

TRADUZIONE DEL TESTO DEL BREVETTO EUROPEO N. 3488443

DAL TITOLO:

"SELEZIONE DI NEOEPITOPI COME TARGET SPECIFICI DELLA MALATTIA PER UNA TERAPIA CON MAGGIORE EFFICACIA"

Depositata il:

** * **

DESCRIZIONE

CAMPO TECNICO

Il presente insegnamento si riferisce a metodi per determinare l'idoneità di un neoepitopo specifico della malattia come un target specifico della malattia e all'utilizzo di tali neoepitopi idonei identificati nell'immunoterapia mirata specificamente al tessuto malato di un paziente, come il tessuto tumorale, che esprime uno o più dei neoepitopi idonei identificati.

BACKGROUND

Il cancro è una delle principali cause di mortalità, pari a 1 su 4 di tutti i decessi. Il trattamento del cancro è stato tradizionalmente basato sulla legge delle medie, che funziona meglio per il maggior numero di pazienti. Tuttavia, a causa dell'eterogeneità molecolare nel cancro, spesso meno del 25% degli individui trattati beneficia delle terapie approvate. La medicina individualizzata basata sul trattamento su misura dei pazienti è considerata una potenziale soluzione a bassa efficacia e dai costi elevati per l'innovazione nello sviluppo di farmaci.

Le immunoterapie del cancro personalizzate stanno emergendo come una potenziale svolta nel trattamento del cancro con il potenziale di trasformare lo standard di cura per i milioni di malati di cancro diagnosticati ogni anno in tutto il mondo. L'aspetto unificante delle immunoterapie del cancro personalizzate consente al sistema immunitario di prendere di mira le anomalie genetiche (mutazioni) tipiche del cancro di un paziente. Tali mutazioni specifiche per la malattia possono codificare per i neoepitopi, i cui neoepitopi sono target specifici della malattia. Le anomalie genetiche più diffuse che affliggono i genomi del cancro che possono essere utilizzati come target specifici della malattia per immunoterapie personalizzate sono variazioni a singolo nucleotide non sinonimo (SNV). Pertanto, l'identificazione precisa ed esaustiva dei SNV di un paziente nelle regioni codificanti del genoma è una fase fondamentale nel processo di produzione di immunoterapie del cancro personalizzate.

Tuttavia, come qui descritto, conoscere l'identità della mutazione specifica della malattia è solo parte del quadro. Piuttosto, il profilo genetico completo di una mutazione richiede la conoscenza del numero esatto di copie del gene contenente la mutazione nella cellula malata, ad esempio, nella cellula tumorale (che include sia gli alleli wild-type sia quelli mutati), il numero di copie dell'allele mutato nella cellula tumorale (indicato qui

come zigosità della mutazione) e il grado di subclonalità della mutazione in un campione di cellule malate, come un campione tumorale. Infatti, le variazioni del numero di copie che si verificano nelle cellule malate sono un componente importante della variazione genetica nelle cellule malate nella maggior parte delle indicazioni della malattia. Inoltre, l'entità delle variazioni del numero di copie, l'identità dei geni interessati dalle variazioni del numero di copie e la precisa composizione genetica delle variazioni del numero di copie è unica per ogni individuo e può variare ampiamente da un individuo all'altro. Vedere, in generale, Shlien e Malkin, 2009, *Genome Med.* 1:62; Yang et al., 2013, *Cell* 153:919-929. La conoscenza precisa di queste caratteristiche genetiche può essere fondamentale per selezionare mutazioni che, quando mirate, sarebbero immuni alla fuga tumorale e quindi potrebbero avere il potenziale di conferire il controllo totale al tumore.

Al fine di massimizzare l'efficacia delle immunoterapie del cancro personalizzate e conferire un controllo duraturo al tumore per la maggior parte dei pazienti trattati, le terapie devono eludere in qualche modo la capacità dei tumori di sfuggire all'immunosorveglianza, ad esempio silenziando l'espressione del target mutato, ad esempio eliminando il gene. Senza affrontare questo problema, le immunoterapie corrono il rischio di ricaduta poiché l'immunoterapia non può mirare alle mutazioni se non sono espresse, ad esempio, eliminate dal genoma. La selezione di neoepitopi adatti che migliorano il controllo del tumore andrebbe a beneficio di tutti gli approcci di immunoterapia personalizzata che prendono di mira i neoepitopi, indipendentemente da come vengono implementati.

Lo stato della tecnica ha suggerito che probabilmente non c'è modo di prevedere epitopi rilevanti con assoluta certezza. Türeci et al. 2016 (*Clinic. Cancer Res.* 22:1885-1896) ha suggerito di utilizzare un poliepitopo per aumentare la probabilità che sia incluso un epitopo adatto (vedere pagina 1890, 2a col., 1° paragrafo di Türeci et al. 2016). Inoltre, il documento dello stato della tecnica Castle et al. 2014 (*Scientific Reports* 4: 1-6) riguarda la questione generale di stabilire se vi sia o meno una correlazione tra la frequenza dell'allele di mutazione del DNA e la frequenza dell'allele di mutazione dell'RNA. Infine, il documento dello stato della tecnica WO2012/159754 descrive un'analisi della variazione del numero di copie utilizzata per identificare possibili neoepitopi. Nessuno di questi documenti dello stato della tecnica, tuttavia, insegna a determinare l'idoneità di un neoepitopo per prendere di mira una cellula malata in base al numero di copie degli alleli che codificano il neoepitopo.

Nella tecnica c'è bisogno di modi in cui selezionare i neoepitopi risultanti da mutazioni specifiche della malattia che si traducono in un maggiore controllo del tumore.

DESCRIZIONE

RIASSUNTO

Il presente insegnamento fornisce modi per superare le carenze nello stato della tecnica fornendo metodi per determinare l'idoneità di un neoepitopo

risultante da una mutazione specifica della malattia in un gene come un target specifico della malattia da cui un tessuto malato non può facilmente sfuggire all'immunosorveglianza, che in caso di cancro si tradurrà in un maggiore controllo del tumore. Una volta identificati i neoepitopi adatti, tali epitopi adatti possono essere utilizzati come target specifici della malattia per indurre una risposta immunitaria specifica in un paziente affetto dalla malattia. Ad esempio, la malattia può essere il cancro e potenzialmente il tumore primario così come le metastasi tumorali che esprimono il neoepitopo adatto possono essere prese di mira per un trattamento più efficace.

Il presente insegnamento si riferisce a un metodo per determinare l'idoneità di un neoepitopo risultante da una mutazione specifica della malattia in un allele in un gene (allele mutato) come un target specifico della malattia comprendente determinare, in una cellula malata o in una popolazione di cellule malate, il numero di copie dell'allele mutato che codifica il neoepitopo. Come qui utilizzato, il numero di copie può anche essere indicato come zigosità in modo tale che, ad esempio, dove il numero di copie dell'allele mutato è 4, l'allele mutato ha una zigosità di 4. Come qui utilizzato, un allele è un sito nel genoma avente una specifica identità nucleotidica, la cui identità può essere la stessa sia sulle copie materne sia paterne del genoma (genotipo omozigote) oppure l'identità può essere diversa sulle copie materne e paterne del genoma (genotipo eterozigote). Un allele mutato è un allele che, a causa di una mutazione specifica della malattia, ha un'identità diversa da quel sito in un genoma normale corrispondente, ad esempio un genoma da una cellula non malata dello stesso individuo (genoma corrispondente), preferibilmente da una cellula non malata dello stesso tipo di tessuto della cellula malata. Un neoepitopo adatto come un target specifico della malattia (neoepitopo adatto) come qui utilizzato è un neoepitopo che, quando preso di mira dal sistema immunitario, ha meno probabilità di avere la sua espressione ridotta o silenziata (ad esempio a causa dell'eliminazione) dal tessuto malato tale per cui è meno probabile che il tessuto malato sia in grado di sfuggire a una risposta, preferibilmente una risposta immunologica generata contro il neoepitopo mediante, ad esempio, la vaccinazione contro il neoepitopo o la somministrazione di cellule immunitarie che sono in grado di prendere di mira (legare) il neoepitopo. In una forma di realizzazione, il numero di copie dell'allele mutato può essere lo stesso del numero di copie del gene che comprende l'allele mutato in modo tale che il presente insegnamento si riferisca anche a un metodo per determinare l'idoneità di un neoepitopo risultante da una mutazione specifica della malattia in un gene come un target specifico della malattia comprendente determinare, in una cellula malata o in una popolazione di cellule malate, il numero di copie del gene avente la mutazione specifica della malattia.

In una forma di realizzazione, un numero elevato di copie dell'allele mutato o del gene avente la mutazione specifica della malattia indica l'idoneità del neoepitopo come un target specifico della malattia, in modo tale che maggiore è il numero di copie dell'allele mutato o del gene avente la mutazione specifica della malattia, maggiore è l'idoneità del neoepitopo come un target specifico della malattia. In una forma di realizzazione, dove

il numero di copie dell'allele mutato o del gene avente la mutazione specifica della malattia nella cellula malata è maggiore di 2, ciò indica l'idoneità del neoepitopo come un target specifico della malattia. In una forma di realizzazione, dove il numero di copie del gene avente la mutazione specifica della malattia è maggiore di 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90 o è maggiore di 100, ciò indica l'idoneità del neoepitopo come target specifico della malattia.

Nei casi in cui non tutte le copie del gene hanno la mutazione, in cui almeno una copia ha l'allele mutato, è preferibile che molte copie del gene abbiano l'allele mutato, cioè è preferibile che una frazione maggiore piuttosto che inferiore delle copie del gene abbia l'allele mutato (zigosità frazionata più alta anziché più bassa). Pertanto, in alcune forme di realizzazione, l'allele mutato si trova in un'elevata frazione di copie del gene di cui almeno una copia ha l'allele mutato (zigosità frazionata), dove la zigosità frazionata è il rapporto tra il numero di copie dell'allele mutato (zigosità dell'allele mutato) sul numero totale di copie del sito nucleotidico a cui si mappa l'allele mutato, in particolare a un genoma di riferimento o un genoma wild-type corrispondente o un genoma abbinato, cioè, genoma wild-type dallo stesso individuo. Maggiore è la zigosità frazionata delle copie dell'allele mutato, maggiore è l'idoneità del neoepitopo come target specifico della malattia. Preferibilmente, la zigosità frazionata può essere maggiore di 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 0,95 e più preferibilmente la zigosità frazionata è 1, cioè tutte le copie del gene nella cellula malata hanno l'allele mutato. Nei casi in cui la zigosità frazionata è 1, non ci sono copie wild-type del gene in modo tale che la cellula malata non possa tornare a esprimere il corrispondente epitopo wild-type. Come qui utilizzato, una frazione di 1 è dove l'ipotesi che la configurazione genetica dell'allele/gene mutato, ad esempio il numero di copie, zigosità, sia la stessa che non può essere confutata dai dati, cioè è statisticamente consistente.

È noto che il tessuto malato, come i tumori, può essere eterogeneo nella composizione genetica e nell'espressione genica in modo tale che sia possibile che non tutte le cellule malate nel tessuto malato abbiano lo stesso numero di copie di un gene e/o del gene di cui almeno una copia ha l'allele mutato (numero totale di copie del gene) e/o copie del gene aventi l'allele mutato, tanto meno la mutazione stessa specifica della malattia. Pertanto, è preferibile che il numero di copie, ad esempio dell'allele mutato e/o della zigosità frazionata e/o il numero totale di copie del sito nucleotidico a cui si mappa l'allele mutato sia lo stesso o simile in un'elevata frazione di cellule malate piuttosto che una bassa frazione di cellule malate nel tessuto malato (una frazione clonale elevata piuttosto che bassa). Maggiore è la frazione di cellule malate aventi lo stesso o un numero simile di copie, ad esempio dell'allele mutato e/o la zigosità frazionata e/o il numero totale di copie del sito nucleotidico a cui si mappa l'allele mutato, maggiore è l'idoneità del neoepitopo come target specifico della malattia. Ad esempio, la frazione clonale può essere almeno 0,3, 0,4, 0,5, 0,6, 0,7, 0,8 o almeno 0,9. In una forma di realizzazione preferita, tutte le cellule malate nel tessuto malato hanno numeri di copie uguali o simili,

cioè la frazione clonale è 1, cioè è statisticamente consistente. Come qui utilizzato, lo stesso numero di copie o simile comprende lo stesso numero di copie o un numero di copie entro il 30%, 25%, 20%, 15%, 10%, 5%, 4%, 3%, 2% o meno del numero di copie, ad esempio numero di copie o numero assoluto di copie con o senza correzione degli errori.

Preferibilmente, una frazione clonale di una mutazione può essere data dalla frazione di cellule malate che hanno la stessa o una simile configurazione genetica della mutazione, in cui una configurazione genetica della mutazione comprende il numero totale di copie del sito nucleotidico a cui si mappa la mutazione e il numero di copie dell'allele mutato. Si dice che una caratteristica è fissa nella popolazione di cellule malate se la caratteristica è presente in tutte le cellule malate a un livello che non può essere confutato statisticamente dai dati disponibili. Preferibilmente, una frazione clonale di 1 significa che la configurazione genetica della mutazione è fissa nella popolazione di cellule malate. Preferibilmente, la configurazione genetica di una mutazione è fissa nella popolazione di cellule malate se la mutazione è fissa nella popolazione di cellule malate e la CNV che interessa il sito codificante la mutazione è fissa nella popolazione di cellule malate. Preferibilmente, se si determina che una mutazione per cui il numero totale di copie del sito nucleotidico a cui si mappa la mutazione è 2 e si trova in una regione bilanciata del genoma malato (tumore) sia fissa nella popolazione di cellule malate, allora la configurazione genetica della mutazione è fissa.

Il gene in cui si trova la mutazione specifica della malattia può trovarsi potenzialmente in qualsiasi gene del genoma. Un tipo preferito di gene in cui si trova una mutazione che determina un neopitopo adatto è un gene la cui espressione determina la trasformazione della cellula in un fenotipo canceroso o la cui mancanza di espressione fa sì che una cellula cancerosa perda il suo fenotipo canceroso, cioè un gene la cui espressione contribuisce alla progressione del tumore. Tali geni sono noti come geni driver. Sono ben noti esempi di geni driver per molti tipi di tumori. Ad esempio, un elenco di 291 geni driver del cancro ad alta affidabilità che agiscono su 3.205 tumori di 12 diversi tipi di cancro è riportato in Tamborero et al., 2013, Comprehensive identification of mutational cancer driver genes across 12 tumor types, Scientific Reports 3:2650. Ulteriori geni driver sono stati identificati utilizzando i metodi descritti in Youn et al., 2011, Identifying cancer driver genes in tumor genome sequencing studies, Bioinformatics 27(2): 175-181, in Sakoparnig et al., 2015, Identification of constrained cancer driver genes based on mutation timing, PLoS Comput. Biol. 11(1):e1004027 e in Forbes et al., 2008, Current protocols in human genetics 10-11. La mutazione specifica della malattia nel gene driver può o meno contribuire al fenotipo canceroso. Preferibilmente, ogni copia del gene driver trovata nella cellula malata ha la mutazione specifica della malattia. Inoltre, preferibilmente, tutte le cellule nel tessuto malato sono cellule malate in cui ogni copia del gene driver ha la mutazione specifica della malattia.

Un altro tipo preferito di gene è un gene essenziale. In una forma di realizzazione, un gene essenziale è un gene che, quando viene silenziato o la

sua espressione viene ridotta (ad esempio venendo eliminato), risulta almeno in una crescita compromessa o in una ridotta idoneità della cellula, preferibilmente in una cellula malata. Tali geni sono qui chiamati geni essenziali. In una forma di realizzazione, un gene essenziale è un gene in cui è presente una riduzione di almeno il 10% nella crescita o una ridotta idoneità della cellula malata dove il gene è silenziato o ha la sua espressione ridotta rispetto a una cellula in cui il gene non è silenziato né è a espressione ridotta. In una forma di realizzazione, la riduzione della crescita o dell'idoneità ridotta è almeno del 20%, 30%, 40%, 50%, 60%, 75%, 80%, 90% o almeno del 95%, più preferibilmente il silenziamento o la riduzione dell'espressione del gene essenziale determina la letalità della cellula malata. Preferibilmente, ogni copia del gene essenziale trovata nella cellula malata ha la mutazione specifica della malattia.

I geni essenziali sono ben noti nella tecnica, ad esempio, un elenco di geni essenziali nell'uomo (ad esempio, nelle linee cellulari umane o dedotto da altri organismi) è descritto in Liao et al., 2008, Proc. Nat. Acad. Sci. USA 105: 6987-6992 e in Georgi et al., 2013, PLoS Genetics 9(5):e1003484, nonché corrispondenti orologi in altri organismi eucarioti come un topo (Liao et al., 2007, Trends Genet. 23:378-381), moscerino della frutta (Spradling et al., 1999, Genetics 153:135-177), *C. elegans* (Kamath et al., 2003, Nature 421:231-237), pesce zebra (Amsterdam et al., 2004, Proc. Natl. Acad. Sci. USA 101:12792-12797), *Arabidopsis thaliana* (Tzafrir et al., 2004, Plant Physiol. 135:1206-1220), lievito (Kim et al., 2010, Nat. Biotechnol. 28:617-623) e così via. Un elenco di geni essenziali derivati da linee cellulari tumorali umane è descritto in Wang et al., 2015, Science 350:1096-1101, e un elenco di geni essenziali può essere trovato nel database dei geni essenziali, DEG5.0 (Zhang et al., 2009, Nucleic Acids Res. 37:D455-D458).

Inoltre, un elenco di geni essenziali la cui eliminazione/silenziamento riduce significativamente l'idoneità di una coorte di linee cellulari può essere generato empiricamente da più tessuti sani e/o da linee cellulari cancerose, quali linee cellulari possono essere derivate da donatori o dal paziente. L'eliminazione/silenziamento dei geni può essere eseguita in modo sperimentale utilizzando varie tecniche di biologia molecolare come la tecnologia CRISPR, l'interferenza dell'RNA e così via, dove viene determinata la sopravvivenza o l'idoneità della cellula con e senza l'espressione del gene essenziale putativo. Un elenco di geni essenziali può anche essere determinato in modo sperimentale da cellule o da una linea cellulare o un elenco di geni essenziali può essere ottenuto mediante approcci bioinformatici. Le cellule o linee cellulari possono essere cellule o linee cellulari malate (cellule tumorali o linee cellulari) o cellule o linee cellulari non malate (sane/normali) e possono essere ottenute da donatori o dal paziente affetto dalla malattia. Preferibilmente, le cellule o le linee cellulari non malate provengono dallo stesso tipo di tessuto della cellula malata e più preferibilmente dallo stesso paziente. Nelle forme di realizzazione in cui la malattia è il cancro, le cellule o le linee cellulari possono essere ottenute dal tumore primario o da qualsiasi metastasi, se presente. Inoltre, un elenco di geni essenziali può essere essenzialmente lo stesso dell'insieme

minimo di geni espressi in un'ampia varietà di tessuti del corpo. Ad esempio, un gene essenziale è un gene che è espresso in un'ampia varietà di tessuti diversi ed è espresso con una soglia RPKM (Minimum reads per Kilobase of transcript Per million mapped Reads) maggiore di 0, preferibilmente maggiore di 0,1, 0,5, 1, 2, 3, 4, 5, 10, 20, 25. Tale elenco di geni essenziali può essere ottenuto analizzando i dati di espressione dell'RNA, (ad esempio RNAseq) ottenuti da un pannello di campioni cellulari ottenuti da almeno 5, 6, 7, 8, 9, 10, 15, 20, 25 o tessuti più diversi. Inoltre, se è disponibile una linea cellulare tumorale del paziente, i geni in cui tutte le copie contengono una mutazione che codifica un neoepitopo possono essere eliminati uno alla volta e possono essere misurati i tassi di crescita di ciascuna linea cellulare modificata. Tale misurazione/analisi può essere eseguita mediante metodi ad alta produttività noti nella tecnica, che consentono lo screening di almeno un gene alla volta, preferibilmente più geni alla volta, al fine di valutarne l'effetto sull'idoneità di un malato o di una cellula non malata. Tali metodi consentono anche la rilevazione di combinazioni di geni sintetiche malate o letali, discusse di seguito. In breve, una libreria di linee cellulari, ciascuna linea cellulare priva di un gene, può essere utilizzata per testare l'eliminazione di uno o più geni candidati, in modo tale da determinare l'effetto sulla cellula dell'eliminazione dei geni.

Il presente insegnamento riguarda inoltre un metodo per determinare l'idoneità di un neoepitopo risultante da una mutazione specifica della malattia in un gene come target specifico della malattia comprendente determinare, in una cellula malata o in una popolazione di cellule malate, il numero di copie del gene, cioè, determinare il numero di copie di un gene in cui almeno una copia del gene ha una mutazione specifica della malattia. Nei casi in cui un gene ha un numero elevato di copie in una cellula malata come una cellula tumorale, ad esempio a causa dell'amplificazione focale, è molto probabile che il gene possa essere un gene driver. Pertanto, un numero elevato di copie di un gene indica l'idoneità del neoepitopo come un target specifico della malattia e maggiore è il numero di copie del gene, maggiore è l'idoneità del neoepitopo come target specifico della malattia. Ad esempio, un numero di copie elevato può essere un numero di copie maggiore di 2, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 o maggiore di 100. Un numero di copie elevato può anche essere un numero di copie maggiore almeno del 50% rispetto al numero di copie del gene in una cellula non malata corrispondente. Un numero di copie elevato può essere anche quando il numero di copie del gene in cui almeno una copia ha la mutazione specifica della malattia è almeno 2x, 3x, 5x, 10x, 15x, 20x, 25x, 30x, 40x, 50x, 60x, 70x, 80x, 90x o almeno 100x maggiore del numero di copie del gene in una cellula non malata corrispondente. A causa delle variazioni del numero di copie che possono essere presenti anche nel genoma normale, il numero di copie del gene nel genoma normale non è necessariamente due. Inoltre, è noto che le amplificazioni focali sono osservate più spesso in alcune malattie rispetto ad altre, come nel glioblastoma dove il gene del recettore del fattore di crescita epidermico è spesso amplificato localmente, quindi questa forma di realizzazione è ben adatta all'utilizzo in quelle malattie.

Inoltre, è preferibile che il numero di copie del gene sia lo stesso o simile in un'elevata frazione di cellule malate piuttosto che in una bassa frazione di cellule malate, in modo tale che maggiore è la frazione di cellule malate aventi lo stesso o simile numero di copie, maggiore è l'idoneità del neoepitopo come un target specifico della malattia. In una forma di realizzazione preferita, tutte le cellule malate nel tessuto malato hanno lo stesso o un numero di copie simile del gene in cui almeno una copia ha la mutazione specifica della malattia, cioè la frazione clonale è 1. Come qui utilizzato, lo stesso numero di copie o simile comprende lo stesso numero di copie o un numero di copie entro il 30%, 25%, 20%, 15%, 10%, 5%, 4%, 3%, 2% o meno del numero di copie, ad esempio numero di copie o numero assoluto di copie con o senza correzione degli errori.

Inoltre, il gene con un numero elevato di copie, in cui almeno una copia ha una mutazione specifica della malattia risultante in un neoepitopo, può preferibilmente essere un gene la cui espressione determina la trasformazione della cellula in un fenotipo canceroso o la cui mancata espressione risulta in una cellula cancerosa che perde il suo fenotipo canceroso, cioè un gene driver come quei geni driver noti nella tecnica o può essere un gene essenziale, ad esempio un gene che quando viene silenziato o la sua espressione viene ridotta, almeno risulta in una crescita ridotta o in una ridotta idoneità della cellula malata.

Il presente insegnamento riguarda anche un metodo per determinare l'idoneità di un neoepitopo risultante da una mutazione specifica della malattia in un gene come un target specifico della malattia comprendente determinare, in una cellula malata o in una popolazione di cellule malate, se il gene che ha la mutazione specifica della malattia è un gene essenziale. In una forma di realizzazione, il gene essenziale è un gene che quando viene silenziato o la sua espressione viene ridotta (ad esempio, mediante eliminazione del gene), almeno risulta in una crescita ridotta o in una ridotta idoneità della cellula malata. In questa forma di realizzazione, dove il gene è un gene essenziale e tutte le copie del gene essenziale hanno la mutazione specifica della malattia (zigosità frazionata di 1) indica l'idoneità del neoepitopo come un target specifico della malattia preferibile. In una forma di realizzazione, un gene essenziale è un gene che è espresso in un'ampia varietà di tessuti diversi ed è espresso con una soglia RPKM (minimum reads per kilobase of transcript per million mapped reads) maggiore di 0, preferibilmente maggiore di 0,1, 0,5, 1, 2, 3, 4, 5, 10, 20, 25. Preferibilmente, tutte le copie del gene essenziale contengono la mutazione. Inoltre, è preferibile che una frazione elevata di cellule malate contenga copie del gene essenziale in cui tutte le copie del gene essenziale hanno la mutazione specifica della malattia (una frazione clonale elevata anziché bassa), in modo tale che maggiore è la frazione di cellule malate contenenti copie del gene essenziale in cui tutte le copie del gene essenziale hanno la mutazione specifica della malattia, maggiore è l'idoneità del neoepitopo come target specifico della malattia. In una forma di realizzazione maggiormente preferita, tutte le cellule malate nel tessuto malato hanno il gene essenziale in cui tutte le copie del gene essenziale hanno la mutazione specifica della malattia, cioè la frazione clonale è 1.

È noto che alcuni geni, quando vengono silenziati individualmente o dove la loro espressione è ridotta individualmente, possono avere solo un piccolo effetto, se presente, sull'idoneità o sulla capacità di crescita della cellula malata. Tuttavia, è stato osservato che quando due di tali geni sono stati entrambi silenziati o dove ciascuna loro espressione è stata ridotta, può risultare in una compromissione della crescita molto più forte, fino alla letalità. Tali combinazioni genetiche sono indicate come sintetiche letali o sintetiche malate/compromesse. Vedere Nijman, 2011, Synthetic lethality: General principles, utility and detection using genetic screens in human cells, FEBS Lett. 585:1-6 per una discussione sui geni malati sintetici e sintetiche letali e sui metodi per identificare tali geni. Poiché entrambi i geni sono necessari per la sopravvivenza della cellula, è improbabile che la cellula silenzi o riduca l'espressione di entrambi i geni. Pertanto, una combinazione adatta di neoepitopi come target specifici della malattia può derivare da mutazioni specifiche della malattia in almeno due geni, i cui geni insieme sono sinteticamente letali o sinteticamente malati. In considerazione di ciò, il presente insegnamento si riferisce inoltre a un metodo per determinare l'idoneità di una combinazione di almeno due neoepitopi risultanti da mutazioni specifiche della malattia in almeno due geni come una combinazione di target specifici della malattia comprendente determinare se una combinazione di almeno due geni aventi ciascuno una mutazione specifica della malattia sono geni sintetiche letali o sintetiche malati. Quando la combinazione di almeno due geni risulta in un fenotipo sintetico letale o malato sintetico, ciò indica che i neoepitopi risultanti sono una combinazione adatta di target specifici della malattia. In una forma di realizzazione preferita, la malattia sintetica determina almeno un effetto maggiore sulla crescita/idoneità cellulare rispetto a quanto ci si aspetterebbe dall'effetto additivo di eliminazione/espressione ridotta di ciascun gene in modo individuale. Questo approccio è favorito dove è presente un numero elevato di neoepitopi adatti poiché maggiore è il numero di neoepitopi, maggiore è il numero di combinazioni che potrebbero essere sintetiche malate o letali. Ad esempio, 10 mutazioni corrispondono a 45 possibili combinazioni, 100 mutazioni corrispondono a 4950 combinazioni e 1000 mutazioni corrispondono a circa 500.000 combinazioni. In alcune forme di realizzazione, almeno due geni hanno ciascuno una zigosità frazionata maggiore anziché inferiore, preferibilmente una zigosità frazionata di 1 e/o ciascuno ha una frazione clonale superiore anziché inferiore, preferibilmente una frazione clonale di 1, entrambi nelle cellule malate e nelle cellule malate nel tessuto malato. Come qui utilizzato, ciascun neoepitopo trovato in una combinazione di neoepitopi adatti è considerato un neoepitopo adatto ai fini del presente insegnamento.

