

ITALIAN PATENT OFFICE

Document No.

102011901955168A1

Publication Date

20121216

Applicant

UNIVERSITA' DEGLI STUDI DI TRIESTE

Title

METODO PER LA PROGNOSI DI TUMORI AL SENO BASATA
SULL'ESPRESSIONE DELL'ENZIMA PROLIL ISOMERARASI PIN1 IN
COMBINAZIONE CON MUTAZIONI DEL GENE TP53.

Descrizione del brevetto per invenzione industriale dal titolo:

Metodo per la prognosi di tumori al seno basata sull'espressione dell'enzima prolil isomerasi Pin1 in combinazione con mutazioni del gene *TP53*

Campo dell'invenzione

L'invenzione è relativa ad un metodo di prognosi di tumori al seno basata sulla determinazione dell'espressione dell'enzima prolil isomerasi Pin1 combinata con il rilevamento della presenza di mutazioni missenso nel gene *TP53* e sull'impiego di una *firma molecolare* data dall'espressione caratteristica di 10 geni associata all'asse Pin1/p53 mutata.

Stato della tecnica

Diverse evidenze hanno messo in luce come le mutazioni puntiformi nel gene *TP53* rappresentino presupposti cruciali nel promuovere la trasformazione tumorale e la formazione delle metastasi. Infatti, tra le più frequenti alterazioni genetiche riscontrate nei tumori umani si contano le mutazioni nel gene soppressore tumorale *TP53* (Soussi T. and Wiman K., 2007). Oltre a essere responsabili di tumori sporadici, nell'uomo le mutazioni ereditarie in *TP53* provocano la sindrome di Li-Fraumeni, un disordine caratterizzato dall'insorgenza precoce di un ampio spettro di tumori (Malkin D. et al., 1990). A causa di questo tipo di mutazioni p53 perde la capacità di legare gli elementi di risposta alla proteina "wild-type" presenti sul DNA, e, in questo modo, p53 mutata perde la capacità di svolgere le sue funzioni di soppressore tumorale e può esercitare un effetto di dominanza negativa sulla proteina "wild-type". Inoltre, diverse proteine p53 mutanti acquisiscono *de novo* una serie di attività funzionali che le rendono capaci di contribuire in maniera significativa al manifestarsi di diversi aspetti della progressione tumorale. In particolare, forti evidenze hanno messo in luce la capacità della proteina mutata di promuovere la migrazione cellulare e la formazione di metastasi (Adorno M. et al., 2009; Caulin C. et al., 2007; Muller P. et al., 2009; Terzian T. et al., 2008).

Per l'acquisizione di nuove funzioni è essenziale che p53 mutata formi complessi proteici aberranti capaci di perturbare l'attività dei partner proteici con cui interagisce. Un esempio di come p53 mutata contribuisca all'instaurarsi del

fenotipo tumorale, influenzando caratteristiche come la migrazione e l'invasione cellulare o l'instabilità genomica, si trova nell'interazione di p53 mutata con il fattore anti-metastatico p63 o con Mre11 e nella loro conseguente inattivazione (Adorno et al., 2009; Song H. et al., 2007).

Diversi studi *in vivo* hanno messo in evidenza come nelle cellule tumorali all'efficiente acquisizione di nuove funzioni da parte di p53 mutata sia associata la presenza di elevati livelli della proteina stessa (Song H. et al., 2007; Soussi T. and Beroud C., 2001; Terzian T. et al., 2008). Queste evidenze vanno a supporto dell'ipotesi secondo cui i segnali di stress potrebbero attivare p53 mutata attraverso meccanismi simili a quelli necessari a stimolare l'azione di p53 "wild-type" (Song H. et al., 2007; Terzian T. et al., 2008). I meccanismi attraverso i quali i segnali oncogenici scatenano l'azione di p53 mutata sono però ancora poco conosciuti.

Tra i fattori potenzialmente in grado di rappresentare elementi di connessione tra i segnali cellulari specifici della trasformazione tumorale e le funzioni svolte da p53 mutata, un candidato particolarmente interessante è l'enzima prolil isomerasi Pin1. Si tratta di un enzima che converte i segnali di fosforilazione presenti sui suoi substrati in cambi conformazionali che comportano una modulazione delle loro funzioni (Lu K.P. and Zhou X.Z., 2007; Yeh E. and Means A.R., 2007). Questo enzima è stato riconosciuto come un regolatore critico delle attività che la proteina p53 "wild-type" esercita nelle cellule sottoposte a stress genotossici (Mantovani F. et al., 2007; Zacchi P. et al., 2002; Zheng H. et al., 2002). Nonostante Pin1 risulti essenziale per lo svolgimento delle funzioni di p53 "wild-type", topi *Pin1*^{-/-} non sviluppano tumori (Atchinson F.W. et al., 2003). Nei tumori umani, invece, Pin1 risulta di frequente sovrespresso (Bao L. et al., 2004) ed è stato dimostrato che in cellule mammarie esso è in grado di favorire gli eventi di trasformazione tumorale che dipendono dalle proteine Her2/Neu/Ras o da Notch1 (Rustighi A. et al., 2009; Wulf G. et al., 2004).

Non è invece noto come durante la trasformazione tumorale il ruolo che Pin1 svolge nei sistemi fisiologici di sorveglianza della cellula venga alterato rendendolo un cruciale amplificatore delle funzioni oncogeniche di p53 mutata.

Le mutazioni in *TP53* rappresentano, in particolare, una delle alterazioni che frequentemente caratterizzano i tumori mammari (Langerød A. et al., 2007).

In generale il carcinoma mammario nel sesso femminile rappresenta non solo la forma di cancro più frequentemente diagnosticata ma anche la principale causa di morte (Jemal A et al., 2011) e il sottotipo *basal-like* la sua forma a più alto rischio di recidiva. Si stima che nel mondo ogni anno le donne alle quali viene diagnosticato superino il milione (Coughlin S. et al., 2009).

Il cancro al seno, di fatto, è una malattia caratterizzata da una forte eterogeneità sotto diversi punti di vista, da quello istologico, a quello della risposta ai trattamenti terapeutici o a quello delle modalità di disseminazione delle metastasi nei diversi distretti corporei (Prat A. et al., 2011).

Negli anni i progressi realizzati nel campo della diagnosi, i programmi di prevenzione e lo sviluppo di nuove strategie terapeutiche hanno portato a una riduzione della mortalità dovuta a questo tipo di tumore (Jemal A. et al., 2009). Tuttavia è stato stimato che ogni anno tra gli Stati Uniti e l'Europa sono più di 120 mila le donne che muoiono per la malattia. (Jemal A. et al., 2009; La Vecchia C. et al., 2010). In parte questo si spiega con il fatto che alla complessità propria di questo tumore non si associa una completa conoscenza della sua biologia e tanto meno la disponibilità di marcatori da utilizzare nella pratica clinica che rispecchino interamente la sua eterogeneità, che permettano di effettuare in ogni caso una prognosi accurata e di determinare la probabilità di successo delle terapie, elementi questi essenziali per guidare le scelte terapeutiche e, dove necessario, lo sviluppo di nuove strategie di intervento.

Attualmente, nella pratica clinica, la diagnosi di carcinoma mammario viene effettuata attraverso l'esame istopatologico del tessuto. La prognosi e la scelta della terapia dipendono da questo tipo di indagine, dalla valutazione di parametri clinici quali le dimensioni del tumore, il suo grado, l'età della paziente e l'invasione dei linfonodi da parte delle cellule tumorali e da una caratterizzazione molecolare del tumore basata sull'espressione dei recettori ormonali, di HER2 e di alcuni marcatori della proliferazione. Questi, che rappresentano metodi standard per effettuare la prognosi, non sempre sono efficaci nel predire la progressione della malattia.

La possibilità di effettuare valutazioni sempre più accurate dei fattori di rischio e di formulare previsioni più precise sul decorso clinico e sull'esito della malattia avrebbero, invece, un notevole impatto dal punto di vista del trattamento dei tumori mammari, poiché permetterebbero di stratificare meglio le pazienti. Questa necessità ha portato all'idea di affiancare agli esami normalmente effettuati una caratterizzazione molecolare più approfondita del tumore. È stato, così, sviluppato (e approvato dalla FDA statunitense) un test per la prognosi di tumori mammari, il MammaPrint[®], basato sull'analisi attraverso la tecnologia del cDNA microarray dell'espressione di più di 70 geni, il cui valore ai fini prognostici è stato dimostrato attraverso gli studi condotti da van't Veer et al. (van't Veer L. et al., 2002).

In questo contesto si inserisce la presente invenzione, offrendo il vantaggio di un'analisi ad elevato valore predittivo dal punto di vista prognostico, basata però sulla determinazione di un numero ridotto di parametri e sull'utilizzo di tecnologie più semplici e quindi di più facile e immediata esecuzione.

Sommario

La presente invenzione si basa su evidenze sperimentali che attestano come Pin1 esalti la trasformazione tumorale in un modello murino della sindrome di Li-Fraumeni ed amplifichi le funzioni oncogeniche di p53 mutata. In linee cellulari umane derivate da carcinoma mammario, gli inventori hanno trovato che le due proteine insieme istruiscono un programma trascrizionale che favorisce l'aggressività del tumore ed in questo programma sono coinvolti 10 geni (che costituiscono una *firma molecolare* associata all'asse Pin1/p53 mutata), di cui 7 sono risultati sufficienti per la valutazione prognostica.

Hanno inoltre trovato che nei tumori mammari umani la sovraespressione di Pin1 combinata con la presenza di mutazioni "missenso" in *TP53* correla con l'esito infausto della malattia.

In un primo aspetto, quindi, l'invenzione concerne un metodo per la prognosi di tumori al seno comprendente la determinazione in campioni biotici dei livelli di espressione di Pin1 combinata con la rilevazione della presenza di mutazioni nel gene *TP53* e/o con la determinazione dei livelli della proteina p53, data la correlazione esistente in campioni tumorali tra la presenza di elevati livelli della proteina p53 e la presenza di mutazioni missenso nella stessa.

In un secondo aspetto, l'invenzione concerne un metodo per la prognosi di tumori al seno che comprende anche la rilevazione dei livelli di espressione della *firma molecolare* associata a Pin1/p53 mutata attraverso la determinazione dell'espressione dei geni *DEPDC1* (Gene ID No 55635, isoforma 1: RefSeq NM_001114120.1, NP_001107592.1; isoforma 2: RefSeq NM_017779.4, NP_060249.2), *CPSF6* (Gene ID No 11052, RefSeq NM_007007.2, NP_008938.2), *C21orf45* (Gene ID No 54069, RefSeq NM_018944.2, NP_061817.1), *CENPA* (Gene ID No 1058, isoforma 1: RefSeq NM_001809.3, NP_001800.1; isoforma 2: RefSeq NM_001042426.1, NP_001035891.1), *FAM64A* (Gene ID No 54478, RefSeq NM_019013.2, NP_061886.2), *CCNE2* (Gene ID No 9134, RefSeq NM_057749.2, NP_477097.1), *BUB1* (Gene ID No 699, RefSeq NM_004336.3, NP_004327.1), *EPB41L4B* (Gene ID No 54566, isoforma 1: RefSeq NM_018424.2, NP_060894.2; isoforma 2: RefSeq NM_019114.3, NP_061987.3), *NCAPH* (Gene ID No 23397, RefSeq NM_015341.3, NP_056156.2), *WDR67* (Gene ID No 93594, isoforma 1: RefSeq NM_145647.3, NP_663622.2; isoforma 2: RefSeq NM_001145088.1, NP_001138560.1). Tra questi 10 geni si è dimostrata essere sufficiente, ai fini prognostici, anche la sola rilevazione dell'espressione dei 7 geni: *DEPDC1*, *CPSF6*, *C21orf45*, *FAM64A*, *EPB41L4B*, *NCAPH*, *WDR67* che quindi possono anche essere considerati come una *firma molecolare* minima. L'utilizzo di questa *firma molecolare*, in particolare, offre anche il vantaggio di stratificare le pazienti con tumore ER+ (estrogeno positivo). Attraverso il metodo proposto dagli inventori, infatti, è possibile individuare all'interno di questo gruppo di casi, tipicamente considerato omogeneo, due diversi sottogruppi che mostrano differenze significative dal punto di vista prognostico. La stratificazione consentita da tale metodo potrebbe portare a una scelta terapeutica diversa rispetto a quella che attualmente si opera in questo casi, permettendo di identificare quelle pazienti che all'interno del gruppo di casi ER+ potrebbero o meno beneficiare della combinazione tra ormono- e chemioterapia.

In un ulteriore aspetto, l'invenzione concerne un kit per l'esecuzione del metodo per la prognosi di tumori basato sulla determinazione dei livelli di pin1 combinata con la rilevazione della presenza di mutazioni nel gene *TP53* e/o con la

determinazione dei livelli della proteina p53 e/o la rilevazione dei livelli di espressione dei geni della *firma molecolare* associata a Pin1/p53 mutata.

Breve descrizione delle figure

Figura 1. La figura mostra come Pin1 esalta la tumorigenesi in topi “*knockin*” per p53 mut. **(A)** Curve di Kaplan-Meier della sopravvivenza in assenza di tumore nelle coorti di topi $p53^{M/M} Pin^{+/+}$ (n=27) e $p53^{M/M} Pin^{-/-}$ (n=22) ($X^2=4$, $P=0,02$, test “long-rank” di Mantel-Haenszel). **(B)** Curve di Kaplan-Meier della sopravvivenza in assenza di tumore nelle coorti di topi $p53^{M/+} Pin^{+/+}$ (n=19) e $p53^{M/+} Pin^{-/-}$ (n=17) ($X^2=5,4$, $P=0,045$, test “long-rank” di Mantel-Haenszel). **(A)** e **(B)** $p53^{+/+} Pin^{+/+}$ (n=22) e $p53^{+/+} Pin^{-/-}$ (n=18). **(C)** Pannello di sinistra: Il grafico a istogrammi mostra la percentuale di animali che sviluppano tumori tra i topi che presentano il genotipo indicato. Pannello di destra: Grafico a torta che rappresenta la percentuale rispetto al totale dei diversi tipi di tumori presenti nei topi $p53^{M/+} Pin^{+/+}$ e $p53^{M/+} Pin^{-/-}$. **(D)** Pannello di sinistra: Grafico a istogrammi che rappresenta la crescita indipendente dall'ancoraggio al substrato di fibroblasti embrionali murini (MEF) trasdotti con un costrutto per l'espressione di H-Ras^{V12}. Le barre di errore indicano la d.s. (n = 3), P relativo alle MEF $p53^{M/M} Pin^{+/+} < 0,01$ in tutti i casi (test- t a due code). Pannello di destra: Il grafico a barre rappresenta il volume dei tumori che hanno avuto origine dalle MEF trasdotte con H-Ras^{V12} indicate, in seguito all'iniezione in topi SCID NOD. Il volume è stato misurato dopo il sacrificio degli animali. Il grafico mostra il volume medio e la d.s. dei tumori derivati da gruppi di 5 topi per ogni tipo cellulare iniettato. $P < 0,01$ (test- t a due code). In basso sono mostrate alcune immagini rappresentative dei tumori. **(E)** Pannello in alto: Analisi per “western blot” dei siti S-P o T-P fosforilati (anticorpo MPM-2) su livelli normalizzati della p53 mutata p53R172H immunoprecipitata da cellule MEF $p53^{M/M} Pin^{+/+}$ infettate con un costrutto per l'espressione di H-Ras^{V12} (+) o con il vettore vuoto (-). NRA: anticorpo non correlato. Pannello in basso: Analisi per “western blot” dei saggi di pull-down effettuati utilizzando la proteina GST-Pin1 ricombinante. I lisati cellulari impiegati nell'analisi sono stati normalizzati come descritto sopra.

Figura 2. La figura mostra: **(A)** Curve della sopravvivenza in assenza di tumore nelle coorti di topi $p53^{+/-} Pin^{+/+}$ (n=25) e $p53^{+/-} Pin^{-/-}$ (n=25). ($X^2 = 0,2$, $P = 0,68$, test

long-rank di Mantel-Haenszel). $p53^{+/+}$ $Pin^{+/+}$ (n=22) e $p53^{+/+}$ $Pin^{-/-}$ (n=18). **(B)** Grafico a istogrammi che mostra in ciascun gruppo la percentuale di topi che ha sviluppato tumori. **(C)** MEF trasdotte con il vettore per l'espressione di H-Ras^{V12} sono state iniettate a livello sottocutaneo in topi NOD SCID e dopo aver sacrificato gli animali è stato misurato il volume dei tumori generatisi. Il grafico mostra il volume medio e la d.s. dei tumori prelevati da 5 topi che avevano subito l'iniezione delle cellule del tipo indicato. Sono mostrate due immagini rappresentative dei tumori rilevati.

Figura 3. La figura mostra i risultati che dimostrano che Pin1 amplifica gli effetti che p53 mutata esercita a favore della migrazione. **(A)** Saggi di migrazione attraverso i pozzetti di cellule umane della linea MDA-MB-231 derivata da tumore mammario trasfettate con i siRNA indicati (RNAi). Le barre di errore indicano la d.s. (n = 3). Sono stati analizzati gli effetti sulla migrazione di due diverse sequenze di siRNA dirette contro Pin1 (Pin1 (A) e Pin1 (B)). **(B)** Saggi di migrazione attraverso i pozzetti di cellule MDA-MB-231 trasfettate con un siRNA di controllo (C) o con il siRNA Pin1 (A), dopo trasduzione di vettori retrovirali per l'espressione delle forme resistenti al siRNA Pin1 HA-Pin1r, HA-Pin1r S67E (inattivo dal punto di vista catalitico) o del vettore vuoto (-). Le barre di errore indicano la d.s. (n = 3). **(C)** Saggi di invasione dei polmoni in seguito a iniezione nella vena caudale di cellule MDA-MB-231 private di Pin1. Pannello di sinistra: Analisi istologica dell'invasione polmonare. Colorazione con ematossilina ed eosina di sezioni rappresentative degli interi lobi polmonari prelevati da topi inoculati con cellule infettate con un vettore per l'espressione di shLacZ (controllo) o con uno per shPin1 (6 topi per ciascun gruppo), le frecce indicano alcune metastasi rappresentative (immagine in alto). Rilevamento per immunoistochimica dell'espressione di Pin1 (magnificazione x12,5) (immagini in basso). I riquadri mostrano un'immagine dettagliata dell'area indicata dalle frecce (magnificazione x200). Pannello di destra: Valutazione computerizzata della percentuale di area di tessuto polmonare occupata da metastasi. I dati sono riportati in istogrammi come percentuali medie $\pm$ SEM (test della somma dei ranghi di Mann-Whitney, $P = 0,009$). **(D)** Saggi di migrazione attraverso i pozzetti di cellule MDA-MB-231 trasfettate con un siRNA specifico per p53 o di controllo (C), dopo trasduzione di

vettori retrovirali pMSCV (-) o pMSCV HA-Pin1. Le barre di errore indicano la d.s. (n = 3).

Figura 4. La figura mostra che Pin1 e p53 mutata riprogrammano l'espressione genica. **(A)** Il diagramma di Venn mostra i geni contemporaneamente perturbati nelle cellule MDA-MB-231 in seguito alla deplezione di Pin1 o a quella di p53 mutata: geni regolati positivamente (freccia che sale); geni regolati negativamente (freccia che scende). **(B)** Curve di sopravvivenza di Kaplan-Meier costruite considerando l'intervallo di tempo tra la diagnosi del tumore primario e quella di metastasi in altri distretti corporei in pazienti con carcinoma mammario classificati sulla base dei dati di espressione dei geni appartenenti alla *firma molecolare* associata a Pin1/p53 mutata. Linea tratteggiata: la curva dei casi in cui i 10 geni sono espressi a elevati livelli; linea continua: la curva dei casi in cui i 10 geni sono espressi a bassi livelli. Pannello di sinistra: Serie di dati di Desmedt (Desmedt C. et al., 2007) ($X^2=25,6$, $P=4.1 \times 10^{-7}$, $n=198$, test "log-rank"). Pannello di destra: Serie di dati di Sotiriou (Sotiriou C. et al., 2006) ($X^2=7,2$, $P=0,0074$, $n=179$, test "log-rank"). **(C)** Grafico a istogrammi che mostra le variazioni nei livelli di espressione dei geni appartenenti alla *firma molecolare* associata a Pin1/p53 mutata e di un gene utilizzato come controllo (*PARP3*) (Gene ID 10039, isoforma 1: RefSeq NM_001003931.2, NP_001003931.2; isoforma 2: RefSeq NM_005485.4, NP_005476.3) che non viene influenzato. Le analisi sono state effettuate tramite qRT-PCR in cellule MDA-MB-231 dopo co-trasduzione dei vettori pMSCV o pMSCV HA-Pin1 con i costrutti codificanti le forme mutanti di p53 indicate in figura o con il vettore vuoto (-), usato come controllo negativo. (n=3, *: $P < 0,001$ se si confrontano i livelli di espressione di ciascun gene della *firma molecolare* associata a Pin1/p53 mutata nelle diverse condizioni indicate, test-t a due code). **(D)** Analisi di ChIP sulla cromatina legata da p53 mutata (IP p53) estratta da cellule MDA-MB-231 dopo trasduzione con vettori per l'espressione di shRNA diretti contro Pin1 endogeno (shPin1), contro p53 mutata (shp53) o contro LacZ per controllo negativo (shLacZ). La cromatina è stata immunoprecipitata con l'anticorpo DO1 o con IgG di topo purificate, per controllo negativo (NRA). La presenza del fattore sui promotori è stata analizzata per qRT-PCR ed è stata quantificata come percentuale della cromatina utilizzata nell'input. (n=3, *: $P <$

0,013 se si confronta nelle condizioni indicate quanto il fattore sia presente sul promotore di ciascun gene appartenente alla *firma molecolare* associata a Pin1/p53 mutata, senza considerare il controllo negativo, test-*t* a due code). **(E)** Analisi di ChIP sulla cromatina legata da Pin1 (IP Pin1) estratta da cellule MDA-MB-231 dopo trasduzione con vettori per l'espressione di shRNA diretti contro la proteina Pin1 endogena (shPin1), contro p53 mutata (shp53) o contro LacZ per controllo negativo (shLacZ). La cromatina è stata immunoprecipitata con un anticorpo policlonale diretto contro Pin1 o con IgG di coniglio purificate per controllo negativo (NRA). La presenza dell'enzima sui promotori è stata analizzata come descritto in (D). Il controllo in (D) e in (E) (gli istogrammi in bianco) è rappresentato da una regione del promotore di DEPDC1 (da -2112 a -1909 rispetto alla sequenza NM_032977.9) non legata né da p53 mutata né da Pin1. Da (C) a (E) le barre di errore indicano la d.s. ($n = 3$, test-*t* a due code).

