



UNIVERSITÀ POLITECNICA DELLE MARCHE
DIPARTIMENTO SCIENZE DELLA VITA E DELL'AMBIENTE

**Corso di Laurea Magistrale
Biologia Molecolare e Applicata**

**Utilizzo della tecnologia NGS per
l'identificazione e il tracciamento delle
varianti SARS-CoV-2**

Use of NGS technology for the identification
and tracking of SARS-CoV-2 variants

Tesi di Laurea Magistrale di:

Filippo Mantellini Faieta

Relatore:

Chiar.mo Prof. Anna La Teana

Correlatore:

Chiar.mo Prof. Liborio Stuppia

Sessione estiva 2023
Anno Accademico 2022-2023

Introduzione

Nel gennaio 2020, un nuovo betacoronavirus, successivamente designato SARS-CoV-2, è stato identificato come agente eziologico di un gruppo di casi di polmonite nella città di Wuhan, provincia di Hubei, Cina. Il COVID 19, malattia causata da questo nuovo patogeno, si è diffusa rapidamente e l'11 marzo 2020, con 118.000 casi segnalati da 110 paesi, l'Organizzazione Mondiale della Sanità (OMS) ha dichiarato la pandemia.

Il sequenziamento dell'intero genoma (WGS) di SARS-CoV-2 ha svolto un ruolo importante dall'inizio della pandemia. In particolare, il WGS ha contribuito a progettare specifiche RT-PCR, strategie antivirali e vaccini. Il sequenziamento di SARS-CoV-2 ha anche permesso di esplorare la trasmissione del lignaggio e si è rivelato utile per valutare l'efficacia delle misure di intervento. Inoltre, la sorveglianza genomica di SARS-CoV-2 ha consentito la caratterizzazione di mutazioni e la valutazione dell'impatto sulla virulenza e sulla trasmissibilità. Il WGS di SARS-CoV-2 viene eseguito a una velocità senza precedenti in tutto il mondo con l'uso di metodi di sequenziamento di nuova generazione (NGS) molto diversi.

La recente esperienza con malattie infettive emergenti, come SARS, MERS, Zika ed Ebola, ha dimostrato che le tecnologie NGS rappresentano potenti strumenti per rintracciare origini, diffusione e catene di trasmissione dei focolai, nonché per monitorare l'evoluzione degli agenti eziologici. Di conseguenza, la pandemia di COVID-19 ha innescato sforzi senza precedenti per lo sviluppo di efficaci strategie di sorveglianza in tempo reale basate sul sequenziamento del genoma del suo agente causale con numerosi genomi SARS-CoV-2 depositati in repository dedicate. Questi dati hanno contribuito a studi sulla dinamica evolutiva del virus e all'identificazione di varianti di potenziale rilevanza clinica.

1. Virus: struttura e classificazione

1.1 Caratteristiche comuni dei virus

Tutti i virus condividono molte caratteristiche comuni: sono entità molto piccole, sono parassiti intracellulari obbligati e hanno un genoma peculiare.

Il più piccolo dei virus ha un diametro di circa 20 nm, sebbene l'influenza e il virus dell'immunodeficienza umana abbiano una dimensione più tipica, circa 100 nm di diametro. Le cellule umane medie hanno un diametro di 10-30 μm (micron), il che significa che sono generalmente da 100 a 1000 volte più grandi dei virus che le stanno infettando. Tuttavia, alcuni virus sono significativamente più grandi di 100 nm. I poxvirus, come il virus vaiolo che causa il vaiolo, possono avvicinarsi a 400 nm di lunghezza, e i filovirus, come il pericoloso virus Ebola e il virus Marburg, hanno un diametro di soli 80 nm ma si estendono in lunghi filamenti che possono raggiungere lunghezze superiori a 1000 nm. Recentemente sono stati scoperti diversi virus molto grandi che infettano le amebe: il megavirus ha un diametro di 400 nm e i pandoravirus hanno una struttura ellittica o ovoidale che si avvicina a 1000 nm di lunghezza. È un errore comune pensare che tutti i virus siano più piccoli dei batteri; la maggior parte dei batteri ha in genere una dimensione di 2000-3000 nm, ma alcuni ceppi di batteri chiamati Micobatteri possono essere 10 volte più piccoli e avvicinarsi alla dimensione di questi grandi virus (**Fig. 1**).

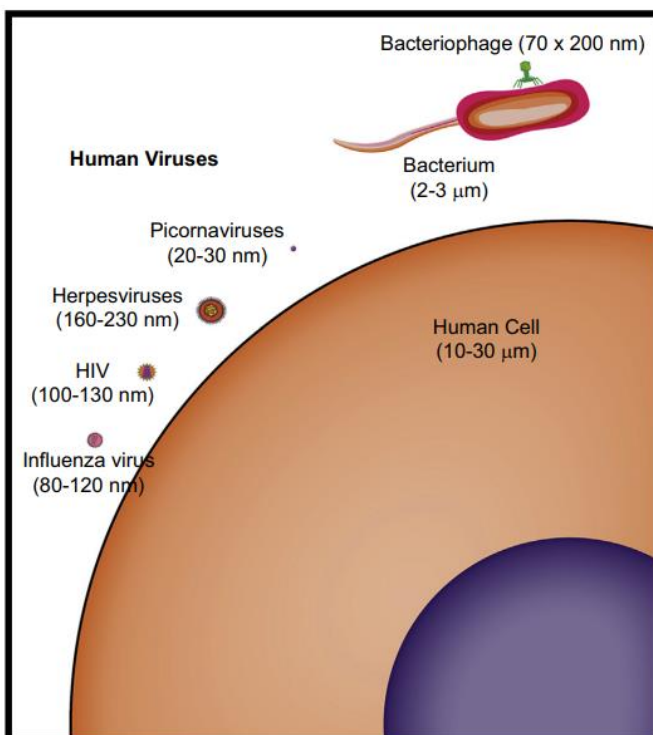


Fig. 1 Confronto dimensionale di virus e cellule. I virus umani possono variare di dimensioni, ma generalmente hanno un diametro compreso tra 20 e 200 nm. In confronto, i batteri hanno generalmente una lunghezza di 2-3 μm e una cellula umana media è di 10-30 μm . Louten J., *Essential Human Virology*, Academic Press, 2017

Inoltre, i virus sono parassiti intracellulari obbligati. Tutti i virus entrano in contatto con la superficie di una cellula e al suo interno si disassemblano e il loro materiale genetico codifica le istruzioni per le proteine che si assembleranno spontaneamente nei nuovi virioni. A differenza delle cellule, che crescono di dimensioni e si dividono equamente in due per replicarsi, i virus utilizzano l'energia e i

macchinari della cellula per creare e assemblare nuovi virioni pezzo per pezzo, completamente da zero ("replica de novo").

Il materiale genetico dei virus può essere composto da DNA o RNA. I genomi non sono necessariamente a doppio filamento; diversi tipi di virus possono anche avere genomi di DNA a filamento singolo (ssDNA) e i virus con genomi di RNA possono essere a filamento singolo o doppio. Analogamente alla dimensione della particella virale, anche quella del genoma può variare notevolmente da virus a virus. Un tipico genoma virale cade nell'intervallo di 7000-20.000 paia di basi (bp) (7-20 kilobasi). I virioni più piccoli saranno naturalmente in grado di contenere meno acido nucleico rispetto ai virioni più grandi, ma grandi virus non hanno necessariamente grandi genomi. Mentre la maggior parte dei virus non contiene molto acido nucleico, alcuni virus a dsDNA hanno genomi molto grandi: gli herpesvirus hanno genomi di 120-200 kb in totale e i pandoravirus hanno i genomi più grandi: fino a 2,5 milioni di basi, che rivaleggiano con le dimensioni del genoma di molti batteri. In confronto, le cellule eucariotiche hanno genomi molto più grandi: un'alga rossa ha il genoma eucariotico più piccolo conosciuto, 8 milioni coppie di basi; una cellula umana contiene oltre 3 miliardi di nucleotidi nel suo materiale ereditario; il genoma più grande mai sequenziato, di oltre 22 miliardi di paia di basi, è quello del pino taeda.

1.2 Struttura dei virus

La particella virale infettiva deve essere rilasciata dalla cellula ospite per infettare altre cellule e individui. Che si tratti di dsDNA, ssDNA, dsRNA o ssRNA, il genoma del virus deve essere protetto durante il processo. Nell'ambiente extracellulare, il virus sarà esposto a enzimi che potrebbero degradare l'acido nucleico. Le sollecitazioni fisiche, come il flusso di aria o liquido, potrebbero anche tagliare in pezzi i filamenti di acido nucleico. Inoltre, i genomi virali sono suscettibili ai danni causati dalle radiazioni ultraviolette o dalla radioattività, allo stesso modo del nostro DNA. Se il genoma del virus è danneggiato, non sarà in grado di produrre virioni.

Per proteggere il fragile acido nucleico da questo ambiente ostile, il virus lo circonda con un involucro proteico, chiamato capside (dal latino capsula, "scatola"). Il capside è composto da uno o più tipi diversi di proteine che si ripetono più e più volte per creare l'intera struttura. Questa struttura ripetitiva forma un capside forte ma leggermente flessibile. In combinazione con le sue piccole dimensioni, il capside è fisicamente molto difficile da rompere e protegge sufficientemente l'acido nucleico al suo interno. Insieme, l'acido nucleico e il capside formano il nucleocapside del virione.

Le proteine del capside si autoassemblano. I primi a dimostrarlo furono H. Fraenkel-Conrat e Robley Williams nel 1955. Separarono il genoma a RNA dalle subunità proteiche del virus del mosaico del tabacco e, quando le rimisero insieme in una provetta, si formarono automaticamente virioni infettivi. Ciò ha indicato che non sono necessarie ulteriori informazioni per assemblare un virus: i componenti fisici si assemblano spontaneamente, principalmente tenuti insieme da forze elettrostatiche e idrofobiche.

La maggior parte dei virus ha anche un envelope che circonda il capside. L'envelope è una membrana lipidica che deriva da una delle membrane cellulari, molto spesso la membrana plasmatica, sebbene possa anche provenire dal reticolo endoplasmatico della cellula, dal complesso di Golgi o persino dalla membrana nucleare, a seconda del virus. Questi virus hanno spesso proteine, chiamate

proteine della matrice, che funzionano per collegare l'envelope al capsido interno. Un virus privo di envelope è noto come virus senza envelope o nudo (**Fig. 2**). Ogni virus possiede anche una proteina di attacco del virus incorporata nel suo strato più esterno, nel capsido, nel caso di un virus nudo, o nell'envelope, nel caso di un virus con membrana. La proteina di attacco del virus è la proteina virale che facilita l'aggancio del virus alla membrana plasmatica della cellula ospite, il primo passo per ottenere l'ingresso in una cellula.

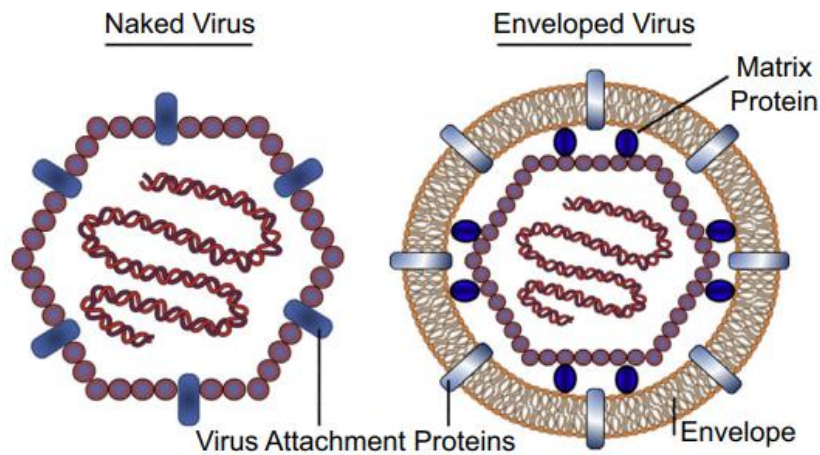


Fig. 2 Confronto tra un virione nudo e uno con envelope. Il capsido di un virione rivestito è avvolto da una membrana lipidica derivata dalla cellula. Le proteine di attacco del virus situate nel capsido o nell'involucro facilitano il legame del virus alla sua cellula ospite. Louten J., *Essential Human Virology*, Academic Press, 2017

1.3 Struttura del capsidico elicoidale

Ogni virus possiede un capsidico proteico per proteggere il suo genoma dall'ambiente ostile. I capsidi virali si presentano prevalentemente in due forme: elicoidale e icosaedrica. Nel caso di un virus elicoidale, l'acido nucleico virale si avvolge in una forma elicoidale e le proteine del capsidico si avvolgono all'interno o all'esterno dell'acido nucleico, formando un lungo tubo o una struttura a bastoncino (**Fig. 3**). Una volta nella cellula, il nucleocapsidico elicoidale si srotola e l'acido nucleico diventa accessibile.

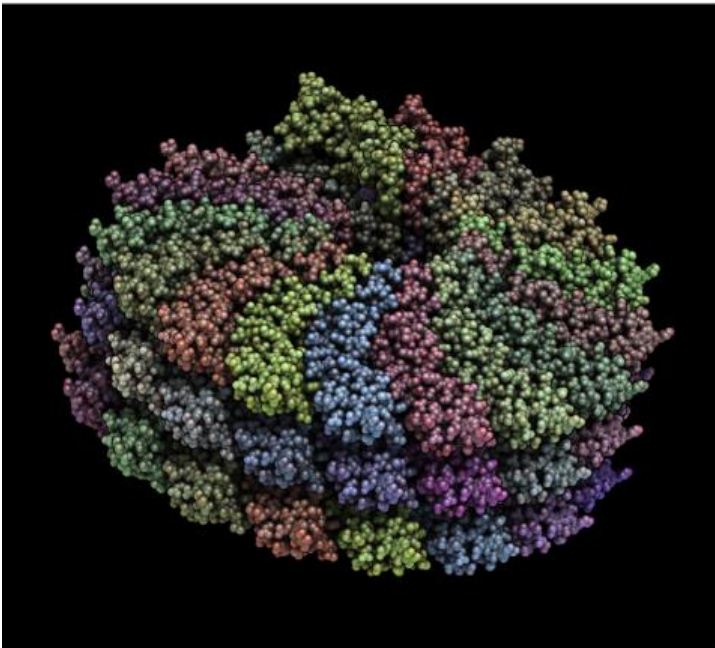


Fig. 3 **Struttura del capsidico elicoidale.** Struttura elicoidale del virus del mosaico del tabacco (TMV). Louten J., *Essential Human Virology*, Academic Press, 2017

Ci sono diversi vantaggi nella formazione di un capsidico elicoidale. Innanzitutto, è richiesto un solo tipo di proteina del capsidico. Questa struttura è semplice e richiede meno energia libera per essere assemblata rispetto a un capsidico composto da più proteine. Inoltre, avere una sola proteina nucleocapsidica significa che è richiesto un solo gene invece di diversi, riducendo così la lunghezza dell'acido nucleico richiesto. Poiché la struttura elicoidale può continuare indefinitamente, non ci sono nemmeno vincoli su quanto acido nucleico può essere impacchettato nel virione: la lunghezza del capsidico corrisponderà alla dimensione dell'acido nucleico avvolto.

I virus elicoidali possono essere dotati di envelope o nudi. Il primo virus descritto, il virus del mosaico del tabacco, è un virus a spirale nudo. In effetti, la maggior parte dei virus delle piante sono elicoidali ed è molto raro che un virus delle piante elicoidale sia dotato di envelope. Al contrario, tutti i virus animali elicoidali sono con membrana. Questi includono virus ben noti come il virus dell'influenza, il virus del morbillo, il virus della parotite, il virus della rabbia e il virus Ebola.

1.4 Struttura del capsidico icosaedrico

Delle due principali strutture del capsidico, l'icosaedro è di gran lunga più diffuso dell'architettura elicoidale. Rispetto a un virus elicoidale in cui le proteine del capsidico si avvolgono attorno all'acido nucleico, i genomi dei virus icosaedrici sono impacchettati completamente all'interno di un capsidico icosaedrico che funge da guscio proteico. Inizialmente si pensava che questi virus fossero sferici, ma i progressi nella microscopia elettronica e nella cristallografia a raggi X hanno rivelato che in realtà avevano una struttura icosaedrica.

Un icosaedro è una forma geometrica con 20 lati (o facce), ciascuno composto da un triangolo equilatero.

Le proteine virali formano ciascuna faccia del capsidico icosaedrico. Tuttavia, le proteine virali non sono triangolari e quindi una sola subunità proteica non è sufficiente per formare l'intera faccia. Pertanto, una faccia è formata da almeno tre subunità proteiche virali assemblate insieme. Queste possono essere tutte la stessa proteina o possono essere tre diverse proteine. Le subunità insieme formano quella che viene chiamata l'unità strutturale, che si ripete per formare il capsidico del virione.

Ma come possono alcuni virus formare capsidici icosaedrici molto grandi? La risposta è la ripetizione. L'unità strutturale può essere ripetuta più e più volte per formare un lato icosaedro più grande. Il numero di unità strutturali che crea ciascun lato è chiamato numero di triangolazione (T), poiché le unità strutturali formano la faccia triangolare dell'icosaedro. Aumentando il numero di unità strutturali identiche su ciascuna faccia, l'icosaedro può diventare progressivamente più grande senza richiedere la produzione di nuove proteine aggiuntive.

Le proteine che compongono l'unità strutturale possono formare strutture tridimensionali note come capsomeri che sono visibili in una micrografia elettronica. Nei virus icosaedrici, i capsomeri generalmente assumono la forma di pentoni (contenenti cinque unità) o esoni (contenenti sei unità) che formano un modello visibile sulla superficie dell'icosaedro.

I capsomeri sono unità morfologiche che derivano dall'interazione delle proteine all'interno delle unità strutturali ripetute.

La ricerca ha dimostrato che le proteine che formano la simmetria icosaedrica richiedono minori quantità di energia, rispetto ad altre strutture, e quindi questa struttura è evolutivamente favorita.

Molti virus che infettano gli animali sono icosaedrici, tra cui il papillomavirus umano, il rinovirus, il virus dell'epatite B e gli herpesvirus (**Fig. 4**). Come le loro controparti elicoidali, anche i virus icosaedrici possono essere nudi o con envelope. Il tipo di acido nucleico virale (dsDNA, ssDNA, dsRNA e ssRNA) non è correlato alla struttura del capsidico; i capsidici virali icosaedrici possono contenere qualsiasi tipo di acido nucleico, a seconda del virus.

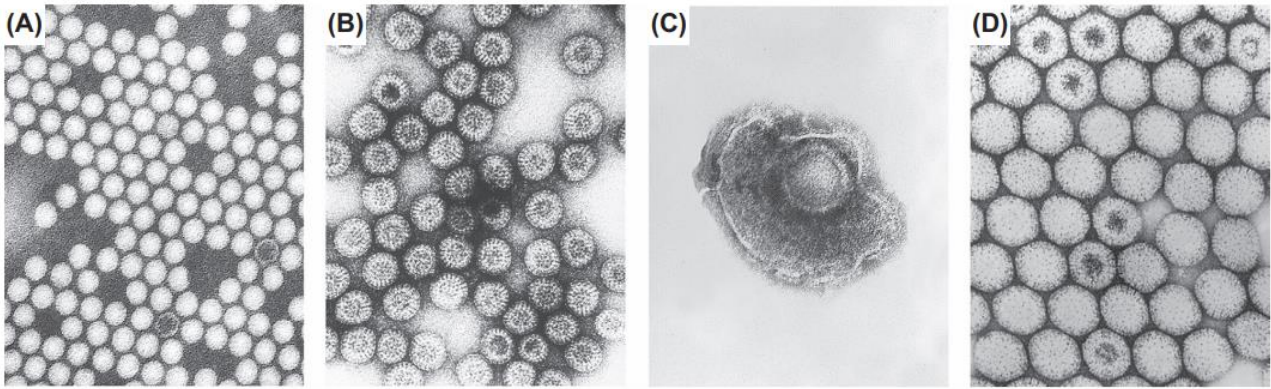


Fig. 4 Micrografie elettroniche di virus icosaedrici. Poliovirus (A), rotavirus (B), virus varicella-zoster (C), il virus che causa la varicella e l'herpes zoster, e reovirus (D). Si noti che C è rivestito. Louten J., *Essential Human Virology*, Academic Press, 2017

1.5 Strutture virali complesse

Alcuni virus hanno un'architettura complessa che non si conforma strettamente a una semplice forma elicoidale o icosaedrica. Poxvirus, geminivirus e molti batteriofagi sono esempi di virus con struttura complessa. I poxvirus, compresi i virus che causano il vaiolo o il vaiolo bovino, sono grandi particelle ovali o a forma di mattone lunghe 200-400 nm. All'interno del complesso virione, un nucleo a forma di manubrio racchiude il DNA virale ed è circondato da due "corpi laterali", la cui funzione è attualmente sconosciuta. Anche i geminivirus mostrano una struttura complessa. Come suggerisce il nome, questi virus che infettano le piante sono composti da due teste icosaedriche unite insieme. Anche molti batteriofagi hanno una struttura complessa, come il batteriofago P2, che ha una testa icosaedrica, contenente l'acido nucleico, attaccata a una guaina cilindrica della coda che facilita il legame del batteriofago alla cellula batterica.

1.6 Classificazione dei virus e tassonomia

La classificazione dei virus è utile poiché consente agli scienziati di contrastare i virus e di rivelare informazioni sui virus appena scoperti confrontandoli con quelli simili; consente, inoltre, agli scienziati di studiare l'origine dei virus e come si sono evoluti nel tempo. Attualmente esistono oltre 2800 diverse specie virali con proprietà molto diverse.

Uno schema di classificazione è stato sviluppato negli anni '70 dal premio Nobel David Baltimore, il quale classifica i virus in base al tipo di genoma e alla strategia di replicazione. Il sistema scompone anche i virus a RNA a filamento singolo in quelli che sono a filamento positivo (+) e a filamento negativo (-). Baltimore ha anche preso in considerazione i virus in grado di effettuare la trascrizione inversa, cosa che le cellule non sono in grado di fare.

Le sette classi sono:

- classe I: virus dsDNA
- classe II: virus ssDNA
- classe III: virus dsRNA
- classe IV: virus ssRNA a senso positivo
- classe V: virus ssRNA a senso negativo
- classe VI: virus a RNA a trascrizione inversa
- classe VII: virus a DNA a trascrizione inversa

Tuttavia, esistono diversi modi in cui i virus possono essere classificati, tra cui la dimensione del virione, la struttura del capsido, il tipo di acido nucleico, le proprietà fisiche, la specie ospite o la malattia causata. Il Comitato internazionale per la tassonomia dei virus (ICTV) dal 1966 è l'unico organismo incaricato di classificare i virus e utilizza la stessa gerarchia tassonomica utilizzata per classificare gli esseri viventi. I virus appartengono a un sistema completamente separato che non rientra nell'albero della vita e sono classificati solo utilizzando ordine, famiglia, genere e specie.

L'ICTV classifica i virus in base a una varietà di caratteristiche diverse con l'intenzione di classificare i virus più simili tra loro. Vengono considerate le proprietà chimiche e fisiche del virus, come il tipo di acido nucleico o il numero di diverse proteine codificate. Le tecnologie attuali ora ci consentono di sequenziare i genomi virali in modo relativamente rapido e semplice, consentendo agli scienziati di confrontare le sequenze di due virus per determinare quanto siano strettamente correlate. Vengono prese in considerazione anche altre proprietà del virione, tra cui la dimensione del virione, la forma del capsido e la presenza o meno di un involucro (**Fig. 5**).

Attualmente l'ICTV ha classificato sette ordini di virus che contengono un totale di 103 famiglie classificate al loro interno. Settantasette famiglie di virus, tuttavia, devono ancora essere assegnate a un ordine, inclusi virus importanti come retrovirus, papillomavirus e poxvirus. Sono stati proposti nuovi ordini ed è probabile che ne verranno creati altri man mano che il processo tassonomico continua.

Virus Taxa Infecting Vertebrates

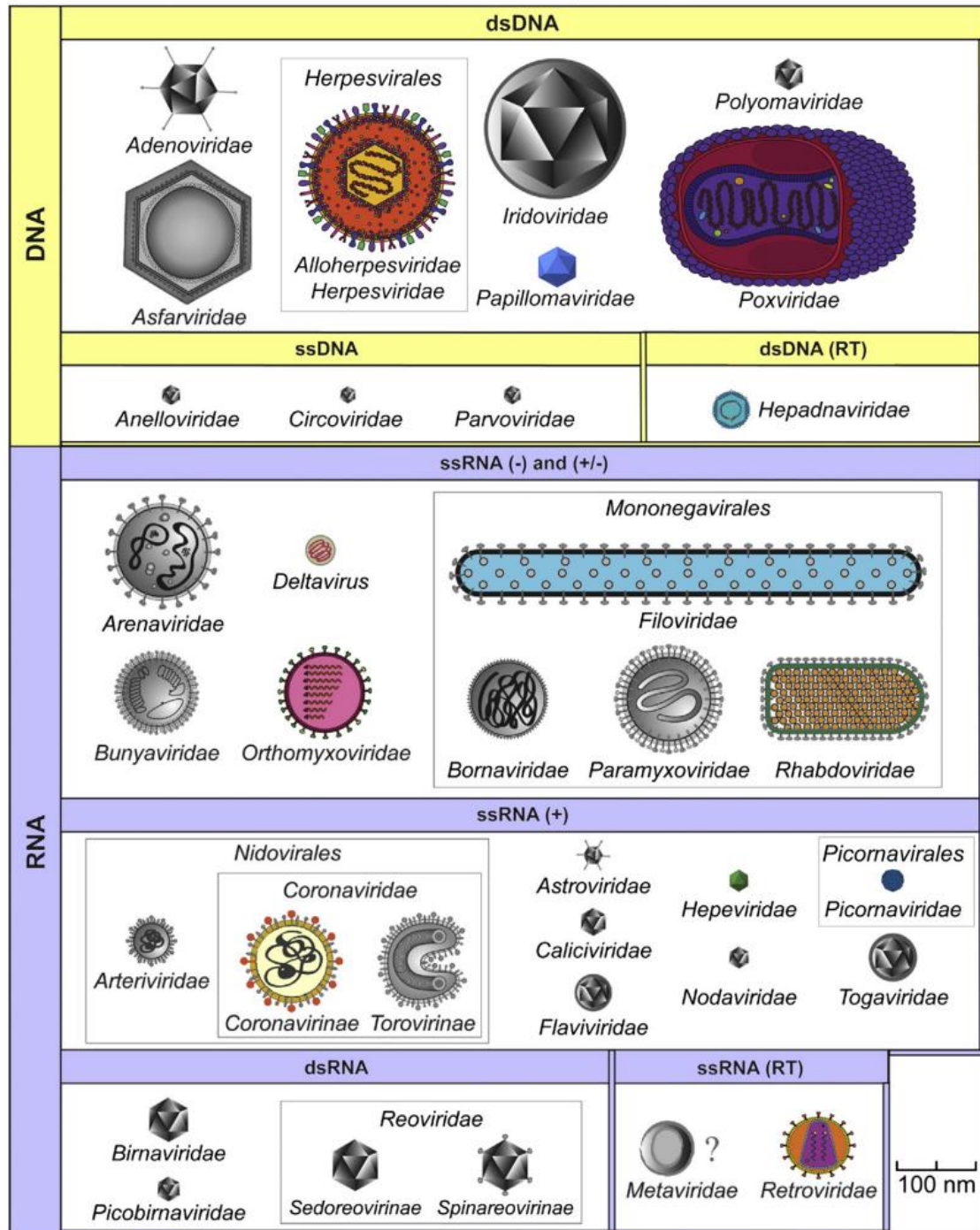


Fig. 5 Taxa di virus che infettano i vertebrati. I virus sono classificati in base al loro tipo di acido nucleico (virus a DNA in riquadri gialli e virus a RNA in riquadri blu) e ulteriormente classificati in base a caratteristiche distintive. Louten J., *Essential Human Virology*, Academic Press, 2017

2. Replicazione virale

Un virus deve subire il processo di replicazione per creare nuovi virioni infettivi in grado di infettare altre cellule del corpo o ospiti successivi. Dopo essere entrato nel corpo, un virus entra in contatto e attraversa la membrana plasmatica di una cellula bersaglio. Al suo interno, rilascia e replica il suo genoma facilitando la produzione delle sue proteine da parte dei ribosomi ospiti. Le particelle virali vengono assemblate da queste molecole biologiche di nuova sintesi e diventano virioni infettivi. Infine, i virioni vengono rilasciati dalla cellula per continuare il processo di infezione.

Le sette fasi della replicazione del virus sono classificate come segue:

1. Attacco
2. Penetrazione
3. *Uncoating*
4. Replicazione
5. Assemblaggio
6. Maturazione
7. Rilascio

Tutti i virus devono eseguire le sette fasi per creare nuovi virioni. Alcune fasi possono svolgersi contemporaneamente ad altre, oppure possono svolgersi non in ordine, a seconda del virus.

2.1 Attacco

Per infettare una cellula, è fondamentale che un virus sviluppi un legame con la cellula ospite. Questa interazione è specifica: il virus contiene una proteina che si adsorbe su un recettore della superficie cellulare. Le molecole del recettore bersaglio sulla superficie cellulare sono necessarie per le funzioni cellulari che i virus sfruttano, solitamente glicoproteine o residui di zucchero/carboidrati presenti sulle glicoproteine o sulla membrana plasmatica. Ad esempio, il rinovirus lega una proteina nota come molecola di adesione intercellulare 1 (ICAM-1), coinvolta nell'attaccamento di una cellula all'altra. I ceppi del virus dell'influenza A si legano agli zuccheri dell'acido sialico presenti alle estremità delle catene di carboidrati cellulari e i virus dell'herpes simplex (HSV) si legano in modo reversibile ai glicosaminoglicani (GAG), come l'eparan solfato, per legarsi alla proteina mediatrice dell'ingresso dell'herpesvirus o nectine sulla superficie cellulare (**Fig. 6**).

Alcuni virus richiedono anche corecettori per infettare le cellule. L'HIV inizialmente si lega a una proteina nota come CD4 sulla superficie dei linfociti T ("cellule T"), ma richiede uno dei due corecettori per continuare il processo di infezione. Gli esseri umani che hanno una versione modificata di CCR5, uno di questi corecettori, sono ampiamente resistenti all'infezione da HIV perché il virus non può utilizzare il CCR5 modificato come corecettore e quindi l'infezione è bloccata. L'infezione di una cellula può essere prevenuta se l'attaccamento del virus può essere inibito e le proteine di attacco del virus sono il bersaglio di molti farmaci antivirali in uso e in fase di sviluppo.

L'adsorbimento comporta forze elettrostatiche opposte sulla proteina del virus e sul recettore della superficie cellulare. La proteina di attacco del virus si trova nella porzione più esterna del virus, poiché è qui che avviene il contatto con la cellula. Essa sporge dall'involucro di un virus con envelope, mentre i virus nudi hanno una o più proteine del capsido che interagiscono con il recettore della

superficie cellulare. Le proteine di attacco virale possono estendersi dalla superficie del virione o possono trovarsi all'interno di "canyon" formati da proteine del capsido. Ad esempio, il 90% dei sierotipi di rhinovirus umani si legano all'ICAM-1 sulla superficie delle cellule. Invece di legarsi all'esterno del capsido del rinovirus, la molecola si innesta in un profondo canyon formato dalle proteine VP1, VP2 e VP3 del rinovirus. Al contrario, il 10% dei rinovirus umani si lega al recettore delle lipoproteine a densità molto bassa (VLDL). Tuttavia, questa interazione non si verifica nei canyon formati dalle proteine virali. Invece, diverse proteine VP1 ai vertici del capsido icosaedrico si legano al recettore. Anche se l'affinità di legame tra la proteina VP1 e il recettore VLDL è bassa, le molteplici proteine VP1 aumentano la forza di legame totale dell'interazione. Questo esempio illustra anche che diversi ceppi dello stesso virus possono sfruttare diversi recettori della superficie cellulare per l'attaccamento.

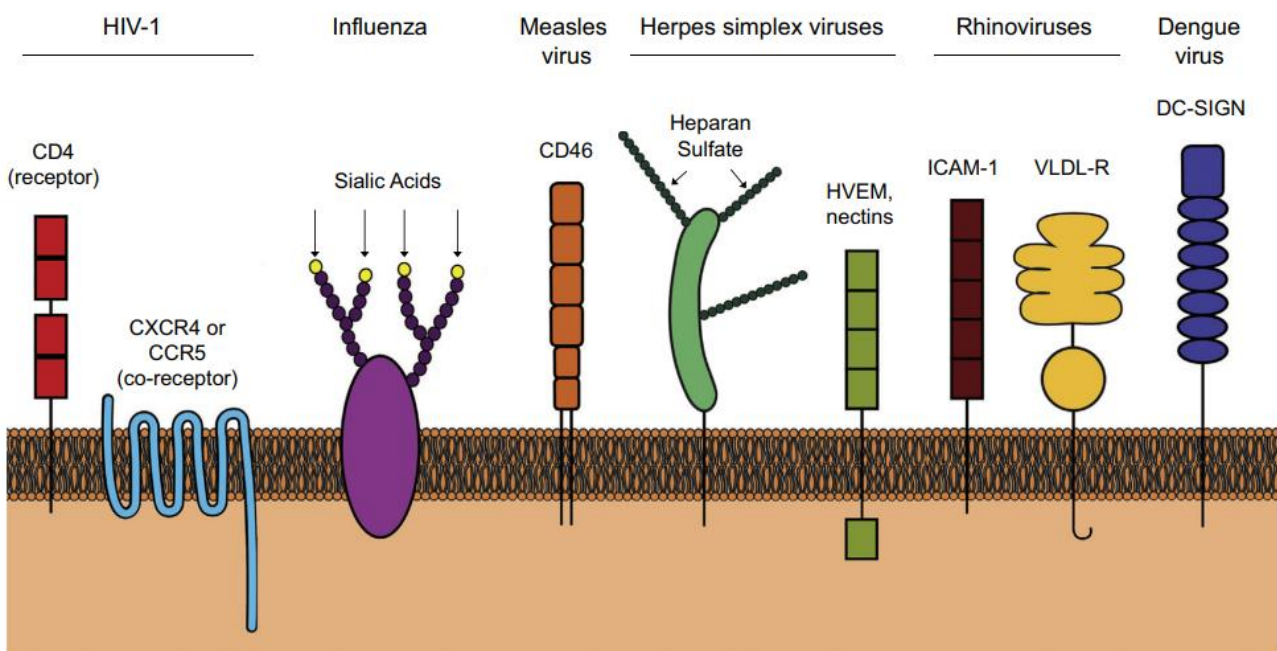


Fig. 6 Recettori sulla superficie cellulare. Diversi virus utilizzano recettori specifici sulla superficie cellulare per l'adsorbimento. L'HIV-1 richiede CD4 come recettore e recettori per le chemochine CCR5 o CXCR4 come corecettori. I virus dell'influenza si legano ai residui terminali di acido sialico presenti sulle glicoproteine della superficie cellulare. I ceppi di laboratorio del virus del morbillo legano il CD46 (sebbene il CD150 sia anche un recettore del virus). Il virus dell'herpes simplex-1 inizialmente si lega all'eparan solfato sui GAG per legare in modo specifico i recettori di ingresso, come HVEM o nectine. Il 90% dei rhinovirus utilizza ICAM-1 come recettore, mentre il 10% utilizza il recettore VLDL. Il virus dengue si attacca tramite DC-SIGN. Il tropismo del virus è determinato da quali cellule del corpo esprimono lo specifico recettore. Louten J., *Essential Human Virology*, Academic Press, 2017

2.2 Penetrazione

La penetrazione si riferisce all'attraversamento della membrana plasmatica da parte del virus. Contrariamente all'attaccamento del virus, la penetrazione richiede energia, sebbene questa sia fornita dalla cellula ospite, non dal virus (**Fig. 7**).

