

# Ruminant microbiome data are skewed and unFAIR, undermining their usefulness for improving sustainable production

Abimael Ortiz-Chura<sup>1</sup>, Milka Popova<sup>1</sup>, and Diego Morgavi<sup>1</sup>

<sup>1</sup>Université Clermont Auvergne, INRAE, VetAgro Sup, UMR Herbivores, F-63122 Saint-Genes-Champanelle, France.

## INTRODUCTION

Ruminant microbiomes differ among species, breeds, body sites, and among geographic locations as the feeding and husbandry conditions are different. This underscores the importance of having a good representation of ruminant microbiomes in their local environment to translate research findings into beneficial approaches.

**Objective:** To examine metadata from ruminant microbiome studies to determine global representativeness and summarize information by ruminant species, body site and geographic location.

## MATERIAL AND METHODS

We analyzed 47,628 sample metadata from ten farmed ruminant species using the International Nucleotide Sequence Database Collaboration database.

```
Search query:
"txid[Organism] AND biosample sra[filter] AND "public"[filter]"
(date of access to the NCBI BioSample database : 28.07.2022)
```

## RESULTS

- ✓ Cattle samples are overrepresented (2-fold) compared to their global population.
- ✓ Sheep and buffalo samples are underrepresented (~2-fold), whereas goat samples are strongly underrepresented (Fig.1).

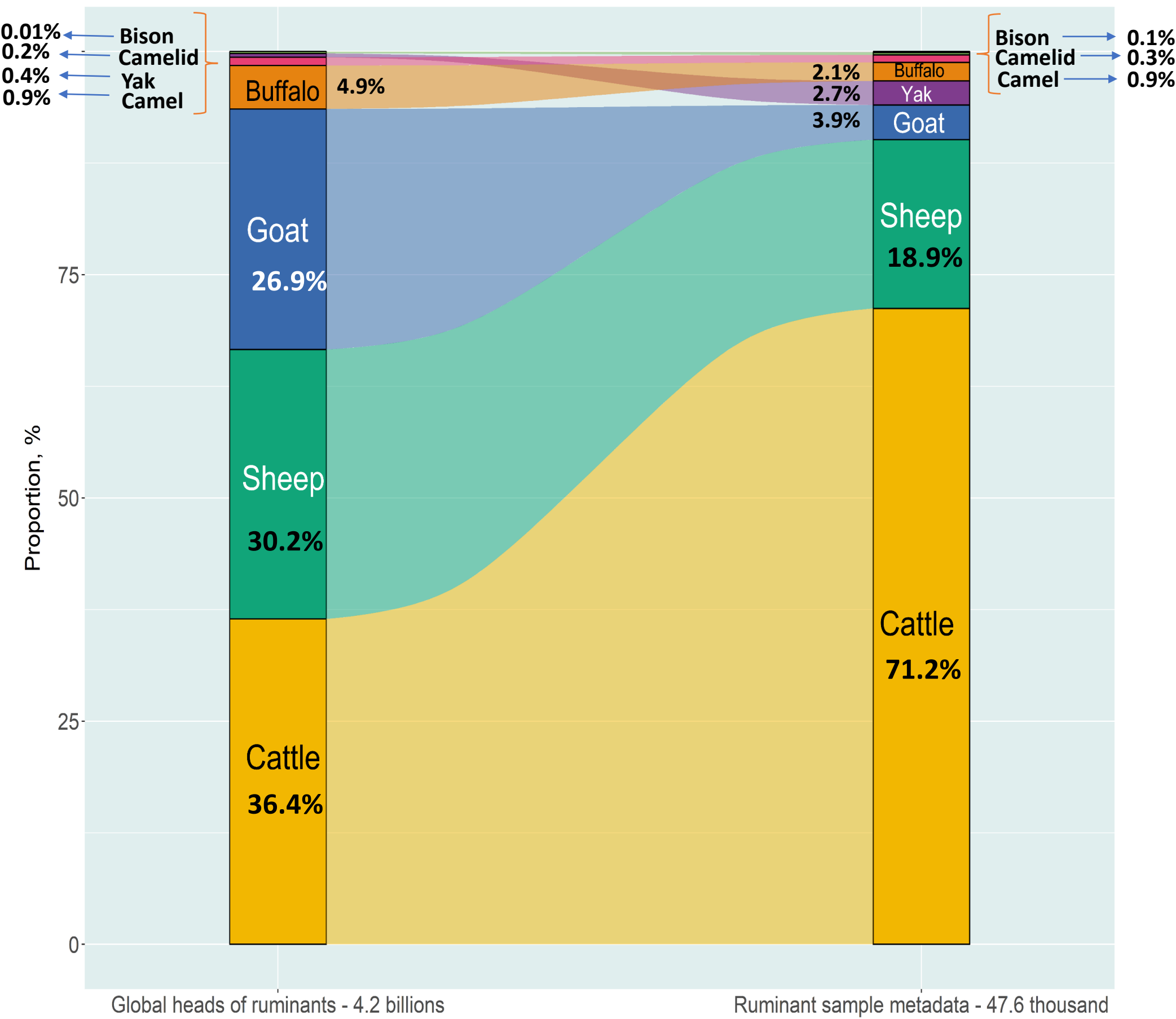


Figure 1. Comparison of proportion (%) between worldwide heads and sample metadata of ruminant species. FAOSTAT database (<https://www.fao.org/faostat/fr/#data/QI>), date of access: 26.10.2022

