

SUPPLEMENTARY TEXT 4

DIET-RESPONSIVE TAXA ARE NOT SPECIFIC TO NATIONALITY

We next checked if the diet-associated taxa differed across the nationalities. As noted above different nationalities were characterized by specific gut microbiome composition at baseline (**figure 1b; Supplementary figure 4a-b**) and different dietary adherence scores (Netherlands and UK having significantly lower scores, followed by Italy with Poland and France having the highest) (**Supplementary figure 7a**) (as also reported by previous studies on this cohort) (18, 19). If nationality-specific differences in diet-associated taxa existed, the performance of the prediction models might vary across nationalities. However, despite differences in the baseline gut microbiome compositions, there was no significant difference in the performance of the model (mean squared errors) for the different nationalities except for Netherlands (where the error rate was significantly high) (**Supplementary figure 7b**). This indicates that the identified diet-associated taxa were similar across most of the nationalities. The higher error-rate for the Netherlands could either be a consequence of lower adherence scores or a different set of diet-responsive taxa in these individuals. We tested the latter possibility by creating two different versions of iterative Random Forest models (two-fold cross validation) for Dutch subjects at baseline. While one version was created by using only the 129 diet-associated markers, the second version was built upon all OTUs besides the diet-associated markers. We observed that the iterative models built using only the 129 diet-associated markers had significantly higher correlation and significantly lower mean squared errors as compared to those obtained using all OTUs besides the diet-associated taxa (**Supplementary figures 7c-d**). This indicated that the diet-responsive taxa were not specific to nationality.