

Validazione di biomarcatori clinici basati su RNA non codificante e microbioma intestinale per il cancro del colon-retto

Introduzione

Il carcinoma del colon-retto (CRC) è la seconda causa di morte per neoplasia nei paesi occidentali sia per l'uomo che per la donna. Un'efficace diagnosi precoce ne aumenta notevolmente la sopravvivenza (a 5 anni è del 95% per gli stadi iniziali e del 5% per quelli avanzati). La colonscopia è la tecnica per eccellenza utilizzata per individuare CRC ed adenomi. Si tratta però di una metodica invasiva, poco accettata dai pazienti, e che può dare complicanze; inoltre è costosa per il sistema sanitario, richiedendo uno staff preparato per l'esecuzione¹. Per superare questi limiti, la ricerca si sta orientando verso l'identificazione di nuovi biomarcatori in grado di discriminare i soggetti sani da quelli con lesioni precancerose o con CRC in *tessutisurrogati* quali ad esempio plasma, siero o feci. Evidenze in letteratura mostrano che alcune specifiche alterazioni genetiche/epigenetiche, determinate configurazioni della popolazione microbica intestinale (microbioma) o di metaboliti da essa prodotti sembrano essere dei candidati ideali sia per la diagnosi che per la prognosi¹.

Gli *small non-coding RNA* (sncRNA), piccole molecole di RNA non codificante le cui funzioni vanno dal controllo trascrizionale, alle modificazioni cromatiniche nel nucleo o alla regolazione della traduzione nel citoplasma, hanno un ruolo rilevante nell'insorgenza di tumori, incluso il CRC. Gli sncRNA sono presenti in tutti i tessuti, inclusi quelli surrogati. Al momento sono stati studiati i microRNA (miRNA), tra i quali ne sono stati identificati alcuni capaci di discriminare CRC e lesioni precancerose in piccole coorti di campioni di plasma²⁻⁴. Non è ancora stato identificato un insieme di sncRNA che consenta una stratificazione dei pazienti rispetto alla lesione che presentano.

È stato dimostrato inoltre che la configurazione del microbioma intestinale è strettamente associata a varie patologie intestinali (ad esempio la rettocolite ulcerosa o il morbo di Crohn), incluso il CRC. Studi preliminari di sequenziamento a bassa risoluzione hanno evidenziato questa associazione⁵⁻⁸, recentemente confermata anche con metodiche di metagenomica di tipo *shotgun* ad alta risoluzione⁸. Poiché il cancro ha un impatto importante sulle vie metaboliche, alterazioni nei livelli dei metaboliti sono altri potenziali biomarcatori nei pazienti affetti da CRC²².

Obiettivo principale del progetto

Questo studio si pone come la continuazione di un progetto finanziato dalla fondazione LILT nel 2015, nel quale si proponeva di studiare la composizione di sncRNA umani e la popolazione microbica in campioni di feci per la diagnosi del CRC. **I risultati incoraggianti** ottenuti ci permettono di proporre un nuovo studio con lo scopo di: **(i) validare i marcatori identificati e stratificare le diverse lesioni tumorali** aumentando il numero di campioni per ogni categoria clinica **(ii) valutare l'efficacia dei marcatori identificati in soggetti ad alto rischio** (pazienti con familiarità per il CRC) **(iii) aggiungere analisi metabolomiche** nel plasma con le quali misurare metaboliti in relazione alla presenza di lesioni pre-cancerose e della patologia tumorale.

Queste analisi ci permetteranno di definire **un pannello di marcatori tumorali** derivati dall'esplorazione dei principali meccanismi coinvolti nello sviluppo del CRC per la messa a punto di nuovi test **non invasivi** per la diagnosi precoce del CRC.

Razionale

Questo progetto presenta molti aspetti essenziali ed innovativi.

