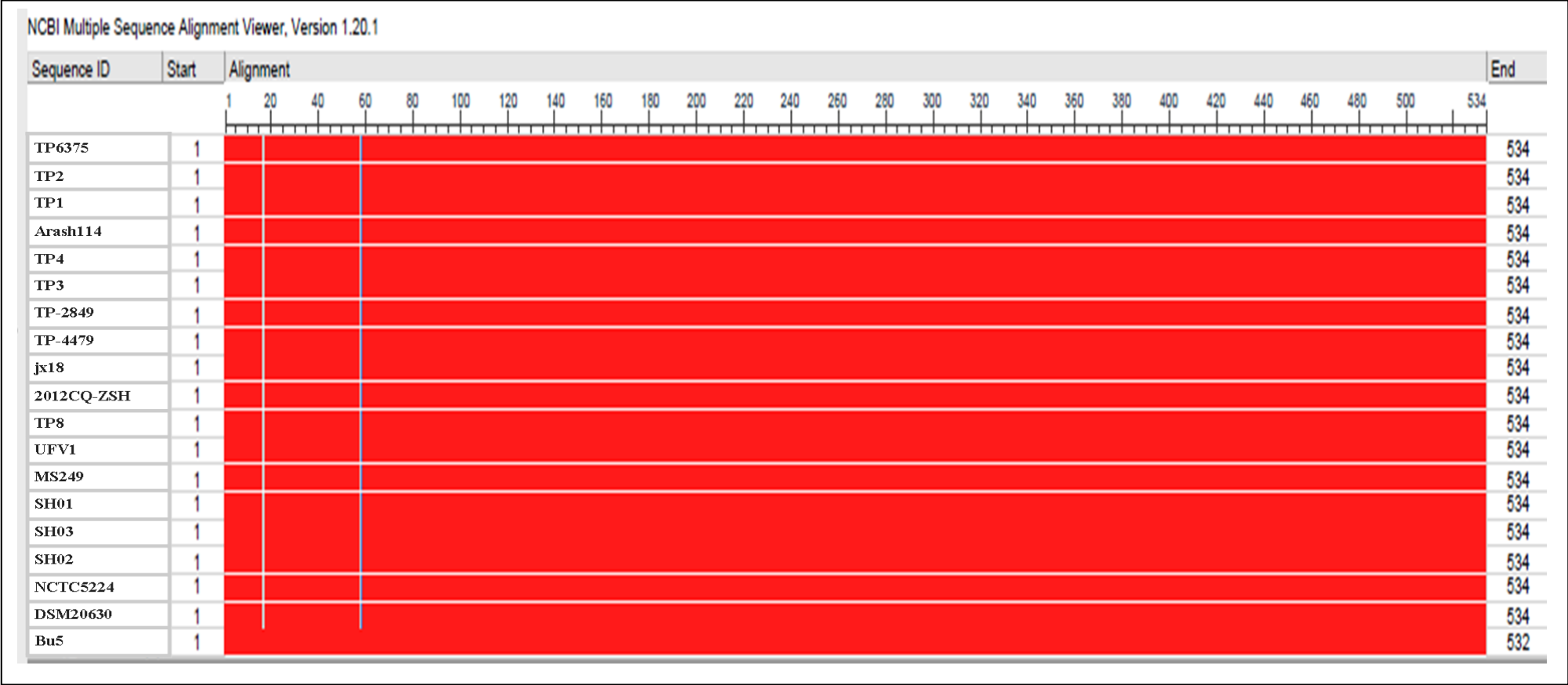
