



## DIPARTIMENTO DI SCIENZE MEDICHE E CHIRURGICHE

<b>TITOLO DEL PROGETTO</b>			
Analisi dei dati genomici (WGS) di coorti di pazienti COVID-19 (Progetto ORCHESTRA)			
ASSEGNO FINANZIATO DA PROGETTO COMPETITIVO <i>(barrare la casella corrispondente)</i>	<input checked="" type="checkbox"/> SI	<input type="checkbox"/> NO	<i>Punti</i>
SE IL FINANZIAMENTO È COMPETITIVO L'ENTE FINANZIATORE	EU H2020		
PROGETTO/ATTIVITÀ A SCOPO COMMERCIALE <i>(es. sperimentazione profit)</i>	<input type="checkbox"/> SI	<input checked="" type="checkbox"/> NO	
CARATTERISTICHE DEL PROGETTO <i>(biomedico/osservazionale/clinico-interventistico/multidisciplinare)</i>	Biomedico		
STATO DI APPROVAZIONE DEL PROGETTO DA PARTE DEL COMITATO ETICO <i>(se necessario per il tipo di studio barrare o evidenziare la casella corrispondente)</i>	<input checked="" type="checkbox"/> Ottenuto	<input type="checkbox"/> Da ottenere	
<b>DESCRIZIONE DEL PROGETTO</b> <i>(max 800 parole)</i>			<i>Punti</i>
<b>Stato dell'Arte e Razionale</b>  Grazie ai progressi di automatizzazione dei protocolli di sequenziamento di nuova generazione ( <i>Next Generation Sequencing</i> , NGS) è ormai possibile processare in parallelo una quantità crescente di genomi con tempi e costi che negli anni vanno a diminuire progressivamente. Tramite questo approccio è possibile raccogliere notevoli quantità di dati per individuo, la cui analisi è oggi possibile grazie al continuo sviluppo di nuovi software bioinformatici che permettono di ricavare diverse informazioni a livello genotipico. In particolare, esistono svariati software utili per l'identificazione di varianti rare o comuni (varianti a singolo nucleotide, piccole inserzioni e delezioni, varianti strutturali) sia germinali che somatiche. Le varianti identificate possono a loro volta essere utilizzate in studi statistici di popolazione per individuare nuovi loci di suscettibilità o associati a una malattia. I principali approcci genomici di NGS utilizzati, <i>Whole Genome Sequencing</i> (WGS) e <i>Whole Exome Sequencing</i> (WES) permettono rispettivamente di identificare varianti a livello dell'intera sequenza genomica o soltanto a livello di regioni codificanti. Rispetto al WES, l'analisi WGS permette un'indagine più completa delle varianti di un individuo, soprattutto nel caso delle varianti strutturali, in cui notoriamente i breakpoints ricadono con maggiore frequenza in regioni non coding. Malgrado il grande potenziale nell'utilizzo del WGS, la grande mole di dati generati rende questo approccio di difficile utilizzo in ambienti non provvisti di una adeguata infrastruttura computazionale. Tuttavia, nuove alternative vengono date da aziende che forniscono servizi, che permettono l'analisi automatizzata e delocalizzata di questi dati in poche ore. La possibilità di utilizzare i dati NGS per svariate tipologie di analisi rende questo approccio di primo livello anche per lo studio di nuove malattie, come nel caso della malattia da COVID-19, dove l'eziopatogenesi e la variabilità nella gravità dei sintomi osservati nella malattia non sono ancora del tutto chiari. E' questo l'obiettivo che l'Università di Bologna persegue come partecipante al progetto europeo "Orchestra", un progetto di ricerca internazionale volto ad affrontare la pandemia di coronavirus e le sue conseguenze in ambito socio-sanitario e che coinvolge 26 partner di 15 paesi. Al progetto afferiscono 10 Working Packages (WP), ognuno con i propri tasks: la presente attività afferisce al WP6 che racchiude gli studi molecolari inclusa la genomica umana. Recenti studi hanno evidenziato una possibile associazione tra varianti genetiche, sia rare che comuni, in geni associati al pathway dell'immunità mediata dall'Interferone di tipo 1 e COVID-19 in forma grave. La pandemia da COVID-19 ha portato			



## DIPARTIMENTO DI SCIENZE MEDICHE E CHIRURGICHE

rapidamente alla necessità di identificare i fattori genetici che predispongono alla diversa manifestazione clinica della malattia, al fine di individuare i pazienti a rischio di una forma grave e stratificare meglio le procedure cliniche da utilizzare. A questo fine, lo studio del genoma di pazienti COVID-19 risulta quindi di grande interesse per poter identificare, tramite un'unica analisi, tutte le possibili tipologie di alterazioni genetiche, sia attualmente note che non.

