



**Richiesta per borsa di studio da attivare ai sensi di quanto disposto dal D.M. n. 1061 del  
10/08/2021**

Il sottoscritto NOVELLETTO Andrea, qualifica ordinario, afferente al Dipartimento di Biologia,  
Interno 4104, email [novelletto@bio.uniroma2.it](mailto:novelletto@bio.uniroma2.it)

CHIEDE

L'attivazione di una borsa di studio di dottorato ai sensi di quanto disposto dal D.M. n. 1061 del  
10/08/2021. A tal fine comunica quanto segue:

La borsa sarà attivata sul seguente corso di dottorato accreditato per il XXXVII ciclo: BIOLOGIA  
CELLULARE E MOLECOLARE – R59

Area per la quale si presenta la richiesta (selezionare solo una delle due):

Innovazione

Green

Tipologia di cofinanziamento (pari ad euro 8000 una tantum):

Nome dell'Ente finanziatore pubblico o privato:

Fondi di ricerca dipartimentali

Progetto di Ricerca (massimo 10.000 battute complessive spazi inclusi) che comprenda

Descrizione del Progetto: In genetica forense la rilevanza dell'accertamento di parentela fra  
campioni di DNA anonimi sta crescendo in parallelo all'aumento dei profili depositati nelle banche dati.  
La difficoltà di identificare la vera parentela fra i donatori di due profili basandosi esclusivamente sul  
DNA (campioni anonimi) aumenta con il numero di generazioni che separano i due. Infatti, la  
somiglianza fra parenti dovuta alla condivisione di alleli identici per discesa (IBD) decade con il numero  
di meiosi che li separa, tendendo però ad un valore non-nullo a causa della condivisione di alleli identici  
solo in stato (IBS) che segregano nella popolazione. Un simile livello di somiglianza è presente quindi  
anche fra qualsiasi coppia di individui.

Attualmente esistono sul mercato diversi kit sviluppati per l'analisi non invasiva della paternità  
ma applicabili anche a parentele più distanti, basati sulla tecnica di sequenziamento di nuova generazione  
(NGS), che permette di analizzare contemporaneamente un elevato numero di marcatori genetici. In un  
recente lavoro il proponente ha dimostrato che anche uno dei più diffusi fra questi kit, comprendente  
>100 marcatori, ha una performance molto scarsa nei casi di parentela oltre il primo grado.

La sensibilità analitica della tecnica NGS utilizzata per i test di paternità/parentela basati su SNPs  
è limitata dal fatto che la tecnica stessa è soggetta ad errori: nel sequenziamento vengono introdotte basi  
azotate errate con una frequenza variabile, ma in genere compresa tra 0,2% e 1%. Inoltre, nei casi forensi  
reali, la quantità di materiale presente nelle tracce è un fattore fondamentale: al diminuire della quantità  
di DNA aumenta il numero di posizioni del genoma in cui non è più possibile determinare il genotipo  
del donatore, ovvero la determinazione del genotipo è inaffidabile. La riduzione dei siti confrontabili fra



individui porta ad una drastica riduzione dei likelihood ratios sui quali si basa l'identificazione delle diverse parentele alternative.

Tutto ciò chiama allo sviluppo di kit con numeri di marcatori di ordini di grandezza superiori e analizzabili su quantità minime di DNA.

Con il progetto si intende sfruttare una nuova classe di marcatori genetici chiamati DNP (double nucleotide polymorphisms) in cui la variabilità interessa due nucleotidi consecutivi, e la cui relativa abbondanza nel genoma umano è stata evidenziata solo di recente. L'obiettivo è quello di produrre un prototipo per test di paternità/parentela molto più sensibile di quelli presenti sul mercato e che sia in grado di fornire risultati conclusivi sulla paternità anche in presenza di concentrazioni molto basse di DNA. Questo risultato può essere raggiunto utilizzando i DNP, la cui analisi dovrebbe essere caratterizzata, su base teorica, da un tasso di errore molto più basso (0,01% vs. 1%) rispetto a quello dei marcatori oggi in uso (SNPs). Al fine di verificare questa ipotesi, fondamentale per il successo del progetto, il tasso di errore dei DNP sarà accuratamente misurato al termine della Task 3.1. Uno step di verifica successivo consisterà nel valutare come la diminuzione del tasso di errore di sequenziamento, ovvero del "rumore di fondo" dell'analisi, incida sulla capacità di identificare alleli in situazioni limite, caratterizzate da una bassissima quantità di DNA. A questo fine, saranno prese in considerazione anche diluizioni estreme del DNA in analisi, ben al di sotto del potere di risoluzione dei kit oggi utilizzati.

