SUPPLEMENTARY INFORMATION

Functional Prediction of Hypothetical Proteins from *Shigella flexneri* and Validation of the Predicted Models by Using ROC Curve Analysis

Md. Amran Gazi[1*], Sultan Mahmud[2], Shah Mohammad Fahim[1], Mohammad Golam Kibria[2], Parag Palit[1], Md. Rezaul Islam[3], Humaira Rashid[2], Subhasish Das[1], Mustafa Mahfuz[1], Tahmeed Ahmeed[1]

[1]Nutrition and Clinical Services Division, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka 1212, Bangladesh, [2]Infectious Diseases Division, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka 1212, Bangladesh, [3]International Max Planck Research School, Grisebachstraße 5, 37077 Göttingen, Germany

https://doi.org/10.5808/GI.2018.16.4.e26

Supplementary Table 1. Scores of conserved domain search for 674 hypothetical proteins of *Shigella flexneri* serotype 2a strain 2457T using CDD-Blast, Pfam, HmmScan, SMART, and Scanprosite tools

Sl No.	Nucleotide ID	Accession ID_Protein	CDD Blast	Pfam	HmmScan	SMART	Scanprosite	Percentage
1	NC_004741.1	WP_000414150.1	0	0	0	0	0	0
2	NC_004741.1	WP_000738723.1	1	1	1	1	0	80
3	NC_004741.1	WP_001102351.1	1	1	1	1	0	80
4	NC_004741.1	WP_000843568.1	1	1	1	1	0	80
5	NC_004741.1	WP_032155592.1	0	0	0	0	0	0
6	NC_004741.1	WP_011110533.1	1	1	1	1	0	80
7	NC_004741.1	WP_032155631.1	0	0	0	0	0	0
8	NC_004741.1	WP_000196533.1	0	0	0	0	0	0
9	NC_004741.1	WP_001255305.1	0	0	0	0	0	0
10	NC_004741.1	WP_001303790.1	0	0	0	0	0	0
11	NC_004741.1	WP_000464383.1	1	0	0	0	0	20
12	NC_004741.1	WP_032155552.1	0	0	0	0	0	0
13	NC_004741.1	WP_005053505.1	1	1	1	1	0	80
14	NC_004741.1	WP_005094211.1	1	1	1	1	0	80
15	NC_004741.1	WP_001347263.1	0	0	0	0	0	0
16	NC_004741.1	WP_000964241.1	1	1	1	1	0	80
17	NC_004741.1	WP_000272188.1	1	1	1	1	0	80
18	NC_004741.1	WP_000417058.1	1	1	1	1	0	80
19	NC_004741.1	WP_005053355.1	1	1	1	1	1	100
20	NC_004741.1	WP_000402248.1	1	1	1	1	0	80
21	NC_004741.1	WP_024259146.1	0	0	0	0	0	0
22	NC_004741.1	WP_001276640.1	1	1	1	1	0	80
23	NC_004741.1	WP_000183806.1	1	1	1	1	0	80
24	NC_004741.1	WP_000343116.1	0	0	0	0	0	0

25	NC_004741.1	WP_000192453.1	1	1	1	1	0	80
26	NC_004741.1	WP_005053303.1	0	0	0	0	0	0
27	NC_004741.1	WP_000627639.1	1	1	1	1	0	80
28	NC_004741.1	WP_032333103.1	0	0	0	0	0	0
29	NC_004741.1	WP_001191885.1	0	0	0	0	0	0
30	NC_004741.1	WP_000092054.1	1	1	1	1	1	100
31	NC_004741.1	WP_001110751.1	0	0	0	0	0	0
32	NC_004741.1	WP_005053269.1	0	0	0	0	0	0
33	NC_004741.1	WP_000556489.1	0	0	0	0	0	0
34	NC_004741.1	WP_000003122.1	0	0	0	1	0	20
35	NC_004741.1	WP_001102097.1	1	1	1	1	0	80
36	NC_004741.1	WP_001201712.1	0	0	0	0	0	0
37	NC_004741.1	WP_000667623.1	0	0	0	0	0	0
38	NC_004741.1	WP_000343515.1	1	0	0	0	0	20
39	NC_004741.1	WP_005100645.1	1	1	1	1	1	80
40	NC_004741.1	WP_000554647.1	0	0	0	0	0	0
41	NC_004741.1	WP_005060548.1	0	0	0	0	0	0
42	NC_004741.1	WP_001142439.1	1	1	1	1	0	80
43	NC_004741.1	WP_032155747.1	0	0	0	0	0	0
44	NC_004741.1	WP_000941942.1	1	1	1	1	0	80
45	NC_004741.1	WP_000942006.1	1	1	1	1	0	80
46	NC_004741.1	WP_000194195.1	1	0	0	0	0	20
47	NC_004741.1	WP_001326928.1	0	1	0	1	0	40
48	NC_004741.1	WP_001382892.1	1	1	1	1	1	100
49	NC_004741.1	WP_000645013.1	0	0	0	0	0	0
50	NC_004741.1	WP_005060811.1	1	1	1	1	0	80
51	NC_004741.1	WP_005053055.1	0	0	0	0	0	0
52	NC_004741.1	WP_005053036.1	1	1	1	1	1	100

53	NC_004741.1	WP_001177122.1	0	0	0	0	0	0
54	NC_004741.1	WP_000680312.1	1	1	1	1	0	80
55	NC_004741.1	WP_005053020.1	1	1	1	1	0	80
56	NC_004741.1	WP_000779831.1	1	1	1	1	1	100
57	NC_004741.1	WP_000136192.1	1	1	1	1	0	80
58	NC_004741.1	WP_032140245.1	0	1	0	1	0	40
59	NC_004741.1	WP_001188905.1	1	1	1	1	0	80
60	NC_004741.1	WP_000970323.1	1	1	1	1	0	80
61	NC_004741.1	WP_000701358.1	1	1	1	1	0	80
62	NC_004741.1	WP_001224555.1	1	1	1	1	0	80
63	NC_004741.1	WP_000752567.1	1	1	1	1	0	80
64	NC_004741.1	WP_000360957.1	1	1	1	1	0	80
65	NC_004741.1	WP_064193753.1	0	0	0	0	0	0
66	NC_004741.1	WP_000460431.1	1	1	1	1	0	80
67	NC_004741.1	WP_000283754.1	1	1	1	1	0	80
68	NC_004741.1	WP_000287805.1	1	0	0	0	0	20
69	NC_004741.1	WP_001301130.1	0	0	0	0	0	0
70	NC_004741.1	WP_005083189.1	0	0	0	0	0	0
71	NC_004741.1	WP_001303843.1	0	0	0	0	0	0
72	NC_004741.1	WP_005049395.1	0	0	0	0	0	0
73	NC_004741.1	WP_005083246.1	0	0	0	0	0	0
74	NC_004741.1	WP_001188346.1	1	1	1	1	0	80
75	NC_004741.1	WP_001053303.1	1	1	1	1	0	80
76	NC_004741.1	WP_000807562.1	1	1	1	1	0	80
77	NC_004741.1	WP_025715253.1	0	0	0	0	0	0
78	NC_004741.1	WP_000405563.1	1	1	1	1	0	80
79	NC_004741.1	WP_001325427.1	0	0	0	0	0	0
80	NC_004741.1	WP_011110552.1	1	1	1	1	1	100

81	NC_004741.1	WP_000153125.1	1	1	1	1	0	80
82	NC_004741.1	WP_000232643.1	1	1	1	1	0	80
83	NC_004741.1	WP_005049464.1	0	0	0	1	0	20
84	NC_004741.1	WP_001030938.1	1	1	1	1	0	80
85	NC_004741.1	WP_000367140.1	1	1	1	1	0	80
86	NC_004741.1	WP_000578172.1	1	1	1	1	0	80
87	NC_004741.1	WP_001044870.1	1	1	1	1	0	80
88	NC_004741.1	WP_001269672.1	1	1	1	1	1	100
89	NC_004741.1	WP_000850550.1	1	1	1	1	0	80
90	NC_004741.1	WP_000073523.1	0	0	0	0	0	0
91	NC_004741.1	WP_005049496.1	0	0	0	0	0	0
92	NC_004741.1	WP_000627468.1	1	1	1	1	0	80
93	NC_004741.1	WP_005060669.1	0	0	0	0	0	0
94	NC_004741.1	WP_000873153.1	1	1	1	1	0	80
95	NC_004741.1	WP_000113500.1	1	1	1	1	0	80
96	NC_004741.1	WP_005020049.1	0	0	0	0	0	0
97	NC_004741.1	WP_005098291.1	0	0	0	0	0	0
98	NC_004741.1	WP_001108106.1	0	0	0	0	0	0
99	NC_004741.1	WP_000454800.1	1	1	1	1	0	80
100	NC_004741.1	WP_047199943.1	1	1	1	1	0	80
101	NC_004741.1	WP_001247854.1	1	1	1	1	1	100
102	NC_004741.1	WP_001343960.1	0	0	0	0	0	0
103	NC_004741.1	WP_005053437.1	0	0	0	0	0	0
104	NC_004741.1	WP_000551132.1	1	1	1	1	0	80
105	NC_004741.1	WP_000266134.1	1	0	0	0	0	20
106	NC_004741.1	WP_005048534.1	1	1	1	1	0	80
107	NC_004741.1	WP_000446914.1	1	1	1	1	0	80
108	NC_004741.1	WP_000871982.1	1	1	1	1	0	80

109	NC_004741.1	WP_000070107.1	1	1	1	1	1	100
110	NC_004741.1	WP_001336078.1	0	1	0	1	0	40
111	NC_004741.1	WP_005048500.1	1	1	1	1	0	80
112	NC_004741.1	WP_000849301.1	1	1	1	1	0	80
113	NC_004741.1	WP_000710620.1	1	1	1	1	0	80
114	NC_004741.1	WP_000188784.1	0	0	0	0	0	0
115	NC_004741.1	WP_000168813.1	1	1	1	1	0	80
116	NC_004741.1	WP_001295900.1	0	0	0	0	0	0
117	NC_004741.1	WP_000681108.1	1	1	1	1	0	80
118	NC_004741.1	WP_001201557.1	1	1	1	1	0	80
119	NC_004741.1	WP_000389260.1	1	1	1	1	0	80
120	NC_004741.1	WP_001303862.1	0	0	0	0	0	0
121	NC_004741.1	WP_001160722.1	1	1	1	1	0	80
122	NC_004741.1	WP_032155760.1	0	0	0	0	0	0
123	NC_004741.1	WP_000687442.1	0	0	0	0	0	0
124	NC_004741.1	WP_029716636.1	1	1	1	1	0	80
125	NC_004741.1	WP_005048249.1	0	0	1	0	0	20
126	NC_004741.1	WP_001118167.1	1	1	1	1	0	80
127	NC_004741.1	WP_000702036.1	1	1	1	1	0	80
128	NC_004741.1	WP_001091985.1	1	1	1	1	0	80
129	NC_004741.1	WP_001005968.1	0	0	0	0	0	0
130	NC_004741.1	WP_005051132.1	1	0	0	0	0	20
131	NC_004741.1	WP_005061679.1	1	1	1	1	0	80
132	NC_004741.1	WP_001039888.1	1	0	1	0	0	40
133	NC_004741.1	WP_000723623.1	1	1	1	1	0	80
134	NC_004741.1	WP_000959226.1	0	0	0	0	0	0
135	NC_004741.1	WP_000350058.1	1	1	1	1	0	80
136	NC_004741.1	WP_000196607.1	0	0	0	0	0	0

