



UNIVERSITÀ DEGLI STUDI DI MILANO - BICOCCA  
Scuola di Scienze  
Dipartimento di Informatica, Sistemistica e Comunicazione  
Corso di Laurea Magistrale in Informatica

## Multiple Sequence Alignment (MSA) di sequenze SARS-CoV2

Alberici Federico - 808058  
Bettini Ivo Junior - 806878  
Traversa Silvia - 816435

Anno Accademico 2019 - 2020

# Indice

<b>SARS-CoV-2</b> . . . . .	<b>2</b>
<b>Obiettivo del progetto</b> . . . . .	<b>3</b>
<b>Sequenze</b> . . . . .	<b>3</b>
<b>Tool</b> . . . . .	<b>3</b>
<b>Analisi verticale</b> . . . . .	<b>4</b>
New York (USA) . . . . .	4
Russia . . . . .	4
Italia . . . . .	4
Spagna . . . . .	4
<b>Analisi verticale</b> . . . . .	<b>4</b>
<b>Appendice A - risultati confronto verticale</b> . . . . .	<b>6</b>
New York (USA) . . . . .	7
Clustal Omega . . . . .	7
Kalign . . . . .	10
MAFFT . . . . .	13
Russia . . . . .	15
Clustal Omega . . . . .	15
Kalign . . . . .	18
MAFFT . . . . .	20
Italia . . . . .	22
Clustal Omega . . . . .	22
Kalign . . . . .	24
MAFFT . . . . .	26
Spagna . . . . .	28
Clustal Omega . . . . .	28
Kalign . . . . .	32
MAFFT . . . . .	37
<b>Appendice B - risultati confronto orizzontale</b> . . . . .	<b>42</b>
Clustal Omega . . . . .	42
Kalign . . . . .	44
MAFFT . . . . .	46

## SARS-CoV-2

I coronavirus sono una famiglia di virus RNA a filamento positivo con aspetto simile a una corona se esaminati al microscopio elettronico (da cui il nome) e noti per causare malattie che vanno dal comune raffreddore a patologie più gravi come la Sindrome respiratoria mediorientale (MERS) o la Sindrome respiratoria acuta grave (SARS).

Nel Dicembre 2019 a Wuhan, in Cina, è stato isolato un nuovo ceppo di coronavirus, denominato SARS-CoV-2 (*Severe Acute Respiratory Syndrome - Coronavirus - 2*), la cui sequenza virale ha un'omologia del 76% con il virus che causò la pandemia della Sars nel 2002. La malattia respiratoria causata da questo nuovo coronavirus è stata chiamata COVID-19.

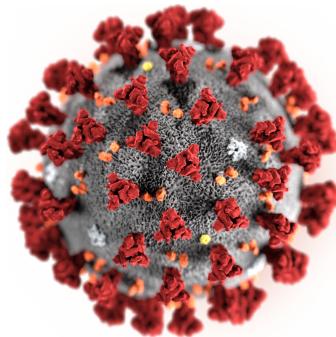


Figure 1: *L'illustrazione, creata dai Centers for Disease Control and Prevention (CDC) statunitense, rivela la morfologia ultrastrutturale del SARS-CoV-2*

Ogni virione (singola particella virale) SARS-CoV-2 ha un diametro di circa 50-200 nanometri. Come altri coronavirus, SARS-CoV-2 presenta quattro proteine strutturali: S (spike), E (inviluppo), M (membrana) e N (nucleocapside). La proteina N contiene il genoma dell'RNA mentre le proteine S, E, M creano insieme l'inviluppo virale. In particolare la proteina spike è quella che permette al virus di attaccarsi alla membrana di una cellula ospite.

Fra le sequenze genomiche SARS-CoV-2 note è stata evidenziata una bassa variabilità, si ritiene dunque che il ceppo sia stato rilevato dalle autorità sanitarie entro poche settimane dalla sua comparsa tra la popolazione umana alla fine del 2019. Dopo il primo caso di infezione noto, avvenuto in Cina verso la fine del 2019, il virus si è diffuso in tutte le province della Cina e in oltre 150 altri paesi in Asia, Europa, Nord America, Sud America, Africa e Oceania. La trasmissione da uomo a uomo di SARS-CoV-2 è stata confermata nel gennaio 2020 e avviene principalmente attraverso goccioline respiratorie da tosse e starnuti entro un raggio di circa 1,8 metri oppure un'altra possibile causa di infezione è il contatto indiretto tramite superfici contaminate.

Il numero di riproduzione di base del virus è stato stimato tra 1.4 e 3.9, il che significa che ogni infezione dal virus dovrebbe causare da 1.4 a 3.9 nuove infezioni nel caso in cui nessun membro della comunità sia immune e non vengano prese misure preventive.

Al 25 maggio 2020, ci sono stati 5.428.605 casi confermati totali di infezione da SARS-CoV-2 nella pandemia in corso, con un totale di decessi attribuiti al virus di 345.375.

## Obiettivo del progetto

L'obiettivo del nostro progetto è riuscire ad allineare le sequenze scaricate (compresa quella che utilizziamo come riferimento per i confronti) con i tool scelti, e produrre in output le variazioni delle sequenze scaricate rispetto alla sequenza di riferimento e creare un'apposita documentazione.

Per poter raggiungere questo obiettivo, abbiamo deciso di dividere la nostra analisi in due parti:

- nella prima parte eseguiamo un'analisi verticale fra la sequenza di riferimento (ossia la prima sequenza che presente nella banca dati raccolta a Wuhan) e delle sequenze prese rispettivamente da Italia, New York, Spagna e Russia. Le sequenze dei singoli paesi sono state prese seguendo un ordine temporale crescente, in modo tale da poter denotare attraverso il confronto delle mutazioni nel tempo.
- nella seconda parte, invece, effettuiamo un confronto orizzontale tra le ultime sequenze raccolte dei vari paesi scelti, in modo da poter individuare la presenza di eventuali mutazioni nello stato corrente del virus, sempre tenendo come riferimento la prima sequenza presente nella banca dati.

## Sequenze

Abbiamo deciso di selezionare alcuni dei paesi con il più alto numero di casi di COVID-19, ossia Stati Uniti (per il quale ci siamo concentrati su uno degli stati a sua volta più colpito, quello di New York), Russia, Spagna e Italia. I casi, al 25 Maggio 2020, sono riportati nella tabella seguente:

<b>Stati Uniti</b>	1,678.477
<b>Russia</b>	353,427
<b>Spagna</b>	235,823
<b>Italia</b>	229,858

Le sequenze analizzate sono state prese dal sito GISAID. Per ogni paese abbiamo preso solo sequenze complete (con più di 29000 basi) e su pazienti tutti di sesso maschile. Abbiamo cercato inoltre di seguire una linea temporale che partisse dalla prima sequenza presente nel database fino all'ultima, con in mezzo sequenze che si distanziavano tra di loro di una settimana.

## Tool

Tra i tool messi a disposizione abbiamo scelto di utilizzare Clustal Omega, Kalign e MAFFT. Abbiamo scelto Clustal Omega poichè sfrutta il modello probabilistico Profilo HMM (Hidden Markov Model), il quale è in grado di incapsulare i cambiamenti evolutivi che si sono verificati in una serie di sequenze correlate ed inoltre questo tool risultava ottimale per sequenze molto lunghe.

Kalign è stato scelto poichè è un tool che si concentra sulle regioni locali e lavora molto velocemente, motivo per il quale ci sembrava particolarmente adatto per le analisi sui singoli stati che volevamo svolgere.

Infine abbiamo scelto MAFFT poichè attraverso l'algoritmo della trasformata di Fourier veloce

ottimizza gli allineamenti in base alle proprietà fisiche, utilizzando sia un allineamento progressivo che iterativo. Anche esso risulta essere ottimale per sequenze di lunghezza medio-grandi.

## **Analisi verticale**

Nell' analisi verticale abbiamo deciso di studiare singolarmente ogni stato, prendendo le sequenze scelte e confrontandole con quella di riferimento. I risultati ottenuti sono visibile nell'appendice A della relazione.

### **New York (USA)**

Nell'eseguire i tre tool non abbiamo ottenuto risultati differenti nonostante i tre approcci differenti di calcolo.

### **Russia**

La Russia, come New York, nonostante i diversi metodi di calcolo non presenta differenze negli output dei tre tool.

### **Italia**

L'esecuzione dei tre tool con le sequenze dell' Italia ha prodotti output praticamente identici tranne per alcune parti.

Clustal e e Mafft presentano sequenze con variazioni di carattere nelle posizioni 29890 della sequenza 3, 29898 della sequenza 4 e 29899 della sequenza 6.

Per Kalign invece si hanno dei cambi di carattere nella posizione 29879 delle sequenze 3, 4 e 6.

### **Spagna**

Nell'eseguire l'analisi con i tre differenti tool abbiamo ottenuto dei risultati identici tranne nelle ultime due sequenze della Spagna. Mentre Kalign e Mafft producono lo stesso output qualcosa cambia con Clustal.

Nell'output di Clustal, nella sesta sequenza spagnola, troviamo dei cambi di carattere nelle posizioni 5175, 6353, 28882, 29687 e 29733. Per quanto riguarda la sequenza sette, invece, si hanno dei cambi di carattere nelle posizioni 8738, 18743 e 22892.

