(y_i = x_i\beta + t_i + e_i\).

For each dataset type, 100 independent replicates were simulated.

Evaluation of the performance of the transmissibility model

The performance of the transmissibility model was assessed by:

1) Its ability to detect non-genetic inheritance. The animal model is a special case of the transmissibility model for which sire and dam coefficients of transmission are fixed to 0.5; it is thus nested in the transmissibility model. After model convergence, the null hypothesis H0 “sire and dam coefficients of transmission are equal to 0.5” was tested against the H1 hypothesis “almost one of the coefficients of transmission (sire or dam) differs from 0.5” by performing a likelihood ratio test of size 5% comparing the transmissibility model with the animal model (mixture \(\chi^2_1, \chi^2_2\)). Rejecting the null hypothesis H0 permits to conclude that the underlying model is not purely additive genetic. The realized type I error is the number of replicates over 100 for which H0 was rejected when analyzing dataset 1, the power is the number of replicates over 100 for which H0 was rejected when analyzing datasets 2 to 4.

2) Its ability to estimate the regression coefficient between different types of relatives. The regression coefficients obtained with the 2 models of estimation were compared
with the expected values assuming the hypothesis of the simulation model. Special
attention was paid to the sire-offspring and dam-offspring regression coefficients (sire
and dam transmissibilities). In addition, sire and dam path coefficients of transmission
and the ratio of transmitted variance to phenotypic variance obtained with the genetic
and transmissibility models were examined in detail to better comprehend how these
models behave with data simulated according to a different model (Eq1).

3) Its ability to predict the transmissible potential of individuals. The sire and dam
“transmissible values” obtained with the transmissibility model ($\tilde{\omega}_s^t$ for sires, $\tilde{\omega}_d^t$ for
dams) and the additive genetic values predicted by the animal model ($0.5\hat{a}$ for sires
and dams) were compared with the simulated transmissible values (weighted sum of
the different sources of inheritance i.e. $0.5a + 0.5(1 - v_s)epi$ for sires, $0.5a +
0.5(1 - v_d)epi + \delta sp$ for dams) for four different groups of animals: all animals with
a phenotype, progeny without a phenotype and both parents known, progeny without a
phenotype and unknown dam, progeny without a phenotype and unknown sire.
Comparison was based on the correlation between the estimated and simulated
transmissible value, and the bias was quantified using the coefficient of regression of
the simulated to the estimated transmissible value.

Data availability

The authors state that all information necessary for performing the simulation study presented
in the manuscript are represented fully within the manuscript. Code source and tutorial to
apply the transmissibility model are available on the zenodo website
10.5281/zenodo.1487869. Additional figures describing expected transmissibility model
estimates in comparison with the true model of inheritance are available on Figshare. Figure
S1 provides the expected estimates in the transmissibility model while the true model is

\[ y_i = x_i \beta + a_i + e_{pi} + e_i. \]

Figure S2 provides the expected estimates in the transmissibility model while the true model is \( y_i = x_i \beta + a_i + e_{pi} + s_{pi} + e_i. \) Figure S3 compares the expected error term variance in the transmissibility model to the mendelian sampling variance.
RESULTS

Practical identifiability and power analysis

The profile likelihoods of the transmissibility model and the model designed to dissociate genetic from non-genetic transmission (i.e. $y_i = a_i + epi_i + e_i$) are shown in Figure 2 for six true values of the sire/dam reset coefficient. The shape of the curve for each parameter to estimate illustrates its practical identifiability. Structural non-identifiable parameters are characterized by a flat profile likelihood that can have a minimum but do not exceed a threshold for increasing and/or decreasing values of the parameter. On one hand, the profile likelihoods of parameters of the model that aims at dissociating genetic from epigenetic effects were, in the case of equal sire and dam epigenetic path coefficient of transmission ($\lambda = 0.5(1 - v_d) = 0.5(1 - v_s)$), flat for increasing values of the relative importance of the transmitted variance ($H_0^2 = \frac{\sigma_a^2 + \sigma_{e_{pi}}^2}{\sigma_a^2 + \sigma_{e_{pi}}^2 + \sigma_e^2}$) and decreasing values of the relative importance of the genetic variance ($r_0 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{e_{pi}}^2}$), and for both increasing and decreasing values of the epigenetic path coefficient of transmission whatever the pedigree size. Difference in the sire and dam epigenetic path coefficient of transmission helps the practical identifiability of the parameters. However, even for large pedigree sizes, the profile likelihood of the genetic variance had difficulties exceeding the $\chi^2$ threshold value. On the other hand, even for small pedigree sizes, the profile likelihood of the proportion of transmitted variance ($\frac{\sigma_t^2}{\sigma_t^2 + \sigma_e^2}$) and path coefficients of the transmissibility model increased substantially in both upwards and downwards directions. These results indicate that all parameters of the transmissibility model are practically identifiable even for small pedigree size while those of the model designed to dissociate genetic from non-genetic transmission effects are not.

The power of the transmissibility model to detect non-genetic inheritance, under the assumption that sire and dam path coefficients of transmission are equal, as a function of the
epigenetic path coefficient of transmission is provided in figure 3 for different combinations of the true genetic and epigenetic variances. Its power to detect non-genetic inheritance increased with pedigree size, proportion of transmitted variance and relative importance of epigenetic variance. For small pedigree sizes, the non-genetic inheritance could only be detected in favorable cases: epigenetic path coefficient of transmission above a certain threshold but significantly less than 0.5 ($0.1 < \lambda < 0.4$), proportion of transmitted variance $> 0.4$ and relative importance of epigenetic variance $> 0.5$. For medium pedigree sizes, the model’s power to detect non-genetic inheritance was more than 80% for many combinations of the true parameters. For large pedigree sizes, the power was close to 100% for all combinations. However, if the true epigenetic path coefficient of transmission was close to 0.5, the model was unable to detect non-genetic inheritance, regardless of the pedigree size, because epigenetic transmission was confounded with genetic inheritance (same mode of transmission). The power of the transmissibility model to detect non-genetic inheritance considering different path coefficients of transmission for the sire and the dam is provided in figure 4. Once again, its power to detect non-genetic inheritance increased with the pedigree size, proportion of transmitted variance, relative importance of epigenetic effects and the distance of the path coefficients to 0.5. In addition, the power to detect non-genetic effects increased with the disequilibrium between sire and dam transmissibility.

**Simulation**

The power of the transmissibility model to detect non-genetic inheritance was 12, 61 and 96% for dataset 2, 3 and 4, respectively. The realized type I error was 4%. The regression coefficients for different relatives were calculated by using the variance component and path coefficients of transmission used for the simulation or estimated using the transmissibility and animal models as described in appendix 4. The sire and dam
transmissibilities \( (b_{OS}, b_{OD}) \) obtained with the animal and transmissibility models are shown in Figure 5. When phenotypes were generated according to the animal model (dataset 1), the sire and dam transmissibilities obtained with the transmissibility and animal models were similar and in line with the simulated transmissibilities. The standard deviations of the transmissibilities were slightly larger for the transmissibility model in comparison with the animal model (0.024 versus 0.021 for sire transmissibility and 0.023 versus 0.021 for dam transmissibility). The transmissibility model provided a good estimation of the dam and sire transmissibilities for datasets 2 to 4 (on average 0.17, 0.17 and 0.22 for the dam and 0.15, 0.14 and 0.09 for the sire for datasets 2 to 4). For datasets 2 to 4, the animal model tended to underestimate the dam transmissibility on the one hand and overestimate the sire transmissibility on the other hand. These under- and over-estimations increased with the proportion of non-genetic heritable variance and with the disequilibrium between sire and dam transmissibilities. The ratio of the regression coefficient between other relatives obtained with the animal and transmissibility models to the expected regression coefficient under the simulated model are presented in Figure 6. Both models provided good estimations for all the regression coefficients for dataset 1. As for the sire and dam transmissibilities, the discrepancy between the regression coefficients estimated with the animal model and the values expected using the simulated model increased with the proportion of non-genetic heritable variance and the disequilibrium between sire and dam transmissibilities (datasets 2 to 4). The greatest difference between estimated and simulated regression coefficients was for paternal half-sibs, paternal half-sibs of the dam - offspring (four times higher than the simulated value) and between paternal half-sibs of the sire and offspring (six times higher than the simulated value) in dataset 4. The regression coefficients obtained with the transmissibility model were still well estimated when the proportion of non-genetic heritable variance increased (datasets 2 and 3). However, with the transmissibility model, an increase in
the proportion of heritable variance from a single parent origin resulted in a bias in the estimation of the regression coefficients between the offspring and its paternal uncle (the average ratios with simulated values were 0.71 and 1.40 for paternal and maternal half sibs of the dam, respectively).

The proportion of transmitted variance ($\tau^2$) and sire and dam path coefficients of transmission estimated with the transmissibility model and heritability $h^2$ obtained with the animal model are presented in table 3 for all datasets. The heritabilities estimated with the animal model were similar for datasets 1 to 3 (0.31 to 0.32) and slightly (but not significantly) higher for dataset 4 (0.36). The general trend was an increase of the proportion of transmitted variance estimated with the transmissibility model from dataset 1 to 4. When data were simulated using the animal model (dataset1), the coefficients of transmission obtained with the transmissibility model were on average slightly lower than 0.5 (0.46 and 0.47 for the sire and the dam path coefficient, respectively) and the proportion of transmitted variance was slightly larger than the simulated heritability (0.35 versus 0.31). In datasets 2 to 4 the estimated sire coefficient of transmission was lower (significantly in dataset 4) than the dam coefficient of transmission in accordance with the lower sire than dam simulated transmissibility.