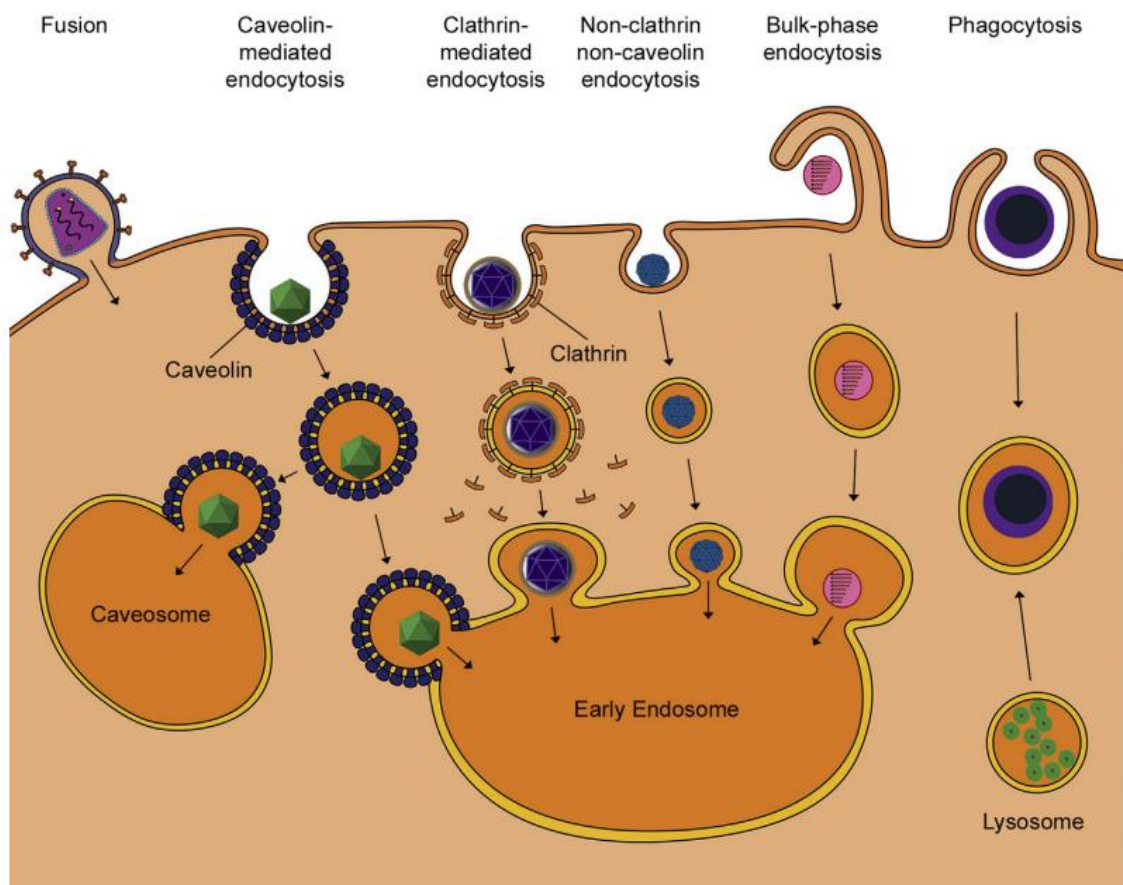


Fig. 7 Penetrazione virale nella cellula. Diversi virus sfruttano vari meccanismi cellulari per ottenere l'ingresso nella cellula dopo aver legato i loro specifici recettori sulla superficie cellulare. Alcuni virus rivestiti subiscono la fusione, che fonde l'involucro virale con la membrana plasmatica. Sia i virus rivestiti sia quelli nudi sfruttano l'endocitosi mediata da recettore nelle fosse rivestite di caveolina o clatrina per ottenere l'ingresso nel citoplasma della cellula. Altri virus ancora subiscono un'endocitosi mediata da recettori che è indipendente sia dalla clatrina che dalla caveolina. Anche l'endocitosi e la fagocitosi in fase *bulk* sono utilizzate dai virus per entrare nella cellula. Louten J., *Essential Human Virology*, Academic Press, 2017

L'endocitosi mediata da recettore è comunemente usata dai virus, sia con envelope che senza, per penetrare nella membrana plasmatica. Quando il pH dell'endosoma diminuisce, le proteine virali cambiano configurazione, il che consente loro di fuoriuscire dall'endosoma. A seconda del virus, ciò può accadere negli endosomi precoci, negli endosomi tardivi o nei lisosomi. La maggior parte dei tipi di virus utilizza l'endocitosi mediata dalla clatrina per entrare nella cellula, inclusi il virus della dengue, il virus dell'epatite C e i reovirus. Alcuni virus ben noti che infettano gli esseri umani, come SV40 e papillomavirus (che causano verruche o cancro cervicale), utilizzano l'endocitosi mediata da caveole; questo è stato scoperto utilizzando un farmaco che ha inibito la formazione di caveole. Il blocco dell'endocitosi mediata da clatrina non ha impedito l'ingresso di questi virus nelle cellule. Altri virus ancora subiscono un'endocitosi mediata da recettori che è indipendente sia dalla clatrina che dalla caveolina.

Anche altre forme di endocitosi, come l'endocitosi in fase *bulk* e la fagocitosi, vengono sfruttate dai virus per entrare nella cellula. Nell'endocitosi in fase *bulk*, la cellula forma una vescicola che inghiotte qualsiasi molecola sia presente nel fluido extracellulare, inclusi i virus. Recentemente, è stato dimostrato che due grandi virus a DNA, HSV-1 e mimivirus, entrano nelle cellule attraverso percorsi simili alla fagocitosi.

Un metodo di penetrazione utilizzato esclusivamente dai virus con envelope è la fusione. La fusione dell'involucro virale può avvenire sulla membrana cellulare o all'interno di vescicole endocitate, come l'endosoma, ed è mediata dalla stessa proteina virale utilizzata dal virus per l'attaccamento o da una diversa proteina virale, a seconda del virus. Ad esempio, l'HIV ha una proteina nota come gp120 che si lega a CD4 e uno dei due corecettori per l'ingresso, CCR5 o CXCR4. Una volta che ciò si verifica, una diversa proteina virale, gp41, fonde l'involucro del virus con la membrana cellulare, rilasciando il nucleocapside nel citoplasma.

2.3 Uncoating

L'*uncoating* si riferisce alla rottura o alla rimozione del capsido e può essere separata o strettamente collegata alla penetrazione. I virus ottengono la rimozione del capsido in una varietà di modi diversi (**Fig. 8**). Ad esempio, i rhinovirus vengono introdotti nella cellula mediante endocitosi mediata da recettori in vescicole rivestite di clatrina. All'interno dell'endosoma acido, il virus si espande di circa il 4% e una delle proteine del capsido, VP1 (proteina virale 1), forma dei pori nell'endosoma che consentono il rilascio dell'RNA del rhinovirus. D'altra parte, il virus dell'influenza ha una proteina virale nota come emoagglutinina (HA) incorporata nell'involucro del virus. L'HA si lega ai residui di acido sialico trovati sulla superficie delle cellule epiteliali respiratorie e la penetrazione avviene attraverso l'endocitosi mediata dal recettore. Il basso pH dell'endosoma provoca un cambiamento conformazionale nella proteina HA virale, rivelando un peptide di fusione che avvicina le due membrane e fonde l'involucro virale con la membrana endosomiale. In questo caso, la proteina HA facilita sia l'attaccamento che la rimozione del capsido. I segmenti dell'RNA virale rilasciati vengono trasportati al nucleo ed entrano attraverso i pori nucleari. Si ritiene che altri capsidi virali, come quelli del poliovirus, non entrino affatto nella cellula: il legame del capsido del poliovirus al recettore della superficie cellulare provoca un cambiamento conformazionale nel virione che crea un poro nella membrana cellulare attraverso il quale il l'RNA virale viene rilasciato nel citoplasma. Al contrario, molti virus rimangono in gran parte intatti dopo la penetrazione. I reovirus non si scoprono

completamente all'interno del citoplasma, fornendo una "base di partenza" per la replicazione del genoma.

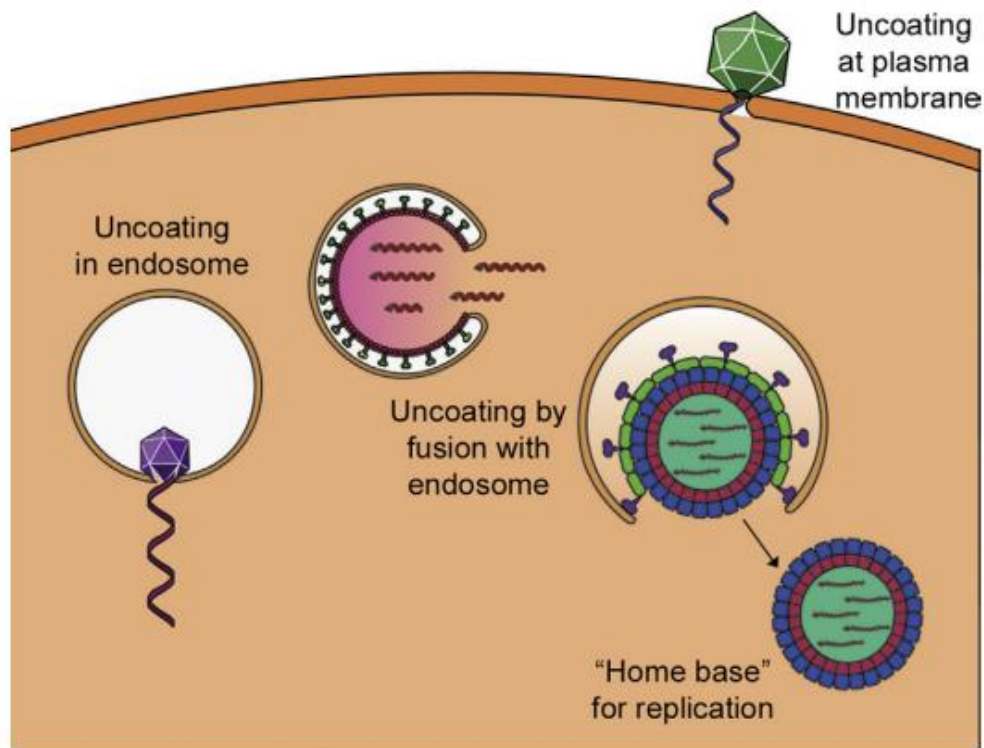


Fig. 8 Uncoating di capsidi virionici. Alcuni virus, inclusi i rinovirus, si espandono per formare pori nell'endosoma attraverso i quali il genoma virale può fuoriuscire. Come il virus dell'influenza, altri virus inducono la fusione dell'involucro del virione con la membrana endosomiale, rilasciando il genoma virale. I capsidi del poliovirus non entrano affatto nella cellula: il legame del capsido al recettore della superficie cellulare induce un cambiamento di conformazione che crea un poro nella membrana attraverso il quale viene trasportato il genoma. Molti virus mantengono un capsido parzialmente intatto nel citosol che funge da "casa base" per la replicazione, come fanno i reovirus. Louten J., *Essential Human Virology*, Academic Press, 2017

Molti herpesvirus infettano i neuroni ma devono replicarsi nel nucleo, che può trovarsi a una certa distanza dal loro sito di ingresso nella membrana plasmatica. Dopo la fusione dell'involucro virale con la membrana plasmatica, i nucleocapsidi intatti di HSV-1 vengono trasportati lungo i microtubuli fino al nucleo. Nel nucleo, il capsido dell'HSV prende contatto con un poro nucleare e il suo DNA viene trasportato nel nucleo (**Fig. 9**).

Altri capsidi ancora sono abbastanza piccoli da passare attraverso i pori nucleari: il capsido dell'epatite B, con un diametro di circa 30 nm, può essere importato intatto attraverso un poro nucleare.

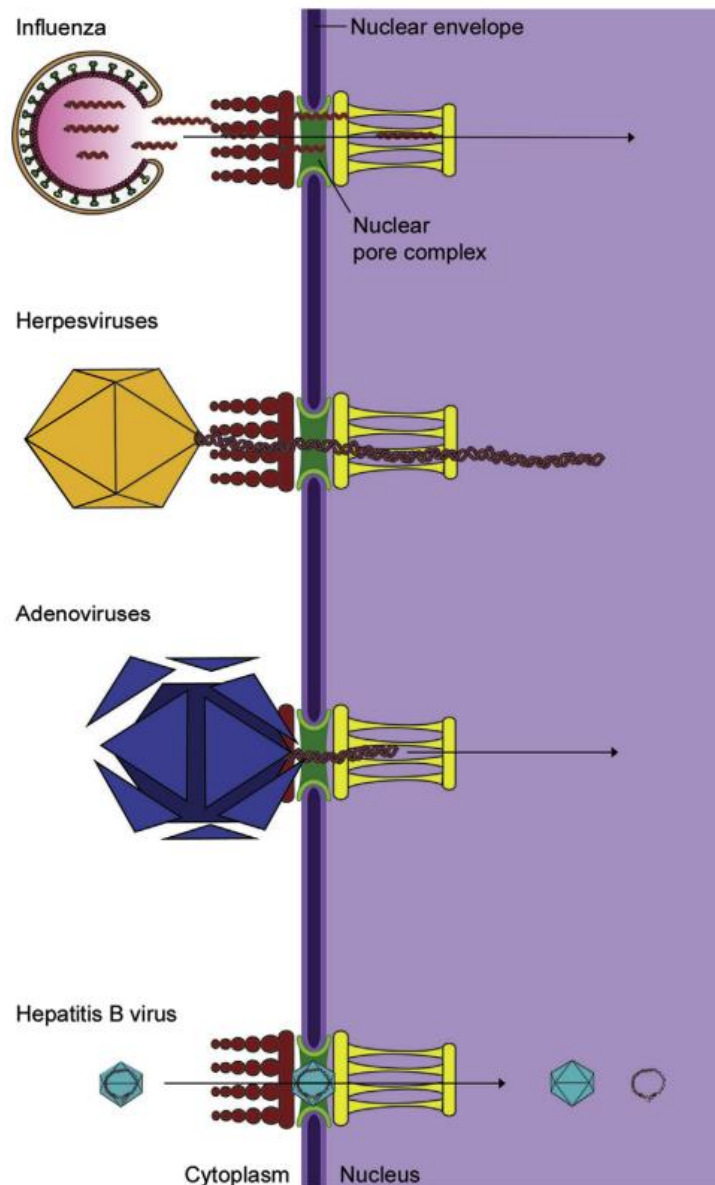


Fig. 9 Trasporto di genomi virali nel nucleo. Diversi virus devono trasportare i loro genomi nel nucleo affinché avvenga la trascrizione e/o la replicazione virale. I segmenti del genoma dell'influenza vengono trasportati attraverso il poro nucleare nel nucleo. I capsidi dell'herpesvirus vengono trasportati lungo i microtubuli fino al poro nucleare, dove avviene la rimozione del rivestimento. I capsidi dell'adenovirus si disassemblano nel poro nucleare e il DNA virale viene trasportato nel nucleo. Altri virus, incluso il virus dell'epatite B, sono abbastanza piccoli da permettere all'intero capside di passare attraverso il poro nucleare. Louten J., *Essential Human Virology*, Academic Press, 2017

2.4 Replicazione

La strategia di replicazione di un virus dipende generalmente dal tipo di acido nucleico che contiene (**Fig. 10**). Molte classi di virus sfruttano le proteine cellulari per replicare i propri genomi. La posizione di queste proteine determinerà quindi generalmente dove avrà luogo la replicazione del genoma virale.

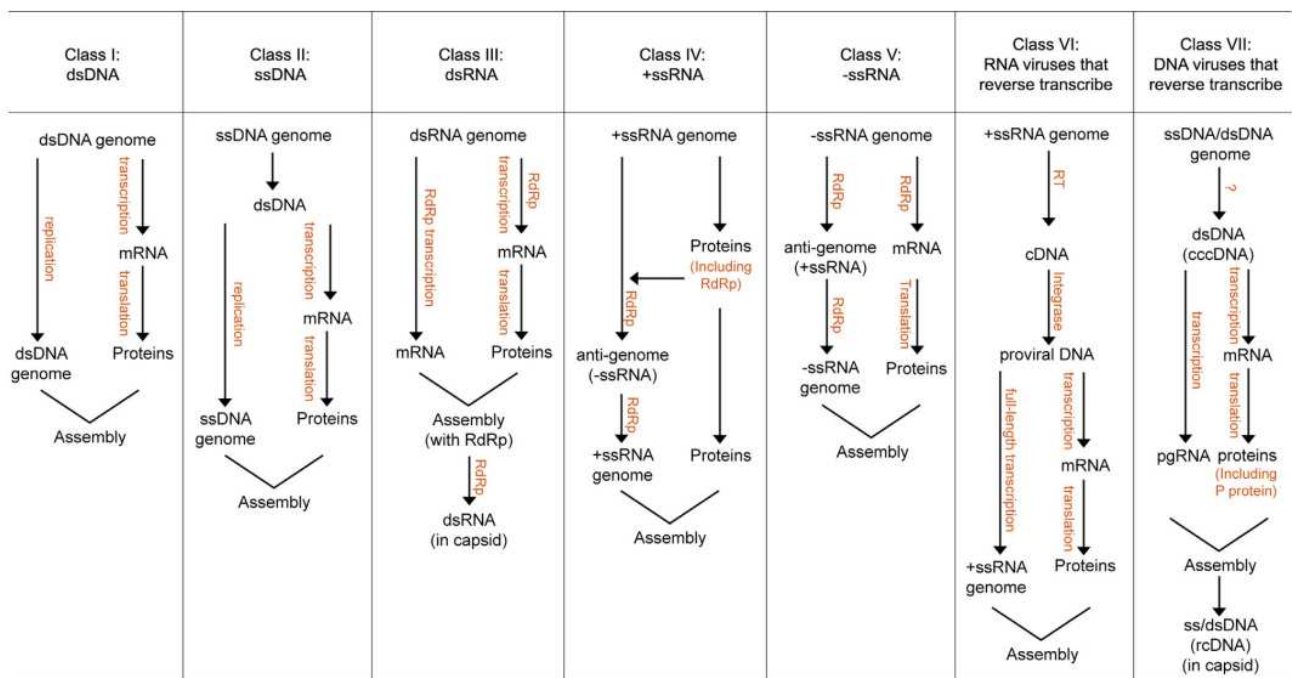


Fig. 10 Le strategie di replicazione dei virus nelle classi di Baltimore. Tutti i virus, indipendentemente dal loro tipo di acido nucleico, devono copiare il genoma e tradurre le proteine per assemblare i virioni nascenti. Louten J., *Essential Human Virology*, Academic Press, 2017

Il sistema di classificazione di Baltimore è utile quando si discutono le strategie generali di replicazione dei virus perché classifica i virus in sette classi in base al loro tipo di genoma (**Tab. 1**):

1. Virus a DNA a doppio filamento
2. Virus a DNA a singolo filamento
3. Virus a RNA a doppio filamento
4. Virus a RNA a senso positivo
5. Virus a RNA a senso negativo
6. Virus a RNA con trascrizione inversa
7. Virus a DNA con trascrizione inversa

Gli acidi nucleici virali si trovano in una varietà di configurazioni. Possono essere lineari o circolari e possono essere segmentati in diversi pezzi più piccoli all'interno del virione, come accade con i virus dell'influenza, o non segmentati come il virus della rabbia, contenente una molecola di acido nucleico che codifica tutti i geni necessari. Le molecole più lunghe sono più soggette a rotture, ma i virus segmentati devono impacchettare tutti i segmenti del genoma in un virione affinché sia

infettivo. Indipendentemente dalla struttura del loro acido nucleico, tutti i virus devono esprimere le loro proteine virali e replicare il loro genoma all'interno della cellula per creare nuovi virioni.

Family	Virus examples
Class I: dsDNA viruses	
Adenoviridae	Adenovirus
Herpesviridae	Herpes simplex virus, Epstein–Barr virus, varicella zoster virus
Papillomaviridae	Human papillomavirus
Polyomaviridae	JC polyomavirus, BK polyomavirus, SV40
Poxviridae	Variola, vaccinia
Class II: ssDNA viruses	
Parvoviridae	Parvovirus B19
Anelloviridae	Torque teno virus
Class III: dsRNA viruses	
Picobirnaviridae	Human picobirnavirus
Reoviridae	Rotavirus
Class IV: +ssRNA viruses	
Astroviridae	Human astrovirus
Caliciviridae	Norwalk virus
Coronaviridae	Human coronavirus
Flaviviridae	Dengue virus, yellow fever virus, West Nile virus, hepatitis C virus
Hepeviridae	Hepatitis E virus
Picornaviridae	Poliovirus, rhinovirus, enterovirus, hepatitis A virus
Togaviridae	Eastern equine encephalitis, Chikungunya virus, rubella virus
Class V: –ssRNA viruses	
Arenaviridae	Lymphocytic choriomeningitis virus, Lassa virus, Machupo virus
Bunyaviridae	Hantavirus, Crimean–Congo hemorrhagic fever virus
Filoviridae	Ebola virus, Marburg virus
Orthomyxoviridae	Influenza A virus, influenza B virus
Paramyxoviridae	Nipah virus, Hendra virus, measles virus, mumps virus
Rhabdoviridae	Rabies virus
Class VI: RNA viruses that reverse transcribe	
Retroviridae	Human immunodeficiency virus-1 and -2
Class VII: DNA viruses that reverse transcribe	
Hepadnaviridae	Hepatitis B virus

Tab. 1 Famiglie di virus umani all'interno di ciascuna classe di replicazione

2.4.1 Classe IV: virus +ssRNA

La classe IV è rilevante in questo contesto in quanto include il SARS-CoV-2, oggetto di studio.

I virus con genomi a ssRNA che possono agire direttamente come mRNA sono noti come virus a RNA a senso positivo (abbreviati +ssRNA) (**Fig. 11A**).

I virus +ssRNA sono più abbondanti di qualsiasi altra classe di virus e infettano un'ampia gamma di specie ospiti. Includono sette diverse famiglie virali umane, inclusi i coronavirus, i flavivirus e i picornavirus, che causano malattie significative negli esseri umani. La loro abbondanza indica che i virus +ssRNA hanno avuto un grande successo evolutivo.

Poiché il genoma dei virus +ssRNA agisce come mRNA, questi virus hanno informazioni genetiche infettive. I loro genomi sono traducibili dai ribosomi dell'ospite e hanno cappucci al 5' (o proteine che agiscono in modo simile a un cappuccio 5') e spesso contengono sequenze di coda poli-A all'estremità 3'. Esperimenti che hanno introdotto solo il genoma del poliovirus nel citoplasma di una cellula hanno portato alla formazione di nuovi virioni, perché la traduzione del genoma è la prima attività che avviene con i genomi +ssRNA. Questo produce tutte le proteine virali necessarie per orchestrare il resto del ciclo di replicazione. Laddove i virus a dsRNA devono portare una RdRp all'interno del virione, i virus +ssRNA codificano una RdRp all'interno del loro genoma. La proteina RdRp viene prodotta immediatamente dopo l'ingresso nella cellula mediante traduzione del genoma virale. È importante notare, tuttavia, che, anche se il virus codifica la propria proteina RdRp, spesso sono necessarie anche proteine cellulari affinché avvenga la replicazione. Ad esempio, nonostante il genoma del poliovirus sia infettivo, non viene replicato quando viene iniettato in un ovocita di rana *Xenopus* (ovulo/uovo) a meno che il citoplasma di una cellula umana non venga iniettato insieme al genoma, indicando che è richiesto almeno un componente cellulare umano per la replicazione del genoma del poliovirus.

Una caratteristica comune dei virus +ssRNA è che il loro genoma infettivo codifica una poliproteina. Ciò fornisce un metodo economico per derivare diverse proteine dalla traduzione di un solo pezzo di mRNA. Nel caso dei picornavirus, il genoma del filamento positivo viene tradotto nella sua interezza, quindi le proteasi scindono la poliproteina in posizioni diverse per creare diverse proteine (**Fig. 11B**). La scissione alternativa di alcune sezioni si traduce in proteine aggiuntive.

Altri virus +ssRNA, come i togavirus che includono il virus della rosolia, iniziano traducendo solo una parte del genoma per creare un insieme iniziale di proteine che dirigono la successiva replicazione del genoma e la traduzione di altre proteine virali. Ciò consente la creazione di "fasi" di replicazione del virus, simili a quanto osservato con la trascrizione genica immediata precoce, precoce e tardiva di alcuni virus a DNA. Anche la creazione di poliproteine accompagna comunemente questo metodo di traduzione e la soppressione della terminazione determina la produzione di diverse catene poliproteiche. Ciò avviene a bassa velocità ma si traduce nella generazione di importanti proteine virali, inclusa la RdRp virale.

Quando le proteine di replicazione dei virus +ssRNA vengono sintetizzate, tendono a raccogliersi in corrispondenza o all'interno di determinate membrane della cellula, creando complessi di replicazione. Ad esempio, le proteine virali del poliovirus rimangono legate alle membrane del reticolo endoplasmatico ruvido (rER) o alle vescicole secretorie. Il poliovirus è un virus senza envelope, quindi la sua funzione sembra essere quella di concentrare le proteine virali in una

posizione della cellula per facilitare meglio i processi di replicazione. Un altro motivo per cui questo potrebbe essersi evoluto è per proteggere il ssRNA virale dalle risposte immunitarie intracellulari.

Il genoma dei virus +ssRNA viene utilizzato per creare un filamento negativo complementare, l'RNA antigenomico, che viene utilizzato come modello per creare molte copie del genoma +ssRNA.

I virus a RNA sono più inclini alla mutazione rispetto ai virus a DNA. Le RdRp non hanno capacità di correzione di bozze. Ciò aumenta il tasso di errore complessivo dell'enzima, da 1 errore per 10⁹ basi per una DNA polimerasi a più di 1 errore per 10⁵ basi per un RdRp, il che si traduce in una minore fedeltà o accuratezza dell'enzima. I virus a RNA hanno alcuni dei più alti tassi di mutazione di tutte le entità biologiche. Le mutazioni generate dalle RdRp possono provocare proteine virali mutate e, successivamente, ceppi leggermente diversi del virus che possono sopravvivere meglio sotto determinate pressioni ambientali.

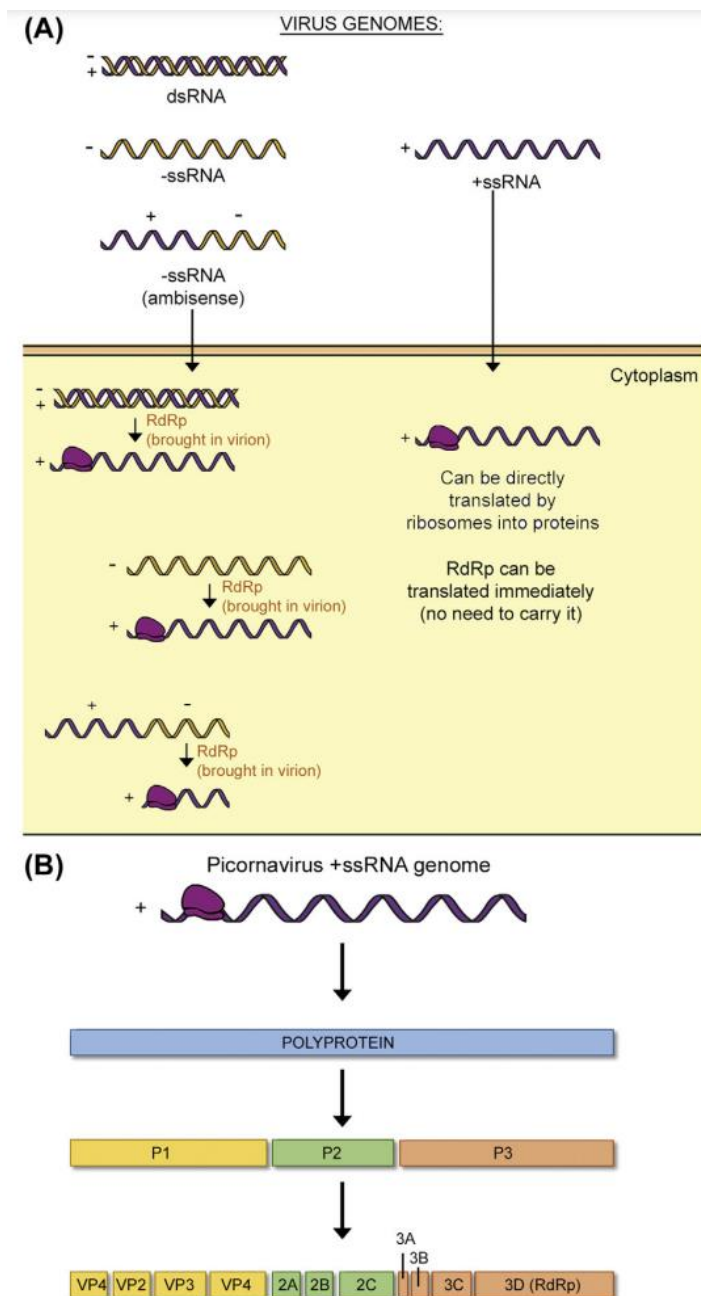


Fig. 11 Dettagli sulla replicazione dei virus a RNA. (A) I virus +ssRNA hanno genomi infettivi che possono essere immediatamente tradotti dai ribosomi. Di conseguenza, non hanno bisogno di portare una RdRp nella cellula. I virus a RNA con genomi dsRNA, -ssRNA o ambisense devono trasportare la propria proteina RdRp nella cellula affinché avvenga la trascrizione. (B) I picornavirus +ssRNA codificano una singola poliproteina che viene scissa più volte per creare tutte le proteine necessarie per la replicazione. Louten J., *Essential Human Virology*, Academic Press, 2017

2.5 Assemblaggio

I virus vengono creati dai componenti appena sintetizzati e, per essere rilasciati dalla cellula, tali componenti devono essere raccolti in un particolare sito della cellula e sottoposti ad assemblaggio per formare una particella virale immatura. Questa fase può spesso avvenire insieme alla maturazione e al rilascio.

La posizione dell'assemblaggio del virione dipenderà dal particolare virus. Può avvenire all'interno del nucleo della cellula, sulla membrana plasmatica o su una varietà di membrane intracellulari, come il complesso del Golgi. La maggior parte dei virus a DNA senza envelope assembla il proprio nucleocapside nel nucleo, poiché quello è il sito della replicazione del genoma. Le proteine virali vengono importate attraverso i pori nucleari per raggiungere il sito di assemblaggio. Tuttavia, una volta assemblati, la maggior parte dei virus a DNA è troppo grande per passare attraverso i pori nucleari. A questo punto, alcuni virus sono in grado di attraversare l'involucro nucleare a doppia membrana, mentre altri inducono la lisi cellulare o l'apoptosi per sfuggire dal nucleo. D'altra parte, i virus con involucri derivati dalla membrana plasmatica di solito si raccolgono lì.

Il genoma di un virus elicoidale è protetto dalla ripetizione delle proteine del capsido. Per questo motivo, le proteine del capsido possono iniziare ad avvolgere il genoma non appena viene copiato (o viceversa, a seconda del virus: il genoma può essere avvolto attorno alle proteine del capsido). Al contrario, alcuni virus icosaedrici completano quasi l'assemblaggio dei loro capsidi prima che venga inserito il genoma al loro interno.

L'assemblaggio spontaneo del capsido, chiamato "autoassemblaggio", si verifica con le proteine del capsido di semplici virus icosaedrici, come i picornavirus e i parvovirus. L'assemblaggio di virus con un'architettura più complessa è orchestrato da una varietà di proteine chaperone virali chiamate proteine di *scaffolding*. Gli herpesvirus e gli adenovirus sono esempi di grandi virus icosaedrici che si assemblano con l'assistenza delle proteine *scaffolding*.

2.6 Maturazione

Dopo che il genoma e altre proteine essenziali sono state impacchettate all'interno del capsido, si verificano le fasi finali della replicazione del virus: maturazione e rilascio. Fino a questo punto, il virione era in fase di formazione e, se la cellula fosse stata aperta a questo punto, i virioni non sarebbero stati in grado di avviare l'infezione di nuove cellule. La maturazione si riferisce ai cambiamenti finali all'interno di un virione immaturo che si traducono in una particella virale infettiva. I cambiamenti strutturali del capsido sono spesso coinvolti e questi possono essere mediati da enzimi dell'ospite o enzimi codificati dal virus. Un buon esempio riguarda la proteina HA dell'influenza. La proteina HA è in grado di legare l'acido sialico dopo essere stata glicosilata (tramite modifica post-traduzionale). Tuttavia, la proteina HA deve essere scissa in due porzioni, HA1 e HA2, per diventare infettiva, perché sebbene la porzione HA1 si leghi al recettore della superficie cellulare, la porzione HA2 è ciò che fonde l'involucro virale alla membrana endosomiale per rilasciare il virus nel citoplasma. Questa scissione viene effettuata dalle proteasi cellulari. Al contrario, la particella centrale dell'HIV è composta da proteine codificate dal gene gag. Il gene viene tradotto in una poliproteina che viene scissa dalla proteasi virale per formare le proteine del capsido, della matrice

e del nucleocapside del virione. In questo caso, la maturazione avviene dopo che il virione è stato rilasciato dalla superficie cellulare ed è necessaria per formare un virione infettivo.

2.7 Rilascio

La fase finale del ciclo di replicazione del virus è il rilascio del virione nell'ambiente extracellulare, dove può continuare il ciclo di infezione con nuove cellule. Il rilascio può avvenire in diversi modi, a seconda del virus. I virus che ottengono l'envelope dalla membrana plasmatica generalmente si assemblano sullo strato interno della membrana plasmatica, incorporando le loro proteine dell'envelope nella membrana plasmatica. Quando le proteine del capsido interagiscono, le proteine virali associate alla membrana fanno sì che la membrana plasmatica inizi a curvarsi attorno al capsido. Questo continua fino a quando la membrana plasmatica non è completamente avvolta attorno al virus, che lascia la cellula. Questo processo è noto come gemmazione (**Fig. 12**).

I virus possono gemmare da qualsiasi sistema di membrane all'interno della cellula, incluso il rER, il complesso del Golgi o persino l'involucro nucleare. In questo caso, il virione, già dotato di envelope, non ha bisogno di gemmare attraverso la membrana plasmatica e generalmente subisce esocitosi per lasciare la cellula.

I virus senza envelope possono uscire dalla cellula anche tramite esocitosi. I virus litici, tuttavia, distruggono la membrana plasmatica e provocano la lisi, o rottura, della cellula. Questo rilascia i virioni nascenti per infettare nuove cellule. Molti virus umani senza envelope vengono rilasciati attraverso la lisi cellulare.

I processi di assemblaggio, maturazione e rilascio sono strettamente collegati, ma tutti sono necessari per creare una progenie di virioni infettivi in grado di continuare il ciclo infettivo.

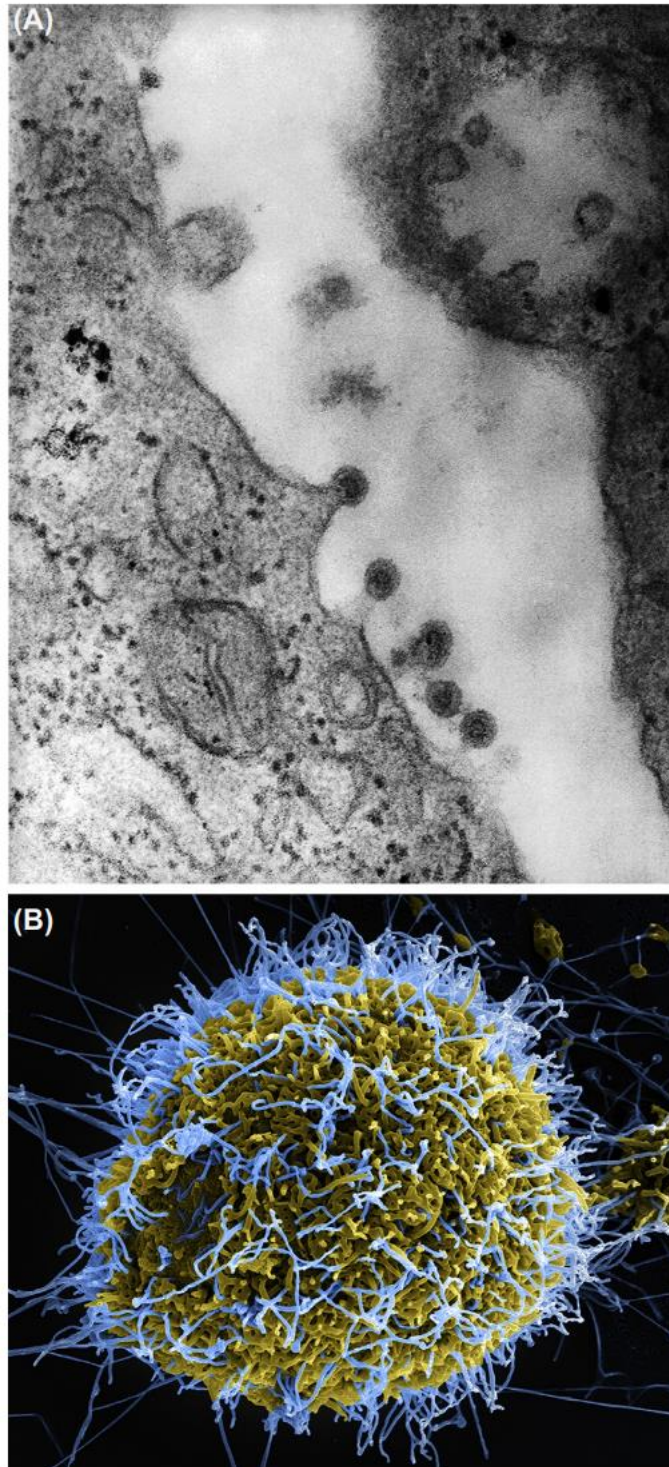


Fig. 12 Gemmazione dei virioni. (A) In questa micrografia elettronica a trasmissione si osservano i virioni del virus della rosolia che gemmano dalla membrana plasmatica dell'ospite. (B) In questa micrografia elettronica a scansione pseudocolorata digitalmente, i virioni elicoidali dell'Ebolavirus (blu) gemmano da una cellula infetta (giallo). Louten J., *Essential Human Virology*, Academic Press, 2017

3. Coronavirus

3.1 Classificazione e filogenesi

I coronavirus sono una vasta famiglia di virus che possono causare malattie sia negli animali che nell'uomo. L'attuale classificazione dei coronavirus riconosce 39 specie in 27 sottogeneri che appartengono alla famiglia Coronaviridae. Di questi, almeno 7 coronavirus sono noti per causare infezioni respiratorie nell'uomo: 229E, NL63, OC43, HKU1, sindrome respiratoria mediorientale da coronavirus (MERS-CoV), sindrome respiratoria acuta grave da coronavirus (SARS-CoV) e sindrome respiratoria acuta grave da coronavirus 2 (SARS-CoV-2). MERS-CoV, SARS-CoV e SARS-CoV-2 appartengono al genere Betacoronavirus e hanno tutti alti tassi di mutazione che si traducono in diversità genetica virale, plasticità e adattabilità per invadere un'ampia gamma di ospiti [2].

