

*Alma Mater Studiorum – Università di Bologna*

**DOTTORATO DI RICERCA IN  
SCIENZE FARMACOLOGICHE E  
TOSSICOLOGICHE, DELLO SVILUPPO E DEL  
MOVIMENTO UMANO**

**XXVIII CICLO**

**2012-2015**

Settore Concorsuale di afferenza: 05/H2 – ISTOLOGIA

Settore Scientifico Disciplinare: BIO/17 – ISTOLOGIA

**Ricerca di polimorfismi genetici e di marcatori  
molecolari associati al carcinoma del colon-retto a  
scopo diagnostico e prognostico**

Presentata da: Dott.ssa Maria Teresa Rodia

**Coordinatore Dottorato**

Prof.ssa Patrizia Hrelia

**Relatore**

Dott.ssa Rossella Solmi

**Esame finale anno 2016**

<b>Introduzione</b>	6
<b>Il Carcinoma del colon-retto (CCR)</b>	7
<b>Rischio di sviluppare il CCR</b>	8
Primo modello: Carcinogenesi sequenziale determinata da alterazioni molecolari	
Secondo modello: Instabilità genetica, microsatelliti	
<b>Diagnosi del CCR</b>	13
Screening della popolazione	
Ricerca del sangue occulto fecale, Colonscopia, Altre modalità di diagnosi	
<b>Classificazione e stadiazione del CCR</b>	15
TNM	
Stadiazione	
Grading	
<b>I - Parte: Polimorfismi a singolo nucleotide (SNP)–Definizione e caratteristiche</b>	
generali	16
Polimorfismi genetici e rischio di carcinoma del colon-retto (CCR)	18
<b>II- Parte: L'mRNA in qualità di marcatore molecolare e utilizzo del saggio qRT-PCR</b>	22
Transcriptome mapper TRAM	23
<b>Scopo della Ricerca</b>	24
<b>Materiali e Metodi</b>	26
Campione di studio	27
<b>I - Parte: Estrazione del DNA da sangue intero periferico</b>	28
Scelta degli SNP	31
Genotipizzazione	31
Reazione di amplificazione mediante PCR	33
Digestione enzimatica	36

Isolamento di RNA e sintesi di cDNA per la valutazione di espressione genica dei geni CDH1, CTNNB1, LEF1, MDR1(ABCB1), ROCK1, ROCK2	40
Real Time qPCR	40
Analisi statistica	41
<b>II-Parte: Ricerca in database e selezione dei marcatori</b>	
Analisi TRAM (Transcriptome Mapper)	44
Estrazione di RNA totale con Trizol LS	45
Progettazione dei primer	46
Retrotrascrizione RT	47
Real time qPCR	48
Analisi statistiche	49
<b>Risultati</b>	50
<b>I – Parte: Ricerca di polimorfismi genetici associati al rischio di carcinoma del colon-retto</b>	
Studio dell’associazione genetica tra CCR ed i polimorfismi selezionati sul gene CDH1 ed analisi dell’espressione di CDH1, CTNNB1, LEF1	51
Studio dell’associazione genetica tra CCR ed i polimorfismi selezionati sul gene <i>ABCB1</i>	53
Analisi bivariata	53
Analisi multivariate	59
Analisi di associazione per aplotipo	62
Studio dell’associazione genetica tra CCR ed i polimorfismi selezionati sui geni <i>ROCK1</i> e <i>ROCK2</i>	65

<b>II – Parte: Ricerca di mRNA marcatori</b>	
Set di dati basati sulla metanalisi TRAM	68
Selezione dei marcatori	69
Analisi quantitativa dei candidati marcatori mRNA	71
Valore diagnostico dei marcatori di mRNA nel sangue per il CCR	74
Valutazione del potenziale diagnostico del pannello dei quattro mRNA	75
<b>Discussione</b>	80
<b>I – Parte: Ricerca di polimorfismi genetici associati al rischio di carcinoma del colon-retto</b>	81
<b>II – Parte: Ricerca di mRNA marcatori</b>	89
<b>Conclusioni</b>	92
<b>Bibliografia</b>	95
<b>Pubblicazioni</b>	103

## ABSTRACT

Il carcinoma del colon-retto (CCR) comprende tutte le neoplasie con sede tra il cieco ed il retto. E' il tumore più frequente nella popolazione italiana. Tra le cause che ne influenzano il rischio sono noti sia fattori esogeni che endogeni.

La componente genetica riveste un ruolo chiave, infatti le cellule tumorali presentano mutazioni ereditate o acquisite a carico di numerosi geni coinvolti nelle principali vie di segnalazione, replicazione e riparazione.

I nostri studi si sono rivolti all'identificazione di modificazioni nell'ambito degli acidi nucleici, mediante analisi di un semplice campione di sangue periferico.

Abbiamo voluto indagare sia sul rischio genetico di contrarre la malattia e dunque sulla predittività, che sul riscontro della patologia e dunque sulla diagnosi e sulla prognosi.

Studi molecolari hanno difatti evidenziato che alcuni polimorfismi a carico di determinati geni possono contribuire allo sviluppo di patologie tumorali.

Il mio lavoro ha mirato, da una parte, a verificare se alcune varianti alleliche dei geni *CDH1*, *ABCBI*, *ROCK1* e *ROCK2* potessero essere considerate fattori di suscettibilità alla patologia, per un possibile impiego in qualità di marcatori predittivi e, dall'altra, ad identificare mRNA marcatori specifici, selezionati tramite un programma bioinformatico utilizzando, per la validazione, campioni di sangue provenienti da pazienti affetti da cancro del colon-retto e controlli sani.

L'indagine condotta sui polimorfismi scelti per i geni indagati, ha evidenziato, per alcuni, un'associazione significativa, che mi ha permesso di raccogliere dati interessanti sull'eziologia del CCR.

Lo studio di mRNA candidati marcatori, mi ha permesso di identificare *TSPAN8*, *LGALS4* e *COLIA2* quali promettenti strumenti di diagnosi e prognosi per il carcinoma del colon-retto. Ulteriori studi saranno necessari per testare i candidati biomarcatori in test di predittività, screening e di follow-up.

# *Introduzione*

## **Il Carcinoma del Colon-Retto**

Il carcinoma del colon-retto (CCR) comprende tutte le neoplasie che hanno sede tra l'inizio del cieco ed il retto.

E' di gran lunga il tumore più frequente nella popolazione italiana, con un picco tra la quinta e la settima decade di età; al secondo posto per mortalità tumorale nell'uomo ed al terzo nella donna [1], con quasi 52.000 diagnosi stimate per il 2015 da parte del Ministero della Salute, sulla base dei dati dell'Associazione Italiana di Oncologia Medica (AIOM) e dell'Associazione Italiana Registri Tumori (AIRTUM) [2].

Nello specifico le sedi maggiormente colpite da questo tipo di cancro sono il retto (nel 39% dei casi) ed il sigma (nel 25% dei casi); il cancro del colon colpisce indifferentemente entrambi i sessi, mentre il cancro del retto ha un'incidenza con un rapporto maschi:femmine di circa 2:1.

Le tendenze di incidenza, mortalità e prevalenza del tumore del colon-retto stimate in Italia non sono omogenee sul territorio italiano. L'incidenza tra gli uomini, dopo un periodo di crescita, tende a stabilizzarsi nelle regioni settentrionali e centrali, mentre continua ad aumentare nelle regioni meridionali, pur presentando valori più bassi. Nelle donne, invece, l'incidenza si è stabilizzata in tutte le aree italiane con livelli minori nel meridione. Nelle regioni settentrionali, centrali e meridionali i tassi di mortalità per 100.000 abitanti sono rispettivamente di 25.2%, 24.7% e 21.8% per i maschi, mentre 14.6%, 14.4%, 13.6% per le femmine.

Nella maggior parte dei casi, circa l'80% , il CCR si verifica in persone che non hanno fattori di rischio ereditario (CCR sporadico), mentre il 20% si verifica in persone che presentano familiarità. La poliposi adenomatosa familiare (FAP) e la sindrome non poliposica familiare (HNPCC) o sindrome di Lynch sono responsabili di circa il 2-5% di questi tumori [3].

Considerando la grande incidenza di questa patologia, sono stati fatti molti studi per chiarire le basi molecolari ed eziologiche, ma le cause rimangono ancora poco chiare.

### **Rischio di sviluppare il CCR**

Tra le cause che influenzano il rischio di sviluppare la neoplasia si conoscono sia fattori esogeni che endogeni:

- Le forme poliposiche: i soggetti con polipi hanno maggiore probabilità di sviluppare CCR [4].
- Familiarità: aumentano il rischio parentele di primo e secondo grado con poliposi adenomatosa familiare (FAP) o cancro coloretale ereditario non poliposico (HPNCC).
- Presenza di processi infiammatori cronici: il morbo di Crohn e la rettocolite ulcerosa [5].
- Pattern ormonale: recenti ricerche volgono l'attenzione al corredo ormonale vista l'incidenza maggiore negli uomini rispetto alle donne.
- Età: l'incidenza di CCR aumenta con l'età, con un picco tra la quinta e la settima decade, raramente presente in soggetti con età inferiore ai 50 anni se non associato a condizioni ereditarie.
- Fumo: il rischio risulta maggiore del 40% nelle donne fumatrici rispetto alle non fumatrici, mentre per gli uomini il rischio risulta del 30% [6].
- Alimentazione: I soggetti con una dieta ricca in grassi e proteine hanno un rischio maggiore, mentre il consumo di fibre e frutta risultano fattori di protezione [7].
- Fattori ambientali

Riveste un ruolo chiave la componente genetica, infatti le cellule presentano mutazioni ereditate o acquisite a carico di numerosi geni coinvolti nelle principali vie di segnalazione, replicazione e riparazione cellulare.

Le cause genetiche alla base della carcinogenesi sono costituite da instabilità cromosomica, microsatellitare ed eventi epigenetici [1]. Intervengono (nell'arco temporale di circa un decennio) diversi momenti patologici caratterizzati da: acquisizione o perdita di funzione di

geni (proto-oncogeni o oncosoppressori), graduale serie di alterazioni istologiche “sequenza adenoma-carcinoma” (Figura 1). Le mutazioni nel gene Adenomatous Polyposis Coli (APC) non sono solo responsabili della FAP, ma svolgono un ruolo cruciale nella maggior parte dei tumori coloretali sporadici.

La FAP è il risultato di una mutazione germinale sul cromosoma 5q21. Gli individui che ereditano un gene mutante dell'oncosoppressore APC hanno un rischio elevato di sviluppare adenomi del colon. Questa malattia si verifica in media in pazienti di età compresa tra i 35 ed i 40, ed è caratterizzata da molteplici polipi adenomatosi coloretali che progrediscono inevitabilmente in cancro. L'HNPCC è caratterizzata invece dall'assenza di polipi e nella linea germinale si verifica, a livello molecolare, instabilità dei microsatelliti dovuta ad alterazioni del meccanismo del “DNA mismatch repair”.

Sono stati individuati due modelli di eziopatogenesi: un primo modello è costituito dalla via classica della carcinogenesi sequenziale (CCR sporadico), dove lo sviluppo del CCR è il risultato di un accumulo di “alterazioni” molecolari, siano esse genetiche e/o epigenetiche, che si manifestano clinicamente a livello morfologico nella progressione temporale da adenoma a carcinoma [8]. Il secondo modello (10-15% della totalità dei casi ereditari e non ereditari) riguarda l'instabilità genetica dei microsatelliti.

#### Primo modello: Carcinogenesi sequenziale determinata da alterazioni molecolari

L'adenoma tubulare rappresenta il precursore del processo neoplastico. In seguito alla sua crescita viene coinvolta la componente villosa con la formazione di un adenoma avanzato (polipo con atipia citologica e strutturale). Gli eventi molecolari più frequenti sono alterazioni del cariotipo, delezioni alleliche, aneuploidia ed inattivazione della proteina p53 e di importanti geni oncosoppressori quali *APC* (Adenomatous Polyposis Coli), *KRAS* (Kirsten Rat Sarcoma), e di geni localizzati sul braccio lungo del cromosoma 18 (Figura 1).

La comparsa dell'adenoma come primo passaggio verso lo sviluppo di CCR è stata correlata all'insorgenza di mutazioni nel gene *APC* (presenti nel 70-80% dei CCR

sporadici). La proteina APC svolge un ruolo chiave nella regolazione dell'adesione cellula-cellula, nella migrazione cellulare, nella segregazione cromosomica e nell'apoptosi nelle cripte del colon. Essa lega e regola  $\beta$ -catenina la quale media il legame con il dominio della porzione citoplasmatica (DST) dell'E-caderina che a sua volta lega il citoscheletro tramite l' $\alpha$ -catenina. Quest'ultima, insieme alle proteine ZO-1- $\alpha$ -actina e vinculina, farebbe da ponte associandosi con i filamenti di actina. In assenza di un ligando WNT (un segnale mitotico) proveniente dall'esterno,  $\beta$ -catenina viene sequestrata in un complesso formato da diverse proteine: APC (Adenomatous Polyposis Coli), glicogeno-sintasi-chinasi 3 $\beta$  (GSK-3 $\beta$ ), fosfoprotein-fosfatasi 2A (PP-2A) e l'axina. Quando questo complesso è assemblato, la GSK-3 $\beta$  catalizza la fosforilazione e l'ubiquitinazione della  $\beta$ -catenina indirizzandola al proteosoma che ne induce la degradazione. Questo meccanismo impedisce quindi che  $\beta$ -catenina vada al nucleo agendo da cofattore di trascrizione. Quando invece il ligando Wnt lega il proprio recettore, "frizzled", viene attivata una proteina intracellulare (Dsh), reclutata poi sulla membrana a destabilizzare il complesso APC, GSK-3 $\beta$ , axina, che non è più in grado di fosforilare la  $\beta$ -catenina. Mutazioni con perdita di funzione APC comportano la mancata fosforilazione della  $\beta$ -catenina e determinano un'aumentata quantità citosolica della proteina che non viene degradata, ma risulta libera di traslocare nel nucleo ed interagire con vari fattori di trascrizione come quelli delle famiglie TCF (T-cell factor)/LEF(lymphoid enhancing factor), che a loro volta attivano la trascrizione di proto-oncogeni come c-Myc, Ciclina D1 e generalmente, geni coinvolti nei meccanismi di crescita ed inibizione dell'apoptosi [9].

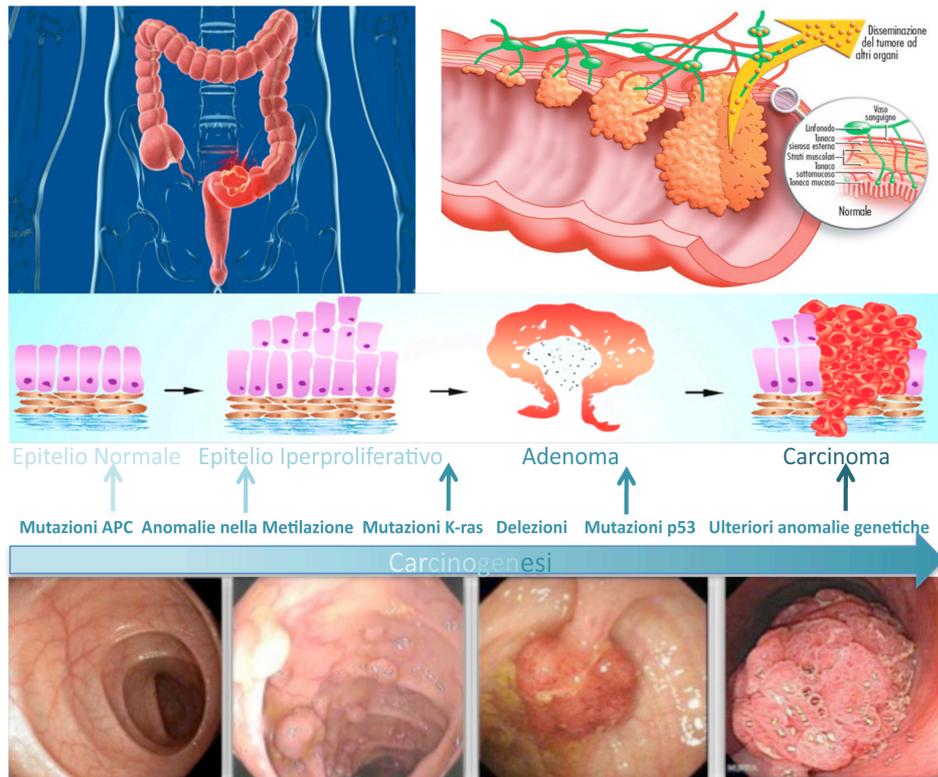
Altre mutazioni che determinano la progressione del CCR sono state viste a carico di Ras (K-Ras, H-Ras e N-Ras), ritrovate nel 40% dei CCR, B-Raf nel 5-10%, PI3K nel 15-25%, PTEN nel 10%, TGF- $\beta$  (fattore di crescita tumorale  $\beta$ ) nel 35%, c-Myc nel 5-10% , p53 nel 70%. In particolare, la perdita di funzione di p53 promuove l'ulteriore accumulo di mutazioni, l'acquisizione di capacità invasive e metastatiche.

Come precedentemente descritto, una perdita della funzione di APC provoca l'attivazione costitutiva della via del segnale Wingless (WNT) e quindi della  $\beta$ -catenina, stimolando la proliferazione cellulare e l'inibizione dell'apoptosi nella cellula della mucosa intestinale. Nella tumorigenesi tali meccanismi stimolano l'espressione di fattori di trascrizione come SNAI 1/2, Twist e ZEB 1/2 che regolano la transizione del fenotipo da epiteliale a mesenchimale (EMT). Le cellule acquisiscono proprietà invasive e di mobilità cellulare metastatiche, con attivazione delle metalloproteinasi di matrice, sovraespressione di N-caderina, vimentina e fibronectina. Il fattore di trascrizione  $\beta$ -catenina, attivato in FAP con APC mutato, aumenta anche l'espressione dell'enzima ciclossigenasi 2 (COX-2) che a sua volta produce elevate quantità della prostaglandina PGE2 con un ruolo centrale nella tumorigenesi intestinale (attivazione della via del segnale NF-kB con aumento della proteina anti-apoptotica Bcl-2). I risultati di tali eventi sono: l'infiammazione, la proliferazione cellulare, la migrazione cellulare e l'angiogenesi. Si ritiene che l'infiammazione cronica contribuisca maggiormente all'iniziazione del CCR sporadico, poiché stimola la produzione di citochine pro-infiammatorie, agenti ossidanti ed enzimi litici. Studi clinici infatti dimostrano che l'utilizzo di farmaci antinfiammatori non steroidei (FANS) che inibiscono in modo selettivo o non selettivo la COX-2, tra cui l'aspirina, costituiscono una buona terapia di prevenzione allo sviluppo del tumore colon-rettale [10].

#### Secondo modello: Instabilità genetica, microsatelliti

L'instabilità genetica dei microsatelliti è dovuta a metilazione del promotore del gene MLH1. Nelle forme ereditarie sono presenti alterazioni nella linea germinale, con il coinvolgimento di diversi geni mutati o alterati. Studi di associazione condotti sull'intero genoma hanno evidenziato alcune varianti di suscettibilità al rischio di CCR, identificando loci sui cromosomi 8q23.3 (EIF3H), 8q24, 10p14, 11q23, 15q13, 16q22, 18q21 (SMAD7), 19q13 e 20p12 [11].

**Figura 1. Carcinogenesi sequenziale: progressione adenoma-carcinoma**



## **Diagnosi del CCR**

### Screening della popolazione

La maggior parte dei tumori colon-rettali, come precedentemente descritto, si sviluppano con la progressione da polipi adenomatosi a tumori maligni nell'arco di diversi anni. Questo rende la diagnosi precoce fondamentale per le possibilità di affrontare la malattia nelle prime fasi del suo sviluppo [12]. Pertanto un regolare programma di diagnosi precoce può prevenire il cancro colon-rettale.

La sopravvivenza generale dei pazienti è condizionata dallo stadio della malattia al momento della diagnosi.

A 5 anni dall'individuazione e cura della malattia si osserva un tasso di sopravvivenza del 90%, quando la lesione era presente in situ, del 68% quando risultava invasa la muscolaris mucosae ed eventuali linfonodi adiacenti e solo del 10% se erano state individuate metastasi a distanza.

Da circa 15 anni in Italia si esegue un "programma di screening per la prevenzione e la diagnosi precoce dei tumori del colon-retto", un invito per le persone tra i 50 e i 69 anni ad eseguire gratuitamente il test per la ricerca del sangue occulto fecale (FOBT) [12-13], come riportato dall'Osservatorio Nazionale degli screening e dal Gruppo Italiano Screening Colon-Rettale (GISCOR). Tale tipo di indagine non presenta rischi e l'individuazione di un sanguinamento occulto può rappresentare un segnale d'allarme. Questo saggio presenta, come del resto tutti i test di popolazione, i parametri di sensibilità (cioè la capacità di individuare tutti i malati) e di specificità (cioè la capacità di individuare solo i malati). La praticità del test si associa ad una sensibilità del 74% che porta all'evidenza di numerosi falsi negativi, individui che pur essendo negativi al sangue occulto presentano in fase di colonscopia la malattia.

## Colonscopia

Una volta che è stato diagnosticato il sangue occulto nelle feci il secondo passaggio per la diagnosi è rappresentato dalla colonscopia, un esame invasivo e con possibilità di rischio per il paziente, che rappresenta l'unico esame in grado di evidenziare lesioni sospette e di eseguire una biopsia per documentarne la natura istologica.

L'esame viene effettuato attraverso l'introduzione per via anale di una sonda, detta colonscopio, munita di microcamera, che mostra in tempo reale l'interno del colon.

## Altre modalità di diagnosi

L'analisi dei livelli di un marcatore tumorale, la glicoproteina dell'antigene carcinoembrionario (CEA), prodotta dalle cellule neoplastiche di origine intestinale, ha rilevato l'esistenza di molti falsi positivi, in quanto è prodotta anche in caso di generici stati infiammatori.

Recentemente i test che rilevano la presenza di emoglobina nel sangue sono stati arricchiti dall'indagine del DNA [14]. Imperiale ed altri [15] hanno proposto un test, Cologuard, (Exact Sciences Corporation, Madison WI), che è stato approvato dalla Food and Drug Administration (FDA), ma l'evidenziazione di falsi positivi ha indotto ad assumere ulteriori valutazioni diagnostiche [16]. Nel 2010, il sistema di CellSearchVR (Veridex, Johnson-Johnson, USA) per il rilevamento di cellule tumorali circolanti (CTC) nel cancro coloretale metastatico (mCCR) con la tecnica dell'immunofluorescenza, ha pure ricevuto l'approvazione della FDA [17-18]. Marshall ed altri [19] invece hanno proposto un test per determinare il rischio di insorgenza al CCR [20] costituito da sette mRNA da utilizzarsi come biomarcatori nel sangue: Annessina A3 (ANXA3), Lectina tipo-C della famiglia D dominio 4 (CLEC4D), Laminina B1 (LMNB1), Prolina Ricca in Gla (acido G-carbossiglutammico) 4 (PRRG4), Fattore di Necrosi Tumorale Proteina Alfa-Indotta 6 (TNFAIP6), Vanin 1 (VNN1) ed il Recettore Beta Interleuchina 2 (IL2RB) (ColonSentry®),

Canada -. Enzo Biochem USA). Il test è stato recentemente approvato dal Dipartimento di Stato della Salute di New York.

### **Classificazione e Stadiazione del CCR**

#### TNM

La classificazione TNM è il sistema internazionale di classificazione dell'evoluzione di un tumore (stadiazione), e considera tre parametri: dimensioni del tumore primitivo (T), coinvolgimento dei linfonodi regionali adiacenti al tumore (N) e presenza di metastasi a distanza (M).

<b>CLASSE</b>	<b>DESCRIZIONE</b>
<b>TX</b>	Tumore primitivo non definibile
<b>T0</b>	Tumore primitivo non evidenziabile
<b>Tis</b>	Carcinoma in situ: intraepiteliale o invasione della lamina propria [comprende cellule tumorali confinate all'interno della membrana basale ghiandolare (intraepiteliale) o della lamina propria (intramucosa) che non raggiungono la sottomucosa]
<b>T1</b>	Tumore che invade la sottomucosa
<b>T2</b>	Tumore che invade la muscolare propria
<b>T3</b>	Tumore con invasione attraverso la muscolare propria nella sottosierosa o nei tessuti pericolici e perirettali non ricoperti da peritoneo
<b>T4</b>	Tumore che invade direttamente altri organi o strutture e/o perfora il peritoneo viscerale
<b>T4a</b>	Tumore che perfora il peritoneo viscerale
<b>T4b</b>	Tumore che invade direttamente altri organi o strutture
<b>NX</b>	Linfonodi regionali non valutabili
<b>N0</b>	Non metastasi nei linfonodi regionali
<b>N1</b>	Metastasi in 1-3 linfonodi regionali
<b>N1a</b>	Metastasi in 1 linfonodo
<b>N1b</b>	Metastasi in 2-3 linfonodi
<b>N1c</b>	Depositi tumorali satelliti nella sottosierosa o nei tessuti non peritonealizzati pericolici e perirettali senza evidenza di metastasi linfonodali regionali
<b>N2</b>	Metastasi in 4 o più linfonodi regionali
<b>N2a</b>	Metastasi in 4-6 linfonodi
<b>N2b</b>	Metastasi in 7 o più linfonodi
<b>MX</b>	Metastasi a distanza non accertabili
<b>M0</b>	Assenza di metastasi a distanza
<b>M1</b>	Metastasi a distanza
<b>M1a</b>	Metastasi confinate ad un organo (fegato, polmone, ovaio, linfonodi extraregionali)
<b>M1b</b>	Metastasi in più di un organo o nel peritoneo

## Stadiazione

Per quanto riguarda la stadiazione, lo stadio è un termine utilizzato convenzionalmente dai medici per descrivere le dimensioni di un tumore e la sua eventuale diffusione ad altri organi. Queste informazioni servono per stabilire qual è il trattamento più appropriato per il singolo caso.

Suddivisione in 4 stadi:

- stadio I: il tumore è circoscritto all'interno della parete intestinale;
- stadio II: il tumore ha invaso tutta la parete intestinale, ma i linfonodi sono indenni;
- stadio III: il tumore ha invaso uno o più linfonodi localizzati vicino all'intestino. I linfonodi sono di solito i primi ad essere raggiunti dalle cellule tumorali se il tumore si diffonde al di fuori dell'intestino;
- stadio IV: il tumore si è diffuso ad altri organi quali il fegato o il polmone.

