

Soluzioni del Capitolo 20

20.1 a. 7; b. 9; c. 3; d. 8; e. 1; f. 12; g. 10; h. 11; i. 2; j. 6; k. 4; l. 5.

20.2 a. Molto spesso le mutazioni di senso sono alleli a perdita di funzione. Si creano costrutti transgenici che esprimono cDNA di tipo selvatico corrispondente a ciascuna delle tre diverse forme di giunzione. Questi costrutti dovrebbero includere sequenze che consentano loro di essere espressi correttamente nei follicoli piliferi [vedi la parte (b) di seguito]. Utilizzando l'iniezione del pronucleo, si trasformano con ogni costrutto i topi mutanti bianchi e si verifica se uno di loro recupera il fenotipo wild-type (grigio). Potrebbe anche essere possibile che sia necessaria una combinazione di due o tre le forme di giunzione per recuperare il fenotipo wild-type.

b. Poiché i costrutti sono basati su cDNA, richiederanno sequenze non presenti nei cDNA, che consentono agli mRNA di essere espressi ed elaborati correttamente. Pertanto, come si osserva nella figura che segue, i costrutti dovranno includere un enhancer specifico del follicolo pilifero e un promotore che possa interagire con l'enhancer. (Idealmente, si dovrebbero utilizzare l'enhancer e il promotore del gene in questione.) Inoltre, è necessario un sito di addizione della coda poli-A per garantire la stabilità e la traduzione degli mRNA.



c. L'iniezione pronucleare provoca l'inserimento di più copie del transgene, in posizioni casuali.

Né il numero di copie né il sito di inserimento possono essere controllati ed entrambi influenzano potenzialmente la quantità di mRNA e quindi la quantità di proteine prodotte. Troppo poche o troppe proteine potrebbero influenzare la capacità del costrutto di recuperare la mutazione.

20.3 a. Incrociare il ceppo transgenico con il tipo selvatico e incrociare fra loro gli animali F_1 , eterozigoti per l'inserimento del transgene. Se la deformità dell'arto non è causata dall'inserimento del transgene, alcuni animali F_2 , omozigoti per l'inserimento del transgene, non mostreranno il fenotipo mutante. Se la deformità dell'arto è dovuta all'inserimento del transgene, l'omozigosi per il transgene dovrebbe essere correlata alla deformità dell'arto.

b. Il transgene si è inserito nel gene e ha causato una mutazione con perdita di funzione. La PCR inversa (vedi Figura 15.29) consente di identificare il gene in cui si è inserito il transgene.

c. Effettuare un test di complementazione. Incrociare i topi mutanti transgenici omozigoti con topi omozigoti per la mutazione *ld*. La progenie è transeterozigote per l'inserzione del transgene e *ld*. Se l'inserzione del transgene è una mutazione a perdita di funzione nello stesso gene che è alterato in *ld* (cioè, se le mutazioni sono alleliche), le mutazioni non completeranno e la progenie mostrerà la deformità dell'arto. Se, invece, si tratta di mutazioni in geni diversi, le mutazioni completeranno (nella progenie non comparirà il fenotipo mutante).

20.4 a. Si crea un costrutto del gene di fusione in cui il promotore *HoxA1* e l'enhancer guidano l'espressione di un cDNA *HoxD4*.

b. Se l'assenza della proteina *HoxD4* è ciò che distingue le ossa occipitali (E) dalle vertebre cervicali (C1 e C2), negli animali trasformati con il gene di fusione *HoxA1::HoxD4*, le ossa occipitali do-

vrebbero essere trasformate in vertebre cervicali. (Cioè, i topi con questo costrutto genico dovrebbero esprimere la proteina HoxD4 nelle cellule che normalmente diventerebbero ossa occipitali, che si svilupperanno invece in vertebre cervicali.)

- 20.5 a.** Si trasformano i moscerini *faf*⁻ con un costrutto in cui le sequenze regolatrici del gene *faf* di *Drosophila* inducono l'espressione di un cDNA *Usp9*. Se il fenotipo occhio ruvido viene recuperato, la proteina umana *Usp9* può effettivamente sostituire funzionalmente la proteina *Faf* del moscerino.
- b.** Se i topi con mutazioni a perdita di funzione *Fam*⁻ hanno occhi mutanti, allora *Fam* deve essere necessario per lo sviluppo dell'occhio nel topo. Più avanti si imparerà come usare il DNA di *Fam* clonato per generare topi con mutazioni a perdita di funzione di *Fam*, che saranno knockout o knockout condizionali.

- 20.6 a.** Le linee di moscerini in cui il costrutto *enhancer trap* si è integrato accanto a un enhancer esprimeranno la GFP generando tessuti fluorescenti verdi. A seconda della natura dell'enhancer in una particolare linea che esprime la GFP, l'espressione della GFP potrebbe essere ubiquitaria in ogni cellula, o limitata a specifici tipi cellulari in particolari fasi di sviluppo.

b. Si opera uno screening per linee che esprimono GFP solo nell'ala. Si utilizza la PCR inversa (vedi Figura 15.29) per amplificare il DNA genomico adiacente al sito di inserimento del transgene *enhancer trap*. Si sequenzia il DNA genomico amplificato e si confronta con la sequenza di riferimento del genoma della *Drosophila* per identificare il gene localizzato nel sito di inserimento. Questo gene dovrebbe essere espresso nell'ala del moscerino perché si trova vicino a un enhancer specifico dell'ala.

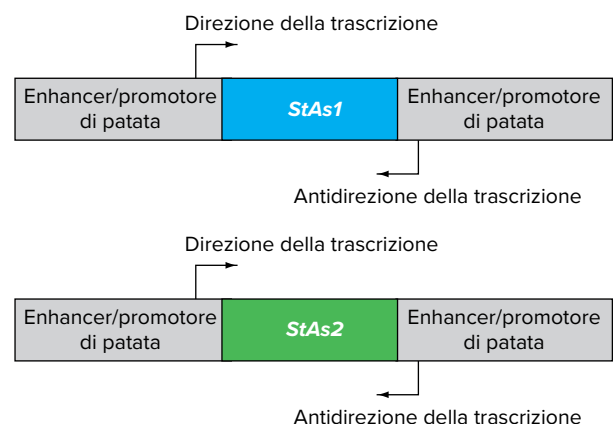
c. Le inserzioni di *enhancer trap* a volte possono situarsi all'interno del gene normalmente regolato dal potenziatore "intrappolato"; in questo modo bloccano la funzione genica e causano una mutazione a perdita di funzione. Gli omozigoti per l'inserzione del transgene potrebbero, quindi, avere un fenotipo mutante (vedi Problema 3.)

- 20.7 a.** Innanzitutto, è necessario identificare il gene che codifica per la proteina antigelo del pesce. Un modo per raggiungere questo obiettivo è purificare la proteina antigelo, determinare la sequenza di amminoacidi di questa proteina con mezzi chimici e quindi esaminare il genoma del pesce, cercando il gene che potrebbe codificarla. Successivamente, si sintetizza il cDNA per il gene. Diverse tecniche permettono di farlo; un metodo consiste

nell'utilizzare primer specifici per il gene e la trascrittasi inversa per amplificare il cDNA specifico da un pool di mRNA totali di pesce. Successivamente, si inserisce il cDNA in un vettore T-DNA che contiene un promotore e un enhancer vegetale ad alta efficienza e si trasfetta il T-DNA ricombinante in *Agrobacterium* (vedi Figura 20.4). Infine, si irrorra l'*Agrobacterium* trasformato sulle cellule vegetali e quindi si fanno crescere embrioni da singole cellule o semi selezionati per l'inserimento del T-DNA. (Se il vettore T-DNA includesse un gene per la resistenza agli erbicidi, si può aggiungere l'erbicida alle colture delle cellule vegetali.) Gli embrioni che sopravvivono alla selezione esprimono il cDNA del gene antigelo del pesce sotto il controllo di un promotore vegetale.

b. I ricercatori infatti hanno trovato la proteina antigelo del pesce nelle piante con l'inserzione di T-DNA, ma la proteina non sembrava funzionare. È probabile che nelle cellule vegetali non siano state apportate le modificazioni post-traduzionali chiave alle proteine del pesce. Una seconda possibilità da considerare è che la proteina antigelo nel pesce non funzioni da sola, ma solo in combinazione con un cofattore (un'altra proteina o anche una piccola molecola) che si trova nei pesci ma non nelle piante.

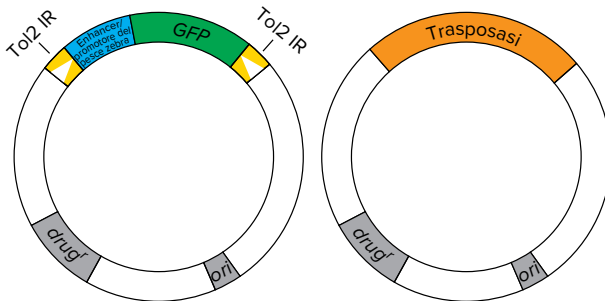
- 20.8 a.** Nella figura che segue, il DNA indicato con *StAs1* e *StAs2* potrebbe rappresentare sia cDNA sia singoli esoni e le frecce indicano il sito di inizio e la direzione della trascrizione. Gli RNA di senso e antiseno, prodotti da ciascun transgene, formano un RNA a doppio filamento che viene processato dal macchinario cellulare in siRNA che funzionano come i miRNA. Le basi dei siRNA si appaiano con gli mRNA *StAs1* e *StAs2*, causandone la degradazione.



