

# Integrative Multi-Cohort Metagenomic Analysis Reveals Microbiome Signatures Predictive of Colorectal Cancer

Xueyao Wang<sup>1,2</sup>, Amir Ata Saei<sup>1,2</sup>

1. Centre for Translational Microbiome Research (CTMR)  
2. Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm, Sweden

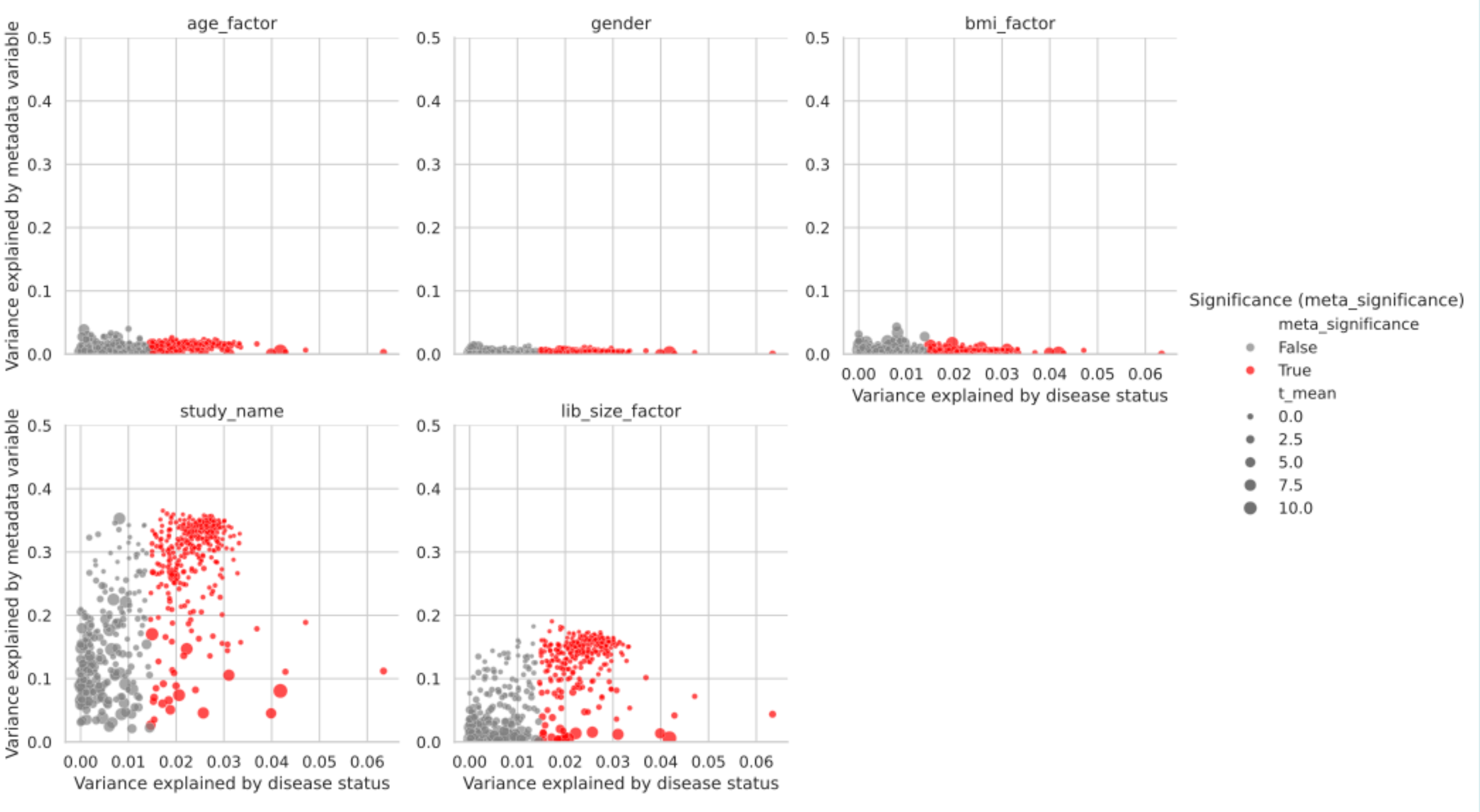
## Backgrounds

- The **gut microbiome**, with its thousands of prokaryotic organisms, is a treasure trove of metabolites that can influence human physiology, including colorectal cancer (CRC), the third leading cause of cancer-related deaths.
- However, most microbiome-CRC models are based on single-cohort data, suffer from poor **generalizability**, and lack **interpretability**.

