

Supplementary Materials

DeepT3 2.0: improving type III secreted effector predictions by an
integrative deep-learning framework

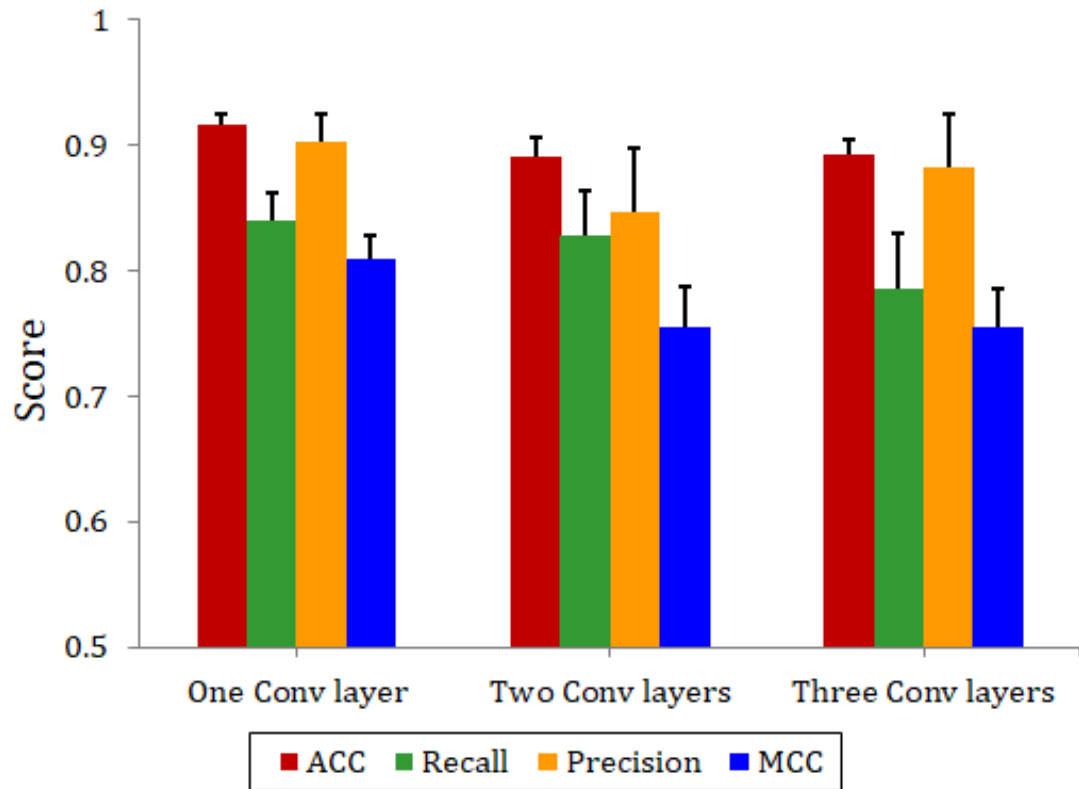
Runyu Jing^{1,2}, Tingke Wen¹, Chengxiang Liao¹, Li Xue³, Fengjuan Liu⁴, Lezheng
Yu^{5,*} and Jiesi Luo^{6,7,8*}

¹School of Cyber Science and Engineering, Sichuan University, Chengdu 610065, China, ²Medical Big Data Center, Sichuan University, Chengdu 610065, China, ³School of Public Health, Southwest Medical University, Luzhou 646000, China, ⁴School of Geography and Resources, Guizhou Education University, Guiyang 550018, China, ⁵School of Chemistry and Materials Science, Guizhou Education University, Guiyang 550018, China, ⁶Department of Pharmacology, School of Pharmacy, Southwest Medical University, Luzhou 646000, China, ⁷Department of Pharmacy, The Affiliated Hospital of Southwest Medical University, Luzhou 646000, China and ⁸Sichuan Key Medical Laboratory of New Drug Discovery and Druggability Evaluation, Luzhou Key Laboratory of Activity Screening and Druggability Evaluation for Chinese Materia Medica, Southwest Medical University, Luzhou 646000, China

