

Matrix Protein 2

10 20 30 40 50 60 70 80 90
A/PR/08/1934 M S L L T E V E T P I R N E W G C R C N G S S D P L A I A A N I G I L H L I L W I L D R L F F K C I Y R R F K Y G L K G P S T E G V P K S M R E E Y R K E Q Q S A V D A D D G H
A/CA/04/2009 M S L L T E V E T P T R S E W E C R C S D S S D P L V I A A N I G I L H L I L W I T D R L F F K C I Y R R F K Y G L K R G P S T E G V P F E S M R E E Y Q Q E Q Q S A V D V D D G H

A/PR/08/1934 F V S I E L E
A/CA/04/2009 F V N I E L E

Non-Structural Protein 1

10 20 30 40 50 60 70 80 90
A/PR/08/1934 M D P N T V S S F Q V D C F L W H V R K R V A D Q E L G D A P F L D R L R R D Q K S L R G R G S T L G L D I E T A T R A G K Q I V E R I L K E E S D E A L K M T M A S V P A S R Y L
A/CA/04/2009 M D S N T M S S F Q V D C F L W H I R K R F A D N G L G D A P F L D R L R R D Q K S L K G R G N T L G L D I E T A T L V G K Q I V E W I L K E E S S E T L R M T I A S V P T S R Y L

100 110 120 130 140 150 160 170 180
A/PR/08/1934 T D M T L E E M S R E W S M L I P K Q K V A G P L C I R M D Q A I M D K N I I L K A N F S V I F D R L E T L I L L R A F T E E G A I V G E I S P L P S L P G H T A E D V K N A V G V
A/CA/04/2009 S D M T L E E M S R D W F M L M P R Q K I I G P L Q V R L D Q A I M E K N I V L K A N F S V I F N R L E T L I L L R A F T E E G A I V G E I S P L P S L P G H T Y E D V K N A V G V

190 200 210 220 230
A/PR/08/1934 L I G G L E W N D N T V R V S E T L Q R F A W R S S N E N G R P P L T F K Q K R E M A G T I R S E V
A/CA/04/2009 L I G G L E W N G N T V R V S E N I Q R F A W R N C D E N G R R S L P F E Q K - - - - -

Non-Structural Protein 2

10 20 30 40 50 60 70 80 90
A/PR/08/1934 M D P N T V S S F Q D I L L R M S K M Q L E S S S E D L N G M I T Q F E S L K I Y R D S L G E A V M R M G D L H S L Q N R N E K W R E Q L G Q K F E E I R W L I E E V R H K L K V T
A/CA/04/2009 M D S N T M S S F Q D I L M R M S K M Q L G S S S E D L N G M V T R F E S L K I Y R D S L G E T V M R M G D L H Y L Q S R N E K W R E Q L G Q K F E E I R W L I E E M R H R L K A T

100 110 120
A/PR/08/1934 E N S F E Q I T F M Q A L H L L L E V E Q E I R T F S F Q L I
A/CA/04/2009 E N S F E Q I T F M Q A L Q L L L E V E Q E I R A F S F Q L I

Polymerase Protein PA

10 20 30 40 50 60 70 80 90
A/PR/08/1934 M E D F V R Q C F N P M I V E L A E K T M K E Y G E D L K I E T N K F A A I C T H L E V C F M Y S D F H F N E Q G E S I I V E L G D P N A L L K H R F E I E G R D R T M A W T V
A/CA/04/2009 M E D F V R Q C F N P M I V E L A E K A M K E Y G E D P K I E T N K F A A I C T H L E V C F M Y S D F H F D E R G E S I I V E S G D P N A L L K H R F E I E G R D R I M A W T V

100 110 120 130 140 150 160 170 180
A/PR/08/1934 V N S I C N T T G A E K P K F L P D L Y D Y K E N R F I E I G V T R R E V H I Y Y L E K A N K I K S E K T H I H I F S F T G E E M A T K A D Y T L D E E S R A R I K T R L F T I R Q
A/CA/04/2009 V N S I C N T T G V E K P K F L P D L Y D Y K E N R F I E I G V T R R E V H I Y Y L E K A N K I K S E K T H I H I F S F T G E E M A T K A D Y T L D E E S R A R I K T R L F T I R Q

190 200 210 220 230 240 250 260 270
A/PR/08/1934 E M A S R G L W D S F R Q S E R G E E T I E E R F E I T G T M R K L A D Q S L P P N F S S L E N F R A Y V D G F E P N G Y I E G K L S Q M S K E V N A R I E P F L K T T P R P L R L
A/CA/04/2009 E M A S R S L W D S F R Q S E R G E E T I E E K F E I T G T M R K L A D Q S L P P N F P S L E N F R A Y V D G F E P N G C I E G K L S Q M S K E V N A K I E P F L R T T P R P L R L

