

Importante studiare il capitolo del DNA ricombinante, sistemi di overespressione in batteri, ingegneria genetica delle piante e manipolazioni genetiche degli animali.

La terapia genica.

Manipolazione genica

Per manipolazione genica si intende la capacità di poter isolare frammenti di dna dal loro organismo ospite e di propagarli in un organismo differente. Quindi la manipolazione genica riguarda l'amplificazione di un tratto di dna di interesse, che viene in tal modo clonato per poi essere analizzato successivamente. L'operazione caratteristica della manipolazione genica è di conseguenza il **clonaggio**. Disponendo di una grande quantità di dna questo può essere sequenziato e si possono studiare le principali regioni regolative dell'espressione genica (promotori, siti di legame dei ribosomi, e **ORF**, ovvero regioni a registro di lettura aperto). Una volta identificato e sequenziato il tratto di dna si può risalire alla sequenza della proteina, applicando le regole del codice genetico. Grazie alla possibilità di clonare ed evidenziare il tratto di dna codificante è stato possibile affiancare accanto agli studi di genetica classica (mendel) quella che viene definita la "**genetica inversa**". Nella genetica classica gli studi iniziano dall'osservazione di un fenotipo alterato, per cercare poi di risalire al gene patologico responsabile di questo fenotipo alterato. La genetica inversa invece procede mediante un'analisi opposta; infatti avendo la possibilità di sequenziare i geni e conoscendo di conseguenza la loro sequenza aminoacidica, si possono introdurre delle mutazioni nel dna per poi andare ad osservare l'impatto che queste mutazioni hanno sul fenotipo

Questa è un'operazione molto importante che spesso viene applicata negli studi riguardanti una determinata proteina di interesse.

Quali sono gli strumenti della genetica inversa: la capacità di poter effettuare **mutagenesi in vitro** e la capacità eventualmente di distruggere in maniera mirata l'espressione del gene di interesse. Grazie a questa importante tecnologia si può, quindi, studiare la funzione di un gene all'interno di un organismo eliminando la sequenza codificante del gene di interesse per poi osservare gli effetti sul fenotipo dell'organismo analizzato. Questi studi vengono effettuati inizialmente in vitro, per poi studiare la funzione del gene in vivo.

Le tecniche di manipolazione genica nascono dalla possibilità di poter overesprimere una determinata proteina, riuscendo a costruire vettori specifici che consentano l'espressione mirata di una proteina, per arrivare poi a studiare gli effetti dell'inattivazione della proteina sull'animale di interesse; esiste inoltre la possibilità di effettuare un **targeting**, ovvero di bersagliare il gene di interesse in varie fasi dello sviluppo per cercare di capire pienamente le modalità di espressione del gene.

Per arrivare a questo c'è stato un percorso..

La nascita della biotecnologia risale alla nascita di un'industria americana, la Genentech, che inizia a manipolare i geni allo scopo di produrre proteine umane (come l'insulina o il plasminogeno) in sistemi procariotici, come batteri, per scopi farmacologici. La Genentech viene fondata nel '76 e produce nel '77 la prima proteina umana; la possibilità di poter produrre una proteina umana è stata molto importante perché ha aperto le porte all'impostazione di qualsiasi tipo di studio in vitro su qualsiasi gene. Una volta overespressa la proteina di interesse in vitro si possono fare molte operazioni, come creare un anticorpo specifico per evidenziare i tessuti in cui è espressa la proteina di interesse, ed una volta stabilita la sua localizzazione si può avere un'idea della funzione di questa proteina.

Negli studi biochimici lisando cellule, si isolano le proteine di interesse, che poi vengono analizzate. Invece la possibilità di poter osservare la localizzazione della proteina in vivo, e di identificare i suoi partner molecolari, permette uno studio più approfondito sulla funzione di questa proteina. In questo modo si possono identificare e analizzare i domini della proteina di interesse, ovvero quelle determinate regioni aminoacidiche che mediano le interazioni della proteina con i suoi principali

partner molecolari. Infatti spesso l'attenzione negli studi si focalizza su un particolare dominio proteico, per poi osservarne la specificità di interazione con il resto del proteoma.

Per arrivare a produrre una proteina umana si sono messe a punto una serie di tecniche e di approcci come l'identificazione dei vettori più adatti da utilizzare, e l'utilizzo di elettroforesi su gel di agarosio per separare i frammenti di dna linearizzati. Altre tecniche importanti sviluppate per ingegnerizzare una proteina di interesse sono ligazione, capacità di ibridare gli acidi nucleici, elettroporazione e trasformazione.

Un ruolo chiave nell'evoluzione delle biotecnologie è stato svolto da *E. coli*, poichè è economico e consente la possibilità di generare e manipolare rapidamente un grande numero di vettori.

Ricordiamo le caratteristiche dei vettori:

I vettori plasmidici sono dei repliconi, ovvero sono in grado di replicarsi in maniera autonoma.

La replicazione autonoma è fondamentale poichè garantisce l'amplificazione del dna di interesse.

Infatti una coltura batterica, trasformata con un plasmide contenente il frammento di dna di interesse, permette di ottenere un grande quantitativo del dna di interesse, successivamente recuperato attraverso la lisi batterica.

Il cDNA della proteina di interesse, che è stato introdotto nei batteri tramite il vettore plasmidico, può essere indotto alla traduzione in vitro, cioè può determinare la sintesi della proteina per cui esso codifica utilizzando le componenti del batterio.

I vettori plasmidici, inoltre, conferiscono alla cellula che li ospita dei vantaggi, come la resistenza agli antibiotici. PBR322 per esempio porta i geni che conferiscono la resistenza all'ampicillina e alla tetraciclina. Questi geni sono molto utili poichè permettono di selezionare soltanto le cellule batteriche che hanno ricevuto il plasmide, poichè saranno le uniche a crescere in presenza di un determinato antibiotico.

Questi vettori possono essere utilizzati anche per trasformare cellule eucariotiche in coltura, e anche in questi casi si possono selezionare soltanto le cellule che hanno acquisito il plasmide mediante introduzione di antibiotici sul terreno di coltura.

Un tipico vettore ha:

- Ori
- resistenza ad un antibiotico
- polilinker: sono siti multipli di clonaggio, in cui è presente la sequenza di riconoscimento di molti enzimi di restrizione. I polinker possiedono, infatti, varie sequenze palindromi, che possono essere riconosciute da vari enzimi di restrizione. Per avere un'efficienza di clonaggio del 100% si utilizzano enzimi di restrizione che siano in grado di determinare la formazione di estremità coesive complementari, sia nel frammento esogeno da clonare che nel vettore plasmidico all'interno del quale deve essere inserito il frammento. Questo si può fare o utilizzando lo stesso enzima di restrizione, in modo tale che i tagli procurati su entrambe le molecole generino delle estremità sporgenti complementari, oppure utilizzando enzimi diversi ma che riconoscono lo stesso sito di restrizione, gli isoschizomeri.

Vettori di espressione

I vettori di espressione consentono la trascrizione dell'inserto, e quindi l'espressione del gene di interesse. Il dna che si inserisce all'interno di vettori di espressione è sempre **c-DNA**, ovvero dna codificante, derivante da retrotrascrizione. Infatti trasformare batteri con inserti che siano dotati anche delle sequenze introniche non è rilevante, poichè lo scopo della trasformazione in cellule batteriche con vettori di espressione è la produzione del gene di interesse. Ci sono dei modi per cui si può massimizzare l'espressione del gene di interesse come proteina.

Fattori che influenzano l'espressione di un gene di interesse:

- Forza del promotore: il livello con cui il promotore è attivo. Inoltre se il plasmide è high copy, ovvero è presente in grande numero all'interno della cellula batterica trasformata, si avrà un grande quantitativo di mRNA prodotto, e quindi il gene di interesse sarà espresso in maniera molto elevata.
- Stabilità del plasmide. Alcuni plasmidi possono essere instabili e vengono talvolta persi nelle varie divisioni successive a cui vanno incontro le cellule batteriche trasformate, e questo porta alla perdita conseguente del frammento di dna di interesse. In genere si usa un plasmide stabile, high copy e con promotori estremamente efficienti. Devono essere utilizzati vettori che siano in grado di produrre un grande quantitativo di proteina, a livello di milligrammi di proteina, a partire da pochi ml di coltura batterica. Tuttavia l'espressione di una proteina eterologa all'interno del batterio potrebbe comportare diversi problemi per la cellula, come il **codon usage**, la scelta dei codoni; infatti nonostante il codice genetico sia universale si è notato che nei sistemi procariotici alcune triplette vengono utilizzate in maniera preferenziale rispetto ad altre, nella codifica di uno stesso aminoacido. Per ovviare a questo sistema sono stati ingegnerizzati i batteri ospiti in modo che esprimano tutti i t-RNA eucariotici. In questo modo si sono arricchite le potenzialità dell'apparato traduzionale batterico.

L'espressione della proteina eterologa può essere tossica per il batterio, quindi è importante un promotore regolabile: i batteri vengono fatti crescere, dopodiché attivando il promotore si induce l'espressione soltanto della proteina di interesse, dopodiché il batterio muore.

Il promotore deve essere, quindi, inducibile.

Fondamentale è il promotore dell'operone LAC, in quanto il gene di interesse da overesprimere risulta sotto il suo controllo.

Nell'operone Lac è presente in maniera costitutiva un repressore associato al promotore, che ne inibisce la trascrizione, e nel momento in cui si fornisce IPTG, questa molecola funge da induttore, e legandosi al repressore lo disattiva dal promotore, riattivando la trascrizione del gene di interesse.

L'aggiunta di IPTG alle cellule batteriche trasformate determina, quindi, l'attivazione dell'espressione della proteina di interesse.

Produzione della somatostatina

La somatostatina è stato il primo ormone peptidico umano prodotto in vitro da batteri ricombinanti. Questa proteina è costituita da 14 aminoacidi. Viene sintetizzata la sequenza oligonucleotidica che codifica per questi 14 aminoacidi. Tuttavia essendo molto piccola, e di conseguenza instabile, questa proteina una volta espressa all'interno di cellule batteriche potrebbe essere facilmente degradata dalle proteasi batteriche. Per prevenire la degradazione da parte di proteasi si è pensato di proteggere la proteina eterologa umana utilizzando una **fusione**, ovvero fondendola con un'altra proteina, di origine batterica. Mediante il processo di fusione due proteine vengono fuse insieme attraverso i rispettivi cDNA, ovvero le rispettive sequenze codificanti, che sono inserite e legate insieme all'interno del vettore. Quindi la sequenza che codifica per la proteina di interesse, sotto forma di cDNA, viene inserita subito a valle della sequenza che codifica per un'altra proteina, la quale per una serie di caratteristiche che possiede si vuole ottenere fusa alla proteina di interesse.

Questa situazione garantisce che nella parte aminoterminale della proteina di fusione è presente una proteina batterica che il batterio riconosce e quindi non degrada, a cui si aggiunge una coda di 14 aa che corrisponde alla somatostatina la quale, in questo stadio fuso, risulta irrilevante per le proteasi batteriche.

Inoltre un elemento molto importante che protegge la proteina dalle proteasi è il **foldig**. La proteina di interesse deve, quindi, essere espressa foldata. Bisogna quindi evitare che la proteina di interesse vada incontro a denaturazione.

Si utilizzò come parte della proteina di fusione batterica la betagalattosidasi, la cui espressione è sottoposto alla regolazione da parte del operone lac. Dopo aver espresso all'interno di batteri

trasformati la proteina di fusione, ovvero betagalattosidasi a cui si è legata la somatostatina, attraverso un taglio con bromuro di cianogeno si separa la somatostatina dalla betagalattosidasi, ed in questo modo la somatostatina viene purificata e quantizzata.

Riassunto dei diversi passaggi:

Si genera un concatenamero di oligonucleotidi della sequenza di interesse, che corrisponde alla proteina di fusione, in cui i diversi oligonucleotidi sintetizzati si legano l'uno all'altro, poichè sono parzialmente complementari e sovrapponibili; la parziale complementarietà che è presente tra gli oligonucleotidi li fa incontrare concatenare in una molecola unica.

Questo concatenamero viene, poi, clonato all'interno di vettore, che è poi inserito all'interno di cellule batteriche. Successivamente le cellule trasformate vengono fatte crescere, ed in seguito alla crescita si aggiunge IPTG, che induce l'espressione della proteina di fusione, che è regolata dal promotore dell'operone lac. Infine, in seguito a lisi cellulare, la somatostatina è separata dalla betagalattosidasi mediante bromuro di cianogeno.

In laboratorio oggi si usa come proteina di fusione la glutatione transferasi, una proteina che si può estrarre molto facilmente dalla coltura, e che garantisce la protezione della proteina di interesse a cui è legata. Inoltre, come si è osservato in laboratorio, la GST rappresenta, per la sua alta affinità al glutatione, un ottimo sistema di selezione della proteina di fusione a partire da un estratto proteico, contenente diverse proteine.

Produzione dell'insulina

La particolarità dell'insulina è che ha dei **legami disolfuro**, che nei batteri non sarebbero consentiti, poichè i batteri non hanno l'ambiente riducente necessario. Per ovviare a questa situazione si overesprimono catene separate, di cui ciascuna catena è legata ad una proteina batterica, la betagalattosidasi; quindi anche in questo caso si ricorre a proteine di fusione. Dopo essere state overesprese come catene separate, mediante un taglio la catena A e la catena B vengono separate dalla betagalattosidasi, e poi riunite insieme in vitro per ristabilire la proteina completa.