Figura 5. **(A)** Validazione per qRT-PCR dei geni bersaglio di Pin1/p53 mutata selezionati. Le variazioni dei livelli di espressione genica sono riportate in termini di quante volte un gene viene differenzialmente espresso (*fold change*) nelle cellule MDA-MB-231 dopo l'abbattimento dei livelli di Pin1 (siPin1) o di p53 mutata (sip53) rispetto a quanto succede nello stesso tipo di cellule trasfettate con il siRNA di controllo. Le barre d'errore indicano la d.s. ($n=3$). **(B)** Effetto prodotto dall'abbattimento dei livelli di Pin1 o di p53 mutata in cellule MDA-MB-468 sull'espressione dei geni bersaglio di Pin1/p53 mutata indicati. I livelli degli mRNA sono stati determinati per qRT-PCR dopo trasfezione dei siRNA indicati. Le variazioni dei livelli di espressione genica sono riportate in termini di quante volte un gene viene differenzialmente espresso dopo l'abbattimento dei livelli di Pin1 (siPin1) o di p53 mutata (sip53) rispetto alla condizione in cui le cellule sono state trasfettate con il siRNA di controllo. Le barre d'errore indicano la d.s. ($n=3$. *: P value < 0,028 se si confronta l'espressione di ciascun gene della *firma molecolare* associata a Pin1/p53 mutata in cellule trasfettate con i siRNA indicati rispetto al controllo, test-*t* a due code).

Figura 6. Da **(A)** a **(D)** analisi di Kaplan-Meier della sopravvivenza di pazienti affette da tumore mammario classificate sulla base dell'espressione dei 10 geni della *firma molecolare* associata a Pin1/p53 mutata. Linea tratteggiata: i casi con

livelli elevati di espressione dei 10 geni. Linea continua: i casi con bassi livelli di espressione dei 10 geni:

(A) Set di dati Desmedt ($X^2=25.7$, $P=4,1 \times 10^{-7}$, $n=198$, test log-rank);

(B) Set di dati Miller ($X^2=9.7$, $P=0,002$, $n=251$, test log-rank);

(C) Set di dati Pawitan ($X^2=7.7$, $P=0,00557$, $n=187$, test log-rank);

(D) Set di dati Sotiriou ($X^2=11.6$, $P=0,00065$, $n=189$, test log-rank).

(E) ed (F): analisi di Kaplan-Meier della sopravvivenza delle pazienti con tumore mammario positivo al recettore dell'estrogeno (ER+). Linea tratteggiata: i casi con elevati livelli di espressione dei 10 geni. Linea continua: i casi con bassi livelli di espressione dei 10 geni:

(E) Set di dati Desmedt ($X^2=14.3$, $P=0.00016$, $n=134$, test log-rank);

(F) Set di dati Miller ($X^2=8.6$, $P=0.00334$, $n=211$, test log-rank).

Figura 7. La figura mostra che nel carcinoma della mammella la presenza di livelli elevati di Pin1 influenza il valore prognostico della presenza di mutazioni missenso in p53. **(A)** Curve di Kaplan-Meier della sopravvivenza di pazienti con tumore alla mammella (limite di osservazione a 10 anni, variabile clinica = OS) valutata in funzione della combinazione della sovraespressione di Pin1 e della presenza di mutazioni missenso in *TP53*. Pin1 basso: tumori che non sovraesprimono Pin1, Pin1 alto: tumori che sovraesprimono Pin1 ($X^2=17,2$, $P < 10^{-4}$, $n=212$, test "log-rank"). **(B)** Analisi multivariata del rischio di morte. L'analisi multivariata è stata realizzata sulla stessa coorte di pazienti analizzata in (A) utilizzando il modello di regressione a rischi proporzionali di Cox. È mostrata la relazione esistente tra sopravvivenza e presenza di livelli elevati di Pin1 e di mutazioni missenso in *TP53* o altri parametri impiegati nella pratica clinica. **(C)** Curve di Kaplan-Meier della sopravvivenza in casi di tumore alla mammella con elevati livelli di espressione di Pin1 (limite di osservazione a 10 anni, variabile clinica = OS), in funzione della presenza di p53 con mutazioni missenso o "wild-type" ($X^2=14,1$, $P < 10^{-4}$, $n=143$, test "log-rank"). **(D)** Curve di Kaplan-Meier della sopravvivenza in casi di tumore alla mammella con bassi livelli di espressione di Pin1 (limite di osservazione a 10 anni, variabile clinica = OS), in funzione della presenza di p53 con mutazioni missenso o "wild-type" ($X^2=0,4$, $P=0,548$, $n=69$, test "log-rank").

Figura 8. **(A)** Analisi di Kaplan-Meier della sopravvivenza (OS) a dieci anni nei casi in cui Pin1 è risultato (linea continua) o non è risultato sovraespresso (linea tratteggiata). ($X^2=0,0148506$, $P=0,903$, $n=212$, test log-rank). **(B)** Analisi di Kaplan-Meier della sopravvivenza (OS) a dieci anni nei casi in cui *TP53* è risultato mutato (linea continua) o “wild type” (linea tratteggiata). ($X^2=21,43$, $P<0,001$, $n=212$, test log-rank). **(C)** Analisi multivariata secondo il modello a rischi proporzionali di Cox realizzata su pazienti sottoposte a trattamento chemioterapico adiuvante a base di antraciclina, considerate in funzione della sovraespressione di Pin1 e della presenza di mutazioni missenso in *TP53* ($n=66$).

Figura 9. Schema che riassume ciò che principalmente emerge sul ruolo di Pin1 e di p53 mutata a favore del fenotipo tumorale aggressivo.

Descrizione dettagliata dell'invenzione

Definizioni

Per gene *TP53* si intende il gene (Gene ID No 7157, RefSeq NM_000546.4, NP_000537.3) codificante per la proteina nativa (o *wild-type*) p53 umana (GenBank accession No ABB80262).

L'enzima prolil isomerasi Pin1 è la proteina (GenBank accession No AAC50492.1) codificata dal gene *Pin1* (Gene ID No 5300, RefSeq NM_006221.2, NP_006212.1).

Per *firma molecolare* minima associata all'asse Pin1/p53 mutata si intende il gruppo costituito dai seguenti sette geni:

DEPDC1 (DEP domain containing 1) identificato con numero di Gene ID No 55635, *Homo sapiens* (isoforma 1: RefSeq NM_001114120.1, NP_001107592.1; isoforma 2: RefSeq NM_017779.4, NP_060249.2);

CPSF6 (cleavage and polyadenylation specific factor 6, 68kDa) identificato con numero di Gene ID No 11052, *Homo sapiens* RefSeq NM_007007.2, NP_008938.2);

C21orf45 (chromosome 21 open reading frame 45) identificato con numero di Gene ID No 54069, *Homo sapiens* RefSeq NM_018944.2, NP_061817.1;

FAM64A (family with sequence similarity 64, member A) identificato con numero di Gene ID No 54478, *Homo sapiens* RefSeq NM_019013.2, NP_061886.2;

EPB41L4B (erythrocyte membrane protein band 4.1 like 4B) identificato con numero di Gene ID No 54566, *Homo sapiens* (isoforma 1: RefSeq NM_018424.2, NP_060894.2; isoforma 2: RefSeq NM_019114.3, NP_061987.3);

NCAPH (non-SMC condensin I complex, subunit H) identificato con numero di Gene ID No 23397, *Homo sapiens* RefSeq NM_015341.3, NP_056156.2;

WDR67 (WD repeat domain 67) identificato con numero di Gene ID No 93594, *Homo sapiens* (isoforma 1: RefSeq NM_145647.3, NP_663622.2; isoforma 2: RefSeq NM_001145088.1, NP_001138560.1).

Per *firma molecolare* associata all'asse Pin1/p53 mutata si intende l'insieme dei geni della *firma molecolare* minima prima indicata e dei seguenti tre geni:

CENPA (centromere protein A) identificato con numero di Gene ID No 1058, *Homo sapiens* (isoforma 1: RefSeq NM_001809.3, NP_001800.1; isoforma 2: RefSeq NM_001042426.1, NP_001035891.1);

CCNE2 (cyclin E2) identificato con numero di Gene ID No 9134, *Homo sapiens* RefSeq NM_057749.2, NP_477097.1

BUB1 (budding uninhibited by benzimidazoles 1 homolog (yeast)) identificato con numero di Gene ID No 699, *Homo sapiens* RefSeq NM_004336.3, NP_004327.1.

Per "score della *firma molecolare*" si intende il punteggio associato ad ogni campione, calcolato utilizzando il valore di espressione dei geni della *firma molecolare* e il valore di espressione di tre ulteriori geni utilizzati come soglia di riferimento e definiti "geni di riferimento".

Per "geni di riferimento" si intendono geni la cui espressione sia paragonabile come intensità ai geni della *firma molecolare* e la cui presenza sia costante nei campioni biotici, come ad esempio:

ERCC6 (excision repair cross-complementing rodent repair deficiency, complementation, group 6) identificato con numero di Gene ID No 2074 *Homo sapiens* RefSeq. NM_000124.2, NP_000115.1;

ADCY3 (adenylate cyclase 3) identificato con numero di Gene ID No 109, *Homo sapiens* RefSeq. NM_004036.3, NP_004027.2;

TUBGCP2 (tubulin, gamma complex associated protein 2) identificato con numero di Gene ID No 10844, *Homo sapiens* RefSeq. NM_006659.2, NP_006650.1;

SOX15 (sex determining region Y-box 15) identificato con numero di Gene ID No 6665 *Homo sapiens* RefSeq. NM_006942.1, NP_008873.1;

RPL10 (ribosomal protein L10) identificato con numero di Gene ID No 6134 *Homo sapiens* RefSeq. NM_006013.3, NP_006004.2;

DGCR14 (DiGeorge syndrome critical region gene 14) identificato con numero di Gene ID No 8220 *Homo sapiens* RefSeq. NM_022719.2, NP_073210.1;

KIAA0586 identificato con numero di Gene ID No 9786 *Homo sapiens* RefSeq. NM_014749.3, NP_055564.3;

ZBTB17 (zinc finger and BTB domain containing 17) identificato con numero di Gene ID No 7709 *Homo sapiens* RefSeq. NM_003443.2, NP_003434.2;

TFIP11 (tuftelin interacting protein 11) identificato con numero di Gene ID 24144 *Homo sapiens* RefSeq. NM_001008697.1, NP_001008697.1;

ZNF672 (zinc finger protein 672) identificato con numero di Gene ID No 79894 *Homo sapiens* RefSeq. NM_024836.1, NP_079112.1;

TXN2 (thioredoxin 2) identificato con numero di Gene ID No 25828 *Homo sapiens* RefSeq. NM_012473.3, NP_036605.2.

I geni di riferimento preferiti sono: *ERCC6*, *ADCY3*, *TUBGCP2*.

Lo score della *firma molecolare* viene declinato in tre livelli: score basso, score medio o score alto. La procedura e l'applicazione dello score della *firma molecolare* sono descritti in dettaglio di seguito nella "Descrizione".

Descrizione

Diversi studi realizzati utilizzando un modello murino di "*knockin*" per p53 mutata hanno messo in luce come, per essere attivata funzionalmente, p53 mutata abbia bisogno della presenza di un contesto oncogenico. Con le evidenze scientifiche alla base della presente invenzione viene fornito uno degli elementi di collegamento che fanno da ponte tra i segnali oncogenici e l'acquisizione del fenotipo aggressivo, dimostrando come nelle cellule tumorali l'enzima Pin1 e p53 mutata vengano integrati in un asse molecolare attivato in seguito alla fosforilazione di p53 mutata a carico di specifici siti di fosforilazione.

Tale conclusione è avvalorata dai risultati ottenuti dalla ricerca eseguita e riportati qui di seguito in dettaglio.

In particolare, depongono a favore dell'importanza di questa isomerasi negli eventi che portano alla trasduzione dei segnali oncogenici e all'esplicitarsi delle attività di p53 mutata le seguenti osservazioni:

- in topi che esprimono p53 mutata, ma non Pin1, la trasformazione tumorale risulta significativamente limitata;
- l'aumento dei fenomeni di tumorigenesi è risultato dipendere strettamente da Pin1 negli esperimenti condotti sulle cellule MEF $p53^{MM}$ trasdotte con H-Ras^{V12};
- la presenza di segnali oncogenici in queste cellule ha provocato un incremento nell'espressione di Pin1 e in parallelo la fosforilazione di p53 mutata in siti S/T-P, portando così le due proteine a interagire.

Sulla base dei risultati riportati di seguito in dettaglio, si può dedurre che Pin1 sia importante in questi sistemi modello come elemento critico di connessione tra la presenza di segnali oncogenici e la completa attivazione di p53 mutata, che si traduce nell'aumento dell'instabilità genomica e nello sviluppo di metastasi.

Si è osservato anche che, nelle cellule di carcinoma mammario, la presenza di segnali aberranti porta alla fosforilazione dei siti S/T-P della proteina p53 mutata favorendo il riconoscimento della proteina da parte di Pin1. Questo evento è necessario affinché si scateni l'effetto di p53 mutata nel promuovere la migrazione e l'invasione cellulare.

I dati prodotti dagli inventori mostrano, inoltre, come l'azione combinata di Pin1 e p53 mutata vada a favore di un fenotipo tumorale aggressivo attraverso l'induzione di un programma trascrizionale specifico. Tra i bersagli di Pin1/p53 mutata sono stati identificati 10 geni che costituiscono una *firma molecolare* associata all'asse Pin1/p53 mutata: *DEPDC1* (Gene ID No 55635), *CPSF6* (Gene ID No 11052), *C21orf45* (Gene ID No 54069), *CENPA* (Gene ID No 1058), *FAM64A* (Gene ID No 54478), *CCNE2* (Gene ID No 9134), *BUB1* (Gene ID No 699), *EPB41L4B* (Gene ID No 54566), *NCAPH* (Gene ID No 23397), *WDR67* (Gene ID No 93594).

È stato trovato che tale *firma molecolare* correla significativamente con l'esito clinico della malattia, come risulta dall'analisi di quattro serie indipendenti di dati

relativi a diversi casi di tumore mammario, su cui è stata eseguita un'analisi di Kaplan-Meier:

- set di dati Desmet (Desmet. C., et al., 2007);
- set di dati Miller (Miller L.D. et al., 2005);
- set di dati Sotiriou (Sotiriou C. et al., 2006);
- set di dati Pawitan (Pawitan Y. Et al., 2005).

Per arrivare a definire questa *firma molecolare*, dunque, gli inventori sono partiti dall'analisi dell'espressione genica globale, in condizioni di silenziamento sia di Pin1 sia di p53 mutata, in un modello cellulare umano di carcinoma mammario triplo-negativo. Successivamente hanno verificato l'importanza del programma trascrizionale indotto da Pin1 e da p53 mutata in tumori primari attraverso l'interrogazione di quattro serie indipendenti di dati relativi a più di 800 casi di tumore al seno. Da tale analisi è risultato che la sovraespressione dei geni maggiormente regolati in condizioni di silenziamento sia di Pin1 sia di p53 mutata (31 geni) veniva riscontrata nei tumori di pazienti aventi una prognosi peggiore. A partire da questo gruppo di 31 geni, è stata ottenuta una *firma molecolare* sulla base del punteggio di correlazione dell'espressione di ciascun gene con i dati clinici, utilizzando un modello a rischi proporzionali di Cox. Per la classificazione dei geni così ottenuta si veda la Tabella 1.

Tabella 1. Classificazione dei geni sulla base della correlazione tra dati di espressione e dati clinici.

<i>Simbolo gene</i>	Gene ID No RefSeq	Punteggio
<i>CPSF6</i>	11052 NM_007007.2, NP_008938.2	9.65
<i>C21orf45</i>	54069 NM_018944.2, NP_061817.1	9.35
<i>CENPA</i>	1058 NM_001809.3, NP_001800.1	8.55
<i>FAM64A</i>	54478 NM_019013.2, NP_061886.2	8.06
<i>CCNE2</i>	9134 NM_057749.2, NP_477097.1	7.44
<i>BUB1</i>	699	5.56

	NM_004336.3, NP_004327.1	
<i>EPB41L4B</i>	54566 isoforma 1: NM_018424.2, NP_060894.2; isoforma 2: NM_019114.3, NP_061987.3	5.23
<i>NCAPH</i>	23397 NM_015341.3, NP_056156.2	5.18
<i>DEPDC1</i>	55635 isoforma 1: NM_001114120.1, NP_001107592.1; isoforma 2: NM_017779.4, NP_060249.2	3.74
<i>WDR67</i>	93594 NM_001145088.1, NP_01138560.1	3.61
<i>NMU</i>	10874 NM_006681.2, NP_006672.1	2.87
<i>FBP1</i>	2203 NM_000507.3, NP_000498.2	2.47
<i>GFRA1</i>	2674 isoforma 1: NM_001145453.1, NP_001138925.1; isoforma 2: NM_005264.4, NP_005255.1	2.16
<i>SRD5A1</i>	6715 NM_001047.2, NP_001038.1	1.99
<i>FBXW11</i>	23291 isoforma 1: NM_033645.2, NP_387449.2; isoforma 2: NM_033644.2, NP_387448.2; isoforma 3: NM_012300.2, NP_036432.2	1.64
<i>IGFBP3</i>	3486 isoforma 1: NM_001013398.1, NP_001013416.1; isoforma 2: NM_000598.4, NP_000589.2	1.64
<i>CHEK1</i>	1111 NM_001274.4, NP_001265.2	1.53
<i>ST6GALNAC5</i>	81849 NM_030965.1, NP_112227.1	1.38
<i>LAMC2</i>	3918 NM_005562.2, NP_005553.2	1.19
<i>CCNA1</i>	8900 isoforma 1: NM_003914.3, NP_003905.1; isoforma 2: NM_001111045.1, NP_001104515.1; isoforma 3: NM_001111046.1, NP_001104516.1	1.08
<i>TNIK</i>	23043	0.98

	isoforma 1: NM_015028.2, NP_055843.1; isoforma 2: NM_001161560.1, NP_001155032.1; isoforma 3: NM_001161561.1, NP_001155033.1; isoforma 4: NM_001161562.1, NP_001155034.1; isoforma 5: NM_001161563.1, NP_001155035.1; isoforma 6: NM_001161564.1, NP_001155036.1; isoforma 7: NM_001161565.1, NP_001155037.1; isoforma 8: NM_001161566.1, NP_001155038.1	
<i>PRSS12</i>	8492 NM_003619.3, NP_003610.2	0.77
<i>HHEX</i>	3087 NM_002729.4, NP_002720.1	0.54
<i>CRISPLD2</i>	83716 NM_031476.3, NP_113664.1	0.52
<i>G3BP1</i>	10146 NM_005754.2, NP_005745.1	0.39
<i>SAMHD1</i>	25939 NM_015474.3, NP_056289.2	0.27
<i>PDLIM4</i>	8572 isoforma 1: NM_003687.3, NP_003678.2; isoforma 2: NM_001131027.1, NP_001124499.1	0.18
<i>LAMB1</i>	3912 NM_002291.2, NP_002282.2	0.17
<i>DKK3</i>	27122 NM_015881.5, NP_056965.3	0.1
<i>LOC151162</i>	151162 NR_024275.1	0
<i>PNMAL1</i>	55228 isoforma 1: NM_018215.3, NP_060685.2; isoforma 2: NM_001103149.1, NP_001096619.1	0

Dalla classificazione dei geni basata sull'associazione ai dati clinici, sono stati selezionati i geni con un punteggio maggiore di 3, ottenendo così una *firma molecolare* associata all'asse Pin1/p53 mutata costituita dai 10 geni *DEPDC1*, *CPSF6*, *C21orf45*, *CENPA*, *FAM64A*, *CCNE2*, *BUB1*, *EPB41L4B*, *NCAPH*, *WDR67*.