137	NC_004741.1	WP_000235193.1	1	0	0	0	0	20
138	NC_004741.1	WP_000224274.1	1	1	1	1	1	100
139	NC_004741.1	WP_001261235.1	1	1	1	1	0	80
140	NC_004741.1	WP_000847791.1	1	1	1	1	0	80
141	NC_004741.1	WP_001301416.1	0	0	0	0	0	0
142	NC_004741.1	WP_001038092.1	1	1	1	1	0	80
143	NC_004741.1	WP_005083611.1	0	0	0		0	0
144	NC_004741.1	WP_000505101.1	1	0	0	0	0	20
145	NC_004741.1	WP_000535353.1	1	0	0	0	0	20
146	NC_004741.1	WP_001143120.1	1	1	1	1	0	80
147	NC_004741.1	WP_000124106.1	0	0	0	0	0	0
148	NC_004741.1	WP_000611853.1	0	0	0	0	0	0
149	NC_004741.1	WP_001297187.1	1	1	1	1	0	80
150	NC_004741.1	WP_032155907.1	0	0	0	0	0	0
151	NC_004741.1	WP_001111218.1	0	0	0	0	0	0
152	NC_004741.1	WP_005047366.1	0	0		0	0	0
153	NC_004741.1	WP_000818776.1	0	0	0	0	0	0
154	NC_004741.1	WP_000749269.1	1	1	1	1	1	100
155	NC_004741.1	WP_000877111.1	1	1	1	1	0	80
156	NC_004741.1	WP_001295962.1	1	1	1	1	0	80
157	NC_004741.1	WP_000587933.1	1	1	1	1	0	80
158	NC_004741.1	WP_001043459.1	1	1	1	1	0	80
159	NC_004741.1	WP_032155646.1	0	0	0	0	0	0
160	NC_004741.1	WP_000103622.1	0	0	0	0	0	0
161	NC_004741.1	WP_005005155.1	0	0	0	0	0	0
162	NC_004741.1	WP_001204964.1	0	0	0	0	1	20
163	NC_004741.1	WP_000770157.1	1	1	1	1	0	80
164	NC_004741.1	WP_000557473.1	0	0	0	0	0	0

165	NC_004741.1	WP_001294167.1	0	0	0	0	0	0
166	NC_004741.1	WP_001132078.1	0	0	0	0	0	0
167	NC_004741.1	WP_000267598.1	1	1	1	1	0	80
168	NC_004741.1	WP_000134107.1	1	1	1	1	0	80
169	NC_004741.1	WP_001005703.1	0	0	0	0	0	0
170	NC_004741.1	WP_029716858.1	1	1	1	1	0	80
171	NC_004741.1	WP_000133415.1	0	0	0	0	0	0
172	NC_004741.1	WP_005047957.1	0	0	0	0	0	0
173	NC_004741.1	WP_032155828.1	0	0	0	0	0	0
174	NC_004741.1	WP_005047951.1	0	0	0	0	0	0
175	NC_004741.1	WP_001295611.1	1	1	1	1	0	80
176	NC_004741.1	WP_000122462.1	0	0	0	0	0	0
177	NC_004741.1	WP_005061990.1	0	0	0	0	0	0
178	NC_004741.1	WP_001125713.1	1	1	1	1	1	100
179	NC_004741.1	WP_000807626.1	1	1	1	1	0	80
180	NC_004741.1	WP_011069340.1	0	0	0	0	0	0
181	NC_004741.1	WP_000280742.1	0	0	0	0	0	0
182	NC_004741.1	WP_001257042.1	1	1	1	1	0	80
183	NC_004741.1	WP_000950192.1	1	1	1	1	0	80
184	NC_004741.1	WP_001169669.1	1	1	1	1	0	80
185	NC_004741.1	WP_000069487.1	0	0	0	0	0	0
186	NC_004741.1	WP_005105319.1	1	1	1	1	0	80
187	NC_004741.1	WP_001303937.1	0	0	0	0	0	0
188	NC_004741.1	WP_000967595.1	1	1	1	1	0	80
189	NC_004741.1	WP_000028536.1	1	1	1	1	0	80
190	NC_004741.1	WP_000807659.1	1	1	1	1	0	80
191	NC_004741.1	WP_001303289.1	0	0	0	0	0	0
192	NC_004741.1	WP_001031530.1	1	1	1	1	0	80

193	NC_004741.1	WP_014640285.1	0	0	0	0	0	0
194	NC_004741.1	WP_000233043.1	0	0	0	0	0	0
195	NC_004741.1	WP_000616081.1	0	0	0	0	0	0
196	NC_004741.1	WP_001288369.1	1	0	0	0	0	0
197	NC_004741.1	WP_001331106.1	1	1	1	1	0	80
198	NC_004741.1	WP_023636694.1	0	0	0	0	1	20
199	NC_004741.1	WP_000825769.1	1	1	1	1	0	80
200	NC_004741.1	WP_032155686.1	1	0	0	0	0	20
201	NC_004741.1	WP_000124119.1	0	0	0	0	0	0
202	NC_004741.1	WP_001296046.1	0	0	0	0	0	0
203	NC_004741.1	WP_005047705.1	1	0	0	0	0	20
204	NC_004741.1	WP_005047713.1	1	1	1	1	0	80
205	NC_004741.1	WP_023517638.1	1	1	1	1	0	80
206	NC_004741.1	WP_011069401.1	1	1	1	1	0	80
207	NC_004741.1	WP_000431885.1	1	1	1	1	0	80
208	NC_004741.1	WP_011069399.1	0	0	0	0	0	0
209	NC_004741.1	WP_001296778.1	0	0	0	0	0	80
210	NC_004741.1	WP_001077956.1	0	0	0	0	0	0
211	NC_004741.1	WP_000554382.1	1	0	0	0	0	20
212	NC_004741.1	WP_005049838.1	1	1	1	1	0	80
213	NC_004741.1	WP_001295499.1	1	1	1	1	0	80
214	NC_004741.1	WP_001043881.1	1	1	1	1	1	100
215	NC_004741.1	WP_001006860.1	1	1	1	1	0	80
216	NC_004741.1	WP_032155836.1	0	0	0	0	0	0
217	NC_004741.1	WP_000156246.1	1	1	1	1	0	80
218	NC_004741.1	WP_001306763.1	0	0	0	0	0	0
219	NC_004741.1	WP_001295493.1	1	1	1	1	1	100
220	NC_004741.1	WP_032155854.1	0	0	0	0	1	20

221	NC_004741.1	WP_000691930.1	1	1	1	1	1	100
222	NC_004741.1	WP_005126892.1	0	0	0	0	0	0
223	NC_004741.1	WP_000138039.1	1	1	1	1	0	80
224	NC_004741.1	WP_001046790.1	1	1	1	1	0	80
225	NC_004741.1	WP_001453023.1	0	0	0	0	0	0
226	NC_004741.1	WP_012602004.1	0	0	0	0	0	0
227	NC_004741.1	WP_000756955.1	1	1	1	1	0	80
228	NC_004741.1	WP_000085238.1	1	1	1	1	0	80
229	NC_004741.1	WP_001215295.1	1	0	0	0	0	20
230	NC_004741.1	WP_000077934.1	0	0	0	0	1	20
231	NC_004741.1	WP_032145487.1	0	1	0	1	0	40
232	NC_004741.1	WP_001297653.1	1	1	1	1	0	80
233	NC_004741.1	WP_000146138.1	1	1	1	1	0	80
234	NC_004741.1	WP_001142445.1	0	0	0	0	0	0
235	NC_004741.1	WP_005050031.1	1	1	1	1	0	80
236	NC_004741.1	WP_042003723.1	0	0	0	0	0	0
237	NC_004741.1	WP_000398613.1	0	0	0	0	0	0
238	NC_004741.1	WP_005062520.1	0	0	0	0	0	0
239	NC_004741.1	WP_000726666.1	1	1	1	1	0	80
240	NC_004741.1	WP_000874243.1	1	1	1	1	0	80
241	NC_004741.1	WP_045177689.1	0	0	0	0	0	0
242	NC_004741.1	WP_001265249.1	1	1	1	1	0	80
243	NC_004741.1	WP_000980987.1	0	0	0	0	0	0
244	NC_004741.1	WP_000214712.1	1	1	1	1	0	80
245	NC_004741.1	WP_001024558.1	1	1	1	1	0	80
246	NC_004741.1	WP_000901367.1	0	0	0	0	0	0
247	NC_004741.1	WP_000258546.1	0	0	0	0	0	0
248	NC_004741.1	WP_000957853.1	0	0	0	0	0	0

249	NC_004741.1	WP_005050130.1	1	1	1	1	0	80
250	NC_004741.1	WP_001295395.1	1	1	1	1	0	80
251	NC_004741.1	WP_000705197.1	1	1	1	1	0	80
252	NC_004741.1	WP_000234660.1	0	0	0	0	0	0
253	NC_004741.1	WP_000520318.1	0	0	1	1	0	40
254	NC_004741.1	WP_000207512.1	0	0	0	0	0	0
255	NC_004741.1	WP_000971490.1	1	0	1	0	0	40
256	NC_004741.1	WP_001240758.1	0	0	0	0	0	0
257	NC_004741.1	WP_000199921.1	0	0	0	0	0	0
258	NC_004741.1	WP_001091024.1	0	0	0	0	0	0
259	NC_004741.1	WP_000113584.1	0	0	0	0	0	0
260	NC_004741.1	WP_000091718.1	0	0	0	0	0	0
261	NC_004741.1	WP_001249851.1	0	0	0	0	0	0
262	NC_004741.1	WP_000233090.1	1	0	0	0	0	20
263	NC_004741.1	WP_000769323.1	1	1	1	1	0	80
264	NC_004741.1	WP_000524868.1	1	1	1	1	0	80
265	NC_004741.1	WP_000597196.1	1	1	1	1	1	100
266	NC_004741.1	WP_032155892.1	1	0	0	0	0	20
267	NC_004741.1	WP_000534313.1	1	1	1	1	0	80
268	NC_004741.1	WP_000212657.1	1	1	1	1	0	80
269	NC_004741.1	WP_000587595.1	1	0	0	0	0	20
270	NC_004741.1	WP_001344535.1	1	0	0	0	0	20
271	NC_004741.1	WP_000528342.1	1	1	1	1	0	80
272	NC_004741.1	WP_001296104.1	1	1	1	1	0	80
273	NC_004741.1	WP_000248636.1	1	1	1	1	1	100
274	NC_004741.1	WP_001301287.1	1	1	1	1	0	80
275	NC_004741.1	WP_032155900.1	0	0	0	0	0	0
276	NC_004741.1	WP_000627104.1	1	0	0	0	1	40