<b>STADIO</b>	<b>DESCRIZIONE</b>
<b>Stadio 0</b>	Tis N0M0
<b>Stadio I</b>	T1N0M0
	T2N0M0
<b>Stadio IIa</b>	T3N0M0
<b>Stadio IIb</b>	T4aN0M0
<b>Stadio IIc</b>	T4bN0M0
<b>Stadio IIIa</b>	T1-2, N1a-c, M0
	T1, N2a, M0
<b>Stadio IIIb</b>	T3, T4a, N1a-c, M0
	T2-3, N2a, M0
	T1-2, N2b, M0
<b>Stadio IIIc</b>	T4a, N2a-b, M0
	T3, N2b, M0
	T4b, N1-2, M0
<b>Stadio IVa</b>	Ogni T, ogni N, M1a
<b>Stadio IVb</b>	Ogni T, ogni N, M1b

Se il tumore si ripresenta dopo il trattamento, si parla di recidiva

## Grading

Il grading invece è un termine utilizzato convenzionalmente per descrivere l'aspetto delle cellule tumorali al microscopio e classificare i tumori in base al grado di differenziazione cellulare e alla rapidità con cui possono crescere. I tumori del colon-retto si classificano secondo tre gradi:

- grado 1: grado basso;
- grado 2: grado intermedio o moderato;
- grado 3: grado alto.

Nei tumori di grado 1 le cellule tumorali sono molto simili alle cellule normali, tendono a crescere lentamente e difficilmente si diffondono ad altri organi. Al contrario, nei tumori di grado 3 sono molto differenziate, tendono a crescere rapidamente e a diffondersi ad altri organi [21-22].

### **I-Parte: Polimorfismi a singolo nucleotide (SNP)–Definizione e caratteristiche generali**

I polimorfismi a singolo nucleotide sono presenti nella popolazione con una frequenza allelica  $> 1\%$  e rappresentano il tipo di variazione genetica più comune nell'uomo (90% dei polimorfismi). Essi sono presenti con varianti alleliche alternative e seguono un'ereditarietà mendeliana. L'ampia diffusione nel corredo genetico (uno ogni 100-300 paia di basi) li rende interessanti negli studi di associazione allelica e fruibili in qualità di marcatori molecolari. Lo SNP può dare origine ad una mutazione sinonima o non-sinonima se presente in una regione codificante di un gene; oppure portare ad alterazioni del prodotto genico se presente in regioni introniche o intergeniche.

Le variazioni nel genoma possono determinare differenti suscettibilità individuali ad una malattia, o differenti risposte farmacologiche. Gli SNP risultano pertanto utili per l'identificazione di "geni malattia" mediante studi di associazione su screening di popolazione. Grazie al progetto Hap-Map nel genoma umano sono stati identificati milioni

di polimorfismi (quattro differenti popolazioni) ed organizzati in blocchi di "linkage disequilibrium".

### **Polimorfismi genetici e rischio di carcinoma del colon-retto (CCR)**

Numerosi studi sono stati condotti per rivelare la potenziale associazione tra varianti genetiche polimorfiche (SNP) e il rischio di CCR [23], sebbene il contributo genetico all'insorgenza del CCR non sia ancora completamente chiarito. Dieta e stile di vita costituiscono oltre il 50% delle basi eziologiche del CCR, ma la suscettibilità individuale ed i fattori che avviano lo sviluppo del tumore e la sua progressione restano motivi fondamentali di ricerca.

Studi molecolari hanno difatti evidenziato che alcuni polimorfismi a carico di determinati geni possono contribuire allo sviluppo di patologie tumorali e ad una variabilità di risposta al trattamento con farmaci antitumorali [24].

Al fine di valutare la possibile correlazione tra alcuni polimorfismi a singolo nucleotide ed il rischio di CCR, abbiamo eseguito un'analisi di associazione caso-controllo in un campione di studio rappresentativo della popolazione italiana. L'obiettivo del lavoro è stato quello di verificare se varianti alleliche dei geni caderina 1 (*CDH1*), Multidrug Resistance Protein 1 (*MDR1*), Rho-Associated Coiled-Coil Containing Protein Kinase 1 (*ROCK1*) e Rho-Associated Coiled-Coil Containing Protein Kinase 2 (*ROCK2*) possano essere considerate fattori di suscettibilità alla patologia, ed essere impiegate come marcatori diagnostici nello studio del cancro del colon-retto.

Di seguito sono riportate le caratteristiche dei geni scelti al fine di evidenziare le motivazioni dell'indagine effettuata.

#### ➤ *CDH1, CTNNB1, LEF1*

Il gene *CDH1*, localizzato sul cromosoma 16q22.1, si estende per 98.31 Kbp e comprende 16 esoni e 15 introni. Il prodotto di *CDH1* è una glicoproteina transmembrana costituita da 882 aminoacidi. L'E-caderina è coinvolta in complessi di adesione nelle cellule epiteliali,

ma anche nella segnalazione intracellulare, insieme alle proteine citoplasmatiche appartenenti alla famiglia delle catenine. Le giunzioni cellula-cellula giocano un ruolo fondamentale nella creazione e nel mantenimento della polarità cellulare e nell'integrità del tessuto epiteliale. Gli eventi che portano ad una riduzione dell'adesione cellula-cellula alterano l'ordine nella gerarchia cellulare, evolvendo verso la distruzione della struttura istologica, con conseguenze potenzialmente patologiche. Una alterazione dell'adesione cellula-cellula è stata osservata nelle fasi iniziali della transizione epitelio-mesenchimale (EMT), che prevede una minor espressione della E-caderina. Questa variazione viene considerata un indicatore di EMT, a sua volta considerata responsabile della progressione e della capacità di metastatizzare di un certo numero di tumori di origine epiteliale [25]. Una espressione minore o assente di E-caderina è indicata come marcatore nel CCR, ad indicare uno stato invasivo, anche se sono stati riportati dati controversi in funzione dell'etnia considerata [26].

Una presenza ridotta del prodotto transmembrana di CDH1 porta ad un'attivazione aberrante del pathway Wnt- $\beta$ -catenina, con un conseguente aumento di attività del fattore di trascrizione  $\beta$ -catenina-TCF [27]. In circostanze normali, il livello citoplasmatico di  $\beta$ -catenina (codificata dal gene CTNNB1) è regolato da un complesso di proteine a cui risulta legata, e l'eccesso di proteina viene eventualmente eliminata dal proteasoma. Se si verifica una perturbazione di questo equilibrio, la  $\beta$ -catenina si accumula nel citoplasma e trasloca al nucleo dove si lega al fattore TCF/LEF attivando una serie di geni bersaglio. In particolare, i bersagli del fattore di trascrizione LEF1, sono coinvolti nella transizione epitelio-mesenchimale, un processo, come già anticipato, considerato fondamentale nell'evoluzione neoplastica [28].

Inoltre sono stati valutati anche i livelli di espressione di due geni correlati, come descritto precedentemente: la beta-catenina (CTNNB1) ed il fattore linfoide enhancer di legame 1

(LEF1). Il gene *CTNNB1* è localizzato sul braccio corto del cromosoma 3p21, ha 16 esoni e si estende per 23.2 kb.

Il gene *LEF1* è localizzato sul braccio lungo del cromosoma 4q25, contiene 12 esoni e si estende almeno 52 kb.

Nel presente studio abbiamo quindi indagato l'ipotesi che polimorfismi presenti nel gene *CDH1* a livello del DNA costitutivo possano influenzare il rischio di CCR. Abbiamo valutato i livelli di espressione di  $\beta$ -catenina, LEF1 e CDH1 in biopsie di alcuni pazienti CCR.

➤ *MDR1 (ABCB1)*

Il gene è localizzato nel braccio lungo del cromosoma 7, contiene 27 esoni ed una sequenza codificante di 96 Kb. MDR1 (Multi Drug Resistance 1) codifica per una glicoproteina (P-gp) costituita da 12 segmenti transmembrana e due siti enzimatici per l'idrolisi dell'ATP. Appartiene ad un gruppo di pompe di efflusso della famiglia ABC (ATP -Binding Cassette) ATP-dipendenti che selettivamente estrudono sostanze tossiche e farmaci dalle cellule. Più in generale, un'alterata funzione della glicoproteina P-gp potrebbe creare una perturbazione dell'equilibrio intra-extra cellulare, tale da portare ad un aumento del rischio di internalizzazione di fattori tossici o potenziali cancerogeni danneggianti il DNA che potrebbero modulare in tal modo la suscettibilità alla trasformazione neoplastica. Nell'epitelio intestinale avvengono le interazioni tra le sostanze introdotte attraverso la dieta ed i prodotti genici coinvolti nel metabolismo. In particolare MDR1 costituisce una sorta di barriera finalizzata a regolare la diffusione di metaboliti e farmaci [29].

Studi molecolari hanno evidenziato che alcuni polimorfismi a livello di geni codificanti proteine della superfamiglia ABC, possono contribuire allo sviluppo di patologie tumorali e ad una variabilità di risposta al trattamento con farmaci antitumorali [24].

➤ *ROCK1 e ROCK2*

Il gene *ROCK1* è localizzato nel cromosoma 18q11.1, contiene 35 esoni ed una sequenza codificante di 160 kDa; *ROCK2* è localizzato nel cromosoma 2p24, contiene 35 esoni ed una sequenza codificante di 160 kDa. Entrambi risultano i principali effettori a valle della famiglia Rho GTPasi, sono chinasi serina-treonina che agiscono come regolatori centrali della dinamica del citoscheletro actina dipendente. Attraverso la fosforilazione di una vasta gamma di substrati bersaglio coinvolti nell'organizzazione del citoscheletro, come la catena leggera della miosina (MLC) [30, 31], MLC fosfatasi (MYPT1) [32, 33], ezrina / radixina / moesina (ERM) [34], l'adducina [35], e la LIM chinasi (LIMK) [36], le ROCK controllano la formazione delle fibre di actina, delle adesioni focali e la contrattilità cellulare basata sulla miosina [37]. Le due isoforme descritte, sebbene codificate da geni diversi, mostrano un'omologia di sequenza amminoacidica del 65% [38]. Nonostante l'ampia omologia nonché la loro promiscuità di substrato, le due isoforme ROCK mostrano funzioni distinte. In particolare, la perdita del gene *ROCK2* è stata associata ad un incremento dell'assemblaggio dei microfilamenti in fibre da stress ed alla formazione di adesioni focali, mentre il gene *ROCK1* sembra necessario per la fagocitosi miosina II dipendente [39-41]. Dal momento che la dinamica del citoscheletro media una serie di processi cellulari coinvolti nell'espressione del fenotipo maligno, come la regolazione della morfologia, l'adesione cellula-cellula, la migrazione e l'invasione, non stupisce che i prodotti dei geni *ROCK1* e *ROCK2* siano coinvolti nell'insorgenza del cancro. La sovraespressione delle ROCK è stata osservata in molti tumori solidi, tra cui il carcinoma epatocellulare [37], l'osteosarcoma [38, 39], il cancro al seno [40], il cancro ai testicoli [41], il cancro della vescica [42], così come nel tumore del colon [43]. Inoltre, diversi studi hanno sottolineato un'attività anti-migratoria e anti-invasiva dovuta ad inibizione delle ROCK [44,45] rafforzando l'associazione tra l'asse Rho-rock ed il cancro [46, 47].

In questo studio sono stati valutati alcuni polimorfismi di queste Rho-chinasi quali potenziali fattori di suscettibilità al cancro coloretale in un campione rappresentativo della popolazione Italiana.

Abbiamo scelto un totale di nove polimorfismi a singolo nucleotide: 4 mappanti nel gene *ROCK1* e 5 mappanti nel gene *ROCK2*.

## **II-Parte: L'mRNA in qualità di marcatore molecolare e utilizzo del saggio qRT-PCR**

La ricerca di marcatori utilizzati come strumento di screening nel sangue dei pazienti rappresenta un attivo campo di ricerca per la diagnosi precoce del cancro del colon-retto. Numerosi progetti includono lo studio di mRNA codificanti, microRNA (miRNA), proteine metaboliche, mutazioni del DNA [48] e marcatori di metilazione. Attualmente, le principali tendenze della ricerca per candidati marcatori mRNA coinvolgono molte prove sperimentali: cellule tumorali circolanti (CTC) [49], cellule staminali tumorali (CSC) [50] e RNA libero circolante (cfRNA) [51]. La diffusione metastatica si verifica molto presto nello sviluppo del tumore; quindi il rilevamento specifico e sensibile delle CTC è diventato di fondamentale importanza per la diagnosi [52] e l'utilizzo della reazione a catena della polimerasi quantitativa (qRT-PCR) è stata recentemente descritta come un buon metodo per la quantificazione delle CTC [53,54].

Gli RNA circolanti liberi (cfRNAs) potrebbero essere degli eccellenti biomarcatori tumorali del sangue, in quanto più informativi, specifici e precisi rispetto ai biomarcatori proteici e molti gruppi di ricerca hanno studiato l'uso potenziale di mRNA circolanti come marcatori per il cancro[55,56]. In generale, la strategia sperimentale è quella di utilizzare la tecnologia dei microarray per il profilo di espressione degli mRNA, che è seguita dalla reazione a catena della polimerasi in tempo reale (qRT-PCR). I campioni utilizzati sono sia mRNA estratti direttamente dal sangue o da frazioni del sangue (siero/plasma)[57].

La nostra ricerca, così come quella di altri gruppi, consiste di un test sul sangue intero senza isolare CTC, CSC o qualsiasi altra frazione del sangue, è molto più semplice da condurre

nella prospettiva di un ampio utilizzo nella pratica oncologica e non è influenzata dalla perdita di CTC o CSC, il cui numero risulta critico per il successo del test, durante la purificazione [58]. L'utilizzo della molecola di RNA tiene conto del fatto che la variazione fenotipica del tumore è associata a cambiamenti nei livelli di mRNA di geni che regolano o influenzano questa variazione.

La ricerca di mRNA marcatori nel sangue ci ha spinto all'utilizzo del saggio qRT-PCR già in ampio uso in diagnostica clinica [59]. La RT-PCR offre diversi vantaggi rispetto ad altri metodi di rilevazione in termini di elevata sensibilità e specificità. Ad oggi, il problema principale è stato quello di identificare mRNA biologicamente adeguati con un chiaro valore soglia dei livelli di espressione tra CCR e soggetti sani. Abbiamo affrontato questo problema sfruttando i dati presenti in letteratura da banche dati on-line di esperimenti di microarray sull'espressione differenziale tra i tessuti [60,61].

Tuttavia, a nostra conoscenza, nessuno studio ha finora affrontato un approccio globale comprendente l'analisi sistematica dell'espressione genica in tessuti CCR ed in cellule del sangue.

### **Transcriptome mapper TRAM**

In particolare abbiamo usato un nuovo approccio basato sul programma bioinformatico Transcriptome Mapper (TRAM) [62,63], volto ad identificare specifici RNA [64] con il più alto rapporto di espressione differenziale tra CCR e campioni di sangue normali e quindi possibilmente adatto per il rilevamento di modifiche relative al sangue dei pazienti con CCR. TRAM è basato su una meta-analisi sistematica su larga scala di tutti i dati disponibili in database ottenuti da esperimenti di microarray.

I migliori RNA candidati, individuati dal nostro approccio di biologia computazionale, tra 38.104 trascritti umani la cui espressione nel sangue è stata confrontata fra CCR (n = 349) e soggetti sani (n = 200), sono stati testati con una qRT-PCR. Per il nostro è stato arruolato studio un gruppo di 134 soggetti tra volontari sani e pazienti con diagnosi di CCR.

# *Scopo della ricerca*

In generale si possono considerare marcatori tumorali tutte quelle molecole che si rivelano diverse dalla normalità in caso di malattia. I nostri studi si sono orientati all'identificazione di modificazioni riguardanti gli acidi nucleici, analizzando un semplice campione di sangue periferico.

Abbiamo voluto indagare sia nel campo del rischio genetico di contrarre la malattia e dunque sulla possibile predittività della malattia che nel campo del riscontro della patologia e dunque sulla possibilità della diagnosi e della prognosi.

#### I Parte: Ricerca di polimorfismi genetici associati al rischio di carcinoma del colon retto.

Le cause del CCR restano materia d'indagine scientifica e molti studi sono condotti al fine di chiarirne i processi molecolari e la suscettibilità individuale.

L'obiettivo di questo filone di ricerca è quello di valutare la possibile correlazione tra polimorfismi a singolo nucleotide (possibilmente presenti in alcuni geni chiave dell'eziologia tumorale) ed il rischio di sviluppare il CCR.

Inoltre, in virtù della marcata presenza nel genoma di SNP che caratterizzano ogni individuo è poi di grande interesse effettuare studi di associazione mirati all'identificazione, in generale, di particolari SNP coinvolti nel rischio di insorgenza del CCR, a scopo predittivo.

Dalla letteratura scientifica abbiamo individuato un set di geni i cui SNP potessero risultare rilevanti nel CCR. L'evidenza di un eventuale rischio può rappresentare una premessa per la messa a punto di uno screening di popolazione ed una terapia preventiva mirata agli individui a rischio, nell'ambito della medicina personalizzata.

#### II Parte: Ricerca di marcatori mRNA associati al carcinoma del colon-retto.

Se il CCR viene rilevato in una fase iniziale, il tasso di sopravvivenza, dal momento dell'intervento terapeutico, aumenta in modo significativo. Partendo da questa evidenza oggettiva, da anni il nostro gruppo di ricerca si occupa di identificare marcatori di CCR ad mRNA, nel sangue. Abbiamo scelto la molecola di mRNA in considerazione del fatto che

la variazione fenotipica del tumore è associata a cambiamenti nei livelli di mRNA di geni che regolano o che subiscono la regolazione. Un test su sangue intero, senza isolare le frazioni del sangue, può risultare uno strumento importante per migliorare sia la diagnosi precoce del cancro del colon-retto che la prognosi in caso di malattia conclamata in corso di progressione.

Il nostro presente progetto di ricerca si è innanzitutto proposto di individuare un pannello di mRNA marcatori la cui combinazione risultasse efficace ed accurata nella diagnosi e prognosi. A tale scopo abbiamo scelto un nuovo approccio metodologico, finalizzato ad identificare mRNA specifici, selezionandoli dai dati disponibili, ottenuti da esperimenti di microarray, tramite un programma bioinformatico, Transcriptome Mapper (TRAM). Sono stati selezionati come candidati mRNA marcatori quelli con il più alto rapporto di espressione differenziale tra il tessuto del cancro coloretale ed i campioni normali di sangue.

Gli mRNA candidati sono stati testati in campioni di sangue provenienti da pazienti affetti da cancro del colon-retto e controlli sani, per verificare la loro capacità di identificare correttamente le persone con il cancro. Per la validazione dei migliori candidati marcatori ottenuti mediante TRAM, nel sangue di 67 pazienti e 67 controlli sani, è stata effettuata una reazione quantitativa a catena della polimerasi in tempo reale (qReal Time PCR). Obiettivi futuri, rispetto ai risultati riscontrati dalla presente nostra ricerca, saranno quelli di verificare la validità dei marcatori individuati nel riconoscimento specifico del carcinoma del colon-retto rispetto ad altre malattie gastro-intestinali e, nell'ambito della diagnosi precoce, la possibilità di poter riconoscere i vari stadi della malattia. Nell'ambito della prognosi, la possibilità di rilevare invece la progressione del carcinoma del colon-retto in pazienti in follow-up.

# *Materiali e Metodi*

## **Campione di studio**

Per i diversi studi, una coorte di pazienti italiani non imparentati, con diagnosi di cancro coloretale è stata reclutata dal Dipartimento di Scienze Mediche e Chirurgiche del Policlinico S. Orsola-Malpighi, dell'Università di Bologna. La diagnosi di cancro coloretale è stata confermata da un esame istopatologico. Il gruppo di controllo era costituito da volontari sani non imparentati, di nazionalità italiana senza pregressa storia clinica di malattia neoplastica. A tutti è stato effettuato un prelievo ematico. Ai pazienti è stato eseguito prima dell'intervento chirurgico primario. Tutti i pazienti arruolati non erano stati preliminarmente trattati con chemioterapia o radioterapia adiuvante. Il campione di sangue intero periferico di pazienti e controlli (circa 2mL) è stato raccolto in provette contenenti l'anticoagulante EDTA. Per ridurre la contaminazione dei campioni con le cellule epiteliali provenienti dalla puntura del prelievo, è stato scartato il primo ml di sangue. Per la realizzazione degli studi condotti su *CDHI*, *CTNNB1* e *LEF1* sono stati inoltre prelevati, per 27 pazienti, campioni biotipici sia del tumore che del tessuto normale circostante. Le informazioni relative ai pazienti reclutati per ciascuna indagine sui diversi geni sono riassunte in Tabelle 1A, 1B, 1C, 1D. Gli studi sono stati approvati dal comitato etico del Policlinico S. Orsola-Malpighi rispettando i principi etici della Dichiarazione di Helsinki per la ricerca medica che coinvolge i soggetti umani. Ciascun partecipante ha aderito al progetto di ricerca scientifica firmando un modulo per il consenso informato.

Per lo studio riguardante i candidati mRNA marcatori abbiamo effettuato l'analisi della potenza statistica dello studio utilizzando il software G \* Potenza 3 per il sistema Macintosh OS X. La dimensione del campione necessario per raggiungere una potenza di 0.80 quando si analizzano i risultati utilizzando il t-test (due code), è stato stimato essere almeno 64 soggetti sani di controllo e 64 pazienti con diagnosi di CCR (per alfa (tipo I) l'errore = 0.05 e assumendo dimensioni medie effetto  $d = 0.5$ ).

Sono stati arruolati per lo studio 67 donatori sani e 67 pazienti.

**Tabella 1.** Informazioni sui pazienti

**1A.** Ricerca di polimorfismi del gene *CDH1* e valutazione dell'espressione tissutale per *CDH1*, *CTNNB1* e *LEF1*

		Casi	Controlli	<i>P</i> value	Biopsie
Età		69±13	40±15	< 0.01	
Sesso	maschi	73	69	0.59	13
	femmine	67	72		14
		<u>140</u>	<u>141</u>		

**1B.** Ricerca di polimorfismi del gene *ABCBI*

		Casi	Controlli	<i>P</i> value
	età	70±13	68±12	0.26
Sesso	maschi	52	49	0.57
	femmine	46	51	
Grado	G1	5	-	
	G2	65	-	
	G3	21	-	
	N.D.	7	-	
Posizione	cieco	1	-	
	c.ascendente	33	-	
	c. trasverso	9	-	
	c.discendente	12	-	
	c. sigma	20	-	
	retto	23	-	
		<u>98</u>	<u>100</u>	

1C. Ricerca di polimorfismi dei geni *ROCK1* e *ROCK2*

		Casi, n (%)
Età, anni	<50	8 (5.8)
	≥50	123 (89.8)
	*N.D.	6 (4.4)
Sesso	Maschi	72 (52.6)
	Femmine	65 (47.4)
		137
Grado	G1	8 (5.8)
	G2	90 (65.7)
	G3	25 (18.2)
	*N.D.	14 (10.2)
Metastasi linfonodali	N0	72 (52.6)
	N1/2	42 (30.7)
	*N.D.	23 (16.8)
Storia familiare	No Familiarità	111 (81)
	Si Familiarità	14 (10)
	*N.D.	12 (8.8)
Storia fumatori	No	79 (57.7)
	Si	16 (11.7)
	Ex Fumatori	31 (22.6)
	*N.D.	11 (8)

\*N.D. Nessun Dato

**1D. Ricerca di mRNA candidati marcatori. Informazioni sui pazienti ed i controlli arruolati nello studio**

		<b>Casi</b>	<b>Controlli</b>	<b>P value</b>
	Età	68±12	65±14	0.40
Sesso	Maschi	34	35	0.42
	Femmine	33	32	
Grado	G1	6	-	
	G2	45	-	
	G3	9	-	
	*N.D.	7	-	
Posizione	Cieco	3	-	
	Colon Ascendente	17	-	
	Colon Trasverso	3	-	
	Colon Discendente	15	-	
	Sigma	10	-	
TNM	Retto	19	-	
	T1N0	13	-	
	T2N0	15	-	
	T2N1	3	-	
	T3N0	13	-	
	T3N1	8	-	
	T3N2	6	-	
	T4N0	3	-	
	T4N1	1	-	
	T4N2	2	-	
Dukes	*N.D.	3	-	
	A	28	-	
	B	14	-	
	C	19	-	
Stadiazione	*N.D.	6	-	
	1	25	-	
	2 A/B	15	-	
	3 A/B/C	18	-	
BMI	*N.D.	9	-	
	≤ 25	26	-	
	>25	32	-	
Storia fumatori	*N.D.	9	-	
	No	35	-	
	Si	8	-	
	Ex Fumatori	22	-	
Storia familiare	*N.D.	2	-	
	No Familiarità	56	-	
	Si Familiarità	8	-	
*N.D. Nessun Dato	*N.D.	3	-	

## **I-Parte**

### **Estrazione del DNA da sangue intero periferico**

Il metodo di estrazione è basato sul protocollo di Higuchi (Higuchi 1989) e prevede l'isolamento del DNA da leucociti.

#### Protocollo di estrazione:

- Utilizzando 0.5 ml di sangue intero vengono aggiunti 0.5 ml di Lysis Buffer, necessario per lisare le membrane plasmatiche di tutti gli elementi figurati del sangue. Le provette vengono agitate vigorosamente a più riprese per favorire la lisi, quindi centrifugate a 13000 giri per 20 sec. Il surnatante viene rimosso facendo attenzione a non aspirare il pellet contenente i nuclei dei leucociti.
- Aggiungendo 1 ml di Lysis Buffer al pellet lo si risospende dapprima invertendo manualmente la provetta più volte e poi agitando fino alla sua completa risospensione.
- Dopo un passaggio in centrifuga a 13000 giri per 20 sec va ripetuto il lavaggio con Lysis Buffer finchè il pellet non risulta più roseo (indice di presenza di emoglobina che può inibire l'attività della Taq polimerasi) ed in seguito ancora una centrifuga a 13000 giri per 30 sec.
- Rimosso il surnatante, il pellet viene risospeso in 0.5 ml di PCR Buffer addizionato di proteinasi K (6 µl 10 mg/ml per ogni ml di PCR Buffer impiegato), una soluzione che ha il compito di lisare gli involucri nucleari e liberare il DNA dei leucociti. A questo punto si susseguono due incubazioni: la prima a 50-60°C per un'ora, per favorire l'azione della proteinasi K, l'altra per 10 minuti a 95°C per inattivare la stessa.
- Infine le provette vengono centrifugate per 10 sec ed il surnatante contenente il DNA è trasferito in un'epENDORF pulita.
- Il DNA è conservato a – 20°C fino al suo utilizzo.

