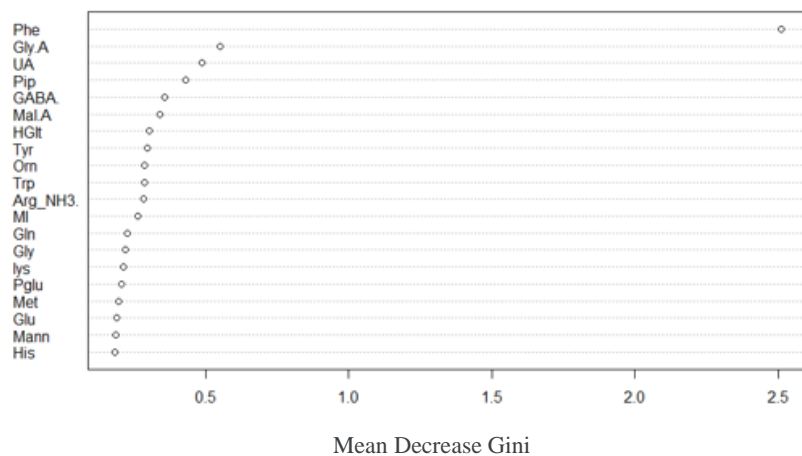
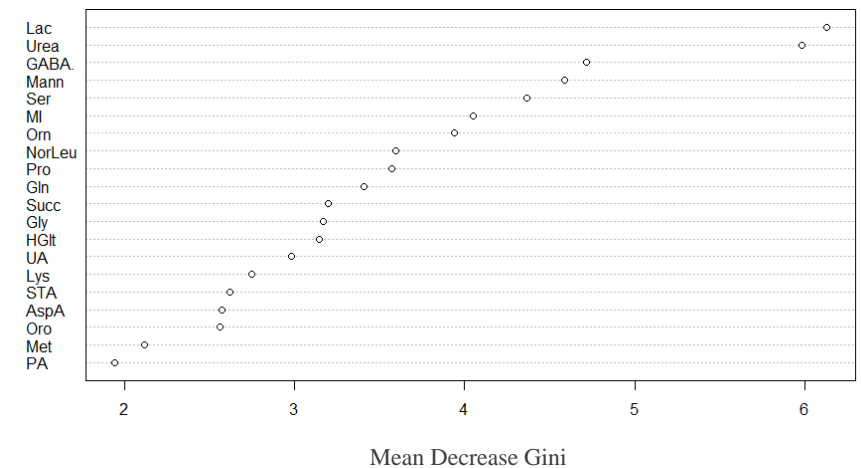


Supplementary Material

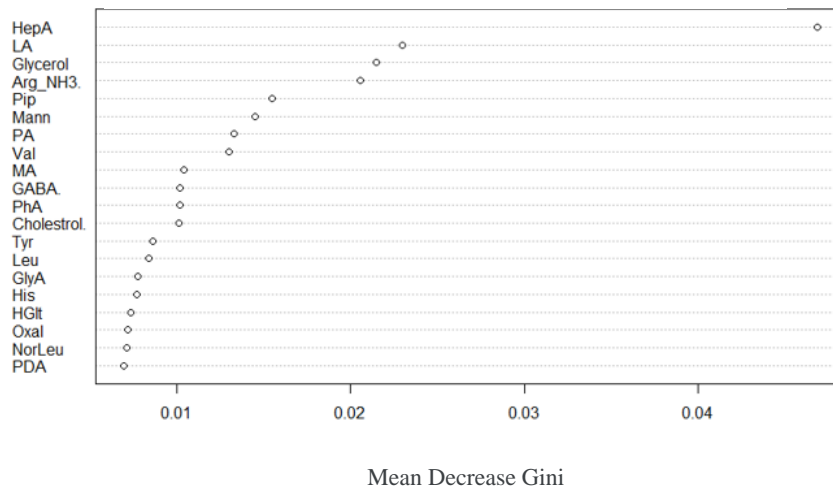
A Variable importance for *Thermus* (brand A)