- b.** Poiché l'asparagina è un amminoacido presente praticamente in ogni proteina della pianta,

è necessaria almeno una piccola quantità di proteine StAs1 e StAs2 affinché la pianta sia vitale. **c.** I transgeni usati come insetticidi erano simili a quelli della parte (a), tranne per il fatto che invece di *StAs1* o *StAs2*, venivano utilizzati i geni del coleottero del Colorado essenziali per la vitalità di questo insetto nocivo. I coleotteri che mangiano le piante di patate ingeriscono gli RNA interferenti, cosicché gli mRNA di uno o più dei loro geni essenziali vengono degradati, provocando la morte dell'insetto. Uno dei principali vantaggi di questo approccio rispetto all'irrorazione delle piante di patate con insetticidi chimici è che sarebbe molto difficile per i coleotteri accumulare mutazioni che li rendessero resistenti agli effetti letali degli RNA interferenti. Qualsiasi mutazione del genere dovrebbe consentire agli insetti di bypassare un pathway biochimico essenziale e questo sarebbe un evento estremamente raro.

- 20.9 a.** Nei diagrammi che seguono, IR = ripetizione invertita e *ori* = origine della replicazione che permette il mantenimento del plasmide in *E. coli*. Un punto chiave è che il plasmide a sinistra che contiene il transgene ha ripetizioni invertite *Tol2* e il plasmide a destra che fornisce la proteina trasposasi non le ha. (Si noti che le sequenze di DNA tra le ripetizioni invertite *Tol2* [contenenti l'enhancer/promotore + la regione codificante GFP] possono essere inserite nel genoma del pesce zebra.)



b. L'uso di un trasposone di pesce medaka (*Tol2*) nel pesce zebra significa che non esiste una fonte endogena di trasposasi *Tol2*. Pertanto, una volta inserito nel genoma del pesce zebra, il transgene sarà stabile.

- 20.10 a.** Le zanzare transgeniche assumono cibo che contiene tetraciclina. **b.** Poiché le zanzare maschio sono omozigoti per il transgene letale, tutta la progenie di ogni femmina selvatica che si accoppia con uno dei maschi transgenici morirà; tutta la prole conterrà il transgene. Presumibilmente, circa il 90% delle femmine selvatiche si è accoppiato con maschi transgenici.

20.11 a. Gli scienziati stanno tentando di eseguire la fecondazione in vitro (IVF) di uova delle femmine viventi di rinoceronte bianco settentrionale con lo sperma scongelato dei quattro diversi maschi. Gli embrioni in via di sviluppo verrebbero quindi impiantati in femmine surrogate di rinoceronte bianco meridionale che darebbero alla luce i rinoceronti bianchi settentrionali.

b. La fecondazione in vitro e la clonazione riproduttiva sono fondamentalmente diverse. Come descritto nella parte (a), nella fecondazione in vitro gli zigoti diploidi di rinoceronte hanno una copia del genoma da un uovo di rinoceronte e un'altra copia del genoma da uno spermatozoo di rinoceronte: la fecondazione ha luogo proprio come avverrebbe naturalmente, tranne che si verifica in una capsula di Petri piuttosto che nel corpo di una femmina di rinoceronte, quindi gli embrioni risultanti devono essere impiantati in una femmina che li partorisca. Nel caso della clonazione, invece, si sostituiscono i nuclei di oociti di rinoceronte con quelli di una cellula somatica; lo zigote risultante avrà lo stesso genoma diploide del rinoceronte che ha donato le cellule somatiche. Gli embrioni clonati avrebbero anche bisogno di essere impiantati in una femmina che li partorisca.

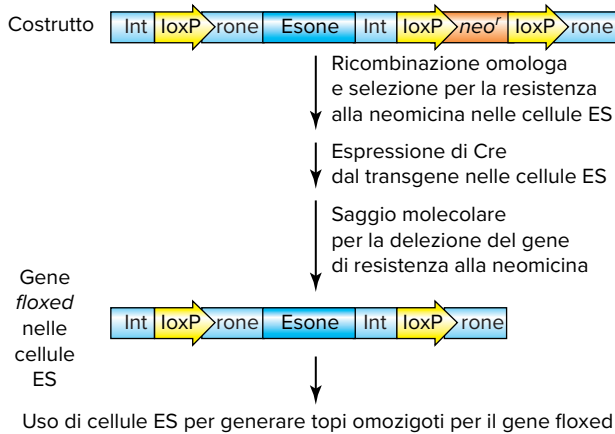
c. Una popolazione di rinoceronti bianchi settentrionali ottenuta attraverso la fecondazione in vitro, come descritto nella parte (a), potrebbe avere difficoltà a sopravvivere a lungo termine a causa della scarsa diversità genetica. Nell'insieme, i geni dell'intera popolazione provengono solo da sei animali diversi.

20.12 a. Alcuni geni nel genoma di topo potrebbero essere aploinsufficienti per una funzione essenziale nelle cellule ES indifferenziate. Le cellule eterozigoti per knockout di questi geni muoiono e non possono essere mantenute come linea cellulare.

b. Alcuni knockout eterozigoti vitali nelle cellule ES totipotenti potrebbero essere aploinsufficienti in tessuti chiave, più avanti nello sviluppo del topo. I topi knockout per questi geni sarebbero letali.

c. *Fam* è aploinsufficiente nelle cellule ES, quindi non è possibile utilizzare le cellule ES per generare topi *Fam*⁻/*Fam*⁺, e tantomeno omozigoti (*Fam*⁻/*Fam*⁻). Per verificare se *Fam* è necessario per lo sviluppo dell'occhio del topo, un ricercatore potrebbe creare un topo knockout condizionale utilizzando il sistema di ricombinazione *Cre/loxP*. Un costrutto che potrebbe essere intro-

dotta nel gene *Fam* del topo mediante ricombinazione omologa nelle cellule ES è mostrato nel diagramma che segue (vedi Figura 20.10a per ulteriori dettagli):

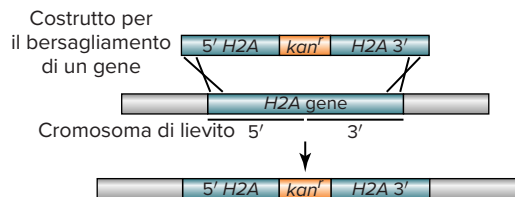


Il topo avrebbe anche un gene *Cre* (esprime la ricombinasi Cre), controllato da un enhancer specifico dell'occhio, come mostrato in Figura 20.10b. L'espressione di *Cre* specifica per l'occhio comporterebbe la delezione dell'esone *floxed* di *Fam* (l'esone azzurro scuro nel diagramma precedente) solo nelle cellule dell'occhio. Pertanto, in un topo eterozigote per un gene *Fam* wild-type e un gene *Fam floxed* ($Fam^+/Fam\ floxed$), gli occhi sarebbero Fam^+/Fam^- (esone deletato), mentre tutte le altre cellule dell'animale sono $Fam^+/Fam\ floxed$ (esone non deletato). Gli occhi di questi topi potrebbero mostrare un fenotipo mutante se Fam^+ è aploinsufficiente negli occhi. Topi che mancano completamente della funzione Fam^+ negli occhi potrebbero essere generati incrociando due eterozigoti $Fam^+/Fam\ floxed$ per ottenere un omozigote $Fam\ floxed/Fam\ floxed$. L'espressione di *Cre* specifica per l'occhio in un tale omozigote comporterebbe la delezione dell'esone *floxed* sia nelle copie del gene *Fam floxed* sia nelle cellule dell'occhio omozigoti $Fam-/Fam^-$.

d. Poiché nelle cellule ES *Fam* è aploinsufficiente per la vitalità, probabilmente è anche aploinsufficiente nella maggior parte delle cellule di topo. Pertanto, sembra probabile che siano necessarie due copie di Fam^+ per la vitalità delle cellule dell'occhio e quindi che gli occhi Fam^+/Fam^- sarebbero assenti perché le cellule oculari eterozigoti morirebbero. Un altro possibile risultato è che i livelli di Fam^+ siano meno essenziali nell'occhio rispetto alle cellule ES, quindi le cellule vivono ma gli occhi non si svi-

luppano correttamente e sono malformati. Ancora, un terzo risultato potrebbe essere che solo parti dell'occhio mancano o si formano in modo anormale, il che significa che la *Fam* è necessaria per il corretto sviluppo di parti specifiche dell'occhio. Infine, è possibile che l'occhio si trasformi in un altro organo, per esempio un'antenna. Si imparerà nel Capitolo 21 che le mutazioni che portano a trasformazioni di una parte del corpo in un'altra sono chiamate mutazioni omeotiche. Ciò significherebbe che *Fam* è un gene omeotico, un gene fondamentale nel prendere decisioni sul destino cellulare.

