27. Genome-wide association study for mid-infrared methane predictions in Walloon dairy cows

H. Atashi[1,2*], H. Wilmot[1,3], S. Vanderick[1], X. Hubin[4], H. Soyeurt[1] and N. Gengler[1]

[1]Gembloux Agro-Bio Tech, University of Liège, 5030 Gembloux, Belgium; [2]Shiraz University, 71441-65186 Shiraz, Iran; [3]National Fund for Scientific Research (F.R.S.-FNRS), Rue d'Egmont 5, 1000 Brussels, Belgium; [4]Elevéo Asbl 5590 Ciney, Belgium; hadi.atashi@uliege.be

Abstract

This study aimed to identify genomic regions associated with two mid infrared-based CH_4 traits [predicted daily CH_4 emission (PME, g/d), and log-transformed predicted CH_4 intensity (LMI)] in Walloon dairy cows. The data consisted of 1,529,282 test-day records from 229,465 cows distributed in 1,530 herds collected from 2006 to 2021. Random regression test-day models were used to estimate variance components. The proportion of genetic variance explained by windows of 50 consecutive SNPs was calculated and regions accounting for at least 1.0% of the total genetic variance were identified. Mean (SD) daily h^2 estimated for PME and LMI were 0.14 (0.05) and 0.24 (0.05), respectively. Two regions on BTA14 (positions 1.86 to 2.12, and 1.48 to 1.68 Mb) were associated with both PME and LMI. A region between 144.38 to 144.46 Mb on BTA1 was associated with PME; and the region between 2.68 and 2.94 Mb on BTA14 was associated with LMI. Results showed potential for genome-enhanced advisory systems to reduce methane emissions.

Introduction

Emissions of greenhouse gases (GHG) have a considerable impact on climate change. Livestock production is a significant contributor to global human-induced GHG emissions in the form of CH_4, N_2O and CO_2 (De Haas *et al.*, 2021). Dairy cattle account for 20% of the global livestock sector's GHG, and over half of this is from enteric CH_4 emissions (Gerber *et al.*, 2013). Although a reduction in CH_4 emission could be achieved by using different CH_4 mitigation strategies like management or feeding (Cottle *et al.*, 2011), a genetic selection of low CH_4 emitting cows can be an effective approach to reduce GHG production from dairy cattle. The genetic variation on CH_4 emission is influenced by many genes with heterogeneous effects (Sypniewski *et al.*, 2021). However, there is currently a gap in the knowledge on the genetic architecture of this new trait (Kandel *et al.*, 2017; Sypniewski *et al.*, 2021). The short-term aim of this study was to estimate the genetic parameters of two MIR-based CH_4 features for the first-parity Walloon dairy cows and to identify their associated genomic regions. The long-term objective is to contribute to genome-enhanced advisory systems to reduce methane emissions.

Materials &methods

Data. Milk samples were collected on first parity cows from 2006 to 2021 by the Walloon Breeding Association (Ciney, Belgium). All milk samples were analysed using Milkoscan FT4000, FT6000, and FT+ (Foss-Electric A/S, Hillerød, Denmark) by the milk laboratory Comité du Lait (Battice, Belgium) to generate the MIR spectral data. Methane emissions (PME, g/d) were predicted from the recorded spectra using the equations developed by Vanlierde *et al.* (2021). To eliminate potential abnormal records, the PME values below the 0.1 percentile and above the 99.9 percentile were deleted (Kandel *et al.*, 2017). Methane emission intensity (PMI, g/ kg of milk) was defined as the ratio of PME divided by the daily milk yield (kg/d) recorded on the same test-day. This PMI was then log-transformed to be normally distributed and called log-transformed CH_4 intensity (LMI). Data were edited to include only cows with known birth date. Records from days in milk (DIM) lower than 5 d and greater than 365 d were eliminated. Age at the first calving (AFC) was restricted to the range of 540 to 1,200 d. Daily milk yield was restricted to the range of 3 to 99 kg. The final dataset consisted of 1,529,282 test-day records on 229,465 cows distributed in 1,530

R.F. Veerkamp and Y. de Haas (eds)

DOI: 10.3920/978-90-8686-940-4_27, © H. Atashi *et al.* 2022

herds collected from 2006 to 2021. More than 95% of cows included were Holstein. Genotypic data of 730,539 SNPs were available for 7,381 animals. SNPs with minor allele frequency less than 5%, and those that did not pass the Hardy-Weinberg equilibrium were excluded. In total, 565,049 SNPs were kept for genomic analyses.

