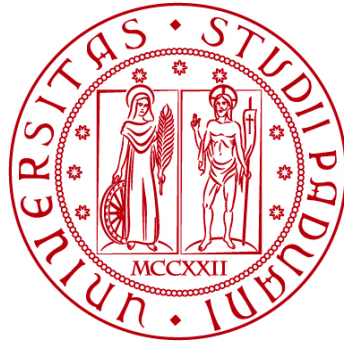


UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI BIOLOGIA

Corso di Laurea in Biotecnologie



ELABORATO DI LAUREA

**EDITING DEL GENE *Hrb87F* MEDIANTE
CRISPR/CAS9 IN CELLULE SR2+**

**Tutor: Prof.ssa Gabriella Margherita Mazzotta
Dipartimento di Biologia**

**Co-tutor: Dott.ssa Chiara Stefanelli
Dipartimento di Biologia**

Laureanda: Giulia Laffranchi

ANNO ACCADEMICO 2025/2026

Sommario

Abstract.....	1
1 INTRODUZIONE.....	2
1.1 Le ribonucleoproteine nucleari eterogenee	2
1.1.1 hnRNP A1.....	3
1.1.2 Hrb87F.....	5
1.2 CRISPR/Cas9.....	6
2 MATERIALI E METODI.....	8
2.1 Linea cellulare	8
2.1.1 Trasfezione cellule.....	8
2.1.2 Selezione con puromicina.....	8
2.2 Estrazione del DNA genomico.....	9
2.2.1 Quantifica con <i>NanoDrop</i> TM	9
2.3 Reazione a catena della polimerasi (PCR).....	9
2.3.1 Disegno dei <i>primer</i>	10
2.3.2 Elettroforesi su gel di agarosio	10
2.3.3 Purifica prodotti di PCR	11
2.4 Sequenziamento Sanger	11
2.4.1 Analisi bioinformatica <i>TIDE</i>	11
2.5 Estrazione delle proteine.....	11
2.5.1 Quantifica tramite Bradford	11
2.6 Western Blot	12
2.6.1 Blottaggio	12
2.6.2 Bloccaggio.....	12
2.6.3 Ibridazione.....	12
2.6.4 Analisi densitometrica tramite <i>ImageJ</i>	13
2.7 Estrazione dell'RNA	13
2.7.1 Spettro per l'integrità dell'RNA (<i>TapeStation</i>).....	13
3 RISULTATI E DISCUSSIONE	14
3.1 Analisi quantitativa e qualitativa del DNA genomico (gDNA)	14
3.2 Verifica dell'amplificazione del DNA target tramite elettroforesi su gel d'agarosio.....	15
3.3 Analisi quantitativa e qualitativa dei prodotti di PCR.....	15
3.4 Valutazione e confronto dell'efficienza di <i>editing</i> delle sgRNA testate	16
3.5 Analisi quantitativa delle proteine.....	18
3.6 Verifica del trasferimento proteico tramite colorazione con Rosso Ponceau	19
3.7 Valutazione dell'espressione proteica mediante Western Blot	20
3.8 Analisi quantitativa e qualitativa dell'RNA.....	21
3.9 Conclusioni.....	21
Bibliografia.....	22

Abstract

Le ribonucleoproteine eterogenee nucleari (hnRNP) svolgono un ruolo chiave nella regolazione e nella maturazione dell'RNA cellulare. Questo progetto di ricerca ha l'obiettivo di caratterizzare l'hnRNP *Hrb87F* in *Drosophila melanogaster*: è stata applicata la tecnologia CRISPR/Cas9 per valutare l'efficacia di diverse guide sgRNA per la generazione di una linea cellulare S2R+ *knock-out*, per poter in seguito valutare gli effetti dell'assenza della proteina sul trascrittoma. Utilizzando il vettore plasmidico pAc-sgRNA-Cas9, contenente la cassetta d'espressione per Cas9 e la resistenza alla puromicina, sono state clonate e trasfettate tre diverse sequenze di sgRNA (sgHrb87F-147, sgHrb87F-170, sgHrb87F-208). Le cellule trasfettate sono state sottoposte a selezione con puromicina per ottimizzare l'efficienza del *knock-out*. I campioni sono stati, poi, raccolti e processati per l'estrazione di DNA, proteine e RNA, per confermare che l'*editing* genomico fosse effettivamente avvenuto.

I risultati ottenuti confermano l'efficacia del sistema CRISPR/Cas9 nell'ingegnerizzazione di cellule S2R+ e pongono le basi per futuri studi di identificazione dell'attività di *Hrb87F* sui suoi target biologici.

1 INTRODUZIONE

1.1 Le ribonucleoproteine nucleari eterogenee

Le ribonucleoproteine nucleari eterogenee (hnRNP, *heterogeneous nuclear ribonucleoproteins*) rappresentano un'ampia famiglia di proteine che legano l'RNA (RBP, *RNA binding proteins*) con molteplici funzioni nel metabolismo degli acidi nucleici, tra cui lo *splicing* alternativo, la stabilizzazione dell'mRNA e la regolazione trascrizionale e traduzionale (Geuens et al., 2016). Storicamente, le hnRNP sono state definite come proteine che si associano ai trascritti nascenti, contribuendo a controllare la maturazione degli RNA nucleari eterogenei (hnRNA/pre-mRNA) in RNA messaggeri (mRNA), stabilizzando gli mRNA durante il loro trasporto cellulare e controllandone la traduzione (Jean-Philippe et al., 2013).

Il ruolo delle hnRNP nella regolazione dell'espressione genica ha suscitato un crescente interesse nella ricerca sulle malattie. Il livello di espressione delle hnRNP risulta alterato in molti tipi di cancro, suggerendo un loro ruolo nella genesi dei tumori. Oltre al cancro, molte hnRNP sono state collegate anche a varie malattie neurodegenerative, come l'atrofia muscolare spinale (SMA), la sclerosi laterale amiotrofica (SLA), la malattia di Alzheimer (AD) e la demenza frontotemporale (FTLD) (Geuens et al., 2016; Stefanelli et al., 2026).

Le hnRNP condividono caratteristiche generali, ma differiscono nella composizione dei domini e nelle proprietà funzionali; infatti, la famiglia delle hnRNP è composta da almeno 20 hnRNP principali abbondanti, denominate da A a U, con pesi molecolari che vanno da 34 kDa a 120 kDa, e da altri membri minori meno abbondanti (Figura 1) (Geuens et al., 2016).

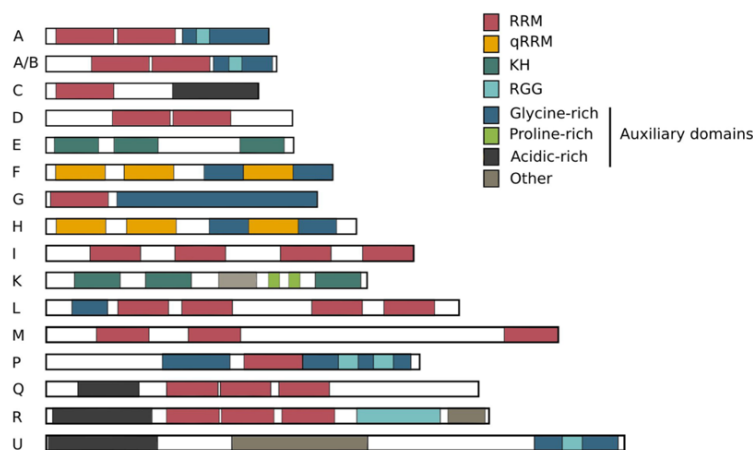


Figura 1 La famiglia delle hnRNP. Le hnRNP sono denominate in ordine alfabetico da hnRNP A1 a hnRNP U. Diversi domini strutturali sono condivisi tra i vari membri della famiglia, il che spiega le simili proprietà di legame all'RNA: RRM (motivo di riconoscimento dell'RNA), qRRM (motivo quasi-RRM), dominio KH (dominio di omologia K), RGG (dominio di legame all'RNA) (Geuens et al., 2016).

Le hnRNP presentano una struttura modulare costituita da uno o più domini di legame all'RNA (RBD, *RNA binding domains*), tra i quali i motivi di riconoscimento dell'RNA (RRM, *RNA recognition motifs*) sono i più comuni. Gli RRM sono dei

domini globulari altamente conservati, lunghi circa 90 amminoacidi, che hanno la capacità di partecipare a interazioni sia specifiche che aspecifiche con gli acidi nucleici. Strutturalmente sono caratterizzati da quattro foglietti β e due α -eliche ($\beta\alpha\beta\beta\alpha\beta$) insieme a due sequenze conservate, l'ottamero RNP1 e l'esamero RNP2, posizionate a circa 30 residui amminoacidici l'una dall'altra. Sono i loop variabili, che collegano i foglietti β , che contribuiscono alla loro specificità di legame all'RNA (Jean-Philippe et al., 2013; Geuens et al., 2016). Altri RBD che contribuiscono alla specificità di legame delle singole hnRNP sono: il quasi-RRM, un dominio ricco di glicina che costituisce il box RGG e il dominio di omologia KH. I box RGG sono caratterizzati da ripetizioni di Arginina-Glicina-Glicina (Arg-Gly-Gly) intervallate da residui aromatici di Fenilalanina (Phe) e Tirosina (Tyr), e sono visti come domini ausiliari, responsabili delle interazioni omologhe ed eterologhe con altre hnRNP. I domini KH sono caratterizzati da una sequenza evolutivamente conservata di circa 70 amminoacidi e strutturalmente presentano un foglietto β antiparallelo a tre filamenti e tre α -eliche ($\beta\alpha\alpha\beta\beta\alpha$). In base alle loro estensioni N- o C-terminali, i domini KH sono suddivisi in due categorie: i domini KH di tipo I presentano un'estensione $\beta\alpha$ al loro C-terminale, mentre i domini KH di tipo II presentano un'estensione $\alpha\beta$ al loro N-terminale (Geuens et al., 2016). Oltre agli RBD, le hnRNP contengono frequentemente domini ausiliari, come domini ricchi di glicina, acidi o ricchi di prolina, che mediano le interazioni proteina-proteina, la localizzazione subcellulare e la specificità funzionale delle singole proteine. Le combinazioni di questi diversi domini aumentano la diversità funzionale delle hnRNP.