## Aims

- To develop and validate a **robust, explainable machine learning (ML) framework** that integrates metagenomic data from diverse cohorts, enabling both CRC prediction and the discovery of microbiome signatures.

## Results



- The factor ‘**study\_center**’ have a predominant impact on species composition

Fig.2 ANOVA-type linear model to check the variance between CRC status and confounding factor

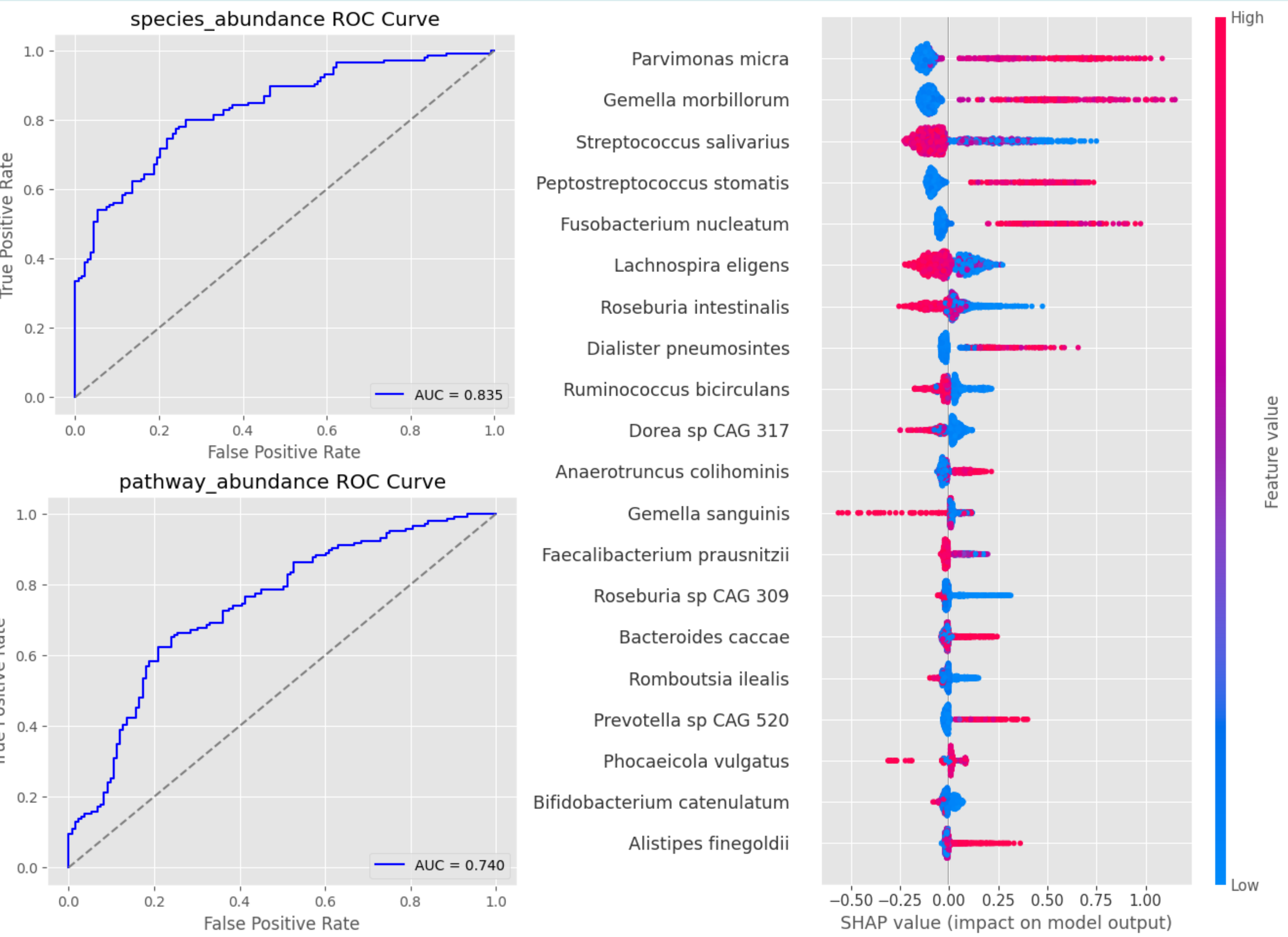


Fig.3 Taxonomic & Functional metagenomic XGBoost classification models generalized across studies

