

## ALLEGATO 1) PIANO DI ATTIVITA'

**TITOLO DEL PROGETTO: Sviluppo di protocolli di sequenziamento massivo di III generazione Oxford Nanopore Technologies per lo studio di genomi virali in ambito di medicina di precisione**

**Docente tutor: Fabio Gentilini**

**Durata: 12 mesi**

### **DESCRIZIONE DEL PROGETTO**

#### **Sviluppo di protocolli di sequenziamento massivo di III generazione Oxford Nanopore Technologies per lo studio di genomi virali in ambito di medicina di precisione**

La medicina di precisione è una disciplina che si sta affermando di pari passo con gli sviluppi tecnologici in particolar modo delle tecniche di sequenziamento di II e III generazione e di analisi massiva di dati (cd Big data) come strumento di diagnostica avanzata in settori differenti: oncologia, genetica, microbiologia. L'obiettivo è consentire diagnosi precise nonché stratificare i pazienti sulla base di caratteristiche peculiari del loro genoma o del genoma degli agenti patogeni che li affliggono in modo da prevedere l'evoluzione della patologia e/o identificare strategie terapeutiche più efficaci. La medicina di precisione si connota, quindi, come una disciplina trasversale a competenze più tradizionali, che si avvale di tecnologie all'avanguardia per una migliore diagnosi e terapia.

Nell'ultima decade le tecniche di sequenziamento di II generazione (SGS) hanno consentito ragguardevoli progressi rispetto al classico sequenziamento con il metodo Sanger. Tuttavia, alcuni limiti intrinseci delle tecniche di SGS, in primo luogo la breve o brevissima lunghezza delle *reads*, ne hanno impedito l'impiego per specifiche ma fondamentali indagini biologiche come il corretto assemblaggio di genomi, l'organizzazione di regioni genomiche complesse, lo studio delle isoforme geniche o lo studio della metilazione, solo per citarne alcune. In tempi più recenti si sono affacciate sul mercato e, in almeno un paio di casi, si stanno diffondendo nella comunità scientifica, due innovative tecniche di sequenziamento cosiddetto di III generazione (TGS). Sebbene profondamente diverse, le due tecniche SMRT della Pacific Bioscience e Nanopore della ONT, sono accomunate dalla possibilità di sequenziamenti di lunghe o lunghissime *reads* (fino a decina di migliaia di basi) con analisi in tempo reale (*real time sequencing*).

A tal riguardo, si precisa che le strumentazioni per il sequenziamento di III generazione ONT è già nelle disponibilità del gruppo di ricerca. Nel corso del precedente assegno di ricerca il Dott. Mion ha acquisito i fondamenti della tecnica TGS Nanopore applicandola nell'ambito del Sequenziamento di *Bordetella bronchiseptica*. Tali competenze saranno opportunamente impiegate per lo studio di isolati di campo di Sars-Cov-2 raccolti da un centro diagnostico della Regione Emilia-Romagna.

Le attività dell'assegno saranno inserite e coordinate in una più ampia collaborazione il cui obiettivo è la messa a punto di un protocollo rapido di sequenziamento completo del

genoma virale partendo da campioni clinici complessi rappresentati da tamponi floccati posti in UTM (Copan diagnostici) dopo prelievo naso-faringeo.

La collaborazione prevede che il centro diagnostico della Regione Emilia-Romagna fornisca i campioni risultati positivi a Sars-CoV-2 mediante tecnica qPCR, processati ed analizzati nell'ambito della diagnostica clinica. Saranno forniti unicamente campioni rappresentati da materiale genetico purificato da campioni contenenti virus inattivato. I campioni saranno retrotrascritti con metodiche diverse e che saranno oggetto del progetto ad opera di un secondo partner dello studio.

Per gli specifici obiettivi del presente progetto, saranno processati unicamente cDNA che saranno sottoposti a sequenziamento con tecnica ONT. Le attività saranno finalizzate alla messa a punto di un metodo di preparazione della library rapido, semplice e robusto che consenta di ottenere la sequenza pressoché completa del genoma virale da matrice complessa-

Le sequenze dei genomi virali saranno poi messe nella disponibilità dei partners per successive analisi bioinformatiche e valutazioni clinico-epidemiologiche.

In particolare si procederà alla valutazione dell'evoluzione delle quasispecie virali all'interno di specifiche categorie di pazienti e la valutazione filogenetica degli isolati di campo.

Il partner si occuperà degli aspetti della valutazione Etica del progetto di ricerca che in nessun caso prevederà il sequenziamento del materiale genetico umano del paziente.

In particolare, obiettivo primario del presente progetto è la messa a punto di protocolli operativi per l'utilizzo e l'applicazione di tecniche di sequenziamento di III generazione ONT nell'ambito della medicina di precisione con ampie ed evidenti ricadute per il SSD VET08.

#### Piano delle attività

Nello specifico le attività che dovranno essere svolte nell'ambito di questo progetto si focalizzeranno sulle seguenti tematiche:

- 1) Messa a punto di protocolli di amplificazione mediante multiplex PCR di genomi virali interi a partire da matrice complessa (acidi nucleici misti virali, batterici e umani) e con bassa/moderata carica virale.
  - a) I protocolli attualmente in uso prevedono l'esecuzione di multiplex PCR disegnati per amplificare frammenti di genomi di circa 400 bp sovrapposti mediante l'impiego di circa 400 coppie di primer. Per la caratteristica intrinseca di ONT, la massimizzazione in termini di resa della flowcell si ottiene con l'analisi di frammenti lunghi o molto lunghi. Tale massimizzazione della resa si intende con la duplice finalità di ridurre gli sprechi e quindi il costo finale dell'analisi anche di un fattore 10, così come quello di ridurre i tempi di analisi. È indubbi, infatti, che per trasferire queste tecniche molecolari in un ambito clinico come quello della medicina di precisione le tecniche devono essere sostenibili da un punto di vista del costo nonché semplici e rapide nell'esecuzione. Il primo obiettivo sarà quello di disegnare multiplex PCR di lunghezza superiore ai 3000 nt e di ridurre al contempo i primers utilizzati da 400 ad un paio di dozzine al massimo. L'assegnista dovrà in primo luogo procedere con il disegno di primer specifici per

multiplex PCR avendo l'accortezza di definirne la specificità con i genomi disponibili di Sars-CoV-2 facendo ampio ricorso ad allineamenti di primers e genomi. L'efficacia sarà poi verificata sperimentalmente ed affinata empiricamente in base ad esperienze consolidate.

- 2) Implementazione dei protocolli di preparazione della library in senso di semplificazione, ripetibilità, robustezza e rapidità.
  - a) Gli amplificati di Sars-CoV-2 saranno poi processati utilizzando i protocolli disponibili per la tecnologia ONT, combinandoli variamente in modo da ridurre i tempi di preparazione e di conseguire risultati replicabili
- 3) Applicazione sui campioni di campo raccolti dai laboratori impegnati nella diagnostica di campo a fini di valutazioni prognostiche, filogenetiche ed evoluzione delle cd quasispecie.
  - a) Infine l'unità di occuperà di fornire ai partner un'analisi preliminare dei dati mediante pipeline bioinformatiche semplificate che forniscono i seguenti dati minimi: file BAM, sequenza consenso, polimorfismi oltre agli indicatori di qualità della sequenza.

Il progetto rientra nelle competenze del SSD VET08.

European Research Council (ERC) di riferimento: LS7\_2 Diagnostic tools