L'espressione di questi geni correla in maniera significativa con l'esito clinico della malattia: nelle pazienti che esprimevano questi geni a livelli superiori rispetto al valore medio di tutti i pazienti, l'intervallo di tempo tra la diagnosi del tumore primario e quella di metastasi in altri distretti corporei (*Time to Distant Metastasis*, TDM) è risultata più breve (2-3 anni a seconda del dataset) e la sopravvivenza ridotta (*Overall Survival*, OS, 2-3 anni a seconda del dataset). L'espressione di questi dieci geni presenta valore prognostico anche considerando esclusivamente le pazienti con tumore mammario con positività al recettore dell'estrogeno. Simili risultati sono stati ottenuti analizzando l'espressione di un gruppo di sette geni costituito da: *C21orf45*, *CPSF6*, *DEPDC1*, *EPB41L4B*, *FAM64A*, *NCAPH* e *WDR67*.

Nel complesso, i dati prodotti dagli inventori indicano che nelle cellule tumorali l'azione combinata di Pin1 e p53 mutata incide profondamente sul programma di espressione genica attivando in maniera diretta un programma trascrizionale che promuove il fenotipo tumorale aggressivo. Gli effetti dell'asse Pin1/p53 mutata hanno, quindi, una rilevanza clinica significativa, come dimostrato attraverso una serie di studi condotti su una coorte di pazienti colpite da carcinoma mammario. Dalle analisi è emerso, infatti, come il valore prognostico della presenza di mutazione in p53 venga rafforzato se a essa si associa la quantificazione dei livelli di Pin1. L'analisi statistica multivariata ha dimostrato che la contemporanea presenza di livelli elevati di espressione di Pin1 (quantificazione del segnale nucleare maggiore di 5, valutato mediante immunistochemica utilizzando il metodo "Quick score" (Detre S. et al., 1995) e di mutazioni missenso in *TP53* rappresenta già di per sé un fattore prognostico indipendente e forte dell'esito clinico della malattia. Lo stato di *TP53* nei tumori che sovraesprimevano Pin1, in effetti, ha permesso di stratificare meglio le pazienti in gruppi caratterizzati da un tempo di sopravvivenza più o meno lungo. Questo effetto è stato osservato anche per un campione di pazienti sottoposte a chemioterapia adiuvante a base di antraciclina.

I dati indicano, quindi, che l'utilizzo del metodo di prognosi proposto dagli inventori e basato sulla quantificazione dei livelli di Pin1 e sulla rilevazione dello stato mutazionale di *TP53* potrebbe offrire significativi vantaggi in fase di

caratterizzazione della malattia. L'analisi combinata di questi due parametri, a differenza di quanto accade considerando solo la presenza di mutazioni nel gene *TP53*, permette di stratificare meglio le pazienti. Un aspetto critico al momento della classificazione della malattia, infatti, è proprio l'identificazione dei casi a elevato rischio di ricorrenza e la capacità di predire (attraverso biomarcatori specifici) la risposta delle pazienti alle terapie. Prerequisito essenziale questo per poter implementare le strategie di cura e guidare future scelte terapeutiche. Da questo punto di vista, in particolare, il metodo basato sulla quantificazione dei livelli di Pin1 e sullo stato mutazionale di *TP53* permette di discriminare nei diversi casi di carcinoma mammario il sottogruppo di pazienti che: i) hanno una minore probabilità di sopravvivenza; ii) che rispondono in maniera inefficace agli interventi terapeutici (e in particolare alla chemioterapia adiuvante a base di antraciclina).

Ai fini prognostici, oltre che mediante la quantificazione dei livelli della proteina Pin1 e la rilevazione dello stato mutazionale di *TP53*, la valenza prognostica dell'asse Pin1/p53 mutata può essere valutata anche dalla combinazione tra la quantificazione dei livelli proteici di Pin1 e di p53, i cui elevati livelli proteici correlano con la presenza di mutazioni missenso nel suo gene (Goh A. M. et al., 2011), ottenendo così un metodo prognostico altrettanto preciso basato però su un unico sistema di indagine, ovvero l'analisi immunohistochimica di entrambe le proteine.

Tali indagini possono essere eseguite su campioni biotici freschi o congelati, o campioni biotici fissati ed inclusi in paraffina rilevando un segnale mediante le tecniche note di immunochimica e/o di biologia molecolare basate sulla determinazione degli acidi nucleici, come la Polymerase Chain Reaction (PCR) e l'analisi mediante particolari kit già in uso per la proteina p53 (ad esempio AmpliChip P53 test) e/o di sequenziamento o altre note ad un tecnico del ramo. La valutazione prognostica è quindi effettuata sulla base di un punteggio derivato dalla quantificazione del segnale rilevato per l'espressione delle proteine pin1 e p53.

In un secondo aspetto tale metodo di prognosi può essere ulteriormente perfezionato integrando la determinazione dei livelli di Pin1 e dello stato mutazionale di *TP53* e/o dei livelli della proteina p53 con l'analisi della *firma*

molecolare associata all'asse Pin1/p53 mutata. Tale ulteriore valutazione, basata sulla rilevazione dei livelli di espressione dei geni appartenenti alla *firma molecolare*, non solo consente di effettuare la prognosi della malattia anche senza conoscere lo stato di Pin1 e p53, ma permette anche la stratificazione delle pazienti con tumore mammario con positività al recettore per gli estrogeni.

I livelli di espressione dei geni di questa *firma molecolare* possono essere monitorati su campioni biotici freschi mediante determinazione dei loro livelli di mRNA mediante PCR. È da considerare, inoltre, l'opportunità di un'analisi più rapida, attraverso il rilevamento dei livelli delle proteine espresse da tali geni, da effettuare tramite immunoistochimica su campioni biologici fissati e inclusi in paraffina. Anche in questo caso la valutazione prognostica può essere effettuata sulla base di un punteggio derivato dalla quantificazione del segnale rilevato per l'espressione dei geni e/o delle proteine espresse da tali geni.

Entrando nello specifico dei due saggi, nel caso in cui i parametri da misurare siano i livelli di espressione della proteina Pin1 e dello stato mutazionale del gene *TP53*, potrà essere eseguita preferibilmente l'analisi immunoistochimica per Pin1 ed il sequenziamento del gene *TP53* sull'RNA estratto dal campione biotico fresco o congelato.

Per la valutazione di tipo immunoistochimico i reagenti per la determinazione dell'espressione di Pin1 sono nello specifico anticorpi poli- e monoclonali anti-Pin1 opportunamente marcati ai fini della loro rilevazione.

Per la rilevazione dello stato mutazionale del gene *TP53* sono necessari oligonucleotici specifici consistenti in sequenze di DNA (primer 5'-3' *forward* e *reverse*) per la reazione di amplificazione, quali ad esempio, gli oligonucleotidi noti di seguito indicati in tabella 2:

Tabella 2: primer 5'-3' *forward* e *reverse* per la rilevazione dello stato mutazionale del gene *TP53*

<i>Simbolo del primer</i>	<i>Sequenza (senso 5'-3')</i>	<i>Regione di appaiamento</i>	<i>Sequenza genomica di riferimento RefSeq</i>
RT2	aatgtcagtctgagtcaggcccttctg	1383-1357	NM_000546.4

			SEQIDNO:1
GIL	tgatgtgtccccggacgatattgaa	325-350	NM_000546.4 SEQIDNO:2
MP9ER	tctcccaggacaggcacaacacg	1037-1014	NM_000546.4 SEQIDNO:3
E6E7F	tttgcgtgtggagtatttgat	797-818	NM_000546.4 SEQIDNO:4
MP6R	ttggcaaaacatcttgttgagggc	605-582	NM_000546.4 SEQIDNO:5

In alternativa e preferibilmente, lo stato mutazionale di *TP53*, anziché mediante sequenziamento potrebbe essere inferito valutando i livelli proteici di p53 (quantificazione del segnale nucleare maggiore di 5, utilizzando il metodo “Quick score” (Detre S. et al., 1995) mediante immunoistochimica con anticorpi poli- e monoclonali anti-p53).

In base al metodo proposto, è possibile suddividere le pazienti con tumore al seno in due categorie aventi migliore (80% senza metastasi a dieci anni) o peggiore prognosi (40% senza metastasi a dieci anni). I casi di quest’ultimo gruppo sono quelli in cui il tumore presenta Pin1 e p53 sovraespressi (per entrambi il punteggio in immunoistochimica dev’essere superiore a 5) oppure quelli in cui il tumore sovraesprime Pin1 (punteggio in immunoistochimica superiore a 5) e presenta mutazioni in *TP53* (identificate tramite sequenziamento).

La determinazione dell’espressione della *firma molecolare* associata a Pin1/p53 mutata può essere effettuata misurando mediante PCR l’espressione degli mRNA dei 10 geni della *firma molecolare* associata a Pin1/p53 mutata su campione biotico fresco o congelato e di geni di riferimento la cui espressione sia di intensità paragonabile ai geni della *firma molecolare* e la cui presenza sia costante nei campioni biotici. Tali geni possono essere scelti tra *ERCC6*, *ADCY3*, *TUBGCP2*, *SOX15*, *RPL10*, *DGCR14*, *KIAA0586*, *ZBTB17*, *TFIP11*, *ZNF672*, *TXN2*. I geni di riferimento preferibilmente sono i tre geni *ERCC6*, *ADCY3*, *TUBGCP2*. I reagenti possono essere i primer necessari per l’amplificazione del mRNA estratto dal campione in esame di seguito riportati in tabella 3:

Tabella 3: primer 5'-3' *forward* e *reverse* per la rilevazione dei 10 geni della *firma molecolare*

<i>Simbolo del gene</i>	<i>Sequenza (Senso 5'-3')</i>	<i>Regione di appaiamento *</i>	<i>RefSeq mRNA di (mRNA ID)</i>
<i>BUB1 forward</i>	ATTCAAGCCACAGAGTGGAGCAG	1354-1376	NM_004336.3 SEQIDNO:6
<i>BUB1 reverse</i>	AGAACTTGTGTTGGCAACCTTATGTG	1398-1423	NM_004336.3 SEQIDNO:7
<i>C21ORF45 forward</i>	GCGACTCGCTGAGCTGGGTG	329-348	NM_018944.2 SEQIDNO:8
<i>C21ORF45 reverse</i>	CCCCGCGCAGCACAAAGTCT	461-480	NM_018944.2 SEQIDNO:9
<i>CCNE2 forward</i>	TGAGCCGAGCGGTAGCTGGT	53-74	NM_057749.2 SEQIDNO:10
<i>CCNE2 reverse</i>	GGGCTGGGGCTGCTGCTTAG	133-152	NM_057749.2 SEQIDNO:11
<i>CENPA forward</i>	CTTCCTCCCATCAACACAGTCG	287-308	NM_001809.3 SEQIDNO:12
<i>CENPA reverse</i>	TGCTTCTGCTGCCTCTTGTAGG	462-483	NM_001809.3 SEQIDNO:13
<i>CPSF6 forward</i>	AGGGGCTGTTCTTGGTGGGG	734-753	NM_007007.2 SEQIDNO:14
<i>CPSF6 reverse</i>	GGCCCAGCTAGAGGAGGAGGC	887-907	NM_007007.2 SEQIDNO:15
<i>DEPDC1 forward</i>	TGGGTATTATCTGCCATGAAGTGCCT	853-878	NM_017779.4 SEQIDNO:16
<i>DEPDC1 reverse</i>	AGGTTGCAGCAAGCCCCAAAATGT	1019-1041	NM_017779.4 SEQIDNO:17
<i>EPB41L4B forward</i>	CGACGGGACCGAAGTGAGCG	571-590	NM_018424.2 SEQIDNO:18
<i>EPB41L4B reverse</i>	CAGTGCGCAACCTGGGCAGA	692-711	NM_018424.2 SEQIDNO:19
<i>FAM64A forward</i>	CTCCAGGCTGCAGCTCGCTC	353-372	NM_00119522 8.1 SEQIDNO:20
<i>FAM64A reverse</i>	CAGCCGGGTGCTCTTCTGGC	519-538	NM_00119522 8.1 SEQIDNO:21

<i>NCAPH forward</i>	GAGGAGCCTGCCCCCTGTCA	2169-2188	NM_015341.3 SEQIDNO:22
<i>NCAPH reverse</i>	TGGGCCTCCTGCTGCTGACT	2325-2344	NM_015341.3 SEQIDNO:23
<i>WDR67 forward</i>	AGGCAACAAGGAGAGCGGCA	108-127	NM_145647.3 SEQIDNO:24
<i>WDR67 reverse</i>	AGCAGTCGCCTGTGCCATCA	240-259	NM_145647.3 SEQIDNO:25
<i>ERCC6 for</i>	TCTCTGTTTCCTTGTGGGCGCTC	106-128	<u>NM_000124.2</u> SEQIDNO:26
<i>ERCC6 rev</i>	CACCCCACAGCAGAGGTGGACA	315-336	<u>NM_000124.2</u> SEQIDNO:27
<i>ADCY3 for</i>	GTGGGCTGGCAGGTCTTCTTTG	704-725	<u>NM_004036.3</u> SEQIDNO:28
<i>ADCY3 rev</i>	GGTGCTTGCGGTCAGCCATGTA	932-953	<u>NM_004036.3</u> SEQIDNO:29
<i>TUBGCP2 for</i>	GACGTCACGCTCACTTCGCCCG	257-278	<u>NM_006659.2</u> SEQIDNO:30
<i>TUBGCP2 rev</i>	TCAGCCCCATCTCCTCCGTGGA	413-434	<u>NM_006659.2</u> SEQIDNO:31

*corrispondente alla sequenza dell'mRNA maturo (dal 5' al 3' sul filamento senso)

Attraverso la quantificazione dell'espressione di questi 10 geni, è possibile suddividere le pazienti con tumore al seno in tre categorie aventi prognosi buona, intermedia o cattiva.

Tale stratificazione è possibile in base ad uno score per la *firma molecolare* ottenuto calcolando i relativi score genici valutando per ognuno dei geni la relativa espressione rispetto ad uno dei geni di riferimento come da seguente tabella 4:

Tabella 4

<i>Firma molecolare</i>	Geni di riferimento
<i>FAM64A</i>	<i>ADCY3</i>
<i>CENPA</i>	<i>ADCY3</i>
<i>C21orf45</i>	<i>ADCY3</i>
<i>NCAPH</i>	<i>ADYC3</i>
<i>BUB1</i>	<i>ERCC6</i>
<i>CCN2</i>	<i>ERCC6</i>

<i>WDR67</i>	<i>ERCC6</i>
<i>DEPDC1</i>	<i>ERCC6</i>
<i>CPSF6</i>	<i>TUBGCP2</i>
<i>EPB41L4B</i>	<i>TUBGCP2</i>

In particolare:

- a) se un gene della *firma molecolare* ha un valore di espressione inferiore di 0,5 su scala logaritmica in base 2 rispetto al valore di espressione corrispondente al proprio gene di riferimento (che equivale a dire che esiste una differenza di 0,5 cicli di RT-PCR), ad esso viene attribuito uno score genico pari a zero.
- b) se un gene della *firma molecolare* ha un valore di espressione superiore di 0,5 su scala logaritmica in base 2 rispetto al valore di espressione corrispondente al proprio gene di riferimento (che equivale a dire che esiste una differenza di 0,5 cicli di RT-PCR), ad esso viene attribuito uno score genico pari a due.
- c) se un gene della *firma molecolare* ha un valore di espressione intermedio ai due valori dei casi a) e b) rispetto al valore di espressione corrispondente al proprio gene di riferimento ad esso viene attribuito uno score genico pari a uno.

La somma degli score genici (S_g) così ottenuti determina il valore dello score della *firma molecolare* dei dieci geni (score firma S_f):

$$S_f = \sum S_g$$

e

- a un valore dello score della *firma molecolare* dei dieci geni (S_f) inferiore a 5, è associata una prognosi favorevole (80% senza metastasi a dieci anni);
- a un valore dello score della *firma molecolare* dei dieci geni (S_f) inferiore a 9 e superiore a 5 si associa una prognosi meno favorevole (60% senza metastasi a dieci anni);
- a un valore superiore a 9 si associa una prognosi meno favorevole (50% senza metastasi a dieci anni).

In alternativa, la determinazione dell'espressione della *firma molecolare* associata a Pin1/p53 mutata potrebbe essere eseguita su campioni biotipici fissati e inclusi in paraffina misurando i livelli dei prodotti proteici dei 7 geni della *firma molecolare* minima e preferibilmente di tutti i 10 geni della *firma molecolare* associata a Pin1/p53 mutata mediante analisi per immunohistochimica. In questo caso i

reagenti sono anticorpi poli- e monoclonali opportunamente marcati contro le proteine prodotte dal gene. Anche in questo caso la quantificazione del segnale nucleare può essere eseguita utilizzando il metodo "Quick score" (Detre S. et al., 1995).

Il kit per l'esecuzione del metodo prognostico oggetto dell'invenzione basato sulla determinazione dei livelli di Pin1 e rilevazione dello stato mutazionale di *TP53* e/o livelli di p53 comprende almeno:

- uno specifico reagente per la determinazione dell'espressione della Pin1 consistente in un anticorpo poli- o monoclonale;
- uno specifico reagente per la determinazione dell'espressione della p53 consistente in un anticorpo poli- o monoclonale; e/o
- almeno un oligonucleotide per l'amplificazione del gene *TP53* derivato sulla sua sequenza di riferimento o sulla sequenza complementare della stessa;
- un foglietto di istruzioni.

Il kit per l'esecuzione del metodo prognostico oggetto dell'invenzione basato sulla determinazione dei livelli di espressione dei geni della *firma molecolare* e/o delle proteine da questi espresse comprende almeno:

- uno specifico reagente per la determinazione dell'espressione delle proteine espresse da almeno i 7 geni della *firma molecolare* minima, e preferibilmente tutti e 10 i geni della *firma molecolare*, consistente in un anticorpo poli- o monoclonale; e/o
- oligonucleotidi per l'amplificazione di almeno i 7 geni della *firma molecolare* minima, e preferibilmente tutti e 10 i geni della *firma molecolare*, e di uno o più geni di riferimento derivati sulle sequenze di riferimento degli stessi o sulle loro sequenze complementari;
- un foglietto per le istruzioni per calcolare lo score della firma molecolare.

La caratterizzazione molecolare del tumore attraverso la rilevazione dei livelli di Pin1 in associazione con la presenza/assenza di p53 mutata e/o attraverso la *firma molecolare* di 10 geni rappresenta un sistema affidabile, sensibile e contenuto nei costi che permette di predire l'esito della malattia, il rischio di

ricorrenza e la risposta delle pazienti ai trattamenti (specialmente alla chemioterapia adiuvante a base di antraciclina), in special modo in quei casi definiti *basal-like*. Questi, a tutt'oggi, rappresentano una sfida in termini di trattamento poiché costituiscono un sottogruppo di tumori al seno che comprende in prevalenza casi triplo-negativi, definiti così perché non esprimono i recettori ormonali e non sovra-esprimono HER2, mostrano tassi di sopravvivenza a cinque anni molto bassi e sono caratterizzati da un'elevata frequenza di recidive, che insorgono nonostante il trattamento con chemioterapia adiuvante.

PARTE SPERIMENTALE

Materiali e metodi

Colture cellulari e costrutti

Le MDA-MB-231 (p53R280K) e le MDA-MB-468 (p53R273H) sono cellule umane derivate da carcinomi mammari del tipo triplo negativo. Le SK-BR-3 (p53R175H) sono cellule umane derivate da tumore alla mammella che sovraesprimono HER2. Tutte le linee cellulari sono state tenute in coltura in terreno Dulbecco's modified Eagle's Medium (DMEM) addizionato con il 10% di FBS, penicillina (100 IU/ml) e streptomycin (100IU/ml). Le cellule MEF sono state generate incrociando topi dal genotipo appropriato e recuperando le cellule dagli embrioni a 13,5 *d.p.c.* come descritto (Zacchi et al., 2002). pLPC e pLPC-H-Ras^{V12} ci sono stati gentilmente fornito da M. Serrano (Efeyan et al. 2009); pGEX-Pin1 per l'espressione di GST-Pin1 e pGEX utilizzato come controllo sono stati già descritti (Zacchi et al., 2002). pcDNA3-p53K280 e pcDNA3-p53K280-4M sono stati creati per mutagenesi sito-specifica a partire dai costrutti pcDNA3-p53 e pcDNA3-Mp53 descritti (Zacchi et al., 2002).

pMSCV-HAPin1 è stato generato subclonando la sequenza codificante specifica per Pin1 dal costrutto pcDNA3-HAPin1 (Mantovani et al., 2007) nel vettore pMSCV. pMSCV-HAPin1r e pMSCV-HAPin1r S67E sono stati generati introducendo, per mutagenesi sito-specifica dei costrutti pcDNA3-HAPin1 e pcDNA3-HAPin1 S67E (Rustighi et al., 2009), mutazioni silenziose (G300A, G303A e T309C, rispetto alla sequenza NM_006221.2) nella regione bersaglio del siRNA (A) specifico per Pin1 e subclonando poi le regioni mutagenizzate nel vettore pMSCV. pLPC-p53K280 e pLPC-p53K280-4M sono stati creati subclonando le

regioni codificanti specifiche per p53K280 e p53K280-4M rispettivamente nel vettore pLPC contenente un *tag* HA.