## Study Design & Methods

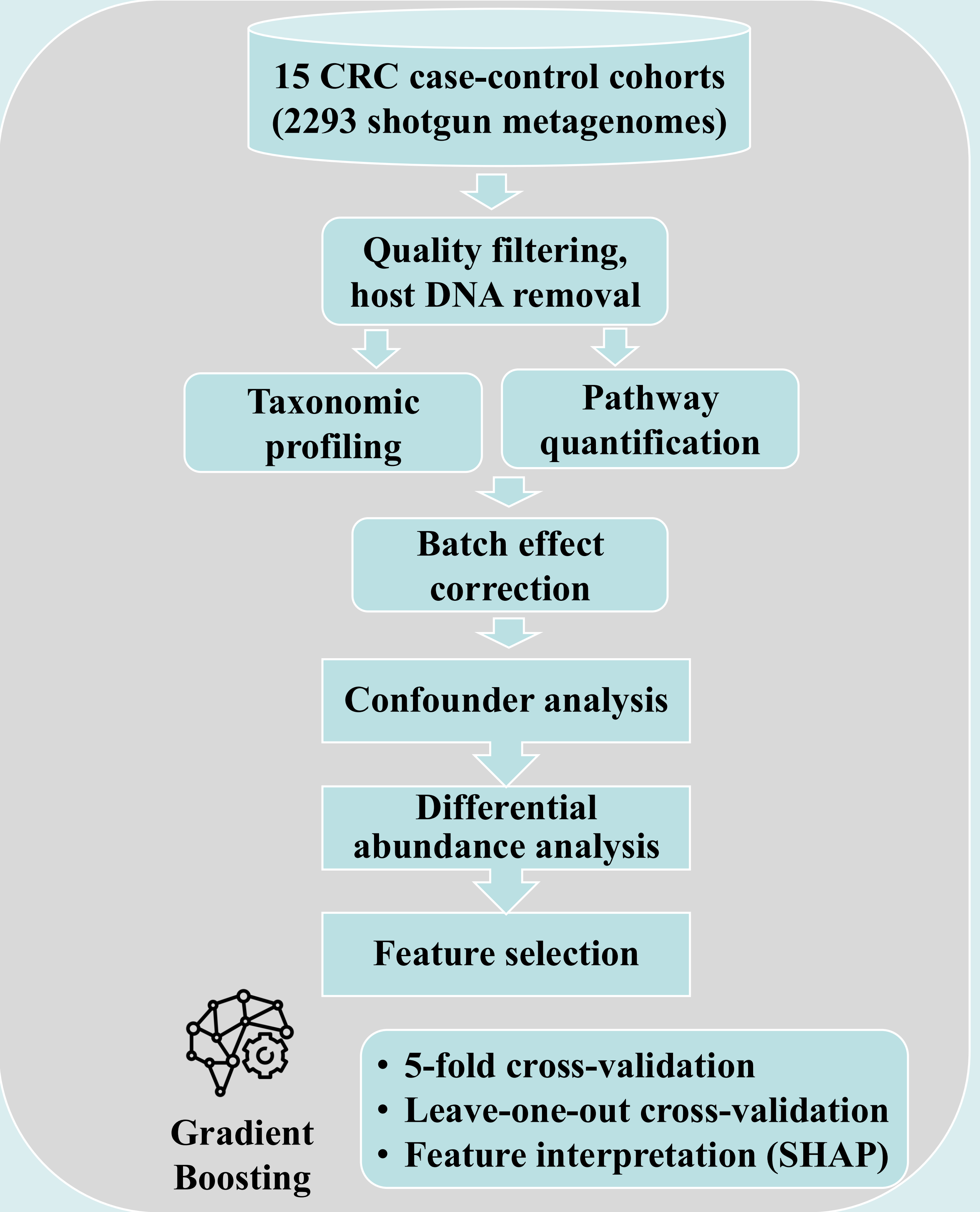


Fig.1 Standardized workflow for metagenomics data analysis in CRC

## Conclusion

- ❑ We demonstrate that integrating global metagenomic CRC datasets with rigorous batch correction yields highly generalizable ML models.
- ❑ Our models achieve high diagnostic performance across diverse populations and reveal microbial signatures associated with CRC.

## Future Directions

- ❑ Prospective validation of top-ranked microbial biomarkers
- ❑ Extend the model by integrating **metabolomic profiles** to enhance mechanistic interpretation and predictive power.



**E-mail:** xueyao.wang@ki.se  
**Visiting address:** Solnavägen 9, Biomedicum C8, 17165 Solna  
**Postal address:** C1 Mikrobiologi, tumör-och cellbiologi, C1 CTMR, 171 77 Stockholm



**Karolinska  
Institutet**