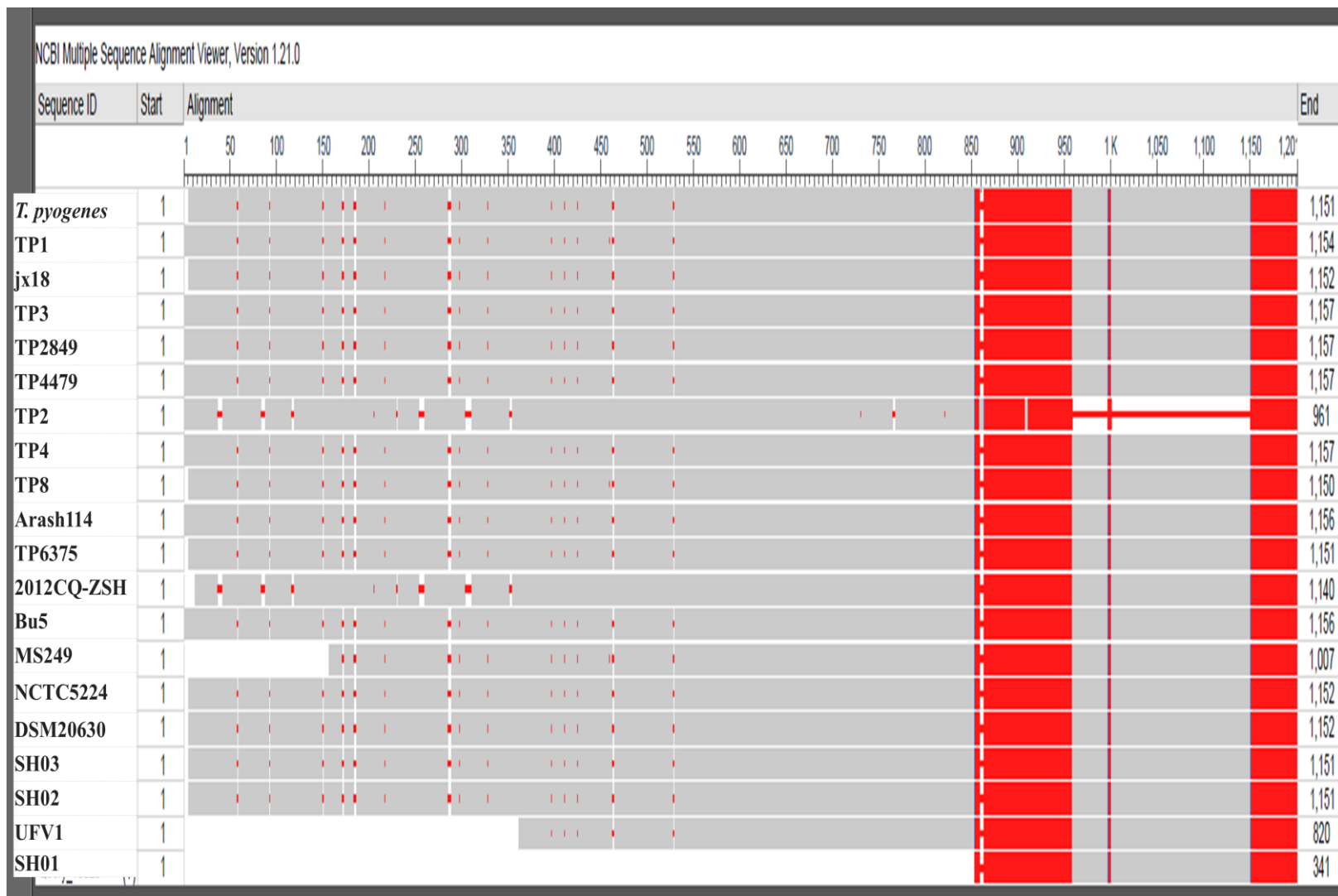