Riassunto delle caratteristiche di un vettore di espressione:

Promotore con ATG (codone di inizio che corrisponde ad un primo aminoacido di Met iniziale). Nel polilinker, ovvero nella regione in cui è presente il sito di restrizione per l'enzima utilizzato, viene inserito il cDNA di interesse, che corrisponde ad una proteina di fusione, in cui una proteina batterica è legata ad una proteina di interesse, che successivamente verrà separata dalla proteina batterica, mediante una reazione di taglio. Tuttavia anzichè utilizzare un metodo chimico, che prevede una reazione di taglio dei legami peptidici attraverso il **bromuro di cianogeno**, per dividere la proteina di fusione nelle due componenti di base, si è introdotto un sistema differente e più specifico, che prevede l'utilizzo di **proteasi**. Mediante questo sistema si estrae la proteina di fusione, e la si inserisce in tubo dove viene aggiunta una proteasi specifica. Questa proteasi, attraverso il riconoscimento di sequenze specifiche bersaglio, è in grado di separare la proteina di interesse dalla proteina batterica a cui è fusa. Per utilizzare questo meccanismo di separazione mediante proteasi si costituisce un concatenamero di cDNA, costituito dal cDNA della proteina batterica a cui è legato subito a valle il cDNA della proteina di interesse; nello spazio interposto tra queste due sequenze, che codificano per le due proteine, si inserisce un tratto di dna che codifica per la sequenza aminoacidica riconosciuta dalla proteasi.

Il vettore di espressione più evoluto è, quindi, costituito da un promotore inducibile, un cDNA che codifica per una proteina di fusione, ed una sequenza riconosciuta da una proteasi, localizzata tra i cDNA della prima proteina e quello della seconda.

Ripetizione delle caratteristiche dei vettori :

- Hanno proprietà di base dei plasmidi
- Alcuni di questi si basano sul batteriofago lambda
- Possono essere a singolo filamento
- Alcuni vettori sono utilizzati per l'introduzione di grandi inserti, fino a 2Mb (Yac)
- **Vettori specializzati, come i vettori di espressione : questi vettori di espressione consentono l'overespressione di proteine ricombinanti** (argomento principale del corso della Gonfloni).

Vantaggi del dna plasmidico

Innanzitutto la disponibilità di un metodo per purificare grandi quantità di dna plasmidico. Questo metodo si basa sul fatto che i vettori sono a basso peso molecolare rispetto al cromosoma batterico. Attraverso delle estrazioni specifiche si riesce ad arricchire soltanto il vettore plasmidico all'interno della soluzione e di eliminare il cromosoma batterico. I vettori inoltre conferiscono all'ospite una resistenza a determinati antibiotici, ovvero fenotipi facilmente selezionabili.

I vettori plasmidici, inoltre, contengono **multiple cloning sites**, ovvero un **polilinker**, un tratto di DNA in cui è presente la sequenza di riconoscimento di molti enzimi di restrizione; questo consente una scelta molto accurata degli enzimi di restrizione da utilizzare per il taglio del dna plasmidico e del dna esogeno da clonare.

I vettori di espressione permettono di overesprimere una proteina di interesse. Per overesprimere una proteina di interesse si usano sistemi inducibili come l'operone lac, un sistema che si basa sul fatto che aggiungendo IPTG questa molecola funziona come un induttore, che si lega al repressore e ne favorisce il cambiamento di conformazione e la sua separazione dall'operatore; attraverso questa separazione il promotore non è più ingombrato dal repressore e si attiva in questo modo l'espressione del gene di interesse. Tuttavia la produzione di proteine eterologhe è spesso tossica per i batteri. Infatti nel caso in cui è presente una situazione di **leakness**, ovvero un'attivazione costitutiva del promotore della proteina eterologa, dopo il raggiungimento di un livello basale di produzione della proteina di interesse, i batteri potrebbero non sopravvivere, ed in questo modo si perderebbero i batteri prima di produrre al loro interno la quantità di proteina eterologa desiderata; l'obiettivo della trasformazione con vettori di espressione è, infatti, la produzione di un grande quantitativo di proteina di interesse, a livello di milligrammi, ben oltre il livello basale. Per ovviare a questo problema si ingegnerizzano dei sistemi inducibili altamente controllati, ed in cui l'espressione della proteina di interesse è sottoposta ad un duplice controllo. In questo modo i batteri possono essere fatti crescere ad una densità elevata, e soltanto dopo questa intensa crescita della coltura batterica si induce un'espressione massiccia della proteina di interesse eterologa.

Evoluzione dal promotore di lac ad un sistema inducibile controllato:

Il sistema bifasico

Si costruisce un vettore di espressione che ha un promotore diverso da lac, e che è attivabile in seguito all'espressione di una specifica polimerasi. Esistono organismi particolari, come i fagi, che sono in grado di infettare i batteri e di indurli alla produzione delle proteine virali. I fagi hanno dei promotori molto efficaci e sono più semplici dei batteri; hanno, infatti, un sistema di trascrizione molto più semplice in cui è sufficiente soltanto una RNA POL. Questa RNA POL riconosce in maniera molto efficace il promotore fagico. Grazie a queste componenti si è pensato di prendere il cDNA codificante per la proteina di interesse, e di clonarlo in frame, insieme ad un'altra proteina per dare origine ad una proteina di fusione, ma che è sottoposta al controllo non più del promotore lac ma del promotore fagico di **T7**. A questo punto una volta inserita la sequenza codificante all'interno del vettore di espressione, ed una volta trasformati i batteri con questo vettore, per essere espressa la proteina di fusione necessita della polimerasi di T7, che è normalmente assente all'interno dei batteri. Per sintetizzare questa rna pol fagica si è pensato di inserire all'interno del cromosoma batterico il gene che codifica per la polimerasi di T7, sottoforma di cDNA. In questo modo si costruisce in laboratorio un ceppo batterico **BL21** che all'interno del cromosoma nella regione DE3 ha inglobato il gene codificante per la polimerasi di T7. Questa polimerasi è inoltre sottoposta al controllo del promotore di Lac; quindi si giunge alla produzione della polimerasi di T7 quando si introduce nel terreno **IPTG**. Grazie all'introduzione nel terreno di IPTG, la RNA POL di T7 è attiva e può di conseguenza attivare l'espressione della proteina di fusione, il cui promotore è sottoposto al suo controllo.

Quindi da una parte, in seguito all'aggiunta di IPTG, si ha un batterio competente che può esprimere T7 RNA POL; dall'altra parte il cDNA della proteina di fusione ha un **T7 promoter**, ovvero un promotore sottoposto a regolazione da parte dell' RNA POL di T7.

In conclusione nella **prima fase** l'inserimento di IPTG porta all'espressione della T7 RNA POL; infine nella **seconda fase** la T7 RNA POL riconosce il suo promotore a monte del cDNA di interesse e trascrive in questo modo la proteina di fusione.

La T7 pol ha un'altissima affinità con il suo promotore, precedentemente inserito a monte del cDNA di interesse; quindi subito dopo l'introduzione di IPTG nel batterio si determina la sintesi di un grande quantitativo della proteina di fusione.

Questo sistema bifasico è molto più sicuro ed efficiente della regolazione da parte dell'operone lac, ed inoltre elimina il problema del **leakness costitutivo**. Infatti se la proteina di interesse fosse regolata dal promotore lac, anche una minima espressione della proteina di fusione ad un livello basale, a causa dell'eventuale presenza di alcune molecole di lattosio nel terreno, porterebbe a morte i batteri prima di produrre un quantitativo sufficiente di proteina. Il doppio controllo, invece, fa sì che la produzione della proteina di fusione sia veloce e massiccia.

Tuttavia nonostante l'alta efficienza di questo sistema bifasico si potrebbero avere delle complicazioni. Infatti una presenza di lattosio nel terreno porterebbe ad un'espressione basale costitutiva della polimerasi di T7, quindi a problemi di leakness. Il persistere di questa situazione farebbe sì che le poche molecole pol fagica, attivate da eventuali residui di lattosio, andrebbero ad attivare un'espressione basale della proteina di fusione, portando a morte i batteri senza un'espressione della proteina a livelli massicci. Per superare questo problema si è modificato ulteriormente il ceppo batterico BL21DE3, che a questo punto è diventato altamente ingegnerizzato allo scopo di una produzione massiccia della proteina di fusione. All'interno del batterio BL21DE3 (cioè batteri BL21 contenenti sequenze che codificano per l' RNA pol di T7 nella regione DE3) è stato inserito un piccolo vettore plasmidico che esprime un lisozima ingegnerizzato. Questo lisozima, che è una peptidasi, è stato ingegnerizzato allo scopo di tagliare soltanto la molecola T7 RNA POL.

Quindi, attraverso questo **triplice** controllo, anche se ci fosse un'espressione basale della polimerasi fagica, questa presenza sarebbe annullata dall'azione del lisozima. Soltanto attraverso l'introduzione massiccia di IPTG si supera l'effetto del lisozima, e si esprime una grande quantità di polimerasi fagica, che a sua volta attiva e trascrive in maniera massiccia il cDNA di interesse.

Bioreattori (classica domanda d'esame insieme alle proteine di fusione)

Un **bioreattore** è un animale modificato geneticamente e che consente la produzione di proteine ricombinanti per molti scopi, a partire da quelli farmaceutici. Un bioreattore è da considerarsi, quindi, una fabbrica di proteine.

Quando si genera un animale transgenico si cerca di veicolare la proteina di interesse in elementi che possono essere prelevati dall'animale senza ucciderlo, come nel latte, nel siero sanguigno o nelle urine.

Uno speciale bioreattore è la pecora:

La pecora viene utilizzata come bioreattore per la produzione di **plasminogeno**, un ormone umano che viene estratto dal latte di questo animale.

Se lo scopo del bioreattore è che la proteina di interesse venga espressa nel latte, si utilizzerà un promotore **tessuto-specifico**, un promotore che è molto attivo nelle cellule responsabili della produzione del latte, e che codifica per una proteina molto abbondante nel latte di pecora, come la **betagalattoglobulina ovina**. Per fare in modo che questo animale esprima la proteina di interesse, il plasminogeno, si ingegnerizza il dna codificante per la proteina di interesse, mettendo questa sequenza di dna sotto il controllo di un promotore che sia molto attivo nel tessuto dell'animale in cui si vuole overesprimere la nostra proteina di interesse.

Si utilizza, quindi, un promotore tessuto-specifico che sia attivo in maniera costitutiva, per arrivare alla produzione di una grande quantità di proteina target.

Si utilizzano le tecniche del dna ricombinante per generare un vettore di espressione adeguato. Questo vettore di espressione viene costruito in laboratorio dove si inserisce il gene di interesse sotto il controllo del promotore della betagalattoglobulina ovina, una proteina presente nel latte di pecora. In seguito ad una reazione di restrizione tra vettore ed inserto (che codifica per l'ormone umano) si effettua un clonaggio, e si ottiene un vettore ricombinante, che contiene il gene di interesse, sottoposto al controllo del promotore della betagalattoglobulina.

A questo punto il vettore viene iniettato all'interno del pronucleo maschile. In questo processo di iniezione l'inserzione del vettore è multipla e casuale e l'animale risultante è completamente **transgenico**, poichè tutte le cellule derivanti dalle divisioni successive dell'ovocita fecondato avranno al loro interno l'inserto, ovvero il **transgene**.

Generalmente i vettori vengono linearizzati, poichè questo facilita la loro inserzione; inoltre l'inserzione è quasi sempre multipla ed il transgene può essere presente, quindi, in copie multiple, in tandem.

A questo punto l'ovocita viene inserito nell'ovidotto di una madre vicaria, che partorirà un animale completamente transgenico, da cui si otterrà latte contenente l'ormone umano.

Questo tipo di operazione è molto grossolana, in quanto essendo multiple e casuali, le inserzioni del transgene potrebbero essere dannose, poichè potrebbero avvenire in zone importanti del genoma dell'organismo manipolato. Questo tipo di manipolazione, infatti, non è utile ai fini dello studio della funzione del transgene, ma soltanto per utilizzare l'organismo manipolato come bioreattore, come una sorta di fabbrica per produrre la proteina di interesse.

Per formare bioreattori si utilizza soltanto una cassetta di dna, ovvero un tratto di dna capace di essere trascritto anche nell'organismo in cui è stato iniettato, poichè è regolato da un promotore dell'organismo ospite che è espresso costitutivamente in un determinato tessuto.

Manipolazione delle piante. (non saranno argomento d'esame).

Le capacità di veicolare dna negli organismi eucariotici possono essere schematizzate in questo modo:

- Trasformare
- microiniettare
- Usare virus
- Bombardare con particelle metalliche di tungsteno rivestite di dna (spesso si usa sui fusti delle piante)

Per ottenere piante transgeniche si usano plasmidi speciali, chiamati **TI**.

Per manipolare le piante si usano batteri speciali, come l'agrobacterium tumefaciens, che attaccano alcuni tipi di piante, e che hanno al loro interno enormi plasmidi difficili da manipolare.

Le piante transgeniche si ottengono mediante l'uso di alcuni plasmidi chiamati **plasmidi TI** (tumor inducing), poichè causano tumori nel colletto delle piante. Il plasmide di questo batterio è molto grande, di **200 kb**, quindi è difficile da manipolare, tagliare e purificare (se si vuole clonare in genere si usano i PUC18, che sono vettori da 3 kb, mentre i vettori espressione arrivano fino a 6-7Kb).