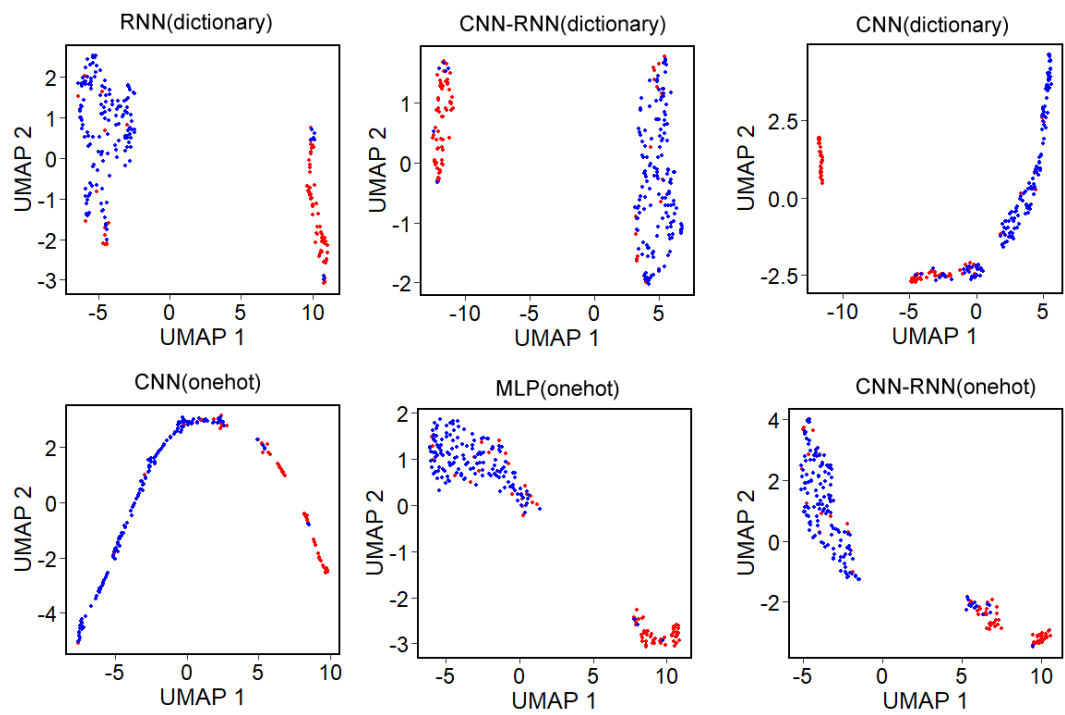
Supplementary Figures



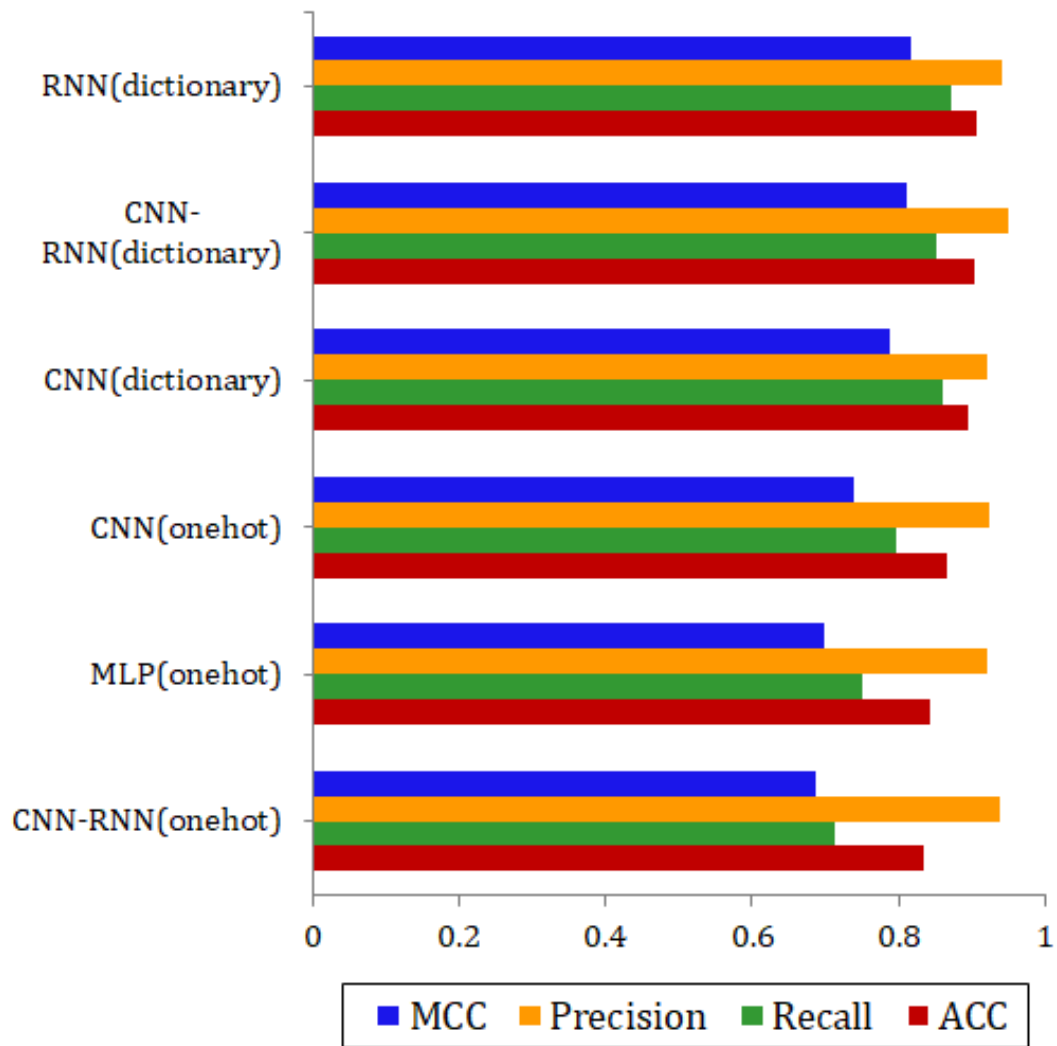
Supplementary Figure 1. Hyperparametric optimization of four deep learning models using the *MCC* values.