### Quantificazione del DNA:

Il DNA estratto è stato quantificato mediante spettrofotometro Thermo Scientific Nanodrop, un sistema che permette di quantificare la presenza di acidi nucleici ed eventuali proteine nel campione in esame. Il Nanodrop sfrutta l'equazione di Lambert-Beer che correla l'assorbanza alla concentrazione  $A = \epsilon \cdot b \cdot c$   $A$  = assorbanza  $\epsilon$  = coefficiente di assorbimento  $b$  = cammino ottico in cm  $c$  = concentrazione dell'analita in molarità (M). La lettura spettrofotometrica degli acidi nucleici permette di misurare attraverso questa equazione concentrazione e qualità di un estratto. Gli acidi nucleici assorbono alla lunghezza d'onda di 260 nm ed il rapporto di assorbanza tra 260 e 280 nm, fornisce invece indicazioni sul grado di contaminazione proteica che deve risultare compreso tra 1.7 e 2.

### **Scelta degli SNP**

I marcatori scelti per i vari studi nell'ambito del mio progetto di Dottorato, sono stati selezionati sia in base ai dati presenti in letteratura, sia utilizzando programmi bioinformatici al fine di scegliere i polimorfismi più informativi che mappano le regioni cromosomiche d'interesse.

### **Genotipizzazione**

Ciascun polimorfismo scelto è stato amplificato mediante PCR utilizzando primer fiancheggiati la sequenza di interesse ed i prodotti di amplificazione sono stati sottoposti a digestione enzimatica con una specifica endonucleasi. La tabella 2 riassume i dettagli dei polimorfismi indagati. I frammenti di restrizione sono stati separati in SDS-PAGE nativa al 10% e visualizzati mediante colorazione al bromuro di etidio, un intercalante del DNA. La separazione dei frammenti mediante elettroforesi su gel di poliacrilammide, ci ha permesso di attribuire gli specifici genotipi ad ogni campione sottoposto ad analisi. Per valutare l'accuratezza delle genotipizzazioni, un terzo dei campioni presi casualmente sono stati testati in cieco da un secondo operatore. Sui dati ottenuti dalla genotipizzazione di ciascuno SNP abbiamo quindi effettuato le analisi di associazione caso-controllo.

**Tabella 2.** Specifiche sugli SNP investigati.

Gene e Polimorfismo ID	Posizione <sup>a</sup> SNP	Cambio Nucleotidico	Sequenza Primer (5'-3')	pb	Enzima di Restrizione
<i>CDH1</i> rs11865026	introne 2 chr16:68774033	T>C	Fw- TCCAGTCTTTGACCACATGAA Rev- TATCCAATCCCCAGCACTCT	228	<i>Bsr I</i>
<i>MDR1</i> rs1128503	esone 12 chr7:87550285	C>T	Fw- TTCACTTCAGTTACCCATCTCG Rev-CATCTCACCATCCCCTCTGT	237	<i>Hae III</i>
<i>MDR1</i> rs2032582	esone 21 chr7:87531302	G>T/A	Fw-TGTTGTCTGGACAAGCACTGA Rev-TGTGAAGACAATGGCCTGAA	298	<i>Gsa I</i>
<i>MDR1</i> rs1045642	esone 26 chr7:87509329	C>T	Fw-GGCTCCGAGCACACCTGGGCA Rev-GGCCAGAGAGGCTGCCACATGCT	220	<i>Dpn II</i>
<i>ROCK1</i> rs35996865	5'-UTR chr18:21112383	T>G	Fw-TACCAGCATTGAGCACCAGT Rev-TTGACTGTATCTTCCGGAGTCA	189	<i>BstX I</i>
<i>ROCK1</i> rs73963110	introne 1 chr18:21098352	T>C	Fw-AGTGGCATTGTCAACCAACA Rev-TTCCCCTATGTCCAGATCA	212	<i>Fnu4HI</i>
<i>ROCK1</i> rs2127958	introne 1 chr18:21073648	C>T	Fw-GAGTCAGGTGAAAACATCATTAGG Rev-GGGTGGTGTCTTCCAAAAAT	190	<i>Dra I</i>
<i>ROCK1</i> rs288980	introne 9 chr18:21029619	C>T	Fw-CCCACCCCAATATCTCCATA Rev-GGAAGACTTGGTCCACATCAA	204	<i>Fat I</i>
<i>ROCK2</i> rs12692437	introne 1 chr2:11340435	T>C	Fw-CTTACATCAACAGTAAAgTCCA* Rev-TTCTTTATGAAGAGTTTGGCATT	151	<i>Hpy166 II</i>
<i>ROCK2</i> rs7563468	introne 1 chr2:11291149	G>A	Fw-AGCCAAACACCCAGTTTCC Rev-CACCTGGCCAACAACAATAA	192	<i>Dpn II</i>
<i>ROCK2</i> rs35768389	esone 16 chr2:11214974	T>A	Fw-TAACTTGGCAGTCTCCAGCA Rev-TGTTGTTTGGTTTTGAATGGA	181	<i>Dpn II</i>
<i>ROCK2</i> rs17463896	introne 17 chr2:11213800	A>G	Fw-TGCTCATGCTGTCTCCCTA Rev-ATGCAGTCCTTGCCCTGAAT	217	<i>Fat I</i>
<i>ROCK2</i> rs16857265	introne 32 chr2:11190286	T>C	Fw-TGCAGATGTACAACAAAATCCA Rev-TTTTGTATAGCAACGGTcAGT**	194	<i>BtsIMut I</i>

\*\*Inserzione di mutazione A&gt;C

\*Inserzione di mutazione T&gt;G

<sup>a</sup>UCSC Genome Browser assembly ID: hg38

## **Reazione di amplificazione mediante PCR**

Ciascuna sequenza contenente il polimorfismo (rinvenibile presso la banca dati del sito <http://www.ncbi.nlm.nih.gov/snp/>) è stata amplificata mediante la reazione a catena della polimerasi (PCR) utilizzando coppie di primer specifici complementari alle sequenze nell'intorno di ciascun marcatore, individuati con l'ausilio del programma Primer3 (<http://bioinfo.ut.ee/primer3-0.4/primer3/>).

Le condizioni adottate per la reazione di amplificazione di ciascun marcatore prevedono i seguenti step:

- denaturazione iniziale: 94°C per 5 minuti
- denaturazione: 94°C per 30 secondi
- annealing: 55°C (variabile a seconda della sequenza indagata) per 30 secondi
- estensione: 72°C per 30 secondi
- step di estensione finale: 72°C per 8 minuti.

Il loop centrale di denaturazione, annealing ed estensione è stato di 35 cicli.

La reazione è avvenuta in un volume finale di 12.5 µl contenente: 2 µl di DNA (pari a circa 100 ng), 1.25 µl di dNTPs 10 mM, 0.2 µl di primer senso 50 µM, 0.2 µl di primer antisenso 50 µM, 0.05 µl di Taq polimerasi (pari a 0.5 Unità), 0.3 µl di MgCl<sub>2</sub> 50 mM e 7.25 µl di H<sub>2</sub>O ultrapura.

La corretta resa della reazione è stata visualizzata mediante elettroforesi su gel di agarosio all'1% in TAE (tris acetato - EDTA) 0.5 X addizionato con bromuro d'etidio, utilizzando 5µl di prodotto di PCR.

## **Digestione enzimatica**

Il prodotto di PCR contenente il polimorfismo d'interesse è stato digerito mediante enzimi di restrizione.

La miscela di reazione per ciascuna specifica reazione è stata realizzata in un volume di 10 µl: 3 µl di prodotto di PCR, 1 µl di Buffer specifico 10 X, 0.1µl (10.000 U/ml) di enzima,

5.9  $\mu\text{l}$  di ddH<sub>2</sub>O per un volume finale di 10  $\mu\text{l}$ ; per alcuni enzimi di restrizione, la miscela è stata addizionata con 0.01  $\mu\text{l}$  di BSA 100 X.

La soluzione è stata immersa in un bagno termico per circa due ore alla temperatura specifica per la massima attività di ciascun enzima.

I prodotti di digestione sono stati visualizzati mediante elettroforesi verticale su gel di poliacrilamide al 10%.

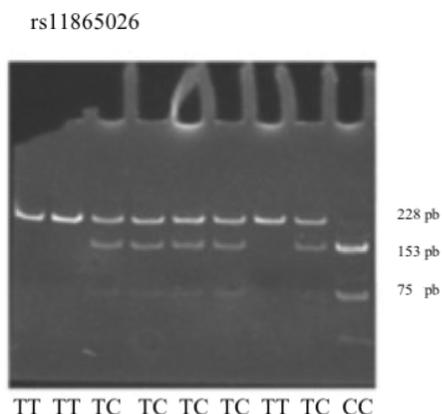
Il gel fatto correre nel mini-protean conteneva: 3.33  $\mu\text{l}$  di una miscela acrilamide/bisacrilamide 29:1 al 30%, 1ml di TBE (tris borato - EDTA) 10 X, 5.55 ml di ddH<sub>2</sub>O, 105  $\mu\text{l}$  di APS (ammonio persolfato) 10%, 6.5  $\mu\text{l}$  TEMED (tetrametilendiammina), per un volume finale di 10 ml (utile per 2 gel).

Dopo il tempo di incubazione della digestione i campioni sono stati mescolati con 2  $\mu\text{l}$  di Loading Buffer 6 X e caricati nei pozzetti del gel. All'interno della camera di corsa è stata versata una soluzione di TBE 1 X. La corsa elettroforetica è stata condotta per circa 1 ora e 30 minuti ad un voltaggio di 85 mV costanti (le tempistiche di corsa possono variare a seconda della lunghezza dei frammenti da separare).

In seguito, i gel separati dai vetri, sono stati posti in agitazione in vaschette contenenti una soluzione tampone TBE 1 X in presenza di bromuro di etidio per una decina di minuti; infine osservati al transilluminatore ai raggi UV.

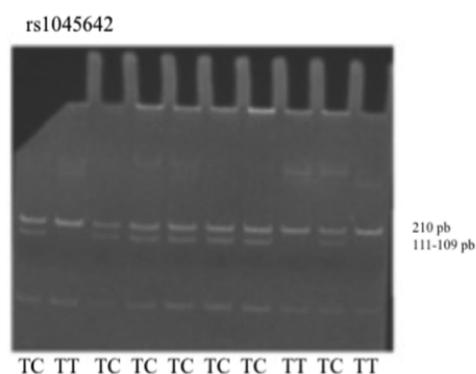
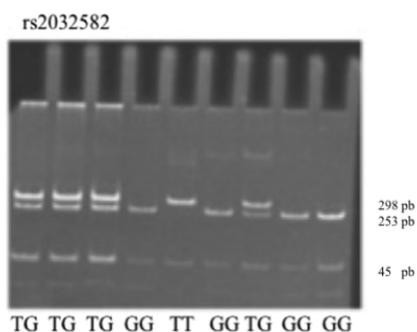
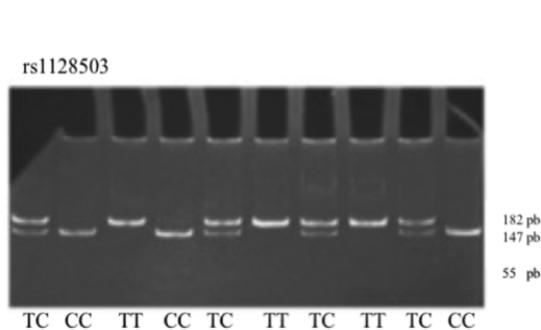
I pattern elettroforetici osservati per ciascun polimorfismo indagato erano quelli indicati nelle figure 2A, 2B, 2C e 2D rispettivamente per gli SNP di *CDHI*, *ABCBI*, *ROCK1* e *ROCK2*.

**Figura 2A.** Per il polimorfismo rs11865026 indagato per *CDHI* presente nell'introne 2: un frammento di 228 pb per il genotipo omozigote ancestrale TT; frammenti di 153 e 75 pb per il genotipo CC e frammenti di 228, 153, 75 pb per l'eterozigote TC.



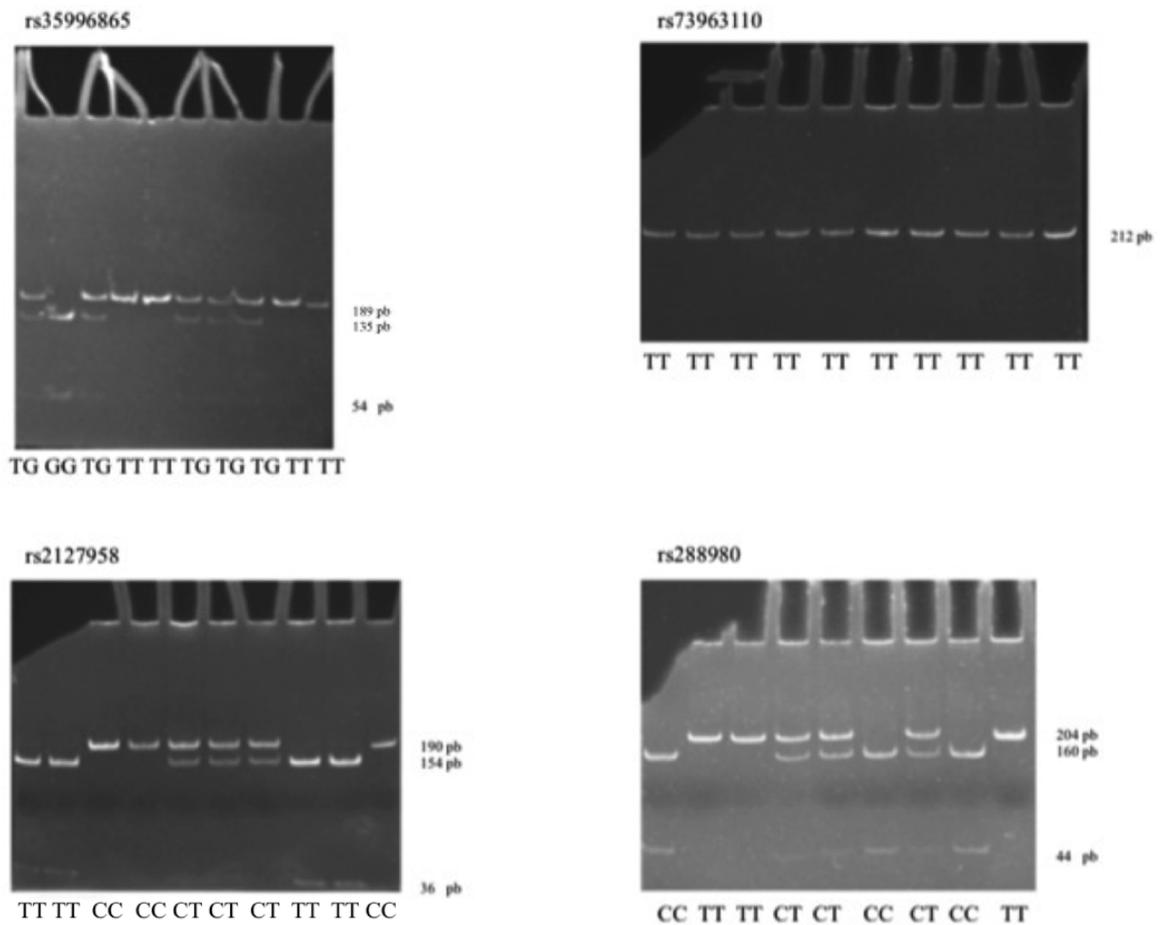
**Figura 2B.** I 3 SNPs genotipizzati per *ABCB1*

- per rs1128503 presente nell'esone 12: frammenti di 182 e 55 pb per il genotipo TT; frammenti di 147, 55 e 35 pb per il genotipo CC e frammenti di 182, 147, 55 e 35 pb per l'eterozigote TC.
- per rs2032582 presente nell'esone 21: un frammento di 298 per il genotipo TT; frammenti di 253 e 45 pb per il genotipo GG e frammenti di 298, 253 e 45 pb per l'eterozigote TG.
- per rs1045642 presente nell'esone 26: un frammento di 220 per il genotipo TT; frammenti di 111 e 109 pb per il genotipo CC e frammenti di 220, 111 e 109 pb per l'eterozigote TC.



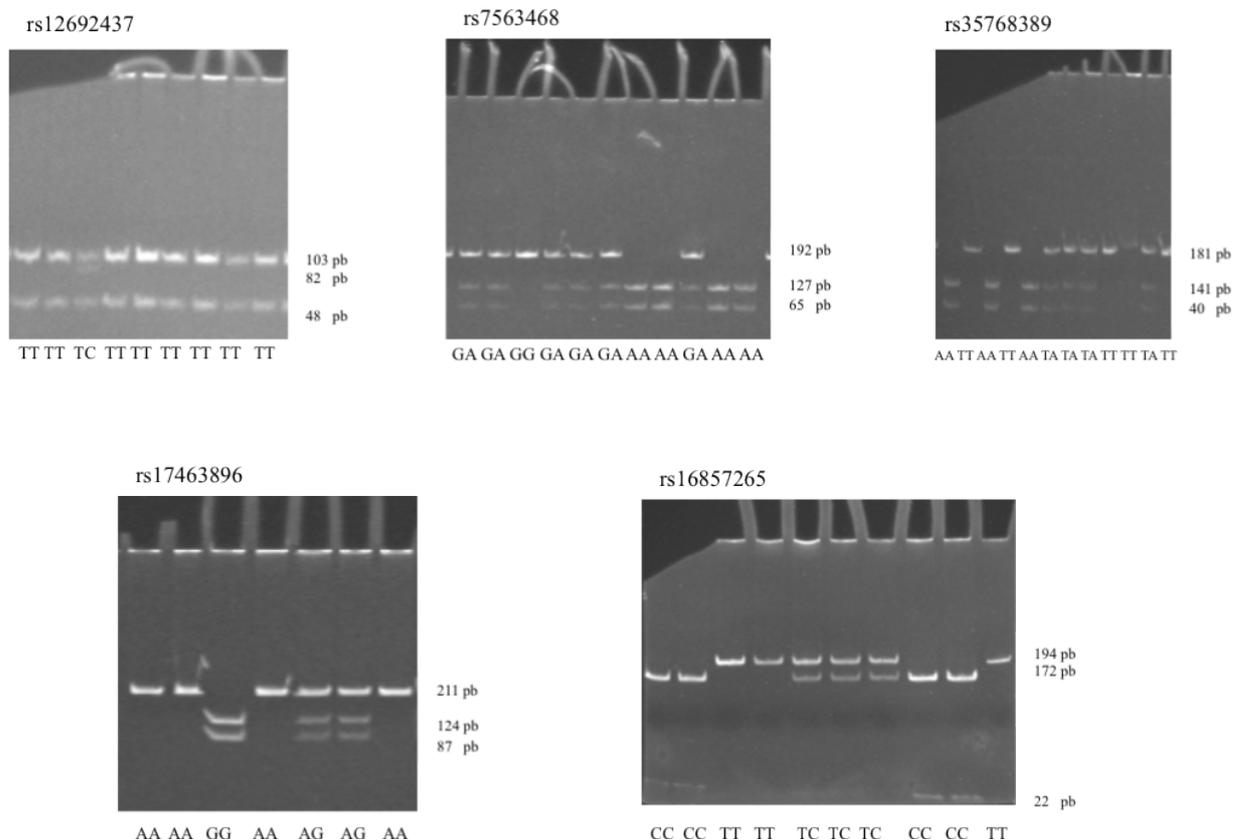
**Figura 2C.** I 4 SNPs indagati per *ROCK1*

- per rs35996865 presente nella regione 5'-UTR: un frammento di 189 per il genotipo omozigote ancestrale TT; frammenti di 135 e 54 pb per il genotipo GG e frammenti di 189, 135, 54 pb per l'eterozigote TG.
- per rs73963110 presente nell'introne 1: un frammento di 212 per il genotipo omozigote ancestrale TT; frammenti di 167 e 45 pb per il genotipo CC e frammenti di 212, 167 e 45 pb per l'eterozigote TC.
- per rs2127958 presente nell'introne 1: un frammento di 190 per il genotipo omozigote ancestrale CC; frammenti di 154 e 36 pb per il genotipo TT e frammenti di 190, 154 e 36 pb per l'eterozigote CT.
- per rs288980 presente nell'introne 9: un frammento di 160 e 44 per il genotipo omozigote ancestrale CC; un frammento di 204 pb per il genotipo TT e frammenti di 204, 160 e 44 pb per l'eterozigote CT.



**Figura 2D.** I 5 SNPs studiati per *ROCK2*

- per rs12692437 presente nell'introne 1: frammenti di 103 e 48 per il genotipo omozigote ancestrale TT; frammenti di 82, 48 e 21 pb per il genotipo CC e frammenti di 103, 82, 48 e 21 pb per l'eterozigote TC.
- per rs7563468 presente nell'introne 1: un frammento di 192 per il genotipo omozigote ancestrale GG; frammenti di 127 e 65 pb per il genotipo AA e frammenti di 192, 127, 65 pb per l'eterozigote GA.
- per rs35768389 presente nell'esone 16: frammenti di 141 e 40 pb per il genotipo omozigote ancestrale TT; un frammento di 181 pb per il genotipo AA e frammenti di 181, 141, 40 pb per l'eterozigote TA.
- per rs17463896 presente nell'introne 17: frammenti di 211 e 6 per il genotipo omozigote ancestrale AA; frammenti di 124, 87 e 6 pb per il genotipo omozigote GG e frammenti di 211, 124, 87, 6 pb per l'eterozigote AG.
- per rs16857265 presente nell'introne 32: un frammento di 194 per il genotipo omozigote ancestrale TT; frammenti di 172 e 22 pb per il genotipo CC e frammenti di 194, 172, 22 pb per l'eterozigote TC.



## **Isolamento di RNA e sintesi di cDNA per la valutazione di espressione genica dei geni CTNNB1, LEF1 e CDH1**

Ogni biopsia prelevata dal tessuto tumorale e dal tessuto normale è stata congelata in azoto liquido e conservata a -80 ° C fino al momento dell'estrazione dell'RNA.

Un'aliquota di circa 100 mg di ogni campione è stata omogeneizzata in 1 ml di Tri Reagent® Solution (Ambion Inc., Austin, TX, USA) utilizzando un omogeneizzatore di tessuto ed incubata a temperatura ambiente per 5 minuti. 0.2 ml di cloroformio sono stati poi aggiunti a ciascun campione. Dopo la centrifugazione per 15 minuti a 12000 g, la fase acquosa contenente l'RNA totale è stata trasferita in una nuova provetta e miscelata con 0,5 ml di isopropanolo. I campioni sono stati incubati a temperatura ambiente per altri 10-15 minuti e poi centrifugati a 12000 g per 10'. Il pellet di RNA ottenuto, è stato lavato con 1 ml di etanolo al 75%. Dopo centrifugazione a 7500 g per 5 ', l'etanolo è stato poi rimosso ed il pellet di RNA lasciato asciugare brevemente all'aria, infine risospeso in 40 µl di acqua libera da RNasi.

La purezza e la concentrazione dell'RNA recuperato sono state determinate mediante lo spettrofotometro NanoDrop 2000 (Thermo Fisher Scientific, Wilmington, Delaware, USA). Di ciascun campione sono state retroscritte a cDNA aliquote di 1 µg di RNA totale (tessuti tumorali e tessuti di controllo) utilizzando il kit SuperScript VILO cDNA Synthesis (Invitrogen, Carlsbad, CA, USA).

### **Real time qPCR**

L'espressione genica è stata valutata mediante Real-Time PCR quantitativa, utilizzando specifici primers disponibili in commercio (Sigma Aldrich, Inc., St. Louis, Mo, USA). Le sequenze dei primer per i geni selezionati sono elencate nella tabella 2.

Sono stati valutati i livelli di espressione genica di tre geni: *CTNNB1*, *LEF1* e *CDH1*, mentre come gene di riferimento abbiamo utilizzato *GAPDH* scelto in seguito a test preliminari eseguiti con tre geni costitutivi. I quattro saggi effettuati hanno mostrato

un'efficienza di reazione ottimale (tra il 98% e il 102%) ed una corretta linearità delle curve seriali su tutta la serie delle 8 diluizioni logaritmiche.

Le reazioni di PCR sono state effettuate in un volume finale di 20 µl contenente 10 µl 2X di Power SYBR® Green PCR Master Mix (Applied Biosystems, Foster City, CA, USA), 400 nM della concentrazione di ciascun primer e 300 nM di cDNA. Le amplificazioni sono state condotte utilizzando lo strumento ABI PRISM 7500 (Applied Biosystems, Foster City, CA, USA). Dopo una fase di denaturazione iniziale a 95 °C per 10 min, l'amplificazione procedeva con 40 cicli in due fasi di 15 s a 95 °C e 60 s a 60 °C. Come passo finale, è stata eseguita una analisi della curva Melting di dissociazione. Tutti gli esperimenti sono stati eseguiti utilizzando dei controlli negativi privi di template per verificare la presenza di eventuale contaminazione.

Dopo la normalizzazione con il gene di riferimento, i livelli di espressione genica di ciascun marcatore nel tessuto tumorale sono stati confrontati con quelli del tessuto normale mediante il metodo del delta / delta Ct [65].

### **Analisi statistica**

I dati ottenuti dalla genotipizzazione dei polimorfismi sono stati elaborati mediante analisi statistica. La distribuzione dei genotipi negli individui affetti da patologia e nel gruppo di controllo è stata testata per la deviazione dall'Equilibrio di Hardy-Weinberg utilizzando il test del  $\chi^2$  di Pearson.

In primo luogo l'associazione genetica è stata studiata in merito al rapporto di verosimiglianza o probabilità impiegando il Software Unphased V3.15 su sistema operativo Windows Vista. Un test specifico è stato poi effettuato per ciascun aplotipo. Si è trattato di un test a punteggio per una differenza di rischio tra un aplotipo e tutti gli altri insieme, diversamente dai rapporti di probabilità, che sono stati calcolati in relazione ad un unico aplotipo di riferimento. Un test di permutazione è stato poi eseguito con 10.000 repliche del campione, per consentire correzioni multiple del test su tutti i test effettuati. In ogni

replicato i valori sono stati mescolati a caso tra tutti i soggetti. In questa analisi abbiamo calcolato i rapporti di probabilità valutando il livello di associazione dei portatori di alleli rari sia in caso di modello dominante che recessivo. La regressione logistica è stata infine utilizzata per le analisi multivariate utilizzando il software SPSS.

L'espressione genica di *CTNNB1*, *LEF1* e *CDH1* è stata studiata su biopsie di 54 campioni provenienti da 27 pazienti (per ciascuno una componente tumorale ed una di tessuto normale). La normalizzazione dei livelli di espressione dei geni nel tessuto tumorale e normale è stata effettuata mediante il T test.