Variance component estimation. The (co)variance components and genomic breeding values for PME and LMI were estimated based on the integration of the random regression test-day model (RR-TDM) into single-step GBLUP procedure (SS RR-TDM) using the following model:

$$y_{ijklm} = \mu + HTD_i + \sum_{b=0}^{4} AS_j \phi_b(t) + \sum_{b=0}^{2} HY_k \phi_b(t) + \sum_{b=0}^{2} pe_l \phi_b(t) + \sum_{b=0}^{2} a_l \phi_b(t) + e_{ijklm} \quad (1)$$

where y_{ijklm} is the test-day record (PME or LMI) on DIM m of cow l, belonging to i^{th} class of HTD, j^{th} class of AS, and k^{th} class of HY, HTD is the fixed effect of herd-test-day, AS is the fixed effect of age-season of calving defined as following: age at calving class (10 classes) × season of calving (four seasons), $\sum_{b=0}^{4} AS_j \phi_b(t)$ is the fixed regression coefficients of the age-season at calving modelled using Legendre polynomials of order 4, $\sum_{b=0}^{2} HY_k \phi_b(t)$, $\sum_{b=0}^{2} pe_l \phi_b(t)$ and $\sum_{b=0}^{2} a_l \phi_b(t)$ are, respectively, the random regression coefficients of herd-year at calving (HY), permanent environment, and additive effects modelled using Legendre polynomials of order 2, and e_{ijklm} is the residual effect.

The variance components were estimated through the Bayesian Gibbs sampling method using a single chain of 100,000 iterations after a burn-in period of 20,000. The vector of genomic estimated breeding values (GEBV) of the included traits for each animal i, was estimated by multiplying the vector of predicted regression coefficients of the additive genetic effect by the matrix of Legendre orthogonal polynomial covariates; that is, $GEBV_i = T\hat{g}_i$, where $\hat{g}_i$ is the vector of predicted regression coefficients of the additive genetic effect for animal i, and T is a matrix of orthogonal covariates associated with the Legendre orthogonal polynomial functions. The SNP effects were estimated individually for the included traits using the postGSf90 software (Aguilar *et al.*, 2014). The animal effects were decomposed into those for genotyped (a_g) and ungenotyped animals (a_n). The animal effects of genotyped animals are a function of the SNP effects, $a_g = Zu$, where Z is a matrix relating genotypes of each locus and u is a vector of the SNP marker effect. The variance of animal effects was assumed as:

$$Var(\mathbf{a_g}) = Var(\mathbf{Zu}) = \mathbf{ZDZ'}\sigma_u^2 = \mathbf{G}\sigma_a^2 \quad (2)$$

where D is a diagonal matrix of weights for variances of markers (D = I), σ_u^2 is the genetic additive genetic variance captured by each SNP marker when the weighted relationship matrix (G) was built with no weight. The SNP effects were obtained using following equation:

$$\hat{u} = \lambda \mathbf{DZ'}\,\mathbf{G^{-1}}\,\hat{a}_g = \mathbf{DZ'}[\mathbf{ZDZ'}]^{-1}\hat{a}_g \quad (3)$$

where λ was defined by VanRaden (2008) as a normalizing constant, as described below:

$$\lambda = \frac{\sigma_u^2}{\sigma_a^2} = \frac{1}{\sum_{i=1}^{M} 2p_i(1-p_i)} \quad (4)$$