Le proteine hnRNP subiscono diverse modifiche post-traduzionali, tra cui fosforilazione, sumoilazione, ubiquitinazione e metilazione, che ne alterano la localizzazione, la specificità di legame all'RNA e l'interazione con altri fattori cellulari. Anche se generalmente le hnRNP hanno una distribuzione nucleare diffusa, tanto che la maggior parte possiede un segnale di localizzazione nucleare (NLS, *nuclear localization signals*) ed è prevalentemente presente nel nucleo in condizioni di omeostasi, molte rimangono legate all'mRNA mentre viene trasportato attraverso i pori nucleari, si associa al ribosoma o viene sequestrato in specifici compartimenti cellulari (Jean-Philippe et al., 2013).

1.1.1 hnRNP A1

L'hnRNP A1 è uno dei membri più abbondanti e ubiquitariamente espressi della famiglia delle hnRNP, fa parte della sottofamiglia hnRNP A/B. Formando complessi con i pre-mRNA e gli mRNA maturi, l'hnRNP A1 svolge molteplici ruoli nella regolazione dell'espressione genica, modulando lo *splicing* alternativo, il *capping*, lo *splicing* e la poliadenilazione dei trascritti nascenti, ma anche la stabilità, l'esportazione attraverso i pori nucleari e la traduzione dei trascritti cellulari e virali (Siculella et al., 2023). Le diverse funzioni svolte dall'hnRNP A1 non si limitano alla biogenesi dell'mRNA, ma si estendono al processamento dei microRNA, al mantenimento dei telomeri e alla regolazione dell'attività dei fattori di trascrizione (Jean-Philippe et al., 2013).

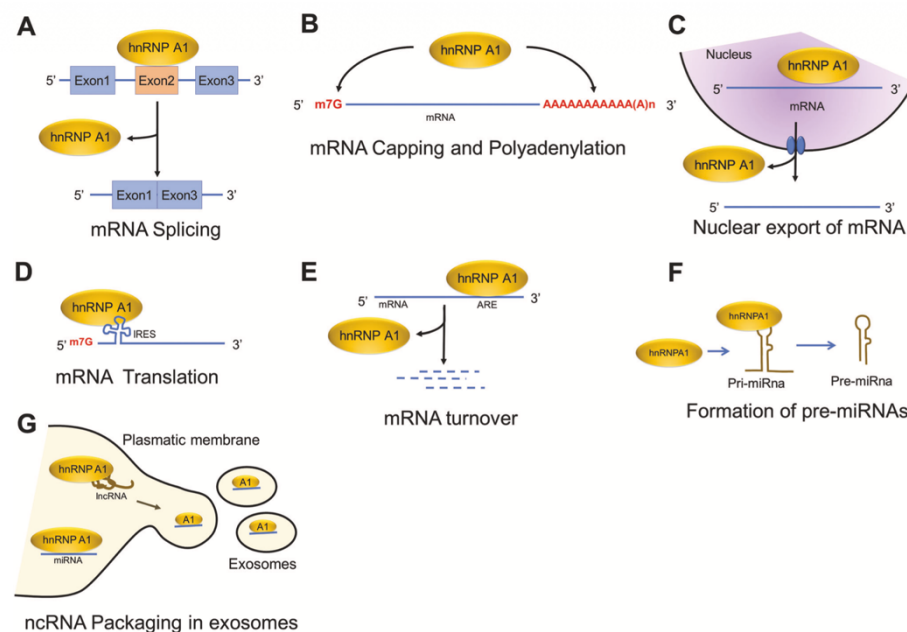


Figura 2 Ruoli di hnRNP A1 nella regolazione dell'mRNA. L'hnRNP A1 contribuisce a vari processi nella biogenesi e nella regolazione dell'RNA. Svolge un ruolo importante nello *splicing*, nel processamento, nella localizzazione nucleare, nella traduzione e nel *turnover* dell'mRNA. Inoltre, è coinvolto nella formazione dei pre-miRNA e nell'impacchettamento di ncRNA in organelli (Siculella et al., 2023).

Sebbene per hnRNP A1 siano state trovate molteplici varianti di trascritto generate per *splicing* alternativo, solo due trascritti sono stati convalidati sperimentalmente: A1-B, l'isoforma a lunghezza intera di 372 amminoacidi (38 kDa), e A1-A, la variante più corta, priva dei residui dal 253 al 303 (320 aa, 34 kDa), che, nella maggior parte dei tessuti, è oltre 20 volte più abbondante rispetto alla proteina a lunghezza intera (Jean-Philippe et al., 2013).

Dal punto di vista strutturale (Figura 3), l'hnRNP A1 può essere divisa in un dominio N-terminale, denominato frammento UP1, che contiene due domini RRM strettamente correlati, seguito da una regione C-terminale altamente flessibile e ricca di glicina (*Gly-rich*). Quest'ultima contiene il dominio di legame all'RNA box RGG, un dominio simil-prionico (PrLD, *prion-like domain*) che, se mutato, induce l'aggregazione dell'hnRNP A1 in fibrille, e una sequenza di indirizzamento nucleare detta M9. La regione ricca di glicina è responsabile del legame cooperativo dell'hnRNP A1 al suo RNA bersaglio e, inoltre, è richiesta per le interazioni omologhe ed eterologhe tra l'hnRNP A1, altre hnRNP e proteine leganti l'RNA. La sequenza M9 è un dominio di *shuttling* nucleo-citoplasma (NS) di 38 amminoacidi, necessario e sufficiente per conferire la localizzazione nucleare tipica di molte hnRNP, ma che funge anche da segnale di esportazione nucleare, consentendo l'esportazione e l'accumulo citoplasmatico dell'hnRNP A1 in risposta a segnali specifici; in particolare, l'attività di *shuttling* nucleo-citoplasma è mediata dall'interazione diretta della sequenza M9 con due recettori di trasporto della famiglia delle carioferine- β , la Trasportina 1 e 2 (Trn1 e Trn2) (Jean-Philippe et al., 2013; Siculella et al., 2023).



Figura 3 Schema della struttura di hnRNP A1. È costituita da due RBD (RRM1 e RRM2) all'estremità N-terminale, e da un dominio ricco di glicina, che contiene il box RGG, la sequenza M9 e il dominio PrLD, all'estremità C-terminale.

L'hnRNP A1 può andare incontro a modifiche post-traduzionali che includono la metilazione, la fosforilazione e la sumoilazione. Molti residui di arginina all'interno del box RGG sono siti di metilazione, che agisce modulando il legame con gli acidi nucleici. È stato dimostrato che diversi residui di serina, sia nei domini RRM che nel dominio ricco di glicina, vengono fosforilati sia *in vitro* che *in vivo*: la proteina chinasi C (PKC) e le chinasi interagenti con le MAP-chinasi (MNK) fosforilano una serie di residui di serina e riducono la capacità dell'hnRNP A1 di legare specifiche sequenze di RNA, regolandone la localizzazione. L'hnRNP A1 include anche un sito di sumoilazione all'interno del secondo RRM e sembra essere sumoilato *in vivo*, sebbene il significato funzionale di questa modifica sia ancora sconosciuto (Jean-Philippe et al., 2013).

Le specificità di legame con RNA, DNA e proteine dell'hnRNP A1 definiscono diverse funzioni nella maturazione dell'RNA eucariotico, nel mantenimento, nell'espressione e nella stabilità del genoma. E proprio date queste sue molteplici funzioni nell'espressione genica e nel metabolismo cellulare, è stato dimostrato che l'hnRNP A1 svolge un ruolo chiave in alcune patologie umane, in particolare nello sviluppo del cancro e delle metastasi, nella neurodegenerazione, in difetti genetici e nella replicazione di agenti patogeni virali. Per quanto riguarda la neurodegenerazione, è noto che hnRNP A1 svolge diversi ruoli chiave nelle funzioni neuronali e, siccome la sua espressione causa cambiamenti nella maturazione dell'RNA, la variazione della sua concentrazione contribuisce all'insorgenza di patologie neurodegenerative, come la malattia di Alzheimer (AD), l'atrofia muscolare spinale (SMA), la sclerosi multipla (MS) e la sclerosi laterale amiotrofica (SLA) (Jean-Philippe et al., 2013).

1.1.2 Hrb87F

Hrb87F, nota anche come Hrp36, è una ribonucleoproteina nucleare eterogenea appartenente alla famiglia di tipo A/B e rappresenta un omologo in *Drosophila melanogaster* della hnRNP A1 umana. È stata identificata grazie all'omologia di sequenza con le proteine hnRNP A/B dei mammiferi; infatti, condivide molte proprietà strutturali e funzionali con le proteine hnRNP A1 di tipo A/B. È costituita da due domini di legame all'RNA, RBD1 e RBD2, all'estremità N-terminale e possiede un dominio ricco di glicina (GRD) C-terminale, contenente due box RGG e un segnale di localizzazione nucleare simile a M9 (Figura 4). L'RBD1 e l'RBD2, associati al GRD, contribuiscono alla regolazione degli RNA in diverse condizioni. L'RBD1 e il GRD sono principalmente responsabili della localizzazione della proteina all'interno del nucleo e, in generale, del legame ai trascritti di RNA nascenti durante la normale crescita, mentre la coppia RBD2-GRD diventa il fattore

regolatore dominante in condizioni di stress, come lo shock termico (Zu et al., 1998; Stefanelli et al., 2026).