Anche se la storia dei CoV è iniziata negli anni '40 [4], i primi CoV umani sono stati identificati negli anni '60 come agenti infettivi di malattie respiratorie lievi denominate HCoV-229E e HCoV-OC43 [5,6]. Lo scoppio di SARS-CoV nel 2002 ha portato a una ricerca attiva sui nuovi coronavirus e ha identificato HCoV-NL63 e HCoV-HKU1 rispettivamente nel 2004 e nel 2005 [7,8]. Inoltre, altri HCoV (coronavirus umani) scoperti fino ad oggi sono MERS-CoV, SARS-CoV e SARS-CoV-2. I primi quattro CoV, ovvero HCoV229E, HCoV-OC43, HCoV-NL63 e HCoV-HKU, sono comunemente dispersi e sovvenzionano circa un terzo del comune raffreddore negli esseri umani [10]. Nei casi più gravi, tuttavia, possono causare bronchiolite e polmonite pericolosa per la vita in individui e bambini immunocompromessi [11,12]. Inoltre, questi CoV sono stati associati a malattie enteriche e neurologiche [13,14,15,16,17]. Tuttavia, SARS-CoV, MERS-CoV e SARS-CoV-2 sono ben noti come HCoV pandemici. Nella tassonomia, HCoV-229E e HCoV-NL63 sono collocati sotto AlphaCoV mentre HCoV-HKU1, HCoV-OC43, SARS-CoV e MERS-CoV sono classificati come BetaCoV; entrambi i gruppi infettano principalmente i mammiferi mentre GammaCoV e DeltaCoV sono specifici per gli uccelli, ma occasionalmente possono anche infettare i mammiferi [9]. In precedenza, il Coronaviridae Study Group (CSG) ha identificato i ceppi SARS-CoV e MERS-CoV in una nuova specie sotto il nuovo sottogruppo informale del genere BetaCoV [18]. Tuttavia, la recente introduzione del rango di sottogenere nella tassonomia dei virus ha stabilito i due sottogruppi informali di SARS-CoV e MERS-CoV rispettivamente come sottogeneri Sarbecovirus e Merbecovirus [19,20]. Sorprendentemente, il SARS-CoV-2 appena emerso differisce dai virus pandemici zoonotici precedentemente riportati, vale a dire SARS-CoV e MERS-CoV; e quindi, la posizione tassonomica di SARS-CoV-2 sotto i sottogeneri Sarbecovirus può essere cambiata sulla base di ulteriori prove [21] (**Fig. 13**).

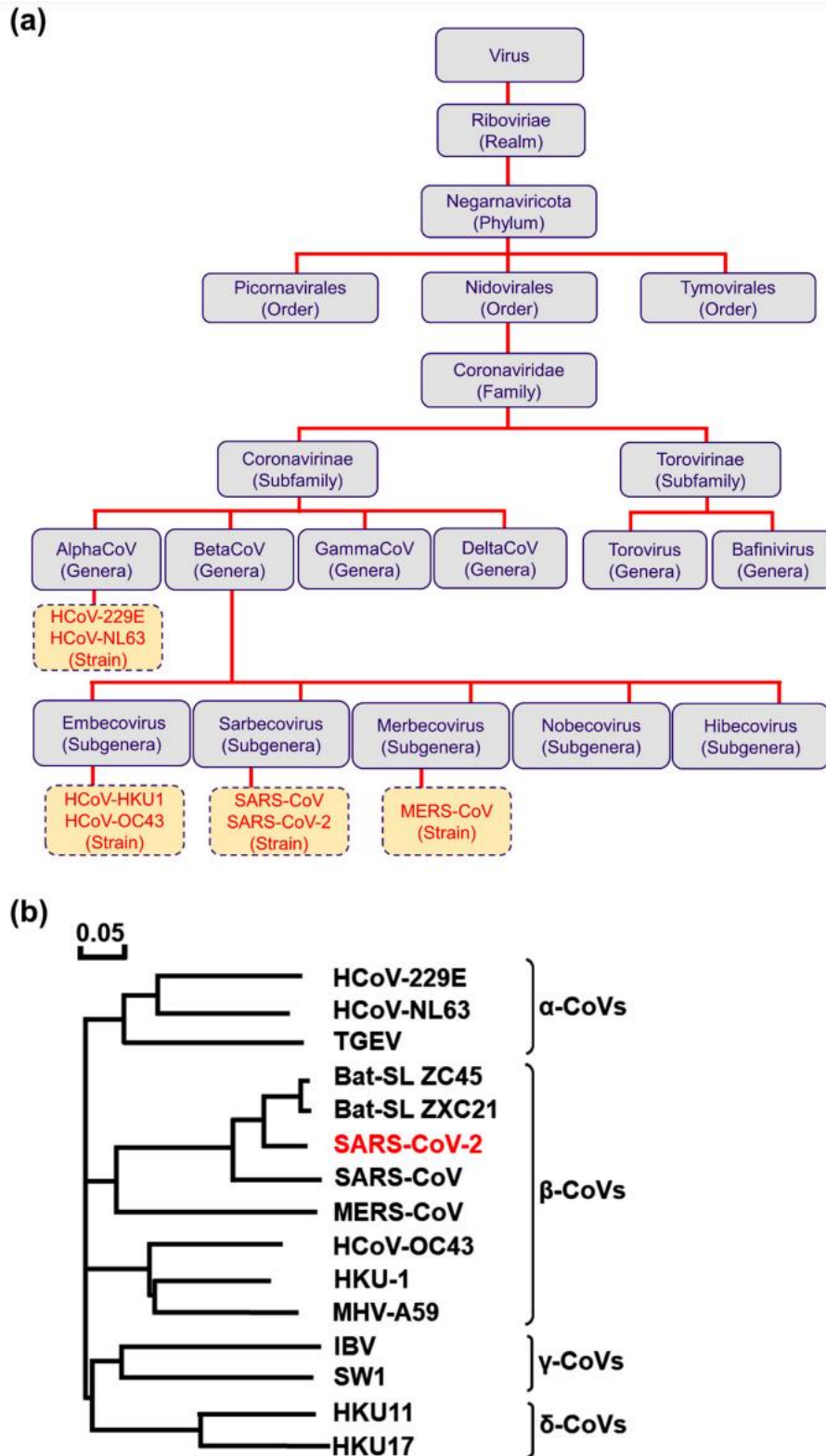


Fig. 13 La classificazione tassonomica **(a)** e le posizioni per i sette HCoV noti e **(b)** l'analisi dell'albero filogenetico dei CoV costruita sulla base del gene S utilizzando il software Molecular Evolutionary Genetics Analysis 6 con il metodo neighbor-joining.

3.2 SARS-CoV-2

All'inizio di dicembre 2019, una polmonite atipica è stata segnalata in un gruppo di pazienti a Wuhan (Cina) ed è stato dimostrato che era causata da un nuovo coronavirus, il SARS-CoV-2 e la relativa malattia indicata come COVID-19 [23]. I serbatoi animali sono probabilmente pipistrelli [24]; è stato proposto che i pangolini possano essere gli ospiti animali che trasmettono il virus all'uomo [25] ma l'eventuale ospite intermedio non è stato ancora identificato in modo definitivo.

L'infezione da SARS-CoV-2 può essere asintomatica (fino al 40% dei casi) o causare un ampio spettro di malattie, da sintomi lievi a malattie potenzialmente letali [26]. Gli individui infetti presentano più comunemente febbre, tosse secca, fiato corto, affaticamento, mialgia, nausea/vomito o diarrea, mal di testa, debolezza, rinorrea, anosmia e ageusia. Le complicanze comuni tra i pazienti ospedalizzati includono polmonite, ARDS (sindrome da distress respiratorio acuto), danno epatico acuto, danno cardiaco, coagulopatia protrombotica, danno renale acuto e manifestazioni neurologiche. I pazienti in condizioni critiche potrebbero anche sviluppare una tempesta di citochine e una sindrome da attivazione dei macrofagi. Comorbidità come ipertensione, diabete, malattie cardiovascolari, malattie polmonari croniche, malattie renali croniche, neoplasie e malattie epatiche croniche erano presenti nel 60-90% dei pazienti ospedalizzati [27]. Sintomi lievi si sono verificati nell'80% dei casi confermati in laboratorio. Circa il 14-19% dei pazienti ha richiesto il ricovero in ospedale e il 3-5% dei casi il trasferimento in terapia intensiva, più comunemente a causa di insufficienza respiratoria ipossiémica. Tra questi, il 29-91% necessitava di ventilazione meccanica invasiva. Complessivamente, la mortalità dei pazienti ospedalizzati con COVID-19 è stata di circa il 15-20%, mentre ha raggiunto il 40% nei pazienti ricoverati in terapia intensiva. Il tasso di mortalità globale stimato variava tra lo 0,25 e il 3,0% [28]. La mortalità era dello 0,02% nei pazienti di età compresa tra 20 e 49 anni, dello 0,5% per i pazienti di età compresa tra 50 e 69 anni e maggiore del 5,4% per i pazienti di età superiore a 80 anni. I bambini con COVID-19 manifestavano sintomi più lievi prevalentemente limitati al tratto respiratorio superiore. Tuttavia, in alcuni bambini con COVID-19 è stata descritta una rara sindrome infiammatoria multisistemica [29]. È stato dimostrato che l'alterata risposta all'interferone di tipo I è coinvolta in pazienti con COVID-19 potenzialmente letale [30,31]. Il COVID-19 si è diffuso a livello globale in pochi mesi portando a oltre 74 milioni di contaminazioni e oltre 1,6 milioni di morti in tutto il mondo al 18 dicembre 2020 [32].

Molteplici misure di sanità pubblica sono state attuate dalla maggior parte dei paesi, come azioni individuali (distanziamento fisico, lavaggio delle mani, uso di mascherine ed etichetta per la tosse), identificazione di cluster (identificazione dei casi, tracciamento dei contatti e isolamento), azioni normative (chiusura delle scuole, chiusura, ordine di soggiorno, chiusura e restrizione dei trasporti pubblici, limiti alle dimensioni degli assembramenti e alla capacità aziendale), chiusure delle frontiere interne e internazionali, restrizioni di viaggio e quarantena forzata. L'obiettivo di queste misure di sanità pubblica è stato ritardare e appiattire la curva epidemica, prevenendo il collasso del sistema sanitario e proteggendo le persone a più alto rischio prima della disponibilità di vaccini e trattamenti sicuri ed efficaci. Tuttavia, contrariamente a SARS-CoV, la diffusione virale nei pazienti con COVID-19 avviene pochi giorni prima dell'insorgenza dei sintomi (pre-sintomatica) e alcuni pazienti rimangono asintomatici durante la diffusione del virus, il che ha reso le misure di isolamento dei casi meno efficienti [33]. In meno di un anno, un primo vaccino basato su una nuova piattaforma costituita da mRNA che codifica la proteina spike del virus (sviluppato da Pfizer/BioNTech) ha ottenuto la designazione per uso di emergenza dalla FDA statunitense. Diversi altri vaccini basati su

virus intero inattivato, vettore virale ricombinante non replicativo, proteina spike ricombinante e particelle simili a virus sono stati sviluppati [34].

3.2.1 Struttura e genoma

SARS-CoV-2 è un virus dotato di envelope con virioni approssimativamente sferici o moderatamente pleomorfi di circa 60-140 nm di diametro [35]. La membrana virale contiene la glicoproteina spike (S) che forma i peplomeri sulla superficie del virione, conferendo al virus la sua “corona” – o morfologia simile a una corona al microscopio elettronico. La glicoproteina di membrana (M) e la proteina dell'involucro (E) forniscono la struttura ad anello. All'interno del virione si trova un nucleocapside elicoidale composto dalla proteina nucleocapside (N) complessata con un genoma di RNA a singolo filamento positivo di circa 30 kb di lunghezza [36] (**Fig. 14**).

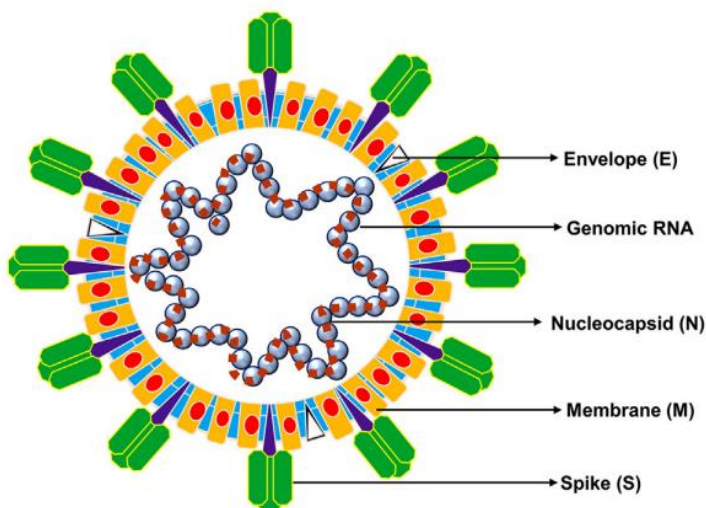


Fig. 14 Struttura tipica di un Coronavirus (80-120 nm) che mostra varie proteine strutturali, cioè S, M, E e RNA genomico impacchettato all'interno della particella dalla proteina N. [3]

I Coronavirus umani, inclusi SARS-CoV, MERS-CoV e SARS-CoV-2, contengono il più grande genoma di RNA a filamento singolo a senso positivo non segmentato di dimensioni comprese tra 26,2 e 31,7 kb con contenuti G+C variabili dal 32% al 43% [37,38]. L'organizzazione del genoma per un Coronavirus è 5'-leader-UTR replicasi/trascrittasi, spike (S), envelope (E), membrana (M), nucleocapside (N) 3'UTR-coda poli (A). Il 5'UTR (regione non tradotta) e il 3'UTR sono coinvolti nelle interazioni inter- e intramolecolari, essenzialmente richieste nelle interazioni RNA-RNA e per il legame di proteine virali e cellulari [39]. L'ORF1a e l'ORF1b occupano i primi due terzi del genoma e producono la poliproteina replicasi/trascrittasi che subisce attività autoproteolitica con l'ausilio della proteinasi simile alla papaina (PLpro) e della proteasi principale simile alla 3C (3CLpro o Mpro) per formare 16 proteine non strutturali (nsp) [40]. PLpro conduce la scissione all'N-terminale della poliproteina replicasi per produrre nsp1, nsp2 e nsp3 necessari per la replicazione virale [41]. Tuttavia, Mpro è la prima proteina che separa autoproteoliticamente la replicasi 1ab della poliproteina (~790 kDa, sequenza di riconoscimento situata a Leu-Gln↓Ser,Ala,Gly; dove i segni ↓ indicano il sito di clivaggio) per produrre enzimi virali maturi e ulteriori nsp a valle scissi a 11 siti per

rilasciare nsp4-nsp16 [42,43]. I successivi frame di lettura sul genoma codificano le quattro principali proteine strutturali; S, E, M e N, che sono proteine comuni a tutti i CoV [44]. Tutte le proteine strutturali e accessorie virali sono decodificate da RNA subgenomici (sg) di CoV in cui le proteine accessorie sono disperse tra i filamenti di ORF ma il loro numero e le loro attività funzionali sono limitate ai ceppi di CoV [40,45] (**Fig. 15**).

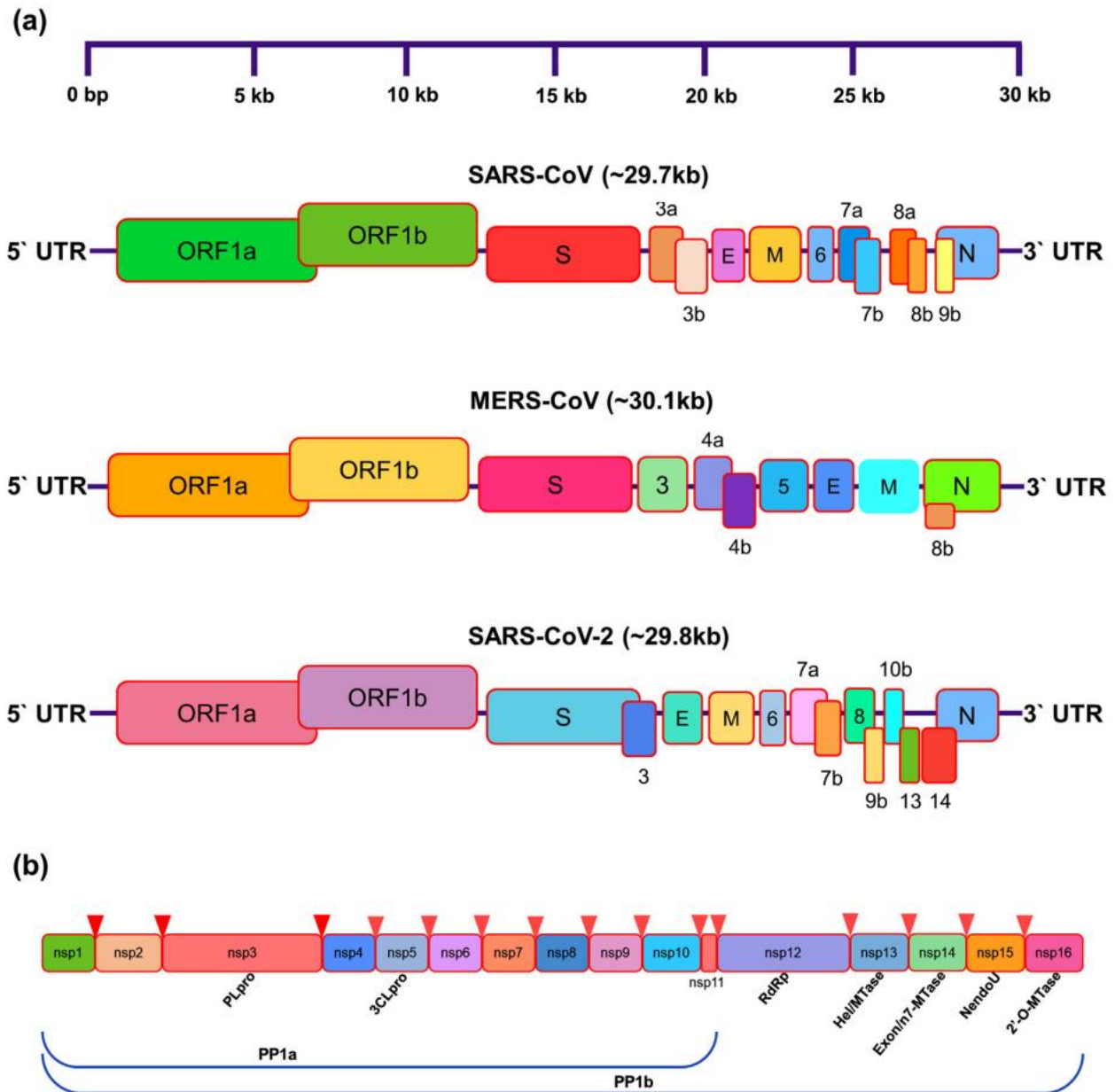


Fig. 15 Rappresentazione schematica per (a) organizzazione del genoma e (b) domini funzionali nel genoma di SARS-CoV, MERS-CoV e SARS-CoV-2. [3]

3.2.2 Proteine virali

3.2.2.1 Proteine non strutturali

Oltre alle proteine strutturali che formano il capsido, il genoma virale codifica molte NSP che svolgono numerosi ruoli nei processi di replicazione e assemblaggio del virus [47]. Queste proteine partecipano alla patogenesi virale modulando la regolazione precoce della trascrizione, l'attività dell'elicasi, l'immunomodulazione, la transattivazione genica e contrastando la risposta antivirale [48,49,50]. La ricerca ha rivelato che le NSP di SARS-CoV-2 sono coinvolte in molti processi biologici, tra cui replicazione del genoma virale, elaborazione delle proteine, trascrizione e proteolisi. Queste proteine sono coinvolte nel legame dell'RNA, nell'attività dell'endopeptidasi, nell'attività della transferasi, nel legame dell'ATP, nel legame dello ione zinco, nell'attività di RNA polimerasi 5'-3' RNA-diretta, attività esoribonucleasica, che produce 5' fosfomonoesteri e attività di metiltransferasi (**Tab. 2**).

List of non-structural proteins in SARS-CoV-2 and their molecular functions.

S. No.	Range	Protein name and ID	Description	Proposed function
1.	1–180	Nsp1 YP_009725297.1	Nsp1 is the N-terminal product of the viral replicase	Leader protein host translation inhibitor. Mediates RNA replication and processing. Involved in mRNA degradation [41].
2.	181–818	Nsp2 YP_009725298.1	Nsp2 is a replicase product essential for proofreading viral replication	Modulation of host cell survival signalling pathway by interacting with host PHB and PHB2 [42].
3.	819–2763	Nsp3 YP_009725299.1	Nsp3 is a papain-like proteinase contains several domains.	Functions as a protease to separate the translated polyprotein into its distinct proteins [43,44].
4.	2764–3263	Nsp4 YP_009725300.1	A membrane-spanning protein contains transmembrane domain 2 (TM2)	Believed to anchor the viral replication-transcription complex to modified ER membranes [45].
5.	3264–3569	Nsp5 YP_009725301.1	3C-like proteinase and main proteinase	Involved in viral polyprotein processing during replication [46].
6.	3570–3859	Nsp6 YP_009725302.1	Putative transmembrane domain	Plays a role in the initial induction of autophagosomes from host endoplasmic reticulum.
7.	3860–3942	Nsp7 YP_009725303.1	Nsp7 is an RNA-dependent RNA polymerase	It forms a hexadecameric super-complex with nsp8 that adopts a hollow cylinder-like structure implicated replication [47,48].
8.	3943–4140	Nsp8 YP_009725304.1	Multimeric RNA polymerase; replicase	It forms a hexadecameric super-complex with nsp7 that adopts a hollow cylinder-like structure implicated replication [47,48].
9.	4141–4253	Nsp9 YP_009725305.1	A single-stranded RNA-binding viral protein	Participate in viral replication by acting as an ssRNA-binding protein [49].
10.	4254–4392	Nsp10 YP_009725306.1	Growth-factor-like protein contains two zinc-binding motifs	In viral transcription by stimulating both nsp14 3'-5' exoribonuclease and nsp16 2'-O-methyltransferase activities. Therefore plays an essential role in viral mRNAs cap methylation [50].
11.	4393–5324	Nsp12 YP_009725307.1	RNA-dependent RNA polymerase (Pol/RdRp)	Responsible for replication and transcription of the viral RNA genome [51].
12.	5325–5925	Nsp13 YP_009725308.1	Zinc-binding domain, NTPase/helicase domain, RNA 5'-triphosphatase	A helicase core domain that binds ATP. Zinc-binding domain is involved in replication and transcription [52,53].
13.	5926–6452	Nsp14 YP_009725309.1	Proofreading Exoribonuclease domain (ExoN/nsp14)	Exoribonuclease activity acting in a 3' to 5' direction and N7-guanine methyltransferase activity.
14.	6453–6798	Nsp15 YP_009725310.1	EndoRNase; nsp15-A1 and nsp15B-NendoU	Mn(2+)-dependent endoribonuclease activity
15.	6799–7096	Nsp16 YP_009725311.1	2'-O-ribose methyltransferase	Methyltransferase that mediates mRNA cap 2'-O-ribose methylation to the 5'-cap structure of viral mRNAs [54].
16.	4393-4405	Nsp11 YP_009725312.1	Made of 13 amino acids (sadaqsflngfav) and identical to the first segment of Nsp12.	Unknown

Tab. 2 Proteine non strutturali del SARS-CoV-2 e relativa funzione. [46]

3.2.2.2 Spike

L'ingresso dei coronavirus nelle cellule ospiti è mediato dalla glicoproteina spike (proteina S) [52,53,54]. Le glicoproteine spike transmembrana formano omotrimeri che sporgono dalla superficie virale. La glicoproteina spike è fondamentale per l'ingresso dei coronavirus, quindi è un attraente bersaglio antivirale. La proteina S è composta da due subunità funzionali, le subunità S1 e S2. La subunità S1 è costituita dal dominio N-terminale (NTD) e dal dominio di legame del recettore (RBD). La funzione della subunità S1 è quella di interagire con il recettore sulla cellula ospite. La subunità S2 contiene il peptide di fusione (FP), la ripetizione dell'eptade 1 (HR1), l'elica centrale (CH), il dominio del connettore (CD), la ripetizione dell'eptade 2 (HR2), il dominio transmembrana (TM) e la coda citoplasmatica (CT) (**Fig. 16A**). La funzione della subunità S2 è quella di fondere le membrane dei virus e delle cellule ospiti. Il sito di scissione al confine tra le subunità S1 e S2 è chiamato sito di scissione della proteasi S1/S2. Per tutti i coronavirus, le proteasi ospiti scindono la glicoproteina spike nel sito di scissione S2 per attivare le proteine che sono fondamentali per fondere le membrane dei virus e delle cellule ospiti attraverso cambiamenti conformazionali irreversibili. I glicani N-legati sono fondamentali per il corretto ripiegamento, la neutralizzazione degli anticorpi e la decorazione estensiva dei trimeri proteici della spike [55,56]. Nel complesso, la struttura della proteina S di SARS-CoV-2 assomiglia a quella del SARS-CoV strettamente correlata. Nella conformazione di prefusione, le subunità S1 e S2 rimangono legate in modo non covalente. Diversi tipi di coronavirus utilizzano domini speciali nella subunità S1 per riconoscere diversi recettori di ingresso. SARS-CoV e SARS-CoV-2, per entrare nelle cellule ospiti, riconoscono il recettore dell'enzima di conversione dell'angiotensina 2 (ACE2) sulle cellule ospiti tramite il dominio di legame al recettore (RBD). La proteina S ha due forme di struttura, lo stato chiuso e lo stato aperto (**Fig. 16B**). Nello stato chiuso, i tre motivi di riconoscimento non sporgono dall'interfaccia formata da tre protomeri di proteine spike. Nello stato aperto, l'RBD è nella conformazione "up". Lo stato aperto è necessario per la fusione del SARS-CoV-2 e delle membrane della cellula ospite, facilitando così l'ingresso del SARS-CoV-2 nelle cellule ospiti [55].

La proteina spike di SARS-CoV-2 contiene un RBD che riconosce specificamente il recettore ACE2. L'RBD è un obiettivo critico per i composti e gli anticorpi antivirali [57]. L'RBD di SARS-CoV-2 comprende due domini strutturali: il nucleo e i sottodomini esterni. Il sottodominio principale è altamente conservato. È composto da cinque foglietti beta disposti in modo antiparallelo e un legame disolfuro tra due foglietti beta. Il sottodominio esterno è principalmente dominato dal loop che è stabilizzato dal legame disolfuro [58]. Il nucleo dell'RBD di SARS-CoV-2 è costituito da cinque fogli beta disposti in modo antiparallelo e collegati da loop e brevi eliche. Tra gli *strand* b4 e b7 antiparalleli c'è il motivo di legame del recettore (RBM) che consiste in loop ed eliche, oltre ai brevi *strand* b5 e b6. L'RBM contiene la maggior parte dei siti di legame per il recettore ACE2. Otto dei nove residui di Cys nell'RBD formano quattro paia di legami disolfuro. Tre legami disolfuro sono nel nucleo dell'RBD, migliorando la stabilizzazione del foglio beta (C336-C361, C379-C432 e C391-C525). Per quanto riguarda il restante legame disolfuro (C480-C488), promuove le connessioni tra i loop nell'RBM. Il dominio della peptidasi all'N-terminale di ACE2 contiene il sito di legame, che è formato da due lobi di RBM e ACE2. L'RBM si lega al piccolo lobo dell'ACE2 sul lato inferiore. La superficie dell'RBM è leggermente concava verso l'interno per fare spazio ad ACE2 [59].

Il confine S1/S2 della proteina spike di SARS-CoV-2 costituisce il sito di taglio per la furina proteasi della cellula ospite simile alla subtilisina, che distingue la proteina S di SARS-CoV-2 da quella di SARS-

CoV. Il sito di taglio della furina comprende quattro residui (P681, R682, R683 e A684) e si trova al confine tra le subunità S1 e S2. Funzionalmente, R682, R683, A684 e R685 costituiscono il sito minimo di scissione della furina polibasica, RXYR, dove X o Y è un'arginina o una lisina caricata positivamente [60]. Tali siti di scissione polibasici non sono presenti nei betacoronavirus del gruppo 2b correlati a SARS-CoV e SARS-CoV trovati negli esseri umani. Ciò può contribuire all'elevata virulenza di SARS-CoV-2 come risultato delle proteasi furiniche necessarie per l'attivazione proteolitica di S, in quanto sono espresse in modo ubiquitario nell'uomo, fornendo tropismo e patogenesi dei tessuti espansi [61]. Inoltre, uno studio ha generato un virus SARS-CoV-2 mutante privo del sito di taglio della furina (dPRRA) nella proteina spike. Il virus mutante mostrava una ridotta elaborazione della proteina spike nelle cellule Vero E6 rispetto al virus SARS-CoV-2 *wild-type*, inoltre aveva anche una ridotta replicazione nelle cellule respiratorie umane Calu3 e causava una forma di malattia attenuata in un modello di patogenesi del criceto. Questi risultati hanno dimostrato un ruolo importante del sito di taglio della furina nella replicazione e nella patogenesi del SARS-CoV-2 [62].

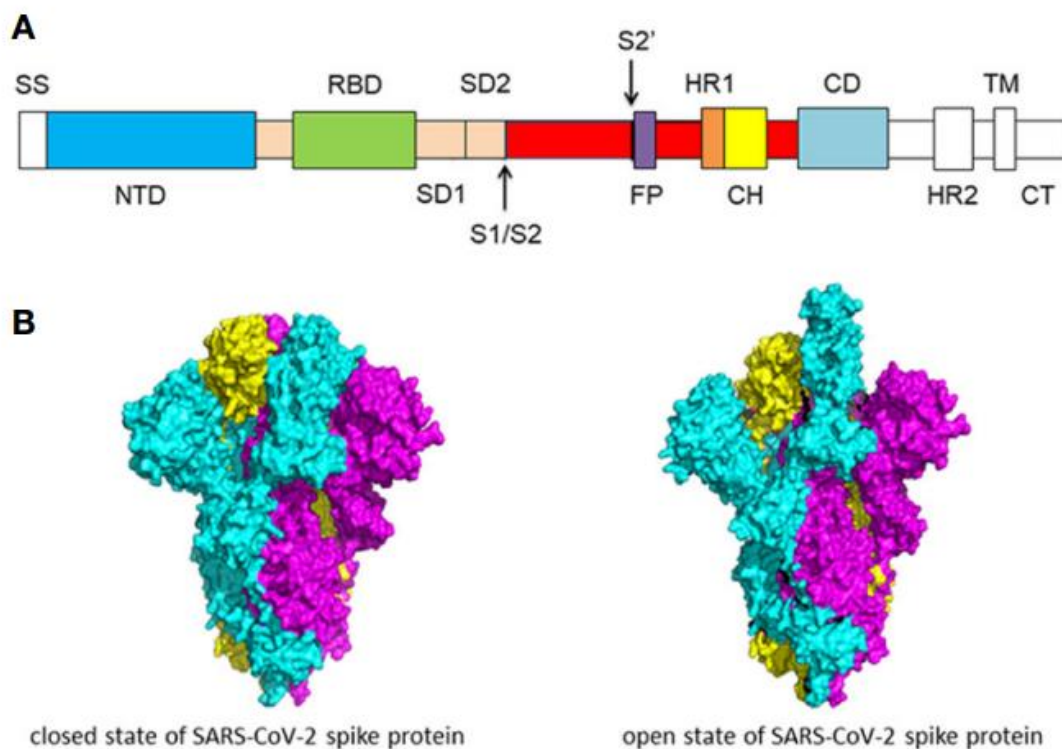


Fig. 16 (A) Schema della struttura primaria della proteina spike di SARS-CoV-2. Domini diversi sono indicati da colori diversi. SS, sequenza singola; NTD, dominio N-terminale; RBD, dominio di legame al recettore; SD1, sottodominio 1; SD2, sottodominio 2; sito di scissione della proteasi S1/S2; sito di clivaggio della proteasi S2'; FP, peptide di fusione; HR1, ripetizione eptade 1; CH, elica centrale; CD, dominio del connettore; HR2, ripetizione eptade 2; TM, dominio transmembrana; CT, coda citoplasmatica. Il sito di scissione della proteasi è indicato dalle frecce. **(B)** Struttura Cryo-EM della proteina spike di SARS-CoV-2. Lo stato chiuso (PDB: 6VXX) della glicoproteina S di SARS CoV-2 (a sinistra) e lo stato aperto (PDB: 6VYB) della proteina spike (a destra). [51]

3.2.2.3 Envelope E

Le proteine dell'envelope (E) sono un gruppo di proteine virali relativamente piccole che sostengono l'assemblaggio e il rilascio dei virioni [40]. Tra le proteine strutturali del SARS-CoV-2, la proteina E è considerata un potenziale bersaglio di farmaci. La proteina E è relativamente piccola (75 aa) e svolge un ruolo significativo nella morfogenesi e nell'assemblaggio virale [63]. La proteina E è nota per agire come le viroporine che si assemblano nella membrana ospite formando pori proteico-lipidici coinvolti nel trasporto di ioni. Le sequenze della proteina E per tutti e quattro i ceppi mostrano regioni altamente conservate tra BAT-CoV, SARS-CoV e SARS-CoV-2 mentre mostrano una leggera variazione nella sequenza con le proteine dell'envelope di MERS-CoV (**Fig. 17**).



Fig. 17 Multiallineamento delle sequenze delle proteine envelope di BAT-CoV, SARS-CoV, SARS-CoV-2 e MERS-CoV. [46]

3.2.2.4 Membrana M

Le proteine M sono proteine strutturali lunghe 222 aminoacidi che funzionano in concomitanza con le proteine E, N e S e svolgono un ruolo importante nel confezionamento dell'RNA [64]. Il tratto conservato di aminoacidi suggerisce un'architettura comune per queste proteine (**Fig. 18**). Le proteine M sono le proteine virali più abbondanti dei CoV e sono coinvolte nel fornire una forma distinta al virus. Il profilo MSA (*Multiple Sequence Alignment*) della proteina M mostra una conservazione della sequenza più elevata tra BAT-CoV, SARS-CoV e SARS CoV-2. Tuttavia, è stata osservata una notevole variazione nella sequenza della proteina M del ceppo MERS-CoV. La presenza di tre domini transmembrana è una caratteristica distintiva delle proteine M.

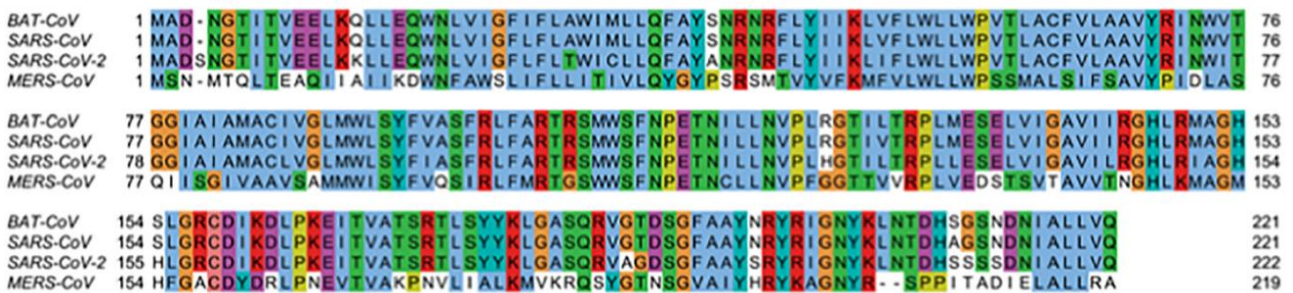


Fig. 18 Multiallineamento delle sequenze delle proteine M di BAT-CoV, SARS-CoV, SARS-CoV-2 e MERS-CoV. [46]

3.2.2.5 Nucleoproteina N

Le proteine del nucleocapside (N) svolgono un ruolo importante nel confezionamento dell'RNA virale nel ribonucleocapside [65]. La proteina N di SARS-CoV-2 è altamente conservata tra i CoV mostrando un'identità di sequenza di circa il 90% con quella di SARS-CoV. Media l'assemblaggio virale interagendo con il genoma virale e la proteina M, coinvolti nell'incremento della trascrizione e della replicazione dell'RNA virale [66]. Pertanto, le proteine N sono considerate potenziali bersagli farmacologici. Le proteine N si legano all'RNA virale attraverso il loro dominio di legame all'RNA lungo ~ 140 aminoacidi, situato nel loro nucleo [40]. Il profilo MSA della proteina N da BAT-CoV, SARS-CoV e SARS-CoV-2 mostra regioni altamente conservate (**Fig. 19**). Sulla base dell'elevata somiglianza di sequenza della proteina N, si può suggerire che gli anticorpi contro la proteina N di SARS-CoV riconoscano probabilmente la proteina N di SARS-CoV-2. Un modello simile è stato osservato per il ceppo MERS-CoV, dove regioni con leggere variazioni di sequenza suggeriscono la sua divergenza nel processo evolutivo.