The correlations between the true and predicted transmissible potentials obtained with the transmissibility and animal models for the different datasets are presented in table 4. When the transmissible potential was mainly of genetic origin (datasets 1 & 2), the correlations between true and predicted sire or dam transmissible potentials were the same for the transmissibility and animal models for all types of animals (with or without phenotypes). Within the subgroup of animals without phenotypes, the correlation between simulated and predicted transmissible potentials was highest when both parents were known. When the transmissible potential was mainly not of genetic origin (datasets 3 & 4), the correlations
between simulated and predicted dam transmissible potentials were slightly (but not significantly) higher for the transmissibility model in comparison with the animal model, except for animals without phenotypes and unknown dam. The biases in the estimation of the transmissible potential obtained with both models are presented in table 5. When the transmissible potential was mainly of genetic origin, there was no important bias in the prediction, whatever the model or animal subtype. When the transmissible potential was mainly not of genetic origin (datasets 3 & 4), its prediction tended to be less biased with the transmissibility model than with the animal model, the bias being higher for the sire transmissible potential than for the dam transmissible potential (on average 0.64 and 0.96 for sires with phenotype; 1.13 and 1.00 for dams with phenotype, 0.52 and 0.97 for sires without phenotype, 0.91 and 1.02 for dams without phenotypes for the animal and transmissibility model, respectively). Within the group of animals without phenotypes, the biases obtained with the animal model were higher for animals (sires or dams) with an unknown dam (ranging from 0.32 to 0.75), whereas the biases obtained with the transmissibility model were relevant for the sire transmissible potential of animals without phenotypes and one unknown parent only (ranging from 0.87 to 1.28) but the standard deviations were significant.

**DISCUSSION**

Analyzing phenotypes with a model designed to disentangle the different sources of inheritance from each other is challenging. This difficulty is well illustrated by the profile likelihoods obtained with a model aimed at dissociating genetic and non-genetic inheritance even if the parameters of such a model are theoretically identifiable. For equal sire and dam epigenetic path coefficient of transmission, on the one hand, all values of $\lambda = 0.5(1 - v_d) = 0.5(1 - v_s)$ resulted in similar profile likelihoods. On the other hand, the profile likelihood was flat for all estimated values of $r = \frac{\sigma^2_a}{\sigma^2_a + \sigma^2_{epi}}$ lower than 0.5 and of $H^2 = \frac{\sigma^2_a + \sigma^2_{epi}}{\sigma^2_a + \sigma^2_{epi} + \sigma^2_e}$ higher.
than 0.4. These results suggest a certain amount of mutual dependency between the estimates for the variance components, probably due to the excessive flexibility of the model via $\lambda$ that compensates the changes in the value of the different variance components in both upwards and downwards directions. If the $\lambda$ estimate is close to 0.5 then the epigenetic effect will mimic genetic effects and compensate a decrease in its variance (i.e. different genetic and epigenetic variances combinations ($r_0$) lead to the same likelihood), and if the $\lambda$ estimate is close to 0, then the epigenetic effect will mimic residual effects (different values of $H_0^2$ lead to the same likelihood). These findings are in line with the conclusions on this model reported by Varona et al. (2015). Identifiability is improved for $H^2$ and $\lambda_d$ when the sire and dam path epigenetic coefficients of transmission differ significantly. However, difficulties to separate genetic and epigenetic variances remain. As detailed in appendix 1, when only pedigree information is available, the possibility to separate the different components in Eq1 arises from the comparison between the covariance between different categories of relatives. Depending on the values of the parameters in Eq1, these covariances can be very similar and a very large amount of specific structured data required to estimate the parameters for the different sources of inheritance. The same difficulties have been encountered when estimating the variance components in models aimed at separating different kinds of genetic effects (maternal and social effects (CANTET and CAPPA 2008; GERMAYR 1992). Because the data structure often does not meet the requirements needed to disentangle the different sources of inheritance (also suggested by Varona et al. (2015) when applying a model aimed at separating genetic and non-genetic heritance with the same sire and dam path coefficients of transmission in a large dataset of 78,209 birth weight records), and because this distinction is not required for selection, we proposed the transmissibility model. This model combines the different inherited factors into a single transmissible value that could prove useful for selection. The objective of this model is not to quantify the different sources of inheritance...
but to take them into account for selection. Less information is needed to estimate the variance components of this model as illustrated by its profile likelihood that shows that, even for small pedigree sizes, the parameters of the transmissibility model were practically identifiable, and justifies its application to real data.

If the objective is to disentangle the different sources of inheritance then additional information is required to ensure the practical identifiability of the variance components of the different inherited factors. To study epigenetic inheritance, high-throughput technologies that quantify DNA methylation are now well established (COULDERY and CAVE 2014). DNA methylation information obtained at the individual level can be used to build the epigenetic relationship matrix $T$ between individuals that will be integrated into the model to explain phenotypic variability. However, it should be noted that if stable epigenetic mutations are in high linkage disequilibrium with the single nucleotide polymorphisms (SNPs) available on the SNP chip for the species studied; this method will not bring anything new to genomic selection (GODDARD and WHITELAW 2014). DNA methylation information from relatives could also be used to estimate the probability of epigenetic marks being erased during vertical transmission and provide an estimation of the path coefficients of transmission for this inherited factor that could be included in the model as a known parameter. Accounting for epigenetic effects in the model could also be performed by adding the level of methylation of each individual (CORTUJO et al. 2014) as a covariate explaining phenotype. For microbial inheritance, 16S rDNA amplicon sequencing is now routinely used for microbial community profiling. Information on operational taxonomic units (OTUs) can then be used to evaluate the “distance” between the microbiota of 2 individuals to compute the microbial relationship matrix $B$ (CAMARINHA-SILVA et al. 2017; LOZUPONE et al. 2011; XIA and SUN 2017). The $B$ matrix is then considered as known in the model and facilitates the estimation of the different variance components. However, recent studies aimed at estimating the microbiota variance
using this approach reported that almost all the genetic variance passes into the microbiota variance (GILBERT et al. 2018). These results suggest that there is still some confusion about the two factors that can be explained by the genetic origin of the variability of the microbiota (CAMARINHA-SILVA et al. 2017; DAVENPORT 2016). Analyzing the two phenotypes (trait of interest, microbiota) with a recursive model (GIANOLA and SORENSEN 2004) would probably help understanding this point. However, further research is needed to implement this kind of model on such a complex trait (beta diversity). To decipher genetic from cultural effects, an experimental design based on partial cross-fostering has been proposed by Danchin et al. (2013). However, it is not clear how this procedure avoids confusion between the microbiota and the culture as their modes of transmission are so close. Extracting the cultural relationship between individuals by social network analysis (SCOTT 2017) would probably be more effective. Nonetheless, this requires direct or indirect recording of animal behavior that can be time-consuming and/or expensive and/or difficult to set up for a large number of animals. At present, it seems unrealistic to try to record all this additional information in order to disentangle the different sources of inheritance in the context of genetic evaluation. The transmissibility model overcomes these difficulties by combining all these inherited factors into a single value: the transmissible potential \( t \).

To complete the practical identifiability study and to better understand the behavior of the transmissibility model and the animal model in regard to genetic and non-genetic inheritance, we evaluated by simulation the performance of the transmissibility model, and compared the results with the animal model. We deliberately used a different model to simulate the data from the models used for the estimations (except for dataset 1 which coincide with an animal model). Indeed, our purpose was to not favor one of the models but to evaluate their performances in the “real” world. Consequently, except for dataset 1, the variance components obtained with the 2 models are not comparable to those used for the simulation.
and the performance of the models were assessed as their ability to match as closely as possible the expected covariances between relatives and to predict the transmissible potential of individuals.

The covariances between different kinds of relatives are used to estimate the parameters of the transmissibility and animal models. On the one hand, the animal model has to fit these different covariances with the major restriction of using equal to 0.5 sire and dam path coefficients of transmission. On the other hand, the transmissibility model offers more flexibility with sire and dam path coefficients of transmission than can vary from 0 to 1, although their sum must always be lower than 1. This higher flexibility resulted in a closer fit of the transmissibility model with the simulated model compared with the animal model in different situations as shown in figures 5 and 6. When data were simulated under the assumption that the covariance between relatives is of genetic origin only, the path transmission coefficients (sire and dam) estimated with the transmissibility model were close but lower than 0.5. This apparent underestimation is the consequence of the bounded parameter space for the sum of the two path coefficients of transmission. Indeed, pure genetic inheritance correspond to a transmissibility model with parameters at the boundary of the parameter space (sum of the sire and dam path coefficients of transmission equal to 1). The distribution of the estimated sum of these two path coefficients over the simulation is thus truncated to the right with a point mass at 1. This problem of parameter estimation at the boundary of parameter space has been reported previously (Kopylev 2012). It is thus not possible to obtain an average estimation of the two parameters at 0.5. However, the transmissibility model was able to correctly estimate the covariance between relatives and correctly assess that the inheritance was of genetic origin only as confirmed by the realized type I error obtained which was close to 5%. The predicted transmissible potentials were similar to the breeding values obtained with the animal model. The transmissibility model is
thus able to model appropriately the genetic transmission of phenotypes. When data were
simulated under the assumption that the covariance between relatives is of both genetic and
non-genetic origin, the transmissibility model was able to correctly estimate sire and dam
transmissibilities in all situations which was not the case of the animal model, especially when
the transmissibilities were very different (dataset 4). On the one hand, the flexibility of the
transmissibility model that results from the possibility of using different values for sire and
dam path coefficients of transmission enables it to adjust the estimated regression between
relatives to the simulated values even if the simulation model is different. Of course, all the
regression coefficients could not be estimated exactly, especially when the sources of
inheritance were diverse and of non-negligible importance (i.e. the simulated model differed
greatly from the estimation model; dataset 4). In such cases, the limitations of the
transmissibility model that arise from the use of equal transmitted variance for both sexes
become apparent. However, the bias in the estimation of part of the regression coefficients is
low and involves coefficients of low values. On the other hand, to fit as close as possible to
the simulated regression coefficients with its intrinsic constraints, the animal model estimated
a genetic variance that resulted in an estimated equal sire and dam transmissibility ($\frac{1}{2} h^2$)
between the sire and dam simulated transmissibilities. This phenomenon is particularly
apparent when the disequilibrium between sire and dam transmissibilities is important. As a
consequence, the estimations of the different regressions between other relatives obtained
with this approach are biased. Thus, the transmissibility model better estimate covariances
between relatives than the animal model in the case of non-genetic inheritance.