## **II – Parte**

### **Ricerca in database e selezione dei marcatori**

L'archivio di genomica funzionale "Espressione Genica Omnibus" (GEO), è stato utilizzato per la ricerca dell'insieme di dati presenti nel database di esperimenti basati sulla tecnologia dei microarrays. Sono stati utilizzati per la ricerca le seguenti parole chiave: "Cancro del colon, cancro del retto e Homo sapiens [organismo]". Inoltre, la ricerca delle parole chiave "Sangue intero [campione di ricerca] ed Homo sapiens [organismo]" sono state utilizzate per recuperare l'insieme di dati nello stesso database, generando quindi un gruppo che serviva come un insieme per evidenziare l'espressione genica specifica per CCR rispetto al confronto differenziale sulle cellule del sangue normale.

I risultati della ricerca sono stati poi filtrati utilizzando i seguenti criteri.

I criteri di inclusione dell'insieme dei dati nell'analisi sono stati:

- esperimenti del tipo "Profili di espressione mediante microarray "
- tessuti primari
- qualsiasi età o sesso del soggetto
- il CCR in ogni fase
- sangue periferico non frazionato
- cellule del sangue non stimulate

-disponibilità dei dati grezzi o dati pre-elaborati per il singolo canale del tessuto di interesse nel caso di esperimenti con due canali.

I criteri di esclusione sono stati:

-campioni di linee cellulari

-tessuti di adenoma non maligni

-cellule metastatiche

-pazienti in trattamento

-sangue del cordone ombelicale

-frazioni del sangue (ad esempio cellule mononucleate del sangue periferico o campioni di sangue sottoposti a riduzione della globina mRNA)

-sangue trattato non-sano o di soggetti pediatrici

-esoni array (ostacolano l'elaborazione dei dati di TRAM a causa di un elevato numero di righe di dati) o piattaforme con sonde suddivise in distinti array per ogni campione (ostacolano la normalizzazione nel campione)

-piattaforme che analizzano un numero atipico di geni (cioè  $<5.000$  o  $> 60.000$ )

Per ottenere una mappa quantitativa del trascrittoma, i valori di ciascun gruppo di dati sono stati linearizzati quando forniti come logaritmi.

Infine, a causa della tempistica del processo di elaborazione TRAM, per ogni serie di GEO abbiamo preso le prime dieci liste di campioni scelti a caso ed i primi 10 campioni elencati quando la dimensione del campione era maggiore di 10 per un totale di 349 campioni con CCR su 2.532 trovati e 200 campioni di sangue di soggetti normali su 958 trovati.

I 349 campioni con CCR presi in considerazione per la nostra meta-analisi sono stati selezionati in modo casuale da 37 studi condotti in diversi paesi (Australia, Cina, Danimarca, Finlandia, Germania, Ungheria, Italia, Giappone, Norvegia, Polonia, Singapore, Corea del Sud, Spagna, Stati Uniti d'America). I dati clinici non erano completamente disponibili per tutti i soggetti e la fascia d'età comprendeva pazienti tra i 42 e gli 88 anni.

I maschi e le femmine sono stati equamente rappresentati e sono stati inclusi tutti i siti del colon e le fasi Dukes.

I 200 campioni di sangue non patologico scelti sono stati ottenuti da 23 studi anche condotti in diversi paesi (Australia, Cina, Danimarca, Germania, India, Italia, Paesi Bassi, Norvegia, Regno Unito, Stati Uniti d'America) su soggetti sani, senza storia clinica di pregressa malattia neoplastica.

### **Analisi TRAM (Transcriptome Mapper)**

Il software TRAM (Transcriptome Mapper) consente l'importazione dei dati di espressione genica registrati nella banca dati NCBI GEO in formato di testo delimitato da tabulazioni. Consente inoltre l'integrazione di tutti i dati decodificando le sonde identificate con simboli corrispondenti a geni tramite i dati presenti nell'analisi UniGene, normalizzando i dati provenienti da più piattaforme utilizzando una normalizzazione intra-campione e inter-campione (normalizzazione in scala quantile).

Abbiamo creato una directory (cartella) per ogni tessuto, contenente tutti i dataset di esempio relativi alla stessa fonte (gruppo 'A' pazienti con CCR, n = 349; gruppo 'B' sangue di soggetti normali, n = 200) e selezionati per lo studio.

Il confronto ci ha permesso di analizzare le mappe del trascrittoma differenziale di 'A' contro il 'B', utilizzando il rapporto tra i valori di espressione medi per ciascun locus.

Abbiamo utilizzato una versione aggiornata di TRAM, per cui una maggiore risoluzione dei geni identificati e database aggiornati di Entrez Gene, Unigene e TRAM 1.1 (giugno 2013, disponibile gratuitamente a <http://apollo11.isto.unibo.it/software>, invece della versione originale del 2011). Sono stati caricati i dati relativi a 10 piattaforme (GPL80, GPL1449, GPL2006, GPL2895, GPL3121, GPL4811, GPL6370, GPL6883, GPL8432, GPL10665). In breve, i valori di espressione genica sono stati assegnati a loci individuali via UniGene, intra-campione, normalizzati come percentuale del valore medio e inter-campione, normalizzati in scala quantile.

## **Estrazione di RNA totale con Trizol LS**

Il sangue intero raccolto in provetta contenente EDTA come anticoagulante è stato trattato per la lisi con TRIzol LS contenente fenolo e guanidina isotiocianato che è un inibitore dell'RNAsi (Invitrogen, Carlsbad, CA, USA), entro un'ora dal momento del prelievo. L'RNA totale è stato estratto da 1mL di sangue secondo il protocollo del produttore.

E' stato aggiunto cloroformio ad ogni provetta, agitando manualmente per 15'.

Dopo un'incubazione di 15 min a T.A. è stata eseguita una centrifugazione a 12.000 rcf per 15 min a 4°C così da ottenere una fase acquosa ed una fase organica separate da un'interfaccia. E' stata prelevata la fase acquosa contenente l'RNA, alla quale è stato aggiunto un pari volume di isopropanolo che ne determina la precipitazione. Dopo 10 minuti di incubazione a T.A., è stata eseguita una seconda centrifugazione di 12.000 rcf per 10 min a 4°C; eliminato il surnatante, il "pellet" è stato lavato facendo due passaggi in etanolo al 75% seguiti da una centrifugazione a 7.600 rcf per 20 min a 4°C.

Dopo i lavaggi, il "pellet" di RNA ottenuto è stato fatto asciugare brevemente a temperatura ambiente, in modo da far evaporare tutto l'alcool residuo, potente inibitore della retrotrascrizione. Infine il pellet di RNA è stato risospeso in 15 microlitri di acqua DEPC RNasi-free, per una concentrazione finale di circa 0,5µg/µl e conservato a -20°C.

La concentrazione di tutti i campioni di RNA è stata quantificata allo spettrofotometro Nanodrop ND-2000 (Thermo Fisher Scientific, Waltham, MA).

### **Progettazione dei primer**

La progettazione dei primer è stata effettuata secondo i parametri suggeriti da Sharrocks (1994), con l'ausilio del programma Primer 3 e Amplify per elaboratori Macintosh.

Di ciascun primer è stata successivamente verificata la specificità, con il confronto diretto mediante l'algoritmo BLAST, con le sequenze dei geni, delle EST umane e dei frammenti genomici contenuti nelle banche dati dell'NCBI. Gli oligonucleotidi progettati, sono stati prenotati mediante un servizio commerciale di sintesi (SIGMA ALDRICH, Milano, Italia), come prodotti liofilizzati. La sequenza di ciascun primer utilizzato ed il gene di riferimento è riportata nella tabella 3. I prodotti liofilizzati sono stati risospesi in acqua bidistillata sterile alla concentrazione di 100 mM. I primer sono stati utilizzati nei diversi esperimenti sempre alla concentrazione finale di 10  $\mu$ M.

**Tabella 3.** Primers utilizzati per la qRT-PCR

Simbolo del Gene	Nome del Gene	RifSeq RNA GenBank Accession No.	Sequenza coppie di Primer (5' → 3') <sup>a</sup>	RT-PCR (pb)
<i>TSPAN8</i>	Tetraspanina 8	NM_004616	ctaagtctgatcgattgtg aacacaattatggcttcctg	103
<i>EPCAM</i>	Molecola di Adesione delle Cellule Epiteliali	NM_002354	gtatgagaaggctgagataaag cttcaaagatgtctctgccc	147
<i>SPINK1</i>	Inibitore Serina Peptidasi , Kazal tipo 1	NM_003122	tgaaaatcggaacgccagac gcggtgacctgatgggattt	89
<i>COL1A2</i>	Collagene, tipo I, alfa 2	NM_000089	gtggttactactggattgac ctgccagcattgatagtttc	184
<i>CDH1</i>	Homo sapiens caderina 1, tipo 1, E-caderina	NM_004360	cagtacaacgaccaacca cacgctgacctctaaggtgg	99
<i>LGALS4</i>	Lectina, legame al Galattoside, Solubile, 4	NM_006149	ttaccctgggtcccggacatt agcctcccgaatatggcac	135
<i>CEACAM6</i>	Antigene Carcinoembrionario collegato alle Molecole di Adesione 6	NM_002483	cacagtctctggaagtgtccc ggccagcactccaatcgt	99
<i>B2M</i>	Beta-2-microglobulina	NM_004048	tgcttgcctgtgaacctgt tgcggcatcttcaaacctccatga	97

<sup>a</sup> Per ogni gene in alto: primer forward; in basso: primer reverse

### Retrotrascrizione (RT)

300 ng di RNA totale sono stati retrotrascritti con il kit RevertAid First Strand cDNA Synthesis (Carlo Erba Reagenti, Milano, Italia), secondo le istruzioni del produttore, in un volume finale di 20 µL in presenza di 10 mM di desossiribonucleotidi trifosfato (dNTP), 100 µM di primer esameri casuali, 200 U di enzima trascrittasi inversa e relativo tampone

di retrotrascrizione 5X e 20U/ $\mu$ l dell'enzima inibitore dell'RNasi. La reazione di retrotrascrizione ha previsto un'incubazione a 42°C per 1 ora, preceduta da un'incubazione di 5 min a 25°C per permettere la sintesi dei primer esameri casuali, di seguito la miscela è passata a 70°C per 5 minuti per porre termine alla reazione. Come controllo negativo è stato incluso un campione d'acqua in tutte le retrotrascrizioni.

### **Real Time qPCR**

Le reazioni di amplificazione sono state ottenute utilizzando il sistema EvaGreen (Bio-Rad, Hercules, CA), secondo le istruzioni del produttore.

Le reazioni PCR in tempo reale sono state eseguite utilizzando lo strumento CFX96 (Bio-Rad, Hercules, CA), realizzate su piastre ottiche da 96 pozzetti (Hard-Shell PCR Plates 96-Well Low-Profile, Bio-Rad) e sigillate con il film ottico adesivo (Microseal,B, Bio-Rad).

Le reazioni eseguite in duplicato, prevedevano un primo step a 95 °C per 10 minuti, seguito da 40 cicli di 95 °C per 15 sec e 60 °C per 1 min (rispettivamente denaturazione e annealing), con analisi della curva di fusione. Ogni ciclo qPCR comprendeva sempre un controllo negativo privo di cDNA, e un controllo positivo di cDNA derivato dalla linea cellulare HT-29, dove i geni di interesse sono notoriamente presenti.

L'efficienza di reazione (E) è stata calcolata dalla pendenza della curva standard generata da diluizioni seriali di 10 volte del calibratore cDNA, secondo la formula:

$$E = [10^{-1 / \text{slope}} - 1] \times 100.$$

I cicli soglia ottenuti in seguito all'impostazione manuale del threshold, sono stati elaborati con excel. Per l'analisi dei dati abbiamo utilizzato il  $\Delta$ CT, che valuta le differenze per ciascun campione tra la media dei replicati cicli soglia (Ct) del gene bersaglio e la media dei replicati cicli soglia (Ct) del gene di riferimento. L'errore è stato valutato come deviazione standard dalla media di tutti i replicati.

## **Analisi statistiche**

Il T-Test Student è stato adottato per il confronto dei livelli di espressione analizzati tra i casi CCR ed i controlli.

L'analisi della curva ROC (Receiver Operating Characteristic) è stata utilizzata per ciascun marcatore per valutare l'accuratezza di un test diagnostico utilizzando tutto il range dei valori ottenuti  $\Delta C_t$  dei pazienti CCR e dei controlli. La curva ROC ci ha permesso anche di identificare il valore soglia ottimale per ciascun marcatore.

Abbiamo calcolato sia l'area sotto la curva che i parametri di sensibilità (la proporzione di veri positivi) e specificità (la proporzione di veri negativi), con corrispondenti intervalli di confidenza al 95% utilizzando la versione MedCalc 16 per le analisi statistiche.

L'area sotto la curva delle ROC può assumere valori compresi tra 0.5 e 1; tanto più la curva si avvicina ad 1, tanto maggiore risulta il potere discriminante del test.

Per l'interpretazione dei valori ottenuti dell'area sotto la curva "AUC", possiamo riferirci alla classificazione di Swets (J. A. Swets e D. M. Green, 1966):

-test non informativo  $AUC=0.5$

-test poco accurato  $0.5 < AUC \leq 0.7$

-test accurato  $0.7 < AUC \leq 0.9$

-test altamente accurato  $0.9 < AUC \leq 1$

-test perfetto  $AUC=1$

In seguito abbiamo effettuato un'analisi discriminante per i marcatori selezionati utilizzando il software statistico SPSS, versione 23. Il gruppo di soggetti sani ed il gruppo dei pazienti sono stati inseriti in una variabile identificativa delle rispettive diagnosi.

I quattro marcatori indipendenti sono stati raggruppati costruendo un'unica variabile. Infine è stata costruita la curva ROC per il pannello dei 4 marcatori e per tutte le altre combinazioni.

# *Risultati*

## Studio dell'associazione genetica tra CCR ed i polimorfismi selezionati sul gene *CDHI*

### Analisi statistica

Per studiare l'associazione genetica tra CCR e quattro polimorfismi selezionati sul gene *CDHI*, è stata eseguita una serie di test di associazione allelica e di aplotipo. La frequenza dei genotipi osservati nei gruppi dei casi e dei controlli è risultata coerente con la legge dell'equilibrio di Hardy-Weinberg (Tabella 4).

**Tabella 4.** Genotipi osservati per i polimorfismi studiati di *CDHI* in pazienti e controlli

dbSNP ID <sup>a</sup>	Posizione genoma <sup>b</sup>	Alleli (A/a) <sup>c</sup>	Gruppi	AA	Aa	aa	MAF <sup>d</sup>	HWE <i>P</i> value <sup>e</sup>
rs11865026	chr16:68774033	T/C	casi	64	55	19	0.34	0.20
			controlli	64	58	19	0.34	0.32
rs11642413	chr16:68790144	A/G	casi	38	64	36	0.49	0.40
			controlli	41	64	36	0.48	0.28
rs10431923	chr16:68839013	G/T	casi	35	72	31	0.49	0.60
			controlli	32	72	37	0.52	0.79
rs13689	chr16:68868272	T/C	casi	99	31	3	0.14	0.42
			controlli	108	30	3	0.13	0.60

<sup>a</sup>NCBI-SNPs data base

<sup>b</sup>UCSC Genome Browser su Human February 2009 assembly

<sup>c</sup>Allele maggiore indicato per primo

<sup>d</sup>Frequenza dell'allele minore

<sup>e</sup>Test del *P*value per verificare l'Equilibrio di Hardy-Weinberg

Le stime delle frequenze aplotipiche indicano un elevato livello di linkage disequilibrium tra i polimorfismi rs11865026, rs11642413 e rs10431923, mentre l' rs13689 è apparso debolmente in linkage ed è stato quindi escluso dalla successiva analisi di associazione aplotipica (Tabella 5).

**Tabella 5.** Linkage disequilibrium tra SNP di *CDHI*.

	rs11865026	rs11642413	rs10431923	rs13689
rs11865026		0.718	0.678	0.399
rs11642413	0.276		0.688	0.534
rs10431923	0.241	0.46		0.402
rs13689	0.013	0.043	0.025	

D' al di sopra della diagonale;  $r^2$  al di sotto della diagonale

La tabella 6 riporta il *P* value di associazione globale ottenuto in 7 test, 4 per ogni SNP e 3 per diverse combinazioni di aplotipi. Un livello nominale di significatività è stato ottenuto per l'associazione tra CCR e le combinazioni di aplotipo rs11642413-rs10431923 (totale *P* value = 0.0063).

Allo scopo di adattare le correzioni sui molteplici test effettuati in questo studio ( $n = 7$ ) sono state testate oltre 50.000 permutazioni. Il valore empirico corretto del *P* value per l'associazione con la combinazione di aplotipo rs11642413-rs10431923 è risultata 0.048.

**Tabella 6.** Associazione tra polimorfismi di *CDHI* e CCR *P* values globali dei test sui rapporti di verosimiglianza per associazione con un singolo SNP o combinazioni di SNPs.

SNP o combinazioni di SNPs	<i>P</i> value Globale
rs11865026	0.93
rs11642413	0.80
rs10431923	0.45
rs13689	0.54
rs11865026-rs11642413	0.98
rs11642413-rs10431923	<b>0.0063</b>
rs11865026-rs11642413-rs10431923	0.27

La tabella 7 riporta i livelli di associazione di ciascun aplotipo all'interno della finestra rs11642413-rs10431923. I due aplotipi rari aumentano il rischio di CCR, dal momento che la loro frequenza è apparsa quasi raddoppiata tra i pazienti CCR.

**Tabella 7.** Stime di frequenze di aplotipo ed effetti di aplotipo per i polimorfismi rs11642413 e rs10431923.

Aplotipo	Casi	Controlli	Freq-Ca	Freq-Co	OR (95% C.I.)
AT	104.6	131.3	0.38	0.47	<i>ref.</i>
GG	106.6	121.3	0.39	0.43	1.103 (0.7653-1.59)
GT	32.4	14.7	0.12	0.05	2.763 (1.372-5.565)
AG	28.4	14.7	0.10	0.05	2.422 (1.183-4.958)

#### **Studio dell'espressione genica di *CTNNB1*, *LEF1*, e *CDH1***

La normalizzazione dei livelli di espressione dei geni osservati nel tessuto tumorale e nel tessuto normale sono stati confrontati mediante T test dei campioni accoppiati. I livelli di espressione nei tessuti tumorali sono risultati significativamente più bassi per *CTNNB1* ( $P$  value = 0.006; fold change? = 0.49 [IC 95% 0.80-0.30]) e per *CDH1* ( $P$  value= 0.009; fold change?= 0.45 [IC 95% 0.80-0.25]) , mentre per *LEF1* non c'erano differenze di espressione ( $P$  value = 0.27).

#### **Studio dell'associazione genetica tra CCR ed i polimorfismi selezionati sul gene *MDR1***

##### **Analisi bivariata**

Le frequenze genotipiche dei casi e dei controlli sono risultati nell'equilibrio di Hardy-Weinberg (tabella 8).

Lo studio di associazione è stato condotto per verificare se varianti alleliche in polimorfismi mappanti in *ABCB1* fossero associate sia riguardo il CCR nel suo complesso che riguardo a sottogruppi individuati sulla base del grado istologico riscontrato, della localizzazione del tumore (cieco, ascendente, trasverso e discendente) o del sesso del paziente (tabelle 9, 10, e 11). Nessuna delle varianti alleliche è risultata essere associata al

rischio di insorgenza del CCR sia nel campione in toto, che stratificato per il grado istologico; tuttavia, è stata riscontrata un'associazione significativa rispetto alla localizzazione nel sigma e in particolare le varianti alleliche dei polimorfismi rs1128503 e rs2032582 sono risultate meno frequenti nei pazienti con il tumore al sigma (valori di  $P = 0.006$  e  $0.023$ , rispettivamente). Per quanto riguarda il genere maschile la stratificazione per genere ha dimostrato che lo SNP rs1128503 è associato al CRC nei maschi ( $P = 0.03$ ;  $OR_{Ret} = 0.49$  (95% CI 0.20-1.18),  $OR_{Rom} = 0.31$  (95% CI 0.09-1.00). L'allele variante, meno frequente nei pazienti di sesso maschile (MAF = 0.32) rispetto ai controlli (MAF = 0.44), è risultato più frequente nelle pazienti (MAF = 0.53); tuttavia, la differenza non si è mostrata significativa rispetto ai controlli di sesso femminile ( $P = 0.12$ ). Nessuna differenza nella frequenza allelica tra i generi è stata osservata all'interno del gruppo di controllo.

Il livello di associazione per rs2032582 è apparso più debole; gli OR sono risultati essere:  $OR_{Ret} = 0.49$  (95% CI 0.18-1.32) e  $OR_{Rom} = 0.22$  (95% CI 0.05-1.09).

L'odds ratio (OR) presente in eterozigosi ( $OR_{Ret}$ ) per la variante allelica rs1128503 nel confronto tra malati e sani è risultata 0.26 (95% CI 0.09-0.80), mentre in omozigosi ( $OR_{Rom}$ ) 0.21 (95% CI 0.04-1.03).

**Tabella 8.** Distribuzione genotipica per i polimorfismi studiati del gene *MDR1* e test d'associazione

SNP	Gruppi	Casi				Assoc. <sup>a</sup>	Dominante	Recessivo
		AA	Aa	aa	MAF <sup>b</sup>	<i>P</i> value	OR (95% C.I.)	OR (95% C.I.)
	Controlli	33	45	22	0.45	<i>ref.</i>	<i>ref.</i>	<i>ref.</i>
1236C>T	CCR tot	37	38	21	0.42	0.57	0.79(0.44-1.4)	0.85(0.40-1.8)
rs1128503	CCR maschi	25	21	6	0.32	0.03	0.53(0.27-1.1)	0.36(0.13-1.0)
	CCR femmine	12	17	15	0.53	0.16	1.31(0.60-2.9)	1.87(0.74-4.8)
	Controlli	32	44	24	0.46	<i>ref.</i>	<i>ref.</i>	<i>ref.</i>
2677G>T/A	CCR tot	32	45	21	0.44	0.75	0.97(0.54-1.8)	0.87(0.41-1.9)
rs2032582	CCR maschi	20	26	6	0.37	0.11	0.75(0.37-1.5)	0.40(0.14-1.2)
	CCR femmine	12	19	15	0.53	0.25	1.33(0.61-2.9)	1.67(0.66-4.2)
	Controlli	29	42	29	0.50	<i>ref.</i>	<i>ref.</i>	<i>ref.</i>
3435C>T	CCR tot	28	45	24	0.48	0.68	1.01(0.54-1.9)	0.86(0.41-1.8)
rs1045642	CCR maschi	17	26	8	0.41	0.15	0.82(0.40-1.7)	0.47(0.18-1.3)
	CCR femmine	11	19	16	0.55	0.39	1.30(0.58-2.9)	1.46(0.58-3.67)

<sup>a</sup>Test per differenze di frequenze alleliche; <sup>b</sup>Frequenza dell'allele minore

**Tabella 9.** Informazioni relative alla genotipizzazione dello SNP rs1128503 mappante nell'esone 12 del gene *MDR1*.

		Genotipo			MAF <sup>b</sup>	P value	
		N <sup>a</sup>	CC	CT			TT
Controlli		100	33	45	22	0.44	<i>ref.</i>
Casi		96	37	38	21	0.42	0.57
	Grado 1	5	2	1	2	0.5	0.76
Casi	Grado 2	64	24	27	13	0.41	0.58
	Grado 3	20	8	7	5	0.43	0.82
	Colon	51	19	21	11	0.42	0.70
Casi	Sigma	21	14	5	2	0.21	0.0056
	Retto	24	4	12	8	0.58	0.085
Controlli	Maschi	49	14	24	11	0.47	<i>ref.</i>
Casi	Maschi	52	25	21	6	0.32	0.027
Controlli	Femmine	51	19	21	11	0.42	<i>ref.</i>
Casi	Femmine	44	12	17	15	0.53	0.12

<sup>a</sup>Numero di individui

<sup>b</sup>Frequenza dell'allele minore

**Tabella 10.** Informazioni relative alla genotipizzazione dello SNP rs2032582 mappante nell'esone 21 del gene *MDR1*.

		Genotipo				MAF <sup>b</sup>	P value
		N <sup>a</sup>	GG	GT	TT		
Controlli		100	32	44	24	0.46	<i>ref.</i>
Casi		98	32	45	21	0.44	0.75
	Grado 1	5	2	1	2	0.5	1.00
Casi	Grado 2	65	20	31	14	0.45	0.91
	Grado 3	21	8	8	5	0.43	0.71
	Colon	52	15	26	11	0.46	0.98
Casi	Sigma	22	12	8	2	0.27	0.023
	Retto	24	5	11	8	0.56	0.20
Controlli	Maschi	49	16	21	12	0.46	<i>ref.</i>
Casi	Maschi	52	20	26	6	0.37	0.18
Controlli	Femmine	51	16	23	12	0.46	<i>ref.</i>
Casi	Femmine	46	12	19	15	0.53	0.32

<sup>a</sup>Numero di individui

<sup>b</sup>Frequenza dell'allele minore

**Tabella 11.** Informazioni relative alla genotipizzazione dello SNP rs1045642 mappante nell'esone 26 del gene *MDR1*.

		Genotipo					MAF <sup>b</sup>	P value
		N <sup>a</sup>	CC	CT	TT			
Controlli		100	29	42	29	0.5	ref.	
Casi		97	28	45	24	0.48	0.68	
	Grado 1	5	2	1	2	0.5	1.00	
Casi	Grado 2	64	17	29	18	0.51	0.89	
	Grado 3	21	8	10	3	0.38	0.16	
	Colon	51	14	27	10	0.46	0.52	
Casi	Sigma	22	8	7	7	0.48	0.78	
	Retto	24	6	11	7	0.52	0.79	
Controlli	Maschi	49	13	20	16	0.53	ref.	
Casi	Maschi	51	17	26	8	0.41	0.092	
Controlli	Femmine	51	16	22	13	0.47	ref.	
Casi	Femmine	46	11	19	16	0.55	0.24	

<sup>a</sup>Numero di individui

<sup>b</sup>Frequenza dell'allele minore

## **Analisi multivariate**

Le statistiche bivariate hanno suggerito che i polimorfismi *ABCB1* da noi studiati possono influenzare la localizzazione del cancro e possono agire in modo diverso nei due sessi.

Abbiamo effettuato pertanto un approccio statistico di analisi multivariata per capire come il genotipo *ABCB1* possa influenzare la localizzazione del CCR ed una differenza di genere tra i pazienti considerando ulteriori fattori di confondimento.