Il costrutto pSRshPin1 è stato creato clonando il siRNA a doppio filamento (A) specifico per Pin1 (si veda sotto) nel vettore pSuper Retro (Rustighi et al., 2009). Il plasmide pSRshp53 ci è stato gentilmente fornito da R. Agami (Drost J et al., 2010). Il vettore pSRshLacZ è stato descritto in Rustighi et al., 2009.

Il vettore lentivirale codificante per il gene reporter della luciferasi di lucciola utilizzato nei saggi per studiare lo sviluppo di metastasi in vivo è stato descritto in Adorno et al., 2009.

Dove non specificato, i vettori base (commercialmente disponibili) impiegati per creare i diversi costrutti sono stati utilizzati come vettori di controllo nei diversi esperimenti. Le sequenze siRNA sono riportate in tabella 5.

Tabella 5: Sequenze dei siRNA utilizzati

siRNA	Sequenza (Senso, 5'-3')	Gene bersaglio	Specie
Pin1 (A)	CGGGAGAGGAGGACUUUGA SEQIDNO:32	Pin1	<i>H. sapiens</i>
Pin1 (B)	GCCAUUUGAAGACGCCUCG SEQIDNO:33	Pin1	<i>H. sapiens</i>
p53	GACUCCAGUGGUAUUCUAC SEQIDNO:34	Tp53	<i>H. sapiens</i>
p53u	GGUGAACCUUAGUACCUAA SEQIDNO:35	Tp53	<i>H. sapiens</i>
Controllo (LacZ)	GUGACCAGCGAAUACCUGU SEQIDNO:36	Beta-D- galattosidasi	<i>E. coli</i>

Ceppi murini e analisi condotte nei topi

Le coorti di topi impiegate in questo studio hanno origine dagli incroci di topi $p53^{M/+}$ o $p53^{+/-}$ con topi $Pin1^{+/-}$, tutti con background genetico del tipo C57BL/6. La caratterizzazione genotipica degli animali è stata effettuata attraverso l'analisi delle reazioni a catena della polimerasi (PCR), come descritto in precedenza (Atchinson F.W. et al., 2003; Lang G.A. et al., 2004). Gli animali che durante le osservazioni mostravano segni di malattia o la presenza evidente di tumori sono stati sacrificati. Gli organi da essi prelevati sono stati tenuti per 24 ore in una soluzione di PBS al 10% di formalina per il fissaggio. I tessuti sono stati inclusi in

paraffina e sezionati in fettine di 5µm di spessore. Prima di essere sottoposti all'analisi patologica sono stati trattati al fine di eliminare i residui di paraffina e sottoposti a colorazione con ematossilina ed eosina. Le procedure che hanno visto coinvolti gli animali sono state realizzate in conformità alle linee guida istituzionali, stilate nel rispetto delle leggi e delle politiche internazionali in materia (UKCCCR, 1989).

Saggi di trasformazione e di tumorigenesi *in vivo*

Per i saggi sono state utilizzati fibroblasti embrionali murini (MEF) da topi $p53^{M/M}$ $Pin1^{+/+}$, $p53^{M/M}$ $Pin1^{-/-}$, $p53^{-/-}$ $Pin1^{+/+}$ e $p53^{-/-}$ $Pin1^{-/-}$ che avevano effettuato un basso numero di passaggi in coltura. Le cellule infettate con il vettore pLPC H-Ras^{V12} o con il vettore vuoto sono state risospese in DMEM addizionato con siero bovino fetale al 10%, penicillina (100 IU/ml) e streptomicina (IU/ml) e contenente agarosio allo 0,3%. Le cellule sono state piastrate alla densità di 1×10^5 per piastra da 60 mm di diametro, sulla superficie di uno strato di terreno che presentava le stesse caratteristiche di quello precedente, fatta eccezione per il fatto di contenere agarosio all'1%. Dopo 14 giorni di coltura è stato contato il numero di colonie che raggiungevano almeno i 100 µm di diametro. Tutti gli esperimenti sono stati condotti in triplicato utilizzando almeno due cloni diversi per genotipo. Per verificare la capacità tumorigenica delle MEF infettate con il costrutto che esprime H-Ras^{V12}, 1×10^6 cellule sono state iniettate sottocute e su entrambi i lati in topi NOD SCID. Dopo due settimane, il volume dei tumori è stato misurato attraverso un calibro, come descritto in precedenza (Rustighi A. et al., 2009).

Saggi di migrazione e di invasività

Per i saggi di cicatrizzazione, le cellule MDA-MB-231—trasfettate sono state seminate in piastre dal diametro di 35 mm e tenute in coltura fino a una confluenza del 90%. Le cellule sono state raschiate dalla superficie utilizzando una punta da 200 µl. La chiusura della lesione è stata monitorata per 36 ore. Per i saggi di migrazione attraverso i pozzetti sono stati utilizzati inserti PET da 24 pozzetti (dimensione dei pori 8,0 µm, Falcon). I saggi di invasività sono stati condotti in inserti PET da 24 pozzetti (dimensione dei pori 8,0 µm, Falcon) utilizzando filtri rivestiti di Matrigel®.

Saggi per studiare lo sviluppo di metastasi *in vivo*

Per i saggi di invasione dei polmoni, cellule di linee cellulari umane MDA-MB-231 derivate da tumori mammari sono state co-trasdotte con un vettore lentivirale codificante per il gene reporter della luciferasi di lucciola e con pSR shLacZ o con pSR shPin1 (sei animali per ciascun gruppo). 10^6 cellule sono state risospese in 100 μ l di PBS e inoculate nella vena caudale di topi SCID. Per la visualizzazione *in vivo*, a tempi diversi dall'iniezione delle cellule sono state acquisite immagini di bioluminescenza utilizzando un dispositivo di acquisizione "*cooled charge-coupled*" montato sopra la camera oscura destinata all'alloggiamento del campione (IVIS Lumina II Imaging System; Caliper Life Sciences, Alameda, CA). Dieci minuti prima di acquisire le immagini veniva somministrata i.p. agli animali anestetizzati una soluzione contenente 150 mg/kg di D-luciferina (Caliper Life Sciences) in PBS. Le analisi istologiche su campioni di tessuto polmonare inclusi in paraffina e colorati con ematossilina ed eosina sono state condotte a un mese dall'iniezione delle MDA-MB-231. Per l'analisi computerizzata del tessuto polmonare è stata valutata l'area occupata dai foci delle metastasi, identificabile nelle sezioni dei lobi polmonari. Il parametro è stato calcolato utilizzando un microscopio Leica DM200 dotato di una macchina fotografica digitale a colori Leica DFC295 e di un software Leica Application Suite (LAS) V3. Il valore di area totale e quello della somma delle aree sono stati, infine, confrontati con il valore dell'area dell'intero lobo polmonare. L'espressione di Pin1 è stata analizzata per immunoistochimica (IHC) utilizzando un anticorpo monoclonale (G-8, sc-46660, Santa Cruz) diretto contro questo enzima.

Analisi dell'espressione genica

L'RNA totale è stato estratto utilizzando il Trizol (Invitrogen) e poi sottoposto a trattamento con DNasi-I (Ambion). Per gli esperimenti con i microarray, gli mRNA di tre campioni biologici differenti per ogni gruppo (siPin1 o sip53) sono stati utilizzati per l'ibridazione dei chip Affymetrix GeneChip Human Genome U133A 2.0. Negli esperimenti di RT-PCR quantitativa, l'mRNA è stato trascritto utilizzando il sistema Superscript III (Invitrogen). La Real Time PCR è stata realizzata utilizzando una SYBR Green PCR master mix e un termociclatore StepOne Plus dell'Applied Biosystems. Le sequenze dei primer (senso 5'-3') per l'analisi dell'espressione genica mediante qRT-PCR per i 10 geni della *firma molecolare* e

per i tre geni di controllo sono riportate nella precedente tabella 3 e corrispondono alle sequenze: BUB 1 *for* SEQIDNO:6; BUB1 *rev* SEQIDNO:7; C21ORF45 *for* SEQIDNO:8; C21ORF45 *rev* SEQIDNO:9; CCNE2 *for* SEQIDNO:10; CCNE2 *rev* SEQIDNO:11; CENPA *for* SEQIDNO:12; CENPA *rev* SEQIDNO:13; CPSF6 *for* SEQIDNO:14; CPSF6 *rev* SEQIDNO:15; DEPDC1 *for* SEQIDNO:16; DEPDC1 *rev* SEQIDNO:17; EPB41L4B *for* SEQIDNO:18; EPB41L4B *rev* SEQIDNO:19; FAM64A *for* SEQIDNO:20; FAM64A *rev* SEQIDNO:21; NCAPH *for* SEQIDNO:22; NCAPH *rev* SEQIDNO:23; WDR67 *for* SEQIDNO:24; WDR67 *rev* SEQIDNO:25; ERCC6 *for* SEQIDNO:26; ERCC6 *rev* SEQIDNO:27; ADCY3 *for* SEQIDNO:28; ADCY3 *rev* SEQIDNO:29; TUBGCP2 *rev* SEQIDNO:30; TUBGCP2 *for* SEQIDNO:31.

Analisi dei dati dei microarray

Annotazioni relative all'espressione genica e pre-processamento: l'intensità dei segnali grezzi di espressione genica è stata normalizzata al livello della sonda utilizzando l'algoritmo RMA, come precedentemente descritto (Irizarry R.A. et al., 2003). L'analisi a basso livello dei dati dei microarray è stata realizzata in ambiente R/Bioconductor utilizzando, in particolare, i packages per microarray limma e affy (<http://www.R-project.org>). Le serie di dati accessibili al pubblico utilizzati (GSE7390, GSE2990, GSE3494, GSE1456) sono state scaricate dal sito web Gene Expression Omnibus (GSE7390, GSE2990, GSE3494) o dal sito degli autori dello studio che li ha prodotti (GSE1456, <http://www.mebki.se/~yudpaw/>). Le annotazioni geniche relative a tutte le piattaforme considerate sono state ottenute dai packages R-Bioconductor metadata (Gentleman R.C. et al., 2004). L'analisi dell'espressione differenziale dei geni di tutti i set di dati considerati in questo studio è stata condotta utilizzando funzioni e metodi implementati in R/Bioconductor (Gentleman R.C. et al., 2004; Smyth G.K., 2004). In breve, individualmente per ciascun gene è stato applicato un modello lineare a effetti fissi per stimare le differenze di espressione tra i gruppi di campioni da confrontare. Per moderare gli errori standard dei valori di M è stato impiegato un approccio empirico bayesiano (Smyth, 2004). Infine, per ciascun gene analizzato sono state ottenute le statistiche t-moderate, i rapporti log-odd di espressione differenziale

(statistiche B) e i valori P grezzi e aggiustati (controllo FDR tramite il metodo di Benjamini e Hochberg).

Analisi dei dati di sopravvivenza e dei microarray per l'individuazione della *firma molecolare*.

Per verificare la presenza o meno di associazione tra l'espressione dei 31 geni riportati nella precedente tabella 1 contemporaneamente più repressi nelle due diverse condizioni sperimentali analizzate e il tempo di sopravvivenza di pazienti con carcinoma alla mammella (Desmedt et al., 2007), è stato utilizzato il modello a rischi proporzionali di Cox. In particolare, questo modello (statistica di Wald) è stato applicato all'espressione di ciascun gene per testarne l'associazione con i dati di sopravvivenza, in modo da produrre un coefficiente o punteggio di associazione con i dati clinici. Dalla classificazione dei geni basata sull'associazione ai dati clinici, sono stati selezionati i geni con un punteggio maggiore di 3, ottenendo così una *firma molecolare* costituita da 10 geni. Per verificare la correlazione esistente tra la *firma molecolare* di 10 geni e i dati clinici nei casi di carcinoma alla mammella, è stato effettuato un test di Mantel Haenszel su gli altri set di dati normalizzati utilizzati (GSE2990, GSE3494, GSE1456) (survival R package).

Dalla matrice di espressione per ogni set di dati sono stati estratti i valori di espressione per i geni componenti la firma dei 10 geni mediante l'uso dell'identificativo univoco probe-set, ottenendo così una sotto-matrice di espressione. Questa matrice è costituita da m righe/geni per n colonne/campioni, con m che va da uno a dieci, e n che va da 1 al numero di campioni presenti nello studio. I geni sono stati normalizzati per assegnare a ciascun gene lo stesso peso, con la seguente procedura:

$$SV_j = \frac{V_j - \bar{V}_j}{\text{mad}(V_j)}$$

quindi per ogni riga/gene j , che porta a:

$$SV_j = \{SX_{j,1}, SX_{j,2}, \dots, SX_{j,n}\}.$$

SV_j = vettore di espressione genica standardizzato rappresentante i valori di espressione genica standardizzati per uno specifico gene/riga j per tutti i campioni del dataset (n).

V_j = vettore di espressione genica rappresentante i valori di espressione genica per uno specifico gene/riga j per tutti i campioni del set di dati (n). $V_j = \{X_{j,1}, X_{j,2}, \dots, X_{j,n}\}$.

$\bar{V}_j$ = media stimata per uno specifico vettore di espressione V_j per uno specifico gene/riga j .

$\text{mad}(V_j)$ = deviazione mediana per uno specifico vettore di espressione V_j per uno specifico gene/riga j .

Per misurare il contributo di ogni gene alla firma dieci geni si è calcolato un vettore gene virtuale (VG) facendo la media di ogni contributo di ogni gene a partire dal vettore di espressione genica standardizzato per tutte i geni/righe m .

$$VG = \{VG_1, VG_2, \dots, VG_n\} \text{ dove } VG_k = \frac{1}{m} \sum_{j=1}^m SX_{j,k}$$

Infine un campione k è stato definito "Alto livello firma dieci geni" se il valore di VG_k era superiore al valore mediano del vettore per il gene virtuale VG, mentre un campione k è stato definito "Basso livello firma dieci geni" se il valore di VG_k era inferiore al valore mediano del vettore per il gene virtuale VG.

Alto livello: $VG_k > \text{mediana di VG}$

Basso livello: $VG_k < \text{mediana di VG}$

Le curve di sopravvivenza sono state costruite utilizzando il metodo di Kaplan-Meier (survival package). Quindi con test log-rank sono state analizzate le differenze di espressione dei 10 geni della *firma molecolare* (come sopra descritto) rispetto alla durata dell'intervallo di tempo tra l'insorgenza del tumore primario e quello di metastasi in altri distretti corporei (se il dato era disponibile) o rispetto alla sopravvivenza.

Immunoprecipitazione della cromatina

La cromatina è stata immunoprecipitata utilizzando l'anticorpo p53 DOI (Santa Cruz) o un anticorpo policlonale diretto contro Pin1 (Zacchi P. et al., 2002). Come controlli negativi sono state utilizzate le IgG purificate dal siero di coniglio o di topo. Il DNA co-immunoprecipitato è stato analizzato per Real Time PCR

utilizzando un termociclatore StepOne Plus (Applied Biosystems) e una SYBR Green Universal PCR master mix (Applied Biosystems). La presenza del fattore sulla cromatina dei promotori analizzati è stata calcolata come percentuale della quantità di cromatina immunoprecipitata utilizzata come input, utilizzando il metodo $2^{-\Delta Ct}$. Le sequenze dei primer sono riportate di seguito in tabella 6:

Tabella 6: primer 5'-3' per i 10 geni della *firma molecolare*

Simbolo del gene	Sequenza	Regione di appaiamento	Sequenza genomica di riferimento
<i>BUB1 (+) for</i>	GGGCTGCCTCACACCCGTTT SEQIDNO:37	8105-8124	NG_012048.1
<i>BUB1 (+) rev</i>	ACACCTGCCCTTCCAGTGGG SEQIDNO:38	8213-8232	NG_012048.1
<i>C21ORF45 (+) for</i>	GGTGCTGAGCAGCTGGTGCAAA SEQIDNO:39	2936-2957	NC_000021.8
<i>C21ORF45 (-) rev</i>	TGCGGGGTACATTACTCCCCT SEQIDNO:40	3093-3114	NC_000021.8
<i>CCNE2 (+) for</i>	AAATCCAGGAGTTGCAGTGG SEQIDNO:41	13891-13910	NT_008046.16
<i>CCNE2 (+) rev</i>	ACTCTCAGGGGCTCCTTCTC SEQIDNO:42	13990-14010	NT_008046.16
<i>CENPA (+) for</i>	ACTTCCCAAGGGTCACACAGTCA SEQIDNO:43	-4444 -4421	NT_022184.15
<i>CENPA (+) rev</i>	TGTGGCAGAATGGGACAGGCG SEQIDNO:44	-4384 -4363	NT_022184.15
<i>CPSF6 (+) for</i>	TGGCAATGAATGCGAGAGGAAGGT SEQIDNO:45	-626 -605	NT_029419.12
<i>CPSF6 (+) rev</i>	CCCCGGGCATTCTGTGAACGG SEQIDNO:46	-451 -430	NT_029419.12
<i>DEPDC1 (+) for</i>	AGGGCCAGGCAGAAAAACCGT SEQIDNO:47	-442 -421	NT_032977.9
<i>DEPDC1 (+) rev</i>	CGGGTTTCCCTGGCGCTGTT SEQIDNO:48	-277 -257	NT_032977.9
<i>DEPDC1 (-) for</i>	GGCTTGGGACGCCTCACGAT SEQIDNO:49	-2112 -2092	NT_032977.9
<i>DEPDC1 (-) rev</i>	GCCCATGAACTGCCTAGCTGTGG SEQIDNO:50	-1932 -1909	NT_032977.9
<i>EPB41L4B (+)</i>	AGTGCAGCAGTGAGACGGGGA	-4317 -4296	NT_008470

<i>for</i>	SEQIDNO:51		
<i>EPB41L4B (+)</i> <i>rev</i>	ACCCCAGAGGCTAGTCCGAAGG SEQIDNO:52	-4231 -4209	NT_008470
<i>FAM64A (+) for</i>	ATCGGGGCGTCTGGTGGGAA SEQIDNO:53	5362-5382	NT_010718.16
<i>FAM64A (+) rev</i>	CGAGGAGGGGAGGCAGGGTC SEQIDNO:54	5564-5584	NT_010718.16
<i>NCAPH (+) for</i>	CCCTGCCTGCATCTGTGCCA SEQIDNO:55	2477-2497	NT_022171.15
<i>NCAPH (+) rev</i>	GCAGGCAAAGAGGTTGCCTTTGT SEQIDNO:56	2569-2591	NT_022171.15
<i>WDR67 (+) for</i>	CACGAGGCTCCCAAAGGGCG SEQIDNO:57	79523- 79503	NT_008046
<i>WDR67 (+) rev</i>	AAAATCCGCCCAAGCGGCCA SEQIDNO:58	79351- 79331	NT_008046

Procedura per la rilevazione dei livelli di Pin1 in campioni di tessuto tumorale

Costruzione dei microarray di tessuto (TMA)

I microarray di tessuto (TMA) utilizzati in questo studio per rilevare i livelli di Pin1 in campioni di tessuto tumorale derivano da una serie storica di campioni di carcinoma mammario, prelevati da 212 tumori primari precedentemente non trattati e non selezionati in altro modo. I TMA sono stati creati utilizzando un *arrayer* specifico manuale (Beecher Instruments Inc), con in media sei sezioni circolari del diametro di 0,6 mm per campione. Uno specialista patologo esperto nello studio della mammella (L.J.) ha marcato per ogni campione le aree tumorali rappresentative su vetrini colorati con ematossilina ed eosina prima di prelevare le sezioni.