- ✓ Most samples originated in Western Europe, North America, Australasia and China but countries with the largest cattle and sheep populations from South America, Africa, South and Western Asia, and Eastern Europe are underrepresented (red color) (Fig. 2).

## CONCLUSION

We highlighted that certain ruminant species and geographical regions are underrepresented in the ruminant-associated microbiome dataset. This is an issue for the development of microbial strategies to meet sustainability challenges in areas with expanding livestock production systems. This study shows that incomplete metadata accompanies ruminant microbiome sequencing data in public repositories, hindering their reusability.

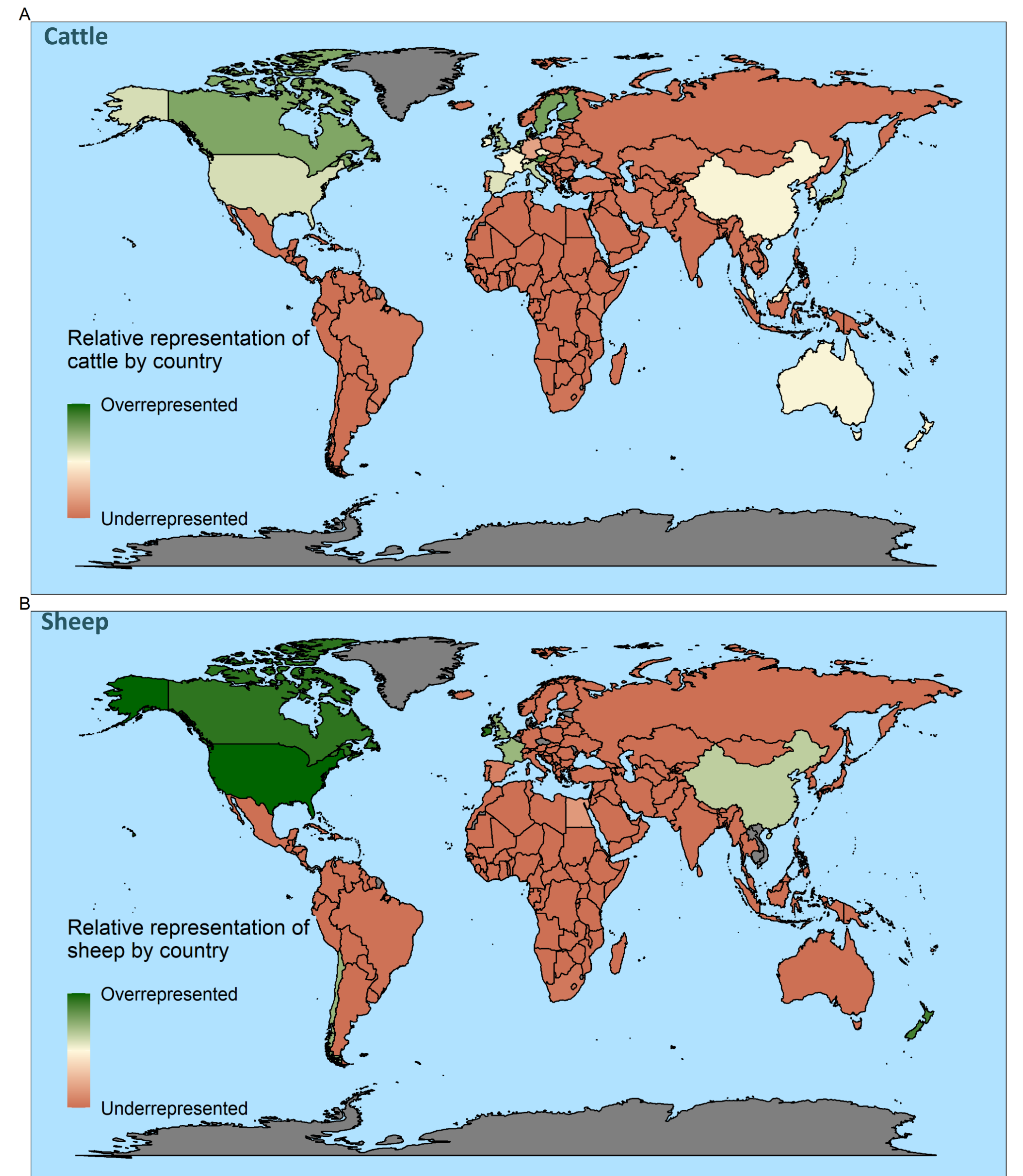


Fig. 2. Data from cattle (A) and sheep (B) associated microbiome relative to abundance of livestock population in the world. Green hues mark countries where microbiome samples are overrepresented, and red hues mark countries that are underrepresented. Countries with no data on cattle or sheep populations in the FAOSTAT database (accessed 26 October 2022) are marked in gray.