Prima di tutto, la regressione logistica è stata utilizzata per verificare se l'esito (0 = controlli sani, 1 = CCR in una posizione anatomica specifica, cioè colon o sigma o retto) poteva essere previsto in base al genotipo rs1128503 ed al genere (tabella 12). Questa analisi ha indicato che il carcinoma del sigma è associato al polimorfismo rs1128503 mentre il genere non ha agito come fattore di confondimento.

In effetti, gli OR per le varianti alleliche dei pazienti rispetto ai controlli, distinti per genere, risultavano invariati in relazione agli OR grezzi ottenuti dall'analisi bivariata.

La seconda analisi multivariata mirava a capire se il rapporto di genere tra i pazienti (0 = CCR pazienti di sesso maschile, 1 = pazienti CCR femmine) potevano essere rilevanti secondo la posizione del tumore (colon, sigma, retto), il grado istologico (G1, G2, G3), l'età della diagnosi (anni), ed il genotipo del marcatore rs1128503 (tabella 13).

Tra queste variabili indipendenti, il fattore che influenzava il genere era il genotipo rs1128503 ( $P = 0.009$ ).

L'OD calcolato sulle femmine che presentavano l'allele variante al polimorfismo rs1128503, rispetto ai controlli, corretti per tutti gli altri fattori, sono  $OR_{Ret} = 0.11$  (95% CI 0.03-0.44) e per  $OR_{Rom} = 0.24$  (95% CI 0.06-0.87). Inoltre, un effetto borderline è stato osservato per il grado istologico ( $P$  value = 0.07), dove l'OR di femmine con tumore in G3, contro G1, corretto per tutti gli altri fattori era 0.26 (95% CI 0.08-0.87).

**Tabella 12.** Analisi di regressione logistica per verificare se il cancro del colon sigma (0 = controlli sani, 1 = colon sigma) possa essere modulato dal genotipo rs1128503 e dal genere.

Variabile indipendente	B <sup>a</sup>	e.s. <sup>b</sup>	Wald <sup>c</sup>	d.f. <sup>d</sup>	P value	OR (95%C.I.)
rs1128503			7.57	2	0.02	
CT vs CC	-1.35	0.57	5.58	1	0.02	0.26 (0.08-0.79)
TT vs CC	-1.53	0.81	3.60	1	0.06	0.22 (0.04-1.05)
Genere (0 = maschi, 1 = femmine)	-0.33	0.50	0.44	1	0.51	0.72 (0.27-1.92)
Costante	-0.70	0.39	3.17	1	0.07	
Modello	$\chi^2 = 8.59$ d.f.=3 $P = 0.035$					

<sup>a</sup>Coefficiente di regressione

<sup>b</sup>Errore standard

<sup>c</sup>Wald test

<sup>d</sup>Gradi di libertà

**Tabella 13.** Analisi di regressione logistica per verificare se il rapporto maschi/femmine nei pazienti (0 = pazienti di sesso maschile, 1 = pazienti di sesso femminile) possa essere modulato dal genotipo rs1128503, grado istologico, localizzazione del tumore, ed età della diagnosi.

Variabile indipendente	B <sup>a</sup>	e.s. <sup>b</sup>	Wald <sup>c</sup>	d.f. <sup>d</sup>	P value	OR (95%C.I.)
rs1128503			9.52	2.00	0.01	
CT vs CC	-2.22	0.72	9.51	1.00	0.00	0.11(0.03-0.45)
TT vs CC	-1.45	0.67	4.67	1.00	0.03	0.24(0.06-0.87)
Grado			5.27	2.00	0.07	
G2 vs G1	-1.91	1.23	2.42	1.00	0.12	0.15(0.01-1.65)
G3 vs G1	-1.33	0.61	4.75	1.00	0.03	0.26(0.08-0.87)
Localizzazione			1.18	2.00	0.55	
sigma vs colon	0.61	0.61	1.02	1.00	0.31	1.85(0.56-6.08)
retto vs colon	0.68	0.75	0.83	1.00	0.36	1.98(0.46-8.57)
Età	-0.01	0.02	0.23	1.00	0.63	
Costante	2.54	1.63	2.43	1.00	0.12	

Modello  $\chi^2= 16.86$  d.f.=7  $P = 0.018$

<sup>a</sup>Coefficiente di regressione

<sup>b</sup>Errore standard

<sup>c</sup>Wald test

<sup>d</sup>Gradi di libertà

### **Analisi di associazione per aplotipo**

Per indagare ulteriormente l'associazione genetica tra polimorfismi di *ABCB1* ed il tumore del sigma è stata eseguita un'analisi di associazione per aplotipo (tabella 14).

L'attuale suscettibilità polimorfica non è risultata tipica.

Il livello più alto di associazione è stato osservato per gli aplotipi rs1128503C-rs2032582G- rs1045642T ( $P$  value = 0.0007). Questi aplotipi sono risultati più frequenti tra i pazienti affetti da cancro del sigma; l'odd ratio rispetto all'aplotipo più frequente è 3.48 (95% CI 1.23-9.82).

Dal momento che in questa analisi sono stati eseguiti numerosi test, è stato condotto un test di permutazione per regolare i livelli di significatività: il valore di  $P$  corretto per l'associazione con l'aplotipo rs1128503C -rs2032582G-rs1045642T è risultato 0.032.

L'aplotipo specifico per genere è stato valutato tra i pazienti CCR.

Una differenza significativa tra pazienti CCR maschi e femmine è risultata evidente per ciascuno degli alleli testati come pure per alcuni aplotipi; valori di  $P$  nominali ed OR ottenuti sono elencati nella tabella 15.

Dopo una correzione per confronti multipli, abbiamo confermato una frequenza significativa dell'allele comune del polimorfismo rs1128503 (corretto  $P$  value = 0.035), nonché degli aplotipi rs1128503C-rs2032582G (corretto  $P$  value = 0.047).

**Tabella 14.** Analisi di associazione allelica ed aplotipica per il cancro al sigma.

M1 <sup>a</sup>	M2 <sup>b</sup>	M3 <sup>c</sup>	Casi	Controlli	Freq-Ca	Freq-Co	OR (95% C.I.)	$\chi^2$	P value
C	-	-	33.0	111.0	0.79	0.56	<i>ref.</i>		
T	-	-	9.0	89.0	0.21	0.45	0.34(0.15-0.75)	7.668	0.0056
-	G	-	32.0	108.0	0.73	0.54	<i>ref.</i>		
-	T	-	12.0	92.0	0.27	0.46	0.44(0.21-0.90)	5.172	0.023
-	-	C	23.0	100.0	0.52	0.50	<i>ref.</i>		
-	-	T	21.0	100.0	0.48	0.50	0.91(0.48-0.76)	0.075	0.78
C	G	-	30.0	97.4	0.71	0.49	<i>ref.</i>	7.104	0.0077
C	T	-	3.0	13.6	0.07	0.07	0.72(0.19-2.68)	0.018	0.89
T	G	-	0.0	10.6	0.00	0.05	non calcolabile	2.237	0.13
T	T	-	9.0	78.4	0.21	0.39	0.37(0.16-0.83)	4.916	0.027
C	-	C	22.0	91.0	0.52	0.48	<i>ref.</i>	0.277	0.60
C	-	T	11.0	16.0	0.26	0.08	2.84(1.16-6.98)	10.560	0.0012
T	-	T	9.0	83.0	0.21	0.44	0.45(0.20-1.03)	7.120	0.0076
-	G	C	21.9	90.3	0.50	0.45	<i>ref.</i>	0.328	0.57
-	G	T	10.1	17.7	0.23	0.09	2.37(0.94-5.92)	7.129	0.0076
-	T	C	1.1	9.7	0.03	0.05	0.49(0.06-4.09)	0.539	0.46
-	T	T	10.9	82.3	0.25	0.41	0.55(0.25-1.20)	4.085	0.043
C	G	C	21.0	83.2	0.50	0.44	<i>ref.</i>	0.540	0.46
C	G	T	9.0	10.3	0.21	0.05	3.48(1.23-9.82)	11.390	0.00074
C	T	C	1.0	7.8	0.02	0.04	0.51(0.06-4.37)	0.289	0.59
C	T	T	2.0	5.7	0.05	0.03	1.38(0.25-7.65)	0.488	0.49
T	G	T	0.0	7.6	0.00	0.04	non calcolabile	1.596	0.21
T	T	T	9.0	75.5	0.21	0.40	0.47(0.20-1.10)	5.173	0.023

<sup>a</sup>M1 1236 C>T rs1128503

<sup>b</sup>M2 2677 G>T/A rs2032582

<sup>c</sup>M3 3435 C>T rs1045642

**Tabella 15.** Analisi di associazione allelica ed aplo tipica per CCR maschi (casi) contro CCR femmine (controlli)

M1 <sup>a</sup>	M2 <sup>b</sup>	M3 <sup>c</sup>	Femmine	Maschi	Freq-Fem	Freq-Masch	OR (95% C.I.)	$\chi^2$	P value
C			41	71	0.47	0.68	<i>ref.</i>	9.22	0.002
T			47	33	0.53	0.32	0.40(0.23-0.73)	9.22	0.002
	G		43	66	0.47	0.63	<i>ref.</i>	5.53	0.019
	T		49	38	0.53	0.37	0.50(0.29-0.90)	5.53	0.019
		C	41	60	0.45	0.59	<i>ref.</i>	3.94	0.047
		T	51	42	0.55	0.41	0.56(0.32-0.99)	3.94	0.047
C	G		34.7	63.8	0.39	0.61	<i>ref.</i>	9.09	0.003
C	T		6.3	7.2	0.07	0.07	0.61(0.19-1.99)	0.00	0.991
T	G		5.3	2.2	0.06	0.02	0.22(0.04-1.21)	1.88	0.171
T	T		41.7	30.8	0.47	0.30	0.4(0.22-0.75)	6.61	0.010
C		C	27.9	52.5	0.32	0.52	<i>ref.</i>	7.68	0.006
C		T	13.2	17.5	0.15	0.17	0.70(0.28-1.75)	0.26	0.607
T		C	11.2	7.5	0.13	0.07	0.35(0.11-1.13)	1.49	0.222
T		T	35.9	24.5	0.41	0.24	0.36(0.18-0.72)	6.78	0.009
	G	C	33.4	51.8	0.36	0.51	<i>ref.</i>	4,41	0.036
	G	T	9.6	13.2	0.10	0.13	0.89(0.33-2.39)	0,26	0.614
	T	C	7.6	8.2	0.08	0.08	0.70(0.22-2.23)	0,02	0.895
	T	T	41.4	28.8	0.45	0.28	0.45(0.24-0.86)	5,90	0.015
C	G	C	27.4	49.0	0.32	0.50	<i>ref.</i>	6.35	0.012
C	G	T	7.4	12.9	0.09	0.13	0.97(0.34-2.81)	0.99	0.320
C	T	T	6.2	4.1	0.07	0.04	0.37(0.10-1.45)	0.71	0.398
T	G	C	3.2	2.1	0.04	0.02	0.37(0.06-2.39)	0.33	0.566
T	T	C	7.4	5.9	0.09	0.06	0.44(0.13-1.58)	0.55	0.458
T	T	T	34.4	24.0	0.40	0.24	0.39(0.19-0.79)	5.33	0.021

<sup>a</sup>M1 1236 C>T rs1128503

<sup>b</sup>M2 2677 G>T/A rs2032582

<sup>c</sup>M3 3435 C>T rs1045642

## **Studio dell'associazione genetica tra CCR ed i polimorfismi selezionati sul gene *ROCK1* e *ROCK2***

### **Analisi statistica**

Per il presente studio, abbiamo genotipizzato un totale di nove polimorfismi a singolo nucleotide: quattro di loro mappanti nel gene *ROCK1* ed altri cinque mappanti in *ROCK2*, al fine di valutare il possibile coinvolgimento di queste due chinasi nella carcinogenesi del colon-retto nell'ambito della popolazione italiana. La genotipizzazione condotta per il polimorfismo rs73963110 di *ROCK1* non ha fornito alcuna prova della presenza della variante allelica nel nostro set di campioni, quindi lo SNP è stato escluso dall'analisi statistica. Uno studio di associazione caso-controllo è stato eseguito per verificare se varianti genetiche in polimorfismi mappanti nei geni *ROCK1* o *ROCK2* fossero associate con i casi CCR in toto. Successivamente abbiamo stratificato i casi CCR in base al genere e abbiamo preso in considerazione i casi che presentavano interessamento linfonodale (tabelle 16 e 17). Le varianti alleliche non sono risultate associate né con la comparsa di CCR né con la presenza di metastasi ai linfonodi regionali, per nessuno dei geni indagati; tuttavia, una significativa associazione è stata osservata nei maschi CCR per una variante presente nel gene *ROCK1*. Infatti l'allele variante di rs35996865 risultava più frequente nei pazienti di sesso maschile ( $P$  value = 0.028) rispetto al gruppo di controllo. L'odds ratio calcolato (OR) per la variante allelica rs35996865 (G) nei portatori eterozigoti rispetto ai non portatori (OR<sub>et</sub>) è stato pari a 1.88 (95% CI 1.02-3.49), mentre per i portatori omozigoti rispetto ai non portatori (OR<sub>om</sub>) non è risultato significativo (OR 1.977; 95% CI 0.76-5.13), probabilmente a causa del numero esiguo degli individui con quel genotipo. Tuttavia, considerando globalmente i genotipi che presentano la variante, l'OR risulta di 1.9 (95% CI 1.07-3.38). L'allele variante è più frequente nei pazienti di sesso maschile (MAF = 0.33) rispetto ai controlli (MAF = 0.23), mentre la frequenza nei pazienti di sesso femminile (MAF = 0.23) risulta simile a quella osservata per i controlli.

Per il gruppo di controllo, nessuna differenza nella frequenza allelica è stata osservata stratificando per genere ( $P$  value = 0.45). Per indagare ulteriormente l'associazione genetica tra i polimorfismi di *ROCK1* o *ROCK2* ed il CCR, è stata effettuata un'analisi di associazione per l'aplotipo. Nessun aplotipo ha mostrato un livello di associazione allelica significativa neppure stratificando per sesso e per interessamento linfonodale (dati non mostrati).

**Tabella 16.** Analisi di associazione al locus *ROCK1*

informazioni SNP			Genotipi				Rischio Relativo (95% CI)		
SNP ID	Alleli <sup>a</sup>	Campioni	11	12	22	MAF <sup>b</sup>	$P$ value <sup>c</sup>	OR <sub>et</sub>	OR <sub>om</sub>
rs35996865	T/G	Controlli	87	42	12	0.23	<i>ref.</i>	<i>ref.</i>	<i>ref.</i>
		Casi	71	54	12	0.28	0.173	1.575 (0.945-2.626)	1.225 (0.519-2.894)
		Casi maschi	33	30	9	0.33	0.0285	1.883 (1.017-3.488)	1.977 (0.763-5.126)
		Casi femmine	38	24	3	0.23	0.9418	1.308 (0.697-2.456)	0.572 (0.153-2.145)
		Casi N0	37	29	5	0.27	0.3606	1.624 (0.882-2.988)	0.98 (0.322-2.979)
		Casi N1/2	20	16	6	0.33	0.0675	1.657 (0.78-3.521)	2.175 (0.728-6.494)
rs2127958	C/T	Controlli	22	72	47	0.41	<i>ref.</i>	<i>ref.</i>	<i>ref.</i>
		Casi	22	56	59	0.36	0.2619	0.778 (0.392-1.545)	1.255 (0.621-2.539)
		Casi maschi	11	30	31	0.36	0.3156	0.833 (0.360-1.930)	1.319 (0.562-3.099)
		Casi femmine	11	26	28	0.37	0.417	0.722 (0.308-1.692)	1.191 (0.503-2.821)
		Casi N0	11	29	32	0.35	0.2528	0.806 (0.347-1.870)	1.362 (0.581-3.192)
		Casi N1/2	8	16	17	0.39	0.732	0.611 (0.231-1.618)	0.995 (0.373-2.653)
rs288980	C/T	Controlli	40	72	29	0.46	<i>ref.</i>	<i>ref.</i>	<i>ref.</i>
		Casi	38	66	33	0.48	0.6239	0.965 (0.554-1.682)	1.198 (0.614-2.336)
		Casi maschi	17	38	17	0.5	0.4456	1.242 (0.623-2.476)	1.379 (0.605-3.147)
		Casi femmine	21	28	16	0.46	0.9918	0.741 (0.373-1.470)	1.051 (0.469-2.356)
		Casi N0	22	34	15	0.45	0.8409	0.859 (0.443-1.663)	0.94 (0.418-2.118)
		Casi N1/2	12	17	13	0.51	0.412	0.787 (0.342-1.812)	1.494(0.596-3.745)

<sup>a</sup>Primo allele ancestrale (in NCBI dbSNP database)

<sup>b</sup>Frequenza dell'allele minore

<sup>c</sup> $P$  value per l'analisi di associazione

<sup>d</sup>N0: assenza di cellule tumorali nei linfonodi regionali

<sup>e</sup>N1/2: presenza di metastasi nei linfonodi regionali

**Tabella 17.** Analisi di associazione al *locus ROCK2*

informazioni SNP			Genotipi					Rischio Relativo (95% CI)	
SNP ID	Alleli <sup>a</sup>	Campioni	11	12	22	MAF <sup>b</sup>	<i>P</i> value <sup>c</sup>	OR <sub>et</sub>	OR <sub>om</sub>
rs12692437	T/C	Controlli	119	21	1	0.08	<i>ref.</i>	<i>ref.</i>	<i>ref.</i>
		Casi	108	26	2	0.11	0.2504	1.364 (0.726-2.565)	2.204 (0.197-24.565)
		Casi maschi	56	15	1	0.12	0.2219	1.518 (0.728-3.165)	2.125 (0.131-34.595)
		Casi femmine	52	11	1	0.1	0.5072	1.199 (0.539-2.665)	2.288 (0.140-37.292)
		Casi N0	61	11	0	0.08	0.8522	1.022 (0.463-2.256)	0.648 (0.026-16.135)
		Casi N1/2	33	8	1	0.12	0.2924	1.374 (0.558-3.382)	3.606 (0.22-59.212)
rs7563468	G/A	Controlli	3	30	108	0.13	<i>ref.</i>	<i>ref.</i>	<i>ref.</i>
		Casi	3	37	95	0.16	0.2892	1.233 (0.232-6.559)	0.880 (0.173-4.462)
		Casi maschi	2	21	48	0.18	0.1802	1.050 (0.161-6.841)	0.667 (0.108-4.120)
		Casi femmine	1	16	47	0.14	0.719	1.600 (0.154-16.661)	1.306 (0.132-12.879)
		Casi N0	1	18	53	0.14	0.7456	1.800 (0.174-18.638)	1.472 (0.15-14.493)
		Casi N1/2	1	10	30	0.15	0.6599	1 (0.093-10.736)	0.833 (0.084-8.304)
rs35768389	T/A	Controlli	140	1	0	0.004	<i>ref.</i>	<i>ref.</i>	<i>ref.</i>
		Casi	136	0	0	0	13.793	0.343 (0.014-8.495)	1.029 (0.020-52.242)
		Casi maschi	72	0	0	0	N.C.	N.C.	N.C.
		Casi femmine	64	0	0	0	N.C.	N.C.	N.C.
		Casi N0	72	0	0	0	N.C.	N.C.	N.C.
		Casi N1/2	42	0	0	0	N.C.	N.C.	N.C.
rs17463896	A/G	Controlli	44	65	32	0.46	<i>ref.</i>	<i>ref.</i>	<i>ref.</i>
		Casi	40	60	36	0.49	0.5115	1.015 (0.584-1.766)	1.238 (0.652-2.348)
		Casi maschi	24	32	16	0.44	0.7987	0.903 (0.470-1.734)	0.917 (0.420-1.999)
		Casi femmine	16	28	20	0.53	0.1657	1.185 (0.575-2.443)	1.719 (0.773-3.824)
		Casi N0	16	34	22	0.54	0.0999	1.438 (0.71-2.916)	1.891 (0.859-4.16)
		Casi N1/2	17	15	10	0.42	0.509	0.597 (0.27-1.32)	0.809 (0.327-1.998)
rs16857265	T/C	Controlli	118	18	5	0.1	<i>ref.</i>	<i>ref.</i>	<i>ref.</i>
		Casi	120	15	1	0.06	0.113	0.819 (0.395-1.702)	0.197 (0.023-1.709)
		Casi maschi	62	9	1	0.08	0.4381	0.952 (0.404-2.242)	0.381 (0.044-3.330)
		Casi femmine	58	6	0	0.05	0.0745	0.678 (0.256-1.800)	0.184 (0.010-3.387)
		Casi N0	63	8	1	0.07	0.3066	0.832 (0.343-2.021)	0.375 (0.043-3.277)
		Casi N1/2	37	5	0	0.06	0.264	0.886 (0.308-2.55)	0.287 (0.016-5.317)

<sup>a</sup>Primo allele ancestrale (in NCBI dbSNP database)

<sup>b</sup>Frequenza dell'allele minore

<sup>c</sup>P value per l'analisi di associazione

<sup>d</sup>N0: assenza di cellule tumorali nei linfonodi regionali

<sup>e</sup>N1/2: presenza di metastasi nei linfonodi regionali

## **Risultati II – Parte: Ricerca di mRNA marcatori**

### **Set di dati basati sulla metanalisi TRAM**

Una meta-analisi sistematica sull'espressione genica differenziale tra CCR e soggetti sani è stata condotta per identificare gli mRNA con il più alto rapporto di espressione nel sangue dei CCR, al fine di selezionare i migliori candidati biomarcatori. Il programma Transcriptome Mapper (TRAM) ci ha permesso la gestione di piattaforme sperimentali [62], a seguito di una ricerca completa, per un totale di 37 serie GEO selezionate per CCR, insieme ad ulteriori 23 serie per il sangue normale. Le serie selezionate comprendevano un totale di 2.532 e 958 campioni, rispettivamente. Un campionamento casuale è stato fatto selezionando le prime 10 liste di campioni elencati per ciascuna serie includendo più di 10 campioni, riducendo così il numero di campioni analizzati 349 CCR e 200 soggetti sani per il sangue normale (rispettivamente, 14% e 21% del totale).

L'analisi dei risultati finali dei geni normalizzati ha permesso l'identificazione dei candidati marcatori, su un totale di 38.104, con un valore di espressione disponibile sia nel gruppo 'A' (CCR) che 'B' (sangue), con il più alto rapporto 'A' / 'B' tra il valore di espressione nei tessuti CCR rispetto alle cellule del sangue. La lista comprende loci sovraespressi in CCR con un rapporto di 32:1, che può essere considerato come una soglia per la scelta dei migliori candidati mRNA e, in questo caso, è prevista una differenza chiaramente identificabile, equivalente in teoria, a 5 cicli di PCR tra casi e controlli durante la fase esponenziale di amplificazione.

## **Selezione dei marcatori**

Una selezione ulteriore dei risultati completi del database TRAM è stata fatta per cercare loci con valori medi di espressione provenienti da una serie di dati superiori al 50% del numero di campioni per ogni pool, cioè > 175 per CCR e > 100 per il sangue. In questo modo, il valore medio di espressione non valutato in una grande frazione di campioni, è stato escluso dalla validazione sperimentale; questo è il caso di espressioni UniGene con sequenze codificate non caratterizzate aventi prefissi "Hs".

Questo filtraggio ha portato alla selezione teorica di 15 migliori candidati (Tabella 18).

Per selezionare i trascritti con alta abbondanza assoluta, considerando che il valore di espressione è espresso come percentuale del valore medio nel profilo di espressione integrato (cioè, 1000 = dieci volte la media), abbiamo ulteriormente escluso l'ultimo locus della lista (da SLC26A3 a KRT20) tranne CEACAM6, che ha un alto valore di espressione assoluto. In seguito abbiamo escluso la Molecola di Adesione delle Cellule 5 Correlata all'Antigene Carcinoembrionale (CEACAM5), il Collagene tipo III Alfa 1 (COL3A1) e Cheratina 18 (KRT18) a causa dell'esistenza di pseudogeni, che non consentono uno specifico disegno dei primer PCR per mRNA e quindi di distinguere tra mRNA e DNA.

I sette candidati marcatori scelti per la valutazione sono stati:

Tetraspanina 8 (TSPAN8), E-Caderina (Epiteliale) Molecola di Adesione delle Cellule Epiteliali (EpCAM), Inibitore Serina Peptidasi Kazal di tipo 1 (SPINK1), Collagene di tipo I Alfa 2 (COL1A2), Caderina 1 (CDH1), Lectina legame al Galattoside Solubile 4 (LGALS4), Antigene Carcinoembrionario collegato alle Molecole di Adesione 6 (CEACAM6) (Tabella 18).

**Tabella 18.** Selezione dei candidati marcatori (i primi 15 loci con il più alto rapporto di espressione 'A'/'B' e con un valore di espressione del punto specifico per locus superiore del 50% per il gruppo 'A' rispetto al gruppo 'B')

Nome del gene	Valore 'A'	Valore 'B'	Rapporto	Posizione	Dati dei punti	Dati dei punti	SD come % di espressione	SD come % di espressione
	<i>Cancro del colon-retto</i>	<i>Sangue normale</i>	'A'/'B'		'A'	'B'	'A'	'B'
<b><i>TSPAN8</i></b>	2,313.03	13.12	176.36	chr12	349	185	67.66	424.15
<b><i>EPCAM</i></b>	2,111.87	13.60	155.27	chr2	354	222	69.86	82.81
<b><i>SPINK1</i></b>	1,086.88	12.68	85.70	chr5	368	215	99.51	107.87
<i>COL3A1</i>	862.27	10.10	85.35	chr2	1291	527	140.54	110.39
<i>CEACAM5</i>	2,074.89	24.44	84.88	chr19	572	315	132.13	144.57
<b><i>COL1A2</i></b>	989.85	12.79	77.42	chr7	767	549	115.82	131.00
<b><i>CDH1</i></b>	825.52	11.87	69.56	chr16	573	455	120.84	105.73
<b><i>LGALS4</i></b>	1,980.67	29.50	67.15	chr19	369	185	77.43	104.96
<i>KRT18</i>	1,719.60	25.69	66.93	chr12	482	318	88.43	111.69
<i>SLC26A3</i>	800.45	12.31	65.01	chr7	369	195	187.48	138.12
<i>REG1A</i>	776.24	12.54	61.91	chr2	346	185	191.25	75.25
<i>FNI</i>	664.80	11.22	59.25	chr2	1588	902	144.43	125.93
<i>LUM</i>	556.84	9.40	59.22	chr12	403	262	108.62	83.50
<b><i>CEACAM6</i></b>	2,245.12	38.49	58.33	chr19	583	274	65.15	480.79
<i>KRT20</i>	731.52	12.58	58.13	chr17	372	182	119.91	85.52

Tessuto di cancro del colon-retto ('A'), sangue normale ('B').