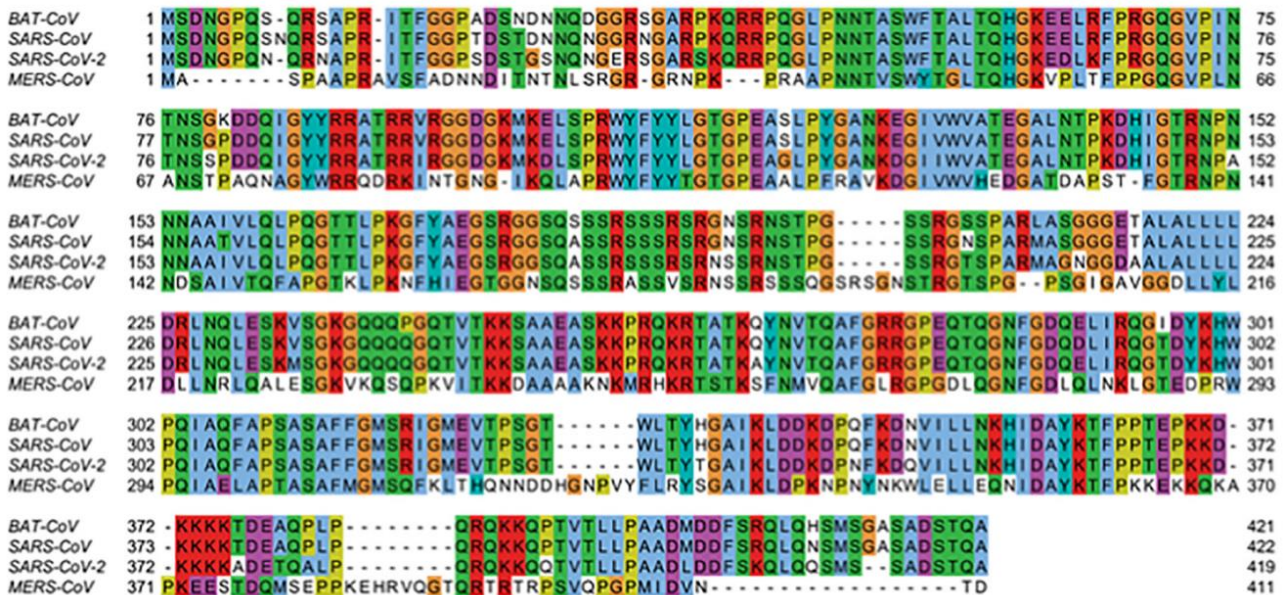


Fig. 19 Profilo MSA della proteina N di BAT-CoV, SARS-CoV, SARS-CoV-2 e MERS-CoV che mostra le regioni di conservazione della sequenza. [46]

3.2.2.6 Poliproteina replicasi

La poliproteina replicasi è un altro enzima essenziale che promuove la scissione dell'RNA ospite e la replicazione dell'RNA virale [67]. Come discusso in precedenza, gli ORF non strutturali 1a e 1ab condividono la maggior parte del contenuto nucleotidico del genoma virale. Le poliproteine replicasi sono proteine multifunzionali che svolgono vari compiti, contribuendo alla patogenesi virale [68]. Tuttavia, il ruolo principale di queste proteine è quello di sostenere la trascrizione e la replicazione dell'RNA virale. Principalmente, queste proteine sono suddivise in varie NSP come nsp1, nsp2, nsp3, nsp4 e proteasi come PLpro e 3CLpro. L'ORF1ab contiene uno specifico dominio RdRp che svolge un ruolo fondamentale nella trascrizione e nella replicazione dell'RNA virale. La poliproteina replicasi è

composta da tre domini, tra cui macro, papaina-like e il dominio proteasi principale. L'MSA della poliproteina replicasi 1a mostra una regione conservata tra tutti e tre i domini ad eccezione della proteasi principale in cui sono presenti minime regioni di conservazione della sequenza. La proteasi principale di SARS-CoV e SARS-CoV-2 mostra una somiglianza di sequenza più elevata (~ 96%) suggerendo un'evoluzione comune. Il dominio principale della proteasi di MERS-CoV non mostra alcuna somiglianza con gli altri tre ceppi virali, suggerendo una chiara divergenza nell'evoluzione. Anche la poliproteina replicasi 1ab contiene i domini di proteasi macro, papaina e principale con un'aggiunta al dominio catalitico RdRp che agisce come una poliproteina legante l'RNA. L'MSA della poliproteina replicasi 1ab suggerisce una maggiore conservazione della sequenza tra tutti e quattro i ceppi di CoV.

3.2.2.7 RdRp

La replicazione di SARS-CoV-2 è dominata da un complesso di replicazione/trascrizione che contiene diverse subunità. Il complesso è composto da proteine virali non strutturali (NSP) e il nucleo del complesso è la RdRp in nsp12. Le funzioni di nsp12 richiedono fattori accessori, tra cui nsp7 e nsp8 in quanto nsp12 da sola ha poca attività. La struttura cristallina del complesso nsp12-nsp7-nsp8 è stata identificata (**Fig. 20A**). L'RNA polimerasi RNA-dipendente, che catalizza la sintesi dell'RNA virale, è una componente critica della replicazione/trascrizione del coronavirus. L'RdRp è un importante bersaglio dei farmaci antivirali. La nsp12 del SARS-CoV-2 contiene un dominio nucleotidil transferasi associato a Nidovirus (NiRAN) e un dominio "mano destra" al C-terminale. Questi due domini sono collegati da un dominio di interfaccia. Un'unica forcina beta si osserva nel dominio di estensione N-terminale. La forcina beta forma stretti contatti per stabilizzare la struttura complessiva. L'RNA polimerasi RNA-dipendente contiene tre sottodomini: dita, palmo e pollice. La struttura beta-hairpin si inserisce nella scanalatura di bloccaggio formata dal sottodominio del palmo e dal dominio NiRAN. Nel dominio del palmo, i motivi della polimerasi A-G, altamente conservati, formano la camera del sito attivo del dominio RdRp di SARS-CoV-2. I motivi RdRp mediano la sintesi dell'RNA diretta dal template in una cavità centrale attraverso quattro percorsi accessibili al solvente caricati positivamente, tra cui il percorso di ingresso del template, il percorso di ingresso del primer, il canale di ingresso degli NTP e il percorso di uscita del filamento nascente [69]. Uno studio recente mostra la struttura cryo-EM del complesso nsp12-nsp7-nsp8 in forma attiva [70] (**Fig. 20B**). Quando è stato aggiunto un substrato minimo di RNA a forcina, il complesso nsp12-nsp7-nsp8 ha mostrato attività di estensione dell'RNA dipendente dall'RNA. La struttura del complesso RdRp-RNA mostra il complesso nsp12-nsp7-nsp8 impegnato con più di due giri di RNA duplex. La struttura dell'RdRp-RNA è simile a quella dell'enzima libero con alcune caratteristiche uniche. Rispetto all'enzima libero, il complesso RdRp-RNA contiene una regione proteica estesa in nsp8 e un RNA sporgente. La subunità nsp12 si lega con il primo giro di RNA tra i sottodomini del pollice e i sottodomini delle dita. Il sottodominio del palmo contiene il sito attivo che è formato da cinque motivi nsp12 A-E. Il motivo C interagisce con l'estremità 3' dell'RNA e include l'acido aspartico 760 e 761. I motivi nsp12 F e G si trovano nel sottodominio delle dita e hanno la funzione di posizionare il template di RNA. Quando il duplex di RNA lascia la fessura della RdRp, forma un secondo turn elicoidale, sporgente dalla superficie di nsp12. Nessun fattore strutturale nella RdRp limiterà l'estensione dell'RNA duplex. Tra le estensioni ad alfa-elica c'è il duplex dell'RNA. Le regioni N-terminali, che si trovano nelle due subunità nsp8 e sono altamente conservate, formano le estensioni alfa-elicoideali. Queste estensioni

di nsp8 utilizzano i residui caricati positivamente per interagire con i backbone dell'RNA. L'nsp8 potrebbe funzionare come dei "poli scorrevoli", scivolando lungo l'RNA sporgente per impedire alla RdRp di dissociarsi prematuramente durante la replicazione. Il sito di legame del trifosfato è conservato. È probabile che i residui D623, S682 e N691 interagiscano con il gruppo 2'-OH del trifosfato (NTP), rendendo l'RdRp specifica per la sintesi dell'RNA rispetto al DNA [71].

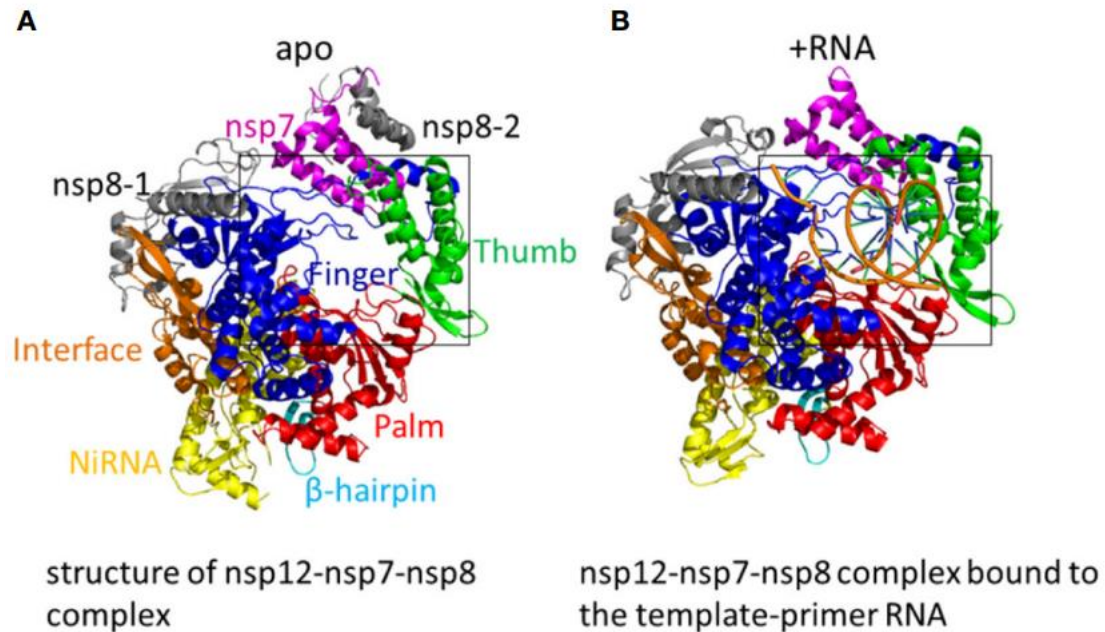


Fig. 20 (A) La struttura del complesso nsp12-nsp7-nsp8. Colori: nsp7, magenta; nsp8-1 e nsp8-2, grigio; β -hairpin, ciano; NiRAN, giallo; l'interfaccia, arancione; il dominio delle dita, blu; il dominio del palmo, rosso; il dominio del pollice, verde. (PDB: 6M71) **(B)** La struttura della RNA polimerasi RNA-dipendente (RdRp) di SARS-CoV-2 in forma attiva. Il complesso nsp12-nsp7-nsp8 si lega all'RNA templatoprimer. (PDB: 7BV2). [51]

3.2.2.8 Proteasi principale

La proteasi principale (Mpro) di SARS-CoV-2 svolge un ruolo fondamentale nel mediare la replicazione e la trascrizione dei geni virali. Mpro idrolizza la poliproteina in almeno undici siti conservati e inizia con la scissione di pp1a e pp1b di Mpro. Considerando l'assenza di omologhi strettamente correlati nell'uomo, insieme all'importanza funzionale della proteasi principale nel ciclo di vita del virus, la proteasi principale è un attraente bersaglio antivirale. La simmetria cristallografica mostra che Mpro forma un omodimero (protomero A e protomero B). Ogni protomero contiene tre sottodomini, vale a dire dominio I, dominio II e dominio III. Un lungo loop collega il dominio II e il dominio III. La fessura tra il dominio I e il dominio II si trova nella tasca di legame del substrato, che presenta i residui della diade catalitica His41 e Cys145 [72]. Come per tutti i coronavirus, i siti attivi di Mpro sono altamente conservati e sono costituiti da quattro siti: S1', S1, S2 e S4. Nel sito S1', il tiolo di una cisteina fissa gli inibitori mediante un legame covalente. Per gli inibitori, il legame covalente è fondamentale per mantenere la loro attività antivirale [73].

3.2.3 Ciclo replicativo SARS-CoV-2

In quanto microrganismo intracellulare obbligato, il coronavirus sfrutta il meccanismo della cellula ospite per la propria replicazione e diffusione. L'ingresso di SARS-CoV-2 nelle cellule ospiti è mediato dalla glicoproteina S transmembrana che forma omotrimeri che sporgono dalla superficie virale [74]. L'RBD di SARS-CoV-2 si lega direttamente al dominio peptidico dell'enzima di conversione dell'angiotensina (ACE2) [74,75,76,77] (**Fig. 21**).



Fig. 21 La struttura complessiva dell'RBD di SARS-CoV-2 legato al recettore ACE2. ACE2 è di colore ciano, il nucleo dell'RBD è di colore verde (PDB: 6MOJ). [51]

ACE2 è una proteina di membrana di tipo I che partecipa alla maturazione dell'angiotensina, un ormone peptidico che controlla la vasocostrizione e la pressione sanguigna [78]. Nel tratto respiratorio ACE2 è ampiamente espresso sulle cellule epiteliali di alveoli, trachea, bronchi, ghiandole sierose bronchiali [79], monociti e macrofagi alveolari [80].

Dopo il legame dell'RBD nella subunità S1 al recettore ACE2, la proteina S di SARS-CoV-2 viene scissa dalla serina proteasi 2 TMPRSS2 transmembrana associata alla superficie cellulare, che attiva il dominio S2 per la fusione della membrana virale e cellulare [81].

Come SARS-CoV e altri coronavirus, SARS-CoV-2 probabilmente entra nelle cellule bersaglio attraverso l'endocitosi mediata da recettore, dove la fusione del virus avvolge le membrane dell'endosoma e porta al rilascio del nucleocapside virale nel citosol della cellula infetta [82].

Dopo il rilascio e la rimozione dell'RNA virale nel citoplasma, la replicazione del coronavirus inizia con la traduzione di ORF1a e ORF1b nelle poliproteine pp1a e pp1ab tramite un meccanismo di frameshifting [83]. Successivamente, le poliproteine pp1a e pp1ab vengono processate da proteasi virali interne, inclusa la proteasi principale Mpro [84]. La scissione della poliproteina produce 15 proteine replicasi mature, che si assemblano in un complesso di replicazione-trascrizione che si occupa della sintesi dell'RNA a filamento negativo. Vengono prodotti sia RNA subgenomici a filamento negativo a lunghezza intera che multipli. Quelli a lunghezza intera fungono da modello per nuovi RNA genomici a lunghezza intera e gli altri dirigono la sintesi degli mRNA subgenomici necessari per esprimere i geni per le proteine strutturali e accessorie che risiedono nel quarto prossimale al 3' del genoma [82]. La replicazione dell'RNA del coronavirus avviene su una rete reticolovescicolare, indotta dal virus, di membrane modificate del reticolo endoplasmatico (ER) [85].

L'assemblaggio dei virioni è eseguito rapidamente con l'accumulo di nuovo RNA genomico e componenti strutturali. I complessi della proteina N con l'RNA del genoma formano strutture elicoidali. Quindi, la proteina M transmembrana, localizzata nelle membrane intracellulari del compartimento intermedio ER-Golgi (ERGIC), interagisce con le altre proteine strutturali virali (proteine S, E ed N) per consentire la gemmazione dei virioni [86,87]. Dopo l'assemblaggio e la gemmazione, i virioni vengono trasportati in vescicole e infine rilasciati per esocitosi (**Fig. 22**).

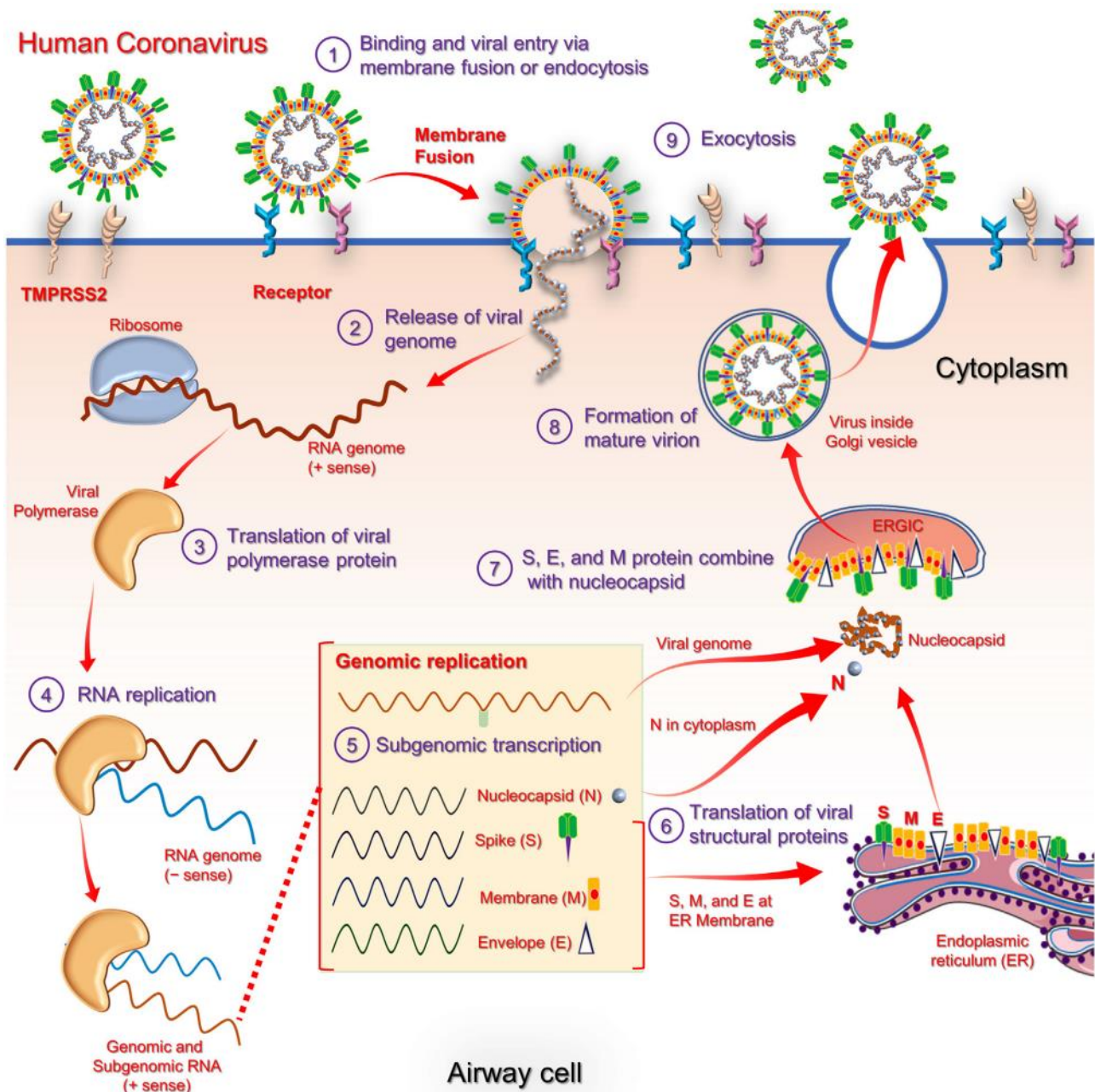


Fig. 22 Una panoramica del ciclo vitale dei coronavirus all'interno della cellula ospite. [3]

3.2.4 Varianti SARS-CoV-2

Le varianti si sviluppano attraverso cambiamenti nucleotidici che emergono naturalmente nel genoma virale durante la replicazione (**Fig. 23**), con questi cambiamenti che si verificano a una velocità maggiore nei virus a RNA rispetto ai virus a DNA [89]. Tuttavia, la velocità con cui si verificano questi cambiamenti nucleotidici nei Coronavirus è significativamente inferiore a quella di altri virus a RNA poiché possiedono un enzima che corregge alcuni errori commessi durante la replicazione [90]. La proteina non strutturale 14 di SARS-CoV-2 (nsp14) contiene attività esoribonucleasica (ExoN) che ha dimostrato di avere un effetto di "correzione di bozze"; inoltre, è stato dimostrato che l'inattivazione di ExoN è dannosa per la replicazione di SARS-CoV-2 e MERS-CoV [91,92]. Le alterazioni genomiche vantaggiose rispetto alla replicazione virale, alla trasmissione e all'evasione immunitaria aumenteranno di frequenza in una popolazione a causa della selezione naturale [93]. L'evasione immunitaria si verifica quando le alterazioni genomiche rendono inefficace la risposta immunitaria a causa della sua incapacità di riconoscere ed eliminare il virus. Al contrario, le alterazioni genomiche che riducono la fitness virale saranno eliminate dalla popolazione dei virus circolanti. Le alterazioni genomiche che hanno scarso effetto sulla fitness virale possono anche aumentare o diminuire di frequenza solo per caso e contribuire al pool di varianti circolanti. Varie opzioni terapeutiche (plasma di convalescenza, anticorpi monoclonali (mAb) e vaccini) nonché fattori ambientali agiscono anche come pressioni selettive per nuove varianti e contribuiscono alla persistenza di quest'ultime. Se le varianti modificano in modo dimostrabile il fenotipo (virulenza e trasmissione) di un virus, vengono chiamate ceppi [89,94].

Oltre agli errori di replicazione, anche le pressioni derivate dall'ospite come l'editing dell'RNA o la modifica dell'RNA possono essere essenziali per la formazione di varianti SARS-CoV-2 (**Fig. 23**). All'analisi del genoma di SARS-CoV-2 sono stati osservati ampi modelli di sostituzione nucleotidica citosina-uracile e adenosina-inosina [95,96]. Questi modelli sono stati interpretati come firme degli enzimi di editing dell'RNA, come gli enzimi APOBEC e ADAR1 [95,96]. Gli ospiti hanno sviluppato questi enzimi di modifica dell'RNA come meccanismi immunitari innati virali sensoriali [97]; tuttavia, i virus possono sfruttare questi meccanismi per il loro potenziale evolutivo. È stato quindi suggerito che la variabilità intra-ospite osservata nei pazienti COVID-19 sia guidata dagli enzimi di modifica dell'RNA dell'ospite [98], ma rimane poco chiaro come SARS-CoV-2 manipoli questi enzimi per introdurre cambiamenti nucleotidici nel suo genoma virale [95,96,97,98].

La ricombinazione si verifica in cellule infette da più varianti, in cui il materiale genetico di due varianti è impacchettato in un singolo virione (**Fig. 23**). È stato dimostrato che due varianti di SARS-CoV-2 possono infettare una persona contemporaneamente [99]. Questi virioni ricombinati possono possedere diverse proprietà patogene con gravi implicazioni per le contromisure SARS-CoV-2, specialmente se i ricombinanti possono sfuggire all'immunità sia naturale che indotta dal vaccino [100].

L'evoluzione virale accelerata e la diffusione prolungata del virus competente per la replicazione sono state osservate in individui immunocompromessi affetti da infezione cronica da SARS-CoV-2 [101,102,103,104]. Diversi studi hanno rivelato una forte pressione selettiva su SARS-CoV-2 durante le terapie immunitarie come i trattamenti con plasma di convalescenza e mAb, che hanno dimostrato di essere associati all'emergere di varianti virali con ridotta suscettibilità agli anticorpi neutralizzanti [101,102,105].

Le varianti SARS-CoV-2 sono classificate dall'OMS in due tipi: varianti di preoccupazione (VOC) e varianti di interesse (VOI). Diverse VOC sono emerse dal ceppo *wild-type* originale isolato a Wuhan dall'inizio dell'epidemia nel dicembre 2019. Secondo il Center for Disease Control (CDC), una VOC ha una maggiore trasmissibilità, una maggiore virulenza, resistenza al vaccino o ha acquisito l'immunità da una precedente infezione e ha la capacità di eludere il rilevamento diagnostico [106].

A marzo 2023, tuttavia, l'OMS ha deciso di rielaborare completamente la classificazione delle varianti sotto monitoraggio, valutando ciascuna di essa singolarmente e in maniera indipendente. Tale decisione è stata presa con l'intento di fornire una classificazione più chiara. Oltre alle varianti VOC e VOI è stata introdotta una terza categoria denominata VUM (varianti sotto monitoraggio). L'OMS, inoltre, ha aggiornato anche le definizioni di queste tre categorie. L'aggiornamento principale mira a rendere più specifica la definizione di variante di preoccupazione (VOC), al fine di includere i principali passaggi evolutivi del SARS-CoV-2 che richiedono importati interventi di sanità pubblica. Infine, in futuro verranno attribuite lettere greche solamente alle varianti di preoccupazione (VOC) e non più alle varianti di interesse (VOI).

La variante Alpha (B.1.1.7) è stata rilevata per la prima volta alla fine di settembre 2020 ed è diventata rapidamente il ceppo predominante nel Regno Unito (UK) [107,108]. La variante Beta (B.1.351), rilevata per la prima volta nell'ottobre 2020, è diventata il ceppo dominante con conseguente seconda ondata in Sud Africa [109]. Allo stesso modo, la variante Gamma (P.1) è stata rilevata in quattro brasiliani che si sono recati in Giappone nel gennaio 2021 [110] ed è stata responsabile della recrudescenza delle infezioni a Manaus, nonostante gli alti livelli di infezione precedente nel paese [111,112]. La variante Delta (B.1.617.2), rilevata per la prima volta nel dicembre 2020, è stata responsabile del massiccio aumento dei casi che hanno causato una seconda ondata in India [113,114] e delle infezioni in più cluster negli Stati Uniti d'America [115]. La variante Omicron (B.1.1.529), designata VOC dall'OMS, è stata rilevata per la prima volta nel novembre 2021 da laboratori di sorveglianza genomica di livello mondiale in Sud Africa ed è stata trovata in molti paesi in tutto il mondo [116,117]. L'emergere di queste varianti è preoccupante poiché possono influenzare la trasmissibilità virale, la virulenza e il tasso di reinfezione sfuggendo all'immunità naturale e a quella indotta dal vaccino [118].

Sono state segnalate molte altre VOI che si prevede influenzino solo la trasmissione, la virulenza e l'immunità acquisita o vaccinale. Le VOI e le varianti monitorate includono Epsilon (B.1.427/B.1.429) identificata in California, Zeta (P.2) identificata in Brasile, Eta (B.1.525) identificata in Nigeria e Regno Unito, Theta (P.3) identificata nelle Filippine, Iota (B.1.526/B.1.526.1) identificata a New York, Kappa (B.1.617) e Delta Plus (B.1.617.2.1) identificate in India, Lambda (C.37) identificata in Perù e Mu (B.1.621) identificata in Colombia [106,119,120].

Le VOC di SARS-CoV-2 che sono emerse condividono firme comuni nella loro proteina S, ma ogni variante mostra anche nuovi cambiamenti. Esistono quattro principali sostituzioni di amminoacidi nell'RBD che sono state al centro di studi che valutano la virulenza e l'evasione immunitaria. Il primo è l'N501Y trovato nel sito di legame per ACE2 dell'RBD ed è comune alle varianti Alpha, Beta, Gamma e Omicron [121,122]. La seconda e la terza sostituzione, E484K/Q/A e K417T/N, sono presenti nei ceppi Beta, Gamma e Omicron [121,122]. Il quarto, L452R, è unico per la variante Delta [121]. Inoltre, la variante Omicron ha tra 26 e 32 cambiamenti amminoacidici nella proteina S che sono al centro delle indagini [123]. Questa variante, pur avendo alcune sostituzioni in comune con Beta e Delta, ha un percorso evolutivo distinto [124].

N501Y è comune tra le varianti Alpha, Beta e Gamma e ha dimostrato di aumentare l'infettività cellulare nei modelli animali [125]. L'asparagina in posizione 501 (N501) è uno dei sei aminoacidi essenziali coinvolti nel legame del recettore ACE2 [126] e la sua sostituzione con tirosina (Y) ha dimostrato di aumentare l'affinità di legame con il recettore dell'ospite [127]. Ciò ha probabilmente contribuito alla maggiore trasmissibilità e alla maggiore carica virale osservata nelle varianti Alpha [108,128] e Beta [109]. L'analisi iniziale basata su studi caso controllo abbinati nel Regno Unito ha indicato che la variante Alpha non era significativamente associata a un aumento del rischio di ospedalizzazione o mortalità negli individui infetti. Tuttavia, ulteriori analisi indicano che la variante è associata a un aumento della gravità e di conseguenza a un aumento del tasso di mortalità fino al 61% [128,129,130]. La variante Alpha contiene anche la sostituzione P681H che si trova in un'area adiacente al sito di scissione della furina che è nota per essere importante nell'infezione e nella trasmissione [131,132]. La delezione DH69/V70 presente anche nella variante Alpha migliora l'infettività virale in vitro ed è stata collegata a determinati kit commerciali che non riescono a rilevare la glicoproteina S [133].

Le sostituzioni amminoacidiche nella proteina S della variante Beta sono più estese che in Alpha presentando dieci cambiamenti con tre nell'RBD. Il complesso di interazione RBD-ACE2 è stato analizzato strutturalmente utilizzando metodi in silico per valutare l'impatto delle sostituzioni N501Y, K417T ed E484K [134]. Il residuo N501 è importante per l'interazione con l'ACE2 mentre non si prevede che K417 ed E484 svolgano un ruolo importante. Gli ultimi due residui possono infatti ridurre l'affinità di legame, il che dimostra che l'aumento della trasmissibilità osservato nella variante Beta è dovuto a N501Y o ad altre alterazioni del virus. Nell'ottobre 2020, la variante Beta rappresentava l'11% del virus SARS-CoV-2 sequenziato in Sud Africa, ma a dicembre di quell'anno l'87% dei casi era dovuto a questa variante [135]. A Cape Town una soglia di 100.000 casi è stata raggiunta il 50% più rapidamente nella seconda ondata con la variante beta rispetto alla prima ondata causata dalla variante *wild-type*.

La variante Gamma è stata responsabile dell'ondata di infezioni in Brasile alla fine del 2020 che ha spinto il sistema sanitario sull'orlo del collasso. Ciò si è verificato nonostante un livello di infezione precedentemente elevato durante la prima ondata, indicando che la variante non era suscettibile all'immunità acquisita naturalmente. Gamma contiene 17 sostituzioni amminoacidiche. La modellazione computazionale ha dimostrato che la variante era 1,4-2,2 volte più trasmissibile rispetto alla variante *wild-type* [111,136]. I rapporti hanno indicato che la variante Gamma può essere più grave con un numero maggiore di individui più giovani che presentano una malattia avanzata e soccombono al virus [137,138].

Una variante dominante con otto modifiche della proteina S, la variante Delta, è stata responsabile della grave seconda ondata che ha colpito l'India all'inizio del 2021 e poi si è diffusa ampiamente anche in altri paesi [114]. Questa variante ha provocato una terza ondata devastante in Sud Africa e ha rapidamente soppiantato la variante Beta entro tre mesi dall'introduzione nel paese: ciò è stato attribuito al vantaggio di trasmissione del 46% rispetto a Beta [139]. Una sostituzione amminoacidica trovata nell'RBD, L452R, e una nel sito di clivaggio della furina della proteina S, P681R, sono state implicate nell'aumentata trasmissione associata a questa variante. È stato dimostrato che la sostituzione P681R, trovata anche nella variante Alpha, aumenta l'attività di fusione con il recettore dell'ospite, il che potrebbe aumentare l'infettività e la velocità di trasmissione [113]. Tuttavia, la mutazione P681R è stata precedentemente identificata in varianti che non aumentavano di

frequenza in modo efficiente, indicando che un numero cumulativo di mutazioni potrebbe essere responsabile di questo aumento della trasmissione. Dati epidemiologici e genomici preliminari hanno dimostrato che la variante Delta è più trasmissibile del ceppo originale e di altre varianti [140,141]. Anche il rischio di ospedalizzazione è risultato doppio rispetto a quello della variante Alpha [142], e i pazienti ospedalizzati con infezione da Delta presentavano una malattia più grave e un tasso di mortalità in ospedale più elevato [143].

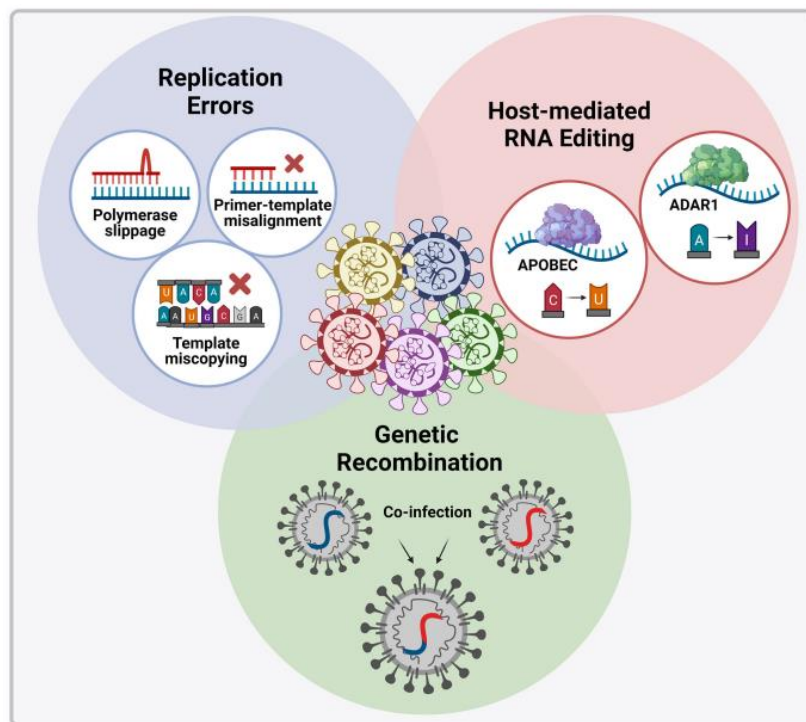


Fig. 23 Meccanismi molecolari che introducono alterazioni genomiche nel genoma di SARS-CoV-2. I cambiamenti dei nucleotidi possono emergere naturalmente nel genoma virale attraverso i vari errori di replicazione mostrati. Oltre agli errori di replicazione, gli enzimi di modifica dell'RNA derivati dall'ospite, come gli enzimi APOBEC e ADAR1 possono introdurre sostituzioni puntiformi [citosina (C) in uracile (U) e adenosina (A) a inosina (I)] nel genoma virale. Infine, se due varianti virali co-infettano la stessa cellula, può verificarsi una ricombinazione per cui il materiale genetico delle due varianti è impacchettato in un singolo virione. [88]

4. Scopo del progetto

Il mio lavoro di ricerca, sviluppato presso il laboratorio di genetica medica del CAST (Center for Advanced Studies and Technology) dell'Università degli Studi "G. d'Annunzio" Chieti – Pescara, sotto la supervisione del Prof. Liborio Stuppia e del Dott. Federico Anaclerio, è incentrato sulla analisi della sequenza del virus SARS-CoV-2 e successiva caratterizzazione dello specifico lignaggio di appartenenza facendo uso di una piattaforma di sequenziamento di nuova generazione (NGS).

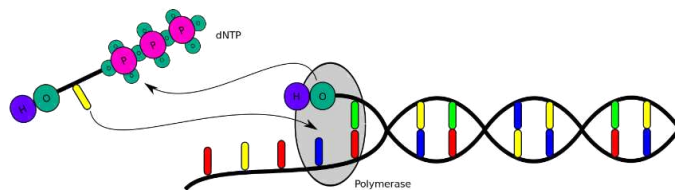
In particolare, sono intervenuto nelle varie fasi del processo di sequenziamento, dalla preparazione dei campioni fino all'analisi dei dati di sequenza virale e alla individuazione di varianti di potenziale rilevanza clinica.

L'NGS ha svolto un ruolo centrale dall'inizio della pandemia. Questa tecnologia ha contribuito a determinare nuovi ospiti, l'evoluzione e i modelli di diffusione del SARS CoV-2, che sono stati fondamentali per la progettazione di strategie efficaci di prevenzione e controllo delle malattie. La velocità di acquisizione e analisi dei dati è stata inconcepibilmente rapida, il che ha reso applicabili i metodi NGS per monitorare e contrastare la diffusione del virus.

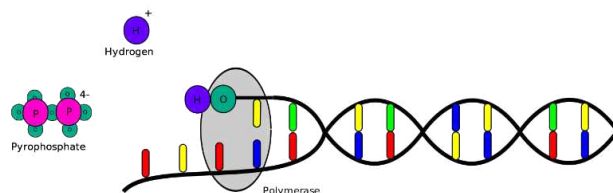
Gli strumenti utilizzati nella diagnostica clinica appartengono a una delle aziende leader nel sequenziamento NGS, la "Thermo Fisher Scientific", e prendono il nome di "ION Chef Instrument" e "S5 Plus Instrument".

4.1 Sequenziamento ionico a semiconduttori

Il sequenziamento ionico a semiconduttori è un metodo di sequenziamento basato sulla rilevazione di ioni idrogeno che vengono rilasciati durante la polimerizzazione del DNA (**Fig. 24**). È un metodo di "sequenziamento per sintesi", durante il quale viene costruito un filamento complementare sulla base della sequenza di un filamento modello.



Polymerase integrates a nucleotide.



Hydrogen and pyrophosphate are released.

Fig. 24 Incorporazione di un dNTP dalla DNA polimerasi e liberazione di uno ione idrogeno più una molecola di pirofosfato.

Un micropozzetto contenente un filamento di DNA modello da sequenziare viene riempito con una singola specie di deossiribonucleotide trifosfato (dNTP). Se il dNTP introdotto è complementare al nucleotide modello principale, viene incorporato nel filamento complementare in crescita. Ciò provoca il rilascio di uno ione idrogeno che attiva un sensore ionico ISFET, che indica che si è verificata una reazione (**Fig. 25**). Se nella sequenza dello stampo sono presenti ripetizioni omopolimeriche, più molecole dNTP saranno incorporate in un singolo ciclo. Ciò comporta un numero corrispondente di ioni idrogeno rilasciati e un segnale elettronico proporzionalmente più alto.

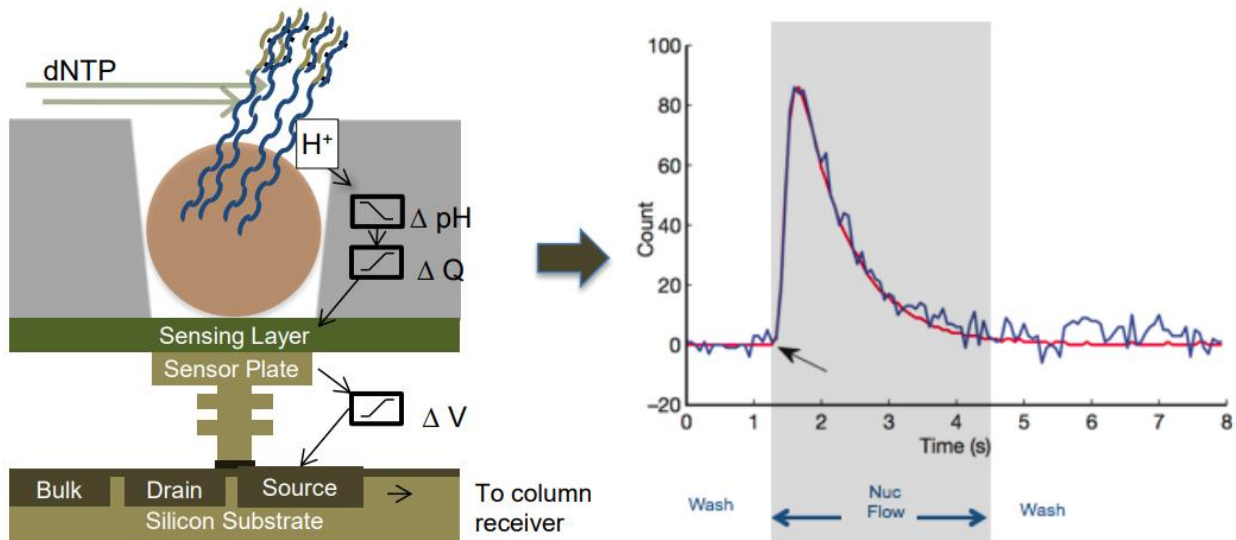


Fig. 25 A sinistra rappresentazione schematica di un dispositivo ISFET (ion-sensitive field-effect transistor) con sopra un micropozzetto e una sfera legata ai frammenti di DNA da sequenziare. A destra variazione del voltaggio proporzionale al cambiamento di pH causato dal rilascio di ioni H^+ .