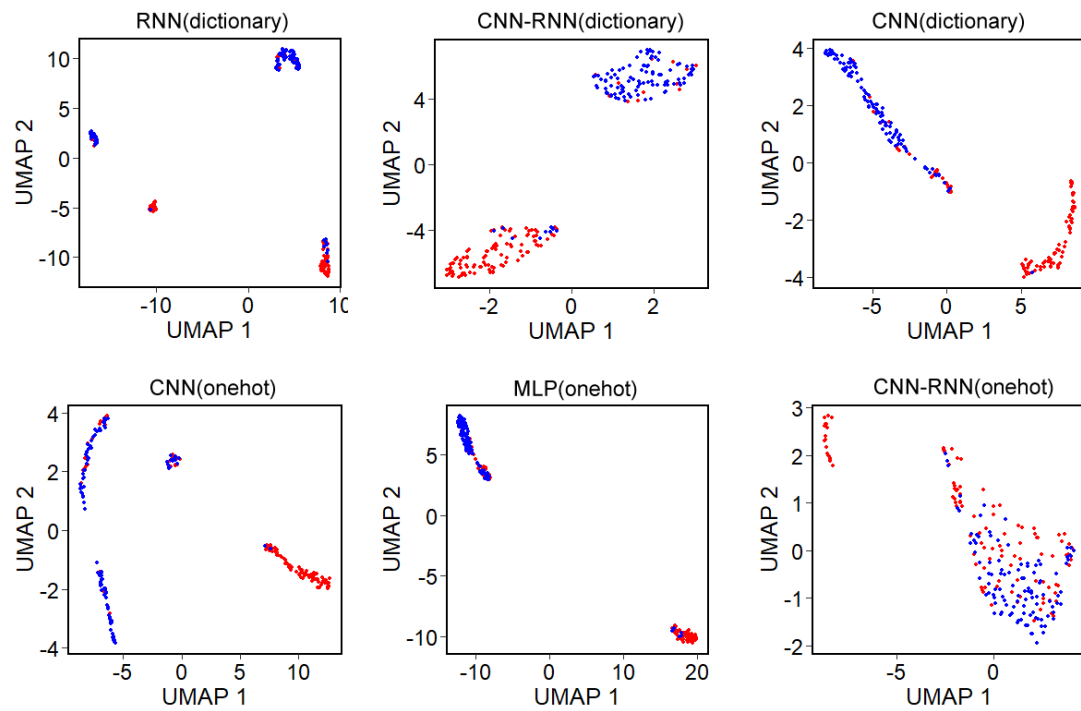
Supplementary Figure 2. Performance comparison of *CNN-RNN(dictionary)* models with different number of convolutional layers.