Nevertheless, the greater flexibility of the transmissibility model compared with the animal
model has certain disadvantages. Indeed, the estimation of the path coefficients of
transmission in the transmissibility model is more time-consuming than parameter estimation
in the animal model. We evaluated that the CPU time for one iteration of convergence on a
linux system and intel®Xeon®E5-2698v3 processor is proportional to the cube of the number of individuals in the pedigree \( CPU(s) = 2 \times \left(\frac{nb\_individauls}{1000}\right)^3 \). To overcome this computation time problem, it is preferable to estimate path coefficients on a representative subset of the data and fix them to these estimated values when applying the transmissibility model to a large dataset.

Correlations between true and predicted transmissible values were similar for the animal model and the transmissibility model. The incorrect estimation of some regressions between relatives with the animal model did not influence these correlations. This is because these correlations were mainly dependent on the information used to build the evaluations and not on the regression coefficient and they were the same for both models. The high correlations (1, 0.98, 0.96 and 0.96 for datasets 1 to 4) between EBV obtained with the animal model and predicted transmissible values obtained with the transmissibility model confirm that, conversely to expected, the animal model is unable to isolate the genetic component of inherited factors; genetic and non-genetic sources of inheritance being integrated in the EBV. Thus, when there are non-genetic sources of inheritance, selection based on EBV obtained with the animal model does not correspond to a pure genetic selection but to a selection on the transmissible values. In this context, the limits of the animal model were observed on the bias. This bias was important and leads to an overestimation or an underestimation of the transmissible values depending on the source of information available: the performance of the paternal or maternal half-siblings, the sire’s or the dam’s own performance. Therefore, in practice, if the animal model performs well when very standardized selection schemes are used, with very homogeneous selection information between candidates and in the absence of overlapping generations, it would not in all other conditions. This was probably the reason for the slight discrepancy between correlations for the two models for dataset 4 when all non-
phenotyped animals are combined in a single set event though the information available for each individual is different, whereas the correlations are similar when the same kind of information is available. In the presence of overlapping generations, the biased estimations obtained do not ensure a fair comparison between candidate animals. In this case, which is common in real-life situations, the transmissibility model will be more effective than the animal model for selecting animals on their transmissible potential to improve the trait of interest in the selected population provided that the non-genetic inheritance of this trait is not negligible. It should be noted that when selection is relaxed, conversely to pure genetic selection, part of the benefit on the transmissible potential achieved by previous selection will theoretically gradually disappear and only genetic progress will be maintained. This proportion depends on the relative importance of the non-genetic effect in the inherited factors. Indeed, non-genetic effects are diluted in future generations (TAL et al. 2010) resulting in a decline to 0 of their average effects when there is no selection. The long term benefit of selection will thus mainly depend on the genetic benefit that will be conserved even when selection is relaxed. Additional calculations on the simulated data showed that the correlation between the true breeding values and the sire or dam estimated transmissible potentials and their bias were in the same range for the transmissible and the animal model (results not shown) indicating that even if selection is relaxed, long term benefit of the transmissibility model will be higher than the one obtained with the animal model. It should be noted that these results were obtained for a restricted set of simulated data (datasets 1 to 4); additional investigations are necessary to insure that same conclusion can be drawn in all situations.

Evidence of the vertical transmission of non-genetic effects has been widely reported in livestock (ABECIA et al. 2007; BRAUNSCHEIG et al. 2012; FEENEY et al. 2014; SANGA 2010). However the quantification of the non-genetic inheritance for the traits under selection is rare.
Paiva et al. (2018a) reported an epigenetic heritability of 0.10 for body weight in meat quails, Difford et al. (2018) estimated simultaneously that significant parts of the variance for CH4 emission in cattle were explained by genetics and the microbiota (0.21 and 0.13, respectively). To our knowledge the part of variance explained by culture/behavioral inheritance has never been reported in the literature for livestock species. The limited amount of data on the quantification of non-genetic inheritance in livestock is probably due to the novelty of the subject for these species and the lack of easy tools for estimating it. The transmissibility model will help to overcome this obstacle by estimating a deviation from the pure genetic inheritance using routinely recorded field data (phenotype and pedigree). Analysis of the power of the transmissibility model to detect non-genetic inheritance demonstrated that its power depends on pedigree size, the relative importance of the non-genetic inheritance and also the difference between the sire and dam path coefficients of transmission. There is nothing we can do for the two last components but we suggest using populations with complex pedigree structures, because more information will be provided for the same computing time.

We designed our transmissibility model to take into account non-genetic sources of inheritance in order to estimate the path coefficients of transmission between parents and offspring and to predict the transmissible potential of individuals. Estimations are based on the covariances between different types of relatives. It should be noted that other phenomena than the non-genetic inherited factors reported here can cause deviation from the law of transmission assumed in the animal model such as the X-chromosomal inheritance, parent-of-origin effects, mitochondrial inheritance etc. (FERNANDO and GROSSMAN 1990; HUTCHISON III et al. 1974; NEUGEBAUER et al. 2010). Disentangling all these phenomena without any information other than pedigrees and phenotypes is challenging (certainly unfeasible) and has never been attempted before. Another phenomenon that could create confusion with non-
genetic effects is the maternal effect. Indeed, maternal genetic effects induce different
covariances between the offspring and the mother and between the offspring and the father
(Appendix 4) that can mimic the transmission of non-genetic effects. To illustrate this point,
we performed an additional simulation of 100 replicates (results not shown) in which the
phenotypes were generated with the same parameters as in dataset 3 plus a maternal effect
with a genetic variance equal to 5 (independent from direct genetic effects). The application
of the transmissibility model to these simulated data resulted in a higher estimated value for
the dam coefficient of transmission than obtained for dataset 3 (0.53 versus 0.44) while the
estimates of the sire coefficient of transmission and $\tau^2$ remained similar to those obtained for
dataset 3. This result illustrates that maternal effects have been wrongly integrated into the
dam transmissible potential. To avoid such confusion between maternal genetic effects and
non-genetic effects, it is possible to include maternal genetic effects in the transmissibility
model as an additional random effect (demonstration of the structural identifiability of the
model is provided in appendix 5). The application of the transmissibility model with maternal
genetic effects to the aforementioned simulated data (i.e. dataset 3 + maternal genetic effects)
resulted in estimates of the sire/dam coefficients of transmission and $\tau^2$ similar to those
obtained with the transmissibility model applied to dataset 3, while the estimates of the
maternal genetic variance were in line with those simulated (4.7 on average). This result
confirms that it is possible to disentangle maternal genetic effects and non-genetic effects with
the transmissibility model.

It should be noted that our transmissibility model assumes that the transmitted variance $\sigma^2_t$ is
the same for all individuals in the population and constant across generations. However,
epigenetic, microbiota and cultural effects are subject to change with environmental
conditions (AGUILERA et al. 2010; SPOR et al. 2011). Taking into account this influence of the
environment on part of the heritable factors should be feasible by considering the transmitted
variance as a function of the environment or of time in a manner similar to that used in the structured antedependence model (SAD) (DAVID et al. 2017) (the transmissibility model is a particular form of SAD model).

Finally, our model combines the different sources of inheritance into a single transmissible value. However, the OWN Fortran program we propose to run the transmissibility model allows the different sources of inheritance to be separated, if required, especially genetic and epigenetic inheritance as described by Varona et al. (2015) or Paiva et al. (2018a; PAIVA et al. 2018b). Furthermore, conversely to the previous methods proposed, the program that we propose has the advantage of 1) using a frequentist method for the estimations (faster than the Bayesian approach), 2) estimating the path coefficients (conversely to the grid search proposed by Paiva et al. (2018a)), and 3) allowing for different path coefficients for sires and dams which has never been proposed before.

ACKNOWLEDGMENTS

The authors would like to thank the three anonymous reviewers for their helpful and constructive comments that greatly contributed to improving the final version of the paper.
LITERATURE CITED

ABECIA L., FONDEVILA M., BALECELLS J. and MCEWAN N., 2007 The effect of lactating rabbit does on the development of the caecal microbial community in the pups they nurture. J Appl Microbiol 103: 557-564. doi: 10.1111/j.1365-2672.2007.03277.x

AGUILERA O., FERNÁNDEZ A. F., MUÑOZ A. and FRAGA M. F., 2010 Epigenetics and environment: a complex relationship. J Appl Physiol 109: 243-251. doi: 10.1152/japplphysiol.00068.2010

BONDURIANSKY R., CREAN A. J. and DAY T., 2012 The implications of nongenetic inheritance for evolution in changing environments. Evol Appl 5: 192-201. doi: 10.1111/j.1752-4571.2011.00213.x

BRAUNSWEG M., JAGANNATHAN V., GUTZWILLER A. and BEE G., 2012 Investigations on transgenerational epigenetic response down the male line in F2 pigs. PLoS One 7: e30583. doi: 10.1371/journal.pone.0030583.