### Obiettivi

Ci si propone di analizzare dati WGS/WES al fine di:

- 1) Identificare la variazione genomica umana, incluse varianti di singolo nucleotide, indel e copy number variants oltre a varianti strutturali e alterazione da espansione, nei pazienti reclutati con infezione da SARS-CoV2 e gradi diversi di gravità clinica di malattia da COVID-19
- 2) Condurre analisi di qualità dei dati e predisposizione per l'armonizzazione dei dati al fine di successive analisi (in particolari per varianti di singolo nucleotide)
- 3) Condurre un'analisi di variant enrichment (case-only o case-control) per l'individuazione di geni di predisposizione a malattia in pazienti COVID-19, di interpretazione funzionale delle varianti identificate su geni candidati, e di variant class burden per identificare classi di varianti di particolare interesse

### Metodologia (descrizione del campione, principali tecniche utilizzate, aspetti biostatistici, fattibilità...)

Presso il laboratorio di Genetica Medica del Dipartimento di Scienze Mediche e Chirurgiche è in corso la raccolta e il sequenziamento di 900 campioni derivanti da sangue periferico di pazienti con infezione da SARS-CoV2 e diversa gravità di sintomi (da molto gravi ad asintomatici). I dati ottenuti verranno analizzati tramite diverse pipeline bioinformatiche già in uso o che verranno appositamente sviluppate al fine di identificare varianti e/o loci di possibile interesse per la malattia con dedicate analisi statistiche.

### Risultati attesi

Tramite l'utilizzo di un'ampia coorte di dati genomici derivanti da pazienti COVID-19, ci si propone di identificare loci associati alla malattia con significatività statistica che possano spiegare l'eziopatogenesi e la variabilità clinica della malattia da COVID-19.

## DESCRIZIONE DELLE ATTIVITÀ DELL'ASSEGNISTA

(per i **nuovi** assegni: max 400 parole; competenze richieste, scansione temporale della formazione, scansione temporale dell'attività, obiettivi primari e secondari)

(per i **rinnovi**: max 600 parole – da integrare con la relazione dell'assegnista; formazione raggiunta, attività effettuata, obiettivi raggiunti/competenze acquisite, formazione ancora da acquisire (se pertinente), scansione temporale dell'attività durante il rinnovo)

Punti

Sarà richiesta al candidato/a una competenza in analisi NGS con conoscenze dei principali programmi di analisi genomica (allineamento, variant calling e data interpreting).

Lo schema temporale delle attività sarà suddiviso come segue:

Mesi 0-6.

Fase 1: Analisi WES/WGS: allineamento delle reads e creazione dei multi-sample VCF per varianti a singolo nucleotide e piccole inserzioni e delezioni e per le varianti strutturali/espansioni.

Mesi 6-12



DIPARTIMENTO DI SCIENZE MEDICHE E CHIRURGICHE

Fase 2: Ricerca e interpretazione di varianti rare/comuni di interesse e analisi statistica (Enrichment) dei dati ottenuti.

Mesi 12-18

Fase 3: Interpretazione di varianti strutturali e analisi delle varianti su DNA mitocondriale.

**SE RINNOVO, SI RICORDA DI ALLEGARE ANCHE LA RELAZIONE DELL'ASSEGNISTA CON LA SUA PRODUZIONE SCIENTIFICA.**

*Scheda attività assistenziale (se prevista)*

<b>ATTIVITÀ ASSISTENZIALI DELL'ASSEGNISTA/ N. ORE SETTIMANA</b>
<b>AZIENDA SANITARIA PRESSO CUI SI SVOLGERÀ L'ATTIVITÀ</b>

Si ricorda che, come previsto dagli Accordi sull'impiego nell'attività assistenziale dei Titolari di assegni di ricerca, sottoscritti tra l'Università di Bologna e le Aziende Ospedaliere di riferimento, una volta stipulato il contratto con il vincitore della selezione, il tutor deve consegnare alla Direzione Medica Ospedaliera la relativa modulistica, nella quale andranno riportate le attività qui segnalate.