Questa tecnologia differisce da altre tecnologie di sequenziamento per sintesi in quanto non vengono utilizzati nucleotidi modificati o misurazioni ottiche.

In natura, l'incorporazione di un desossiribonucleoside trifosfato (dNTP) in un filamento di DNA in crescita comporta la formazione di un legame covalente e il rilascio di pirofosfato e di uno ione idrogeno caricato positivamente [144,145,146]. Un dNTP sarà incorporato solo se è complementare al nucleotide presente nel modello. Il sequenziamento ionico a semiconduttori sfrutta questa caratteristica determinando se uno ione idrogeno viene rilasciato fornendo una singola specie di dNTP alla reazione.

I micropozzetti su un chip semiconduttore, che contengono ciascuno molte copie di una molecola di DNA modello a filamento singolo da sequenziare e la DNA polimerasi, vengono inondati in sequenza con i quattro dNTP non modificati [144,147,148]. Se un dNTP introdotto è complementare al successivo nucleotide spaiato sul filamento stampo, viene incorporato nel filamento complementare in crescita dalla DNA polimerasi [149]. Se il dNTP introdotto non è complementare non c'è incorporazione e nessuna reazione biochimica. Lo ione idrogeno che viene rilasciato nella reazione modifica il pH della soluzione, che viene rilevato da un ISFET [144,145,147]. Le molecole di dNTP non

attaccate vengono lavate via prima del ciclo successivo quando viene introdotta una diversa specie di dNTP [147].

Sotto lo strato di micropozzetti c'è uno strato sensibile agli ioni, al di sotto del quale si trova un sensore di ioni ISFET [150]. Tutti gli strati sono contenuti all'interno di un chip semiconduttore CMOS, simile a quello utilizzato nell'industria elettronica [150,151].

Ciascun chip contiene una serie di micropozzetti con i corrispondenti rivelatori ISFET [147]. Ogni ione idrogeno rilasciato attiva quindi il sensore di ioni ISFET. La serie di impulsi elettrici trasmessi dal chip a un computer viene tradotta in una sequenza di DNA, senza che sia richiesta alcuna conversione intermedia del segnale [147,152]. Poiché gli eventi di incorporazione dei nucleotidi sono misurati direttamente dall'elettronica, si evita l'uso di nucleotidi marcati e misurazioni ottiche [150,151]. L'elaborazione del segnale e l'assemblaggio del DNA possono quindi essere eseguiti tramite software.

I principali vantaggi del sequenziamento ionico a semiconduttori sono l'alta velocità di sequenziamento e i bassi costi iniziali e operativi [148,152]. Ciò è stato reso possibile evitando nucleotidi modificati e misurazioni ottiche.

Poiché il sistema registra gli eventi naturali di incorporazione dei nucleotidi mediati dalla polimerasi, il sequenziamento può avvenire in tempo reale. In realtà, la velocità di sequenziamento è limitata dal ciclo dei nucleotidi attraverso il sistema [153]. Se i chip semiconduttori vengono migliorati (come previsto dalla legge di Moore), il numero di letture per chip (e quindi per esecuzione) dovrebbe aumentare [152].

Il costo per l'acquisizione di un sequenziatore da Ion Torrent Systems Inc. al momento del lancio era di circa \$ 50.000 USD, escluse le apparecchiature per la preparazione dei campioni e un server per l'analisi dei dati [148,152,154]. Il costo per corsa è anche significativamente inferiore a quello dei metodi di sequenziamento automatizzati alternativi, circa \$ 1.000 [148,155].

Se sul filamento modello (filamento da sequenziare) sono presenti ripetizioni omopolimeriche dello stesso nucleotide (ad es. TTTT), vengono incorporati più nucleotidi introdotti e vengono rilasciati più ioni idrogeno in un singolo ciclo. Ciò si traduce in una maggiore variazione del pH e in un segnale elettronico proporzionalmente maggiore [152]. Questa è una limitazione del sistema in quanto è difficile enumerare lunghe ripetizioni. Questa limitazione è condivisa da altre tecniche che rilevano aggiunte di singoli nucleotidi come il pirosequenziamento [156]. I segnali generati da un numero di ripetizioni elevato sono difficili da differenziare dalle ripetizioni di un numero simile ma diverso; ad esempio, le ripetizioni omografiche di lunghezza 7 sono difficili da differenziare da quelle di lunghezza 8.

Un'altra limitazione di questo sistema è la breve lunghezza delle *read* rispetto ad altri metodi di sequenziamento come il sequenziamento di Sanger o il pirosequenziamento. *Read* di lunghezza maggiore sono vantaggiose per l'assemblaggio del genoma de novo. I sequenziatori a semiconduttore Ion Torrent producono *read* con una lunghezza media di circa 400 nucleotidi [145,148].

Il rendimento è attualmente inferiore a quello di altre tecnologie di sequenziamento ad alto rendimento, anche se gli sviluppatori sperano di migliorarlo aumentando la densità dei chip [145].

4.1.1 Ion Torrent workflow

Il processo di sequenziamento prevede varie fasi: preparazione libreria, amplificazione clonale, isolamento delle sfere con i frammenti da sequenziare, caricamento del chip, sequenziamento e analisi dei dati (Fig. 26).

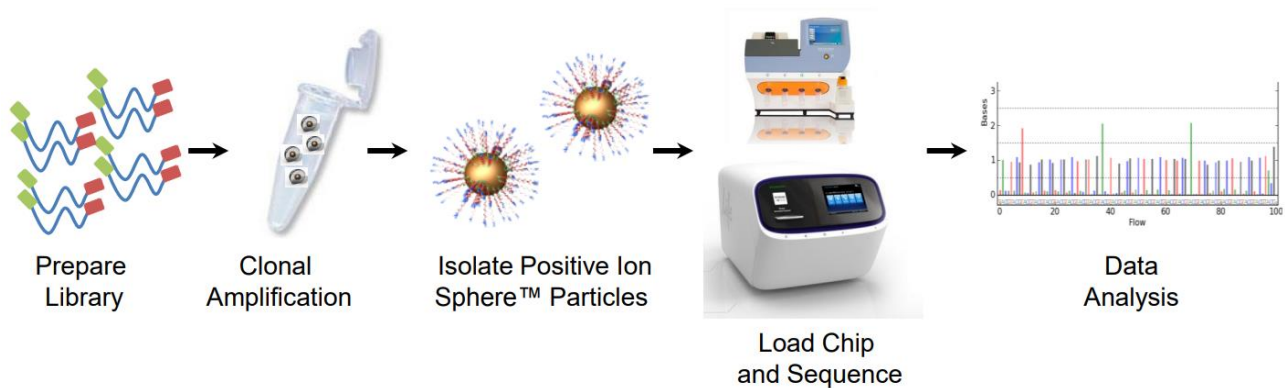


Fig. 26 Flusso di lavoro del sistema NGS Ion Torrent.

La preparazione di una libreria rappresenta il primo passo per il sequenziamento di nuova generazione. Prima che i campioni di DNA o RNA possano essere sequenziati, gli acidi nucleici devono essere isolati, frammentati, bisogna renderne piatte (blunt) le estremità e vanno legati in modo covalente agli adattatori utilizzando metodi di ligazione. La complessità di una libreria NGS ben preparata deve riflettere in modo completo e accurato la complessità del campione, ma può comportare anche delle distorsioni (*bias*) introdotte durante la sua preparazione. Un obiettivo chiave nella preparazione di una libreria di DNA o RNA per il sequenziamento di nuova generazione è quello di massimizzare la complessità riducendo al contempo le distorsioni introdotte da PCR e amplificazione, poiché la qualità della libreria può avere un forte impatto sui risultati del NGS.

Una volta ottenuta la libreria si può procedere con la fase di amplificazione in emulsione (ePCR). La PCR in emulsione viene utilizzata per amplificare clonalmente i modelli di DNA sulla superficie delle sfere, che vengono successivamente collocate nei micropozzetti dove si verifica la reazione di sequenziamento.

Un'amplificazione clonale ideale dovrebbe portare a delle sfere con tanti frammenti della stessa sequenza. Possono verificarsi anche casi in cui ci sono delle sfere senza frammenti legati o sfere con frammenti di sequenza differente (*"mixed" read*). Proprio per questo la concentrazione del DNA nella libreria va ottimizzata in modo da avere un'amplificazione il più efficiente possibile. Un input di DNA troppo basso si traduce in poche sfere disponibili per il sequenziamento, al contrario un input troppo alto può portare ad una grande quantità di *"mixed" read*.

Dopo avere isolato le sfere si può procedere alla fase di sequenziamento vera e propria.

Con il termine "flow" si fa riferimento all'esposizione del chip ad un particolare dNTP seguito da una fase di lavaggio. Invece, con "ciclo" si intende il passaggio dei 4 dNTP all'interno del chip (un ciclo corrisponde a 4 flow consecutivi).

All'interno dei micropozzetti del chip per ogni base della sequenza da sequenziare viene ripetuto un ciclo, in questo modo verrà rilevato il dNTP corretto che si andrà ad appaiare al filamento modello presente sulla sfera.

I segnali elettrici ottenuti tramite il chip vengono riportati in uno ionogramma (**Fig. 27**). Nel grafico per ogni "flow" avremo un segnale di voltaggio che sarà proporzionale al numero di nucleotidi aggiunto dalla polimerasi. Ovviamente un segnale nullo sta ad indicare che non è avvenuta nessuna incorporazione.

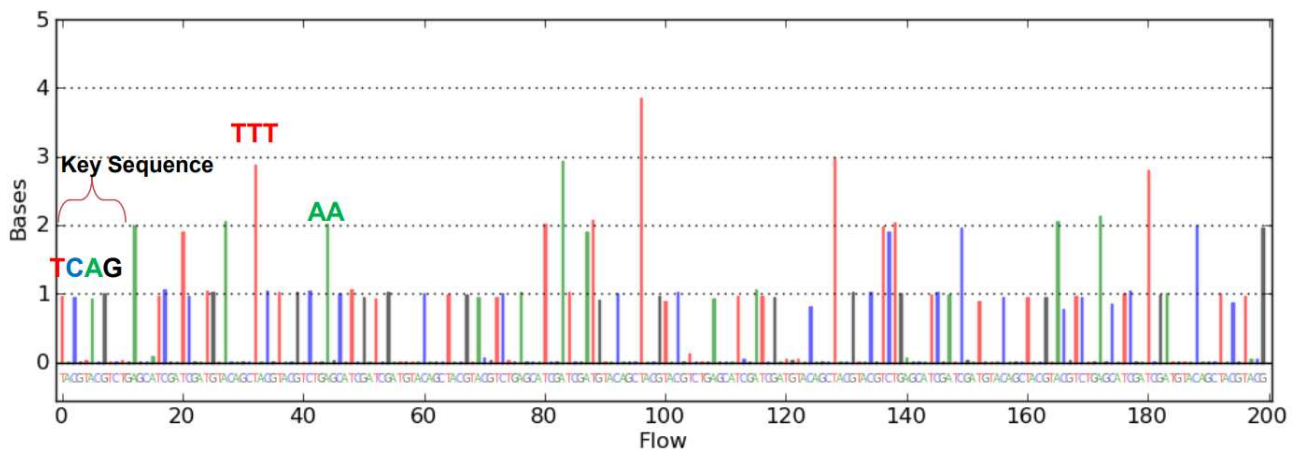


Fig. 27 Ionogramma. L'altezza di ogni barra colorata rappresenta il numero di nucleotidi incorporati per ogni "flow".

Il laboratorio del CAST di Chieti è dotato di un sequenziatore "Ion GeneStudio S5 Plus" più lo strumento "Ion Chef" entrambi della "Thermo Fisher Scientific" (**Fig. 28**). Lo strumento Ion Chef è una stazione robotizzata che, partendo da una libreria di DNA con gli opportuni adattatori, esegue: l'amplificazione in emulsione, l'arricchimento delle sfere con un elevato carico di template e il caricamento di tali sfere nel chip. Il risultato finale sarà un chip completamente caricato e pronto per il montaggio sul sequenziatore.

Questo sistema supporta un numero di read che va da 2-3 milioni fino a 100-130 milioni. La lunghezza delle read può arrivare a 600 bp, mentre i tempi di sequenziamento variano tra le 2.5 e le 4 ore. Ovviamente questi dati sono molto variabili e dipendono sia dal modello di sequenziatore sia dal chip utilizzato.

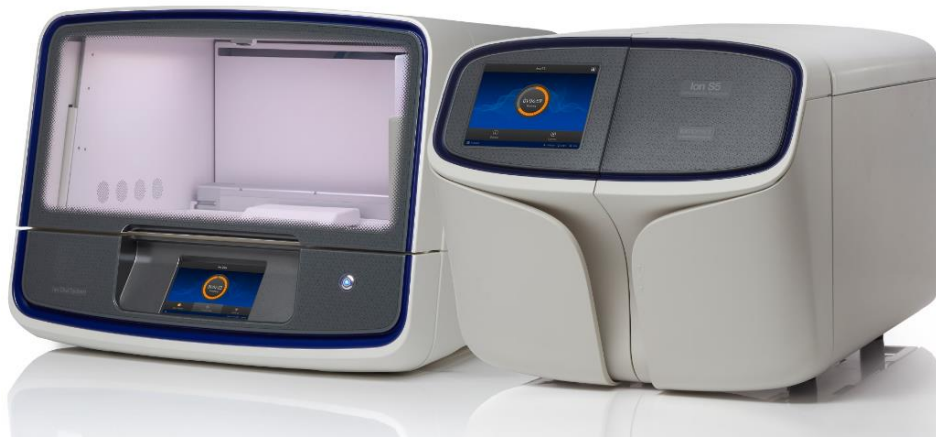


Fig. 28 A destra sequenziatore “*Ion GeneStudio S5 Plus*”, a sinistra strumento “*Ion Chef*”.

5. Materiali e metodi

Per realizzare il lavoro sono stati presi un totale di 232 campioni in un periodo compreso tra aprile 2022 e aprile 2023. I campioni da sequenziare sono stati scelti in base a vari criteri come: cicli di amplificazione, provenienza geografica dei pazienti, età dei pazienti, provenienza del campione (reparto ospedaliero, drive-in ecc.) ed anche positività in RT-PCR ai 3 target principali (ORF1ab, gene S e gene N).

I tamponi arrivati in laboratorio, prima di poter essere sequenziati, andavano incontro ad una fase di estrazione dell'RNA e poi alla qRT-PCR. Successivamente i campioni opportuni venivano recuperati dalle piastre con l'RNA estratto in precedenza. Tali campioni subivano un singolo ciclo di retrotrascrizione in termociclatore per poter essere convertiti in cDNA prima di procedere con la preparazione delle librerie e il successivo sequenziamento.

5.1 Estrazione RNA e qRT-PCR

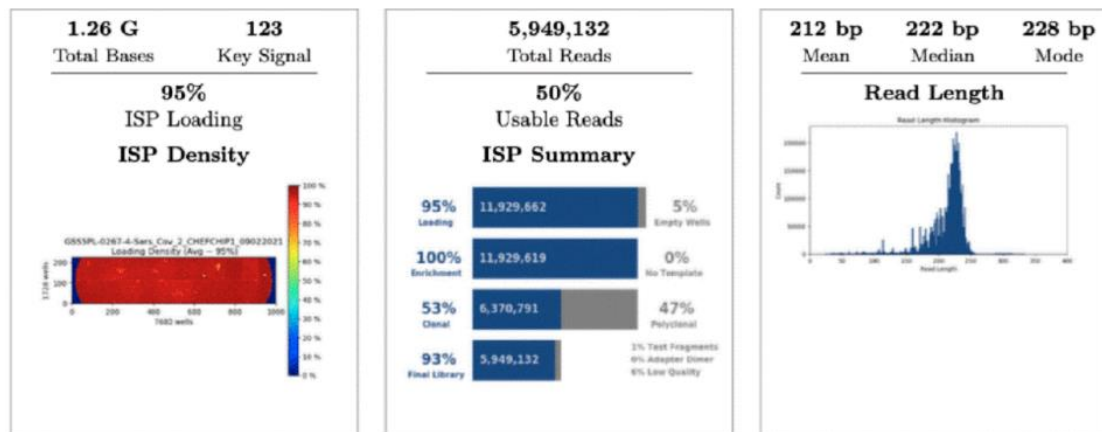
Per effettuare l'analisi, i campioni nasofaringei nell'*Universal transport medium* (UTM) sono stati aperti sotto cappa di biosicurezza e sono stati utilizzati 200 µL per l'estrazione dell'RNA utilizzando il kit di isolamento degli acidi nucleici MagMAX Viral/Pathogen II (MVP II) e il processore automatico di particelle magnetiche KingFisher (Thermo Fisher Scientific, Waltham, MA, USA), come indicato nelle istruzioni del produttore. Quindi, 10 µL dell'RNA estratto sono stati sottoposti a RT-PCR real-time (qRT-PCR) utilizzando il test TaqPath™ COVID-19 CE-IVD RT-PCR Kit (Thermo Fisher Scientific, Waltham, MA, USA) seguendo le istruzioni del kit. Il test QuantStudio 5 Real-Time PCR System (DX) è stato utilizzato per l'analisi qRT-PCR (Thermo Fisher Scientific, Waltham, MA, USA). Il test TaqPath™ prende di mira tre diverse regioni genomiche virali: i geni ORF1ab, N e S. Un risultato negativo valido per SARS-CoV-2 è stato determinato dall'amplificazione dell'unico controllo interno MS2. Un campione è stato considerato positivo in presenza di amplificazione di almeno due dei tre geni bersaglio. In particolare, i campioni positivi sono stati classificati come "S positivi" quando sono stati rilevati tutti e tre i target o quando sono stati rilevati solo due target, compreso S. D'altra parte, i campioni sono stati definiti "S negativi" quando il target S non è stato amplificato [157].

5.2 Retrotrascrizione e Next-Generation Sequencing (NGS)

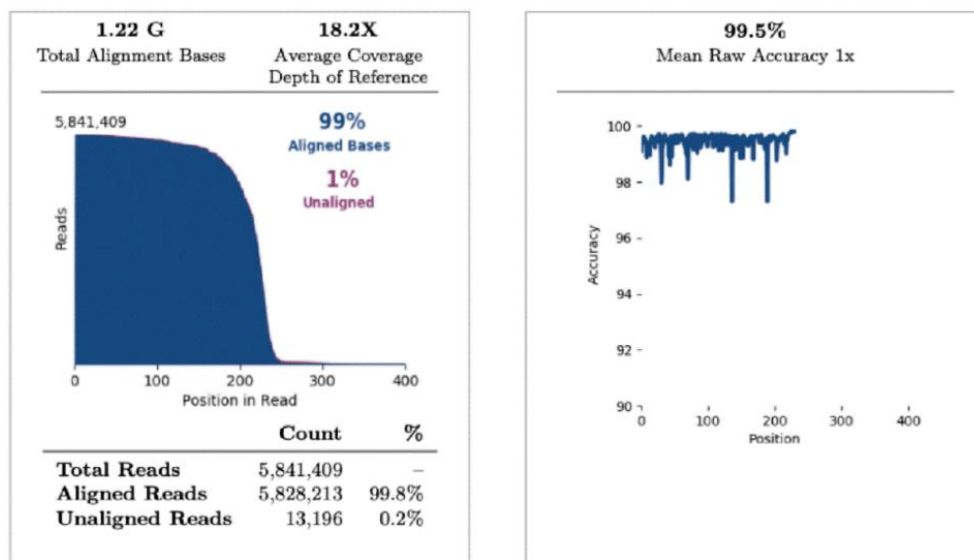
Per il sequenziamento dell'intero genoma virale, l'RNA totale è stato retrotrascritto utilizzando Invitrogen SuperScript VILOTM cDNA Synthesis Kit (Thermo Fisher Scientific, Waltham, MA, USA). I campioni sono stati analizzati utilizzando il sistema Ion Torrent S5 (Thermo Fisher Scientific, Waltham, MA, USA) dopo la preparazione della libreria, consistente nella frammentazione e nel legame dell'adattatore sui prodotti PCR e nell'amplificazione clonale. Le librerie di cDNA sono state quindi preparate utilizzando il pannello di ricerca Ion AmpliSeq SARS-CoV-2 (Thermo Fisher Scientific, Waltham, MA, USA). Dopo la quantificazione delle librerie di cDNA con il sistema Real-Time Step One PCR (Thermo Fisher Scientific, Waltham, MA, USA), i campioni di particelle di sfere ioniche (ISP) sono stati caricati su un chip Ion 520™ con lo Ion Chef (Thermo Fisher Scientific, Waltham, MA, USA). Il

sequenziamento è stato eseguito utilizzando i reagenti di sequenziamento Ion S5™ (Thermo Fisher Scientific, Waltham, MA, USA). Per l'analisi dei dati NGS sono stati utilizzati la piattaforma Torrent Suite 5.14.0 e plug-in specifici. Tutte le sequenze analizzate hanno mostrato una precisione di allineamento superiore al 96% e una copertura di base superiore a 20× (Fig. 29). Il software pangolin è stato utilizzato per l'assegnazione dei lignaggi SARS-CoV-2. Tutte le sequenze sono state quindi inviate come file FASTA su gisaid.org, che fornisce accesso aperto ai dati genomici su SARS-CoV-2 [157].

Run Summary



Alignment Summary (aligned to SarsCov2)

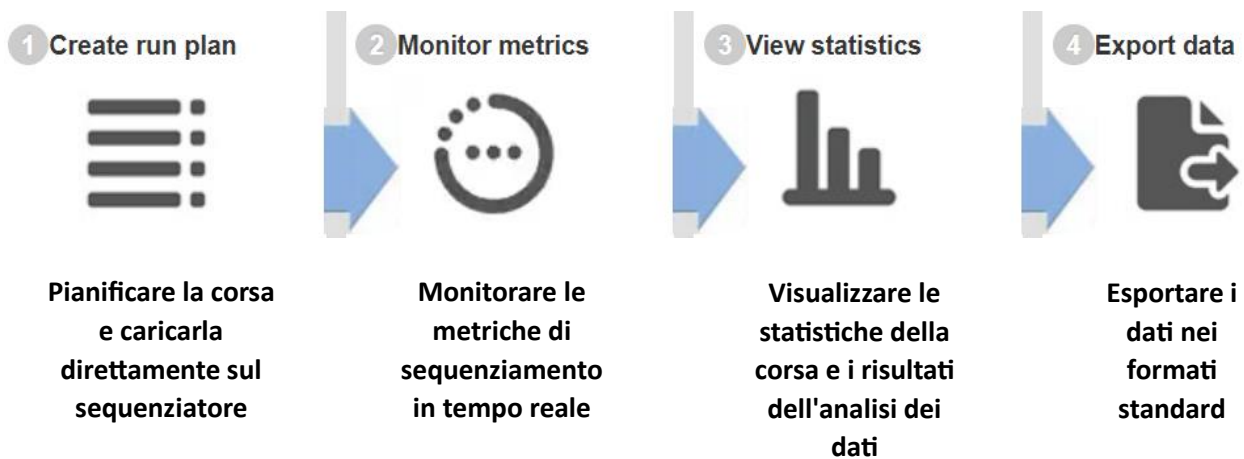


	Alignment Quality		
	AQ17	AQ20	Perfect
Total Number of Bases [Mbp]	1.18 G	1.12 G	898 M
Mean Length [bp]	206	199	165
Longest Alignment [bp]	376	376	348
Mean Coverage Depth	17.6	16.7	13.4

Fig. 29 NGS run summary. Rappresentazione di un rapporto di analisi dei dati NGS sul sequenziamento di SARS-CoV-2 [157].

5.3 Analisi bioinformatiche

Per programmare, monitorare e analizzare i dati relativi al sequenziamento è stato usato il software “S5 Torrent Server”.



Il software è diviso in tre sezioni: “Plan”, “Monitor” e “Data”.

Nella prima sezione (“Plan”), è possibile impostare tutti i parametri relativi alla corsa di sequenziamento:

- caratteristiche del template
- applicazione che si andrà ad utilizzare
- kit usati (ad esempio il modello di chip)
- plugin
- scegliere i nomi dei campioni
- referenze dei *barcode*
- descrizione dei campioni caricati nel chip

Nella seconda sezione (“Monitor”) si può seguire in tempo reale l’andamento del sequenziamento.

Infine, nella sezione “Data” è possibile analizzare i dati e visualizzare il report della corsa NGS (**Fig. 30, 31**). Nel report sono riportati una serie di parametri importanti per valutare la corsa di sequenziamento.

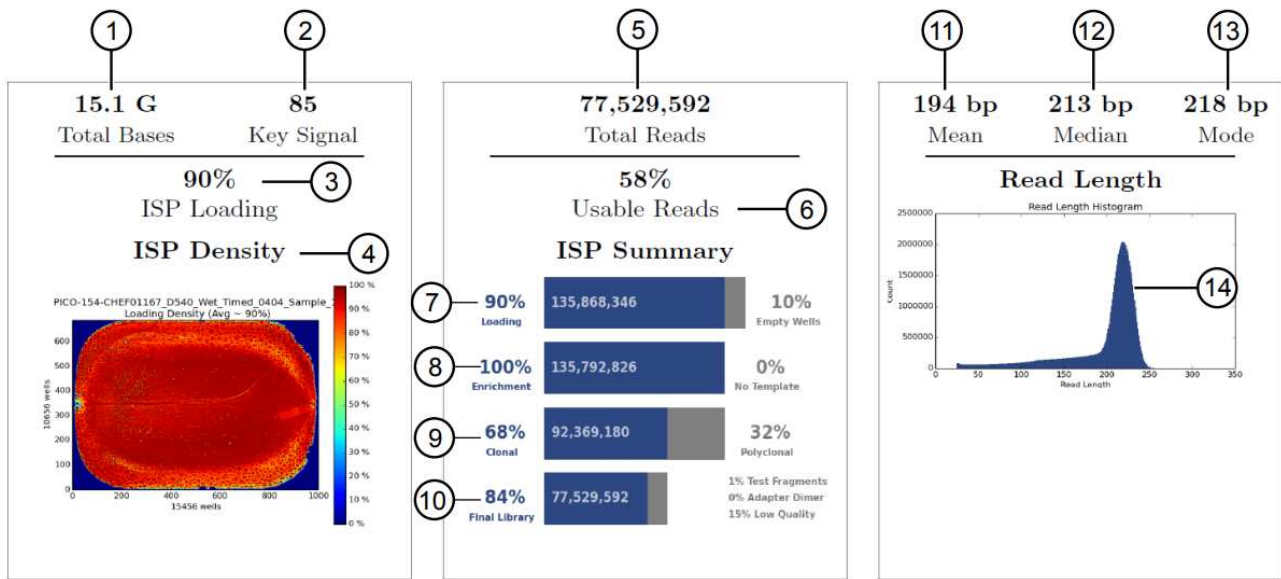


Fig. 30 Run report prima dell'allineamento.

1. **Total Bases.** Il numero di coppie di basi filtrate e tagliate riportate nel file BAM di output.
2. **Key Signal.** Il segnale medio per tutti le ISP della libreria con chiave della libreria (TCAG).
3. **ISP Loading.** La percentuale di pozzetti del chip che contengono una particella sferica ionica (ISP). Il valore percentuale considera solo i pozzetti indirizzabili.
4. **ISP Loading Density.** Una rappresentazione visiva della distribuzione del caricamento dei pozzetti sulla superficie del chip. Il colore rosso indica le aree ad alto caricamento e il blu indica le aree a basso caricamento.
5. **Total Reads.** Il numero totale di read filtrate e tagliate indipendentemente dalla lunghezza riportata nel file BAM di output.
6. **Usable Reads.** La percentuale di ISP della libreria che superano i filtri policlonali, di bassa qualità e primer-dimer. Questa percentuale viene calcolata dividendo le ISP della libreria finali per le ISP della libreria.
7. **ISP Summary-Loading.** La percentuale di pozzetti del chip che contengono una ISP. Il valore percentuale considera i pozzetti indirizzabili.
8. **ISP Summary-Enrichment.** Il numero previsto di ISP che hanno un segnale chiave identico al segnale chiave della libreria o al segnale chiave del frammento di test (TF). Il valore di arricchimento percentuale riportato è il numero di pozzetti caricati con live ISP che sono ISP della libreria o ISP TF. Questo numero è calcolato dividendo i pozzetti con live ISP per il numero di pozzetti caricati con ISP.
9. **ISP Summary-Clonality.** La percentuale di ISP clonali (tutti le library ISP e TF ISP che sono clonali, non policlonali). Una ISP è clonale se tutti i suoi frammenti di DNA sono clonati da un singolo template originale. Tutti i frammenti su tale ISP sono identici e rispondono all'unisono mentre ogni nucleotide scorre a sua volta attraverso il chip. Questa percentuale viene calcolata dividendo il numero di ISP con un singolo template di DNA per il numero di pozzetti totali.
10. **ISP Summary-Final Library.** La percentuale di read che superano tutti i filtri e che vengono registrate nel file BAM di output. Questo valore può essere diverso dalle read totali a causa di tecnicismi associati al taglio delle letture oltre un requisito minimo che fa sì che le letture totali siano leggermente inferiori alla libreria finale.
11. **Mean Read Length.** La lunghezza media delle read che hanno passato il quality check.
12. **Median Read Length.** La lunghezza mediana delle read che hanno passato il quality check.
13. **Mode Read Length.** La moda della lunghezza delle read che hanno passato il quality check.
14. **Read Length Histogram.** L'istogramma della lunghezza delle read è un istogramma delle lunghezze tagliate di tutte le read presenti nei file di output.

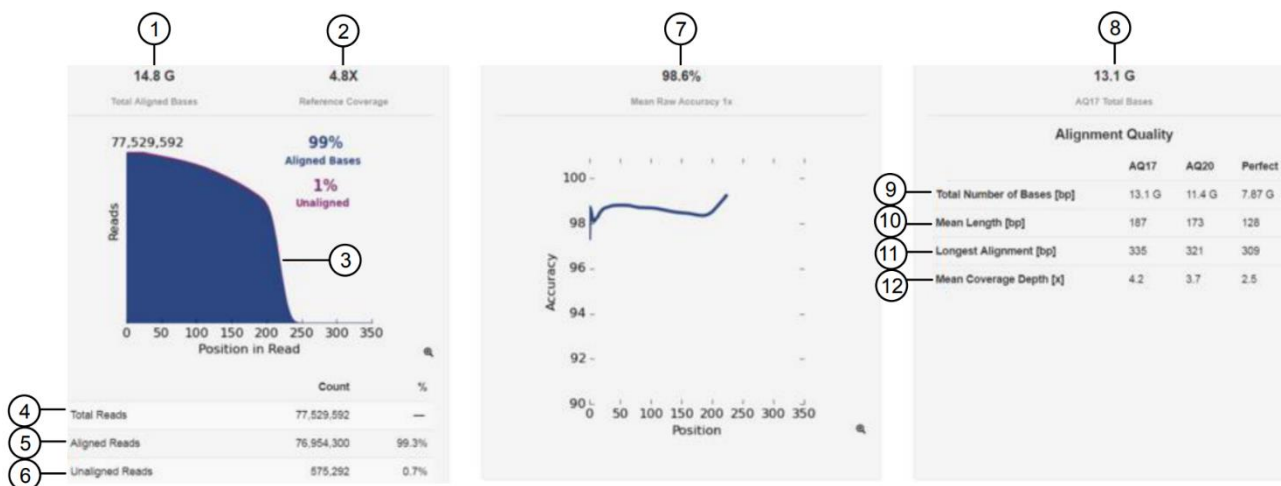


Fig. 31 Run report dopo l'allineamento. In questo caso le read sono allineate ad una sequenza genomica di riferimento.

- Total Aligned Bases.** Il numero di coppie di basi filtrate e tagliate riportate nel file BAM di output che sono allineate alla sequenza di riferimento, escludendo la chiave della libreria, i barcode e le sequenze dell'adattatore al 3'.
- Reference Coverage.** Il rapporto tra il totale delle basi allineate diviso per il numero di basi nella sequenza di riferimento. Il reference coverage non tiene conto dell'arricchimento effettuato per amplificare selettivamente un sottoinsieme della sequenza di riferimento.
- Alignment plot.** Un grafico del numero di letture allineate (blu) e non allineate (viola) per posizione in una sequenza allineata.
- Total Reads.** Il numero totale di read dopo il filtraggio.
- Aligned Reads.** Il numero di read allineate alla sequenza di riferimento espresso come conteggio totale e percentuale delle read allineate totali.
- Unaligned Reads.** Il numero di read che non si allineano alla sequenza di riferimento espresso come conteggio totale e percentuale delle read totali.
- Mean Raw Accuracy 1x.** L'accuratezza grezza media in ogni singola posizione di base in una read calcolata come $1 - (\text{errori totali nella base sequenziata}) / \text{basi totali sequenziate}$.
- AQ17 Total Bases.** Il numero totale di basi su tutte le posizioni allineate con un tasso di errore del 2% o inferiore.
- Total Number of Bases (bp).** Il numero totale di basi su tutte le posizioni che si allineano con un determinato tasso di errore. (AQ17 tasso di errore $\leq 2\%$, AQ20 $\leq 1\%$, Perfetto = nessun errore misurabile).
- Mean Length (bp).** La lunghezza media, in coppie di basi, per le read allineate a un dato tasso di errore.
- Longest Alignment (bp).** La lunghezza massima delle read a un dato tasso di errore.
- Mean Coverage Depth (x).** Il rapporto tra le basi allineate totali a un dato tasso di errore e le dimensioni della regione target.

Per espandere le capacità di analisi del software sono stati usati diversi plug-in bioinformatici. Il plug-in in campo informatico è un programma non autonomo che interagisce con un altro programma per ampliarne o estenderne le funzionalità originarie.

Il plug-in “*coverageAnalysis*” è stato utilizzato per visualizzare statistiche e grafici che descrivono il livello di copertura della sequenza prodotto per le regioni genomiche mirate (**Fig. 32**). I risultati del riepilogo della corsa si basano sul tipo di libreria selezionato durante la configurazione del plug-in (nel nostro caso la libreria è “AmpliSeq DNA”).

SARS_{CoV2c}*coverageAnalysis*

Library type: AmpliSeq DNA

Target regions: Coronavirus_v2.2.20200206.designed

Barcode Name	Sample	Mapped Reads	Filtered Reads	Target Reads	Mean Depth	Uniformity
IonCode_0125	I27-39	565,795	0.00%	99.81%	3,939	96.70%
IonCode_0126	I27-51	637,623	0.00%	99.99%	4,375	95.92%
IonCode_0127	I27-88	664,862	0.00%	99.64%	4,623	97.01%
IonCode_0128	I28-02	784,165	0.00%	99.95%	5,430	96.92%
IonCode_0129	I28-12	590,510	0.00%	99.97%	4,033	96.20%
IonCode_0130	I28-25	606,991	0.00%	99.95%	4,229	93.48%
IonCode_0131	I28-63	529,690	0.01%	98.93%	3,628	96.57%
IonCode_0132	I29-31	383,346	0.00%	99.94%	2,638	97.34%

Fig. 32 Plug-in coverageAnalysis.

Il plug-in “*pangolin*” è stato usato per risalire alle varianti sequenziate (**Fig. 33**). Questo plug-in si collega direttamente al sito cov-lineages.org e sfrutta la nomenclatura “PANGO”, utilizzata da ricercatori e agenzie di sanità pubblica in tutto il mondo per monitorare la trasmissione e la diffusione del SARS-CoV-2, comprese le varianti preoccupanti (VOC).

Barcode Name	Sample	Reads	Lineage	Prob	Scorpio	Class	Pango QC	Note	Cons. Length	N Pct	Avg Depth	<100x	QC Status	Spike: N Pct	Spike: Avg Depth	Spik <100
IonCode_0125	I27-39	567,477	XBB.1.5	0.3333333333333333	Omicron (BA.2-like)	VOC-22JAN-01	v0.1.10	False	29,744	0.2% (56Ns)	3742	1%	Passed	0.0% (0Ns)	3101	1.4%
IonCode_0126	I27-51	638,220	EG.1	0.0	Omicron (BA.2-like)	VOC-22JAN-01	v0.1.10	False	29,743	0.1% (40Ns)	4117	1.2%	Passed	1.0% (40Ns)	3066	5%
IonCode_0127	I27-88	667,393	XBB.1.9.1	0.0	Omicron (BA.2-like)	VOC-22JAN-01	v0.1.10	False	29,744	0.0% (0Ns)	4393	1.4%	Passed	0.0% (0Ns)	3570	5%
IonCode_0128	I28-02	785,494	EG.1	0.0	Omicron (BA.2-like)	VOC-22JAN-01	v0.1.10	False	29,743	0.0% (0Ns)	5139	0.5%	Passed	0.0% (0Ns)	3906	1.4%
IonCode_0129	I28-12	591,403	EG.1	0.0	Omicron (BA.2-like)	VOC-22JAN-01	v0.1.10	False	29,744	0.2% (57Ns)	3820	0.7%	Passed	0.0% (1Ns)	2806	1.4%

Fig. 33 Plug-in pangolin.

5.4 Protocolli

5.4.1 Estrazione RNA

1. Operazioni preliminari

- Determinare il numero di reazioni richieste sulla base dei campioni da processare, più un controllo negativo per piastra.
- Preparare etanolo all'80% usando etanolo assoluto e acqua nucleasi-free (non trattata con DEPC) per il numero di reazioni richieste, sufficiente per 1 mL a reazione, più un 10%.
- Etichettare il lato di ogni piastra "KingFisher 96 plate" (4):

Label	Number of plates
Sample plate	1
Wash 1	1
Wash 2	1
Elution plate	1

- Etichettare anche la piastra con i puntali (KingFisher 96 KF microplate).
- Contrassegnare il controllo negativo nella piastra o inserirlo nell'ultima posizione.