20.13 a. Il diagramma che segue mostra un costrutto per eliminare il gene *H2A* del lievito. I rettangoli azzurri possono essere sequenze del gene *H2A* a ciascuna estremità, come mostrato, o al 5' e 3' del gene *H2A* (non mostrato). La ricombinazione omologa tra il gene *H2A* del genoma di lievito e il costrutto lineare porta alla sostituzione dell'allele normale con l'allele contenente il gene *kan^r* al centro del registro di lettura aperto. Questo allele dovrebbe essere non funzionale (allele nullo).



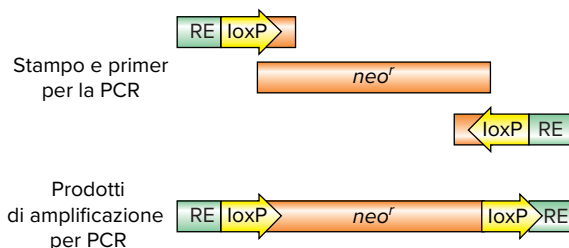
b. (I) Con la tecnologia del DNA ricombinante si genera un plasmide contenente il gene *kan^r* inserito nel registro di lettura aperto di *H2A* come nella parte (a). (II) Si creerà quindi un frammento lineare contenente il gene *H2A* interrotto, come indicato nel diagramma. È possibile creare questo frammento lineare ritagliandolo dal plasmide con enzimi di restrizione o mediante PCR, amplificando le sequenze desiderate del plasmide. (III) Si trasforma poi il lievito con il frammento lineare di DNA e si selezionano le colonie per la resistenza alla kanamicina. (IV) Si analizza la regione genomica *H2A* della singola colonia utilizzando la PCR e il sequenziamento del prodotto PCR per verificare l'inserimento tramite ricombinazione omologa.

c. Ci si aspetta che il gene *H2A* sia essenziale, perché il suo prodotto proteico è una parte essenziale della struttura dei nucleosomi (vedi Capitolo 12). Quindi si esegue il targeting del gene in un ceppo diploide con l'obiettivo di eliminare sol-

tanto un singolo allele. (Il targeting genetico è sufficientemente inefficiente, per cui è improbabile eliminare entrambe le copie del gene contemporaneamente in tutte le cellule.) In un tale diploide, resterà una copia del gene di tipo selvatico per impedire la morte della cellula di lievito bersagliata con successo.

d. Se una cellula aploide knockout per *H2A* fosse vitale, si dovrebbe concludere che, contrariamente alle aspettative, *H2A* non è un gene essenziale. La spiegazione più probabile sarebbe che il genoma del lievito contenga un altro gene che produce una proteina molto simile all'*H2A*, che può sostituire funzionalmente la proteina *H2A* nei nucleosomi.

- 20.14** L'idea è di sintetizzare una coppia di primer per PCR che amplificherà lo stampo del gene *neo^r* e allo stesso tempo aggiungerà siti *loxP* alle due estremità del gene. Ciò può essere ottenuto sintetizzando primer lunghi che contengono i 34 nucleotidi del sito *loxP* e anche ~20 nucleotidi del gene *neo^r*. Inoltre, l'estremità 5' di ciascun primer conterrà un sito di restrizione (RE) che non si trova altrove nel prodotto di amplificazione della PCR, ma si trova nell'introne in cui verrà ligato questo frammento di DNA. Come mostrato nel diagramma seguente, il prodotto PCR avrà lo stesso sito di restrizione su entrambe le estremità (RE) e siti *loxP* che fiancheggiano il gene *neo^r*.



Ora si può tagliare il prodotto della PCR con l'enzima di restrizione e ligarlo in un plasmide contenente l'introne, tagliato con lo stesso enzima di restrizione che lascia le stesse estremità coesive:



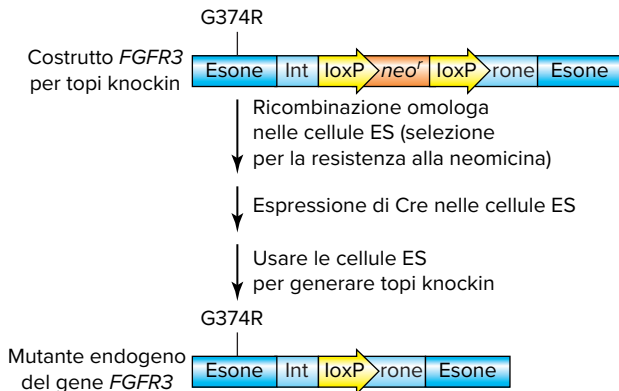
- 20.15 a.** Poiché la LCA è dovuta alla perdita di funzione del gene *RPE65*, è possibile utilizzare un topo knockout per generare un modello animale di LCA. Come per la maggior parte delle mutazioni a perdita di funzione, il fenotipo LCA è recessivo e compare solo negli omozigoti o nei transeterozigoti per due alleli diversi. Pertanto, il modello murino dovrebbe essere omozigote per l'allele

knockout. Se questi topi knockout sono ciechi, è probabile che *RPE65* funzioni in modo simile nei topi e nell'uomo.

b. La procedura da usare per creare un modello murino per la sindrome dell'X fragile dipenderebbe dall'aspetto della malattia che si vuole studiare. Poiché la malattia è causata dalla perdita della funzione del gene *FMR-1*, un knockout potrebbe riprodurre il fenotipo della malattia. Il gene *FMR-1* si trova anche nei topi sul cromosoma X murino. I topi maschi emizigoti per il knockout e i topi femmina omozigoti per il knockout mostrano infatti un fenotipo che ricorda molto i sintomi della sindrome dell'X fragile osservati nei bambini. D'altra parte, se si vuole studiare come nascono nuovi alleli mutanti da alleli pre-mutazionali o come l'espansione del numero di ripetizioni CGG inibisce l'espressione di *FMR-1*, si può generare un knockin con ripetizioni CGG aggiuntive.

c. La malattia di Huntington è causata da un allele a guadagno di funzione del gene *HD* che produce una proteina con una regione polyQ eccessivamente lunga. Per studiare il fenotipo della malattia, come funziona la proteina mutante o il processo mediante il quale un allele pre-mutazionale evolve in un allele della malattia, sarebbe necessario un knockin che aggiunge ripetizioni CAG alla regione codificante del gene del topo nella stessa posizione in cui queste ripetizioni si trovano nell'*HD* umano. Tali modelli murini knockin con 150 ripetizioni CAG sono stati effettivamente creati e mostrano neurodegenerazione a esordio tardivo. Le aziende farmaceutiche possono utilizzare questi topi per stabilire se i farmaci candidati alleviano i sintomi della malattia, sebbene i modelli siano difficili da usare perché l'esordio è piuttosto tardivo rispetto alla durata di vita del topo.

- 20.16 a.** L'obiettivo è quello di inserire nel gene *FGFR3* murino la mutazione missenso (G380R) che causa l'acondroplasia quando è presente nel gene umano. Si scopre che la mutazione corrispondente nel gene *FGFR3* del topo è G374R. (La mutazione è un singolo cambiamento di base: da GGG [Gly] ad AGG [Arg]. Le proteine murine e umane sono di dimensioni leggermente diverse, il che spiega la differenza nella numerazione.) Il costrutto knockin e la mutazione finale che produce nel gene *FGFR3* murino sono mostrati nel diagramma che segue, basato sulla Figura 21.11:



b. (I) Progettare uno sgRNA che tagli il genoma vicino al sito del codone da mutare e un clone di DNA contenente la mutazione. (Il clone di DNA deve avere i cosiddetti “bracci di omologia” di ~500 bp a sinistra e a destra della mutazione che corrispondono al DNA genomico di tipo selvatico che circonda il sito che si desidera mutare. Questi bracci di omologia permetteranno il verificarsi della ricombinazione omologa a sinistra e a destra della sequenza mirata, generando la formazione di un allele knockin come mostrato in basso a destra in Figura 20.13.)

(II) Iniettare lo sgRNA, il gene *Cas9* e i DNA lineari a doppio filamento contenenti la mutazione puntiforme in uova di topo appena fecondate.

(III) Successivamente, impiantare gli embrioni iniettati in madri surrogate.

(IV) Quando questi animali nascono, testare i cuccioli di topo per le cellule con la mutazione usando la PCR. Poiché i topi trattati che risultano positivi alla mutazione sono probabilmente un mosaico per cellule mutanti e normali, è necessario incrociare questi topi con topi normali. Se le cellule della linea germinale portano la mutazione, i gameti che producono genereranno eterozigoti non a mosaico nella generazione successiva.