277	NC_004741.1	WP_000124121.1	0	0	1	0	0	20
278	NC_004741.1	WP_000018633.1	1	1	1	1	0	80
279	NC_004741.1	WP_012817775.1	0	0	0	0	0	0
280	NC_004741.1	WP_032155659.1	0	0	0	0	0	0
281	NC_004741.1	WP_001380520.1	1	1	1	1	0	80
282	NC_004741.1	WP_000879272.1	1	0	0	0	0	20
283	NC_004741.1	WP_000168747.1	1	1	1	1	0	80
284	NC_004741.1	WP_000275187.1	0	0	0	0	0	0
285	NC_004741.1	WP_005047608.1	1	1	1	1	0	80
286	NC_004741.1	WP_000155622.1	0	0	0	0	0	0
287	NC_004741.1	WP_001024930.1	1	1	1	1	0	80
288	NC_004741.1	WP_005084198.1	0	0	0	0	0	0
289	NC_004741.1	WP_001039885.1	1	0	1	0	0	40
290	NC_004741.1	WP_000930145.1	1	1	1	1	0	80
291	NC_004741.1	WP_000009987.1	0	0	0	0	0	0
292	NC_004741.1	WP_000245528.1	1	1	1	1	0	80
293	NC_004741.1	WP_005049830.1	1	0	0	0	0	20
294	NC_004741.1	WP_001173294.1	1	1	1	0	0	60
295	NC_004741.1	WP_000930141.1	1	1	1	1	0	80
296	NC_004741.1	WP_001007942.1	0	0	0	0	0	0
297	NC_004741.1	WP_000082749.1	1	1	1	1	0	80
298	NC_004741.1	WP_005048633.1	0	0	0	0	0	0
299	NC_004741.1	WP_001028876.1	0	0	0	0	0	0
300	NC_004741.1	WP_000755956.1	1	1	1	1	1	100
301	NC_004741.1	WP_001099210.1	1	1	1	1	0	80
302	NC_004741.1	WP_000586688.1	1	0	0	0	0	20
303	NC_004741.1	WP_000457719.1	1	1	1	1	0	80
304	NC_004741.1	WP_001030133.1	0	0	0	0	0	0

305	NC_004741.1	WP_005063152.1	1	1	1	1	0	80
306	NC_004741.1	WP_000455174.1	1	1	1	1	0	80
307	NC_004741.1	WP_001103659.1	1	0	0	0	0	20
308	NC_004741.1	WP_042791229.1	0	0	0	0	0	0
309	NC_004741.1	WP_000082120.1	1	1	1	1	0	80
310	NC_004741.1	WP_001297814.1	0	0	0	0	0	0
311	NC_004741.1	WP_001237866.1	1	1	1	1	1	100
312	NC_004741.1	WP_000377229.1	1	1	1	1	0	80
313	NC_004741.1	WP_032155628.1	0	0	0	0	0	0
314	NC_004741.1	WP_000106474.1	1	1	1	1	0	80
315	NC_004741.1	WP_000118898.1	1	1	1	1	0	80
316	NC_004741.1	WP_000230645.1	0	0	0	0	0	0
317	NC_004741.1	WP_024259260.1	0	0	0	0	0	0
318	NC_004741.1	WP_005048789.1	1	1	1	1	0	80
319	NC_004741.1	WP_000431460.1	1	0	0	0	0	20
320	NC_004741.1	WP_032155863.1	1	0	0	0	0	20
321	NC_004741.1	WP_032155819.1	0	0	0	0	0	0
322	NC_004741.1	WP_001039899.1	0	0	0	0	0	0
323	NC_004741.1	WP_064716611.1	1	1	1	1	0	80
324	NC_004741.1	WP_000594909.1	0	0	0	0	1	20
325	NC_004741.1	WP_001062338.1	1	1	1	1	0	80
326	NC_004741.1	WP_001265248.1	1	1	1	1	0	80
327	NC_004741.1	WP_000152435.1	0	0	0	0	0	0
328	NC_004741.1	WP_032155691.1	1	0	0	0	0	20
329	NC_004741.1	WP_001343759.1	0	0	0	0	0	0
330	NC_004741.1	WP_000466572.1	1	0	0	0	0	20
331	NC_004741.1	WP_061440266.1	1	0	0	0	0	20
332	NC_004741.1	WP_001016348.1	1	1	1	1	0	80

333	NC_004741.1	WP_000450409.1	1	1	1	1	0	80
334	NC_004741.1	WP_005088730.1	0	0	1	0	0	20
335	NC_004741.1	WP_000282151.1	1	1	1	1	0	80
336	NC_004741.1	WP_001243860.1	1	1	1	1	0	80
337	NC_004741.1	WP_011110604.1	0	0	0	0	0	0
338	NC_004741.1	WP_011069433.1	0	0	0	0	0	0
339	NC_004741.1	WP_000055830.1	0	0	0	0	0	0
340	NC_004741.1	WP_011069434.1	1	0	1	0	0	40
341	NC_004741.1	WP_000454701.1	1	1	1	1	1	100
342	NC_004741.1	WP_000489605.1	1	0	0	0	0	20
343	NC_004741.1	WP_000003197.1	1	1	1	1	1	100
344	NC_004741.1	WP_000929408.1	1	1	1	1	0	80
345	NC_004741.1	WP_014532286.1	0	0	0	0	0	0
346	NC_004741.1	WP_001324860.1	0	0	0	0	1	20
347	NC_004741.1	WP_000830460.1	1	0	1	0	0	40
348	NC_004741.1	WP_011110605.1	1	1	1	1	0	80
349	NC_004741.1	WP_000261596.1	0	0	0	0	0	0
350	NC_004741.1	WP_005049040.1	0	0	0	0	0	0
351	NC_004741.1	WP_005098884.1	1	1	1	1	0	80
352	NC_004741.1	WP_005049034.1	1	1	1	1	0	80
353	NC_004741.1	WP_024259269.1	1	1	1	1	0	80
354	NC_004741.1	WP_000636931.1	0	0	0	0	0	0
355	NC_004741.1	WP_000380421.1	0	0	0	0	0	0
356	NC_004741.1	WP_001294399.1	0	0	0	0	0	0
357	NC_004741.1	WP_005049026.1	1	1	1	1	0	80
358	NC_004741.1	WP_001087240.1	1	0	1	0	1	60
359	NC_004741.1	WP_005049020.1	1	1	1	1	1	100
360	NC_004741.1	WP_048814497.1	1	1	1	1	1	100

361	NC_004741.1	WP_011069443.1	0	0	0	0	0	0
362	NC_004741.1	WP_000691708.1	0	0	0	0	0	0
363	NC_004741.1	WP_001295452.1	1	1	1	1	0	80
364	NC_004741.1	WP_032155550.1	0	0	0	0	0	0
365	NC_004741.1	WP_001308773.1	0	0	0	0	0	0
366	NC_004741.1	WP_000182053.1	1	1	1	1	0	80
367	NC_004741.1	WP_000202798.1	1	1	1	1	0	80
368	NC_004741.1	WP_001135673.1	1	1	1	1	0	80
369	NC_004741.1	WP_001303596.1	0	0	0	0	0	0
370	NC_004741.1	WP_001296837.1	0	0	0	0	0	0
371	NC_004741.1	WP_001225855.1	1	1	1	1	0	80
372	NC_004741.1	WP_001104543.1	1	1	1	1	0	80
373	NC_004741.1	WP_005046832.1	1	0	0	0	0	20
374	NC_004741.1	WP_001215763.1	1	1	1	1	0	80
375	NC_004741.1	WP_032083391.1	0	0	0	0	0	0
376	NC_004741.1	WP_000301054.1	1	1	1	1	1	100
377	NC_004741.1	WP_001009396.1	1	0	0	0	0	20
378	NC_004741.1	WP_000140529.1	1	1	1	1	0	80
379	NC_004741.1	WP_000070619.1	1	1	1	1	0	80
380	NC_004741.1	WP_001446945.1	1	1	1	1	0	80
381	NC_004741.1	WP_000525371.1	1	1	1	1	0	80
382	NC_004741.1	WP_000426116.1	1	1	1	1	0	80
383	NC_004741.1	WP_000106622.1	1	1	1	1	0	80
384	NC_004741.1	WP_001115612.1	0	0	0	0	0	0
385	NC_004741.1	WP_001274496.1	1	1	1	1	0	80
386	NC_004741.1	WP_000559763.1	1	1	1	1	0	80
387	NC_004741.1	WP_005070009.1	1	1	1	1	0	80
388	NC_004741.1	WP_000937783.1	0	0	0	0	0	0

389	NC_004741.1	WP_000937210.1	1	1	1	1	0	80
390	NC_004741.1	WP_000825597.1	1	1	1	1	0	80
391	NC_004741.1	WP_000867638.1	0	0	0	0	1	20
392	NC_004741.1	WP_009008053.1	0	0	0	0	0	0
393	NC_004741.1	WP_000639883.1	1	1	1	1	0	80
394	NC_004741.1	WP_000490072.1	1	1	1	1	0	80
395	NC_004741.1	WP_042188255.1	0	0	0	0	0	0
396	NC_004741.1	WP_000826512.1	1	1	1	1	0	80
397	NC_004741.1	WP_000201413.1	1	1	1	1	0	80
398	NC_004741.1	WP_000719924.1	1	0	1	0	0	40
399	NC_004741.1	WP_001107736.1	0	0	0	0	0	0
400	NC_004741.1	WP_000555795.1	0	0	0	0	0	0
401	NC_004741.1	WP_000338539.1	0	0	0	0	0	0
402	NC_004741.1	WP_001373377.1	0	0	0	0	0	0
403	NC_004741.1	WP_000806589.1	0	0	0	0	0	0
404	NC_004741.1	WP_001507728.1	1	1	1	1	0	80
405	NC_004741.1	WP_001308835.1	0	0	0	0	0	0
406	NC_004741.1	WP_001349976.1	1	0	0	0	0	20
407	NC_004741.1	WP_000339447.1	1	1	1	1	0	80
408	NC_004741.1	WP_001244758.1	1	0	0	0	0	20
409	NC_004741.1	WP_000017552.1	0	0	0	0	0	0
410	NC_004741.1	WP_000076001.1	1	1	1	1	0	80
411	NC_004741.1	WP_000755178.1	0	0	1	0	0	20
412	NC_004741.1	WP_000743213.1	0	0	0	0	0	0
413	NC_004741.1	WP_000131871.1	0	0	0	0	0	0
414	NC_004741.1	WP_000211355.1	1	1	1	1	0	80
415	NC_004741.1	WP_000213809.1	0	0	0	0	0	0
416	NC_004741.1	WP_001162384.1	1	1	1	1	0	80

417	NC_004741.1	WP_005047279.1	1	1	1	1	0	80
418	NC_004741.1	WP_001094726.1	1	1	1	1	0	80
419	NC_004741.1	WP_000266171.1	1	1	1	1	1	100
420	NC_004741.1	WP_000951754.1	1	0	0	0	0	20
421	NC_004741.1	WP_000647601.1	1	0	0	0	0	20
422	NC_004741.1	WP_012135949.1	0	0	0	0	0	0
423	NC_004741.1	WP_001303621.1	0	0	0	0	0	0
424	NC_004741.1	WP_032155618.1	0	0	0	0	0	0
425	NC_004741.1	WP_001330697.1	0	0	0	0	0	0
426	NC_004741.1	WP_001212392.1	1	1	1	1	0	80
427	NC_004741.1	WP_000589825.1	1	1	1	1	1	100
428	NC_004741.1	WP_000284119.1	0	0	0	0	0	0
429	NC_004741.1	WP_000491410.1	1	1	1	1	0	80
430	NC_004741.1	WP_000281320.1	1	1	1	1	0	80
431	NC_004741.1	WP_000483311.1	1	0	0	0	0	20
432	NC_004741.1	WP_001287454.1	1	1	1	1	0	80
433	NC_004741.1	WP_001307965.1	0	0	0	0	0	0
434	NC_004741.1	WP_000493764.1	1	0	0	0	0	20
435	NC_004741.1	WP_032155822.1	0	0	0	0	0	0
436	NC_004741.1	WP_000611930.1	0	0	0	0	0	0
437	NC_004741.1	WP_001224024.1	1	0	0	0	0	20
438	NC_004741.1	WP_001288227.1	0	0	0	0	0	0
439	NC_004741.1	WP_000444999.1	1	1	1	1	0	80
440	NC_004741.1	WP_000206987.1	1	1	1	1	0	80
441	NC_004741.1	WP_001393510.1	0	0	0	0	0	0
442	NC_004741.1	WP_000203905.1	1	1	1	1	0	80
443	NC_004741.1	WP_000184250.1	1	1	1	1	0	80
444	NC_004741.1	WP_001078387.1	1	1	1	1	0	80

