

PERCHÉ IL COVID-19 COLPISCE IN MODI DIVERSI?*Alessandra Renieri***Dipartimento di Biotecnologie Mediche, Università degli Studi di Siena
24 febbraio 2021**

Sono veramente contenta di essere qui, di parlare, una volta tanto, alla popolazione generale e non solo in un convegno scientifico. Sono contenta di poter presentare i dati che abbiamo acquisito negli ultimi 12 mesi lavorando tantissimo non solo con il mio gruppo, ma anche con una collaborazione nell'ambito dell'Università di Siena, dell'Azienda Ospedaliera Universitaria Senese, con tantissimi professionisti dell'ASL Sud Est e con più di 40 ospedali che collaborano a questo consorzio GEN-COVID che coordino, nonché professori e studenti di dottorato, essenziali nell'ambito della intelligenza artificiale. Una collaborazione con Cineca ci ha permesso anche di poter depositare, in maniera congrua, i grandi dati che vengono generati dal sequenziamento del genoma umano.

Iniziamo da un fatto che è sotto gli occhi di tutti, ovvero che la maggioranza dei soggetti infetti da COVID-19 sono in realtà asintomatici o paucisintomatici, mentre una piccolissima percentuale che si può stimare intorno all'1,8%, quindi una piccola percentuale dell'insieme ma rilevante, in quanto stiamo parlando di una pandemia, avrà un COVID grave. Quello che abbiamo fatto è di fenotipizzare oltre 2000 soggetti infetti. Fenotipizzare vuol dire raccogliere nella maniera più accurata le caratteristiche cliniche, anche perché, ormai si sa che il COVID è una malattia sistemica e quindi colpisce non solo il polmone ma anche tantissimi altri organi e apparati, a partire dall'apparato cardiaco e tantissimi altri.

I pazienti ospedalizzati sono stati classificati secondo una gravità di trattamento, quelli intubati, quelli con ossigeno sotto pressione, supporto del solo ossigeno senza pressione e poi soggetti non ospedalizzati che sono paucisintomatici o asintomatici. Un questionario, con più di 160 domande specifiche, è stato messo a punto da più di 20 medici appartenenti a 7 differenti specialità, perché un po' tutte le specialità mediche concorrono a cercare di capire questa malattia così complessa, dall'ospedale pilota che ovviamente è stata l'Azienda Ospedaliera Universitaria Senese, all'Asl Sud Est e, poi, come vi ho detto, più di 40 ospedali italiani. Una volta raccolto il materiale biologico, è stato sequenziato il DNA, che è stato poi indicato con sigle diverse a seconda della tipologia di sequenziamento effettuata. Con la sigla HLA si indica il sequenziamento sul sistema HLA che è un sistema estremamente importante nella risposta immunitaria. Con la sigla WES si indica l'analisi dell'Esoma completo cioè di tutti i geni del nostro patrimonio genetico che sono oltre 20.000 e con la sigla GWAS si indica l'analisi di marcatori sparsi per tutto il genoma anche nelle regioni dove non sono presenti i geni.

L'arruolamento non vi nascondo che è stato un percorso abbastanza complesso e laborioso, con approvazione successiva in sequenza, di tutti gli ospedali, da parte dei comitati etici, raccolta della scheda clinica, raccolta dei prelievi che venivano fatti e vengono tuttora fatti a gruppi di pazienti. Quindi c'è un lavoro preparatorio estremamente importan-

te.