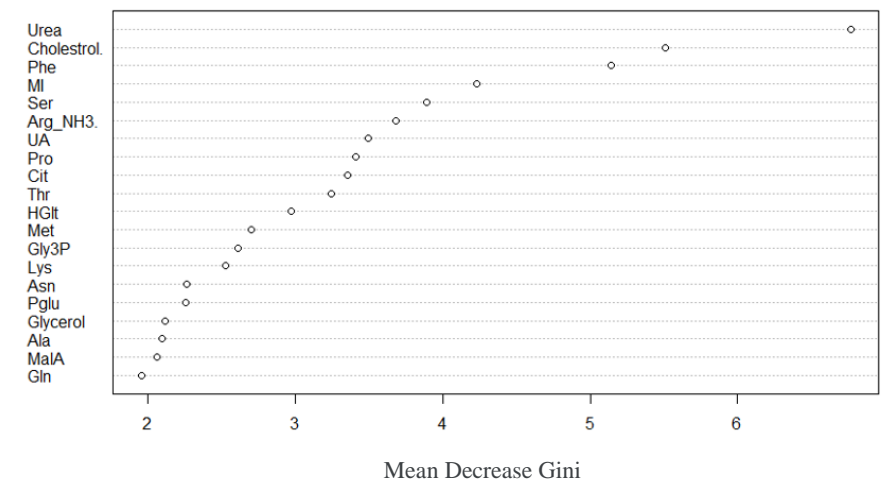
B Variable importance for *Streptococcus* (brand B)



C Variable importance for *Lactococcus* (brand B)



D Variable importance for *Streptococcus* (brand C)



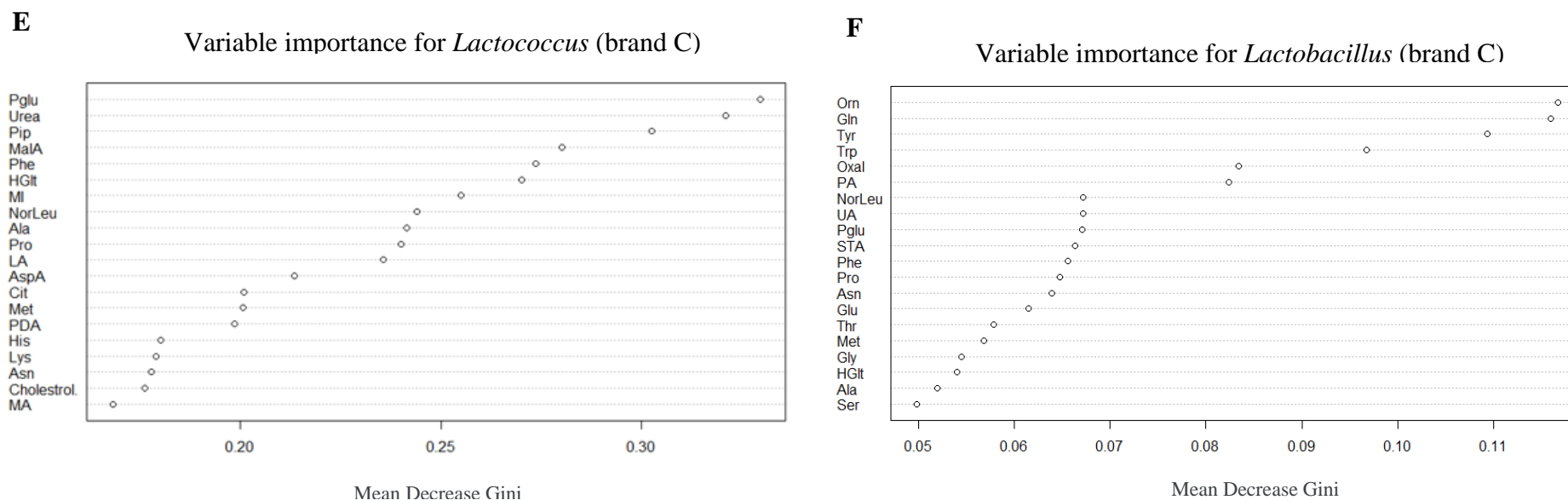


Figure S1. Random forest regression model to predict the association between cheese bacterial genera and metabolites. Variable importance plots showed which metabolites (response data) in the random forests optimised models had the highest association with the relative abundance of: (A) *Thermus* (brand A), (B) *Streptococcus* (brand B), (C) *Lactococcus* (brand B), (D) *Streptococcus* (brand C), (E) *Lactococcus* (brand C) and (F) *Lactobacillus* (brand C). The higher the Mean Gini value the more important the variable in the optimized model.