



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

SARS-CoV-2	2
Obiettivo del progetto	3
Sequenze	3
Tool	3
Analisi verticale	4
New York (USA)	4
Russia	4
Italia	4
Spagna	4
Analisi verticale	4
Appendice A - risultati estesi	5
New York (USA)	6
Clustal Omega	6
Kalign	9
MAFFT	12
Russia	14
Clustal Omega	14
Kalign	17
MAFFT	19
Italia	21
Clustal Omega	21
Kalign	23
MAFFT	25
Spagna	27
Clustal Omega	27
Kalign	31
MAFFT	35

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.

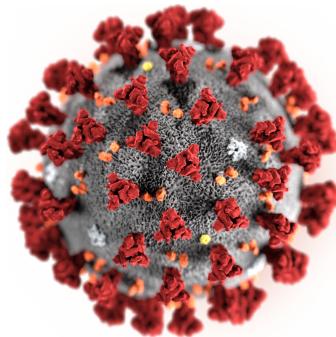


Figure 1: *L'illustrazione, creata dai Centers for Disease Control and Prevention (CDC) statunitense, rivela la morfologia ultrastrutturale del SARS-CoV-2*

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:

Stati Uniti	1,678.477
Russia	353,427
Spagna	235,823
Italia	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

New York (USA)

Russia

Italia

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

Appendice A - risultati estesi

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 TTAGGAGAACGACAAAAAA 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 ATAGCTTCTTAGGAGAACGACA 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
					TGAACAATGCTAGGGAGAGC TGCCTATATGGAAGAGCCCTA ATGTGTAAAATTAAATTAGTA GTGCTATCCCCATGTGATTTA ATAGCTTCTTAGGAGAACATGAC AAAAAAAAAAAAAAA AAAAAAAAAAAAAA	
4	USA	EPI_ISL_421594	29763	139		
5	USA	EPI_ISL_426309	0	25	ATTAAGGTTATACCTTCCCA 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	TAATAGCTCTTAGGAGAACATG 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	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	ATTAAGGTTATACCTCCC AGGTAACAAACCAACCAACT TTCGATCTCTGTAGATCTGT TCTCTAACGAACCTTAAAAT CTGTGTGGCTGTCACTCGGCT GCATGCTTAGTGCACTCACGC AGTATAATTAA	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	CTGACACACGTTATGTGCTCATG GATGGCTCTATTATTCAATTCCCT AACACCTACCTTGAAGGTTCTGT TAGAGTGGTAACAACCTTTGATT CTGAGTACT	NNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN
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	ATTAAAGGTTTACCTTCCCAG 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	CCATGTGATTTAACAGCTCTTAG GAGAATGACAAAAAAAAAAAAAA AAAAAAAAAAAAAAA	—————
3_USA—EPI_ISL_421365	0	55	ATTAAAGGTTTACCTTCCCAGGT ACAAACCAACCAACTTCGATCTCTT 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	AGCTGCCTATGGAAAGAGCC CTAATGTGAAATTAAATTAA GTAGTGCTATCCCCATGTGATT TTAATAGCTTCTAGGAGAATG ACAAAAAAA AAAAAAA	—————
4_USA—EPI_ISL_421594	0	55	ATTAAAGGTTTACCTTCCC GGTAACAAACCAACCAACTTT CGATCTCTGTA	—————
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
			TGAACAATGCTAGGGAGAGCTG CCTATATGGAAGAGGCCCTAATGT GTAAAATTAAATTAGTAGTGCTA TCCCCATGTGATTAAATAGCTTC TTAGGAGAATGACAAAAAAA AAAAAAAAAAAAAAA	_____
4_USA—EPI_ISL_421594	29763	139		_____
5_USA—EPI_ISL_426309	0	25	ATTAAGGTTATACCTCCAGGT	_____
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	TAATAGCTCTTAGGAGAAC CAAAAAAAA AAAAAAA	_____
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
			ATTAAGGTTATAACCTCCC AGGTAACAAACCAACCAACT TTCGATCTCTGTAGATCTGTT CTCTAACGAACCTTAAAATC TGTGTGGCTGTCACTCGGCTG CATGCTTAGTGCACTCACGCA GTATAATTAA	NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN
9_USA—EPI_ISL_435547	0	136		
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 TTTCCTAACACACCTACCTGAA GGTTCTGTTAGAGTGGTAACA ACTTTGATTCTGAGTACT	NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN
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	ATTAAGGTTTACCTCCCA 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	AAAAAAAAAAAAAA	————
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	AAAAAAA 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	AAAAAAA 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	GACAAAAAAAAAAAAAAA AAAAA AAAAA	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 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 AAAAAAA	NNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNN

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	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	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

