

Table S4. Species with sequenced genomes that were the closest match with annotated 454 reads from Rothamsted Park Grass soil. Percentage of annotated reads and standard deviation among the 13 pyrosequenced DNA pools that matched with each genome with a similarity at 96% or better and an E value of 10^{-5} or smaller).

Class	Species	Distribution (%)	SD (%)
Alphaproteobacteria	Bradyrhizobium japonicum USDA 110	0.086	0.041
Alphaproteobacteria	Bradyrhizobium sp. BTAi1	0.031	0.008
Acidobacteria	Solibacter usitatus Ellin6076	0.022	0.009
Alphaproteobacteria	Rhodopseudomonas palustris BisB18	0.019	0.006
Gammaproteobacteria	Pseudomonas fluorescens PfO-1	0.013	0.010
Alphaproteobacteria	Nitrobacter hamburgensis X14	0.013	0.003
Alphaproteobacteria	Rhodopseudomonas palustris BisA53	0.011	0.003
Betaproteobacteria	Burkholderia fungorum	0.009	0.008
Alphaproteobacteria	Rhodopseudomonas palustris BisB5	0.008	0.002
Gammaproteobacteria	Pseudomonas fluorescens SBW25	0.007	0.006
Alphaproteobacteria	Mesorhizobium loti MAFF303099	0.007	0.005
Bacteroidetes	Flavobacterium johnsonia johnsoniae UW101	0.007	0.018
Alphaproteobacteria	Nitrobacter winogradskyi Nb-255	0.006	0.002
Alphaproteobacteria	Nitrobacter sp. Nb-311A	0.005	0.001
Alphaproteobacteria	Rhodopseudomonas palustris HaA2	0.005	0.001
Gammaproteobacteria	Pseudomonas fluorescens Pf-5	0.005	0.005
Alphaproteobacteria	Rhizobium leguminosarum bv. viciae 3841	0.005	0.003
Actinobacteria (class)	Streptomyces avermitilis MA-4680	0.005	0.006
Acidobacteria	Acidobacteria bacterium Ellin345	0.004	0.003
Alphaproteobacteria	Rhodopseudomonas palustris CGA009	0.004	0.001
	Total 20 first detected species	0.273	0.070
	Total annotated sequences	0.354	0.089