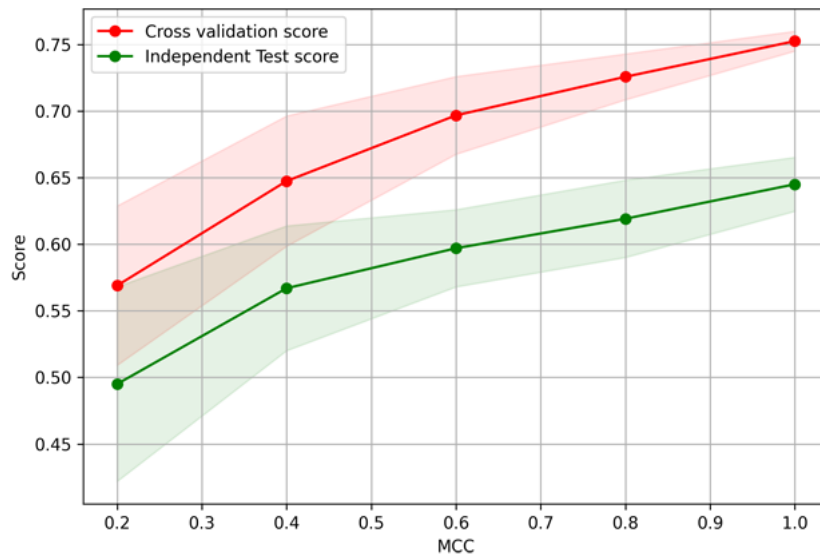
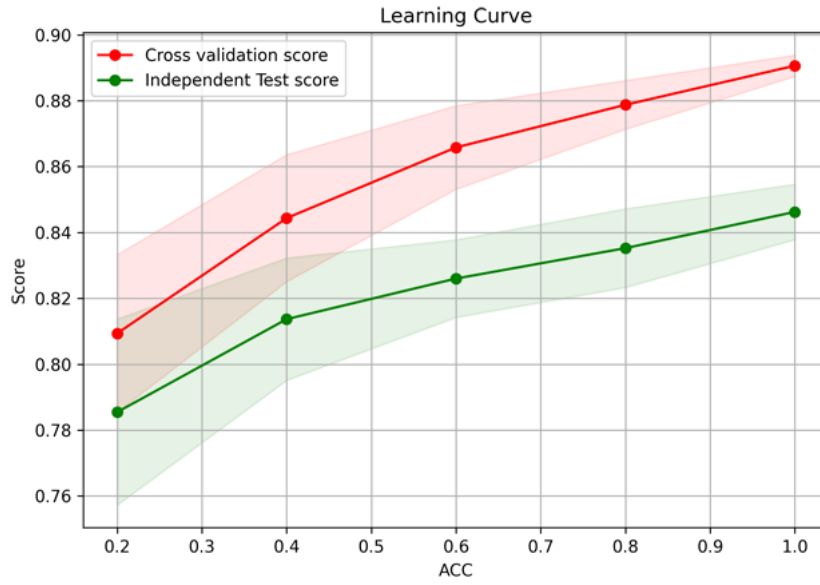
Supplementary Figure 3. Comparison of the last hidden layer representations of the test set for the six deep learning models.



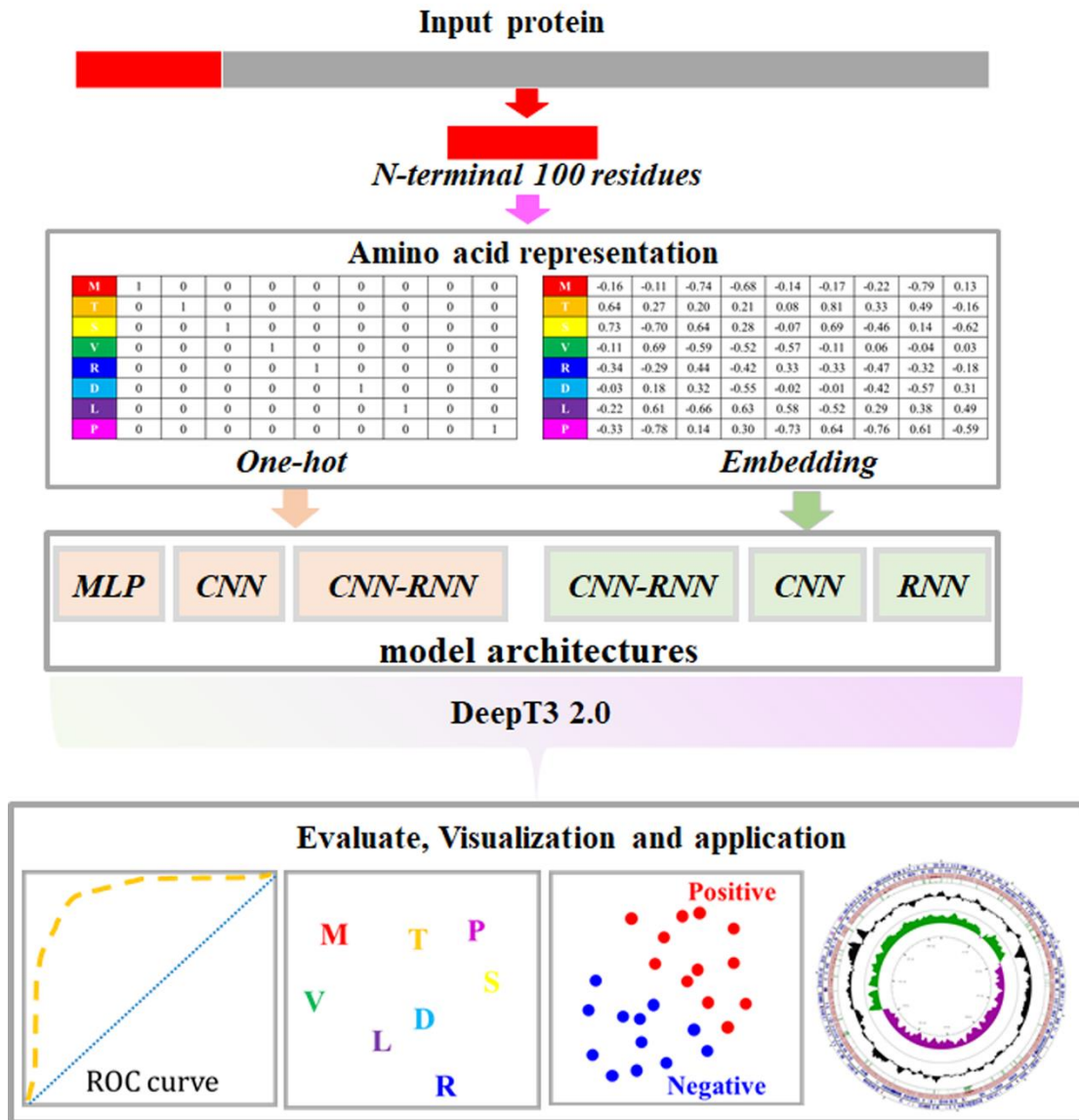
Supplementary Figure 4. Performance comparison of different models on the Bastion3 independent set.