Come indicato nel presente documento, il numero di copie del gene, nella cellula malata o non malata, può essere il numero di copie relativo, ma è preferibilmente il numero assoluto di copie e più preferibilmente è il numero assoluto di copie normalizzato contro una ploidia, ad esempio, la ploidia del genoma della cellula malata, cioè il numero di copie del genoma. Ancor più preferibilmente, il numero di copie relativo, assoluto e normalizzato viene corretto per errore.

Quando si stima il numero assoluto di copie, ad esempio di un allele o di un gene mutato o della sua zigosità, la stima può essere imprecisa ed è opportuno correggere il numero assoluto di copie per tenere conto delle fonti di errore. Nelle forme di realizzazione dove viene utilizzato il sequenziamento di nuova generazione per ottenere informazioni sulla sequenza da genomi ed esomi, le fonti di errore possono includere: una distorsione nella purezza stimata del campione di tessuto malato come un campione di tumore, una distorsione in qualsiasi parametro stimato richiesto per far derivare la purezza e/o i numeri assoluti di copie, errori stocastici dovuti alla copertura finita del campione da sequenziare, capacità di rilevamento limitata a causa di una bassa purezza, una bassa frazione clonale e così via.

In una forma di realizzazione dove il numero di copie assolute viene determinato utilizzando segmenti eterozigoti bilanciati contenenti un SNP eterozigote, come quello descritto nella domanda di brevetto internazionale PCT intitolata "Tumor Modeling Based on Primary Balanced Heterozygous Segments" depositata in data pari alla presente, un errore nel numero assoluto di copie di un segmento può propagarsi ad altri parametri stimati, come il numero assoluto di copie dell'allele mutato, ad esempio un SNV, che codifica un neoepitopo, la zigosità dell'allele mutato, la frazione clonale e così via. Poiché tali parametri stimati a valle hanno implicazioni cliniche per determinare l'idoneità di un neoepitopo come target specifico della malattia per un paziente come descritto nel presente documento, è desiderabile correggere gli errori nel numero assoluto di copie per ottenere il valore più accurato del numero assoluto di copie. Inoltre, poiché le mutazioni che codificano i neoepitopi possono essere prioritarie per la loro idoneità per poter essere incluse in un vaccino da somministrare al paziente utilizzando i criteri qui discussi, e poiché una frazione di zigosità di 1 è qualitativamente migliore di una frazione di zigosità inferiore a 1, in particolare quando il gene che contiene la mutazione è un gene essenziale, è vantaggioso avere le stime più accurate dei numeri assoluti di copie e di qualsiasi parametro da essi derivato.

In una forma di realizzazione preferita, il numero assoluto di copie di un allele mutato, ad esempio un SNV, e/o la zigosità di un SNV può essere corretto per errore. In una forma di realizzazione, il numero assoluto di copie di un SNV viene prima corretto per errore, e quindi la zigosità dell'SNV viene corretta per riflettere il numero assoluto di copie corretto per errore dell'SNV. Una volta che il numero assoluto di copie dell'SNV e/o la zigosità dell'SNV sono corretti per errore, la stima della frazione clonale può essere corretta anche per riflettere i numeri assoluti corretti per errore, che include determinare se la frazione clonale è statisticamente coerente con un valore di 1.

Il numero assoluto di copie di un SNV è preferibilmente dato dal numero assoluto di copie del segmento a cui corrisponde l'SNV. I numeri assoluti di copie di tutti i segmenti, ad esempio, nel genoma del tumore malato possono essere corretti per errore, incluso il numero assoluto di copie dell'SNV. Una volta che i numeri assoluti di copie di tutti i segmenti in un genoma sono corretti per errore, è possibile calcolare una ploidia corretta per errore in base ai numeri assoluti di copie corretti per errore di segmenti, ad esempio nel genoma del tumore malato.

In una forma di realizzazione specifica, se il numero assoluto di copie di un SNV viene corretto per errore in modo tale che il nuovo numero assoluto di copie differisca dal numero assoluto di copie originale, ciò può essere considerato un'indicazione che il numero assoluto di copie stimato dell'SNV non è affidabile.

Un esempio di correzione degli errori dei numeri assoluti di copie dei segmenti è la correzione dell'errore di parità, che comprende la correzione di un numero assoluto di copie dispari di un segmento in un numero assoluto di copie pari se il segmento si trova in una regione bilanciata. Una regione bilanciata è una regione, ad esempio, del genoma del tumore malato in cui gli alleli materni e paterni all'interno della regione hanno subito un'amplificazione uguale (bilanciata) o entrambi gli alleli materni e paterni non hanno subito alcuna amplificazione.

La decisione di correggere per errore il numero assoluto di copie dispari del segmento al numero assoluto di copie pari maggiore più vicino o al numero assoluto di copie pari inferiore più vicino può dipendere dalle letture della malattia e dalle letture normali di mappatura del segmento e dal confronto con i limiti previsti che definiscono il numero assoluto di copie di un segmento. In cui una lettura normale è una lettura relativa al campione normale sequenziato e una lettura della malattia è una lettura relativa al campione malato sequenziato. In particolare, quando la malattia è il cancro, una lettura del tumore è una lettura relativa al campione di tumore sequenziato.

Ad esempio, in una correzione degli errori di parità del primo tipo, se CN_{mut} è il numero assoluto di copie nel genoma malato del segmento a cui si mappa la mutazione, il numero assoluto di copie di un segmento previsto per avere un valore di CN_{mut} , può essere corretto in $CN_{mut}+1$ se $r > \rho_{th}$ e in $CN_{mut}-1$ se $r < \rho_{th}$, dove r è la malattia rispetto al normale rapporto di conta della lettura del segmento (il rapporto del numero della malattia, ad esempio letture del tumore di mappatura del segmento rispetto al numero di letture normali di mappatura del segmento), in cui ρ_{th} è il limite di decisione previsto, il cui valore dipende anche dalla purezza del campione di tessuto della malattia, ad esempio campione del tumore.

Un numero di copie specifico dell'allele di un segmento è il numero di copie nel genoma malato dell'allele materno o dell'allele paterno del segmento. Quando il segmento contiene un SNP eterozigote (segmento eterozigote), l'SNP eterozigote può essere utilizzato per determinare il numero di copie specifico dell'allele del segmento. Un segmento eterozigote può essere assegnato a un nodo preferito, in cui un nodo può essere definito come una combinazione unica di un numero assoluto di copie del segmento eterozigote e un numero di copie specifico dell'allele del segmento eterozigote. I nodi pari sono un sottoinsieme di nodi per cui il numero assoluto di copie del segmento è pari. Se un segmento eterozigote contiene più di un SNP eterozigote, il gruppo di due o più SNP eterozigoti può essere rappresentato da un singolo elemento del gruppo, oppure le frequenze alleliche di tutti gli elementi del gruppo possono essere mediate, oppure può essere presa una media, a condizione che la frequenza allelica di ciascun SNP eterozigote sia calcolata in modo coerente per l'allele che ha il numero maggiore o minore di copie nel genoma malato.

Una correzione degli errori di parità del primo tipo di un segmento eterozigote può comportare la ricerca del nodo pari più probabile che corrisponda al segmento eterozigote, ad esempio, sulla base di un quadro di massima verosimiglianza date le letture misurate della malattia (ad esempio tumore) e delle letture normali di mappatura del segmento.

In una correzione degli errori di parità del secondo tipo, vengono identificati i segmenti a monte e a valle più vicini che non richiedono correzione degli errori di parità, preferibilmente entro 10 Mb, 5 Mb, 1 Mb del segmento contenente l'SNV. Se il numero assoluto di copie di entrambi i segmenti a monte e a valle più vicini sono identici, il numero assoluto di copie del segmento contenente l'SNV viene modificato nel numero assoluto di copie del segmento più vicino. Generalmente, la correzione degli errori di parità del secondo tipo è preferita alla correzione degli errori di parità del primo tipo, a meno che non possa essere implementata poiché non possono essere identificati opportuni segmenti adiacenti, nel qual caso può essere applicata la correzione degli errori di parità del primo tipo.

Al posto o in aggiunta alla correzione degli errori di parità possono essere introdotte anche altre forme di correzione degli errori di un numero assoluto di copie di un segmento. Ad esempio, metodi che considerano il numero assoluto di copie dei segmenti nelle immediate vicinanze del gene contenente la mutazione nel genoma malato, in cui il numero assoluto di copie del segmento contenente l'SNV viene modificato nella modalità dei numeri assoluti di copie dei segmenti vicini, preferibilmente se il cambio nel numero assoluto di copie non è maggiore di 3, 2 e preferibilmente di 1, e preferibilmente se la maggior parte dei segmenti vicini (50%, 60%, 70%, 80%, 90%, 100%) hanno un numero assoluto di copie uguale alla modalità.

È possibile combinare diversi schemi di correzione degli errori per i numeri assoluti di copie. In una forma di realizzazione preferita, la prima correzione degli errori di parità viene applicata come primo strato di correzione degli errori, al di sopra del quale possono essere applicati ulteriori metodi di correzione degli errori.

Come qui utilizzato, un segmento può essere una regione predeterminata del genoma, ad esempio, predeterminata in base a un genoma di riferimento. Un segmento può estendersi su un gene, ad esempio, come definito in un genoma di riferimento a cui sono allineate le letture. Un segmento può anche essere un frammento di un gene, un esone, un'unione di esoni o l'unione di esoni associati all'interno di un dato gene. Un segmento può anche essere un altro insieme di regioni predeterminate in un genoma di riferimento (con o senza introni), o in un altro insieme di regioni predeterminate in un genoma di riferimento basato sul genoma normale. In forme di realizzazione specifiche, un segmento può essere una regione del genoma di riferimento con un dato numero costante di copie e/o un dato numero di copie specifico dell'allele, ad esempio nel genoma, tumore malato o in alternativa un frammento di un gene con un dato numero costante di copie e/o un numero di copie specifico dell'allele, ad

esempio nel genoma del tumore malato. Un segmento può essere definito per includere o escludere gli introni.

Un numero di copie di un segmento in un dato genoma (ad esempio nel genoma normale o nel genoma del tumore) può essere definito come la frequenza totale della sequenza nucleotidica del segmento nel genoma, ignorando le variazioni causate dai SNP e/o SNV e/o altri cambiamenti associati al cancro come, ma non solo, mutazioni, inserimenti, eliminazioni e/o altre varianti genetiche correlate al cancro. Preferibilmente le diverse copie del segmento nel dato genoma hanno la stessa lunghezza o quasi la stessa lunghezza.

Il numero di copie di un segmento in un genoma può significare il numero di copie fisiche del segmento in una cellula contenente il genoma. Un numero assoluto di copie di un segmento nel genoma normale può essere definito come il numero di copie fisiche di un dato segmento in una cellula sana. Un numero assoluto di copie di un segmento, ad esempio nel genoma, tumore malato può essere definito come il numero di copie fisiche di un dato segmento, ad esempio in un tumore, cellula malati. Il numero di copie di un segmento nel genoma può essere indicato come il numero assoluto di copie del segmento in detto genoma. Un numero di copie può significare un numero assoluto di copie.

In una forma di realizzazione specifica, se solo una parte di un segmento viene amplificata o eliminata in un genoma, allora tale copia parziale del segmento può essere contata come una copia del segmento o non contata come una copia del segmento. In una forma di realizzazione preferita, le copie del segmento che si estendono per meno del 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10% o 5% della lunghezza del segmento possono essere ignorate.

Un genoma di riferimento viene utilizzato per mappare le letture e fornire un sistema di coordinate per il genoma normale e, ad esempio per il genoma, tumore malato in cui un sistema di coordinate può comprendere fornire un numero di cromosomi, una posizione del nucleotide nel cromosoma, nonché la direzionalità della lettura, dove la posizione nel cromosoma è indicata dalla linea.

Un genoma di riferimento può essere basato sul genoma di uno o più elementi della stessa specie di un soggetto che fornisce il campione di tessuto malato, oppure può essere basato sul genoma normale del soggetto.

Un campione di tessuto malato, come un campione di tumore, può anche comprendere contaminazione dal genoma normale, in particolare il genoma normale dello stesso paziente da cui è stato prelevato il campione, e/o in caso di eterogeneità intratumorale, comprendere anche più di un genoma del tumore. La purezza, la purezza del campione del tumore, la purezza del tumore e la purezza del campione sono tutti considerati termini equivalenti, indicando preferibilmente la frazione di cellule del tumore presenti in un campione del tumore. Per contaminazione normale si intende preferibilmente la frazione di cellule normali presenti nel campione tumorale e può essere data da uno meno la purezza.

Normalizzando contro la ploidia dei controlli cellulari per la presenza di copie di un gene dovute a eventi di duplicazione del genoma. In una

forma di realizzazione, il numero assoluto di copie può essere normalizzato rispetto alla ploidia del genoma, la cui ploidia è la media del numero assoluto di copie di tutti i segmenti in un dato genoma, è una data cellula ponderata per la lunghezza di ciascun segmento. In una forma di realizzazione, il numero assoluto di copie può essere normalizzato rispetto alla ploidia del cromosoma che contiene il gene di interesse mutato (che comprende la mutazione), la cui ploidia è la media del numero assoluto di copie di tutti i segmenti su un dato cromosoma in una data cellula ponderata per la lunghezza di ciascun segmento sul cromosoma. In una forma di realizzazione, il numero assoluto di copie può essere normalizzato rispetto alla ploidia di una regione vicina del cromosoma che contiene il gene di interesse mutato, la cui ploidia è la media di ciascun segmento in una data regione in una data cellula ponderata per la lunghezza di ogni segmento della regione. La regione vicina può trovarsi entro una distanza predeterminata dal gene che ha la mutazione specifica della malattia, ad esempio entro 100 megabasi (Mb), 75 Mb, 50 Mb, 25 Mb, 10 Mb, 5 Mb, 4 Mb, 3 Mb, 2 Mb o 1 Mb del gene che ha la mutazione specifica della malattia. Il numero di copie di un segmento può essere calcolato regolarmente mediante metodi noti nella tecnica, sia in modo sperimentale sia in modo computazionale. Ad esempio, i brevetti EP n. 2 198 292 B1 e EP 2 002 016 B1 descrivono metodi per determinare, rispettivamente, il numero di copie relativo e la frequenza del numero di copie delle sequenze di acido nucleico. Inoltre, la pubblicazione della domanda di brevetto EP n. 2 835 752 A e la pubblicazione della domanda di brevetto internazionale n. WO 2014/014497 e WO 2014/138153 descrivono anche metodi per determinare le variazioni del numero di copie. Vedere anche, Machado et al., 2013, Copy Number Variation of Fc Gamma Receptor Genes in HIV-Infected and HIV-Tuberculosis Co-Infected Individuals in Sub-Saharan Africa, PLoS, 8(11):e78165. Altri metodi includono l'utilizzo di FACS, FISH o altri metodi basati sulla fluorescenza, cariotipo spettrale (SKY) e PCR digitale. Un segmento può anche essere un gene.

La mutazione specifica della malattia può essere qualsiasi mutazione che si traduca nell'espressione di un neoepitopo, preferibilmente sulla superficie della cellula malata. In particolare, la mutazione può essere un evento indel o di fusione genica oppure può essere una variazione di un singolo nucleotide (mutazione puntiforme). Preferibilmente, le mutazioni specifiche della malattia sono mutazioni non sinonime, preferibilmente mutazioni non sinonime di proteine espresse in un tumore o in una cellula cancerosa. È possibile utilizzare qualsiasi metodo noto nella tecnica per determinare mutazioni specifiche della malattia, e in particolare metodi che utilizzano dati di sequenziamento di nuova generazione per determinare eventuali cambi tra il genoma/esoma delle cellule malate rispetto al genoma/esoma delle corrispondenti cellule wild-type non malate. Ad esempio, Carter et al., 2012, Absolute quantification of somatic DNA alterations in human cancer, Nature Biotechnology 30:413-421; Cibulskis et al., 2013, Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples, Nature Biotechnology 31:213-219 e Li e Li, 2014, A general framework for analyzing tumor subclonality using SNP array and DNA sequencing data, Genome Biology 15:473-495 descrivono metodi

non solo per identificare mutazioni specifiche della malattia, ma descrivono anche metodi per determinare il numero di copie del gene, zigosità frazionata e subclonalità frazionata della zigosità e zigosità frazionata. Un altro metodo per determinare il numero di copie, come il numero assoluto di copie, riguarda l'utilizzo di segmenti del genoma, ciascun segmento contenente almeno un polimorfismo a singolo nucleotide eterozigote (SNP) e i cui segmenti sono bilanciati (uguale numero di ciascuna versione dell'SNP eterozigote) e condividono un numero comune di copie (numero primario di copie) che è preferibilmente il numero assoluto di copie più frequentemente osservato tra tutti i segmenti bilanciati del genoma, come descritto nella domanda di brevetto internazionale PCT intitolata "Tumor Modeling Based on Primary Balanced Heterozygous Segments" depositata in data pari alla presente. Inoltre, oltre a determinare il numero assoluto di copie, questa domanda può anche determinare zigosità, zigosità frazionata e subclonalità dell'allele mutato o del gene di cui almeno una copia comprende l'allele mutato. Inoltre, la metodologia ivi descritta esegue anche la correzione degli errori per i numeri assoluti di copie, il che migliora l'accuratezza dei numeri assoluti di copie e delle zigosità e dei parametri ivi derivati, come subclonalità, ploidia e così via.

Generalmente, un numero totale di copie del sito nucleotidico a cui si mappa l'allele mutato può significare il numero assoluto di copie della mutazione, che può significare il numero assoluto di copie dell'SNV, in particolare quando la mutazione è un SNV. In generale, il numero assoluto di copie della mutazione può essere preferibilmente dato dal numero assoluto di copie di un segmento a cui si mappa la mutazione (in cui il numero assoluto di copie si trova, ad esempio nel genoma del tumore malato).

Generalmente, un numero di copie di un allele mutato che codifica un neoepitopo può significare un numero assoluto di copie di un allele mutato di una mutazione, che può significare un numero assoluto di copie dell'allele alternativo di un SNV (una zigosità di un SNV), in cui l'allele alternativo dell'SNV è l'allele mutato, in particolare quando la mutazione è un SNV.

Preferibilmente, quando la mutazione non è un SNV, il numero assoluto di copie di un allele mutato di una mutazione può essere stimato in modo simile al metodo applicato per l'SNV.

Generalmente, un numero di copie di un gene può significare il numero assoluto di copie di un segmento, in cui il segmento può essere il gene o può comprendere il gene. Un numero di copie di un gene può significare un numero assoluto di copie di un gene.

Preferibilmente, la malattia può essere qualsiasi malattia in cui si desidera una risposta immunitaria contro la cellula/tessuto malato, come una cellula infettata da un virus. Preferibilmente, la malattia è cancro.

I metodi dell'insegnamento possono comprendere l'ulteriore fase di determinare l'utilizzabilità/appropriatezza dei neoepitopi adatti identificati dai metodi dell'insegnamento come target specifici della malattia adatti all'utilizzo in un metodo per fornire una risposta immunitaria contro il

neoepitopo adatto, come inclusione del neoepitopo adatto in un vaccino anticancro. Pertanto, ulteriori fasi possono comportare uno o più dei seguenti: determinare l'antigenicità e/o l'immunogenicità del neoepitopo adatto; valutare se il neoepitopo adatto è espresso sulla superficie della cellula malata; la capacità di un peptide comprendente il neoepitopo adatto di essere presentato come un epitopo presentato da MHC; determinare l'efficacia di espressione del neoepitopo adatto da un acido nucleico codificante; determinare se i neoepitopi adatti previsti, in particolare quando presenti nel loro contesto di sequenza naturale, ad esempio quando affiancati da sequenze di amminoacidi, affiancano anche detti neoepitopi nella proteina naturale, e quando espressi in cellule presentanti l'antigene sono in grado di stimolare cellule T, come T cellule del paziente aventi la specificità desiderata.

Una volta determinato che il neoepitopo è adatto/appropriato per l'utilizzo come un target in considerazione della sua antigenicità/immunogenicità, capacità di essere espresso, capacità di presentarsi come epitopo presentato da MHC ecc., i neoepitopi idonei identificati possono essere classificati, cioè, prioritari, sul loro potenziale di non essere ridotto o eliminato dalla cellula malata, che è meno probabile che il tessuto malato possa sfuggire alla presa di mira del neoepitopo. Ad esempio, una priorità inizia con il neoepitopo "migliore", che è codificato da un gene essenziale, in cui tutte le copie del gene essenziale hanno la mutazione che codifica il neoepitopo, seguito da una coppia di geni sintetici malati o letali, in cui ogni copia di ciascun gene ha la mutazione che codifica il neoepitopo, seguita da un neoepitopo codificato da un gene driver noto con un numero assoluto di copie molto elevato in cui tutte le copie del gene hanno la mutazione, seguito da un neoepitopo codificato da un gene che non è noto per essere un gene driver con un numero assoluto di copie molto elevato e con un'elevata zigosità, seguito da un neoepitopo codificato da un gene con un numero di copie elevato (zigosità) e così via.

In una forma di realizzazione, un neoepitopo codificato da un gene essenziale con una zigosità frazionata di 1 è preferito ad altri neoepitopi che non sono codificati da un gene essenziale. In una forma di realizzazione, tra due geni essenziali che codificano neoepitopi con una zigosità frazionata di 1, è preferito il neoepitopo codificato dal gene avente un numero assoluto di copie maggiore. In una forma di realizzazione, tra due geni essenziali che codificano i neoepitopi, è preferito il neoepitopo codificato dal gene che porta a un'idoneità inferiore quando eliminato. In una forma di realizzazione, tra geni che codificano neoepitopi in cui tutti i geni hanno una zigosità frazionata di 1, sono preferiti i neoepitopi codificati dai geni aventi un numero assoluto di copie maggiore. In una forma di realizzazione in cui la zigosità frazionata è inferiore a 1, allora i neoepitopi codificati da geni con un'elevata zigosità sono preferiti ai geni con un'elevata zigosità frazionata, e se la zigosità è uguale o simile, allora i geni con un numero assoluto di copie elevato sono preferiti a quelli con un'elevata zigosità frazionata (10 copie dell'allele mutato/20 copie totali del sito nucleotidico sono migliori di 3/4 a causa della maggiore zigosità; 10/100 sono migliori di 10/20 perché il primo può essere un gene driver; 9/100 sono migliori

di 10/20 perché la zigosità è simile ma il primo potrebbe essere un gene driver). In una forma di realizzazione, un neoepitopo codificato da un gene driver, in cui la mutazione specifica della malattia è responsabile della trasformazione della cellula in un fenotipo canceroso è preferito a quelli in cui la mutazione non ha un ruolo nella trasformazione della cellula nel fenotipo canceroso. Inoltre, è preferibile che i neoepitopi abbiano una frazione clonale maggiore anziché inferiore.

Ulteriori forme di realizzazione del presente insegnamento riguardano l'utilizzo dei metodi per determinare l'idoneità di un neoepitopo come target specifico della malattia per la produzione di un medicinale, come un vaccino, ad esempio un vaccino anticancro personalizzato. Il vaccino può essere derivato da uno o più neoepitopi adatti o da una combinazione di neoepitopi adatti individuati dai metodi dell'insegnamento. In una forma di realizzazione preferita, il vaccino comprende un peptide o un polipeptide comprendente uno o più neoepitopi adatti o una combinazione di neoepitopi adatti identificati dai metodi dell'insegnamento o un acido nucleico che codifica detto peptide o polipeptide.

In particolare, può essere fornito un vaccino ricombinante che quando somministrato a un paziente fornisce preferibilmente una raccolta di epitopi presentati da MHC di cui almeno uno è un neoepitopo adatto o almeno due dei quali sono una combinazione adatta di neoepitopi identificati con i metodi del presente insegnamento, come 2 o più, 5 o più, 10 o più, 15 o più, 20 o più, 25 o più, 30 o più e preferibilmente fino a 60, fino a 55, fino a 50, fino a 45, fino a 40, fino a 35 o fino a 30 epitopi presentati da MHC. La presentazione di questi epitopi da parte delle cellule di un paziente, in particolare le cellule che presentano l'antigene, si traduce in cellule T che prendono di mira gli epitopi quando sono legate a MHC e quindi, il tumore del paziente, preferibilmente il tumore primario così come le metastasi tumorali, che esprimono antigeni da cui gli epitopi delle cellule T sono derivati e presentano gli stessi epitopi sulla superficie delle cellule tumorali.

I metodi del presente insegnamento sono anche utili nella produzione di cellule immunitarie ricombinanti che esprimono un recettore dell'antigene mirato a un neoepitopo adatto o a un neoepitopo in una combinazione di neoepitopi adatti. Preferibilmente, le cellule immunitarie sono cellule T e il recettore dell'antigene è un recettore della cellula T.

Il presente insegnamento si riferisce anche a un metodo per fornire una cellula immunitaria ricombinante mirata a un neoepitopo adatto o a un epitopo in una combinazione di neoepitopi adatti, detto metodo comprendendo la trasfettazione di una cellula immunitaria con un recettore dell'antigene ricombinante mirato al neoepitopo adatto o a un epitopo in una combinazione di epitopi adatti identificati dai metodi del presente insegnamento per determinare l'idoneità di un neoepitopo come un target specifico della malattia, nonché per cellule immunitarie ricombinanti prodotte mediante tali metodi.

Il presente insegnamento fornisce anche metodi per prendere di mira una popolazione cellulare o un tessuto che esprime uno o più neoepitopi. Ad

esempio, un anticorpo diretto contro uno o più dei neoepitopi può essere utilizzato per prendere di mira le cellule o i tessuti che esprimono uno o più neoepitopi identificati mediante i metodi qui descritti. In una forma di realizzazione, il presente insegnamento fornisce metodi per fornire una risposta immunitaria a una popolazione di cellule target o a un tessuto target che esprime uno o più neoepitopi in un mammifero, detto metodo comprendendo la somministrazione al mammifero (a) di una o più cellule immunitarie che esprimono uno o più recettori antigenici mirati a uno o più neoepitopi; (b) la somministrazione di un acido nucleico che codifica uno o più dei neoepitopi o (c) la somministrazione di un peptide o polipeptide comprendente uno o più dei neoepitopi, in cui i neoepitopi sono identificati secondo i metodi dell'insegnamento per determinare l'idoneità di un neoepitopo come un target specifico della malattia. In una forma di realizzazione, il metodo per fornire una risposta immunitaria a una popolazione di cellule target o a un tessuto target che esprime uno o più neoepitopi in un mammifero comprende le fasi di (i) determinare, in una cellula malata o in una popolazione di cellule malate, il numero di copie di un allele mutato in un gene che codifica un neoepitopo (una mutazione specifica della malattia) e (ii) somministrare (a) una cellula immunitaria che esprime un recettore dell'antigene mirato al neoepitopo risultante dalla mutazione specifica della malattia; (b) somministrare un acido nucleico che codifica il neoepitopo risultante dalla mutazione specifica della malattia o (c) somministrare un peptide o un polipeptide comprendente il neoepitopo risultante dalla mutazione specifica della malattia.

In una forma di realizzazione, il metodo per fornire una risposta immunitaria a una popolazione di cellule target o a un tessuto target che esprime uno o più neoepitopi in un mammifero comprende le fasi di (i) determinare, in una cellula malata o in una popolazione di cellule malate, il numero di copie di un gene in cui almeno una copia del gene ha una mutazione specifica della malattia che risulta in un neoepitopo e (ii) somministrare (a) una cellula immunitaria che esprime un recettore dell'antigene mirato al neoepitopo risultante dalla mutazione specifica della malattia; (b) somministrare un acido nucleico che codifica il neoepitopo risultante dalla mutazione specifica della malattia o (c) somministrare un peptide o un polipeptide comprendente il neoepitopo risultante dalla mutazione specifica della malattia. In una forma di realizzazione, il metodo per fornire una risposta immunitaria a una popolazione di cellule target o a un tessuto target che esprime uno o più neoepitopi in un mammifero comprende le fasi di (i) determinare, in una cellula malata o in una popolazione di cellule malate, se un gene avente una mutazione specifica della malattia risultante in un neoepitopo è un gene essenziale e (ii) somministrare (a) una cellula immunitaria che esprime un recettore dell'antigene mirato al neoepitopo risultante dalla mutazione specifica della malattia; (b) somministrare un acido nucleico che codifica il neoepitopo risultante dalla mutazione specifica della malattia o (c) somministrare un peptide o un polipeptide comprendente il neoepitopo risultante dalla mutazione specifica della malattia. Preferibilmente, tutte le copie del gene essenziale hanno la mutazione specifica della malattia, cioè la zigosità frazionata è 1.