445	NC_004741.1	WP_032155588.1	0	0	0	0	0	0
446	NC_004741.1	WP_000860229.1	0	0	0	0	0	0
447	NC_004741.1	WP_000242461.1	1	0	0	0	0	20
448	NC_004741.1	WP_000379402.1	0	0	0	0	0	0
449	NC_004741.1	WP_005099217.1	0	0	0	0	0	0
450	NC_004741.1	WP_005099219.1	0	0	0	0	0	0
451	NC_004741.1	WP_000971492.1	1	0	0	0	0	20
452	NC_004741.1	WP_005051685.1	1	1	1	1	1	100
453	NC_004741.1	WP_000528349.1	0	0	0	0	0	0
454	NC_004741.1	WP_001010156.1	0	0	0	0	0	0
455	NC_004741.1	WP_005051767.1	1	1	1	1	0	80
456	NC_004741.1	WP_011110620.1	0	1	1	0	0	40
457	NC_004741.1	WP_001094817.1	1	1	1	1	0	80
458	NC_004741.1	WP_000745204.1	1	1	1	1	0	80
459	NC_004741.1	WP_000984792.1	1	1	1	1	0	80
460	NC_004741.1	WP_005064025.1	0	1	0	1	0	40
461	NC_004741.1	WP_000338035.1	0	0	0	0	0	0
462	NC_004741.1	WP_000291751.1	0	0	0	0	0	0
463	NC_004741.1	WP_001323220.1	0	0	0	0	0	0
464	NC_004741.1	WP_032142224.1	0	0	0	0	0	0
465	NC_004741.1	WP_001128940.1	1	1	1	1	0	80
466	NC_004741.1	WP_001013320.1	1	1	1	1	0	80
467	NC_004741.1	WP_000271035.1	1	1	1	1	0	80
468	NC_004741.1	WP_001195464.1	1	1	1	1	0	80
469	NC_004741.1	WP_011069510.1	0	0	0	0	0	0
470	NC_004741.1	WP_005093820.1	0	0	0	0	1	20
471	NC_004741.1	WP_000261147.1	0	0	0	0	0	0
472	NC_004741.1	WP_005051842.1	0	0	0	0	0	0

473	NC_004741.1	WP_001069724.1	1	1	1	1	0	80
474	NC_004741.1	WP_005093816.1	0	0	0	0	0	0
475	NC_004741.1	WP_005051844.1	0	0	0	0	0	0
476	NC_004741.1	WP_001387238.1	1	1	1	1	1	100
477	NC_004741.1	WP_000692350.1	1	1	1	1	0	80
478	NC_004741.1	WP_000761715.1	0	0	0	0	0	0
479	NC_004741.1	WP_000772029.1	1	1	1	1	0	80
480	NC_004741.1	WP_000340141.1	0	0	0	0	0	0
481	NC_004741.1	WP_000853257.1	1	1	1	1	0	80
482	NC_004741.1	WP_000248097.1	1	1	1	1	1	100
483	NC_004741.1	WP_000984979.1	0	0	0	0	0	0
484	NC_004741.1	WP_000339534.1	1	0	0	0	0	20
485	NC_004741.1	WP_005051896.1	0	0	0	0	0	0
486	NC_004741.1	WP_001059136.1	1	1	1	1	0	80
487	NC_004741.1	WP_024167679.1	0	0	0	0	0	0
488	NC_004741.1	WP_001298764.1	0	0	0	0	0	0
489	NC_004741.1	WP_000691640.1	1	1	1	1	0	80
490	NC_004741.1	WP_000848528.1	1	1	1	1	1	100
491	NC_004741.1	WP_000527661.1	0	0	0	0	0	0
492	NC_004741.1	WP_001701108.1	0	0	0	0	0	0
493	NC_004741.1	WP_000442868.1	1	1	1	1	0	80
494	NC_004741.1	WP_001406537.1	0	0	0	0	0	0
495	NC_004741.1	WP_024259304.1	0	0	0	0	0	0
496	NC_004741.1	WP_001298386.1	0	0	0	0	0	0
497	NC_004741.1	WP_032140301.1	0	0	0	0	0	0
498	NC_004741.1	WP_000942538.1	1	1	1	1	0	80
499	NC_004741.1	WP_000016819.1	1	1	1	1	0	80
500	NC_004741.1	WP_000422149.1	1	1	1	1	0	80

501	NC_004741.1	WP_005050960.1	1	1	1	1	0	80
502	NC_004741.1	WP_000031415.1	1	1	1	1	0	80
503	NC_004741.1	WP_000785722.1	1	1	1	1	0	80
504	NC_004741.1	WP_000096080.1	1	1	1	1	0	80
505	NC_004741.1	WP_000732225.1	1	1	1	1	0	80
506	NC_004741.1	WP_005089560.1	0	0	0	0	0	0
507	NC_004741.1	WP_005050890.1	0	0	0	0	0	0
508	NC_004741.1	WP_001343556.1	1	1	1	1	0	80
509	NC_004741.1	WP_000449030.1	1	0	0	0	0	20
510	NC_004741.1	WP_000189314.1	1	1	1	1	1	100
511	NC_004741.1	WP_001346700.1	0	0	0	0	0	0
512	NC_004741.1	WP_000620405.1	1	1	1	1	0	80
513	NC_004741.1	WP_005077180.1	0	0	0	0	0	0
514	NC_004741.1	WP_005050708.1	1	1	1	1	0	80
515	NC_004741.1	WP_001028769.1	1	1	1	1	0	80
516	NC_004741.1	WP_001303690.1	0	0	0	0	0	0
517	NC_004741.1	WP_011110633.1	0	0	0	0	0	0
518	NC_004741.1	WP_001061203.1	0	0	0	0	0	0
519	NC_004741.1	WP_001326891.1	1	0	0	0	0	20
520	NC_004741.1	WP_024259309.1	1	1	1	1	0	80
521	NC_004741.1	WP_005050690.1	0	0	0	0	0	0
522	NC_004741.1	WP_000460680.1	1	1	1	1	0	80
523	NC_004741.1	WP_000757326.1	0	1	1	1	0	60
524	NC_004741.1	WP_000595564.1	1	1	1	1	0	80
525	NC_004741.1	WP_000487766.1	0	0	0	0	0	0
526	NC_004741.1	WP_024166609.1	0	0	0	0	0	0
527	NC_004741.1	WP_005050602.1	0	0	0	0	0	0
528	NC_004741.1	WP_011069598.1	1	1	1	1	0	80

529	NC_004741.1	WP_000155673.1	1	1	1	1	0	80
530	NC_004741.1	WP_001014565.1	1	1	1	1	0	80
531	NC_004741.1	WP_001324833.1	0	0	0	0	0	0
532	NC_004741.1	WP_000719886.1	1	1	1	1	0	80
533	NC_004741.1	WP_001112357.1	1	0	0	0	0	20
534	NC_004741.1	WP_000169147.1	1	1	1	1	0	80
535	NC_004741.1	WP_032155643.1	0	0	0	0	0	0
536	NC_004741.1	WP_000627171.1	1	1	1	1	0	80
537	NC_004741.1	WP_023517643.1	0	1	0	1	0	40
538	NC_004741.1	WP_046201574.1	1	0	0	0	0	20
539	NC_004741.1	WP_000940102.1	1	1	1	1	0	80
540	NC_004741.1	WP_011110634.1	0	0	0	0	0	0
541	NC_004741.1	WP_001442985.1	0	1	0	1	0	40
542	NC_004741.1	WP_000660586.1	0	0	0	0	0	0
543	NC_004741.1	WP_000797352.1	1	1	1	1	0	80
544	NC_004741.1	WP_000655986.1	1	1	1	1	0	80
545	NC_004741.1	WP_000802226.1	1	1	1	1	0	80
546	NC_004741.1	WP_000591073.1	0	0	0	0	0	0
547	NC_004741.1	WP_000510376.1	0	0	0	0	0	0
548	NC_004741.1	WP_014334093.1	0	1	0	1	0	40
549	NC_004741.1	WP_001086388.1	0	0	0	0	0	0
550	NC_004741.1	WP_001295676.1	1	1	1	1	0	80
551	NC_004741.1	WP_032155602.1	0	0	0	0	0	0
552	NC_004741.1	WP_024259324.1	0	0	1	0	0	20
553	NC_004741.1	WP_000256409.1	1	1	1	1	0	80
554	NC_004741.1	WP_032142137.1	0	0	0	0	0	0
555	NC_004741.1	WP_000893994.1	0	0	0	0	0	0
556	NC_004741.1	WP_000115988.1	0	0	0	0	0	0

557	NC_004741.1	WP_001295264.1	1	1	1	1	0	80
558	NC_004741.1	WP_005052859.1	1	1	1	1	0	80
559	NC_004741.1	WP_001277142.1	1	1	1	1	0	80
560	NC_004741.1	WP_024259323.1	1	0	0	0	0	20
561	NC_004741.1	WP_000032578.1	1	1	1	1	0	80
562	NC_004741.1	WP_001127088.1	0	0	0	0	0	0
563	NC_004741.1	WP_000841001.1	1	1	1	1	0	80
564	NC_004741.1	WP_000336276.1	0	0	0	0	0	0
565	NC_004741.1	WP_000456043.1	0	0	0	0	0	0
566	NC_004741.1	WP_005051995.1	1	1	1	1	0	80
567	NC_004741.1	WP_000454294.1	0	0	0	0	0	0
568	NC_004741.1	WP_032155619.1	0	0	0	0	0	0
569	NC_004741.1	WP_001113432.1	1	1	1	1	0	80
570	NC_004741.1	WP_000703959.1	1	1	1	1	0	80
571	NC_004741.1	WP_000511292.1	0	0	0	0	0	0
572	NC_004741.1	WP_005052029.1	1	1	1	1	0	80
573	NC_004741.1	WP_000772934.1	1	0	0	0	1	40
574	NC_004741.1	WP_005052034.1	1	1	1	1	0	80
575	NC_004741.1	WP_032155621.1	0	1	0	1	0	40
576	NC_004741.1	WP_000190670.1	1	0	0	0	0	20
577	NC_004741.1	WP_045178171.1	1	1	1	1	0	80
578	NC_004741.1	WP_005052068.1	1	1	1	1	0	80
579	NC_004741.1	WP_000542440.1	1	1	1	1	0	80
580	NC_004741.1	WP_032155816.1	0	0	0	0	0	0
581	NC_004741.1	WP_011069564.1	0	0	0	0	0	0
582	NC_004741.1	WP_001304210.1	0	0	0	0	0	0
583	NC_004741.1	WP_001004881.1	0	0	0	0	0	0
584	NC_004741.1	WP_045178164.1	0	0	0	0	0	0

