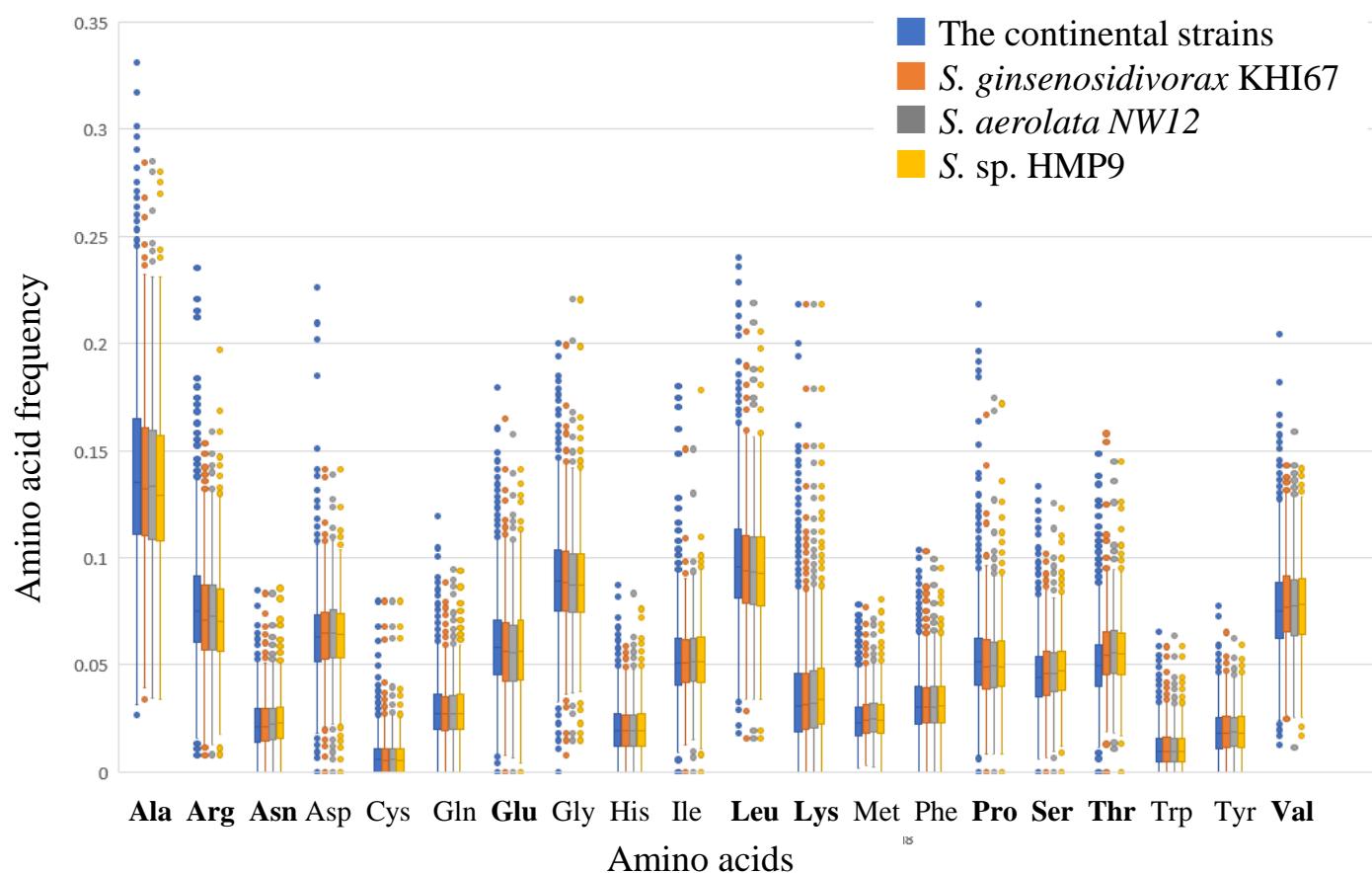


(A)



(B)

Amino acid	The continental strains ^{*1}	Mean Freq.				Significance test result (adjusted p-value)		
		S. ginsenosidivorax KHI67	S. aerolata NW12	S. sp. HMP9	The continental strains vs	S. ginsenosidivorax KHI67	S. aerolata NW12	S. sp. HMP9
Ala	0.1386	0.1352	0.1345	0.1327		0.054	0.023	0.13
Arg	0.0521	0.0730	0.0729	0.0716		9.50E-05	5.90E-05	1.70E-07
Asn	0.0355	0.0229	0.0233	0.0237		0.54	0.23	0.43
Asp	0.0224	0.0634	0.0636	0.0633		0.075	0.075	0.86
Cys	0.0767	0.0072	0.0073	0.0072		0.96	0.96	0.96
Gln	0.0321	0.0285	0.0288	0.0290		0.78	0.92	0.31
Glu	0.0619	0.0564	0.0562	0.0570		0.0015	0.0015	0.016
Gly	0.0288	0.0896	0.0884	0.0885		0.8	0.79	0.79
His	0.0593	0.0200	0.0201	0.0202		0.84	0.84	0.84
Ile	0.0893	0.0520	0.0529	0.0529		0.93	0.33	0.33
Leu	0.0521	0.0957	0.0951	0.0945		0.05	0.015	0.61
Lys	0.0202	0.0364	0.0366	0.0382		0.62	0.62	0.23
Met	0.0979	0.0255	0.0261	0.0254		0.02951	0.00098	0.84
Phe	0.0072	0.0324	0.0322	0.0323		0.9	0.9	0.9
Pro	0.0245	0.0503	0.0507	0.0505		0.053	0.15	0.27
Ser	0.0450	0.0464	0.0469	0.0480		0.0287	0.0028	9.30E-05
Thr	0.0504	0.0559	0.0564	0.0561		<2e-16	<2e-16	2.20E-13
Trp	0.0186	0.0113	0.0112	0.0112		0.93	0.93	0.93
Tyr	0.0761	0.0190	0.0189	0.0190		0.97	0.97	0.97
Val	0.0114	0.0791	0.0779	0.0786		0.00062	0.067	0.081

*1. The continental strains are *S. taxi* ATCC 55669, *S. hengshuiensis* WHSC-8, *S. sanxanigenens* NX02, *S. wittichii* RW1 and *S. sp.* MM-1.

Supplementary Figure 7. Distribution of amino acid frequencies in housekeeping genes. (A) Boxplot of amino acid frequencies in housekeeping genes of HMP9 and closely related strains. Bold amino acids: Amino acids with significant differences in Figure 9A. (B) The mean frequencies of amino acids and adjusted p-values of significance tests. The t tests for the continental strains vs HMP9 and closely related strains were performed for each amino acid. Yellow column: significant difference. Bold amino acids: Amino acids with significant differences in Figure 9A.