In una forma di realizzazione, il metodo per fornire una risposta immunitaria a una popolazione di cellule target o a un tessuto target che esprime uno o più neoepitopi in un mammifero comprende le fasi di (i) determinare, in una cellula malata o in una popolazione di cellule malate, se una combinazione di almeno due geni, ciascuno avente una mutazione specifica della malattia risultante in un neoepitopo, sono geni sintetici letali o sintetici malati e (ii) somministrare (a) una o più cellule immunitarie che esprimono uno o più recettori dell'antigene mirati a uno o più neoepitopi risultanti dalle mutazioni specifiche della malattia di almeno due geni; (b) somministrare un acido nucleico che codifica uno o più neoepitopi risultanti dalle mutazioni specifiche della malattia di almeno due geni o (c) somministrare un peptide o un polipeptide comprendente uno o più neoepitopi risultanti dalle mutazioni specifiche della malattia di almeno due geni. Preferibilmente, tutti i neoepitopi risultanti dalle mutazioni specifiche della malattia di almeno due geni sono presi di mira dalle cellule immunitarie somministrate, codificanti dall'acido nucleico somministrato, o compresi all'interno del peptide o di un polipeptide somministrato.

Inoltre, la risposta immunitaria può essere fornita a un mammifero avente una malattia, disturbo o condizione associata all'espressione del neoepitopo risultante dalla mutazione specifica della malattia, in modo tale che la malattia, disturbo o condizione venga trattata o prevenuta. Preferibilmente, la malattia, il disturbo o la condizione è cancro.

Preferibilmente, le cellule immunitarie sono cellule T e i recettori dell'antigene sono recettori della cellula T e la risposta immunitaria è una risposta immunitaria mediata da cellule T. Più preferibilmente, la risposta immunitaria è una risposta immunitaria antitumorale e la popolazione di cellule target o il tessuto target che esprime uno o più neoepitopi adatti sono cellule tumorali o tessuto tumorale.

Altre caratteristiche e vantaggi della presente invenzione saranno resi evidenti dalla seguente descrizione dettagliata e dalle rivendicazioni.

L'invenzione è definita dalle rivendicazioni allegate.

DESCRIZIONE DETTAGLIATA DELL'INVENZIONE

Sebbene il presente insegnamento sia descritto in dettaglio di seguito, è importante comprendere che questo insegnamento non è limitato alle particolari metodologie, protocolli e reagenti descritti nel presente documento, poiché questi possono variare. Si comprenderà altresì che la terminologia impiegata nel presente documento ha solo lo scopo di descrivere particolari forme di realizzazione e non intende essere limitativa della portata del presente insegnamento che verrà limitato solamente dalle rivendicazioni allegate. A meno che non sia diversamente definito, tutti i termini tecnici e scientifici utilizzati nel presente documento hanno lo stesso significato come comunemente inteso da un soggetto mediamente esperto nella tecnica.

Di seguito, verranno descritti gli elementi del presente insegnamento. Questi elementi sono elencati con forme di realizzazione specifiche, tuttavia,

si dovrebbe comprendere che possono essere combinati in qualsiasi modo e in qualsiasi numero per creare ulteriori forme di realizzazione. Le forme di realizzazione preferite variamente descritte non dovrebbero essere interpretate in modo da limitare il presente insegnamento solo le forme di realizzazione esplicitamente descritte. Questa descrizione dovrebbe essere intesa come supporto e comprende forme di realizzazione che combinano le forme di realizzazione esplicitamente descritte con qualsiasi numero degli elementi descritti e/o preferiti. Inoltre, qualsiasi permutazione e combinazione di tutti gli elementi descritti in questa domanda dovrebbe essere considerata descritta dalla descrizione della presente domanda a meno che il contesto non indichi diversamente. Ad esempio, se in una forma di realizzazione preferita il neoepitopo ha un'elevata zigosità piuttosto che un'elevata zigosità frazionata, e in una forma di realizzazione preferita il neoepitopo risulta da una mutazione in un gene essenziale, allora in una forma di realizzazione preferita, il neoepitopo adatto ha un'elevata zigosità frazionata e risulta da una mutazione in un gene essenziale, e in una forma di realizzazione più preferita la zigosità frazionata è uguale a 1.

Preferibilmente, i termini utilizzati nel presente documento sono definiti come descritto in "A multilingual glossary of biotechnological terms: (IUPAC Recommendations)", H. G. W. Leuenberger, B. Nagel, and H. Kölbl, Eds., (1995) Helvetica Chimica Acta, CH-4010 Basel, Switzerland.

La pratica del presente insegnamento impiegherà, se non diversamente indicato, metodi convenzionali di biochimica, biologia cellulare, immunologia e tecniche di DNA ricombinante che sono spiegati nella letteratura nel campo (confrontare, ad esempio, Molecular Cloning: A Laboratory Manual, 2a edizione, J. Sambrook et al. eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1989).

In tutta questa specifica e nelle rivendicazioni che seguono, a meno che il contesto non richieda diversamente, la parola "comprendere" e le variazioni come "comprende" e "comprendente" implicheranno l'inclusione di un elemento indicato, numero intero o fase o gruppo di elementi, numeri interi o fasi ma non l'esclusione di qualsiasi altro elemento, numero intero o fase o gruppo di elementi, numeri interi o fasi sebbene in alcune forme di realizzazione tale altro elemento, numero intero o fase o gruppo di elementi, numeri interi o fasi possano essere esclusi, l'oggetto consiste nell'inclusione di un elemento indicato, numero intero o fase o gruppo di elementi, numeri interi o fasi. Gli articoli "un", "uno", "il" e riferimenti simili utilizzati nel contesto della descrizione dell'invenzione (specialmente nel contesto delle rivendicazioni) devono essere interpretati in modo da comprendere sia il singolare sia il plurale, a meno che non sia diversamente indicato nel presente documento o chiaramente contraddetto dal contesto. La citazione di intervalli di valori nel presente documento è semplicemente intesa a fungere da metodo abbreviato di riferimento singolo a ciascun valore separato che rientra nell'intervallo. Se non diversamente indicato nel presente documento, ciascun singolo valore è incorporato nella specifica come se fosse citato singolarmente nel presente documento.

Tutti i metodi descritti nel presente documento possono essere eseguiti in un qualsiasi ordine idoneo, salvo diversamente indicato nel presente

documento o diversamente chiaramente contraddetto dal contesto. Ad esempio, la determinazione del fatto che il neoepitopo sia un target specifico della malattia, determinando il numero di copie del gene codificante, può essere determinata prima, dopo o in concomitanza con la determinazione che il gene è un gene driver o un gene essenziale, oppure può essere determinato prima, dopo o in concomitanza con la determinazione che il neoepitopo è espresso sulla superficie della cellula o induce una risposta immunitaria soddisfacente tale che sarebbe adatto in un vaccino.

L'utilizzo di uno qualsiasi o di tutti gli esempi, oppure di un linguaggio esemplificativo (ad esempio, "come") fornito nel presente documento viene inteso semplicemente per illustrare meglio l'insegnamento e non pone una limitazione alla portata dell'insegnamento, salvo diversamente rivendicato. Nessun linguaggio nella specifica deve essere interpretato come a indicare elementi non rivendicati fondamentali per la messa in pratica dell'insegnamento.

Il presente insegnamento prevede la terapia di malattie, inclusa l'immunoterapia e la radioterapia, in particolare del cancro, prendendo di mira i neoepitopi ("neoepitopi adatti") che sono espressi solo all'interno o sulle cellule malate e che hanno la caratteristica di essere espressi da geni che sono meno probabilmente silenziati dalla cellula malata, in modo tale che la cellula malata abbia meno probabilità di sfuggire all'immunosorveglianza attraverso il neoepitopo mirato. L'immunoterapia può essere effettuata con metodi immunoterapeutici attivi e/o passivi. Ad esempio, in una forma di realizzazione un anticorpo o altra molecola mirata specificamente a un neoepitopo e coniugata a un agente tossico in grado di distruggere la cellula che esprime il neoepitopo può essere utilizzata secondo il presente insegnamento per prendere di mira e distruggere quella cellula.

L'insegnamento è specificamente rivolto all'identificazione di tali neoepitopi idonei come target specifici della malattia in immunoterapia. Una volta identificati i neoepitopi adatti, possono essere utilizzati nei vaccini al fine di indurre una risposta immunitaria contro il neoepitopo, in particolare, inducendo e/o attivando cellule effettrici appropriate come le cellule T che riconoscono il neoepitopo adatto identificato, in particolare quando presentato nel contesto di MHC, tramite un recettore dell'antigene appropriato, come un recettore della cellula T o un recettore della cellula T artificiale, che si traduce nella morte della cellula malata che esprime il neoepitopo adatto. In alternativa, o in aggiunta, possono essere somministrate cellule immunitarie che riconoscono il neoepitopo adatto identificato attraverso un appropriato recettore dell'antigene, il che comporterà anche la morte delle cellule che esprimono il neoepitopo adatto.

Gli approcci immunoterapeutici secondo l'insegnamento includono l'immunizzazione con un peptide o un polipeptide contenente il neoepitopo adatto, ii) acido nucleico che codifica il peptide o il polipeptide contenente il neoepitopo adatto, iii) cellule ricombinanti che codificano il peptide o il polipeptide contenente il neoepitopo adatto, iv) virus ricombinanti che codificano il peptide o il polipeptide contenente il neoepitopo e v) cellule

che presentano l'antigene pulsate con il peptide o il polipeptide contenente il neoepitopo o trasfettate con acidi nucleici che codificano il peptide o il polipeptide. Altri approcci immunoterapeutici secondo l'insegnamento includono il trasferimento di vi) recettori della cellula T che riconoscono il neoepitopo e vii) cellule effettrici che codificano i recettori (come le cellule T) che riconoscono il neoepitopo, in particolare quando presentate nel contesto di MHC.

Il termine "mutazione specifica della malattia" nel contesto del presente insegnamento si riferisce a una mutazione somatica che è presente nell'acido nucleico di una cellula malata ma assente nell'acido nucleico di una corrispondente cellula normale, non malata. La malattia può essere cancro, pertanto, il termine "mutazione specifica del tumore" o "mutazione specifica del cancro" si riferisce a una mutazione somatica che è presente nell'acido nucleico di un tumore o di una cellula cancerosa ma assente nell'acido nucleico di una corrispondente cellula normale, cioè cellula non tumorale o non cancerosa. I termini "mutazione specifica del tumore" e "mutazione tumorale" e i termini "mutazione specifica del cancro" e "mutazione del cancro" sono qui utilizzati in modo intercambiabile.

Come qui utilizzato, un polimorfismo a singolo nucleotide (SNP) è un sito nel genoma normale in cui almeno uno dei due alleli (materno o paterno) ha un'identità diversa da quella del genoma normale, o rispetto, ad esempio, a un genoma di riferimento.

Come qui utilizzato, un polimorfismo a singolo nucleotide eterozigote (SNP eterozigote) è definito come un sito nel genoma normale in cui i due alleli (gli alleli materno e paterno) hanno un'identità diversa.

Come qui utilizzato, il termine "zigosità frazionata" si riferisce alla frazione del numero di copie di un gene avente una mutazione specifica della malattia in considerazione del numero totale di copie del gene, indipendentemente dal fatto che il gene presenti la mutazione o meno. Ad esempio, se ci sono un totale di 20 copie del gene e 10 copie hanno la mutazione specifica della malattia, la zigosità frazionata è 0,5. Se tutte le copie del gene hanno la mutazione specifica della malattia, la zigosità frazionata è 1. La zigosità frazionata può essere almeno 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9 o almeno 0,95. Nel contesto del presente insegnamento, è preferita una zigosità frazionata maggiore piuttosto che una zigosità frazionata inferiore. In una forma di realizzazione, la zigosità frazionata di un allele mutato che codifica un epitopo, preferibilmente un neoepitopo, è il rapporto tra il numero di copie dell'allele mutato rispetto al numero totale di copie del sito nucleotidico a cui si mappa l'allele mutato, ad esempio in un genoma di riferimento.

Come qui utilizzato, il termine "frazione clonale" si riferisce alla frazione del numero di cellule malate che contengono la stessa mutazione specifica della malattia nello stesso gene e le sue caratteristiche genetiche come numero di copie e zigosità frazionata in considerazione del numero totale di cellule malate, indipendentemente dal fatto che le cellule malate abbiano o meno la stessa mutazione nello stesso gene. Questo termine può

applicarsi anche al tessuto tumorale, in quanto la frazione clonale è la frazione di cellule malate nel tessuto tumorale che contiene la stessa mutazione specifica della malattia nello stesso gene e le sue caratteristiche genetiche, come il numero di copie in considerazione del numero totale di cellule nel tessuto tumorale. Ad esempio, in un campione ottenuto da un tumore, in cui solo la metà del numero totale di cellule tumorali ha la stessa mutazione nello stesso gene, la frazione clonale è 0,5. La frazione clonale può essere almeno 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9 o almeno 0,95. Nel contesto del presente insegnamento, è preferita una frazione clonale maggiore piuttosto che una frazione clonale inferiore. La maggiormente preferita è una frazione clonale di 1, in cui tutte le cellule malate hanno la stessa mutazione specifica della malattia nello stesso gene. "Frazione clonale", "clonalità frazionata" e "subclonalità frazionata" sono qui utilizzate in modo intercambiabile.

Il termine "amplificazione focale" si riferisce a un'amplificazione, o aumento del numero di copie, di una parte di un genoma, ad esempio, l'amplificazione di uno o più geni localizzati insieme sullo stesso cromosoma, che risulta in un numero di copie maggiore di 2, preferibilmente maggiore di 5, 10, 15, 20, 25, 50, 75, 100 per questa parte del genoma, purché non si sia verificato alcun evento di eliminazione per la stessa parte del genoma. Pertanto, ai fini del presente insegnamento, i geni che sono amplificati focalmente nelle cellule malate sono quei geni che possono avere un numero di copie aumentato rispetto al numero di copie wild-type di 2 o al numero di copie wild-type di 1 per quei geni sui cromosomi X e Y nei maschi. L'amplificazione focale è distinta dagli eventi di duplicazione e/o amplificazione dell'intero genoma.

Il termine "risposta immunitaria" si riferisce a una risposta corporea integrata a un antigene e preferibilmente si riferisce a una risposta immunitaria cellulare o a una risposta immunitaria cellulare nonché umorale. La risposta immunitaria può essere protettiva/preventiva/profilattica e/o terapeutica.

"Indurre una risposta immunitaria" può significare che non c'era risposta immunitaria contro un particolare antigene prima dell'induzione, ma può anche significare che c'era un certo livello di risposta immunitaria contro un particolare antigene prima dell'induzione e dopo l'induzione detta risposta immunitaria è migliorata. Pertanto, "indurre una risposta immunitaria" include anche "potenziare una risposta immunitaria". Preferibilmente, dopo aver indotto una risposta immunitaria in un soggetto, detto soggetto è protetto dallo sviluppo di una malattia come una malattia tumorale o la condizione della malattia è migliorata inducendo una risposta immunitaria. Ad esempio, una risposta immunitaria contro un antigene associato al tumore può essere indotta in un paziente malato di cancro o in un soggetto a rischio di sviluppare una malattia cancerosa. Indurre una risposta immunitaria in questo caso può significare che la condizione della malattia del soggetto è migliorata, che nel soggetto non si sviluppano metastasi o che il soggetto a rischio di sviluppare una malattia tumorale non sviluppa una malattia cancerosa.

Una "risposta immunitaria cellulare", una "risposta cellulare", una "risposta cellulare contro un antigene" o un termine simile intende includere una

risposta cellulare diretta a cellule caratterizzate dalla presentazione di un antigene con MHC di classe I o II. La risposta cellulare si riferisce a cellule chiamate cellule T o linfociti T che agiscono come "helper" o "killer". Le cellule T helper (chiamate anche cellule T CD4+) svolgono un ruolo centrale regolando la risposta immunitaria e le cellule killer (chiamate anche cellule T citotossiche, cellule T citolitiche, cellule T CD8+ o CTL) uccidono le cellule malate come le cellule cancerose, impedendo la produzione di più cellule malate. Preferibilmente, di una risposta di CTL anti-tumorale contro cellule tumorali che esprimono uno o più antigeni espressi da uno o più tumori e preferibilmente che presentano tali antigeni espressi dal tumore con MHC di classe I.

Un "antigene" secondo l'insegnamento copre qualsiasi sostanza, preferibilmente un peptide o una proteina, che è un target e/o induce una risposta immunitaria come una reazione specifica con anticorpi o linfociti T (cellule T). Preferibilmente, un antigene comprende almeno un epitopo come un epitopo delle cellule T. Preferibilmente, un antigene nel contesto del presente insegnamento è una molecola che, opzionalmente dopo l'elaborazione, induce una reazione immunitaria, che è preferibilmente specifica per l'antigene (incluse le cellule che esprimono l'antigene). L'antigene o un suo epitopo delle cellule T è preferibilmente presentato da una cellula, preferibilmente da una cellula presentante l'antigene che include una cellula malata, in particolare una cellula cancerosa, nel contesto delle molecole MHC, che si traduce in una risposta immunitaria contro l'antigene (comprese le cellule che esprimono l'antigene).

Preferibilmente, un antigene nel contesto del presente insegnamento è una molecola che, opzionalmente dopo l'elaborazione, induce una reazione immunitaria, che è preferibilmente specifica per l'antigene. Secondo il presente insegnamento, può essere utilizzato qualsiasi antigene adatto, che è un candidato per una reazione immunitaria, in cui la reazione immunitaria può essere sia una reazione immunitaria umorale sia cellulare. Nel contesto del presente insegnamento, l'antigene è preferibilmente presentato da una cellula, preferibilmente da una cellula che presenta l'antigene, nel contesto delle molecole MHC, che può tradursi in una reazione immunitaria contro l'antigene. Un antigene è preferibilmente un prodotto che corrisponde o è derivato da un antigene presente in natura. Tali antigeni naturali possono includere o possono essere derivati da allergeni, virus, batteri, funghi, parassiti e altri agenti infettivi e patogeni o un antigene può anche essere un antigene tumorale. Secondo il presente insegnamento, un antigene può corrispondere a un prodotto presente in natura, ad esempio una proteina virale, o una parte di essa. In forme di realizzazione preferite, l'antigene è un polipeptide superficiale, cioè un polipeptide visualizzato naturalmente sulla superficie di una cellula, un agente patogeno, un batterio, un virus, un fungo, un parassita, un allergene o un tumore. L'antigene può suscitare una risposta immunitaria contro una cellula, un patogeno, un batterio, un virus, un fungo, un parassita, un allergene o un tumore.

Il termine "antigene associato alla malattia" o "antigene specifico della malattia" è utilizzato nel suo senso più ampio per riferirsi a qualsiasi

antigene associato o specifico per una malattia. Tale antigene è una molecola che contiene epitopi che stimoleranno il sistema immunitario dell'ospite a produrre una risposta immunitaria specifica dell'antigene cellulare e/o una risposta anticorpale umorale contro la malattia. L'antigene associato alla malattia può quindi essere utilizzato per scopi terapeutici. Gli antigeni associati alla malattia sono preferibilmente associati all'infezione da microbi, tipicamente antigeni microbici, o associati al cancro, tipicamente tumori.

Il termine "patogeno" si riferisce a materiale biologico patogeno in grado di provocare malattie in un organismo, preferibilmente in un organismo vertebrato. Gli agenti patogeni includono microrganismi come batteri, organismi eucarioti unicellulari (protozoi), funghi e virus.

Nel contesto del presente insegnamento, il termine "antigene del tumore" o "antigene associato al tumore" si riferisce a proteine che sono in condizioni normali espresse in modo specifico in un numero limitato di tessuti e/o organi o in specifici stadi di sviluppo, ad esempio, l'antigene del tumore può essere in condizioni normali espresse in modo specifico nel tessuto dello stomaco, preferibilmente nella mucosa gastrica, negli organi riproduttivi, ad esempio, nei testicoli, nel tessuto trofoblastico, ad esempio, nella placenta o nelle cellule germinali, e sono espressi o espressi in modo anormale in uno o più tessuti tumorali o cancerosi. In questo contesto, "un numero limitato" significa preferibilmente non più di 3, più preferibilmente non più di 2. Gli antigeni del tumore nel contesto del presente insegnamento includono, ad esempio, antigeni di differenziazione, preferibilmente antigeni di differenziazione specifici di tipo cellulare, vale a dire proteine che sono in condizioni normali espresse in modo specifico in un certo tipo di cellula in un determinato stadio di differenziazione, antigeni di cancro/testicoli, vale a dire proteine che in condizioni normali sono espresse in modo specifico nei testicoli e talvolta nella placenta, e antigeni specifici della linea germinale. Nel contesto del presente insegnamento, l'antigene del tumore è preferibilmente associato alla superficie cellulare di una cellula cancerosa ed è preferibilmente non espresso o solo raramente espresso nei tessuti normali. Preferibilmente, l'antigene tumorale o l'espressione aberrante dell'antigene tumorale identifica le cellule tumorali. Nel contesto del presente insegnamento, l'antigene del tumore che è espresso da una cellula cancerosa in un soggetto, ad esempio in un paziente affetto da una malattia cancerosa, è preferibilmente un'auto-proteina in detto soggetto. Nelle forme di realizzazione preferite, l'antigene del tumore nel contesto del presente insegnamento è espresso in condizioni normali, nello specifico in un tessuto o organo che non è essenziale, vale a dire tessuti o organi che quando danneggiati dal sistema immunitario non portano alla morte del soggetto, o in organi o strutture del corpo che non sono o sono solo difficilmente accessibili dal sistema immunitario. Preferibilmente, la sequenza amminoacidica dell'antigene del tumore è identica tra l'antigene del tumore che è espresso nei tessuti normali e l'antigene del tumore che è espresso nei tessuti cancerosi.

Secondo l'insegnamento, i termini "antigene tumorale", "antigene espresso dal tumore", "antigene canceroso" e "antigene espresso dal cancro" sono equivalenti e sono qui utilizzati in modo intercambiabile.

I termini "epitopo", "peptide dell'antigene", "epitopo dell'antigene", "peptide immunogenico" e "MHC legante l'antigene" sono qui utilizzati in modo intercambiabile e si riferiscono a un determinante antigenico in una molecola come un antigene, cioè a una parte o a un frammento di un composto immunologicamente attivo che è riconosciuto dal sistema immunitario, ad esempio, che è riconosciuto da una cellula T, in particolare quando presentato nel contesto delle molecole MHC. Un epitopo di una proteina comprende preferibilmente una porzione continua o discontinua di detta proteina ed è preferibilmente tra 5 e 100, preferibilmente tra 5 e 50, più preferibilmente tra 8 e 30, più preferibilmente tra 10 e 25 amminoacidi di lunghezza, ad esempio, l'epitopo può essere preferibilmente di 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o 25 amminoacidi di lunghezza. Secondo l'insegnamento un epitopo può legarsi a molecole MHC come le molecole MHC sulla superficie di una cellula e quindi, può essere un "MHC legante un peptide". Il termine "complesso maggiore di istocompatibilità" e l'abbreviazione "MHC" includono molecole MHC di classe I e di classe II e si riferiscono a un complesso di geni che si trova in tutti i vertebrati. Le proteine o le molecole MHC sono importanti per la segnalazione tra i linfociti e le cellule che presentano l'antigene o le cellule malate nelle reazioni immunitarie, in cui le proteine o le molecole MHC si legano ai peptidi e le presentano per il riconoscimento da parte dei recettori di cellule T. Le proteine codificate dall'MHC sono espresse sulla superficie delle cellule e mostrano sia autoantigeni (frammenti di peptidi dalla cellula stessa) sia non autoantigeni (ad esempio, frammenti di microrganismi invasori) a una cellula T. Le porzioni immunogeniche preferite si legano a una molecola MHC di classe I o di classe II. Come qui utilizzato, si dice che un peptide "si lega a" una molecola MHC di classe I o di classe II se tale legame è rilevabile utilizzando qualsiasi saggio noto nella tecnica. Il termine "MHC legante il peptide" si riferisce a un peptide che si lega a una molecola MHC di classe I e/o una molecola MHC di classe II. Nel caso di complessi di MHC di classe I/peptide, i peptidi leganti sono tipicamente lunghi 8-10 amminoacidi sebbene peptidi più lunghi o più corti possano essere efficaci. Nel caso di complessi di MHC di classe II/peptide, i peptidi leganti sono tipicamente lunghi 10-25 amminoacidi e sono in particolare lunghi 13-18 amminoacidi, mentre peptidi più lunghi e più corti possono essere efficaci.

Come qui utilizzato, il termine "neoepitopo" si riferisce a un epitopo che non è presente in un riferimento come una normale cellula non cancerosa o germinale ma si trova nelle cellule malate, come cellule cancerose. Ciò include, in particolare, le situazioni in cui in una normale cellula non cancerosa o germinale si trova un epitopo corrispondente, tuttavia, a causa di una o più mutazioni in una cellula cancerosa la sequenza dell'epitopo viene modificata in modo da risultare nel neoepitopo. Inoltre, un neoepitopo può non solo essere specifico per le cellule malate, ma può anche essere specifico per il paziente affetto dalla malattia. Poiché i neoepitopi e i neoepitopi adatti identificati dai metodi dell'insegnamento sono sottoinsiemi degli epitopi, la presente descrizione relativa agli epitopi in generale come target immunologici si applica ugualmente ai neoepitopi e ai neoepitopi adatti.

In una forma di realizzazione particolarmente preferita dell'insegnamento, un epitopo o un neoepitopo è un epitopo della cellula T. Come qui utilizzato, il termine "epitopo delle cellule T" si riferisce a un peptide che si lega a una molecola MHC in una configurazione riconosciuta da un recettore delle cellule T. Tipicamente, gli epitopi delle cellule T sono presentati sulla superficie di una cellula che presenta l'antigene.

Come qui utilizzato, il termine "previsione di modificazioni di amminoacidi immunogenici" si riferisce a una previsione se un peptide comprendente tale modificazione di amminoacidi sarà immunogenico e quindi utile come epitopo, in particolare un epitopo della cellula T, nella vaccinazione.

Secondo l'insegnamento, un epitopo di cellule T può essere presente in un vaccino come parte di un'entità più grande come una sequenza di vaccino e/o un polipeptide comprendente più di un epitopo di cellule T. Il peptide presentato o l'epitopo delle cellule T viene prodotto dopo un'adeguata elaborazione.

Gli epitopi delle cellule T possono essere modificati in uno o più residui che non sono essenziali per il riconoscimento del TCR o per il legame con MHC. Tali epitopi delle cellule T modificati possono essere considerati immunologicamente equivalenti.

Preferibilmente, l'antigene o il peptide se riconosciuto da un recettore di cellule T è in grado di indurre in presenza di opportuni segnali di costimolazione, l'espansione clonale della cellula T che trasporta il recettore di cellule T che riconosce l'antigene o il peptide.

Preferibilmente, un epitopo di cellule T comprende una sequenza amminoacidica sostanzialmente corrispondente alla sequenza amminoacidica di un frammento di un antigene. Preferibilmente, detto frammento di un antigene è il peptide MHC di classe I e/o di classe II presentato.

L'epitopo cellulare AT secondo l'insegnamento si riferisce preferibilmente a una porzione o frammento di un antigene che è in grado di stimolare una risposta immunitaria, preferibilmente una risposta cellulare contro l'antigene o cellule caratterizzate dall'espressione dell'antigene e preferibilmente dalla presentazione dell'antigene come cellule malate, in particolare cellule cancerose. Preferibilmente, tale peptide è in grado di stimolare una risposta cellulare contro una cellula caratterizzata dalla presentazione di un antigene di classe I MHC e preferibilmente è in grado di stimolare un linfocita T citotossico reattivo all'antigene (CTL).

In alcune forme di realizzazione l'antigene è un autoantigene, in particolare un antigene del tumore. Gli antigeni del tumore e le loro determinazioni sono noti alla persona esperta.

Il termine "immunogenicità" si riferisce all'efficacia relativa nell'indurre una risposta immunitaria che è preferibilmente associata a trattamenti terapeutici, come trattamenti contro il cancro. Come qui utilizzato, il termine "immunogenico" si riferisce alla proprietà di avere immunogenicità.

