

Name	Nb aa	MW (kDa)	Access Number (P)	Organism Species	Identity	Tissue Chr Gene locus	Function
AaegCWPdAN4X4	482	51.4	XP_021703255	<i>Aedes aegypti</i>	DAN4 isoform X4	Whole pupae, tegument Chr2	Cell wall protein
AaquTIF2	351	37.7	XP_050101586	<i>An. aquasalis</i>	Translation initiation factor IF-2-like isoform X3	Chr2	“OS-D-like” Translation initiator
AaquWAS/WASL	418	45.4	XP_050101585	<i>Anopheles aquasalis</i>	WAS/WASL-interacting protein family member 1-like isoform X2	Chr2 LOC126581753	“OS-D-like” Mucin-2
AEalbSamkC	408	44.3	XP_029726842	<i>Ae. albopictus</i>	SamkC isoform X9	All body Chr2	Ser/Thr-protein kinase
AglaCWPX3	243	26.7	XP_018563027	<i>Anoplophora glabripennis</i>	CWP3-like isoform X3	Larvae LOC108904830	“OS-D” Integrity of cell wall Stress response
ANalbLRRX4	489	52.6	XP_035778647	<i>Anopheles albimanus</i>	LRR-like protein 5 isoform X4	Adult female, whole body Chr2	“OS-D” Leucine-rich repeat extensin
AcrMucin2	548	58.6	XP_052865497	<i>An. cruzii</i>	Mucin2-like isoform 1	Tegument Chr3	Mucin
AcrWAS/WASL	396	43.6	XP_052865501	<i>An. cruzii</i>	WAS/WASL isoform X4 UL36	Tegument Chr3	WAS/WASL-interacting protein
AmerMucin2	448	47.8	XP_041777817	<i>An. merus</i>	Mucin-2-like isoform X5	Adults, whole body Chr3R	Mucin
AmerMucin5AC	452	48.7	XP_041777815	<i>An. merus</i>	Sec-31-like isoform X3	Adults, whole body	Protein transport protein
AMS2711750	96	9.6	KPK83846	<i>Bacteroides</i> sp. SM23_62_1 Bacteria Bacteroidia Bacteroidales	Hypothetical protein AMS27_11750	Sediment metagenome	Hypothetical
AplanCSP	325	37.3	CAB3231652	<i>Arctia plantaginis</i>	Unnamed	LOCUS4521	“OS-D”-like
CAUAvd	136	16.1	DAK32723	<i>Heunggongvirae</i> Caudoviricetes, Viruses	*ctUaT18 Avd-like	Human metagenome	DGR-protein
CparMraZ	164	18.7	HHE32374	<i>Chlorobaculum parvum</i> Bacteria Chlorobiia Chlorobiales	MraZ	Hydrothermal vent metagenome	Transcriptional regulator
CsecThap1	126	14.4	XP_023715540	<i>Cryptotermes secundus</i>	Tha p 1	Workers, whole body	Allergen
CUPHUBBP	152	15.8	WP_035823962	<i>Cupriavidus</i> sp. SK-4 Bacteria β-Proteobacteria; Burkholderiales	DNA-binding protein HU family		Histone-like DBP Chromosome condensation
EbalCSP3	213	24.2	QIS77190	<i>Episyrphus balteatus</i>	“CSP”		“OS-D”-like