'Valore': media dell'espressione del gene normalizzato in tutti i campioni; 'Dati dei punti': numero dei punti con un valore di espressione per il locus; 'SD': deviazione standard espressa come percentuale media del valore di espressione. Il nome del gene in grassetto: selezionato per la validazione sperimentale.

### **Analisi quantitativa dei candidati marcatori mRNA**

L'RNA estratto di ciascun campione (pazienti e soggetti sani) è stato valutato per la qualità, e l'analisi di espressione dei candidati marcatori mediante PCR quantitativa è stata normalizzata utilizzando il gene *B2M* come riportato da Hamm et al. [104]. La curva standard e l'efficienza per ogni coppia di primers è mostrata in figura 3 e tabella 19.

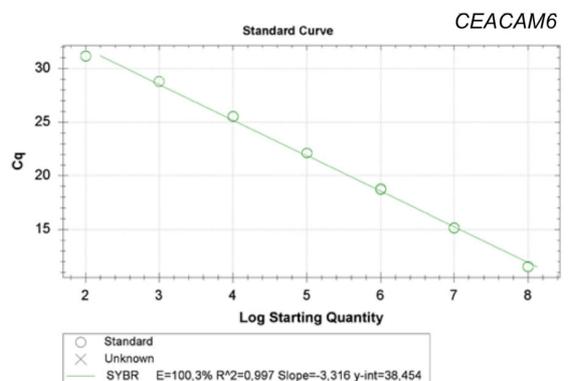
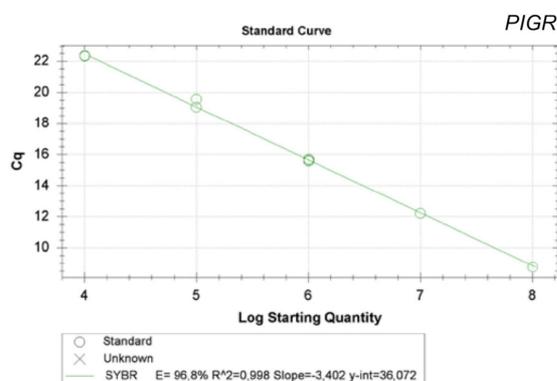
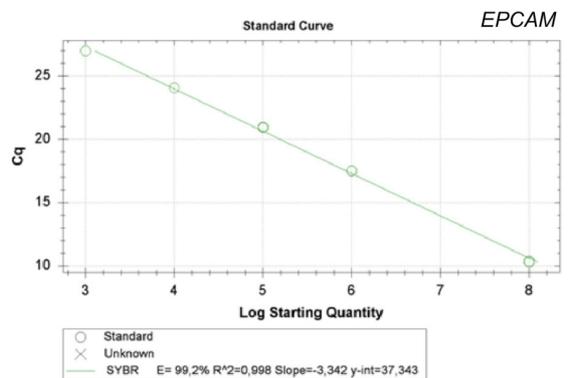
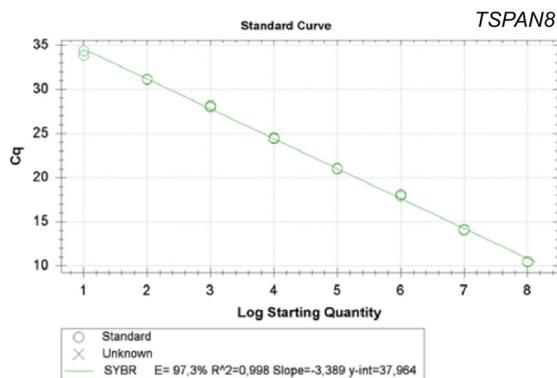
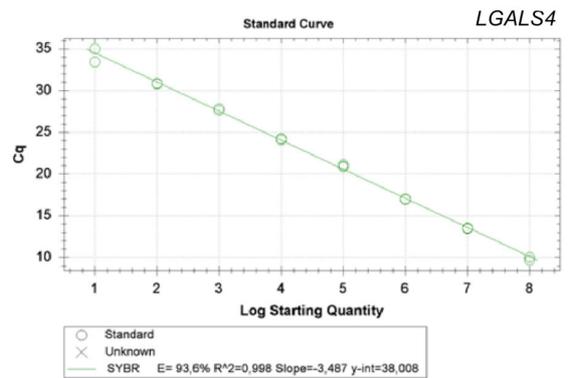
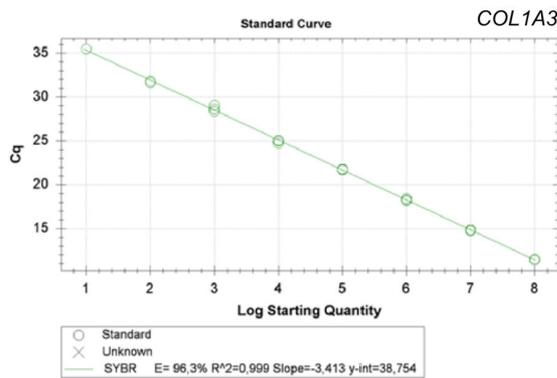
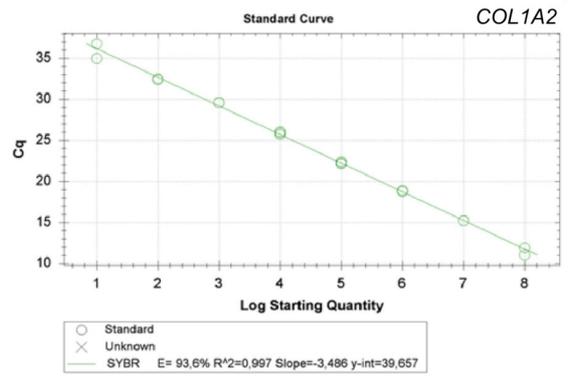
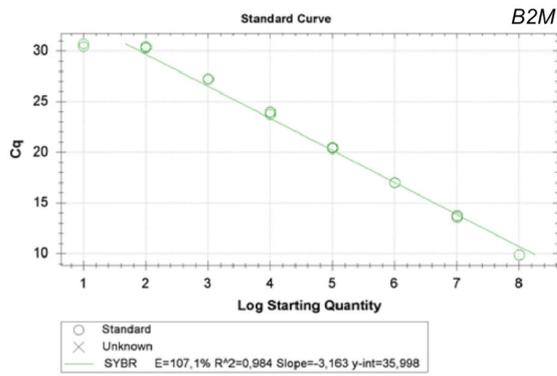
I geni testati hanno mostrato un unico picco nell'analisi della curva di melting e tutti i controlli negativi non hanno mostrato nessun valore di amplificazione rilevabile, convalidando così l'amplificazione specifica.

I livelli di espressione di mRNA normalizzati nel sangue, indicati come Delta CT (Ciclo soglia) sono riportati in tabella 20. Il livello di espressione relativa dei marcatori candidati sembra essere estremamente specifico per questi geni. L'espressione di quattro dei sette geni analizzati: *TSPAN8*, *LGALS4*, *COL1A2*, *CEACAM6*, ha portato differenze statisticamente significative tra casi e controlli.

**Tabella 19.** Efficienza dei primers (%)

	EFFICIENZA(%)	PENDENZA
<i>B2M</i>	107.1	-3.163
<i>COL1A2</i>	93.6	-3.486
<i>COL1A3</i>	96.3	-3.413
<i>LGALS4</i>	93.6	-3.487
<i>TSPAN8</i>	97.3	-3.389
<i>CEACAM6</i>	100.3	-3.316
<i>EPCAM</i>	99.2	-3.342
<i>PIGR</i>	96.8	-3.4

**Figura 3.** Curve di efficienza dei primers.



## Valore diagnostico dei marcatori di mRNA nel sangue per il CCR

Per valutare l'accuratezza diagnostica in termini di specificità e sensibilità dei candidati marcatori, è stata eseguita l'analisi ROC (Receiver Operating Curve). In particolare, l'analisi dei singoli marcatori è stata eseguita utilizzando il Software MedCalc. Questo programma utilizza test statistici che permettono di ottenere la sensibilità e specificità dei dati ROC, e l'indice Youden che fornisce il valore di soglia che minimizza la probabilità di trovare falsi positivi e falsi negativi per determinare una soglia per ciascun marcatore, come riportato nella Tabella 20. Una volta selezionati i migliori marcatori mediante analisi ROC, abbiamo integrato i valori dei 4 candidati marcatori (*TSPAN8*, *LGALS4*, *CEACAM6* e *COL1A2*), in un'unica variabile tramite analisi discriminante utilizzando il Software SPSS.

Più in dettaglio SPSS ha calcolato il potere discriminante dell'uso simultaneo dei migliori punteggi ottenuti con MedCalc.

Quindi abbiamo costruito un modello di logistica multipla (regressione logistica binaria); in questo modello la variabile dipendente è la presenza / assenza della malattia (normale Vs cancro) e la variabile indipendente è il livello di marcatori circolanti utilizzati per costruire la curva ROC, creando così un grafico di probabilità. Attraverso questa analisi, abbiamo esaminato se l'uso simultaneo dei marcatori in esame possa avere un potere discriminante aggiuntivo rispetto all'analisi dei singoli marcatori.

Le più alte precisioni diagnostiche vicino ad 1 sono state trovate per *TSPAN8* (AUC 0.751), *LGALS4* (AUC 0.746), *COL1A2* (AUC 0.718) e *CEACAM6* (AUC 0.632).

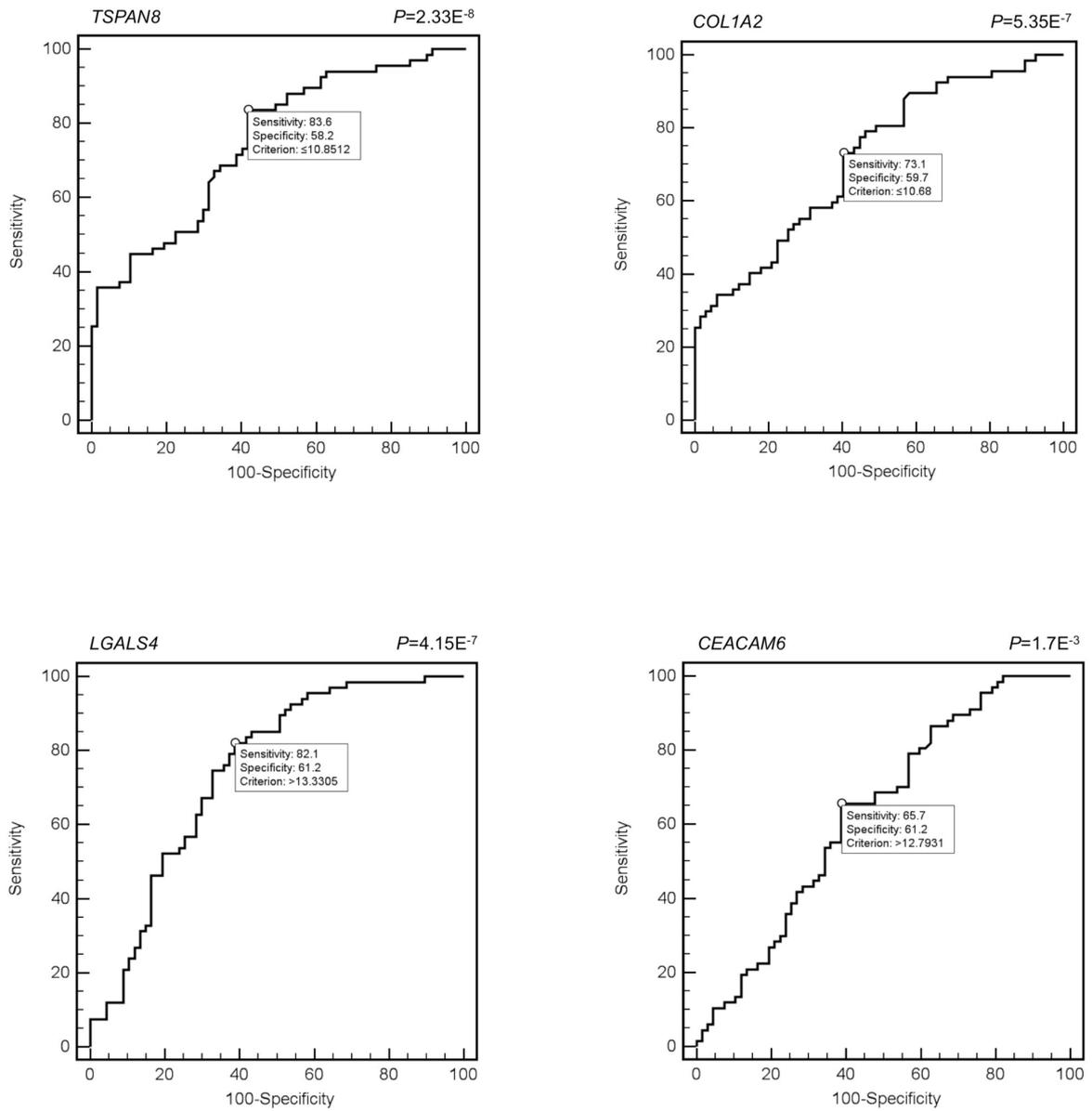
L'elaborazione grafica corrispondente per ciascun candidato marcatore è riportata in Figura 4.

**Tabella 20.** Livelli di espressione mRNAs dei marcatori indicati. Sono riportati i valori statistici tra CCR e controlli, le AUC derivate dalle curve ROC, sensibilità e specificità.

	Numero (n) CCR	CCR Media $\Delta Ct \pm SD$	Numero Controlli	Controlli Media $\Delta Ct \pm SD$	P value	AUC	Sensibilità %	Specificità %
<i>TSPAN8</i>	67	9.41 $\pm$ 2.00	67	11.33 $\pm$ 1.72	0.00000002	0.751	83.6	58.2
<i>EPCAM</i>	39	11.23 $\pm$ 1.36	36	11.83 $\pm$ 1.23	0.08	0.631	-	-
<i>SPINK 1</i>	39	11.88 $\pm$ 2.87	36	11.85 $\pm$ 2.59	0.9	0.503	-	-
<i>COL1A2</i>	67	9.59 $\pm$ 2.14	67	11.45 $\pm$ 1.92	0.0000005	0.718	73.1	59.7
<i>CDHI</i>	39	9.9 $\pm$ 0.9	36	9.5 $\pm$ 1.08	0.2	0.581	-	-
<i>LGALS4</i>	67	14.43 $\pm$ 1.28	67	12.89 $\pm$ 1.97	0.0000004	0.746	82.1	61.2
<i>CEACAM6</i>	67	13.23 $\pm$ 1.24	67	12.34 $\pm$ 1.89	0.0017	0.632	65.7	61.2
Pannello <i>TSPAN8 + LGALS4</i>	67	-	67	-		0.861	92.54	67.16

'CCR': carcinoma del colon-retto; ' $\Delta Ct$ ': delta del ciclo soglia; 'Ct': ciclo soglia; 'SD': deviazione standard ; 'AUC': area sotto la curva derivata dalle curve ROC; P-value: 2-way Anova.

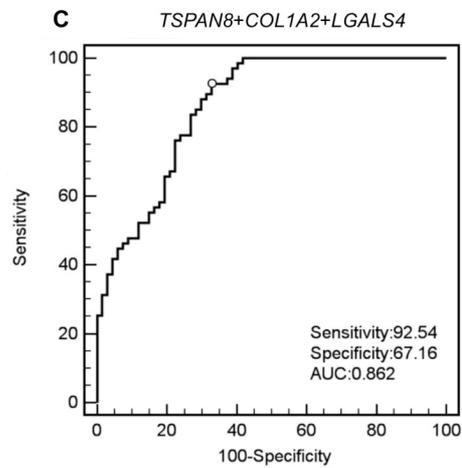
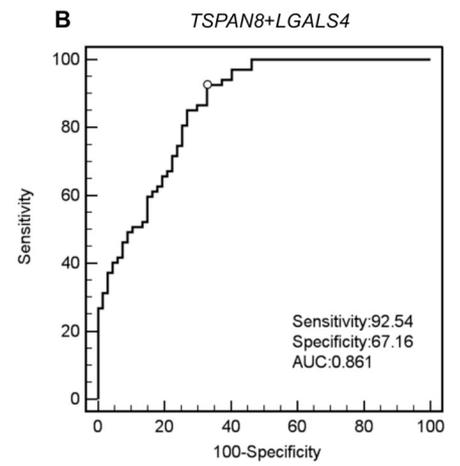
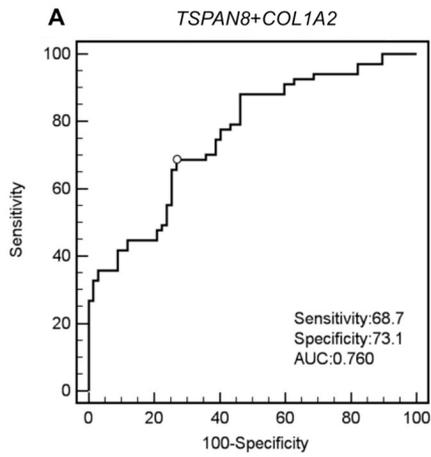
**Figura 4.** Elaborazione grafica dell'analisi ROC per *TSPAN8*, *COL1A2*, *LGALS4* e *CEACAM6*.



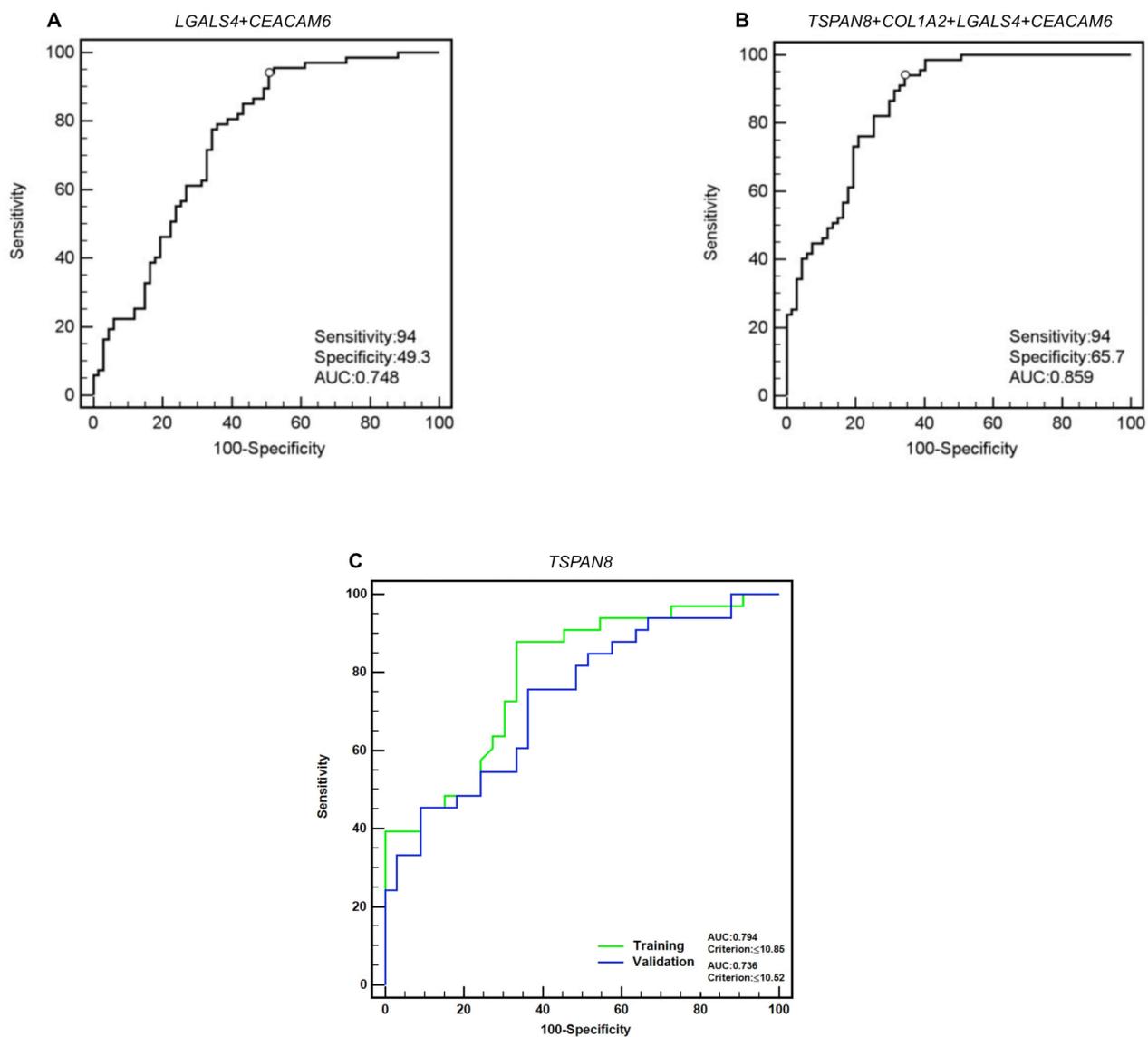
### **Valutazione del potenziale diagnostico del pannello dei quattro mRNA**

Per valutare il potenziale uso dei candidati mRNA selezionati come pannello diagnostico per il CCR, l'analisi ROC è stata effettuata sui dati convalidati impostati per ogni possibile combinazione. La combinazione *TSPAN8* e *COL1A2*, ha mostrato una sensibilità al 68.7% e una specificità del 73.1% (Fig. 5A). La combinazione *LGALS4* e *TSPAN8* ha migliorato la sensibilità al 92.54% con specificità del 67.16% ed un valore predittivo positivo (PPV) del 76.25% e valore predittivo negativo (VPN) di 88.25%. Per i singoli geni, la specificità ha raggiunto circa il 61% e 58.2% rispettivamente, mentre la sensibilità misurava rispettivamente 82.1% e 83.6% (Fig. 5B). La combinazione di *TSPAN8* con *LGALS4* e *COL1A2* ha restituito la stessa sensibilità e la specificità del *TSPAN8* in combinazione con *LGALS4* (Fig. 5C e Fig 6).

**Figura 5.** Elaborazione grafica dell'analisi ROC per la combinazione **A** *TSPAN8* e *COL1A2*, **B** *TSPAN8* e *LGALS4*, **C** *TSPAN8*, *COL1A2* e *LGALS4*.



**Figura 6.** Elaborazione grafica dell'analisi ROC per la combinazione **A** *LGALS4* e *CEACAM6*, **B** *TSPAN8*, *COL1A2*, *LGALS4* e *CEACAM6* **C** Training e Validation per *TSPAN8*.



# *Discussione*

## **I – Parte: Ricerca di polimorfismi genetici associati al rischio di carcinoma del colon-retto**

### Studio dell'associazione genetica tra CCR ed i polimorfismi selezionati sul gene *CDH1* ed analisi dell'espressione di *CDH1*, *CTNNB1*, *LEF1*

L'E-caderina è una proteina soppressore del tumore. E' stato osservato che una minor espressione della E-caderina, tipicamente coinvolta nella transizione epitelio mesenchimale (EMT), determina un rilascio della  $\beta$ -catenina dal complesso E-caderina /  $\beta$ -catenina.

Così, la  $\beta$ -catenina risulta libera di traslocare al nucleo, con conseguente attivazione del complesso fattore di trascrizione TCF / LEF, in grado di attivare la EMT. Tutte queste circostanze sono collegate ad una attivazione aberrante del pathway di Wnt, a sua volta implicato in processi patologici come la tumorigenesi del colon [66]. In conseguenza di quanto descritto, i polimorfismi presenti nel gene *CDH1* sono diventati oggetto di studi volti a verificare l'associazione tra varianti alleliche ed il rischio di CCR. Attualmente i risultati si mostrano contrastanti [67,68].

Le nostre analisi di associazione, effettuate su 4 comuni polimorfismi mappanti nel gene *CDH1* hanno suggerito che nessuno dei polimorfismi indagati è associato al CCR.

Tuttavia, è stata ottenuta, l'evidenza di associazione con la combinazione di aplotipo tra rs11642413-rs10431923, suggerendo che i polimorfismi di *CDH1* possono influenzare lo sviluppo di CCR. Sebbene il *P* value corretto per tutti i test di associazione eseguiti fosse vicino alla soglia di significatività (*P* value = 0.048), i due aplotipi risultati associati aumentano molto il rischio di CCR, con indici di odds ratio rispettivamente di 2.8 e 2.5.

Abbiamo inoltre valutato, i livelli di espressione di *CDH1* e di due geni correlati *CTNNB1* e *LEF1* in biopsie di tessuto tumorale e di tessuto normale di 27 pazienti. È stato interessante notare una significativa riduzione dei trascritti genici di *CDH1* e *CTNNB1* nel tessuto tumorale rispetto al tessuto normale. I dati presenti in letteratura segnalano sia una

diminuzione dell'espressione sia una inattivazione somatica di *CDH1* (causata da mutazioni e da metilazioni del promotore) in un certo numero di tumori epiteliali [69].

Per la E-caderina si prospetta un ruolo di soppressore. *CTNNB1* è invece considerato un oncogene, ma i nostri dati, contrariamente a quanto atteso, mostrano un livello di espressione ridotta del gene nel cancro. D'altra parte, sono già stati descritti ridotti livelli di espressione della E-caderina e delle catenine nel cancro endometriale [70] e nei carcinomi epatocellulari [71]. Questo può essere spiegato considerando il doppio ruolo della  $\beta$ -catenina. Infatti, la  $\beta$ -catenina può essere sia una componente dell'apparato giunzionale cellulare, sia un fattore di trascrizione. Nei normali strati epiteliali la  $\beta$ -catenina ancora la E-caderina alla actina corticale del citoscheletro al fine di creare correttamente giunzioni aderenti. L'integrità delle giunzioni aderenti risulta essenziale per l'inibizione da contatto, per la regolazione della crescita cellulare e per l'adesione tra le cellule epiteliali.

I ridotti livelli di espressione di *CDH1* osservati, potrebbero essere determinati da un reclutamento inferiore di  $\beta$ -catenina nelle giunzioni cellulari con conseguente aumento della frazione di beta-catenina impiegata nel nucleo. [72]. Gli effetti potrebbero sia esacerbare la tumorigenicità che diminuire l'espressione di beta-catenina, probabilmente a causa di un meccanismo di feed-back negativo. Mutazioni germinali o somatiche a carico del gene *CTNNB1* sono state osservate solo in un piccolo sottoinsieme di CCR [73,74]. La maggior parte delle mutazioni osservate in *CTNNB1* inibiscono la sua regolazione da parte di Adenomatous Polyposis Coli (APC) evitando la degradazione della  $\beta$ -catenina da parte del proteosoma. Le conseguenze potrebbero essere la stabilizzazione del cosiddetto pool libero della  $\beta$ -catenina nel citoplasma ed eventualmente la sua traslocazione al nucleo. È quindi più probabile che ci sia una modulazione epigenetica dell'espressione di *CTNNB1*, eventualmente influenzata dalla presenza della E-caderina a livello della superficie cellulare: una quantità ridotta di questa proteina transmembrana, potrebbe reclutare una minor quantità di  $\beta$ -catenina. Questa ipotesi rimane da valutare auspicabilmente in un

modello *in vitro*. Ciò potrebbe essere realizzato dosando diverse quantità di beta-catenina in base alla sua posizione (la  $\beta$ -catenina coinvolta nelle giunzioni aderenti, la  $\beta$ -catenina che risulta libera nel citoplasma e la  $\beta$ -catenina che opera nel nucleo in relazione all'espressione variabile di CDH1).