Figura 4 Schema della struttura di *Hrb87F*. È costituito da due RBD (RBD1 e RBD2) all'estremità N-terminale, e da un dominio ricco di glicina (GRD), che contiene due box RGG e una sequenza simile a M9, all'estremità C-terminale.

In generale, *Hrb87F* ha molti ruoli in comune con altre hnRNP; svolge funzioni regolatorie nell'espressione genica grazie alla sua capacità di legare un'ampia gamma di trascritti di RNA, intervenendo in processi come lo *splicing* dell'RNA, l'esportazione nucleare e la regolazione della risposta allo stress. Partecipa allo *splicing* alternativo, in particolare l'RBD1 è coinvolto nell'*exon-skipping* (Zu et al., 1998). È stato studiato che la deplezione di *Hrb87F* causa non solo la disgregazione dell'organizzazione degli *omega speckles* (Hsr ω), ma anche un ritardo nello sviluppo, una riduzione della fecondità femminile e un'intolleranza allo stress nei moscerini, portando alla luce il ruolo di *Hrb87F* nella risposta allo stress. Inoltre, *Hrb87F* è coinvolta nel mantenimento dei telomeri e, in particolare, della regolazione dell'*array* di retrotrasposoni (HeT-A, TART e TAHRE, chiamati collettivamente HTT), che in *Drosophila* regola la lunghezza dei telomeri (Singh & Lakhota, 2016).

1.2 CRISPR/Cas9

Il sistema CRISPR/Cas9 (*Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated protein 9*) è stato scoperto come meccanismo di immunità adattativa utilizzato dai batteri come difesa dalle invasioni fagiche ed è stato successivamente studiato e utilizzato come tecnologia di ingegneria genetica (Hsu et al., 2014).

L'acronimo CRISPR (*clustered regularly interspaced short palindromic repeats*) si riferisce a dei segmenti di DNA costituiti da brevi sequenze ripetute, ciascuna delle quali è seguita da una porzione di DNA detta spaziatore (*spacer*). La trascrizione di queste unità ripetizione-spaziatore genera due RNA non codificanti: un RNA CRISPR (crRNA), che contiene le sequenze guida che derivano dagli spaziatori utili a indirizzare la nucleasi, e un crRNA trans-attivatore (tracrRNA), che è complementare alla sequenza ripetuta. Quando un virus tenta di invadere una cellula procariotica per la seconda volta, il sistema crRNA:tracrRNA riconosce il DNA virale come estraneo e, dopo l'appaiamento con il DNA del virus, recluta l'endonucleasi Cas9 per generare una rottura a doppio filamento (DSB, *Double-Strand Break*) nel sito di riconoscimento, portando quindi al silenziamento del gene virale (Sun et al., 2016). A livello strutturale e molecolare, questo sistema si basa sull'interazione tra due componenti, un RNA guida (sgRNA, *single-guide RNA*) e l'endonucleasi Cas9. Quest'ultima agisce come una vera e propria forbice molecolare ed è caratterizzata dalla presenza di due domini nucleasici distinti e funzionalmente indipendenti, HNH che taglia il filamento complementare di DNA che si appaia con l'RNA guida e RuvC che taglia il filamento non complementare. Il reclutamento dell'endonucleasi sul corretto sito bersaglio è compito dell'RNA

guida, che contiene una sequenza specifica di circa 20 nucleotidi, ma la sequenza bersaglio verrà riconosciuta solo se è seguita da una sequenza PAM (*protospacer-adjacent motif*) all'estremità 5' (Hsu et al., 2014).

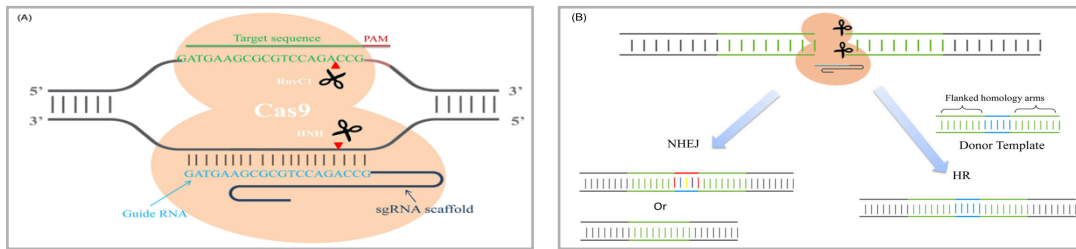


Figura 5 (A) Schema del reclutamento di Cas9 da parte di un sgRNA per il riconoscimento del sito bersaglio. Un sgRNA recluta l'endonucleasi Cas9 a tagliare il DNA in un sito specifico quando la sequenza bersaglio è seguita da una sequenza PAM. Il dominio Ruv RuvC di Cas9 taglia il filamento non complementare, mentre il dominio HNH taglia il filamento complementare del DNA. Il sito di taglio, di solito, si verifica 3 pb a monte della sequenza PAM (Sun et al., 2016). **(B) Editing genomico.** Una volta formate le rotture a doppio filamento, le cellule attivano il processo di riparazione NHEJ, inducendo piccole inserzioni o delezioni casuali nel sito di taglio. Se viene fornito uno stampo di DNA omologo, le cellule riparano il DNA tramite ricombinazione omologa (HR) (Sun et al., 2016).

Una volta reclutata sul sito bersaglio, la Cas9 si attiva e crea una rottura a doppio filamento (DSB) nel DNA d'interesse (Hsu et al., 2014), in seguito alla quale le cellule attivano i loro processi di riparazione *Non-Homologous End Joining* (NHEJ). Questo meccanismo di riparo introduce mutazioni casuali di inserzione/delezione (*indel*) delle basi nel sito di taglio, che possono portare una modifica nel *reading frame* del gene con conseguenti cambiamenti nell'espressione del gene bersaglio. Se, invece, viene fornito uno stampo di DNA omologo, le cellule riparano il danno al DNA tramite ricombinazione omologa (HR, *Homologous recombination*); perciò, il sistema CRISPR/Cas9 può essere utilizzato nell'editing genomico, anche per la produzione di modelli *knock-down*, *knock-in* e *knock-out* (Sun et al., 2016).

Una criticità dell'utilizzo del sistema CRISPR/Cas9 è la presenza di *off-target*, cioè dei tagli in un punto differente rispetto al sito bersaglio; infatti, la Cas9 è in grado di legarsi ai siti di *editing* e indurre la rottura del DNA anche in presenza di appaiamenti imperfetti (*mismatches*) tra l'sgRNA e il DNA genomico, soprattutto nelle regioni più distanti dalla PAM.

Il sistema CRISPR/Cas9 è stato ampiamente implementato in organismi modello e linee cellulari; in particolare, in *Drosophila melanogaster*, l'*editing* genomico viene ottenuto tramite iniezione embrionale di Cas9 e sgRNA, o tramite l'espressione stabile nella linea germinale di Cas9 sotto il controllo di promotori tessuto-specifici. Nel 2013, Bassett et al. hanno dimostrato l'applicazione del sistema di *editing* genico CRISPR/Cas9 alle linee cellulari di *Drosophila*, introducendolo tramite un plasmide che combinava il promotore U6, per guidare la trascrizione dell'sgRNA, e un promotore dell'actina 5C per codificare la Cas9 nelle cellule S2. I risultati hanno mostrato un'alta efficienza nell'induzione di DSB e tassi elevati di mutazione per NHEJ. Inoltre, hanno applicato il metodo mediante ricombinazione omologa, in cui uno stampo di DNA donatore con bracci di circa 1 kb ha fornito la massima efficienza di integrazione (Bassett et al., 2013).

2 MATERIALI E METODI

2.1 Linea cellulare

Linea cellulare S2R+ di *Drosophila melanogaster*, mantenuta alla temperatura di 25°C, senza CO₂, in terreno *Schneider's Drosophila Medium* (Biowest) con 10% di siero fetale bovino (FBS). Le cellule della linea S2R+ derivano dalla linea embrionale S2 (*Schneider* 2), mostrano un fenotipo appiattito e sono fortemente aderenti alle piastre di coltura. La loro caratteristica principale è l'espressione endogena dei recettori della famiglia *Frizzled* che le rende sensibili alla via di trasduzione del segnale di *Wingless* (Wg) (Yanagawa et al., 1998).

2.1.1 Trasfezione cellule

Per la trasfezione delle cellule è stato utilizzato il plasmide pAc-sgRNA-Cas9 (*Addgene ID: 49330*), che permette l'espressione della proteina Cas9 e resistenza alla puromicina prodotte come mRNA policistronico (sequenza T2A) (Figura 6). Le tre guide (sgHrb87F-147, sgHrb87F-170 e sgHrb87F-208) sono state precedentemente clonate nel vettore (Tabella 1).

Nome	Sequenza	PAM	Strand
sgHrb87F-147	CCCAAGACGAAGCGCTCTC G	CGG	Forward
sgHrb87F-170	ATTGTCGATCATGTACGACT	GGG	Reverse
sgHrb87F-208	GTGCGTCCATCGATCTTG TG	TGG	Reverse

Tabella 1 Sequenza delle tre guide.

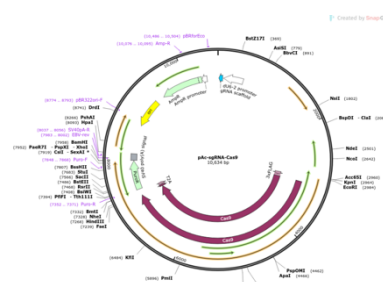


Figura 6 Plasmide pAc-sgRNA-Cas9 (*Addgene ID: 49330*).

Sono state piastrate 8×10^5 cellule per pozzetto in una piastra da 12 pozzetti. Il giorno seguente sono state effettuate tre trasfezioni indipendenti con 300 ng di ciascun plasmide utilizzando l'*Effectene Transfection Reagent* (Qiagen). Seguendo il protocollo, 300 ng di DNA sono stati portati a un volume di 75 μ L con *Buffer EC*, sono stati aggiunti 2.4 μ L di *Enhancer* e la reazione è stata lasciata in incubazione per 5 minuti a temperatura ambiente per permettere la condensazione del DNA. Aggiunti 6 μ L di *Effectene*, una seconda incubazione di 10 minuti a temperatura ambiente permette la formazione dei complessi di trasfezione. Infine, sono stati aggiunti 400 μ L di terreno ai complessi e sono stati addizionati alle cellule.