Immunomarcatura dei TMA

Le sezioni (dello spessore di 4 micron) tagliate dai blocchetti di TMA sono state stese su vetrini Superfrost® (VWR International Ltd) e seccate per 1 ora a 60°C prima che fosse rimossa la paraffina in Histoclear (National Diagnostics). Le sezioni sono state poi re-idratate per immersione in una serie di soluzioni a percentuale gradualmente decrescente di alcool. Le sezioni sono state incubate in una soluzione tampone contenente Acido citrico 10 mM a pH 6,0. Sono state poi riscaldate al forno a microonde per 15 min prima di essere immunomarcate

utilizzando il sistema di processamento Dako Autostainer e il kit Vectastain® ABC (Vector Labs), secondo il protocollo fornito dalla casa produttrice. In breve, le sezioni sono state bloccate per 20 min utilizzando siero normale di capra addizionato al 10% (v/v) con una soluzione stock di avidina (Vector Labs) e poi incubate per 1 ora in una soluzione contenente l'anticorpo primario (anticorpo monoclonale diretto contro Pin1 (G-8, sc-46660), diluizione 1:250) e addizionata al 10% (v/v) con una soluzione stock di biotina (Vector Labs), in modo da ridurre la marcatura aspecifica di fondo. Le sezioni sono state, quindi, incubate per 30 min in presenza di un anticorpo secondario universale biotinilato (Vectastain® ABC kit, Vector Labs) e successivamente per altri 30 min in presenza del reagente Vectastain® Elite ABC. Diaminobenzidina liquida (DAB) (DAKO) è stata utilizzata come cromogeno, incubandovi le sezioni per 5 min. Queste poi sono state controcolorate con una soluzione di ematossilina di Mayer. Durante l'immunomarcatura, tra ogni passaggio e il successivo i vetrini sono stati lavati brevemente in tampone salino a base di Tris (TBS) a pH 7,6. In ogni esperimento sono stati inclusi campioni che sarebbero risultati certamente marcati positivamente e controlli negativi, allestiti utilizzando tampone TBS al posto della soluzione contenente l'anticorpo primario.

Analisi al microscopio digitale virtuale e valutazione dei TMA

I vetrini di TMA immunomarcati sono stati esaminati (Scancope XT, Aperio Technologies) utilizzando un obiettivo a ingrandimento 40x e catalogati all'interno del database "Aperio Spectrum Plus + TMA" (version 9.0.748.1521). L'analisi dei TMA relativi ai casi di carcinoma mammario è stata realizzata alla cieca. La classificazione dei campioni è stata sottoposta a verifica da parte dello stesso patologo che ha condotto le prime valutazioni, per mezzo di un'analisi standard al microscopio ottico, utilizzando un microscopio Nikon Eclipse E600. Sono state analizzate solo le sezioni circolari contenenti tessuto tumorale; criteri di esclusione delle sezioni dall'analisi erano dimensioni inferiori al 20% dell'atteso, la sovrapposizione delle sezioni, la presenza prevalente di tessuto adiposo, di stroma o di tessuto mammario epiteliale normale. Per la valutazione sono stati, quindi, sufficienti 212 campioni tumorali. Sono stati considerati solo i tumori maligni invasivi mentre i tumori *in situ*, l'epitelio circostante e lo stroma sono stati

ignorati. È stata valutata la positività alla marcatura con gli anticorpi in ogni sezione che aveva caratteristiche adeguate ed è stato assegnato un punteggio utilizzando il metodo "Quick Score" (Detre S. et al., 1995). In breve, è stata calcolata la percentuale di cellule positive e in base a questa è stato assegnato un punteggio su una scala da 1 a 6: 0 - 4 % = 1; , 5 - 19 % = 2; ; 20 - 39 % = 3, 40 - 59 % = 4, 60 - 79 % = 5 e 80 - 100 % = 6. È stata valutata l'intensità media dei segnali provenienti dalle cellule positive alla marcatura e ad essa è stato assegnato un punteggio da 0 a 3; nessun segnale = 0, segnale di marcatura debole = 1, segnale intermedio = 2 e segnale forte = 3. Il punteggio Quick Score è stato poi ottenuto moltiplicando il punteggio relativo alla percentuale di cellule marcate e il punteggio relativo all'intensità del segnale di marcatura, per un valore massimo raggiungibile di 18. Dato che per la maggior parte dei campioni più sezioni sono risultate disponibili per la valutazione, è stato generato un punteggio Quick Score totale unico per ciascun tumore e quindi per ciascuna paziente. È stato annotato anche il tipo di localizzazione cellulare, citoplasmatica o nucleare, del segnale di immunomarcatura e quindi ciascun compartimento cellulare è stato considerato separatamente, ottenendo così due punteggi Quick Score per ciascuna paziente.

Procedura per analizzare lo stato di *TP53* su campioni di tessuto tumorale

Lo stato di *TP53* è stato analizzato tramite sequenziamento utilizzando come materiale di partenza l'RNA totale estratto da campioni di tessuto mammario tumorale congelato. L'RNA totale è stato estratto utilizzando l'RNA assay kit della QIAGEN. La qualità dell'acido nucleico recuperato è stata valutata tramite la piattaforma microfluidica Agilent-bioanalyzer. I campioni che all'analisi davano un rapporto tra le assorbanze a 280 nm e 260 nm inferiore a 1,2 venivano scartati. La reazione di retrotrascrizione è stata condotta a partire da 500ng di RNA totale utilizzando primer specifici per p53 in modo da amplificare tutte le sue isoforme.

Primer per l'analisi dello stato mutazionale di *TP53* tramite sequenziamento

Primer	Sequenza 5'-3'	Regione di appaiamento	Sequenza genomica di riferimento
RT2	Aatgtcagtctgagtcaggcccttctg SEQIDNO:1	1383-1357	NM_000546.4

GIL	Tgatgctgtccccggacgatattgaa SEQIDNO:2	325-350	NM_000546.4
MP9ER	Tctcccaggacaggcacaacacg SEQIDNO:3	1037-1014	NM_000546.4
E6E7F	ttgcgtgtggagtatttgat SEQIDNO:4	797-818	NM_000546.4
MP6R	Ttgcaaaacatctgttgagggc SEQIDNO:5	605-582	NM_000546.4

Metodi per la determinazione della presenza della *firma molecolare* via qRT-PCR

Dalla matrice di espressione per il dataset di Desmedt et al., 2007 (GSE7390) sono stati estratti i valori di espressione per i geni che mostravano variazione più bassa (deviazione standard inferiore a 0,4) tra i campioni; tra questi geni sono stati individuati tre geni aventi valori assoluti di espressione compresi nella variazione di espressione assoluta dei dieci geni. E' stato poi costruito uno score per la *firma molecolare* sommando i relativi score genici valutando per ognuno dei geni la relativa espressione rispetto ad uno dei geni di riferimento come da seguente tabella:

"Firma molecolare"	"Geni di riferimento"
FAM64A	ADCY3
CENPA	ADCY3
C21orf45	ADCY3
NCAPH	ADYC3
BUB1	ERCC6
CCN2	ERCC6
WDR67	ERCC6
DEPDC1	ERCC6
CPSF6	TUBGCP2
EPB41L4B	TUBGCP2

In particolare:

- a) se un gene della *firma molecolare* ha un valore di espressione inferiore di 0,5 su scala logaritmica in base 2 rispetto al valore di espressione corrispondente al

proprio gene di riferimento (che equivale a dire che esiste una differenza di 0,5 cicli di RT-PCR), ad esso viene attribuito uno score genico pari a zero.

b) se un gene della *firma molecolare* ha un valore di espressione superiore di 0.5 su scala logaritmica in base 2 rispetto al valore di espressione corrispondente al proprio gene di riferimento (che equivale a dire che esiste una differenza di 0,5 cicli di RT-PCR), ad esso viene attribuito uno score genico pari a due.

c) se un gene della *firma molecolare* ha un valore di espressione intermedio ai due valori dei casi a) e b) rispetto al valore di espressione corrispondente al proprio gene di riferimento ad esso viene attribuito uno score genico pari a uno.

La somma degli score genici (S_g) determina il valore dello score della *firma molecolare* dei dieci geni (S_f):

$$S_f = \sum S_g$$

A seconda del valore di S_f è possibile valutare la prognosi della paziente sulla base dei seguenti criteri:

a un valore dello score della *firma molecolare* dei dieci geni (S_f) inferiore a 5, è associata una prognosi favorevole (80% senza metastasi a dieci anni);

- a un valore dello score della *firma molecolare* dei dieci geni (S_f) inferiore a 9 e superiore a 5 si associa una prognosi meno favorevole (60% senza metastasi a dieci anni);
- a un valore superiore a 9 si associa una prognosi meno favorevole (50% senza metastasi a dieci anni).

Analisi dei dati di sopravvivenza su TMA

La sopravvivenza delle pazienti è stata calcolata utilizzando come intervallo temporale quello che va dalla data di inizio del trattamento alla data di morte o, nel caso in cui la paziente fosse ancora in vita, alla data in cui è stato effettuato il censimento. Le curve di sopravvivenza sono state costruite utilizzando il metodo di Kaplan-Meier. Le differenze tra le varie caratteristiche tumorali sono state stimate attraverso il test "log-rank". Il trattamento chemioterapico è stato effettuato seguendo un protocollo classico a base di antraciclina.

Analisi multivariata secondo il modello a rischi proporzionali di Cox

Per calcolare il rischio di morte nei 212 casi di tumore considerati nei TMA, è stata eseguita un'analisi utilizzando il modello di regressione a rischi proporzionali di

Cox. È stata esaminata la relazione esistente tra la sopravvivenza delle pazienti, la presenza di mutazioni missenso in *TP53* e i livelli di Pin1 e lo stato di una serie di altri parametri predittivi comunemente utilizzati nella pratica clinica, tra cui le dimensioni del tumore, la presenza o meno di recettori per gli estrogeni (stato di ER) o per il progesterone (stato di PGR), lo stato dei linfonodi, i livelli di espressione di HER2, la presenza di infiltrazione a livello dei linfonodi e il grado del tumore. Il modello di regressione a rischi proporzionali di Cox è stato applicato in prima istanza alle variabili cliniche a disposizione. Successivamente sono state inserite le variabili relative alla presenza di mutazioni missenso in *TP53* e ai livelli di Pin1. I parametri clinici sono stati considerati nella seguente maniera:

- Dimensioni del tumore: variabile continua;
- Grado del tumore: grado 1, grado 2, grado 3;
- Infiltrazione ai linfonodi: negativo se minore o uguale a 4 (0), positivo se > 4 (1);
- Stato di ER: negativo se minore o uguale a 4 (1), positivo se > 4 (0);
- Stato di PGR: negativo se minore o uguale a 4 (1), positivo se > 4 (0);
- Stato di HER2: positivo alla colorazione o negativo (0,1);
- Stato di Pin1: colorazione nucleare positiva per Pin1 > 5 (1), negativa se minore o uguale a 5 (0).

Risultati

Pin1 esalta la trasformazione tumorale in un modello murino di “knockin” per p53 mutata

Al fine di comprendere il tipo di influenza che Pin1 esercita su p53 mutata durante la tumorigenesi, sono stati incrociati topi “knockin” per p53 mutata – portatori di un allele “knockin” codificante per la proteina p53R172H ($p53^{M/+}$) (Lang G.A. et al., 2004) – con topi $Pin1^{+/-}$ (Atchinson F.W. et al., 2003). Sono state generate e analizzate due coorti di topi derivati da questi incroci, composte da animali con genotipo selvatico o knock-out per Pin1 e portatori di uno o due alleli *knockin* per p53 mutata ($p53^{M/+} Pin1^{+/+}$, $p53^{M/+} Pin1^{-/-}$, $p53^{M/+} Pin1^{+/-}$ e $p53^{M/M} Pin1^{+/+}$, $p53^{M/M} Pin1^{-/-}$). I dati di sopravvivenza relativi ai topi $p53^{M/M} Pin1^{+/-}$ e $p53^{M/+} Pin1^{+/-}$ sono risultati paragonabili a quelli ottenuti in altri studi già pubblicati (Lang G.A. et al., 2004). In entrambe le coorti alle quali appartenevano i topi privi di Pin1, la

mediana della sopravvivenza in assenza di tumore è risultata più elevata (Figura 1A e 1B). Nella coorte dei topi $p53^{M/M}$, l'assenza di Pin1 ha prodotto solo la variazione della frequenza di alcuni tipi di tumore (Tabella 7).

Tabella 7: Spettro dei tumori nella coorte di topi $p53^{M/M}$

	$p53^{M/M}$ Pin1 ^{+/+} (n=27)	$p53^{M/M}$ Pin1 ^{-/-} (n=22)
Neoplasie emopoietiche linfoidi	61 % (20/33)	42 % (11/26)
Linfoma timico	(16)	(7)
Linfoma multicentrico	(2)	(2)
Linfoma follicolare splenico	(2)	(2)
Neoplasie emopoietiche non-linfoidi	9 % (3/33)	23 % (6/26)
Sarcoma isiocitico	(2)	(6)
Leucemia mieloide	(1)	(-)
Carcinomi	3 % (1/33)	4 % (1/26)
Carcinoma prostatico	(1)	(1)
Sarcomi	27 % (9/33)	23 % (6/26)
Osteosarcoma	(0)	(1)
Emangiosarcoma	(6)	(4)
Sarcoma dei tessuti molli indifferenziato	(3)	(1)
Altri	0 % (0/33)	8 % (2/26)
Tumore del neuroectoderma primitivo	(-)	(1)
Nefroblastoma	(-)	(1)
Tumori totali	33	26

Considerando, invece, la coorte di topi $p53^{M/+}$ Pin1^{-/-} rispetto a quelli $p53^{M/+}$ Pin1^{+/+}, è stata osservata una notevole riduzione della frequenza di tumori, un ridotto numero di linfomi e la completa assenza di carcinomi (Figura 1C, Tabella 8).

Tabella 8: Spettro dei tumori nelle coorti di topi $p53^{M/+}$ e $p53^{+/-}$

	p53^{M/+} Pin1^{+/+}	p53^{M/+} Pin1^{-/-}	p53^{+/-} Pin1^{+/+}	p53^{+/-} Pin1^{-/-}
Neoplasie emopoietiche linfoidi	18 % (4/22)	12.5 % (1/8)	15 % (3/20)	20 % (4/20)
Linfoma timico	(1)	(1)	(1)	(-)
Linfoma multicentrico	(3)	(-)	(2)	(4)
Neoplasie emopoietiche non-linfoidi	14 % (3/22)	25 % (2/8)	13 % (3/20)	25% (5/20)
Sarcoma isiocitico	(2)	(5)	(3)	(5)
Leucemia mieloide	(1)	(-)	(-)	(-)
Carcinomi	27 % (6/22)	0 % (0/8)	0 % (0/20)	5 % (1/20)
Cholangiocarcinoma	(1)	(-)	(-)	(-)
Epatocarcinoma	(1)	(-)	(-)	(-)
Adenocarcinoma intestinale	(1)	(-)	(-)	(-)
Carcinoma prostatico indifferenziato	(1)	(-)	(-)	(-)
Carcinoma mammario indifferenziato	(1)	(-)	(-)	(-)
Carcinoma uroteliale della pelvi renale	(1)	(-)	(-)	(-)
Carcinoma gastrico delle cellule squamose	(-)	(-)	(-)	(1)
Sarcomi	36 % (8/22)	62.5 % (5/8)	70 % (14/20)	50 % (10/20)
Osteosarcoma	(3)	(2)	(6)	(3)
Emangiosarcoma	(2)	(1)	(-)	(1)
Sarcoma indifferenziato dei tessuti molli	(3)	(2)	(7)	(6)
Rhabdosarcoma	(-)	(-)	(1)	(-)
Altri	5 % (1/22)	0 % (0/8)	0 % (0/20)	0 % (0/20)
Adenoma pituitario	(1)	(-)	(-)	(-)

Tumori totali	22	8	20	20

Sono stati considerati anche i topi con genotipo $p53^{+/-} Pin1^{+/+}$ e $p53^{+/-} Pin1^{-/-}$, i quali non hanno fatto registrare alcuna differenza nella sopravvivenza in assenza di tumore, nella frequenza o nello spettro dei tumori (Figure 2A e 2B, Tabella 3).

Nei topi, quindi, all'assenza di Pin1 è associata una riduzione dei fenomeni di tumorigenesi solo in un contesto in cui p53 mutata risulta espressa. Da questi dati si può dedurre, quindi, che Pin1 e p53 mutata cooperano durante gli eventi che portano alla trasformazione tumorale.

Circa i meccanismi che sottendono a questa interazione genetica tra Pin1 e p53 mutata è emerso che, in campioni di tessuto tumorale provenienti da topi $p53^{M/+} Pin1^{+/+}$, Pin1 co-immunoprecipitava con p53. Dato che l'enzima riconosce specificamente sui suoi substrati i motivi S/T-P fosforilati, queste osservazioni fanno supporre che nelle cellule tumorali Pin1 possa legare p53 mutata, regolandone la funzione in risposta ai segnali di fosforilazione prodotti dallo stress oncogenico.

L'effetto di Pin1 sulla capacità di p53 mutata di favorire in cellule primarie la trasformazione tumorale indotta da Ras (Lang G.A. et al., 2004) sono stati valutati su fibroblasti embrionali (MEF) prelevati da topi $p53^{M/M} Pin1^{+/+}$, $p53^{M/M} Pin1^{-/-}$, $p53^{-/-} Pin1^{+/+}$ e $p53^{-/-} Pin1^{-/-}$ e analizzati in saggi volti a rilevarne la capacità di crescita indipendente dall'ancoraggio al substrato e il potenziale tumorigenico *in vivo* in seguito a introduzione nelle cellule di H-Ras^{V12} per mezzo di vettori retrovirali. Le cellule $p53^{M/M} Pin1^{+/+}$ si sono rivelate capaci di dare origine al doppio delle colonie in soft agar (Figura 1D, a sinistra) e, una volta iniettate sottocute in topi immunocompromessi, sono state in grado di sviluppare tumori di dimensioni due volte più grandi rispetto a quelli originati dalle cellule $p53^{M/M} Pin1^{-/-}$ o $p53^{-/-}$ (Figura 1D a destra e Figura 2C).

In cellule MEF $p53^{M/M} Pin1^{+/+}$ trasdotte con H-Ras^{V12}, in accordo con i risultati ottenuti in studi precedenti, è stato osservato che la sovraespressione della proteina H-Ras^{V12} produceva un aumento nei livelli di p53 mutata e di Pin1 (Ryo A. et al., 2002). Inoltre, gli esperimenti di "immunoblotting", dove sono stati

utilizzati anticorpi specifici che riconoscono residui amminoacidici S/T-P fosforilati, hanno mostrato che la sovraespressione di H-Ras^{V12} portava a un aumento dei livelli di fosforilazione di p53 mutata, proprio a livello dei residui amminoacidici che costituiscono i motivi di riconoscimento e legame da parte di Pin1 (Figura 1E, pannello superiore). In accordo con quanto osservato, l'interazione di p53 mutata con Pin1 risultava potenziata (Figura 1E, pannello inferiore).

Nell'insieme, queste evidenze sperimentali fanno supporre che p53 mutata abbia bisogno di Pin1 per concretizzare la sua capacità di promuovere la trasformazione cellulare e la progressione tumorale e rivelano per l'enzima un ruolo di trasduttore cruciale dei segnali oncogenici che portano all'attivazione delle funzioni acquisite da p53 mutata.

Pin1 potenzia l'effetto di p53 mutata a favore della migrazione cellulare e dell'invasività in linee cellulari tumorali umane

Anche in cellule tumorali umane delle linee MDA-MB-231 e SK-BR-3 derivate da tumori mammari è stato osservato che l'enzima lega la proteina p53 mutata endogena. Per capire se Pin1 fosse in grado in questo contesto di favorire le funzioni oncogeniche di p53 mutata è stato valutato il ruolo di Pin1 nell'induzione della capacità migratoria e dell'invasività cellulare, aspetti critici questi del fenotipo metastatico che si manifesta in seguito all'azione di p53 mutata (Adorno M. et al., 2009; Muller P. et al., 2009). Abbattendo, attraverso RNAi, i livelli di espressione di p53 mutata o in alternativa di Pin1, si è ottenuta una riduzione significativa della capacità delle cellule MDA-MB-231 di migrare e di essere invasive (Figura 3A). L'effetto esercitato da Pin1 sulla migrazione era specifico e si esplicava necessariamente in presenza della sua attività enzimatica. Reintroducendo, infatti, Pin1 nelle cellule attraverso un costrutto resistente al silenziamento (Pin1r) è stato possibile recuperare la capacità di queste di migrare. Ciò, tuttavia, non avveniva se a essere fornito era un costrutto codificante per un mutante di Pin1 (Pin1rS67E) inattivo dal punto di vista catalitico (Figura 3B). L'influenza dell'enzima nel promuovere la capacità di invasione e la migrazione è stata, poi, valutata *in vivo* iniettando nelle vene di topi immunocompromessi cellule MDA-MB-231 private di Pin1 o cellule di controllo. Le cellule carenti dell'enzima hanno mostrato una capacità notevolmente ridotta di invadere i polmoni (Figura 3C).

Il passo successivo, dunque, è stato chiedersi se Pin1 sia in grado di produrre gli effetti osservati in maniera epistatica, modulando le funzioni di p53 mutata. Come mostrato nella Figura 3D, la sovraespressione di Pin1 nelle cellule MDA-MB-231 provoca un aumento significativo della migrazione e della capacità invasiva di queste cellule, tuttavia se in concomitanza vengono abbattuti i livelli di p53 mutata, gli effetti osservati scompaiono quasi completamente.