The percentage of genetic variance explained by the i^{th} genomic region was estimated as:

$$\frac{Var(a_i)}{\sigma_a^2} \times 100\% = \frac{Var(\sum_{j=1}^{50} Z_j \hat{u}_j)}{\sigma_a^2} \times 100 \quad (5)$$

where a_i is the genetic value of the i^{th} region that consists of 50 adjacent SNPs, σ_a^2 is the total genetic variance, Zj is the vector of the SNP content of the j^{th} SNP for all individuals, and $\hat{u}_j$ is the marker effect of the j^{th} SNP within the i^{th} region. The results were presented by the proportion of variance explained by each window of 50 adjacent SNPs (with an average size of 165 kb) and windows explaining for at least 1.0% of the total additive genetic variance were identified.

Results

Means (SD) for PME, PMI and LMI were 326.8 (67.65) g/d, 14.97 (5.70) g CH_4/kg of milk yield, and was 2.64 (0.35), respectively. Predicted CH_4 and LMI reached a maximum of 341.6 g/d and 2.83 by DIM 80 and 344, respectively (Figure 1-A). Heritability estimates for PME and LMI were relatively stable across lactation, with a mean h^2 of 0.14 and 0.24, and peak h^2 of 0.21 (DIM 186) and 0.30 (DIM 211), respectively (Figure 1-B, Table 1). Two regions on BTA14 (positions 1.86 to 2.12, and 1.48 to 1.68 Mb) were associated with both PME and LMI. A region between 144.38 to 144.46 Mb on BTA1 was associated with PME; and the region between 2.68 and 2.94 Mb on BTA14 was associated with LMI (Figure 1-C, and 1-D). Functional genes including *LY6D*, *GML*, *CYHR1*, *PPP1R16A*, *ARHGAP39*, *ZNF7*, *OPLAH*, *MAF1*, and *SPATC1* were identified as potential candidate genes for PME and LMI (Assembly UMD3.1).

Table 1. Mean (SD) phenotypic values and heritability estimates for considered milk mid-infrared-based CH_4 features.

	Phenotypic values		Heritability		
Trait	**Mean**	**SD**	**Mean**	**SD**	**range**
PME[1]	326.8	67.65	0.14	0.05	0.06-0.21
LMI[2]	2.64	0.36	0.24	0.05	0.16-0.30

[1] Methane emission (PME, g/d) predicted from the recorded milk MIR spectra.

[2] Log-transformed CH_4 intensity (LMI) based on the ratio of PME divided by the daily milk yield (kg/d).

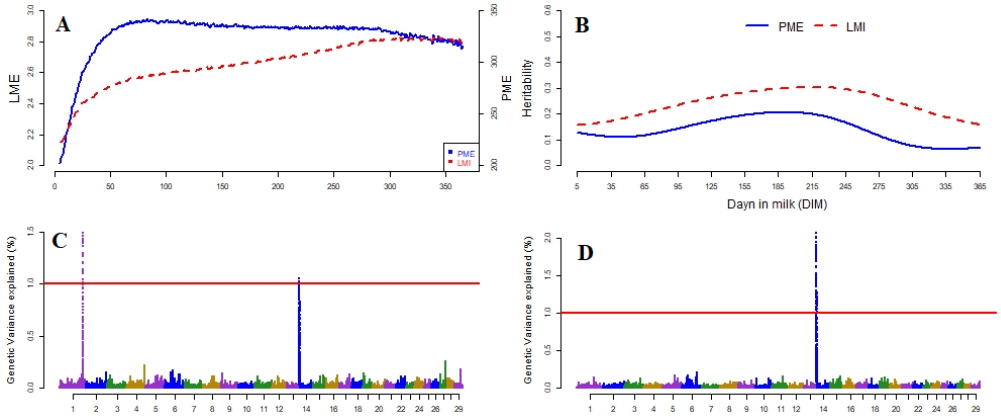

Figure 1. The lactation curves for PME (solid) and LMI (dashed) (A). Heritability estimates for PME (solid) and LMI (dashed) throughout lactation (B). Manhattan plots for PME (C) and LMI (D). The x-axis indicates the BTA number and the y-axis indicates the percentage of additive genetic variance explained by the window of 50 SNPs. The PME is the CH_4 emission (g/d) and the LMI is the log transformed of CH_4 intensity (g CH_4/kg of milk yield).