Questo plasmide ha una regione importante che si chiama **T-DNA**, e che ha la speciale capacità di inserirsi nel cromosoma della piante, come se fosse una sorta di trasposone. Quindi si sfrutta la capacità del T-DNA di inserirsi nel cromosoma delle piante.

Questo plasmide viene manipolato, disarmandolo, ovvero eliminando diverse sequenze del T-DNA ed introducendo all'interno della stessa regione un gene di interesse, che potrebbe essere per esempio un gene per la resistenza all'erbicida. Il gene di interesse mediante le capacità di trasferimento della regione T-DNA, riesce ad essere inserito all'interno della pianta. In questo modo si formano piante transgeniche capaci di resistere a stress di vario tipo.

I plasmidi TI conferiscono al batterio un vantaggio metabolico poichè hanno dei geni che consentono la sintesi delle **opine**, delle sostanze che consentono al batterio di avere una fonte di azoto.

Da un plasmide TI molto grande, di circa 200Kb, si passa ad un plasmide disarmato, che ha soltanto il **braccio L** ed in cui sono state eliminate diverse sequenze di dna, come quella relativa alla formazione del tumore e il **braccio R**. Il braccio L ed il braccio R sono le estremità di sinistra e di destra del T-DNA. Attraverso coniugazione batterica il plasmide **disarmato** viene fatto ricombinare con un vettore intermedio, in modo da ottenere un plasmide **cointegrato** che abbia al suo interno il gene di interesse e le due braccia L ed R che devono essere trasferite come T-DNA. Il vettore cointegrato contiene anche due geni per la resistenza alla **kanamicina** e alla **spectinomicina** (provenienti dal vettore intermedio), utili per selezionare i plasmidi ricombinanti.

L'inserzione del T-DNA in questo caso è casuale ma è un evento unico, e avviene in un solo punto. Quindi la pianta che si genera non è completamente transgenica, ma è un eterozigote, poichè il T-DNA è stato inserito soltanto in uno dei due cromosomi omologhi. Per avere una pianta completamente transgenica, ovvero che in entrambe le copie cromosomiche abbia inserito il T-DNA, è necessario fare un autofecondazione, dove per un 25% dei casi si darà origine ad un omozigote per la regione T-DNA.

(ragionare sulla figure sul griffith a pag 441 e 442)

La manipolazione genetica degli animali (cap 20 dello Strachan)

La scorsa lezione sono stati analizzati i bioreattori, ed i gruppi che si possono utilizzare per esprimere proteine eterologhe nei batteri; si è poi visto come si manipolano geneticamente gli animali e si sono accennate le piante transgeniche.

La scorsa volta si sono, inoltre, analizzati i diversi passaggi che portano alla formazione di un bioreattore, un animale completamente transgenico, inserendo il transgene nei pronuclei, e da questo ottenendo un oocita fecondato transgenico, il quale trasmetterà il transgene a tutte le cellule che da esso avranno origine, e l'organismo adulto sarà di conseguenza completamente transgenico. In questa lezione si studierà in dettaglio la manipolazione delle cellule staminali, e come da esse si arriva alla produzione di un animale transgenico. Manipolando cellule staminali embrionali totipotenti e inserendo queste cellule modificate all'interno di una blastocisti si giunge alla formazione di un embrione **chimerico**, che contiene due popolazioni di cellule differenti, una derivante dalla blastocisti ed una derivante dalle cellule embrionali transgeniche impiantate. In questo modo l'embrione chimerico se reinserito in una madre vicaria può dare origine ad un **animale chimerico**, un mosaico di diversi tipi cellulari, in cui alcune cellule contengono il transgene mentre altre non lo contengono. Se il transgene è presente in parte delle cellule della linea germinale potrà essere trasmesso a parte della progenie. Poiché le cellule embrionali transgeniche possono dar vita a parte della linea germinale, a partire da una chimera, attraverso incroci successivi si può ottenere un animale completamente transgenico.

Manipolando, invece, direttamente l'oocita fecondato, attraverso una microiniezione nei pronuclei, si ottiene un animale completamente transgenico.

Sistema della microiniezione: si preparano delle pipette di vetro estremamente sottili, di cui una è utilizzata per trattenere l'oocita, mentre l'altra pipetta, da microiniezione, ha una punta finissima che viene utilizzata per penetrare prima nell'oocita fecondato e poi nel pronucleo maschile, al cui interno è pipettata una soluzione acquosa del DNA contenente il transgene. Dopo l'estrazione dalla micropipetta gli oociti sopravvissuti vengono, infine, impiantati negli ovidotti.

Quando si vuole soltanto **aggiungere** un gene si modifica direttamente l'oocita fecondato, modificandolo attraverso l'aggiunta di un transgene. Se, invece, si vuole effettuare una **mutagenesi mirata** di un gene specifico, l'operazione da svolgere diventa più complessa e si deve iniziare a lavorare con cellule staminali embrionali.

Quindi le manipolazioni dell'oocita, essendo più grossolane, vengono utilizzate generalmente soltanto allo scopo di produrre proteine di interesse attraverso bioreattori, poiché l'inserzione del transgene in questi casi è casuale e a volte dannosa. Al contrario se si vuole studiare la funzione di un gene in maniera approfondita, bisogna lavorare su cellule staminali embrionali, eliminando la copia originaria del gene di interesse nell'animale modello utilizzato, e inserendo al suo posto il transgene che si vuole analizzare. Quest'operazione avviene attraverso una **ricombinazione omologa** tra l'allele originario dell'animale modello con l'allele alterato che si vuole inserire al suo posto.

Per produrre animali transgenici si può fare una microiniezione del transgene nell'oocita fecondato, all'interno del pronucleo maschile, ottenendo un numero variabile di copie del transgene ad inserzione casuale. Quest'operazione causa, inoltre, delezioni e riarrangiamenti del dna genomico limitrofo all'inserzione. La microiniezione è, quindi, un'operazione molto grossolana.

Per studiare al meglio la funzione di un gene di interesse bisogna inventare un modo per inserire il transgene nella maniera piu "fisiologica" possibile, e per farlo si sono evoluti dei vettori, come i retrovirus ricombinanti, che sono in grado di entrare all'interno di cellule staminali embrionali in singola copia, senza alterare il dna genomico fiancheggiante all'inserzione; questi vettori devono inoltre eliminare tutti gli elementi virali di regolazione che potrebbero avere degli effetti negativi. In sintesi il transgene di interesse deve essere introdotto all'interno delle cellule bersaglio senza danni, e una volta manipolate, queste cellule devono essere introdotte in una blastocisti, ottenendo un embrione chimerico, che darà origine ad un animale chimerico. Per studiare la funzione del transgene l'animale deve essere fatto incrociare con animali chimerici dello stesso tipo, allo scopo di ottenere una linea pura; infatti soltanto attraverso una linea pura si può studiare in maniera approfondita la funzione del gene di interesse, poichè in questo modo il gene è espresso in tutte le cellule dell'organismo, che di conseguenza risulta completamente transgenico.

Formazione di animali chimerici

Nella formazione di animali chimerici un allele di interesse wild type viene sostituito con un allele alterato per osservarne la funzione; a questo scopo si compie una **ricombinazione omologa**, nel locus del gene di interesse. Per studiare approfonditamente la funzione di un gene questo può essere inattivato, ed un modo di inattivarlo è quello di inserire al posto di un suo esone particolare una **cassetta neo**, ovvero un gene che se espresso conferisce la resistenza ad un antibiotico. Quest'operazione è molto utile perchè quando si modificano le cellule staminali con un vettore al cui interno è presente l'esone che si vuole inattivare, e all'interno di questo esone è presente la cassetta neo, si possono selezionare soltanto le cellule che hanno ricevuto l'inserito, poichè saranno resistenti ad un determinato antibiotico.

L'obiettivo è quello di ottenere un animale completamente transgenico utilizzando la metodologia delle cellule staminali. Per inattivare l'espressione di un gene di interesse si fa in modo che al posto di un esone bersaglio venga inserita una cassetta neo. Mediante tecniche di ingegneria genetica si genera un vettore che possiede al posto dell'esone bersaglio una cassetta neo che consente l'espressione di una proteina che conferisce resistenza alla **neomicina**. Tuttavia nelle regioni fiancheggianti di questa cassetta neo vengono mantenute delle regioni di omologia con l'esone bersaglio. Si mantengono, quindi, delle regioni di omologia a destra e a sinistra della cassetta neo per permettere una successiva ricombinazione omologa nel gene di interesse proprio in queste due regioni che fiancheggiano l'esone bersaglio.

All'interno di una piastra petri le cellule staminali embrionali vengono trattate con vettori linearizzati, per favorire l'evento di ricombinazione omologa.

Le cellule vengono trattate nella piastra con vettori. Tuttavia, a differenza del lievito, negli animali la ricombinazione omologa non è molto efficiente, quindi l'operazione di sostituzione dell'esone deve essere molto precisa.

Teoricamente potrebbero verificarsi tre situazioni differenti in seguito a ricombinazione omologa:

- potrebbe avvenire una ricombinazione omologata mirata, con un doppio scambio nelle due regioni di omologia fiancheggianti l'esone bersaglio. Al posto dell'esone bersaglio viene inserita la cassetta neo e lo scambio risulta perfetto.
- L'inserzione dell'esone potrebbe essere casuale, e avvenire in qualunque parte del cromosoma, e questo potrebbe essere dannoso per le cellule.
- L'inserzione potrebbe non avvenire affatto.

Per risolvere questo problema si è inventato un sistema di **doppia selezione**: accanto alla cassetta neo, che conferisce resistenza alla neomicina, si introduce un altro marcatore, il gene **TK**, un gene molto importante che codifica per la **timidin-kinasi**, una proteina dell'herpes. L'inserimento del **ganciclovir** attiva la timidin kinasi che a sua volta fosforila il ganciclovir; il ganciclovir nello stato trifosfato si converte in un potente inibitore della DNA polimerasi, e questo porta le cellule a morte. Per ottenere soltanto eventi di ricombinazione omologa mirata, ovvero doppi scambi nelle regioni di omologia fiancheggianti la cassetta neo, si selezionano le cellule in presenza contemporanea sia di ganciclovir che di neomicina. Infatti se è avvenuta un inserzione casuale il vettore ha inserito nel cromosoma anche il gene TK, e le cellule di conseguenza, nonostante siano resistenti alla neomicina, in presenza di ganciclovir muoiono. L'obiettivo, invece, è quello di selezionare soltanto l'evento che ha consentito l'inserzione della cassetta neo, ma che contemporaneamente ha provocato la perdita del gene TK, presente vicino alla cassetta neo. Infatti soltanto un evento di ricombinazione omologa mirata, con un doppio scambio nelle due regioni di omologia fiancheggianti l'esone bersaglio, potrebbe causare il contemporaneo inserimento della cassetta neo e la perdita del gene TK fiancheggiante la cassetta neo.

La ricombinazione omologa dipende dalle regioni di omologia: maggiori sono queste regioni più è alta è la probabilità di ricombinazione omologa.

ripetizione..

L'obiettivo è quello di ottenere un animale transgenico e la strategia che si utilizza è quella della manipolazione delle cellule staminali embrionali, per inattivare un gene di interesse al fine di studiarne la funzione. A questo scopo il locus di interesse viene marcato e allo stesso tempo inattivato sostituendo un suo esone con una cassetta neo, cioè una cassetta che conferisce una resistenza alla neomicina, ma alle cui estremità vengono mantenute delle regioni di omologia con lo stesso esone che ha sostituito, ovvero con l'esone con cui deve ricombinare nella cellula ospite. Grazie alla proteina che conferisce resistenza alla neomicina (codificata dalla cassetta neo) l'inserzione del gene nel cromosoma della cellula embrionale staminale può essere riconosciuta selezionando in coltura le cellule che sopravvivono in presenza di neomicina. Le cellule resistenti saranno quelle che hanno inglobato la cassetta neo. Tuttavia per selezionare ulteriormente soltanto le cellule staminali embrionali che hanno inserito la cassetta neo in corrispondenza del sito esatto, ovvero in corrispondenza dell'esone bersaglio nel gene di interesse che si vuole disattivare mediante una ricombinazione omologa mirata, si sfrutta un ulteriore sistema di selezione basato sulla presenza di un gene tk che fianeggia la cassetta neo, introdotta nel vettore linearizzato. Essendo la presenza del gene tk letale in presenza di ganciclovir, soltanto le cellule che avranno subito un evento di ricombinazione omologa sopravviveranno ad una presenza contemporanea sul terreno sia di neomicina che di ganciclovir; ovvero soltanto quelle cellule che hanno avuto un doppio scambio nelle due regioni di omologia fiancheggianti la cassetta neo, e che avranno di conseguenza (con il secondo scambio) eliminato il gene tk che fianeggia la cassetta neo, mentre avranno acquisito nel cromosoma la cassetta neo al posto dell'esone bersaglio, e la proteina sarà in questo modo inattivata.

Ennesima ripetizione..

Il dna può essere introdotto in una cellula per trasduzione virale e mediante l'uso di liposomi, delle vescicole artificiali che sono in grado di veicolare nella cellula il dna che è presente al loro interno. Queste cellule devono essere selezionate una volta manipolate e per selezionarle si utilizzano dei marcatori. L'obiettivo è quello di selezionare soltanto le cellule in cui è avvenuta ricombinazione omologa. Per ricombinazione omologa si intende che l'esone originale all'interno del gene di interesse è stato sostituito precisamente con un esone ingegnerizzato in laboratorio, ed in cui è stata

eliminata una parte codificante per la proteina di interesse, ed al suo posto è stata introdotta una cassetta neo, una sequenza codificante per una proteina che conferisce resistenza alla neomicina.