Con Kalign e Mafft, nella sequenza sei, abbiamo dei cambi di caratteri nelle posizioni 8738, 18743 e 22892. Nella sequenza sette i cambi di carattere sono nelle posizioni 5175, 6353, 28882, 29687 e 29733.

## **Analisi verticale**

Nell'analisi verticale abbiamo, invece, deciso di confrontare le ultime sequenze che erano disponibili con quella di riferimento. Anche questi risultati sono consultabili nell'appendice B della

relazione

Il risultato che abbiamo ottenuto non differisce fra i tre tool, tutte le sequenze hanno subito gli stessi cambiamenti nonostante si trovassero in aree geografiche diverse. Questo risultato ottenuto in base alle sequenze da noi scelte, sottolinea come nonostante il virus sia molto mutevole, questo potrebbero non dipendere dal clima o il luogo in cui si trova.

## **Appendice A - risultati confronto verticale**

## New York (USA)

Clustal Omega

ID_SEQ	POS	LENGTH	REFERENCE	MUTATION
1_USA—EPI_ISL_420793	240	1	C	T
1_USA—EPI_ISL_420793	1058	1	C	T
1_USA—EPI_ISL_420793	3036	1	C	T
1_USA—EPI_ISL_420793	14407	1	C	T
1_USA—EPI_ISL_420793	23402	1	A	G
1_USA—EPI_ISL_420793	25562	1	G	T
1_USA—EPI_ISL_420793	29881	21	AAAAAAAAAAAAAAAAAAAAAA	
2_USA—EPI_ISL_426291	0	30	ATTAAGGTTTACCTCCAG GTAACAA	
2_USA—EPI_ISL_426291	240	1	C	T
2_USA—EPI_ISL_426291	1058	1	C	T
2_USA—EPI_ISL_426291	3036	1	C	T
2_USA—EPI_ISL_426291	8769	1	C	Y
2_USA—EPI_ISL_426291	14407	1	C	T
2_USA—EPI_ISL_426291	23402	1	A	G
2_USA—EPI_ISL_426291	25562	1	G	T
2_USA—EPI_ISL_426291	29835	67	CCATGTGATTAAAGCTTC TTAGGAGAATGACAAAAAA AAAAAAAAAAAAAAA AAAAAAA	
3_USA—EPI_ISL_421365	0	55	ATTAAGGTTTACCTCCC AGGTAACAAACCAACCAACT TTCGATCTCTGTA	
3_USA—EPI_ISL_421365	2675	1	C	T
3_USA—EPI_ISL_421365	8781	1	C	T
3_USA—EPI_ISL_421365	22605	1	A	T
3_USA—EPI_ISL_421365	23524	1	C	T
3_USA—EPI_ISL_421365	28143	1	T	C
3_USA—EPI_ISL_421365	28877	1	G	A
3_USA—EPI_ISL_421365	29741	1	G	A
3_USA—EPI_ISL_421365	29780	122	AGCTGCCTATGGAGAGCCC TAATGTGAAAATTAAATTAGT AGTGCATCCCCATGTGATTAA ATAGCTTCTTAGGAGAATGACA AAAAAAAAAAAAAAA AAAAAAA	
4_USA—EPI_ISL_421594	0	55	ATTAAGGTTTACCTCCAG GTAACAAACCAACCAACTTCG ATCTCTGTA	
4_USA—EPI_ISL_421594	240	1	C	T
4_USA—EPI_ISL_421594	1058	1	C	T
4_USA—EPI_ISL_421594	3036	1	C	T
4_USA—EPI_ISL_421594	14407	1	C	T

4_USA—EPI_ISL_421594	23125	1	G	T
4_USA—EPI_ISL_421594	23402	1	A	G
4_USA—EPI_ISL_421594	25562	1	G	T
4_USA—EPI_ISL_421594	29763	139	TGAACAATGCTAGGGAGAGC TGCCTATATGGAAGAGCCCTA ATGTGTAAAATTAATTTAGTA GTGCTATCCCCATGTGATTAA ATAGCTTCTTAGGAGAATGAC AAAAAAAAAAAAAAA AAAAAAAAAAA	_____
5_USA—EPI_ISL_426309	0	25	ATTAAAGGTTTATACCTTCCCA GGT	_____
5_USA—EPI_ISL_426309	240	1	C	T
5_USA—EPI_ISL_426309	1058	1	C	T
5_USA—EPI_ISL_426309	3036	1	C	T
5_USA—EPI_ISL_426309	14407	1	C	T
5_USA—EPI_ISL_426309	23402	1	A	G
5_USA—EPI_ISL_426309	25562	1	G	T
5_USA—EPI_ISL_426309	29846	56	TAATAGCTTCTTAGGAGAATG ACAAAAAAA AAAAAAA	_____
6_USA—EPI_ISL_428767	0	9	ATTAAAGGT	NNNNNNNN
6_USA—EPI_ISL_428767	211	1	T	A
6_USA—EPI_ISL_428767	232	1	C	T
6_USA—EPI_ISL_428767	240	1	C	T
6_USA—EPI_ISL_428767	1058	1	C	T
6_USA—EPI_ISL_428767	2535	1	C	T
6_USA—EPI_ISL_428767	2841	1	C	T
6_USA—EPI_ISL_428767	3036	1	C	T
6_USA—EPI_ISL_428767	14407	1	C	T
6_USA—EPI_ISL_428767	18936	1	G	A
6_USA—EPI_ISL_428767	23402	1	A	G
6_USA—EPI_ISL_428767	25562	1	G	T
7_USA—EPI_ISL_430418	0	1	A	N
7_USA—EPI_ISL_430418	240	1	C	T
7_USA—EPI_ISL_430418	1058	1	C	T
7_USA—EPI_ISL_430418	3036	1	C	T
7_USA—EPI_ISL_430418	14407	1	C	T
7_USA—EPI_ISL_430418	23402	1	A	G
7_USA—EPI_ISL_430418	25562	1	G	T
8_USA—EPI_ISL_435521	0	10	ATTAAAGGTT	NNNNNNNNNN
8_USA—EPI_ISL_435521	240	1	C	T
8_USA—EPI_ISL_435521	1058	1	C	T
8_USA—EPI_ISL_435521	3036	1	C	T
8_USA—EPI_ISL_435521	14407	1	C	T

8_USA—EPI_ISL_435521	15573	1	G	A
8_USA—EPI_ISL_435521	23402	1	A	G
8_USA—EPI_ISL_435521	25562	1	G	T
			ATTAAGGTTTACCTCCC AGGTAACAAACCAACCAACT TTCGATCTCTGTAGATCTGT TCTCTAACGAACCTTAAAAT CTGTGTGGCTGTCACTCGGCT GCATGCTTAGTGCACTCACGC AGTATAATTAA	NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN
9_USA—EPI_ISL_435547	0	136		NNNNNNNNNNNNNNNNNNNN
9_USA—EPI_ISL_435547	145	3	ACT	NNN
9_USA—EPI_ISL_435547	240	1	C	T
9_USA—EPI_ISL_435547	1058	1	C	T
9_USA—EPI_ISL_435547	3036	1	C	T
			CTGACACACGTTATGTGCTCATG GATGGCTCTATTATTCAATTCCCT AACACCTACCTTGAAGGTTCTGT TAGAGTGGTAACAACTTTGATT CTGAGTACT	NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN
9_USA—EPI_ISL_435547	9113	102		NNNNNNNNNNNNNNNNNNNN
9_USA—EPI_ISL_435547	11915	1	C	T
9_USA—EPI_ISL_435547	14407	1	C	T
9_USA—EPI_ISL_435547	18997	1	C	T
9_USA—EPI_ISL_435547	23402	1	A	G
9_USA—EPI_ISL_435547	25562	1	G	T
9_USA—EPI_ISL_435547	27058	1	C	T
9_USA—EPI_ISL_435547	29539	1	G	A
9_USA—EPI_ISL_435547	29901	1	A	N

ID_SEQ	POS	LENGTH	REFERENCE	MUTATION
1_USA—EPI_ISL_420793	240	1	C	T
1_USA—EPI_ISL_420793	1058	1	C	T
1_USA—EPI_ISL_420793	3036	1	C	T
1_USA—EPI_ISL_420793	14407	1	C	T
1_USA—EPI_ISL_420793	23402	1	A	G
1_USA—EPI_ISL_420793	25562	1	G	T
1_USA—EPI_ISL_420793	29881	21	AAAAAAAAAAAAAAAAAAAAAA	—————
2_USA—EPI_ISL_426291	0	30	ATTAAAGGTTTATACCTTCCC AGGTAACAA	—————
2_USA—EPI_ISL_426291	240	1	C	T
2_USA—EPI_ISL_426291	1058	1	C	T
2_USA—EPI_ISL_426291	3036	1	C	T
2_USA—EPI_ISL_426291	8769	1	C	Y
2_USA—EPI_ISL_426291	14407	1	C	T
2_USA—EPI_ISL_426291	23402	1	A	G
2_USA—EPI_ISL_426291	25562	1	G	T
2_USA—EPI_ISL_426291	29835	67	CCATGTGATTTAATAGCTT CTTAGGAGAACGACAAAAAA AAAAAAAAAAAAAAAAAAAAAA AAAAAAA	————— ————— ————— —————
3_USA—EPI_ISL_421365	0	55	ATTAAAGGTTTATACCTTCC CAGGTAAACAAACCAACCAAC TTTCGATCTCTT GTA	————— ————— —————
3_USA—EPI_ISL_421365	2675	1	C	T
3_USA—EPI_ISL_421365	8781	1	C	T
3_USA—EPI_ISL_421365	22605	1	A	T
3_USA—EPI_ISL_421365	23524	1	C	T
3_USA—EPI_ISL_421365	28143	1	T	C
3_USA—EPI_ISL_421365	28877	1	G	A
3_USA—EPI_ISL_421365	29741	1	G	A
3_USA—EPI_ISL_421365	29780	122	AGCTGCCTATATGGAAGAGCC CTAATGTGTAATTAATTAA GTAGTGCTATCCCCATGTGATT TTAATAGCTTCTTAGGAGAACG ACAAAAAAA AAAAAAA	————— ————— ————— ————— —————
4_USA—EPI_ISL_421594	0	55	ATTAAAGGTTTATACCTTCCC GGTAACAAACCAACCAACTTT CGATCTCTTGTAA	————— ————— —————
4_USA—EPI_ISL_421594	240	1	C	T
4_USA—EPI_ISL_421594	1058	1	C	T
4_USA—EPI_ISL_421594	3036	1	C	T
4_USA—EPI_ISL_421594	14407	1	C	T

