Simulation of host-microbiome evolution throughout a divergent selection experiment

Simulation of host-microbiome evolution throughout a divergent selection experiment

C. Casto-Rebollo[1*], I. Pocrnic[2], G. Gorjanc[2], N. Ibáñez-Escriche[1]

[1]Institute for Animal Science and Technology, Universitat Politècnica de València, C/Camino de Vera s/n, València, 46022, Spain; [2]The Roslin Institute, The University of Edinburgh, Easter Bush Campus, Edinburgh, EH25 9RG, UK; *cricasre@posgrado.upv.es

Abstract

The microbiome composition influences the host response to selection and shapes complex phenotypes. It is a multifactorial complex trait in which the microbial inheritance, the host-genome, and the own microbial interactions influence its variability. The expensive sequence-based techniques limit the availability of empirical data. Thus, other approaches are necessary to evaluate strategies for microbiome studies. This study aimed to develop a simulator following a long-term phenotype selection experiment in rabbits, considering both genome and microbiome inheritance. The result showed genetic, microbiome, and phenotypic trends across 13 generations. Differences in the microbiome contribution to the phenotype were shown only in the scenario with a high host genetic effect of the microbiome. This is a preliminary work and further improvements to the simulator will be shared with the research community.

Introduction

Each year, new research underpins the importance of the effect of microbiota on animals' phenotypes suggesting that this contribution is on par with the host genome (Camarinha-Silva *et al.*, 2017; Difford *et al.*, 2018). However, how the microbiome composition shapes complex phenotypes and mediates the host response to selection remains unclear (Henry *et al.*, 2021).

Animal's microbiome composition is established by both the transmission mode (microbial inheritance) and the host genome (microbial heritability). Microbial inheritance could be vertical or horizontal (Koskella *et al*., 2017), depending on whether and in what quantity the microbiota in offspring came from their parents (Parental microbiota; PM) or their environment (Environmental microbiota; EM). Thus, the transmission mode plays a crucial role in the individual pool of available microbes (Dominguez-Bello *et al*., 2010). On the other hand, microbial heritability could influence the dynamic nature of the microbiome in the host, so microbial composition can be considered as another quantitative trait of an animal (Pérez-Enciso *et al*., 2021). Furthermore, other factors such as the microbial genome and the microbial interactions also shape the microbiome composition, which is clearly a multifactorial complex trait.

The expensive sequence-based techniques limit the availability of empirical data to study microbiome composition at scale. Moreover, the lack of consensus on theoretical models and analytic methodologies constrains comparisons between experiments and learning how various microbiome-related factors impact animal's phenotypes. Here we take a simulation approach to gain understanding of the key principles in studying the effect of the microbiome on phenotypes and selection response. To this end we have developed a simulator following a long-term phenotype selection experiment in rabbits (Blasco *et al.*, 2017), considering both genome and microbiome inheritance. The overall

aim is to develop a tool for testing different data collection strategies and statistical models to further the study of microbiome in animal breeding.

Materials & Methods

The developed simulator allows users to simulate a host genome, microbiome composition, and animal's phenotype as a function of genome, microbiome, and environment. In this study, we used this simulator to mimic a phenotypic divergent selection experiment modulated by the host-microbiome (Blasco *et al.*, 2017). Moreover, three scenarios were implemented to study microbial heritability and its effect on the phenotype. The simulation was repeated 10 times.

Genome simulation. The simulation of 1000 diploid genomes in the base population was performed with the AlphaSimR R package (Gaynor *et al.*, 2021), using the genome and demographic history of rabbits (*Oryctolagus cuniculus*). We simulated 21 chromosomes with a length of 1.28^{10} base pairs, 10,000 segregating sites, and 100 Quantitative Trait Loci (QTL) per chromosome. The allele substitution effects of loci QTLs ($\boldsymbol{\alpha}$) were sampled from gamma distribution, $\boldsymbol{\alpha}$~Γ(k=0.2, θ=1). In the base population, the genetic values (gv) were distributed with mean 0 and genetic variance of 0.81 with a targeted heritability of 0.13 (for litter size from Blasco *et al.*, (2017)).