Nel complesso questi risultati indicano che il legame di Pin1 a p53 mutata scatena così l'azione di p53 mutata a favore della migrazione cellulare.

L'asse Pin1/p53 mutata attiva uno specifico programma trascrizionale che favorisce il fenotipo tumorale aggressivo

Per cercare di comprendere in che modo Pin1 e p53 mutata possano globalmente alterare l'espressione genica, è stato analizzato il profilo trascrizionale di cellule MDA-MB-231 dopo deplezione di Pin1 o di p53 mutata. In entrambi i casi si è riscontrato un effetto sulla trascrizione di geni appartenenti alle stesse categorie funzionali, fatto che suggerisce come Pin1 e p53 mutata siano coinvolti nella regolazione di processi cellulari simili, tra cui la proliferazione, la motilità, l'organizzazione dinamica del citoscheletro, il metabolismo e la trasduzione del segnale. In particolare, sono stati identificati 386 geni che, in seguito all'abbattimento dei livelli di Pin1 o di p53 mutata, risultano sovraespressi e 303 che nelle stesse condizioni subiscono, invece, una regolazione negativa. Sulla base di questo risultato, si può supporre che un gruppo di geni siano regolati dall'azione combinata di queste due proteine (Figura 4A).

L'effetto prodotto dalla deplezione di Pin1 o di p53 mutata sull'espressione dei geni selezionati è stato confermato per qRT-PCR (Figura 5A).

Il programma trascrizionale indotto dall'azione combinata di p53 mutata e di Pin1 include geni che svolgono un ruolo rilevante nel determinare l'aggressività tumorale e potrebbe quindi influenzare l'andamento prognostico dei tumori mammari. Per verificare quest'ipotesi è stata condotta un'analisi su tumori primari alla ricerca di una correlazione tra l'espressione dei geni indotti da Pin1 e da p53 mutata (cioè quei geni che negli esperimenti in cui sono stati delineati i profili trascrizionali sono risultati regolati negativamente) e la prognosi della malattia, utilizzando quattro serie indipendenti di dati relativi a un totale di più di 800 casi di

tumore al seno (Desmedt C. et al., 2007; Miller L.D. et al., 2005; Pawitan Y. et al., 2005; Sotiriou C. et al., 2006). Tra i 303 geni regolati negativamente sono stati selezionati per l'analisi il sottogruppo (costituito da 31 geni) dei geni maggiormente repressi in entrambi le condizioni di silenziamento. Il set di dati Desmedt (Desmedt C. et al., 2007) è stato utilizzato come "set di addestramento" per la costruzione di un sistema di classificazione dei geni basato sulla correlazione con i dati clinici (vedere Tabella 1).

Dalla classificazione dei geni basata sull'associazione ai dati clinici, sono stati selezionati i geni con un punteggio maggiore di 3, ottenendo così una *firma molecolare* associata a Pin1/p53 mutata costituita dai 10 geni riportati nella tabella 1 precedentemente riportata.

L'espressione di questi geni correla in maniera significativa con l'esito clinico della malattia: nelle pazienti che esprimevano questi geni a elevati livelli l'intervallo di tempo tra la diagnosi del tumore primario e quella di metastasi in altri distretti corporei (*Time to Distant Metastasis*, TDM) è risultata più breve e la sopravvivenza ridotta (*Overall Survival*, OS) (Figure 4B e 6A-D). L'espressione di questi dieci geni presenta valore prognostico anche considerando esclusivamente le pazienti con tumore mammario con positività al recettore dell'estrogeno (Figure 6E e 6F). Simili risultati sono stati ottenuti analizzando l'espressione di un gruppo di sette geni costituito da: *C21orf45*, *CPSF6*, *DEPDC1*, *EPB41L4B*, *FAM64A*, *NCAPH* e *WDR67* (dati non mostrati).

Il passo successivo è stato quello di verificare se questi dieci geni fossero regolati dall'azione combinata di Pin1 e p53 mutata. Come mostrato nella Figura 4C, in cellule MDA-MB-231 private della proteina p53 mutata endogena, l'introduzione di un costrutto codificante per p53K280, ma non di uno per l'espressione della proteina p53K280-4M (una forma modificata della p53 mutata endogena (p53K280), incapace di legare Pin1, in cui l'amminoacido alanina ha preso il posto dei residui di serina o di treonina presenti a livello dei quattro siti S/T-P (S33, S46, T81 e S315) di fosforilazione e di legame per Pin1), ha indotto l'espressione di tutti questi geni. Tuttavia, il dato ancora più importante è che, quando in maniera ectopica Pin1 è stato espresso insieme a p53K280, l'espressione di questi geni è

stata fortemente indotta. Se invece Pin1 veniva sovraespresso da solo o insieme a p53K280-4M, non si osservava alcun effetto su di essi (Figura 4C).

Per analizzare in dettaglio il meccanismo che sottende l'attivazione trascrizionale mediata dall'azione combinata di Pin1 e p53 mutata, sono stati condotti esperimenti di immunoprecipitazione della cromatina (ChIP) in cellule MDA-MB-231. Come mostrato nella Figura 4D, p53 mutata è reclutata in maniera specifica sui promotori di tutti i geni appartenenti alla firma molecolare associata a Pin1/p53 mutata. In seguito al silenziamento di Pin1, però, il suo reclutamento è risultato ridotto di circa due volte. D'altra parte anche Pin1 è presente sugli stessi promotori, ma secondo modalità che dipendono da p53 mutata (Figura 4E).

Nel complesso, i dati indicano che questi geni sono bersagli diretti dell'azione di induzione trascrizionale svolta dall'asse Pin1/p53 mutata e che Pin1 potrebbe essere necessario affinché p53 mutata interagisca in maniera appropriata con diversi siti funzionali presenti sulla cromatina. Riguardo a ciò, un'analisi più approfondita di queste regioni promotrici ha rivelato la presenza di siti di riconoscimento e legame per diversi fattori trascrizionali noti per la loro capacità di interagire con p53 mutata. Tra questi ci sono Ets-1, NF-Y, Sp1 e VDR (Brosh R. and Rotter V., 2009) (dati non mostrati).

La maggior parte di questi geni sono risultati regolati da Pin1 e da p53 mutata anche in cellule MDA-MB-231 che esprimono come forma endogena mutante la proteina p53H273 (Figura 5B). Ciò fa supporre che Pin1 possa influenzare l'attività trascrizionale di altre proteine p53 mutanti.

Per cercare di comprendere quale ruolo funzionale svolgano questi dieci geni come mediatori degli effetti esercitati da Pin1/p53 mutata sulla migrazione cellulare, è stato analizzato cosa succede alle cellule MDA-MB-231 in seguito al loro silenziamento e sono stati identificati sei geni il cui silenziamento ha prodotto come effetto una riduzione della capacità di migrazione delle cellule. L'abbattimento dei livelli di *DEPDC1* ha prodotto l'effetto più marcato, con conseguenze anche sulla capacità invasiva di queste cellule.

Nel complesso questi risultati indicano che Pin1 agisce in concerto con p53 mutata nel riprogrammare l'espressione genica delle cellule tumorali, attivando uno specifico programma trascrizionale di cui fa parte una serie di geni importanti

nella specificazione delle caratteristiche migratorie e di invasività delle cellule. Questi geni non erano stati identificati, finora, in qualità di bersagli diretti dell'azione trascrizionale di p53 mutata.

La sovraespressione di Pin1 e la presenza di mutazioni in p53 influenzano l'esito clinico dei tumori mammari

Il passo successivo nelle analisi è stato quello di valutare nel tumore alla mammella il tipo di associazione esistente tra l'espressione di Pin1 e lo stato mutazionale di p53 e l'esito clinico della malattia. Sono stati condotti saggi di immunoistochimica quantitativa al fine di valutare l'espressione di Pin1 e di p53 in carcinomi mammari primari. Lo stato di p53 è stato analizzato tramite sequenziamento diretto e sono stati inclusi in questo studio solo quei tumori risultati portatori di alleli selvatici o di alleli con mutazioni missenso, per un totale di 212 campioni. Pin1 si è rivelato sovraespresso in 144 su 212 casi (68%), mentre in 46 su 212 (22%) sono state riscontrate mutazioni missenso in *TP53* (Tabella 9).

Tabella 9: Set di dati dei TMA

	p53 wild type	p53 mutata	Totale
Pin1 elevato	111	32	143
Pin1 basso	55	14	69
Totale	166	46	212

Non è emersa alcuna associazione tra la probabilità di sopravvivenza e la sovraespressione di Pin1 (Figura 8A). Tuttavia, in accordo con quanto osservato in studi precedenti (Langerød A. et al., 2007; Olivier M. et al., 2006) è stata riscontrata una riduzione della sopravvivenza nei casi in cui erano presenti mutazioni missenso in p53 (Figura 8B). Da sottolineare è il fatto che quando è stata valutata la probabilità di sopravvivenza in funzione della combinazione tra i livelli di espressione di Pin1 e la presenza di mutazioni missenso in p53, è emerso che la probabilità di sopravvivenza era significativamente ridotta in quei casi di tumore che esprimevano elevati livelli di Pin1 (Pin1 alto) e p53 mutata, rispetto a quelli in cui l'espressione di p53 mutata si associava alla presenza di bassi livelli di Pin1 (Pin1 basso) o a quelli in cui la proteina p53 era "wild type" (Figura 7A). L'analisi multivariata, ottenuta utilizzando il modello a rischi proporzionali di Cox, ha dimostrato, inoltre, che la presenza di una sovraespressione di Pin1 in

combinazione con quella di mutazione in p53 costituiva un fattore indipendente di predizione dell'esito clinico della malattia (Figura 7B).

L'evidenza del fatto che il valore prognostico della presenza di mutazione in p53 dipenda dai livelli di espressione di Pin1 è ulteriormente rafforzata dall'osservazione che la correlazione tra la presenza di mutazione in p53 e la riduzione della probabilità di sopravvivenza è presente solo nei casi in cui l'espressione di Pin1 è elevata (Figure 7C e 7D). Si è osservato che, in pazienti sottoposte a trattamento chemioterapico adiuvante a base di antraciclina, l'espressione di Pin1 e la presenza di mutazioni missenso in p53 sono risultate associate a esito infausto della malattia. Anche in questo sottogruppo di pazienti la combinazione tra sovraespressione di Pin1 e presenza di mutazione in p53 si è rivelata un fattore prognostico indipendente (Figura 8C).

Nel complesso, quindi, i dati prodotti dagli inventori indicano che nelle cellule tumorali l'azione combinata di Pin1 e p53 mutata incide profondamente sul programma di espressione genica attivando in maniera diretta un programma trascrizionale che promuove il fenotipo tumorale aggressivo (Figura 9).

Referenze

- Adorno, M., Cordenonsi, M., Montagner, M., Dupont, S., Wong, C., Hann, B., Solari, A., Bobisse, S., Rondina, M. B., Guzzardo, V. et al. (2009). A Mutant-p53/Smad Complex Opposes p63 to Empower TGFb-Induced Metastasis. *Cell* **137**, 87-98.
- Atchison, F. W., Capel, B., and Means, A. R. (2003). Pin1 regulates the timing of mammalian primordial germ cell proliferation. *Development* **130**, 3579-3586.
- Bao, L., Kimzey, A., Sauter, G., Sowadski, J. M., Lu, K. P., and Wang, D. G. (2004). Prevalent overexpression of prolyl isomerase Pin1 in human cancers. *Am. J. Pathol.* **164**, 1727-1737.
- Brosh, R. and Rotter, V. (2009). When mutants gain new powers: news from the mutant p53 field. *Nat. Rev. Cancer* **9**, 701-713.
- Caulin, C., Nguyen, T., Lang, G. A., Goepfert, T. M., Brinkley, B. R., Cai, W., Lozano, G., and Roop, D. R. (2007). An inducible mouse model for skin cancer reveals distinct roles for gain- and loss-of-function p53 mutations. *J. Clin. Invest.* **117**, 1893-1901.

- Coughlin, S.S., Ekwueme, D.U. (2009). Breast cancer as a global health concern. *Cancer Epidemiol.* 33, 315-8.
- Desmedt, C., Piette, F., Loi, S., Wang, Y., Lallemand, F., Haibe-Kains, B., Viale, G., De Lorenzi, M., Zhang, Y., Sagatchian D'Assignies, M. et al. (2007). Strong Time Dependence of the 76-Gene Prognostic Signature for Node-Negative Breast Cancer Patients in the TRANSBIG Multicenter Independent Validation Series. *Clin. Cancer Res.* 13, 3207-3214.
- Detre, S., Sacconi Jotti, G., and Dowsett, M. (1995). "quickscore" method for immunohistochemical semiquantitation: validation for oestrogen receptor in breast carcinomas. *J. Clin. Pathol.* 48, 876-878.
- Drost, J., Mantovani, F., Tocco, F., Elkon, R., Comel, A., Holstege, H., Kerkhoven, R., Jonkers, J., Voorhoeve, P.M., Agami, R., Del Sal, G. (2010). BRD7 is a candidate tumour suppressor gene required for p53 function. *Nat Cell Biol.* 12, 380-9.
- Efeyan, A. Murga, M., Martinez-Pastor, B., Ortega-Molina, A., Soria, R., Collado, M., Fernandez-Capetillo, O. and Serrano M. (2009). Limited Role of Murine ATM in Oncogene-Induced Senescence and p53-Dependent Tumor Suppression. *PLoS ONE.* 4 (5): e5475.
- Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., and Gaultier, L. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 5, R80.
- Guo, A. M., Coffill, C. R., and Lane, D. (2011). The role of mutant p53 in human cancer. *J of Pathology.* 223, 116-126.
- Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., and Speed, T. P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4, 249-264.
- Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., and Thun, M.J. (2009). Cancer statistics. *CA Cancer J Clin.* 59, 225-49.
- Jemal, A., Bray, F., Center, M.M., Ferlay, J., Ward, E., and Forman, D. (2011). Global cancer statistics. *CA Cancer J Clin.* 61, 69–90.

- La Vecchia, C., Bosetti, C., Lucchini, F., Bertuccio, P., Negri, E., Boyle, P., and Levi, F. (2010). Cancer mortality in Europe, 2000-2004, and an overview of trends since 1975. *Ann Oncol.* *21*, 1323-60.
- Lang, G. A., Iwakuma, T., Suh, Y., Liu, G., Rao, A., Parant, J. M., Valentin-Vega, Y. A., Terzian, T., Caldwell, L. C., Strong, L. C. et al. (2004). Gain of Function of a p53 Hot Spot Mutation in a Mouse Model of Li-Fraumeni Syndrome. *Cell* *119*, 861-872.
- Langerød, A., Zhao, H., Borgan, O., Nesland, J. M., Bukholm, I. R. K., Ikeda, T., Kåresen, R., Børresen-Dale, A. L., and Jeffrey, S. S. (2007). *TP53* mutation status and gene expression profiles are powerful prognostic markers of breast cancer. *Breast Cancer Res.* *9*, R30
- Lu, K. P. and Zhou, X. Z. (2007). The prolyl isomerase PIN1: a pivotal new twist in phosphorylation signalling and disease. *Nat. Rev. Mol. Cell Biol.* *8*, 904-916.
- Malkin, D., Li, F. P., Strong, L. C., Fraumeni, J. F. Jr., Nelson, C. E., Kim, D. H., Kassel, J., Gryka, M. A., Bischoff, F. Z., Tainsky, M. A. et al. (1990). Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* *250*, 1233-1238.
- Mantovani, F., Tocco, F., Girardini, J. E., Smith, P., Gasco, M., Lu, X., Crook, T., and Del Sal, G. (2007). The prolyl isomerase Pin1 orchestrates p53 acetylation and dissociation from the apoptosis inhibitor iASPP. *Nat. Struct. Mol. Biol.* *14*, 912-920.
- Miller, L. D., Smeds, J., George, J., Vega, V. B., Vergara, L., Ploner, A., Pawitan, Y., Hall, P., Klaar, S., Liu, E. T. et al. (2005). An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc. Natl. Acad. Sci. USA.* *102*, 13550-13555.
- Muller, P., Caswell, P. T., Doyle, B., Iwanicki, M. P., Tan, E. H., Karim, S., Lukashchuk, N., Gillespie, D. A., Ludwig, R. L., Gosselin, P. et al. (2009). Mutant p53 Drives Invasion by Promoting Integrin Recycling. *Cell* *139*, 1327-1341.
- Olivier, M., Langerød, A., Carrieri, P., Bergh, J., Klaar, S., Eyfjord, J., Theillet, C., Rodriguez, C., Lidereau, R., Bieche, I. et al. (2006). The Clinical Value

of Somatic TP53 GeneMutations in 1,794 Patients with Breast Cancer. *Clin. Cancer. Res.* 12, 1157-1167.

- Pawitan, Y., Bjöhle, J., Amler, L., Borg, A. L., Egyhazi, S., Hall, P., Han, X., Holmberg, L., Huang, F., Klaar, S. et al. (2005). Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. *Breast Cancer Res.* 7, R953-964.
- Prat, A., Perou, C.M. (2011). Deconstructing the molecular portraits of breast cancer. *Mol Oncol.* 5, 5-23.
- Rustighi, A., Tiberi, L., Soldano, A., Napoli, M., Nuciforo, P., Rosato, A., Kaplan, F., Capobianco, A., Pece, S., Di Fiore, P. P. et al. (2009). The prolyl-isomerase Pin1 is a Notch1 target that enhances Notch1 activation in cancer. *Nat. Cell Biol.* 11, 133-142.
- Ryo, A., Liou, Y. C., Wulf, G., Nakamura, M., Lee, S. W., and Lu, K. P. (2002). PIN1 Is an E2F Target Gene Essential for Neu/Ras-Induced Transformation of Mammary Epithelial Cells. *Mol. Cell. Biol.* 22, 5281-5295.
- Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3, Article 3.
- Song, H., Hollstein, M., and Xu, Y. (2007). p53 gain-of-function cancer mutants induce genetic instability by inactivating ATM. *Nature Cell Biology* 9, 573-580.
- Sotiriou, C., Wirapati, P., Loi, S., Harris, A., Fox, S., Smeds, J., Nordgren, H., Farmer, P., Praz, V., Haibe-Kains, B. et al. (2006). Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J. Natl. Cancer Inst.* 98, 262-272.
- Soussi, T. and Beroud, C. (2001). Assessing TP53 status in human tumours to evaluate clinical outcome. *Nature Reviews on Cancer* 1, 233-240.
- Soussi, T. and Wiman, K. G. (2007). Shaping Genetic Alterations in Human Cancer: The p53 Mutation Paradigm. *Cancer Cell* 12, 303-312.
- Terzian, T., Suh, Y., Iwakuma, T., Post, S. M., Neumann, M., Lang, G. A., Van Pelt, C. S., and Lozano, G. (2008). The inherent instability of mutant p53 is alleviated by Mdm2 or p16INK4a loss. *Genes & Dev.* 22, 1337-1344.

- UKCCCR (1989). UKCCCR guidelines for the welfare of animals in experimental neoplasia. *Cancer Metastasis Rev.* 8, 82-88.
- van't Veer, L. J., Dai, H., van de Vijver, M. J., He, Y. D., Hart, A. A., Mao, M., Peterse, H. L., van der Kooy, K., Marton, M. J., Witeveen, A. T. et al. (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415, 530-536.
- Wulf, G., Garg, P., Liou, Y. C., Iglehart, D., and Lu, K. P. (2004). Modeling breast cancer in vivo and ex vivo reveals an essential role of Pin1 in tumorigenesis. *EMBO J.* 23, 3397-3407.
- Yeh, E. and Means, A. R. (2007). PIN1, the cell cycle and cancer. *Nat. Rev. Cancer* 7, 381-388.
- Zacchi, P., Gostissa, M., Uchida, T., Salvagno, C., Avolio, F., Voliniak, S., Ronai, Z., Blandino, G., Schneider, C., and Del Sal, G. (2002). The prolyl isomerase Pin1 reveals a mechanism to control p53 functions after genotoxic insults. *Nature* 419, 853-857.
- Zheng, H., You, H., Zhou, X. Z., Murray, S. A., Uchida, T., Wulf, G., Tang, X., Lu, K. P., and Xiao, Z.-X. (2002). The Prolyl isomerase Pin1 is a novel regulator of p53 in genotoxic response. *Nature* 419, 849-853.