4_USA—EPI_ISL_421594	23125	1	G	T
4_USA—EPI_ISL_421594	23402	1	A	G
4_USA—EPI_ISL_421594	25562	1	G	T
			TGAACAAATGCTAGGGAGAGCTG CCTATATGGAAGAGCCCTAATGT GTAAAATTAATTAGTAGTGCTA TCCCCATGTGATTAAATAGCTTC TTAGGAGAATGACA AAAAAAAAAAAAA AAAAAAAAAAAAA	
4_USA—EPI_ISL_421594	29763	139		
5_USA—EPI_ISL_426309	0	25	ATTAAGGTTATACCTTCCCAGGT	
5_USA—EPI_ISL_426309	240	1	C	T
5_USA—EPI_ISL_426309	1058	1	C	T
5_USA—EPI_ISL_426309	3036	1	C	T
5_USA—EPI_ISL_426309	14407	1	C	T
5_USA—EPI_ISL_426309	23402	1	A	G
5_USA—EPI_ISL_426309	25562	1	G	T
			TAATAGCTTCTTAGGAGAATGA CAAAAAAAAAAAAA AAAAAAAAAAAA	
5_USA—EPI_ISL_426309	29846	56		
6_USA—EPI_ISL_428767	0	9	ATTAAGGT	NNNNNNNNN
6_USA—EPI_ISL_428767	211	1	T	A
6_USA—EPI_ISL_428767	232	1	C	T
6_USA—EPI_ISL_428767	240	1	C	T
6_USA—EPI_ISL_428767	1058	1	C	T
6_USA—EPI_ISL_428767	2535	1	C	T
6_USA—EPI_ISL_428767	2841	1	C	T
6_USA—EPI_ISL_428767	3036	1	C	T
6_USA—EPI_ISL_428767	14407	1	C	T
6_USA—EPI_ISL_428767	18936	1	G	A
6_USA—EPI_ISL_428767	23402	1	A	G
6_USA—EPI_ISL_428767	25562	1	G	T
7_USA—EPI_ISL_430418	0	1	A	N
7_USA—EPI_ISL_430418	240	1	C	T
7_USA—EPI_ISL_430418	1058	1	C	T
7_USA—EPI_ISL_430418	3036	1	C	T
7_USA—EPI_ISL_430418	14407	1	C	T
7_USA—EPI_ISL_430418	23402	1	A	G
7_USA—EPI_ISL_430418	25562	1	G	T
8_USA—EPI_ISL_435521	0	10	ATTAAGGTT	NNNNNNNNNN
8_USA—EPI_ISL_435521	240	1	C	T
8_USA—EPI_ISL_435521	1058	1	C	T
8_USA—EPI_ISL_435521	3036	1	C	T
8_USA—EPI_ISL_435521	14407	1	C	T

8_USA—EPI_ISL_435521	15573	1	G	A
8_USA—EPI_ISL_435521	23402	1	A	G
8_USA—EPI_ISL_435521	25562	1	G	T
9_USA—EPI_ISL_435547	0	136	ATTAAAGGTTATACCTTCCC AGGTAAACAACCAACCAACT TTCGATCTCTTAGATCTGTT CTCTAAACGAACTTAAAATC TGTGTGGCTGTCACTCGGCTG CATGCTTAGTGCACTCACGCA GTATAATTAA	NNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN
9_USA—EPI_ISL_435547	145	3	ACT	NNN
9_USA—EPI_ISL_435547	240	1	C	T
9_USA—EPI_ISL_435547	1058	1	C	T
9_USA—EPI_ISL_435547	3036	1	C	T
9_USA—EPI_ISL_435547	9113	102	CTGACACACGTTATGTGCTC ATGGATGGCTCTATTATTCAA TTTCCTAACACCTACCTTGAA GGTTCTGTTAGAGTGGTAACA ACTTTGATTCTGAGTACT	NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN
9_USA—EPI_ISL_435547	11915	1	C	T
9_USA—EPI_ISL_435547	14407	1	C	T
9_USA—EPI_ISL_435547	18997	1	C	T
9_USA—EPI_ISL_435547	23402	1	A	G
9_USA—EPI_ISL_435547	25562	1	G	T
9_USA—EPI_ISL_435547	27058	1	C	T
9_USA—EPI_ISL_435547	29539	1	G	A
9_USA—EPI_ISL_435547	29901	1	A	N

## MAFFT

ID_SEQ	POS	LENGTH	REFERENCE	MUTATION
1_USA—EPI_ISL_420793	240	1	C	T
1_USA—EPI_ISL_420793	1058	1	C	T
1_USA—EPI_ISL_420793	3036	1	C	T
1_USA—EPI_ISL_420793	14407	1	C	T
1_USA—EPI_ISL_420793	23402	1	A	G
1_USA—EPI_ISL_420793	25562	1	G	T
1_USA—EPI_ISL_420793	29881	21	AAAAAAAAAAAAAAAAAAAAAA	—————
2_USA—EPI_ISL_426291	0	30	ATTAAAGGTTATACCTTCCCAG GTAACAA	—————
2_USA—EPI_ISL_426291	240	1	C	T
2_USA—EPI_ISL_426291	1058	1	C	T
2_USA—EPI_ISL_426291	3036	1	C	T
2_USA—EPI_ISL_426291	8769	1	C	Y
2_USA—EPI_ISL_426291	14407	1	C	T
2_USA—EPI_ISL_426291	23402	1	A	G
2_USA—EPI_ISL_426291	25562	1	G	T
2_USA—EPI_ISL_426291	29835	67	CCATGTGATTTAATAGCTTC TTAGGAGAACATGACAAAAAAA AAAAAAAAAAAAAAA AAAAAAA	————— ————— ————— ——
3_USA—EPI_ISL_421365	0	55	ATTAAAGGTTATACCTTCCC AGGTAACAAACCAACCAACTT TCGATCTCTTGTA	————— ————— —————
3_USA—EPI_ISL_421365	2675	1	C	T
3_USA—EPI_ISL_421365	8781	1	C	T
3_USA—EPI_ISL_421365	22605	1	A	T
3_USA—EPI_ISL_421365	23524	1	C	T
3_USA—EPI_ISL_421365	28143	1	T	C
3_USA—EPI_ISL_421365	28877	1	G	A
3_USA—EPI_ISL_421365	29741	1	G	A
3_USA—EPI_ISL_421365	29780	122	AGCTGCCTATATGGAAGAGCC CTAATGTGAAATTAATTAA GTAGTGCTATCCCCATGTGATT TTAATAGCTTCTTAGGAGAATG ACAAAAAAA AAAAAAA	————— ————— ————— ————— —————
4_USA—EPI_ISL_421594	0	55	ATTAAAGGTTATACCTTCCC GGTAACAAACCAACCAACTT CGATCTCTTGTA	————— ————— —————
4_USA—EPI_ISL_421594	240	1	C	T
4_USA—EPI_ISL_421594	1058	1	C	T
4_USA—EPI_ISL_421594	3036	1	C	T
4_USA—EPI_ISL_421594	14407	1	C	T
4_USA—EPI_ISL_421594	23125	1	G	T