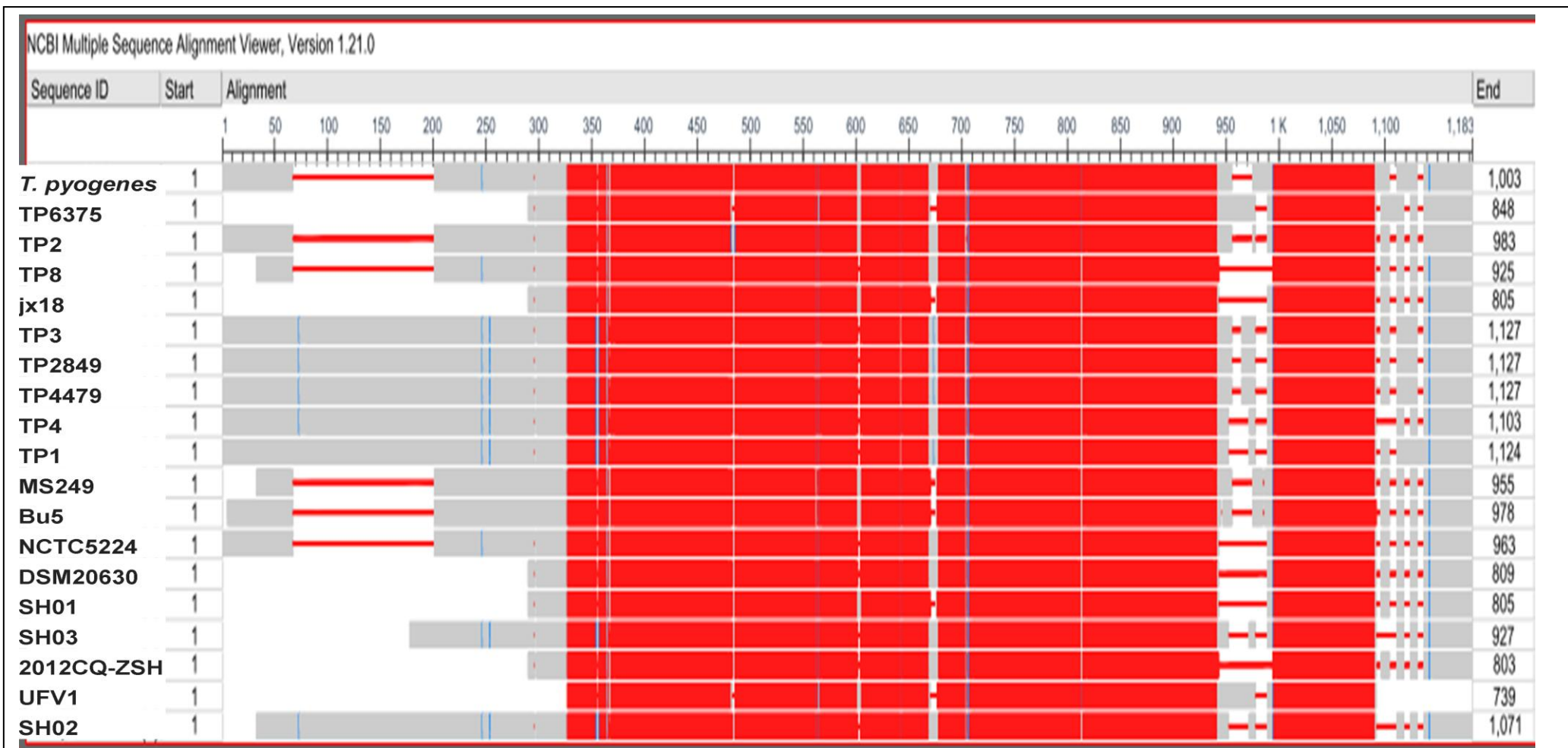
**Supplementary Figure 2 (a-h)** Multiple sequence alignment of crucial virulence factors such as pyolysin (plo), collagen-binding protein A (*cbpa*), neuraminidases (*nanH* and *nanP*), and fimbriae (*fimA*, *fimC*, *fimE*, and *fimJ*) which were identified in investigated *T.pyogenes* genomesby NCBI COBALT. Only alignment positions with no gaps will be colored. Red depicts highly conserved positions and blue depicts lower conservation. This method highlights highly conserved and less conserved amino acid positions on the basis of relative entropy