Discussion

We aimed to estimate genetic parameters and identify genomic regions associated with CH_4 emission in Walloon dairy cows. The mean PME found (326.8 g/d) was in the range reported by previous studies (Pickering *et al.*, 2015; Kandel *et al.*, 2017). Moderate h^2 were estimated for PME and LMI in agreement with the literature (Kandel *et al.*, 2017; Sypniewski *et al.*, 2021). Three genomic regions were found on BTA14 which, when combined, explained 2.10 and 3.35% of the total genetic variance of PME and LMI, respectively. The genomic regions and genes associated with PME and LMI have been previously reported to be associated with milk yield and composition in dairy cows (Bennewitz *et al.*, 2004; Atashi *et al.*, 2020). However, this is the first report on the association between these regions and CH_4 emission in cattle. In conclusion, the moderate h^2 found indicates that the studied traits can be introduced in the breeding programs. Further studies must be conducted to evaluate the relationship of our mid-infrared-based CH_4 traits to direct CH_4 traits and the effect of their inclusion in current breeding objectives. Moreover, large database could be used for genome-enhanced consulting systems to reduce methane emissions.

Acknowledgements

Project supported by the Walloon Government (SPW-DGARNE, Namur, Belgium) through the WALLeSmart Project (D31-1392). H. Wilmot, as Research Fellow acknowledges the support of the Fonds de la Recherche Scientifique – FNRS (Brussels, Belgium).

References

Aguilar I., Misztal I., Tsuruta S., Legarra A. and Wang, H. (2014). Proc. of the 10th WCGALP, Vancouver, Canada.

Atashi H., Salavati M., De Koster J., Ehrlich J., Crowe, M. *et al.* (2020). J. Anim. Breed. Genet. 137, 292-304. https://doi.org/10.1111/jbg.12442

Bennewitz J., Reinsch N., Paul S., Looft C., Kaupe B. *et al.* (2004). J. Dairy Sci. 87, 431-442. https://doi.org/10.3168/jds.s0022-0302(04)73182-3

Cottle D.J., Nolan J.V., and Wiedemann S.G. (2011). Anim. Prod. Sci. 51: 491-514. https://doi.org/10.1071/AN10163

De Haas Y., Veerkamp R., de Jong G., and Aldridge M. (2021). Animal 15: 100294. https://doi.org/10.1016/j.animal.2021.100294

Gerber P.J., Steinfeld H., Henderson B., Mottet A., Opio, *et al. (*2013). Tackling climate change through livestock – A global assessment of emissions and mitigation opportunities. Food and Agriculture Organization of the United Nations (FAO), Rome.

Kandel P., Vanrobays M.L., Vanlierde A., Dehareng F., Froidmont E. *et al.* (2017). J. Dairy Sci. 100: 5578-5591. https://doi.org/10.3168/jds.2016-11954

Pickering NK., Chagunda M.G., Banos G., Mrode R., McEwan J.C. *et al.* (2015). J. Anim. Sci. 93, 11-20.https://doi.org/10.2527/jas.2014-8302

Sypniewski M., Strabel T., and Pszczola M. (2021). Animals 11: 3175. https://doi.org/10.3390/ani11113175.

Vanlierde A, Dehareng F, Gengler N, Froidmont E, McParland S. et al (2021). J. Sci. Food Agric. 101(8), 3394-3403. https://doi.org/10.1002/jsfa.10969

VanRaden P.M. (2008). J Dairy Sci. 91: 4414-4423. https://doi.org/10.3168/jds.2007-0980