4_USA—EPI_ISL_421594	23402	1	A	G
4_USA—EPI_ISL_421594	25562	1	G	T
4_USA—EPI_ISL_421594	29763	139	TGAACAAATGCTAGGGAGAGC TGCCTATATGGAAGAGCCCTA ATGTGTAAAATTAAATTTAGTA GTGCTATCCCCATGTGATTTA ATAGCTTCTTAGGAGAATGAC AAAAAAAAAAAAAAA AAAAAAAAAAAAAA	_____
5_USA—EPI_ISL_426309	0	25	ATTAAGGTTTACCTTCCA GGT	_____
5_USA—EPI_ISL_426309	240	1	C	T
5_USA—EPI_ISL_426309	1058	1	C	T
5_USA—EPI_ISL_426309	3036	1	C	T
5_USA—EPI_ISL_426309	14407	1	C	T
5_USA—EPI_ISL_426309	23402	1	A	G
5_USA—EPI_ISL_426309	25562	1	G	T
5_USA—EPI_ISL_426309	29846	56	TAATAGCTCTAGGAGAATG ACAAAAAAA AAAAAAA	_____
6_USA—EPI_ISL_428767	0	9	ATTAAGGT	NNNNNNNN
6_USA—EPI_ISL_428767	211	1	T	A
6_USA—EPI_ISL_428767	232	1	C	T
6_USA—EPI_ISL_428767	240	1	C	T
6_USA—EPI_ISL_428767	1058	1	C	T
6_USA—EPI_ISL_428767	2535	1	C	T
6_USA—EPI_ISL_428767	2841	1	C	T
6_USA—EPI_ISL_428767	3036	1	C	T
6_USA—EPI_ISL_428767	14407	1	C	T
6_USA—EPI_ISL_428767	18936	1	G	A
6_USA—EPI_ISL_428767	23402	1	A	G
6_USA—EPI_ISL_428767	25562	1	G	T
7_USA—EPI_ISL_430418	0	1	A	N
7_USA—EPI_ISL_430418	240	1	C	T
7_USA—EPI_ISL_430418	1058	1	C	T
7_USA—EPI_ISL_430418	3036	1	C	T
7_USA—EPI_ISL_430418	14407	1	C	T
7_USA—EPI_ISL_430418	23402	1	A	G
7_USA—EPI_ISL_430418	25562	1	G	T
8_USA—EPI_ISL_435521	0	10	ATTAAGGTT	NNNNNNNN
8_USA—EPI_ISL_435521	240	1	C	T
8_USA—EPI_ISL_435521	1058	1	C	T
8_USA—EPI_ISL_435521	3036	1	C	T
8_USA—EPI_ISL_435521	14407	1	C	T
8_USA—EPI_ISL_435521	15573	1	G	A

8_USA—EPI_ISL_435521	23402	1	A	G
8_USA—EPI_ISL_435521	25562	1	G	T
9_USA—EPI_ISL_435547	0	136	ATTAAAGGTTTATAACCTTCCC AGGTAAACAAACCAACCAAC TTTCGATCTCTTAGATCTG TTCTCTAACGAACCTTAAA TCTGTGTGGCTGTCACTCGGC TGCATGCTTAGTGCACTCACG CAGTATAATTAA	NNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN
9_USA—EPI_ISL_435547	145	3	ACT	NNN
9_USA—EPI_ISL_435547	240	1	C	T
9_USA—EPI_ISL_435547	1058	1	C	T
9_USA—EPI_ISL_435547	3036	1	C	T
9_USA—EPI_ISL_435547	9113	102	CTGACACACGTTATGTGCTCA TGGATGGCTCTATTATTCAATT TCCTAACACACCTACCTTGAAGG TTCTGTTAGAGTGGTAACAAAC TTTGATTCTGAGTACT	NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN
9_USA—EPI_ISL_435547	11915	1	C	T
9_USA—EPI_ISL_435547	14407	1	C	T
9_USA—EPI_ISL_435547	18997	1	C	T
9_USA—EPI_ISL_435547	23402	1	A	G
9_USA—EPI_ISL_435547	25562	1	G	T
9_USA—EPI_ISL_435547	27058	1	C	T
9_USA—EPI_ISL_435547	29539	1	G	A
9_USA—EPI_ISL_435547	29901	1	A	N

## Russia

Clustal Omega

ID_SEQ	POS	LENGTH	REFERENCE	MUTATION
1_Russia—EPI_ISL_428860	240	1	C	T
1_Russia—EPI_ISL_428860	3036	1	C	T
1_Russia—EPI_ISL_428860	14407	1	C	T
1_Russia—EPI_ISL_428860	18567	1	C	T
1_Russia—EPI_ISL_428860	23402	1	A	G
1_Russia—EPI_ISL_428860	26529	1	A	G
2_Russia—EPI_ISL_421275	240	1	C	T
2_Russia—EPI_ISL_421275	3036	1	C	T
2_Russia—EPI_ISL_421275	4001	1	C	T
2_Russia—EPI_ISL_421275	10096	1	G	A
2_Russia—EPI_ISL_421275	11082	1	G	T
2_Russia—EPI_ISL_421275	13535	1	C	T
2_Russia—EPI_ISL_421275	14407	1	C	T
2_Russia—EPI_ISL_421275	23402	1	A	G
2_Russia—EPI_ISL_421275	23730	1	C	T

2_Russia—EPI_ISL_421275	27405	1	C	T
2_Russia—EPI_ISL_421275	28880	3	GGG	AAC
2_Russia—EPI_ISL_421275	29890	12	AAAAAAAAAAAAA	———
3_Russia—EPI_ISL_427337	240	1	C	T
3_Russia—EPI_ISL_427337	3036	1	C	T
3_Russia—EPI_ISL_427337	14407	1	C	T
3_Russia—EPI_ISL_427337	23402	1	A	G
4_Russia—EPI_ISL_428852	240	1	C	T
4_Russia—EPI_ISL_428852	3036	1	C	T
4_Russia—EPI_ISL_428852	10968	1	C	T
4_Russia—EPI_ISL_428852	14407	1	C	T
4_Russia—EPI_ISL_428852	23402	1	A	G
4_Russia—EPI_ISL_428852	28880	3	GGG	AAC
5_Russia—EPI_ISL_427334	0	8	ATTAAAGG	NNNNNNNN
5_Russia—EPI_ISL_427334	75	180	TTTAAAATCTGTGTGGCTGT CACTCGGCTGCATGCTTAGT GCACTCACCGCAGTATAATT ATAACTAATTACTGTCGTTG ACAGGACACGAGTAACTCGT CTATCTTCTGCAGGCTGCTTA CGGTTTGTCCGTGTTGCAG CCGATCATCAGCACATCTAGG TTTCGTCCGGGTGTGACC	NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN
5_Russia—EPI_ISL_427334	3036	1	C	T
5_Russia—EPI_ISL_427334	14407	1	C	T
5_Russia—EPI_ISL_427334	23402	1	A	G
5_Russia—EPI_ISL_427334	26749	1	C	T
5_Russia—EPI_ISL_427334	28880	3	GGG	AAC
5_Russia—EPI_ISL_427334	29866	36	GACAAAAAAAAAAAAAAA AAAAAAA	NNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN
6_Russia—EPI_ISL_430096	0	8	ATTAAAGG	NNNNNNNN
6_Russia—EPI_ISL_430096	240	1	C	T
6_Russia—EPI_ISL_430096	3036	1	C	T
6_Russia—EPI_ISL_430096	14407	1	C	T
6_Russia—EPI_ISL_430096	23402	1	A	G
6_Russia—EPI_ISL_430096	26749	1	C	T
6_Russia—EPI_ISL_430096	28880	3	GGG	AAC
6_Russia—EPI_ISL_430096	29872	30	AAAAAAAAAAAAAAA AAAAAAA	NNNNNNNNNNNNNNNNNN NNNNNNNNNNNN
7_Russia—EPI_ISL_430108	240	1	C	T
7_Russia—EPI_ISL_430108	834	1	C	T
7_Russia—EPI_ISL_430108	3036	1	C	T
7_Russia—EPI_ISL_430108	14407	1	C	T
7_Russia—EPI_ISL_430108	20267	1	A	G

7_Russia—EPI_ISL_430108	23402	1	A	G
7_Russia—EPI_ISL_430108	29870	32	AAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAA	NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNN

ID_SEQ	PO	LENGTH	REFERENCE	MUTATION
1_Russia—EPI_ISL_428860	240	1	C	T
1_Russia—EPI_ISL_428860	3036	1	C	T
1_Russia—EPI_ISL_428860	14407	1	C	T
1_Russia—EPI_ISL_428860	18567	1	C	T
1_Russia—EPI_ISL_428860	23402	1	A	G
1_Russia—EPI_ISL_428860	26529	1	A	G
2_Russia—EPI_ISL_421275	240	1	C	T
2_Russia—EPI_ISL_421275	3036	1	C	T
2_Russia—EPI_ISL_421275	4001	1	C	T
2_Russia—EPI_ISL_421275	10096	1	G	A
2_Russia—EPI_ISL_421275	11082	1	G	T
2_Russia—EPI_ISL_421275	13535	1	C	T
2_Russia—EPI_ISL_421275	14407	1	C	T
2_Russia—EPI_ISL_421275	23402	1	A	G
2_Russia—EPI_ISL_421275	23730	1	C	T
2_Russia—EPI_ISL_421275	27405	1	C	T
2_Russia—EPI_ISL_421275	28880	3	GGG	AAC
2_Russia—EPI_ISL_421275	29890	12	AAAAAAAAAAAA	_____
3_Russia—EPI_ISL_427337	240	1	C	T
3_Russia—EPI_ISL_427337	3036	1	C	T
3_Russia—EPI_ISL_427337	14407	1	C	T
3_Russia—EPI_ISL_427337	23402	1	A	G
4_Russia—EPI_ISL_428852	240	1	C	T
4_Russia—EPI_ISL_428852	3036	1	C	T
4_Russia—EPI_ISL_428852	10968	1	C	T
4_Russia—EPI_ISL_428852	14407	1	C	T
4_Russia—EPI_ISL_428852	23402	1	A	G
4_Russia—EPI_ISL_428852	28880	3	GGG	AAC
5_Russia—EPI_ISL_427334	0	8	ATTAAAGG	NNNNNNNN
5_Russia—EPI_ISL_427334	75	180	TTTAAATCTGTGTGGCTGT CACTCGGCTGCATGCTTAGT GCACTCACGCAGTATAATTA ATAACTAATTACTGTCGTG ACAGGACACGAGTAACTCG TCTATCTCTGCAGGCTGCT TACGGTTCGTCCGTGTTGC AGCCGATCATCAGCACATCT AGGTTTCGTCCGGGTGTGACC	NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN
5_Russia—EPI_ISL_427334	3036	1	C	T
5_Russia—EPI_ISL_427334	14407	1	C	T
5_Russia—EPI_ISL_427334	23402	1	A	G
5_Russia—EPI_ISL_427334	26749	1	C	T