2. Preparazione piastre di lavorazione

- Preparare le piastre secondo la tabella seguente. Coprire le piastre con parafilm e conservarle a temperatura ambiente per un massimo di un'ora.

Plate ID	Plate position	Plate type	Reagent	Volume per well
Wash 1 Plate	2	KingFisher™ Deepwell 96 Plate	Wash Buffer	500 µL
Wash 2 Plate	3		80% Ethanol	1,000 µL
Elution Plate	4		Elution Solution	50 µL
Tip Comb Plate	5	Place a KingFisher™ 96 tip comb for DW magnets in a KingFisher™ 96 KF microplate		

3. Preparazione Binding Bead Mix

- Vortexare le biglie magnetiche per assicurarsi che la soluzione sia omogenea.
- Per il numero di reazioni richieste, preparare la "Binding Bead Mix" secondo la tabella seguente.

Component	Volume per well ^[1]
Binding Solution	265 µL
Total Nucleic Acid Magnetic Beads	10 µL
Total volume per well	275 µL

- Mescolare bene per inversione, conservare a temperatura ambiente.

4. Preparazione piastra campioni (Sample plate)

- Aggiungere 5 µl di Proteinasi K ad ogni pozzetto della piastra etichettata come “Sample plate”.
- Aggiungere 200 µl di campione ad ogni pozzetto.
- Aggiungere 200 µl di acqua nucleasi-free (non trattata con DEPC) al pozzetto del controllo negativo.
- Invertire la Binding Bead Mix 5 volte per mescolare, poi aggiungere 275 µl ad ogni pozzetto (anche il pozzetto con il controllo negativo).
- Aggiungere 5 µl di controllo del fago MS2 ad ogni pozzetto.

5. Elaborazione campioni

- Iniziare la corsa, poi caricare le piastre preparate in precedenza in posizione quando richiesto dallo strumento.
- Dopo che la corsa è completata (circa 22 minuti), rimuovere subito la “Elution Plate” dallo strumento e coprirla immediatamente con parafilm per prevenire l’evaporazione. I campioni sono eluiti in 50 µl di “Elution Solution”.

5.4.2 qRT-PCR

1. Preparazione reazioni RT-PCR

- Scongellare i reagenti.
- Spinnare delicatamente i reagenti, poi centrifugare brevemente per raccogliere il liquido sul fondo della provetta.
- Diluire il “TaqPath COVID-19 Control” (1×10^4 copie/ µl) ad uno stock di lavoro di 25 copie/ µl:
 - Pipettare 98 µl di “Dilution Buffer” in una provetta da microcentrifuga, poi aggiungere 2 µl di TaqPath COVID-19 Control. Mescolare bene e poi centrifugare brevemente.
 - Pipettare 87.5 µl di Dilution Buffer in una seconda provetta da microcentrifuga, poi aggiungere 12.5 µl della diluizione fatta in precedenza. Mescolare bene e poi centrifugare brevemente.
- Preparare le Mix di reazione:

- Per ogni corsa, unire i componenti seguenti sufficienti per il numero di campioni di RNA da testare più un controllo positivo e uno negativo.
Tutti i volumi includono un eccesso del 10% per tener conto degli errori delle pipette.

Component	Volume per RNA sample or control	Volume for n RNA samples plus 2 controls	Volume for 94 RNA Samples plus 2 controls
TaqPath™ 1-Step Multiplex Master Mix (No ROX™) (4X)	6.25 µL	$6.875 \times (n + 2)$ µL	660 µL
COVID-19 Real Time PCR Assay Multiplex	1.25 µL	$1.375 \times (n + 2)$ µL	132 µL
Nuclease-free Water	7.50 µL	$8.25 \times (n + 2)$ µL	792 µL
Total Reaction Mix volume	15.0 µL	—	1584 µL

- Preparare la piastra di reazione:
 - Pipettare 15.0 µL del Mix di reazione fatto in precedenza in ogni pozzetto di una piastra “MicroAmp Fast Optical 96-Well Reaction Plate” con barcode. Possono essere usate anche piastre senza barcode.
 - Spinnare delicatamente la piastra contenente i campioni di RNA purificato e il controllo negativo, poi centrifugare brevemente per raccogliere il liquido sul fondo della piastra.
 - Aprire la piastra con l’RNA e il controllo negativo. Aggiungere i campioni di RNA, il controllo negativo o il controllo positivo ad ogni pozzetto della piastra di reazione secondo la tabella seguente.
 - Chiudere la piastra completamente con il “MicroAmp Optical Adhesive Film”.
 - Spinnare la piastra alla massima velocità per 10-30 secondi con una pressione media.
 - Centrifugare la piastra di reazione per 1-2 minuti a $\geq 650 \times g$ (≥ 650 RCF) per rimuovere le bolle e raccogliere il liquido sul fondo della piastra.

Component	Volume per reaction		
	RNA Sample reaction	Positive Control reaction	Negative Control reaction
Reaction Mix	15.0 µL	15.0 µL	15.0 µL
Purified sample RNA (from RNA extraction)	10.0 µL	—	—
Positive Control (diluted TaqPath™ COVID-19 Control, from step 3)	—	2.0 µL	—
Nuclease-free Water	—	8.0 µL	—
Purified Negative Control (from RNA extraction)	—	—	10.0 µL
Total volume	25.0 µL	25.0 µL	25.0 µL

2. Analisi campioni

- La piastra di reazione è analizzata con lo strumento “QuantStudio 5 Dx Real-Time PCR System”.
- Impostare il volume a 25 µL e il seguente protocollo termico.

Step	Temperature	Time	Number of cycles
UNG incubation	25°C	2 minutes	1
Reverse transcription	53°C	10 minutes	1
Activation	95°C	2 minutes	1
Denaturation	95°C	3 seconds	40
Anneal / extension	60°C	30 seconds	

- Verificare che le sonde fluorescenti e i rispettivi target siano corretti.

Reporter dye	Detector	Quencher
FAM	ORF1ab	None
VIC	N gene	None
ABY	S gene	None
JUN	MS2	None

- Infine, iniziare l'analisi.

5.4.3 NGS Covid – ION CHEF & ION S5

1. Retrotrascrizione

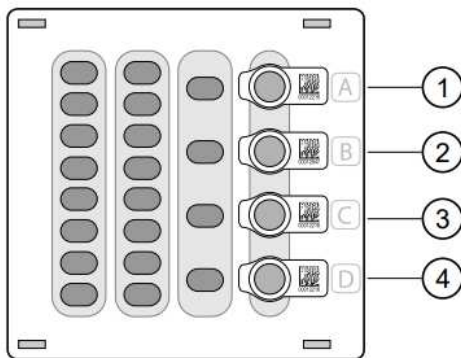
- Prelevare piastre "elution" dal frigo con l'RNA estratto.
- Rinominare eppendorf da 0.2 mL con i codici dei pazienti.
- Spinnare le due provette dell'NGS reverse transcription kit contenenti la retrotrascrittasi e il buffer.
- Inserire 7 µl di RNA + 3 µl dell'NGS reverse transcription kit nell'eppendorf da 0.2 mL.
Buffer è al 5x → 2 µl
Enzima è al 10x → 1 µl
- Spinnare eppendorf da 0.2 mL per eliminare tutte le bolle d'aria.
- Posizionare le eppendorf nel termociclatore per effettuare la retrotrascrizione.
10 minuti → 25 °C
10 minuti → 50 °C
5 minuti → 85 °C
- Dopo aver finito il ciclo il termociclatore rimane a 10 °C ed i campioni possono rimanere al suo interno per un massimo di 15 ore.
- Mettere i 10 µl in piastra (IonCode 96-well PCR plate) con 5 µl di acqua nucleasi-free per un totale di 15 µl in ogni pozzetto. I campioni vanno inseriti nella prima colonna della piastra (A1-H1). La piastra ha 8 barcode, quindi è importante mantenere l'ordine corretto dei campioni.
- Centrifugare brevemente la piastra per eliminare tutte le bolle d'aria.

2. Preparazione libreria sistema Ion Chef

- Almeno un'ora prima prendere le 2 cartucce per l'NGS. Una è la "Reagents cartridge" ed è conservata a -20 °C, mentre l'altra è la "Solutions cartridge" (contiene le biglie per l'amplificazione) e si trova a +4 °C.
- Scongelare i pool di primer 1 e 2.
- Unire 120 µl di pool e 180 µl di acqua e poi trasferire 150 µl nelle eppendorf della cartuccia "Reagents cartridge".

Posizione A → Pool 1

Posizione B → Pool 2



- ① Position A (150 µL Primer Pool 1 at 2X concentration)
- ② Position B (150 µL Primer Pool 2 at 2X concentration)
- ③ Position C (Empty tube)
- ④ Position D (Output tube)

- Caricare lo strumento con la Reagents cartridge, la Solutions cartridge, la Tip cartridge, l'Enrichment cartridge, la 96-well PCR plate (che contiene il cDNA) e il frame per la PCR.
- Settare lo strumento:
 - Primer pool = 2
 - I cicli di amplificazione target dipendono dalla quantificazione del numero delle copie virali fatto dalla qPCR

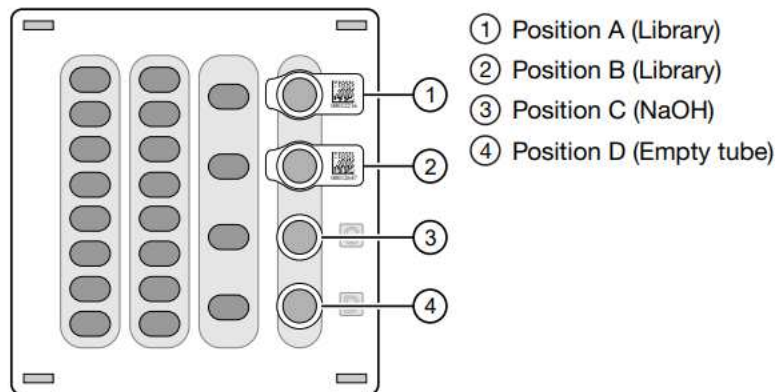
Tier	Viral copy number	Number of amplification cycles
Low	50-1500	27
Medium	1500-50000	21
High	50000-1500000	16

- Tempo di annealing ed estensione = 4 minuti

- Dopo circa 6 ore e 46 minuti avremo la libreria (il tempo varia a seconda dei cicli di amplificazione).
- Alla fine, avremo 700 µl di libreria con barcode nella eppendorf in posizione D della Reagents cartridge. Le librerie possono essere conservate a 4-8 °C per un mese, mentre per periodi più lunghi da -30 a -10 °C.

3. Templato

- Scongelare un'ora prima le due cartucce a -20 °C: "Ion S5 sequencing reagents" e "Ion 510 520 530 Chef reagents".
- Prendere cartuccia "Ion S5 chef solutions" a temperatura ambiente.
- Oltre alle cartucce inserire plasticheria nello strumento.
- Scongelare le 2 eppendorf a -20 °C con le librerie.
- Inserire 25 µl per ogni libreria nelle due eppendorf della cartuccia Ion 510 520 530 Chef reagents. La cartuccia ha altre due eppendorf, una contiene NaOH e l'ultima è vuota.



- Inserire il chip (senza guanti) a incastro nello strumento.
- Far partire il templato sullo strumento Chef

4. Sequenziamento

- Inizializzare il sequenziatore e cambiare i consumabili, rovesciare i flaconi 3 volte.
 - 1 flacone cleaning
 - 1 flacone wash
 - 1 flacone waste
- Lasciare chip della corsa precedente nel sequenziatore (serve allo strumento per tararsi).
- Alla fine del templato rimuovere i chip dallo Chef. I chip possono essere subito inseriti nel sequenziatore o essere conservati a +4 °C.
- Dopo aver selezionato "RUN" sul sequenziatore, togliere il vecchio chip e inserire il nuovo.
- Il tempo di sequenziamento è di circa 2 ore e 45 per ogni chip (ci sono due chip, uno per ogni libreria).

6. Risultati

6.1 Varianti rilevate

L'analisi delle sequenze dei 232 campioni sequenziati tra aprile 2022 e aprile 2023 ha rivelato che la variante Omicron è stata l'unica ad essere riscontrata.

La variante Omicron del SARS-CoV-2, chiamata lignaggio B.1.1.529, è la variante del coronavirus SARS-CoV-2 con la più alta prevalenza cumulativa globale nell'anno 2022 [158]. È stata rilevata per la prima volta in Botswana in un campione raccolto il 9 novembre 2021 e successivamente l'Organizzazione Mondiale della Sanità l'ha designata come VOC.

Del lignaggio B.1.1.529 sono state identificate 5 varianti (BA.1, BA.2, BA.3, BA.4, BA.5) con oltre 75 sotto-lignaggi dovuti a mutazioni puntuali. La variante BA.1 (B.1.1.529.1 e B.1.1.529.1.1) ha prevalso ampiamente (>95%) dagli ultimi mesi del 2021 fino a febbraio 2022, essendosi diffusa rapidamente anche fuori dal Sud Africa. La variante BA.2 ha cominciato a prevalere sulla BA.1 da marzo 2022, mentre la BA.3 è stata rilevata raramente. Le varianti BA.4 e BA.5, con molte mutazioni in comune con la BA.2, sono emerse da aprile 2022.

La variante Omicron ha un numero insolitamente alto di mutazioni, molte delle quali sono nuove o colpiscono il peplomero, conosciuto come proteina spike. Queste caratteristiche la rendono preoccupante per quanto riguarda la sua trasmissibilità, l'efficacia del sistema immunitario e dei vaccini.

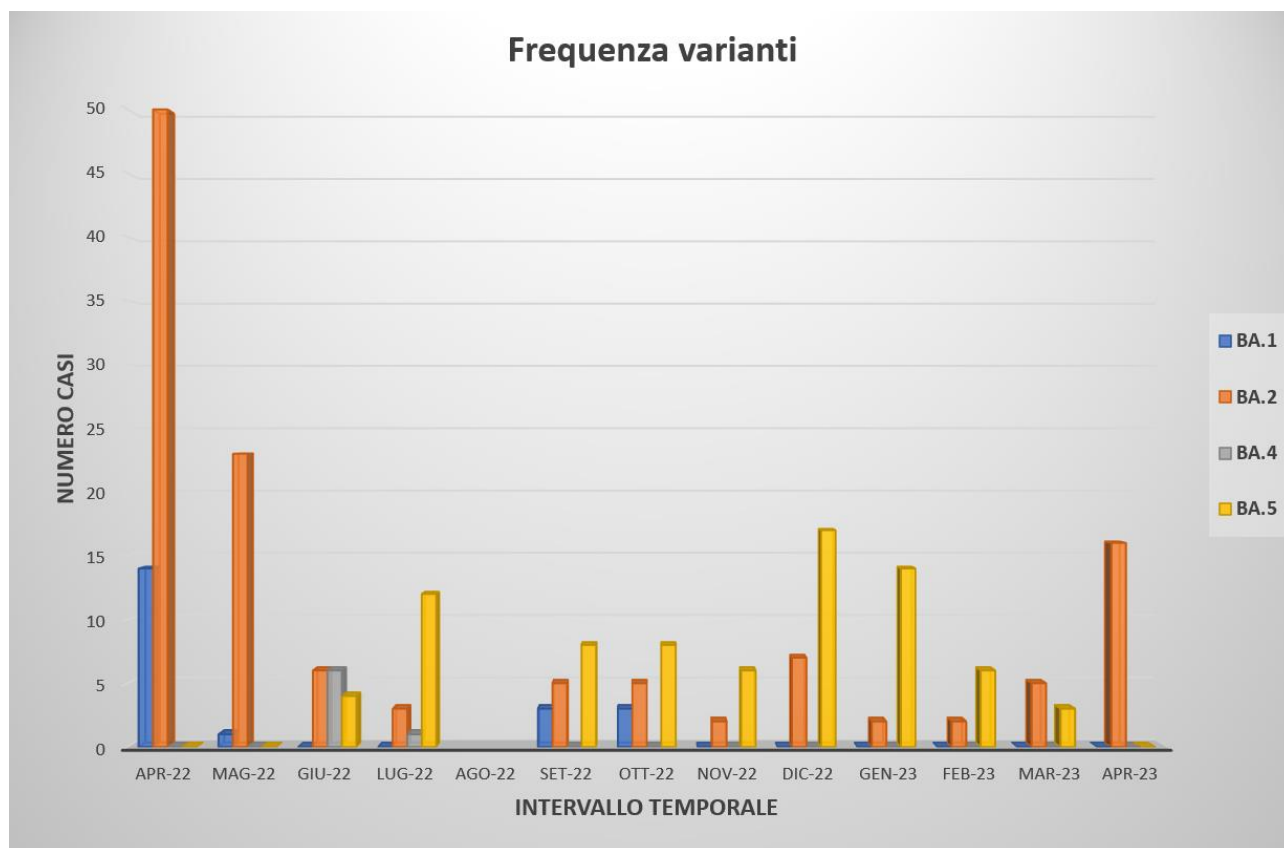


Grafico 1 Frequenza varianti nell'intervallo temporale aprile 2022-aprile 2023.

Nel grafico 1 è possibile osservare la frequenza di ogni variante nell'intervallo di tempo di riferimento. Nel grafico sono rappresentate le varianti BA.1, BA.2, BA.4, BA.5 che corrispondono a Omicron 1, 2, 4, 5. La variante Omicron 3 non è presente nel grafico in quanto non è mai stata rilevata.

Per rendere più leggibile e chiaro il grafico tutti i sotto-lignaggi rilevati sono stati raggruppati all'interno delle quattro classi rappresentate.

Leggendo l'istogramma si può notare che la variante Omicron 1 viene quasi immediatamente sostituita dalle varianti BA.2 e BA.5, le quali persistono fino a marzo e aprile 2023.

La variante Omicron 4 è poco rappresentata in quanto è stata rilevata solo in sette casi (6 a giugno e 1 a luglio 2022).

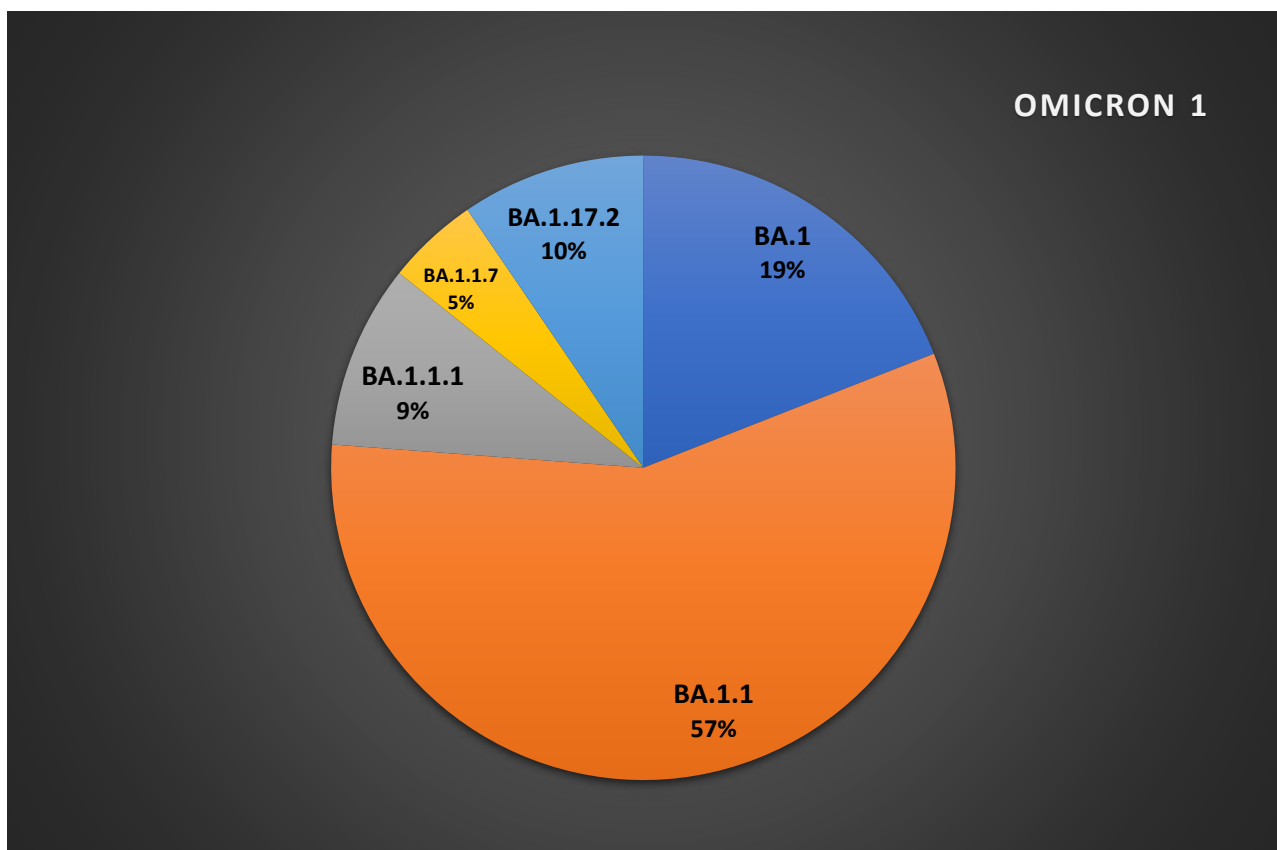


Grafico 2 Distribuzione sotto-lignaggi Omicron 1

Il virus Sars-CoV-2, come tutti i virus a RNA, è caratterizzato da un elevato tasso di mutazione, fino a un milione di volte superiore a quello dell'uomo. Questo alto tasso è correlato sia all'evoluzione virale che alla modulazione della virulenza, tratti che possono consentire un più vantaggioso adattamento del virus al suo ospite.

Per rappresentare meglio la complessità dei sotto-lignaggi rilevati in un anno di sequenziamento, il grafico 2 riporta il totale dei casi di Omicron 1 con le percentuali delle varianti sequenziate. Dato che Omicron 1 ad aprile 2022 aveva già fatto il suo corso, la situazione risulta abbastanza semplice, con solo 4 varianti rilevate oltre alla BA.1.

Le varianti seguono la nomenclatura “PANGO” presente sul sito cov-lineages.org.

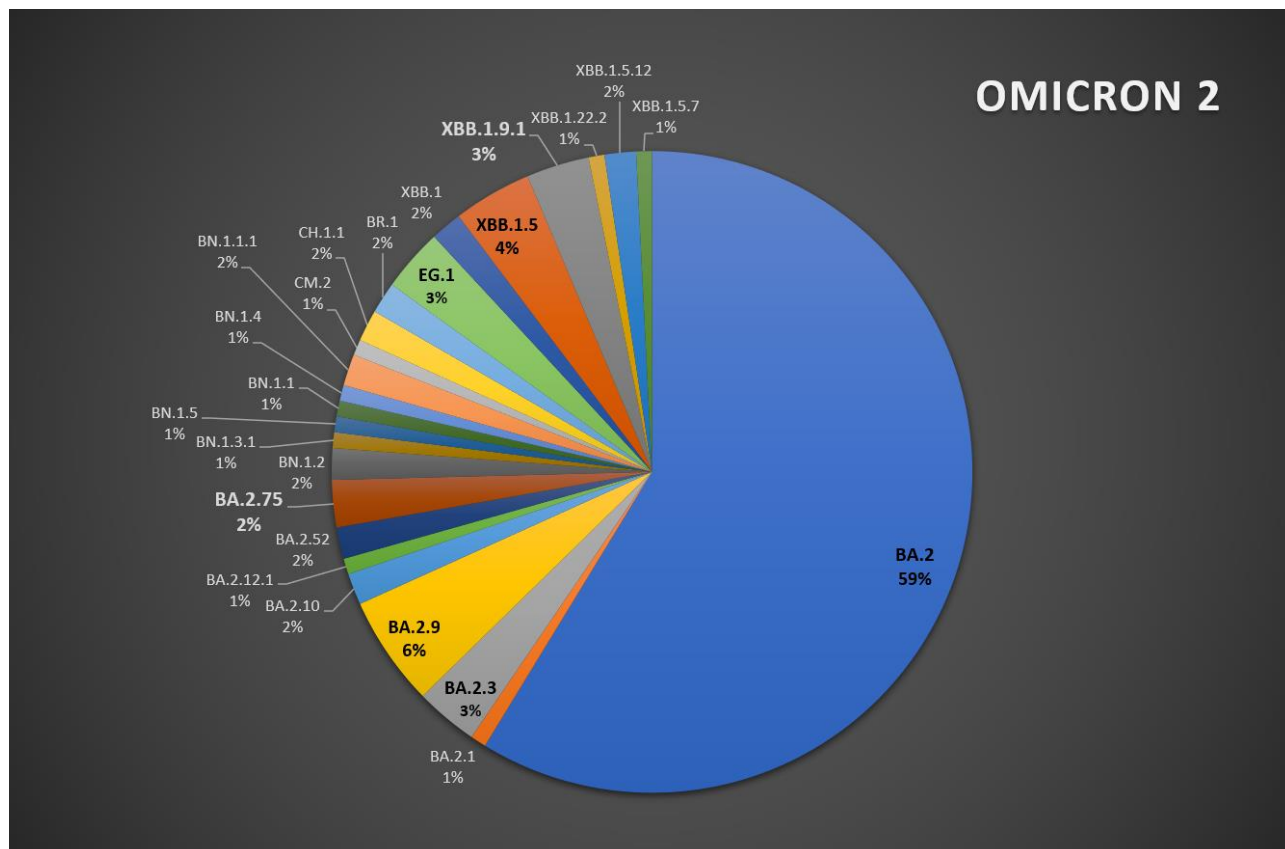


Grafico 3 Distribuzione sotto-lignaggi Omicron 2

Per Omicron 2 su un totale di 111 casi sono state rilevate 23 sotto-varianti oltre alla BA.2 (grafico 3). I lignaggi BN, CH, BR, EG e CM sono sempre riconducibili a Omicron 2, ma è stata usata una nomenclatura differente per ridurne la lunghezza (ad esempio BN.1.2 sta per B.1.1.529.2.75.5.1.2). La variante BA.2.75, rinominata “Centaurus”, è stata segnalata a maggio 2022 ed è stata associata ad una maggiore fusogenicità e ad una replicazione più efficiente nelle cellule polmonari.

Nel grafico sono presenti anche vari lignaggi della variante XBB. Un solo caso con variante XBB è stato rilevato a ottobre 2022, molto probabilmente è stato uno dei primi casi a livello italiano ed europeo dato che la sequenza non era ancora depositata sul sito GISAID.

XBB è una variante ricombinante di BA.2.10.1 e BA.2.75.2 (due sotto-lignaggi di Omicron 2) con mutazioni aggiuntive sulla proteina spike. È nota anche con il nome di “variante Gryphon”.

Particolarmente interessante è il lignaggio XBB.1.5, rinominato “variante Kraken”. L'OMS ha dichiarato che fra ottobre 2022 e gennaio 2023, questa variante è stata segnalata in 38 Paesi e che la maggior parte delle segnalazioni arrivava dagli Stati Uniti. La variante è presente anche in Europa; difatti, a partire dal 12 gennaio 2023, l'ECDC (il Centro europeo per il controllo e la prevenzione delle malattie) l'ha inserita nella lista delle varianti di interesse (VOI) secondo la sua classificazione (che si differenzia da quella dell'Organizzazione Mondiale della Sanità). A marzo 2023 la variante Kraken sembrava essere quella prevalente a livello della maggior parte delle Regioni dell'OMS (stima effettuata sulla base delle segnalazioni ricevute dai diversi Stati). Da studi preliminari sembra emergere una maggiore capacità di fuga immunitaria e una maggiore trasmissibilità di XBB.1.5 rispetto a XBB.

Anche il lignaggio XBB.1.9.1, chiamato “Hyperion”, ha destato interesse e ad aprile 2023 è stato inserito dell'OMS nella lista delle varianti sotto monitoraggio (VUM).

Per Omicron 4 non è presente un grafico in quanto sono stati rilevati solo 7 casi e tutti di BA.4.

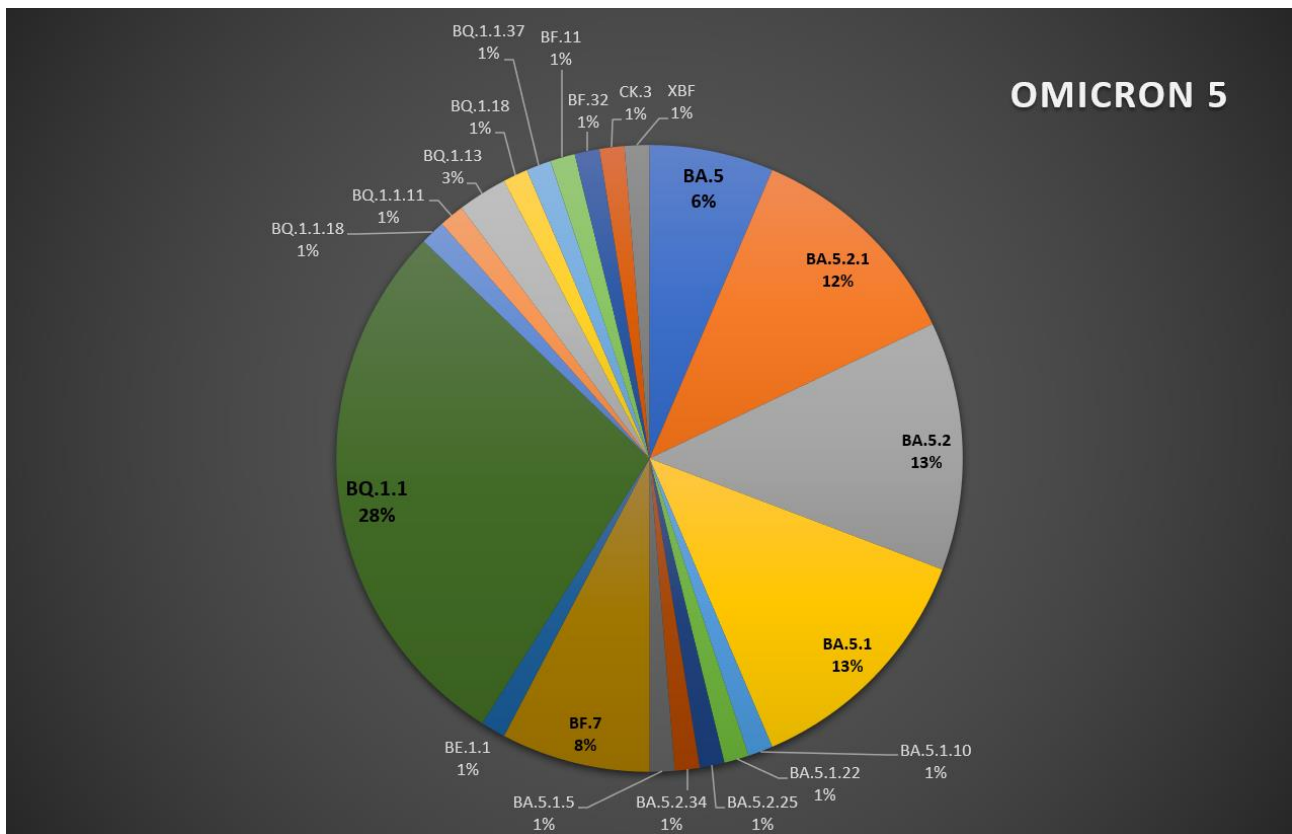


Grafico 4 Distribuzione sotto-lignaggi Omicron 5

Sono stati sequenziati un totale di 77 casi di Omicron 5 con 20 sotto-varianti oltre alla BA.5 (grafico 4). La variante più rappresentata è la BQ.1.1 (alias di B.1.1.529.5.3.1.1.1.1.1.1) che è stata chiamata "variante Cerberus". Secondo il report dell'ECDC, la variante Cerberus sembra essere emersa in Africa centrale o occidentale prima di diffondersi nel resto del mondo. È stato visto che questa variante possiede una maggior capacità di resistere al trattamento con alcuni anticorpi monoclonali impiegati in terapia rispetto ad Omicron 5.

Un solo caso con variante XBF è stato rilevato a marzo 2023. La variante XBF, chiamata "variante Bythos", è una variante ricombinante di BA.5.2.3 e CJ.1 (quest'ultimo è un sottolignaggio di BA.2.75.3). Nel rapporto del 1° marzo 2023, l'Organizzazione Mondiale della Sanità fa sapere che la nuova variante rappresenta l'1,2% delle varianti sequenziate a livello globale. Si ritiene che XBF possieda un maggior vantaggio di trasmissione rispetto ad altre sottovarianti attualmente circolanti. A inizio marzo 2023 è stata segnalata la sua presenza anche in Italia con una trentina di casi accertati.

6.2 Deposito dati

I dati di sequenziamento dovevano essere depositati periodicamente su tre piattaforme, una regionale, una nazionale e una internazionale.

Il deposito regionale avveniva sul sito dell'IZS (Istituto Zooprofilattico Sperimentale) di Teramo che aveva implementato una piattaforma per la visualizzazione delle sequenze genomiche di SARS-CoV-2 in Abruzzo (https://sarscov2.izs.it/j_test_sarscov2/#/).

Il deposito nazionale avveniva nel database dell'Istituto Superiore di Sanità sulla piattaforma per il progetto I-Co-Gen di IRIDA ARIES (<https://irida.iss.it>) (Fig. 34). IRIDA (Integrated Rapid Infectious Disease Analysis) ARIES (Advanced Research Infrastructure for Experimentation in GenomicS) è un'infrastruttura progettata per la raccolta, analisi automatica dei dati e scambio di informazioni derivanti dalla caratterizzazione genomica degli agenti infettivi. È stata pubblicata per fornire agli operatori di sanità gli strumenti necessari per utilizzare i dati di caratterizzazione genomica dei microrganismi a supporto della sorveglianza delle malattie infettive.

🏠 / Progetti

Abruzzo - SARS-CoV-2

Campioni Metadati Analitici Analisi Esportazioni NCBI Attività Recenti Impostazioni

Azioni sui Campioni Esporta **Invia al Centro Analisi**

Seleziona Tutti Deseleziona Search: cast **T** **Svuota**

Nessun campione selezionato

Nome	Descrizione	Ospedale	Codice Interno	Progetto	Data Campione	Caricato Il	Modificato Il
<input type="checkbox"/> hCoV-19-Italy-Abruzzo-CAST-I2931		CAST - Università D'Annunzio Chieti	hCoV-19/ITA/ABR/I2931/2023	Abruzzo - SARS-CoV-2	17 apr 2023, 02:00	2 mag 2023, 16:38	2 mag 2023, 16:42
<input type="checkbox"/> hCoV-19-Italy-Abruzzo-CAST-I2863		CAST - Università di Chieti	hCoV-19/ITA/ABR/I2863/2023	Abruzzo - SARS-CoV-2	17 apr 2023, 02:00	2 mag 2023, 16:37	2 mag 2023, 16:41
<input type="checkbox"/> hCoV-19-Italy-Abruzzo-CAST-I2825		CAST - Università D'Annunzio Chieti	hCoV-19/ITA/ABR/I2825/2023	Abruzzo - SARS-CoV-2	17 apr 2023, 02:00	2 mag 2023, 16:36	2 mag 2023, 16:40

Fig. 34 Schermata della piattaforma I-Co-Gen. Nell'elenco ci sono tre sequenze di hCoV-19 (human coronavirus 19) depositate dal CAST di Chieti in data 2 maggio 2023.

Il deposito internazionale avveniva sul sito GISAID (<https://gisaid.org>) (Fig. 35). L'iniziativa GISAID promuove la rapida condivisione dei dati di tutti i virus influenzali e del coronavirus che causa il COVID-19. Ciò include la sequenza genetica e i relativi dati clinici ed epidemiologici associati a virus umani e dati geografici e specie-specifici associati a virus aviari e altri virus animali, per aiutare i ricercatori a capire come i virus si evolvono e si diffondono durante epidemie e pandemie.

© 2008 - 2023 | Terms of Use | Privacy Notice | Contact

GISAID

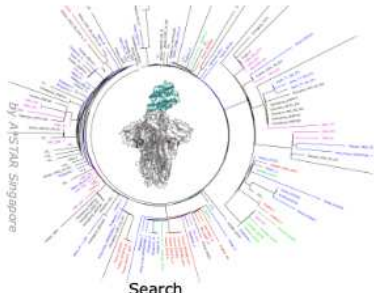
Registered Users | EpiFlu™ | **EpiCoV™** | EpiPox™ | My Profile

EpiCoV™ | Search | Downloads | Upload


Pandemic coronavirus causing COVID-19

A previously unknown human coronavirus (hCoV-19) was first detected in late 2019 in patients in the City of Wuhan, who suffered from respiratory illnesses including atypical pneumonia, an illness that has become known as coronavirus disease (COVID-19). The coronavirus originated from an animal host and is closely related to the virus responsible for the Severe Acute Respiratory Syndrome coronavirus (SARS).


On 10. January 2020, the first virus genomes and associated data were publicly shared via GISAID. The World Health Organization announced on 11. March 2020 the first coronavirus pandemic. As the pandemic progresses, scientists from around the globe are tracking the virus and its genome sequences to ensure optimal virus diagnostic tests, to track and trace the ongoing outbreak and to identify potential intervention options. Several analyses to assist with these efforts are offered here, including sequence alignments, diagnostic primer and probe coordinates, 3D protein models, drug targets, phylogenetic trees and many more.




[Search](#)



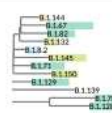
Audacity




AudacityInstant




BLAST




CoVizu^e




Emerging Variants




Lineage Frequency




Official GISAID




PrimerChecker



Submission tracker



Spike glycoprotein



Important note: In the [GISAID EpiFlu™ Database Access Agreement](#), you have accepted certain terms and conditions for viewing and using data regarding influenza viruses. To the extent the Database contains data relating to non-influenza viruses, the viewing and use of these data is subject to the same terms and conditions, and by viewing or using such data you agree to be bound by the terms of the [GISAID EpiFlu™ Database Access Agreement](#) in respect of such data in the same manner as if they were data relating to influenza viruses.

v2.5.1

Fig. 35 Schermata del sito GISAID. Oltre al deposito delle sequenze il sito offre anche altri servizi come: allineamenti di sequenze, coordinate di primer diagnostici e di sonde, modelli proteici 3D, target farmacologici, alberi filogenetici e altri.