### Metadata of Cattle:

- ✓ Microbiomes from the gastrointestinal tract were the most frequently studied (~84% of the total).
- ✓ Dairy systems dominated over meat systems with Holstein as the main breed (70% of samples). Most samples come from cows (Fig. 3).
- ✓ More than 40% of the samples lacked basic information..

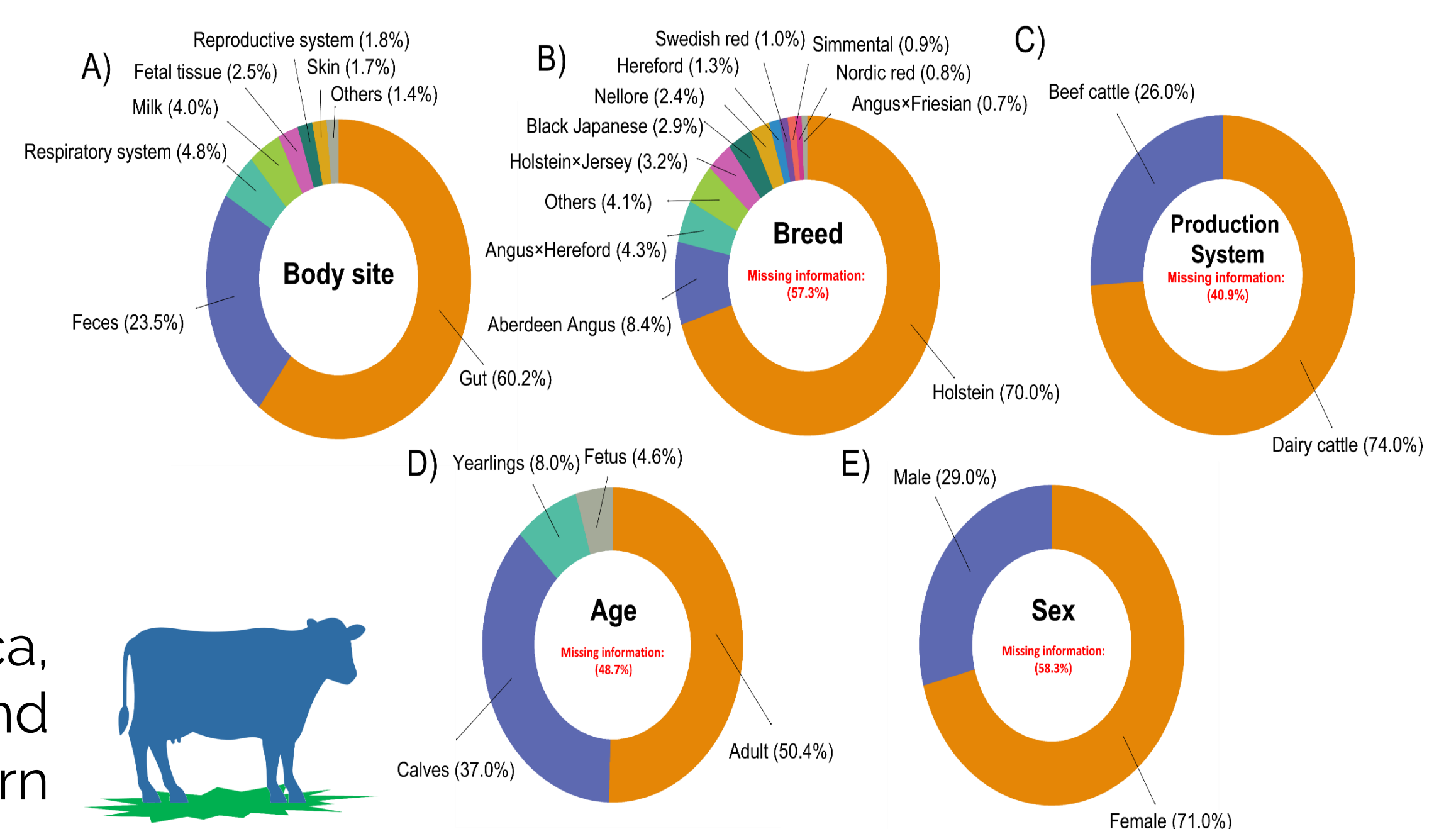


Fig. 3. Cattle sample distribution according to five different categories: body site (A), breed (B), production system (C), age (D), and sex (E).