Il vettore di targeting è linearizzato, perché questo favorisce gli eventi di ricombinazione omologa. L'esone bersaglio, che appartiene al gene di interesse che si vuole inattivare per analizzarne la funzione, è stato modificato in laboratorio, e al posto della sua normale regione codificante è stata inserita una regione di dna che codifica per una proteina che conferisce resistenza ad un antibiotico, la cassetta neo. (la ricombinazione omologa dipende dalle regioni di omologia: maggiori sono queste regioni più è alta è la probabilità di ricombinazione omologa). Il gene della neomicina è molto piccolo, quindi a destra e sinistra del gene neo si inseriscono delle consistenti regioni di omologia con l'esone bersaglio. Per guidare questa ricombinazione dopo aver preparato il vettore, lo si fonde con i liposomi, dopo che il dna è stato complessato con i liposomi questi possono essere introdotti nelle cellule staminali attraverso pinocitosi, e mediante questo inserto le cellule della coltura che hanno ricevuto il vettore potranno sopravvivere in presenza di neomicina. Tuttavia se l'inserzione è avvenuta in un punto casuale con molte probabilità le cellule avranno acquisito anche il gene TK fiancheggiante la cassetta neo, e questo sarebbe letale per le cellule in presenza di ganciclovir. Al contrario con un'inserzione esatta in corrispondenza del locus bersaglio originario, mediante una ricombinazione omologa, le regioni di omologia fiancheggianti la cassetta neo favoriscono un doppio scambio che permette da una parte di acquisire l'esone bersaglio modificato (inattivando così il gene di interesse e raggiungendo l'obiettivo) mentre dall'altra il gene TK viene eliminato dal secondo scambio, e le cellule sopravviveranno in presenza di ganciclovir.

Si inferisce attraverso questo sistema di doppia selezione che le cellule staminali embrionali totipotenti sopravvissute ad una presenza contemporanea di neomicina e ganciclovir corrispondono a cellule che hanno acquisito l'inserto nella posizione esatta, e che quindi sono state inattivate correttamente nel gene di interesse da una mutagenesi mirata in vitro.

Quando si effettua manipolazione genica degli animali si può:

- 1) utilizzare questa tecnica per studiare la funzione di un gene di interesse
- 2) si possono generare degli animali modello per lo studio di patologie

Per ottenere un animale transgenico si può:

- 1) agire sull'ovocita fecondato, ottenendo un animale completamente transgenico ma in cui l'inserzione del transgene è probabilmente casuale.
- 2) agire sulle staminali embrionali: si modifica il dna in maniera mirata, e poi le cellule manipolate vengono introdotte nella blastocisti nelle prime fasi di sviluppo embrionale, ottenendo un'alta probabilità di ottenere un animale chimerico, ma che poi, con reincroci, permette di riottenere una linea pura, soltanto nel caso in cui le cellule staminali manipolate che sono state introdotte nella blastocisti abbiano dato origine a parte della linea germinale.

Si può pensare di inattivare un gene, sempre attraverso una ricombinazione omologa, ovvero attraverso l'introduzione di un esone modificato al posto dell'esone funzionale, ma utilizzando questa volta un marcatore differente. L'esone modificato può contenere al posto di un gene per la resistenza ad un antibiotico, una cassetta contenente una sequenza (in cDNA) che codifica per una proteina che dà fluorescenza, ed in questo modo gli animali transgenici ottenuti saranno fluorescenti!

Questo è un modo diretto per osservare direttamente l'animale che è stato modificato.

Nella maggior parte dei casi l'eliminazione di un intero esone di una proteina (sostituito con una cassetta che codifica per tutt'altro) porta alla completa inattivazione della proteina.

Quindi per disattivare un gene spesso basta sostituire un esone con un gene cassetta, con una sorta di etichetta che permette la selezione delle cellule correttamente alterate nella regione di interesse.

Ci sono 2 tipi di vettori:
(figura 21.4 strachan)

Un **vettore d'inserzione** modifica il locus di interesse mediante un singolo evento di ricombinazione reciproca, che causa l'inserzione del marcatore ma anche dell'intera sequenza vettoriale fiancheggiante. Con un vettore di inserzione viene attivato un meccanismo anche detto "**colpo e fuga**" o **hit&run**, in cui il gene viene inattivato dall'inserimento del marcatore, come nel caso della cassetta neo. Questo è il metodo più sicuro per determinare mutazioni **knock-out**. Tuttavia è un metodo poco preciso poiché si inseriscono nel dna della cellula trasdotta molte inutili sequenze di dna vettoriale oltre al marcatore.

Un **vettore di sostituzione**, invece, è molto più mirato e specifico rispetto ad un vettore di inserzione, ed è utilizzato per inattivare un gene di interesse, sostituendo una sequenza codificante, come un esone, con un esone modificato ed ingegnerizzato contenente un marcatore (come la neomicina), attraverso una ricombinazione omologa, mediante *due eventi* di scambio nelle regioni di omologia fiancheggianti l'esone; questo doppio scambio consente l'acquisizione soltanto dell'esone modificato e non della sequenza vettoriale, che invece viene tagliata in un punto esterno alla regione omologa.

Se si vuole fare una sostituzione altamente mirata che permette lo scambio di singoli aminoacidi il procedimento è più complicato; in questo caso si vuole determinare una mutazione più specifica che consente di studiare soltanto il cambiamento di un singolo aminoacido all'interno di un esone bersaglio.

Si vuole sostituire un esone bersaglio con una copia dello stesso esone, ma che contiene una singola mutazione puntiforme, che determina il cambiamento di un singolo aminoacido, per studiarne gli effetti sul gene e sulla proteina corrispondente.

A questo scopo si effettua una **doppia sostituzione**. (fig 21.5 strachan).

Per determinare il cambiamento di un singolo nucleotide si utilizza un vettore **knock-out** di sostituzione, un vettore linearizzato che possiede il gene TK all'esterno della regione di omologia ed una cassetta **HPRT** nello spazio interposto tra i due siti di omologia dell'esone bersaglio che si vuole sostituire; all'interno delle regioni di omologia avvengono i crossover con la regione bersaglio, nel gene di interesse.

Il gene HPRT codifica per un enzima umano, il quale una volta introdotto nel topo può essere misurato e selezionato sulla base dell'affinità con un determinato substrato che viene inserito nelle cellule. Grazie a questo inserto si possono selezionare soltanto le cellule **murine** che hanno introdotto il gene umano, le quali saranno in grado di produrre l'enzima umano e di interagire con il substrato introdotto.

Il primo passaggio consiste quindi nell'introduzione di una cassetta HPRT al posto di un esone bersaglio nel gene di interesse, utilizzando un vettore di sostituzione knock-out.

Come si intuisce dalla figura 21.5, in caso non di una doppia ma di una singola ricombinazione tra vettore knock-out ed esone bersaglio, la cellula acquisirebbe anche le regioni vettoriali fiancheggianti, tra cui il gene TK, e quindi in presenza di gancicovir la cellula non sopravviverebbe. Questo sistema permette di selezionare le cellule che hanno ricevuto soltanto HPRT dal vettore, mediante un doppio scambio nelle due regioni di omologia con l'esone, fiancheggianti HPRT.

Successivamente le cellule vengono di nuovo trattate con un ulteriore vettore di sostituzione, che possiede l'esone bersaglio normalmente fiancheggiato dalle sue caratteristiche regioni di omologia, ma che è alterato con la mutazione puntiforme desiderata. In questo modo si inserisce al posto della cassetta HPRT l'esone originario del gene di interesse alterato in un singolo nucleotide, per analizzare gli effetti di una mutazione puntiforme e di una conseguente singola mutazione aminoacidica nella funzione del gene di interesse. Infine le cellule che erano state precedentemente selezionate per la loro capacità di interagire con il substrato grazie alla cassetta HPRT, ora verranno controselezionate per verificare che siano state effettivamente modificate con il secondo vettore di sostituzione, e che abbiano subito un doppio evento di ricombinazione, eliminando la cassetta HPRT e inserendo al suo posto l'esone bersaglio modificato con la mutazione puntiforme desiderata, che si vuole ottenere in queste cellule.

Quindi in conclusione basta selezionare soltanto le cellule che hanno perso la capacità di produrre l'enzima umano e che di conseguenza non sono più in grado di interagire con lo specifico substrato che è stato introdotto nella coltura.

La natura è ridondante: L'inattivazione di un singolo gene spesso non basta per capire la funzione della proteina corrispondente, perché spesso un altro gene è in grado di complementare e di sopperire alla funzione mancante del gene che è stato silenziato; quindi spesso accade che mutazioni knock-out non hanno effetti fenotipici evidenti proprio a causa della natura ridondante del genoma, a causa della presenza nel genoma di almeno un altro gene che è in grado di svolgere correttamente la funzione del gene che è stato inattivato.

Queste tecniche di ricombinazione omologa sono molto complesse e laboriose e purtroppo hanno una frequenza molto bassa poiché la ricombinazione omologa è un evento raro; si è pensato, quindi, di evolvere questa tecnica e utilizzare sistemi molto più mirati per far avvenire una ricombinazione sito-specifica. Si è cercato di utilizzare non più vettori con grandi regioni di omologia per un sito specifico, sperando poi di ottenere una ricombinazione omologa, ma piuttosto si è cercato di forzare in qualche modo la ricombinazione. (nella prossima lezione questi sistemi molto più efficaci saranno maggiormente approfonditi)

Terapia genica

Applicazioni delle tecnologie della mutagenesi mirata nella terapia genica.

Con terapia genica si intende l'insieme di procedimenti svolti a fini terapeutici, per curare una malattia modificando geneticamente le cellule dei pazienti, allo scopo di sostituire il gene alterato che causa la patologia, con il corrispondente sano.

Se si può miratamente distruggere un gene mediante l'inserimento di una cassetta di neomicina, allora allo stesso modo si può pensare di curare una malattia sostituendo la parte danneggiata di un determinato gene.

Possibilità della terapia genica:

- alterare un oocita fecondato. (non applicato)
- compiere una terapia genica sulla linea germinale (mai sperimentata sull'uomo poiché illegale)
- terapia genica somatica (l'unica forma attualmente consentita ed in sperimentazione)

Nella **terapia genica somatica** si modificano geneticamente le cellule somatiche per cercare di curare una determinata patologia o almeno per alleviarla il più possibile. Mediante questa terapia vengono prelevate cellule somatiche dal paziente affetto; queste cellule vengono poi opportunamente trattate **ex vivo** in coltura, manipolate e reiniettate nel tessuto del paziente in cui è manifestata fenotipicamente la patologia. La terapia genica può agire anche **in vivo**, operando direttamente nei tessuti del paziente, trasferendo mediante vettori appropriati la copia sana di un determinato allele all'interno delle cellule affette, per sostituirlo alla forma alterata o per addizionarlo al genoma in caso di totale assenza del locus.

La terapia genica è nata allo scopo di curare determinate malattie di cui si conoscevano le basi molecolari e può essere applicata per la cura dei tumori, delle malattie ereditarie (che spesso sono dovute alla mancanza o all'espressione di una forma alterata di un gene), delle malattie infettive e delle malattie del sistema immunitario.

Sono state studiate diverse strategie, come la possibilità di combattere patologie dovute a geni alterati, i quali agiscono come caratteri mendeliani dominanti, cercando abolire l'espressione dell'allele mutante.

Si può pensare anche di eliminare un allele alterato sostituendolo con l'allele sano. E' necessario stabilire innanzitutto il gene da trasferire e le modalità del suo trasferimento; bisogna quindi utilizzare il vettore più adatto, che sia in grado di veicolare l'allele sano nelle cellule che presentano lo stesso allele nella forma alterata che causa la patologia. Bisogna stabilire, inoltre, il tessuto bersaglio del paziente su cui si vuole applicare la terapia genica. Una volta stabilito il tessuto bersaglio si può stabilire una strategia appropriata:

- 1) in vivo: si applica la terapia genica direttamente nel paziente
- 2) ex vivo: si prelevano cellule dal paziente, si trattano in coltura, si manipolano e si reiniettano nel tessuto del paziente

Il primo esempio di terapia genica è stato applicato a dei piccoli topi little-little. Si è iniettato all'interno di una madre vicaria un oocita fecondato manipolato, con all'interno un transgene che contiene un ormone della crescita di ratto, e che è sotto il controllo di un promotore della metalotioneina. In questo modo fornendo metallo il topo aumenta di dimensioni poichè esprime l'ormone codificato dal transgene.

Nella decisione della strategia di terapia genica da applicare bisogna porsi alcune domande:

- 1) si può identificare il gene responsabile della malattia?
- 2) si possono selezionare le cellule bersaglio per mantenerle in coltura e manipolarle?
- 3) qual'è il tipo di sistema più efficace per compiere un trasferimento genico?
- 4) quante cellule devono acquisire il gene?
- 5) il gene espresso in eccesso è dannoso?

E' importante constatare che le cellule manipolate una volta reintrodotte nel tessuto hanno una vita limitata, per cui la terapia genica somatica è un approccio locale e temporaneo che permette, almeno per ora, soltanto di alleviare gli effetti di una patologia e non di estirparla completamente da un tessuto.