2.1.2 Selezione con puromicina

Le cellule trasfettate, dopo 3 giorni, sono state messe in selezione in terreno con 5 μ g/mL di puromicina per 7 giorni cambiando terreno ogni 2 giorni. Le cellule trasfettate e selezionate sono state raccolte in *Eppendorf* da 1.5 mL per tre estrazioni: $\frac{1}{4}$ per il DNA, $\frac{1}{2}$ per le proteine e $\frac{1}{4}$ per l'RNA. Le cellule sono state centrifugate a 16.000 xg per 10 secondi per formare il pellet cellulare utilizzato nei passaggi successivi.

2.2 Estrazione del DNA genomico

Per l'estrazione del DNA genomico (gDNA) è stato seguito il protocollo del *Wizard® Genomic DNA Purification Kit (Promega)*. Al pellet cellulare sono stati aggiunti 600 µL di *Nuclei Lysis Solution* e 3 µL di *RNase Solution* mischiando per inversione; la miscela è stata incubata in termoblocco a 37°C per 15 minuti e, successivamente, lasciata raffreddare per 5 minuti a temperatura ambiente. Aggiunti 200 µL di *Protein Precipitation Solution* e mischiato tramite vortex, il campione è stato messo in ghiaccio per 5 minuti e centrifugato a 16.000 xg per 4 minuti, così da permettere alle proteine di precipitare. Recuperato il surnatante, che a questo punto contiene il DNA, sono stati aggiunti 600 µL di isopropanolo freddo mischiando per inversione. Dopo un'incubazione di 2 ore a -20°C, il campione è stato centrifugato a 16.000 xg per 15 minuti a 4°C precipitando il DNA. Eliminato accuratamente il surnatante, il pellet è stato lavato con 600 µL di etanolo (EtOH) al 70%, centrifugando a 16.000 xg per 5 minuti a 4°C. Rimosso l'etanolo, il pellet è stato asciugato all'aria. Una volta asciugato, il pellet è stato reidratato in 11 µL di acqua *Nuclease-free*.

2.2.1 Quantifica con *NanoDrop™*

1 µL degli acidi nucleici estratti è stato utilizzato per l'analisi tramite *NanoDrop™ ND-1000 UV-Vis (Thermo Fisher Scientific)*, uno spettrofotometro in grado di quantificare e valutare la purezza di DNA e RNA.

Lo strumento restituisce un profilo di assorbimento tra 200 e 320 nm e il valore di assorbanza a 260 nm (A_{260}) viene utilizzato per determinare la concentrazione di acido nucleico tramite la Legge di *Lambert-Beer*. Altri due valori di assorbanza importanti sono a 280 nm (A_{280}) e 230 nm (A_{230}). Il rapporto A_{260}/A_{280} indica l'eventuale contaminazione da proteine, il cui valore ideale è circa 1.8-2.0; il rapporto A_{260}/A_{230} , rileva la contaminazione da sali, carboidrati o fenoli, il cui valore ideale è compreso tra 2.0 e 2.2.

2.3 Reazione a catena della polimerasi (PCR)

In seguito alla quantifica del gDNA, i campioni di DNA sono stati preparati per la PCR (*Polymerase Chain Reaction*), una tecnica introdotta da Kary Mullis nel 1985, che permette di amplificare specifici frammenti di DNA presenti nel campione.

In preparazione alla PCR, a 35.6 µL di acqua sono stati addizionati: 10 µL di *buffer (Colorless GoTaq Reaction Buffer 5X)*, 1 µL di dNTPs (10 mM), 1 µL di *primer forward* (10 µM), 1 µL di *primer reverse* (10 µM), 1 µL di DNA (precedentemente diluito a 40 ng/µL) e 0.4 µL di *GoTaq DNA Polymerase* (5 U/µL) (*Promega*).

La reazione è stata effettuata nel termociclatore *Eppendorf Mastercycler X40 (Eppendorf)*. Dopo un primo step di denaturazione a 95°C per 5 minuti, seguono tre step ripetuti per 35 volte: una prima fase di denaturazione a 95°C per 30 secondi, una seconda fase di *annealing* a 58.2°C per 30 secondi per permettere l'appaiamento dei *primer* rispettando la loro temperatura di *melting* (T_m), e una terza fase a 72°C per 1 minuto per l'allungamento da parte della Taq Polimerasi.

Infine, una ulteriore fase di allungamento di 5 minuti al termine dei 35 cicli permette la completa amplificazione dei campioni.

2.3.1 Disegno dei *primer*

Il disegno dei *primer* per la PCR è stato eseguito utilizzando il software *Primer BLAST* (NCBI), mettendo come stampo la sequenza genomica di riferimento del gene Hrb87F di *Drosophila melanogaster*. I parametri termodinamici e strutturali sono stati impostati in modo da garantire un'elevata efficienza di amplificazione e prevenire la formazione di strutture secondarie aspecifiche o di dimeri di *primer*. I criteri di selezione prevedono: una lunghezza dei *primer* intorno ai 20 nucleotidi, una temperatura di *melting* (T_m) ideale vicina ai 60°C, con una differenza di temperatura massima tra i due *primer* inferiore ai 3°C e un contenuto di GC di circa il 50%.

La coppia di *primer* selezionata (Hrb87F_ctrl_KOF3 e Hrb87F_ctrl_KOR3) è in grado di generare un amplicone atteso di 993 pb. Inoltre, un secondo *primer reverse*, interno alla sequenza amplificata dalla prima coppia, è stato selezionato per il sequenziamento Sanger (Hrb87F_ctrl_KOR) (Tabella 2).

Nome primer	Sequenza (5' → 3')	Lunghezza (bp)	T_m (°C)	Contenuto GC (%)
Hrb87F_ctrl_KOF3 (Forward)	TGAAATTCACCCGCA GATCACC	22	61.20	50.00
Hrb87F_ctrl_KOR3 (Reverse)	TAAGGAGCTGGACG GTTGTT	20	58.95	50.00
Hrb87F_ctrl_KOR	ATGTTACGCTCACG ATCTGG	21	60.74	52.38

Tabella 2 Sequenze e proprietà termodinamiche dei *primer* utilizzati per la PCR e il sequenziamento.

2.3.2 Elettroforesi su gel di agarosio

Per verificare che sia effettivamente avvenuta l'amplificazione tramite PCR è stata svolta un'elettroforesi su gel d'agarosio 1% con *MIDORI Green* (NIPPON Genetics). 0.5 g di agarosio sono stati disciolti in 50 mL di tampone TAE (Trias-Acetato-EDTA) in microonde; il composto è stato lasciato leggermente raffreddare sotto cappa chimica e sono stati aggiunti 3 µL di *MIDORI Green* (NIPPON Genetics), un colorante che ha una emissione di fluorescenza a 530 nm una volta che è legato ai gruppi fosfato del DNA ed eccitato con raggi UV. Il gel è stato fatto solidificare in un'apposita vaschetta dotata di pettine, per creare i pozzetti che sono poi stati caricati con 5 µL di prodotto di PCR, addizionato a 5 µL di acqua e 2 µL di *loading buffer* (Green GoTaq Reaction Buffer 5X). Nel primo pozzetto è stato caricato il *ladder GeneRuler 1 kb DNA Ladder* (Invitrogen), un marcatore di peso molecolare che presenta delle bande a pesi molecolari noti, con cui si possono confrontare le bande lasciate dai campioni in corsa. Il gel è stato poi visualizzato al transilluminatore UV, *Molecular Imager Gel Doc XR+ System* (Bio-Rad, US).

2.3.3 Purifica prodotti di PCR

I prodotti di PCR sono stati purificati utilizzando *QIAquick PCR Purification Kit* (Qiagen). Secondo protocollo, a 45 μL di prodotti di PCR sono stati aggiunti 225 μL di *Buffer PB* e successivamente 10 μL di sodio acetato 3M pH 5.0, per rendere il pH della soluzione meno basico. Il campione è stato trasferito nell'apposita colonnina e centrifugato per 1 minuto a 16.000 xg. Alla colonnina sono stati aggiunti 750 μL di *Buffer PE* ed è stata svolta un'ulteriore centrifuga per 1 minuto a 16.000 xg per il lavaggio. Infine, sono stati aggiunti 30 μL di acqua *Nuclease-free* per eluire il DNA dalla colonnina tramite un'ultima centrifuga di 1 minuto a 16.000 xg.

2.4 Sequenziamento Sanger

Parte dell'amplificato di PCR è stato poi preparato per essere spedito alla ditta *Eurofins Genomics* che si occupa del sequenziamento: 5 μL di DNA a 6 ng/ μL sono stati uniti a 2.5 μL di acqua e 2.5 μL di *primer* Hrb87F_ctrl_KOR (10 μM) per il sequenziamento.

2.4.1 Analisi bioinformatica TIDE

Per valutare quantitativamente l'efficienza dell'*editing* genomico e caratterizzare il profilo delle mutazioni (inserzioni o delezioni, *indels*), è stato utilizzato il *software TIDE* (*Tracking of Indels by Decomposition*) (Brinkman et al., 2014). Nel *software* sono stati caricati e confrontati i file di sequenziamento Sanger relativi al campione di controllo (*wild-type*) e ai campioni editati. Poi, è stata fornita la sequenza dell'RNA guida (sgRNA) utilizzato, con la quale l'algoritmo ha potuto confrontare i cromatogrammi quantificando le deviazioni del segnale a valle del sito di taglio della nucleasi Cas9. Il *software* ha restituito una stima della percentuale complessiva dell'efficienza di *editing* e il *pattern* delle modificazioni introdotte nel target.