280 290 300 310 320 330 340 350 360
A/PR/08/1934 P N G P P C S Q R S K F L L M D A L K L S I E D P S H E G E G I P L Y D A I K C M R T F F G W K E P N V V K P H E K G I N P N Y L S W K Q V L A E L Q D I E N E E K I P K T K N M
A/CA/04/2009 P D G R L C H Q R S K F L L M D A L K L S I E D P S H E G E G I P L Y D A I K C M K T F F G W K E P N I V K P H E K G I N P N Y L M A W K Q V L A E L Q D I E N E E K I P R T K N M

370 380 390 400 410 420 430 440 450
A/PR/08/1934 K K T S Q L K W A L G E N M A P E K V D F D D C K D V G D L K Q Y D S D E P E L R S L A S W I Q N E F N K A C E L T D S S W I E L D E I G E D V A P I E H I A S M R R N Y F T S E V
A/CA/04/2009 K R T S Q L K W A L G E N M A P E K V D F D D C K D V G D L K Q Y D S D E P E P R S L A S W V Q N E F N K A C E L T D S S W I E L D E I G E D V A P I E H I A S M R R N Y F T A E V

460 470 480 490 500 510 520 530 540
A/PR/08/1934 S H C R A T E Y I M K G V Y I N T A L L N A S C A A M D D F L I P M I S K R C T K E G R R K T N L Y G F I I K G R S H L R N D T D V V N F V S M E F S L T D P R L E P H K W E K Y
A/CA/04/2009 S H C R A T E Y I M K G V Y I N T A L L N A S C A A M D D F L I P M I S K R C T K E G R R K T N L Y G F I I K G R S H L R N D T D V V N F V S M E F S L T D P R L E P H K W E K Y

550 560 570 580 590 600 610 620 630
A/PR/08/1934 C V L E I G D M L L R S A I G Q V S R P M F L Y V R T N G T S K I K M K W G M E M R R C L L Q S L Q Q I E S M I E A E S S V K E K D M T K E F F E N K S E T W P I G E S P K G V E E
A/CA/04/2009 C V L E I G D M L L R T A I G Q V S R P M F L Y V R T N G T S K I K M K W G M E M R R C L L Q S L Q Q I E S M I E A E S S V K E K D M T K E F F E N K S E T W P I G E S P R G V E E

640 650 660 670 680 690 700 710
A/PR/08/1934 S S I G K V C R T L L A K S V F N S L Y A S P Q L E G F S A E S R K L L I V Q A L R D N L E P G T F D L G G L Y E A I E E C L I N D P W V L L N A S W F N S F L T H A L S
A/CA/04/2009 G S I G K V C R T L L A K S V F N S L Y A S P Q L E G F S A E S R K L L I V Q A L R D N L E P G T F D L G G L Y E A I E E C L I N D P W V L L N A S W F N S F L T H A L K

Polymerase Protein PB1

10 20 30 40 50 60 70 80 90
AFPR08/1934 MDVNP TLLFLKVP AQNAISTTFFPYTGDPYPYSHGTGTGYTMDTVNRTHQYSEKARWTTNTTETGAPQLNPIDGGLPEDNEPSSGYAQTDCVLE
ACA04/2009 MDVNP TLLFLKVP AQNAISTTFFPYTGDPYPYSHGTGTGYTMDTVNRTHQYSEKARWTTNTTETGAPQLNPIDGGLPEDNEPSSGYAQTDCVLE

100 110 120 130 140 150 160 170 180
AFPR08/1934 AMAFLEESHPIGFENSCLEETMEVQQTRVDKLTQGRQTYDWTLNRNQPAATALANTIEVFRSNGLTANESGRLIDFLKDVMEISMKKEEMG
ACA04/2009 AMAFLEESHPIGFENSCLEETMEVQQTRVDKLTQGRQTYDWTLNRNQPAATALANTIEVFRSNGLTANESGRLIDFLKDVMEISMNKEELE

190 200 210 220 230 240 250 260 270
AFPR08/1934 ITTHFQRKRRVRDNMTKKMVTQRTIGKRRKQRLNKRSYLIRALTLNMTKDAERGKLRRAIATPGMQIRGFVYFVETLARSICEKLEQSG
ACA04/2009 ITTHFQRKRRVRDNMTKKMVTQRTIGKRRKQRLNKRGYLIRALTLNMTKDAERGKLRRAIATPGMQIRGFVYFVETLARSICEKLEQSG

280 290 300 310 320 330 340 350 360
AFPR08/1934 LPVGGNEKKAKLANVVRKMMTNSQDTELSLTI TGDNTKWNENQNPRLMFLAMITYMTRNQPEWFRNVLSTAPIMFSNKMARLGKGYMFESK
ACA04/2009 LPVGGNEKKAKLANVVRKMMTNSQDTELSLFTITGDNTKWNENQNPRLMFLAMITYMTRNQPEWFRNVLSTAPIMFSNKMARLGKGYMFESK