Alcune malattie, come la **fibrosi cistica**, non possono prevedere trattamenti di terapia genica ex vivo. La fibrosi cistica è una patologia che riguarda l'epitelio polmonare ed un'approccio efficace per combatterla consiste nell'utilizzare una terapia genica in vivo, utilizzando come vettori gli adenovirus, introdotti nell'organismo mediante aerosol. Una volta giunti nei polmoni questi vettori

si inseriscono nelle cellule dell'epitelio polmonare e permettono l'introduzione del transgene, il quale può ripristinare la funzione corretta del gene alterato nella fibrosi cistica.

Tipi di vettori che si utilizzano per il trasferimento genico

Non esiste il sistema di trasferimento genico ideale; ciascuno presenta i propri limiti e i propri vantaggi; inoltre la tipologia scelta dipende molto dalla natura del tessuto bersaglio. I vettori che si utilizzano per la terapia genica possono andare incontro a diversi problemi. Tuttavia i vettori virali si sono dimostrati molto efficienti nel trasferimento di geni in cellule umani, mediante **trasduzione**.

I **retrovirus** sono virus ad RNA che possiedono una funzione di trascrittasi inversa, che li rende capaci di sintetizzare una forma di dna complementare. Questi virus si integrano nel dna soltanto quando le cellule sono in attiva divisione, e di conseguenza quando la membrana nucleare è dissolta. Questi vettori, quindi, entrano molto efficacemente nelle cellule e possono integrarsi al DNA; tuttavia questi virus integrandosi si inseriscono casualmente nel DNA. I retrovirus sono, inoltre, molto difficili da produrre in quantità elevate in coltura, e possono veicolare inserti grandi soltanto fino ad **8Kb**. Nonostante queste limitazioni sono considerati comunque i vettori più promettenti per veicolare dna alle cellule bersaglio. Inoltre la capacità di trasdurre soltanto cellule in divisione può essere molto vantaggiosa nella terapia genica dei tumori di tessuti normalmente composti da cellule che non si dividono; le cellule cancerose potrebbero essere così uccise in modo selettivo, preservando le cellule sane.

In un retrovirus si lasciano le estremità 5' e 3' LTR, si eliminano i geni virali non indispensabili all'inserzione, come relativi al capsido e si inserisce al loro posto il gene terapeutico.

Gli adenovirus, che sono virus a DNA che producono infezioni delle vie respiratorie, nel processo di infezione non si integrano al dna ma esistono in forma episomica, non integrata, e quindi l'espressione del gene inserito può andare perduta con la divisione cellulare. Gli adenovirus vengono veicolati all'interno delle cellule bersaglio riconoscendo determinati recettori, mediante **endocitosi mediata da recettore**.*

Per essere utilizzati come vettori questi virus vengono modificati, mantenendo i geni virali indispensabili per l'inserzione nelle cellule bersaglio ma eliminando i geni che codificano per funzioni dannose e per le proteine relative al capsido, ed al loro posto viene inserito il gene di interesse che si vuole veicolare nelle cellule bersaglio.

Gli adenovirus possono essere prodotti in quantità elevate in coltura e possono infettare un'ampia gamma di tipi cellulari, comprese le cellule che non si dividono. Inoltre essendo virus grandi hanno anche la capacità di accettare grandi inserti di dna esogeno, fino a **35Kb**; tuttavia la delezione di geni virali può essere controproducente poiché in questo modo gli adenovirus resistono meno alle cellule immunitarie e perdono parzialmente la loro grande efficienza di trasduzione. Un'altro svantaggio degli adenovirus è che possono causare infiammazioni croniche, come accade quando sono utilizzati nella terapia genica contro la fibrosi cistica.

*Per modificare un adenovirus in modo ad hoc si cerca di lavorare sul suo tropismo (sulla sua capacità di riconoscere recettori specifici) cercando di modificare il suo involucro esterno in modo tale che riconosca soltanto alcuni tipi di recettori; in questo modo si è operata una selezione sul tipo di cellule bersaglio.

Le ricerche sui vettori virali sono molto intense, allo scopo di costruire il vettore più adatto per un determinato tessuto bersaglio.

Come i retrovirus gli adenovirus possono avere delle importanti applicazioni contro le cellule tumorali. Le cellule tumorali spesso espongono sulla loro superficie di membrana dei recettori anomali rispetto al fenotipo sano di quel tessuto. Sfruttando la conoscenza della struttura di questi recettori alterati, vettori virali possono essere ingegnerizzati allo scopo di riconoscerli e di penetrare nelle cellule cancerose procurandone la morte.

Virus adeno-associati. I virus adeno-associati sono un gruppo di piccoli virus a DNA a singolo filamento. Sono molto utili poichè sono in grado di integrarsi con il cromosoma in un punto specifico, che non è dannoso. Tuttavia non riescono ad effettuare un'infezione senza la collaborazione di un virus helper co-infettante, che ne facilita l'ingresso nella cellula bersaglio. Un altro svantaggio che hanno questi vettori è che possono accettare soltanto piccoli inserti, grandi fino a 4,5Kb.

Herpesvirus. Sono vettori che consentono infezioni dei neuroni e sono in grado di portare inserti grandi fino a 20Kb; tuttavia non si integrano al DNA e quindi l'espressione a lungo termine dei geni trasferiti non è possibile.

Immunoterapia ed effetto bystander.

Nasce per curare le malattie ereditarie, ma nelle cellule tumorali ha molto successo.

Adenovirus modificati potrebbero essere indirizzati all'interno di cellule tumorali riconoscendo gli specifici recettori alterati, e introducendo al loro interno, come transgene, geni che causano morte cellulare, come TK, un potente inibitore della DNA POL, in presenza di ganciclovir. Inoltre attraverso le **gap junctions**, il ganciclovir passa alle cellule circostanti provocando quindi anche la morte delle cellule che non hanno ricevuto l'inserto.

Immunodeficienza.

Portiamo l'esempio della **SCID**, in cui mancano i linfociti T funzionali.

Nei pazienti manca l'enzima **ADA** necessario al metabolismo della tirosina. Essendo una patologia che interessa i linfociti è stato possibile compiere una terapia genica **ex vivo**, prelevando i linfociti del paziente. Questi linfociti sono stati messi in coltura e trattati con dei retrovirus che avevano come inserto il gene ADA. Al paziente è stato allo stesso tempo dato anche l'enzima ADA nel sangue; quindi attraverso questa azione combinata, in cui da una parte l'enzima mancante veniva fornito direttamente nel sangue, e dall'altra veniva prodotto attraverso i linfociti ingegnerizzati che sono stati manipolati e reintrodotti nell'organismo, il paziente seppure con dei cicli continui di terapia può vivere una vita più o meno normale.

Un esempio di trattamento di terapia genica **in vivo** è quello della **fibrosi cistica**. La fibrosi cistica è una malattia autosomica recessiva, che prevede un deficit nel trasporto di ioni Cl. Poichè il tessuto bersaglio corrisponde alle cellule dell'epitelio polmonare l'unica opzione possibile è la terapia genica in vivo, che prevede l'utilizzo di vettori adenovirali, oppure dei liposomi, che sono delle membrane artificiali. Purtroppo i pazienti hanno sviluppato una infiammazione cronica e generalmente questo approccio ha fallito. Tuttavia se soltanto il 10% delle cellule trasdotte riuscisse a ripristinare il canale normale del cloro, si potrebbero dare vantaggi importanti alle cellule adiacenti poichè gli ioni Cl verrebbero pompate dalla cellula resa sana dalla trasduzione. Questo potrebbe dare grandi benefici al paziente.

Un'altra malattia genetica trattata mediante terapia genica è l'**ipercolesterolemia familiare**, una patologia dominante. Questa patologia è stata trattata attraverso una terapia genica **ex vivo**.

Nei pazienti affetti i geni che codificano per i recettori per le lipoproteine sono mutati e non vengono espressi. Nel 92' per un paziente omozigote sono stati prelevati gli epatociti, si sono fatti crescere in coltura, sono stati infettati con retrovirus, e poi reinseriti nel sistema venoso portale del paziente. Chiaramente quest'intervento è soltanto temporaneo, poichè le cellule hanno una vita di circa 18 mesi, per cui periodicamente il trattamento dev'essere ripetuto.

Il successo della terapia genica dipende da una serie di fattori:

Essa dipende innanzitutto dall'efficacia dei vettori, e dal tipo di **eredità**. Le eredità più adatte sono quelle recessive, perchè i malati hanno un prodotto genico molto limitato, per cui anche bassi livelli di espressione dei geni reintrodotti possono produrre un effetto positivo. Basta pensare al caso della fibrosi cistica, in cui anche il 10% delle cellule trasdotte può dare un grande sollievo al paziente. Al contrario le malattie dominanti sono le meno adatte per la terapia genica, poichè basta la mutazione su un singolo allele per avere un forte effetto fenotipico patologico. In questi casi non ci sono molte probabilità di stabilire una strategia curativa adeguata. Infatti se la malattia è dominante cioè indica che è insorta una mutazione in un locus genico tale da produrre una proteina alterata, che ha perso la propria funzione originaria, e quest'alterazione prevale sull'altra copia del gene, che invece è sana. In questi casi l'unica possibilità è quella di impedire l'espressione del gene mutato; si può impedire l'espressione di un gene mutato a vari livelli: a livello di DNA, a livello dell'mRNA, e a livello della proteina.

Un altro elemento molto importante per il successo della terapia genica è anche la natura della mutazione; nei casi di una mutazione che comporta perdita di funzione, si può pensare di aumentare semplicemente il numero di copie geniche, veicolando nella cellula bersaglio copie multiple del gene di interesse.

Anche l'accessibilità delle cellule bersaglio è molto importante poichè più sono accessibili più è facile manipolarle in vivo con vettori virali.

Altra cosa rilevante è la dimensione del DNA codificante dell'inserito, poichè minori sono le dimensioni del gene di interesse da veicolare, maggiore è la facilità con cui questo può essere inserito nel vettore e poi trasferito nella cellula bersaglio.

Infine fondamentale è il controllo dell'espressione del gene di interesse. A volte, come negli adenovirus, il gene terapeutico veicolato è dotato di un promotore autonomo, in modo che la sua espressione sia inducibile e tale da garantire un certo effetto positivo.

Le cellule staminali vengono trattate con **vettore di targeting**, vettori che possono portare alla distruzione di un gene di interesse. Si erano considerati precedentemente diversi casi di vettori, i vettori di inserzione, di sostituzione, e di doppia sostituzione.

Per introdurre delle mutazioni ci sono, quindi, 2 possibilità:

- 1) **hit&run**: si usa un singolo **vettore di inserzione**, perché si inattiva il gene in maniera netta, "brutale", ed al posto di un esone si inserisce un marcatore, come una cassetta codificante per esempio per una proteina che conferisce resistenza alla neomicina.
- 2) **tag&exchange (marcatura e scambio)**: Si utilizzano **vettori di sostituzione**. Si marca la regione di interesse scambiandola con un altro gene (come HPRT), e poi si fa avvenire una seconda sostituzione. Con lo scambio si introduce la mutazione desiderata, eliminando contemporaneamente i marcatori. Infatti l'operazione migliore consiste nell'inattivazione di un gene inserendo al posto di un esone corretto un esone modificato, ma tuttavia senza

introdurre nell'organismo delle ulteriori sequenze di dna estraneo, ovvero le sequenze del vettore fiancheggianti l'esone modificato. Quindi l'ideale di un gene knockout sarebbe poter inattivare un gene senza dover introdurre al suo interno le sequenze del vettore, a partire dai marcatori, ma semplicemente sostituendo l'esone bersaglio di interesse con un esone modificato, che può differire anche per una singola mutazione puntiforme.

(fig 21.4 strachan)

Quindi nel caso 1 di un vettore di inserzione si disattiva un gene facendo avvenire un singolo evento di ricombinazione tra il gene bersaglio e la parte linearizzata del vettore, che possiede una regione di omologia con il gene bersaglio.

Un vettore d'inserzione, quindi, modifica il locus di interesse mediante un singolo evento di ricombinazione reciproca, che causa l'inserzione del marcatore ma anche dell'intera sequenza vettoriale fiancheggiante. Questa disattivazione causa, quindi, l'inserzione di molte sequenze prive di interesse, di dna estraneo appartenente al vettore.

Invece nel caso di un vettore di sostituzione il procedimento è più mirato rispetto a quello di un vettore di inserzione, ed avvenendo un doppio crossing over, l'inserimento di sequenze vettoriali nel cromosoma della cellula trasdotta è limitato alla regione del marcatore. Inoltre le cellule che hanno avuto l'inserzione saranno selezionabili in quanto resistenti alla neomicina (poichè la regione del marcatore codifica per una proteina che conferisce resistenza ad un antibiotico).

(fig 21.5 strachan)

Nel caso, invece, di una doppia sostituzione l'obiettivo è introdurre al posto di un esone bersaglio, un esone modificato che differisce per un singolo nucleotide. La sostituzione deve essere quindi estremamente mirata e annullare totalmente la possibilità di inserire altre sequenze vettoriali all'interno del gene bersaglio.