5_Russia—EPI_ISL_427334	28880	3	GGG	AAC
5_Russia—EPI_ISL_427334	29866	36	GACAAAAAAAAAAAAAA AAAAA	NNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN
6_Russia—EPI_ISL_430096	0	8	ATTAAAGG	NNNNNNNN
6_Russia—EPI_ISL_430096	240	1	C	T
6_Russia—EPI_ISL_430096	3036	1	C	T
6_Russia—EPI_ISL_430096	14407	1	C	T
6_Russia—EPI_ISL_430096	23402	1	A	G
6_Russia—EPI_ISL_430096	26749	1	C	T
6_Russia—EPI_ISL_430096	28880	3	GGG	AAC
6_Russia—EPI_ISL_430096	29872	30	AAAAAAAAAAAAAAA AAAAA	NNNNNNNNNNNNNNNN NNNNNNNNNN
7_Russia—EPI_ISL_430108	240	1	C	T
7_Russia—EPI_ISL_430108	834	1	C	T
7_Russia—EPI_ISL_430108	3036	1	C	T
7_Russia—EPI_ISL_430108	14407	1	C	T
7_Russia—EPI_ISL_430108	20267	1	A	G
7_Russia—EPI_ISL_430108	23402	1	A	G
7_Russia—EPI_ISL_430108	29870	32	AAAAAAAAAAAAAAA AAAAA	NNNNNNNNNNNNNNNN NNNNNNNNNNNNNN

ID_SEQ	POS	LENGTH	REFERENCE	MUTATION
1_Russia—EPI_ISL_428860	240	1	C	T
1_Russia—EPI_ISL_428860	3036	1	C	T
1_Russia—EPI_ISL_428860	14407	1	C	T
1_Russia—EPI_ISL_428860	18567	1	C	T
1_Russia—EPI_ISL_428860	23402	1	A	G
1_Russia—EPI_ISL_428860	26529	1	A	G
2_Russia—EPI_ISL_421275	240	1	C	T
2_Russia—EPI_ISL_421275	3036	1	C	T
2_Russia—EPI_ISL_421275	4001	1	C	T
2_Russia—EPI_ISL_421275	10096	1	G	A
2_Russia—EPI_ISL_421275	11082	1	G	T
2_Russia—EPI_ISL_421275	13535	1	C	T
2_Russia—EPI_ISL_421275	14407	1	C	T
2_Russia—EPI_ISL_421275	23402	1	A	G
2_Russia—EPI_ISL_421275	23730	1	C	T
2_Russia—EPI_ISL_421275	27405	1	C	T
2_Russia—EPI_ISL_421275	28880	3	GGG	AAC
2_Russia—EPI_ISL_421275	29890	12	AAAAAAAAAAAA	_____
3_Russia—EPI_ISL_427337	240	1	C	T
3_Russia—EPI_ISL_427337	3036	1	C	T
3_Russia—EPI_ISL_427337	14407	1	C	T
3_Russia—EPI_ISL_427337	23402	1	A	G
4_Russia—EPI_ISL_428852	240	1	C	T
4_Russia—EPI_ISL_428852	3036	1	C	T
4_Russia—EPI_ISL_428852	10968	1	C	T
4_Russia—EPI_ISL_428852	14407	1	C	T
4_Russia—EPI_ISL_428852	23402	1	A	G
4_Russia—EPI_ISL_428852	28880	3	GGG	AAC
5_Russia—EPI_ISL_427334	0	8	ATTAAAGG	NNNNNNNN
5_Russia—EPI_ISL_427334	75	180	TTTAAATCTGTGGCTGTC ACTCGGCTGCATGCTTAGTGC ACTCACGCAGTATAATTAA CTAATTACTGTCGTTGACAGGA CACGAGTAACTCGTCTATCTC TGCAGGCTGCTTACGGTTCGT CCGTGTTGCAGCCGATCATCAG CACATCTAGGTTCGTCCGGGT GTGACC	NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN
5_Russia—EPI_ISL_427334	3036	1	C	T
5_Russia—EPI_ISL_427334	14407	1	C	T
5_Russia—EPI_ISL_427334	23402	1	A	G
5_Russia—EPI_ISL_427334	26749	1	C	T
5_Russia—EPI_ISL_427334	28880	3	GGG	AAC

5_Russia—EPI_ISL_427334	29866	36	GACAAAAAAAAAAAAAA AAAAA	NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN
6_Russia—EPI_ISL_430096	0	8	ATTAAAGG	NNNNNNNN
6_Russia—EPI_ISL_430096	240	1	C	T
6_Russia—EPI_ISL_430096	3036	1	C	T
6_Russia—EPI_ISL_430096	14407	1	C	T
6_Russia—EPI_ISL_430096	23402	1	A	G
6_Russia—EPI_ISL_430096	26749	1	C	T
6_Russia—EPI_ISL_430096	28880	3	GGG	AAC
6_Russia—EPI_ISL_430096	29872	30	AAAAAAA AAAAAAA	NNNNNNNNNNNNNNNN NNNNNNNNNN
7_Russia—EPI_ISL_430108	240	1	C	T
7_Russia—EPI_ISL_430108	834	1	C	T
7_Russia—EPI_ISL_430108	3036	1	C	T
7_Russia—EPI_ISL_430108	14407	1	C	T
7_Russia—EPI_ISL_430108	20267	1	A	G
7_Russia—EPI_ISL_430108	23402	1	A	G
7_Russia—EPI_ISL_430108	29870	32	AAAAAAA AAAAAAA	NNNNNNNNNNNNNNNN NNNNNNNNNNNN

## Italia

## Clustal Omega

ID_SEQ	POS	LENGTH	REFERENCE	MUTATION
1.Italy—EPI_ISL_412974	11083	1	G	T
1.Italy—EPI_ISL_412974	26144	1	G	T
2.Italy—EPI_ISL_412973	241	1	C	T
2.Italy—EPI_ISL_412973	3037	1	C	T
2.Italy—EPI_ISL_412973	14408	1	C	T
2.Italy—EPI_ISL_412973	23403	1	A	G
2.Italy—EPI_ISL_412973	29867	2	TG	NN
3.Italy—EPI_ISL_417423	0	1	-	G
3.Italy—EPI_ISL_417423	241	1	C	T
3.Italy—EPI_ISL_417423	3037	1	C	T
3.Italy—EPI_ISL_417423	14408	1	C	T
3.Italy—EPI_ISL_417423	23403	1	A	G
3.Italy—EPI_ISL_417423	29891	13	AAAAAAAAAAAAAA	-----
4.Italy—EPI_ISL_418256	42	1	T	Y
4.Italy—EPI_ISL_418256	44	2	CG	NN
4.Italy—EPI_ISL_418256	241	1	C	T
4.Italy—EPI_ISL_418256	3037	1	C	T
4.Italy—EPI_ISL_418256	14408	1	C	T
4.Italy—EPI_ISL_418256	23403	1	A	G
4.Italy—EPI_ISL_418256	28881	3	GGG	AAC
4.Italy—EPI_ISL_418256	29899	5	AAAAA	—
5.Italy—EPI_ISL_419254	1	25	ATTAAGGTTATACCTTCCCAGGT	-----
5.Italy—EPI_ISL_419254	241	1	C	T
5.Italy—EPI_ISL_419254	3037	1	C	T
5.Italy—EPI_ISL_419254	14408	1	C	T
5.Italy—EPI_ISL_419254	23403	1	A	G
5.Italy—EPI_ISL_419254	29860	44	AGGAGAATGACAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAA	-----
6.Italy—EPI_ISL_428854	241	1	C	T
6.Italy—EPI_ISL_428854	3037	1	C	T
6.Italy—EPI_ISL_428854	14408	1	C	T
6.Italy—EPI_ISL_428854	23403	1	A	G
6.Italy—EPI_ISL_428854	29009	1	G	A
6.Italy—EPI_ISL_428854	29900	4	AAAA	—
7.Italy—EPI_ISL_435152	1	37	ATTAAGGTTATACCTTCC CAGGTAACAAACCAACC	NNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN
7.Italy—EPI_ISL_435152	241	1	C	T
7.Italy—EPI_ISL_435152	3037	1	C	T
7.Italy—EPI_ISL_435152	5986	1	C	T
7.Italy—EPI_ISL_435152	6866	1	A	W