370 380 390 400 410 420 430 440 450
AFPR08/1934 SMKLRITQIPAEMLASIDLKYFNDESTKKEIEKIRPLLIDGTASLSPGMMMGFMNMLSTVLGVSILNLGQKRYTKTIYWWDGLQSSDDFALI
ACA04/2009 RMKLRITQIPAEMLASIDLKYFNDESTKKEIEKIRPLLIDGTASLSPGMMMGFMNMLSTVLGVSILNLGQKRYTKTIYWWDGLQSSDDFALI

460 470 480 490 500 510 520 530 540
AFPR08/1934 VNAPNHEG IQAGVDRFYRTCKLHGINMSKKKSYINRGTGTFEFTSFFYRYGFVANFSMELPSFGVSGSNESADMSIGVTVIKNNMINNDLG
ACA04/2009 VNAPNHEG IQAGVDRFYRTCKLVGINMSKKKSYINKGTGTFEFTSFFYRYGFVANFSMELPSFGVSGVNESADMSIGVTVIKNNMINNDLG

550 560 570 580 590 600 610 620 630
AFPR08/1934 PATAQMALQLFIKDYRYTYRCHRGDTQIQTRRSFEIKKLWEDTRSKAGLLVSDGGPNLYNIRNLHIEVCLKWEMLMDEYDQGRRLCNPLNP
ACA04/2009 PATAQMALQLFIKDYRYTYRCHRGDTQIQTRRSFEIKKLWEDTQSKVGLLVSDGGPNLYNIRNLHIEVCLKWEMLMDDYRGRRLCNPLNP

640 650 660 670 680 690 700 710 720
AFPR08/1934 FVSHKEI ESMNNAVMMPAHGPAKNMEYDAVATTHSWIPKRNRSILNTSQRQVLEDEQMYQRCCNLFKFKFPSSSYRRPVGISSMVEAMVS
ACA04/2009 FVSHKEIDSMNNAVMMPAHGPAKSMYDAVATTHSWIPKRNRSILNTSQRQVLEDEQMYQRCCNLFKFKFPSSSYRRPVGISSMVEAMVS

730 740 750
AFPR08/1934 RARIDARIDFESGRIKKEEFTEIMKICSTIEELRRQK
ACA04/2009 RARIDARVDFESGRIKKEEFSEIMKICSTIEELRRQK

Polymerase Protein PB2

10 20 30 40 50 60 70 80 90
AFPR08/1934 MERIKELRNLMSSQSRTRILTCTVDHMAI IKKYTSGRQEKNPALRMKWMAMKYPITADKRIEMIPERNEQQGTLWSKMNDAGSDRVM
ACA04/2009 MERIKELRDLMSQSRTRILTCTVDHMAI IKKYTSGRQEKNPALRMKWMAMRYPITADKRIEMIPERNEQQGTLWSKNTDAGSDRVM

100 110 120 130 140 150 160 170 180
AFPR08/1934 VSPLAVTWNRRNGPMTNTVHYPKIYKTYFERVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVFPNEVGARILTSE
ACA04/2009 VSPLAVTWNRRNGPTTSTVHYPKVYKTYFEKVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVFPNEVGARILTSE

190 200 210 220 230 240 250 260 270
AFPR08/1934 SQLTITKEKKEELQDCKISPLMVAYMLERELVRKTRFLPVAGGTSVYIEVLHLTQGTQWQMYTPGGEVKNDVDQSLIAARNIVRRA
ACA04/2009 SQLAITKEKKEELQDCKIAPLMVAYMLERELVRKTRFLPVAGGTGSVYIEVLHLTQGTQWQMYTPGGEVKNDVDQSLIAARNIVRRA

280 290 300 310 320 330 340 350 360
AFPR08/1934 AVSADPLASLLEMCHSTQIGGIRMVDI LKQNPTEEQAVGICKAAMGLRISSSSFGGFTFKRTSGSSVKKEEEVLTGNLQTLKIRVHEGY
ACA04/2009 AVSADPLASLLEMCHSTQIGGIRMVDI LRQNPTEEQAVDICKAAIGLRISSSSFGGFTFKRTSGSSVKKEEEVLTGNLQTLKIRVHEGY

370 380 390 400 410 420 430 440 450
AFPR08/1934 EEF TMVGRRA TAILRKATRRLIQLIVSGRDEQSIAEAIIVAMVFSQEDCMIKAVRGDLNFVNRRANQRLNPMHQLLRHFQKDAKVLFNQWVG
ACA04/2009 EEF TMVGRRA TAILRKATRRLIQLIVSGRDEQSIAEAIIVAMVFSQEDCMIKAVRGDLNFVNRRANQRLNPMHQLLRHFQKDAKVLFNQWVG