2.5 Estrazione delle proteine

Il pellet cellulare è stato lisato mediante l'uso di 30 μL di *RIPA buffer* con l'aggiunta di un inibitore delle proteasi per 40 minuti in ghiaccio. Successivamente, il lisato cellulare è stato centrifugato a 16.000 xg per 10 minuti a 4°C, ed è stato prelevato il surnatante contenente le proteine. Questo è stato poi nuovamente centrifugato a 16.000 xg per 3 minuti a 4°C e poi sono stati recuperati circa 30 μL di surnatante. 1 μL è stato utilizzato per la quantifica tramite Bradford (paragrafo 2.5.1). Sono stati poi preparati 10 μL per ciascun campione con 10 μg di estratto proteico, 3.5 μL di *NuPAGE LDS Sample Buffer* (Invitrogen) e 0.7 μL di DTT 1M. Dopo aver mischiato i campioni tramite vortex, questi sono stati conservati a -80°C. Il giorno seguente i campioni sono stati denaturati in termoblocco a 70°C per 10 minuti e rimessi in ghiaccio.

2.5.1 Quantifica tramite Bradford

Per misurare l'OD dei campioni, sono stati uniti 240 μL di acqua con 60 μL di colorante Bradford e 1 μL di estratto proteico e sono stati lasciati risposare per qualche minuto. L'OD è stato misurato a 595 nm con il *BioTek Epoch Microplate Spectrophotometer* (Agilent) e tramite la retta di taratura con *Bovine Serum*

Albumin (BSA) (0 ng/ μ L, 0.1 ng/ μ L, 0.25 ng/ μ L, 0.5 ng/ μ L, 0.75 ng/ μ L, 1 ng/ μ L, 1.5 ng/ μ L e 2 ng/ μ L) è stato possibile stimare la concentrazione di proteine totali in ciascun campione.

2.6 Western Blot

Le proteine estratte sono state separate mediante elettroforesi su gel di poliacrilammide *NuPAGE 4-12% Bis-Tris Gel (Invitrogen)* nell'apposita cameretta di corsa *X Cell II Blot Module* e *X Cell Sure Lock* riempita con il *running buffer* e 500 μ L di antiossidante *NuPAGE Antioxidant (Invitrogen)*. Oltre ai campioni sono stati caricati 6 μ L di marker *Prestained Protein SHARPMASSTM VI (EuroClone)* e la corsa è stata effettuata applicando una differenza di potenziale di 100 V.

2.6.1 Blottaggio

Alla fine della corsa elettroforetica, è stato effettuato il trasferimento delle proteine su una membrana di nitrocellulosa tramite *elettroblotting*. Il *sandwich* di trasferimento è stato formato assemblando in ordine: tre spugne, un pezzo di carta, il gel, la membrana di nitrocellulosa, un altro pezzo di carta e altre tre spugne all'interno dell'*XCell II Blot Module (Invitrogen)*; il tutto è stato sempre mantenuto bagnato con il *transfer buffer*. Dopo aver montato la cameretta con il *sandwich*, la parte interna è stata riempita contemporaneamente con *transfer buffer* e l'esterno con acqua milliQ. Il trasferimento è stato eseguito impostando il voltaggio a 30 V per circa un'ora e mezza.

2.6.2 Bloccaggio

La membrana è stata estratta dal *sandwich* e colorata con rosso Ponceau per visualizzare le bande e asciugata all'aria. Una volta ritagliata, la membrana è stata reidratata con TBST e incubata 1 ora in agitazione a temperatura ambiente nella soluzione di bloccaggio (5% latte in TBST).

2.6.3 Ibridazione

Dopo aver rimosso la soluzione di bloccaggio, è stato aggiunto l'anticorpo primario, *mouse α -Hrb87F*, e la membrana è stata messa in agitazione a velocità sostenuta a 4°C *overnight*. Il giorno seguente sono stati fatti tre lavaggi di 10 minuti l'uno con TBST in agitazione a temperatura ambiente e, al termine di questi, è stato aggiunto l'anticorpo secondario, *goat α -mouse*, coniugato con l'enzima perossidasi del rafano e la membrana è stata nuovamente messa in agitazione a temperatura ambiente per un'ora e mezza. Successivamente sono stati svolti altri tre lavaggi di 10 minuti l'uno in agitazione a temperatura ambiente.

La membrana è stata visualizzata all'*UVItec Mini HD9 (Eppendorf)* con l'utilizzo di *LiteAblot® TURBO Extra Sensitive Chemiluminescent Substrate (EuroClone)* che fornisce alla perossidasi il suo substrato che, una volta ossidato, emette un segnale luminoso rilevato dallo strumento. In seguito, la membrana è stata lavata per tre volte con TBST, in agitazione per 10 minuti a temperatura ambiente e poi è stato aggiunto l'anticorpo *housekeeping, mouse α -Hsp70*, lasciato in agitazione a temperatura ambiente per 2 ore. Infine, una volta ripetuti i lavaggi e l'incubazione con l'anticorpo secondario *goat α -mouse*, la membrana è stata lavata e poi visualizzata nuovamente al macchinario.

2.6.4 Analisi densitometrica tramite *ImageJ*

L'analisi quantitativa dell'espressione proteica, a seguito della visualizzazione della membrana al macchinario, è stata effettuata utilizzando il *software* di elaborazione di immagini *ImageJ* (NIH). Le immagini acquisite tramite chemiluminescenza sono state analizzate per misurare l'intensità del segnale delle bande della proteina d'interesse (*Hrb87F*) e della proteina di riferimento (*housekeeping*, *Hsp70*). Per ogni campione il valore densitometrico di *Hrb87F* è stato normalizzato rispetto a quello del rispettivo controllo endogeno, così da poter calcolare l'espressione relativa della proteina e, infine, i dati ottenuti sono stati confrontati con il campione di controllo non trasfettato (NT), considerato uguale a 1, per valutare la riduzione dei livelli proteici nelle cellule editate.

2.7 Estrazione dell'RNA

Al pellet delle cellule trasfettate con la guida sgHrb87F-147 è stato aggiunto 1 mL di *TriplexXtractor Reagent (GRiSP)* ed è stato lasciato incubare per 5 minuti a temperatura ambiente; sono stati aggiunti 200 µL di cloroformio mescolando per inversione per 15 secondi e incubando per 3 minuti a temperatura ambiente. Il campione è stato centrifugato a 16.000 xg per 15 minuti a 4°C, permettendo la separazione di due fasi, di cui è stata prelevata la fase acquosa superiore contenente l'RNA. A questa sono stati aggiunti 500 µL di isopropanolo, mescolando per inversione, ed è stata lasciata a -20°C *overnight* per permettere l'aggregazione dell'RNA. Il giorno seguente, la miscela è stata centrifugata a 16.000 xg per 10 minuti a 4°C, per far precipitare il pellet a cui, dopo aver rimosso il surnatante, è stato aggiunto 1 mL di EtOH al 75% e poi centrifugata nuovamente a 16.000 xg per 5 minuti a 4°C. Rimosso l'etanolo, il pellet è stato asciugato all'aria e, una volta asciutto, è stato risospeso in 11 µL di acqua *Nuclease-free* e conservato a -80°C. Anche in questo caso, con 1 µL di campione è stata svolta una quantifica al *NanoDrop*TM (paragrafo 2.2.1).

2.7.1 Spettro per l'integrità dell'RNA (*TapeStation*)

La qualità dell'RNA totale estratto dai campioni è stata verificata tramite un'analisi quantitativa e qualitativa mediante il sistema *TapeStation (Agilent Technologies)*. Questo strumento permette di analizzare il profilo di migrazione dei campioni di RNA, separando le molecole in base al loro peso molecolare; in particolare, il *software* elabora i dati elettroforetici, valutando il rapporto tra le bande dell'RNA ribosomiale 28S e 18S, e restituisce un valore numerico detto RIN (*RNA Integrity Number*), che va da 1, completamente degradato, a 10, integro.

3 RISULTATI E DISCUSSIONE

In questo capitolo vengono presentati e discussi i risultati ottenuti dall'applicazione del sistema CRISPR/Cas9, finalizzato alla generazione di una linea cellulare *knock-out* per il gene *Hrb87F* in cellule S2R+ di *Drosophila melanogaster*. La strategia sperimentale ha previsto inizialmente la trasfezione delle cellule con il vettore plasmidico pAc-sgRNA-Cas9, ingegnerizzato con tre diverse sequenze guida (sgHrb87F-147, sgHrb87F-170 e sgHrb87F-208). Per ottimizzare l'efficienza del *knock-out* e selezionare la popolazione cellulare editata, i campioni sono stati sottoposti a una fase di selezione con puromicina.

L'efficacia dell'*editing* genomico è stata successivamente validata attraverso un'analisi a livello del gDNA tramite sequenziamento Sanger e scomposizione bioinformatica *TIDE*, con lo scopo di quantificare l'efficienza di taglio e determinare il profilo delle inserzioni e delezioni (*indels*) indotte da ciascun sgRNA, e a livello proteico tramite western blot, per verificare l'effettiva riduzione dei livelli di espressione della proteina target Hrb87F. Il processo di analisi ha permesso di confrontare l'azione delle tre guide e di individuare quella migliore per futuri esperimenti.

3.1 Analisi quantitativa e qualitativa del DNA genomico (gDNA)

Per poter procedere con le analisi molecolari, il DNA genomico (gDNA) estratto dai campioni è stato quantificato tramite *NanoDrop*TM.

Campione	Concentrazione (ng/ μ L)	Rapporto A260/280	Rapporto A260/230	Volume di eluizione (μ L)
147-puro	108.4	1.94	2.15	25
170-puro	43.4	1.96	2.06	25
208-puro	69.2	1.94	2.05	20

Tabella 3 Risultati della quantificazione e valutazione della purezza del gDNA estratto dai campioni.

