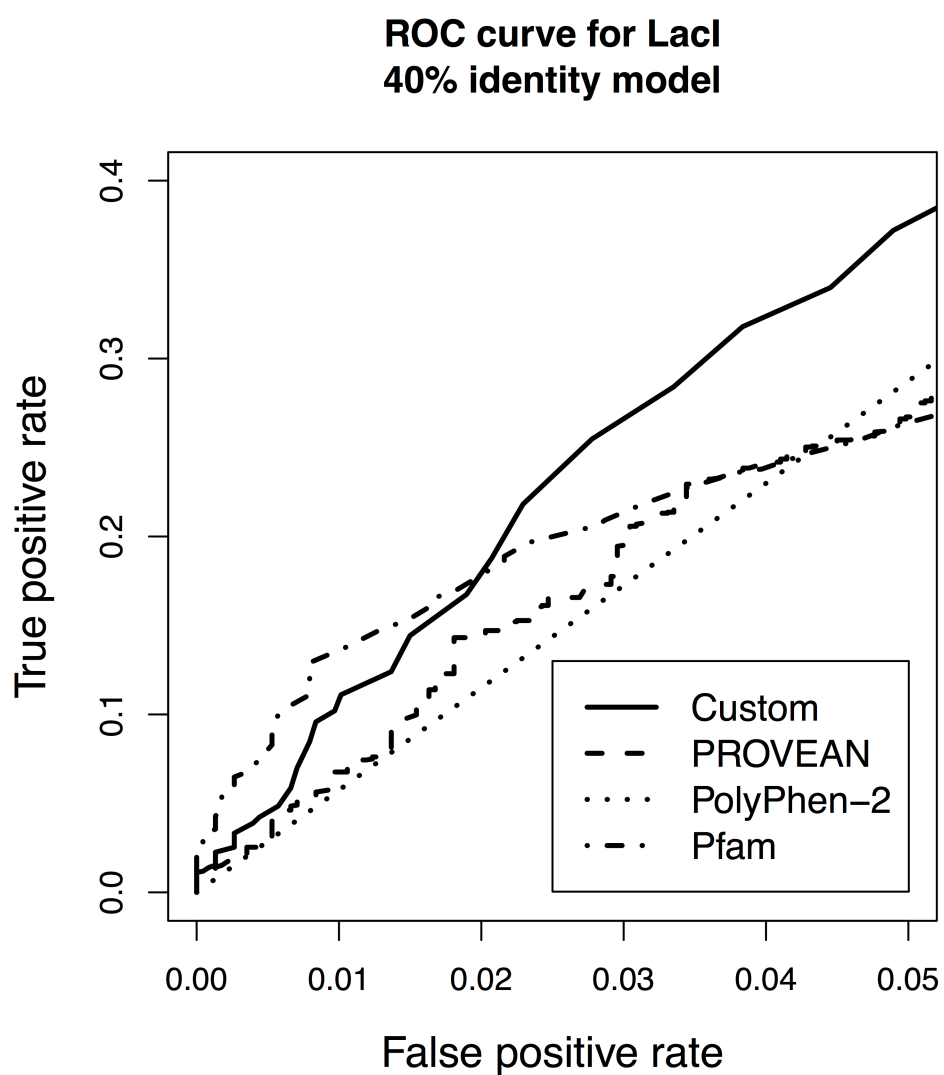
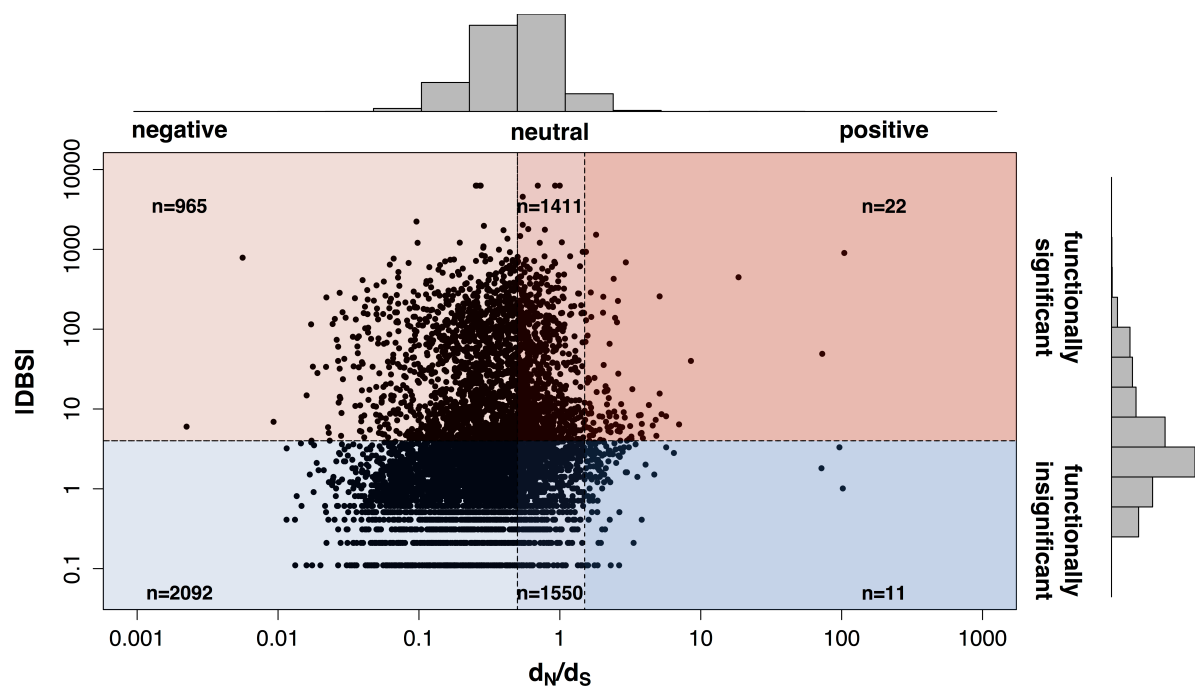