585	NC_004741.1	WP_001390447.1	0	0	0	0	0	0
586	NC_004741.1	WP_005052132.1	0	0	0	0	0	0
587	NC_004741.1	WP_011069562.1	1	0	0	0	0	20
588	NC_004741.1	WP_000344113.1	0	0	0	0	0	0
589	NC_004741.1	WP_000999840.1	1	0	0	0	0	20
590	NC_004741.1	WP_000924289.1	1	1	1	1	0	80
591	NC_004741.1	WP_000621323.1	1	1	1	1	0	80
592	NC_004741.1	WP_001297375.1	1	1	1	1	1	100
593	NC_004741.1	WP_000483856.1	1	1	1	1	0	80
594	NC_004741.1	WP_000665677.1	1	1	1	1	0	80
595	NC_004741.1	WP_000517100.1	0	0	0	0	0	0
596	NC_004741.1	WP_000479627.1	1	1	1	1	0	80
597	NC_004741.1	WP_000332751.1	0	0	0	0	0	0
598	NC_004741.1	WP_001331222.1	0	1	0	1	0	40
599	NC_004741.1	WP_011069558.1	1	1	1	1	0	80
600	NC_004741.1	WP_000858193.1	1	1	1	1	1	100
601	NC_004741.1	WP_001296808.1	1	1	1	1	0	80
602	NC_004741.1	WP_000576411.1	1	1	1	1	0	80
603	NC_004741.1	WP_001296791.1	1	1	1	1	1	100
604	NC_004741.1	WP_000198578.1	0	0	0	0	0	0
605	NC_004741.1	WP_032155594.1	0	0	0	0	0	0
606	NC_004741.1	WP_001063318.1	1	1	1	1	0	80
607	NC_004741.1	WP_001328969.1	0	0	0	0	0	0
608	NC_004741.1	WP_001205330.1	0	0	0	0	1	20
609	NC_004741.1	WP_000020617.1	0	0	0	0	0	0
610	NC_004741.1	WP_005064932.1	0	0	0	0	0	0
611	NC_004741.1	WP_000555608.1	0	0	0	0	0	0
612	NC_004741.1	WP_000751953.1	0	0	0	0	0	0

613	NC_004741.1	WP_011069548.1	0	0	0	0	0	0
614	NC_004741.1	WP_032155607.1	0	0	0	0	0	0
615	NC_004741.1	WP_000643692.1	1	1	1	1	0	80
616	NC_004741.1	WP_005052620.1	1	1	1	1	0	80
617	NC_004741.1	WP_000042900.1	1	1	1	1	0	80
618	NC_004741.1	WP_000778795.1	1	1	1	1	1	100
619	NC_004741.1	WP_005093467.1	0	1	0	1	0	40
620	NC_004741.1	WP_001181212.1	0	0	0	0	0	0
621	NC_004741.1	WP_000907005.1	1	0	0	1	0	40
622	NC_004741.1	WP_005097678.1	1	0	0	0	0	20
623	NC_004741.1	WP_011069541.1	1	0	0	1	0	40
624	NC_004741.1	WP_001303701.1	1	0	0	0	0	20
625	NC_004741.1	WP_005052731.1	0	0	0	0	0	0
626	NC_004741.1	WP_001055752.1	0	0	0	0	0	0
627	NC_004741.1	WP_001254807.1	1	1	1	1	0	80
628	NC_004741.1	WP_005052744.1	1	0	0	0	1	40
629	NC_004741.1	WP_000497332.1	1	1	1	1	0	80
630	NC_004741.1	WP_000847163.1	1	1	1	1	0	80
631	NC_004741.1	WP_001303699.1	0	0	0	0	0	0
632	NC_004741.1	WP_001148908.1	1	1	1	1	0	80
633	NC_004741.1	WP_000907085.1	1	1	1	1	0	80
634	NC_004741.1	WP_001007729.1	1	1	1	1	0	80
635	NC_004741.1	WP_000786137.1	0	0	0	0	0	0
636	NC_004741.1	WP_005065417.1	1	1	1	1	0	80
637	NC_004741.1	WP_001295738.1	0	0	0	0	0	0
638	NC_004741.1	WP_000062539.1	1	1	1	1	0	80
639	NC_004741.1	WP_000242065.1	1	1	0	1	0	60
640	NC_004741.1	WP_032155811.1	0	0	0	0	0	0

641	NC_004741.1	WP_000121001.1	1	1	1	1	0	80
642	NC_004741.1	WP_024259336.1	0	0	0	0	0	0
643	NC_004741.1	WP_032155640.1	0	0	0	0	0	0
644	NC_004741.1	WP_001243676.1	1	1	1	1	0	80
645	NC_004741.1	WP_001351186.1	1	0	0	0	0	20
646	NC_004741.1	WP_011110644.1	1	1	1	1	0	80
647	NC_004741.1	WP_000079652.1	1	1	1	1	0	80
648	NC_004741.1	WP_001243871.1	1	1	1	1	0	80
649	NC_004741.1	WP_011069606.1	0	0	0	0	0	0
650	NC_004741.1	WP_000166281.1	1	1	1	1	0	80
651	NC_004741.1	WP_001296688.1	0	0	0	0	0	0
652	NC_004741.1	WP_005053984.1	1	1	1	1	0	80
653	NC_004741.1	WP_000937635.1	0	0	0	0	0	0
654	NC_004741.1	WP_001119485.1	1	1	1	1	0	80
655	NC_004741.1	WP_001205243.1	1	1	1	1	1	100
656	NC_004741.1	WP_000044756.1	0	0	0	0	0	0
657	NC_004741.1	WP_001008046.1	1	1	1	1	0	80
658	NC_004741.1	WP_001238362.1	1	1	1	1	1	100
659	NC_004741.1	WP_005134385.1	1	0	0	0	0	20
660	NC_004741.1	WP_001243705.1	1	1	1	1	0	80
661	NC_004741.1	WP_000943980.1	1	1	1	1	1	100
662	NC_004741.1	WP_005053837.1	1	0	0	0	0	20
663	NC_004741.1	WP_000492914.1	1	1	1	1	0	80
664	NC_004741.1	WP_032155818.1	0	0	0	0	0	0
665	NC_004741.1	WP_000895690.1	1	0	0	0	0	20
666	NC_004741.1	WP_000132640.1	1	1	1	1	1	100
667	NC_004741.1	WP_000467859.1	1	1	1	1	0	80
668	NC_004741.1	WP_005053796.1	0	0	0	0	0	0

669	NC_004741.1	WP_000538192.1	1	1	1	1	0	80
670	NC_004741.1	WP_001338213.1	1	1	1	1	0	80
671	NC_004741.1	WP_000604352.1	1	1	1	1	0	80
672	NC_004741.1	WP_000494556.1	0	0	0	0	0	0
673	NC_004741.1	WP_000007444.1	0	0	0	0	0	0
674	NC_004741.1	WP_001303782.1	0	0	0	0	0	0

Note: 0 = 0%, 1 = 25%.

Supplementary Table 2. List of predicted physicochemical parameters of 39 hypothetical proteins

Sl. No.	Accession ID_Protein	No. of amino acids	MW	PI	Extinction coefficient	Instability index	Classification	Alphabatic index	Grand average of hydropathicity (GRAVY)
1	WP_005053355.1	274	29970.4	7.62	24325	28.39	Stable	84.01	-0.016
2	WP_000092054.1	364	40443.3	9.61	51005	47.89	Unstable	79.67	-0.384
3	WP_001382892.1	179	19590.3	5.28	2980	35.06	Stable	101.23	-0.143
4	WP_005053036.1	192	20906	9.04	7450	35.73	Stable	95.05	-0.062
5	WP_000779831.1	190	19441.2	7.87	6990	45.39	Unstable	96.58	0.172
6	WP_011110552.1	108	12039.7	7.61	8730	66.94	Unstable	77.78	-0.544
7	WP_001269672.1	193	21386.6	8.73	11460	31.1	Stable	90.98	-0.238
8	WP_001247854.1	619	69683.8	5.5	107425	33.35	Stable	78.24	-0.458
9	WP_000070107.1	377	42056.8	7.71	52035	31.62	Stable	124.14	0.611
10	WP_000224274.1	369	40593.3	7.03	40950	37.16	Stable	82.22	-0.191
11	WP_000749269.1	191	20942.5	5.57	16960	11.14	Stable	75.6	-0.436
12	WP_001125713.1	108	12371.5	9.16	7450	53.58	Unstable	82.13	-0.624
13	WP_001043881.1	165	18093.6	4.66	11585	35.37	Stable	101.09	0.133
14	WP_001295493.1	114	12493.2	4.96	20970	30.77	Stable	97.63	0.024
15	WP_000691930.1	84	8942.38	7.66	12740	41.42	Unstable	81.31	-0.167
16	WP_000597196.1	155	15601.7	9.36	2980	24.38	Stable	94.77	0.114
17	WP_000248636.1	370	39841.2	8.49	79075	32.75	Stable	143.05	1.029
18	WP_000755956.1	275	30284.6	5.68	30035	24.35	Stable	84.04	-0.331
19	WP_001237866.1	107	11755.5	6.56	6210	30.08	Stable	97.57	0.277
20	WP_000454701.1	527	59450.2	5.15	37930	32.47	Stable	116.7	0.177
21	WP_000003197.1	219	24222.7	5.17	26595	42.24	Unstable	85.98	-0.055
22	WP_005049020.1	153	16568.1	7.7	7450	23.39	Stable	82.94	-0.176
23	WP_048814497.1	243	27279.8	4.72	22585	43.49	Unstable	104.32	-0.255

24	WP_000301054.1	216	25324.2	9.31	51005	44.31	Unstable	77.64	-0.482
25	WP_000266171.1	1033	117109	5.58	244955	42.42	Unstable	83.71	-0.343
26	WP_000589825.1	160	17240.6	8.71	18450	39.1	Stable	82.44	-0.271
27	WP_005051685.1	251	26572.5	10.11	34505	23.04	Stable	76.18	-0.279
28	WP_001387238.1	158	17740.4	6.76	8480	28.12	Stable	108.04	-0.073
29	WP_000248097.1	82	9417.1	4.02	6990	44.78	Unstable	135.37	0.266
30	WP_000848528.1	85	9513.04	8.8	4720	42.38	Unstable	94.24	-0.107
31	WP_000189314.1	100	11241.9	9.84	11460	36.12	Stable	84	-0.544
32	WP_001297375.1	222	25258.4	7.84	13075	32.89	Stable	101.49	-0.086
33	WP_000858193.1	113	12552.2	9.33	14565	22.75	Stable	129.38	1.041
34	WP_001296791.1	232	25960.7	4.58	65890	36.83	Stable	74.87	-0.38
35	WP_000778795.1	127	14544.5	6.59	26470	47.26	Unstable	85.98	-0.412
36	WP_001205243.1	294	32718.5	4.96	50795	48.91	Unstable	98.91	-0.096
37	WP_001238362.1	177	19911.7	8.88	35535	27.45	Stable	78.25	-0.194
38	WP_000943980.1	387	45038.7	4.64	94685	46.38	Unstable	83.7	-0.434
39	WP_000132640.1	113	12294.1	8.64	10095	36.84	Stable	87.26	-0.168

MW, molecular weight; GRAVY, grand average of hydropathicity.