Come mostrato nella Tabella 3, l'estrazione ha prodotto rese comprese tra 43.4 ng/ μ L e 108.4 ng/ μ L, fornendo una quantità di materiale ampiamente sufficiente per l'utilizzo del DNA come template nella successiva reazione di PCR.

Per quanto riguarda i parametri qualitativi, i campioni hanno mostrato un elevato grado di purezza; il rapporto A260/280, compreso tra 1.94 e 1.96, indica l'assenza di contaminazioni proteiche significative e, allo stesso modo, il rapporto A260/230, con valori compresi tra 2.05 e 2.15, indica l'assenza di residui organici e sali derivanti dai reagenti utilizzati durante l'estrazione.

3.2 Verifica dell'amplificazione del DNA target tramite elettroforesi su gel d'agarosio

Il gDNA estratto e quantificato è stato utilizzato come template per l'amplificazione tramite PCR della regione d'interesse e, per verificare l'effettiva amplificazione, i prodotti di PCR sono stati sottoposti a corsa elettroforetica su gel d'agarosio.

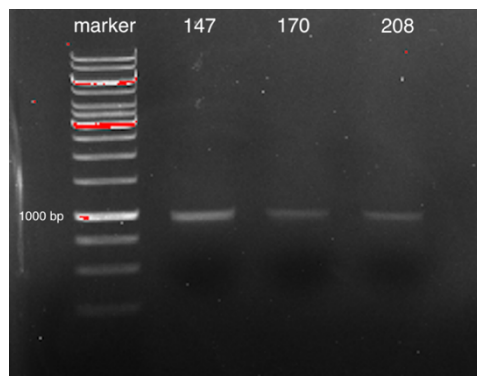


Figura 7 Elettroforesi su gel d'agarosio dei prodotti di PCR. Nel primo pozzetto a sinistra è stato caricato il marcatore di peso molecolare, seguito dai tre campioni amplificati (147, 170 e 208).

L'analisi del gel (Figura 7) evidenzia la presenza di una singola banda netta e ben definita per ciascuno dei campioni. Confrontando la distanza di migrazione dei campioni con il marcatore di peso molecolare, si osserva che la banda dei campioni è allineata con la banda di riferimento che corrisponde a 1000 bp. L'assenza di bande multiple o di segnali di fondo significativi (*smear*) conferma l'elevata specificità dei *primer* utilizzati e l'ottimizzazione delle condizioni di reazione.

3.3 Analisi quantitativa e qualitativa dei prodotti di PCR

Per rimuovere *primer* residui, dNTPs, enzimi e sali del *buffer* di reazione, i frammenti di PCR amplificati sono stati sottoposti a un processo di purificazione, necessario per evitare interferenze nelle fasi successive. I prodotti di PCR purificati sono stati quantificati tramite *NanoDrop*TM (Tabella 4).

Campione	Concentrazione (ng/ μ L)	Rapporto A260/280	Rapporto A260/230	Volume di eluizione (μ L)
147-puro	20.1	1.97	2.29	30
170-puro	23.1	1.84	1.13	30
208-puro	15.9	1.91	1.71	30

Tabella 4 Risultati della quantificazione e valutazione della purezza dei prodotti di PCR purificati.

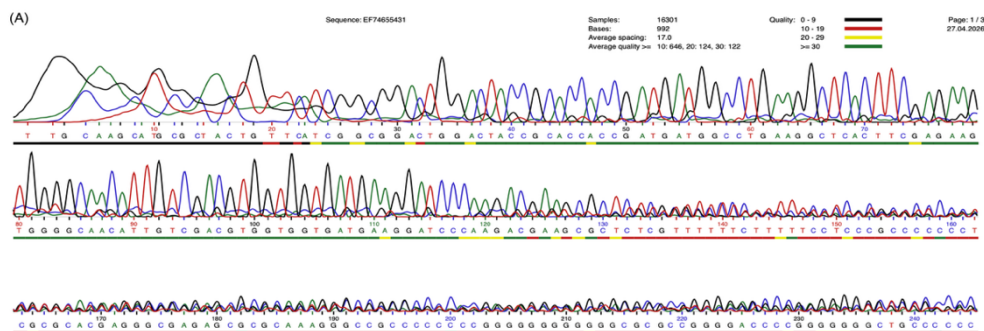
L'analisi quantitativa ha evidenziato concentrazioni comprese tra 15.9 ng/ μ L e 23.1 ng/ μ L che, considerando l'eluizione in un volume di 30 μ L, rappresentano una quantità adeguata di DNA target per procedere con le fasi successive.

Per quanto riguarda i parametri di purezza, il rapporto A260/280 ha mostrato per tutti i campioni dei valori ottimali, compresi tra 1.84 e 1.97, confermando la rimozione dell'enzima Taq polimerasi e l'assenza di contaminazioni proteiche; invece, il rapporto A260/230 mostra dei buoni valori per i campioni 147-puro e 208-puro (2.29 e 1.71), mentre il campione 170-puro presenta un valore più basso (1.13) che suggerisce la probabile presenza di una minima contaminazione, derivante probabilmente dai tamponi di lavaggio utilizzati per la purificazione. Essendo però dei quantitativi molto bassi, questi residui non compromettono la qualità e l'efficienza delle successive reazioni.

3.4 Valutazione e confronto dell'efficienza di *editing* delle sgRNA testate

Per identificare la strategia di *knock-out* più efficace per il gene Hrb87F, è stata condotta una valutazione comparativa delle tre diverse sequenze di sgRNA (sgHrb87F-147, sgHrb87F-170 e sgHrb87F-208) precedentemente trasfettate. L'efficienza di taglio e il profilo delle mutazioni (inserzioni e delezioni, *indels*) generate nella linea cellulare sono stati determinati mediante sequenziamento Sanger della regione target. I tracciati elettroforetici ottenuti sono stati successivamente analizzati tramite il software *TIDE* (*Tracking of Indels by Decomposition*).

L'analisi dei cromatogrammi ha rivelato una netta differenza qualitativa tra i campioni. I tracciati relativi alle guide sgHrb87F-147 e sgHrb87F-170 (Figura 8) mostrano il profilo tipico di una popolazione cellulare editata con successo: a valle del sito di taglio atteso per Cas9, si osserva un drastico decadimento della qualità di lettura. Questo fenomeno è una conseguenza diretta della riparazione del DNA mediata dal meccanismo *Non-Homologous End Joining* (NHEJ), che introduce mutazioni casuali variabili da cellula a cellula; la sovrapposizione dei diversi frammenti genera un segnale misto che il sequenziatore interpreta come rumore di fondo. Al contrario, il cromatogramma relativo alla guida sgHrb87F-208 (Figura 8) ha mostrato un tracciato lineare con un'elevata qualità di lettura lungo tutta la sequenza, suggerendo l'assenza di modificazioni strutturali significative nella popolazione cellulare.



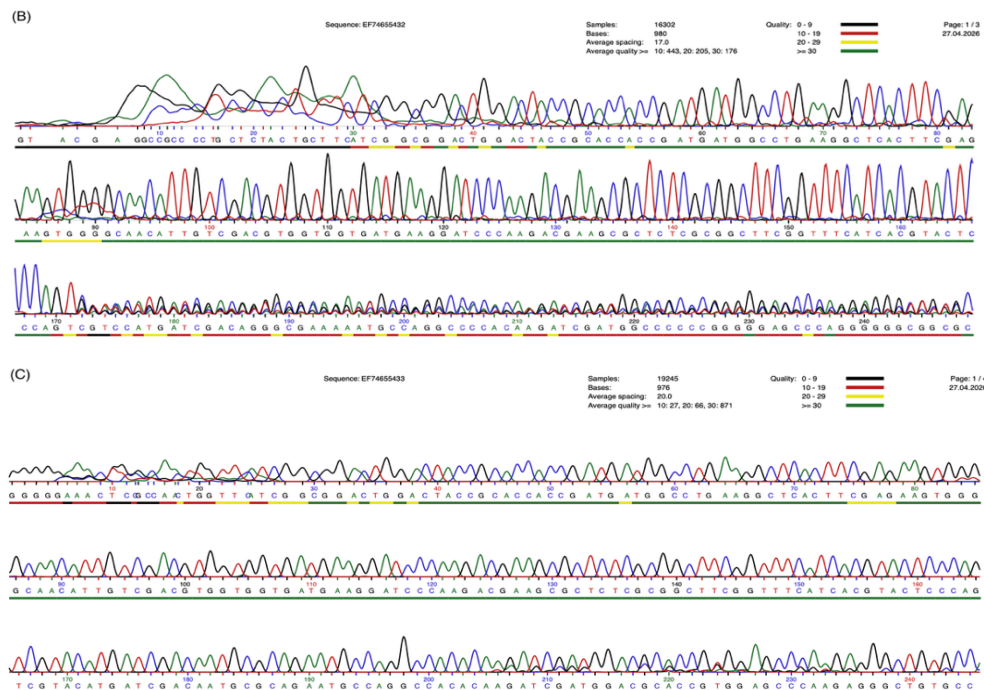


Figura 8 Cromatogramma delle regioni target. Rappresentazione dei cromatogrammi ottenuti dal sequenziamento Sanger per le tre guide. (A) sgHrb87F-147 e (B) sgHrb87F-170 mostrano il caratteristico segnale elettroforetico misto e il calo della qualità di lettura delle basi a valle del sito teorico di taglio di Cas9. (C) sgHrb87F-208 evidenzia un tracciato lineare a elevata risoluzione e stabilità qualitativa lungo l'intera sequenza analizzata, indicativo dell'assenza di *editing* significativo.