7_Italy—EPI_ISL_435152	6869	1	A	W
7_Italy—EPI_ISL_435152	14408	1	C	T
7_Italy—EPI_ISL_435152	23403	1	A	G
7_Italy—EPI_ISL_435152	28881	3	GGG	AAC
7_Italy—EPI_ISL_435152	29701	1	G	T
7_Italy—EPI_ISL_435152	29879	25	AAAAAAAAAAAAAA AAAAAAAAAAAAAA	_____
8_Italy—EPI_ISL_436729	1	9	ATTAAAGGT	NNNNNNNNN
8_Italy—EPI_ISL_436729	31	1	A	G
8_Italy—EPI_ISL_436729	241	1	C	T
8_Italy—EPI_ISL_436729	3037	1	C	T
8_Italy—EPI_ISL_436729	3045	1	C	T
8_Italy—EPI_ISL_436729	6449	1	C	T
8_Italy—EPI_ISL_436729	6863	1	A	M
8_Italy—EPI_ISL_436729	6866	1	A	W
8_Italy—EPI_ISL_436729	6869	1	A	W
8_Italy—EPI_ISL_436729	14408	1	C	T
8_Italy—EPI_ISL_436729	19677	1	G	R
8_Italy—EPI_ISL_436729	21627	1	C	Y
8_Italy—EPI_ISL_436729	23403	1	A	G
8_Italy—EPI_ISL_436729	25459	1	G	K
8_Italy—EPI_ISL_436729	28881	3	GGG	AAC
8_Italy—EPI_ISL_436729	29874	30	AAAAAAAAAAAAAA AAAAAAAAAAAAAA	_____

ID_SEQ	POS	LENGTH	REFERENCE	MUTATION
1_Italy—EPI_ISL_412974	11083	1	G	T
1_Italy—EPI_ISL_412974	26144	1	G	T
2_Italy—EPI_ISL_412973	241	1	C	T
2_Italy—EPI_ISL_412973	3037	1	C	T
2_Italy—EPI_ISL_412973	14408	1	C	T
2_Italy—EPI_ISL_412973	23403	1	A	G
2_Italy—EPI_ISL_412973	29867	2	TG	NN
3_Italy—EPI_ISL_417423	0	1	-	G
3_Italy—EPI_ISL_417423	241	1	C	T
3_Italy—EPI_ISL_417423	3037	1	C	T
3_Italy—EPI_ISL_417423	14408	1	C	T
3_Italy—EPI_ISL_417423	23403	1	A	G
3_Italy—EPI_ISL_417423	29879	13	AAAAAAAAAAAAAA	—————
4_Italy—EPI_ISL_418256	42	1	T	Y
4_Italy—EPI_ISL_418256	44	2	CG	NN
4_Italy—EPI_ISL_418256	241	1	C	T
4_Italy—EPI_ISL_418256	3037	1	C	T
4_Italy—EPI_ISL_418256	14408	1	C	T
4_Italy—EPI_ISL_418256	23403	1	A	G
4_Italy—EPI_ISL_418256	28881	3	GGG	AAC
4_Italy—EPI_ISL_418256	29879	5	AAAAA	—
5_Italy—EPI_ISL_419254	1	25	ATTAAGGTTTAT ACCTTCCCAGGT	—————
5_Italy—EPI_ISL_419254	241	1	C	T
5_Italy—EPI_ISL_419254	3037	1	C	T
5_Italy—EPI_ISL_419254	14408	1	C	T
5_Italy—EPI_ISL_419254	23403	1	A	G
5_Italy—EPI_ISL_419254	29860	44	AGGAGAATGACAAAAAA AAAAAAAAAAAAAAAAAAAAAA	—————
6_Italy—EPI_ISL_428854	241	1	C	T
6_Italy—EPI_ISL_428854	3037	1	C	T
6_Italy—EPI_ISL_428854	14408	1	C	T
6_Italy—EPI_ISL_428854	23403	1	A	G
6_Italy—EPI_ISL_428854	29009	1	G	A
6_Italy—EPI_ISL_428854	29879	4	AAAA	—
7_Italy—EPI_ISL_435152	1	37	ATTAAGGTTTACCTTC CCAGGTAACAAACCAACC	NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN
7_Italy—EPI_ISL_435152	241	1	C	T
7_Italy—EPI_ISL_435152	3037	1	C	T
7_Italy—EPI_ISL_435152	5986	1	C	T
7_Italy—EPI_ISL_435152	6866	1	A	W

7_Italy—EPI_ISL_435152	6869	1	A	W
7_Italy—EPI_ISL_435152	14408	1	C	T
7_Italy—EPI_ISL_435152	23403	1	A	G
7_Italy—EPI_ISL_435152	28881	3	GGG	AAC
7_Italy—EPI_ISL_435152	29701	1	G	T
7_Italy—EPI_ISL_435152	29879	25	AAAAAAAAAAAA AAAAAAAAAAAAAA	—————
8_Italy—EPI_ISL_436729	1	9	ATTAAAGGT	NNNNNNNNN
8_Italy—EPI_ISL_436729	31	1	A	G
8_Italy—EPI_ISL_436729	241	1	C	T
8_Italy—EPI_ISL_436729	3037	1	C	T
8_Italy—EPI_ISL_436729	3045	1	C	T
8_Italy—EPI_ISL_436729	6449	1	C	T
8_Italy—EPI_ISL_436729	6863	1	A	M
8_Italy—EPI_ISL_436729	6866	1	A	W
8_Italy—EPI_ISL_436729	6869	1	A	W
8_Italy—EPI_ISL_436729	14408	1	C	T
8_Italy—EPI_ISL_436729	19677	1	G	R
8_Italy—EPI_ISL_436729	21627	1	C	Y
8_Italy—EPI_ISL_436729	23403	1	A	G
8_Italy—EPI_ISL_436729	25459	1	G	K
8_Italy—EPI_ISL_436729	28881	3	GGG	AAC
8_Italy—EPI_ISL_436729	29874	30	AAAAAAAAAAAAAA AAAAAAAAAAAAAA	—————

ID_SEQ	POS	LENGTH	REFERENCE	MUTATION
1_Italy—EPI_ISL_412974	11083	1	G	T
1_Italy—EPI_ISL_412974	26144	1	G	T
2_Italy—EPI_ISL_412973	241	1	C	T
2_Italy—EPI_ISL_412973	3037	1	C	T
2_Italy—EPI_ISL_412973	14408	1	C	T
2_Italy—EPI_ISL_412973	23403	1	A	G
2_Italy—EPI_ISL_412973	29867	2	TG	NN
3_Italy—EPI_ISL_417423	0	1	-	G
3_Italy—EPI_ISL_417423	241	1	C	T
3_Italy—EPI_ISL_417423	3037	1	C	T
3_Italy—EPI_ISL_417423	14408	1	C	T
3_Italy—EPI_ISL_417423	23403	1	A	G
3_Italy—EPI_ISL_417423	29890	13	AAAAAAAAAAAAAA	-----
4_Italy—EPI_ISL_418256	42	1	T	Y
4_Italy—EPI_ISL_418256	44	2	CG	NN
4_Italy—EPI_ISL_418256	241	1	C	T
4_Italy—EPI_ISL_418256	3037	1	C	T
4_Italy—EPI_ISL_418256	14408	1	C	T
4_Italy—EPI_ISL_418256	23403	1	A	G
4_Italy—EPI_ISL_418256	28881	3	GGG	AAC
4_Italy—EPI_ISL_418256	29898	5	AAAAA	—
5_Italy—EPI_ISL_419254	1	25	ATTAAGGTTATA CCTTCCCAGGT	—————
5_Italy—EPI_ISL_419254	241	1	C	T
5_Italy—EPI_ISL_419254	3037	1	C	T
5_Italy—EPI_ISL_419254	14408	1	C	T
5_Italy—EPI_ISL_419254	23403	1	A	G
5_Italy—EPI_ISL_419254	29859	44	AGGAGAATGACAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAA	—
6_Italy—EPI_ISL_428854	241	1	C	T
6_Italy—EPI_ISL_428854	3037	1	C	T
6_Italy—EPI_ISL_428854	14408	1	C	T
6_Italy—EPI_ISL_428854	23403	1	A	G
6_Italy—EPI_ISL_428854	29009	1	G	A
6_Italy—EPI_ISL_428854	29899	4	AAAA	—
7_Italy—EPI_ISL_435152	1	37	ATTAAGGTTATACCTTCC CAGGTAACAAACCAACC	NNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN
7_Italy—EPI_ISL_435152	241	1	C	T
7_Italy—EPI_ISL_435152	3037	1	C	T
7_Italy—EPI_ISL_435152	5986	1	C	T
7_Italy—EPI_ISL_435152	6866	1	A	W

7_Italy—EPI_ISL_435152	6869	1	A	W
7_Italy—EPI_ISL_435152	14408	1	C	T
7_Italy—EPI_ISL_435152	23403	1	A	G
7_Italy—EPI_ISL_435152	28881	3	GGG	AAC
7_Italy—EPI_ISL_435152	29701	1	G	T
7_Italy—EPI_ISL_435152	29878	25	AAAAAAAAAAAA AAAAAAAAAAAAAA	—————
8_Italy—EPI_ISL_436729	1	9	ATTAAAGGT	NNNNNNNNN
8_Italy—EPI_ISL_436729	31	1	A	G
8_Italy—EPI_ISL_436729	241	1	C	T
8_Italy—EPI_ISL_436729	3037	1	C	T
8_Italy—EPI_ISL_436729	3045	1	C	T
8_Italy—EPI_ISL_436729	6449	1	C	T
8_Italy—EPI_ISL_436729	6863	1	A	M
8_Italy—EPI_ISL_436729	6866	1	A	W
8_Italy—EPI_ISL_436729	6869	1	A	W
8_Italy—EPI_ISL_436729	14408	1	C	T
8_Italy—EPI_ISL_436729	19677	1	G	R
8_Italy—EPI_ISL_436729	21627	1	C	Y
8_Italy—EPI_ISL_436729	23403	1	A	G
8_Italy—EPI_ISL_436729	25459	1	G	K
8_Italy—EPI_ISL_436729	28881	3	GGG	AAC
8_Italy—EPI_ISL_436729	29873	30	AAAAAAAAAAAAAAA AAAAAAAAAAAAAA	—————