Ad esempio, il termine "modifica immunogenica" quando utilizzato nel contesto di un peptide, polipeptide o proteina si riferisce all'efficacia di

detto peptide, polipeptide o proteina nell'indurre una risposta immunitaria che è causata da e/o diretta contro detta modifica. Preferibilmente, il peptide, polipeptide o proteina non modificato non induce una risposta immunitaria, induce una risposta immunitaria diversa o induce un livello diverso, preferibilmente un livello inferiore, di risposta immunitaria.

Secondo l'insegnamento, il termine "immunogenicità" o "immunogenico" si riferisce preferibilmente alla relativa efficacia nell'indurre una risposta immunitaria biologicamente rilevante, in particolare una risposta immunitaria utile per la vaccinazione. Pertanto, in una forma di realizzazione preferita, una modificazione dell'amminoacido o un peptide modificato è immunogenico se induce una risposta immunitaria contro la modificazione target in un soggetto, la cui risposta immunitaria può essere vantaggiosa per scopi terapeutici o profilattici.

"Elaborazione dell'antigene" o "elaborazione" si riferisce alla degradazione di un polipeptide o di un antigene nei prodotti di elaborazione, che sono frammenti di detto polipeptide o antigene (ad esempio, la degradazione di un polipeptide in peptidi) e l'associazione di uno o più di questi frammenti (ad esempio, tramite legame) con molecole MHC per la presentazione da parte delle cellule, preferibilmente cellule che presentano l'antigene a cellule T specifiche.

Le "cellule presentanti l'antigene" (APC) sono cellule che presentano frammenti peptidici di antigeni proteici in associazione con molecole MHC sulla loro superficie cellulare. Alcuni APC possono attivare le cellule T specifiche dell'antigene.

Le cellule che presentano l'antigene professionale sono molto efficienti nell'internalizzare l'antigene, sia per fagocitosi sia mediante endocitosi mediata dal recettore, e quindi visualizzano un frammento dell'antigene, legato a una molecola MHC di classe II, sulla loro membrana. La cellula T riconosce e interagisce con il complesso della molecola MHC di classe II dell'antigene sulla membrana della cellula che presenta l'antigene. Un ulteriore segnale costimolatorio viene quindi prodotto dalla cellula che presenta l'antigene, portando all'attivazione della cellula T. L'espressione di molecole costimolatorie è una caratteristica distintiva delle cellule che presentano l'antigene professionale.

I principali tipi di cellule che presentano l'antigene professionale sono le cellule dendritiche, che hanno la più ampia gamma di presentazione dell'antigene e sono probabilmente le cellule più importanti che presentano l'antigene, i macrofagi, le cellule B e determinate cellule epiteliali attivate.

Le cellule dendritiche (DC) sono popolazioni di leucociti che presentano antigeni catturati nei tessuti periferici alla cellula T tramite i percorsi di presentazione dell'antigene MHC di classe II e I. È ben noto che le cellule dendritiche sono potenti induttori delle risposte immunitarie e l'attivazione di queste cellule è una fase fondamentale per l'induzione dell'immunità antitumorale. Le cellule dendritiche sono opportunamente classificate come cellule "immature" e "mature", che possono essere utilizzate come un modo semplice per discriminare tra due fenotipi ben

caratterizzati. Tuttavia, questa nomenclatura non dovrebbe essere interpretata per escludere tutte le possibili fasi intermedie di differenziazione. Le cellule dendritiche immature sono caratterizzate come cellule che presentano l'antigene con un'elevata capacità di assorbimento e di elaborazione dell'antigene, che è correlata con l'alta espressione del recettore Fc γ e del recettore di mannosio. Il fenotipo maturo è tipicamente caratterizzato da una minore espressione di questi marcatori, ma un'elevata espressione di molecole della superficie cellulare responsabili dell'attivazione della cellula T come MHC di classe I e II, molecole di adesione (ad esempio CD54 e CD11) e molecole costimolatorie (ad esempio, CD40, CD80, CD86 e 4-1 BB).

La maturazione delle cellule dendritiche è indicata come lo stato di attivazione delle cellule dendritiche a cui tali cellule dendritiche che presentano l'antigene portano all'innescamento delle cellule T, mentre la presentazione da parte delle cellule dendritiche immature risulta in tolleranza. La maturazione della cellula dendritica è principalmente causata da biomolecole con caratteristiche microbiche rilevate da recettori innati (DNA batterico, RNA virale, endotossina ecc.), citochine pro-infiammatorie (TNF, IL-1, IFN), legatura di CD40 sulla superficie della cellula dendritica mediante CD40L e sostanze rilasciate da cellule sottoposte a morte cellulare stressante. Le cellule dendritiche possono essere derivate coltivando cellule del midollo osseo in vitro con citochine, come il fattore di stimolazione delle colonie di granulociti macrofagi (GM-CSF) e il fattore alfa di necrosi tumorale.

Le cellule che presentano l'antigene non professionale non esprimono costitutivamente le proteine MHC di classe II richieste per l'interazione con le cellule T naïve; questi sono espressi solo su stimolazione delle cellule che presentano l'antigene non professionale da parte di alcune citochine come IFN γ .

"Cellule che presentano l'antigene" possono essere caricate con peptidi presentati da MHC di classe I trasducendo le cellule con acido nucleico, preferibilmente RNA, che codifica un peptide o un polipeptide comprendente il peptide da presentare, ad esempio un acido nucleico che codifica l'antigene.

In alcune forme di realizzazione, una composizione farmaceutica dell'insegnamento comprendente un veicolo di rilascio del gene che prende di mira una cellula dendritica o una cellula che presenta l'antigene può essere somministrata a un paziente, traducendosi nella trasfezione che si verifica in vivo. La trasfezione in vivo di cellule dendritiche, ad esempio, può generalmente essere eseguita utilizzando qualsiasi metodo noto nella tecnica, come quelli descritti nel documento WO 97/24447 o nell'approccio della pistola genica descritto da Mahvi et al., *Immunology and cell Biology* 75:456-460, 1997.

Il termine "cellula che presenta l'antigene" include anche le cellule target.

"Cellula target" indica una cellula che è il target per una risposta immunitaria come una risposta immunitaria cellulare. Le cellule target includono cellule che presentano un antigene o un epitopo dell'antigene, vale a dire un frammento peptidico derivato da un antigene e includono qualsiasi cellula indesiderabile come una cellula cancerosa. In forme di realizzazione preferite, la cellula target è una cellula che esprime un antigene come qui descritto e preferibilmente presenta detto antigene con MHC di classe I.

Il termine "porzione" si riferisce a una frazione. Rispetto a una struttura particolare come una sequenza di amminoacidi o una proteina, il termine "porzione" di essa può indicare una frazione continua o discontinua di detta struttura. Preferibilmente, una porzione di una sequenza di amminoacidi comprende almeno l'1%, almeno il 5%, almeno il 10%, almeno il 20%, almeno il 30%, preferibilmente almeno il 40%, preferibilmente almeno il 50%, più preferibilmente almeno il 60%, più preferibilmente almeno il 70%, ancora più preferibilmente almeno l'80% e più preferibilmente almeno il 90% degli amminoacidi di detta sequenza amminoacidica. Preferibilmente, se la porzione è una frazione discontinua, detta frazione discontinua è composta da 2, 3, 4, 5, 6, 7, 8 o più parti di una struttura, ciascuna parte essendo un elemento continuo della struttura. Ad esempio, una frazione discontinua di una sequenza di amminoacidi può essere composta da 2, 3, 4, 5, 6, 7, 8 o più, preferibilmente non più di 4 parti di detta sequenza di amminoacidi, in cui ciascuna parte comprende preferibilmente almeno 5 amminoacidi continui, almeno 10 amminoacidi continui, preferibilmente almeno 20 amminoacidi continui, preferibilmente almeno 30 amminoacidi continui della sequenza amminoacidica.

I termini "parte" e "frammento" sono utilizzati nel presente documento in modo intercambiabile e si riferiscono a un elemento continuo. Ad esempio, una parte di una struttura come una sequenza di amminoacidi o una proteina si riferisce a un elemento continuo di detta struttura. Una porzione, una parte o un frammento di una struttura comprende preferibilmente una o più proprietà funzionali di detta struttura. Ad esempio, una porzione, una parte o un frammento di un epitopo, peptide o proteina è preferibilmente immunologicamente equivalente all'epitopo, peptide o proteina da cui è derivato. Nel contesto del presente insegnamento, una "parte" di una struttura come una sequenza di amminoacidi comprende, preferibilmente consiste almeno nel 10%, almeno nel 20%, almeno nel 30%, almeno nel 40%, almeno nel 50%, almeno nel 60%, almeno nel 70%, almeno nell'80%, almeno nell'85%, almeno nel 90%, almeno nel 92%, almeno nel 94%, almeno nel 96%, almeno nel 98%, almeno nel 99% dell'intera struttura o sequenza amminoacidica.

Il termine "cellula immunoreattiva" nel contesto del presente insegnamento si riferisce a una cellula che esercita funzioni effettrici durante una reazione immunitaria. Una "cellula immunoreattiva" è preferibilmente in grado di legare un antigene o una cellula caratterizzata dalla presentazione di un antigene o di un peptide dell'antigene derivato da un antigene e che media una risposta immunitaria. Ad esempio, tali cellule secernono citochine e/o chemochine, secernono anticorpi, riconoscono le cellule cancerose e opzionalmente eliminano tali cellule. Ad esempio, le cellule

immunoreattive comprendono le cellule T (cellule T citotossiche, cellule T helper, cellule T infiltranti il tumore), cellule B, cellule natural killer, neutrofili, macrofagi e cellule dendritiche. Preferibilmente, nel contesto del presente insegnamento, "cellule immunoreattive" sono cellule T, preferibilmente cellule T CD4+ e/o CD8+.

Preferibilmente, una "cellula immunoreattiva" riconosce un antigene o un peptide dell'antigene derivato da un antigene con un certo grado di specificità, in particolare se presentato nel contesto di molecole MHC come sulla superficie di cellule che presentano l'antigene o cellule malate come le cellule cancerose. Preferibilmente, detto riconoscimento consente alla cellula che riconosce un antigene o un peptide dell'antigene derivato da detto antigene di essere sensibile o reattivo. Se la cellula è una cellula T helper (cellula T CD4+) portatrice di recettori che riconoscono un antigene o un peptide dell'antigene derivato da un antigene nel contesto delle molecole MHC di classe II, tale capacità di risposta o reattività può comportare il rilascio di citochine e/o l'attivazione dei linfociti CD8+ (CTL) e/o delle cellule B. Se la cellula è un CTL, tale capacità di risposta o reattività può comportare l'eliminazione di cellule presentate nel contesto di molecole MHC di classe I, vale a dire cellule caratterizzate dalla presentazione di un antigene con MHC di classe I, ad esempio, tramite apoptosi o una lisi cellulare mediata dalla perforina. La capacità di risposta a CTL può includere un flusso di calcio sostenuto, divisione cellulare, produzione di citochine come IFN- γ e TNF- α , sovraregolazione dei marcatori di attivazione come CD44 e CD69 e la distruzione citolitica specifica di cellule target che esprimono l'antigene. La capacità di risposta a CTL può anche essere determinata utilizzando un reporter artificiale che indica accuratamente la capacità di risposta a CTL. Tale CTL che riconosce un antigene o un peptide dell'antigene derivato da un antigene ed è sensibile o reattivo è anche qui definito "CTL sensibile all'antigene". Se la cellula è una cellula B, tale capacità di risposta può comportare il rilascio di immunoglobuline.

I termini "cellula T" e "linfocita T" sono utilizzati nel presente documento in modo intercambiabile e includono cellule T helper (cellule T CD4+) e cellule T citotossiche (CTL, cellule T CD8+) che comprendono cellule T citolitiche.

Le cellule T appartengono a un gruppo di globuli bianchi noti come linfociti e svolgono un ruolo centrale nell'immunità cellulo-mediata. Possono essere distinti da altri tipi di linfociti, come le cellule B e le cellule natural killer per la presenza di uno speciale recettore sulla loro superficie cellulare chiamato recettore di cellule T (TCR). Il timo è il principale organo responsabile della maturazione delle cellule T. Sono stati scoperti diversi sottoinsiemi di cellule T, ognuno con una funzione distinta.

Le cellule T helper aiutano altre cellule di globuli bianchi nei processi immunologici, compresa la maturazione delle cellule B in plasmacellule e l'attivazione di cellule T citotossiche e macrofagi, tra le altre funzioni. Queste cellule sono anche note come cellule T CD4+, poiché esprimono la proteina CD4 sulla loro superficie. Le cellule T helper si attivano quando vengono presentate con antigeni peptidici mediante molecole MHC di

classe II, che sono espresse sulla superficie delle cellule che presentano l'antigene (APC). Una volta attivate, si dividono rapidamente e secernono piccole proteine chiamate citochine che regolano o aiutano la risposta immunitaria attiva.

Le cellule T citotossiche distruggono le cellule viralmente infette e le cellule tumorali e sono anche implicate nel rigetto del trapianto. Queste cellule sono anche note come cellule T CD8+, poiché esprimono la glicoproteina CD8 sulla loro superficie. Queste cellule riconoscono i loro obiettivi legandosi all'antigene associato alle molecole MHC di classe I, che sono presenti sulla superficie di quasi tutte le cellule del corpo.

La maggior parte delle cellule T ha un recettore della cellula T (TCR) esistente come un complesso di diverse proteine. L'effettivo recettore della cellula T è composto da due catene peptidiche separate, che sono prodotte dai geni indipendenti alfa e beta ($TCR\alpha$ e $TCR\beta$) del recettore della cellula T e sono chiamati catene α - e β -TCR. $\gamma\delta$ Le cellule T (cellule T gamma delta) rappresentano un piccolo sottoinsieme di cellule T che possiedono un distinto recettore della cellula T (TCR) sulla loro superficie. Tuttavia, nelle cellule T $\gamma\delta$, il TCR è costituito da una catena γ e una catena δ . Questo gruppo di cellule T è molto meno comune (2% delle cellule T totali) rispetto alle cellule T $\alpha\beta$.

Secondo l'insegnamento, il termine "recettore dell'antigene" include recettori naturali come il recettore della cellula T nonché recettori ingegnerizzati, che conferiscono una specificità arbitraria come la specificità di un anticorpo monoclonale su una cellula effettrice immunitaria come una cellula T. In questo modo, un gran numero di cellule T specifiche dell'antigene può essere generato per il trasferimento cellulare adottivo. Pertanto, un recettore dell'antigene secondo l'insegnamento T può essere presente sulle cellule T, ad esempio al posto o in aggiunta al recettore della cellula T della cellula T stessa. Tali cellule T non richiedono necessariamente l'elaborazione e la presentazione di un antigene per il riconoscimento della cellula target, ma piuttosto possono riconoscere preferibilmente con specificità qualsiasi antigene presente su una cellula target. Preferibilmente, detto recettore dell'antigene è espresso sulla superficie delle cellule. Ai fini del presente insegnamento, le cellule T comprendenti un recettore dell'antigene sono comprese dal termine "cellula T" come qui utilizzato. Nello specifico, secondo l'insegnamento, il termine "recettore dell'antigene" include recettori artificiali comprendenti una singola molecola o un complesso di molecole che riconoscono, cioè si legano a, una struttura target (ad esempio un antigene) su una cellula target come una cellula cancerosa (ad esempio legando un sito di legame dell'antigene o un dominio di legame dell'antigene a un antigene espresso sulla superficie della cellula target) e può conferire specificità a una cellula effettrice immunitaria come una cellula T che esprime detto recettore dell'antigene sulla superficie della cellula. Preferibilmente, il riconoscimento della struttura target da parte di un recettore dell'antigene si traduce nell'attivazione di una cellula effettrice immunitaria che esprime detto recettore dell'antigene. Un recettore dell'antigene può comprendere una o più unità proteiche, dette unità proteiche comprendendo uno o più domini come qui descritto. Secondo l'insegnamento, un "recettore dell'antigene" può anche essere un "recettore dell'antigene chimerico (CAR)", un "recettore

chimerico della cellula T" o un "recettore artificiale della cellula T".

Un antigene può essere riconosciuto da un recettore dell'antigene attraverso qualsiasi dominio di riconoscimento dell'antigene (qui indicato anche semplicemente come "domini") in grado di formare un sito di legame dell'antigene come attraverso porzioni di anticorpi che legano l'antigene e recettori della cellula T che possono risiedere sopra catene peptidiche uguali o diverse. In una forma di realizzazione, i due domini che formano un sito di legame dell'antigene sono derivati da un'immunoglobulina. In una forma di realizzazione, i due domini che formano un sito di legame dell'antigene sono derivati da un recettore della cellula T. Particolarmente preferiti sono i domini variabili dell'anticorpo, come frammenti variabili a catena singola (scFv) derivati da anticorpi monoclonali e domini variabili del recettore della cellula T, in particolare catene singole alfa e beta di TCR. In effetti, quasi tutto ciò che lega un determinato target con alta affinità può essere utilizzato come dominio di riconoscimento dell'antigene.

Il primo segnale di attivazione delle cellule T è fornito dal legame del recettore di cellule T a un breve peptide presentato dal complesso principale di istocompatibilità (MHC) su un'altra cellula. Questo garantisce che venga attivata solo una cellula T con un TCR specifico per quel peptide. La cellula partner è solitamente una cellula di presentazione dell'antigene professionale (APC), di solito una cellula dendritica nel caso di risposte naive, sebbene le cellule B e i macrofagi possano essere APC importanti. I peptidi presentati alle cellule T CD8+ dalle molecole MHC di classe I hanno una lunghezza di 8-10 amminoacidi; i peptidi presentati alle cellule T CD4+ dalle molecole MHC di classe II sono più lunghi, poiché le estremità della fessura di legame della molecola MHC di classe II sono aperte.

Secondo il presente insegnamento, una molecola è in grado di legarsi a un target se ha un'affinità significativa per detto target predeterminato e si lega a detto target predeterminato in saggi standard. "Affinità" o "affinità di legame" è spesso misurata dalla costante di dissociazione di equilibrio (KD). Una molecola non è (sostanzialmente) in grado di legarsi a un target se non ha affinità significativa con detto target e non si lega in modo significativo a detto target nei saggi standard.

I linfociti T citotossici possono essere generati in vivo incorporando un antigene o un peptide dell'antigene in cellule che presentano l'antigene in vivo. L'antigene o il peptide dell'antigene può essere rappresentato come proteina, come DNA (ad esempio all'interno di un vettore) o come RNA. L'antigene può essere elaborato per produrre il partner del peptide della molecola MHC, mentre il frammento può essere presentato senza la necessità di ulteriore elaborazione. Quest'ultimo è il caso in particolare, se questi possono legarsi alle molecole MHC. In generale, è possibile la somministrazione a un paziente mediante iniezione intradermica. Tuttavia, l'iniezione può essere effettuata anche intranodalmente in un linfonodo (Maloy et al., 2001, Proc Natl Acad Sci USA 98:3299-303). Le cellule risultanti presentano il complesso di interesse e sono riconosciute da linfociti T citotossici autologhi, che poi si propagano.

L'attivazione specifica delle cellule T CD4+ o CD8+ può essere rilevata in vari modi. I metodi per rilevare l'attivazione della cellula T specifica includono il rilevamento della proliferazione di cellule T, la produzione di citochine (ad esempio, linfocine) o la generazione di attività citolitica. Per le cellule T CD4+, un metodo preferito per rilevare l'attivazione della cellula T specifica è il rilevamento della proliferazione delle cellule T. Per le cellule T CD8+, un metodo preferito per rilevare l'attivazione della cellula T specifica è il rilevamento della generazione di attività citolitica.

Per "cellula caratterizzata dalla presentazione di un antigene" o "cellula che presenta un antigene" o espressioni simili si intende una cellula come una cellula malata, ad esempio una cellula cancerosa, o una cellula che presenta l'antigene che presenta l'antigene che esprime o un frammento derivato da detto antigene, ad esempio mediante l'elaborazione dell'antigene, nel contesto di molecole MHC, in particolare molecole MHC di classe I. Allo stesso modo, il termine "malattia caratterizzata dalla presentazione di un antigene" denota una malattia che coinvolge cellule caratterizzate dalla presentazione di un antigene, in particolare con MHC di classe I. La presentazione di un antigene da parte di una cellula può essere effettuata trasferendo la cellula con un acido nucleico come l'RNA che codifica l'antigene.

Per "frammento di un antigene presentato" o espressioni simili si intende che il frammento può essere presentato da MHC di classe I o di classe II, preferibilmente MHC di classe I, ad esempio quando aggiunto direttamente alle cellule che presentano l'antigene. In una forma di realizzazione, il frammento è un frammento che è naturalmente presentato dalle cellule che esprimono un antigene.

Il termine "immunologicamente equivalente" significa che la molecola immunologicamente equivalente come la sequenza di amminoacidi immunologicamente equivalente presenta le stesse o essenzialmente le stesse proprietà immunologiche e/o esercita gli stessi o essenzialmente gli stessi effetti immunologici, ad esempio, rispetto al tipo di effetto immunologico come l'induzione di una risposta immunitaria umorale e/o cellulare, la forza e/o la durata della reazione immunitaria indotta o la specificità della reazione immunitaria indotta. Nel contesto del presente insegnamento, il termine "immunologicamente equivalente" è preferibilmente utilizzato rispetto agli effetti o alle proprietà immunologiche di un peptide o di una variante del peptide utilizzata per l'immunizzazione. Ad esempio, una sequenza amminoacidica è immunologicamente equivalente a una sequenza amminoacidica di riferimento se detta sequenza amminoacidica quando esposta al sistema immunitario di un soggetto induce una reazione immunitaria avente una specificità di reagire con la sequenza amminoacidica di riferimento.

Il termine "funzioni effettrici immunitarie" nel contesto del presente insegnamento include qualsiasi funzione mediata da componenti del sistema immunitario che si traduce, ad esempio, nella distruzione di cellule tumorali o nell'inibizione della crescita tumorale e/o nell'inibizione dello sviluppo del tumore, inclusa l'inibizione della diffusione del tumore e delle metastasi. Preferibilmente, le funzioni effettrici immunitarie nel contesto del presente insegnamento sono funzioni effettrici mediate dalla cellula T. Tali funzioni comprendono nel caso di una cellula T helper (cellula T

CD4+) il riconoscimento di un antigene o di un peptide dell'antigene derivato da un antigene nel contesto di molecole MHC di classe II da parte di recettori della cellula T, il rilascio di citochine e/o l'attivazione di linfociti CD8+ (CTL) e/o cellule B e nel caso di CTL il riconoscimento di un antigene o di un peptide dell'antigene derivato da un antigene nel contesto di molecole MHC di classe I da parte di recettori della cellula T, l'eliminazione di cellule presentate nel contesto di molecole MHC di classe I, ad esempio, tramite apoptosi o lisi cellulare mediata dalla perforina, produzione di citochine come IFN- γ e TNF- α e la distruzione citolitica specifica delle cellule target che esprimono l'antigene.

Il termine "complesso maggiore di istocompatibilità" e l'abbreviazione "MHC" includono molecole MHC di classe I e MHC di classe II e si riferiscono a un complesso di geni che si trova in tutti i vertebrati. Le proteine o le molecole MHC sono importanti per la segnalazione tra i linfociti e le cellule che presentano l'antigene o le cellule malate nelle reazioni immunitarie, in cui le proteine o le molecole MHC si legano ai peptidi e le presentano per il riconoscimento da parte dei recettori di cellule T. Le proteine codificate dall'MHC sono espresse sulla superficie delle cellule e mostrano sia autoantigeni (frammenti di peptidi dalla cellula stessa) sia non autoantigeni (ad esempio, frammenti di microrganismi invasori) a una cellula T.

La regione MHC è divisa in tre sottogruppi, classe I, classe II e classe III. Le proteine di MHC di classe I contengono una catena α e β -microglobulina (non fa parte dell'MHC codificata dal cromosoma 15). Presentano frammenti di antigene alle cellule T citotossiche. Sulla maggior parte delle cellule del sistema immunitario, in particolare sulle cellule che presentano l'antigene, le proteine MHC di classe II contengono catene α e β e presentano frammenti dell'antigene alle cellule T helper. La regione MHC di classe III codifica per altri componenti immunitari, come i componenti del complemento e alcuni che codificano le citochine.

L'MHC è sia poligenico (ci sono diversi geni MHC di classe I e MHC di classe II) che polimorfico (ci sono più alleli di ogni gene).

Come qui utilizzato, il termine "aplotipo" si riferisce agli alleli HLA trovati su un cromosoma e alle proteine codificate in tal modo. L'aplotipo può anche riferirsi all'allele presente in qualsiasi locus all'interno dell'MHC. Ogni classe di MHC è rappresentata da diversi loci: ad esempio, HLA-A (Antigene leucocitario umano-A), HLA-B, HLA-C, HLA-E, HLA-F, HLA-G, HLA-H, HLA-J, HLA-K, HLA-L, HLA-P e HLA-V per la classe I e HLA-DRA, HLA-DRB1-9, HLA-DQA1, HLA-DQB1, HLA-DPA1, HLA-DPB1, HLA-DMA, HLA-DMB, HLA-DOA e HLA-DOB per la classe II. I termini "allele di HLA" e "allele di MHC" sono qui utilizzati in modo intercambiabile.

Le MHC mostrano un polimorfismo estremo. All'interno della popolazione umana ci sono, in ogni locus genetico, un gran numero di aplotipi che comprendono alleli distinti. Diversi alleli polimorfici di MHC, sia di classe I sia di classe II, hanno specificità peptidiche differenti, in quanto ogni allele codifica proteine che si legano ai peptidi esibendo particolari schemi di sequenza.

Nel contesto del presente insegnamento, una molecola MHC è preferibilmente una molecola HLA.

Nel contesto del presente insegnamento, il termine "peptide legante MHC" include peptidi o peptidi leganti MHC di classe I e/o classe II che possono essere trattati per produrre peptidi leganti MHC di classe I e/o classe II. Nel caso di complessi MHC/peptide di classe I, i peptidi leganti sono tipicamente 8-12, preferibilmente 8-10 amminoacidi lunghi sebbene peptidi più lunghi o più corti possano essere efficaci. Nel caso di complessi MHC/peptide di classe II, i peptidi leganti sono tipicamente 9-30, preferibilmente lunghi 10-25 amminoacidi e sono in particolare lunghi 13-18 amminoacidi, mentre i peptidi più lunghi e più corti possono essere efficaci.

Un "peptide dell'antigene" si riferisce preferibilmente a una porzione o frammento di un antigene che è in grado di stimolare una risposta immunitaria, preferibilmente una risposta cellulare contro l'antigene o le cellule caratterizzate dall'espressione dell'antigene e preferibilmente dalla presentazione dell'antigene come cellule malate, in particolare cellule cancerose. Preferibilmente, un peptide dell'antigene è in grado di stimolare una risposta cellulare contro una cellula caratterizzata dalla presentazione di un antigene di MHC di classe I e preferibilmente è in grado di stimolare un linfocita T citotossico reattivo all'antigene (CTL). Preferibilmente, i peptidi dell'antigene sono peptidi presentati da MHC di classe I e/o di classe II o possono essere elaborati per produrre peptidi presentati da MHC di classe I e/o di classe II. Preferibilmente, i peptidi dell'antigene comprendono una sequenza amminoacidica sostanzialmente corrispondente alla sequenza amminoacidica di un frammento di un antigene. Preferibilmente, detto frammento di un antigene è il peptide MHC di classe I e/o di classe II presentato. Preferibilmente, un peptide dell'antigene comprende una sequenza amminoacidica sostanzialmente corrispondente alla sequenza amminoacidica di tale frammento e viene elaborato per produrre tale frammento, cioè un peptide presentato da MHC di classe I e/o di classe II derivato da un antigene.

Se un peptide deve essere presentato in modo diretto, cioè senza elaborazione, in particolare senza scissione, possiede una lunghezza che è adatta a legarsi a una molecola MHC, in particolare una molecola MHC di classe I e preferibilmente è lungo 7-20 amminoacidi, più preferibilmente 7-12 amminoacidi, più preferibilmente 8-11 amminoacidi, in particolare 9 o 10 amminoacidi.