- **L'utilizzo di tessuti surrogati** (sangue e feci) permette di eseguire analisi minimamente invasive e con costi ridotti. Inoltre, le feci, a diretto contatto con le lesioni in esame, riflettono l'insieme di alterazioni correlate con l'insorgenza della patologia tumorale anche nelle fasi iniziali.
- **L'analisi integrata del microbioma intestinale, di RNA non codificanti e del metaboloma** utilizzando tecnologie di ultima generazione ed approcci computazionali ad alta risoluzione come

quelli sviluppati dal nostro team di ricerca. Un tale approccio complementare non è mai stato provato, ma ha grandissime potenzialità di generare biomarcatori altamente predittivi.

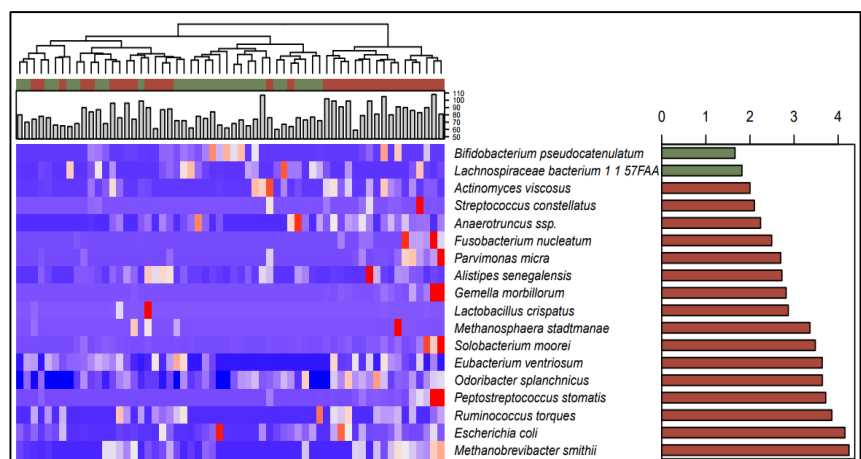
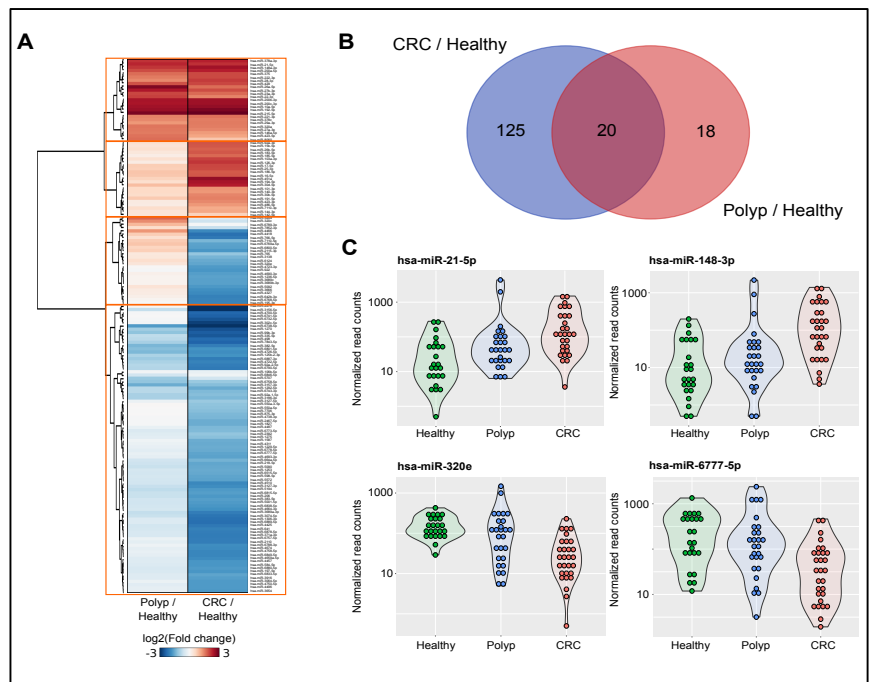
- **L'analisi di soggetti ad alto rischio di sviluppare CRC**, in particolare di pazienti sottoposti frequentemente a colonscopia perché affetti da poliposi adenomatosa familiare (FAP), causata da mutazioni ereditarie nel gene *APC*, che determina la continua comparsa di numerosi polipi i quali posso evolvere in CRC. L'uso di metodiche poco costose e non invasive per identificare lesioni nelle loro fasi iniziali migliorerebbe la qualità della vita dei soggetti FAP.
- **L'impiego di tecnologie di sequenziamento di nuova generazione (Next Generation Sequencing - NGS)** per le quali i proponenti del progetto hanno una provata esperienza.
- **La creazione di un team di ricercatori multidisciplinari.** I ricercatori coinvolti nel progetto, sono specializzati in diverse discipline (biologia, informatica, clinica e biologia computazionale) e costituiscono un gruppo eterogeneo capace di affrontare tutti gli aspetti del progetto.