Supplementary Table 3. List of predicted sub-cellular localization of 39 hypothetical proteins

S. No.	Accession No.	Sub-cellular localization			Signal peptide (Signal P)	Secretory protein (Secretome P)	Trans membrane helices prediction		
		CELLO	PSORT B	PSLpred			HMMTOP	TMHMM	SOSUI
1	WP_005053355.1	Periplasmic	Unknown	Periplasmic	Yes	Yes	No	No	Soluble
2	WP_000092054.1	Periplasmic	Unknown	Periplasmic	Yes	Yes	1 TM Helices	No	Soluble
3	WP_001382892.1	Periplasmic/ Extracellular	Unknown	Outer membrane	Yes	Yes	No	No	Soluble
4	WP_005053036.1	Periplasmic	Unknown	Periplasmic	Yes	Yes	1 TM Helices	No	Membrane, 1 TM helix
5	WP_000779831.1	Periplasmic	Periplasmic	Periplasmic	Yes	Yes	No	No	Soluble
6	WP_011110552.1	Periplasmic	Unknown	Periplasmic	Yes	Yes	No	No	Membrane, 1 TM helix
7	WP_001269672.1	Periplasmic	Outer membrane	Periplasmic	No	Yes	2 TM Helices	No	Membrane, 1 TM helix
8	WP_001247854.1	Cytoplasmic	Cytoplasmic	Cytoplasmic	No	No	No	No	Soluble
9	WP_000070107.1	Inner membrane	Cytoplasmic membrane	Inner-membrane	No	No	6 TM Helices	6 TM Helices	membrane, 6 TM helix
10	WP_000224274.1	Periplasmic	Unknown	Cytoplasmic	No	No	No	No	Soluble
11	WP_000749269.1	Periplasmic	Unknown	Periplasmic	Yes	No	1 TM Helices	No	Soluble

12	WP_001125713.1	Cytoplasmic	Cytoplasmic	Periplasmic	No	Yes	No	No	Soluble
13	WP_001043881.1	Cytoplasmic	Unknown	Cytoplasmic	No	No	No	No	Soluble
14	WP_001295493.1	Cytoplasmic	Cytoplasmic	Cytoplasmic	No	No	No	No	Soluble
15	WP_000691930.1	Periplasmic	Unknown	Periplasmic	Yes	Yes	No	No	Membrane, 1 TM helix
16	WP_000597196.1	Extracellular	Outer membrane	Extracellular	No	Yes	No	No	Membrane, 1 TM helix
17	WP_000248636.1	Inner membrane	Cytoplasmic membrane	Inner membrane	No	No	9 TM Helices	10 TM Helices	8 TM Helices
18	WP_000755956.1	Periplasmic	Unknown	Periplasmic	No	Yes	No	No	Soluble
19	WP_001237866.1	Cytoplasmic	Unknown	Cytoplasmic	No	No	No	No	Membrane, 1 TM helix
20	WP_000454701.1	Inner membrane	Cytoplasmic membrane	InnerMembrane	No	No	7 TM Helices	7 TM Helices	Membrane, 7 TM helix
21	WP_000003197.1	Cytoplasmic	Unknown	Cytoplasmic	No	No	No	No	Soluble
22	WP_005049020.1	Periplasmic	Unknown	Periplasmic	Yes	Yes	No	No	Membrane, 1 TM helix
23	WP_048814497.1	Cytoplasmic/ Outer membrane	Extracellular	Extracellular	No	No	No	No	Soluble
24	WP_000301054.1	Cytoplasmic	Cytoplasmic	Cytoplasmic	No	No	No	No	Soluble
25	WP_000266171.1	Cytoplasmic	Unknown	Cytoplasmic	No	No	No	No	Soluble

26	WP_000589825.1	Periplasmic	Outer membrane	Outer membrane	No	Yes	No	No	Membrane, 1 TM helix
27	WP_005051685.1	Extracellular	Outer membrane	Extracellular	No	yes	No	1 TM Helices	Soluble
28	WP_001387238.1	Cytoplasmic	Cytoplasmic	Cytoplasmic	No	No	No	No	Soluble
29	WP_000248097.1	Cytoplasmic	Unknown	Cytoplasmic	No	No	No	No	Soluble
30	WP_000848528.1	Periplasmic	Outer membrane	Cytoplasmic	No	No	1 TM Helices	No	Membrane, 1 TM helix
31	WP_000189314.1	Cytoplasmic/Periplasmic	Unknown	Outer membrane	No	No	No	No	Soluble
32	WP_001297375.1	Cytoplasmic	Cytoplasmic	Cytoplasmic	No	No	No	No	Soluble
33	WP_000858193.1	Inner membrane	Cytoplasmic membrane	Inner membrane	No	No	4 TM Helices	4 TM Helices	Membrane, 3 TM helix
34	WP_001296791.1	Extracellular	Extracellular	Extracellular	No	Yes	1 TM Helices	1 TM Helices	Soluble
35	WP_000778795.1	Cytoplasmic	Unknown	Cytoplasmic	No	No	No	No	Soluble
36	WP_001205243.1	Cytoplasmic	Extracellular	Cytoplasmic	No	No	No	1 TM Helices	Soluble
37	WP_001238362.1	Periplasmic	Outer membrane	Cytoplasmic	Yes	Yes	No	1 TM Helices	Membrane, 1 TM helix
38	WP_000943980.1	Cytoplasmic	Cytoplasmic	Cytoplasmic	No	No	No	No	Soluble
39	WP_000132640.1	Cytoplasmic	Unknown	Cytoplasmic	No	yes	No	No	Soluble

Supplementary Table 4. List of annotated functions of 39 hypothetical proteins from *Shigella flexneri* using CDD-BLAST, Pfam, HmmScan, SMART, Scanprosite, PS2-v2, and STRING

Sl. No	Acc ID	Functional domain (BLAST,Pfam, HmmScan, SMART,Scanprosite)	Templates	Domain in (PS)2-v2	Predicted functional partner (STRING)
1	WP_005053355.1	Peptidase, C92 family	No template	Error	Minor fimbrial subunit, D-mannose specific adhesin
2	WP_000092054.1	DUF1615/Lipoprotein	1m9iA	Same	Microcin B17 transporter
3	WP_001382892.1	DUF3251/lipoprotein YajI/immunoglobulin like domain	2jwyA	Same	Hypothetical protein SF0234/ATP synthase
4	WP_005053036.1	Lipoprotein_16/Uncharacterized lipoprotein	2iqiF	Same	Regulatory protein AmpE
5	WP_000779831.1	Lipoprotein chaperone (YscW)	No template	Error	Universal stress protein UspB
6	WP_011110552.1	YbfN-like lipoprotein	No template	Error	Hypothetical protein ybfM
7	WP_001269672.1	LPS-assembly lipoprotein RlpB (LptE)	2r76A	Same	LPS assembly outer membrane complex protein LptD
8	WP_001247854.1	Topoisomerases, DnaG-type primases, Hedgehog/Intein domain	No template	Error	DNA-directed RNA polymerase subunit beta
9	WP_000070107.1	ATP-binding cassette transporter	2dyrA	**OXIDOREDUCTASE**	ATP-binding protein ybhF_2
10	WP_000224274.1	MOSC beta barrel domain/2Fe-2S iron-sulfur cluster binding domain	2piaA	Same	Fe/S biogenesis protein NfuA
11	WP_000749269.1	YceI-like domain	1y0gA	Same	yceJ Cytochrome
12	WP_001125713.1	YcgL domain	2h7aA	Same	Hypothetical protein ycgN

13	WP_001043881.1	GAF domain	1vhmB	Same	Hypothetical protein; RNA chaperone proQ
14	WP_001295493.1	Endoribonuclease L-PSP/YjgFfamily	1qd9A	Same	D-amino acid dehydrogenase small subunit
15	WP_000691930.1	Domain of unknown function (DUF333)	2pqcA	**TRANSFERASE**	Hypothetical protein yeaP
16	WP_000597196.1	Glycine zipper 2TM domain	No template	Error	Flagellar fliJ protein
17	WP_000248636.1	AI-2E family transporter/permease	2jlnA	Same	Glutamine amidotransferase/anthranilate phosphoribosyltransferase
18	WP_000755956.1	SPFH domain / Band 7 family	3bk6A	Same	Integrase
19	WP_001237866.1	YecR-like lipoprotein	No template	Error	Glycosyl transferase
20	WP_000454701.1	TerC family/Transporter associated domain/CBS domain	2yvyA	Same	Glutamate synthase
21	WP_000003197.1	von Willebrand factor type A domain	1atzB	Same	Chaperonin
22	WP_005049020.1	Uncharacterized lipoprotein YehR	2joeA	Same	Transporter
23	WP_048814497.1	Leucine rich repeat protein/NEL or novel E3 ligase domain	3cvrA	Same	Aerobic respiration control sensor protein ArcB
24	WP_000301054.1	Lipopolysaccharide kinase (Kdo/WaaP)	1blxA	Same	Lipopolysaccharide core heptose(I) kinase RfaP
25	WP_000266171.1	Tetratricopeptide repeat (TPR)	No template	Error	NAGC-like transcriptional regulator
26	WP_000589825.1	Outer membrane protein (ompA) like domain/membrane lipoprotein	2k1sA	Same	Hypothetical protein SF2663

27	WP_005051685.1	LysM (lysin-like motif)/ Peptidase family M23	2gu1A	Same	Beta-hexosaminidase
28	WP_001387238.1	DNA repair protein RadC-like JAB domain	2qlcA	Same	Hypothetical protein SF2995
29	WP_000248097.1	Carrier protein (CP) domain and phosphopantetheine attachment site	1x3oA	Same	Class II aminotransferase
30	WP_000848528.1	Lipoprotein leucine-zipper	1jcdB	Same	Porin
31	WP_000189314.1	GIY-YIG nuclease domain	1zg2A	Same	Hypothetical protein yhbP
32	WP_001297375.1	DNA repair protein RadC-like JAB domain	No template	Error	DNA mismatch repair protein MutS;
33	WP_000858193.1	yiaA/B two helix domain	1oedA	Same	Hypothetical protein yiaA
34	WP_001296791.1	Autotransporter beta-domain	No template	Error	Biofilm formation regulatory protein BssR
35	WP_000778795.1	Acetyltransferase (GNAT) domain	2k5tA	Same	Aspartate alpha-decarboxylase
36	WP_001205243.1	Xylose isomerase-like TIM barrel (AP_endonuc_2)	1k77A	Same	3-ketoacyl-ACP reductase
37	WP_001238362.1	Lipocalin-like domain	1qwdA	Same	Sugar nucleotide epimerase
38	WP_000943980.1	Glutathionylspermidine synthase	No template	Error	Nicotinate phosphoribosyltransferase
39	WP_000132640.1	Toxin SymE/SpoVT-AbrB domain	1ve0A	Same	Hypothetical protein SF1670

Note: Proteins with discrepant results are shown in bold.