Se un peptide è parte di un'entità più grande comprendente sequenze aggiuntive, ad esempio di una sequenza di vaccino o di polipeptide e deve essere presentato dopo l'elaborazione, in particolare dopo la scissione, il peptide prodotto dall'elaborazione ha una lunghezza che è adatta a legarsi a una molecola MHC, in particolare una molecola MHC di classe I, e preferibilmente è lunga 7-20 amminoacidi, più preferibilmente 7-12 amminoacidi, più preferibilmente 8-11 amminoacidi, in particolare 9 o 10 amminoacidi. Preferibilmente, la sequenza del peptide che deve essere presentata dopo l'elaborazione è derivata dalla sequenza amminoacidica di un antigene, cioè, la sua sequenza corrisponde in modo sostanziale ed è preferibilmente completamente identica a un frammento di un antigene. Pertanto, una MHC legante un peptide comprende una sequenza che

sostanzialmente corrisponde ed è preferibilmente completamente identica a un frammento di un antigene.

I peptidi aventi sequenze amminoacidiche sostanzialmente corrispondenti a una sequenza di un peptide che è presentato da MHC di classe I possono differire per uno o più residui che non sono essenziali per il riconoscimento del TCR del peptide come presentato da MHC di classe I o per il legame peptidico a MHC. Tali peptidi sostanzialmente corrispondenti sono anche in grado di stimolare un CTL reattivo all'antigene e possono essere considerati immunologicamente equivalenti. I peptidi aventi sequenze amminoacidiche diverse da un peptide presentato sui residui che non influenzano il riconoscimento del TCR ma migliorano la stabilità del legame con MHC possono migliorare l'immunogenicità del peptide dell'antigene e possono essere indicati nel presente documento come "peptidi ottimizzati". Utilizzando la conoscenza esistente su questi residui che può avere maggiori probabilità di influenzare il legame all'MHC o al TCR, può essere impiegato un approccio razionale alla progettazione di peptidi sostanzialmente corrispondenti. I peptidi risultanti che sono funzionali sono contemplati come peptidi antigeni.

Un peptide dell'antigene quando presentato da MHC dovrebbe essere riconoscibile da un recettore della cellula T. Preferibilmente, il peptide dell'antigene se riconosciuto da un recettore della cellula T è in grado di indurre in presenza di opportuni segnali di costimolazione, l'espansione clonale della cellula T che trasporta il recettore della cellula T che riconosce il peptide dell'antigene. Preferibilmente, i peptidi dell'antigene, in particolare se presentati nell'ambito di molecole MHC, sono in grado di stimolare una risposta immunitaria, preferibilmente una risposta cellulare contro l'antigene da cui sono derivati o cellule caratterizzate dall'espressione dell'antigene e preferibilmente caratterizzate dalla presentazione dell'antigene. Preferibilmente, un peptide dell'antigene è in grado di stimolare una risposta cellulare contro una cellula caratterizzata dalla presentazione dell'antigene con MHC di classe I e preferibilmente è in grado di stimolare un CTL reattivo all'antigene. Tale cellula preferibilmente è una cellula target.

Il termine "genoma" si riferisce alla quantità totale di informazioni genetiche nei cromosomi di un organismo o di una cellula.

Il termine "esoma" si riferisce a una parte del genoma di un organismo formato da esoni, che codificano porzioni di geni espressi. L'esoma fornisce il modello genetico utilizzato nella sintesi di proteine e altri prodotti genici funzionali. È la parte funzionalmente più rilevante del genoma e, quindi, è più probabile che contribuisca al fenotipo di un organismo. Si stima che l'esoma del genoma umano comprenda l'1,5% del genoma totale (Ng et al., 2008, PLoS Gen., 4(8): 1-15).

Il termine "trascrittoma" si riferisce all'insieme di tutte le molecole di RNA, inclusi mRNA, rRNA, tRNA e altri RNA non codificanti prodotti in una cellula o in una popolazione di cellule. Nel contesto del presente insegnamento per trascrittoma si RNAseq si intende l'insieme di tutte le molecole di RNA prodotte in una cellula, in una popolazione di cellule, preferibilmente una popolazione di cellule cancerose, o tutte le cellule di un

dato individuo in un certo momento.

Un "acido nucleico" è preferibilmente acido desossiribonucleico (DNA) o acido ribonucleico (RNA), più preferibilmente RNA, più preferibilmente RNA trascritto in vitro (IVT RNA) o RNA sintetico. Gli acidi nucleici includono DNA genomico, cDNA, mRNA, molecole prodotte in modo ricombinante e sintetizzate chimicamente. Un acido nucleico può essere presente come una molecola a filamento singolo o a doppio filamento e lineare o chiusa circolarmente in modo covalente. Si può isolare un acido nucleico. Il termine "acido nucleico isolato" significa che l'acido nucleico (i) è stato amplificato in vitro, ad esempio tramite reazione a catena della polimerasi (PCR), (ii) è stato prodotto in modo ricombinante mediante clonazione, (iii) è stato purificato, ad esempio, mediante scissione e separazione mediante elettroforesi su gel, o (iv) è stato sintetizzato, ad esempio, mediante sintesi chimica. Un acido nucleico può essere impiegato per l'introduzione, cioè la trasfezione di, cellule, in particolare, sotto forma di RNA che può essere preparato mediante trascrizione in vitro da uno stampo di DNA. L'RNA può inoltre essere modificato prima dell'applicazione stabilizzando le sequenze, il capping e la poliadenilazione.

Il termine "materiale genetico" si riferisce ad acido nucleico isolato, DNA o RNA, una sezione di una doppia elica, una sezione di un cromosoma, o l'intero genoma di un organismo o di una cellula, in particolare il suo esoma o trascrittoma.

Il termine "mutazione" si riferisce a un cambio o a una differenza nella sequenza di acido nucleico (sostituzione, aggiunta o eliminazione nucleotidica) nel genoma malato rispetto a un riferimento, e preferibilmente un corrispondente genoma normale. Una "mutazione somatica" può verificarsi in qualsiasi cellula del corpo eccetto le cellule germinali (sperma e uovo) e quindi non viene trasmessa ai bambini. Queste alterazioni possono (ma non sempre) causare il cancro o altre malattie. Preferibilmente una mutazione è una mutazione non sinonima. Il termine "mutazione non sinonima" si riferisce a una mutazione, preferibilmente una sostituzione nucleotidica, che si traduce in un cambio amminoacidico come una sostituzione amminoacidica nel prodotto di traslazione, che preferibilmente si traduce nella formazione di un neoepitopo.

Il termine "variante/variazione a singolo nucleotide" (SNV) si riferisce a una differenza nella sequenza dell'acido nucleico in un particolare sito (allele) quando si confronta un genoma di una cellula malata, come una cellula tumorale, e un genoma di una cellula normale, non malata preferibilmente abbinata (corrispondente) o un genoma di riferimento. Come qui utilizzato, il termine mutazione comprende preferibilmente un SNV.

Un evento di variazione del numero di copie (CNV) nel genoma malato (tumore) è un evento di variazione del numero di copie somatico che si verifica solo nelle cellule malate ed è definito come un cambio nel numero di copie degli alleli materni e/o paterni di una regione del genoma malato (tumore) rispetto a un genoma normale abbinato, dove l'alternanza colpisce preferibilmente una regione del genoma che si estende per circa

1 kb o più.

Il termine "mutazione" include mutazioni puntiformi, indel, fusioni, cromotrips e modifiche dell'RNA.

Il termine "indel" descrive una classe di mutazione speciale, definita come una mutazione risultante in un inserimento ed eliminazione colocalizzate e in un guadagno o in una perdita netta di nucleotidi. Nelle regioni codificanti del genoma, a meno che la lunghezza di un indel non sia un multiplo di 3, producono una mutazione frameshift.

Indel può essere contrastato con una mutazione puntiforme, dove un Indel inserisce ed elimina i nucleotidi da una sequenza, una mutazione puntiforme è una forma di sostituzione che sostituisce uno dei nucleotidi.

Le fusioni possono generare geni ibridi formati da due geni precedentemente separati. Può verificarsi come risultato di una traslocazione, eliminazione interstiziale o inversione cromosomica. Spesso i geni di fusione sono oncogeni. I geni di fusione oncogeni possono portare a un prodotto genico con una funzione nuova o diversa dai due partner di fusione. In alternativa, un proto-oncogene è fuso a un forte promotore, e quindi la funzione oncogenica è impostata per funzionare da una sovraregolazione causata dal forte promotore del partner di fusione a monte. I trascritti di fusione oncogena possono anche essere causati da eventi di trans-splicing o read-through.

Il termine "cromotrips" si riferisce a un fenomeno genetico mediante cui regioni specifiche del genoma vengono distrutte e quindi ricostituite tramite un singolo evento devastante.

Il termine "modifica dell'RNA" o "modificazione dell'RNA" si riferisce a processi molecolari in cui il contenuto di informazioni in una molecola di RNA viene alterato attraverso un cambio chimico nella composizione della base. La modifica dell'RNA include modifiche nucleosidiche come la citidina (C) in uridina (U) e l'adenosina (A) in deaminazioni di inosina (I), nonché aggiunte e inserzioni di nucleotidi non modellati. La modifica dell'RNA negli mRNA altera efficacemente la sequenza amminoacidica della proteina codificata in modo che differisca da quella prevista dalla sequenza del DNA genomico.

Il termine "firma di mutazione del cancro" si riferisce a un insieme di mutazioni che sono presenti nelle cellule tumorali rispetto alle cellule di riferimento non cancerose.

Un "riferimento" nel contesto del presente insegnamento può essere utilizzato per correlare e confrontare i risultati ottenuti da un campione di tumore. Tipicamente il "riferimento" può essere ottenuto sulla base di uno o più campioni normali, in particolare campioni che non sono affetti da una malattia cancerosa, ottenuti da un paziente o da uno o più individui diversi, preferibilmente individui sani, in particolare individui della stessa specie. Un "riferimento" può essere determinato empiricamente testando un numero sufficientemente elevato di campioni normali.

Il termine "genoma di riferimento" si riferisce a un genoma che fornisce un sistema di coordinate per il genoma normale e il genoma malato. Un genoma di riferimento viene utilizzato per mappare le letture e fornire un sistema di coordinate per il genoma normale e il genoma del tumore, in cui il sistema di coordinate consente di fornire il numero di cromosomi, una posizione del nucleotide nel cromosoma e la direzionalità della lettura. Un genoma di riferimento può essere basato sul genoma di uno o più elementi della stessa specie del soggetto che fornisce il campione malato, oppure può essere basato sul genoma normale del soggetto (un genoma corrispondente).

Qualsiasi metodo di sequenziamento adatto può essere utilizzato nel contesto del presente insegnamento per identificare mutazioni specifiche della malattia, essendo preferite le tecnologie di sequenziamento di nuova generazione (NGS), e opzionalmente in combinazione con array SNP per ottenere informazioni sul numero assoluto di copie. I metodi di sequenziamento di terza generazione potrebbero sostituire la tecnologia NGS in futuro per accelerare la fase di sequenziamento del metodo. A scopo di chiarimento: i termini "Next Generation Sequencing" o "NGS" nel contesto del presente insegnamento indicano tutte le nuove tecnologie di sequenziamento ad alta produttività che, in contrasto con la metodologia di sequenziamento "convenzionale" nota come chimica di Sanger, leggono i modelli di acido nucleico in modo casuale parallelamente lungo l'intero genoma suddividendo l'intero genoma in piccoli pezzi. Tali tecnologie NGS (note anche come tecnologie di sequenziamento massivamente parallele) sono in grado di fornire informazioni sulla sequenza di acido nucleico di un intero genoma, esoma, trascrittoma (tutte le sequenze trascritte di un genoma) o metiloma (tutte le sequenze metilate di un genoma) in tempi molto brevi, ad esempio entro 1-2 settimane, preferibilmente entro 1-7 giorni o più preferibilmente entro meno di 24 ore e consentono, in linea di principio, approcci di sequenziamento di cellule singole. Più piattaforme NGS disponibili in commercio o menzionate in letteratura possono essere utilizzate nel contesto del presente insegnamento, ad esempio quelle descritte in dettaglio in Zhang et al., 2011, The impact of next-generation sequencing on genomics, *J. Genet Genomica* 38(3):95-109 o in Voelkerding et al., 2009, Next generation sequencing: From basic research to diagnostics, *Clinical chemistry* 55:641-658. Esempi non limitativi di tali tecnologie/piattaforme NGS sono

1) La tecnologia di sequenziamento per sintesi nota come pirosequenziamento implementata ad esempio nel GS-FLX 454 Genome Sequencer™ della società associata a Roche 454 Life Sciences (Branford, Connecticut), descritta per la prima volta in Ronaghi et al., 1998, A sequencing method based on real-time pyrophosphate, *Science* 281:363-365. Questa tecnologia utilizza una PCR in emulsione in cui le sfere di legame al DNA a filamento singolo vengono incapsulate mediante un vigoroso vortex in micelle acquose contenenti reagenti PCR circondate da olio per l'amplificazione PCR in emulsione. Durante il processo di pirosequenziamento, la luce emessa dalle molecole di fosfato durante l'incorporazione dei nucleotidi viene registrata mentre la polimerasi sintetizza il filamento di DNA.

- 2) Gli approcci di sequenziamento per sintesi sviluppati da Solexa (ora parte di Illumina Inc., San Diego, California) basati su terminatori di coloranti reversibili e implementati ad esempio nell'Illumina/Solexa Genome Analyzer™ e nell'Illumina HiSeq 2000 Genome Analyzer™. In questa tecnologia, tutti e quattro i nucleotidi vengono aggiunti simultaneamente in frammenti di cluster oligo-innescati nei canali delle cellule a flusso insieme alla DNA polimerasi. L'amplificazione a ponte estende i filamenti del cluster con tutti e quattro i nucleotidi marcati in modo fluorescente per il sequenziamento.
- 3) Approcci di sequenziamento per legatura, ad esempio implementati nella piattaforma SOLid™ di Applied Biosystems (ora Life Technologies Corporation, Carlsbad, California). In questa tecnologia, un pool di tutti i possibili oligonucleotidi di lunghezza fissa viene etichettato in base alla posizione sequenziata. Gli oligonucleotidi vengono ricotti e ligati; la legatura preferenziale da parte della DNA ligasi per le sequenze corrispondenti si traduce in un segnale informativo del nucleotide in quella posizione. Prima del sequenziamento, il DNA viene amplificato mediante PCR in emulsione. Le sfere risultanti, ciascuna contenente solo copie della stessa molecola di DNA, vengono depositate su un vetrino. Come secondo esempio, la piattaforma Polonator™ G.007 di Dover Systems (Salem, New Hampshire) impiega anche un approccio di sequenziamento per legatura utilizzando una PCR in emulsione a matrice casuale, basata su sfere, per amplificare i frammenti di DNA per il sequenziamento parallelo.
- 4) Tecnologie di sequenziamento a molecola singola, come ad esempio implementate nel sistema PacBio RS di Pacific Biosciences (Menlo Park, California) o nella piattaforma HeliScope™ di Helicos Biosciences (Cambridge, Massachusetts). La caratteristica distintiva di questa tecnologia è la sua capacità di sequenziare singole molecole di DNA o RNA senza amplificazione, definita come sequenziamento del DNA Single-Molecule Real Time (SMRT). Ad esempio, HeliScope utilizza un sistema di rilevamento della fluorescenza altamente sensibile per rilevare direttamente ogni nucleotide mentre viene sintetizzato. Un approccio simile basato sul trasferimento di energia di risonanza di fluorescenza (FRET) è stato sviluppato da Visigen Biotechnology (Houston, Texas). Altre tecniche a molecola singola basate sulla fluorescenza provengono da US Genomics (GeneEngine™) e Genovoxx (AnyGene™).
- 5) Nanotecnologie per il sequenziamento di singole molecole in cui vengono utilizzate varie nanostrutture, ad esempio disposte su un chip per monitorare il movimento di una molecola di polimerasi su un singolo filamento durante la replicazione. Esempi non limitativi di approcci basati su nanotecnologie sono la piattaforma GridON™ di Oxford Nanopore Technologies (Oxford, Regno Unito), le piattaforme di sequenziamento dei nano-pori assistito da ibridazione (HANS™) sviluppate da Nabsys (Providence, Rhode Island) e la piattaforma proprietaria di sequenziamento del DNA basata sulla ligasi con tecnologia DNA nanoball (DNB) chiamata legatura combinatoria sonda-ancoraggio (cPAL™).

6) Tecnologie basate sulla microscopia elettronica per il sequenziamento di singole molecole, ad esempio quelle sviluppate da LightSpeed Genomics (Sunnyvale, California) e Halcyon Molecular (Redwood City, California)

7) Sequenziamento dei semiconduttori ionici basato sulla rilevazione degli ioni di idrogeno che vengono rilasciati durante la polimerizzazione del DNA. Ad esempio, Ion Torrent Systems (San Francisco, California) utilizza un array ad alta densità di pozzi micro-lavorati per eseguire questo processo biochimico in modo massicciamente parallelo. Ciascun pozzetto contiene un modello di DNA diverso. Sotto i pozzetti c'è uno strato sensibile agli ioni e sotto di esso un sensore ionico proprietario.

In una forma di realizzazione, se si è verificata una mutazione specifica della malattia può essere determinata con un metodo relativo a determinare che un sito nel genoma normale è coerente con un genotipo omozigote come riflesso da un allele normale e da tre alleli di rumore e da una distribuzione del rumore ideale e a dichiarare una mutazione in cui il sito corrispondente nel genoma del tumore è incoerente con il genotipo omozigote e una distribuzione del rumore ideale, in cui le letture sono coerenti con una distribuzione del rumore ideale se le letture mappano ciascuno degli alleli del rumore con una probabilità di un terzo di un tasso di errore per base, come descritto nella domanda di brevetto internazionale PCT intitolata "Highly Accurate Mutation Detection, In Particular for Personalized Therapeutics" depositata in data pari alla presente, la cui descrizione è qui incorporata nella sua interezza per riferimento.

Preferibilmente, le preparazioni di DNA e RNA servono come materiale di partenza per NGS. Tali acidi nucleici possono essere facilmente ottenuti da campioni come materiale biologico, ad esempio da tessuti tumorali freschi, congelati istantaneamente o fissati in formalina inclusi in paraffina (FFPE) o da cellule isolate di recente o da CTC presenti nel sangue periferico dei pazienti. Il DNA genomico normale non mutato o l'RNA può essere estratto da tessuto somatico normale, tuttavia le cellule germinali sono preferite nel contesto del presente insegnamento. Il DNA germinale o l'RNA sono estratti dalle cellule mononucleate del sangue periferico (PBMC) in pazienti con neoplasie non ematologiche. Sebbene gli acidi nucleici estratti dai tessuti FFPE o da singole cellule isolate di recente siano altamente frammentati, sono adatti alle applicazioni NGS.

Nella letteratura sono descritti diversi metodi NGS mirati per il sequenziamento dell'esoma (per la revisione vedere, ad esempio, Teer e Mullikin, 2010, Human Mol Genet 19(2):R145-51), che possono essere utilizzati insieme al presente insegnamento. Molti di questi metodi (descritti ad esempio come cattura del genoma, partizionamento del genoma, arricchimento del genoma ecc.) utilizzano tecniche di ibridazione e includono approcci di ibridazione basate su array (ad esempio Hodges et al., 2007, Nat. Genet. 39:1522-1527) e a base liquida (ad esempio Choi et al., 2009, Proc. Natl. Acad. Sci USA 106:19096-19101). Sono disponibili anche kit commerciali per la preparazione del campione di DNA e la successiva cattura dell'esoma: ad esempio, Illumina Inc. (San Diego, California) offre il kit di preparazione del campione di DNA TruSeq™ e il kit di

arricchimento dell'esoma TruSeq™ Exome Enrichment Kit.

Al fine di ridurre il numero di risultati falsi positivi nel rilevare mutazioni somatiche specifiche del cancro o differenze di sequenza quando si confronta, ad esempio, la sequenza di un campione di tumore con la sequenza di un campione di riferimento come la sequenza di un campione di linea germinale, si preferisce determinare la sequenza in repliche di uno o entrambi questi tipi di campioni. Pertanto, è preferibile che la sequenza di un campione di riferimento come la sequenza di un campione di linea germinale sia determinata due volte, tre volte o più. In alternativa o in aggiunta, la sequenza di un campione di tumore viene determinata due, tre o più volte. Può anche essere possibile determinare la sequenza di un campione di riferimento come la sequenza di un campione di linea germinale e/o la sequenza di un campione di tumore più di una volta determinando almeno una volta la sequenza nel DNA genomico e determinando almeno una volta la sequenza in RNA di detto campione di riferimento e/o di detto campione di tumore. Ad esempio, determinando le variazioni tra le repliche di un campione di riferimento come un campione della linea germinale, è possibile stimare il tasso atteso di mutazioni somatiche di falsi positivi (FDR) come quantità statistica. Le ripetizioni tecniche di un campione dovrebbero generare risultati identici e qualsiasi mutazione rilevata in questo "confronto stesso vs stesso" è un falso positivo. In particolare, per determinare il tasso di false scoperte per il rilevamento di mutazioni somatiche in un campione di tumore rispetto a un campione di riferimento, può essere utilizzata una ripetizione tecnica del campione di riferimento come riferimento per stimare il numero di falsi positivi. Inoltre, varie metriche relative alla qualità (ad esempio copertura o qualità SNP) possono essere combinate in un unico punteggio di qualità utilizzando un approccio di apprendimento automatico. Opzionalmente, per una data variazione somatica possono essere conteggiate tutte le altre variazioni con un punteggio di qualità superiore, il che consente una classificazione di tutte le variazioni in un set di dati.

Nel contesto del presente insegnamento, il termine "RNA" si riferisce a una molecola che comprende residui ribonucleotidici ed è preferibilmente composta interamente o sostanzialmente da residui ribonucleotidici. "Ribonucleotide" si riferisce a un nucleotide con un gruppo idrossile nella posizione 2' di un gruppo β -D-ribofuranosile. Il termine "RNA" comprende RNA a doppio filamento, RNA a filamento singolo, RNA isolato come RNA parzialmente o completamente purificato, RNA essenzialmente puro, RNA sintetico e RNA generato in modo ricombinante come RNA modificato che differisce dall'RNA naturale per aggiunta, eliminazione, sostituzione e/o alterazione di uno o più nucleotidi. Tali alterazioni possono includere l'aggiunta di materiale non nucleotidico, come all'estremità o alle estremità di un RNA o internamente, ad esempio, su uno o più nucleotidi dell'RNA. I nucleotidi nelle molecole di RNA possono anche comprendere nucleotidi non standard, come nucleotidi non presenti in natura o nucleotidi o deossinucleotidi sintetizzati chimicamente. Questi RNA alterati possono essere indicati come analoghi o analoghi dell'RNA naturale.

Il termine "RNA" include e si riferisce preferibilmente a "mRNA". Il termine "mRNA" significa "RNA messaggero" e si riferisce a un "trascritto" che viene generato utilizzando uno stampo di DNA e codifica un peptide o polipeptide. Tipicamente, un mRNA comprende un 5'-UTR, una regione codificante della proteina e un 3'-UTR. L'mRNA possiede solo un'emivita limitata nelle cellule e in vitro. Nel contesto del presente insegnamento, l'mRNA può essere generato mediante trascrizione in vitro da uno stampo di DNA. La metodologia di trascrizione in vitro è nota alla persona esperta. Ad esempio, è disponibile in commercio una varietà di kit di trascrizione in vitro.

La stabilità e l'efficienza di traslazione dell'RNA possono essere modificate secondo necessità. Ad esempio, l'RNA può essere stabilizzato e la sua traslazione aumentata da una o più modifiche aventi effetti stabilizzanti e/o aumentando l'efficienza di traduzione dell'RNA. Tali modifiche sono descritte, ad esempio, in PCT/EP2006/009448. Per aumentare l'espressione dell'RNA utilizzato nelle forme di realizzazione del presente insegnamento, esso può essere modificato all'interno della regione codificante, cioè la sequenza che codifica il peptide o la proteina espressi, preferibilmente senza alterare la sequenza del peptide o della proteina espressi, in modo da aumentare il contenuto di GC per aumentare la stabilità dell'mRNA e per eseguire un'ottimizzazione del codone e, quindi, migliorare la traslazione nelle cellule.

Il termine "modifica" nel contesto dell'RNA come utilizzato secondo il presente insegnamento include qualsiasi modifica dell'RNA che non è naturalmente presente in detto RNA.

In una forma di realizzazione dell'insegnamento, l'RNA utilizzato secondo l'insegnamento non presenta 5'-trifosfati non coperti. La rimozione di tali 5'-trifosfati non coperti può essere ottenuta trattando l'RNA con una fosfatasi.

L'RNA secondo l'insegnamento può avere ribonucleotidi naturali o sintetici modificati per aumentarne la stabilità e/o diminuire la citotossicità. Ad esempio, in una forma di realizzazione, nell'RNA utilizzato secondo l'insegnamento, la 5-metilcitosina è sostituita parzialmente o completamente, preferibilmente completamente, alla citosina. In alternativa o in aggiunta, in una forma di realizzazione, nell'RNA utilizzato secondo l'insegnamento la pseudouridina è sostituita parzialmente o completamente, preferibilmente completamente, all'uridina.

In una forma di realizzazione, il termine "modifica" si riferisce alla fornitura di un RNA con un rivestimento di 5' o un analogo di rivestimento di 5'.

Il termine "rivestimento di 5'" si riferisce a una struttura di rivestimento che si trova all'estremità 5' di una molecola di mRNA e generalmente consiste in un nucleotide guanosico collegato all'mRNA tramite un insolito legame trifosfato da 5' a 5'. In una forma di realizzazione, questa guanosina è metilata nella posizione 7. Il termine "rivestimento di 5' convenzionale" si riferisce a un rivestimento di 5' di RNA presente in natura, preferibilmente al rivestimento di 7-metilguanosina (m7G). Nel contesto del presente insegnamento, il termine "rivestimento di 5'" include un analogo di rivestimento di 5' che assomiglia alla struttura del rivestimento dell'RNA ed è modificato per possedere la capacità di stabilizzare l'RNA

se collegato a esso, preferibilmente in vivo e/o in una cellula.

Fornire un RNA con un rivestimento di 5' o un analogo di rivestimento di 5' può essere ottenuto mediante trascrizione in vitro di un modello di DNA in presenza di detto rivestimento di 5' o un analogo di rivestimento di 5', in cui detto rivestimento di 5' è cotrascrizionalmente incorporato nel filamento di RNA generato, oppure l'RNA può essere generato, ad esempio, mediante trascrizione in vitro, e il rivestimento di 5' può essere attaccato all'RNA post-trascrizionalmente utilizzando enzimi di copertura, ad esempio, enzimi di copertura di virus di vaiolo bovino.

L'RNA può comprendere ulteriori modifiche. Ad esempio, un'ulteriore modifica dell'RNA utilizzato nel presente insegnamento può essere un'estensione o un troncamento della coda di poli(A) naturale o un'alterazione delle regioni 5' o 3' non traslate (UTR) come l'introduzione di un UTR che non è correlato alla regione codificante di detto RNA, ad esempio, lo scambio del 3'-UTR esistente con o l'inserimento di una o più, preferibilmente due copie di un 3'-UTR derivato da un gene globinico, come alfa2-globina, alfa1-globina, beta-globina, preferibilmente beta-globina, più preferibilmente beta-globina umana.

L'RNA con una sequenza poli-A non mascherata viene traslato in modo più efficiente rispetto all'RNA con una sequenza poli-A mascherata. Il termine "coda poli(A)" o "sequenza poli-A" si riferisce a una sequenza di residui di adenile (A) che si trova tipicamente all'estremità 3' di una molecola di RNA e "sequenza poli-A non mascherata" significa che la sequenza poli-A all'estremità 3' di una molecola di RNA termina con una A della sequenza poli-A e non è seguita da nucleotidi diversi da A situati all'estremità 3', cioè a valle, della sequenza poli-A. Inoltre, una lunga sequenza poli-A di circa 120 coppie di basi si traduce in una stabilità ottimale del trascritto e in un'efficienza di traslazione dell'RNA.

Pertanto, al fine di aumentare la stabilità e/o l'espressione dell'RNA utilizzato secondo il presente insegnamento, può essere modificato in modo da essere presente insieme a una sequenza poli-A, preferibilmente avente una lunghezza da 10 a 500, più preferibilmente da 30 a 300, ancor più preferibilmente da 65 a 200 e specialmente da 100 a 150 residui di adenosina. In una forma di realizzazione particolarmente preferita, la sequenza poli-A ha una lunghezza di circa 120 residui di adenosina. Per aumentare ulteriormente la stabilità e/o l'espressione dell'RNA utilizzato secondo l'insegnamento, la sequenza poli-A può essere smascherata.

