

Predicting potential anti-codon shifts (alloaccepter shifts) in TryTryp tRNA models

After investigating each of the sequences in previous report, I listed some of the most possible anti-codon shifts here.

geneid	func	func_score	pot_shifts	pot_shifts_s	tseac	araac	shifts:	substitution
PconfusumCUL13_Pcon_scaffold00007_3_45	P	-144	H	-40	GGG	(ggg)	H(GUG) > P(GGG)	1
TbruceiTREU927_Tb927_11_bin_v5.1_73	R	-94	I	-81	CCG	(ccg)	I > R(ccg)	3
TbruceiTREU927_Tb927_11_bin_v5.1_74	H	-122	W	-105	GTG	(gtg)	W(CCA) > H (GTG)	3
TbruceiTREU927_Tb927_11_bin_v5.1_76	P	-139	R	-118	TGG	(tgg)	R(TCG) > P(TGG)	1
TcruziCLBrener_Tcruzi_11019_2	I	-116	M W	-102 -102	GAT	(gat)	M(CAT) > I(GAT)	1
TcruziCLBrener_Tcruzi_20084_17	C	-122	S	-116	GCA	(gca)	S(GCT/GGA) > C(GCA)	1
TcruziCLBrener_Tcruzi_22215_21	M	-143	C	-86	CAT	(cat)	C > M(CAT)	3
TcruziCLBrener_Tcruzi_22215_22	M	-103	I M	-103 -103	CAT	(cat)	I(GAT) > M(CAT)	1
TcruziCLBrener_Tcruzi_22215_24	P	-144	W	-97	TGG	(tgg)	H(ATG) > P(TGG)	2
TcruziCLBrener_Tcruzi_22215_25	S	-136	T	-116	TGA	(tga)	T(TGT) > S(TGA)	1
TcruziCLBrenerNon-Esmeraldo-like_TcChr21-P_8	V	-58	G	-31	CAC	(cac)	G(CCC) > V(CAC)	1
TcruziEsmeraldo_ANOX01006378.1_11	G	-140	H	-115	GCC	(gcc)	I(GAT)/H(GTG) > G(GCC)	2
TcruziEsmeraldo_KB205213.1_31	V	-58	G	-31	CAC	(cac)	G(CCC) > V(CAC)	1
TcruziEsmeraldo_KB205435.1_59	V	-58	G	-31	CAC	(cac)	G(CCC) > V(CAC)	1
TcruziEsmeraldo_KB205928.1_89	V	-58	G	-31	CAC	(cac)	G(CCC) > V(CAC)	1
TcruziJRcl4_AODP01003975.1_2	S	-144	R	-110	GGA	(gga)	R(GCG) > S(GGA)	1
TcruziJRcl4_AODP01004654.1_4	V	-107	I	-15	AAC	(aac)	I(AAT) > V(AAC)	1
TcruziTulac12_KB851446.1_107	V	-58	G	-31	CAC	(cac)	G(CCC) > V(CAC)	1
TvixaxY486_TvY486_bin_archived_contig_1183_36	K	-110	R	-94	TTT	(ttt)	R(TCT)/I(TAT) > K(TTT)	1
TvixaxY486_TvY486_bin_archived_contig_3413_48	H	-123	V	-121	GTG	(gtg)	V(GTC) > H(GTG)	1
TvixaxY486_TvY486_bin_archived_contig_3413_51	K	-110	R	-94	TTT	(ttt)	R(TCT)/I(TAT) > K(TTT)	1
TvixaxY486_TvY486_bin_archived_contig_4539_66	N	-118	R	-106	GTT	(gtt)	R(TCT/CCT) > N(GTT)	3

func: Functions predicted by gene-finders

func_score: The score of the functions in column "Func" from the classifier.

pot_shifts : Profile models with the highest score from the classifier.

pot_shifts_s : The highest score from the classifier (score of models listed in column pot_shifts)

tseac/araac : TSE/ARA anti-codon

shift: Most possible anti-codon shifts

substitution : Minimum number of substitutions needed for the shifts

For example: tRNA sequence with geneid "TcruziJRcl4_AODP01004654.1_4" is predicted as model V with anti-codon AAC. However, it is very similar to model I from classifier, and the best anti-codon distance from model I to model V(with anti-codon AAC) is only 1 substitution. Also, the sample size of these two models is relatively high, so the models are relatively reliable.

Model:	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	X	Y	Z
frequency:	211	64	105	160	104	228	80	171	183	336	97	125	200	162	347	228	218	241	52	76	78	76