I Campioni sono stati raccolti e trattati a partire dai primi mesi del 2020 fino ad ora, Febbraio del 2021. Parte del sangue è stato trasformato in DNA che è stabile indefinitivamente. Plasma e il siero sono stati raccolti per testare i marcatori che si possono trovare in questi fluidi biologici e anche le cellule sono state conservate in modo da poterle rimettere in coltura. L'arruolamento è iniziato negli ospedali, ma poi sono stati anche fondamentali l'arruolamento nelle USCA e nei dipartimenti di prevenzione. Di estrema rilevanza è stata l'acquisizione dell'apparecchio di Sequenziamento NovaSeq6000, inaugurato il 23 di gennaio del 2020, un'operazione importante non solo per il finanziamento da parte del Ministero dell'Università e della Ricerca ma anche per aver messo Siena in una condizione all'avanguardia per quanto riguarda la parte tecnologica, non solo sul territorio italiano, ma anche a livello mondiale. Con questo apparecchio, in 12-24 ore, si può sequenziare l'intero set di geni di ogni individuo, di un gruppo di oltre 200 pazienti. E' quindi un apparecchio altamente performante, non solo per quanto riguarda la velocità e la capacità, ma anche per quanto riguarda la risoluzione. Tanto che i dati che abbiamo acquisito potranno essere anche successivamente rianalizzati non solo per le mutazioni, chiamate costitutive, presenti cioè in tutte le cellule del soggetto, ma anche per eventuali mutazioni somatiche, cioè presenti solamente in parte delle cellule degli individui, che molto spesso concorrono alla diversità di gravità di malattia.

L'ETÀ È IL MAGGIORE PREDITTORE DI GRAVITÀ

E' sotto gli occhi di tutti, che l'età è il maggiore predittore di gravità. E quindi è ben noto che un soggetto di vent'anni avrà una malattia asintomatica o paucisintomatica e via via che si passa ai 40 ai 60 e agli 80, probabilmente ci sarà bisogno di un supporto di ossigeno o anche di ossigeno sotto pressione. Quindi analizzando le caratteristiche cliniche e di trattamento di questi 2000 pazienti abbiamo delineato quello che è l'andamento più frequente della gravità di malattia rispetto all'età. Analizzando sempre la solita casistica di pazienti, ci siamo però anche resi conto che, al di là di quello che era l'andamento atteso, nella maggioranza dei soggetti esisteva un andamento o meno grave o più grave dell'atteso, fino ad arrivare alla categoria quinta dell'OMS, che sono i deceduti, quali ovviamente hanno un andamento estremamente diverso dall'atteso, estremamente grave. Altri soggetti, quasi centenari, nonostante la loro età, se la sono cavata con un piccolo supplemento di ossigeno senza pressione; quindi indubbiamente questi sono dei soggetti che hanno un andamento diverso dall'atteso in base alla propria genetica, al proprio patrimonio genetico. In particolare, se si confrontano i maschi con le femmine, si vede che l'andamento atteso nei maschi è più grave che nelle femmine, perciò quando si va a trattare un individuo o comunque a ipotizzare, predire, la gravità di una delle malattie di un

individuo, oltre all'età, deve essere preso in considerazione il sesso, prima ancora di parlare di genetica in senso più stretto, quindi della presenza delle sequenze caratteristiche nei 20.000 geni del nostro patrimonio genetico.

UNA PARTE DEI GRAVI HA MALATTIA MENDELIANA

Considerando tutta la popolazione infetta, il COVID grave è presente in un sottoinsieme di questi e, un sottoinsieme di questi soggetti e con COVID grave, presenta una malattia mendeliana. Il nome mendeliano deriva da Mendel che alla fine dell'Ottocento, in particolare 155 anni fa, ha postulato determinate leggi di trasmissione che poi, sono state successivamente interpretate e abbiamo acquisito che uno solo dei 20.000 geni del nostro genoma, se mutato, può dare luogo a una malattia che viene definita mendeliana o monogenica. Quindi, come in tante altre malattie complesse, dal Parkinson, dalle malattie cardiovascolari alle malattie neurologiche degenerative, esiste un sottoinsieme di pazienti che risponde al modello di malattia mendeliana. Sono i cosiddetti malati rari. Anche nel caso del COVID-19, esistono dei malati gravi rari, ovvero dovuti alla mutazione in un solo gene. In questo caso il modello mendeliano combina la mutazione in un gene con l'infezione perché esiste ovviamente il fattore ambientale, che è la presenza del virus, cioè l'impatto del virus sulle mucose, sulle cellule del soggetto. Quindi, come, attraverso gli insegnamenti di Mendel, abbiamo imparato a seguire la segregazione di malattie come l'emofilia, che è una delle malattie che tutti conosciamo come malattia mendeliana, così abbiamo dimostrato che anche per il COVID esiste una modalità di trasmissione di questo genere. L'emofilia ha la caratteristica di essere presente solo nei soggetti di sesso maschile, perché il gene sta sul cromosoma X mentre le femmine sono solamente delle portatrici asintomatiche. Quindi, come nel caso dell'emofilia, anche per il COVID-19 esistono delle famiglie che abbiamo studiato, dove solamente il maschio è gravemente affetto, i soggetti che non hanno la mutazione, anche se vengono a contatto con il virus, non sono gravi, sono praticamente asintomatici, mentre chi ha preso la mutazione è grave, per cui viene trattato anche in età giovanile con il casco a pressione continua o alternante d'ossigeno. Esiste quindi un modello che è molto simile a quello dell'emofilia e che quindi possiamo chiamare eredità di tipo mendeliano nel COVID-19, che ci permette di predire con una precisione molto accurata, la gravità dei familiari, laddove tutti divengono infetti.