## Spagna

Clustal Omega





6_Spain—EPI_ISL_444975	6353	1	C	T
6_Spain—EPI_ISL_444975	14407	1	C	T
6_Spain—EPI_ISL_444975	20267	1	A	G
6_Spain—EPI_ISL_444975	23402	1	A	G
6_Spain—EPI_ISL_444975	28882	1	G	A
6_Spain—EPI_ISL_444975	29687	1	G	C
6_Spain—EPI_ISL_444975	29733	1	G	C
7_Spain—EPI_ISL_444976	240	1	C	T
7_Spain—EPI_ISL_444976	3036	1	C	T
7_Spain—EPI_ISL_444976	8738	1	C	T
7_Spain—EPI_ISL_444976	14407	1	C	T
7_Spain—EPI_ISL_444976	18743	1	C	T
7_Spain—EPI_ISL_444976	20267	1	A	G
7_Spain—EPI_ISL_444976	22892	1	A	G
7_Spain—EPI_ISL_444976	23402	1	A	G

ID_SEQ	PO	LENGTH	REFERENCE	MUTATION
1_Spain—EPI_ISL_418251	240	1	C	T
1_Spain—EPI_ISL_418251	3036	1	C	T
1_Spain—EPI_ISL_418251	8781	1	C	T
1_Spain—EPI_ISL_418251	11243	1	T	C
1_Spain—EPI_ISL_418251	14407	1	C	T
1_Spain—EPI_ISL_418251	21262	1	C	T
1_Spain—EPI_ISL_418251	23402	1	A	G
1_Spain—EPI_ISL_418251	29780	122	AGCTGCCTATATGGAA	_____
			GAGCCCTAATGTGTA	_____
			AAATTAATTTAGTA	_____
			GTGCTATCCCCATGT	_____
			GATTTTAATAGCTTC	_____
			TTAGGAGAACATGACAA	_____
			AAAAAAAAAAAAAAA	_____
2_Spain—EPI_ISL_419240	8781	1	AAAAAAA	_____
			AAAAAAA	_____
			C	T
			T	A
			C	T
			G	T
			T	C
2_Spain—EPI_ISL_419240	28143	1	C	T
			CTATATGGAAGAGCC	_____
			CTAATGTGTAAAATT	_____
			AATTTAGTAGTGCT	_____
			ATCCCCATGTGATTT	_____
			TAATAGCTTCTTAGG	_____
			AGAATGACA	_____
2_Spain—EPI_ISL_419240	28656	1	AAAAAAA	_____
			AAAAAAA	_____
			C	T
			T	C
			C	T
			CTATATGGAAGAGCC	_____
			CTAATGTGTAAAATT	_____
2_Spain—EPI_ISL_419240	28862	1	AATTTAGTAGTGCT	_____
			ATCCCCATGTGATTT	_____
			TAATAGCTTCTTAGG	_____
			AGAATGACA	_____
			AAAAAAA	_____
			AAAAAAA	_____
			CTTTGT	_____
3_Spain—EPI_ISL_419685	0	54	ATTAAAGGTTTATACC	_____
			TTCCCAGGTAACAAAC	_____
3_Spain—EPI_ISL_419685	2000	1	CAACCAACTTTCGATC	_____
			TCTTG	_____
3_Spain—EPI_ISL_419685	2000	1	G	A







				AGCTTCTTAGGAGAA	
5_Spain—EPI_ISL_436324	29850	52		TGACAAAAAAAAAAAAA	
				AAAAAAAAAAAAAAA	
				AAAAAAA	
6_Spain—EPI_ISL_444976	240	1	C		T
6_Spain—EPI_ISL_444976	3036	1	C		T
6_Spain—EPI_ISL_444976	8738	1	C		T
6_Spain—EPI_ISL_444976	14407	1	C		T
6_Spain—EPI_ISL_444976	18743	1	C		T
6_Spain—EPI_ISL_444976	20267	1	A		G
6_Spain—EPI_ISL_444976	22892	1	A		G
6_Spain—EPI_ISL_444976	23402	1	A		G
7_Spain—EPI_ISL_444975	240	1	C		T
7_Spain—EPI_ISL_444975	3036	1	C		T
7_Spain—EPI_ISL_444975	5175	1	A		G
7_Spain—EPI_ISL_444975	6353	1	C		T
7_Spain—EPI_ISL_444975	14407	1	C		T
7_Spain—EPI_ISL_444975	20267	1	A		G
7_Spain—EPI_ISL_444975	23402	1	A		G
7_Spain—EPI_ISL_444975	28882	1	G		A
7_Spain—EPI_ISL_444975	29687	1	G		C
7_Spain—EPI_ISL_444975	29733	1	G		C

## MAFFT

ID_SEQ	POS	LENGTH	REFERENCE	MUTATION
1_Spain—EPI_ISL_418251	240	1	C	T
1_Spain—EPI_ISL_418251	3036	1	C	T
1_Spain—EPI_ISL_418251	8781	1	C	T
1_Spain—EPI_ISL_418251	11243	1	T	C
1_Spain—EPI_ISL_418251	14407	1	C	T
1_Spain—EPI_ISL_418251	21262	1	C	T
1_Spain—EPI_ISL_418251	23402	1	A	G
1_Spain—EPI_ISL_418251 29780 122			AGCTGCCCTATATGGA AGAGCCCTAATGTGT AAAATTAAATTTAGT AGTGCTATCCCCATG TGATTTTAATAGCTT CTTAGGGAGAATGACA AAAAAAAAAAAAAAA AAAAAAAAAAAAAAA AA	_____
2_Spain—EPI_ISL_419240	8781	1	C	T
2_Spain—EPI_ISL_419240	9476	1	T	A
2_Spain—EPI_ISL_419240	14804	1	C	T
2_Spain—EPI_ISL_419240	25978	1	G	T
2_Spain—EPI_ISL_419240	28143	1	T	C
2_Spain—EPI_ISL_419240	28656	1	C	T
2_Spain—EPI_ISL_419240	28862	1	C	T
2_Spain—EPI_ISL_419240 29786 116			CTATATGGAAGAGCC CTAATGTGTAAAATT AATTTAGTAGTGCT ATCCCCATGTGATTT TAATAGCTTCTTAGG AGAATGACAAAAAAA AAAAAAAAAAAAAAA AAAAAAA	_____
3_Spain—EPI_ISL_419685	0	54	ATTAAAGGTTTATACCT TCCCAGGTAACAAACCA ACCAACTTCGATCTCTTGT	_____
3_Spain—EPI_ISL_419685	2000	1	G	A





			TAAACATACTGACTTT TCAAGTGAAATCATA GGATACAAGGCATT GATGGTGGTGTCACT CGTGACATAGCATCT ACAGATACTTGTTTT GCTAACAAACATGCT GATTTGACACATGG TTTAGCCAGCGTGGT GGTAGTTATACTAA	NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN
4_Spain—EPI_ISL_436301	8655	150	ACAAAGCTTGCCCCAT TGATTGCTGCAGTC TAACAAGAGAAGTGG GTTTGTCTGCCTG GTTTGCCTGGCACG	NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN
4_Spain—EPI_ISL_436301	8807	74	GCACA	NNNN
4_Spain—EPI_ISL_436301	8888	5	C	T
4_Spain—EPI_ISL_436301	14407	1	T	C
4_Spain—EPI_ISL_436301	17125	1	A	G
4_Spain—EPI_ISL_436301	20267	1	A	G
4_Spain—EPI_ISL_436301	23402	1	A	W
4_Spain—EPI_ISL_436301	27577	1	G	K
4_Spain—EPI_ISL_436301	29850	52	AGCTTCTTAGGAGAA TGACAAAAAAA AAAAAAA AAAAAA	_____
5_Spain—EPI_ISL_436324	0	47	ATTAAGGTTTATAC CTTCCCAGGTAA ACCAACCAACTTCG AT	_____
5_Spain—EPI_ISL_436324	240	1	C	T
5_Spain—EPI_ISL_436324	3036	1	C	T
5_Spain—EPI_ISL_436324	13379	227	TCTGCGGTATGTGGAA AGGTTATGGCTGTAG TTGTGATCAACTCCG CGAACCCATGCTTCA GTCAGCTGATGCACA ATCGTTTTAACGG GTTTGCCTGTAAAGT GCAGCCCGTCTTACA CCGTGCGGCACAGGC ACTAGTACTGATGTC GTATACAGGGCTTT GACATCTACAATGAT AAAGTAGCTGGTTTT GCTAAATTCTAAAA ACTAATTGTTGTCGCT	NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN
5_Spain—EPI_ISL_436324	14407	1	C	T
5_Spain—EPI_ISL_436324	22403	1	A	W
5_Spain—EPI_ISL_436324	23402	1	A	G
5_Spain—EPI_ISL_436324	29143	1	C	T