Risultati preliminari

Abbiamo analizzato i profili di espressione di sncRNA nelle feci di 80 soggetti suddivisi in tre categorie: sani (n=24), con polipi (n=27) o con CRC (n=29). Analisi preliminari eseguite coi 163 miRNA disregolati in polipi e CRC rispetto agli individui sani hanno evidenziato 4 gruppi di profili di regolazione (Figura 1A).

Nei due confronti effettuati 20 miRNA sono comuni (Figura 1B) e tra questi sono rilevabili profili di espressione differenti. In Figura 1C, sono riportati due miRNA (miR-21-5p e miR-148-3p) la cui espressione aumenta al progredire della lesione, e due miRNA (miR-320e e miR-6777-5p) con andamento opposto.

Per gli stessi soggetti abbiamo anche svolto un'analisi preliminare dei dati metagenomici rianalizzando con le metodiche da noi sviluppate i dati ad alta risoluzione ottenuti da Zeller et al¹ per identificare organismi microbici a livello di ceppo²⁻⁴. Questi nuovi metodi, che possono trovare marcatori a livelli tassonomici e funzionali molto specifici, hanno dato una predittività media dei marcatori specie-specifici superiore all'85% (Figura 2).



Metodi

Pazienti arruolati nello studio

Soggetti con sintomatologia suggestiva per patologia del colon e del retto che afferiscono alla Clinica Santa Rita di Vercelli per sottoporsi ad una colonscopia (resp. G. Clerico). Criteri di inclusione: 1) Soggetti caucasici di entrambi i sessi con età > 30 anni che non abbiano avuto precedente diagnosi di CRC. 2) Pazienti che abbiano eseguito una colonscopia completa. I pazienti vengono classificati in: sani, con patologia coloretale di tipo infiammatorio, con lesioni adenomatose, con CRC (nuova diagnosi). Il disegno dello studio è stato approvato dal Comitato Etico locale. Al momento sono stati raccolti oltre 100 campioni biologici, con relativi dati clinici e demografici dei pazienti. Per avere un numero sufficientemente rappresentativo per ognuna delle categorie sopra elencate si prevede di raccogliere 200 campioni. Verranno arruolati 30 soggetti con FAP in sorveglianza endoscopica attiva dopo colectomia profilattica presso la Fondazione IRCCS Istituto Nazionale dei Tumori di Milano (INT, resp. M. Gariboldi) dai quali si preleveranno campioni di sangue e feci a differenti intervalli di tempo. I pazienti partecipano ad uno studio per la prevenzione della comparsa di polipi (approvato dal Comitato Etico di INT).

Raccolta e trattamento dei campioni

Campioni di feci ottenuti prima di iniziare la preparazione alla colonscopia sono raccolti in tubi contenenti una soluzione stabilizzante per DNA e RNA e conservati a -80°C. Vengono anche raccolti 5 ml di sangue dai quali verrà isolato il plasma che verrà conservato a -80°C.