I dati di sequenziamento sono stati elaborati quantitativamente con l'algoritmo *TIDE* per mappare lo spettro delle *indels* (Figura 9). L'analisi ha confermato l'efficacia biologica delle prime due guide. Nello specifico, la sgHrb87F-147 ha mostrato l'efficienza di taglio più elevata, pari al 58.2%, inducendo prevalentemente delezioni di 2 paia di basi (-2 bp, 12.2%) e di un singolo paio di basi (-1 bp, 8.8%). Entrambe le modificazioni rappresentano mutazioni *frameshift* (*out of frame*) in grado di alterare il *reading frame* del gene, portando all'introduzione di codoni di stop prematuri e garantendo un efficace *knock-out* proteico. Risultati significativi sono stati ottenuti anche con la guida sgHrb87F-170, che ha mostrato un'efficienza totale del 57.3%. Tuttavia, l'analisi dello spettro delle *indels* ha evidenziato una criticità, poiché oltre a una delezione *frameshift* di 1 bp (20.7%), è stata rilevata una percentuale significativa di delezioni pari a 3 paia di basi (-3 bp, 14.1%). Siccome il codice genetico è organizzato in triplette, una delezione multipla di 3 bp risulta *in-frame*, comportando una semplice rimozione di un singolo amminoacido senza alterare la restante sequenza della proteina. Questo tipo di mutazione comporta il rischio che il prodotto proteico conservi parzialmente o totalmente la propria stabilità e funzionalità biologica, compromettendo l'efficacia del *knock-out*. Per quanto riguarda la guida sgHrb87F-208, invece, l'analisi *TIDE* ha stimato un'efficienza di *editing* estremamente bassa, pari al 9.6%, con l'86.2% della popolazione cellulare che ha mantenuto la sequenza *wild-type*. Per quest'ultimo campione e la guida sgHrb87F-170, i coefficienti di

determinazione stimati dal *software* sono risultati elevati ($R^2= 0.96$ e $R^2= 0.93$ rispettivamente); mentre il campione sgHrb87F-147 presenta un valore inferiore ($R^2= 0.76$) indicando una minore affidabilità della stima di *editing* ottenuta.

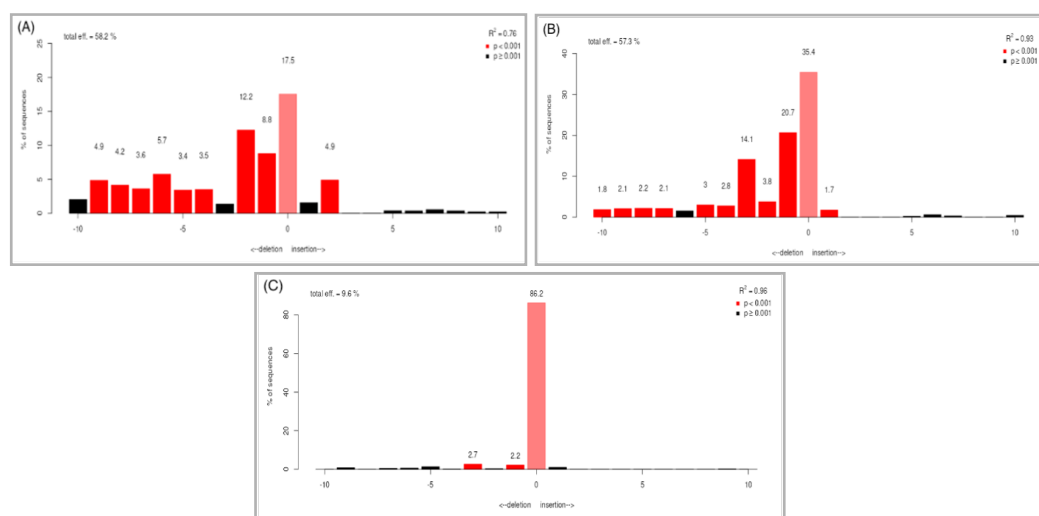


Figura 9 Istogrammi derivanti dall'analisi *TIDE* relativi alla distribuzione percentuale e alla dimensione delle inserzioni e delezioni (*indels*) nella linea cellulare trattata. (A) Profilo di *editing* della guida sgHrb87F-147 (efficienza totale: 58.2%; $R^2= 0.76$). (B) Profilo di *editing* della guida sgHrb87F-170 (efficienza totale: 57.3%; $R^2= 0.93$). (C) Profilo della guida inefficiente sgHrb87F-208, caratterizzato da una netta prevalenza di sequenze non modificate (efficienza totale: 9.6%; $R^2= 0.96$). Le bande di colore rosso indicano le mutazioni che presentano una rilevanza statistica significativa ($p < 0.001$).

RNA guida	Efficienza totale (%)	Principali <i>indels</i> rilevate (%)	Affidabilità del modello (R^2)
sgHrb87F-147	58.2%	-2 bp (12.2%); -1 bp (8.8%); -6 bp (5.7%)	0.76
sgHrb87F-170	57.3%	-1 bp (20.7%); -3 bp (14.1%); -2 bp (3.8%)	0.93
sgHrb87F-208	9.6%	-3 bp (2.7%); -1 bp (2.2%)	0.96

Tabella 5 Riassunto dell'efficienza di *editing* e delle principali *indels* identificate tramite analisi *TIDE* per le tre sgRNA testate.

In conclusione, il confronto diretto delle tre strategie di *targeting* (Tabella 5) evidenzia che sia la guida sgHrb87F-147 che la sgHrb87F-170 sono in grado di indurre l'*editing* del gene target in percentuali significative, a differenza di sgHrb87F-208. Tuttavia, la sgHrb87F-147 si configura come la migliore, poiché presenta la maggiore efficienza di taglio e un profilo di mutazioni caratterizzato prevalentemente da delezioni *out of frame* potenzialmente distruttive per la funzionalità di *Hrb87F*, essenziali per l'ottenimento di un *knock-out* completo.

3.5 Analisi quantitativa delle proteine

La quantificazione per determinare la concentrazione proteica dei campioni estratti, per poter standardizzare le quantità da utilizzare per la successiva analisi

di western blot, è stata effettuata tramite saggio colorimetrico di Bradford. Per risalire alla concentrazione dei campioni è stata costruita una retta di taratura utilizzando concentrazioni note di *Bovine Serum Albumin* (BSA), comprese tra 0.1 e 2 $\mu\text{g}/\mu\text{L}$. L'analisi di regressione lineare ha restituito l'equazione $y = 0.3483x + 0.0215$ (Figura 10) e il relativo coefficiente di determinazione ($R^2 = 0.9876$) indica un'elevata linearità della risposta colorimetrica nel range di concentrazioni analizzato e garantisce l'affidabilità dell'interpolazione.

I valori di assorbanza letti per i campioni analizzati sono stati interpolati sull'equazione della retta per ricavare le concentrazioni proteiche ($\mu\text{g}/\mu\text{L}$); i risultati della quantifica, riassunti nella Tabella 6, indicano che le concentrazioni proteiche ottenute sono adeguate a procedere alle fasi successive.

Campione	Assorbanza (OD-bianco)	Concentrazione ($\mu\text{g}/\mu\text{L}$)
147	0.756	2.11
170	0.878	2.46
208	0.822	2.30
NT	0.850	2.38

Tabella 6 Concentrazione proteica dei campioni calcolata tramite interpolazione sulla retta di taratura BSA.

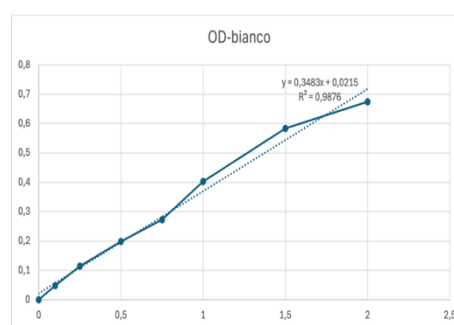


Figura 10 Retta di taratura per il saggio Bradford ottenuta con standard di BSA. L'equazione della retta e il valore di R^2 evidenziano una buona linearità del saggio.

3.6 Verifica del trasferimento proteico tramite colorazione con Rosso Ponceau

Dopo la quantifica, i campioni proteici sono stati separati mediante elettroforesi su gel di poliacrilammide (SDS-PAGE) e trasferiti su una membrana di nitrocellulosa per procedere con l'analisi di western blot. Per accertare l'effettivo trasferimento delle proteine dal gel alla membrana, questa è stata sottoposta a una colorazione totale utilizzando il Rosso Ponceau.

Come si nota dalla Figura 11 (A), la colorazione ha rivelato un pattern a bande multiple e definite in tutti i pozzetti caricati con il campione, confermando l'efficacia della corsa elettroforetica e del trasferimento delle proteine sulla membrana. Inoltre, il trasferimento del marcatore di peso molecolare, che si può notare dalle bande colorate, in particolare quella rossa a circa 75 kDa e quella verde a circa 25 kDa, fornisce il riferimento per poter stimare il peso molecolare delle proteine d'interesse.

3.7 Valutazione dell'espressione proteica mediante Western Blot

Per verificare se l'*editing* genomico riscontrato a livello del DNA si sia tradotto in un'effettiva diminuzione del prodotto proteico, è stata condotta un'analisi dell'espressione di *Hrb87F* mediante western blot (Figura 11). Gli estratti proteici totali, ricavati dalle cellule trasfettate con le tre sgRNA (sgHrb87F-147, sgHrb87F-170 e sgHrb87F-208) sono stati analizzati in parallelo con un controllo di cellule non trasfettate (NT). La proteina Hsp70 è stata utilizzata come controllo interno (*housekeeping*) per consentire la normalizzazione del segnale e confermare il caricamento omogeneo dei campioni.