S1 Fig. Illustration of DBS calculation on a per-residue basis. This illustration uses two proteins that have accumulated a number of non-synonymous changes between *E. coli* and *S. Enteritidis*, showing a protein that scored a high DBS (*E. coli* gene b0589, *S. Enteritidis* gene SEN0560) and a protein that scored a low DBS (*E. coli* gene b0064, *S. Enteritidis* gene SEN0105). The high scoring protein has a number of positions with high positive DBS which strongly influence the overall DBS for the protein. The low scoring protein has DBS values that are smaller in magnitude and have little individual impact on the overall score. DBS values in the positive and negative directions are of a similar magnitude, so cancel out to some extent.



S2 Fig. Closer view of the ROC curve for the LacI predictions, FPR up to 5%. Pfam performs well to begin with, then performance relative to other methods declines with increasingly permissive scoring thresholds.



S3 Fig. Distribution of |DBS| vs dN/dS for orthologous genes in human and chimpanzee, filtered for |DBS|>0, dN>0 and dS>0. Cutoffs set at 0.5 and 1.5 for dN/dS, and Benjamini-Hochberg adjusted p-value of 0.05 for DBS. Correlation between |DBS| and dN/dS is 0.20. Correlation between |DBS| and dN is 0.39.

Supplementary Table 1: Number of human variants from the humsavar database that could be scored by each predictive method.

Method	Number of variants scored
Pfam	51247
Treefam	51122
Custom	21772
FATHMM	61820
PolyPhen	61601
PROVEAN	61820

Supplementary Table 2: GO terms showing enrichment/depletion in human polymorphisms with DBS between 5 and 10 when scored against the Treefam HMMs. Data generated using AmiGO 2 (<http://amigo.geneontology.org/rte>).