Il primo passaggio di una doppia sostituzione consiste nell'introduzione di una cassetta HPRT al posto dell'esone bersaglio nel gene di interesse, utilizzando un vettore di sostituzione knock-out. Successivamente le cellule vengono di nuovo trattate con un ulteriore vettore di sostituzione, che possiede l'esone bersaglio normalmente fiancheggiato dalle sue caratteristiche regioni di omologia, ma che è alterato con la mutazione puntiforme desiderata. In questo modo si inserisce al posto della cassetta HPRT l'esone originario del gene di interesse alterato in un singolo nucleotide, per poi analizzare gli effetti di questa mutazione puntiforme e di conseguenza di una conseguente singola mutazione aminoacidica nella funzione del gene di interesse.

Il passaggio intermedio, di modificazione attraverso HPRT, è fondamentale perchè è l'unico modo per selezionare le cellule modificate nell'esone bersaglio, altrimenti sostituendo direttamente l'esone bersaglio normale con quello mutato in un singolo nucleotide, non si potrebbero distinguere le cellule modificate dalle cellule normali, poichè in assenza di marcatori inseriti, le cellule differiscono solo per un esone e al suo interno soltanto per un singolo nucleotide! Per questo è necessario selezionarle prima con HPRT e nel secondo passaggio fare una controselezione per eliminare da una parte quelle che ancora mantengono la capacità di legare il substrato di HPRT, e per evidenziare dall'altra quelle che non legano più il substrato, poichè ciò indica che hanno acquisito tramite ricombinazione l'esone modificato con la mutazione puntiforme desiderata (anche chiamato esone*, star).

Applicazioni di Genetica

Abbiamo fino adesso fatto il capitolo 21 dello strachan (non abbiamo fatto gli animali transcromosomici, né la transgenesi mediata da YAC). Ci sono alcuni capitoli che riguardano il *knock in* e *knock out* che sono argomenti specifici e quindi non sono nell'esame. Gli animali transgenici e i bioreattori sì. Il capitolo 22 della terapia genica, c'è un paragrafo sui vettori oncovirali (definizioni di adenovirus (virus a dna) e retrovirus (virus a rna)) che è da leggere e ricordare. Andiamo a studiare le tecniche della terapia genica classica (ex vivo e in vivo), e cos'è un vettore episomico o un vettore che si integra nel cromosoma; come viene preso l'adenovirus da una cellula mediante un recettore ed endocitosi, e ci fermiamo. Infine c'è da sapere cosa è ADA, la terapia genica fatta sulla scid, la fibrosi cistica e la ipercolesterolemia familiare.

Ricapitoliamo gli argomenti.

Come abbiamo detto la volta scorsa manipolare geneticamente gli animali consente di studiare la funzione genica di un gene, e quindi posso creare modelli animali di malattie.

Ci sono due sistemi che si possono andare a manipolare:

1. l'oocita fecondato
2. cellule staminali embrionali (che sono totipotenti, e quindi possono dare vita sia a una linea somatica che a una germinale).

Quando si manipola un oocita fecondato nel dna si inserisce il nostro vettore e quindi l'animale è completamente transgenico e l'inserzione è molto spesso multipla. Con una micropipetta infilo in un pronucleo il vettore linearizzato che va ad inserirsi nel cromosoma, ho una integrazione. Queste integrazioni sono spesso multiple. L'oocita manipolato viene poi messo in una madre vicaria. Otteniamo un animale completamente transgenico; questo tipo di manipolazione è grossolana: non ho controllato dove si è inserito il mio dna, ci sono molte copie e non so dove si vanno a mettere, (potrei addirittura andare a targettare una zona importante per l'animale). Questa tecnica grossolana la uso per generare magari *bioreattori*, animali che producono proteine farmaco.

Se voglio studiare la funzione di un gene devo mettermi in un contesto più fisiologico possibile. Per studiare la funzione di un gene posso:

- inattivare completamente il gene
- modifico il gene per inserire una mutazione, fare un *knock in*, cioè inserire una mutazione nel gene per vedere qual è l'effetto di quella mutazione (ad esempio se voglio dimostrare che nella data malattia è coinvolto il dato gene che ha la data mutazione ricreo la stessa situazione in un modello animale).

Quando vado a studiare una situazione fisiologica vado a manipolare le cellule staminali embrionali del topo con *vettori di targeting*; seleziono le cellule che hanno ricevuto il mio transgene e poi le reinietto in una blastocisti, che va messa in una mamma vicaria. Alcune di queste cellule possono andare a colonizzare la linea germinale, quindi ottengo un animale mosaico, chimerico, eterozigote, in cui alcune cellule che hanno e attraverso reincroci riesco a ottenere una linea pura, che è importante nel caso in cui devo studiare una mutazione recessiva (magari l'omozigosi è letale; magari se fai knock out di un gene questo non ha effetto perché la ridondanza dei geni e le famiglie geniche ti mascherano il knock out).

Quali vettori per il gene targeting.

Sostituzione genica: intendo quando faccio una ricombinazione omologa, cioè nel locus del mio gene sostituisco un gene alterato. L'ideale sarebbe non avere nessun pezzo di vettore, ma sostituire

solo la cassetta del gene senza altri elementi di disturbo. In questo caso quindi è proprio una sostituzione, una ricombinazione omologa.

Addizione genica: se invece non uso la ricombinazione omologa ma infilo un pezzo di transgene nelle cellule, sto inserendo più copie di uno stesso gene e mi chiedo cosa succede (analisi molto grossolana). Nell'addizione genica potrei anche ad andare ad inserire il mio transgene in zone importanti (magari per la regolazione...), posso creare dei grossi danni.

L'espressione di un gene si può distruggere (fare il knock out) andando ad utilizzare vettori diversi di inserzione o di sostituzione; quando vogliamo applicare queste manipolazioni alla terapia genica, sostituiamo un allele mutante con uno selvatico per migliorare una patologia (e questa è la terapia genica).

I vettori che garantiscono l'introduzione di mutazioni sono di due tipi. In un caso abbiamo sostituzione dell'intero vettore, comprensivo di marcatore, e lo otteniamo con un semplice o doppio scambio. Abbiamo inserito negli esoni del nostro gene dei pezzi che provengono dal vettore e il gene marcatore. Se voglio fare una sostituzione più precisa, al limite di una sostituzione amminoacidica, marchiamo la zona intorno all'esone che vogliamo mutagenizzare. Marcare vuol dire che al posto dell'esone (ad es) 8 del gene, metto con un vettore di knock out un gene marcatore (per es HPRT); le cellule, che all'inizio erano HPRT⁻. Vogliamo che ci sia una ricombinazione omologa tra le regioni di omologia a monte e a valle della box 8; ottengo cellule HPRT⁺ tk⁻ (è il sistema di selezione positivo/negativa: se voglio che avvenga un evento di ricombinazione omologa utilizzo due marcatori, uno tk a una estremità del vettore che mi assicura di poter controselezionare tutte le volte che ho inserzione casuale del mio vettore, perché in questo caso insieme al mio marcatore viene inserito anche il gene tk, che però è un gene suicida perché se aggiungiamo nel mezzo ganciclovir la cellula muore). Quindi per fare in modo di avere una ricombinazione omologa dobbiamo selezionare sia per il marcatore che per l'altro all'estremità.

(Il vettore di targeting ha nell'esone di interesse il marcatore neo e il tk a una estremità. Quando tratto le staminali con questo vettore posso avere tre situazioni: nessuna inserzione, inserzione casuale (e se è casuale è altamente probabile che si inserisca l'intero vettore, e quindi accanto al marcatore neo ci sia anche il marcatore tk nel cromosoma) oppure una ricombinazione mirata nelle regioni di omologia, in cui avviene uno scambio preciso in cui al posto dell'esone X c'è la cassetta neo e nel contempo c'è la perdita di tk, per cui selezionando queste cellule sia in presenza di neomicina che di ganciclovir ottengo soltanto le cellule che hanno avuto l'inserzione voluta).

In questo modo prendiamo solo le cellule che hanno sostituito nel locus di interesse l'HPRT e perso contemporaneamente tk. Abbiamo così marcato la zona che vogliamo mutagenizzare e poi facciamo avvenire una seconda ricombinazione omologa usando un vettore che ha le regioni di omologia e l'esone 8 modificato e selezioniamo contro queste cellule, quindi torniamo al fenotipo originario HPRT⁻, di tutte quelle che hanno perso il gene nella zona.

In breve: Se voglio che ci sia ricombinazione omologa uso un vettore con due marcatori: uno mi seleziona per le cellule che hanno ricevuto l'inserto (magari è un enzima che fa qualcosa di particolare, come HPRT), e uno che mi assicura che non ci sia inserzione casuale, e questo è una sequenza suicida all'estremità dell'inserto (che se è inserzione casuale entra, se è precisa nel punto dove deve andare non entra).

Nelle cellule di mammifero la ricombinazione omologa è un evento molto raro, bisogna in qualche modo forzarlo.

La mutagenesi mirata avviene tramite ricombinazione omologa. Essendo un evento molto raro aumenta quando il grado di omologia di sequenza tra il dna introdotto e il gene bersaglio è molto elevata. Maggiore è il grado di omologia più elevata sarà la frequenza.

Per semplificare l'identificazione degli eventi di ricombinazione omologa i vettori contengono un marcatore, es il gene neo che permette selezione a favore delle cellule che hanno incorporato il dna estraneo. Per semplificare faccio un trucco: faccio sì che il gene neo sia automaticamente inserito nel mio locus di interesse così posso selezionare direttamente le cellule che hanno incorporato il dna estraneo. È fondamentale avere un'etichetta all'interno del gene.

Una volta si inducevano mutazioni e si andava a guardare il fenotipo dell'animale per capire a posteriori quale gene era mutato, invece con questa tecnologia inseriamo una sequenza nota di dna estraneo, transgene, abbiamo un vantaggio sulle mutazioni indotte, perché lasciamo un "etichetta", il gene per esempio della neomicina (il gene neo), da cui posso partire per andare a caratterizzare quello che c'è intorno mediante speciali PCR (race pcr) e in questo modo posso "camminare" lungo il locus in esame e capisco cosa c'è intorno e le funzioni di queste regioni.

Importante ricordarsi che una volta che avviene ricombinazione omologa si ottiene un animale chimerico e questo va incrociato al fine di ottenere una linea pura.

Cosa è possibile studiare con questa tecnologia.

Si può pensare di studiare gli elementi di regolazione che agiscono in *cis*. Possiamo andare a guardare con un gene indicatore (tipo lacZ) gli elementi in *cis* (stavolta al posto di neo mettiamo lacZ), come la funzione genica di un determinato gene andando a vedere cosa succede quando lo inattivo miratamente.

Terapia genica.

Per la terapia genica avevamo visto che si può manipolare l'animale. Non si può curare un uovo fecondato né fare una terapia genica della linea germinale; quello che si può fare è una terapia genica della linea somatica su staminali e iniettarlo nell'animale.

Bisogna decidere se usare un sistema *in vivo* (direttamente agiamo sull'individuo) o *ex vivo* (agisco su cellule in coltura).

È importante ricordare che tipo di vettori si utilizzano:

- Retrovirus: virus a rna che retrotrascrivono l'rna a dna complementare, che si va ad inserire nel genoma.
 - contro: inserzioni molto piccole. Si inseriscono casualmente, provocando problemi.
- Adenovirus: virus a dna. Non si integrano. Sono comodi per una terapia genica di particolari tessuti come l'epitelio respiratorio nella fibrosi cistica, perché hanno un grande tropismo per questi tipi di cellule.
- Adenovirus associati: sono meglio degli adenovirus perché si integrano, e anche in un punto specifico (non causando nessun tipo di problema di inserzione casuale come nei retrovirus).
- Liposomi: altro modo per veicolare dna, non sono invasivi. L'adenovirus può causare infiammazione cronica, il liposoma no. Questi veicolano dna all'interno di una cellula perché si fondono con la membrana.

La costruzione di un vettore.

Vediamo come funziona la costruzione di un vettore pensando ai retrovirus. Questi hanno delle regioni che codificano per il capsido, e poi hanno bisogno di trascrittasi inversa per sintetizzare l'rna in dna complementare. In laboratorio si levano tutte le proteine strutturali del virus per avere spazio per inserire il nostro gene terapeutico. Tutte le cassette di inserzione del gene terapeutico possono avere promotori molto specifici, anche inducibili e tessuto specifici. Un modo semplice per avere espressione elevata della proteina di interesse nelle cellule bersaglio è mettere a monte della sequenza della proteina un promotore tessuto specifico o inducibile. Si usano promotori inducibili con tetraciclina. Ha molta affinità per le cellule in divisione, in cui si infila quando la membrana plasmatica si apre.

Per l'adenovirus è importante ricordare che, a differenza dei retrovirus, attacca cellula anche non in divisione usando dei recettori. Quindi per manipolare e migliorare i vettori si ingegnerizzano i recettori o le proteine esterne del virus in modo da poter dividere e modificare il prodotto.

Topo little-little.

Il primo caso di terapia genica applicato a un mammifero. In questo caso la terapia genica prevedeva di modificare il gene del controllo dell'ormone della crescita causando aumento della dimensione del topolino. In quel caso si manipolava un oocita fecondato il costruito aveva il gene dell'ormone della crescita sotto il controllo del promotore della metallo tioneina. L'oocita modificato veniva poi impiantato in una mamma vicaria. Il risultato era che la presenza di copie di questo gene determinava un aumento del peso: se davi metallo il topo diventava più grosso.

Il caso SCID.

Altro esempi di terapia genica (in questo caso ex vivo) applicata è quello della SCID, i cui affetti non sono in grado di fare l'enzima Ada per cui si prendono cellule in coltura che sono espanse e trattate con vettori retrovirali per poi essere reiniettate nel paziente. Questo sistema non dura tutta la vita ma solo per la durata della vita delle cellule e quindi va rifatta.