Supplementary Figure 5. Comparison of the last hidden layer representations of the Bastion3 independent set for the six deep learning models.



Supplementary Figure 6. Learning curves of DeepT3 2.0 using the *ACC* and *MCC* scores.



Supplementary Figure 7. The overall flow chart of DeepT3 2.0.

Supplementary Tables

Supplementary Table 1. DeepT3 2.0 predicted 200 putative effector candidates in the *Chlamydia trachomatis* genome

ID	Name
WP_009873042.1	SctW family type III secretion system gatekeeper subunit CopN
WP_009871965.1	inclusion-associated protein
WP_009871941.1	SycD/LcrH family type III secretion system chaperone Scc2
WP_009872104.1	cadmium-translocating P-type ATPase
WP_009872009.1	exodeoxyribonuclease V subunit beta
WP_009871579.1	inclusion membrane protein IncB
WP_010725191.1	autotransporter domain-containing protein
WP_010725099.1	DUF1669 domain-containing protein
WP_009873324.1	aminotransferase class V-fold PLP-dependent enzyme
WP_009872236.1	DUF720 domain-containing protein
WP_010725201.1	(d)CMP kinase
WP_010725149.1	IncA family protein
WP_010725128.1	IncA family protein
WP_009873080.1	IncA family protein
WP_009871620.1	DUF5399 domain-containing protein
WP_009872237.1	DUF720 domain-containing protein
WP_010725368.1	CT847 family type III secretion system effector
WP_010725374.1	type III secretion system translocon subunit CopB2
WP_010725373.1	type III secretion system translocon subunit CopD2
WP_009871944.1	type III secretion system translocon subunit CopD
WP_009871943.1	type III secretion system translocon subunit CopB
WP_009871695.1	endopeptidase La
WP_009872625.1	nucleotide exchange factor GrpE

WP_010725304.1	type III secretion system effector protein
WP_010725032.1	type III secretion system effector protein
WP_009872258.1	deubiquitinase
WP_009872257.1	deubiquitinase
WP_009871466.1	inclusion membrane protein IncA
WP_009871465.1	inclusion membrane protein IncG
WP_010725172.1	chorismate synthase
WP_010724989.1	signal recognition particle protein
WP_010725336.1	tRNA (adenosine(37)-N6)-dimethylallyltransferase MiaA
WP_010725333.1	alanine--tRNA ligase
WP_009871719.1	3-phosphoshikimate 1-carboxyvinyltransferase
WP_010725163.1	DNA polymerase III subunit gamma/tau
WP_009871828.1	DNA repair protein RecO
WP_010725310.1	trigger factor
WP_009872047.1	protein arginine kinase
WP_009871354.1	ribonuclease HIII
WP_010725202.1	type III secretion system actin-recruiting effector Tarp
WP_010725041.1	4-alpha-glucanotransferase
WP_009871788.1	uroporphyrinogen-III synthase
WP_009871687.1	HPr family phosphocarrier protein
WP_009871783.1	UPF0158 family protein
WP_009873064.1	ABC transporter ATP-binding protein
WP_009871480.1	class I SAM-dependent methyltransferase
WP_009873051.1	phosphatidylserine/phosphatidylglycerophosphate/cardiolipin synthase family protein
WP_009872088.1	hypothetical protein
WP_009871795.1	hypothetical protein
WP_009872189.1	insulinase family protein
WP_009872094.1	hypothetical protein