Microbiome simulation. We simulated a parental (PM) and an environmental microbiome (EM) to model both aspects of microbiome inheritance. The species abundances (x_k) in PM and EM were sampled from a negative binomial distribution, x_k~NB (r=2, q=0.0001). The total number of species was 1,000 in EM and 600 in PM (with overlap between them). Both, PM, and EM, were simulated with a total species abundance of 10^8. The average of species abundance (μ_k) varied depending on the combination of the species in the individuals due to the microbiome inheritance. Hence, a simulation of 100,000 microbiome profiles was done to obtain μ_k in the founder population, considering that the microbiome inheritance was 100% of the microbial species from the PM (assuming a vertical transmission from the dam due to the delivery mode) and 15% from the EM. The stability of microbial species was computed as a coefficient of variation (CV), sampled from a uniform distribution U(0.01, 2). This CV allowed us to compute the species variance ($\sigma^2_{m_k}$) and simulate the microbiome for each animal in the base population (PM_i), sampling the abundance of each microbial species k from a normal distribution N(μ_k, $\sigma^2_{m_k}$). A total of 35 microbial species were assigned an effect on the phenotype, following a gamma distribution, ω_k~Γ(k=0.2, θ=5) (Pérez-Enciso *et al.*, 2021). The individual's microbiome value (mv) in the base population was distributed with mean 0 and microbiome variance (σ^2_m) of 0.81, to match the genetic variance (Pérez-Enciso *et al.*, 2021).

Phenotype simulation. The phenotype was simulated as:

$$y_i = \mu_P + \sum_{j=1}^{n} z_{ij}\alpha_j + \sum_{k=1}^{m} x_{ik}\omega_k + e_i \quad (1)$$

where y_i is the phenotype of individual i; μ_P is the phenotypic mean in the base population; $\sum_{j=1}^{n} z_{ij}\alpha_j$ is the gv_i of individual i, where z_{ij} is the genotype of individual i for SNP j and α_j is the allele substitution effect of the SNP j; $\sum_{k=1}^{m} x_{ik}\omega_k$ is the mv_i of individual i, where x_{ik} is the abundance of species k in individual i (see equation 2) and ω_k is the effect of species k on the host phenotype; and e_i is the residual of inidividual i distributed as N(0, σ^2_e), were σ^2_e is the residual variance.

Inheritance. A divergent selection based on the female's phenotypes was simulated according to Blasco *et al.* (2017). In the base population, 125 individuals with the highest phenotype and 125 with the lowest phenotype were selected to initiate divergent populations. Breeding males were full-sibs of the best 25 breeding females from each divergent population. Each breeding male was mated with five breeding females, avoiding a close relationship among them. Thirteen discrete generations of selection were simulated simultaneously for the two divergent populations. In each generation, 125 breeding females and 25 breeding males were selected in each population. The breeding females were selected as in the base population, and the breeding males were males from the offspring of the best female for each sire (best mating). Mating and genome inheritance were simulated using the AlphaSimR package (Gaynor *et al.*, 2021). Microbiome inheritance was 100% from the PM and 15% from the EM. The transmission EM was the same for all animals and the two divergent populations, but different between generations (EM_g). Hence, the microbiome composition in the offspring was computed using equation 2.

Scenarios considered. Three scenarios were simulated to study how microbial heritability affected response to selection. The first scenario assumed the absence of microbial heritability (NMH), so there was no host genome effect on its microbiome. In the second and third scenarios, microbial heritability was included in the simulation. The second scenario assumed intermediate microbial heritability (IMH) from 0.25 to 0.50 and the third scenario assumed high microbial heritability (HMH) from 0.5 to 0.9. Consequently, the species abundance was considered as another quantitative trait simulated following Pérez-Enciso *et al.* (2021):

$$x_{ik} = \mu_k + \sum_{j=1}^{n} z_{ij}\beta_{jk} + e_{ik} \quad (2)$$

where x_{ik} is the abundance of the specie k in the individual i; μ_k is the average abundance of the specie k in the base population; $\sum_{j=1}^{n} z_{ij}\beta_{jk}$ is the genetic value (gv_{ik}) of individual i for the abundance of specie k, z_{ij} is the genotype of SNP j for individual i, β_{jk} being the substitution allele effect of SNP j for the species k; and e_{ik} is the residual of individual i for the species k distributed as N(0, $\sigma^2_{e_k}$), where $\sigma^2_{e_k}$ is residual variance for the species k. The gv_{ik} were distributed with mean 0 and microbial genetic variance (σ^2_{mg}) according to each microbial heritability. We assumed that 10% of the total number of species (100) were under the genetic control of 10% of the QTLs (210). From these 100 species, 18 of them influence the phenotype (50% of total species influencing the phenotype).