Inoltre, l'incorporazione di due o più regioni 3' non traslate (UTR) nella regione 3' non traslata di una molecola di RNA può comportare un miglioramento dell'efficienza di traslazione. Un effetto sinergico può essere ottenuto incorporando due o più di tali regioni 3' non traslate. Le regioni 3' non traslate possono essere autologhe o eterologhe rispetto all'RNA in cui vengono introdotte. In una particolare forma di realizzazione, la regione 3' non traslata è derivata dal gene della β -globina umana.

Una combinazione delle modificazioni sopra descritte, ovvero l'incorporazione di una sequenza poli-A, lo smascheramento di una sequenza poli-A

e l'incorporazione di una o più regioni 3' non traslate, ha un'influenza sinergica sulla stabilità dell'RNA e sull'aumento dell'efficienza di traslazione.

Il termine "stabilità" dell'RNA si riferisce alla "emivita" dell'RNA. "Emivita" si riferisce al periodo di tempo necessario per eliminare metà dell'attività, di quantità o di numero di molecole. Nel contesto del presente insegnamento, l'emivita di un RNA è indicativa per la stabilità di detto RNA. L'emivita dell'RNA può influenzare la "durata dell'espressione" dell'RNA. Ci si può aspettare che l'RNA con una lunga emivita venga espresso per un periodo di tempo prolungato.

Naturalmente, se si desidera diminuire la stabilità e/o l'efficienza di traslazione dell'RNA, è possibile modificare l'RNA in modo da interferire con la funzione degli elementi come descritto sopra aumentando la stabilità e/o l'efficienza di traslazione dell'RNA.

Il termine "espressione" è utilizzato nel suo significato più generale e comprende la produzione di RNA e/o peptidi o polipeptidi, ad esempio mediante trascrizione e/o traslazione. Per quanto riguarda l'RNA, il termine "espressione" o "traslazione" si riferisce in particolare alla produzione di peptidi o polipeptidi. Comprende inoltre l'espressione parziale di acidi nucleici. Inoltre, l'espressione può essere transitoria o stabile.

Il termine espressione include anche una "espressione anomala" o "espressione anormale". "Espressione anomala" o "espressione anormale" significa che l'espressione è alterata, preferibilmente aumentata, rispetto a un riferimento, ad esempio uno stato in un soggetto che non presenta una malattia associata all'espressione anomala o anormale di una certa proteina, ad esempio, un antigene del tumore. Un aumento dell'espressione si riferisce a un aumento almeno del 10%, in particolare almeno del 20%, almeno del 50% o almeno del 100%, o più. In una forma di realizzazione, l'espressione si trova solo in un tessuto malato, mentre l'espressione in un tessuto sano viene repressa.

Il termine "espresso in modo specifico" significa che una proteina è essenzialmente espressa solo in uno specifico tessuto o organo. Ad esempio, un antigene tumorale espresso in modo specifico nella mucosa gastrica significa che detta proteina è espressa principalmente nella mucosa gastrica e non è espressa in altri tessuti o non è espressa in misura significativa in altri tipi di tessuti o organi. Pertanto, una proteina espressa esclusivamente nelle cellule della mucosa gastrica e in misura significativamente minore in qualsiasi altro tessuto, come quello dei testicoli, è espressa in modo specifico nelle cellule della mucosa gastrica. In alcune forme di realizzazione, un antigene tumorale può anche essere specificamente espresso in condizioni normali in più di un tipo di tessuto o organo, come in 2 o 3 tipi di tessuto o organi, ma preferibilmente in non più di 3 diversi tipi di tessuto o organo. In questo caso, l'antigene tumorale viene quindi espresso in modo specifico in questi organi. Ad esempio, se un antigene tumorale è espresso in condizioni normali preferibilmente in misura approssimativamente uguale nel polmone e nello stomaco, detto antigene tumorale è espresso in modo specifico nel polmone e nello stomaco.

Nel contesto del presente insegnamento, il termine "trascrizione" si riferisce a un processo, in cui il codice genetico in una sequenza di DNA viene

trascritto in RNA. Successivamente, l'RNA può essere traslato in proteina. Secondo il presente insegnamento, il termine "trascrizione" comprende "trascrizione in vitro", in cui il termine "trascrizione in vitro" si riferisce a un processo in cui l'RNA, in particolare l'mRNA, è sintetizzato in vitro in un sistema privo di cellule, preferibilmente utilizzando estratti cellulari appropriati. Preferibilmente, i vettori di clonazione vengono applicati per la generazione di trascritti. Questi vettori di clonazione sono generalmente designati come vettori di trascrizione e sono compresi dal termine "vettore". L'RNA utilizzato nel presente insegnamento è preferibilmente RNA trascritto in vitro (IVT-RNA) e può essere ottenuto mediante trascrizione in vitro di uno stampo di DNA appropriato. Il promotore per il controllo della trascrizione può essere qualsiasi promotore per qualsiasi RNA polimerasi. Esempi particolari di RNA polimerasi sono le RNA polimerasi T7, T3 e SP6. Preferibilmente, la trascrizione in vitro è controllata da un promotore T7 o SP6. Un modello di DNA per la trascrizione in vitro può essere ottenuto clonando un acido nucleico, in particolare cDNA, e introducendolo in un vettore appropriato per la trascrizione in vitro. Il cDNA può essere ottenuto mediante trascrizione inversa dell'RNA.

Il termine "traslazione" si riferisce al processo nei ribosomi di una cellula mediante cui un filamento di RNA messaggero dirige l'assieme di una sequenza di amminoacidi per formare un peptide o un polipeptide.

Le sequenze di controllo dell'espressione o sequenze regolatrici, che nel contesto del presente insegnamento possono essere collegate in modo funzionale con un acido nucleico, possono essere omologhe o eterologhe rispetto all'acido nucleico. Una sequenza codificante e una sequenza regolatrice sono collegate insieme "funzionalmente" se sono legate insieme in modo covalente, in modo che la trascrizione o traslazione della sequenza codificante sia sotto il controllo o sotto l'influenza della sequenza regolatrice. Se la sequenza codificante deve essere tradotta in una proteina funzionale, con collegamento funzionale di una sequenza regolatrice con la sequenza codificante, l'induzione della sequenza regolatrice porta a una trascrizione della sequenza codificante, senza causare uno spostamento del frame di lettura nella sequenza codificante o incapacità della sequenza codificante di essere tradotta nella proteina o nel peptide desiderato.

Il termine "sequenza di controllo dell'espressione" o "sequenza regolatrice" comprende, nel contesto dell'insegnamento, promotori, sequenze leganti i ribosomi e altri elementi di controllo, che controllano la trascrizione di un acido nucleico o la traslazione di un RNA derivato. In certe forme di realizzazione dell'insegnamento, le sequenze regolatrici possono essere controllate. La struttura esatta delle sequenze regolatrici può variare in funzione della specie o del tipo di cellula, ma generalmente comprende sequenze 5' non trascritte e sequenze 5' e 3' non traslate che sono coinvolte rispettivamente nell'inizio della trascrizione e della traslazione, come TATA box, sequenza di rivestimento, sequenza CAAT e simili. Più specificatamente, le sequenze regolatrici 5' non trascritte comprendono una regione del promotore che include una sequenza del promotore per il controllo trascrizionale dell'acido nucleico collegato in modo funzionale. Le sequenze regolatrici possono anche comprendere sequenze di

potenziamento o sequenze di attivatori a monte.

Preferibilmente, l'RNA che deve essere espresso in una cellula viene introdotto in detta cellula. In una forma di realizzazione dei metodi secondo l'insegnamento, l'RNA che deve essere introdotto in una cellula è ottenuto mediante trascrizione in vitro di un modello di DNA appropriato.

Termini come "RNA in grado di esprimere" e "RNA codificante" sono utilizzati in modo intercambiabile nel presente documento e rispetto a un particolare peptide o polipeptide significano che l'RNA, se presente nell'ambiente appropriato, preferibilmente all'interno di una cellula, può essere espresso per produrre detto peptide o polipeptide. Preferibilmente, l'RNA è in grado di interagire con il macchinario di traslazione cellulare per fornire il peptide o il polipeptide che è in grado di esprimere.

Termini come "trasferimento", "introduzione" o "trasfezione" sono qui usati in modo intercambiabile e si riferiscono all'introduzione di acidi nucleici, in particolare acidi nucleici esogeni o eterologhi, in particolare RNA in una cellula. Secondo il presente insegnamento, la cellula può far parte di un organo, un tessuto e/o un organismo. Secondo il presente insegnamento, la somministrazione di un acido nucleico viene ottenuta come acido nucleico nudo o in combinazione con un reagente di somministrazione. Preferibilmente, la somministrazione di acidi nucleici è sotto forma di acidi nucleici nudi. Preferibilmente, l'RNA viene somministrato in combinazione con sostanze stabilizzanti come gli inibitori della RNasi. Il presente insegnamento prevede anche l'introduzione ripetuta di acidi nucleici nelle cellule per consentire un'espressione sostenuta per periodi di tempo prolungati.

Le cellule possono essere trasfettate con qualsiasi vettore a cui l'RNA può essere associato, ad esempio formando complessi con l'RNA o formando vescicole in cui l'RNA è racchiuso o incapsulato, con conseguente maggiore stabilità dell'RNA rispetto all'RNA nudo. I vettori utili includono, ad esempio, vettori contenenti lipidi come lipidi cationici, liposomi, in particolare liposomi cationici, micelle e nanoparticelle. I lipidi cationici possono formare complessi con acidi nucleici caricati negativamente. Può essere utilizzato qualsiasi lipide cationico.

Preferibilmente, l'introduzione di RNA che codifica un peptide o polipeptide in una cellula, in particolare in una cellula presente in vivo, determina l'espressione di detto peptide o polipeptide nella cellula. In particolari forme di realizzazione, è preferito il targeting degli acidi nucleici a particolari cellule. In tali forme di realizzazione, un vettore che viene applicato per la somministrazione dell'acido nucleico a una cellula (per esempio, un retrovirus o un liposoma), mostra una molecola di targeting. Ad esempio, può essere incorporata una molecola come un anticorpo che è specifico per una proteina superficiale della membrana sulla cellula target o un ligando per un recettore sulla cellula target nel trasportatore di acido nucleico o può essere legata a esso. Nel caso in cui l'acido nucleico venga somministrato dai liposomi, le proteine che si legano a una proteina della membrana superficiale associata all'endocitosi possono essere incorporate nella formulazione del liposoma per consentire il targeting e/o

l'assorbimento. Tali proteine comprendono proteine del capsido di suoi frammenti che sono specifici per un particolare tipo di cellula, anticorpi contro proteine che sono interiorizzate, proteine che prendono di mira una posizione intracellulare ecc.

Il termine "cellula" o "cellula ospite" si riferisce preferibilmente a una cellula intatta, cioè una cellula con una membrana intatta che non ha rilasciato i suoi normali componenti intracellulari come enzimi, organelli o materiale genetico. Una cellula intatta è preferibilmente una cellula vitale, cioè una cellula vivente in grado di svolgere le sue normali funzioni metaboliche. Preferibilmente, detto termine si riferisce a qualsiasi cellula che può essere trasformata o trasfettata con un acido nucleico esogeno. Il termine "cellula" include cellule procariotiche (ad esempio, E. coli) o cellule eucariotiche (ad esempio, cellule dendritiche, cellule B, cellule CHO, cellule COS, cellule K562, cellule HEK293, cellule HELA, cellule di lievito e cellule di insetti). L'acido nucleico esogeno può essere trovato all'interno della cellula (i) liberamente disperse come tale, (ii) incorporato in un vettore ricombinante, o (iii) integrato nel genoma della cellula ospite o nel DNA mitocondriale. Le cellule di mammiferi sono particolarmente preferite per il trasferimento adottivo, come le cellule di esseri umani, topi, criceti, maiali, capre e primati. Le cellule possono essere derivate da un gran numero di tipi di tessuto e includono cellule primarie e linee cellulari. Esempi specifici includono cheratinociti, leucociti del sangue periferico, cellule staminali del midollo osseo e cellule staminali embrionali. In una forma di realizzazione preferita, la cellula è una cellula che presenta l'antigene, in particolare una cellula dendritica, un monocita o un macrofago.

Una cellula che comprende una molecola di acido nucleico esprime preferibilmente il peptide o la proteina codificata dall'acido nucleico.

Il termine "espansione clonale" si riferisce a un processo in cui un'entità specifica viene moltiplicata. Nel contesto del presente insegnamento, il termine è preferibilmente utilizzato nel contesto di una risposta immunologica in cui i linfociti sono stimolati da un antigene, proliferano e il linfocita specifico che riconosce detto antigene è amplificato. Preferibilmente, l'espansione clonale porta alla differenziazione dei linfociti.

Termini come "ridurre" o "inibire" come utilizzati nel presente documento significano la capacità di causare una diminuzione complessiva nel livello preferibilmente del 5% o più, del 10% o più, del 20% o più, più preferibilmente del 50% o più e più preferibilmente del 75% o maggiore. Il termine "inibire" o frasi simili includono un'inibizione completa o essenzialmente completa, cioè una riduzione a zero o essenzialmente a zero.

Termini come "aumentare" o "potenziare" si riferiscono preferibilmente a un aumento o miglioramento di circa almeno il 10%, preferibilmente almeno il 20%, preferibilmente almeno il 30%, preferibilmente almeno il 40%, preferibilmente almeno il 50%, preferibilmente almeno l'80%, preferibilmente almeno il 100%, preferibilmente almeno il 200% e in particolare almeno il 300%. Questi termini possono anche riferirsi a un aumento, miglioramento, promozione o prolungamento da zero o da un livello non misurabile o non rilevabile a un livello superiore a zero o un livello che è misurabile o rilevabile.

Il termine "peptide" secondo l'insegnamento si riferisce a sostanze comprendenti due o più, preferibilmente 3 o più, preferibilmente 4 o più, preferibilmente 6 o più, preferibilmente 8 o più, preferibilmente 10 o più, preferibilmente 13 o più, preferibilmente 16 o più, preferibilmente 21 o più e preferibilmente fino a 8, 10, 20, 30, 40 o 50, in particolare 100 amminoacidi uniti in modo covalente da legami peptidici. Il termine "polipeptide" o "proteina" si riferisce a grandi peptidi, preferibilmente a peptidi con più di 100 residui amminoacidici, ma in generale i termini "peptidi", "polipeptidi" e "proteine" sono sinonimi e vengono utilizzati nel presente documento in modo intercambiabile. Secondo l'insegnamento, il termine "modificazione" o "cambio di sequenza" rispetto a peptidi, polipeptidi o proteine si riferisce a un cambio di sequenza in un peptide, polipeptide o proteina rispetto a una sequenza parentale come la sequenza di un peptide, polipeptide o proteina wild-type. Il termine include varianti di inserimento di amminoacidi, varianti di aggiunta di amminoacidi, varianti di eliminazione di amminoacidi e varianti di sostituzione di amminoacidi, preferibilmente varianti di sostituzione di amminoacidi. Tutti questi cambiamenti di sequenza secondo l'insegnamento possono potenzialmente creare nuovi epitopi.

Le varianti di inserimento di amminoacidi comprendono inserzioni di uno o due o più amminoacidi in una particolare sequenza amminoacidica.

Le varianti di addizione di amminoacidi comprendono fusioni amminoacidiche e/o carbossiterminali di uno o più amminoacidi, come 1, 2, 3, 4, 5 o più amminoacidi.

Le varianti di eliminazione di amminoacidi sono caratterizzate dalla rimozione di uno o più amminoacidi dalla sequenza, ad esempio mediante rimozione di 1, 2, 3, 4, 5 o più amminoacidi.

Le varianti di sostituzione degli amminoacidi sono caratterizzate dal fatto che almeno un residuo nella sequenza viene rimosso e un altro residuo viene inserito al suo posto.

Secondo l'insegnamento, una modifica o un peptide modificato utilizzato per testare nei metodi dell'insegnamento può essere derivato da una proteina comprendente una modifica.

Il termine "derivato" significa secondo l'insegnamento che una particolare entità, in particolare una particolare sequenza di peptidi, è presente nell'oggetto da cui è derivata. Nel caso di sequenze amminoacidiche, in particolare regioni di sequenza particolari, "derivato" in particolare significa che la sequenza amminoacidica rilevante è derivata da una sequenza amminoacidica in cui è presente.

Gli agenti, le composizioni e i metodi qui descritti possono essere utilizzati per trattare un soggetto con una malattia, ad esempio, una malattia caratterizzata dalla presenza di cellule malate che esprimono un antigene e che presentano un peptide dell'antigene. Le malattie particolarmente preferite sono le malattie del cancro. Gli agenti, le composizioni e i metodi descritti nel presente documento possono anche essere utilizzati per

l'immunizzazione o la vaccinazione per prevenire una malattia qui descritta.

Un tale agente è un vaccino come un vaccino contro il cancro progettato sulla base di neoepitopi adatti che resistono alla fuga immunitaria identificati dai metodi del presente insegnamento.

Secondo l'insegnamento, il termine "vaccino" si riferisce a una preparazione farmaceutica (composizione farmaceutica) o un prodotto che in seguito alla somministrazione induce una risposta immunitaria, in particolare una risposta immunitaria cellulare, che riconosce e attacca un agente patogeno o una cellula malata come una cellula cancerosa. Un vaccino può essere utilizzato per la prevenzione o il trattamento di una malattia. Il termine "vaccino anticancro personalizzato" o "vaccino anticancro individualizzato" riguarda un particolare malato di cancro e significa che un vaccino anticancro è adattato alle esigenze o circostanze speciali di un singolo malato di cancro.

I vaccini anticancro forniti secondo l'insegnamento quando somministrati a un brevetto forniscono uno o più epitopi della cellula T adatti a stimolare, innescare e/o espandere le cellule T specifiche per il tumore del paziente. Le cellule T sono preferibilmente dirette contro le cellule che esprimono antigeni da cui derivano gli epitopi delle cellule T. Perciò, i vaccini descritti nel presente documento sono preferibilmente in grado di indurre o promuovere una risposta cellulare, preferibilmente l'attività della cellula T citotossica, contro una malattia del cancro caratterizzata dalla presentazione di uno o più neoantigeni associati al tumore di classe I. Poiché un vaccino qui fornito avrà come target mutazioni specifiche del cancro, sarà specifico per il tumore del paziente.

Nel contesto del presente insegnamento, un vaccino si riferisce a un vaccino che quando somministrato a un paziente fornisce preferibilmente uno o più epitopi della cellula T (neoepitopi, neoepitopi adatti, combinazione di neoepitopi adatti identificati nel presente documento), come 2 o più, 5 o più, 10 o più, 15 o più, 20 o più, 25 o più, 30 o più e preferibilmente fino a 60, fino a 55, fino a 50, fino a 45, fino a 40, fino a 35 o fino a 30 epitopi della cellula T, che incorporano modificazioni di amminoacidi o peptidi modificati predetti come epitopi adatti. La presentazione di questi epitopi da parte delle cellule di un paziente, in particolare le cellule che presentano l'antigene, dà luogo preferibilmente a cellule T che prendono di mira gli epitopi quando sono legate a MHC e quindi, il tumore del paziente, preferibilmente il tumore primario così come le metastasi tumorali, che esprimono antigeni da cui gli epitopi delle cellule T sono derivati e presentano gli stessi epitopi sulla superficie delle cellule tumorali.

I metodi dell'insegnamento possono comprendere l'ulteriore fase di determinare l'utilizzabilità delle modificazioni amminoacidiche identificate o dei peptidi modificati contenenti un neoepitopo adatto identificato nel presente documento per la vaccinazione anticancro. Le fasi successive possono coinvolgere una o più delle seguenti: (i) valutare se le modifiche si trovano in epitopi presentati da MHC noti o previsti, (ii) test in vitro e/o in silico se le modifiche si trovano in epitopi presentati da MHC, ad esempio testare se le modifiche fanno parte di sequenze peptidiche che sono

trasformate in e/o presentate come epitopi presentati da MHC e (iii) test in vitro se gli epitopi modificati previsti, in particolare quando presenti nel loro contesto di sequenza naturale, ad esempio quando sono affiancati da sequenze di amminoacidi che affiancano anche detti epitopi nella proteina naturale e quando espresse in cellule presentanti l'antigene sono in grado di stimolare le cellule T come le cellule T del paziente aventi la specificità desiderata. Tali sequenze di affiancamento possono ciascuna comprendere 3 o più, 5 o più, 10 o più, 15 o più, 20 o più e preferibilmente fino a 50, fino a 45, fino a 40, fino a 35 o fino a 30 amminoacidi e possono affiancare la sequenza epitopica N-terminale e/o C-terminale.

I peptidi modificati determinati secondo l'insegnamento possono essere classificati in base alla loro utilizzabilità come epitopi per la vaccinazione contro il cancro. Pertanto, in un aspetto, il metodo di insegnamento comprende un processo analitico manuale o basato su computer in cui i peptidi modificati identificati vengono analizzati e selezionati per la loro utilizzabilità nel rispettivo vaccino da fornire. In una forma di realizzazione preferita, detto processo analitico è un processo basato su algoritmo computazionale. Preferibilmente, detto processo analitico comprende determinare e/o classificare gli epitopi secondo una previsione della loro capacità di essere immunogenici.

I neoepitopi identificati secondo l'insegnamento e forniti in un vaccino sono preferibilmente presenti sotto forma di un polipeptide comprendente detti epitopi (neoepitopi, neoepitopi adatti, neoepitopi trovati nella combinazione di neoepitopi adatti identificati nel presente documento) come un polipeptide poliepitopico o un acido nucleico, in particolare RNA, che codifica detto polipeptide. Inoltre, i neoepitopi possono essere presenti nel polipeptide sotto forma di una sequenza vaccinale, cioè presenti nel loro contesto di sequenza naturale, ad esempio affiancati da sequenze amminoacidiche che affiancano anche detti epitopi nella proteina presente in natura. Tali sequenze di affiancamento possono ciascuna comprendere 5 o più, 10 o più, 15 o più, 20 o più e preferibilmente fino a 50, fino a 45, fino a 40, fino a 35 o fino a 30 amminoacidi e possono affiancare la sequenza epitopica N-terminale e/o C-terminale. Pertanto, una sequenza di vaccino può comprendere 20 o più, 25 o più, 30 o più, 35 o più, 40 o più e preferibilmente fino a 50, fino a 45, fino a 40, fino a 35 o fino a 30 amminoacidi. In una forma di realizzazione, gli epitopi e/o le sequenze del vaccino sono allineati nel polipeptide dalla testa alla coda.

In una forma di realizzazione, gli epitopi/neoepitopi adatti identificati nel presente documento e/o le sequenze del vaccino sono distanziate da linker, in particolare da linker neutri. Il termine "linker" utilizzato nel contesto del presente insegnamento si riferisce a un peptide aggiunto tra due domini peptidici come epitopi o sequenze di vaccino per collegare detti domini peptidici. Non vi è alcuna limitazione particolare per quanto riguarda la sequenza del linker. Tuttavia, è preferibile che la sequenza del linker riduca l'ostacolo sterico tra i due domini peptidici, sia ben tradotta e supporti o consenta l'elaborazione degli epitopi. Inoltre, il linker non dovrebbe avere o dovrebbe avere solo pochi elementi di sequenza immunogenica. I linker preferibilmente non dovrebbero creare epitopi non endogeni come quelli generati dalla sutura di giunzione tra epitopi

adiacenti, che potrebbero generare reazioni immunitarie indesiderate. Pertanto, il vaccino polipitopico dovrebbe preferibilmente contenere sequenze di linker che sono in grado di ridurre il numero di epitopi di giunzione di legame MHC indesiderati. Hoyt et al. (EMBO J. 25(8), 1720-9, 2006) e Zhang et al. (J. Biol. Chem., 279 (10), 8635-41, 2004) hanno dimostrato che le sequenze ricche di glicina compromettono l'elaborazione proteosomiale e quindi l'uso di sequenze di linker ricche di glicina agisce per ridurre al minimo il numero di peptidi contenuti nel linker che possono essere processati dal proteasoma. Inoltre, è stato osservato che la glicina inibisce un forte legame nelle posizioni del solco di legame dell'MHC (Abastado et al., 1993, J. Immunol. 151(7):3569-75). Schlessinger et al., 2005, Proteins 61(1): 115-26 avevano scoperto che gli amminoacidi glicina e serina inclusi in una sequenza amminoacidica si traducono in una proteina più flessibile che viene traslata ed elaborata in modo più efficiente dal proteasoma, consentendo un migliore accesso agli epitopi codificati. Ciascun linker può comprendere 3 o più, 6 o più, 9 o più, 10 o più, 15 o più, 20 o più e preferibilmente fino a 50, fino a 45, fino a 40, fino a 35 o fino a 30 amminoacidi. Preferibilmente il linker è arricchito in amminoacidi glicina e/o serina. Preferibilmente, almeno il 50%, almeno il 60%, almeno il 70%, almeno l'80%, almeno il 90% o almeno il 95% degli amminoacidi del linker sono glicina e/o serina. In una forma di realizzazione preferita, un linker è sostanzialmente composto dagli amminoacidi glicina e serina. In una forma di realizzazione, il linker comprende la sequenza amminoacidica (GGS)a(GSS)b(GGG)c(SSG)d(GSG)e in cui a, b, c, d ed e sono indipendentemente un numero selezionato da 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 e in cui a + b + c + d + e sono diversi da 0 e preferibilmente sono 2 o più, 3 o più, 4 o più o 5 o più. In una forma di realizzazione, il linker comprende una sequenza come qui descritta includendo le sequenze di linker descritte negli esempi come la sequenza GGSGGGGSG.

In una forma di realizzazione particolarmente preferita, un polipeptide che incorpora uno o più neoepitopi adatti identificati dai metodi del presente documento, come un polipeptide poliepitopico, viene somministrato a un paziente sotto forma di un acido nucleico, preferibilmente RNA come RNA trascritto in vitro o sintetico, che può essere espresso nelle cellule di un paziente come le cellule che presentano l'antigene per produrre il polipeptide. Il presente insegnamento prevede anche la somministrazione di uno o più polipeptidi multiepitopici che ai fini del presente insegnamento sono compresi dal termine "polipeptide poliepitopico", preferibilmente sotto forma di un acido nucleico, preferibilmente RNA come RNA trascritto in vitro o sintetico, che può essere espresso nelle cellule di un paziente come le cellule che presentano l'antigene per produrre uno o più polipeptidi. In caso di somministrazione di più di un polipeptide multiepitopico i neoepitopi adatti forniti dai diversi polipeptidi multiepitopici possono essere differenti o parzialmente sovrapposti. Una volta presente nelle cellule di un paziente come le cellule che presentano l'antigene, il polipeptide secondo l'insegnamento viene elaborato per produrre i neoepitopi identificati secondo l'insegnamento. La somministrazione di un vaccino fornito secondo l'insegnamento può fornire epitopi di classe II presentati da MHC che sono in grado di suscitare una risposta delle cellule T

helper CD4+ contro le cellule che esprimono antigeni da cui derivano gli epitopi presentati da MHC. In alternativa o in aggiunta, la somministrazione di un vaccino fornito secondo l'insegnamento può fornire neoepitopi presentati da MHC di classe I che sono in grado di suscitare una risposta della cellula T CD8+ contro cellule che esprimono antigeni da cui derivano neoepitopi presentati da MHC. Inoltre, la somministrazione di un vaccino fornito secondo l'insegnamento può fornire uno o più neoepitopi (inclusi neoepitopi noti e neoepitopi adatti identificati secondo l'insegnamento) nonché uno o più epitopi non contenenti mutazioni somatiche specifiche del cancro ma essendo espresse da cellule cancerose e preferibilmente inducendo una risposta immunitaria contro le cellule cancerose, preferibilmente una risposta immunitaria specifica per il cancro. In una forma di realizzazione, la somministrazione di un vaccino fornito secondo l'insegnamento fornisce neoepitopi che sono epitopi di classe II presentati da MHC e/o sono in grado di suscitare una risposta di cellule T helper CD4+ contro cellule che esprimono antigeni da cui derivano gli epitopi presentati da MHC nonché epitopi non contenenti mutazioni somatiche specifiche del cancro che sono epitopi di classe I presentati dall'MHC e/o sono in grado di suscitare una risposta delle cellule T CD8+ contro le cellule che esprimono gli antigeni da cui derivano gli epitopi presentati dall'MHC. In una forma di realizzazione, gli epitopi che non contengono mutazioni somatiche specifiche del cancro sono derivati da un antigene tumorale. In una forma di realizzazione, i neoepitopi e gli epitopi che non contengono mutazioni somatiche specifiche del cancro hanno un effetto sinergico nel trattamento del cancro. Preferibilmente, un vaccino fornito secondo l'insegnamento è utile per la stimolazione poliepitopica delle risposte delle cellule T citotossiche e/o helper.