Supplementary Table 5. List of annotated functions of 39 hypothetical proteins from *Shigella flexneri* using MOTIF, Interproscan, CATH, SUPERFAMILY, and ProtoNet

Sl No.	Acc ID	MOTIF	INTERPROSCAN	CATH	SUPERFAMILY	ProtoNet
1	WP_005053355.1	Papain-like amidase enzyme, YaeF/YiiX, C92 family	Papain-like amidase enzyme, YaeF/YiiX, C92 family	Lipoprotein/Uncharacterized protein	Cysteine proteinases YiiX-like	Cluster 3674930 Proteobacteria
2	WP_000092054.1	Protein of unknown function (DUF1615)	Protein of unknown function DUF1615	No hit	GFP-like	Cluster 4109548 Protein of unknown function DUF1615
3	WP_001382892.1	Protein of unknown function (DUF3251)	Domain of unknown function DUF3251	Hypothetical lipoprotein yajI	Phase 1 flagellin	Cluster 3711586 2JWY
4	WP_005053036.1	Uncharacterized lipoprotein	Uncharacterised protein family, YajG	No hit	Phase 1 flagellin	Cluster 4028813 Uncharacterized lipoprotein
5	WP_000779831.1	lipoprotein chaperone (YscW)	No result	No hit	No result	Cluster 4131069 Proteobacteria
6	WP_011110552.1	YbfN-like lipoprotein	YbfN-like lipoprotein	No hit	No result	Cluster 4085534 Lipoprotein
7	WP_001269672.1	Lipopolysaccharide-assembly LptE	LPS-assembly lipoprotein LptE	LPS-assembly lipoprotein LptE	No result	Cluster 3965977 Rare lipoprotein B
8	WP_001247854.1	Toprim-like	DNA primase/Toprim domain	DNA primase/helicase	DNA primase/helicasecore	Cluster 3410389 DNA helicase, DnaB-like
9	WP_00007010	ABC-2 family	ABC-2 transporter	membrane transport	MFS general	Cluster 4114591

	7.1	transporter protein		permease YbhS/ATP-binding protein	substrate transporter	ABC-2
10	WP_000224274.1	MOSC domain/ 2Fe-2S iron-sulfur cluster	MOSC, N-terminal beta barrel	MOSC domain/ 2Fe-2S iron-sulfur cluster	MOSC N-terminal domain-like	Cluster 4155424 MOSC, N-terminal beta barrel
11	WP_000749269.1	YceI-like domain	YceI-like domain	YceI-like domain	YceI-like domain	Cluster 4314345 YceI-like
12	WP_001125713.1	YcgL domain	YcgL domain	No hit	YcgL-like	Cluster 4083593 YcgL domain
13	WP_001043881.1	GAF domain	GAF domain	GAF domain	GAF domain	Cluster 4085038 GAF
14	WP_001295493.1	Endoribonuclease L-PSP	YjgF/L-PSP	Endoribonuclease L-PSP family	YjgF/L-PSP	Cluster 4054994 YjgF/L-PSP
15	WP_000691930.1	Domain of unknown function (DUF333)	Domain of unknown function (DUF333)	No hit	No result	Cluster 4079210 Protein of unknown function DUF333
16	WP_000597196.1	Glycine zipper 2TM domain	Glycine zipper 2TM domain	No hit	No result	Cluster 4073223 Glycine zipper 2TM domain
17	WP_000248636.1	AI-2E family transporter	AI-2E family transporter	No hit	No result	Cluster 4074174 AI-2E family transporter
18	WP_000755956.1	SPFH domain / Band 7 family	SPFH domain / Band 7 family	No hit	Band 7/SPFH domain	Cluster 3793474 Band 7/SPFH domain
19	WP_001237866.1	YecR-like lipoprotein	YecR-like lipoprotein	No hit	SRCR-like	Cluster 3743846 Enterobacteriales
20	WP_000454701.1	TerC family/Transporter	TerC family/Transporter	TerC family/Transporter	CBS-domain pair/transporter-	Cluster 4019065 membrane protein

		associated domain/CBS domain	associated domain/CBS domain	associated domain/CBS domain	associated domain	TerC
21	WP_000003197.1	von Willebrand factor type A domain	TerY/vWA-like	von Willebrand factor type A domain	vWA-like	Cluster 4002958 TerY/vWA-like
22	WP_005049020.1	Protein of unknown function (DUF1307)/YehR-like	Protein of unknown function (DUF1307)/YehR-like	Putative lipoprotein YehR	YehR-like	Cluster 4046970 Protein of unknown function (DUF1307)/YehR-like
23	WP_048814497.1	Leucine Rich repeats	LRR-containing bacterial E3 ligase	leucine rich repeat protein/E3 ligase domain	Leucine Rich repeats	Cluster 4154032 Protein binding
24	WP_000301054.1	Lipopolysaccharide kinase (Kdo/WaaP)	Lipopolysaccharide kinase	Lipopolysaccharide kinase (Kdo/WaaP)	Lipopolysaccharide kinase (Kdo/WaaP)	Cluster 3990101 Lipopolysaccharide kinase
25	WP_000266171.1	Tetratricopeptide repeat	Tetratricopeptide-like domain	TPR repeat-containing protein	TPR-like	Cluster 4040666 Tetratricopeptide region
26	WP_000589825.1	OmpA family	OmpA-like domain	outer membrane lipoprotein	OmpA-like	Cluster 4198784 Outer membrane protein
27	WP_005051685.1	Peptidase family M23/LysM domain	Peptidase M23/LysM domain	Peptidase M23	Peptidoglycan hydrolase LytM	Cluster 4141397 Peptidase M23B
28	WP_001387238.1	RadC-like JAB domain	RadC protein	DNA repair protein RadC	JAB1/MPN domain	Cluster 4114260 RadC protein
29	WP_000248097.1	Phosphopantetheine attachment site	Phosphopantetheine binding ACP	Acyl carrier protein	Acyl-carrier protein (ACP)	Cluster 4146821 Phosphopantetheine

						-binding
30	WP_000848528.1	Lipoprotein leucine-zipper	Murein-lipoprotein	Major outer membrane lipoprotein	Outer membrane lipoprotein	Cluster 4066376 Murein-lipoprotein
31	WP_000189314.1	GIY-YIG catalytic domain	GIY-YIG endonuclease	No hit	GIY-YIG endonuclease	Cluster 4157077 GIY-YIG endonuclease
32	WP_001297375.1	RadC-like JAB domain	RadC-like JAB domain	DNA repair protein RadC	RuvA domain 2-like/JAB1/MPN domain	Cluster 4403730 RadC-like JAB domain
33	WP_000858193.1	yiaA/B two helix domain	YiaAB two helix	No hit	No result	Cluster 4440457 YiaAB two helix
34	WP_001296791.1	Autotransporter beta-domain	Autotransporter, YhjY	No hit	Autotransporter	Cluster 3853611 Autotransporter, YhjY
35	WP_000778795.1	Acetyltransferase (GNAT) domain	Acyl-CoA N-acyltransferase	Putative N-acetyltransferase	Acyl-CoA N-acyltransferases (Nat)	Cluster 4355736 N-acetyltransferase activity
36	WP_001205243.1	Xylose isomerase-like TIM barrel	Xylose isomerase-like TIM barrel	Putative hydroxypyruvate isomerase	Xylose isomerase-like	Cluster 4100779 Xylose isomerase-type TIM barrel
37	WP_001238362.1	Lipocalin-like domain	Lipocalin, ApoD type	Outer membrane lipoprotein Blc	Lipocalins	Cluster 4145424 Lipocalin
38	WP_000943980.1	Glutathionylspermidine synthase preATP-grasp	Glutathionylspermidine synthase, pre-ATP-grasp	No hit	Glutathione synthetase ATP-binding domain-like	Cluster 4243753 Glutathionylspermidine synthase
39	WP_000132640.1	Toxin SymE, type I toxin-antitoxin system	Type I toxin-antitoxin system, SymE toxin	No hit	No result	Cluster 4040297 Toxin SymE-like

Supplementary Table 6. List of annotated functions of 25 proteins with known function from *Shigella flexneri* using BLAST, Pfam, HmmScan, SMART, and Scanprosite for receiver operating characteristic analysis

Sl no	Acc ID protein	Protein name	BLAST	Pfam	HmmScan	SMART	Scanprosite
1	WP_000241642.1	Homoserine kinase	Homoserine kinase 1 (5)	Homoserine kinase 1 (5)	Homoserine kinase 1 (5)	Homoserine kinase 1 (5)	Homoserine kinase 1 (5)
2	WP_000809168.1	Protein hokC	Protein hokC 1 (5)	Protein hokC 1 (5)	Protein hokC 1 (5)	Protein hokC 1 (5)	Protein hokC 1 (5)
3	WP_001286897.1	Isoleucine--tRNA ligase	Isoleucine--tRNA ligase 1 (5)	Isoleucine--tRNA ligase 1 (5)	Isoleucine--tRNA ligase 1 (5)	Isoleucine--tRNA ligase 1 (5)	Isoleucine--tRNA ligase 1 (5)
4	WP_000124415.1	Ferrichrome porin FhuA	Ferrichrome outer membrane transporter 1 (4)	TonB dependent receptor 0 (2)	TonB dependent receptor 0 (2)	TonB dependent receptor 0 (2)	TonB dependent receptor 0 (2)
5	WP_001183183.1	MFS transporter	MFS transporter 1 (5)	MFS transporter 1 (5)	MFS transporter 1 (5)	MFS transporter 1 (5)	MFS transporter 1 (5)
6	WP_001230481.1	Ail/Lom family protein	OM_channels super family 1 (3)	Ail/Lom protein 1 (5)	Ail/Lom protein 1 (5)	Ail/Lom protein 1 (5)	Virulence outer membrane protein 1 (3)
7	WP_001287126.1	Glutamine--tRNA ligase	Glutaminyl-tRNA synthetase 1 (5)	tRNA synthetases 1 (5)	tRNA synthetases 1 (5)	tRNA synthetases 1 (5)	Aminoacyl-transfer RNA synthetases 1 (5)
8	WP_001295442.1	Flagellar L-ring protein	Flagellar basal body L-	Flagellar L-ring protein 1	Flagellar L-ring protein 1	Flagellar L-ring protein 1	PROKAR_LIPOPROTEIN 0 (3)

			ring protein 1 (4)	(5)	(5)	(5)	
9	WP_000130034.1	D-alanine--D-alanine ligase	D-alanine--D-alanine ligase 1 (5)	D-alanine--D-alanine ligase 1 (5)	D-alanine--D-alanine ligase 1 (5)	D-alanine--D-alanine ligase 1 (5)	D-alanine--D-alanine ligase 1 (5)
10	WP_000197853.1	Alanine racemase	Alanine racemase 1 (5)	Alanine racemase 1 (5)	Alanine racemase 1 (5)	Alanine racemase 1 (5)	Alanine racemase 1 (5)
11	WP_000569431.1	Ribonuclease HII	Ribonuclease HII 1 (5)	Ribonuclease HII 1 (5)	Ribonuclease HII 1 (5)	Ribonuclease HII 1 (5)	No hit 0 (5)
12	WP_000901098.1	VOC family protein	VOC family protein 1 (5)	VOC family protein 1 (5)	VOC family protein 1 (5)	VOC family protein 1 (5)	VOC family protein 1 (5)
13	WP_001260712.1	Proline--tRNA ligase	Proline--tRNA ligase 1 (5)	Proline--tRNA ligase 1 (5)	Proline--tRNA ligase 1 (5)	Proline--tRNA ligase 1 (5)	Proline--tRNA ligase 1 (5)
14	WP_000051892.1	Integrase	Integrase 1 (5)	Integrase 1 (5)	Integrase 1 (5)	Integrase 1 (5)	No hit 0 (5)
15	WP_001120449.1	Oxidoreductase	Oxidoreductase 1 (5)	Oxidoreductase 1 (5)	Oxidoreductase 1 (5)	Oxidoreductase 1 (5)	Oxidoreductase 1 (5)
16	WP_000460136.1	LysR family transcriptional regulator	LysR family transcriptional regulator 1 (5)	LysR family transcriptional regulator 1 (5)	LysR family transcriptional regulator 1 (5)	LysR family transcriptional regulator 1 (5)	LysR family transcriptional regulator 1 (5)
17	WP_001018618.1	Flavodoxin-1	Flavodoxin-1 1 (5)	Flavodoxin-1 1 (5)	Flavodoxin-1 1 (5)	Flavodoxin-1 1 (5)	Flavodoxin-1 1 (5)
18	WP_000773301.1	Acyl-CoA esterase	Acyl-CoA esterase 1 (5)	Acyl-CoA esterase 1 (5)	Acyl-CoA esterase 1 (5)	Acyl-CoA esterase 1 (5)	No hit 0 (5)
19	WP_000201488.1	DNA-packaging	DNA-packaging	DNA-packaging	DNA-packaging	DNA-packaging	No hit 0 (5)