BRIGHT M., and BULGHERESI S., 2010 A complex journey: transmission of microbial symbionts. Nat Rev Microbiol 8: 218. doi: 10.1038/nrmicro2262

CAMARINHA-SILVA A., MAUZHAMMER M., WELLMANN R., VITAL M., PREUSS S. et al., 2017 Host genome influence on gut microbial composition and microbial prediction of complex traits in pigs. Genetics: 117.200782. doi: 0.1534/genetics.117.200782

CANKAR K., KRAIGHER H., RAVNIKAR M. and RUPNIK M., 2005 Bacterial endophytes from seeds of Norway spruce (Picea abies L. Karst). FEMS Microbiol Lett 244: 341-345. doi: 10.1016/j.femsle.2005.02.008

CANTET R. J. C., and CAPP A. P., 2008 On identifiability of (co)variance components in animal models with competition effects. J Anim Breed Genet 125: 371-381. doi: 10.1111/j.1439-0388.2008.00743.x

CHARLESWORTH D., BARTON N. H. and CHARLESWORTH B., 2017 The sources of adaptive variation. Proc R Soc B 284: 20162864. doi: 10.1098/rspb.2016.2864

CORTUO S., WARDENAAR R., COLOME-TATCHE M., JOHANNES F. and COLOT V., 2014 Genome-wide analysis of DNA methylation in Arabidopsis using MeDIP-chip, pp. 125-149 in Plant Epigenetics and Epigenomics. Springer.

COULDERY C., and CAVE V., 2014 Assessing DNA methylation levels in animals: choosing the right tool for the job. Anim Genet 45 Suppl 1: 15-24. doi: 10.1111/age.12186

DANCHIN É., CHARMAINTIER A., CHAMPAIGNE F. A., MESOUDI A., PUJOL B. et al., 2011 Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. Nat Rev Genet 12: 475. doi: 10.1038/nrg3028

DANCHIN É., PUJOL B. and WAGNER R. H., 2013 The Double Pedigree: A Method for Studying Culturally and Genetically Inherited Behavior in Tandem. PLoS One 8: e61254. doi: 10.1371/journal.pone.0061254

DAVENPORT E. R., 2016 Elucidating the role of the host genome in shaping microbiome composition. Gut Microbes 7: 178-184. doi: 10.1080/19490976.2016.1155022

DAVID I., GARREAU H., BALMISE E., BILLON Y. and CANARIO L., 2017 Multiple-trait structured antedependence model to study the relationship between litter size and birth weight in pigs and rabbits. Genet Sel Evol 49: 11. doi: 10.1186/s12711-017-0288-3

DIFFORD G. F., Plichita D. R., LØVENDAHL P., Lassen J., NOEL S. J. et al., 2018 Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows. PLoS Genet 14: e1007580. doi: 10.1371/journal.pgen.1007580

EICHLER E. E., FLINT J., GIBSON G., KONG A., LEAL S. M. et al., 2010 Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet 11: 446-450. doi: 10.1038/nrg2809

FALCONER D. S., 1974 Introduction à la génétique quantitative. Masson et cie.

FEENEY A., NILSSON E. and SKINNER M. K., 2014 Epigenetics and transgenerational inheritance in domesticated farm animals. J Anim Sci Biotechno 5: 48. doi: 10.1186/2049-1891-5-48

FELDMAN M. W., and CAVALLE-SFORZA L. L., 1975 Models for cultural inheritance: A general linear model. Ann Hum Biol 2: 215-226. doi: 10.1080/03014467500000791
FERNANDO R., and GROSSMAN M., 1990 Genetic evaluation with autosomal and X-chromosomal inheritance. Theor Appl Genet 80: 75-80. doi: 10.1007/BF00224018

FISHER R. A., 1918 The correlation between relatives on the supposition of Mendelian inheritance.

Trans R Soc Edin 52: 399-433. doi:

GERSTMAYR S., 1992 Impact of the data structure on the reliability of the estimated genetic parameters in an animal model with maternal effects. J Anim Breed Genet 109: 321-336. doi: 10.1111/j.1439-0388.1992.tb00412.x

GIANOLA D., and SORENSEN D., 2004 Quantitative genetic models for describing simultaneous and recursive relationships between phenotypes. Genetics 167: 1407-1424. doi:

10.1534/genetics.103.025734

GILBERT H., LAGARRIGUE S., VERSCHUREN L. M. G., ZEMB O., VELASCO M. et al., 2018 Effect of gut microbiota on production traits, interaction with genetics, pp. 204-204 in 69. Annual Meeting of the European Association of Animal Production. Academic Publishers, Dubrovnik.

GODDARD M. E., and WHITEIW E., 2014 The use of epigenetic phenomena for the improvement of sheep and cattle. Front Genet 5: 247. doi: 10.3389/fgene.2014.00247

HUTCHISON III C. A., NEWBOLD J. E., POTTER S. S. and EDGELL M. H., 1974 Maternal inheritance of mammalian mitochondrial DNA. Nature 251: 536. doi:

JABLONKA E., and LAMB M. J., 2008 Soft inheritance: challenging the modern synthesis. Genet Mol Biol 31: 389-395. doi: 10.1590/S1415-47572008000300001

JABLONKA E., and LAMB M. J., 2014 Evolution in four dimensions, revised edition: Genetic, epigenetic, behavioral, and symbolic variation in the history of life. MIT press.

KOPYLEV L., 2012 Constrained parameters in applications: Review of issues and approaches. ISRN Biomath. doi: 10.5402/2012/872956

LEY R. E., PETERTSON D. A. and GORDON J. I., 2006 Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell 124: 837-848. doi: 10.1016/j.cell.2006.02.017

LOZUPONE C., LLADSER M. E., KNIGHTS D., STOMBAUGH J. and KNIGHT R., 2011 UniFrac: an effective distance metric for microbial community comparison. ISME J 5: 169. doi: 10.1038/ismej.2010.133

MAMELI M., 2004 Nongenetic selection and nongenetic inheritance. Br J Philos Sci 55: 35-71. doi: 10.1093/bjps/55.1.35

MANOLIO T. A., COLLINS F. S., COX N. J., GOLDSTEIN D. B., HINDORF L. A. et al., 2009 Finding the missing heritability of complex diseases. Nature 461: 747. doi: 10.1038/nature08494

MARCHESI J. R., ADAMS D. H., FAVA F., HERMES G. D., HIRSCHFIELD G. M. et al., 2015 The gut microbiota and host health: a new clinical frontier. Gut 65: 330-339. doi: 10.1136/gutjnl-2015-309990

MISKA E. A., and FERGUSON-SMITH A. C., 2016 Transgenerational inheritance: Models and mechanisms of non–DNA sequence–based inheritance. Science 354: 59-63. doi: 10.1126/science.aaf4945

NEUGEBAUER N., LUTHER H. and REINSCH N., 2010 Parent-of-origin effects cause genetic variation in pig performance traits. Animal 4: 672-681. doi: 10.1017/s1751731109991625

O’DEA R. E., NOBLE D. W. A., JOHNSON S. L., HESSELSON D. and NAKAGAWA S., 2016 The role of non-genetic inheritance in evolutionary rescue: epigenetic buffering, heritable bet hedging and epigenetic traps. Environ Epigenet 2: dvv014-dvv014. doi: 10.1093/heap/dvv014

PAIVA J., DE RESENDE M., RESENDE R., DE OLIVEIRA H., SILVA H. et al., 2018a Transgenerational epigenetic variance for body weight in meat quails. J Anim Breed Genet 135: 178-185. doi: 10.1111/jbmg.12329

PAIVA J., DE RESENDE M., RESENDE R., OLIVEIRA H., SILVA H. et al., 2018b A note on transgenerational epigenetics affecting egg quality traits in meat-type quails. Br Poult Sci 59: 624-628. doi: 10.1080/00071668.2018.1514582

QUADRANA L., and COLOT V., 2016 Plant transgenerational epigenetics. Annu rev Genet 50: 467-491. doi: 10.1146/annurev-genet-120215-035254

RAFFA J. D., and THOMPSON E. A., 2016 Power and effective study size in heritability studies. Stat Biosci 8: 264-283. doi: 10.1007/s12561-016-9143-2
Rauje A., Kreutz C., Maiwald T., Bachmann J., Schilling M. et al., 2009 Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics 25: 1923-1929. doi: 10.1093/bioinformatics/btp358

Rossiter M., 1996 Incidence and consequences of inherited environmental effects. Annu Rev Ecol Evol Syst 27: 451-476. doi: 10.1146/annurev.ecolsys.27.1.451

Sandoval-Motta S., Aldana M., Martínez-Romero E. and Frank A., 2017 The human microbiome and the missing heritability problem. Front Genet 8: 80. doi: 10.3389/fgene.2017.00080

Sanga U., 2010 Maternal effects in transmission of self-medicative behavior from mother to offspring in sheep. pp. in Wildland Resources. Utah State University.

Scott J., 2017 Social network analysis. Sage publication Ltd.

Searle S. R., 1982 Matrix algebra useful for statistics. John Wiley and sons, New york.

Shreiner A. B., Kao J. Y. and Young V. B., 2015 The gut microbiome in health and in disease. Curr Opin Gastroenterol 31: 69. doi: 10.1097/MOG.0000000000000139

Sommer F., and Bäckhed F., 2013 The gut microbiota—masters of host development and physiology. Nat Rev Microbiol 11: 227. doi: 10.1038/nrmmicro2974

Spor A., Koren O. and Ley R., 2011 Unravelling the effects of the environment and host genotype on the gut microbiome. Nat Rev Microbiol 9: 279. doi: 10.1038/nrmmicro2540

Sylvain F.-É., and Derome N., 2017 Vertically and horizontally transmitted microbial symbionts shape the gut microbiota ontogenesis of a skin-mucus feeding discus fish progeny. Sci Rep 7: 5263. doi: 10.1038/s41598-017-05662-w

Tal O., Kisdi E. and Jablonka E., 2010 Epigenetic contribution to covariance between relatives. Genetics 184: 1037-1050. doi: 10.1534/genetics.109.112466

Turner T. R., James E. K. and Poole P. S., 2013 The plant microbiome. Genome Biol 14: 209. doi: 10.1186/gb-2013-14-6-209

Van Otterdijk S. D., and Michels K. B., 2016 Transgenerational epigenetic inheritance in mammals: how good is the evidence? FASEB J 30: 2457-2465. doi: 10.1096/fj.201500083