**(a) Multiple sequence alignment of pyolysin (plo) proteins identified in investigated *T. pyogenes* genomesby NCBI COBALT.**



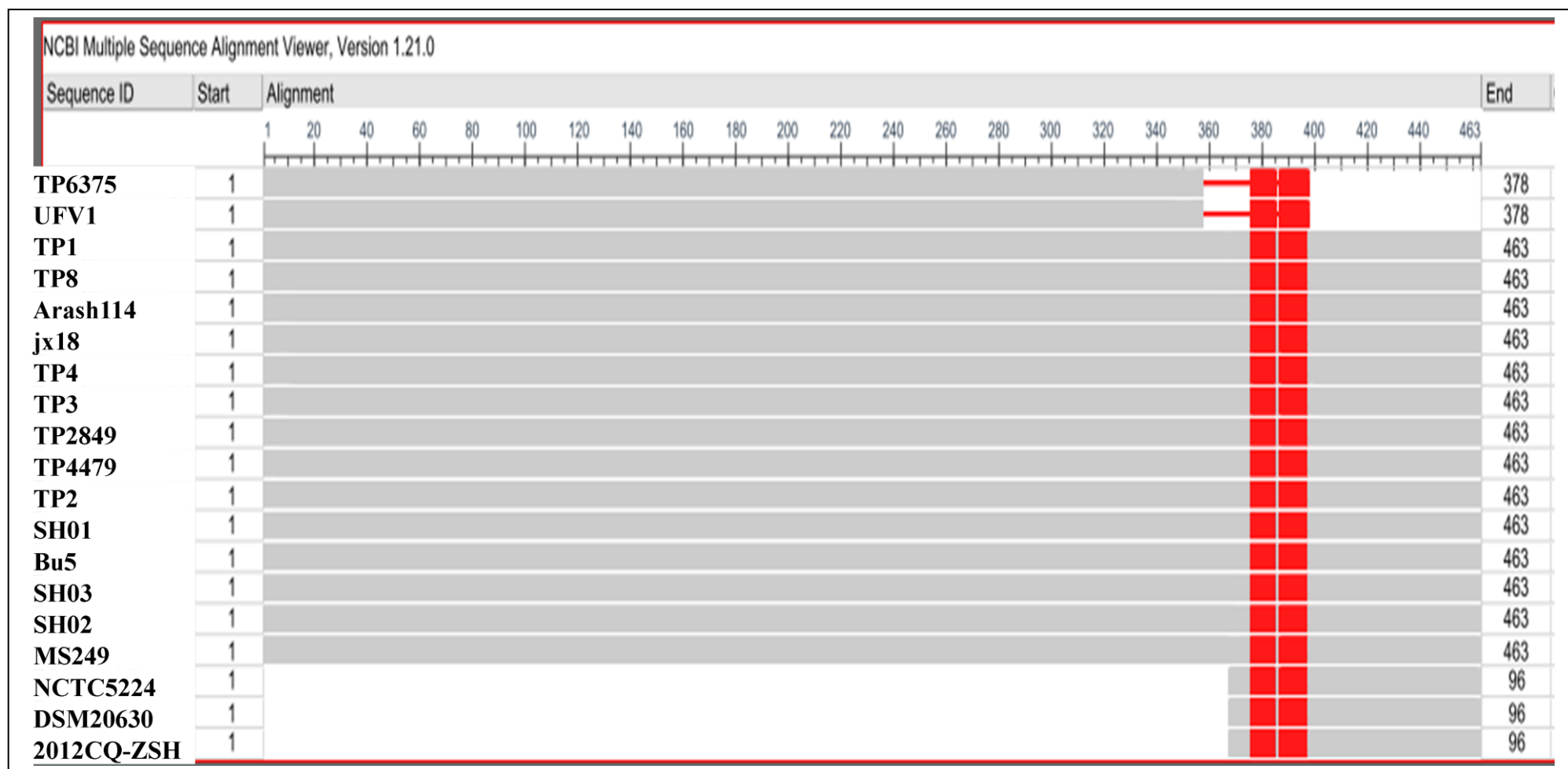
**(b) Multiple sequence alignment of collagen-binding protein A (*cbpa*) proteins identified in investigated *T. pyogenes* genomes by NCBI COBALT.**



(c) Multiple sequence alignment of neuraminidases H (*nanH*) proteins identified in investigated *T. pyogenes* genomes by NCBI COBALT.



(d) Multiple sequence alignment of neuraminidases P (*nanP*) proteins identified in the investigated *T. pyogenes* genomes by NCBI COBALT.



(e) Multiple sequence alignment of *fimA* proteins identified in the investigated *T. pyogenes* genomes by NCBI COBALT.



(f) Multiple sequence alignment of *fimC* proteins identified in the investigated *T. pyogenes* genomes by NCBI COBALT.

NCBI Multiple Sequence Alignment Viewer, Version 1.20.1

Sequence ID	Start	Alignment	End
		1 50 100 150 200 250 300 350 400 450 500 550 580	
TP6375	1		583
Arash114	1		583
TP3	1		584
TP2849	1		584
TP4479	1		584
TP2	1		583
TP1	1		583
TP8	1		584
2012CQ-ZSH	1		584
jx18	1		581
TP4	1		581
UFV1	1		583
SH03	1		583
SH01	1		584
SH02	1		584
Bu5	1		585
MS249	1		530
NCTC5224	1		584
DSM20630	1		584

(g) Multiple sequence alignment of *fimE* proteins identified in the investigated *T. pyogenes* genomes by NCBI COBALT.

NCBI Multiple Sequence Alignment Viewer, Version 1.20.1

Sequence ID	Start	Alignment	End
		1 50 100 150 200 250 300 350 400 450 500 550 600 647	
TP6375	1		600
Arash114	1		612
jx18	1		545
TP4	1		558
TP3	1		558
TP2849	1		558
TP4479	1		558
2012CQ_ZSH	1		542
TP2	1		552
TP8	1		539
TP1	1		555
SH01	1		542
NCTC5224	1		540
DSM20630	1		540
SH03	1		582
SH02	1		582
MS249	1		539

**(h) Multiple sequence alignment of *fimJ* proteins identified in the investigated *T. pyogenes* genomes by NCBI COBALT.**