Kalign

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	AGCTGCCTATATGGAAGAGC	_____
			CCTAATGTGTAAAATTAATTT	_____
			TAGTAGTGCTATCCCCATGTG	_____
			ATTTTAATAGCTTCTTAGGAG	_____
			AATGACAAAAAA AAAAAAAAAAAA	_____
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	CTATATGGAAGAGCCCTAAT	_____
			GTGTAAAATTAATTTAGTAG	_____
			TGCTATCCCCATGTGATTAA	_____
			ATAGCTTCTTAGGAGAATGAC	_____
			AAAAAAAAAAAA AAAAAAAAAAAA	_____
3_Spain—EPI_ISL_419685	0	54	ATTAAAGGTTATACCTTCC	_____
			CAGGTAACAAACCAACCAA	_____
			CTTCGATCTCTTGT	_____
3_Spain—EPI_ISL_419685	2000	1	G	A
3_Spain—EPI_ISL_419685	5296	290	ACTGCATTGTTAACACTCCA	NNNNNNNNNNNNNNNNNNNN
			ACAAATAGAGTGAAGTTA	NNNNNNNNNNNNNNNNNNNN
			ATCCACCTGCTCTACAAGATG	NNNNNNNNNNNNNNNNNNNN
			CTTATTACAGAGCAAGGGCTG	NNNNNNNNNNNNNNNNNNNN
			GTGAAGCTGCTAACTTTGTG	NNNNNNNNNNNNNNNNNNNN
			CACTTATCTTAGCCTACTGTAA	NNNNNNNNNNNNNNNNNNNN
			TAAGACAGTAGGTGAGTTAGG	NNNNNNNNNNNNNNNNNNNN
			TGATGTTAGAGAAACAATGAGT	NNNNNNNNNNNNNNNNNNNN
			TACTTGTTCACCATGCCAATT	NNNNNNNNNNNNNNNNNNNN
			AGATTCTGCAAAAGAGTCTTG	NNNNNNNNNNNNNNNNNNNN
3_Spain—EPI_ISL_419685	8781	1	AACGTGGTGTGAAAACCTGTG	NNNNNNNNNNNNNNNNNNNN
			GACAAACAGCAGACAACCCTAA	NNNNNNNNNNNNNNNNNNNN
3_Spain—EPI_ISL_419685	9476	1	GGGTGTAGAAGCTGTATGTACA	NNNNNNNNNNNNNNNNNNNN
			TGGGCACACT	NNNNNNNNNNNNNNNNNN

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	MUTATI
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	AGCTGCCTATATGGAAGAGC	
			CCTAATGTGTAAAATTAAATT	
2_Spain—EPI_ISL_419240	8781	1	TAGTAGTGCTATCCCCATGTG	
			ATTTTAATAGCTTCTAGGAG	
2_Spain—EPI_ISL_419240	9476	1	AATGACAAAAAA	
			AAAAAAAAAAAAAA	
2_Spain—EPI_ISL_419240	14804	1	A	
			2_Spain—EPI_ISL_419240	
2_Spain—EPI_ISL_419240	25978	1	G	
			2_Spain—EPI_ISL_419240	
2_Spain—EPI_ISL_419240	28143	1	T	
			2_Spain—EPI_ISL_419240	
2_Spain—EPI_ISL_419240	28656	1	C	
			2_Spain—EPI_ISL_419240	
2_Spain—EPI_ISL_419240	28862	1	T	
			2_Spain—EPI_ISL_419240	
2_Spain—EPI_ISL_419240	29786	116	CTATATGGAAGAGCCCTAAT	
			GTTGAAAATTAAATTAGTAG	
3_Spain—EPI_ISL_419685	0	54	TGCTATCCCCATGTGATTAA	
			TAGCTTCTTAGGAGAACATGACA	
3_Spain—EPI_ISL_419685	2000	1	AAAAAAAAAAAAAA	
			AAAAAAAAAA	
3_Spain—EPI_ISL_419685	5296	290	ATTAAAGGTTTATAACCTCCCA	
			GGTAACAAACCAACCAACTTT	
3_Spain—EPI_ISL_419685	8781	1	CGATCTCTTGT	
			ACTGCATTGTTAACACTCCAA	NNNNNNNNNNNN
3_Spain—EPI_ISL_419685	9476	1	CAAATAGAGTTGAAGTTAAC	NNNNNNNNNNNN
			CACCTGCTCTACAAGATGCTTA	NNNNNNNNNNNN
3_Spain—EPI_ISL_419685	11243	1	TTACAGAGCAAGGGCTGGTGA	NNNNNNNNNNNN
			GCTGCTAACCTTGTGCACTTAT	NNNNNNNNNNNN
3_Spain—EPI_ISL_419685	14407	1	CTTAGCCTACTGTAATAAGACA	NNNNNNNNNNNN
			GTTAGGTGAGTTAGGTGATGTTA	NNNNNNNNNNNN
3_Spain—EPI_ISL_419685	21262	1	GAGAAACAAATGAGTTACTTGT	NNNNNNNNNNNN
			TCAACATGCCAATTAGATTCTT	NNNNNNNNNNNN
3_Spain—EPI_ISL_419685	23402	1	GCAAAAGAGTCTGAACGTGGT	NNNNNNNNNNNN
			GTGTAAAACCTGTGGACAAACAG	NNNNNNNNNNNN
3_Spain—EPI_ISL_419685	28143	1	CAGACAACCCTTAAGGGTGTAG	NNNNNNNNNNNN
			AAGCTGTTATGTACATGGGCACAC	NNNNNNNNNNNN
3_Spain—EPI_ISL_419685	28656	1	CT	NNNNNNNN
			3_Spain—EPI_ISL_419685	
3_Spain—EPI_ISL_419685	28862	1	C	T
			3_Spain—EPI_ISL_419685	
3_Spain—EPI_ISL_419685	29786	116	T	A
			3_Spain—EPI_ISL_419685	

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