Rivendicazioni

1. Metodo per la prognosi di tumori al seno comprendente la determinazione dei livelli di espressione di Pin1 e la rilevazione dello stato mutazionale del gene *TP53* e/o la determinazione dei livelli di p53 in un campione.
2. Metodo per la prognosi di tumori al seno comprendente la determinazione di una *firma molecolare* consistente nella quantificazione in un campione dell'espressione del panel di geni: *DEPDC1* (Gene ID No 55635, isoforma 1: RefSeq NM_001114120.1, NP_001107592.1; isoforma 2: RefSeq NM_017779.4, NP_060249.2), *CPSF6* (Gene ID No 11052, RefSeq NM_007007.2, NP_008938.2), *C21orf45* (Gene ID No 54069, RefSeq NM_018944.2, NP_061817.1), *FAM64A* (Gene ID No 54478, RefSeq NM_019013.2, NP_061886.2), *EPB41L4B* (Gene ID No 54566, isoforma 1: RefSeq NM_018424.2, NP_060894.2; isoforma 2: RefSeq NM_019114.3, NP_061987.3), *NCAPH* (Gene ID No 23397, RefSeq NM_015341.3, NP_056156.2), *WDR67* (Gene ID No 93594, isoforma 1: RefSeq NM_145647.3, NP_663622.2; isoforma 2: RefSeq NM_001145088.1, NP_001138560.1).
3. Metodo secondo la rivendicazione 2, comprendente l'ulteriore determinazione dell'espressione dei geni *CENPA* (Gene ID No 1058, isoforma 1: RefSeq NM_001809.3, NP_001800.1; isoforma 2: RefSeq NM_001042426.1, NP_001035891.1), *CCNE2* (Gene ID No 9134, RefSeq NM_057749.2, NP_477097.1), *BUB1* (Gene ID No 699, RefSeq NM_004336.3, NP_004327.1).
4. Metodo secondo una delle rivendicazioni da 1 a 3, in cui il campione è una biopsia di tumore del seno fresca o congelata o fissata in paraffina o un acido nucleico estratto dalla stessa.
5. Metodo secondo la rivendicazione 1, in cui la determinazione dei livelli di espressione di Pin1 e la rilevazione dello stato mutazionale del gene *TP53* e/o dei livelli di p53 è eseguita con tecniche molecolari e/o immunologiche.
6. Metodo secondo le rivendicazioni 2 e 3, in cui la determinazione dei livelli di espressione dei geni della firma è eseguita con tecniche molecolari e/o immunologiche.

7. Metodo secondo le rivendicazioni 5 e 6, in cui le tecniche molecolari sono scelte tra PCR, analisi su microarray e sequenziamento.
8. Metodo secondo le rivendicazioni 5 e 6, in cui l'espressione della proteina e/o genica è misurata con almeno un reagente di determinazione selezionato nel gruppo consistente di:
 - un oligonucleotide per l'amplificazione dell'acido nucleico da determinare derivato sulla sequenza del gene in esame o sulla sua sequenza complementare;
 - un anticorpo specifico marcato.
9. Metodo per la prognosi di tumori al seno secondo una delle rivendicazioni da 1 a 8, in cui la valutazione prognostica è eseguita sulla base di un punteggio derivato dalla quantificazione di un segnale rilevato per l'espressione delle proteine e/o dei geni in esame.
10. Metodo per la prognosi di tumori al seno secondo la rivendicazione 9, in cui quando il segnale di espressione è rilevato con metodi immunologici sui livelli delle proteine Pin1 e p53 il punteggio discriminante per la prognosi è inferiore o superiore a 5.
11. Metodo per la prognosi di tumori al seno secondo la rivendicazione 9, in cui quando il segnale di espressione è rilevato con metodi molecolari sui livelli di geni della firma molecolare il punteggio (score genico S_f) è ottenuto dalla somma degli score genici (score firma S_g) di ogni singolo gene

$$S_f = \sum S_g$$

calcolato valutando l'espressione relativa di ogni gene della *firma molecolare* con l'espressione di geni di riferimento.

12. Metodo per la prognosi di tumori al seno secondo la rivendicazione 11, in cui i geni di riferimento sono scelti tra *ERCC6* (Gene ID No 2074, RefSeq. NM_000124.2, NP_000115.1), *ADCY3* (Gene ID No 109, RefSeq. NM_004036.3, NP_004027.2), *TUBGCP2* (Gene ID No 10844, RefSeq. NM_006659.2, NP_006650.1), *SOX15* (Gene ID No 6665 RefSeq. NM_006942.1, NP_008873.1), *RPL10* (Gene ID No 6134 Homo sapiens RefSeq. NM_006013.3, NP_006004.2), *DGCR14* (Gene ID No 8220 RefSeq. NM_022719.2, NP_073210.1), *KIAA0586* (Gene ID No 9786

- RefSeq. NM_014749.3, NP_055564.3), ZBTB17 (Gene ID No 7709 NM_003443.2, NP_003434.2), *TFIP11* (Gene ID No 24144 RefSeq. NM_001008697.1, NP_001008697.1), *ZNF672* (Gene ID No 79894 RefSeq. NM_024836.1, NP_079112.1), *TXN2* (Gene ID No 25828 RefSeq. NM_012473.3, NP_036605.2).
13. Metodo per la prognosi di tumori al seno secondo la rivendicazione 11, in cui i geni di riferimento sono preferibilmente *ERCC6* (Gene ID 2074, RefSeq. NM_000124.2, NP_000115.1), *ADCY3* (Gene ID 109, RefSeq. NM_004036.3, NP_004027.2), *TUBGCP2* (Gene ID 10844, RefSeq. NM_006659.2, NP_006650.1).
14. Metodo per la prognosi di tumori al seno secondo la rivendicazione 11, in cui a un valore di S_f inferiore a 5 è associata una prognosi favorevole, a un valore di S_f inferiore a 9 e superiore a 5 è associata una prognosi intermedia e ad un valore di S_f superiore a 9 è associata una prognosi meno favorevole.
15. *Firma molecolare* minima consistente nel set di geni *DEPDC1* (Gene ID No 55635, isoforma 1: RefSeq NM_001114120.1, NP_001107592.1; isoforma 2: RefSeq NM_017779.4, NP_060249.2), *CPSF6* (Gene ID No 11052, RefSeq NM_007007.2, NP_008938.2), *C21orf45* (Gene ID No 54069, RefSeq NM_018944.2, NP_061817.1), *FAM64A* (Gene ID No 54478, RefSeq NM_019013.2, NP_061886.2), *EPB41L4B* (Gene ID No 54566, isoforma 1: RefSeq NM_018424.2, NP_060894.2; isoforma 2: RefSeq NM_019114.3, NP_061987.3), *NCAPH* (Gene ID No 23397, RefSeq NM_015341.3, NP_056156.2), *WDR67* (Gene ID No 93594, isoforma 1: RefSeq NM_145647.3, NP_663622.2; isoforma 2: RefSeq NM_001145088.1, NP_001138560.1) per uso in un metodo di prognosi in tumori al seno basata sull'espressione di detti geni.
16. *Firma molecolare* secondo la rivendicazione 15, in cui al panel dei geni sono aggiunti gli ulteriori geni *CENPA* (Gene ID No 1058, isoforma 1: RefSeq NM_001809.3, NP_001800.1; isoforma 2: RefSeq NM_001042426.1, NP_001035891.1), *CCNE2* (Gene ID No 9134, RefSeq

NM_057749.2, NP_477097.1), *BUB1* (Gene ID No 699, RefSeq NM_004336.3, NP_004327.1).

17. *Firma molecolare* secondo le rivendicazioni 15 o 16 per l'uso di un metodo di prognosi basato sull'espressione dei geni di tale firma in tumori al seno (positivi per l'estrogeno) ER+.
18. Kit per un metodo di prognosi di tumori al seno su un campione bioptico da una paziente con tumore al seno comprendente almeno:
 - uno specifico reagente per la determinazione dell'espressione della Pin1 consistente in un anticorpo poli- o monoclonale;
 - uno specifico reagente per la determinazione dell'espressione della p53 consistente in un anticorpo poli- o monoclonale; e/o
 - almeno un oligonucleotide per l'amplificazione del gene *TP53* derivato sulla sua sequenza di riferimento o sulla sequenza complementare della stessa;
 - un foglietto di istruzioni per calcolare lo score.
19. Kit per un metodo di prognosi di tumori al seno su un campione bioptico da una paziente con tumore al seno comprendente almeno:
 - uno specifico reagente per la determinazione dell'espressione delle proteine espresse da almeno i 7 geni della *firma molecolare* minima consistente in un anticorpo poli- o monoclonale; e/o
 - oligonucleotidi per l'amplificazione di almeno i 7 geni della *firma molecolare* minima e di uno o più geni di riferimento derivati sulle sequenze di riferimento degli stessi o sulle loro sequenze complementari;
 - un foglietto per le istruzioni per calcolare lo score della firma molecolare.
20. Kit per un metodo di prognosi di tumori al seno su un campione bioptico da una paziente con tumore al seno comprendente almeno:
 - uno specifico reagente per la determinazione dell'espressione delle proteine espresse dai 10 geni della *firma molecolare* consistente in un anticorpo poli- o monoclonale; e/o

- oligonucleotidi per l'amplificazione dei 10 geni della *firma molecolare* e uno o più geni di riferimento derivati sulle sequenze di riferimento degli stessi o sulle loro sequenze complementari;
- un foglietto per le istruzioni per calcolare lo score della firma molecolare.

Claims

1. A method for breast cancer prognosis comprising the detection in a sample of Pin1 expression levels and *TP53* gene mutation status and/or p53 levels.
2. A method for breast cancer prognosis comprising the detection of a *molecular signature* consisting in the detection in a sample, of the expression levels of the gene panel: *DEPDC1* (Gene ID No 55635, isoform 1: RefSeq NM_001114120.1, NP_001107592.1; isoform 2: RefSeq NM_017779.4, NP_060249.2), *CPSF6* (Gene ID No 11052, RefSeq NM_007007.2, NP_008938.2), *C21orf45* (Gene ID No 54069, RefSeq NM_018944.2, NP_061817.1), *FAM64A* (Gene ID No 54478, RefSeq NM_019013.2, NP_061886.2), *EPB41L4B* (Gene ID No 54566, isoform 1: RefSeq NM_018424.2, NP_060894.2; isoform 2: RefSeq NM_019114.3, NP_061987.3), *NCAPH* (Gene ID No 23397, RefSeq NM_015341.3, NP_056156.2), *WDR67* (Gene ID No 93594, isoform 1: RefSeq NM_145647.3, NP_663622.2; isoform 2: RefSeq NM_001145088.1, NP_001138560.1).
3. A method for breast cancer prognosis according to claim 2, comprising the further detection of gene expression of *CENPA* (Gene ID No. 1058, isoform 1: RefSeq NM_001809, NP_663622.2; isoform 2: RefSeq NM_001042426.1, NP_001035891.1), *CCNE2* (Gene ID No 9134, RefSeq NM_057749.2, NP_477097.1), *BUB1* (Gene ID No 699, RefSeq NM_004336.3, NP_004327.1).
4. A method for breast cancer prognosis according to one of the claims from 1 to 3, wherein the sample is a fresh or frozen or paraffin embedded breast cancer biopsy, or a nucleic acid extracted from the same.
5. A method for breast cancer prognosis according to claim 1, wherein the detection of Pin1 expression levels and *TP53* gene mutation status and/or p53 levels is performed with molecular and/or immunologic techniques.
6. A method for breast cancer prognosis according to claims 2 and 3, wherein the detection of the expression levels of signature genes is performed with molecular and/or immunologic techniques.
7. A method for breast cancer prognosis according to claims 5 and 6, wherein

the molecular techniques are selected from PCR, microarray analysis and sequencing.

8. A method for breast cancer prognosis according to claims 5 and 6, wherein the protein and/or gene expression is measured with at least one detection reagent selected in a group consisting of:
 - an oligonucleotide for amplifying the nucleic acid to be detected, derived from the sequence of the pertaining gene or from its complementary sequence;
 - a specific labeled antibody.
9. A method for breast cancer prognosis according to one of the claims from 1 to 8, wherein the prognostic evaluation is performed considering a score derived from the quantification of a signal detected for the expression of pertaining proteins and/or genes.
10. A method for breast cancer prognosis according to claim 9, wherein when the expression signal of Pin1 and p53 protein levels is detected with immunological methods the discriminant score for the prognosis is less or more than 5.
11. A method for breast cancer prognosis according to claim 9, wherein when the expression signal of *molecular signature* gene levels is detected by molecular methods the score (signature score S_f) is obtained from the sum of gene scores (S_g) of each gene

$$S_f = \sum S_g$$

calculated by assessing the relative expression of each molecular signature gene with the expression of reference genes.

12. A method for breast cancer prognosis according to claim 11, wherein the reference genes are selected from: *ERCC6* (Gene ID No 2074, RefSeq. NM_000124.2, NP_000115.1), *ADCY3* (Gene ID No 109, RefSeq. NM_004036.3, NP_004027.2), *TUBGCP2* (Gene ID No 10844, RefSeq. NM_006659.2, NP_006650.1), *SOX15* (Gene ID No 6665 RefSeq. NM_006942.1, NP_008873.1), *RPL10* (Gene ID No 6134 Homo sapiens RefSeq. NM_006013.3, NP_006004.2), *DGCR14* (Gene ID No 8220 RefSeq. NM_022719.2, NP_073210.1), *KIAA0586* (Gene ID No 9786

- RefSeq. NM_014749.3, NP_055564.3), ZBTB17 (Gene ID No 7709 NM_003443.2, NP_003434.2), *TFIP11* (Gene ID No 24144 RefSeq. NM_001008697.1, NP_001008697.1), *ZNF672* (Gene ID No 79894 RefSeq. NM_024836.1, NP_079112.1), *TXN2* (Gene ID No 25828 RefSeq. NM_012473.3, NP_036605.2).
13. A method for breast cancer prognosis according to claim 11, wherein the reference genes are: *ERCC6* (Gene ID 2074, RefSeq. NM_000124.2, NP_000115.1), *ADCY3* (Gene ID 109, RefSeq. NM_004036.3, NP_004027.2), *TUBGCP2* (Gene ID 10844, RefSeq. NM_006659.2, NP_006650.1).
 14. A method for breast cancer prognosis according to claim 11, wherein a S_f value less than 5 is associated with a favorable prognosis, a S_f value less than 9 and more than 5 is associated with an intermediate prognosis and a S_f value more than 9 is associated with a less favorable prognosis.
 15. A minimum *molecular signature* consisting of the set of genes *DEPDC1* (Gene ID No. 55635, isoform 1: RefSeq NM_001114120.1, NP_001107592.1; isoform 2: RefSeq NM_017779.4, NP_060249.2), *CPSF6* (Gene ID No 11052, RefSeq NM_007007.2, NP_008938.2), *C21orf45* (Gene ID No 54069, RefSeq NM_018944.2, NP_061817.1), *FAM64A* (Gene ID No 54478, RefSeq NM_019013.2, NP_061886.2), *EPB41L4B* (Gene ID No 54566, isoform 1: RefSeq NM_018424.2, NP_060894.2; isoform 2: RefSeq NM_019114.3, NP_061987.3), *NCAPH* (Gene ID No 23397, RefSeq NM_015341.3, NP_056156.2), *WDR67* (Gene ID No 93594, isoform 1: RefSeq NM_145647.3, NP_663622.2; isoform 2: RefSeq NM_001145088.1, NP_001138560.1) for the use in a method of breast cancer prognosis based on the expression of these genes.
 16. A *molecular signature* according to claim 15, wherein the additional genes *CENPA* (Gene ID No. 1058, isoform 1: RefSeq NM_001809, NP_663622.2; isoform 2: RefSeq NM_001042426.1, NP_001035891.1), *CCNE2* (Gene ID No 9134, RefSeq NM_057749.2, NP_477097.1), *BUB1* (Gene ID No 699, RefSeq NM_004336.3, NP_004327.1) are added to the gene panel.
 17. A *molecular signature* according to claims 15 or 16, for the use of a method

of prognosis based on gene expression of such signature in ER+ (estrogen positive) breast tumors.

18. A kit for a method of breast cancer prognosis on a biopsy sample from a breast cancer patient comprising at least:
 - a specific reagent for detecting the expression of Pin1, consisting of a poly- or monoclonal antibody;
 - a specific reagent for detecting the expression of p53, consisting of a poly- or monoclonal antibody; and/or,
 - at least one oligonucleotide for amplifying the *TP53* gene derived from its reference sequence or from the complementary sequence thereof;
 - a leaflet of instructions to calculate the score.
19. A kit for a method of breast cancer prognosis on a biopsy sample from a breast cancer patient comprising at least:
 - a specific reagent for detecting the expression of proteins expressed from at least the 7 genes of the minimum *molecular signature* consisting of a poly- or monoclonal antibody; and/or
 - oligonucleotides for amplifying at least the 7 genes of the minimum *molecular signature* and one or more reference genes derived from either the reference sequences thereof or their complementary sequences;
 - a leaflet of instructions to calculate the molecular signature score.
20. A kit for a method of breast cancer prognosis on a biopsy sample from a breast cancer patient comprising at least:
 - a specific reagent for detecting the expression of proteins expressed from the 10 genes of the *molecular signature* consisting of a poly- or monoclonal antibody; and/or
 - oligonucleotides for amplifying the 10 genes of the *molecular signature* and one or more reference genes derived from either the reference sequences thereof or their complementary sequences;
 - a leaflet of instructions to calculate the molecular signature score.