• Ricombinazione sito-specifica.

Finora abbiamo visto che è possibile manipolare gli animali in modo molto efficiente con ricombinazione omologa. Affinché la ricombinazione omologa abbia frequenza più elevata si possono estendere le regioni di omologia (maggiori sono e maggiore sarà la frequenza di ricombinazione). C'è però la possibilità di utilizzare un sistema diverso. Questa si basa sulla disponibilità di ricombinanti e le dimensioni e specificità delle sequenze bersaglio. La ricombinazione omologa avviene grazie a proteinasi endogene e c'è bisogno di lunghe regioni di omologia; la sito-specifica è specializzata perché possiamo utilizzare enzimi particolari che non appartengono alla cellula eucariotica, e sono capaci di riconoscere sequenze specifiche di piccole dimensioni.

Si può pensare di utilizzare delle *ricombinasi*. Ci sono degli enzimi che sono capaci di riconoscere delle sequenze bersaglio, e in qualche modo consentire la ricombinazione tra due sequenze bersaglio. Si possono ingegnerizzare le ricombinasi e le sue sequenze bersaglio. Questa tecnica si usa perché riduce la manipolazione genica il più possibile (non hai bisogno di grandissimi regioni di omologia per avere una frequenza elevata di ricombinazione nella cellule staminali). Si possono ingegnerizzare i costrutti in modo che con sequenze specifiche riconosciute da un enzima che quando le vede è capace di tagliare e far avvenire una ricombinazione.

Normalmente si usa la ricombinasi III che appartiene al *batteriofago P1*, oppure la FLP (flippasi derivato dal plasmide del lievito).

Fissiamoci su III e le sequenze che sa riconoscere.

Queste ricombinasi funzionano molto bene in sistemi eucariotici eterologhi. La ricombinasi III riconosce bene le sequenze LOXP. Lunghe 34 bp. È costituita da unità ripetute di 13 basi. In laboratorio possiamo inserire queste sequenze a monte e a valle del nostro gene, per es della neomicina. Nel momento in cui vogliamo inattivare il gene mediante inserzione della cassetta neo, ma poi vogliamo che la cassetta neo venga excisa dal nostro costrutto, perché vogliamo inattivare il locus ma non vogliamo portarci dietro neo con il marcatore. Se il marcatore neo lo fiancheggiamo con i siti LOXP e se nel nostro sistema cellulare qualcuno boh, questa è capace di boh excidere completamente il frammento e quindi otteniamo l'excisione della cassetta neo.

CRE riconosce sequenze ripetute e invertite LOXP che sono 21.7 del libro c'è una sequenza di 8 paia di basi e attorno con orientamento invertito ci sono le unità ripetute di 13 nucleotidi. CRE sta per Causa di REcombinazione e media la ricombinazione tra i siti LOXP. Quello che succede è che se hanno opportuno orientamento la sequenza compresa tra i siti LOXP viene excisa.

Ripetiamo meglio: quello che si è cercato di fare è limitare il problema di ottimizzare la ricombinazione omologa senza ogni volta portarsi grandi regioni di omologia per favorire l'evento. Un modo per bypassare il problema è usare delle ricombinasi e le sequenze riconosciute da queste ricombinasi. Le ricombinasi sono proteine molto semplici e hanno la capacità di riconoscere sequenze specifiche bersaglio molto piccole di dna. Le ricombinasi sono capaci di excidere, provocare ricombinazione, il dna compreso tra due LOXP (perché questi si giustappongono). Mediante una ricombinazione sito-specifica usando sequenze bersaglio precise e ricombinasi (che non sono eucariotiche ma funzionano bene in sistema eucariotico).

Per far avvenire una ricombinazione di questo tipo sitospecifica ci servono le ricombinasi e la possibilità di inserire nel gene bersaglio le sequenze riconosciute dalla ricombinasi. Nel caso di CRE, le sequenze riconosciute si chiamano LOXP, che molto brevi e fatte di 34 basi a unità ripetute di 13.

La presenza dei siti loxP in tandem ci consente di ottenere le seguenti situazioni. CRE excide la regione tra i due siti oppure scambia la regione di dna in base a come sono messe le sequenze loxp. Stiamo parlando di un sistema che prevede l'utilizzo di una ricombinasi e di un sito loxp. Se vogliamo fare una genetica fine, supponiamo che abbiamo inserito nel gene che ci interessa il transgene: supponiamo di avere un topolino CRE, cioè che esprime la ricombinasi CRE (*o tre, non sono sicuro*) per cui è stato ingegnerizzato, e quindi ha nelle sue cellule una cassetta con il c-DNA di CRE sotto un promotore (quindi il locus esprime CRE). Accanto abbiamo un altro topo che ha il locus A, che è il nostro locus di interesse, e lo abbiamo ingegnerizzato in modo che il locus A del nostro gene di interesse sia loxato, cioè abbia le estremità con i siti LOXP. Quando il topolino viene incrociato abbiamo la ricombinasi CRE, che può essere tessuto specifica perché ho ingegnerizzato i topi a esprimere CRE soltanto in un tessuto. Dove CRE è espressa avverrà l'excisione, quindi il nostro locus A verrà completamente abolito. Rimarrà solamente il sito LOXP.

Con questo sistema, solo quando CRE è presente il tessuto viene knockato. Questa tecnologia permette di studiare la funzione dei geni andando a guardare l'organismo in distretti diversi. Inattivi il gene quando vuoi tu e solamente in distretti precisi. Si può modulare l'inattivazione di un gene utilizzando la ricombinasi III, che si ottiene perché avendo i siti LOXP il gene può essere exciso, nel momento e nel tessuto specifico. In questo modo si realizza una analisi fine, ad esempio perché invece di fare una inattivazione precoce puoi andare a inattivare il gene successivamente e soltanto in un distretto particolare. In questo modo capisci se ha un effetto in un tessuto e in un altro.

Se il gene ha funzioni diverse, nel cervello una e nel fegato un'altra, a me interessa che l'animale viva per osservare quello che succede nel fegato e nel cervello. Con la possibilità di decidere dove spegnere il gene puoi districare i vari compartimenti e guardare meglio le varie situazioni. Tu li puoi anche usare in vivo, ma anche vitro.

Avendo a che fare con famiglie di geni se tu fai il knockout di un gene che però ha una importanza fondamentale non riesci ad avere l'animale; allo stesso modo, se supponiamo che abbiamo fatto knockout e l'animale sta bene non possiamo concludere nulla, magari stai guardando in una finestra temporale troppo stretta e magari il gene di cui hai fatto knockout entra in gioco più tardi.

Quindi a questi studi di knockout vanno sempre associati degli studi di espressione che ci dicono quanto è espresso un gene, perché quello mi sa dire quanto e quando nello sviluppo è espresso il gene.

Se vogliamo studiare quanto è espresso un gene devi prendere l'm-RNA. Andiamo a guardare il pool di gene espressi. Ciascun gene può essere rappresentato con un'etichetta. Una breve sequenza è unica per un gene, quindi quello che si fa è purificare tutto l'm-RNA da una cellula e mediante oligonucleotidi di T (primer che vanno ad allinearsi soltanto nella coda di poly-A) e un enzima iniziano a copiare. Producono quindi c-DNA perché stanno facendo trascrittasi inversa, la quale non deve copiare tutto il dna ma basta un piccolo pezzetto e costruiscono così delle etichette per ciascun m-RNA. Per correlare l'analogo dell'm-RNA al gene lo sequenzi, perché è complementare al gene.

Quindi con il sistema *anti gene expression* da un pool di m-RNA a sintetizzare piccoli pezzi di c-DNA di ciascun gene, li legano tutti assieme come fosse una lunga stringa e li mettono nel sequenziatore. La sequenza breve è l'etichetta, rappresenta un singolo gene: es ala-glu-ecc è la stringa amminoacidica specifica di una proteina, quindi ogni piccola sequenza rappresenta un gene. quello che poi fanno è quantificare il trascrittoma. In questo modo hanno una idea dei geni espressi nella cellula. In questo modo si può anche confrontare la cellula cancro con una cellula normale, andando a guardare proprio come cambia il pattern di espressione.

Questa tecnica solo guardando su un totale di proteine espresse ci fa vedere differenze sul profilo di espressione, che è la chiave di lettura per capire la funzione di un gene.

Riassumendo: Tu riesci a prendere dalla cellula solo gli m-RNA poly-A (solo quelli tradotti quindi) e da questi metti nella miscela oligo-d-T (i primer) e l'enzima della trascrittasi inversa, così ti produci piccoli c-DNA a partire da questi m-RNA. Se ad esempio la globina è espressa 500 volte più della nucleoporina avrai 500 molecole di m-RNA in più di globina che di nucleoporina, quindi 500 etichette di globina più presenti nel concatenamero rispetto alla nucleoporina. Quando li vai a sequenziare riesci a quantificare e ad avere un'idea del profilo di espressione dei geni. In questo modo puoi confrontare cellule, e capire come alcune mutazioni possono cambiare il profilo di espressione e quindi a farti un'idea di quella che può essere la funzione della proteina mutata.

Northern e Southern Blot.

Un'altra tecnica per andare a capire come è espresso un gene, data la sequenza e quindi la possibilità di costruire primer posso usare il Northern (rna) o il Southern (dna). Un altro modo è usare tecniche in cui tu marchi con oligo radioattivo la sequenza che vuoi andarti a guardare, fai correre l'rna e poi vai sopra con la sequenza radioattiva e impressioni una lastra.

Nel caso di esperimenti sul topo puoi prendere tessuti a tempi diversi, fai estrazione di rna totale, lo corri su un gel e ci fai passare sopra un probe radioattivo per vedere se c'è un tempo preciso in cui questo gene è acceso, se c'è una diversa cinetica di attivazione.

• Terapia genica basata sull'inibizione mirata.

Abbiamo detto che le malattie autosomiche recessive sono quelle che hanno più possibilità di successo in terapia genica, perché basta poco per avere un miglioramento.

Se abbiamo malattie ereditarie dominanti dovremo arrivare a inibire il gene mutato, ma per farlo bisogna veicolare degli oligo nel dna e non è facile.

Vedremo invece come si blocca l'espressione di un gene senza mutarlo. Si può bloccare a livello:

- di dna, bloccando la sua trascrizione
- di rna, e lì ci sono vari step (maturazione, trasporto, attacco ai ribosomi)
- della proteina, bloccando la maturazione o usando anticorpi per riconoscere e sequestrare la proteina. (degli anticorpi che riconoscono la proteina mutata (nel senso che non è wild type) e la sequestrano (questi si chiamano "*intracorpi*"). Questo degli intracorpi può essere un modo per silenziare l'espressione di una proteina in una cellula, ingegnerizzando la cellula in modo che la cellula produca l'anticorpo contro la proteina di interesse e la sequestri così essa non può funzionare.

Strategie ablative di terapia genica.

Questo di "spegnere" l'espressione di un gene mutato può essere tentata come una *strategia ablativa di terapia genica*. Questa si può eseguire a livello di dna, di rna e di proteina, quindi si può fare:

- mediate da acidi nucleici
- mediate da proteine

Mediate da acidi nucleici: si somministrano oligonucleotidi antisenso, che sono complementari alle sequenze di senso e quindi ci si legano.

Questa è la tecnica del silenziamento, che si fa di routine in laboratorio; si comprano oligonucleotidi antisenso che bloccano l'espressione della proteina di interesse. Dobbiamo a questo punto inserire gli oligo nella cellula: usiamo i liposomi o l'elettroporazione, sistema molto efficace con cui si possono veicolare gli acidi nucleici nelle cellule. Possiamo inoltre usare dei plasmidi che hanno l'oligonucleotide antisenso anziché una sequenza codificante per una proteina che viene prodotto come m-RNA.

Le tecniche sono egualmente valide: possiamo dare oligonucleotidi oppure vettori in grado di veicolare sequenze che venendo trascritte diventano antisenso con il gene di interesse.

Quando si forma il duplex questo viene tagliato e non viene prodotta la proteina. Il silenziamento quindi va utilizzato e misurato, si fa prima o dopo con oligo antisenso di controllo che non dovrebbero impattare sull'espressione della proteina di interesse. Una volta che abbiamo ottenuto tutto questo sistema cellulare pronto potete poi validare i risultati scientifici in presenza di oligo antisenso.

Come sono fatti gli oligo antisenso.

In genere sono 12-40 bp, sono specifiche per un tratto del gene bersaglio, si appaiano all'm-RNA complementare, formano degli ibridi duplex che bloccano la traduzione o per ingombro sterico o perché inibiscono splicing ecc e causano la degradazione dell'm-RNA.

È importante quando scegliamo gli oligo fare attenzione alla lunghezza, perché potrebbero formare strutture secondarie.

Mediate da proteine: Oppure si possono usare anticorpi intracellulari, o addirittura si possono manipolare anche usare i ribozimi, molecole autocatalitiche che tagliano l'rna, quindi si può sopprimere l'espressione di un gene utilizzando ribozimi che vanno a riconoscere e a tagliare. Potrei ingegnerizzare la cellula con anticorpi che vanno a legare la proteina e la bloccano, e quindi potrebbe essere un modo per curare una malattia ereditaria dominante.

Trasduzione di transgeni che codificano dominanti negativi.