5_Spain—EPI_ISL_436324	29850	52	AGCTTCTTAGGAGAA TGACAAAAAAAAAAAAA AAAAAAAAAAAAAAA AAAAAAA	
6_Spain—EPI_ISL_444976	240	1	C	T
6_Spain—EPI_ISL_444976	3036	1	C	T
6_Spain—EPI_ISL_444976	8738	1	C	T
6_Spain—EPI_ISL_444976	14407	1	C	T
6_Spain—EPI_ISL_444976	18743	1	C	T
6_Spain—EPI_ISL_444976	20267	1	A	G
6_Spain—EPI_ISL_444976	22892	1	A	G
6_Spain—EPI_ISL_444976	23402	1	A	G
7_Spain—EPI_ISL_444975	240	1	C	T
7_Spain—EPI_ISL_444975	3036	1	C	T
7_Spain—EPI_ISL_444975	5175	1	A	G
7_Spain—EPI_ISL_444975	6353	1	C	T
7_Spain—EPI_ISL_444975	14407	1	C	T
7_Spain—EPI_ISL_444975	20267	1	A	G
7_Spain—EPI_ISL_444975	23402	1	A	G
7_Spain—EPI_ISL_444975	28882	1	G	A
7_Spain—EPI_ISL_444975	29687	1	G	C
7_Spain—EPI_ISL_444975	29733	1	G	C

## Appendice B - risultati confronto orizzontale

### Clustal Omega

ID_SEQ	POS	LENGTH	REFERENCE	MUTATION
Italy—EPI_ISL_436729	0	9	ATTAAGGT	NNNNNNNNN
Italy—EPI_ISL_436729	30	1	A	G
Italy—EPI_ISL_436729	240	1	C	T
Italy—EPI_ISL_436729	3036	1	C	T
Italy—EPI_ISL_436729	3044	1	C	T
Italy—EPI_ISL_436729	6448	1	C	T
Italy—EPI_ISL_436729	6862	1	A	M
Italy—EPI_ISL_436729	6865	1	A	W
Italy—EPI_ISL_436729	6868	1	A	W
Italy—EPI_ISL_436729	14407	1	C	T
Italy—EPI_ISL_436729	19676	1	G	R
Italy—EPI_ISL_436729	21626	1	C	Y
Italy—EPI_ISL_436729	23402	1	A	G
Italy—EPI_ISL_436729	25458	1	G	K
Italy—EPI_ISL_436729	28880	3	GGG	AAC
Italy—EPI_ISL_436729	29872	30	AAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	—————
Russia—EPI_ISL_430108	240	1	C	T
Russia—EPI_ISL_430108	834	1	C	T
Russia—EPI_ISL_430108	3036	1	C	T
Russia—EPI_ISL_430108	14407	1	C	T
Russia—EPI_ISL_430108	20267	1	A	G
Russia—EPI_ISL_430108	23402	1	A	G
Russia—EPI_ISL_430108	29870	32	AAAAAAAAAAAAAAAAAAA AAAAAAAAAAA	NNNNNNNNNNNNNNNNNN NNNNNNNNNN
Spain—EPI_ISL_444975	240	1	C	T
Spain—EPI_ISL_444975	3036	1	C	T
Spain—EPI_ISL_444975	5175	1	A	G
Spain—EPI_ISL_444975	6353	1	C	T
Spain—EPI_ISL_444975	14407	1	C	T
Spain—EPI_ISL_444975	20267	1	A	G
Spain—EPI_ISL_444975	23402	1	A	G
Spain—EPI_ISL_444975	28882	1	G	A
Spain—EPI_ISL_444975	29687	1	G	C
Spain—EPI_ISL_444975	29733	1	G	C

USA—EPI_ISL_435547	0	136	ATTAAGGTTATACCTTCC CAGGTAAACAAACCAACCAA CTTCGATCTCTTAGATCT GTTCTCTAACGAACTTAAA ATCTGTGGCTGTCACTCGG CTGCATGCTTAGTGCAC GCAGTATAATTAA	NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN
USA—EPI_ISL_435547	145	3	ACT	NNN
USA—EPI_ISL_435547	240	1	C	T
USA—EPI_ISL_435547	1058	1	C	T
USA—EPI_ISL_435547	3036	1	C	T
USA—EPI_ISL_435547	9113	102	CTGACACACGTTATGTGCTC ATGGATGGCTCTATTATTCAA TTTCCTAACACCTACCTTGAA GGTTCTGTAGAGTGGTAACA ACTTTGATTCTGAGTACT	NNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN NN
USA—EPI_ISL_435547	11915	1	C	T
USA—EPI_ISL_435547	14407	1	C	T
USA—EPI_ISL_435547	18997	1	C	T
USA—EPI_ISL_435547	23402	1	A	G
USA—EPI_ISL_435547	25562	1	G	T
USA—EPI_ISL_435547	27058	1	C	T
USA—EPI_ISL_435547	29539	1	G	A
USA—EPI_ISL_435547	29901	1	A	N

Kalign

USA—EPI_ISL_435547	145	3	ACT	NNN
USA—EPI_ISL_435547	240	1	C	T
USA—EPI_ISL_435547	1058	1	C	T
USA—EPI_ISL_435547	3036	1	C	T
USA—EPI_ISL_435547	9113	102	CTGACACACGTTATGTGCTC ATGGATGGCTCTATTATTCAA TTTCCTAACACCTACCTTGAA GGTTCTGTAGAGTGGTAACA ACTTTGATTCTGAGTACT	NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN NN
USA—EPI_ISL_435547	11915	1	C	T
USA—EPI_ISL_435547	14407	1	C	T
USA—EPI_ISL_435547	18997	1	C	T
USA—EPI_ISL_435547	23402	1	A	G
USA—EPI_ISL_435547	25562	1	G	T
USA—EPI_ISL_435547	27058	1	C	T
USA—EPI_ISL_435547	29539	1	G	A
USA—EPI_ISL_435547	29901	1	A	N

## MAFFT

ID_SEQ	POS	LENGTH	REFERENCE	MUTATION
Italy—EPI_ISL_436729	0	9	ATTTAAAGGT	NNNNNNNNN
Italy—EPI_ISL_436729	30	1	A	G
Italy—EPI_ISL_436729	240	1	C	T
Italy—EPI_ISL_436729	3036	1	C	T
Italy—EPI_ISL_436729	3044	1	C	T
Italy—EPI_ISL_436729	6448	1	C	T
Italy—EPI_ISL_436729	6862	1	A	M
Italy—EPI_ISL_436729	6865	1	A	W
Italy—EPI_ISL_436729	6868	1	A	W
Italy—EPI_ISL_436729	14407	1	C	T
Italy—EPI_ISL_436729	19676	1	G	R
Italy—EPI_ISL_436729	21626	1	C	Y
Italy—EPI_ISL_436729	23402	1	A	G
Italy—EPI_ISL_436729	25458	1	G	K
Italy—EPI_ISL_436729	28880	3	GGG	AAC
Italy—EPI_ISL_436729	29872	30	AAAAAAAAAAAAAAA AAAAAAAAAAAAAAA	————— —————
Russia—EPI_ISL_430108	240	1	C	T
Russia—EPI_ISL_430108	834	1	C	T
Russia—EPI_ISL_430108	3036	1	C	T
Russia—EPI_ISL_430108	14407	1	C	T
Russia—EPI_ISL_430108	20267	1	A	G
Russia—EPI_ISL_430108	23402	1	A	G
Russia—EPI_ISL_430108	29870	32	AAAAAAAAAAAAAAA AAAAAAAAAAAAAAA	NNNNNNNNNNNNNN NNNNNNNNNNNNNNNN
Spain—EPI_ISL_444975	240	1	C	T
Spain—EPI_ISL_444975	3036	1	C	T
Spain—EPI_ISL_444975	5175	1	A	G
Spain—EPI_ISL_444975	6353	1	C	T
Spain—EPI_ISL_444975	14407	1	C	T
Spain—EPI_ISL_444975	20267	1	A	G
Spain—EPI_ISL_444975	23402	1	A	G
Spain—EPI_ISL_444975	28882	1	G	A
Spain—EPI_ISL_444975	29687	1	G	C
Spain—EPI_ISL_444975	29733	1	G	C

			ATTAAGGTTTAC CTTCCCAGGTACAA ACCAACCAACTTCG ATCTCTTGTAGATCTG TTCTCTAACGAACCTT TAAAATCTGTGTGGCT GTCACTCGGCTGCATG CTTAGTGCACTCACGC AGTATAATTAA	NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN
USA—EPI_ISL_435547	0	136	ACT	NNN
USA—EPI_ISL_435547	145	3	C	T
USA—EPI_ISL_435547	240	1	C	T
USA—EPI_ISL_435547	1058	1	C	T
USA—EPI_ISL_435547	3036	1	C	T
			CTGACACACGTTATG TGCTCATGGATGGCT CTATTATTCAATTCC TAACACCTACCTTGA AGGTTCTGTTAGAGT GGTAACAACCTTTGA TTCTGAGTACT	NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN
USA—EPI_ISL_435547	9113	102	C	T
USA—EPI_ISL_435547	11915	1	C	T
USA—EPI_ISL_435547	14407	1	C	T
USA—EPI_ISL_435547	18997	1	C	T
USA—EPI_ISL_435547	23402	1	A	G
USA—EPI_ISL_435547	25562	1	G	T
USA—EPI_ISL_435547	27058	1	C	T
USA—EPI_ISL_435547	29539	1	G	A
USA—EPI_ISL_435547	29901	1	A	N