Varona L., Munilla S., Mouresan E. F., González-Rodríguez A., Moreno C. et al., 2015 A Bayesian model for the analysis of transgenerational epigenetic variation. G3-Genes Genom Genet 5: 477-485. doi: 10.1534/g3.115.016725

Visscher P. M., and Goddard M. E., 2015 A general unified framework to assess the sampling variance of heritability estimates using pedigree or marker-based relationships. Genetics 199: 223-232. doi: 10.1534/genetics.114.171017

Xia Y., and Sun J., 2017 Hypothesis testing and statistical analysis of microbiome. Genes Dis 4: 138-148. doi: 10.1016/j.gendis.2017.06.001
Table 1: Pedigree structure of the three populations considered for the practical identifiability study

|                      | Small                  | Medium                | Large                 |
|----------------------|------------------------|-----------------------|-----------------------|
| Number of paternal  | 25 * 30 half-sibs      | 250 * 30 half-sibs    | 500 * 100 half-sibs   |
| Number of maternal  | 25 * 30 half-sibs      | 250 * 30 half-sibs    | 500 * 100 half-sibs   |
| Number of sire-offspring | 1,000                | 5,000                | 25,000               |
| Number of dam-offspring | 1,000                | 5,000                | 25,000               |
| Number of uncle-nephew | 1,000                | 5,000                | 25,000               |
| Number of aunt-nephew | 1,000                | 5,000                | 25,000               |

Table 2: Parameters of the simulations

|                      | Dataset 1 | Dataset 2 | Dataset 3 | Dataset 4 |
|----------------------|-----------|-----------|-----------|-----------|
| \( \sigma^2_e \)     | 75        | 75        | 75        | 75        |
| \( \sigma^2_a \)     | 34.357    | 33        | 5         | 5         |
| \( \sigma^2_{epi} \) | 0         | 17        | 45        | 45        |
| \( \frac{1}{2}(1 - v_s) \) | 0         | 0.059     | 0.333     | 0.200     |
| \( \frac{1}{2}(1 - v_d) \) | 0         | 0.265     | 0.411     | 0.475     |
| \( \sigma^2_{sp} \)  | 0         | 2         | 2         | 10.099    |
| \( \delta \)         | 0         | 0.7       | 0.7       | 0.7       |
| \( \frac{\sigma^2_a + \sigma^2_{epi} + \sigma^2_{sp}}{\sigma^2_{a} + \sigma^2_{epi} + \sigma^2_{sp} + \sigma^2_{e}} \) | 0.314     | 0.409     | 0.409     | 0.445     |
| \( \frac{\sigma^2_a}{\sigma^2_a + \sigma^2_{epi} + \sigma^2_{sp}} \) | 1.000     | 0.635     | 0.096     | 0.083     |
| Sire transmissibility | 0.157     | 0.138     | 0.138     | 0.085     |
| Dam transmissibility  | 0.157     | 0.176     | 0.176     | 0.229     |

Phenotypes were simulated under the following model: 

\[ y_i \sim N(\mathbf{x}_i \mathbf{\beta} + a_i + epi_i + sp_i, \sigma^2_e) \]
Table 3: Proportion of transmitted variance and coefficients of transmission estimated with the transmissibility model and heritability estimated with the animal model for each dataset.

|                      | Dataset 1     | Dataset 2     | Dataset 3     | Dataset 4     |
|----------------------|---------------|---------------|---------------|---------------|
| **Animal model**     |               |               |               |               |
| \( h^2 \)            | 0.32 ± 0.04   | 0.31 ± 0.05   | 0.31 ± 0.05   | 0.36 ± 0.05   |
| \( \tau^2 \)         | 0.35 ± 0.06   | 0.36 ± 0.06   | 0.40 ± 0.09   | 0.42 ± 0.09   |
| **Transmissibility model** |       |               |               |               |
| Sire coefficient of transmission | 0.46 ± 0.06 | 0.42 ± 0.06 | 0.35 ± 0.06 | 0.22 ± 0.06 |
| Dam coefficient of transmission | 0.47 ± 0.07 | 0.47 ± 0.07 | 0.44 ± 0.09 | 0.55 ± 0.10 |
Table 4: Correlation (Mean ± sd) between predicted and true transmissible potential for the different datasets

| Model                                      | Dataset 1       | Dataset 2       | Dataset 3       | Dataset 4       |
|--------------------------------------------|-----------------|-----------------|-----------------|-----------------|
| Animals with phenotype                      |                 |                 |                 |                 |
| Animal Sire                                | 0.71 ± 0.03     | 0.68 ± 0.03     | 0.67 ± 0.03     | 0.61 ± 0.03     |
| Dam                                        | 0.71 ± 0.03     | 0.71 ± 0.02     | 0.68 ± 0.03     | 0.70 ± 0.02     |
| Transmissibility                           | 0.71 ± 0.03     | 0.67 ± 0.03     | 0.68 ± 0.03     | 0.60 ± 0.03     |
| Animals without phenotypes                 |                 |                 |                 |                 |
| Animal Sire                                | 0.42 ± 0.05     | 0.40 ± 0.06     | 0.32 ± 0.05     | 0.28 ± 0.05     |
| Dam                                        | 0.42 ± 0.05     | 0.39 ± 0.05     | 0.32 ± 0.05     | 0.31 ± 0.05     |
| Transmissibility                           | 0.42 ± 0.05     | 0.39 ± 0.06     | 0.32 ± 0.05     | 0.28 ± 0.05     |
| Animals without phenotypes, both parents known |               |                 |                 |                 |
| Animal Sire                                | 0.52 ± 0.07     | 0.49 ± 0.08     | 0.39 ± 0.08     | 0.34 ± 0.08     |
| Dam                                        | 0.52 ± 0.07     | 0.47 ± 0.07     | 0.39 ± 0.07     | 0.38 ± 0.07     |
| Transmissibility                           | 0.51 ± 0.07     | 0.48 ± 0.08     | 0.39 ± 0.08     | 0.34 ± 0.09     |
| Animals without phenotypes, dam known      |                 |                 |                 |                 |
| Animal Sire                                | 0.37 ± 0.07     | 0.34 ± 0.07     | 0.29 ± 0.07     | 0.30 ± 0.07     |
| Dam                                        | 0.37 ± 0.07     | 0.35 ± 0.07     | 0.31 ± 0.08     | 0.39 ± 0.07     |
| Transmissibility                           | 0.37 ± 0.07     | 0.34 ± 0.07     | 0.30 ± 0.08     | 0.30 ± 0.07     |
| Animals without phenotypes, sire known     |                 |                 |                 |                 |
| Animal Sire                                | 0.38 ± 0.09     | 0.36 ± 0.09     | 0.26 ± 0.09     | 0.19 ± 0.09     |
| Dam                                        | 0.38 ± 0.09     | 0.33 ± 0.09     | 0.25 ± 0.09     | 0.14 ± 0.08     |
| Transmissibility                           | 0.38 ± 0.09     | 0.35 ± 0.09     | 0.26 ± 0.09     | 0.18 ± 0.09     |
|                                            | 0.38 ± 0.09     | 0.32 ± 0.09     | 0.25 ± 0.09     | 0.14 ± 0.07     |
Table 5: Bias (Mean ± sd) in the estimation of the transmissible potential for the different datasets

| Model                                      | Model                                      | Dataset 1 | Dataset 2 | Dataset 3 | Dataset 4 |
|--------------------------------------------|--------------------------------------------|-----------|-----------|-----------|-----------|
| Animals with phenotype Animal               | Sire                                       | 1.01 ± 0.07 | 0.89 ± 0.08 | 0.81 ± 0.07 | 0.46 ± 0.04 |
|                                            | Dam                                        | 1.01 ± 0.07 | 1.06 ± 0.09 | 1.03 ± 0.09 | 1.23 ± 0.10 |
| Transmissibility                           | Sire                                       | 1.04 ± 0.13 | 0.98 ± 0.13 | 0.98 ± 0.15 | 0.94 ± 0.24 |
|                                            | Dam                                        | 1.03 ± 0.13 | 1.04 ± 0.13 | 1.00 ± 0.12 | 0.99 ± 0.10 |
| All animals without phenotypes Animal      | Sire                                       | 0.99 ± 0.14 | 0.89 ± 0.14 | 0.65 ± 0.12 | 0.38 ± 0.08 |
|                                            | Dam                                        | 0.99 ± 0.14 | 0.97 ± 0.15 | 0.83 ± 0.15 | 0.98 ± 0.18 |
| Transmissibility                           | Sire                                       | 1.10 ± 0.23 | 1.10 ± 0.26 | 1.02 ± 0.31 | 0.91 ± 0.34 |
|                                            | Dam                                        | 1.08 ± 0.24 | 1.07 ± 0.23 | 1.04 ± 0.30 | 1.00 ± 0.27 |
| Animals without phenotypes, both parents known Animal | Sire                                       | 0.98 ± 0.15 | 0.89 ± 0.16 | 0.65 ± 0.15 | 0.38 ± 0.10 |
|                                            | Dam                                        | 0.98 ± 0.15 | 0.96 ± 0.17 | 0.83 ± 0.18 | 0.98 ± 0.22 |
| Transmissibility                           | Sire                                       | 1.09 ± 0.24 | 1.10 ± 0.28 | 1.01 ± 0.33 | 0.91 ± 0.37 |
|                                            | Dam                                        | 1.08 ± 0.24 | 1.07 ± 0.25 | 1.03 ± 0.33 | 1.00 ± 0.29 |
| Animals without phenotypes, dam known Animal | Sire                                       | 1.02 ± 0.22 | 0.88 ± 0.20 | 0.70 ± 0.20 | 0.44 ± 0.12 |
|                                            | Dam                                        | 1.02 ± 0.22 | 1.03 ± 0.22 | 0.93 ± 0.25 | 1.33 ± 0.28 |
| Transmissibility                           | Sire                                       | 1.14 ± 0.31 | 1.05 ± 0.29 | 1.01 ± 0.35 | 0.87 ± 0.36 |
|                                            | Dam                                        | 1.15 ± 0.44 | 1.11 ± 0.35 | 1.07 ± 0.38 | 1.03 ± 0.32 |
| Animals without phenotypes, sire known Animal | Sire                                       | 1.01 ± 0.27 | 0.92 ± 0.25 | 0.63 ± 0.24 | 0.32 ± 0.16 |
|                                            | Dam                                        | 1.01 ± 0.27 | 0.95 ± 0.27 | 0.75 ± 0.28 | 0.55 ± 0.31 |
| Transmissibility                           | Sire                                       | 1.16 ± 0.40 | 1.25 ± 0.51 | 1.12 ± 0.59 | 1.28 ± 0.86 |
|                                            | Dam                                        | 1.13 ± 0.34 | 1.11 ± 0.37 | 1.05 ± 0.50 | 1.01 ± 0.55 |
Figure 1: Path coefficient diagram describing the transmission of the different inherited factors in livestock species.

\[ y_s = \mu + a_s + epi_s + \text{mic}_s + \text{cult}_s + e_s \]
\[ y_d = \mu + a_d + epi_d + \text{mic}_d + \text{cult}_d + e_d \]
\[ y_o = \mu + a_o + epi_o + \text{mic}_o + \text{cult}_o + e_o \]
\[ y_i = \mu + a_i + epi_i + \text{mic}_i + \text{cult}_i + e_i \]

Indices s, d, o and i refer to sire, dam, offspring and comate i, respectively.

