

RESUMEN

El análisis del control genético del hospedador sobre su microbiota ha sido señalado recientemente como un tema prometedor en diferentes campos de estudio. La relación entre el holobionte hospedador-microbioma y los fenotipos en el ganado lechero podría conducir a nuevos conocimientos en los programas de selección genética. El análisis de la microbiota implica lidiar con bases de datos composicionales, entre otras problemáticas relacionadas con la inferencia sobre la causalidad de los microbios sobre los rasgos de interés, el efecto de un solo microbio frente al efecto de toda la microbiota en los caracteres relevantes, o la agregación de toda la microbiota para abordar la implementación en los programas de selección genética. Las técnicas óptimas para analizar este tipo de datos no están bien definidas aún. Por lo tanto, se deben utilizar diferentes enfoques estadísticos para tratar mejor este tipo de datos. Dentro de esta tesis doctoral, se realizó la estimación y análisis a través de diferentes enfoques estadísticos con el objetivo de desentrañar el control genético del hospedador sobre la microbiota en ganado lechero. Además, se analizó el rasgo de concentración de metano como un fenotipo potencial para ser incluido en el programa de mejora de ganado lechero español. Un total de 437 vacas de 14 granjas ganaderas comerciales de ganado lechero de la región Norte de España (País Vasco, Cantabria, Navarra y Gerona) participaron en el proyecto METALGEN, que se articuló en tres subproyectos. Esta tesis se desarrolló dentro del subproyecto uno (La base (meta)-genómica del binomio animal – microbioma, y su relación con la eficiencia alimentaria y emisiones de metano en la raza Holstein-Frisona) del proyecto METALGEN, liderado por INIA. Los animales fueron genotipados por CONAFE usando los chips de baja densidad EuroG 10K y EuroG LD 12k (Illumina, San Diego, California, EE. UU.), Y luego imputados al chip Bovine 50k SNP (Illumina, San Diego, California, EE. UU.) usando su población de referencia para obtener genotipos que contienen 54.609 SNP. La composición microbiana de cada vaca se obtuvo a partir de la secuenciación completa del metagenoma de muestras de contenido ruminal utilizando un dispositivo MinION de Oxford Nanopore Technologies. La concentración de metano se midió con el monitor de gas infrarrojo Guardian® NG de Edinburgh Sensors durante las visitas de las vacas al sistema automatizado de ordeño. El metano fue semanalmente promediado a partir de los picos de cada vaca durante un periodo de 2 a 3 semanas. Concentración (ppm CH₄) e intensidad (ppm CH₄/ kg leche) fueron utilizados como fenotipos.

Se estimaron factores de riesgo para concentración de metano e intensidad de metano utilizando variables fenotípicas de producción de leche (kg), grasa láctea (%), proteína láctea (%), caracteres de conformación (n=18); escalas GEBVs para concentración de metano e intensidad de metano previamente generadas a partir de un ssGBLUP; así como la transformación logarítmica centrada (CLR) de la composición de la microbiota a nivel de filo (n=86) y género (n=1240), como factores de riesgo potenciales. Las vacas fueron clasificadas en cuartiles (baja, media-baja, media-alta y alta), según registros individuales de las mediciones de metano y la producción de leche promedio en el caso de intensidad de metano. Un abordaje de modelo umbral, bajo un marco Bayesiano de Cadenas de Markov Monte Carlo (MCMC), fue utilizado para determinar factores de riesgo para ser clasificada como alta para ambos caracteres. Las vacas de alta producción tuvieron menor probabilidad de ser clasificadas en la categoría alta para intensidad de metano. Los caracteres de estructura y capacidad corporal fueron positivamente correlacionados a una probabilidad incrementada de estar en la categoría más alta de concentración de metano, mientras que caracteres de buena conformación mamaria fueron factores protectores en contra de estar en la categoría alta de intensidad de metano. Mayores GEBVs también fueron factores protectores para ambos caracteres, reduciendo la probabilidad de ser clasificadas como altas emisoras de metano. Mayor abundancia relativa de la mayoría de los eucariotas (principalmente protozoos ciliados y hongos) y algunas arqueas (*Methanobrevibacter spp.*, *Methanothermus spp.* y *Methanosphaera spp.*) fueron factores de riesgo para ser clasificadas en la categoría alta.

El abordaje de factores de riesgo, utilizando GEBVs, permitió asociar el efecto genético del hospedador sobre los fenotipos, así como el efecto de la microbiota sobre los fenotipos; sin embargo, careció de información con relación al efecto genético del hospedador sobre el microbioma. Con el objetivo de abordar esta limitación, se propuso un conjunto de modelos de ecuaciones estructurales (SEM) de tipo recursivo dentro de un marco de Cadenas de Markov Monte Carlo (MCMC) para analizar conjuntamente la relación hospedador-metagenoma-fenotipo. Se estableció un modelo bivariado no-recursivo como punto de referencia. La abundancia relativa de microorganismos ruminales (RA) y CH₄ fueron incluidos como fenotipos en ambos sets de modelos mixtos animales bivariados (recursivos y no-recursivos). La genética del hospedador fue incluida dentro de los modelos a través de la matriz de relaciones genómicas entre animales, permitiendo estimar componentes de varianza. La heredabilidad de CH₄ se estimó en 0,12 ± 0,01 en ambos modelos, recursivo y bivariado. Asimismo, las estimaciones de