20.17 a. Per esprimere GFP, il costrutto di “gene trapping” deve integrarsi all’interno dell’introne di un gene. Il costrutto è costituito essenzialmente da un esone che, inserendosi tra gli esoni di un gene del topo, può essere incorporato in un mRNA maturo. A causa del suo sito di addizione della coda poli-A, il costrutto diventerà l’esone finale del gene alterato.

b. Solo una frazione di topi con il transgene integrato in un introne esprimerà GFP. Questa affermazione è vera anche se assumiamo che il transgene venga sempre a trovarsi all’interno dell’mRNA prodotto dal trascritto primario del gene

alterato. Ci sono diversi motivi per cui solo una frazione dei topi esprimerà GFP.

In primo luogo, il costrutto può inserirsi in due diversi orientamenti rispetto al promotore. Solo l’orientamento in cui il registro di lettura aperto della GFP si ritrova nel trascritto può comportare l’espressione della proteina GFP.

In secondo luogo, la posizione del transgene rispetto ad altri esoni è importante. Se il transgene si inserisce tra gli esoni che contengono entrambi sequenze 5’ UTR, allora l’ATG del transgene segnerà l’inizio della traduzione della GFP che così verrà prodotta. Se il transgene si inserisce tra gli esoni codificanti, la traduzione inizierà a monte e l’ATG del cDNA della GFP deve essere in frame con l’ORF dell’esone precedente affinché la proteina GFP possa essere tradotta; la probabilità che ciò accada è 1 su 3 (se l’orientamento è corretto). Se il transgene si inserisce tra gli esoni che contengono entrambi sequenze 3’ UTR, non verrà prodotto alcun GFP.

Infine, se le ORF sono a registro, gli amminoacidi della GFP saranno attaccati al terminale C della proteina codificata dall’esone/i precedente/i. Come parte di una proteina ibrida, la GFP può o meno ripiegarsi correttamente e quindi la fluorescenza può essere rilevabile o meno.

c. Se il transgene si inserisce in modo tale che la proteina GFP sia prodotta senza altri amminoacidi attaccati a essa [vedi parte (b)], le informazioni ottenute con l’enhancer trapping e con il gene trapping sono sostanzialmente simili; GFP riporterebbe il pattern di espressione degli elementi regolativi del gene che lo ha intrappolato. Tuttavia, esiste una sottile differenza tra i pattern di espressione rivelati dai due metodi. Se il costrutto di trapping genico si inserisce tra esoni che contengono entrambi sequenze 5’ UTR, l’mRNA che esprime GFP avrà sequenze 5’ UTR dal gene del topo. Le sequenze 5’ UTR potrebbero influenzare la traduzione o la stabilità dell’mRNA per GFP, forse in modo tessuto-specifico. Un costrutto di enhancer trapping ha il proprio 5’ UTR, quindi tali effetti non possono essere osservati.

Se la GFP si inserisce in modo tale da costituire il C terminale di una proteina ibrida [vedi parte (b)], potrebbero essere potenzialmente rivelate informazioni sulla localizzazione subcellulare della proteina, specialmente se la porzione N-terminale della proteina ibrida contiene tutta o la maggior parte della normale proteina. Se la regio-

ne N-terminale della proteina è in grado di raggiungere la sua posizione normale nella cellula (per esempio, un particolare organello), la GFP attaccata mostrerà fluorescenza in quel punto. In una linea enhancer trapping, la GFP non sarebbe espressa come una proteina ibrida, quindi queste informazioni non sarebbero disponibili.

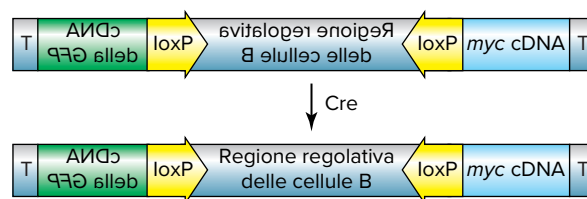
d. Sia gli inserimenti di enhancer trapping sia quelli di gene trapping possono generare mutazioni nei geni dei loro siti di inserimento. Tuttavia, è più probabile che una mutazione sia generata dal gene trapping che non dall'enhancer trapping. Un costrutto di enhancer trapping può esprimere GFP senza che il transgene si inserisca nel gene, mentre un costrutto di gene trapping deve inserirsi in un introne del gene affinché venga prodotta GFP. Il codone di arresto con cui termina l'ORF della GFP interromperà l'ulteriore traduzione dell'mRNA e quindi impedirà la traduzione di eventuali amminoacidi codificati dagli esoni a valle del sito di inserimento.

20.18 a. Se l'ipotesi è corretta, l'introduzione nel genoma del topo, mediante iniezione nel pronucleo, di un transgene in cui la regione regolativa dei linfociti B è fusa con il cDNA di *myc* dovrebbe aumentare la frequenza di induzione del tumore nel sistema immunitario del topo.

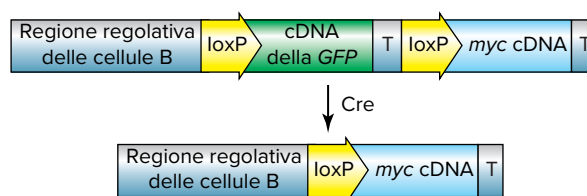
b. I topi transgenici descritti nella parte (a) esprimerebbero il cDNA di *myc* nei linfociti B per tutta la vita dell'organismo. In questa parte del problema, stiamo esplorando come attivare la sovra-espressione di *myc* a partire da un momento definito nello sviluppo dell'organismo.

Un modo per sovra-esprimere *Myc* nei linfociti B a partire da 1 settimana di età è generare un transgene, come quello mostrato nel diagramma che segue, che può essere introdotto nel genoma del topo mediante iniezione nel pronucleo. (Nel diagramma la regione regolativa delle cellule B contiene sia un promotore sia un enhancer specifico delle cellule B; T = sequenza di terminazione della trascrizione.) Se i topi contengono anche un costrutto *hs::Cre*, uno shock termico a una settimana di vita indurrà l'espressione di *Cre* che farà capovolgere l'enhancer delle cellule B, in modo da far esprimere *myc* anziché GFP. Il gene GFP in questo costrutto ha due scopi. In primo luogo, è un controllo positivo per garantire che i topi siano stati effettivamente trasformati dal costrutto. In secondo luogo, serve ad assicurarsi che lo shock termico abbia funzionato come desiderato se le cellule B smettono di

esprimere GFP e quindi perdono la loro fluorescenza.



Un altro costrutto che funzionerebbe ha siti *loxP* con lo stesso orientamento che fiancheggiano GFP; il cDNA GFP (che è stato inizialmente trascritto nei linfociti B) verrà eliminato quando *Cre* viene espresso:



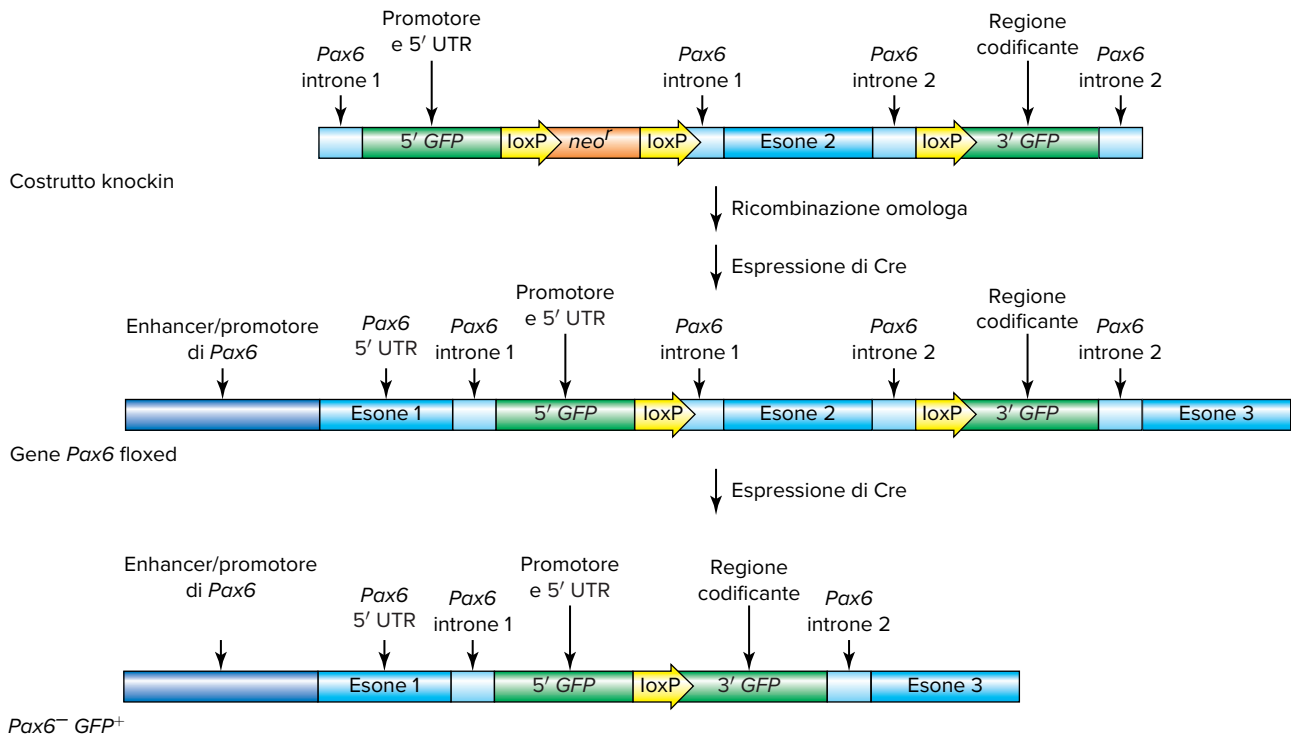
20.19 a. Per generare un topo che esprime normalmente *Pax6* ovunque tranne che negli occhi, si dovrebbero generare cellule ES con un gene *Pax6* floxed (tra due siti *loxP*), utilizzando la procedura mostrata in Figura 20.10, e un transgene con un *enhancer specifico dell'occhio del topo::Cre*. In alternativa, si potrebbe introdurre nei topi un transgene che esprime la sequenza complementare inversa del cDNA di *Pax6* guidato da un enhancer specifico dell'occhio. Quest'ultimo costrutto inibirà l'espressione endogena di *Pax6* nell'occhio per mezzo di RNAi. Cioè, le cellule dell'occhio avranno RNA *Pax6* a doppio filamento che può essere maturato in siRNA *Pax6*, che a sua volta porterà alla degradazione degli mRNA di *Pax6*.