Results and Discussion

The simulation showed expected divergent trends throughout the 13 generations of selection (Figure 1). Although there were 35 microbial species with effect on host phenotype, and 18 of them under host genetic control, no differences were observed in phenotype trends between the scenarios NMH and IMH (Figure 1C). Differences in the phenotype trends were only shown with the scenario HMH (Figure 1C), that is when microbial heritability was high, from 0.5 to 0.9. In this scenario, microbiome value trend diverged more than with the NMH and IMH scenarios (Figure 1B). These preliminary results suggested that, under the assumed simulation, a high microbial heritability is necessary to observe a significant contribution of microbiome to response to selection. Likewise, the microbiome composition is multifactorial, and it has a huge

variability within and between individuals. Another scenario considering only the effect of the host genome on the phenotype is necessary to fully assess the degree of microbiome contribution on the phenotype.

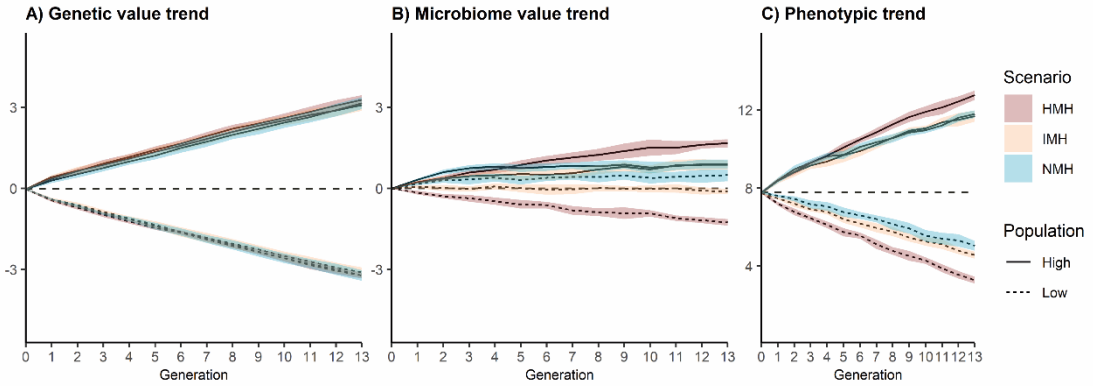

Figure 1. Genetic, microbiome, and phenotypic trends throughout a divergent selection experiment. Averages and 95% quantiles over replicates are shown for **A)** genetic value, **B)** microbiome value, and **C)** phenotype value. HMH: High microbial heritability, IMH: Intermediate microbial heritability, NMH: No microbial heritability. High (dashed line) and Low are the two divergent populations.

Conclusion

In this study, a simulator of the modulation of a host-microbiome evolution has been developed. The result showed genetic, microbiome, and phenotypic trends across 13 generations. Differences in the microbiome effect on the host phenotype were shown only in the scenario with high microbial heritability. This could be due to the reduction of the impact of the environmental microbiome. This is a preliminary work, and further improvements to the simulator will be shared with the research community.

References

Blasco A., Martínez-Álvaro M., García ML., Ibáñez-Escriche N., and Argente A. (2017) Genet Sel Evol 49, 48. https://doi.org/10.1186/s12711-017-0323-4

Camarinha-Silva A., Maushammer M., Wellmann R., Vital M., Preuss S., *et al.* (2017) Genetics 206:1637–1644. https://doi.org/10.1534/genetics.117.200782

Difford G. F., Plichta D.R., Løvendahl P., Lassen J., Noel S.J., *et al.* (2018) PLoS Genet 14(10):e1007580. https://doi.org/10.1371/journal.pgen.1007580

Dominguez-Bello M.G., Costello E.K., Contreras M., Magris M., Hidalgo G., *et al.* (2010) Proc Natl Acad Sci U S A 29;107(26):11971-5. https://doi.org/10.1073/pnas.1002601107.

Gaynor R.C., Gorjanc G., and Hickey J.M. (2021) G3 (Bethesda) 9;11(2):jkaa017. https://doi.org/10.1093/g3journal/jkaa017.

Henry L.P., Bruijning M., Forsberg S.K.G., and Ayroles J.F. (2021) Nat Commun 12, 5141 (2021). https://doi.org/10.1038/s41467-021-25315-x

Koskella B., Hall L.J., and Metcalf C.J.E. (2017) Nat Ecol Evol 1:1606–1615. https://doi.org/10.1038/s41559-017-0340-2