#### Studio dell'associazione genetica tra CCR ed i polimorfismi selezionati sul gene *MDR1 ABCB1*

La funzione principale del prodotto del gene *ABCB1* è di estrarre una gamma di composti xenobiotici tossici dalle cellule, proteggendo così l'organismo da esposizioni potenzialmente dannose. *ABCB1*, nonché la glicoproteina-P, conferisce protezione limitando l'assorbimento di xenobiotici lungo il tratto gastrointestinale e stimolando l'escrezione di questi composti dal fegato e dai reni [75,76].

L'attività di questa pompa di efflusso può essere influenzata, in parte, dai livelli di espressione del gene *ABCB1*. Attualmente, almeno un polimorfismo, rs1045642 mappante nell'esone 26, è stato associato ad una riduzione della quantità di prodotto finale del gene [77,78].

A questo polimorfismo viene attribuito un coinvolgimento nello sviluppo di diversi tipi di cancro, come il glioma, il cancro al seno e tumori epiteliali renali [79,80,81] così come nel tumore del colon [82,83]. Molti polimorfismi comuni del gene *ABCB1* sono stati studiati da diversi autori a causa della loro potenziale capacità di modulare il trasporto mediato dal prodotto del gene *ABCB1*. In particolare, in uno studio funzionale condotto in vivo la variante C3435T (rs1045642) è stata correlata con un'alterata espressione e attività di *ABCB1*[77] Inoltre è stata trovata un'associazione con l'aumento del rischio di sviluppare CCR in pazienti con un'età inferiore di 50 anni ed in anziani non fumatori [84]. In aggiunta, le varianti C3435T (rs1045642) e G-rs3789243-A sono state associate ad una modulazione del rischio di CCR, in combinazione con il consumo di carni rosse lavorate (industriali) [85]. Al fine di valutare la possibile correlazione tra polimorfismi *ABCB1* e il rischio di

CCR, abbiamo eseguito un'analisi di associazione caso-controllo in un campione di studio rappresentativo della popolazione italiana, costituito da 98 pazienti affetti da cancro del colon retto e 100 soggetti sani, definiti controlli.

Abbiamo quindi genotipizzato tre SNP mappanti in tre diversi esoni del gene *ABCB1*: 1236C>T polimorfismo rs1128503 (esone 12), 2677GA polimorfismo rs2032582 (esone 21) e 3435C>T polimorfismo rs1045642 (esone 26).

Il nostro lavoro indica che *ABCB1* ha un ruolo nel CCR, in particolare nell'insorgenza del carcinoma del sigma. Infatti, le varianti alleliche dei polimorfismi rs1128503 e rs2032582 risultavano meno frequenti nei pazienti con tumori localizzati nel sigma. Poiché alcuni aplotipi hanno mostrato una forte associazione allelica, potremmo speculare che l'allele di suscettibilità non sia stato ancora genotipizzato.

Inoltre, abbiamo dimostrato che polimorfismi di *ABCB1* sono differentemente rappresentati nei maschi rispetto alle femmine.

Va sottolineato che il tasso di incidenza del cancro al colon varia notevolmente a seconda del genere (maschi : femmine = 2 : 1). Il genere potrebbe riflettere anche una maggior influenza di alcuni fattori di rischio magari legati alle esposizioni professionali, alla dieta o allo stile di vita. Ad esempio, la popolazione Danese nei portatori omozigoti per il polimorfismo C3435T ha mostrato un aumentato rischio di sviluppare CCR in relazione al consumo di carne rossa [85]. Anche i diversi livelli ormonali possono influenzare l'espressione genica.

È noto, per esempio, che ormoni steroidei, quali il progesterone modulano l'espressione di P-gp in alcuni tessuti, come dimostrato in un modello murino [86,87].

In questo studio abbiamo evidenziato che gli alleli ancestrali sono più frequenti nei maschi CCR, mentre le varianti alleliche sono più rappresentate nelle femmine.

La frequenza allelica nei controlli è risultata simile nei due generi ed a livelli intermedi tra pazienti maschi e femmine.

Alcuni SNP possono alterare l'espressione di ABCB1 sia influenzando sequenze regolatrici che sono bersaglio di ormoni steroidei sia alterando la capacità di ABCB1 di processare specifiche sostanze il cui apporto potrebbe essere influenzato dal genere.

Un'ipotesi simile è di Nakamura e colleghi [88], che suggeriscono un'influenza della variante del polimorfismo rs1045642 in *ABCB1*, nelle donne, sui livelli sierici di cortisolo e di aldosterone.

Sulla base dei nostri risultati si potrebbe ipotizzare che, anche se ormoni come il progesterone non sono trasportati da P-gp [89], potrebbero comunque modulare l'efficienza della pompa di efflusso, determinando la variabilità legata al genere osservata nel nostro campione di studio.

I nostri dati, anche se relativi a diversi marcatori mappanti sullo stesso gene, sono in accordo con quelli riportati da Sainz e colleghi [90]. Anche loro hanno evidenziato un diverso livello di rischio al CCR nei pazienti che presentano una specifica variante polimorfica in *ABCB1*, dipendente dal genere.

Tuttavia, data la complessità sia dei meccanismi coinvolti nello sviluppo del CCR che dei meccanismi mediati dagli ormoni, risulta necessario uno studio più completo, con ampliamento della casistica, per comprendere il vero ruolo dei polimorfismi *ABCB1* nel CCR.

#### Studio dell'associazione genetica tra CCR ed i polimorfismi selezionati sul gene *ROCK1* e *ROCK2*

Dato l'ampio spettro di attività del citoscheletro influenzate dai prodotti dei geni *ROCK*, non è sorprendente che l'espressione e l'attività di queste chinasi sia stata ampiamente investigata nel cancro coloretale. Recentemente, Sari e colleghi [91] hanno suggerito che polimorfismi genetici presenti nei geni *ROCK1* e *ROCK2* possano modificare la suscettibilità individuale al cancro del colon-retto nella popolazione Turca.

Hanno dimostrato che l'espressione dei prodotti proteici di ROCK1 e di ROCK2 è marcatamente incrementata nei pazienti CCR rispetto ai soggetti sani. Inoltre, è stato dimostrato che le loro inibizioni, determinate mediante l'impiego di Y-27632, un composto che agisce contemporaneamente sulle due isoforme chinasiche, riducono significativamente la motilità delle cellule di adenocarcinoma del colon [92] e che il knockdown di ROCK2 riduce l'invasività in linee cellulari epiteliali intestinali ad alta malignità [91]. Al contrario, diverse segnalazioni hanno dimostrato che l'inibizione della Rho-chinasi stimola la proliferazione, la progressione e la migrazione delle cellule tumorali del colon [93,94]. Pertanto, il contributo preciso delle ROCK chinasi nella progressione ed invasione del cancro del colon-retto resta ancora controverso ed ha bisogno di maggiori chiarimenti.

Molto poco si conosce circa il rapporto tra le varianti genetiche di *ROCK1* e *ROCK2* e la suscettibilità al CCR, ed al momento solo un altro gruppo di ricerca ha indagato una possibile associazione tra i polimorfismi in *ROCK1* e *ROCK2* ed il rischio di sviluppare CCR [95]. Le varianti alleliche non sono risultate associate né con la comparsa di CCR né con la presenza di metastasi ai linfonodi regionali, per nessuno dei geni indagati; tuttavia, una significativa associazione è stata osservata nei casi CCR maschi per una variante presente nel gene *ROCK1*. Degno di nota, il fatto che nessuna evidenza di associazione è stata riscontrata per la variante rs35768389 in *ROCK2* o per la variante rs73963110 in *ROCK1*, che invece sono state descritte in precedenza come rilevanti per lo sviluppo del CCR in una coorte Turca [95]. In particolare, nel nostro studio sono state evidenziate delle differenze statisticamente significative nelle frequenze genotipiche ed alleliche per quanto riguarda il polimorfismo rs35768389 in *ROCK2* tra il gruppo di controllo ed il gruppo di pazienti, mentre la genotipizzazione del polimorfismo rs73963110 in *ROCK1* non ha fornito alcuna prova della presenza della variante allelica in tutto il nostro set di dati. Questi risultati contrastanti potrebbero essere giustificabili in considerazione del fatto che l'influenza dei fattori genetici sulla malattia del cancro al colon possono differire in modo

significativo tra popolazioni di diverse etnie. La discrepanza può anche dipendere dalle diverse frequenze alleliche osservate per i diversi polimorfismi, nell'ambito delle diverse popolazioni studiate. In particolare, le frequenze alleliche dei marcatori menzionati nella coorte italiana raggiungono valori più coerenti con il database NCBI rispetto a quelle riscontrate nella popolazione turca. Quando lo studio di associazione caso-controllo è stato effettuato al fine di valutare l'associazione tra varianti genetiche selezionate nei geni *ROCK1* o *ROCK2* con CCR, stratificando per genere, abbiamo scoperto che l'allele variante del polimorfismo rs35996865 mappante in *ROCK1* era più frequente nei pazienti di sesso maschile ( $P$  value = 0.028) rispetto al gruppo di controllo. A nostra conoscenza, ciò rappresenta la prima evidenza di una possibile correlazione tra varianti genetiche polimorfiche di *ROCK1* e l'aumento del rischio nei pazienti CCR di sesso maschile. Sappiamo, dai dati di incidenza, che il rischio di CCR presenta dimorfismo sessuale, con un tasso significativamente più alto nei maschi rispetto alle femmine (2:1) [96]. È stato proposto che gli ormoni maschili svolgano un ruolo chiave nello sviluppo di adenomi del colon. Inoltre, diversi studi descrivono un legame tra androgeni e la via di segnalazione ROCK [97,98]. D'altra parte, le citochine sono state associate alla diffusione delle cellule tumorali umane che hanno origine epiteliale. L'interleuchina 6 (IL-6) induce la progressione del carcinoma gastrico attraverso l'attivazione della segnalazione c-Src / RhoA / ROCK [99]. Nello specifico, gli estrogeni inibiscono la produzione di IL-6 e la minore concentrazione di questo ormone nei maschi sembra giustificare la disparità di insorgenza di carcinoma epatocellulare, rispetto alle femmine [100]. Alla luce di queste considerazioni, proponiamo che questo rapporto tra ormoni sessuali e la segnalazione RhoA / ROCK possa essere influenzata da differenze di genere. Sorprendentemente, la variante G dello SNP rs35996865 in *ROCK1*, che risulta associata all'insorgenza del CCR nei maschi, mappa sul 5'-UTR, una regione putativa regolatoria che potrebbe essere influenzata da una modulazione a monte diversa, a seconda del sesso. Come ultima

considerazione, il polimorfismo rs35996865 è stato recentemente associato al rischio di sviluppare la sindrome metabolica in una popolazione turca [101] ed in una popolazione cinese al rischio di indurre il carcinoma renale a cellule chiare [102]. In realtà ci sono pochi dati presenti in letteratura su questo SNP. Non stupisce un suo ruolo nell'obesità o come fattore di rischio nel cancro coloretale oppure in un altro tipo di carcinoma.

Tenendo conto del coinvolgimento delle isoforme *ROCK* nell'invasione tumorale, abbiamo analizzato la relazione tra i polimorfismi selezionati e la presenza di metastasi nei linfonodi regionali. Il nostro studio di associazione allelica non supporta l'ipotesi che vede le varianti selezionate in *ROCK1* ed in *ROCK2* come fattori di rischio al CCR ed alla diffusione dei linfonodi. Tuttavia, considerando il ruolo cardine svolto dalle *ROCK1* e *2* nella migrazione [103], con molta cautela si deve escludere il loro ruolo biologico nell'invasività del cancro del colon. Crediamo che ulteriori indagini in pazienti che presentano metastasi in organi distali potrebbero contribuire a chiarire meglio il reale coinvolgimento dei polimorfismi *ROCK* nell'influenzare l'invasione ed il processo metastatico delle cellule tumorali del colon-retto.

Questi risultati complessivi indicano un potenziale ruolo minore per i polimorfismi delle Rho-chinasi nello sviluppo del CCR nella popolazione italiana, limitatamente al sottogruppo di pazienti di sesso maschile. Va considerato che, sebbene i marcatori selezionati rappresentino la grande maggioranza dei polimorfismi con una  $MAF > 0.05$ , il livello di informatività dei polimorfismi più rari è molto bassa. Pertanto, l'assenza di associazione con il CCR per il nostro campione in esame non esclude necessariamente un ruolo di *ROCK1* o *2* nell'eziologia del CCR, dal momento che altri polimorfismi rari potrebbero rappresentare fattori di suscettibilità. Inoltre, due delle varianti genotipizzate, scelte dai dati presenti in letteratura, che erano state precedentemente descritte nella predisposizione al CCR [95], non sono risultate polimorfiche nel nostro campione di studio, perdendo la loro informatività.

## II – Parte: Ricerca di mRNA marcatori

Un test di screening sul sangue intero potrebbe rivelarsi un valore aggiunto importante per migliorare gli screening sul CCR. Abbiamo scelto i candidati marcatori con il software TRAM, un potente programma di bioinformatica, che ha permesso senza una conoscenza a priori, il confronto completo tra l'espressione genica differenziale del tumore del colon, rispetto ai tessuti normali del sangue. TRAM ci ha anche permesso di analizzare ed integrare i dati di espressione genica relativi a tutte le molecole di RNA umane descritte in letteratura. In particolare, l'analisi bioinformatica in cieco, ha rilevato che alcuni geni selezionati sono già noti come marcatori in carcinomi solidi. Inoltre, l'utilizzo del sangue intero come oggetto di diagnosi, senza alcun tentativo di frazionamento, permette l'individuazione di entrambe le molecole di mRNA o RNA liberi circolanti (cfRNA). Quest'ultime molecole sono state descritte nel sangue dei pazienti con CCR come riflesso della gravità del tumore circolante [55]. Tali molecole infatti, sono stabili nel sangue e hanno un'emivita variabile da 15 minuti a diverse ore [105,106,107]. Nel nostro studio abbiamo rilevato sia molecole di RNA provenienti dalla frazione circolante libera che dalle CTCs. In queste condizioni la nostra analisi ha confermato tre dei sette geni scelti come potenziali biomarcatori nel sangue dei pazienti con CCR. L'espressione dei geni *TSPAN8*, *LGALS4* e *COL1A2*, nel sangue in toto, può rivelarsi informativa nella diagnosi del CCR.

La combinazione di *TSPAN8* e *LGALS4* mostra valori promettenti di sensibilità 92.5% e specificità 67.16%, in competizione con il diffuso test del sangue occulto nelle feci (FOBT), o il test immunochimico sulle feci (FIT) (74% e 95%) e Cologuard (92% e 87%) [108,109]. A nostra conoscenza, questo è il primo studio che esamina l'espressione di *TSPAN8* mRNA nel sangue per la diagnosi del CCR. Poco si conosce circa le proteine TSPAN8 o Co-029 nel cancro. Studi precedenti hanno descritto questa proteina come marcatore di invasività che permette alle cellule di melanoma di attraversare la membrana basale, portando così ad un'invasione cutanea ed alla progressione metastatica. Quindi

*TSPAN8* è considerato un marcatore promettente nella diagnosi precoce, almeno nel melanoma [110].

Gli altri biomarcatori comprendono glicoproteine di superficie, come CEACAM6 e geni stromali quali COL1A2 e LGALS4. Quest'ultimo regola la motilità delle cellule sul collagene I, cooperando con il complesso E-caderina / p120-catenina (p120ctn).

COL1A2, la proteina più abbondante nel corpo umano, aumenta la sintesi nel tumore del colon e rivela un'espressione notevolmente superiore nei tumori in stadio II, suggerendo che la sua attivazione è un evento precoce nella tumorigenesi del CCR [111,112].

Nel nostro studio, l'espressione di *TSPAN8* così come per *COL1A2* è significativamente inferiore nel sangue di soggetti sani rispetto ai pazienti.

Al contrario, è interessante notare che il livello di espressione di *LGALS4* ha tendenza opposta, con livelli più elevati nei soggetti sani rispetto ai pazienti. Questo risulta coerente con la letteratura che mostra, per *LGALS4*, elevata espressione nel piccolo intestino, colon e retto, mentre in condizioni di tumori coloretali i livelli di espressione diminuiscono.

È stata suggerita quindi una sua funzione di soppressore del tumore inibendo la proliferazione cellulare.

Inoltre, il livello di espressione inferiore per *LGALS4* mRNA può rappresentare un evento precoce nella carcinogenesi del colon [113]. Noi sosteniamo che geni la cui espressione nel tumore del colon è molto diversa da quella della mucosa normale possano rappresentare dei potenziali marcatori biologici di progressione delle lesioni maligne [114].

È interessante notare che, combinando *TSPAN8* e *LGALS4*, che mostrano espressione speculare nel cancro e/o nel sangue normale (*TSPAN8* è più elevato nel sangue CCR, mentre *LGALS4* è più basso), si rivelino valori promettenti di sensibilità e specificità rispetto ai singoli marcatori. Mentre, *TSPAN8* in combinazione con *COL1A2* mostra una curva ROC con un miglioramento di specificità al 73.1%.

Senza dubbio, i nostri risultati richiedono un'estensione del campione di studio con una

coorte più ampia, che potrebbe anche consentire una stratificazione in gradi, comprendente pazienti con altre malattie gastrointestinali.

Sarebbe interessante anche verificare se la diversa stadiazione tumorale e l'invasività si traducessero con diversi livelli di espressione dei biomarcatori del pannello. Finora i nostri risultati suggeriscono che il pannello, che riflette l'eterogeneità della malattia, è più efficiente nel diagnosticare il CCR rispetto all'utilizzo del singolo marcatore [115,116].

Coi nostri risultati dimostriamo inoltre che sia geni stromali che glicoproteine di membrana risultano cruciali nel profilo di espressione genica del cancro. In effetti, le interazioni tra stroma e parenchima sono sempre più riconosciute nei tumori come fattori importanti per la biologia e l'esito clinico [117,118].

# *Conclusioni*

Per quanto riguarda le nostre analisi di associazione effettuate su 4 comuni polimorfismi mappanti sul gene *CDHI* abbiamo rilevato che nessuno è associato al CCR, ma l'evidenza di associazione con la combinazione di aplotipo tra rs11642413-rs10431923, suggerisce che i polimorfismi di *CDHI* possono influenzare lo sviluppo di CCR.

Per quanto invece riguarda i livelli di espressione di *CDHI* e di due geni correlati, *CTNNB1* e *LEF1*, in biopsie di tessuto tumorale e di tessuto normale di 27 pazienti, abbiamo notato una significativa riduzione dei trascritti genici di *CDHI* e *CTNNB1* nel tessuto tumorale rispetto al tessuto normale. Ciò potrebbe essere dovuto a cause somatiche o epigenetiche piuttosto che ad un'anomalia nella linea germinale.

Abbiamo quindi genotipizzato tre SNP mappanti in tre diversi esoni del gene *MDR1* (*ABCB1*). *ABCB1* sembra avere un ruolo nel CCR, in particolare nell'insorgenza del carcinoma del sigma. Infatti, le varianti alleliche dei polimorfismi rs1128503 e rs2032582 risultano meno frequenti in tali pazienti. Inoltre, abbiamo dimostrato che polimorfismi di *ABCB1* sono differenzialmente rappresentati nei maschi rispetto alle femmine.

Suggeriamo che alcune delle differenze di incidenza del cancro del colon nei pazienti di sesso maschile rispetto al femminile siano dovute al differente pattern ormonale genere-dipendente coinvolto nella modulazione ormone/dipendente di *ABCB1* che determina un ruolo diverso dello stesso polimorfismo.

Nella genotipizzazione di quattro polimorfismi di *ROCK1* e cinque di *ROCK2* le varianti alleliche non sono risultate associate né con la comparsa di CCR né con la presenza di metastasi ai linfonodi regionali, per nessuno dei geni indagati.

Una significativa associazione è stata osservata nei casi CCR maschi per il polimorfismo rs35996865 nel gene *ROCK1*.

Ciò rappresenta la prima evidenza di una possibile correlazione tra varianti genetiche polimorfiche di *ROCK1* e l'aumento del rischio nei pazienti CCR di sesso maschile.

Per quanto riguarda la ricerca di mRNA marcatori un test di screening sul sangue intero

potrebbe rivelarsi un valore aggiunto importante per migliorare la diagnosi precoce sul carcinoma del colon-retto. L'espressione dei geni *TSPAN8*, *LGALS4* e *COLIA2*, nel sangue in toto, può rivelarsi informativa nella diagnosi del CCR.

La combinazione di *TSPAN8* e *LGALS4* mostra valori promettenti di sensibilità 92.5% e specificità 67.16%. E' interessante notare come essi evidenzino espressione speculare nel cancro e/o nel sangue normale (*TSPAN8* è più elevato nel sangue CCR, mentre *LGALS4* è più basso). Inoltre, *TSPAN8* in combinazione con *COLIA2* mostra una curva ROC con un miglioramento di specificità al 73.1%. A nostra conoscenza, questo è il primo studio che esamina l'espressione di *TSPAN8* mRNA nel sangue, per la diagnosi del CCR.

In sintesi, il nostro studio preliminare su mRNA estratto da sangue in toto, dimostra che il pannello *TSPAN8*, *LGALS4* e *COLIA2* è uno strumento affidabile per rilevare la presenza di CCR, considerando i livelli di sensibilità e specificità evidenziati.

Tutti i nostri risultati hanno un valore preliminare e richiedono conferme con un'estensione del campione di studio.

# ***Bibliografia***

1. Kasper, Braunwald E, Fauci and Harrison. Harrison - Principi di Medicina Interna (2007) New York - Milano.
2. Linee guida Tumori del colon retto, Aiom, Edizione 2015; Associazione Italiana Registri Tumori, 2015.
3. Assistenza Nazionale Tumori.
4. Gonçalves BM, Fontainhas V, Caetano AC, Ferreira A, Gonçalves R, Bastos P, Rolanda C. Oncological outcomes after endoscopic removal of malignant colorectal polyps. *Rev Esp Enferm Dig.* 2013 Sep;105(8):454-61.
5. Basseri RJ, Basseri B, Vassilaki ME, Melmed GY, Ippoliti A, Vasiliauskas EA, Fleshner PR, Lechago J, Hu B, Berel D, Targan SR, Papadakis KA. Colorectal cancer screening and surveillance in Crohn's colitis. *J Crohns Colitis.* 2012 Sep;6(8):824-9. doi: 10.1016/j.crohns.2012.01.005. Epub 2012 Feb 5.
6. Peppone LJ, Mahoney MC, Cummings KM, Michalek AM, Reid ME, Moysich KB, Hyland A. Colorectal cancer occurs earlier in those exposed to tobacco smoke: implications for screening. *J Cancer Res Clin Oncol.* 2008 Jul;134(7):743-51. doi: 10.1007/s00432-007-0332-8. Epub 2008 Feb 9.
7. Jedrychowski W, Maugeri U, Popiela T, Kulig J, Sochacka-Tatara E, Pac A, Sowa A, Musial A. Case-control study on beneficial effect of regular consumption of apples on colorectal cancer risk in a population with relatively low intake of fruits and vegetables. *Eur J Cancer Prev.* 2010 Jan;19(1):42-7. doi: 10.1097/CEJ.0b013e3283333d0cc.
8. Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, et al. Genetic alterations during colorectal-tumor development. *The New England journal of medicine.* 1988;319(9):525-32.
9. Fearon, E. R. Molecular genetics of colorectal cancer. *Annu. Rev. Pathol.* 6, 479–507 (2011).
10. Bujanda L, Sarasqueta C, Cosme A, Hijona E, Enriquez-Navascues JM, Placer C, et al. Evaluation of alpha 1-antitrypsin and the levels of mRNA expression of matrix metalloproteinase 7, urokinase type plasminogen activator receptor and COX-2 for the diagnosis of colorectal cancer. *PloS one.* 2013;8(1):e51810.
11. Hutter CM1, Chang-Claude J, Slattery ML, Pflugeisen BM, Lin Y, Duggan D, Nan H, Lemire M, Rangrej J, Figueiredo JC, Jiao S, Harrison TA, Liu Y, Chen LS, Stelling DL, Warnick GS, Hoffmeister M, Küry S, Fuchs CS, Giovannucci E, Hazra A, Kraft P, Hunter DJ, Gallinger S, Zanke BW, Brenner H, Frank B, Ma J, Ulrich CM, White E, Newcomb PA, Kooperberg C, LaCroix AZ, Prentice RL, Jackson RD, Schoen RE, Chanock SJ, Berndt SI, Hayes RB, Caan BJ, Potter JD, Hsu L, Bézieau S, Chan AT, Hudson TJ, Peters U. Characterization of gene-environment interactions for colorectal cancer susceptibility loci. *Cancer Res.* 2012 Apr 15;72(8):2036-44. doi: 10.1158/0008-5472.CAN-11-4067. Epub 2012 Feb 24.
12. Etzioni R, Urban N, Ramsey S, McIntosh M, Schwartz S, Reid B, et al. The case for early detection. *Nature reviews Cancer.* 2003;3(4):243-52.
13. Fung KY, Nice E, Priebe I, Belobrajdic D, Phatak A, Purins L, et al. Colorectal cancer biomarkers: to be or not to be? Cautionary tales from a road well travelled. *World journal of gastroenterology : WJG.* 2014;20(4):888-98.
14. Rengucci C, Maio GD, Menghi M, Scarpi E, Guglielmo S, Fusaroli P, et al. Improved Stool DNA Integrity Method for Early Colorectal Cancer Diagnosis. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology.* 2014.
15. Imperiale TF, Ransohoff DF, Itzkowitz SH, Levin TR, Lavin P, Lidgard GP, et al. Multitarget stool DNA testing for colorectal-cancer screening. *The New England journal of medicine.* 2014;370(14):1287-97.