b. Affinché le cellule dell'occhio *Pax6* esprimano la GFP, è necessario utilizzare uno speciale ceppo knockout condizionale che include un gene GFP diviso in due, per rimuovere specificamente dalle cellule dell'occhio il gene *Pax6*. Il costrutto knock-in mostrato nella parte superiore della figura che segue genererebbe il gene *Pax6* floxed presentato nella parte inferiore della figura. Si noti che il gene *GFP* all'interno del gene *Pax6* floxed è diviso in due parti: una regione 5', contenente un promotore basale del topo e il 5' UTR, e una regione 3' con i codoni per gli amminoacidi della GFP. Ogni pezzo del gene GFP è all'interno di un introne diverso; l'esone 2 interrompe i due introni. Prima dello shock termico, un topo con il gene

Pax6 floxed esprimerebbe una proteina *Pax6* normale, perché tutte le alterazioni del gene sono all'interno degli introni, quindi le sequenze GFP e loxP verrebbero eliminate dal trascritto primario. Un tale topo non sarebbe in grado di produrre GFP prima dello shock termico perché le sequenze GFP sono interrotte.

I topi omozigoti per l'allele floxed hanno anche un transgene che esprime Cre solo nell'occhio.

La ricombinazione catalizzata da Cre eliminerà solo nell'occhio l'esone 2 dal gene floxed. *Pax6* verrà inattivato nelle cellule dell'occhio. Contemporaneamente, il promotore e l'ORF di GFP vengono riuniti (l'enhancer *Pax6* endogeno lavorerà con il promotore del topo che guida l'espressione di GFP), in modo che le cellule senza funzione *Pax6* mostrino fluorescenza verde. (Vedi Figura 20.10 per maggiori dettagli.)



20.20 a. Alcune malattie umane non mostrano sintomi fino a un'età avanzata. Poiché la durata della vita del topo è più breve, i topi invecchiano più velocemente. Pertanto, a seconda della biochimica e della fisiologia alla base della malattia, è possibile che i topi omozigoti per il knockout muoiano per cause naturali prima che si manifestino gli effetti del knockout. Come discusso nella parte (c) del Problema 14, questo fenomeno ha di fatto ridotto l'utilità dei modelli murini per la malattia di Huntington.

b. L'omozigosi per alcune mutazioni con perdita di funzione negli esseri umani potrebbe non essere letale, ma potrebbe causare fenotipi di malattie negli individui che nascono. L'omozigosi per knockout del gene corrispondente potrebbe essere letale in utero nei topi. Pertanto, i topi omozigoti potrebbero non manifestare i sintomi della malattia perché questi topi potrebbero non nascere mai. Un modo per aggirare questo pro-

blema è generare un modello murino per la malattia utilizzando un knockout condizionale in cui il gene può essere inattivato in un tessuto specifico e solo dopo la nascita.

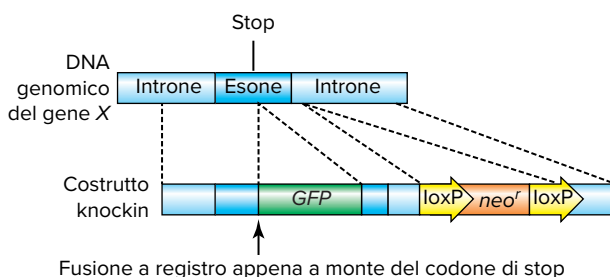
c. Se i topi hanno diverse copie di un gene, è probabile che un knockout solo per una di queste copie non abbia effetto, perché le altre copie potrebbero fornire il prodotto genico. Due strategie potrebbero consentire ai ricercatori di affrontare questo problema. In primo luogo, tutte le copie del gene potrebbero essere eliminate in sequenza per ottenere un modello della malattia umana a perdita di funzione. In alternativa, un transgene che esprime un RNAi potrebbe potenzialmente diminuire o eliminare l'espressione di tutte le copie del gene se si potesse progettare un singolo siRNA omologo a tutti questi geni.

d. I topi di diverse linee consanguinee potrebbero avere alleli distinti di geni modificatori (soppressori e potenziatori), con conseguente diversa

espressività di un fenotipo mutante knockout; infatti, il fenotipo mutante potrebbe non essere osservato affatto in alcuni background genetici. Se un particolare topo knockout non mostra un fenotipo mutante, potrebbe essere utile incrociarlo con un ceppo mutante diverso e quindi rigenerare gli omozigoti. Mescolare due diversi background genetici potrebbe eliminare alcuni soppressori e rivelare un fenotipo simile a una malattia umana.

e. I topi knockout in cui l'espressione dei geni vicini è inibita potrebbero mostrare fenotipi completamente estranei alla malattia umana studiata. Sarebbe quindi difficile sapere quali fenotipi sono associati al gene knock-out e quali al gene vicino. Possibili modi per affrontare questa situazione includono la creazione di diversi tipi di knockout che non dovrebbero influenzare i geni adiacenti (per esempio, utilizzando il targeting genetico per introdurre una mutazione non senso nel gene di interesse) o l'inibizione della funzione di questo gene mediante l'interferenza a RNA.

- 20.21 a.** La figura che segue illustra un costrutto knockin che consentirebbe di marcare con la GFP il prodotto proteico a lunghezza intera del gene in questione. Si possono ottenere le sequenze codificanti GFP dal cDNA e inserirle a registro nell'ORF del gene, appena a monte del codone di stop. In questo modo, il gene alterato codificherebbe per una proteina di fusione in cui tutti gli amminoacidi del prodotto genico sono seguiti al C-terminale dalla GFP. Il gene *neo^r* è richiesto come marcatore selezionabile per identificare le cellule ES che hanno integrato il costrutto knockin mediante ricombinazione omologa.



b. La strategia knockin ha il vantaggio che le sequenze regolatrici del gene *X* non hanno bisogno di essere caratterizzate perché sono già presenti nel gene *X* nel cromosoma del topo. Supponiamo di aver impiegato invece un transgene che esprime una proteina di fusione Proteina X-GFP. Per

garantire che la proteina di fusione sia prodotta in quelle cellule in cui il gene *X* è normalmente espresso, il transgene deve contenere un frammento di DNA con le sequenze regolatrici del gene *X*. Sarebbe necessario condurre diversi esperimenti per identificare un tale frammento di DNA.

- 20.22 a.** Le delezioni possono essere effettuate utilizzando CRISPR/Cas9, introducendo contemporaneamente due diversi sgRNA che indirizzano la scissione di Cas9 nei punti di rottura desiderati delle delezioni. Cas9 creerà due rotture a doppio filamento nel DNA che, se vengono riparate da NHEJ, portano spesso alla delezione del DNA compreso tra i due punti di rottura.

b. Un esone del gene *PPO* potrebbe essere eliminato come descritto nella parte (a). Ciascuno dei due sgRNA sarà complementare a un'estremità dell'esone.

- 20.23 a.** Si può clonare il mammut lanoso mediante trasferimento nucleare di cellule somatiche. Si trasferiscono i nuclei dalle cellule di mammut negli oociti enucleati di elefanti asiatici, quindi si impiantano gli oociti con nuclei di cellule somatiche in madri surrogate di elefanti asiatici. (Vedi, per i dettagli, il riquadro Gli strumenti della genetica "Clonazione mediante trasferimento di nuclei di cellule somatiche".) Per quanto ne sappiamo, nessuna cellula vivente di mammut lanoso è stata mai recuperata, quindi questa parte della domanda è ipotetica.

b. (I) In primo luogo, si identificano i geni coinvolti nella struttura dei peli. Questo è molto più facile a dirsi che a farsi. Si potrebbero cercare i geni noti per influenzare la struttura dei peli in altri animali e verificare se sono diversi tra mammut ed elefante secondo modalità che potrebbero avere un senso. Assumiamo nei passaggi successivi che è effettivamente possibilmente identificare tali geni, ma sicuramente non sarà facile.