6.3 Mutazioni rilevate

La variante Omicron ha un numero di mutazioni molto elevato rispetto agli altri lignaggi precedenti. Le mutazioni di maggior interesse sono quelle che colpiscono la proteina Spike e in particolar modo l'RBD (Receptor Binding Domain).

Tra le mutazioni principali riscontrate nella proteina Spike di Omicron ci sono: D614G, E484A, N501Y, Q493K, K417N, S477N, Y505H e G496S. Sette di queste mutazioni si trovano nella regione RBD, mentre la D614G è situata fuori da tale regione. Mutazioni significative sono state riscontrate anche nell'RBM (E484A, N501Y, Q493K, S477N, Y505H, T478K, N440K, G446S, Q496S e Q498R) [159] (**Fig. 36**).

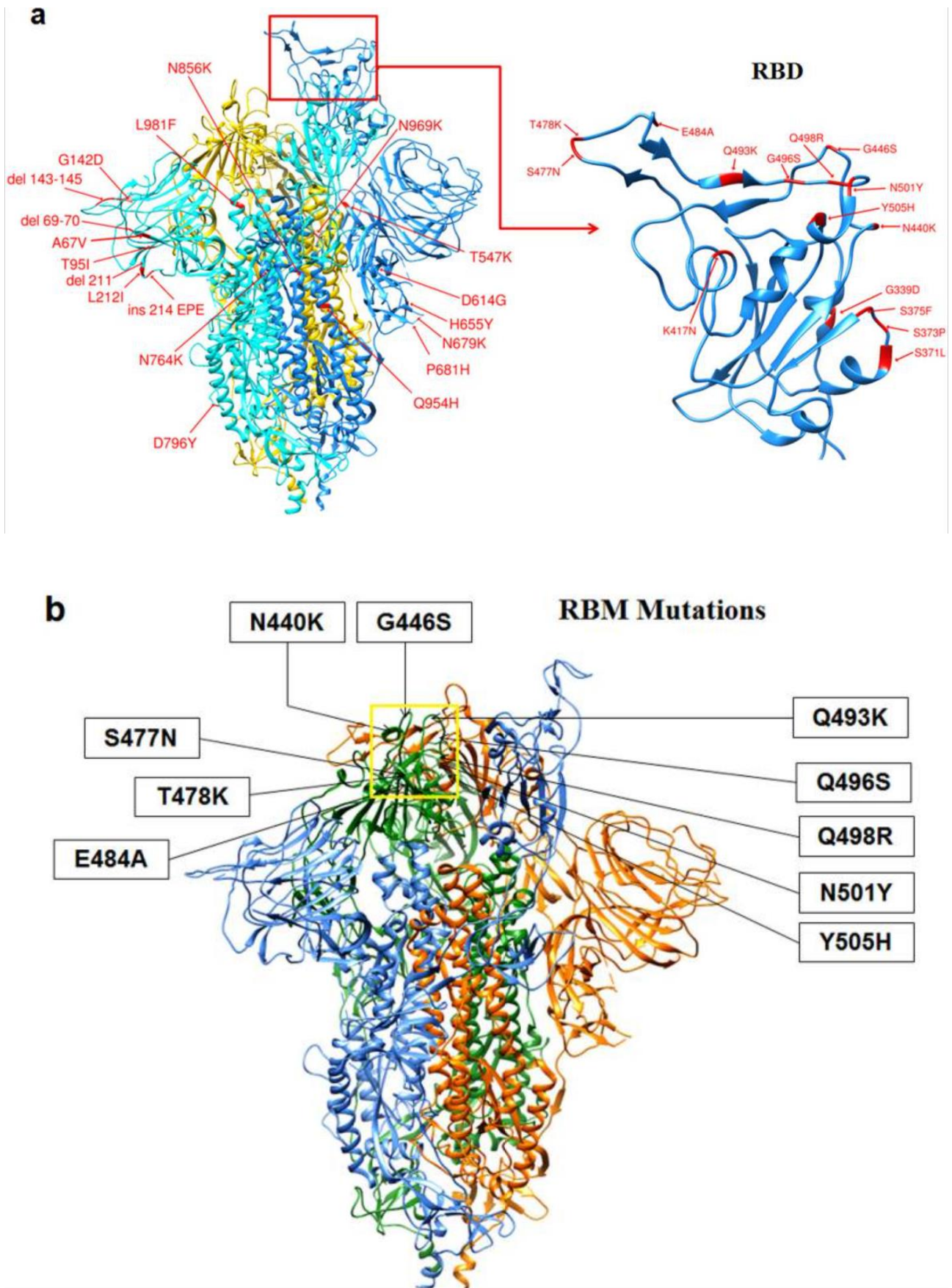


Fig. 36 Mutazioni mappate sulla glicoproteina S di Omicron. (A) Insieme delle mutazioni presenti su Spike con evidenziato il dominio RBD. (B) Mutazioni sull'RBM. [159]

Uno studio ha analizzato i punteggi $\Delta\Delta G$, le interazioni interatomiche, la fluttuazione e le analisi di deformazione per comprendere le proprietà di queste mutazioni (D614G, E484A, N501Y, Q493K, K417N, S477N, Y505H, G496S) [159]. Il punteggio $\Delta\Delta G$ illustra come le mutazioni puntiformi modificano la stabilità della proteina [160].

Tutte queste mutazioni, con vari meccanismi, vanno ad influenzare l'infettività e l'evasione immunitaria:

- **D614G** causa un aumento della flessibilità della catena proteica principale che ha mostrato un alto tasso di infettività con una migliore affinità di legame della proteina S al recettore ACE2. Questa mutazione aumenta la fitness virale e amplifica le proprietà di replicazione e trasmissione del virus [161,162,163,164].
- **E484A** aumenta la flessibilità molecolare ed ha un ruolo cruciale nel promuovere l'infettività virale, la trasmissibilità e/o l'antigenicità. Questa mutazione riduce la potenza di neutralizzazione degli anticorpi [165,166].
- **N501Y** causa un aumento della flessibilità molecolare e delle interazioni elettrostatiche tra la proteina S e ACE2 mediante un forte legame idrogeno vicino al sito di mutazione. Questo legame influenza la fusione del virus con le cellule ospiti e porta all'aumento della velocità di trasmissione virale [167,168,169,170].
- **Q493K** diminuisce leggermente la flessibilità molecolare, contribuisce notevolmente alle energie di legame dell'RBD e del recettore ACE2 e quasi raddoppia il potenziale elettrostatico del complesso proteina S-ACE2 [171,172].
- **K417N** causa un aumento dell'energia libera di legame del complesso RBD-ACE2. Tuttavia, questa singola mutazione tende anche a destabilizzare questo complesso attraverso un legame più lento e una dissociazione più veloce [162,167,170,173].
- **S477N** aumenta l'affinità di legame del dominio RBD per il recettore ACE2. Inoltre, contribuisce al meccanismo di evasione immunitaria [167,174,175].
- **Y505H** causa una diminuzione della stabilità della proteina Spike e un maggior rischio di malattia virale [176].
- **G496S** cambia la forma del motivo di legame al recettore (RBM) e genera un'interferenza sterica per il legame degli anticorpi (Abs) all'RBD della Spike, provocando la perdita delle interazioni Ab-RBD [177,178].

7. Discussione

Analizzando i dati di sequenziamento è stato possibile constatare l'altissimo tasso di mutazione del virus SARS-CoV-2. Questa proprietà non solo ha reso particolarmente difficile il tracciamento delle varianti a livello globale, ma anche lo studio delle caratteristiche associate ad ognuna di esse. Nel periodo di riferimento (aprile 2022 – aprile 2023) sono stati rilevati un totale di 51 lignaggi differenti.

Un enorme sforzo è stato fatto per implementare una nomenclatura dinamica al fine di classificare i lignaggi del SARS-CoV-2. Per questo scopo è stato sviluppato il software PANGOLIN (*Phylogenetic Assignment of Named Global Outbreak Lineages*) disponibile da aprile 2020.

La variante Omicron, con i suoi lignaggi e sotto-lignaggi, è stata l'unica ad essere rilevata durante il periodo di sequenziamento. Uno studio ha rilevato che la variante Omicron ha il più alto numero di mutazioni rispetto alle altre VOC (Alpha, Gamma e Delta) ed anche la sua proteina Spike risulta essere la più mutata. La proteina Spike di Omicron è quella con più mutazioni rispetto a tutte le proteine strutturali (E, M, N) ed anche alle proteine non strutturali (NSP). In particolare, le due regioni più variabili della glicoproteina sono il dominio RBD e al secondo posto il dominio N-terminale (NTD) [159] (**Fig. 37**). Questi dati non sorprendono in quanto, essendo coinvolta con il legame al recettore ACE2, la proteina S è sottoposta ad una maggiore pressione selettiva.

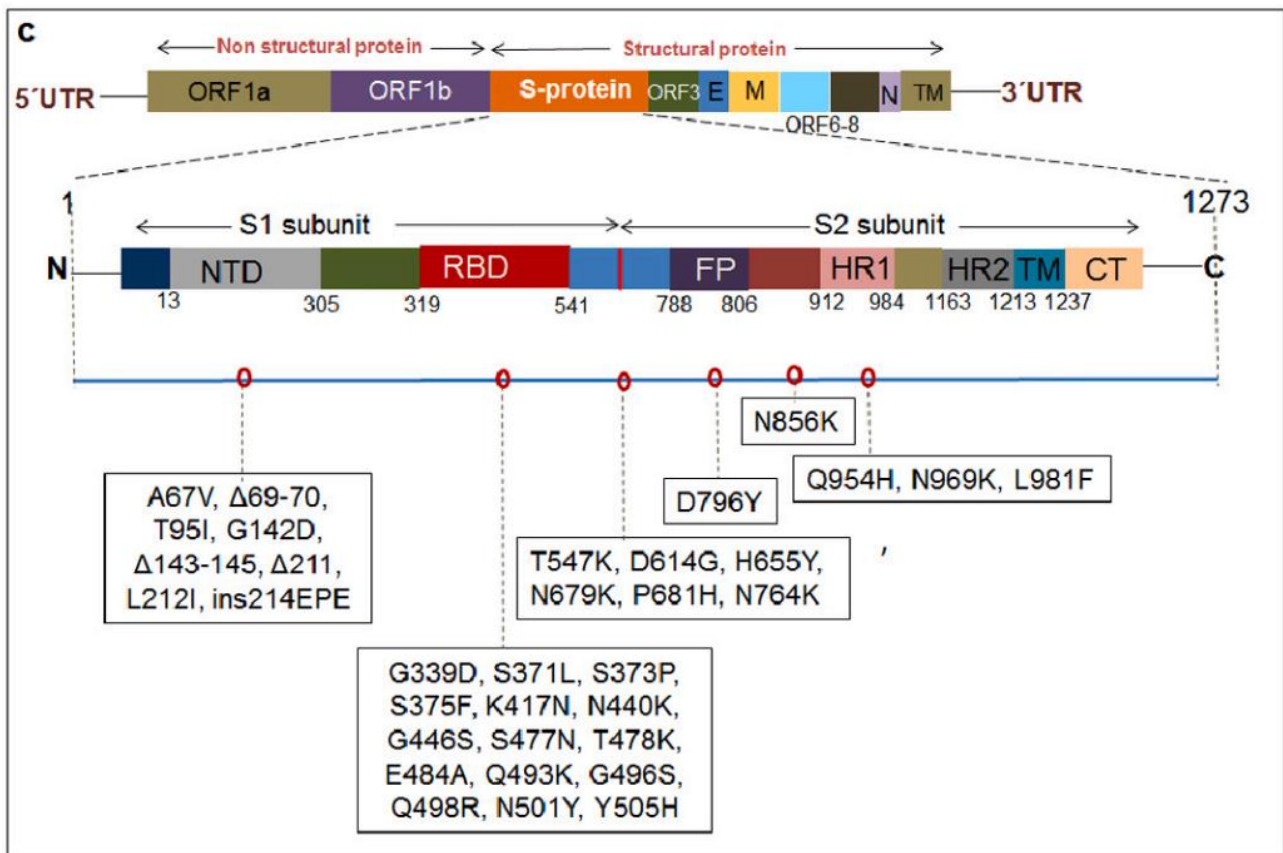


Fig. 37 Il diagramma schematico mostra le mutazioni nelle diverse regioni delle subunità S1 e S2 nella glicoproteina Spike in Omicron. [159]

A causa di queste caratteristiche la variante Omicron si è rivelata estremamente diffusiva ed in poco tempo ha sostituito la precedente variante Delta. Analizzando i rapporti periodici rilasciati dall'Istituto Superiore di Sanità (ISS) si può avere una visione più chiara della situazione. Le informazioni presentate nei rapporti derivano dai dati inviati dalle Regioni relativi a una diagnosi di infezione da virus SARS-CoV-2. Questi dati sono elaborati dall'ISS integrando i dati microbiologici ed epidemiologici e restituendo la distribuzione nel tempo e nello spazio delle varianti di interesse per la sanità pubblica notificate al Sistema di Sorveglianza Integrata COVID-19 e le caratteristiche delle persone che hanno contratto una infezione causata da una di queste varianti.

I rapporti periodici includono anche un'analisi dettagliata dei dati di caratterizzazione genomica degli isolati SARS-CoV-2 che le Regioni caricano sulla piattaforma I-Co-Gen, con particolare riferimento alle varianti virali identificate.

Nel rapporto del 29 aprile 2022 l'analisi delle sequenze depositate in I-Co-Gen negli ultimi 45 giorni mostra che la variante Omicron rappresenta il 99.99% dei sequenziamenti depositati, il lignaggio BA.2 risulta essere il più frequente (62.8% dei sequenziamenti totali), seguito dal BA.2.9 (12.9%), dal BA.1.1 (9.5%) e dal BA.2.3 (5.1%) (**Fig. 38**). Inoltre, il Sistema di Sorveglianza Integrata riporta 5.533 casi confermati di infezione da SARS-CoV-2 sequenziati/genotipizzati nel marzo 2022. Nel periodo 13 marzo – 26 aprile 2022 sono stati rilevati 4175 casi di Omicron e 7 casi di Delta (B.1.617.2). Inoltre, la prevalenza di casi dovuti ad infezione da variante Omicron si osserva in tutte le fasce di età (2201 casi nelle femmine e 1974 casi nei maschi).

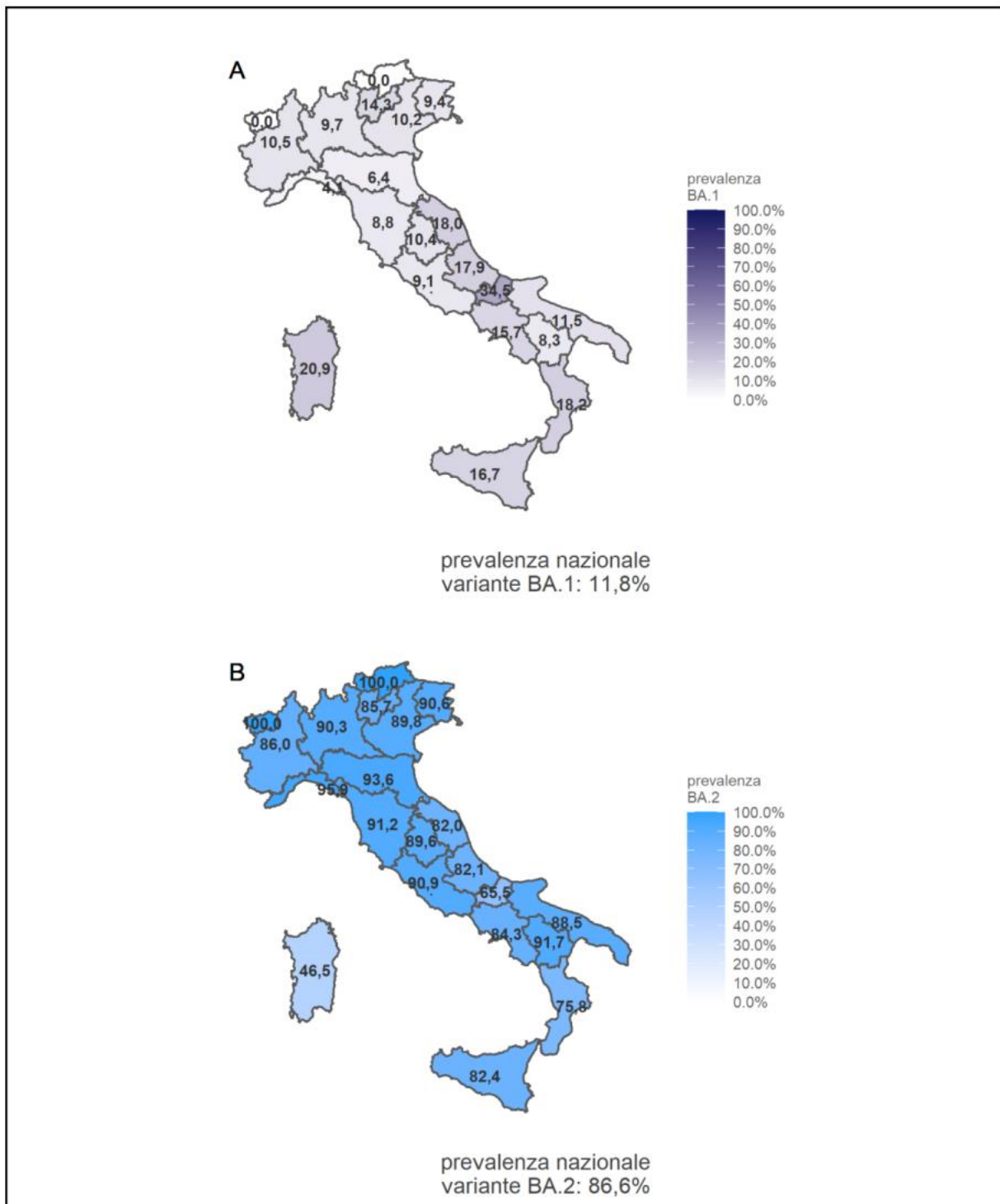


Fig. 38 Prevalenza della variante omicron - sottolignaggio BA.1 (A) e sottolignaggio BA.2 (B) - in Italia, indagine rapida del 4 aprile 2022. Dalla mappa si evince la circolazione predominante del sottolignaggio BA.2 (range: 46.5% - 100%) rispetto al sottolignaggio BA.1 (range: 0% - 34.5%). Nella mappa sono inclusi tutti i sotto-lignaggi.

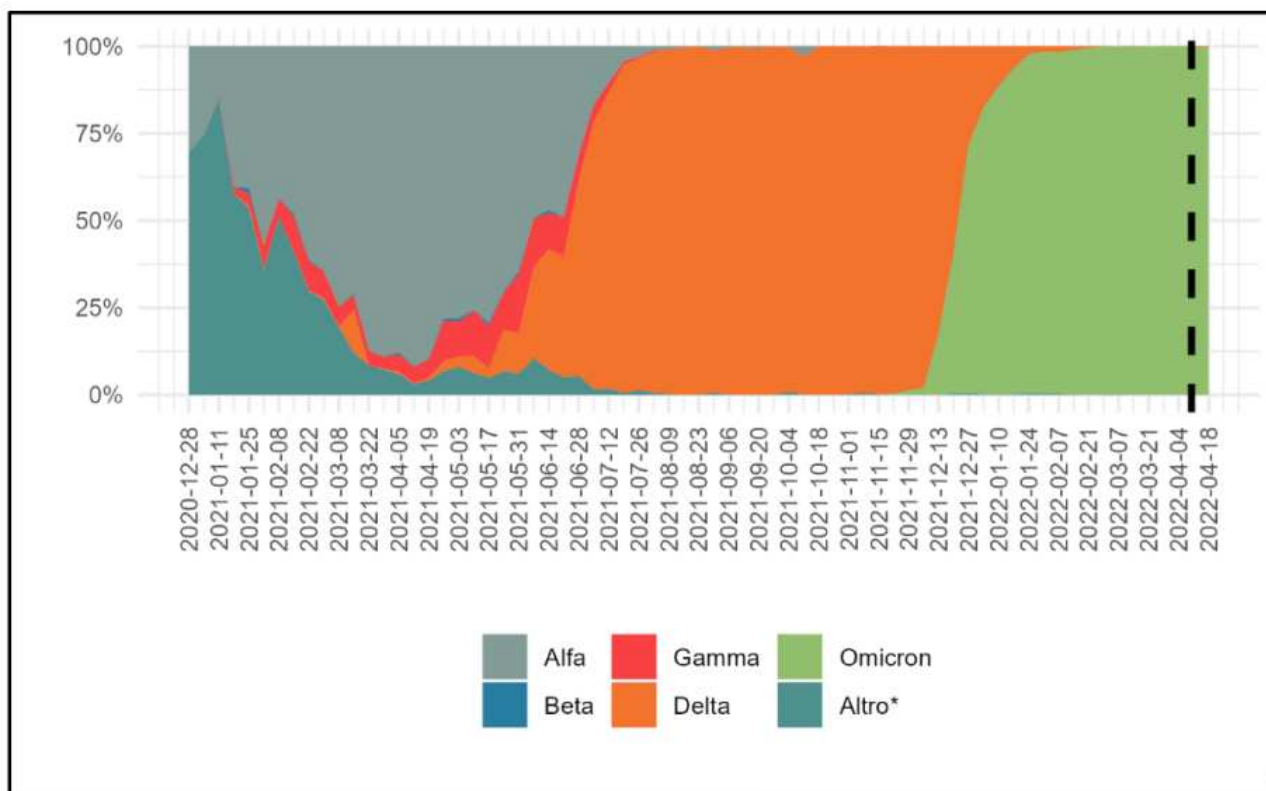


Grafico 5 Andamento delle principali varianti identificate mediante sequenziamento dal software della piattaforma I-Co-Gen, per settimana di campionamento (28 dicembre 2020 - 26 aprile 2022).

Come si evince dal grafico 5, il numero di sequenze attribuibili alla variante Omicron inizia ad aumentare rapidamente già dall'ultima settimana di dicembre 2021 per poi divenire predominante.

Dato che da aprile 2022 in poi la variante Omicron è stata quella principale su tutto il territorio italiano, è interessante analizzare anche il report di aprile 2023 per vedere come si è evoluta la situazione e come si sono ramificati e diffusi i vari sotto-lignaggi.

Nel rapporto del 28 aprile 2023 i dati di sequenziamento delle ultime sei settimane mostrano che la variante XBB risulta predominante (76.4%) con 62 sotto-lignaggi identificati. Il lignaggio XBB.1.5 (variante Kraken) è quello maggioritario (30.9%), mentre al secondo posto c'è XBB.1.9.1 (Hyperion) con una frequenza pari al 14.5%. Risultano anche 8 sequenziamenti attribuibili a XBB.1.16.

Nel periodo 13 marzo – 23 aprile 2023 sono state identificate 199 sequenze riconducibili alla variante BA.2.75 (variante Centaurus), 286 sequenze di BQ.1, di cui la BQ.1.1 (variante Cerberus) era quella prevalente con 92 sequenze depositate, e 1730 sequenze del lignaggio XBB (**Fig. 39**).

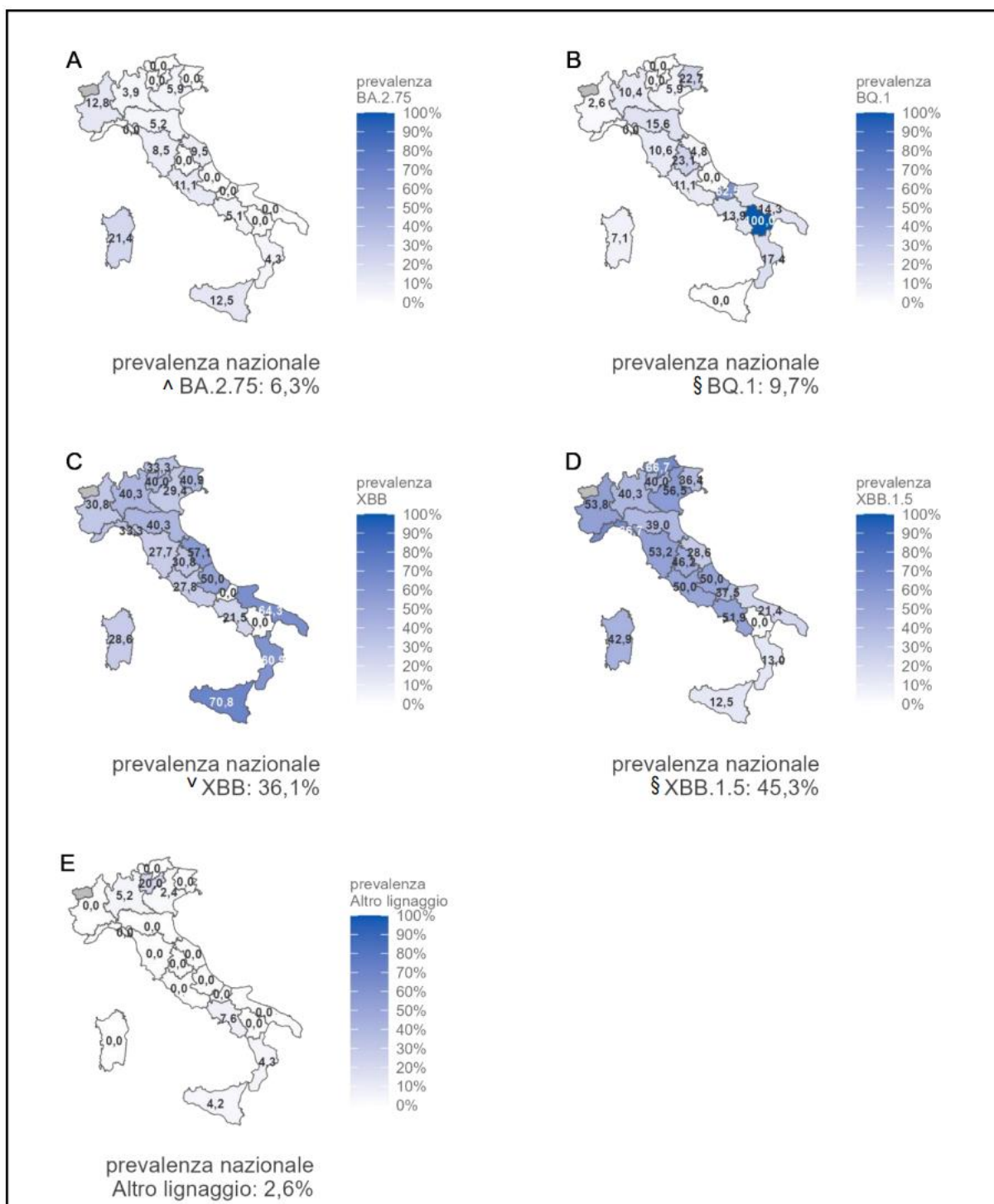


Fig. 39 Prevalenza di BA.2.75, BQ.1, XBB, XBB.1.5 e altri lignaggi in Italia, indagine rapida del 4 aprile 2023. Nelle percentuali riportate per ogni regione sono inclusi tutti i sotto-lignaggi relativi a BA.2.75, BQ.1, XBB (escluso XBB.1.5) e XBB.1.5.

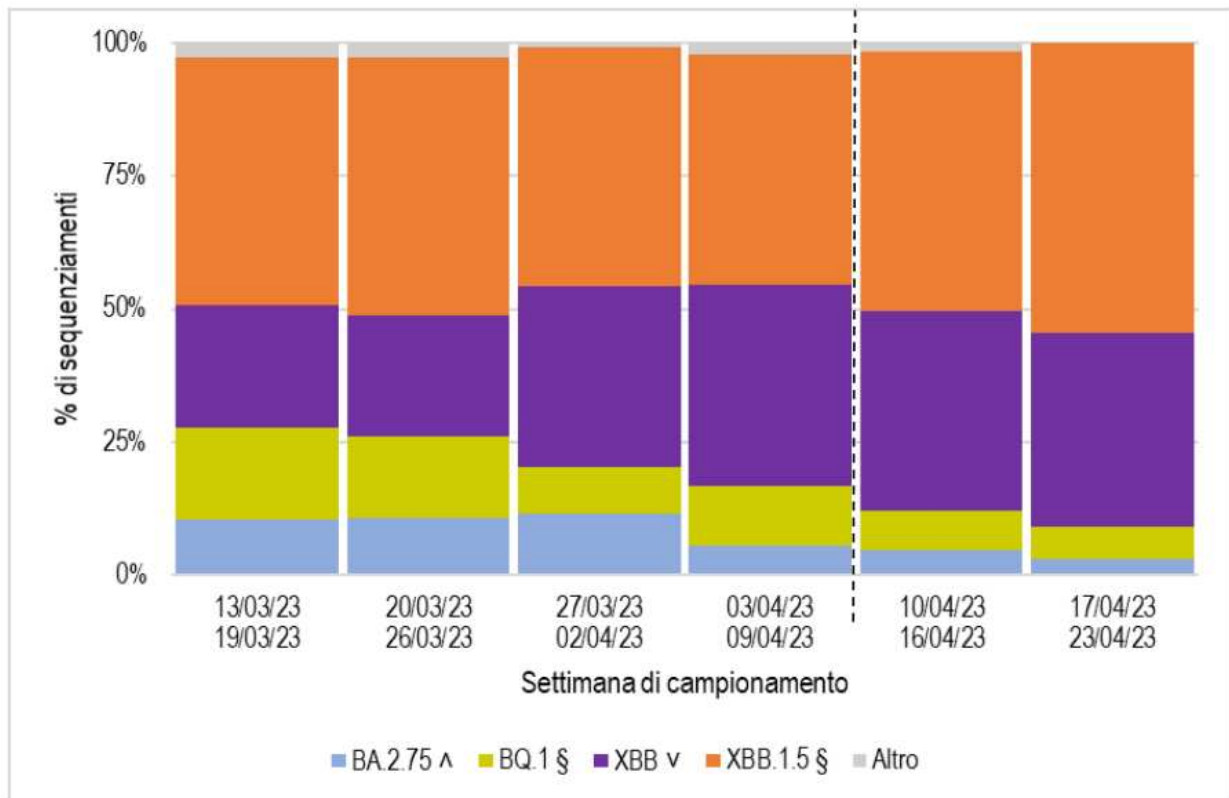


Grafico 6 Distribuzione percentuale dei principali lignaggi identificati mediante sequenziamento genomico dal software della piattaforma I-Co-Gen, per settimana di campionamento (ultime sei settimane).

Nel grafico 6 si può notare una diminuzione di sequenziamenti attribuibili a BQ.1 e BA.2.75 in favore delle varianti XBB e XBB.1.5.

È interessante notare come quando si presenti una qualche mutazione anche minimamente vantaggiosa per il virus, la variante che porta tale modifica emerga immediatamente rispetto alle altre. Ovviamente questo si riflette anche sul numero di sequenze depositate e quindi sul numero di casi positivi al SARS-CoV-2. Ad esempio, il lignaggio XBB.1.5 (variante Kraken) porta una mutazione aggiuntiva sulla proteina Spike (F486P). Questa modifica presente nell'RBD sembrerebbe aumentare ulteriormente l'affinità del virus al recettore ACE2 andando di conseguenza ad influenzare i tassi di trasmissione e infezione.

8. Conclusioni

L'epidemia di SARS-COV-2 ha fornito uno slancio alla ricerca e ha portato alla rapida adozione della tecnologia NGS per affrontare l'urgente sfida di tracciare le varianti nel tentativo di ridurre la diffusione del virus e di monitorare la risposta ai vaccini.

I metodi NGS sono stati rapidamente adattati al paradigma SARS-CoV-2 e hanno dimostrato di essere applicabili a un'ampia varietà di questioni biologiche associate. Il tasso di produzione e analisi dei dati è stato senza precedenti e sarebbe stato inconcepibile solo pochi anni fa.

In pochi mesi i dati della sequenza genomica hanno consentito la ricostruzione del probabile tempo di spillover di SARS-CoV-2 nella popolazione umana, lo sviluppo di sistemi per la classificazione dei ceppi virali, fondamentali per monitorare la diffusione del virus, e l'identificazione di siti nel genoma di SARS-CoV-2 che potrebbero essere sotto l'influenza di varie pressioni selettive.

In poco tempo è stata prodotta una grandissima mole di dati e la sfida è stata proprio quella di gestirli ed integrarli al fine di migliorare la comprensione delle malattie infettive e i rimedi atti a contrastarle.

Uno degli effetti positivi della pandemia è stato sicuramente quello di rafforzare e sviluppare ulteriormente il dialogo e la cooperazione tra i vari centri di ricerca e ha contribuito a valorizzare il lavoro dei laboratori più piccoli.