16. Robertson DJ, Dominitz JA. Stool DNA and colorectal-cancer screening. *The New England journal of medicine*. 2014;370(14):1350-1.
17. Cristofanilli M, Hayes DF, Budd GT, Ellis MJ, Stopeck A, Reuben JM, et al. Circulating tumor cells: a novel prognostic factor for newly diagnosed metastatic breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2005;23(7):1420-30.
18. Cohen SJ, Punt CJ, Iannotti N, Saidman BH, Sabbath KD, Gabrail NY, et al. Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2008;26(19):3213-21.
19. Marshall KW, Mohr S, Khettabi FE, Nossova N, Chao S, Bao W, et al. A blood-based biomarker panel for stratifying current risk for colorectal cancer. *International journal of cancer Journal international du cancer*. 2010;126(5):1177-86.
20. Novak DJ, Liew GJ, Liew CC. GeneNews Limited: bringing the blood transcriptome to personalized medicine. *Pharmacogenomics*. 2012;13(4):381-5.
21. Osservatorio nazionale screening: [www.osservatorionazionalecreening.it](http://www.osservatorionazionalecreening.it).
22. Gruppo Italiano Screening COloRettale: [www.giscor.it](http://www.giscor.it).
23. Li L., Lv L., Liang Y., Shen X., Zhou S., Zhu J., Ma R. Association of 8q23-24 region (8q23.3 loci and 8q24.21 loci) with susceptibility to colorectal cancer: a systematic and updated meta-analysis. *Int J Clin Exp Med*. 2015 Nov 15;8(11):21001-13. eCollection 2015.
24. Fardel O., Lecureur V., Guillouzo A. The P-glycoprotein multidrug transporter. *Gen Pharmacol*. 1996 Dec;27(8):1283-91.
25. Arias A.M. (2001) Epithelial mesenchymal interactions in cancer and development. *Cell* 105(4): 425-431.
26. He X., Z. Chen, M. Jia & X. Zhao (2013) Downregulated E-cadherin expression indicates worse prognosis in Asian patients with colorectal cancer: evidence from meta-analysis. *PLoS One* 8(7): e70858.
27. Behrens J. (2005) The role of the Wnt signalling pathway in colorectal tumorigenesis. *Biochem. Soc. Trans.* 33(Pt 4): 672-675.
28. Kalluri R. & R.A. Weinberg (2009) The basics of epithelial-mesenchymal transition. *J. Clin. Invest.* 119(6): 1420-1428.
29. Thiebaut F, Tsuruo T, Hamada H, Gottesman MM, Pastan I, Willingham MC: Cellular localization of the multidrug-resistance gene product P-glycoprotein in normal human tissues. *Proc Natl Acad Sci U S A* 1987, 84:7735-7738.
30. Amano M, Ito M, Kimura K, Fukata Y, Chihara K, Nakano T, Matsuura Y, Kaibuchi K: Phosphorylation and activation of myosin by rho-associated kinase (rho-kinase). *The Journal of biological chemistry* 1996;271:20246-20249.
31. Kureishi Y, Kobayashi S, Amano M, Kimura K, Kanaide H, Nakano T, Kaibuchi K, Ito M: Rho-associated kinase directly induces smooth muscle contraction through myosin light chain phosphorylation. *The Journal of biological chemistry* 1997;272:12257-12260.
32. Kawano Y, Fukata Y, Oshiro N, Amano M, Nakamura T, Ito M, Matsumura F, Inagaki M, Kaibuchi K: Phosphorylation of myosin-binding subunit (mbs) of myosin phosphatase by rho-kinase in vivo. *The Journal of cell biology* 1999;147:1023-1038.
33. Kimura K, Ito M, Amano M, Chihara K, Fukata Y, Nakafuku M, Yamamori B, Feng J, Nakano T, Okawa K, Iwamatsu A, Kaibuchi K: Regulation of myosin phosphatase by rho and rho-associated kinase (rho-kinase). *Science* 1996;273:245-248.
34. Matsui T, Maeda M, Doi Y, Yonemura S, Amano M, Kaibuchi K, Tsukita S, Tsukita S: Rho-kinase phosphorylates cooh-terminal threonines of ezrin/radixin/moesin (erm)

- proteins and regulates their head-to-tail association. *The Journal of cell biology* 1998;140:647-657.
35. Fukata Y, Oshiro N, Kinoshita N, Kawano Y, Matsuoka Y, Bennett V, Matsuura Y, Kaibuchi K: Phosphorylation of adducin by rho-kinase plays a crucial role in cell motility. *The Journal of cell biology* 1999;145:347-361.
  36. Maekawa M, Ishizaki T, Boku S, Watanabe N, Fujita A, Iwamatsu A, Obinata T, Ohashi K, Mizuno K, Narumiya S: Signaling from rho to the actin cytoskeleton through protein kinases rock and lim-kinase. *Science* 1999;285:895-898.
  37. Riento K, Ridley AJ: Rocks: Multifunctional kinases in cell behaviour. *Nature reviews Molecular cell biology* 2003;4:446-456.
  38. Nakagawa O, Fujisawa K, Ishizaki T, Saito Y, Nakao K, Narumiya S: Rock-i and rock-ii, two isoforms of rho-associated coiled-coil forming protein serine/threonine kinase in mice. *FEBS letters* 1996;392:189-193.
  39. Yoneda A, Multhaupt HA, Couchman JR: The rho kinases i and ii regulate different aspects of myosin ii activity. *The Journal of cell biology* 2005;170:443-453.
  40. Yoneda A, Ushakov D, Multhaupt HA, Couchman JR: Fibronectin matrix assembly requires distinct contributions from rho kinases i and -ii. *Molecular biology of the cell* 2007;18:66-75.
  41. Lock FE, Ryan KR, Poulter NS, Parsons M, Hotchin NA: Differential regulation of adhesion complex turnover by rock1 and rock2. *PloS one* 2012;7:e31423.
  42. Wong CC, Wong CM, Tung EK, Man K, Ng IO: Rho-kinase 2 is frequently overexpressed in hepatocellular carcinoma and involved in tumor invasion. *Hepatology* 2009;49:1583-1594.
  43. Liu X, Choy E, Hornicek FJ, Yang S, Yang C, Harmon D, Mankin H, Duan Z: Rock1 as a potential therapeutic target in osteosarcoma. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society* 2011;29:1259-1266.
  44. Zucchini C, Manara MC, Pinca RS, De Sanctis P, Guerzoni C, Sciandra M, Lollini PL, Cenacchi G, Picci P, Valvassori L, Scotlandi K: Cd99 suppresses osteosarcoma cell migration through inhibition of rock2 activity. *Oncogene* 2014;33:1912-1921.
  45. Lane J, Martin TA, Watkins G, Mansel RE, Jiang WG: The expression and prognostic value of rock i and rock ii and their role in human breast cancer. *International journal of oncology* 2008;33:585-593.
  46. Kamai T, Yamanishi T, Shirataki H, Takagi K, Asami H, Ito Y, Yoshida K: Overexpression of rhoa, rac1, and cdc42 gtpases is associated with progression in testicular cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2004;10:4799-4805.
  47. Kamai T, Tsujii T, Arai K, Takagi K, Asami H, Ito Y, Oshima H: Significant association of rho/rock pathway with invasion and metastasis of bladder cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2003;9:2632-2641.
  48. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Science translational medicine*. 2014;6(224):224ra24.
  49. Paterlini-Brechot P, Benali NL. Circulating tumor cells (CTC) detection: clinical impact and future directions. *Cancer letters*. 2007;253(2):180-204.
  50. Langan RC, Mullinax JE, Raiji MT, Upham T, Summers T, Stojadinovic A, et al. Colorectal cancer biomarkers and the potential role of cancer stem cells. *Journal of Cancer*. 2013;4(3):241-50.
  51. Schwarzenbach H. Circulating nucleic acids as biomarkers in breast cancer. *Breast cancer research : BCR*. 2013;15(5):211.
  52. Barbazan J, Muinelo-Romay L, Vieito M, Candamio S, Diaz-Lopez A, Cano A, et al. A multimarker panel for circulating tumor cells detection predicts patient outcome and

- therapy response in metastatic colorectal cancer. *International journal of cancer Journal international du cancer*. 2014;135(11):2633-43.
53. Lagoudianakis EE, Katakaki A, Manouras A, Memos N, Papadima A, Derventzi A, et al. Detection of epithelial cells by RT-PCR targeting CEA, CK20, and TEM-8 in colorectal carcinoma patients using OncoQuick density gradient centrifugation system. *The Journal of surgical research*. 2009;155(2):183-90.
  54. Gervasoni A, Monasterio Munoz RM, Wengler GS, Rizzi A, Zaniboni A, Parolini O. Molecular signature detection of circulating tumor cells using a panel of selected genes. *Cancer letters*. 2008;263(2):267-79.
  55. Gonzalez-Masia JA, Garcia-Olmo D, Garcia-Olmo DC. Circulating nucleic acids in plasma and serum (CNAPS): applications in oncology. *OncoTargets and therapy*. 2013;6:819-32.
  56. De Maio G, Rengucci C, Zoli W, Calistri D. Circulating and stool nucleic acid analysis for colorectal cancer diagnosis. *World journal of gastroenterology : WJG*. 2014;20(4):957-67.
  57. Ganepola GA, Nizin J, Rutledge JR, Chang DH. Use of blood-based biomarkers for early diagnosis and surveillance of colorectal cancer. *World journal of gastrointestinal oncology*. 2014;6(4):83-97.
  58. Smirnov DA, Zweitzig DR, Foulk BW, Miller MC, Doyle GV, Pienta KJ, et al. Global gene expression profiling of circulating tumor cells. *Cancer research*. 2005;65(12):4993-7.
  59. Bustin SA, Mueller R. Real-time reverse transcription PCR and the detection of occult disease in colorectal cancer. *Molecular aspects of medicine*. 2006;27(2-3):192-223.
  60. Lauriola M, Ugolini G, Rosati G, Zanotti S, Montroni I, Manaresi A, et al. Identification by a Digital Gene Expression Displayer (DGED) and test by RT-PCR analysis of new mRNA candidate markers for colorectal cancer in peripheral blood. *International journal of oncology*. 2010;37(2):519-25.
  61. Solmi R, De Sanctis P, Zucchini C, Ugolini G, Rosati G, Del Governatore M, et al. Search for epithelial-specific mRNAs in peripheral blood of patients with colon cancer by RT-PCR. *International journal of oncology*. 2004;25(4):1049-56.
  62. Lenzi L, Facchin F, Piva F, Giulietti M, Pelleri MC, Frabetti F, et al. TRAM (Transcriptome Mapper): database-driven creation and analysis of transcriptome maps from multiple sources. *BMC genomics*. 2011;12:121.
  63. Piovesan A, Vitale L, Pelleri MC, Strippoli P. Universal tight correlation of codon bias and pool of RNA codons (codonome): The genome is optimized to allow any distribution of gene expression values in the transcriptome from bacteria to humans. *Genomics*. 2013;101(5):282-9.
  64. Caracausi M, Vitale L, Pelleri MC, Piovesan A, Bruno S, Strippoli P. A quantitative transcriptome reference map of the normal human brain. *Neurogenetics*. 2014;15(4):267-87.
  65. Livak K.J., Schmittgen T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>(-Delta Delta C(T)) Method. *Methods*. 2001 Dec;25(4):402-8.
  66. Burgess A.W., M.C. Faux, M.J. Layton & R.G. Ramsay (2011) Wnt signaling and colon tumorigenesis--a view from the periphery. *Exp. Cell Res*. **317**(19): 2748-2758.
  67. Wang Y., H. Yang, L. Li, H. Wang, C. Zhang & X. Xia (2011) E-cadherin (CDH1) gene promoter polymorphism and the risk of colorectal cancer : a meta-analysis. *Int. J. Colorectal Dis*. 27(2): 151-158.
  68. Geng P., Y. Chen, J. Ou, X. Yin, R. Sa & H. Liang (2012) The E-cadherin (CDH1) - C160A polymorphism and colorectal cancer susceptibility: a meta-analysis. *DNA Cell Biol*. 31(6): 1070-1077.

69. Berx G. & F. van Roy (2009) Involvement of members of the cadherin superfamily in cancer. *Cold Spring Harb. Perspect. Biol.* 1(6): a003129. Burgess A.W., M.C. Faux, M.J.
70. Moreno-Bueno G., D. Hardisson, D. Sarrío et al. (2003) Abnormalities of E- and P-cadherin and catenin (beta-, gamma-catenin, and p120ctn) expression in endometrial cancer and endometrial atypical hyperplasia. *J. Pathol.* 199(4): 471-478.
71. Zhai B., H.X. Yan, S.Q. Liu, L. Chen, M.C. Wu & H.Y. Wang (2008) Reduced expression of E-cadherin/catenin complex in hepatocellular carcinomas. *World J. Gastroenterol.* 14(37): 5665-5673.
72. Fuchs S.Y., A.V. Ougolkov, V.S. Spiegelman & T. Minamoto (2005) Oncogenic beta-catenin signaling networks in colorectal cancer. *Cell Cycle* 4(11): 1522-1539.
73. Kitaeva M.N., L. Grogan, J.P. Williams et al. (1997) Mutations in beta-catenin are uncommon in colorectal cancer occurring in occasional replication error-positive tumors. *Cancer Res.* 57(20): 4478-4481.
74. Sygut A., K. Przybyłowska, T. Ferenc et al. (2012) Genetic variations of the CTNNA1 and the CTNNB1 genes in sporadic colorectal cancer in Polish population. *Pol. Przegl. Chir.* 84(11): 560-564.
75. Terao T, Hisanaga E, Sai Y, Tamai I, Tsuji A: Active secretion of drugs from the small intestinal epithelium in rats by P-glycoprotein functioning as an absorption barrier. *J Pharm Pharmacol* 1996, 48:1083-1089.
76. Tanigawara Y: Role of P-glycoprotein in drug disposition. *Ther Drug Monit* 2000, 22:137-140.
77. Hoffmeyer S, Burk O, von Richter O, Arnold HP, Brockmoller J, John A, Cascorbi I, Gerloff T, Roots I, Eichelbaum M, Brinkmann U: Functional polymorphisms of the human multidrug-resistance gene: multiple sequence variations and correlation of one allele with P-glycoprotein expression and activity in vivo. *Proc Natl Acad Sci U S A* 2000, 97:3473-3478.
78. Hitzl M, Drescher S, van der Kuip H, Schaffeler E, Fischer J, Schwab M, Eichelbaum M, Fromm MF: The C3435T mutation in the human MDR1 gene is associated with altered efflux of the P-glycoprotein substrate rhodamine 123 from CD56+ natural killer cells. *Pharmacogenetics* 2001, 11:293-298.
79. Wang J, Wang B, Bi J, Li K, Di J: MDR1 gene C3435T polymorphism and cancer risk: a meta-analysis of 34 case-control studies. *J Cancer Res Clin Oncol* 2012, 138:979-989.
80. Miller KL, Kelsey KT, Wiencke JK, Moghadassi M, Miike R, Liu M, Wrensch M: The C3435T polymorphism of MDR1 and susceptibility to adult glioma. *Neuroepidemiology* 2005, 25:85-90.
81. Siegmund M, Brinkmann U, Schaffeler E, Weirich G, Schwab M, Eichelbaum M, Fritz P, Burk O, Decker J, Alken P, et al: Association of the P-glycoprotein transporter MDR1(C3435T) polymorphism with the susceptibility to renal epithelial tumors. *J Am Soc Nephrol* 2002, 13:1847-1854.
82. He T, Mo A, Zhang K, Liu L: ABCB1/MDR1 gene polymorphism and colorectal cancer risk: a meta-analysis of case-control studies. *Colorectal Dis* 2013, 15:12-18.
83. Kurzawski M, Drozdziak M, Suchy J, Kurzawski G, Bialecka M, Gornik W, Lubinski J: Polymorphism in the P-glycoprotein drug transporter MDR1 gene in colon cancer patients. *European journal of clinical pharmacology* 2005, 61:389-394.
84. Osswald E, John A, Laschinski G, Arjomand-Nahad F, Malzahn U, Kirchheiner J, Gerloff T, Meisel C, Mrozikiewicz PM, Chernov J, et al: Association of MDR1 genotypes with susceptibility to colorectal cancer in older non-smokers. *European journal of clinical pharmacology* 2007, 63:9-16.
85. Andersen V, Ostergaard M, Christensen J, Overvad K, Tjønneland A, Vogel U: Polymorphisms in the xenobiotic transporter Multidrug Resistance 1 (MDR1) and

- interaction with meat intake in relation to risk of colorectal cancer in a Danish prospective case-cohort study. *BMC cancer* 2009, 9:407.
86. Piekarczyk RL, Cohen D, Horwitz SB: Progesterone regulates the murine multidrug resistance *mdr1b* gene. *J Biol Chem* 1993, 268:7613-7616.
  87. Yang CP, DePinho SG, Greenberger LM, Arceci RJ, Horwitz SB: Progesterone interacts with P-glycoprotein in multidrug-resistant cells and in the endometrium of gravid uterus. *J Biol Chem* 1989, 264:782-788.
  88. Nakamura T, Okamura N, Yagi M, Omatsu H, Yamamori M, Kuwahara A, Nishiguchi K, Horinouchi M, Okumura K, Sakaeda T: Effects of ABCB1 3435C>T genotype on serum levels of cortisol and aldosterone in women with normal menstrual cycles. *Genetics and molecular research : GMR* 2009, 8:397-403.
  89. Ueda K, Okamura N, Hirai M, Tanigawara Y, Saeki T, Kioka N, Komano T, Hori R: Human P-glycoprotein transports cortisol, aldosterone, and dexamethasone, but not progesterone. *J Biol Chem* 1992, 267:24248-24252.
  90. Sainz J, Rudolph A, Hein R, Hoffmeister M, Buch S, von Schonfels W, Hampe J, Schafmayer C, Volzke H, Frank B, et al: Association of genetic polymorphisms in ESR2, HSD17B1, ABCB1, and SHBG genes with colorectal cancer risk. *Endocrine-related cancer* 2011, 18:265-276.
  91. Vishnubhotla R, Sun S, Huq J, Bulic M, Ramesh A, Guzman G, Cho M, Glover SC: Rock-ii mediates colon cancer invasion via regulation of mmp-2 and mmp-13 at the site of invadopodia as revealed by multiphoton imaging. *Laboratory investigation; a journal of technical methods and pathology* 2007;87:1149-1158.
  92. Paduch R, Kandefler-Szerszen M, Szuster-Ciesielska A, Plewka K: Transforming growth factor-beta1 modulates metalloproteinase-2 and -9, nitric oxide, rhoa and alpha-smooth muscle actin expression in colon adenocarcinoma cells. *Cell biology international* 2010;34:213-223.
  93. Adachi S, Yasuda I, Nakashima M, Yamauchi T, Yoshioka T, Okano Y, Moriwaki H, Kozawa O: Rho-kinase inhibitor upregulates migration by altering focal adhesion formation via the akt pathway in colon cancer cells. *European journal of pharmacology* 2011;650:145-150.
  94. Makrodouli E, Oikonomou E, Koc M, Andera L, Sasazuki T, Shirasawa S, Pintzas A: Braf and ras oncogenes regulate rho gtpase pathways to mediate migration and invasion properties in human colon cancer cells: A comparative study. *Molecular cancer* 2011;10:118.
  95. Sari I, Berberoglu B, Ozkara E, Oztuzcu S, Camci C, Demiryurek AT: Role of rho-kinase gene polymorphisms and protein expressions in colorectal cancer development. *Pathobiology : journal of immunopathology, molecular and cellular biology* 2013;80:138-145.
  96. Abotchie PN, Vernon SW, Du XL: Gender differences in colorectal cancer incidence in the united states, 1975-2006. *Journal of women's health* 2012;21:393-400.
  97. Gonzalez-Montelongo MC, Marin R, Gomez T, Diaz M: Androgens differentially potentiate mouse intestinal smooth muscle by nongenomic activation of polyamine synthesis and rho kinase activation. *Endocrinology* 2006;147:5715-5729.
  98. Vignozzi L, Morelli A, Filippi S, Ambrosini S, Mancina R, Luconi M, Mungai S, Vannelli GB, Zhang XH, Forti G, Maggi M: Testosterone regulates rhoa/rho-kinase signaling in two distinct animal models of chemical diabetes. *The journal of sexual medicine* 2007;4:620-630; discussion 631-632.
  99. Lin MT, Lin BR, Chang CC, Chu CY, Su HJ, Chen ST, Jeng YM, Kuo ML: Il-6 induces ags gastric cancer cell invasion via activation of the c-src/rhoa/rock signaling pathway. *International journal of cancer Journal international du cancer* 2007;120:2600-2608.

100. Naugler WE, Sakurai T, Kim S, Maeda S, Kim K, Elsharkawy AM, Karin M: Gender disparity in liver cancer due to sex differences in myd88-dependent il-6 production. *Science* 2007;317:121-124.
101. Tabur S, Oztuzcu S, Oguz E, Korkmaz H, Eroglu S, Ozkaya M, Demiryurek AT: Association of rho/rho-kinase gene polymorphisms and expressions with obesity-related metabolic syndrome. *European review for medical and pharmacological sciences* 2015;19:1680-1688.
102. Zhao R, Liu K, Huang Z, Wang J, Pan Y, Huang Y, Deng X, Liu J, Qin C, Cheng G, Hua L, Li J, Yin C: Genetic variants in caveolin-1 and rhoa/rock1 are associated with clear cell renal cell carcinoma risk in a chinese population. *PloS one* 2015;10:e0128771.
103. Kale VP, Hengst JA, Desai DH, Amin SG, Yun JK: The regulatory roles of rock and mrck kinases in the plasticity of cancer cell migration. *Cancer letters* 2015;361:185-196.
104. Hamm, A., *et al.* Tumour-educated circulating monocytes are powerful candidate biomarkers for diagnosis and disease follow-up of colorectal cancer. *Gut* (2015)
105. Fleischhacker M, Schmidt B. Circulating nucleic acids (CNAs) and cancer-a survey. *Biochimica et biophysica acta.* 2007;1775(1):181-232.
106. Schwarzenbach H, Hoon DS, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. *Nature reviews Cancer.* 2011;11(6):426-37.
107. Tsui NB, Ng EK, Lo YM. Stability of endogenous and added RNA in blood specimens, serum, and plasma. *Clinical chemistry.* 2002;48(10):1647-53.
108. Rengucci C, Maio GD, Menghi M, Scarpi E, Guglielmo S, Fusaroli P, et al. Improved Stool DNA Integrity Method for Early Colorectal Cancer Diagnosis. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology.* 2014.
109. Imperiale TF, Ransohoff DF, Itzkowitz SH. Multitarget stool DNA testing for colorectal-cancer screening. *The New England journal of medicine.* 2014;371(2):187-8.
110. Berthier-Vergnes O, Kharbili ME, de la Fouchardiere A, Pointecouteau T, Verrando P, Wierinckx A, et al. Gene expression profiles of human melanoma cells with different invasive potential reveal TSPAN8 as a novel mediator of invasion. *British journal of cancer.* 2011;104(1):155-65.
111. Bode MK, Karttunen TJ, Makela J, Risteli L, Risteli J. Type I and III collagens in human colon cancer and diverticulosis. *Scandinavian journal of gastroenterology.* 2000;35(7):747-52.
112. Zou X, Feng B, Dong T, Yan G, Tan B, Shen H, et al. Up-regulation of type I collagen during tumorigenesis of colorectal cancer revealed by quantitative proteomic analysis. *Journal of proteomics.* 2013;94:473-85.
113. Satelli A, Rao PS, Thirumala S, Rao US. Galectin-4 functions as a tumor suppressor of human colorectal cancer. *International journal of cancer Journal international du cancer.* 2011;129(4):799-809.
114. Rechreche H, Mallo GV, Montalto G, Dagorn JC, Iovanna JL. Cloning and expression of the mRNA of human galectin-4, an S-type lectin down-regulated in colorectal cancer. *European journal of biochemistry / FEBS.* 1997;248(1):225-30.
115. Fung KY, Nice E, Priebe I, Belobrajdic D, Phatak A, Purins L, et al. Colorectal cancer biomarkers: to be or not to be? Cautionary tales from a road well travelled. *World journal of gastroenterology : WJG.* 2014;20(4):888-98.
116. Winegarden N. Microarrays in cancer: moving from hype to clinical reality. *Lancet.* 2003;362(9394):1428.

117. Fukino K, Shen L, Patocs A, Mutter GL, Eng C. Genomic instability within tumor stroma and clinicopathological characteristics of sporadic primary invasive breast carcinoma. *Jama*. 2007;297(19):2103-11.
118. Patocs A, Zhang L, Xu Y, Weber F, Caldes T, Mutter GL, et al. Breast-cancer stromal cells with TP53 mutations and nodal metastases. *The New England journal of medicine*. 2007;357(25):2543-51.

## **PUBBLICAZIONI**

Systematic large-scale meta-analysis identifies a panel of two mRNAs as blood biomarkers for colorectal cancer detection.

**Rodia MT**, Ugolini G, Mattei G, Montroni I, Zattoni D, Ghignone F, Veronese G, Marisi G, Lauriola M, Strippoli P, Solmi R. *Oncotarget*. 2016 March 16. doi: 10.18632/oncotarget.8108. PMID: 26993598

Possible gender-related modulation by the ROCK1 gene in colorectal cancer susceptibility. Zucchini C, Martinelli M, De Sanctis P, **Rodia MT**, Mattei G, Ugolini G, Montroni I, Ghignone F, Solmi R. *Pathobiology*. 2015;82(6):252-8. doi: 10.1159/000439405. Epub 2015 Oct 20. PMID: 26562026.

CDH1 polymorphisms and low expression of E-cadherin and beta-catenin in colorectal cancer patients. Martinelli M, Palmieri A, **Rodia MT**, Cura F, Scapoli L, Ugolini G, Montroni I, De Sanctis P, Solmi R. *J Biol Regul Homeost Agents*. 2015 Jul-Sep;29 (3 Suppl 1): 89-96. PMID: 26511186.

Human Multidrug Resistance 1 gene polymorphisms and Idiopathic Pulmonary Fibrosis. Martinelli M, Scapoli L, Pacilli AM, Carbonare P, Girardi A, Mattei G, **Rodia MT**, Solmi R. *J Res Med Sci*. 2015 Jan;20 (1): 93-6. PMID: 25767528.

Colorectal cancer susceptibility: apparent gender-related modulation by ABCB1 gene polymorphisms. Martinelli M, Scapoli L, Cura F, **Rodia MT**, Ugolini G, Montroni I, Solmi R. *J Biomed Sci*. 2014 Sep 4;21:89. doi: 10.1186/s12929-014-0089-8. PMID: 25355168.

## **TITOLARITA' DI BREVETTI**

Diagnosi cancro colon-rettale.

Numero di Deposito: 102015000016638 (21/05/2015) **Maria Teresa Rodia**, Rossella Solmi, Mattia Lauriola, Giampaolo Ugolini, Isacco Montroni, Gabriella Mattei.