		protein FI	protein FI 1 (5)	protein FI 1 (5)	protein FI 1 (5)	protein FI 1 (5)	
20	WP_005049594.1	Terminase	Terminase 1 (5)	Terminase 1 (5)	Terminase 1 (5)	Terminase 1 (5)	No hit 0 (5)
21	WP_000537402.1	Thioredoxin-disulfide reductase	Thioredoxin-disulfide reductase 1 (5)	Thioredoxin-disulfide reductase 1 (5)	Thioredoxin-disulfide reductase 1 (5)	Thioredoxin-disulfide reductase 1 (5)	Thioredoxin-disulfide reductase 1 (5)
22	WP_000109301.1	MFS transporter	MFS transporter 1 (5)	MFS transporter 1 (5)	MFS transporter 1 (5)	MFS transporter 1 (5)	MFS transporter 1 (5)
23	WP_005047463.1	Porin OmpA	Porin OmpA 1 (5)	Porin OmpA 1 (5)	Porin OmpA 1 (5)	Porin OmpA 1 (5)	Porin OmpA 1 (5)
24	WP_001247604.1	YjbF family lipoprotein	YjbF family lipoprotein 1 (5)	YjbF family lipoprotein 1 (5)	YjbF family lipoprotein 1 (5)	YjbF family lipoprotein 1 (5)	YjbF family lipoprotein 1 (5)
25	WP_014532269.1	DUF333 domain-containing protein	Domain of unknown function (DUF333) 1 (5)	Domain of unknown function (DUF333) 1 (5)	Domain of unknown function (DUF333) 1 (5)	Domain of unknown function (DUF333) 1 (5)	PROKAR_LIPOPROTEIN 0 (3)

True positive and true negative are denoted by “1” and “0”.

Integers in parentheses denote the confidence level.

Supplementary Table 7. List of annotated functions of 25 proteins with known function from *Shigella flexneri* using MOTIF, Interproscan, CATH, SUPERFAMILY, and ProtoNet for receiver operating characteristic analysis

Sl No.	Acc ID protein	Protein name	MOTIF	INTERPROSCAN	CATH	SUPERFAMILY	ProtoNet
1	WP_000241642.1	Homoserine kinase	Homoserine kinase 1 (5)	Homoserine kinase 1 (5)	Homoserine kinase 1 (5)	Homoserine kinase 1 (5)	Homoserine kinase 1 (5)
2	WP_000809168.1	Protein hokC	hok_gef 1 (3)	hok_gef 1 (3)	0 (5)	hok_gef 0 (5)	hok_gef 1 (3)
3	WP_001286897.1	Isoleucine--tRNA ligase	Isoleucine--tRNA ligase 1 (5)	Isoleucine--tRNA ligase 1 (5)	Isoleucine--tRNA ligase 1 (5)	Isoleucine--tRNA ligase 1 (5)	Isoleucine--tRNA ligase 1 (5)
4	WP_000124415.1	Ferrichrome porin FhuA	TonB-dependent Receptor 0 (2)	TonB-dependent siderophore receptor 0 (2)	TonB-dependent siderophore receptor 0 (2)	TonB-dependent siderophore receptor 0 (2)	TonB-dependent siderophore receptor 0 (2)
5	WP_001183183.1	MFS transporter	MFS transporter 1 (5)	MFS transporter 1 (5)	MFS transporter 1 (5)	MFS transporter 1 (5)	MFS transporter 1 (5)
6	WP_001230481.1	Ail/Lom family protein	Ail/Lom family protein 1 (5)	Ail/Lom family protein 1 (5)	Ail/Lom family protein 1 (5)	Ail/Lom family protein 1 (5)	Ail/Lom family protein 1 (5)
7	WP_001287126.1	Glutamine--tRNA ligase	Glutamine--tRNA ligase 1 (5)	Glutamine--tRNA ligase 1 (5)	Glutamine--tRNA ligase 1 (5)	Glutamine--tRNA ligase 1 (5)	Glutamine--tRNA ligase 1 (5)
8	WP_001295442.1	Flagellar L-ring protein	Flagellar L-ring protein 1 (5)	Flagellar L-ring protein 1 (5)	0 (5)	Flagellar L-ring protein 0 (5)	Flagellar L-ring protein 1 (5)

9	WP_000130034.1	D-alanine--D-alanine ligase	D-alanine--D-alanine ligase 1 (5)	D-alanine--D-alanine ligase 1 (5)	D-alanine--D-alanine ligase 1 (5)	D-alanine--D-alanine ligase 1 (5)	D-alanine--D-alanine ligase 1 (5)
10	WP_000197853.1	Alanine racemase	Alanine racemase 1 (5)	Alanine racemase 1 (5)	Alanine racemase 1 (5)	Alanine racemase 1 (5)	Alanine racemase 1 (5)
11	WP_000569431.1	Ribonuclease HII	Ribonuclease HII 1 (5)	Ribonuclease HII 1 (5)	Ribonuclease HII 1 (5)	Ribonuclease HII 1 (5	Ribonuclease HII 1 (5)
12	WP_000901098.1	VOC family protein	VOC family protein 1 (5)	VOC family protein 1 (5)	VOC family protein 1 (5)	VOC family protein 1 (5)	VOC family protein 1 (5)
13	WP_001260712.1	Proline--tRNA ligase	Proline--tRNA ligase 1 (5)	Proline--tRNA ligase 1 (5)	Proline--tRNA ligase 1 (5)	Proline--tRNA ligase 1 (5)	Proline--tRNA ligase 1 (5)
14	WP_000051892.1	Integrase	Integrase 1 (5)	Integrase 1 (5)	Integrase 1 (5)	Integrase 1 (5)	Integrase 1 (5)
15	WP_001120449.1	Oxidoreductase	Oxidoreductase 1 (5)	Oxidoreductase 1 (5)	Oxidoreductase 1 (5)	Oxidoreductase 1 (5)	Oxidoreductase 1 (5)
16	WP_000460136.1	LysR family transcriptional regulator	LysR family transcriptional regulator 1 (5)	LysR family transcriptional regulator 1 (5)	LysR family transcriptional regulator 1 (5)	LysR family transcriptional regulator 1 (5)	LysR family transcriptional regulator 1 (5)
17	WP_001018618.1	Flavodoxin-1	Flavodoxin_1, 3, 4, 5 1 (3)	Flavodoxin, long chain 1 (3)	Short-chain flavodoxin YkuP 1 (3)	Flavodoxin, long chain 1 (3)	Flavodoxin, long chain 1 (3)
18	WP_000773301.1	Acyl-CoA esterase	Alpha/beta hydrolase fold 1 (3)	Alpha/beta hydrolase fold, alpha/beta hydrolase fold, 1 1 (3)	Esterase YbfF 1 (5)	Alpha/beta hydrolase fold, alpha/beta hydrolase fold, 1 (3)	Alpha/beta hydrolase fold-1 1 (3)
19	WP_000201488.1	DNA-packaging	DNA-packaging	DNA-packaging protein FI 1 (5)	0 (5)	DNA-packaging protein FI 0 (5)	DNA-packaging

		protein FI	protein FI 1 (5)				protein FI 1 (5)
20	WP_005049594.1	Terminase	Phage terminase large subunit (GpA) 1 (3)	Bacteriophage lambda, GpA 1 (3)	0 (5)	Bacteriophage lambda, GpA 0 (5)	Phage terminase GpA 1 (3)
21	WP_000537402.1	Thioredoxin-disulfide reductase	Pyridine nucleotide-disulphide oxidoreductase 1 (3)	Pyridine nucleotide-disulphide oxidoreductase, class-II 1 (4)	Thioredoxin reductase 1 (4)	Pyridine nucleotide-disulphide oxidoreductase, class-II 1 (3)	Thioredoxin-disulfide reductase 1 (5)
22	WP_000109301.1	MFS transporter	MFS transporter 1 (5)	MFS transporter 1 (5)	MFS transporter 1 (5)	MFS transporter 1 (5)	MFS transporter 1 (5)
23	WP_005047463.1	Porin OmpA	Porin OmpA 1 (5)	Porin OmpA 1 (5)	Porin OmpA 1 (5)	Porin OmpA 1 (5)	Porin OmpA 1 (5)
24	WP_001247604.1	YjbF family lipoprotein	YjbF family lipoprotein 1 (5)	YjbF family lipoprotein 1 (5)	YjbF family lipoprotein 1 (5)	YjbF family lipoprotein 1 (5) YjbF family lipoprotein 1 (5)	YjbF family lipoprotein 1 (5)
25	WP_014532269.1	DUF333 domain-containing protein	DUF333 domain-containing protein 1 (5)	DUF333 domain-containing protein 1 (5)	0 (5)	DUF333 domain-containing protein 1 (5)	DUF333 domain-containing protein 1 (5)

True positive and true negative are denoted by “1” and “0.”

Integers in parentheses denote the confidence level.

Supplementary Table 8. List of predicted virulence factors of 39 hypothetical proteins by using VICMPred and Virulentpred

Sl. No.	Acc ID_Protein	VICMpred	Virulentpred
1	WP_005053355.1	Cellular process	Virulent
2	WP_000092054.1	Cellular process	Virulent
3	WP_001382892.1	Information and storage	Virulent
4	WP_005053036.1	Cellular process	Virulent
5	WP_000779831.1	Cellular process	Virulent
6	WP_011110552.1	Information and storage	Virulent
7	WP_001269672.1	Metabolism Molecule	Virulent
8	WP_001247854.1	Virulence factors	Non-Virulent
9	WP_000070107.1	Metabolism Molecule	Non-Virulent
10	WP_000224274.1	Cellular process	Non-Virulent
11	WP_000749269.1	Virulence factors	Virulent
12	WP_001125713.1	Cellular process	Virulent
13	WP_001043881.1	Cellular process	Non-Virulent
14	WP_001295493.1	Metabolism Molecule	Non-Virulent
15	WP_000691930.1	Cellular process	Virulent
16	WP_000597196.1	Metabolism Molecule	Virulent
17	WP_000248636.1	Metabolism Molecule	Non-Virulent
18	WP_000755956.1	Metabolism Molecule	Non-Virulent
19	WP_001237866.1	Cellular process	Virulent
20	WP_000454701.1	Metabolism Molecule	Non-Virulent
21	WP_000003197.1	Cellular process	Virulent
22	WP_005049020.1	Cellular process	Non-Virulent
23	WP_048814497.1	Cellular process	Virulent
24	WP_000301054.1	Metabolism Molecule	Non-Virulent
25	WP_000266171.1	Metabolism Molecule	Non-Virulent
26	WP_000589825.1	Cellular process	Virulent
27	WP_005051685.1	Cellular process	Virulent
28	WP_001387238.1	Cellular process	Virulent
29	WP_000248097.1	Cellular process	Virulent
30	WP_000848528.1	Cellular process	Virulent
31	WP_000189314.1	Cellular process	Virulent
32	WP_001297375.1	Metabolism Molecule	Non-Virulent
33	WP_000858193.1	Cellular process	Non-Virulent
34	WP_001296791.1	Cellular process	Virulent
35	WP_000778795.1	Cellular process	Virulent

36	WP_001205243.1	Cellular process	Non-Virulent
37	WP_001238362.1	Cellular process	Non-Virulent
38	WP_000943980.1	Cellular process	Non-Virulent
39	WP_000132640.1	Cellular process	Non-Virulent