DIFETTI NEL GENE *TLR7* CAUSANO RIDOTTA RISPOSTA IMMUNITARIA E COVID GRAVE NEI MASCHI

Uno dei geni che, per esempio, segue questo tipo di andamento, si chiama *TLR7*. Abbiamo appena ricevuto la notizia dell'accettazione di un lavoro, che è il decimo che abbiamo pubblicato in questi 12 mesi, grazie appunto al lavoro collaborativo e a questo avanzamento di conoscenze e risultati. Quindi mutazioni singole in questo gene causano quello che possiamo definire COVID-19 di tipo Mendeliano. Il *TLR7* che cos'è? È un sensore che interviene nel riconoscimento della presenza di RNA a singolo filamento, come è appunto il genoma virale del coronavirus. È ovvio che difetti importanti in questo gene impediscono alla cellula di riconoscere che c'è la presenza di un RNA estraneo e quindi impediscono poi di attivare tutto il sistema a valle di attivazione dell'immunità innata degli Interferoni che è appunto la prima barriera contro l'infezione. E quindi è normale che in questi casi un gene singolo faccia tutto questo danno, e provochi una malattia complessa come quella di COVID-19.

DIFETTI NEL GENE *ACE-2* CAUSANO RIDOTTA ENTRATA DEL VIRUS E MALATTIA LIEVE

Abbiamo identificato anche l'effetto opposto: a volte ci sono dei geni importanti come *ACE-2* che, nel caso del COVID-19, è diventato un attore importante in quanto è il recettore del virus sulle cellule. Ci sono anche dei difetti che diventano un vantaggio, come sapevamo per l'HIV, in cui, un difetto nel recettore, rendeva le persone immuni da AIDS. Anche nel COVID-19, esistono dei rarissimi soggetti sia maschi che femmine che hanno un andamento di malattia molto, molto lieve perché hanno un difetto in *ACE-2* e quindi è difettiva anche la funzione recettoriale rispetto al virus. Ci sono stati dei novantenni che non hanno avuto bisogno di nessun supporto respiratorio e se la sono cavata molto bene con l'infezione virale. Un'altra femmina di 92 anni e un maschio di 80 hanno avuto bisogno solo di un supporto di ossigeno senza pressione. Quindi è un po' la genetica mendeliana, con meccanismo patogenetico opposto, cioè di protezione, che rende la malattia più mite rispetto a renderla più severa come era l'esempio precedente.

A parte, quindi, questi rari casi di eredità mendeliana cosa succede a tutto il resto dei soggetti COVID-19 gravi, che non possono essere spiegati con la mutazione in non solo gene?