\( a = \) genetic effects, \( epi = \) epigenetic effect, \( \text{mic} = \) microbiota effect, \( \text{cult} = \) cultural effect, \( y = \) phenotype, \( e = \) residual.
Figure 2: Profile likelihood versus parameters to estimate in a model designed to disentangle genetic and epigenetic effects and in the transmissibility model.

The true model is $y_i = a_i + epi_i + e_i$. The values of the true model are $H^2 = \frac{\sigma_a^2 + \sigma_{epi}^2}{\sigma_a^2 + \sigma_{epi}^2 + \sigma_e^2} = 0.4$, $r = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{epi}^2} = 0.5$, and 6 different combinations for the epigenetic path coefficient of transmission are considered.
The two models of estimation are \( y_i = a_i + epi_i + e_i \) (mod1) and the transmissibility model
\[ y_i = t_i + e_i. \]

First four left panels: profile likelihoods for the parameters \( H_0^2, r_0, \lambda_{s0} \) and \( \lambda_{d0} \) for mod1.

Three last panels: profile likelihoods for \( r_0^2, \omega_{s0} \) and \( \omega_{d0} \)

Top to bottom: small, medium and large pedigree sizes; colors of lines: different values of the combination \( \lambda_s, \lambda_d \)

Dotted horizontal line: the \( \chi^2 \) threshold value.
Figure 3: Power to detect non-genetic inheritance with a transmissibility model considering equal path coefficients of transmission for the sire and the dam

The true model is $y_i = x_i \beta + a_i + epi_i + e_i$ considering same sire and dam epigenetic path coefficients of transmission ($\lambda = \lambda_s = \lambda_d$). Values of the true model are $H^2 = \frac{\sigma_a^2 + \sigma_{epi}^2}{\sigma_a^2 + \sigma_{epi}^2 + \sigma_e^2}$

$0.2, 0.4 \text{ or } 0.6$, $r = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{epi}^2} = 0.3, 0.5 \text{ or } 0.7$ and varying values for the path coefficient of transmission $\lambda \in [0.05, 0.5]$
The model of estimation is the transmissibility model $y_i = x_i \beta + t_i + e_i$ that considers same path coefficients of transmission for the sire and the dam ($\omega_d = \omega_s = \omega$).

The null hypothesis $H_0$ is $\omega = 0.5$, the alternative hypothesis is $H_1 \omega \neq 0.5$. 
Figure 4: Power to detect non-genetic inheritance with the transmissibility model considering different path coefficients of transmission for the sire and the dam.

The true model is \( y_i = x_i \beta + a_i + epi_i + e_i \). Values of the true model are \( H^2 = \frac{\sigma_a^2 + \sigma_{epi}^2}{\sigma_a^2 + \sigma_{epi}^2 + \sigma_e^2} = 0.2, 0.4 \) or \( 0.6 \), \( r = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{epi}^2} = 0.3, 0.5 \) or \( 0.7 \) and varying values for the sire and dam epigenetic path coefficient of transmission \( \lambda_s \in [0.05, 0.5] \) and \( \lambda_d \in [0.05, 0.5] \).

The model of estimation is the transmissibility model \( y_i = x_i \beta + t_i + e_i \).
The null hypothesis $H_0$ is $\omega_d = \omega_s = 0.5$, the alternative hypothesis is $H_1 \omega_d \neq 0.5$ or $\omega_s \neq 0.5$. 
Figure 5: Sire and dam transmissibilities estimated using the animal and transmissibility models for the different datasets.

Dotted lines show the simulated values.
Figure 6: ratio of regression coefficient between relatives to simulated regression coefficient estimated by the animal and transmissibility models for the different datasets.

The types of relatives are: 1: paternal grand dam – offspring, 2: maternal grand dam – offspring, 3: paternal half-sib of the sire (paternal uncle 1) – offspring, 4: maternal half-sib of the sire (paternal uncle 2) – offspring, 5: paternal half-sib of the dam (maternal uncle 1) – offspring, 6: maternal half-sib of the dam (maternal uncle 2) – offspring, 7: paternal half-sibs, 8: full-sibs.
Appendix 1: demonstration of the structural identifiability of the parameters in Eq1

\[ y_i = x_i \beta + a_i + e_{pi} + s_{pi} + e_i, e \sim \text{MVN}(0, I \sigma^2_e) \]

\[ a_i = 0.5a_s + 0.5a_d + e_{a,i}, a \sim \text{MVN}(0, A \sigma^2_a) \]

\[ e_{pi} = 0.5(1 - v_s)e_{pi}s + 0.5(1 - v_d)e_{pi}d + e_{epi,i} \]

\[ = \lambda_s e_{pi}s + \lambda_d e_{pi}d + e_{epi,i}, e_{pi} \sim \text{MVN}(0, T \sigma^2_{epi}) \]

\[ s_{pi} = \delta s_{pi}d + e_{sp,i}, s_{pi} \sim \text{MVN}(0, \Sigma \sigma^2_{sp}) \]

Let’s consider

\[ r = \frac{\sigma^2_a}{\sigma^2_a + \sigma^2_{epi}} \]

\[ \theta = \frac{\sigma^2_{sp}}{\sigma^2_a + \sigma^2_{epi} + \sigma^2_{sp}} \]

\[ \alpha = \sigma^2_a + \sigma^2_{epi} + \sigma^2_{sp} \]

6 parameters describing the different sources of inheritance have to be estimated:

\[ \sigma^2_a, \sigma^2_{sp}, \sigma^2_{epi}, \lambda_s, \lambda_d, \delta. \]

They can be obtained using the covariances between relatives described in table 1

Table 1: list of covariances between phenotypes of relatives useful for identifiability of the parameters in Eq1

| Covariance between    | Formula for the covariance given Eq 1 |
|-----------------------|---------------------------------------|
| COVs-O                | \( \frac{1}{2} r(1 - \theta) + \lambda_s (1 - \theta)(1 - \theta) \) \alpha |
| COVd-O                | \( \frac{1}{2} r(1 - \theta) + \lambda_d (1 - \theta)(1 - \theta) + \delta \theta \) \alpha |
| COVPGS-O              | \( \frac{1}{4} r(1 - \theta) + \lambda_s^2 (1 - \theta)(1 - \theta) \) \alpha |
| COVMGS-O              | \( \frac{1}{4} r(1 - \theta) + \lambda_s \lambda_d (1 - \theta)(1 - \theta) \) \alpha |
| COVMGD-O              | \( \frac{1}{4} r(1 - \theta) + \lambda_d^2 (1 - \theta)(1 - \theta) + \delta^2 \theta \) \alpha |
| COVPUGD-O             | \( \frac{1}{8} r(1 - \theta) + \lambda_s^3 (1 - \theta)(1 - \theta) \) \alpha |
| COVPUGD-O             | \( \frac{1}{8} r(1 - \theta) + \lambda_s \lambda_d^2 (1 - \theta)(1 - \theta) \) \alpha |

**Estimation of \( \lambda_s \)**

\[ 2 \text{COV}_{PGS-O} - \text{COV}_{S-O} = [(1 - r)(1 - \theta) \lambda_s (2 \lambda_s - 1)] \alpha \]
Thus \( \lambda_s = \frac{2\text{COV}_{PU-O} - \text{COV}_{PGS-0}}{2\text{COV}_{PGS-O} - \text{COV}_{S-0}} \)

• Estimation of \( \lambda_d \)

\[
2\text{COV}_{MGS-0} - \text{COV}_{S-0} = [(1 - r)(1 - \theta)\lambda_s(2\lambda_d - 1)]\alpha
\]

\[
2\text{COV}_{PUGD-O} - \text{COV}_{MGS-0} = [(1 - r)(1 - \theta)\lambda_s\lambda_d(2\lambda_d - 1)]\alpha
\]

Thus, \( \lambda_d = \frac{2\text{COV}_{PUGD-O} - \text{COV}_{MGS-0}}{2\text{COV}_{MGS-O} - \text{COV}_{S-0}} \)

• Estimation of \( \sigma_a^2, \sigma_{epi}^2 \)

\[
\frac{2\text{COV}_{PGS-O}}{\text{COV}_{S-0}} = \frac{r + 4\lambda_s^2 (1 - r)}{r + 2\lambda_s (1 - r)} \quad \text{thus} \quad r = \frac{4\lambda_d (\lambda_s - \frac{\text{COV}_{PGS-O}}{\text{COV}_{S-0}})}{4\lambda_d (\lambda_s - \frac{\text{COV}_{PGS-O}}{\text{COV}_{S-0}}) + 2\frac{\text{COV}_{PGS-O}}{\text{COV}_{S-0}} - 1}
\]