GO Term	Background frequency	Sample frequency	Expected	+/-	P-value
molecular function (GO:0003674)	16178	446	366.50	+	4.66E-16
ion binding (GO:0043167)	5913	209	134.00	+	6.27E-10
catalytic activity (GO:0003824)	5495	197	124.50	+	1.42E-09
anion binding (GO:0043168)	2602	112	58.95	+	4.24E-08
molecular transducer activity (GO:0060089)	1829	86	41.44	+	2.45E-07
protein kinase activity (GO:0004672)	593	41	13.44	+	1.25E-06
signal transducer activity (GO:0004871)	1614	77	36.57	+	1.47E-06
phosphotransferase activity, alcohol group as acceptor (GO:0016773)	698	45	15.81	+	1.52E-06
ATP binding (GO:0005524)	1443	71	32.69	+	2.13E-06
carbohydrate derivative binding (GO:0097367)	2162	93	48.98	+	3.57E-06
receptor activity (GO:0004872)	1504	72	34.08	+	5.03E-06
adenyl ribonucleotide binding (GO:0032559)	1479	71	33.51	+	5.90E-06
adenyl nucleotide binding (GO:0030554)	1486	71	33.67	+	7.16E-06
purine ribonucleoside triphosphate binding (GO:0035639)	1773	79	40.17	+	1.76E-05
kinase activity (GO:0016301)	816	47	18.49	+	2.06E-05
binding (GO:0005488)	13616	368	308.50	+	2.11E-05
purine ribonucleoside binding (GO:0032550)	1782	79	40.37	+	2.19E-05
purine ribonucleotide binding (GO:0032555)	1815	80	41.12	+	2.22E-05
purine nucleoside binding (GO:0001883)	1785	79	40.44	+	2.35E-05
ribonucleoside binding (GO:0032549)	1786	79	40.46	+	2.41E-05
signaling receptor activity (GO:0038023)	1286	63	29.14	+	2.51E-05
transferase activity, transferring phosphorus-containing groups (GO:0016772)	963	52	21.82	+	2.61E-05
purine nucleotide binding (GO:0017076)	1824	80	41.33	+	2.75E-05
protein serine/threonine kinase activity (GO:0004674)	437	32	9.90	+	2.85E-05
nucleoside binding (GO:0001882)	1796	79	40.69	+	3.06E-05
ribonucleotide binding (GO:0032553)	1830	80	41.46	+	3.17E-05
small molecule binding (GO:0036094)	2538	101	57.50	+	3.43E-05
oxygen binding (GO:0019825)	47	11	1.07	+	4.11E-05
oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen (GO:0016712)	30	9	0.68	+	1.04E-04
transmembrane signaling receptor activity (GO:0004888)	1185	58	26.85	+	1.06E-04
organic cyclic compound binding (GO:0097159)	5763	181	130.60	+	9.77E-04
steroid hydroxylase activity (GO:0008395)	30	8	0.68	+	1.40E-03
transferase activity (GO:0016740)	2146	84	48.62	+	1.46E-03
heme binding (GO:0020037)	134	15	3.04	+	1.61E-03
nucleotide binding (GO:0000166)	2276	87	51.57	+	2.35E-03
nucleoside phosphate binding (GO:1901265)	2277	87	51.59	+	2.39E-03
tetrapyrrole binding (GO:0046906)	143	15	3.24	+	3.55E-03
iron ion binding (GO:0005506)	167	16	3.78	+	5.06E-03
arachidonic acid epoxygenase activity (GO:0008392)	16	6	0.36	+	5.36E-03
arachidonic acid monooxygenase activity (GO:0008391)	16	6	0.36	+	5.36E-03
heterocyclic compound binding (GO:1901363)	5690	175	128.90	+	7.07E-03
cation binding (GO:0043169)	4059	133	91.96	+	9.06E-03
monooxygenase activity (GO:0004497)	101	12	2.29	+	1.15E-02
oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen (GO:0016705)	165	15	3.74	+	1.94E-02
metal ion binding (GO:0046872)	3984	128	90.26	+	3.92E-02

Supplementary Table 3: *P* values for sign test of whether there is an enrichment for positive DBS values.

	S.Enteritidis	S. Typhimurium	S. Heidelberg	S. Gallinarum	S. Typhi	S. Paratyphi
S.Enteritidis						
S. Typhimurium	9.57E-01					
S. Heidelberg	9.97E-01*	1.59E-01				
S. Gallinarum	3.07E-32	5.77E-22	1.07E-23			
S. Typhi	4.08E-12	2.83E-20	3.52E-16	6.92E-01		
S. Paratyphi	1.28E-06	2.64E-13	2.08E-10	9.91E-01*	9.22E-01	

* Significant enrichment for negative DBS values (*P* = 0.01 for Enteritidis-Heidelberg and 0.005 for Gallinarum-Paratyphi)

Supplementary Table 4: Putative LOF mutations in genes involved in anaerobic metabolism of *S. Enteritidis* and *S. Gallinarum* (according to the classification in Table S7 of Nuccio and Bäumlér, 2014, using our scoring method). Newly identified LOFs are in bold.

Locus tag (E)	Locus tag (G)	Pfam domain	DBS	p-value	Gene name	Gene product	Nature of mutation
SEN2024	SG2051	CbiJ	34.1	4.08e-55	cbiJ	cobalt-precorrin-6a reductase	Truncation
SEN2371	SG2419	Thiolase_N	29.3	9.55e-41	yfcY	putative 3-ketoacyl-CoA thiolase	Truncation
SEN2038	SG2067	Dehydratase_LU	16	5.5e-12	pduC	glycerol dehydratase large subunit	Point mutations
SEN2334	SG2381	BPD_transp_1	14.5	8.31e-10	hisM	histidine transport system permease	Point mutations
SEN3490	SG3763	Ldh_2	10.9	1.94e-05	yiaK	putative carboxylic acid dehydrogenase	Small insertion
SEN1458	SG1526	Aldedh	8.1	0.00674	ydcW	putative aldehyde dehydrogenase	Point mutation
SEN0506	SG0537	Gly_kinase	6.9	0.0427	glxK	glycerate kinase	Point mutations