WP_009871688.1	hypothetical protein
WP_010725165.1	HEAT repeat domain-containing protein
WP_009871736.1	HEAT repeat domain-containing protein
WP_010724983.1	PhoH family protein
WP_010725193.1	Pmp family polymorphic membrane protein autotransporter adhesin
NP_220390.1	deubiquitinase [Chlamydia trachomatis D/UW-3/CX]
NP_220389.1	deubiquitinase [Chlamydia trachomatis D/UW-3/CX]
NP_220385.1	hypothetical protein CT_863 [Chlamydia trachomatis D/UW-3/CX]
NP_220383.1	type III secretion system translocator CopB2 [Chlamydia trachomatis D/UW-3/CX]
NP_220382.1	type III secretion system translocator CopD2 [Chlamydia trachomatis D/UW-3/CX]
NP_220370.1	hypothetical protein CT_849 [Chlamydia trachomatis D/UW-3/CX]
NP_220369.1	hypothetical protein CT_848 [Chlamydia trachomatis D/UW-3/CX]
NP_220368.1	hypothetical protein CT_847 [Chlamydia trachomatis D/UW-3/CX]
NP_220333.1	inclusion membrane protein [Chlamydia trachomatis D/UW-3/CX]
NP_220326.1	insulinase/protease [Chlamydia trachomatis D/UW-3/CX]
NP_220287.1	hypothetical protein CT_768 [Chlamydia trachomatis D/UW-3/CX]
NP_220285.1	tRNA dimethylallyltransferase [Chlamydia trachomatis D/UW-3/CX]
NP_220268.1	alanine--tRNA ligase [Chlamydia trachomatis D/UW-3/CX]
NP_220246.1	metal transporting ATPase [Chlamydia trachomatis D/UW-3/CX]
NP_220240.1	aminotransferase [Chlamydia trachomatis D/UW-3/CX]
NP_220237.1	hypothetical protein CT_718 [Chlamydia trachomatis D/UW-3/CX]
NP_220231.1	type III secretion system effector protein [Chlamydia trachomatis D/UW-3/CX]
NP_220226.1	trigger factor [Chlamydia trachomatis D/UW-3/CX]
NP_220214.1	hypothetical protein CT_695 [Chlamydia trachomatis D/UW-3/CX]
NP_220213.1	hypothetical protein CT_694 [Chlamydia trachomatis D/UW-3/CX]
NP_220194.1	protein-arginine kinase [Chlamydia trachomatis D/UW-3/CX]
NP_220175.1	hypothetical protein CT_656 [Chlamydia trachomatis D/UW-3/CX]

NP_220160.1 hypothetical protein CT_642 [Chlamydia trachomatis D/UW-3/CX]

NP_220157.1 exodeoxyribonuclease V subunit beta [Chlamydia trachomatis D/UW-3/CX]

NP_220139.1 hypothetical protein CT_622 [Chlamydia trachomatis D/UW-3/CX]

NP_220138.1 hypothetical protein CT_621 [Chlamydia trachomatis D/UW-3/CX]

NP_220137.1 hypothetical protein CT_620 [Chlamydia trachomatis D/UW-3/CX]

NP_220135.1 hypothetical protein CT_618 [Chlamydia trachomatis D/UW-3/CX]

NP_220114.1 hypothetical protein CT_598 [Chlamydia trachomatis D/UW-3/CX]

NP_220094.1 type III secretion system translocator CopD

NP_220093.1 type III secretion system translocator CopB [Chlamydia trachomatis D/UW-3/CX]

NP_220091.1 type III secretion low calcium response chaperone [Chlamydia trachomatis D/UW-3/CX]

NP_220087.1 type II secretion system protein D [Chlamydia trachomatis D/UW-3/CX]

NP_220065.1 hypothetical protein CT_550 [Chlamydia trachomatis D/UW-3/CX]

NP_220044.1 hypothetical protein CT_529 [Chlamydia trachomatis D/UW-3/CX]

NP_219997.1 hypothetical protein CT_483 [Chlamydia trachomatis D/UW-3/CX]

NP_219983.1 DNA repair protein RecO [Chlamydia trachomatis D/UW-3/CX]

