

Journal of Applied Biology & Biotechnology

Available online at: <https://jabonline.in/>

Expression and characterization of global consensus nonstructural protein 1 (NS1) of four dengue virus serotypes in *Escherichia coli*

Nguyen Le Cu¹, Luong Ngoc Nguyen², Chung Van Huynh¹, Tuan Van Le³, Anh Nguyen Tuan Tran¹, Dat Tuan Dang¹

¹Buon Ma Thuot Medical University, Buon Ma Thuot, Vietnam.

²College of Sciences, Hue University, Hue, Vietnam.

³Tay Nguyen Institute of Hygiene and Epidemiology, Tan An, Vietnam.

doi: <https://doi.org/10.7324/JABB.2024.188233>

SUPPLEMENTARY MATERIAL

Supplementary table 1: A summary of the global consensus Non-structural protein 1 of four Dengue serotypes.

Serotype	Number of database sequences used to create the consensus	Accession number of the protein sequences corresponding to the consensus	Number of sequences with sequence identity >99%/Total number of sequences found by BLASTp	Accession number of the encoding genes
DENV1	1726	AYP74685.1	1747/3846	OR349256
DENV2	1415	AFK65750.1	1062/3601	OR349258
DENV3	809	AHG23254.1	603/1783	OR349260
DENV4	212	AFX65870.1	56/809	OR349262

Supplementary table 2: Linear epitopes of NS1 for DENV1 to DENV4 validated using the Immune Epitope Database (IEDB) at <https://www.iedb.org>, aligned with the consensus NS1 sequences. Any mismatches between the epitopes and the consensus sequences are highlighted in bold and underscored.

Protein	Linear epitope sequences (starting position - ending position on polyprotein)	Percentage identity to the consensus NS1 sequences
DENV1-NS1	QYKFQADSPK <u>K</u> LSAA (806-820)	93%
	SPKRLSAAI (813-820)	100%
	QISNELN <u>Y</u> ILLENDM (845-859)	93%
	KKMIRPQPMEHKYSWKSWSGKAKIIGAD <u>I</u> QN (876-905)	97%
	KIIGAD <u>I</u> QNTT <u>F</u> IID (897-911)	93%
	YGGPISQHNY (1022-1031)	100%
DENV2-NS1	QYKFQPESPSKLASA (806-820)	100%
	EGICGIRSVTRLENL (827-841)	100%
	CGIRSVTRLENLMWK (830-844)	100%
	DYGFVFTTNIWLKL (932-946)	100%
	YGFVFTTNIWLKLR (933-947)	100%
	GFG <u>F</u> FTTNIWLKLRE (934-948)	93%
	GFGVFTTNIWLKLKE (934-948)	100%
	GFGVFTTNIWLKLRE (934-948)	100%
	FGVFTTNIWLKLREK (935-949)	100%
	GVFTTNIWLKLKEKQ (936-950)	100%
	NRAVHADMGYWIESA (966-980)	100%
HADMGYWIESALNDT (970-984)	100%	
	LNDTWK <u>I</u> EKASFIEV (981-995)	93%
DENV3-NS1	IFVTNEVHTWTEQYKFQA (792-809)	100%
	TEQYKFQADSPKRLA (802-816)	100%
	QYKFQADSPKRLATA (804-818)	100%

	CGIRSTTRMENLLWKQIA (828-845)	100%
	QIANELNYILWENNIK (843-858)	100%
	ANELNYILWENNIK (843-859)	100%
	IKLTVVVGDI (857-866)	100%
	GDIIGVLEQGKRTLTPQPMELK (864-885)	100%
	GKAKIVTAETQNSSFIID (892-909)	100%
	GPNTPECPSASRAWNVWE (910-927)	100%
	NFLEVEDYGF (925-934)	100%
	VEDYGFVFTTNIWLKLREVYT (928-949)	100%
	FGVFTTNIWLKLR <u>D</u> V (933-947)	93%
	FGVFTTNIWLKLREV (933-947)	100%
	FG <u>M</u> FTTNIW <u>M</u> KFREG (934-948)	79%
	HADMGYWIESQKNGS (968-982)	100%
	HADMGYWIESSKN <u>Q</u> T (969-983)	92%
DENV4-NS1	N <u>F</u> LEVEDYGF (925-934)	90%
	FGMFTTNIWMKFREG (934-948)	100%
	HADMGYWIESSKNQ <u>T</u> (969-983)	100%

Supplementary Figure 1.

NC

PC