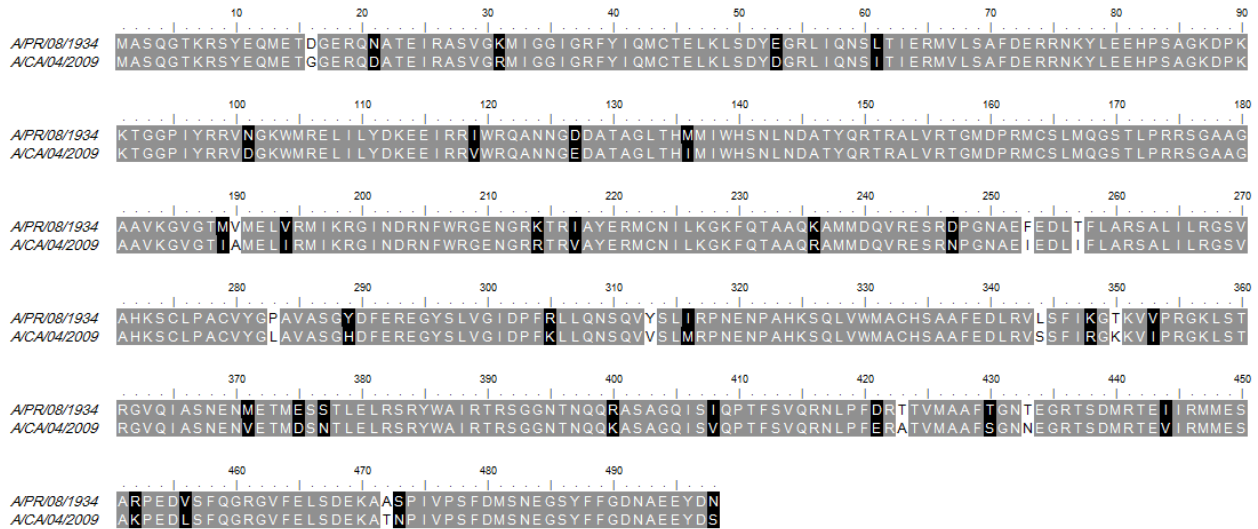
460 470 480 490 500 510 520 530 540
AFPR08/1934 VEPIDNVNMGMI GILPDMTPSI EMSMRGVRI SKMGVDEYSSTERVVVSDRFLVRVDRQGNVLLSPEEVSETQGTTEKLTITYSSMMWEIN
ACA04/2009 IESIDNVNMGMI GILPDMTPSI EMSLRGIRVSKMGVDEYSSTERVVVSDRFLVRVDRQGNVLLSPEEVSETQGTTEKLTITYSSMMWEIN

550 560 570 580 590 600 610 620 630
AFPR08/1934 GPESVLVNTYQWIRNWEITVKIQWSQNP TMLYNKMEFEPFQSLVPKAI RGGYSGFVRTLFQQMRDVLGTFDTAQIKLLPFAAAPPKQSR
ACA04/2009 GPESVLVNTYQWIRNWEITVKIQWSQNP TMLYNKMEFEPFQSLVPKAI TRSRYSGFVRTLFQQMRDVLGTFDTVQIKLLPFAAAPPEQSR

640 650 660 670 680 690 700 710 720
AFPR08/1934 MQFSSFTVNVVRSQMRILVVRGNSPVFNYNKATKRLTVLGKDAQTLTEDPDEGTASVESAVLRGFLILGKEDRRYGPALSINELSNLAKGE
ACA04/2009 MQFSSLTVNVVRSGLRILVVRGNSPVFNYNKATKRLTVLGKDAQALTLEDPDEGTSVESAVLRGFLILGKEDKRYGPALSINELSNLAKGE

730 740 750
AFPR08/1934 KANVLI GGGDVVLMKRRDSSILTDSQTATKRIRMAIN
ACA04/2009 KANVLI GGGDVVLMKRRDSSILTDSQTATKRIRMAIN

Nucleoprotein



Viral Proteins	Percent Homology
Hemagglutinin	81.42
Neuraminidase	83.48
Matrix protein 1	94.84
Matrix protein 2	86.6
Non-structural protein 1	82.65
Non-structural protein 2	87.6
Polymerase protein PA	95.95
Polymerase protein PB1	94.85
Polymerase protein PB2	95.13
Nucleoprotein	91.57

Supplemental Figure 1. Protein sequences for different viral proteins were retrieved from NCBI for A/California/04/2009 and A/Puerto Rico/8/1934 H1N1 virus isolates. The sequence alignment was performed by ClustalW using default parameters and edited using BioEdit. Identical and similar amino acid residues are represented in gray and black background respectively. The percent identity was calculated using Clustal Omega. Matrix used for alignment by ClustalW: BLOSUM62