heredabilidad para la abundancia relativa de los taxones se superpusieron entre los modelos y variaron entre 0.08 y 0.48. Las correlaciones genéticas entre la composición microbiana y el CH₄ variaron de -0,76 a 0,65 en el modelo bivariado no recursivo y de -0,68 a 0,69 en el modelo recursivo. Independientemente del modelo estadístico utilizado, se estimaron consistentemente correlaciones genéticas positivas con metano para los 7 géneros pertenecientes al filo Ciliophora, así como para los géneros pertenecientes a los filos Euryarchaeota (*Methanobrevibacter sp.*), Chytridiomycota (*Neocallimastix sp.*) y Fibrobacteres (*Fibrobacter sp.*).

El análisis de taxones únicos informó acerca de relaciones independientes entre el fenotipo y un microorganismo a la vez, así como lo concerniente al efecto genético del hospedador sobre un microorganismo dado. Sin embargo, careció de información exhaustiva acerca del efecto de toda la microbiota sobre el fenotipo y el efecto genético sobre el microbioma completo. Para abordar el efecto conjunto de toda la microbiota sobre el fenotipo, la microbiota debe tomar una forma conformable para ser incluida dentro de modelos animales genéticos. Para esto, doce matrices de relación de microbiota (K) fueron construidas a partir de diferentes métricas de distancia del microbioma, con el objetivo de comparar su desempeño dentro de un marco de estimación de componentes de varianza para CH₄ y toda la microbiota. Análisis de simulación (n = 1000) y datos reales fueron desarrollados considerando cuatro modelos posibles: un modelo genómico aditivo (GBLUP), un modelo de microbioma (MBLUP), un modelo de efectos genéticos y microbioma (HBLUP) y un modelo de efectos de interacción genético, microbioma y genético × microbioma (HiBLUP). Todos los modelos se implementaron dentro de un marco Bayesiano utilizando el paquete BGLR en R. Las (Ks) de escalado multidimensional (MDS), el análisis de redundancia (RDA) y el análisis de correspondencia restringida (CCA) funcionaron mejor en la simulación para estimar la heredabilidad y la microbiabilidad. Los mismos métodos para construir Ks estuvieron entre los modelos más plausibles en los datos reales, de acuerdo con el criterio de información de desviación (DIC). El DIC también fue utilizado para obtener el modelo más plausible, que fue el HiBLUP. Un nuevo término “*Holobiabilidad*” fue definido para referirse a la proporción de la varianza atribuible a los efectos del holobionte hospedador-microbioma. Las estimaciones a partir de datos reales usando HiBLUP variaron dependiendo de la K utilizada y estuvieron entre 0.15-0.17, 0.15-0.21 y 0.42-0.59 para heredabilidad, microbiabilidad y holobiabilidad, respectivamente

A pesar de que el uso de Ks permitió estimar el efecto del microbioma completo sobre el fenotipo, faltó información concerniente al efecto genético del hospedador sobre el microbioma completo. Para lidiar con este problema, el microbioma completo no solo necesitaba ser conformable, como en forma de K, sino que necesitaba estar agregado en vectores, para poder ser tratados como un fenotipo. El conjunto de datos de microbioma fue agregado a través de análisis de componentes principales (PCA), en pocos componentes principales (PCs) que fueron utilizados como aproximaciones del metagenoma central. El PCA permitió condensar la enorme y difusa información taxonómica y funcional del metagenoma en unos pocos PC. Se aplicaron modelos animales bivariados utilizando estos PC y la producción de metano como fenotipos. La variabilidad condensada en estos PC está controlada por el genoma de la vaca, con estimaciones de heredabilidad para el primer PC de $\sim 0,30$ en todos los niveles taxonómicos, con una gran probabilidad ($> 83\%$) de que la distribución posterior sea $> 0,20$ y con un intervalo de mayor densidad posterior al 95% (95% HPD) no conteniendo cero. La mayoría de las estimaciones de correlación genética entre PC1 y metano fueron grandes ($\geq 0,70$), con la mayor parte de la distribución posterior ($> 82\%$) siendo $> 0,50$ y con su 95% HPD no conteniendo cero.

Estos resultados sugieren que todo el metagenoma del rumen regula recursivamente las emisiones de metano en las vacas lecheras, y que tanto el CH₄ como las composiciones de la microbiota están parcialmente controladas por el genotipo del hospedador. El CH₄ fue positivamente asociado con la abundancia relativa (RA) de eucariotas (protozoos y hongos) en Filo, Clase, Orden, Familia y Genero. Las lecturas largas con nanoporos permitieron la caracterización del metagenoma central del rumen usando secuenciación del metagenoma completo, y las variables agregadas (PC) propuestas podrían ser usadas en programas de mejora de animales para reducir las emisiones de metano en las generaciones futuras.