(II) Si può utilizzare la tecnica CRISPR/Cas9 per modificare gli alleli di quei geni nelle cellule in coltura degli elefanti asiatici per renderli simili a quelli dei mammut lanosi.

(III) Quindi, tramite il trasferimento nucleare di cellule somatiche, si genera un elefante asiatico il cui genoma contiene i geni dei peli di mammut lanoso. Se l'esperimento ha successo, l'elefante asiatico risultante avrà peli più folti rispetto ai normali elefanti asiatici. Naturalmente, se le differenze nella struttura dei peli dipendono da più di una variante, potrebbe essere necessario ripe-

tere questo passaggio per molti o tutti i geni per creare un elefante con una struttura dei peli indistinguibile da quella dei mammut.

c. Tutte le differenze di sequenza genica tra l'elefante asiatico e il mammut, che influenzano il fenotipo, potrebbero essere modificate nel genoma dell'elefante. (Studi preliminari indicano che esistono più di 1500 differenze di questo tipo.) I nuclei con il genoma di elefante modificato potrebbero essere utilizzati per generare un animale attraverso il trasferimento nucleare di cellule somatiche. Un tale animale sarebbe in teoria simile al mammut lanoso estinto.

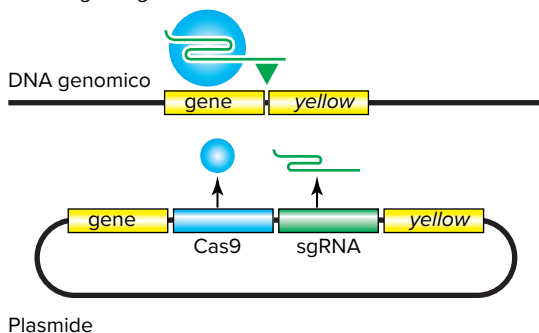
Alterare più di 1500 sequenze rappresenta un'ardua sfida tecnologica. Un'altra sfida tecnica è mettere incinta un numero sufficiente di surrogati di elefanti asiatici per ottenere prole sana e vitale. L'idea della de-estinzione crea problemi etici perché l'animale che viene generato è in realtà un ibrido, mai effettivamente esistito,

tra un elefante e un mammut lanoso; non possiamo sapere esattamente quali saranno le sue caratteristiche e come sarebbe la sua vita. Inoltre, la de-estinzione degli animali potrebbe avere effetti ecologici non previsti.

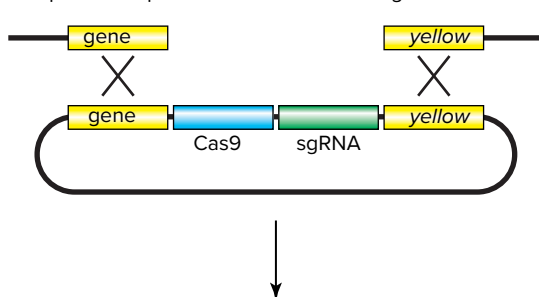
20.24 Una rottura a doppio filamento (DSB) perfettamente riparata è ancora un substrato per CRISPR/Cas9, quindi Cas9 continuerà a generare DSB fino a quando, per caso, non si verificherà una riparazione imperfetta che distrugge il sito target dell'sgRNA.

20.25 a. L'allele mutante *yellow* nello spermatozoo conteneva un'inserzione dei geni *Cas9* e *sgRNA*. La loro espressione ha portato a un DSB all'interno del gene *yellow* selvatico, che è stato ereditato nella F_1 dalla madre. La ricombinazione omologa per riparare il DSB porta all'inserimento dei geni *Cas9* e *sgRNA* nella copia materna. Il risultato è che le cellule somatiche non hanno più un gene *yellow*⁺ funzionante.

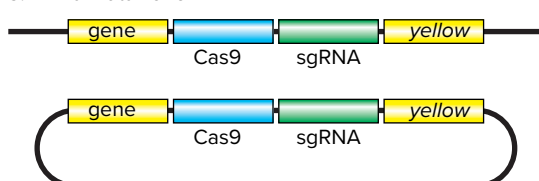
1. Bersaglio tagliato



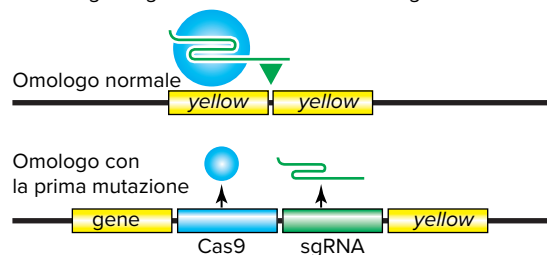
2. Riparazione per ricombinazione omologa



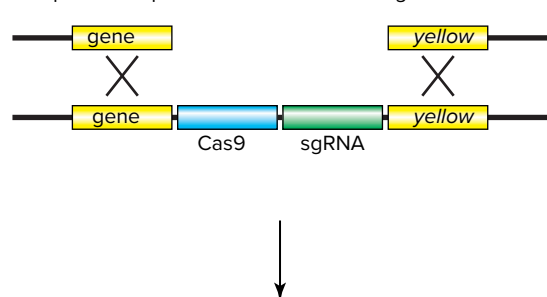
3. Prima mutazione



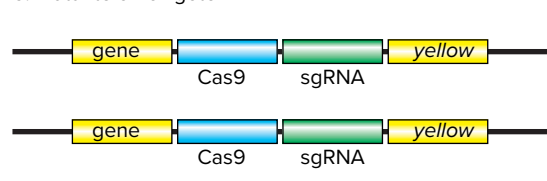
4. Bersaglio tagliato sul cromosoma omologo



5. Riparazione per ricombinazione omologa



6. Mutante omozigote



b. La trasmissione di un allele *yellow* mutante con un'inserzione del gene *Cas9/sgRNA* risulta in una progenie omozigote con tutti (o quasi tutti) i gameti mutanti per *yellow*. Quindi, come visto nella parte (a), la progenie risultante da questi gameti sarà omozigote per l'allele mutante *yellow* sia nelle cellule somatiche sia in quelle della linea germinale.

c. Gli "alleli di resistenza" si generano quando la rottura del doppio filamento viene riparata in modo errato dal NHEJ, portando a inserzioni o delezioni che distruggono il sito target. Questi potrebbero generare un allele mutante *yellow* resistente se la mutazione cambia il registro di lettura o colpisce amminoacidi cruciali per la funzione del gene *yellow*. In alternativa, l'allele di resistenza potrebbe funzionare come un allele *yellow*⁺ wild-type se la mutazione non influisce sulla funzione del gene.

20.26 a. L'sgRNA nel costrutto del gene drive illustrato di seguito indirizza Cas9 a operare una rottura a doppio filamento al confine introne 4/esone 5.



b. Il costrutto del gene drive si inserirebbe per ricombinazione omologa al confine introne 4/esone 5, distruggendo il sito accettore di splicing. L'esone 5 femminile specifico verrà saltato e non verrà incorporato nell'mRNA di *dsx*.

c. Nonostante il fatto che le femmine omozigoti per l'inserzione del gene drive siano sterili, il gene drive può diffondersi nella popolazione attraverso i maschi, perché la produzione di Dsx-M non è influenzata dal gene drive. È importante che il gene *dsx* sia su un autosoma. Se *dsx* fosse sul cromosoma X in *Anopheles gambiae*, solo i maschi originali che portano il costrutto del gene drive che vengono rilasciati nella popolazione potrebbero trasmettere l'X mutante; la loro progenie femminile sarebbe sterile e la loro progenie maschile avrebbe l'Y del padre, non l'X del padre. Poiché *dsx* è autosomico, la progenie maschile dei maschi con il gene drive, originariamente rilasciati, fornirà una fonte costante di maschi mutagenici.

d. Gli effetti ecologici dell'eradicazione delle zanzare non possono essere conosciuti con certezza in anticipo e una volta che viene rilasciato un gene drive letale, non si può tornare indietro.