Il vaccino fornito secondo l'insegnamento può essere un vaccino ricombinante.

Un altro tipo di agente è una cellula immunitaria, come una cellula T che esprime un recettore della cellula T o una cellula T che esprime in modo ricombinante un recettore della cellula T, o che esprime un recettore della cellula T (CAR) artificiale o chimerico, il cui recettore è mirato a un antigene, ad esempio un neoepitopo adatto identificato dai metodi del presente insegnamento come un target adatto specifico della malattia, preferibilmente dove tale neoepitopo è espresso sulla superficie di una cellula in un complesso con molecole MHC. Preferibilmente, una volta che la cellula immunitaria riconosce l'antigene mediante il legame recettore-antigene, la cellula immunitaria (cellula immunoreattiva) viene stimolata, innescata e/o espansa o esercita funzioni effettrici delle cellule immunoreattive come descritto sopra.

Il termine "cellula T specifica dell'antigene" o termini simili si riferisce a una cellula T che riconosce un antigene, ad esempio un neoepitopo adatto complessato all'interno di molecole MHC di classe I e dopo il legame a detto antigene esercita preferibilmente funzioni effettrici delle cellule T come descritto sopra. Le cellule T e altre cellule linfoidi sono considerate specifiche per l'antigene se le cellule target distruggono le cellule target che esprimono l'antigene. La specificità delle cellule T può essere valutata utilizzando una qualsiasi varietà di tecniche standard, ad esempio,

all'interno di un test di rilascio di cromo o di un test di proliferazione. In alternativa, può essere misurata la sintesi di linfocine (come l'interferone γ).

I recettori della cellula T e altri recettori dell'antigene sono descritti sopra. Il termine "CAR" (o "recettore dell'antigene chimerico") si riferisce a un recettore artificiale comprendente una singola molecola o un complesso di molecole che riconosce, cioè si lega a, una struttura target (ad esempio un antigene) su una cellula target come un cellula cancerosa (ad esempio, legando un dominio di legame all'antigene a un antigene espresso sulla superficie della cellula target) e può conferire specificità a una cellula immunitaria effettrice come una cellula T che esprime detto CAR sulla superficie cellulare. Preferibilmente, il riconoscimento della struttura target da parte di un CAR determina l'attivazione di una cellula immunitaria effettrice che esprime detto CAR. Un CAR può comprendere una o più unità proteiche, dette unità proteiche comprendendo uno o più domini come qui descritti. Il termine "CAR" non include i recettori della cellula T.

In una forma di realizzazione, un frammento variabile a catena singola (scFv) derivato da un anticorpo monoclonale è fuso nella transmembrana CD3-zeta e all'endodominio. Tali molecole determinano la trasmissione di un segnale zeta in risposta al riconoscimento da parte di scFv del suo target dell'antigene su una cellula target e alla distruzione della cellula target che esprime l'antigene del target. I domini di riconoscimento dell'antigene che possono anche essere utilizzati includono tra gli altri le catene singole alfa e beta del recettore di cellule T (TCR). In effetti, quasi tutto ciò che lega un determinato target con alta affinità può essere utilizzato come dominio di riconoscimento dell'antigene.

Dopo il riconoscimento dell'antigene, i recettori si raggruppano e un segnale viene trasmesso alla cellula. A questo proposito, un "dominio di segnalazione di cellule T" è un dominio, preferibilmente un endodominio, che trasmette un segnale di attivazione alla cellula T dopo che l'antigene è stato legato. Il componente dell'endodominio più comunemente utilizzato è CD3-zeta.

La terapia adattiva di trasferimento cellulare con cellule T ingegnerizzate con CAR che esprimono recettori chimerici dell'antigene è una promettente modalità di terapia, poiché le cellule T modificate con CAR possono essere ingegnerizzate per mirare praticamente a qualsiasi antigene espresso su cellule malate, ad esempio antigeni tumorali. Preferibilmente, l'antigene tumorale è un neoepitopo risultante da una mutazione specifica del tumore identificata dai metodi del presente insegnamento come un target adatto specifico del tumore. Ad esempio, le cellule T del paziente possono essere ingegnerizzate in modo genetico (modificate geneticamente) per esprimere un CAR specificamente diretto verso i neoepitopi specifici dell'antigene complessati con molecole MHC sulla superficie delle cellule tumorali del paziente, quindi infuse nuovamente nel paziente.

Un CAR può sostituire la funzione di un recettore della cellula T e, in particolare, può conferire reattività come attività citolitica a una cellula come una cellula T. Tuttavia, in contrasto al legame del recettore della cellula T a un complesso MHC del peptide dell'antigene come descritto sopra, un

CAR può legarsi a un antigene, in particolare quando espresso sulla superficie cellulare.

Secondo l'insegnamento, i CAR possono generalmente comprendere tre domini. Il primo dominio è il dominio di legame che riconosce e lega l'antigene. Il secondo dominio è il dominio della costimolazione. Il dominio di costimolazione serve a migliorare la proliferazione e la sopravvivenza dei linfociti citotossici dopo il legame del CAR a una parte mirata. L'identità del dominio di costimolazione è limitata solo in quanto ha la capacità di migliorare la proliferazione cellulare e la sopravvivenza dopo il legame della parte mirata mediante il CAR. I domini di costimolazione adatti includono CD28, CD137 (4-1BB), un elemento della famiglia di recettori del fattore di necrosi tumorale (TNF), CD134 (OX40), un elemento della superfamiglia dei recettori TNFR e CD278 (ICOS), una molecola costimolatrice della superfamiglia CD28 espressa su cellule T attivate. Il terzo dominio è il dominio di segnalazione di attivazione (o dominio di segnalazione di cellule T). Il dominio di segnalazione di attivazione serve per attivare i linfociti citotossici al legame di CAR all'antigene. L'identità del dominio di segnalazione di attivazione è limitata solo in quanto ha la capacità di indurre l'attivazione del linfocita citotossico selezionato dopo il legame dell'antigene da parte del CAR. Domini di segnalazione di attivazione adatti includono la catena CD3 [zeta] di cellule T e il recettore Fc [gamma].

I CAR possono comprendere tre domini, insieme sotto forma di una proteina di fusione. Tali proteine di fusione comprenderanno generalmente un dominio legante, uno o più domini di costimolazione e un dominio di segnalazione di attivazione, collegati in una direzione da N-terminale a C-terminale. Tuttavia, i CAR non sono limitati a questa disposizione e altre disposizioni sono accettabili e includono un dominio legante, un dominio di segnalazione di attivazione e uno o più domini di costimolazione. Si comprenderà che poiché il dominio di legame deve essere libero di legare l'antigene, il posizionamento del dominio di legame nella proteina di fusione sarà generalmente tale da ottenere la visualizzazione della regione all'esterno della cellula. Allo stesso modo, poiché i domini di costimolazione e attivazione del segnale servono a indurre l'attività e la proliferazione dei linfociti citotossici, la proteina di fusione mostrerà generalmente questi due domini all'interno della cellula. I CAR possono includere elementi aggiuntivi, come un peptide di segnale per garantire la corretta esportazione della proteina di fusione sulla superficie cellulare, un dominio di transmembrana per garantire che la proteina di fusione sia mantenuta come proteina di membrana integrale e un dominio cerniera (o regione spaziatrice) che conferisce flessibilità al dominio di legame e consente un forte legame all'antigene.

Le cellule utilizzate insieme ai CAR e ad altri recettori dell'antigene artificiali sono preferibilmente cellule T, in particolare linfociti citotossici, preferibilmente selezionati tra cellule T citotossiche, cellule natural killer (NK) e cellule killer attivate da linfocine (LAK). Dopo l'attivazione, ciascuno di questi linfociti citotossici innesca la distruzione delle cellule target. Ad esempio, le cellule T citotossiche innescano la distruzione delle cellule target con uno o entrambi i seguenti mezzi. In primo luogo, dopo l'attivazione le cellule T rilasciano citotossine come perforina, granzimi e

granulicina. La perforina e la granulicina creano pori nella cellula target, mentre i granzimi entrano nella cellula e innescano una cascata di caspasi nel citoplasma che induce l'apoptosi (morte cellulare programmata) della cellula. In secondo luogo, l'apoptosi può essere indotta tramite l'interazione del ligando Fas-Fas tra le cellule T e le cellule target. I linfociti citotossici saranno preferibilmente cellule autologhe, sebbene possano essere utilizzate cellule eterologhe o allogeniche.

Un dominio di legame per un antigene che può essere presente all'interno di un CAR ha la capacità di legarsi (mirare) a un antigene, cioè la capacità di legarsi (mirare) a un epitopo presente in un antigene, preferibilmente la capacità di legarsi (mirare) a un neoepitopo identificato con i metodi del presente insegnamento come un target specifico della malattia dove il neoepitopo viene presentato nel contesto di MHC sulla superficie della cellula. Preferibilmente, un dominio di legame per un antigene è specifico per l'antigene.

Un altro tipo di agente è una cellula immunitaria, come una cellula linfoide, caricata con un peptide contenente un neoepitopo adatto identificato con i metodi del presente insegnamento. In una forma di realizzazione preferita, la cellula linfoide è una cellula dendritica. Nell'ambito del presente insegnamento, cellule linfoidi preferibilmente isolate dal paziente da trattare vengono incubate con un antigene da prendere di mira e le cellule incubate vengono quindi somministrate al paziente dove viene indotta una risposta immunitaria alle cellule che esprimono l'antigene. Pertanto, un peptide comprendente un epitopo adatto può essere incubato con cellule dendritiche e le cellule incubate possono essere somministrate per indurre una risposta immunitaria contro cellule che esprimono il neoepitopo adatto.

Il termine "ricombinante" nel contesto del presente insegnamento significa "realizzato attraverso l'ingegneria genetica". Preferibilmente, una "entità ricombinante" come un polipeptide ricombinante nel contesto del presente insegnamento non si verifica naturalmente, e preferibilmente è il risultato di una combinazione di entità come sequenze di amminoacidi o acidi nucleici che non sono combinate in natura. Ad esempio, un polipeptide ricombinante nel contesto del presente insegnamento può contenere diverse sequenze di amminoacidi come neo-epitopi o sequenze di vaccini derivate da differenti proteine o differenti porzioni della stessa proteina fuse insieme, ad esempio, da legami peptidici o linker appropriati.

Il termine "presente in natura" come utilizzato nel presente documento si riferisce al fatto che un oggetto può essere trovato in natura. Ad esempio, un peptide o un acido nucleico che è presente in un organismo (compresi i virus) e che può essere isolato da una fonte in natura e che non è stato intenzionalmente modificato dall'uomo in laboratorio, è reperibile in natura.

Secondo l'insegnamento, il termine "malattia" si riferisce a qualsiasi stato patologico, incluso la malattia del cancro, in particolare quelle forme di cancro descritte nel presente documento.

Il termine "normale" si riferisce a uno stato sano o alle condizioni di un soggetto o tessuto sano, cioè condizioni non patologiche, in cui "sano"

preferibilmente significa non canceroso.

"Malattie che coinvolgono cellule che esprimono un antigene" significa che viene individuata l'espressione dell'antigene nelle cellule di un tessuto o organo malato. L'espressione nelle cellule di un tessuto o organo malato può aumentare rispetto allo stato in un tessuto o organo sano. Un aumento si riferisce a un aumento almeno del 10%, in particolare almeno del 20%, almeno del 50%, almeno del 100%, almeno del 200%, almeno del 500%, almeno del 1000%, almeno del 10000% o anche di più. In una forma di realizzazione, l'espressione si trova solo in un tessuto malato, mentre l'espressione in un tessuto sano viene repressa. Secondo l'insegnamento, le malattie che coinvolgono o sono associate alle cellule che esprimono un antigene includono le malattie del cancro.

Il cancro (termine medico: neoplasia maligna) è una classe di malattie in cui un gruppo di cellule mostra una crescita incontrollata (divisione oltre i limiti normali), invasione (intrusione e distruzione dei tessuti adiacenti) e, talvolta, metastasi (diffuse in altre posizioni nel corpo tramite linfa o sangue). Queste tre proprietà maligne dei tumori li differenziano dai tumori benigni, che sono autolimitanti e non invadono né metastatizzano. La maggior parte dei cancri forma un tumore ma alcuni, come la leucemia, no.

Il tumore maligno è essenzialmente sinonimo di cancro. La malignità, neoplasia maligna e tumore maligno sono essenzialmente sinonimi di cancro. Secondo l'insegnamento, il termine "tumore" o "malattia tumorale" si riferisce a una crescita anormale di cellule (chiamate cellule neoplastiche, cellule tumorigeniche o cellule tumorali) che formano preferibilmente un rigonfiamento o una lesione. Per "cellula tumorale" si intende una cellula anormale che cresce per una proliferazione cellulare rapida e incontrollata e continua a crescere dopo che gli stimoli che hanno avviato la nuova crescita cessano. I tumori mostrano una mancanza parziale o totale di organizzazione strutturale e coordinazione funzionale con il tessuto normale e di solito formano una massa distinta di tessuto, che può essere benigna, pre-maligna o maligna.

Un tumore benigno è un tumore a cui mancano tutte e tre le proprietà maligne di un cancro. Quindi, per definizione, un tumore benigno non cresce in modo illimitato e aggressivo, non invade i tessuti circostanti e non si diffonde nei tessuti non adiacenti (metastatizza).

Il neoplasma è una massa anormale di tessuto a causa della neoplasia. La neoplasia (nuova crescita in greco) è la proliferazione anormale delle cellule. La crescita delle cellule supera e non è coordinata con quella dei tessuti normali che la circondano. La crescita persiste nella stessa maniera eccessiva anche dopo la cessazione degli stimoli. Di solito provoca un nodulo o un tumore. I neoplasmi possono essere benigni, pre-maligni o maligni.

"Crescita di un tumore" o "crescita tumorale" nel contesto del presente insegnamento si riferisce alla tendenza di un tumore di aumentare le sue dimensioni e/o alla tendenza delle cellule tumorali di proliferare.

Ai fini del presente insegnamento, i termini "cancro" e "malattia da cancro" sono usati in modo intercambiabile con i termini "tumore" e "malattia tumorale".

I cancri sono classificati in base al tipo di cellula che assomiglia al tumore e, quindi, al tessuto che si presume sia l'origine del tumore. Questi sono rispettivamente l'istologia e la posizione.

Il termine "cancro" secondo l'insegnamento comprende leucemie, seminomi, melanomi, teratomi, linfomi, neuroblastomi, gliomi, cancro del retto, cancro dell'endometrio, cancro del rene, cancro del surrene, cancro della tiroide, cancro del sangue, cancro della pelle, cancro del cervello, cancro cervicale, cancro intestinale, cancro del fegato, cancro del colon, cancro dello stomaco, cancro dell'intestino, cancro della testa e del collo, cancro gastrointestinale, cancro dei linfonodi, cancro dell'esofago, cancro del colon-retto, cancro del pancreas, cancro dell'orecchio, naso e gola (ORL), cancro del seno, cancro della prostata, cancro dell'utero, cancro dell'ovaio e cancro ai polmoni e loro metastasi. Esempi di questi sono carcinomi polmonari, carcinomi mammari, carcinomi della prostata, carcinomi del colon, carcinomi delle cellule renali, carcinomi cervicali o metastasi dei tipi di cancro o tumori descritti sopra. Il termine cancro secondo l'insegnamento comprende anche metastasi cancerose e ricadute del cancro.

Per "metastasi" si intende la diffusione delle cellule cancerose dalla sua sede originaria a un'altra parte del corpo. La formazione di metastasi è un processo molto complesso e dipende dal distacco delle cellule maligne dal tumore primario, dall'invasione della matrice extracellulare, dalla penetrazione delle membrane basali endoteliali per entrare nelle cavità corporee e nei vasi sanguigni e quindi, dopo essere state trasportate dal sangue, dall'infiltrazione di organi target. Infine, la crescita di un nuovo tumore, cioè un tumore secondario o metastatico, nel sito target dipende dall'angiogenesi. Le metastasi tumorali si verificano spesso anche dopo la rimozione del tumore primario perché le cellule o i componenti del tumore possono rimanere e sviluppare un potenziale metastatico. In una forma di realizzazione, il termine "metastasi" secondo l'insegnamento si riferisce a "metastasi a distanza" che si riferisce a una metastasi che è lontana dal tumore primario e dal sistema linfonodale regionale.

Le cellule di un tumore secondario o metastatico sono come quelle del tumore originale. Ciò significa, ad esempio, che, se il cancro ovarico metastatizza al fegato, il tumore secondario è costituito da cellule ovariche anormali, non da cellule epatiche anormali. Il tumore nel fegato viene quindi chiamato cancro ovarico metastatico, non cancro al fegato.

Il termine "cellule tumorali circolanti" o "CTC" si riferisce a cellule che si sono staccate da un tumore primario o da metastasi tumorali e circolano nel flusso sanguigno. Le CTC possono costituire i semi per la successiva crescita di ulteriori tumori (metastasi) in diversi tessuti. Le cellule tumorali circolanti si trovano a frequenze nell'ordine di 1-10 CTC per ml di sangue intero in pazienti con malattia metastatica. Sono stati sviluppati metodi di ricerca per isolare CTC. Nella tecnica sono stati descritti diversi metodi di ricerca per isolare le CTC, ad esempio tecniche che utilizzano

il fatto che le cellule epiteliali esprimono comunemente la proteina di adesione cellulare EpCAM, che è assente nelle cellule del sangue normali. La cattura immunomagnetica basata su granuli comporta il trattamento di campioni di sangue con anticorpi contro EpCAM che è stato coniugato con particelle magnetiche, seguito dalla separazione delle cellule contrassegnate in un campo magnetico. Le cellule isolate vengono quindi colorate con l'anticorpo verso un altro marker epiteliale, la citocheratina, nonché un comune marker leucocitario CD45, in modo da distinguere le CTC rare dai globuli bianchi contaminanti. Questo solido approccio semi-automatizzato identifica le CTC con una resa media di circa 1 CTC/ml e una purezza dello 0,1% (Allard et al., 2004, Clin Cancer Res 10:6897-6904). Un secondo metodo per isolare le CTC utilizza un dispositivo di cattura CTC basato su microfluidica che prevede il flusso di sangue intero attraverso una camera incorporata con 80.000 micropost che sono stati resi funzionali mediante rivestimento con anticorpo a EpCAM. Le CTC vengono quindi colorate con anticorpi secondari contro citocheratina o marker tissutali specifici, come PSA nel cancro alla prostata o HER2 nel cancro al seno e vengono visualizzate mediante scansione automatizzata di micropost su più piani lungo coordinate tridimensionali. I chip CTC sono in grado di identificare cellule tumorali circolanti positive alla citocherizzazione in pazienti con una resa media di 50 cellule/ml e una purezza compresa di 1-80% (Nagrath et al., 2007, Nature 450:1235-1239). Un'altra possibilità per isolare le CTC è l'utilizzo del test CellSearch™ Circulating Tumor Cell (CTC) di Veridex, LLC (Raritan, NJ) che cattura, identifica e conta le CTC in una provetta di sangue. Il sistema CellSearch™ è una metodologia approvata dalla Food and Drug Administration (FDA) statunitense per il conteggio delle CTC nel sangue intero che si basa su una combinazione di etichettatura immunomagnetica e microscopia digitale automatizzata. Ci sono altri metodi per isolare i CTC descritti in letteratura che possono essere tutti usati insieme al presente insegnamento.

Una ricaduta o una recidiva si verificano quando una persona viene nuovamente colpita da una condizione che l'ha colpita in passato. Ad esempio, se un paziente ha sofferto di una malattia tumorale, ha ricevuto un trattamento di successo di detta malattia e sviluppa nuovamente detta malattia, detta nuova malattia di sviluppo può essere considerata come ricaduta o recidiva. Tuttavia, secondo l'insegnamento, una ricaduta o una recidiva di una malattia tumorale può, ma non necessariamente, verificarsi nel sito della malattia tumorale originale. Così, ad esempio, se un paziente ha sofferto di un tumore ovarico e ha ricevuto un trattamento di successo, una ricaduta o una recidiva può essere il verificarsi di un tumore ovarico o il verificarsi di un tumore in un sito diverso dall'ovaio. Una ricaduta o recidiva di un tumore include anche situazioni in cui un tumore si verifica in un sito diverso da quello del tumore originale, nonché nel sito del tumore originale. Preferibilmente, il tumore originale per cui il paziente ha ricevuto un trattamento è un tumore primario e il tumore in un sito diverso dal sito del tumore originale è un tumore secondario o metastatico.

Il termine "risposta immunitaria" si riferisce a una reazione del sistema immunitario come a organismi immunogenici, come batteri o virus, cellule o sostanze. Il termine "risposta immunitaria" include la risposta immunitaria innata e la risposta immunitaria adattativa. Preferibilmente, la risposta

immunitaria è correlata a un'attivazione di cellule immunitarie, a un'induzione della biosintesi di citochine e/o alla produzione di anticorpi.

Si preferisce che la risposta immunitaria indotta dalle composizioni del presente insegnamento comprenda le fasi di attivazione di cellule che presentano l'antigene, come cellule dendritiche e/o macrofagi, la presentazione di un antigene o di un suo frammento da parte di dette cellule che presentano l'antigene e l'attivazione di cellule T citotossiche a causa di questa presentazione.

Il termine "cellule immunitarie" si riferisce alle cellule del sistema immunitario coinvolte nella difesa del corpo di un individuo. Il termine "cellule immunitarie" comprende tipi specifici di cellule immunitarie e i loro precursori inclusi leucociti comprendenti macrofagi, monociti (precursori dei macrofagi), granulociti come neutrofilo, eosinofilo e basofilo, cellule dendritiche, mastociti e linfociti come cellule B, cellule T cellule e cellule natural killer (NK). Macrofagi, monociti (precursori dei macrofagi), neutrofilo, cellule dendritiche e mastociti sono cellule fagocitiche.

Il termine "immunoterapia" si riferisce al trattamento di una malattia o condizione inducendo, potenziando o eliminando una risposta immunitaria. Le immunoterapie progettate per suscitare o amplificare una risposta immunitaria sono classificate come immunoterapie di attivazione, mentre le immunoterapie che riducono o eliminano sono classificate come immunoterapie di soppressione. Il termine "immunoterapia" include l'immunizzazione dell'antigene o la vaccinazione dell'antigene, o l'immunizzazione del tumore o la vaccinazione del tumore. Il termine "immunoterapia" si riferisce anche alla manipolazione delle risposte immunitarie in modo tale che le risposte immunitarie inadeguate siano modulate in quelle più appropriate nel contesto di malattie autoimmuni come l'artrite reumatica, le allergie, il diabete o la sclerosi multipla.

Il termine "immunizzazione" o "vaccinazione" descrive il processo di somministrazione di un antigene a un individuo con lo scopo di indurre una risposta immunitaria, ad esempio, per ragioni terapeutiche o profilattiche.

Con "trattare" si intende somministrare un composto o una composizione come descritto nel presente documento a un soggetto al fine di prevenire o eliminare una malattia, inclusa la riduzione delle dimensioni di un tumore o del numero di tumori in un soggetto; arrestare o rallentare una malattia in un soggetto; inibire o rallentare lo sviluppo di una nuova malattia in un soggetto; diminuire la frequenza o la gravità dei sintomi e/o delle recidive in un soggetto che attualmente ha o che ha precedentemente avuto una malattia; e/o prolungare, cioè aumentare la durata della vita del soggetto. In particolare, il termine "trattamento di una malattia" include la cura, la riduzione della durata, il miglioramento, la prevenzione, il rallentamento o l'inibizione della progressione o il peggioramento, o la prevenzione o il ritardo dell'insorgenza di una malattia o dei suoi sintomi.

Per "essere a rischio" si intende un soggetto, vale a dire un paziente, che viene identificato come avente una probabilità superiore al normale di sviluppare una malattia, in particolare un cancro, rispetto alla popolazione generale. Inoltre, un soggetto che ha avuto o che ha attualmente una malattia, in particolare il cancro, è un soggetto che ha un rischio maggiore di sviluppare una malattia, in quanto tale soggetto può continuare a

sviluppare una malattia. I soggetti che attualmente hanno o che hanno avuto un cancro hanno anche un aumentato rischio di metastasi del cancro.

Una somministrazione profilattica di un'immunoterapia, ad esempio, una somministrazione profilattica della composizione dell'insegnamento, protegge preferibilmente il ricevente dallo sviluppo di una malattia. Una somministrazione terapeutica di un'immunoterapia, ad esempio, una somministrazione terapeutica della composizione dell'insegnamento, può portare all'inibizione del progresso/crescita della malattia. Ciò comprende la decelerazione del progresso/crescita della malattia, in particolare un'interruzione della progressione della malattia, che preferibilmente porta all'eliminazione della malattia.

L'immunoterapia può essere eseguita utilizzando una qualsiasi di una varietà di tecniche, in cui gli agenti forniti nel presente documento funzionano per rimuovere le cellule malate da un paziente. Tale rimozione può avvenire come risultato del potenziamento o dell'induzione di una risposta immunitaria in un paziente specifico per un antigene o in una cellula che esprime un antigene.

All'interno di alcune forme di realizzazione, l'immunoterapia può essere l'immunoterapia attiva, in cui il trattamento si basa sulla stimolazione in vivo del sistema immunitario dell'ospite endogeno per reagire contro le cellule malate con la somministrazione di agenti che modificano la risposta immunitaria (come polipeptidi e acidi nucleici come forniti nel presente documento).

Gli agenti e le composizioni forniti nel presente documento possono essere utilizzati da soli o in combinazione con regimi terapeutici convenzionali come chirurgia, irradiazione, chemioterapia e/o trapianto di midollo osseo (autologo, singenico, allogenico o non correlato).

Il termine "in vivo" si riferisce alla situazione in un soggetto.

I termini "soggetto", "individuo", "organismo" o "paziente" si riferiscono a vertebrati, in particolare mammiferi. Ad esempio, i mammiferi nel contesto del presente insegnamento sono esseri umani, primati non umani, mammiferi domestici come cani, gatti, pecore, bovini, capre, maiali, cavalli ecc., animali da laboratorio come topi, ratti, conigli, porcellini d'india ecc. così come animali in cattività come gli animali degli zoo. I termini si riferiscono anche ai vertebrati non mammiferi come gli uccelli (in particolare uccelli domestici come polli, anatre, oche, tacchini) e ai pesci (in particolare pesci d'allevamento, ad esempio salmoni o pesci gatto). Il termine "animale" come qui utilizzato include anche gli animali.

Il termine "autologo" è utilizzato per descrivere tutto ciò che è derivato dallo stesso soggetto. Ad esempio, "trapianto autologo" si riferisce a un trapianto di tessuto o di organi derivati dallo stesso soggetto. Tali procedure sono vantaggiose perché superano la barriera immunologica che altrimenti porta al rigetto.

Il termine "eterologo" è utilizzato per descrivere qualcosa costituito da più elementi differenti. Ad esempio, il trasferimento del midollo osseo di un individuo in un individuo diverso costituisce un trapianto eterologo. Un gene eterologo è un gene derivato da una fonte diversa dal soggetto.

Come parte della composizione per un'immunizzazione o una vaccinazione, preferibilmente uno o più agenti come descritti nel presente documento vengono somministrati insieme a uno o più adiuvanti per indurre una risposta immunitaria o per aumentare una risposta immunitaria. Il termine "adiuvante" si riferisce a composti che prolungano o potenziano o accelerano una risposta immunitaria. La composizione del presente insegnamento esercita preferibilmente il suo effetto senza aggiunta di adiuvanti. Tuttavia, la composizione della presente domanda può contenere qualsiasi adiuvante noto. Gli adiuvanti comprendono un gruppo eterogeneo di composti come emulsioni oleose (ad esempio, adiuvanti di Freund), composti minerali (come allume), prodotti batterici (come la tossina di Bordetella pertussis), liposomi e complessi immunostimolanti. Esempi di adiuvanti sono monofosforil-lipide-A (MPL SmithKline Beecham). Saponine come QS21 (SmithKline Beecham), DQS21 (SmithKline Beecham; WO 96/33739), QS7, QS17, QS18 e QS-L1 (So et al., 1997, Mol. Cells 7: 178-186), adiuvanti di Freund incompleti, adiuvanti di Freund completi, vitamina E, montanide, allume, oligonucleotidi CpG (Krieg et al., 1995, Nature 374: 546-549) e varie emulsioni acqua in olio preparate da oli biodegradabili come squalene e/o tocoferolo.