EREDITÀ POST-MENDELIANA

A questo punto, dobbiamo parlare di eredità post-mendeliana. Ci dobbiamo evolvere da Mendel con l'aiuto prezioso dei colleghi che sono esperti di intelligenza artificiale, con i quali abbiamo collaborato per l'analisi dei dati generati dal sequenziatore. Ci siamo dovuti evolvere, nel senso che abbiamo dovuto imparare a leggere la trasmissibilità di un carattere non seguendo un gene solo, ma seguendo più di un gene. Due fratelli possono ricevere in maniera casuale dai genitori una quantità maggiore di geni che inducono alla gravità o una quantità inferiore e quindi fare la differenza tra loro. I due fratelli possono avere una estrinsecazione clinica di COVID-19 più grave, se hanno ricevuto più geni della gravità o più lieve, se hanno ricevuto meno geni della gravità. Quindi, se entrambi questi fratelli si infettano, l'uno richiederà di essere intubato, l'altro sarà un paucisintomatico o asintomatico. Abbiamo dunque cominciato a vedere quali sono questi geni. In particolare, il recettore degli Androgeni perché il Testosterone che aumenta con la pubertà nei maschi, regola tutti i geni come il *TLR7* che avete visto prima, o il *TLR3*. Regola quindi quei sensori di RNA virale, abbassandone l'espressione. Questo è uno dei motivi per cui i maschi sono più gravemente affetti, perché innalzando il Testosterone, si abbassa tutta la immunità innata. Sono infatti note le proprietà immunomodulatrici del Testosterone. Quindi, in questo caso l'effetto è un effetto non troppo vantaggioso. Peraltro, sappiamo che c'è una differenza di genere, non solo nel COVID-19, ma in generale, per esempio, nelle cause di mortalità che sono completamente diverse in frequenza tra i maschi e le femmine. È noto che i maschi muoiono di più per malattie infettive, mentre è noto che le femmine muoiono di più per malattie autoimmuni o autoinfiammatorie. Quindi c'è una suscettibilità diversa a seconda del genere e questa diversa suscettibilità è dovuta al fatto che dopo la pubertà aumentano gli androgeni nei maschi e gli estrogeni nelle femmine, che sono potenti modulatori di un numero incredibile di geni, alcuni abbassandoli, altri alzandoli. È importante dunque valutare il recettore degli androgeni: ci sono dei soggetti in cui il recettore degli androgeni funziona di più, altri in cui funziona di meno, e quelli in cui funziona di meno sono quelli in cui questo fattore collabora alla

gravità della malattia. Altri *TLR*, che come abbiamo detto stimolano gli Interferoni, altri, come l'*ApoBeck1* è un gene importantissimo perché questo è il sensore che individua l'RNA virale, anzi, non solo l'individua ma lo taglia o meglio, lo modifica, cioè fa "editing", un termine tecnico che significa un cambiamento delle basi del DNA; la T si trasforma in U, cambiando di fatto il codice genetico, modificando così le proteine prodotte dal virus che rimane difettivo e non si può incapsidare e quindi espandere nella cellula ospite. Quindi ovvio che si parla di polimorfismi funzionali, quindi di difetti parziali, non di difetti importanti come quelli delle malattie mendeliane, ma di difetti parziali in cui il gene un po' funziona, un po' no. E tra l'altro sono polimorfismi, la parola polimorfismo vuol dire che sono presenti in un numero molto grande di persone, non sono come appunto i geni mendeliani, che sono rari; stiamo parlando di variazione sul genoma che è una variazione comune e la combinazione di variazioni comuni è quella che poi determina la pendenza dell'ago della bilancia da una parte o dall'altra.

IGENI DI GRAVITÀ E DI "MITIGHEZZA" SONO DIVERSI NEI MASCHI E NELLE FEMMINE

Un dato molto interessante di cui siamo rimasti anche noi sconcertati appena l'abbiamo visto, ci aspettavamo infatti che il COVID-19 facesse parte della medicina di genere, ma non a questo livello. Il nostro dato indica che i geni di gravità o mitighezza sono completamente diversi nei maschi e nelle femmine. Che esperimento abbiamo fatto? Abbiamo testato questi geni in una coorte mista all'interno di tutta la nostra coorte, costituita sia da soggetti di sesso maschile che da soggetti di sesso femminile. Abbiamo estratto con delle tecniche di intelligenza artificiale i geni di interesse dai 20.000 geni ed abbiamo dimostrato sia la presenza dei geni di gravità che di quelli di mitighezza, una parola non troppo brillante, ma che significa che sono più protettivi rispetto alla gravità. Se poi andiamo a fare lo stesso esperimento separando la coorte dei maschi e delle femmine, ci ritroviamo gli stessi geni, ma il singolo gene non è mai in entrambi i grafici, quindi vuol dire che i geni che sono rilevanti nei maschi sono completamente diversi dai geni che sono rilevanti nelle femmine, quindi da questo esperimento in poi abbiamo deciso di trattare sempre la coorte in maniera separata, i maschi da sé, le femmine da sé.