20.27 a. Nel sistema immunitario batterico, la parte rossa dell'sgRNA mostrato in Figura 20.13 rap-

presenta il crRNA e la parte verde il tracrRNA. Il crRNA fornisce la specificità del bersaglio; il tracrRNA lega sia Cas9 sia parte del crRNA, portando così Cas9 al bersaglio.

b. L'assenza del sito PAM nel locus CRISPR batterico impedisce la scissione del cromosoma batterico da parte di Cas9. Il sistema immunitario si è evoluto in modo tale che solo le sequenze Spacer del genoma virale adiacenti a un sito PAM, ma non il sito PAM stesso, sono incorporate nel locus CRISPR. Questo punto è cruciale, perché il sistema CRISPR/Cas9 deve essere in grado di distruggere i genomi dei batteriofagi patogeni (che hanno il sito PAM adiacente agli Spacer) ma, contemporaneamente, di non distruggere il genoma della cellula batterica stessa (che manca di siti PAM adiacenti agli Spacer all'interno del locus CRISPR). Il meccanismo batterico per selezionare le sequenze da integrare nel locus CRISPR è molto complesso e coinvolge almeno altri due enzimi (Cas1 e Cas2) non discussi nel testo.

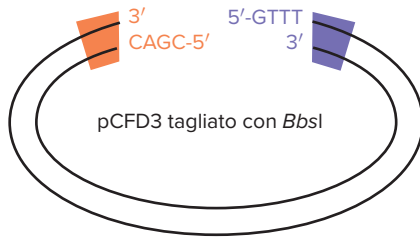
20.28 a. I due siti PAM sono mostrati in rosso nel diagramma seguente:

MetThrAsnSerTyrAspIleHisSer
5'...GTTAAAGTATGACTAACAGCTACGACATACACAGTTGgtgaggttggcatc...3'

Si deve utilizzare il primo sito PAM (quello che si trova parzialmente nella sequenza che codifica per gli amminoacidi). Si noti che con un sgRNA che indirizza Cas9 a questo sito PAM, la scissione avverrebbe nella sequenza codificante. Le indel che si verificano in questa posizione potrebbero creare una mutazione frameshift che fungerebbe da allele knockout. Al contrario, la scissione vicino al secondo sito PAM (quello completamente all'interno dell'introne) creerebbe indel che è improbabile alterino l'espressione genica o il prodotto proteico a meno che l'indel non sia abbastanza grande da rimuovere anche parte della sequenza codificante o il sito donatore di splicing.

b. Dato che un codone è di 3 bp, ci si aspetta che circa 1/3 delle delezioni casualmente mantenga il registro di lettura aperto. I 2/3 che causano mutazioni frameshift produrranno quasi certamente alleli nulli (a meno che siano molto vicine al termine dell'ORF).

c. Le protrusioni al 5' lasciate da *BbsI* sono diverse tra loro: 5'-CGAC e 5'-GTTT. Si vedrà nella parte (d) che questo fatto è importante per legare un inserto nel vettore con un orientamento specifico.



d. I due oligo che potrebbero essere rinaturati e ligati nel vettore tagliato mostrato nella parte (c) sono:

5' **GTGACAGCTACGACATACACAGT** 3'
3' TGTCGATGCTGTATGTGTCA**CAAA** 5'

Si noti che gli oligonucleotidi devono avere 4 basi sporgenti al 5' complementari alle estremità al vettore pCFD3 tagliato con *BbsI* nella parte (c). L'sgRNA non può includere il sito PAM; gli oligonucleotidi contengono ciascuno 20 bp a monte del sito PAM; questi sono mostrati in blu come nella sequenza mostrata nella parte (a).

e. Cas9 scinderà il gene *NiPp1* tra i codoni His e Ser a monte del primo sito PAM nel diagramma della parte (a):

5'...GTTAAAGT**ATGACTAACAGCTACGACATACAC** AGT**G**tgagtt**ggc**atc...3'
MetThrAsnSerTyrAspIleHis Ser
PAM

f. (I) Si iniettano i plasmidi, prodotti nella parte (d), nell'estremità posteriore degli embrioni precoci del moscerino (le cellule germinali primordiali sono in tale sede). La miscela di iniezione deve contenere anche una fonte di proteina Cas9: si può iniettare la proteina stessa, o un RNA che può essere tradotto in proteina Cas9, o un DNA in cui la regione codificante Cas9 è posta a valle di un promotore forte. Nelle cellule germinali, la proteina Cas9 e l'sgRNA agiranno insieme per tagliare il gene *NiPp1* nella posizione indicata nella parte (e), e il NHEJ produrrà mutazioni indel in quel sito. È importante notare che in cellule germinali diverse di un singolo animale iniettato si verificheranno mutazioni indel diverse.

(II) Quando gli iniettati diventano adulti, si incrociano con moscerini non iniettati.

(III) Quindi, si prendono diversi discendenti F_1 e si incrociano singolarmente con moscerini non iniettati. È fondamentale prendere la prole F_1 singolarmente perché ogni F_1 potrebbe avere una mutazione diversa (o nessuna mutazione). Facendo questo secondo incrocio, si isola di fatto un singolo cromosoma mutagenizzato.

(IV) Se si è fortunati, alcuni discendenti F_2 di alcune linee saranno eterozigoti per una mutazio-

ne nulla di *NiPp1*. È possibile identificare le linee in cui ciò si è verificato utilizzando un test PCR (si amplifica il gene *NiPp1* di uno o pochi moscerini in ciascuna linea, per verificare se è stata indotta una nuova mutazione indel).

(V) I maschi e le femmine eterozigoti F_1 possono essere incrociati tra loro per generare omozigoti mutanti *NiPp1*.

[Come nota tecnica, sarebbe vantaggioso utilizzare cromosomi bilanciatori nei passaggi II-IV in modo da poter distinguere i cromosomi mutagenizzati che sono stati sottoposti all'azione di Cas9 dai cromosomi non mutagenizzati.]

g. Insieme al plasmide, iniettare negli embrioni molecole di DNA a doppio filamento lineare con il codone Thr (ACT) cambiato in un codone Ala (GCT). Queste molecole a doppio filamento devono contenere sequenze a sinistra e a destra del codone in questione. In alcune cellule della linea germinale, la riparazione del DSB da parte di Cas9 avverrà per ricombinazione omologa, con conseguente sostituzione del codone Thr con il codone Ala (vedi Figura 20.13, in basso a destra).

20.29 a. I vettori retrovirali inseriscono copie di DNA a doppio filamento del gene terapeutico nel nucleo delle cellule bersaglio e queste copie vengono quindi incorporate mediante la trascrittasi/endonucleasi inversa in un cromosoma. I vettori AAV forniscono alle cellule bersaglio copie del gene che rimangono extracromosomiche.

b. Il vantaggio dei vettori retrovirali è che il gene terapeutico è stabile in un cromosoma. Al contrario, quando si utilizzano vettori AAV, poiché il gene terapeutico è extracromosomico, il DNA alla fine si degrada e la terapia genica non è stabile e deve essere ripetuta. Tuttavia, lo svantaggio principale dei vettori retrovirali è che possono causare mutazioni quando il gene terapeutico si integra nel cromosoma. Se l'evento di integrazione interrompe un gene, il risultato è una mutazione a perdita di funzione. L'evento di integrazione potrebbe in alternativa provocare una mutazione a guadagno di funzione se il forte enhancer nella lunga ripetizione terminale del vettore retrovirale (LTR) provoca la sovra-espressione di un gene vicino. Un altro inconveniente dei vettori retrovirali è che le cellule umane tendono a creare una risposta immunitaria più forte ai retrovirus rispetto agli AAV.

20.30 L'alterazione del genoma delle cellule germinali umane solleva così tante questioni etiche che è illegale nella maggior parte dei Paesi. Una delle

principali preoccupazioni è che se le alterazioni del genoma sono trasmesse attraverso la linea germinale colpiscono ogni cellula del corpo della progenie risultante e possono verosimilmente produrre conseguenze fenotipiche impreviste nell'individuo e nei suoi discendenti per le generazioni a venire. Queste modifiche potrebbero non essere facilmente annullate. Inoltre, molti esperti di etica sono preoccupati per le implicazioni dell'alterazione dei genomi per migliorare specifiche caratteristiche dei figli e dei loro discendenti: le modificazioni sono davvero apportate per aiutare la progenie, o riflettono invece quelli che sono i desideri dei genitori? Come si può tracciare il confine?

20.31 a. L'allele mutante *RPE65* dominante deve essere una mutazione a guadagno di funzione perché gli alleli *RPE65* a perdita di funzione che causano LCA sono recessivi rispetto al tipo selvatico; *RPE65* non è aploinsufficiente.

b. Il tipo di terapia genica utilizzata per l'LCA non funzionerebbe per la retinite pigmentosa causata dall'allele dominante *RPE65*. La strategia di terapia genica per LCA è la sostituzione degli alleli a perdita di funzione con una copia del gene di tipo selvatico. Un gene terapeutico per la retinite pigmentosa dovrebbe in qualche modo inibire la funzione dell'allele mutante dominante *RPE65* o del suo prodotto genico.