Let’s consider \( a = \alpha(1 - \theta) \)

\[
a = \frac{\text{COV}_{S-0}}{0.5r + \lambda_s (1 - r)}
\]

But \( a = \sigma_a^2 + \sigma_{epi}^2 \), thus \( \sigma_a^2 = ar \) and \( \sigma_{epi}^2 = a(1 - r) \)

• Estimation of \( \delta \)

\[
\delta^2 \theta a = \text{COV}_{MGD-O} - a\left(0.25r + \lambda_d^2 (1 - r)\right)
\]

\[
\delta \theta a = \text{COV}_{D-O} - a\left(0.5r + \lambda_d (1 - r)\right)
\]

, thus \( \delta = \frac{\text{COV}_{MGD-O} - a\left(0.25r + \lambda_d^2 (1 - r)\right)}{\text{COV}_{D-O} - a\left(0.5r + \lambda_d (1 - r)\right)} \)

• Estimation of \( \sigma_{sp}^2 \)

\[
\sigma_{sp}^2 = \theta a = \frac{\text{COV}_{D-O} - a\left(0.5r + \lambda_d (1 - r)\right)}{\delta}
\]
Appendix 2: demonstration of the structural identifiability of the parameters in the transmissibility model

\[ y_i = x_i \beta + t_i + e_i, e \sim MVN(0, I \sigma_e^2) \]
\[ t_i = \omega_s t_s + \omega_d t_d + \varepsilon_{a,i}, t \sim MVN(0, M \sigma_t^2) \]

Three parameters describing the different sources of inheritance have to be estimated: \( \sigma_t^2, \omega_s \) and \( \omega_d \).

The following covariances between relatives are useful for estimating the aforementioned parameters:

\[ COV_{S-O} = \omega_s \sigma_t^2 \]
\[ COV_{D-O} = \omega_d \sigma_t^2 \]
\[ COV_{PGS-O} = \omega_s^2 \sigma_t^2 \]

- Estimation of \( \omega_s \)
  \[ \omega_s = \frac{COV_{PGS-O}}{COV_{S-O}} \]

- Estimation of \( \sigma_t^2 \)
  \[ \sigma_t^2 = \frac{COV_{S-O} \omega_s}{COV_{PGS-O}} \]

- Estimation of \( \omega_d \)
  \[ \omega_d = \frac{COV_{D-O}}{\sigma_t^2}, \text{ thus} \]
  \[ \omega_d = \frac{COV_{D-O} \cdot COV_{PGS-O}}{COV_{S-O}^2} \]
Appendix 3: practical identifiability and power analysis

We assume that the true model is the model developed in Eq 1, single parent transmitted value and fixed effects excluded:

\[ y = a + epi + e \] (1)

with \( V(a) = A\sigma^2_a, V(epi) = T\sigma^2_{epi} \)

- log likelihood computation

According to Visscher and Goddard (2015) or Raffa and Thompson (2016), the symmetric \( A \) matrix can be decomposed and written as:

\[ A = U\Delta U \]

with \( UU' = U'U = I \) and \( \Delta \) is a diagonal matrix containing eigenvalues of \( A \).

Consider now simplified relationships summarized as \( N \) independent families subdivided into \( N_c \) families with \( n_c \) individuals of homogeneous relationship \( c \) \( (N = \sum N_c) \). The total number of animals is \( n = \sum n_c N_c \). In that case, the \( A \) matrix is block-diagonal and the eigenvalues \( \delta_i \) of each block are \( n_c - 1 \) eigenvalues of value \( (1 - c) \) and 1 eigenvalue of value \( (1 + c(n_c - 1)) \), and \( U \) is block diagonal with Helmert matrices, the coefficients of which depend only on the size of the family \( n_c \) (SEARLE 1982). So, the \( T \) matrix, which follows the same pattern as \( A \), i.e. blocks of \( N \) homogeneous families with epigenic relationships, can also be decomposed and written as:

\[ T = U\Pi U' \]

where the \( \Pi \) matrix is a diagonal matrix of eigenvalues \( d_i \) of \( T \), and \( U \) is the same block Helmert matrix. The eigenvalues \( d_i \) are directly linked to the reset coefficients \( \nu_s \) and \( \nu_d \).

Assuming one ancestor that links the animals of the family with \( k_s \) and \( k_d \) paternal and maternal paths, the \( d_i \) are:

\[ 
\begin{align*}
\text{1. (n_c - 1) eigenvalues of } & \left( 1 - \left( \frac{1}{2} (1 - \nu_s) \right)^{k_s} \left( \frac{1}{2} (1 - \nu_d) \right)^{k_d} \right) \\
\text{2. 1 eigenvalue of } & \left( 1 + (n_c - 1) \left( \frac{1}{2} (1 - \nu_s) \right)^{k_s} \left( \frac{1}{2} (1 - \nu_d) \right)^{k_d} \right)
\end{align*}
\]

Define \( y^* = U^{-1}y = U'y \), then

\[ y^* = U'a + U'epi + U'e = a^* + epi^* \]

With

\[ V(y^*) = U'AU\sigma^2_a + U'TU\sigma^2_{epi} + U'U\sigma^2_e = \Delta\sigma^2_a + \Pi\sigma^2_{epi} + I\sigma^2_e \]
Let’s consider

\[
r = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{epi}^2} \quad \text{and} \quad H^2 = \frac{\sigma_a^2 + \sigma_{epi}^2}{\sigma_a^2 + \sigma_{epi}^2 + \sigma_e^2}
\]

\[
V(y^*) = (\Delta r H^2 + \Pi (1 - r) H^2 + I (1 - H^2)) \sigma^2
\]

with \( \sigma^2 \) the total variance assumed known.

The log likelihood with respect to \( H^2, r \) and \( \nu_s, \nu_d \) can then be expressed as:

\[
\log L_1 = -\frac{1}{2} \left[ n \log(\sigma^2) + \sum \log((\delta_i - 1)r H^2 + (d_i(\nu_s, \nu_d) - 1) H^2 (1 - r) + 1) + (1/\sigma^2) \sum y_i^2 / ((\delta_i - 1)r H^2 + (d_i(\nu_s, \nu_d) - 1) H^2 (1 - r) + 1) \right]
\]

Similarly, we can derive the log likelihood under the transmissibility model:

\[
y = t + e \quad (2)
\]

with \( V(t) = M \sigma_t^2 \), as:

\[
\log L_2 = -\frac{1}{2} \left[ n \log(\sigma^2) + \sum \log((d_i'(\omega_s, \omega_d) - 1) \tau^2 + 1) + (1/\sigma^2) \sum y_i^2 / ((d_i'(\omega_s, \omega_d) - 1) \tau^2 + 1) \right]
\]

With \( d_i' \) the eigenvalues of the \( M \) matrix and \( \tau^2 = \frac{\sigma_t^2}{\sigma_t^2 + \sigma_e^2} \)

Parameter estimates

We computed parameter estimates assuming the true model was the genetic and epigenetic model (1) and model used for estimation either the same (1) or the transmissibility model (2).

Parameter estimates were computed using

\[
E(y_i^2) = [(\delta_i - 1)r H^2 + (d_i(\nu_s, \nu_d) - 1) H^2 (1 - r) + 1] \sigma^2
\]

with the true parameters \( H^2, r \) and \( \nu_s, \nu_d \) and then maximizing the expected log likelihood \( E(\log L_1) \), considered as a function of a point identified by subscript 0:

\[
E(\log L_1) = -\frac{1}{2} \left[ n \log(\sigma^2) + \sum \log((\delta_i - 1)r_0 H_0^2 + (d_i(\nu_{s,0}, \nu_{d,0}) - 1) H_0^2 (1 - r_0) + 1) + \sum \frac{(\delta_i - 1)r H^2 + (d_i(\nu_s, \nu_d) - 1) H^2 (1 - r) + 1}{((\delta_i - 1)r_0 H_0^2 + (d_i(\nu_{s,0}, \nu_{d,0}) - 1) H_0^2 (1 - r_0) + 1)} \right]
\]

with the constraints \( 0 \leq H_0^2 \leq 1, 0 \leq r_0 \leq 1, 0 \leq \nu_{s,0} \leq 0.5, 0 \leq \nu_{d,0} \leq 0.5 \). So that:
For the transmisslity model, parameter estimates were computed using the same true model (1), with the same true parameters and $E(y_i^2)$ and maximizing $E(\log L_2)$:

$$E(\log L_2) = -\frac{1}{2} \sum \log((d_i'(\omega_{s,0}, \omega_{d,0}) - 1)\tau_{0}^2 + 1)$$

$$+ \sum \frac{(\delta_i - 1)rH^2 + (d_i'(\nu_s, \nu_d) - 1)H^2(1 - r) + 1}{((d_i'(\omega_{s,0}, \omega_{d,0}) - 1)\tau_{0}^2 + 1)}$$

with the constraints $0 \leq \tau_{0}^2 \leq 1, 0 \leq \omega_{s,0} + \omega_{d,0} \leq 1$. So that

$$\{\hat{\tau}^2, \hat{\omega}_s, \hat{\omega}_d\} = \arg\max(E(\log L_2))$$

**Profile Likelihood**

We studied the practical identifiability using Profile Likelihood (Raue et al. 2009) for each parameter of the epigenetic-genetic model and transmissibility model using expected likelihood ratio tests:

$$ELRT_1 = 2 \left(E(\log L_1(\hat{H}_0^2, r_0, \nu_{s,0}, \nu_{d,0})) - E(\log L_1(H_0^2, \hat{r}, \hat{\nu}_s, \hat{\nu}_d))\right)$$

for model 1, and as follows for the transmissibility model:

$$ELRT_2 = 2 \left(E(\log L_2(\tau_{0}^2, \omega_{s,0}, \omega_{d,0})) - E(\log L_2(\hat{\tau}^2, \hat{\omega}_s, \hat{\omega}_d))\right)$$

The profile likelihoods for each parameter $\theta_0 = \{H_0^2, r_0, \nu_{s,0}, \nu_{d,0}\}$ in the definition domain are:

$$PL_1(\theta_{0,i}) = \min_{i \neq j}(ELRT_1(\theta_0))$$

And for the transmissibility model $\theta_0 = \{\tau_{0}^2, \omega_{s,0}, \omega_{d,0}\}$:

$$PL_2(\theta_{0,i}) = \min_{i \neq j}(ELRT_2(\theta_0))$$

The profile likelihood can then be plotted as a function of the parameters to check for structural identifiability.