Identificazione di sncRNA

Lo studio può contare su un protocollo di lavoro già ottimizzato⁵. La complessità e la quantità di sncRNA saranno caratterizzate con tecnologia NGS (resp. AG. Naccarati, B. Pardini, M. Gariboldi). La preparazione di library per small RNA sequencing sarà svolta seguendo il protocollo NebNext per Illumina. Frammenti di 135-160 bp saranno sequenziati su piattaforma Illumina. I dati generati dal sequenziamento NGS saranno analizzati con una pipeline di analisi ottimizzata per l'espressione quantitativa dei sncRNA (Resp. Cordero) (Ferrero et al, 2017). I dati raccolti aiuteranno ad identificare un pannello di sncRNA legati alla presenza di lesioni precancerose/cancerose.

Caratterizzazione del microbioma intestinale dai campioni fecali

Il sequenziamento metagenomico di tipo "shotgun" sarà eseguito presso la *facility* di NGS del CIBIO su piattaforma Illumina HiSeq utilizzando materiali e protocolli Nextera XT per la preparazione delle librerie (resp. N. Segata), con metodiche validate nel nostro laboratorio^{6,7}. Genereremo un minimo di 5Gb per campione con sequenziamento *paired-end*. Tale profondità di sequenziamento è sufficiente a coprire con 1X di coverage un genoma batterico di 5MB presente ad un'abbondanza di 0.01%. Analizzeremo i dati con i nostri approcci che rappresentano lo stato dell'arte nel campo della metagenomica⁸. Inoltre caratterizzeremo i campioni dal punto di vista funzionale, tassonomico, filogenetico^{2,9}.

Analisi metabolomica

L'analisi metabolomica verrà eseguita presso la fondazione Edmund Mach (Resp. Donati) tramite una analisi target su un insieme predefinito di metaboliti di interesse.

Analisi statistiche e potenza dello studio

Saranno costruite curve ROC e le conseguenti AUC calcolate per valutare la specificità e la sensibilità dei marcatori identificati. Un metodo basato su un classificatore *semi-supervisionato* sarà applicato per prevedere le relazioni tra i biomarcatori candidati e le diverse covariate (sesso, età, stadio del tumore, grado, ecc.) e applicheremo anche varianti degli strumenti che abbiamo sviluppato appositamente per la ricerca di biomarcatori¹⁰. Per il microbioma applicheremo tutte le tecniche statistiche che abbiamo di recente descritto¹⁰.

REFERENZE

- 1 Zeller, G. *et al.* Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol Syst Biol* **10**, 766, doi:10.15252/msb.20145645 (2014).
- 2 Truong, D. T. *et al.* MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nature methods* **12**, 902-903, doi:10.1038/nmeth.3589 (2015).
- 3 Segata, N., Bornigen, D., Morgan, X. C. & Huttenhower, C. PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. *Nature communications* **4**, 2304, doi:10.1038/ncomms3304 (2013).
- 4 Segata, N. *et al.* Metagenomic microbial community profiling using unique clade-specific marker genes. *Nature methods* **9**, 811-814, doi:10.1038/nmeth.2066 (2012).
- 5 Cordero, F., Beccuti, M., Arigoni, M., Donatelli, S. & Calogero, R. A. Optimizing a massive parallel sequencing workflow for quantitative miRNA expression analysis. *PLoS One* **7**, e31630, doi:10.1371/journal.pone.0031630 (2012).
- 6 Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207-214, doi:10.1038/nature11234 (2012).
- 7 Segata, N. *et al.* Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. *Genome biology* **13**, R42, doi:10.1186/gb-2012-13-6-r42 (2012).
- 8 Segata, N. *et al.* Computational meta'omics for microbial community studies. *Mol Syst Biol* **9**, 666, doi:10.1038/msb.2013.22 (2013).
- 9 Abubucker, S. *et al.* Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS computational biology* **8**, e1002358, doi:10.1371/journal.pcbi.1002358 (2012).
- 10 Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome biology* **12**, R60, doi:10.1186/gb-2011-12-6-r60 (2011).