Possono inoltre essere somministrate altre sostanze che stimolano una risposta immunitaria del paziente. È possibile, ad esempio, usare le citochine in una vaccinazione, per via delle loro proprietà regolatrici sui linfociti. Tali citochine comprendono, ad esempio, interleuchina-12 (IL-12) che ha dimostrato di aumentare le azioni protettive dei vaccini (vedere Hall, 1995, IL-12 al punto di unione, Science 268:1432-1434), GM-CSF e IL-18.

Esistono numerosi composti che migliorano una risposta immunitaria e che quindi possono essere utilizzati in una vaccinazione. Detti composti comprendono molecole costimolanti fornite sotto forma di proteine o acidi nucleici come B7-1 e B7-2 (rispettivamente CD80 e CD86).

Secondo l'insegnamento, un "campione di tumore" è un campione come un campione corporeo contenente cellule tumorali o cancerose come cellule tumorali circolanti (CTC), in particolare un campione di tessuto, inclusi fluidi corporei, e/o un campione cellulare. Secondo l'insegnamento, un "campione non tumorale" è un campione come un campione corporeo non contenente cellule tumorali o cancerose come cellule tumorali circolanti (CTC), in particolare un campione di tessuto, inclusi fluidi corporei, e/o un campione cellulare. Tali campioni corporei possono essere ottenuti nel modo convenzionale come mediante biopsia tissutale, inclusa biopsia con punch, e prelevando sangue, aspirato bronchiale, espettorato, urina, feci o altri fluidi corporei. Secondo l'insegnamento, il termine "campione" include anche campioni elaborati come frazioni o isolati di campioni biologici, ad esempio acido nucleico o isolati cellulari.

Gli agenti, i vaccini e le composizioni terapeutamente attivi descritti nel presente documento possono essere somministrati tramite qualsiasi via convenzionale, inclusa l'iniezione o l'infusione. La somministrazione può essere effettuata, ad esempio, per via orale, endovenosa, intraperitoneale, intramuscolare, sottocutanea o transdermica. In una forma di realizzazione, la somministrazione viene effettuata intranodalmente come mediante

iniezione in un linfonodo. Altre forme di somministrazione prevedono la trasfezione in vitro di cellule che presentano l'antigene come le cellule dendritiche con acidi nucleici qui descritte seguita dalla somministrazione delle cellule che presentano l'antigene.

Gli agenti descritti nel presente documento vengono somministrati in quantità efficaci. Una "quantità efficace" si riferisce alla quantità che ottiene una reazione desiderata o un effetto desiderato da solo o insieme a ulteriori dosi. Nel caso del trattamento di una particolare malattia o di una particolare condizione, la reazione desiderata si riferisce preferibilmente all'inibizione del decorso della malattia. Questo comprende il rallentamento del progresso della malattia e, in particolare, l'interruzione o l'inversione del progresso della malattia. La reazione desiderata in un trattamento di una malattia o di una condizione può anche essere un ritardo dell'insorgenza o una prevenzione dell'insorgenza di detta malattia o di detta condizione.

Una quantità efficace di un agente descritta nel presente documento dipenderà dalla condizione da trattare, dalla gravità della malattia, dai parametri individuali del paziente, inclusi l'età, la condizione fisiologica, le dimensioni e il peso, la durata del trattamento, il tipo di terapia di accompagnamento (se presente), lo specifico percorso di somministrazione e fattori simili. Di conseguenza, le dosi somministrate degli agenti qui descritti possono dipendere da diversi di tali parametri. Nel caso in cui una reazione in un paziente sia insufficiente con una dose iniziale, possono essere utilizzate dosi più elevate (o dosi effettivamente più elevate ottenute con un percorso di somministrazione diverso e più localizzato).

Il termine "farmaceuticamente accettabile" si riferisce alla non tossicità di un materiale che non interagisce con l'azione del componente attivo della composizione farmaceutica.

Le composizioni farmaceutiche del presente insegnamento possono contenere sali, tamponi, agenti conservanti, trasportatori e opzionalmente altri agenti terapeutici. Preferibilmente, le composizioni farmaceutiche del presente insegnamento comprendono uno o più trasportatori, diluenti e/o eccipienti farmaceuticamente accettabili.

Il termine "eccipiente" intende indicare tutte le sostanze della composizione farmaceutica che non sono principi attivi come, ad esempio leganti, lubrificanti, addensanti, agenti tensioattivi, conservanti, emulsionanti, tamponi, agenti aromatizzanti o coloranti.

Il termine "diluente" si riferisce a un agente diluente e/o assottigliante. Inoltre, il termine "diluente" include qualsiasi mezzo o più mezzi di sospensione fluida, liquida o solida e/o di miscelazione.

Il termine "trasportatore" si riferisce a uno o più riempitivi o diluenti solidi o liquidi compatibili, che sono adatti a una somministrazione a un essere umano. Il termine "trasportatore" si riferisce a un componente organico o inorganico naturale o sintetico che viene combinato con un componente attivo per facilitare l'applicazione del componente attivo. Preferibilmente, i componenti del trasportatore sono liquidi sterili come acqua o oli,

inclusi quelli derivati da olio minerale, animale o vegetale, come olio di arachidi, olio di semi di soia, olio di sesamo, olio di semi di girasole ecc.

Soluzioni saline e destrosio acquoso e soluzioni di glicerina possono anche essere utilizzate come composti trasportatori acquosi.

Trasportatori o diluenti farmaceuticamente accettabili per l'utilizzo terapeutico sono ben noti nella tecnica farmaceutica e sono descritti, ad esempio, in Remington's Pharmaceutical Sciences, Mack Publishing Co. (a cura di A. R. Gennaro 1985). Esempi di trasportatori adatti includono, ad esempio, carbonato di magnesio, stearato di magnesio, talco, zucchero, lattosio, pectina, destrina, amido, gelatina, gomma adragante, metilcellulosa, sodio carbossimetilcellulosa, una cera a basso punto di fusione, burro di cacao e simili. Esempi di diluenti idonei includono etanolo, glicerolo e acqua.

Trasportatori, eccipienti o diluenti farmaceutici possono essere selezionati rispetto al percorso di somministrazione previsto e alla pratica farmaceutica standard. Le composizioni farmaceutiche del presente insegnamento possono comprendere, o essere in aggiunta a, uno o più trasportatori, uno o più eccipienti o uno o più diluenti, qualsiasi legante adatto, uno o più lubrificanti, uno o più agenti di sospensione, uno o più agenti di rivestimento e/o uno o più agenti solubilizzanti. Esempi di leganti adatti includono amido, gelatina, zuccheri naturali come glucosio, lattosio anidro, lattosio a libero flusso, beta lattosio, dolcificanti del mais, gomme naturali e sintetiche come acacia, adragante o alginato di sodio, carbossimetilcellulosa e polietilenglicole. Esempi di lubrificanti idonei includono oleato di sodio, stearato di sodio, stearato di magnesio, benzoato di sodio, acetato di sodio, cloruro di sodio e simili. Conservanti, stabilizzanti, coloranti e persino agenti aromatizzanti possono essere forniti nella composizione farmaceutica. Esempi di conservanti includono benzoato di sodio, acido sorbico ed esteri di acido p-idrossibenzoico. Possono anche essere utilizzati antiossidanti e agenti di sospensione.

In una forma di realizzazione, la composizione è una composizione acquosa. La composizione acquosa può opzionalmente comprendere soluti, ad esempio sali. In una forma di realizzazione, la composizione è sotto forma di una composizione liofilizzata. Una composizione liofilizzata è ottenibile mediante liofilizzazione di una rispettiva composizione acquosa.

Gli agenti e le composizioni forniti nel presente documento possono essere utilizzati da soli o in combinazione con altri regimi terapeutici come chirurgia, irradiazione, chemioterapia e/o trapianto di midollo osseo (autologo, singenico, allogenico o non correlato).

Il presente insegnamento è descritto in dettaglio ed è illustrato dalle figure e dagli esempi, che sono utilizzati solo a scopo illustrativo e non intendono essere limitativi. Grazie alla descrizione e agli esempi, ulteriori forme di realizzazione che sono ugualmente incluse nell'insegnamento sono accessibili al lavoratore specializzato.

FIGURE

Figure 1a e 1b Campione di glioblastoma con elevata amplificazione focale del gene del recettore del fattore di crescita dell'epidermide (EGFR).

Figura 1a: Una rappresentazione grafica dei geni locali che circondano EGFR sul cromosoma 7. Figura 1b: Elenco delle 12 variazioni del nucleotide singolo con i più elevati numeri assoluti di copie.

Figura 2 Un elenco di un numero di geni che hanno una mutazione specifica della malattia di un campione di melanoma ordinato per zigosità (Cx).

Figura 3 Un elenco di un numero di geni di un campione di melanoma in cui tutte le copie hanno la mutazione specifica della malattia (la zigosità frazionata è uguale a 1), di cui tre geni sono geni essenziali nell'uomo o considerati geni essenziali negli esseri umani.

Abbreviazioni per le figure: chr_pos posizione cromosomica; NC numero assoluto di copie; Cx zigosità; CE correzione degli errori del numero assoluto di copie; VAF variante frequenza allelica; rho percentuale stimata di cellule tumorali contenenti l'allele mutato (SNV); FLRT+u punteggio di affidabilità nella mutazione; la Classificazione del sito è la classe di affidabilità della mutazione; gene essenziale, Y se il gene risulta essere essenziale.

ESEMPIO 1 Prendere di mira le mutazioni specifiche della malattia in geni con un numero elevato di copie

Le informazioni genomiche per il campione di glioblastoma (Chin et al., 2008, Comprehensive genomic characterization defines human glioblastoma genes and core pathways, Nature 455:1061-1068) sono state analizzate cercando geni con un numero elevato di copie e in cui almeno una copia del gene conteneva una mutazione specifica della malattia. L'analisi della qualità ha indicato che c'era un'alta correttezza di assegnazione nel numero di copie per 11.574 geni analizzati singolarmente e la ploidia del genoma del campione è stata determinata essere 1,95. La figura 1a mostra una rappresentazione grafica dei geni locali che circondano il recettore del fattore di crescita dell'epidermide (EGFR) sul cromosoma 7, che è un gene driver noto ed è stato un target per il trattamento. È stato dimostrato che l'EGFR in questo genoma aveva un numero assoluto di copie corretto per errore pari a 76, di cui 13 copie contenevano la variazione di singolo nucleotide specifica della malattia. La figura 1b fornisce un elenco dei geni in questo campione genomico con il numero assoluto di copie più elevato. Ci sono quattro geni aggiuntivi che hanno un numero assoluto di copie maggiore di 2. Infatti, l'amplificazione di EGFR è una nota caratteristica genetica dei glioblastomi primari (Benito et al., 2009, Neuropathology 30 (4):392-400) e questo gene è stato considerato come un target per il trattamento (Taylor, 2012, Curr Cancer Drug Targets. Mar; 12(3): 197-209).

ESEMPIO 2 Prendere di mira mutazioni specifiche della malattia con un'elevata zigosità

Un esoma ottenuto da un campione di cellule di melanoma da un tumore in un essere umano è stato analizzato cercando geni in cui almeno una copia del gene ha una mutazione specifica della malattia e osservando il numero di copie del gene che ha la mutazione specifica della malattia così

come il numero totale di copie del gene, indipendentemente dal fatto che il gene abbia la mutazione o meno. La figura 2 fornisce un elenco di geni ordinati per zigosità in cui la mutazione specifica della malattia si trova su più copie del gene. Ad esempio, la mutazione specifica della malattia nel gene OXGR1 ha la più elevata zigosità (4) e in particolare ha anche la più alta zigosità frazionata di $4/5$ o 0,8 poiché ci sono un totale di 5 copie del gene OXGR1 e 4 di queste copie contengono la mutazione specifica della malattia. L'elenco fornisce 10 geni aggiuntivi in cui 3 copie del gene su un totale di 4 copie hanno la mutazione, indicando che la mutazione specifica della malattia in questi geni ha una zigosità frazionata di $3/4$ o 0,75. I restanti geni elencati hanno mutazioni che hanno una zigosità frazionata di $2/3$ o 0,66 poiché 2 copie su un totale di 3 copie hanno la mutazione.

ESEMPIO 3 Prendere di mira le mutazioni specifiche della malattia presenti in tutte le copie di un gene essenziale

Un esoma ottenuto da un campione di cellule di melanoma di un tumore in un essere umano è stato analizzato cercando geni in cui tutte le copie del gene hanno la stessa mutazione specifica della malattia e determinando quale di questi geni è un gene essenziale. I geni sono stati determinati come essenziali deducendo la loro essenzialità negli esseri umani dalla consapevolezza che sono essenziali nei topi (Georgi et al., 2013, From mouse to human: evolutionary genomics analysis of human orthologs of essential genes, PLoS Genetics 9 (5):e1003484; Liao et al., 2007, Mouse duplicate genes are as essential as singletons, Trends Genet. 23:378-381). La figura 3 elenca un numero di geni in cui tutte le copie del gene hanno la stessa mutazione specifica della malattia. Inoltre, i tre geni evidenziati sono stati determinati come essenziali deducendo la loro essenzialità dai dati sui topi.

RIVENDICAZIONI

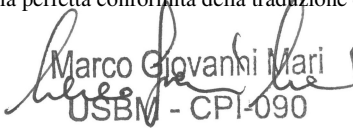
1. Metodo per determinare l' idoneità di un neoepitopo risultante da una mutazione specifica della malattia per un allele in un gene (allele mutato) come target specifico della malattia comprendente determinare, in una cellula malata o in una popolazione di cellule malate, il numero di copie dell'allele mutato che codifica il neoepitopo, in cui un numero di copie dell'allele mutato, che è maggiore di 2, indica l' idoneità del neoepitopo come target specifico della malattia, in cui la malattia è cancro e in cui maggiore è il numero di copie dell'allele mutato, maggiore è l' idoneità del neoepitopo come target specifico della malattia.
2. Metodo secondo la rivendicazione 1, in cui un numero di copie dell'allele mutato che è maggiore di 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90 o è maggiore di 100 indica l' idoneità del neoepitopo come target specifico della malattia.
3. Metodo secondo la rivendicazione 1 o 2, in cui il rapporto tra il numero di copie dell'allele mutato rispetto al numero totale di copie del sito nucleotidico a cui si mappa la mutazione, è maggiore di 0,5, in cui, preferibilmente, il rapporto è 1.
4. Metodo secondo la rivendicazione 3, in cui maggiore è il rapporto tra il numero di copie dell'allele mutato rispetto al numero totale di copie del sito nucleotidico a cui si mappa la mutazione, maggiore è l' idoneità del neoepitopo come target specifico della malattia.
5. Metodo secondo una qualsiasi delle rivendicazioni da 1 a 4, in cui il gene è un gene driver la cui espressione determina la trasformazione della cellula in un fenotipo canceroso o la cui mancanza di espressione fa sì che una cellula cancerosa perda il suo fenotipo canceroso.
6. Metodo secondo una qualsiasi delle rivendicazioni da 1 a 5, in cui il gene è un gene essenziale, in cui, preferibilmente, il gene essenziale è un gene che, quando viene silenziato o la sua espressione è ridotta, risulta almeno in una crescita compromessa o in una ridotta idoneità di una cellula in cui è espresso il gene essenziale, preferibilmente la cellula malata o il gene essenziale è espresso in un'ampia varietà di tessuti diversi ed è espresso con una soglia RPKM minima maggiore di 0.
7. Metodo secondo una qualsiasi delle rivendicazioni da 1 a 6, in cui la mutazione specifica della malattia è una singola variazione del nucleotide.
8. Metodo secondo una qualsiasi delle rivendicazioni da 1 a 7, in cui il neoepitopo è identificato mediante un metodo che comprende il sequenziamento del genoma o di una sua porzione di una cellula malata.
9. Metodo per fornire un vaccino comprendente identificare un neoepitopo adatto o una combinazione di neoepitopi adatti secondo il metodo di una qualsiasi delle rivendicazioni da 1 a 8.
10. Metodo secondo la rivendicazione 9, in cui il vaccino comprende un peptide o polipeptide comprendente uno o più neoepitopi adatti o

una combinazione di neoepitopi adatti o un acido nucleico che codifica detto peptide o polipeptide.

11. Metodo secondo la rivendicazione 10, in cui l'acido nucleico è RNA.

* * * * *

Si attesta la perfetta conformità della traduzione che precede.


Marco Giovanni Mari
USBM - CPI-090

LEGENDA DELLE TAVOLE DI DISEGNO

TAVOLE DA 1/4 A 4/4

FIGURE DA 1a A 3

"Figure" = "Figura"

TAVOLA 1/4

FIGURA 1a

"CN legend" = "Legenda NC"

"CN" = "NC"

"undetermined CN" = "NC indeterminato"

"position (bp)" = "posizione (bp)"

TAVOLE DA 2/4 A 4/4

FIGURE DA 1b A 3

"chr_pos (0 based)" = "chr_pos (in base 0)"

"gene" = "gene"

"CN w/ EC" = "NC con CE"

"Site classification" = "Classificazione del sito"

"Clone type" = "Tipo clone"

"wt" = "in peso"

TAVOLA 2/4

FIGURA 1b

"FIXED" = "FISSO"

"not fixed (>4 std)" = "non fisso (>4 std)"

TAVOLA 4/4

FIGURA 3

"Essential" = "Essenziale"

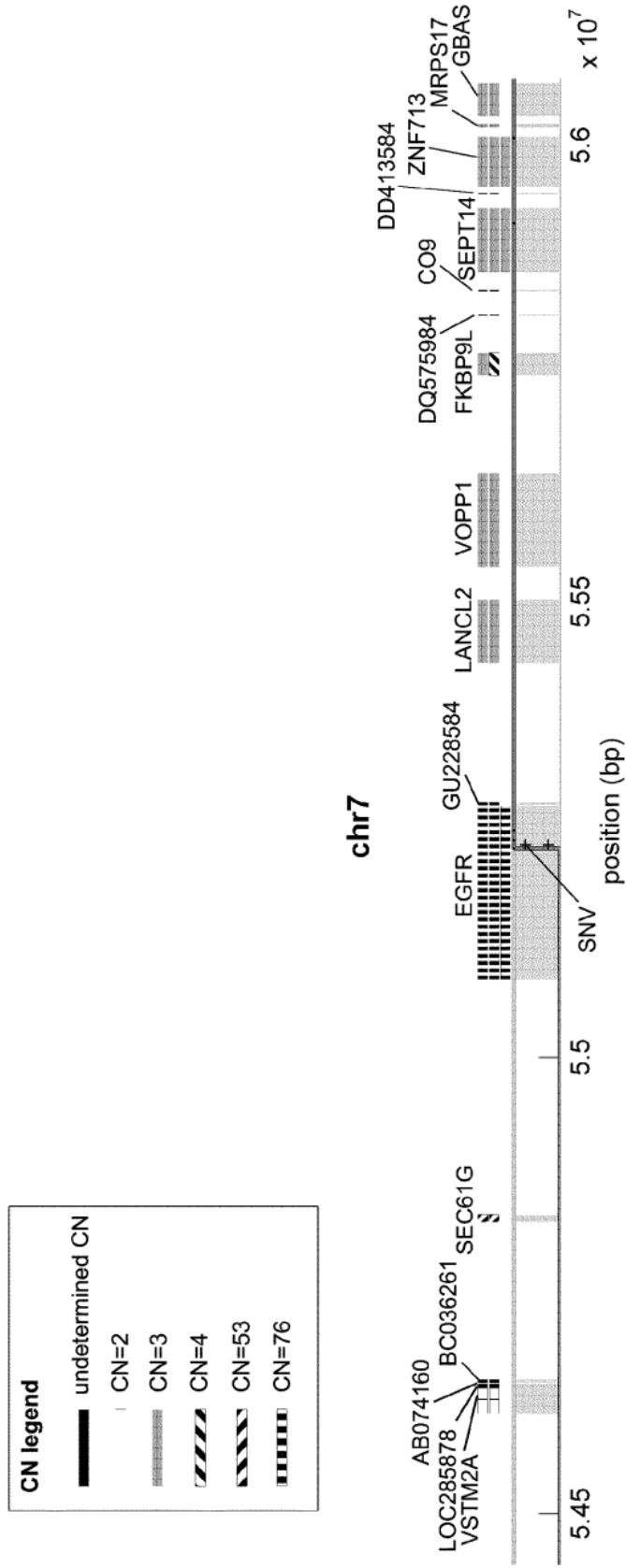
"Y" = "Sì"

* * * * *

Si attesta la perfetta conformità della traduzione che precede.


Marco Giovanni Mari
USBM - CPI-090

Figure 1a



chr_pos (0 based)	gene	CN w/ EC	Cx	VAF=Nx_Q30/(Nx_Q30+Nz_Q30)	rho	FIRTu	Site classification	Clone type	wt mut nt
chr7_55233108	EGFR	76	13	0,165	0,968	-284,984	SOMATIC_HI	?	G A
chr3_50387202	NPRL2	4	1	0,262	1,126	-80,215	SOMATIC_HI	FIXED	G A
chr7_100677038	MUC17	3	1	0,26	0,858	-216,9	SOMATIC_HI	FIXED	C T
chr16_1412528	GNPTG	3	2	0,545	0,9	-64,0632	SOMATIC_HI	?	C T
chr17_46607714	HOXB1	3	1	0,286	0,943	-20,1572	SOMATIC_HI	FIXED	C A
chr1_154148651	TPM3	2	1	0,247	0,568	-34,0489	SOMATIC_HI	not fixed (>4 std)	G A
chr1_161206280	NR1T3, TOMM40L	2	1	0,224	0,515	-21,4224	SOMATIC_HI	not fixed (>4 std)	C T
chr2_198950755	PLCL1	2	1	0,437	1,005	-63,2453	SOMATIC_HI	FIXED	G A
chr2_219255986	SIC11A1	2	1	0,392	0,9	-54,2955	SOMATIC_HI	FIXED	G A
chr2_220422919	OBSL1	2	1	0,361	0,83	-14,37	SOMATIC_HI	FIXED	T C
chr3_45077082	CLEC3B	2	1	0,329	0,756	-26,0748	SOMATIC_HI	FIXED	C T
chr3_137796432	DZIF1L	2	1	0,428	0,984	-114,485	SOMATIC_HI	FIXED	C T

Figure 1b

chr_pos (0 based)	gene	CN w/ EC	Cx	VAF= Nx_Q30/(Nx_Q 30+Nz_Q30)	rho	FLRT+u	Site classification	Clone type	wt	nt	mut	nt
chr13_96437970	OxGR1	5	4	0,743	0,948	-915,97	SOMATIC_HI	?	G			A
chr13_19694928	GJB6	4	3	0,675	0,923	-328,977	SOMATIC_HI	?	G			A
chr13_75293476	LMO7	4	3	0,749	1,025	-795,252	SOMATIC_HI	?	C			T
chr15_43349693	SLC28A2	4	3	0,689	0,943	-290,407	SOMATIC_HI	?	G			A
chr15_52647370	UNC13C	4	3	0,679	0,929	-165,72	SOMATIC_HI	?	C			T
chr15_64514504	MAP2K1	4	3	0,736	1,008	-206,449	SOMATIC_HI	?	A			C
chr15_71819607	C15orf59	4	3	0,679	0,929	-1035,86	SOMATIC_HI	?	A			C
chr15_72751030	EDC3	4	3	0,789	1,079	-314,949	SOMATIC_HI	?	C			T
chr15_77373587	ANKRD34C	4	3	0,706	0,966	-1559,58	SOMATIC_HI	?	G			A
chr15_78256958	FAH	4	3	0,66	0,903	-187,055	SOMATIC_HI	?	C			T
chr15_97489976	KIAA0353, SYNM	4	3	0,714	0,977	-845,185	SOMATIC_HI	?	C			A
chr14_57007978	C14orf105	3	2	0,698	1,084	-522,497	SOMATIC_HI	?	C			T
chr15_82964944	ZSCAN2	3	2	0,703	1,091	-418,868	SOMATIC_HI	?	C			T
chr17_59374140	SCN4A	3	2	0,654	1,015	-241,549	SOMATIC_HI	?	G			A
chr17_65683550	KCNJ2	3	2	0,636	0,988	-419,633	SOMATIC_HI	?	A			T
chr17_76839795	SLC38A10	3	2	0,627	0,973	-172,683	SOMATIC_HI	?	C			T
chr17_76839796	SLC38A10	3	2	0,634	0,984	-173,545	SOMATIC_HI	?	C			T
chr17_77439455	THOC4	3	2	0,653	1,014	-456,673	SOMATIC_HI	?	G			A
chr20_42341837	GDAP1L1	3	2	0,672	1,043	-465,294	SOMATIC_HI	?	G			A
chr20_43919430	ACOT8	3	2	0,664	1,032	-290,253	SOMATIC_HI	?	C			T

Figure 2

chr_pos (0 based)	gene	CN w/ EC	Cx	NAF=Nx_Q30/(Nx_Q30+Nz_Q30)	rho	FLRTtu	Site classification	Clone type	wt nt	mut nt	Essential
chr2_179534236	CCDC141	4	4	0,929	1,004	-22,8242	SOMATIC_HI	?	G	A	
chr7_12375969	AK027618	4	4	0,97	1,049	-633,922	SOMATIC_HI	?	C	T	
chr7_92598706	SAMD9L	4	4	0,917	0,991	-936,399	SOMATIC_HI	?	G	A	
chr7_134295381	CALD1,OK/sw-cl.14	4	4	0,934	1,01	-149,83	SOMATIC_HI	?	G	A	Y
chr7_140099604	BRAF	4	4	0,964	1,043	-97,1667	SOMATIC_HI	?	A	T	Y
chr7_150020820	GIMAP1,GIMAP2	4	4	0,929	1,005	-944,995	SOMATIC_HI	?	G	A	
chr10_50202438	C10orf71	2	2	0,881	1,024	-339,266	SOMATIC_HI	?	G	A	
chr10_96272140	TEC1D12	2	2	0,857	0,997	-70,1344	SOMATIC_HI	?	C	T	
chr10_96474217	CYP2C18,CYP2C19	2	2	0,834	0,97	-346,631	SOMATIC_HI	?	C	T	
chr10_96530337	CYP2C19	2	2	0,752	0,874	-141,696	SOMATIC_HI	?	C	T	
chr10_123960572	TACC2	2	2	0,794	0,923	-86,3343	SOMATIC_HI	?	C	T	
chr10_134068510	PWWP2B	2	2	0,905	1,052	-32,7025	SOMATIC_HI	?	C	T	
chr11_77405330	KCTD14	2	2	0,838	0,975	-135,186	SOMATIC_HI	?	C	T	
chr5_13863324	DNAH5	2	2	0,93	1,082	-69,8589	SOMATIC_HI	?	G	A	Y
chr5_27024137	GDH9	2	2	0,815	0,947	-35,6228	SOMATIC_HI	?	G	A	
chr5_35101295	PRLR	2	2	0,93	1,081	-322,993	SOMATIC_HI	?	C	T	
chr5_140167962	PCDHAL,PCDHA2,PCDHA3,PCDHA4	2	2	0,836	0,972	-235,309	SOMATIC_HI	?	T	C	
chr5_140533761	PCDHB7	2	2	0,845	0,983	-235,171	SOMATIC_HI	?	G	A	
chr6_46929663	GPR116	2	2	0,887	1,032	-214,582	SOMATIC_HI	?	C	T	
chr8_17107646	ZDHHC2	2	2	0,881	1,025	-150,986	SOMATIC_HI	?	T	C	
chr8_22323567	SLC39A14	2	2	0,852	0,991	-38,2677	SOMATIC_HI	?	C	T	
chr9_21961095	CDKN2A,MTAP	2	2	0,851	0,989	-94,7646	SOMATIC_HI	?	C	A	

Figure 3