20.32 a. Si cerca il sito PAM (su entrambi i filamenti di DNA) più vicino alla sequenza GTG che si intende cambiare in GAG. Si trova una sequenza 5'AGG (mostrata in blu nel diagramma sottostante) appena a monte del codone *Hb β^S* mutante (la base mutante è sottolineata e in grassetto). La sequenza sgRNA è mostrata in rosso e le frecce indicano i siti di taglio di Cas9.



b. Oltre all'sgRNA, per cambiare il codone GTG in GAG è necessario aggiungere Cas9 ai precursori dei globuli rossi e anche un DNA a doppio filamento contenente una porzione del gene *Hb β^A* che include il codone GAG. Il DNA genomico dell'*Hb β^A* subirà una ricombinazione omologa con il gene *Hb β^S* nel sito di taglio di Cas9 e sostituirà l'allele *Hb β^S* mutante.

20.33 a. Un problema per la progettazione di un siRNA che impedisca l'espressione dell'allele mutante in

modo specifico è che i siRNA possono legarsi e inibire l'espressione di mRNA senza una perfetta complementarità di sequenza. Le mutazioni dominanti negli alleli della malattia sono spesso modificazioni di una singola base che alterano un codone. Pertanto, è possibile che un siRNA con perfetta complementarità per l'allele della malattia sia ancora in grado di inibire l'espressione dell'allele normale, dato che la differenza è solo di una singola coppia di basi.

Un modo in cui gli scienziati cercano di aggirare questo problema è cercare differenze silenziose di coppie di basi tra l'allele normale e quello mutante che possano essere discriminate da un RNA interferente. Se si trovassero tali polimorfismi, allora potrebbe essere possibile progettare RNA interferenti che differiscono dall'allele wild-type nel paziente in più di una posizione.

Il problema è ancora più difficile nel caso di alleli dominanti rari che causano la sovra-espressione di un normale prodotto proteico; qui, gli mRNA prodotti dall'allele mutante e da quello di tipo selvatico potrebbero essere identici. Il problema potrebbe essere aggirato solo se gli mRNA prodotti dai due alleli non fossero in effetti identici a causa dei polimorfismi silenziosi.

b. Le malattie da espansione di triplette, come la malattia di Huntington, rappresentano sempre una sfida per la progettazione di un RNA interferente che possa colpire solo l'allele mutante. Il motivo è che la differenza tra l'allele normale e quello della malattia non è una sostituzione di base, ma una differenza nel numero di copie di una ripetizione di triplette. Poiché la regione di ripetizione delle triplette negli alleli normali è solitamente più lunga di un RNA interferente (~21-30 nt), è impossibile progettare un RNA interferente che si appai solo con un allele mutante con una regione di ripetizione delle triplette espansa. Come descritto nella parte (a), un possibile modo per aggirare questo problema è considerare le differenze di sequenza silenziose tra gli alleli wild-type e mutanti.

c. CRISPR/Cas9 non sarebbe utile per accorciare le ripetizioni del trinucleotide CAG negli alleli della malattia di Huntington perché queste unità ripetute non contengono sequenze NGG PAM. Pertanto, Cas9 non può tagliare al centro del tratto ripetuto.

Se la sequenza univoca immediatamente adiacente alla regione di ripetizione contenesse una sequenza PAM, con un approccio alternativo, si potrebbe forse usare Cas9 per tagliare alla fine

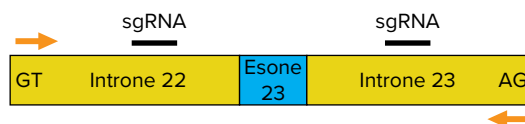
della regione di ripetizione. Ma questa idea ha anche potenziali ostacoli se l'unico modo per riparare la rottura del doppio filamento richiedesse il NHEJ. Eventuali delezioni di grandi dimensioni che potessero accorciare il tratto ripetuto potrebbero estendersi in entrambe le direzioni, danneggiando altre parti del gene. Si potrebbe tuttavia sostituire la regione di ripetizione del trinucleotide con una contenente meno ripetizioni CAG introducendo un costrutto appropriato per creare un knockin mediante ricombinazione omologa.

20.34 a. Il Cas9/sgRNA riconosce 20 basi complementari allo sgRNA e anche il sito PAM (5' NGG). Pertanto, vengono specificate 22 basi (20 per l'sgRNA e 2 per il sito PAM). In un singolo filamento di DNA, la frequenza di una qualsiasi di queste sequenze da 22 nt è 1 su 4^{22} , cioè 1 su più di 7×10^{25} sequenze di nucleotidi! (In altre parole, esistono più di 70 quadrilioni di diverse sequenze da 22 nt, e quella che questo complesso Cas9/sgRNA riconosce è una di esse.) Pertanto, la possibilità che una particolare sequenza da 22 nt appaia in una sequenza casuale di DNA a filamento singolo, cioè 3 miliardi di nucleotidi, è $(3 \times 10^9) \times 1/4^{22}$. La probabilità di trovare una particolare sequenza di 22 nt è due volte più frequente nel DNA a doppio filamento: $2 \times (3 \times 10^9) \times 1/4^{22} = 0,000341$, che è $\ll 1$ ($\sim 1/2930$). In breve, ci sono poche possibilità che si possa trovare una corrispondenza perfetta per una particolare sequenza di 22 nt più di una volta in una sequenza delle dimensioni del genoma umano. Naturalmente, quando gli scienziati progettano sgRNA, usano BLAST per cercare nella sequenza di riferimento umana le corrispondenze che possono tradursi in effetti fuori bersaglio, che in realtà, come si vedrà nella parte (b) di questo problema, non devono essere necessariamente esatte.

b. Quante sequenze differenti corrisponderanno per caso a 19 delle 20 basi dell'sgRNA? La mancata corrispondenza può verificarsi in una qualsiasi delle 20 basi, quindi ci sono 20 diverse posizioni in cui può verificarsi la mancata corrispondenza. Se la mancata corrispondenza è nella base n. 1, sono possibili 3 basi diverse alla base n. 1 e le restanti 19 basi devono corrispondere (hanno solo una opzione). Quindi esistono 3 diverse sequenze di 20 nt che hanno una mancata corrispondenza alla base n. 1, e lo stesso vale per le sequenze di 20 nt che hanno una mancata corrispondenza in una qualsiasi delle altre 19 basi: esistono $3 \times 20 = 60$ diverse sequenze che hanno

una mancata corrispondenza in una delle 20 basi. Si ricordi che anche le due basi definite nel sito PAM devono corrispondere all'sgRNA: la sequenza che l'sgRNA riconosce è quindi lunga 22 nt. Pertanto, utilizzando la stessa logica della parte (a), la possibilità dell'esistenza nel genoma umano di una sequenza di 22 nt con un singolo mismatch in una delle 20 basi dell'sgRNA che potrebbero essere prese di mira da questo sgRNA insieme a Cas9 è $2 \times 60 \times (3 \times 10^9) \times 1/4^{22} \approx 0,02046$ o $\sim 1/49$. Sebbene questa probabilità sia molto più alta di quella che calcolata nella parte (a), c'è ancora solo una piccola possibilità di un effetto fuori bersaglio per uno specifico sgRNA.

20.35 a. Come mostra il diagramma che segue, gli sgRNA possono ibridarsi con un filamento o l'altro del DNA all'interno di ciascun introne appena a monte di una sequenza PAM. Il sito di giunzione 5' (GT) nell'introne 22 e il sito di giunzione 3' (AG) nell'introne 23 devono essere conservati; inoltre, non deve essere eliminata la sequenza di diramazione del lazo. Le rotture del doppio filamento in entrambi gli introni, seguite dalla riparazione tramite NHEJ, comporteranno la delezione dell'esone 23.



b. Come indicato dalle frecce arancioni nel diagramma, i primer PCR dovrebbero affiancare gli sgRNA. La delezione dell'esone 23 può essere rilevata da un prodotto PCR più piccolo del normale. Oppure, se nel DNA di tipo selvatico la regione tra i primer della PCR è troppo grande per essere amplificata, solo il DNA genomico con la delezione potrà dare un prodotto di PCR.

c. La proteina Dmd senza gli amminoacidi codificati dall'esone 23 conserva almeno qualche funzione. Questo risultato indica che questi amminoacidi non sono necessari affinché la proteina funzioni almeno in una certa misura. Sorge quindi la domanda sul perché questa procedura di eliminazione dell'esone non abbia prodotto una cura efficace. La ragione più ovvia è che la delezione dell'esone 23 si è verificata solo nel 10% delle cellule muscolari, quindi il 90% delle cellule aveva ancora la mutazione non senso nell'esone 23. È anche possibile che la proteina Dmd, priva degli amminoacidi codificati dall'esone 23, possa funzionare solo in modo inefficiente o che lo splicing del trascritto con l'esone eliminato sia inefficiente.

d. No; la strategia dell'eliminazione dell'esone non funzionerebbe per le mutazioni puntiformi all'interno degli esoni che codificano per gli aminoacidi cruciali per la funzione proteica o il corretto ripiegamento delle proteine.

e. Quando l'esone 22 viene unito all'esone 24, il registro di lettura deve essere preservato. Ciò significa che i registri di lettura dell'estremità 3' dell'esone 22, delle estremità 5' e 3' dell'esone 23 e dell'estremità 5' dell'esone 24 devono essere tutti uguali.