**Power computation**

The power to detect non-genetic inheritance was calculated for the transmissibility model only.

For the transmissibility model assuming equal path coefficients of transmission for the sire and the dam ($\omega = \omega_s = \omega_d$), the power was calculated to test the hypothesis $\omega \neq 0.50$ with:
\[ 1 - F \left( \chi_\alpha; df; ELRT_2(\tau^2, 0.50, 0.50) \right) \]

For the transmissibility model assuming different path coefficients of transmission for the sire and the dam, the power was also calculated to test the hypothesis that at least one of \( \omega_s, \omega_d \) is different than 0.50 with:

\[
\max \left[ 1 - F \left( \chi_\alpha; df; ELRT_2(\tau^2, 0.50, \min(\hat{\omega}_d, 0.50)) \right), \right.
\]

\[
\left. 1 - F \left( \chi_\alpha; df; ELRT_2(\tau^2, \min(\hat{\omega}_s, 0.50), 0.50) \right) \right]
\]

where \( F \) is the cumulative distribution function of a non central \( \chi^2 \) evaluated at \( \chi_\alpha \) the \( (1 - \alpha) \) quantile for a \( \chi^2 \) distribution with one degree of freedom and non centrality parameter zero. The number of degrees of freedom (\( df \)) was set to 1 to deal with pointwise confidence intervals.
## Appendix 4: Regression Coefficient for Different Types of Relatives Depending on the Model

| Covariance between offspring and | Formula for the regression coefficients | Simulated value | Transmissibility model | Animal model | Maternal model |
|---------------------------------|----------------------------------------|-----------------|------------------------|--------------|----------------|
| Sire                            | $y_i = x_i \beta + a_i + epi_i + sp_i + e_i$ | $\frac{1}{2} \sigma_a^2 + \frac{1}{2} \lambda_s \sigma_{epi}^2$ | $\omega_s \sigma_t^2$ | $\frac{1}{2} \sigma_u^2$ | $\frac{1}{2} \sigma_u^2 + \frac{1}{2} \sigma_{um}^2$ |
| Dam                             | $y_i = x_i \beta + t_i + e_i$ | $\frac{1}{2} \sigma_d^2 + \frac{1}{4} \lambda_d \sigma_{epi}^2 + \delta \sigma_{sp}^2$ | $\omega_d \sigma_t^2$ | $\frac{1}{2} \sigma_u^2$ | $\frac{1}{2} \sigma_u^2 + \frac{5}{4} \sigma_{um}^2 + \frac{1}{2} \sigma_m^2$ |
| Paternal grand dam             | $y_i = x_i \beta + t_i + e_i$ | $\frac{1}{4} \sigma_a^2 + \frac{1}{4} \lambda_s \lambda_d \sigma_{epi}^2$ | $\omega_s \omega_d \sigma_t^2$ | $\frac{1}{4} \sigma_u^2$ | $\frac{1}{4} \sigma_u^2 + \frac{1}{8} \sigma_{um}^2$ |
| Maternal grand dam             | $y_i = x_i \beta + u_i + m_d + e_i$ | $\frac{1}{4} \sigma_a^2 + \frac{1}{4} \lambda_s \lambda_d \sigma_{epi}^2$ | $\omega_s \omega_d \sigma_t^2$ | $\frac{1}{4} \sigma_u^2$ | $\frac{1}{4} \sigma_u^2 + \frac{5}{4} \sigma_{um}^2 + \frac{1}{4} \sigma_m^2$ |
| Paternal half-sib of the sire (paternal uncle 1) | $y_i = x_i \beta + u_i + m_d + e_i$ | $\frac{1}{8} \sigma_a^2 + \frac{1}{8} \lambda_s \lambda_d \sigma_{epi}^2$ | $\omega_s \omega_d \sigma_t^2$ | $\frac{1}{8} \sigma_u^2$ | $\frac{1}{8} \sigma_u^2 + \frac{1}{4} \sigma_{um}^2$ |
| Maternal half-sib of the sire (paternal uncle 2) | $y_i = x_i \beta + u_i + m_d + e_i$ | $\frac{1}{8} \sigma_a^2 + \frac{1}{8} \lambda_s \lambda_d \sigma_{epi}^2$ | $\omega_s \omega_d \sigma_t^2$ | $\frac{1}{8} \sigma_u^2$ | $\frac{1}{8} \sigma_u^2 + \frac{1}{4} \sigma_{um}^2$ |
| Paternal half-sib of the dam (maternal uncle 1) | $y_i = x_i \beta + u_i + m_d + e_i$ | $\frac{1}{8} \sigma_a^2 + \frac{1}{8} \lambda_s \lambda_d \sigma_{epi}^2$ | $\omega_s \omega_d \sigma_t^2$ | $\frac{1}{8} \sigma_u^2$ | $\frac{1}{8} \sigma_u^2 + \frac{1}{4} \sigma_{um}^2$ |
| Maternal half-sib of the dam (maternal uncle 2) | $y_i = x_i \beta + u_i + m_d + e_i$ | $\frac{1}{8} \sigma_a^2 + \frac{1}{8} \lambda_s \lambda_d \sigma_{epi}^2$ | $\omega_s \omega_d \sigma_t^2$ | $\frac{1}{8} \sigma_u^2$ | $\frac{1}{8} \sigma_u^2 + \frac{1}{4} \sigma_{um}^2$ |
| Paternal half-sib               | $y_i = x_i \beta + u_i + m_d + e_i$ | $\frac{1}{8} \sigma_a^2 + \frac{1}{8} \lambda_s \lambda_d \sigma_{epi}^2$ | $\omega_s \omega_d \sigma_t^2$ | $\frac{1}{8} \sigma_u^2$ | $\frac{1}{8} \sigma_u^2 + \frac{1}{4} \sigma_{um}^2$ |
| Full-sib                       | $y_i = x_i \beta + u_i + m_d + e_i$ | $\frac{1}{2} \sigma_a^2 + \frac{1}{4} \lambda_s \lambda_d \sigma_{epi}^2 + \delta^2 \sigma_{sp}^2$ | $\omega_s \omega_d \sigma_t^2$ | $\frac{1}{2} \sigma_u^2$ | $\frac{1}{4} \sigma_u^2 + \sigma_{um} + \sigma_m^2$ |
Appendix 5: demonstration of the structural identifiability of the parameters in a transmissibility model with maternal effects

\[ y_i = x_i \beta + t_i + m_d + e_i, \quad e \sim MVN(0, I \sigma^2) \]

\[ t_i = \omega_s t_s + \omega_d t_d + \epsilon_{a,i} \]

\[
\begin{bmatrix} t \\ m \end{bmatrix} \sim MVN \left( 0, \begin{bmatrix} M \sigma_t^2 & \sigma_{t,m} \\ \sigma_{t,m} & A \sigma_m^2 \end{bmatrix} \right).
\]

Five parameters describing the different sources of inheritance have to be estimated: \( \sigma_t^2, \omega_s, \omega_d, \sigma_m^2 \) and \( \sigma_{t,m} \). They can be obtained using the covariances between relatives described in table 2.

Table 2: list of covariances between phenotypes of relatives useful for parameter identifiability in the transmissibility model with maternal effects.

| Covariance between          | Formula for the covariance                   |
|-----------------------------|----------------------------------------------|
| COV_{S,O} Sire-offspring    | \( \omega_s \sigma_t^2 + \omega_d \sigma_{t,m} \) |
| COV_{D,O} Dam-offspring     | \( \omega_d \sigma_t^2 + (\omega_d^2 + 1) \sigma_{t,m} + 0.5 \sigma_m^2 \) |
| COV_{PGD,O} Paternal grand dam-offspring | \( \omega_s \omega_d \sigma_t^2 + \omega_s \omega_d^2 \sigma_{t,m} \) |
| COV_{PU-O} Paternal half-sib of the sire (paternal uncle 1) - offspring | \( \omega_s^2 \sigma_t^2 \) |
| COV_{PHS} Paternal half-sibs | \( \omega_s^2 \sigma_t^2 \) |

- Estimation of \( \omega_s \)
  \[ \omega_s = \frac{\text{COV}_{Pu-O}}{\text{COV}_{PHS}} \]

- Estimation of \( \omega_d \)
  \[ \omega_d = \frac{\text{COV}_{PGD-O}}{\text{COV}_{S-O}} \]

- Estimation of \( \sigma_t^2 \)
  \[ \sigma_t^2 = \frac{(\text{COV}_{PHS})^3}{(\text{COV}_{Pu-O})^2} \]

- Estimation of \( \sigma_{t,m} \)
  \[ \sigma_{t,m} = \frac{\text{COV}_{S-O} \text{COV}_{PHS} (\text{COV}_{S-O} \text{COV}_{Pu-O} - (\text{COV}_{PHS})^2)}{\text{COV}_{PGD-O} (\text{COV}_{Pu-O})^4} , \text{thus} \]

- Estimation of \( \sigma_m^2 \)
  \[ \sigma_m^2 = 2[\text{COV}_{D-O} - (\omega_d \sigma_t^2 + (\omega_d^2 + 1) \sigma_{t,m})] \]