Figura 1

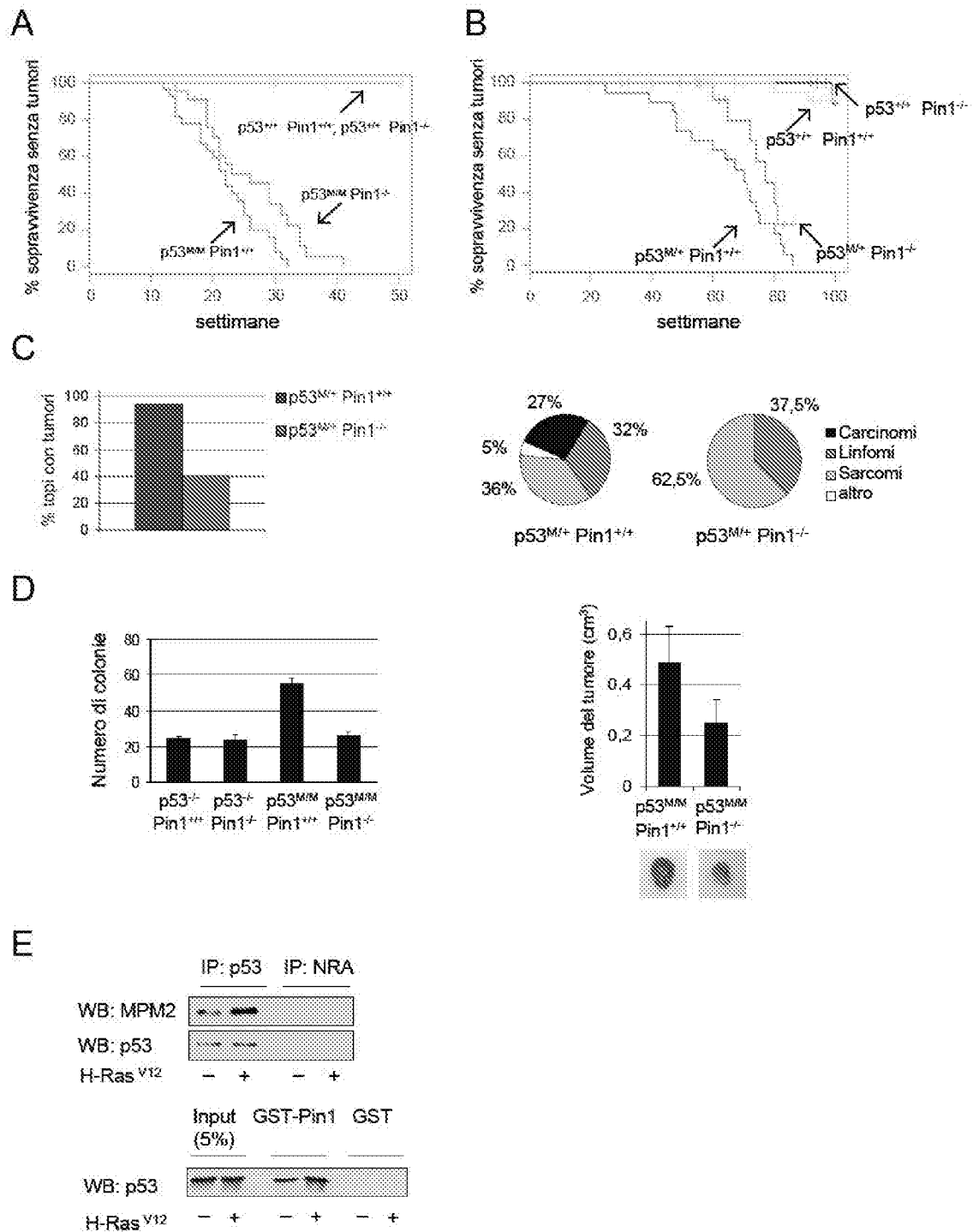


Figura 2

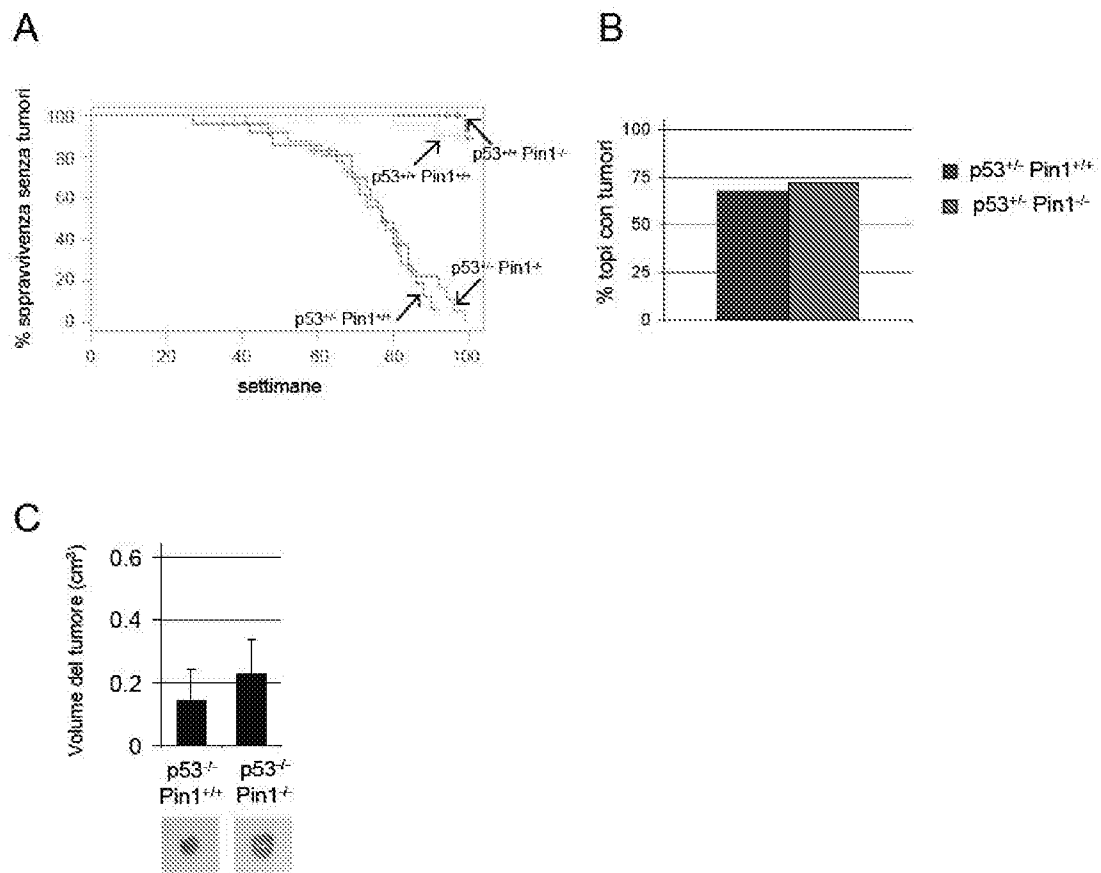


Figura 3

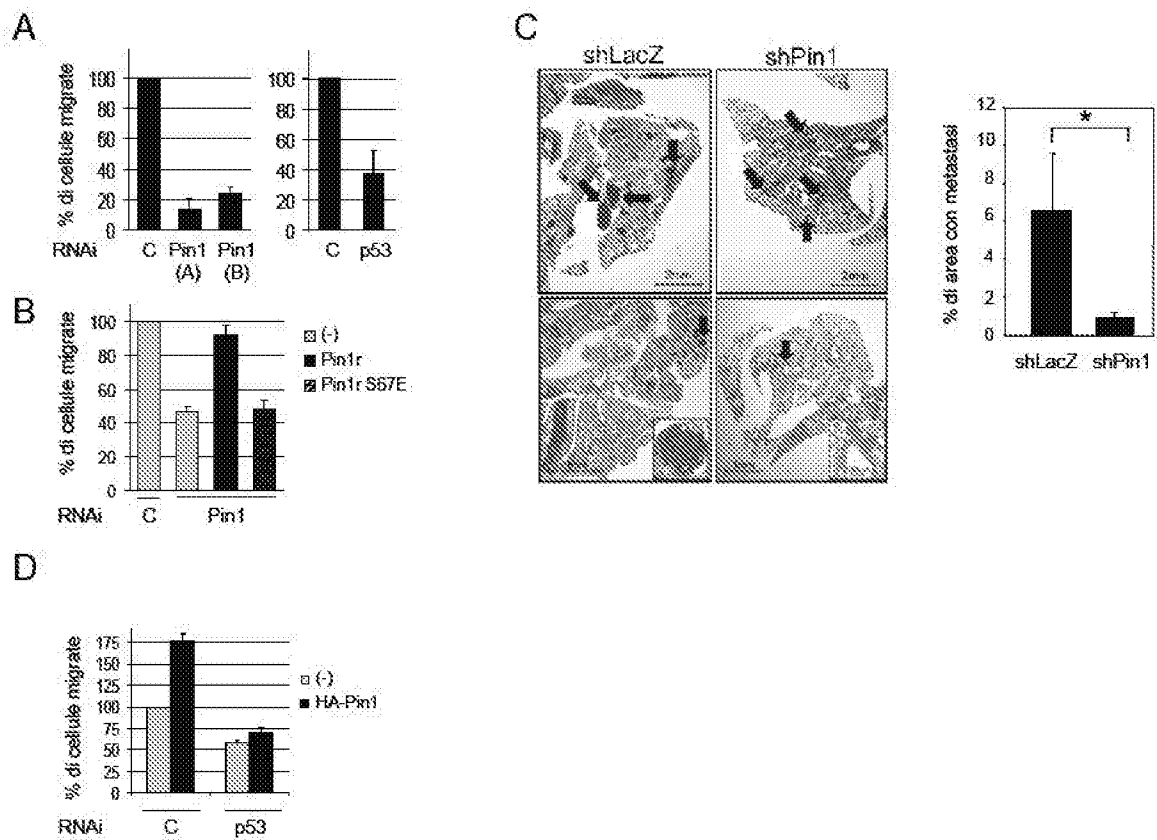


Figura 4

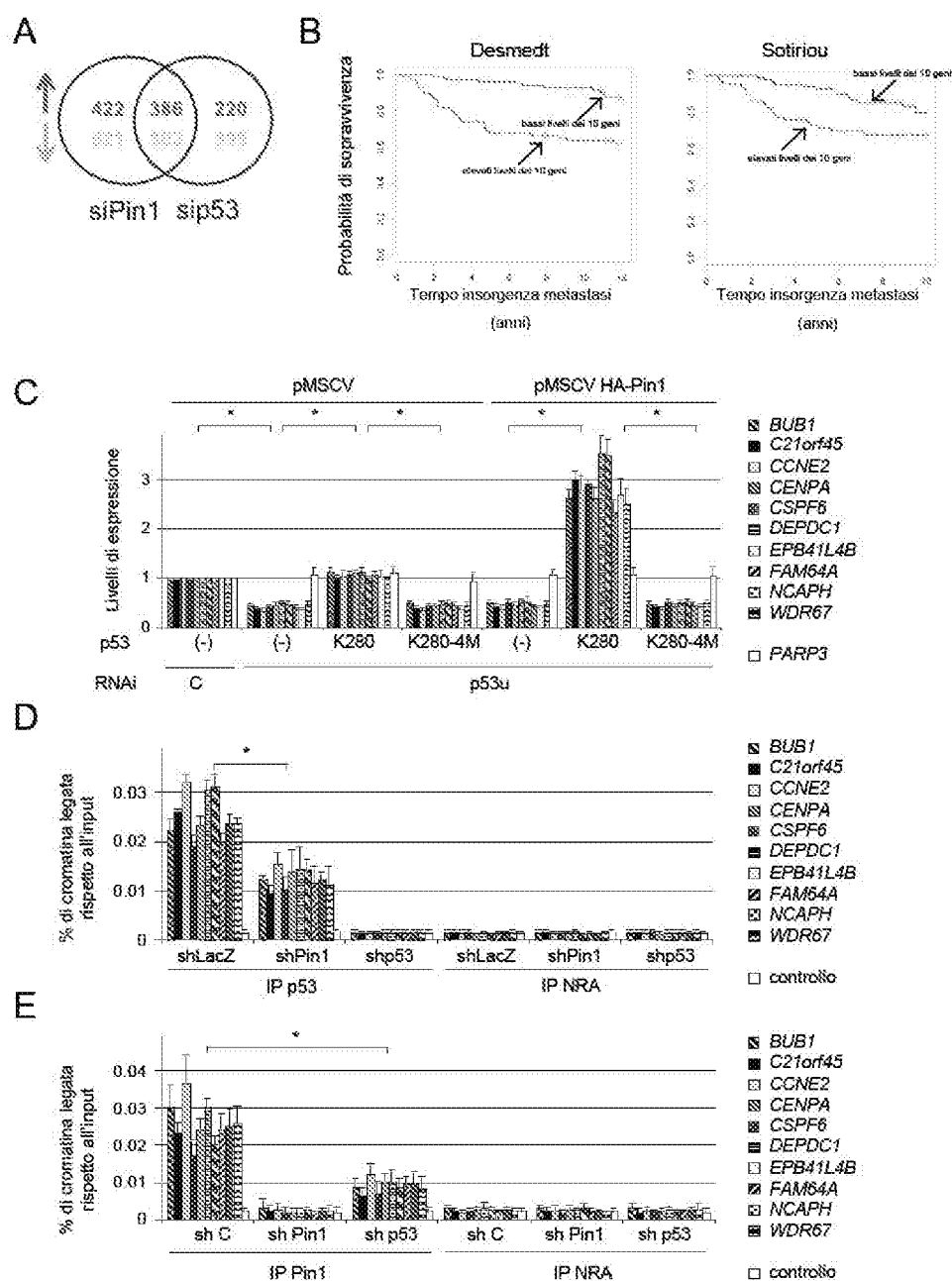
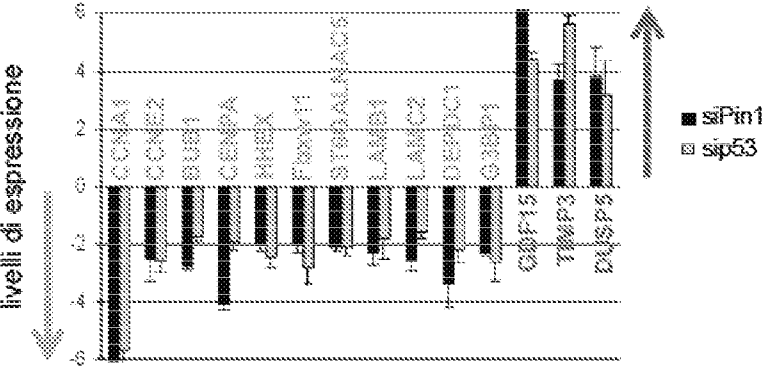


Figura 5

A



B

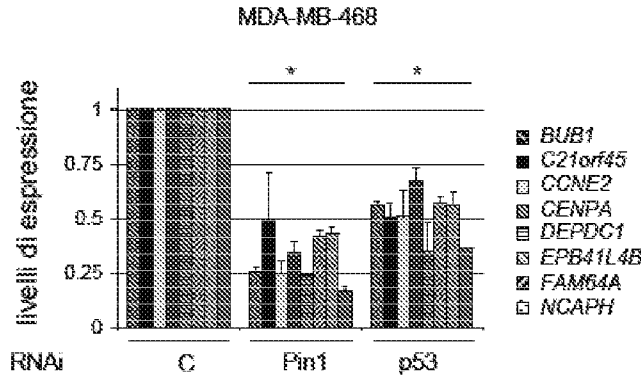


Figura 6

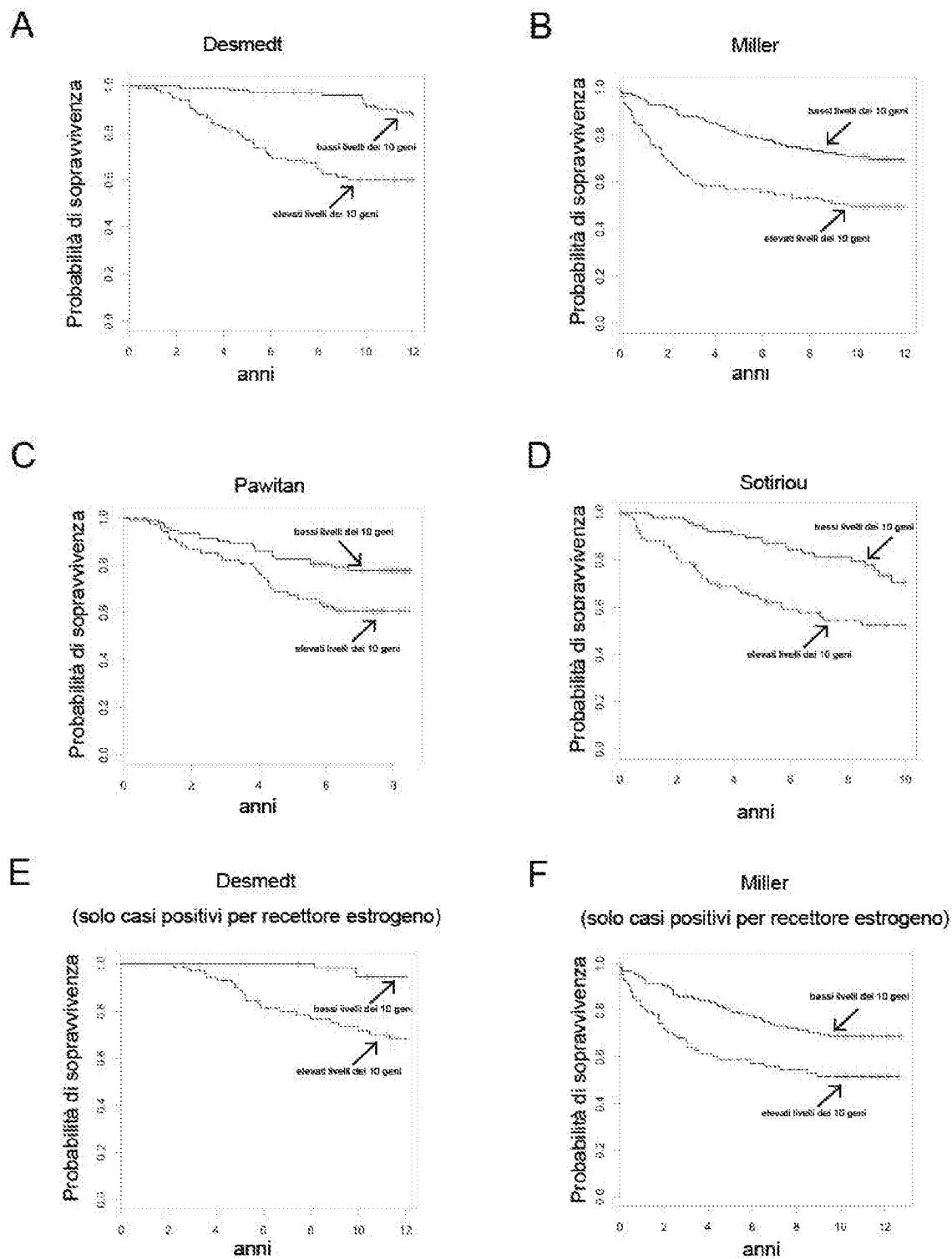


Figura 7

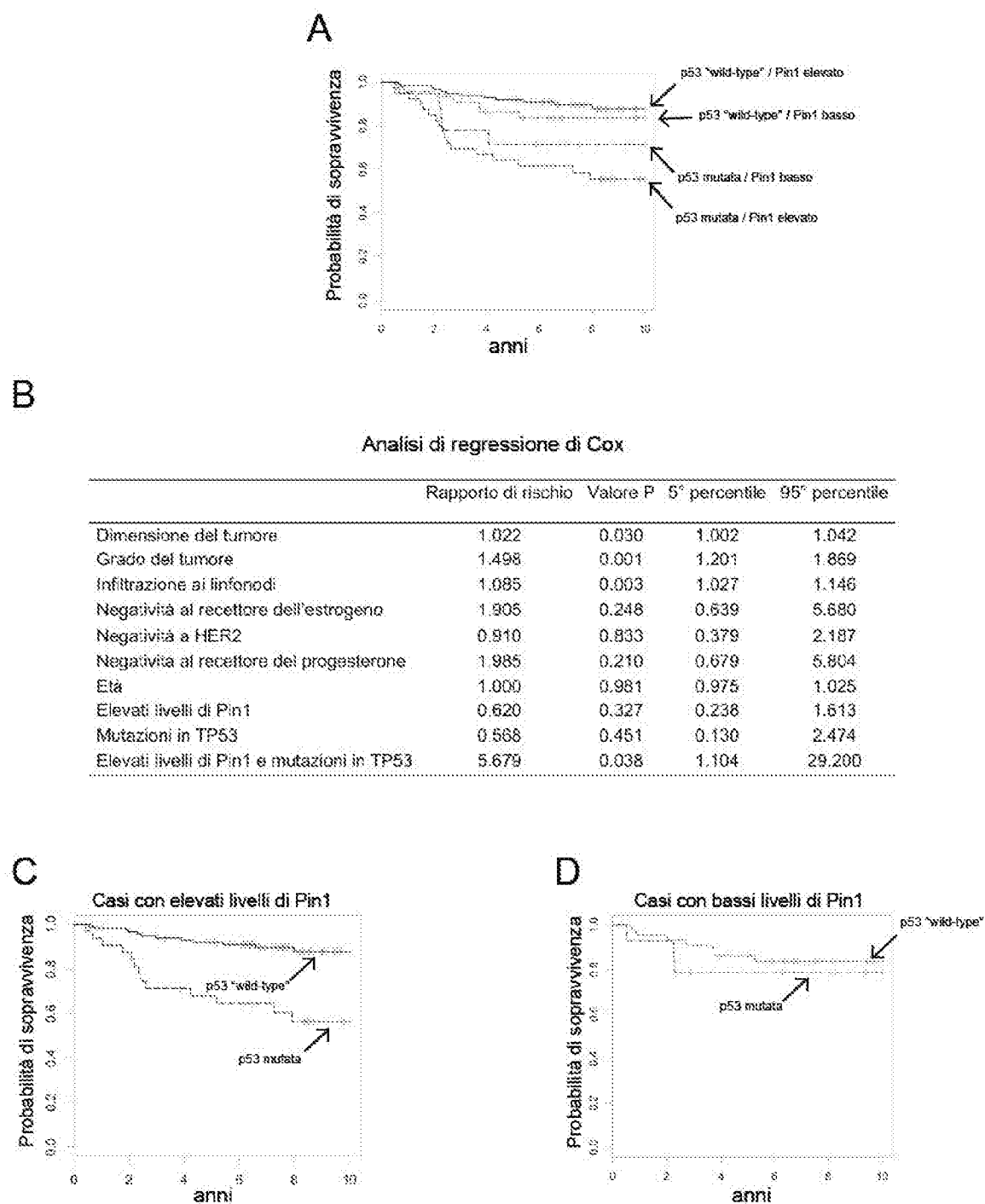
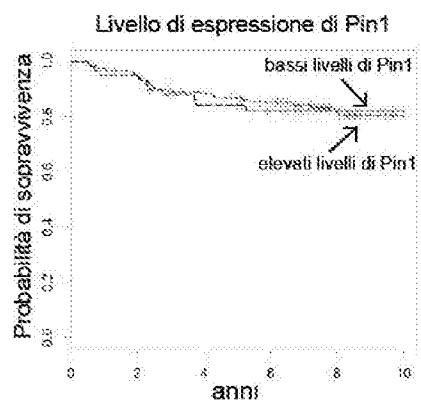
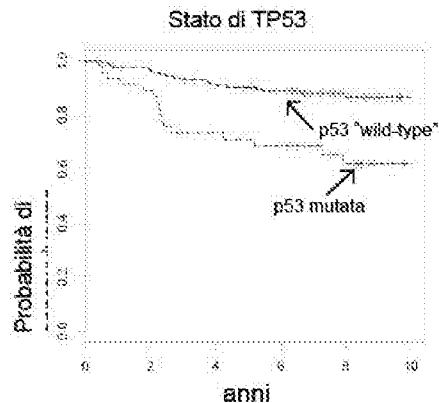


Figura 8

A



B



C

Analisi di regressione di Cox
(solo pazienti sottoposti a chemioterapia)

	Rapporto di rischio	Valore P	5° percentile	95° percentile
Dimensione del tumore	1.025	0.219	0.985	1.066
Grado del tumore	2.576	0.221	0.566	11.729
Infiltrazione ai linfonodi	1.050	0.241	0.967	1.140
Negatività al recettore dell'estrogeno	1.487	0.604	0.332	6.670
Negatività a HER2	0.697	0.533	0.224	2.168
Negatività al recettore del progesterone	0.920	0.045	0.676	0.996
Età	0.980	0.373	0.938	1.024
Elevati livelli di Pin1	0.855	0.848	0.174	4.207
Mutazioni in TP53	0.264	0.278	0.024	2.925
Elevati livelli di Pin1 e mutazioni in TP53	16.484	0.039	1.151	236.322

Figura 9

