In silico prediction and in vitro assessment of microbial substrate utilisation among recently identified health-associated gut taxa

Cathy Lordan (Teagasc, Fermoy, Co. Cork, Ireland and University College Cork, Cork, Ireland), Aaron M Walsh (Teagasc, Fermoy, Co. Cork, Ireland), Dinesh Thapa (Teagasc, Fermoy, Co. Cork, Ireland), R. Paul Ross (APC Microbiome Ireland, Cork, Ireland), Paul D. Cotter (Teagasc, Fermoy, Co. Cork, Ireland and APC Microbiome Ireland, Cork, Ireland).

Introduction: Targeting more recently identified beneficial bacteria, including strict anaerobes, such as *Akkermansia muciniphila*, *Faecalibacterium prausnitzii* and *Eubacterium rectale*, through supplementation with different substrates, can enhance their growth and/or activity. Identification of substrates through in silico approaches can aid in elucidating which substrates may be best appropriate. However, validating these in silico-based tools is pertinent to developing robust genome-based predictions.

Methods: Seven strains from *A. muciniphila*, *E. rectale*, *F. prausnitzii*, *R. inulinivorans* and *R. bromii* were sequenced. The phenotypic microbial trait analyser, Traitar, and metabolic modelling tool, CarveMe, were used to predict substrate utilisation. Five simple sugars identified in both tools were chosen to be evaluated in vitro through growth experiments. Additionally, CAZymes were detected in the genomes using dbCAN2 to further elucidate carbohydrate degradation capacities on a genomic-level.

Results: Mean accuracy of the predictions acquired from Traitar (71.4%) and CarveMe (62.9%) were determined, although strain variability was observed. *R. bromii* 6883 had the capacity to consume all carbohydrates tested, whereas this was not the case for *R. bromii* X-30. CAZymes reflected strain-level differences between two *R. bromii* strains. A. muciniphila YL44 had the highest number of detected CAZymes, while *R. bromii* 6883 had the lowest. CAZyme number did not reflect the number of substrates consumed.

Discussion: Overall, both tools were fairly accurate in their predictions based on growth curves performed, demonstrating experimental validation is important in evaluating the accuracy of predictive in silico tools, while also highlighting the potential for identifying substrates that could enhance the growth and/or activity of microbes. Strain-level sensitivity of these tools shows potential with computational-based predictions for microbes well-studied and those less-studied.