NP_219978.1 hypothetical protein CT_465 [Chlamydia trachomatis D/UW-3/CX]

NP_219969.1 translocated actin-recruiting phosphoprotein [Chlamydia trachomatis D/UW-3/CX]

NP_219965.1 cytidylate kinase [Chlamydia trachomatis D/UW-3/CX]

NP_219962.1 hypothetical protein CT_449 [Chlamydia trachomatis D/UW-3/CX]

NP_219954.1 hypothetical protein CT_442 [Chlamydia trachomatis D/UW-3/CX]

NP_219952.1 hypothetical protein CT_440 [Chlamydia trachomatis D/UW-3/CX]

NP_219945.1 hypothetical protein CT_433 [Chlamydia trachomatis D/UW-3/CX]

NP_219941.1 hypothetical protein CT_429 [Chlamydia trachomatis D/UW-3/CX]

NP_219924.1 outer membrane protein PmpC [Chlamydia trachomatis D/UW-3/CX]

NP_219923.1 outer membrane protein PmpB [Chlamydia trachomatis D/UW-3/CX]

NP_219905.1 stress response protein GrpE [Chlamydia trachomatis D/UW-3/CX]

NP_219894.1 hypothetical protein CT_384 [Chlamydia trachomatis D/UW-3/CX]

NP_219893.1 hypothetical protein CT_383 [Chlamydia trachomatis D/UW-3/CX]

NP_219882.1 pyruvoyl-dependent arginine decarboxylase [Chlamydia trachomatis D/UW-3/CX]

NP_219877.1 chorismate synthase [Chlamydia trachomatis D/UW-3/CX]

NP_219875.1 phosphoshikimate 1-carboxyvinyltransferase [Chlamydia trachomatis D/UW-3/CX]

NP_219874.1 hypothetical protein CT_365 [Chlamydia trachomatis D/UW-3/CX]

NP_219867.1 hypothetical protein CT_358 [Chlamydia trachomatis D/UW-3/CX]

NP_219866.1 hypothetical protein CT_357 [Chlamydia trachomatis D/UW-3/CX]

NP_219857.1 hypothetical protein CT_350 [Chlamydia trachomatis D/UW-3/CX]

NP_219852.1 hypothetical protein CT_345 [Chlamydia trachomatis D/UW-3/CX]

NP_219851.1 Lon protease [Chlamydia trachomatis D/UW-3/CX]

NP_219845.1 hypothetical protein CT_338 [Chlamydia trachomatis D/UW-3/CX]

NP_219844.1 phosphocarrier protein Hpr [Chlamydia trachomatis D/UW-3/CX]

NP_219841.1 DNA polymerase III subunit gamma and tau [Chlamydia trachomatis D/UW-3/CX]

NP_219832.1 hypothetical protein CT_326.1 [Chlamydia trachomatis D/UW-3/CX]

NP_219805.1 hypothetical protein CT_300 [Chlamydia trachomatis D/UW-3/CX]

NP_219793.1 hypothetical protein CT_288 [Chlamydia trachomatis D/UW-3/CX]

NP_219778.1 hypothetical protein CT_273 [Chlamydia trachomatis D/UW-3/CX]

NP_219738.1 inclusion membrane protein C [Chlamydia trachomatis D/UW-3/CX]

NP_219737.1 inclusion membrane protein B [Chlamydia trachomatis D/UW-3/CX]

NP_219734.1 inclusion membrane protein [Chlamydia trachomatis D/UW-3/CX]

NP_219733.1 inclusion membrane protein [Chlamydia trachomatis D/UW-3/CX]

NP_219728.1 inclusion membrane protein [Chlamydia trachomatis D/UW-3/CX]

NP_219718.1 inclusion membrane protein [Chlamydia trachomatis D/UW-3/CX]

NP_219684.1 ABC transporter ATP-binding protein [Chlamydia trachomatis D/UW-3/CX]

NP_219669.1 hypothetical protein CT_166 [Chlamydia trachomatis D/UW-3/CX]

NP_219668.1 hypothetical protein CT_165 [Chlamydia trachomatis D/UW-3/CX]

NP_219667.1 hypothetical protein CT_164 [Chlamydia trachomatis D/UW-3/CX]

NP_219664.1 hypothetical protein CT_161 [Chlamydia trachomatis D/UW-3/CX]

NP_219660.1 phospholipase [Chlamydia trachomatis D/UW-3/CX]

NP_219658.1 phosphatidylcholine-hydrolyzing phospholipase D [Chlamydia trachomatis D/UW-3/CX]

NP_219657.1 hypothetical protein CT_154 [Chlamydia trachomatis D/UW-3/CX]

