

**Table S1.** Quality statistics of all *E. coli* isolate genomes sequenced in this study

Isolate ID	Read quality parameters			Length <sup>d</sup>	De novo assembly parameters		
	Average read quality <sup>a</sup>	Average read coverage <sup>b</sup>	GC percentage <sup>c</sup>		N50 <sup>e</sup>	Nr contigs <sup>f</sup>	Average <i>de novo</i> coverage <sup>g</sup>
MF3R	35,6	434	50,6	5404153	212336	215	188
MF4HI	35,7	352	50,5	5442577	212336	218	187
EF115	35,7	713	50,6	5404545	212336	211	188
EF3	35,7	625	50,6	5405355	212336	212	188
EF40	35,7	490	50,7	5403543	212336	215	188
EF70	35,7	720	50,7	5434420	212336	229	187
EM68	35,7	289	50,8	5402725	212336	207	188
EM86	35,7	704	50,7	5434612	212336	231	187
ES107	35,8	624	51,1	5403648	212336	215	179
ES18	35,7	577	50,6	5432987	212336	231	187
ES91	35,7	488	50,6	5403705	212336	218	188
EF114	35,7	577	50,6	5379215	205777	219	189
EF123	35,7	638	50,6	5375682	205777	222	189
EF22	35,7	536	50,7	5374271	205777	211	189
EF24	35,7	620	50,6	5379348	205777	218	189
EF25	35,7	412	50,6	5376967	205777	211	189
EF42	35,7	648	50,7	5376994	212327	218	189
EF61	35,7	585	50,7	5375140	205777	226	189
EM12	35,7	602	50,7	5375845	212327	218	189
EM62	35,7	613	50,6	5378980	205777	221	189
ES105	35,7	532	50,6	5376589	205777	219	189
MB1H	35,6	379	50,5	5431836	212233	215	187
MD3H	35,6	503	50,7	5401494	212233	215	188
PD2R	35,6	501	50,5	5431149	212233	218	187
MB1HI	35,7	543	50,5	5431980	212233	213	187
EF19	35,7	393	50,6	5401090	212233	207	189
EF21	35,7	633	50,6	5402014	212233	216	188
EF34	35,4	289	50,5	5401854	212233	211	188
EF38	35,7	697	50,6	5401610	212233	222	188
EF41	35,7	667	50,6	5400763	212233	210	188
EF43	35,7	289	50,6	5313933	236432	239	191
EM114	35,7	531	50,7	5314730	205778	248	191
ES190	35,7	423	50,6	5312705	205778	240	191
PF2Br	35,6	489	50,6	5428494	149950	226	187
PA2H	35,7	396	50,6	5435346	150153	231	187
EM34	35,7	289	50,7	5404507	150153	231	188
EM44	35,7	617	50,7	5402857	150153	220	188
ES8	35,7	641	50,6	5404082	150103	229	188
PA2R	35,6	445	50,6	5372178	160593	250	189
PG1Fo	35,6	413	50,7	5344545	160593	228	190
PC4R	35,6	471	50,5	5343527	173667	239	190
PC4Br	35,7	437	50,7	5343841	160593	230	190
MB5R	35,6	477	50,5	5352210	160594	223	190
PE1R	35,7	522	50,6	5363855	205777	219	190

<sup>a</sup> The average read quality should be > 30; <sup>b</sup> the expected coverage for each base should be above 30; <sup>c</sup> the GC percentage should be  $\pm 50\%$ ; <sup>d</sup> the length of the assembled sequence should be between 3.9 Mb and 6.5 Mb; <sup>e</sup> the length of the median contig should be above > 52100; <sup>f</sup> the number of contigs in the assemble sequence should be <450; <sup>g</sup> the average *de novo* coverage should be > 30.