EbalCSP4	321	35.3	QIS77191	<i>E. balteatus</i>	“CSP”		“OS-D”-like
EbalDDBG0285119 X1	371	41.1	XP_055839111	<i>E. balteatus</i>	DDB_G0285119 isoform X1	Chr1	“OS-D-like”
EcorRho	361	39.4	XP_055921371	<i>Eupeodes corollae</i>	Rho17 isoform X3	Chr3	Rho GTPase-activator
HhatDBPXRE	154	18.0	EFC99453	<i>Hungatella hathewayi</i> DSM 13479 Bacteria, Clostridia, Eubacteriales	DNA-binding helix-turn-helix protein	Cytoplasm, CLOSTHATH_02328	Transcriptional regulator XRE-family
IpodCSP	263	30.7	CAH2042437	<i>Iphiclides podalirius</i>	Unnamed	Chr14 IPOD504_LOCUS3820	“OS-D”-like
LborTetR/AcrR	170	19.2	WP_011671486 WP_061220709	<i>Leptospira borgpetersenii</i> Bacteria Spirochaetia Leptospirales	TetR/AcrR		Transcriptional regulator Regulation of DNA-templated transcription
LlonGTA1	475	52.6	WP_058530090	<i>Legionella londiniensis</i> Bacteria g-Proteobacteria Legionellales	Mannose-1-P-guanylyl transferase/ mannose-6-phosphate isomerase		Mannose Glycosyl transferase family A
LlonGTA2	486	54.0	KTD19891	<i>L. londiniensis</i> Bacteria g-Proteobacteria Legionellales	Phospho mannose isomerase GDP mannose pyrophosphorylase	Llon_ctg033	P-mannose Glycosyl transferase family A
OvarDBPXRE	143	16.3	BAL01528	<i>Oscillibacter valericigenes</i> Sjm18-20, Bacteria, Clostridia, Eubacteriales	DNA-binding protein	OBV_43290	Transcriptional regulator XRE-family Nucleotide binding
PaegJg6675	217	25.0	CAH2266975	<i>Pararge aegeria aeg.</i>	Jg6675	PAEG_Locus 25571	“OS-D”-like
PlutMraZ	153	17.3	KZK74918 WP_303680859 WP_011358831	<i>Pelodictyon luteolum</i> (<i>Chlorobium luteolum</i>) Bacteria Chlorobiia Chlorobiales	Division/cell wall cluster transcriptional repressor MraZ	Lake water metagenome	Transcriptional regulator
PpyrMucin5AC	338	37.7	XP_031351203	<i>Photinus pyralis</i>	Mucin-5AC -like	Adult male, whole body LOC116176644	“OS-D-like” Mucin
SFRURICE011779	265	30.8	KAF9818733	<i>Spodoptera frugiperda</i>	A10/OS-D	Adult males, whole body	“OS-D-like”
SfulWP223199665	67	15.5	WP_223199665	<i>Solihabitans fulvus</i> Bacteria, Actinomycetes,	A10/OS-D		“OS-D-like”

				Pseudonocardiales			
TechAcrR	193	20.7	SEF50864	<i>Thermonospora echinospora</i> Bacteria, Actinomycetota, Actinomycetes, Streptosporangiales	**AcrR family DNA-binding protein	SAMN0448971 2_101243	DNA-binding transcriptional regulator
RtorDBPXRE1	138	15.6	WP_044998036	<i>Ruminococcus torques</i> Bacteria Clostridia, Eubacteriales	Helix-turn- helix Transcriptional regulator		Transcriptional regulator XRE-family Nucleotide binding
RtorDBPXRE2	146	16.7	EDK23754	<i>Ruminococcus torques</i> Bacteria, Clostridia, Eubacteriales	DNA-binding helix-turn-helix protein	Contig63	Transcriptional regulator XRE-family Nucleotide binding
WP016226835	135	15.7	EOS77856 WP_016226835	<i>Lachnospiraceae bacterium</i> 10-1 Bacteria, Clostridia, Eubacteriales	Hypothetical	C819_00471	Hypothetical
WP020776583	163	18.9	WP_020776583	<i>Leptospira meyeri</i> Bacteria Spirochaetia Leptospirales	Hypothetical		Hypothetical
WP004787694	163	19.0	WP_004787694	<i>L. meyeri</i> Bacteria Spirochaetia Leptospirales	Hypothetical		Hypothetical
WsmiSamkC	520	56.2	XP_055540454	<i>Wyeomyia smithii</i>	Serine/threonine -protein kinase SamkC isoform X4	Adult male, whole body Chr3 LOC129727069	Ser/Thr kinase SamkC

Table S3. Sequences producing significant PAUP-Tree between 4CSP and intracellular proteins. 4CSP-related intracellular proteins are CWP, LRR, Rho, SamkC, Sec31, TIF, WAS-WASL, and DBPs (transcriptional regulators of the XRE-family).

*Viral Accessory variability determinant (Avd) proteins, such as ctUaT18, are a subset of Diversity Generating Retroelements (DGR) genetic cassettes that mutate target genes only in a specific way to produce hypervariable proteins.

**Bacterial AcrR is a single-component repressor of the antibiotic resistance and lipid transport genes.

Chr: Chromosome.