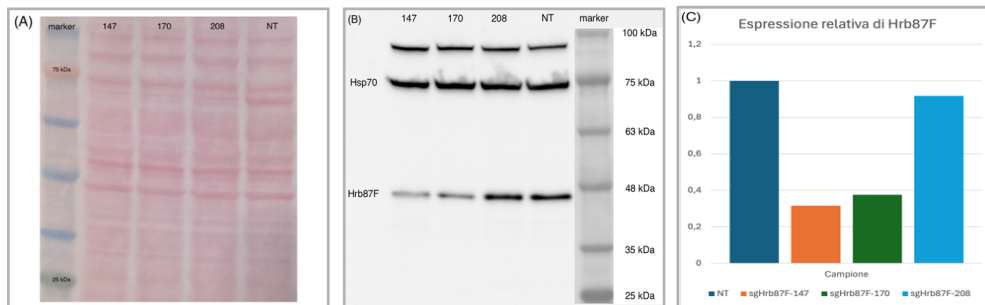


Figura 11 (A) Membrana colorata con Rosso Ponceau post-trasferimento. Nel pozzetto corrispondente al *marker* sono ben distinguibili le bande di riferimento, tra cui la banda rossa a circa 75 kDa e la banda verde a circa 25 kDa. **(B) Analisi dell'espressione proteica di Hrb87F mediante Western Blot.** Immunoblottaggio su estratti proteici totali da cellule S2R+ trasfettate con le tre sgRNA e da cellule di controllo non trasfettate (NT). I segnali relativi alla proteina Hsp70 confermano l'omogeneità del caricamento, mentre le bande inferiori mostrano l'espressione di *Hrb87F* a seguito dell'*editing* genomico. **(C) Livelli relativi di Hrb87F.** Istogramma dei valori di espressione della proteina Hrb87F nei diversi campioni, normalizzati rispetto al controllo interno Hsp70. I valori sono stati calcolati mediante l'analisi dell'intensità delle bande con il software *ImageJ*, assumendo il controllo non trasfettato (NT) come valore di riferimento uguale a 1.

L'analisi densitometrica delle bande, eseguita tramite il software *ImageJ* (NIH) e rielaborata graficamente (Figura 11), ha mostrato una riduzione dei livelli della proteina target nei campioni trattati con le prime due guide. Nello specifico, le cellule trasfettate con la guida sgHrb87F-147 hanno mostrato il decremento più significativo, registrando un livello di espressione relativa pari a circa 0.32 rispetto al controllo NT (posto pari a 1). Un risultato simile è stato osservato per il campione relativo alla guida sgHrb87F-170, in cui l'espressione di *Hrb87F* si è ridotta a un valore di circa 0.38. Al contrario, il campione trattato con la guida sgHrb87F-208 ha mantenuto livelli di espressione proteica quasi inalterati, pari a 0.92, paragonabili al controllo non trasfettato.

I risultati ottenuti confermano i dati discussi in precedenza a livello genomico. La maggiore riduzione di espressione mostrata da sgHrb87F-147 è coerente con la efficienza di taglio (58.2%) e con un profilo mutageno caratterizzato prevalentemente da delezioni *out of frame*. La frazione di proteina residua rilevata nel campione sgHrb87F-170 (0.38), leggermente superiore rispetto alla guida 147, può essere spiegata dalla percentuale di mutazioni *in-frame* (-3 bp, 14.1%) identificate dall'analisi *TIDE*. Tali delezioni permettono la sintesi di varianti

proteiche prive di un amminoacido ma strutturalmente stabili, che continuano a essere riconosciute dall'anticorpo primario *mouse* α -Hrb87F. Infine, l'elevata intensità della banda proteica relativa alla sgHrb87F-208 conferma l'inefficacia di questa guida nel generare un *knock-out*.

3.8 Analisi quantitativa e qualitativa dell'RNA

Per poter valutare a livello trascrizionale l'impatto delle mutazioni introdotte dal sistema CRISPR/Cas9, trovate durante l'analisi del DNA genomico, è stato estratto l'RNA totale dalle cellule S2R+. A seguito dell'estrazione, la quantità e la qualità dell'RNA ottenuto sono state valutate tramite analisi elettroforetica *TapeStation* (Agilent Technologies). L'analisi quantitativa ha confermato una buona resa, con una concentrazione totale di 41.7 ng/ μ L. Dal punto di vista qualitativo, l'indice di integrità (RIN) ha registrato il valore massimo di 10.0, indicando una quasi completa integrità della molecola; tuttavia, analizzando il profilo elettroforetico è stata riscontrata una degradazione a carico del frammento ribosomiale superiore; infatti, il sistema non ha rilevato il tipico picco della subunità 28S e di conseguenza il rapporto 28S/18S. Il picco relativo all'rRNA 18S risulta definito e l'acido nucleico compreso nella regione di interesse principale (tra 200 e 6000 nucleotidi) rappresenta circa il 90% dell'intero campione analizzato (Figura 12).

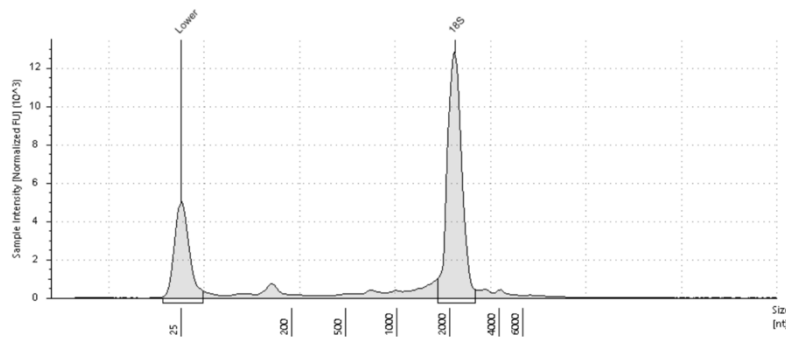


Figura 12 Profilo elettroforetico (*TapeStation*) dell'RNA estratto. Il tracciato illustra l'andamento del campione: sono visibili i segnali del *Lower Marker* (25 nt) e della subunità ribosomiale 18S (2144 nt). L'assenza del picco relativo all'rRNA 28S è attribuibile ad una degradazione del frammento ribosomiale superiore. L'analisi ha misurato una concentrazione totale di 41.7 ng/ μ L e un parametro di integrità (RIN) pari a 10.0.

3.9 Conclusioni

L'insieme dei risultati descritti conferma l'efficacia della tecnologia CRISPR/Cas9 nell'ingegnerizzazione delle cellule S2R+ al fine di generare una linea *knock-out* per il gene *Hrb87F*. L'analisi dei dati prodotti dal sequenziamento genomico e dell'espressione proteica ha permesso di valutare le tre sequenze guida utilizzate; in particolare, la sgHrb87F-147 si è dimostrata la strategia migliore, poiché garantisce il maggior tasso di mutazioni *frameshift* e la conseguente perdita della sintesi proteica. L'effettivo funzionamento di questa linea cellulare editata pone le basi per le successive fasi del progetto di ricerca; il modello cellulare ottenuto potrà essere utilizzato per visualizzare e analizzare gli effetti dell'assenza di proteina Hrb87F sul trascrittoma, permettendo, in futuro, l'identificazione della sua azione sui suoi target biologici in *Drosophila melanogaster*.

Bibliografia

1. Andrew R. Bassett, Charlotte Tibbit, Chris P. Ponting, Ji-Long Liu; Mutagenesis and homologous recombination in *Drosophila* cell lines using CRISPR/Cas9. *Biol Open* 15 January 2014; 3 (1): 42–49. doi: <https://doi.org/10.1242/bio.20137120>
2. Eva K. Brinkman, Tao Chen, Mario Amendola, Bas van Steensel, Easy quantitative assessment of genome editing by sequence trace decomposition, *Nucleic Acids Research*, Volume 42, Issue 22, 16 December 2014, Page e168, <https://doi.org/10.1093/nar/gku936>
3. Geuens T, Bouhy D, Timmerman V. The hnRNP family: insights into their role in health and disease. *Hum Genet.* 2016 Aug;135(8):851-67. doi: 10.1007/s00439-016-1683-5. Epub 2016 May 23. PMID: 27215579; PMCID: PMC4947485.
4. Hsu P, Lander E, Zhang F. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell*, 157, 1262-1278
5. Jean-Philippe J, Paz S, Caputi M. hnRNP A1: the Swiss army knife of gene expression. *Int J Mol Sci.* 2013 Sep 16;14(9):18999-9024. doi: 10.3390/ijms140918999. PMID: 24065100; PMCID: PMC3794818.
6. Siculella, L., Giannotti, L., Di Chiara Stanca, B. *et al.* A comprehensive understanding of hnRNP A1 role in cancer: new perspectives on binding with noncoding RNA. *Cancer Gene Ther* 30, 394–403 (2023). <https://doi.org/10.1038/s41417-022-00571-1>
7. Singh AK, Lakhotia SC. The hnRNP A1 homolog Hrb87F/Hrp36 is important for telomere maintenance in *Drosophila melanogaster*. *Chromosoma.* 2016 Jun;125(3):373-88. doi: 10.1007/s00412-015-0540-y. Epub 2015 Sep 16. PMID: 26373285
8. Stefanelli C, Colaianni D, De Pittà C and Mazzotta GM (2026) From development to disease: the neuronal role of hnRNPs through the lens of *Drosophila*. *Front. Cell Dev. Biol.* 14:1819111. doi: 10.3389/fcell.2026.1819111
9. Sun L, Lutz BM, Tao YX. The CRISPR/Cas9 system for gene editing and its potential application in pain research. *Transl Perioper Pain Med.* 2016;1(3):22-33. PMID: 27500183; PMCID: PMC4971521.
10. Yanagawa S, Lee JS, Ishimoto A. Identification and characterization of a novel line of *Drosophila* Schneider S2 cells that respond to wingless signaling. *J Biol Chem.* 1998 Nov 27;273(48):32353-9. doi: 10.1074/jbc.273.48.32353. PMID: 9822716.
11. Zu K, Sikes ML, Beyer AL. Separable roles in vivo for the two RNA binding domains of *Drosophila* A1-hnRNP homolog. *RNA.* 1998 Dec;4(12):1585-98. doi: 10.1017/s135583829898102x. PMID: 9848655; PMCID: PMC1369727.