Obiettivi formativi: acquisizione di competenze di alta specializzazione nel campo della genetica umana e forense. In particolare: esplorazione e interrogazione di banche dati della diversità genetica umana, analisi di dati di sequenziamento del DNA di tipo NGS, saggi sperimentali di validazione, apprendimento di metodologie statistiche da applicarsi in test di ipotesi riguardanti l'accertamento di parentela.

Attività previste:

Task 1.1: individuazione e selezione di 1000 DNP attraverso analisi bioinformatica di genomi completi e banche dati di frequenza e distribuzione geografica di varianti genetiche.

Task 1.2: scelta e preparazione di campioni di DNA per la validazione sperimentale della reazione di targeting (PCR), con speciale attenzione alla quantificazione e diluizione crescente del materiale; verranno utilizzati campioni con genoma già sequenziato ad alta profondità.

Task 2.1: arricchimento delle regioni di interesse (1000 DNP) nei campioni di DNA genomico di cui sopra mediante PCR ad alta multiplexing, disegnata con software apposito.

Task 2.2: sequenziamento ad elevata depth (2000x) del materiale amplificato.

Task 3.1 allineamento delle sequenze e DNP calling.

Task 4.1 validazione statistica del sistema su coppie di individui con relazione di parentela nota (già disponibili nel laboratorio del proponente)..

Attinenza del progetto all'area indicata: Il progetto parte dal TRL2 (concetto della tecnologia formulato) e si prefigge di superare le limitazioni imposte dall'elevato error rate dell'NGS negli attuali test di paternità mediante l'utilizzazione di nuovi marcatori genetici (DNP). Il passo successivo (TRL3) consiste nello sviluppo, mediante estrazione bioinformatica dalle banche dati di diversità del genoma umano, di un set di almeno 1000 marcatori genetici che permettano, in base a considerazioni di tipo



teorico, di raggiungere un livello di significatività statistica sufficiente per discriminare fra un certo numero di relazioni di parentela in qualsiasi coppia di campioni biologici. Su scala nazionale, il Programma Operativo Nazionale (PON) “Ricerca e Innovazione” attuale (2014-2020) si pone tra gli obiettivi principali proprio la realizzazione di metodi clinici avanzati ed innovativi, inserendosi in un contesto di ricerca europeo focalizzato sull'utilizzo delle grandi potenzialità dell'NGS in questo ambito. Più specificatamente, una delle linee principali della Challenge “Health, demographic change and wellbeing” in Horizon 2020 è proprio incentrata sullo sfruttamento dell'NGS nei test di routine. Inoltre, uno degli aspetti fondamentali di H2020 è il sostegno all'innovazione tecnologica che possa contribuire al benessere delle persone, mettendo a punto test su larga scala a livello pre-commerciale e creando al contempo posti di lavoro altamente qualificati. Infine la proposta soddisfa le raccomandazioni di H2020 per progetti che “tengano in appropriata considerazione gli approcci interdisciplinari e, se rilevanti, l'uso delle conoscenze degli stakeholders nel contenuto in ricerca e innovazione”.

Risultati attesi: Il progetto intende dare innanzitutto una proof of concept del fatto che i DNP hanno caratteristiche più adatte ad un'analisi di parentela rispetto agli SNP attualmente utilizzati. Per questo motivo il prodotto finale del progetto sarà un prototipo che abbia superato i collaudi descritti nel Task 4.1. A questo stadio la proprietà intellettuale si applicherà alla particolare lista di marcatori (DNP) elaborata nell'ambito del progetto, che potrà essere adeguatamente tutelata (brevettata). Si osserva che la fase finale di valorizzazione dei risultati, fino ad arrivare alla produzione del kit completo rientrerà fra le attività del Dottorato, e comporterà lo sviluppo di un software che assolva le funzioni computazionali inizialmente svolte in modo one-off. Un simile software rappresenterà certamente un ulteriore componente suscettibile di brevetto. Il progetto si colloca a pieno titolo nella AdS Scienze della vita, per la sua forte connotazione in genetica e genomica. Si deve far notare, tuttavia, che esso rappresenta un deciso avanzamento di tipo diagnostico e, in quanto tale, rientra anche nella AdS Sicurezza per la Societal Challenge Health, Demographic change and Wellbeing. Lo sviluppo di un kit come quello proposto nel presente progetto rappresenta anche un punto cruciale del triangolo della conoscenza “Education, Research, Innovation”, permettendo una forte sinergia tra Università e le aziende del settore sul territorio regionale e nazionale (che hanno un particolare interesse per l'Innovazione).

Azienda pubblica o privata coinvolta nazionale o straniera in cui si prevede di far svolgere il periodo obbligatorio da 6 a 12 mesi previsto dal Decreto Ministeriale:

Salugene s.r.l. semplificata, Spin-off universitario, Un. Tor Vergata

**Firma**