Bibliografía

1. Esteban Ortiz-Prado, Katherine Simbaña-Rivera, Lenin Gómez- Barreno, Mario Rubio-Neira, Linda P. Guaman, Nikolaos C Kyriakidis, Claire Muslin, Ana María Gómez Jaramillo, Carlos Barba-Ostria, Doménica Cevallos-Robalino, Hugo Sanches-SanMiguel, Luis Unigarro, Rasa Zalakeviciute, Naomi Gadian, Andrés López-Cortés. *Clinical, molecular, and epidemiological characterization of the SARS-CoV-2 virus and the Coronavirus Disease 2019 (COVID-19), a comprehensive literature review.* Diagnostic Microbiology and Infectious Disease 2020. <https://doi.org/10.1016/j.diagmicrobio.2020.115094>.
2. Walls AC, Park Y-J, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell 2020.
3. Nikhil Kirtipala, Shiv Bharadwajb, Sang Gu Kangb. *From SARS to SARS-CoV-2, insights on structure, pathogenicity and immunity aspects of pandemic human coronaviruses.* Infection, Genetics and Evolution 2020. <https://doi.org/10.1016/j.meegid.2020.104502>.
4. Cheever, F.S., Daniels, J.B., et al., 1949. A murine virus (JHM) causing disseminated encephalomyelitis with extensive destruction of myelin. J. Exp. Med.
5. Hamre, D., Procknow, J.J., 1966. A new virus isolated from the human respiratory tract. Proc. Soc. Exp. Biol. Med.
6. McIntosh, K., Dees, J.H., Becker, W.B., Kapikian, A.Z., Chanock, R.M., 1967. *Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease.* Proc. Natl. Acad. Sci. U. S. A.
7. van der Hoek, L., Pyrc, K., Jebbink, M.F., Vermeulen-Oost, W., Berkhout, R.J., Wolthers, K.C., Wertheim-van Dillen, P.M., Kaandorp, J., Spaargaren, J., Berkhout, B., 2004. Identification of a new human coronavirus. Nat. Med.
8. [Woo, P.C., Lau, S.K., Chu, C.M., Chan, K.H., Tsoi, H.W., Huang, Y., Wong, B.H., Poon, R.W., Cai, J.J., Luk, W.K., Poon, L.L., Wong, S.S., Guan, Y., Peiris, J.S., Yuen, K.Y., 2005. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J. Virol.](#)
9. [Woo, P.C.Y., Lau, S.K.P., Lam, C.S.F., Lau, C.C.Y., Tsang, A.K.L., Lau, J.H.N., Bai, R., Teng, J.L.L., Tsang, C.C.C., Wang, M., Zheng, B.J., Chan, K.H., Yuen, K.Y., 2012. Discovery of seven novel mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of alphacoronavirus and betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus. J. Virol.](#)
10. [Kanwar, A., Selvaraju, S., Esper, F., 2017. Human coronavirus-HKU1 infection among adults in Cleveland, Ohio. Open Forum Infect Dis 4.](#)
11. [Pene, F., Merlat, A., Vabret, A., Rozenberg, F., Buzyn, A., Dreyfus, F., Cariou, A., Freymuth, F., Lebon, P., 2003. Coronavirus 229E-related pneumonia in immunocompromised patients. Clinical Infectious Diseases Official Publication Infectious Diseases Society of America 37.](#)
12. [Walsh, E.E., Shin, J.H., Falsey, A.R., 2013. Clinical impact of human coronaviruses 229E and OC43 infection in diverse adult populations. J. Infect. Dis. 208.](#)
13. [Arbour, N., Ekande, S., Cote, G., Lachance, C., Chagnon, F., Tardieu, M., Cashman, N.R., Talbot, P.J., 1999. Persistent infection of human oligodendrocytic and neuroglial cell lines by human coronavirus 229E. J. Virol. 73.](#)
14. [Arbour, N., Day, R., Newcombe, J., Talbot, P.J., 2000. Neuroinvasion by human respiratory coronaviruses. J. Virol. 74.](#)
15. [Jacomy, H., Fragoso, G., Almazan, G., Mushynski, W.E., Talbot, P.J., 2006. Human coronavirus OC43 infection induces chronic encephalitis leading to disabilities in BALB/ C mice. Virology 349.](#)
16. [Smuts, H., 2008. Human coronavirus NL63 infections in infants hospitalised with acute respiratory tract infections in South Africa. Influenza Other Respir. Viruses 2.](#)

17. [Vabret, A., Mourez, T., Gouarin, S., Petitjean, J., Freymuth, F., 2003. An outbreak of coronavirus OC43 respiratory infection in Normandy, France. Clin. Infect. Dis. 36.](#)
18. [van Boheemen, S., de Graaf, M., Lauber, C., Bestebroer, T.M., Raj, V.S., Zaki, A.M., Osterhaus, A.D.M.E., Haagmans, B.L., Gorbalenya, A.E., Snijder, E.J., Fouchier, R.A.M., 2012. Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. Mbio 3.](#)
19. [de Groot, R.J., Baker, S.C., Baric, R.S., Brown, C.S., Drosten, C., Enjuanes, L., Fouchier, R.A.M., Galiano, M., Gorbalenya, A.E., Memish, Z.A., Perlman, S., Poon, L.L.M., Snijder, E.J., Stephens, G.M., Woo, P.C.Y., Zaki, A.M., Zambon, M., Ziebuhr, J., 2013. Middle east respiratory syndrome coronavirus \(MERS-CoV\): announcement of the coronavirus study group. J. Virol. 87.](#)
20. [Gorbalenya, A.E., Snijder, E.J., Spaan, W.J.M., 2004. Severe acute respiratory syndrome coronavirus phylogeny: toward consensus. J. Virol. 78.](#)
21. [Gorbalenya, A.E., Baker, S.C., Baric, R.S., de Groot, R.J., Drosten, C., Gulyaeva, A.A., Haagmans, B.L., Lauber, C., Leontovich, A.M., Neuman, B.W., Penzar, D., Perlman, S., Poon, L.L.M., Samborskiy, D.V., Sidorov, I.A., Sola, I., Ziebuhr, J., Coronaviridae Study Group of the International Committee on Taxonomy of V, 2020. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat. Microbiol. 5.](#)
22. Jocelyne Piret and Guy Boivin. *Pandemics Throughout History*. Frontiers in Microbiology 2021. doi: 10.3389/fmicb.2020.631736
23. Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., et al. (2020). A novel coronavirus from patients with pneumonia in China, 2019. N. Engl. J. Med. 382. doi:10.1056/NEJMoa2001017
24. Lau, S. K. P., Luk, H. K. H., Wong, A. C. P., Li, K. S. M., Zhu, L., He, Z., et al. (2020). Possible bat origin of severe acute respiratory syndrome coronavirus 2. Emerg. Infect. Dis. 26. doi:10.3201/eid2607.200092
25. Lam, T. T., Jia, N., Zhang, Y. W., Shum, M. H., Jiang, J. F., Zhu, H. C., et al. (2020). Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. Nature 583. doi: 10.1038/s41586-020-2169-0
26. Wiersinga, W. J., Rhodes, A., Cheng, A. C., Peacock, S. J., and Prescott, H. C. (2020). Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): a review. JAMA 324. doi: 10.1001/jama.2020.12839
27. Richardson, S., Hirsch, J. S., Narasimhan, M., Crawford, J. M., McGinn, T., Davidson, K. W., et al. (2020). Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. JAMA 323. doi: 10.1001/jama.2020.6775
28. Wilson, N., Kvalsvig, A., Barnard, L. T., and Baker, M. G. (2020). Case-fatality risk estimates for COVID-19 calculated by using a lag time for fatality. Emerg. Infect. Dis. 26. doi:10.3201/eid2606.200320
29. Rowley, A. H., Shulman, S. T., and Arditi, M. (2020). Immune pathogenesis of COVID-19-related multisystem inflammatory syndrome in children (MIS-C). J. Clin. Invest. 130. doi: 10.1172/JCI143840
30. Bastard, P., Rosen, L. B., Zhang, Q., Michailidis, E., Hoffmann, H. H., Zhang, Y., et al. (2020). Auto-antibodies against type I IFNs in patients with lifethreatening COVID-19. Science 370:eabd4585. doi: 10.1126/science.abd4585
31. Zhang, Q., Bastard, P., Liu, Z., Le Pen, J., Moncada-Velez, M., Chen, J., et al. (2020). Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. Science 370:eabd4570. doi: 10.1126/science.abd4570
32. World Health Organization [WHO] (2020). WHO Coronavirus Disease (COVID-19) Dashboard. Available online at: <https://covid19.who.int/> (accessed December 18, 2020).
33. Ganyani, T., Kremer, C., Chen, D., Torneri, A., Faes, C., Wallinga, J., et al. (2020). Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data. March 2020. Euro Surveill 25:2000257. doi: 10.2807/1560-7917.ES.2020.25.17.2000257

34. Batty, C. J., Heise, M. T., Bachelder, E. M., and Ainslie, K. M. (2020). Vaccine formulations in clinical development for the prevention of severe acute respiratory syndrome coronavirus 2 infection. *Adv. Drug. Deliv. Rev.* S0169- 409X. doi: 10.1016/j.addr.2020.12.006
35. Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of the SARSCoV-2 by full-length human ACE2. *Science* 2020.
36. Gralinski LE, Menachery VD. Return of the coronavirus: 2019-nCoV. *Viruses* 2020.
37. [Mousavizadeh, L., Ghasemi, S., Immunology Infection, 2020. Genotype and phenotype of COVID-19: Their roles in pathogenesis.](#)
38. [Yang, H.T., Bartlam, M., Rao, Z.H., 2006. Drug design targeting the main protease, the Achilles' heel of coronaviruses. *Curr Pharm Design* 12.](#)
39. [Yang, D., Leibowitz, J.L., 2015. The structure and functions of coronavirus genomic 3' and 5' ends. *Virus Res.* 206.](#)
40. [Fehr, A.R., Perlman, S., 2015. Coronaviruses: an overview of their replication and pathogenesis. *Methods in Molecular Biology Clifton, N.J.* 1282.](#)
41. [Harcourt, B.H., Jukneliene, D., Kanjanahaluethai, A., Bechill, J., Severson, K.M., Smith, C.M., Rota, P.A., Baker, S.C., 2004. Identification of severe acute respiratory syndrome coronavirus replicase products and characterization of papain-like protease activity. *J. Virol.* 78.](#)
42. Yang, H.T., Xie, W.Q., Xue, X.Y., Yang, K.L., Ma, J., Liang, W.X., Zhao, Q., Zhou, Z., Pei, D.Q., Ziebuhr, J., Hilgenfeld, R., Yuen, K.Y., Wong, L., Gao, G.X., Chen, S.J., Chen, Z., Ma, D.W., Bartlam, M., Rao, Z.H., 2005. Design of wide-spectrum inhibitors targeting coronavirus main proteases. [published correction appears in *PLoS Biol.* 2005 Nov;3(11):e428]. *PLoS Biol.* 3 (10). <https://doi.org/10.1371/journal.pbio.0030324>.
43. [Zhang, L.L., Lin, D.Z., Sun, X.Y.Y., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K., Hilgenfeld, R., 2020. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved alpha-ketoamide inhibitors. *Science* 368 409-+.](#)
44. [Snijder, E.J., Bredenbeek, P.J., Dobbe, J.C., Thiel, V., Ziebuhr, J., Poon, L.L.M., Guan, Y., Rozanov, M., Spaan, W.J.M., Gorbalenya, A.E., 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J. Mol. Biol.* 331.](#)
45. [Fehr, A.R., Athmer, J., Channappanavar, R., Phillips, J.M., Meyerholz, D.K., Perlman, S., 2015. The nsp3 macrodomain promotes virulence in mice with coronavirus-induced encephalitis. *J. Virol.* 89.](#)
46. Ahmad Abu Turab Naqvia, Kisa Fatimab, Taj Mohammada, Urooj Fatimac, Indrakant K. Singhd, Archana Singhe, Shaikh Muhammad Atiff, Gururao Hariprasadg, Gulam Mustafa Hasanh, Md. Imtaiyaz Hassana. *Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach.* BBA - Molecular Basis of Disease 2020. <https://doi.org/10.1016/j.bbadis.2020.165878>
47. B. Xue, D. Blocquel, J. Habchi, A.V. Uversky, L. Kurgan, V.N. Uversky, S. Longhi, Structural disorder in viral proteins, *Chem. Rev.* 114 (2014).
48. Entedar A.J. Alsadi, I.M. Jones, Membrane binding proteins of coronaviruses, *Future Virol* 14 (2019).
49. C. Tang, Z. Deng, X. Li, M. Yang, Z. Tian, Z. Chen, G. Wang, W. Wu, W.H. Feng, G. Zhang, Helicase of Type 2 porcine reproductive and respiratory syndrome virus strain hv reveals a unique structure, *Viruses* 12 (2020) 215.
50. C. Muller, F.W. Schulte, K. Lange-Grunweller, W. Obermann, R. Madhugiri, S. Pleschka, J. Ziebuhr, R.K. Hartmann, A. Grunweller, Broad-spectrum antiviral activity of the eIF4A inhibitor silvestrol against corona- and picornaviruses, *Antivir. Res.* 150 (2018).
51. Mei-Yue Wang, Rong Zhao, Li-Juan Gao, Xue-Fei Gao, De-Ping Wang and Ji-Min Cao. *SARS-CoV-2: Structure, Biology, and Structure-Based Therapeutics Development.* *Front. Cell. Infect. Microbiol.* 10:587269. doi: 10.3389/fcimb.2020.587269

52. Li, W., Moore, M. J., Vasilieva, N., Sui, J., Wong, S. K., Berne, M. A., et al. (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* 426. doi: 10.1038/nature02145
53. Li, F., Li, W., Farzan, M., and Harrison, S. C. (2005). Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Sci. (N. Y. N.Y.)* 309. doi: 10.1126/science.1116480
54. Li, F. (2016). Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annu. Rev. Virol.* 3. doi: 10.1146/annurev-virology-110615-042301
55. Walls, A. C., Park, Y. J., Tortorici, M. A., Wall, A., McGuire, A. T., and Velesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 181. doi: 10.1016/j.cell.2020.02.058
56. Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C. L., Abiona, O., et al. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Sci. (N. Y. N.Y.)* 367. doi: 10.1126/science.abb2507
57. Letko, M., Marzi, A., and Munster, V. (2020). Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat. Microbiol.* 5. doi: 10.1038/s41564-020-0688-y
58. Wang, Q., Zhang, Y., Wu, L., Niu, S., Song, C., Zhang, Z., et al. (2020). Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* 181. doi: 10.1016/j.cell.2020.03.045
59. Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., et al. (2020). Structure of the SARSCoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* 581. doi: 10.1038/s41586-020-2180-5
60. Li, W. (2020). Delving deep into the structural aspects of a furin cleavage site inserted into the spike protein of SARS-CoV-2: A structural biophysical perspective. *Biophys. Chem.* 264, 106420. doi: 10.1016/j.bpc.2020.106420
61. Sternberg, A., and Naujokat, C. (2020). Structural features of coronavirus SARSCoV-2 spike protein: Targets for vaccination. *Life Sci.* 257, 118056. doi: 10.1016/j.lfs.2020.118056
62. Johnson, B. A., Xie, X., Kalveram, B., Lokugamage, K. G., Muruato, A., Zou, J., et al. (2020). Furin Cleavage Site Is Key to SARS-CoV-2 Pathogenesis. *Preprint Server Biol.* doi: 10.1101/2020.08.26.268854
63. S. Li, L. Yuan, G. Dai, R.A. Chen, D.X. Liu, T.S. Fung, Regulation of the ER stress response by the Ion Channel activity of the infectious bronchitis coronavirus envelope protein modulates Virion release, apoptosis, viral fitness, and pathogenesis, *Front. Microbiol.* 10 (2020) 3022.
64. T. Tang, M. Bidon, J.A. Jaimes, G.R. Whittaker, S. Daniel, Coronavirus membrane fusion mechanism offers as a potential target for antiviral development, *Antiviral Research* (2020) 104792.
65. B.W. Neuman, M.J. Buchmeier, Supramolecular architecture of the coronavirus particle, *Adv. Virus Res.* 96 (2016).
66. Y. Cong, M. Ulasli, H. Schepers, M. Mauthe, P. V'kovski, F. Kriegenburg, V. Thiel, C.A. de Haan, F. Reggiori, Nucleocapsid Protein recruitment to replication-transcription complexes plays a crucial role in coronaviral life cycle, *Journal of Virology* 94 (2020).
67. R.L. Graham, J.S. Sparks, L.D. Eckerle, A.C. Sims, M.R. Denison, SARS coronavirus replicase proteins in pathogenesis, *Virus Res.* 133 (2008).
68. Z. Ruan, C. Liu, Y. Guo, Z. He, X. Huang, X. Jia, T. Yang, Potential Inhibitors Targeting RNA-Dependent RNA Polymerase Activity (NSP12) of SARS-CoV-2, (2020).
69. Gao, Y., Yan, L., Huang, Y., Liu, F., Zhao, Y., Cao, L., et al. (2020). Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Sci. (N. Y. N.Y.)* 368. doi: 10.1126/science.abb7498
70. Yin, W., Mao, C., Luan, X., Shen, D. D., Shen, Q., Su, H., et al. (2020). Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV2 by remdesivir. *Sci. (N. Y. N.Y.)* 368. doi: 10.1126/science.abc1560
71. Hillen, H. S., Kokic, G., Farnung, L., Dienemann, C., Tegunov, D., and Cramer, P. (2020). Structure of replicating SARS-CoV-2 polymerase. *Nature* 584. doi: 10.1038/s41586-020-2368-8

72. Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., et al. (2020). Structure of M (pro) from SARS-CoV-2 and discovery of its inhibitors. *Nature* 582. doi: 10.1038/s41586-020-2223-y
73. Yang, H., Xie, W., Xue, X., Yang, K., Ma, J., Liang, W., et al. (2005). Design of widespectrum inhibitors targeting coronavirus main proteases. *PLoS Biol.* 3, e324. doi: 10.1371/journal.pbio.0030324
74. Zhang L, Lin D, Sun X, Curth U, Drosten C, Sauerhering L, et al. Crystal structure of SARSCoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science* 2020a:eabb3405.
75. Wan Y, Shang J, Graham R, Baric RS, Li F. Receptor recognition by novel coronavirus from Wuhan: An analysis based on decade-long structural studies of SARS. *J Virol* 2020;94 (7), e00127-20.
76. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 2020a;579: 270–3.
77. Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, et al. A new coronavirus associated with human respiratory disease in China. *Nature* 2020c;579:265–9.
78. Kuba K, Imai Y, Rao S, Gao H, Guo F, Guan B, et al. A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. *Nat Med* 2005;11:875–9
79. Xu H, Zhong L, Deng J, Peng J, Dan H, Zeng X, et al. High expression of ACE2 receptor of 2019-nCoV on the epithelial cells of oral mucosa. *Int J Oral Sci* 2020a;12, 8.
80. Forrest ARR, Kawaji H, Rehli M, Baillie JK, MJL De Hoon, Haberle V, et al. A promoter-level mammalian expression atlas. *Nature* 2014;507:462–70.
81. de Wilde AH, Snijder EJ, Kikkert M, van Hemert MJ. Host factors in coronavirus replication. In: *Current Topics in Microbiology and Immunology*. Springer Verlag; 2018.
82. Knoops K, Kikkert M, Van Den Worm SHE, Zevenhoven-Dobbe JC, Van Der Meer Y, Koster AJ, et al. SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS Biol* 2008;6:1957–74.
83. Weiss SR, Navas-Martin S. Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. *Microbiol Mol Biol Rev* 2005;69:635–64.
84. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 Novel Coronavirus: implications for virus origins and receptor binding. *Lancet* 2020;395:565–74.
85. Neuman BW, Kiss G, Kunding AH, Bhella D, Baksh MF, Connelly S, et al. A structural analysis of M protein in coronavirus assembly and morphology. *J Struct Biol* 2011;174:11–22.
86. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 Novel Coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 2020a;395:507–13.
87. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 Novel Coronavirus in Wuhan, China. *Lancet* 2020a;395:497–506.
88. Mistry P, Barmania F, Mellet J, Peta K, Strydom A, Viljoen IM, James W, Gordon S and Pepper MS (2022). *SARS-CoV-2 Variants, Vaccines, and Host Immunity*. *Front. Immunol.* 12:809244. doi: 10.3389/fimmu.2021.809244
89. Luring AS, Hodcroft EB. Genetic Variants of SARS-CoV-2-What do They Mean? *JAMA* (2021) 325(6):529–31. doi: 10.1001/jama.2020.27124
90. Denison MR, Graham RL, Donaldson EF, Eckerle LD, Baric RS. Coronaviruses: An RNA Proofreading Machine Regulates Replication Fidelity and Diversity. *RNA Biol* (2011) 8(2):270–9. doi: 10.4161/rna.8.2.15013
91. Ogando NS, Zevenhoven-Dobbe JC, van der Meer Y, Bredenbeek PJ, Posthuma CC, Snijder EJ. The Enzymatic Activity of the Nsp14 Exoribonuclease Is Critical for Replication of MERS-CoV and SARS-CoV2. *J Virol* (2020) 94(23):e01246–20. doi: 10.1128/jvi.01246-20
92. Moeller NH, Shi K, Demir Ö, Banerjee S, Yin L, Belica C, et al. Structure and Dynamics of SARS-CoV-2 Proofreading Exoribonuclease ExoN. *bioRxiv* (2021) 2021.04.02.438274. doi: 10.1101/2021.04.02.438274

93. Karlsson EK, Kwiatkowski DP, Sabeti PC. Natural Selection and Infectious Disease in Human Populations. *Nat Rev Genet* (2014) 15(6):379–93. doi: 10.1038/nrg3734
94. Mascola JR, Graham BS, Fauci AS. SARS-CoV-2 Viral Variants-Tackling a Moving Target. *JAMA* (2021) 325(13):1261–2. doi: 10.1001/jama.2021.2088
95. Di Giorgio S, Martignano F, Torcia MG, Mattiuz G, Conticello SG. Evidence for Host-Dependent RNA Editing in the Transcriptome of SARS-CoV-2. *Sci Adv* (2020) 6(25):eabb5813. doi: 10.1126/sciadv.abb5813
96. Weber S, Ramirez CM, Weiser B, Burger H, Doerfler W. SARS-CoV-2 Worldwide Replication Drives Rapid Rise and Selection of Mutations Across the Viral Genome: A Time-Course Study Potential Challenge for Vaccines and Therapies. *EMBO Mol Med* (2021) 13:e14062. doi: 10.15252/emmm.202114062
97. Mourier T, Sadykov M, Carr MJ, Gonzalez G, Hall WW, Pain A. HostDirected Editing of the SARS-CoV-2 Genome. *Biochem Biophys Res Commun* (2021) 538:35–9. doi: 10.1016/j.bbrc.2020.10.092
98. Pathak AK, Fatihi S, Abbas T, Uppili B, Mishra GP, Ghosh A, et al. IntraHost Variability in Global SARS-CoV-2 Genomes as Signatures of RNA Editing: Implications in Viral and Host Response Outcomes. *bioRxiv* (2020) 2020.12.09.417519. doi: 10.1101/2020.12.09.417519
99. Francisco RDS Jr., Benites LF, Lamarca AP, de Almeida LGP, Hansen AW, Gularte JS, et al. Pervasive Transmission of E484K and Emergence of VUINP13L With Evidence of SARS-CoV-2 Co-Infection Events by Two Different Lineages in Rio Grande do Sul, Brazil. *Virus Res* (2021) 296:198345. doi: 10.1016/j.virusres.2021.198345
100. Haddad D, John SE, Mohammad A, Hammad MM, Hebbar P, Channanath A, et al. SARS-CoV-2: Possible Recombination and Emergence of Potentially More Virulent Strains. *PLoS One* (2021) 16(5):e0251368. doi: 10.1371/journal.pone.0251368
101. Choi B, Choudhary MC, Regan J, Sparks JA, Padera RF, Qiu X, et al. Persistence and Evolution of SARS-CoV-2 in an Immunocompromised Host. *N Engl J Med* (2020) 383(23):2291–3. doi: 10.1056/NEJMc2031364
102. Kemp SA, Collier DA, Datir RP, Ferreira IATM, Gayed S, Jahun A, et al. SARS-CoV-2 Evolution During Treatment of Chronic Infection. *Nature* (2021) 592(7853):277–82. doi: 10.1038/s41586-021-03291-y
103. Khatamzas E, Rehn A, Muenchhoff M, Hellmuth J, Gaitzsch E, Weiglein T, et al. Emergence of Multiple SARS-CoV-2 Mutations in an Immunocompromised Host. *medRxiv* (2021) 2021.01.10.20248871. doi: 10.1101/2021.01.10.20248871
104. Lynch M, Macori G, Fanning S, O'Regan E, Hunt E, O'Callaghan D, et al. Genomic Evolution of SARS-CoV-2 Virus in Immunocompromised Patient, Ireland. *Emerg Infect Dis* (2021) 27(9):2499–501. doi: 10.3201/eid2709.211159
105. Tarhini H, Recoing A, Bridier-Nahmias A, Rahi M, Lambert C, Martres P, et al. Long-Term Severe Acute Respiratory Syndrome Coronavirus 2 (SARSCoV-2) Infectiousness Among Three Immunocompromised Patients: From Prolonged Viral Shedding to SARS-CoV-2 Superinfection. *J Infect Dis* (2021) 223(9):1522–7. doi: 10.1093/infdis/jiab075
106. Atlanta GUSDoHaHS. Centers for Disease Control and Prevention. SARS-CoV-2 Variant Classifications and Definitions Centers for Disease Control and Prevention. Atlanta, Georgia, United States of America: Centers for Disease Control and Prevention (2021).
107. Volz E, Mishra S, Chand M, Barrett JC, Johnson R, Geidelberg L, et al. Assessing Transmissibility of SARS-CoV-2 Lineage B.1.1.7 in England. *Nature* (2021) 593:266–9. doi: 10.1038/s41586-021-03470-x
108. Rambaut A, Loman N, Pybus O, Barclay W, Barrett J, Carabelli A, et al. Preliminary Genomic Characterisation of an Emergent SARS-CoV-2 Lineage in the UK Defined by a Novel Set of Spike Mutations. (2020). Available at: <https://virological.org/t/preliminary-genomic-characterisation-of-anemergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spikemutations/563>.

109. Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J, et al. Detection of a SARS-CoV-2 Variant of Concern in South Africa. *Nature* (2020) 592:438–43. doi: 10.1038/s41586-021-03402-9
110. Fujino T, Nomoto H, Kutsuna S, Ujiie M, Suzuki T, Sato R, et al. Novel SARS-CoV-2 Variant in Travelers From Brazil to Japan. *Emerg Infect Dis* (2021) 27(4):1243–5. doi: 10.3201/eid2704.210138
111. Faria NR, Mellan TA, Whittaker C, Claro IM, Candido DDS, Mishra S, et al. Genomics and Epidemiology of the P.1 SARS-CoV-2 Lineage in Manaus, Brazil. *Science* (2021) 372(6544):815–21. doi: 10.1126/science.abh2644
112. Felipe N, Cristiano da C, Valdinete N, Victor S, André C, Fernanda N, et al. Three SARS-CoV-2 Reinfection Cases by the New Variant of Concern (VOC) P.1/501y.V3. *Res Square* (2021). doi: 10.21203/rs.3.rs-318392/v1
113. Ferreira I, Datir R, Papa G, Kemp S, Meng B, Rakshit P, et al. SARS-CoV-2 B.1.617 Emergence and Sensitivity to Vaccine-Elicited Antibodies. *bioRxiv* (2021) 2021.05.08.443253. doi: 10.1101/2021.05.08.443253
114. Cherian S, Potdar V, Jadhav S, Yadav P, Gupta N, Das M, et al. Convergent Evolution of SARS-CoV-2 Spike Mutations, L452R, E484Q and P681R, in the Second Wave of COVID-19 in Maharashtra, India. *Microorganisms* (2021) 9(7):1542. doi: 10.3390/microorganisms9071542
115. Brown CM, Vostok J, Johnson H, Burns M, Gharpure R, Sami S, et al. Outbreak of SARS-CoV-2 Infections, Including COVID-19 Vaccine Breakthrough Infections, Associated With Large Public Gatherings—Barnstable County, Massachusetts, July 2021. *MMWR Morb Mortal Wkly Rep* (2021) 70:1059–62. doi: 10.15585/mmwr.mm7031e2
116. WHO. Classification of Omicron (B.1.1.529): SARS-CoV-2 Variant of Concern. (2021). Available at: [https://www.who.int/news/item/26-11-2021-classification-of-omicron-\(b.1.1.529\)-sars-cov-2-variant-of-concern](https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern).
117. NCID. Frequently Asked Questions for the B.1.1.529 Mutated SARS-CoV-2 Lineage in South Africa. (2021). Available at: <https://www.nicd.ac.za/frequently-asked-questions-for-the-b-1-1-529-mutated-sars-cov-2-lineagein-south-africa/>.
118. Burki T. Understanding Variants of SARS-CoV-2. *Lancet* (2021) 397 (10273):462. doi: 10.1016/S0140-6736(21)00298-1
119. O’Toole Á, Scher E, Underwood A, Jackson B, Hill V, McCrone JT, et al. Pangolin: Lineage Assignment in an Emerging Pandemic as an Epidemiological Tool. (2021). Available at: <https://cov-lineages.org/>.
120. WHO. Tracking SARS-CoV-2 Variants (2021). Available at: <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>.
121. Raman R, Patel KJ, Ranjan K. COVID-19: Unmasking Emerging SARS-CoV-2 Variants, Vaccines and Therapeutic Strategies. *Biomolecules* (2021) 11(7):993. doi: 10.3390/biom11070993
122. GISAID. CoVariants Shared Mutations. (2021). Available at: <https://covariants.org/shared-mutations>.
123. NCID. SARS-CoV-2 Sequencing Update 26 November 2021. (2021). Available at: https://www.nicd.ac.za/wp-content/uploads/2021/11/Update-of-SAsequencing-data-from-GISAID-26-Nov_Final.pdf.
124. NCID. Frequently Asked Questions for the B.1.1.529 Mutated SARS-CoV-2 Lineage in South Africa. (2021). Available at: <https://www.nicd.ac.za/frequently-asked-questions-for-the-b-1-1-529-mutated-sars-cov-2-lineagein-south-africa/>.
125. Gu H, Chen Q, Yang G, He L, Fan H, Deng YQ, et al. Adaptation of SARSCoV-2 in BALB/c Mice for Testing Vaccine Efficacy. *Science* (2020) 369 (6511):1603–7. doi: 10.1126/science.abc4730
126. Starr TN, Greaney AJ, Hilton SK, Ellis D, Crawford KHD, Dingens AS, et al. Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* (2020) 182 (5):1295–310.e20. doi: 10.1016/j.cell.2020.08.012

127. Ali F, Kasry A, Amin M. The New SARS-CoV-2 Strain Shows a Stronger Binding Affinity to ACE2 Due to N501Y Mutant. *Med Drug Discov* (2021) 10:100086. doi: 10.1016/j.medidd.2021.100086
128. Davies NG, Abbott S, Barnard RC, Jarvis CI, Kucharski AJ, Munday JD, et al. Estimated Transmissibility and Impact of SARS-CoV-2 Lineage B.1.1.7 in England. *Science* (2021) 372(6538):eabg3055. doi: 10.1126/science.abg3055
129. Grint DJ, Wing K, Williamson E, McDonald HI, Bhaskaran K, Evans D, et al. Case Fatality Risk of the SARS-CoV-2 Variant of Concern B.1.1.7 in England, 16 November to 5 February. *Euro Surveill* (2021) 26(11):2100256. doi: 10.2807/1560-7917.Es.2021.26.11.2100256
130. Horby P, Huntley C, Davies N, Edmunds J, Ferguson N. G M. NERVTAG Paper on COVID-19 Variant of Concern B.1.1.7 (2021). Available at: <https://www.gov.uk/government/publications/nervtag-paper-on-covid-19-variantofconcern-b117>.
131. Hoffmann M, Kleine-Weber H, Pöhlmann S. A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. *Mol Cell* (2020) 78(4):779–84.e5. doi: 10.1016/j.molcel.2020.04.022
132. Peacock TP, Goldhill DH, Zhou J, Baillon L, Frise R, Swann OC, et al. The Furin Cleavage Site of SARS-CoV-2 Spike Protein Is a Key Determinant for Transmission Due to Enhanced Replication in Airway Cells. *Nat Microbiol* (2020) 6(7):899–909. doi: 10.1038/s41564-021-00908-w
133. Meng B, Kemp SA, Papa G, Datir R, Ferreira IATM, Marelli S, et al. Recurrent Emergence of SARS-CoV-2 Spike Deletion H69/V70 and Its Role in the Alpha Variant B.1.1.7. *Cell Rep* (2021) 35(13):109292. doi: 10.1016/j.celrep.2021.109292
134. Villoutreix BO, Calvez V, Marcelin AG, Khatib AM. In Silico Investigation of the New UK (B.1.1.7) and South African (501y.V2) SARS-CoV-2 Variants With a Focus at the ACE2-Spike RBD Interface. *Int J Mol Sci* (2021) 22 (4):1695. doi: 10.3390/ijms22041695
135. Abdool Karim SS, de Oliveira T. New SARS-CoV-2 Variants - Clinical, Public Health, and Vaccine Implications. *N Engl J Med* (2021) 384(19):1866– 8. doi: 10.1056/NEJMc2100362
136. Resende PC BJ, Vasconcelos RHT. Spike E484K Mutation in the First SARS-CoV-2 Reinfection Case Confirmed in Brazil, 2020. (2021). Available at: <https://virological.org/t/spike-e484k-mutation-in-the-first-sars-cov-2-reinfection-case-confirmed-in-brazil-2020/584>.
137. Rivers M. Why Are More Young People Getting Sick With Covid-19 in Brazil? Atlanta, Georgia, United States of America: CNN (2021). Available at: <https://edition.cnn.com/2021/03/24/americas/brazil-youth-covid-19-intl/latam/index.html>.
138. Peet C. “A Miracle I Survived”: Younger Brazilians Hit by COVID Surge. Doha, Qatar: Al Jazeera (2021). Available at: <https://www.aljazeera.com/news/2021/3/28/miracle-i-survived-younger-brazilians-hit-covid-surge>.
139. Tegally H, Wilkinson E, Althaus CL, Giovanetti M, San JE, Giandhari J, et al. Rapid Replacement of the Beta Variant by the Delta Variant in South Africa. *medRxiv* (2021) 2021.09.23.21264018. doi: 10.1101/2021.09.23.21264018
140. Campbell F, Archer B, Laurenson-Schafer H, Jinnai Y, Konings F, Batra N, et al. Increased Transmissibility and Global Spread of SARS-CoV-2 Variants of Concern as at June 2021. *Eurosurveillance* (2021) 26(24):2100509. doi: 10.2807/1560-7917.ES.2021.26.24.2100509
141. PHE. SARS-CoV-2 Variants of Concern and Variants Under Investigation in England (2021). Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1001358/Variants_of_Concern_VOC_Technical_Briefing_18.pdf.
142. Sheikh A, McMenemy J, Taylor B, Robertson C. SARS-CoV-2 Delta VOC in Scotland: Demographics, Risk of Hospital Admission, and Vaccine Effectiveness. *Lancet* (2021) 397(10293):2461–2. doi: 10.1016/S0140-6736(21)01358-1
143. Khedar RS, Mittal K, Ambaliya HC, Mathur A, Gupta JB, Sharma KK, et al. Greater Covid-19 Severity and Mortality in Hospitalized Patients in Second (Delta Variant) Wave Compared to the First:

- Single Centre Prospective Study in India. medRxiv (2021) 2021.09.03.21263091. doi: 10.1101/2021.09.03.21263091
144. Bio-IT World, Davies, K. [Powering Preventative Medicine Archived](#) 2016-06-06 at the [Wayback Machine](#). Bio-IT World 2011
 145. Rusk, N. (2011). ["Torrents of sequence"](#). Nat Meth 8(1): 44-44.
 146. Purushothaman, S, Toumazou, C, Ou, C-P [Protons and single nucleotide polymorphism detection: a simple use for the ion sensitive field effect transistor](#)
 147. [Pennisi, E](#) (2010). "Semiconductors inspire new sequencing technologies". *Science*. **327** (5970):1190. [Bibcode:2010Sci...327.1190P](#). [doi:10.1126/science.327.5970.1190](#). PMID [20203024](#)
 148. Perkel, J., ["Making contact with sequencing's fourth generation" Archived](#) 2013-12-27 at the [Wayback Machine](#). Biotechniques, 2011.
 149. Alberts B, [Molecular Biology of the Cell](#). 5th Edition ed. 2008, New York: Garland Science.
 150. [Ion Torrent Official Webpage Archived](#) 2012-11-06 at the [Wayback Machine](#).
 151. Karow, J. (2009) [Ion Torrent Patent App Suggests Sequencing Tech Using Chemical-Sensitive Field-Effect Transistors](#). In Sequence.
 152. Bio-IT World, Davies, K. [It's "Watson Meets Moore" as Ion Torrent Introduces Semiconductor Sequencing Archived](#) 2015-08-02 at the [Wayback Machine](#). Bio-IT World 2010.
 153. Eid, J., et al., ["Real-time DNA sequencing from single polymerase molecules"](#). *Science*, 2009. 323(5910): p. 133-8.
 154. Karow, J. (2010) [Ion Torrent Systems Presents \\$50,000 Electronic Sequencer at AGBT](#). In Sequence.
 155. Karow, J. (2009) [At AGBT, Ion Torrent Customers Provide First Feedback; Life Tech Outlines Platform's Growth](#). In Sequence.
 156. Metzker, M.L., ["Emerging technologies in DNA sequencing"](#). *Genome Res*, 2005.
 157. Anaclerio, F.; Ferrante, R.; Mandatori, D.; Antonucci, I.; Capanna, M.; Damiani, V.; Tomo, P.D.; Ferrante, R.; Ranaudo, M.; De Laurenzi, V.; et al. Different Strategies for the Identification of SARS-CoV-2 Variants in the Laboratory Practice. *Genes* 2021, 12, 1428. <https://doi.org/10.3390/genes12091428>
 158. [Classification of Omicron \(B.1.1.529\): SARS-CoV-2 Variant of Concern](#), su [who.int](#).
 159. Chiranjib Chakraborty, Manojit Bhattacharya, Ashish Ranjan Sharma, Bidyut Mallik. *Omicron (B.1.1.529) - A new heavily mutated variant: Mapped location and probable properties of its mutations with an emphasis on S-glycoprotein*. International Journal of Biological Macromolecules 2022. <https://doi.org/10.1016/j.ijbiomac.2022.07.254>
 160. Capriotti E., Fariselli P., Rossi I., Casadio R. A three-state prediction of single point mutations on protein stability changes. *BMC Bioinformatics*. 2008;9(2):1–9.
 161. Korber B., Fischer W.M., Gnanakaran S., Yoon H., Theiler J., Abfalterer W., et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell*. 2020;182(4):812–827.
 162. Chakraborty C., Saha A., Sharma A.R., Bhattacharya M., Lee S.S., Agoramoorthy G. D614G mutation eventuates in all VOI and VOCin SARS-CoV-2: is it part of the positive selection pioneered by Darwin? *Mol. Ther.-Nucleic Acids*. 2021;26:237–241.
 163. Liu S., Huynh T., Stauff C.B., Wang T.T., Luan B. Structure-function analysis of resistance to bamlanivimab by SARS-CoV-2 variants kappa, delta, and lambda. *J. Chem. Inf. Model*. 2021;61(10):5133–5140.
 164. Benvenuto D., Benedetti F., Demir A.B., Ciccozzi M., Zella D. Analysis of three mutations in italian strains of SARS-CoV-2: implications for pathogenesis. *Chemotherapy*. 2021;66(1–2):33–37.

165. Yang W.T., Huang W.H., Liao T.L., Hsiao T.H., Chuang H.N., Liu P.Y. SARS-CoV-2 E484K mutation narrative review: epidemiology, immune escape, clinical implications, and future considerations. *Infect. Drug Resist.* 2022;15:373.
166. Thakur S., Sasi S., Pillai S.G., Nag A., Shukla D., Singhal R., Phalke S., Velu G.S.K. SARS-CoV-2 mutations and their impact on diagnostics, therapeutics and vaccines. *Front. Med.* 2022;9
167. Chakraborty C., Bhattacharya M., Sharma A.R. Present variants of concern and variants of interest of severe acute respiratory syndrome coronavirus 2: their significant mutations in S-glycoprotein, infectivity, re-infectivity, immune escape and vaccines activity. *Rev. Med. Virol.* 2021;32:1–14. e2270.
168. Chakraborty C., Sharma A.R., Bhattacharya M., Agoramoorthy G., Lee S.-S. Evolution, mode of transmission, and mutational landscape of newly emerging SARS-CoV-2 variants. *MBio.* 2021;12(4)
169. Ali F., Kasry A., Amin M. The new SARS-CoV-2 strain shows a stronger binding affinity to ACE2 due to N501Y mutant. *Med. Drug Discov.* 2021;10:100086.
170. Laffeber C., de Koning K., Kanaar R., Lebbink J.H. Experimental evidence for enhanced receptor binding by rapidly spreading SARS-CoV-2 variants. *J. Mol. Biol.* 2021;433(15)
171. Woo H.G., Shah M. bioRxiv; 2021. Omicron: A Heavily Mutated SARS-CoV-2 Variant Exhibits Stronger Binding to ACE2 and Potently Escape Approved COVID-19 Therapeutic Antibodies.
172. Huang K., Zhang Y., Hui X., Zhao Y., Gong W., Wang T., et al. Q493K and Q498H substitutions in spike promote adaptation of SARS-CoV-2 in mice. *EBioMedicine.* 2021;67
173. Fratev F. N501Y and K417N mutations in the spike protein of SARS-CoV-2 Alter the interactions with both hACE2 and human-derived antibody: a free energy of perturbation retrospective study. *J. Chem. Inf. Model.* 2021;61(12):6079–6084.
174. Souza P.F., Mesquita F.P., Amaral J.L., Landim P.G., Lima K.R., Costa M.B., et al. The spike glycoproteins of SARS-CoV-2: a review of how mutations of spike glycoproteins have driven the emergence of variants with high transmissibility and immune escape. *Int. J. Biol. Macromol.* 2022;208(2022):105–125.
175. Singh A., Steinkellner G., Köchl K., Gruber K., Gruber C.C. Serine 477 plays a crucial role in the interaction of the SARS-CoV-2 spike protein with the human receptor ACE2. *Sci. Reports.* 2021;11(1):1–11.
176. Kumar S., Thambiraja T.S., Karuppanan K., Subramaniam G. Omicron and Delta variant of SARS-CoV-2: a comparative computational study of spike protein. *J. Med. Virol.* 2022;94(4):1641–1649.
177. Kannan S.R., Spratt A.N., Sharma K., Chand H.S., Byrareddy S.N., Singh K. Omicron SARS-CoV-2 variant: unique features and their impact on pre-existing antibodies. *J. Autoimmun.* 2022;126
178. Wang L., Cheng G. Sequence analysis of the emerging SARS-CoV-2 variant Omicron in South Africa. *J. Med. Virology.* 2022;94(4):1728–1733.

Louten J., *Essential Human Virology*, Academic Press, 2017