FORMULA DEL MODELLO POST-MENDELIANO

Un nostro lavoro non ancora uscito per la pubblicazione ma che, grazie ad una opportunità offerta da Med-Archive, è in preprint, ci illustra un modello nuovo che abbiamo definito modello post-mendeliano attraverso una formula che produce di fatto uno score, quindi un punteggio, un valore che abbiamo chiamato Score Poligenico Integrato. La formula di fatto consiste nel contare i polimorfismi funzionali della mitighezza, sottrarci il numero di polimorfismi funzionali della gravità, per quanto riguarda i polimorfismi comuni, e poi sommare la stessa storia per quanto riguarda le varianti rare, cioè il numero di varianti rare della "mitighezza" meno il numero di varianti rare della gravità, in più è stato introdotto un fattore F moltiplicativo per aumentare il peso delle varianti rare perché si sa che in genetica una variante rara che colpisce un gene, quindi la proteina, avrà un impatto molto più importante sulla funzione della proteina stessa rispetto appunto al polimorfismo funzionale, cioè quelli che prima ho chiamato difetti parziali. Per gli appassionati di statistica e intelligenza artificiale, possiamo dire che tutto quello che abbiamo fatto ha una precisione di circa il 70%. Il modello è ovviamente ancora incompleto ma già ci dà dei

buoni risultati, potendo predire al 70% la gravità, non tenendo conto di fattori importanti quali l'età e il sesso, quindi, al di là dell'età e del sesso, basandoci solo sulla genetica. Perché è importante fare tutto ciò? È importante anche perché in questo modo si va a capire meglio i meccanismi fisiopatologici di malattia.

MEDICINA DI GENERE E PERSONALIZZATA IN COVID-19

Avrete tutti seguito la storia dell'idrossiclorochina che all'inizio veniva consigliata, poi non veniva più somministrata. Ci sono stati tanti trial, alcuni sono ancora in corso, poi l'EMA ha preso posizione che non c'erano evidenze di efficacia di questo farmaco. Come mai, altalene avanti indietro, nell'efficacia? Perché si mettono tutti i pazienti insieme, non solo si mettono i maschi insieme alle femmine, ma si mettono anche i soggetti maschi insieme ad altri maschi che hanno una genetica diversa. Nel caso di un polimorfismo nel gene che diminuisce l'autofagia, essendo l'idrossiclorochina un farmaco che inibisce l'autofagia, dare questo farmaco a questi soggetti non solo non è efficace ma può essere anche dannoso. Quindi diciamo in generale, a parte questo esempio, che tutto quello che noi troviamo ha una ricaduta su quello che può essere la terapia. È abbastanza assurdo che, dopo tutti questi mesi, l'unico farmaco attualmente ritenuto efficace sia un farmaco antichissimo che è appunto il cortisone. Questo è un po' l'esempio della sconfitta della medicina, soprattutto della parte terapeutica e quindi noi riteniamo che tenere in conto di più dei meccanismi fisiopatologici a partire da quella che è la genetica dell'ospite, che in questa malattia, più di altre malattie complesse, ha una parte estremamente importante, ci possa aiutare a fare una medicina migliore, per terapie migliori, magari in affiancamento al cortisone nell'ambito del concetto di quella che è appunto la medicina personalizzata di cui si fa tanto parlare ma che quasi mai viene utilizzata, rimanendo un concetto un pochino astratto.

Con questo ho concluso ma mi piace ribadire il lavoro del mio preziosissimo gruppo, un gruppo misto di ricerca, di assistenza. D'altra parte l'interazione della ricerca con l'assistenza è veramente l'anima non solo della nostra disciplina, la genetica medica, ma in generale di ogni avanzamento scientifico, a partire anche da cose banali. Durante il primo lockdown, tutti i grandi istituti di ricerca europei e americani erano completamente fermi, noi, avendo fatto la scelta di posizionare il sequenziatore in ospedale, mentre eravamo aperti e abbiamo potuto lavorare e, a tuttora, abbiamo i dati che mettiamo liberamente a disposizione della comunità scientifica. La casistica più ampia a livello mondiale su questo argomento.

Con questo vi ringrazio.

Genetica Medica Siena