NP_219650.1 hypothetical protein CT_147 [Chlamydia trachomatis D/UW-3/CX]

NP_219647.1 hypothetical protein CT_144 [Chlamydia trachomatis D/UW-3/CX]

NP_219646.1 hypothetical protein CT_143 [Chlamydia trachomatis D/UW-3/CX]

NP_219645.1 hypothetical protein CT_142 [Chlamydia trachomatis D/UW-3/CX]

NP_219644.1 preprotein translocase SecA [Chlamydia trachomatis D/UW-3/CX]

NP_219642.1 oligopeptide ABC transporter substrate-binding protein [Chlamydia trachomatis D/UW-3/CX]

NP_219641.1 microsomal dipeptidase [Chlamydia trachomatis D/UW-3/CX]

NP_219636.1 methyltransferase [Chlamydia trachomatis D/UW-3/CX]

NP_219622.1 inclusion membrane protein A [Chlamydia trachomatis D/UW-3/CX]

NP_219621.1 inclusion membrane protein G [Chlamydia trachomatis D/UW-3/CX]

NP_219608.1 hypothetical protein CT_105 [Chlamydia trachomatis D/UW-3/CX]

NP_219592.1 low calcium response protein E [Chlamydia trachomatis D/UW-3/CX]

NP_219590.1 4-alpha glucanotransferase [Chlamydia trachomatis D/UW-3/CX]

NP_219586.1 hypothetical protein CT_083 [Chlamydia trachomatis D/UW-3/CX]

NP_219585.1 hypothetical protein CT_082 [Chlamydia trachomatis D/UW-3/CX]

NP_219553.1 hypothetical protein CT_050 [Chlamydia trachomatis D/UW-3/CX]

NP_219527.1 signal recognition particle GTPase [Chlamydia trachomatis D/UW-3/CX]

NP_219517.1 ATPase [Chlamydia trachomatis D/UW-3/CX]

NP_219510.1 ribonuclease HII [Chlamydia trachomatis D/UW-3/CX]

NP_219502.1 hypothetical protein CT_875 [Chlamydia trachomatis D/UW-3/CX]

WP_010725097.1 phospholipase D endonuclease

WP_010725256.1	type II secretion system protein GspD
WP_010725170.1	hypothetical protein
WP_010725152.1	hypothetical protein
WP_010725105.1	hypothetical protein
WP_010725093.1	hypothetical protein
WP_010725088.1	oligopeptide-binding protein
WP_009871806.1	hypothetical protein
WP_009871718.1	hypothetical protein
WP_009871696.1	hypothetical protein
WP_009871580.1	inclusion membrane protein C
WP_009872659.1	hypothetical protein
WP_010725175.1	pyruvoyl-dependent arginine decarboxylase AaxB
WP_010725087.1	dipeptidase
WP_010725107.1	hypothetical protein
WP_010725106.1	hypothetical protein
WP_009872407.1	protein translocase
WP_009871988.1	hypothetical protein
WP_009871893.1	hypothetical protein
WP_009871845.1	hypothetical protein
WP_009871823.1	hypothetical protein
WP_009871735.1	hypothetical protein
WP_009871674.1	hypothetical protein
WP_009871560.1	hypothetical protein
WP_009871489.1	hypothetical protein
WP_009871488.1	hypothetical protein
WP_009872823.1	hypothetical protein
WP_009871914.1	hypothetical protein
WP_009871431.1	hypothetical protein
WP_010725376.1	hypothetical protein

WP_010725353.1	hypothetical protein
WP_010725337.1	hypothetical protein
WP_010725305.1	hypothetical protein
WP_010725287.1	hypothetical protein
WP_010725278.1	hypothetical protein
WP_010725277.1	hypothetical protein
WP_010725275.1	hypothetical protein
WP_010725169.1	hypothetical protein
WP_010725130.1	hypothetical protein
WP_010725103.1	hypothetical protein
WP_010725090.1	hypothetical protein
WP_010725058.1	hypothetical protein
WP_010725007.1	hypothetical protein
WP_010724957.1	hypothetical protein

Supplementary Table 2. GO terms for the predicted effectors

GO ID	Description
GO:0000338	protein deneddylation
GO:0004843	thiol-dependent ubiquitin-specific protease activity
GO:0005576	extracellular region
GO:0006508	proteolysis
GO:0008233	peptidase activity
GO:0008234	cysteine-type peptidase activity
GO:0009405	pathogenesis
GO:0016020	membrane
GO:0016021	integral component of membrane
GO:0016579	protein deubiquitination
GO:0016787	hydrolase activity
GO:0019784	NEDD8-specific protease activity
GO:0043657	host cell
GO:0042169	SH2 domain binding