Trasduzione vuol dire veicolare. Supponiamo di avere una proteina che quando presenta una mutazione specifica in un sito attivo diventa costitutivamente attiva, deregolata. Una proteina come un enzima vive in un equilibrio con la funzione catalitica attiva solo in determinate circostanze. Supponiamo di avere una malattia che è la mutazione puntiforme di un enzima che lo rende costitutivamente attivo, e quindi sia questa la causa della malattia. Per spegnere l'azione di questa proteina impazzita potremmo veicolare nella cellula dei transgeni, dei c-dna che possano curare questo effetto.

La proteina X è mutata, e questo la rende troppo attiva, con conseguente problema per la cellula. Per bloccarla a livello di proteina possiamo utilizzare degli anticorpi specifici che la bloccano; ma esiste un'altra strada. Una proteina la possiamo immaginare con tante tasche che interagiscono con le proteine che si orientano, fanno il passaggio di gruppi funzionali e poi se ne vanno. Questo succede normalmente con tutte le modificazioni, che sono regolate: devono esserci e poi spegnersi. Ma se abbiamo una mutazione per cui l'enzima si attacca sempre alle proteine e non può essere più regolato, potrei fornire io un inibitore; ma posso produrre una proteina che possa bloccare la funzione di un enzima deregolato.

Creo un dominante negativo: una stessa molecola cataliticamente incompetente, lega le proteine e le toglie all'altra proteina impazzita: compete nel legame perché ha gli stessi domini, è identica ma non conclude l'attività catalitica (es è una fosforilasi che non attacca il fosfato). Quindi sottrarre alla proteina/enzima impazzita le proteine che modificherebbe, e questo dà un piccolo giovamento alla cellula aiutandola a non avere più tutte le proteine modificate.

AlexR

Riassumendo: Introduci nella cellula per overespressione la stessa proteina che compete con l'altra però è incompetente dal punto di vista enzimatico, per cui sottrae substrato alla proteina impazzita. È una sempre questione di equilibrio.

Essendo una mutazione dominante vuol dire che un allele wild type ce lo dovrebbe avere, che non ce la fa a superare questo problema.

Mancano gli ultimi due minuti ma mi sembrano importanti.

Riparo della mutazione puntiforme.

Posso introdurre in una cellula un chimeric oligomet e un enzima

Genetica molecolare 14.05.07

Se vogliamo studiare la struttura, l'espressione e la funzione di geni umani possiamo usare colture ed estratti cellulari. Possiamo andare a guardare i vari trascritti o caratterizzare un cDNA speciale. Se vogliamo clonare il cDNA di una proteina di interesse possiamo prepararci delle librerie di cDNA e da qui prendere il gene di interesse. Oppure possiamo partire dalla proteina facendo la sequenza primaria, anche se questa è una via più tortuosa. Possiamo anche fare un mappaggio fisico per vedere dove è posizionato il nostro gene di interesse. Possiamo sfruttare l'omologia di geni di una certa famiglia genica per trovarne uno in particolare. Una speciale retrotrascrittasi viene usata per la retrotrascrizione. Se vogliamo arricchire un'estremità 5' partiamo con una PCR disegnando degli oligo antisense, poi usiamo un enzima speciale che è capace di attaccare una cDNA di adenine. Si usa un oligo dT per preparare un frammento. Queste PCR sono le **race PCR**.

Come possiamo avere sequenze di cDNA di lunghezza completa? Normalmente usiamo primer di sequenze brevi di T che si legano alla coda di poli A. Se l'RNA è grande non è detto che riusciremo ad avere un cDNA completo. In questo caso si usano sia oligo dT che esanucleotidi casuali. Casualmente un oligo si va ad appaiare ad un mRNA complementare. Queste sequenze sono degenerate e non fisse. La trascrittasi inversa può quindi estendere la sequenza. Le sequenze sono di cDNA variabile.

Ottenuta la sequenza di cDNA, avremo la sequenza AA della proteina tradotta. Dobbiamo però trovare l'ORF corretto. A volte non vengono trovati degli ORF. Questo può essere dovuto al fatto che il prodotto proteico sia corto o che l'RNA svolga altri ruoli. Un cDNA avrà un sito di arresto della traduzione con la presenza di codoni di terminazione. Ci sarà anche un sito di adenilazione. Poi ci sarà anche il sito di inizio AUG. Nel caso in cui abbiamo una sequenza AA parziale, possiamo usarla per andare a fare un confronto con un database proteico. Possiamo vedere se appartiene ad una famiglia proteica e vedere i domini della proteina in questione.

Se vogliamo vedere la genomica di un organismo, possiamo andare a guardare sia tutto il genoma intero, sia la struttura, sia il proteoma che sarebbe il profilo di espressione dei geni. Come le varie proteine di un organismo hanno a che fare l'una con l'altra. Con *genomica funzionale intendiamo proteomica*, ovvero come vengono espressi i vari geni di una cellula. Quando parliamo di genomica funzionale andiamo a guardare il proteoma, ovvero andiamo a guardare l'insieme dei trascritti, ovvero le ORF che sono le sequenze che possono andare incontro a traduzione. Se vogliamo studiare un gene possiamo eseguire delle tecniche di inattivazione genica. O il knock out è letale, e quindi il gene è essenziale, oppure il gene non ha alcun effetto, quindi il *gene sarà detto orfano* senza che si capisca la sua funzione. Possiamo studiare tutte le interazioni geniche con altri partner molecolari. Spesso il KO non dà la risposta desiderata perché la natura è ridondante.

Possiamo andare ad arricchire la nostra conoscenza di un gene vedendo con quali proteine il nostro "gene" tradotto va ad interagire. I domini proteici sono regioni capaci di folding autonomo. Ciascun dominio è un'unità a se stante. Se lo prendiamo e lo stacciamo mettendolo in tubo di plastica insieme a tutte le proteine di una cellula, esso è capace di riconoscere le proteine con le quali normalmente interagisce. Questo è molto importante. Una proteina lineare di 600AA ha vari domini, unità a se stanti. Una volta espresso questo dominio possiamo produrre anticorpi di questa proteina tramite conigli. Il coniglio fa un siero con anticorpi. Grazie all'anticorpo possiamo andare a guardare dove sta questa proteina nelle cellule. I domini proteici sono producibili in larga scala, possono essere messi su chips, i così detti *domain chips*.

Se vogliamo studiare un'interazione proteina-proteina dobbiamo ragionare in termini di domini. Spesso i domini sono stati caratterizzati per legare residui particolari. Gli Sh3 sono domini piccoli di 50-60 AA che riconoscono regioni ricche in prolina che sono presenti sulle proteine. Una proteina che possiede un Sh3 riconosce delle proteine con regioni ricche in prolina. Le regioni ricche in prolina sono le superfici complementari sulle quali si attacca questo dominio. Questi domini legano le loro regioni specifiche in maniera reversibile con affinità bassa. Le proteine si possono legare e poi distaccarsi. Tutte e due per incontrarsi devono stare nello stesso comparto

cellulare. L'interazione è reversibile. L'anticorpo è la molecola principe per studiare una proteina. L'affinità è elevatissima tra un anticorpo ed una proteina. L'anticorpo riconosce l'antigene e poi attira le cellule del sistema immunitario specifiche. L'affinità tra domini e sequenze bersaglio normalmente è bassa. Una volta trovata l'interazione possiamo costruire dei mutanti per studiarle.

Doppio ibrido: avremo una parte che lega il DNA ed una parte che lega la RNA pol. Si genera una ricostruzione artificiale in modo da avere un fattore di trascrizione diviso con DNA binding domain e Activation domain. Nel doppio ibrido vogliamo capire quali sono i domini che possono interagire con i domini della nostra proteina. Si prendono dei fattori di trascrizione di lievito che sono fatti da una parte che lega DNA ed una parte che si lega alla RNA pol attivando la trascrizione. Noi possiamo dividere queste due proteine ed attaccare alle loro estremità un cDNA, uno di X ed un altro di Y. Alla fine di Y avremo il dominio dell'activation domain. Se X ed Y interagiscono, il fattore di trascrizione si "rifirma" con le sue due parti (DNA binding domain e Activation domain) in modo tale che avvenga la trascrizione del gene indicatore, ad esempio lac Z. Soltanto nei cloni in cui si ha un'interazione, avremo la trascrizione del gene attivatore e quindi, ad esempio, la colonia sarà blu. Presa la colonia blu andiamo a sequenziare il nostro cDNA. Il doppio ibrido consente di trovare delle coppie di domini interagenti sfruttando il fatto che questi, quando interagiscono, permettono la trascrizione di un gene indicatore.

Strumento della genomica funzionale può essere la capacità di trovare mappe di interazione tra domini e proteine. Con questa tecnica, una proteina che normalmente è costituita da dei domini adiacenti, viene divisa in due geni che sono a loro volta fusi con frammenti di cDNA. Si ottiene la funzione solo quando le proteine fuse ai vari domini sono capaci di interagire fra di loro. Il metodo si basa sul fatto che i fattori di trascrizione si possono dividere in due parti separate. Nessuno dei due domini espressi da soli è in grado di attivare l'espressione di un gene reporter. Si ingegnerizzano dei vettori di espressione del lievito in modo tale che il DNA binding domain è fuso con la proteina esca (dominio da studiare). In ciascun vettore avremo un dominio del fattore di trascrizione DNA binding fuso con la nostra esca, nell'altro caso avremo il dominio di attivazione fuso con una *proteina bersaglio* (avremo tante proteine bersaglio e quindi tanti vettori in cui l'unica cosa che cambia è il cDNA che si va a legare all'activation domain). Solo se c'è un'interazione fra le due proteine avremo l'attivazione della trascrizione del gene reporter lac Z. Se l'interazione è modulata da modificazioni post traduzionali, noi non potremo vedere queste interazioni. Questa interazione vale solo nel caso in cui possiamo pescare domini interagenti senza che questi siano modulati da modificazioni post-traduzionali che nel lievito non esistono in maniera completa. Le interazioni che si studiano devono essere quindi anche validate in vivo.

In questi anni è stato possibile sequenziare gran parte del DNA degli organismi. E' però molto importante capire i profili di espressione dei geni. Per analizzare un gene dobbiamo capire l'espressione e la regolazione dell'espressione, e poi capire la sua funzione. In modo diretto possiamo eliminare il gene o inibirlo, e vedere i cambiamenti fenotipici. Se abbiamo fatto il KO del nostro gene senza vedere nessun effetto, possiamo supporre che ci siano altri geni della stessa famiglia genica. In una delle tecnologie per studiare l'espressione di un gene si mettono su chips dei piccoli nucleotidi in cui ciascuno rappresenta un clone di DNA. Da una parte abbiamo le etichette dei nostri geni (nucleotidi). Dalla cellule bersaglio produciamo cDNA marcato che poi va ad ibridarsi con i nostri nucleotidi. A questo punto possiamo andare a guardare l'intensità del segnale per vedere la quantità espressa di DNA.

Un'altra tecnica è quella del **sage**. Quando facciamo il microarray al cDNA, prendiamo il cDNA, lo marchiamo e lo facciamo ibridare. Non possiamo però modulare il tempo e il periodo di sviluppo. Altre tecniche permettono ciò. Sage sta per *serial analysis gene expression* che è il metodo per analizzare la quantità assoluta di trascritto espresso. Questo metodo si basa su una serie di etichette uniche per il nostro gene.

Per studiare l'espressione di un singolo gene si fa il *northern blot*. Il sage è invece un'analisi seriale di tutti i nostri geni. Si parte da una popolazione di RNA producendo delle sintesi parziali di cDNA. I cDNA di ciascun gene sono rappresentati da semplici cassettoni speciali (etichette). Questa tecnica

da un valore assoluto, un numero preciso, perché abbiamo contato i vari mRNA. Possiamo quindi analizzare il profilo di espressione in diverse condizioni. Per purificare l'RNA delle cellule si usa la complementarità della coda di adenina. Una volta separato l'mRNA, si mettono insieme alla trascrittasi inversa così da produrre cDNA. A questo punto avremo enzimi speciali capaci di tagliare delle sequenze brevi che rimuovono delle estremità. Altri enzimi riconoscono delle sequenze e tagliano a 20 bp da queste sequenze. In questo modo isoliamo queste etichette per tagliarle. Prima prepariamo le estremità separandole. Poi tagliamo a 15 bp a valle con altri enzimi. Poi cuciamo insieme tutto per poi sequenziali... Avremo quindi un lungo concatamero che può essere sequenziato.

Nel caso in cui vogliamo vedere l'espressione differenziale nel tempo, possiamo utilizzare la **tecnica di visualizzazione differenziale dell'mRNA**. Con questa tecnica andiamo a guardare delle variazioni di espressioni tra cellule. Vediamo se ci sono trascritti prodotti in un tempo speciale. Scegliamo di guardare a campione degli mRNA speciali agendo sull'oligo dT. Si fissa l'oligo dT in modo da amplificare una frazione del trascrittoma. Vogliamo guardare variazioni nel pattern di espressione di geni. Sintetizziamo degli oligonucleotidi e fissiamo una coppia, ad esempio GC, successiva all'ultima T. Se mettiamo trascrittasi inverse avremo il cDNA solo di quei mRNA con basi complementari al nostro oligo. In questo modo amplifichiamo solo una popolazione specifica di mRNA. In questa tecnica si usa un esamero scelto a caso senza conoscere nulla sulla sequenza dell'intero mRNA. Questa tecnica ci permette di vedere la variazione nell'espressione dell'mRNA a differenti tempi.