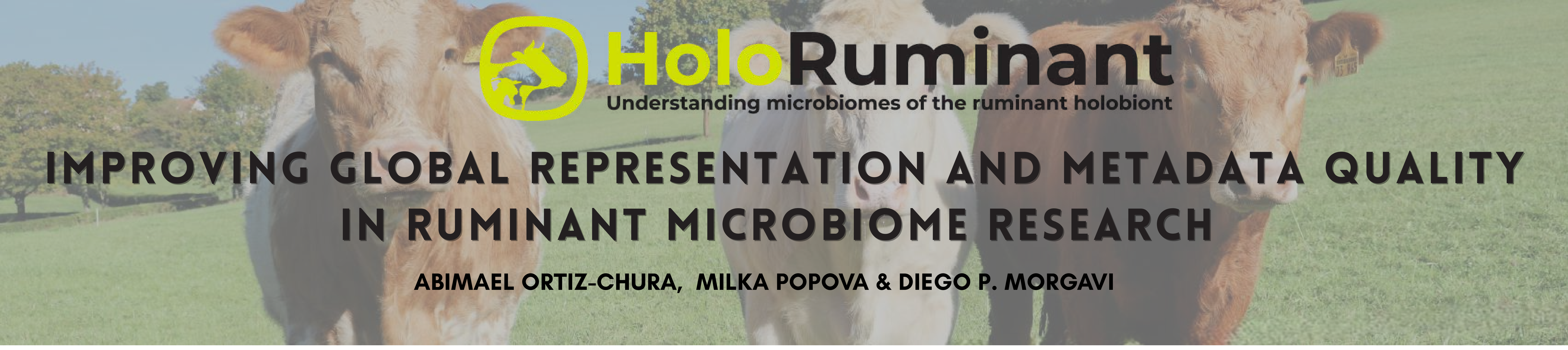

HoloRuminant
Understanding microbiomes of the ruminant holobiont

IMPROVING GLOBAL REPRESENTATION AND METADATA QUALITY IN RUMINANT MICROBIOME RESEARCH

ABIMAEL ORTIZ-CHURA, MILKA POPOVA & DIEGO P. MORGAVI

EXECUTIVE SUMMARY

The ruminant microbiome plays a vital role in livestock productivity, health, and environmental impact. Yet, data from farmed ruminant microbiomes are skewed: 71% of samples come from cattle, and over 90% of all samples originate from Western Europe, North America, Australasia, and China. Key regions with large ruminant populations such as South America, Africa, and parts of Asia are underrepresented. Furthermore, over 40% of existing samples lack basic metadata (e.g., breed, age, sex). This undermines efforts to develop microbiome-based strategies for sustainable livestock production. Addressing these gaps requires coordinated international action, promoting FAIR (Findable, Accessible, Interoperable, Reusable) data principles, and ensuring better representation of species, breeds, and production environments in global microbiome databases.

INTRODUCTION

Livestock contributes significantly to global food security and livelihoods, especially through ruminant animals like cattle, sheep, goats, and buffalo. Ruminant microbiome research is expanding, offering pathways to reduce greenhouse gas emissions, improve feed efficiency, and support animal health. However, current microbiome datasets are dominated by samples from developed countries and dairy breeds, limiting their applicability in regions where livestock productivity is lower, climate challenges are greater, and local breeds predominate. Without action, microbiome research risks excluding the very regions and systems where its benefits could be most impactful.

KEY FINDINGS

DATA SKEW AND UNDERREPRESENTATION:

71.2% of samples come from cattle, with goats, sheep, and buffalo notably underrepresented despite their global importance. Samples from Africa, South America, and parts of Asia are scarce.

METADATA QUALITY ISSUES:

Over 40% of samples lack critical host information such as breed, age, and sex. The taxonomic identifier in repositories is often inconsistent, making it difficult to find, reuse, and integrating data across

FAIR PRINCIPLES NOT CONSISTENTLY APPLIED:

Despite international guidelines, metadata standards are not uniformly adopted in ruminant microbiome research, particularly in public repositories. This limits transparency, reproducibility, and the development of AI-driven solutions

This project has received funding from the European Union's Horizon 2020. research and innovation programme under grant agreement № 101000213

POLICY RECOMMENDATIONS

① **PROMOTE FAIR DATA STANDARDS:**

Funding agencies and journals should require compliance with FAIR principles and the use of recognized ontologies (e.g., Animal Trait Ontology, MIxS standards) as a condition for project funding and publication.

② **SUPPORT INCLUSIVE DATA COLLECTION INITIATIVES:**

International and regional organisations (e.g., FAO, CGIAR) should fund targeted microbiome sampling campaigns focusing on underrepresented regions (Africa, Latin America, Asia) and species (goats, buffalo, indigenous breeds).

③ **REQUIRE MINIMUM METADATA FIELDS:**

Public repositories must mandate core metadata fields: species, breed, age, sex, production system, and sampling environment

④ **BUILD CAPACITY IN UNDERREPRESENTED REGIONS:**

Invest in training programs on microbiome sampling, data management plan, and FAIR compliance for researchers and institutions in developing countries.

⑤ **FOSTER INTERNATIONAL COLLABORATION:**

Encourage cross-border projects to harmonize data collection and metadata protocols, leveraging platforms such as the International Nucleotide Sequence Database Collaboration (INSDC).

IMPLEMENTATION PRIORITIES: STRENGTHENING GLOBAL MICROBIOME KNOWLEDGE GAPS

- **Address Representation Bias:** Prioritise microbiome studies on indigenous breeds, minor ruminant species, and production systems in regions highly exposed to climate change and food security challenges.
- **Expand Body Site Sampling:** Promote microbiome research beyond the gastrointestinal tract, targeting respiratory, skin, reproductive systems, and milk microbiomes linked to animal health and disease.
- **Integrate Longitudinal Study Designs:** Encourage the collection of time-series microbiome data to capture dynamic changes associated with feeding, health status, and environmental stressors.
- **Facilitate Ontology Harmonisation:** Support the development and uptake of livestock-specific ontologies for breeds, production systems, and health traits, ensuring consistency in metadata quality across international repositories.
- **Encourage Cross-Sector Collaboration:** Align microbiome data standards with public health and environmental monitoring initiatives, supporting a One Health approach.

CONCLUSION

The global livestock sector cannot afford to leave large regions and species behind in microbiome research. Improving the representativeness and metadata quality of ruminant microbiome data is both feasible and necessary to support sustainable livestock production worldwide. Coordinated international efforts, policy support, and adherence to FAIR principles are essential next steps.

READ FULL ARTICLE HERE

This project has received funding from the European Union's Horizon 2020. research and innovation programme under grant agreement № 101000213