

# PRINCIPI DI BIOLOGIA MOLECOLARE

- **Assemblaggio nucleosomi**

Nell'assemblaggio del nucleosoma si nota che il tetramero H3-H4 è il primo a legarsi al DNA nella parte più alta del nucleosoma stesso, successivamente il dimerico H2A-H2B si lega dalla parte inferiore e dalla parte superiore. L'ottamero così formato costituisce il core del nucleosoma attorno al quale è avvolto il DNA. Una volta assemblati i nucleosomi, si possono ottenere due forme di cromatina a seconda della compattatura: l'eucromatina e l'eterocromatina.

L'eterocromatina è una struttura altamente compattata che non permette la trascrizione mentre l'eucromatina, risultando rilassata, permette il processo di trascrizione: la differenza risiede nella presenza dell'istone H1 nell'eterocromatina che, legandosi al linker, permette una maggiore compattazione della struttura, definita come fibra a 30nm. La fibra a 30nm esiste in due modelli, il solenoide e lo zig-zag: il primo molto più compatto mentre il secondo molto più presente in natura.

- **Ciclo cellulare**

Il ciclo cellulare è composto di due grandi fasi: interfase e mitosi. L'interfase è costituita dalla fase G1 dove la cellula produce gli organelli e le proteine necessari per l'accrescimento della massa cellulare e prepara i complessi enzimatici necessari per la fase successiva. Nella fase S avviene la duplicazione del materiale genetico. A livello delle origini di replicazione, i due filamenti nucleotidici che costituiscono il DNA vengono separati grazie all'azione di un complesso multiproteico chiamato pre-Loading Complex. Nella fase G2 la cellula continua l'accrescimento della massa ed inizia a prepararsi alla successiva mitosi allestendo altre strutture indispensabili per la divisione, successivamente si ha la mitosi, anch'essa divisa in fasi.

- **Complesso del mediatore**

Il complesso del mediatore è un insieme di proteine che prendono contatto con i fattori di trascrizione e la polimerasi, oltre che con le

proteine attivatrici leganti sequenze di DNA regolatorie: l'insieme dei contatti va a regolare la trascrizione. Poiché il promotore è l'innesco della trascrizione e le sequenze attivatrici o inibenti possono essere lontane dallo stesso, esse per trasmettere il comando alla polimerasi dovranno sfruttare il complesso del mediatore.

### o **DNA polimerasi**

La DNA polimerasi è un enzima con direzione di attività  $5' \rightarrow 3'$  e la sua struttura è paragonata ad una mano con un pugno semichiuso dove vengono identificate tre regioni: palmo, pollice e dita. Nella parte alta del palmo si trova il sito catalitico dell'enzima, la catalisi, mentre le dita favoriscono la reazione ed il pollice mantiene correttamente la polimerasi sul DNA: mentre dita e pollice sono costituite da  $\alpha$ -eliche, il palmo ha una conformazione a  $\beta$ -foglietto. Nel palmo si ha la formazione del legame fosfodiesterico tra i due nucleotidi ed è di fondamentale importanza la presenza di due ioni bivalenti quali Mg o Zn: mentre il primo si coordina con l'ossigeno in  $3'$  aumentandone l'affinità per il fosfato, il secondo si coordina con i fosfati permettendo il rilascio del pirofosfato. Oltre all'azione diretta dei due ioni, la polimerasi riesce a formare legami a idrogeno con il solco minore del DNA stabilizzando il complesso. Le dita si muovono aprendosi e chiudendosi durante il processo di replicazione ed hanno un ruolo fondamentale nello stabilizzare la catalisi: quando non c'è sintesi le dita sono aperte, quando invece il DNA si posiziona sul sito catalitico le dita si chiudono e fanno una rotazione di  $40^\circ$  che determina il loro posizionamento sopra il sito catalitico. L' $\alpha$ -elica contiene residui amminoacidici (tirosina, lisina e arginina) che si vanno a posizionare e creare legami con i gruppi fosfato del dNTP entrante stabilizzandolo e favorendo la catalisi; una volta effettuato il legame al filamento, le dita si riaprono e il DNA scorre passando al nucleotide successivo. Le dita, aprendosi e chiudendosi, permettono una curvatura del DNA a  $90^\circ$  consentendo un corretto movimento del DNA sulla polimerasi e di esporre un solo nucleotide per volta impedendo il salto di alcuni nucleotidi. Il pollice si occupa

principalmente di mantenere nella corretta posizione l'innesco ed il sito attivo, stabilizzando inoltre il complesso innesco e DNA polimerasi. La DNA polimerasi presenta anche attività esonucleasica che le permette di riparare eventuali errori di compilazione. Infatti, se si inserisce un nucleotide errato, l'attività della polimerasi rallenta per la sua diminuita affinità verso il DNA, il DNA scorrerà sul sito esonucleasico che taglia e rimuove il nucleotide errato: a questo punto il filamento ritorna sul sito catalitico e l'attività può riprendere. Esistono tre tipologie di polimerasi: le polimerasi  $\alpha$ ,  $\delta$  e  $\epsilon$ . Poiché la polimerasi  $\alpha$  ha una bassa processività diminuisce la sua affinità verso il DNA e viene sostituita dalla polimerasi  $\epsilon$  o  $\delta$ : l'alta processività di queste ultime è dovuta anche all'aggancio dello sliding-clamp sul filamento.

- **Elicasi**

L'elicasi è l'enzima coinvolto nel processo di apertura della forca replicativa. Sfrutta l'energia dell'idrolisi dell'ATP per separare la forca replicativa andando a rompere i legami a idrogeno tra le basi.

- **Esoni**

Gli esoni sono le sequenze legate insieme dopo l'escissione nello splicing dell'RNA e sono solitamente le sequenze espresse dell'RNA.

- **Frammenti di Okazaki**

Poiché la direzione della forcella di replicazione e della DNA polimerasi è  $5' \rightarrow 3'$ , appare evidente che il filamento con direzione  $3' \rightarrow 5'$  sarà copiato con facilità e celermente (filamento veloce o leading strand) mentre per l'altro filamento il processo risulterà più complesso. Il filamento  $5' \rightarrow 3'$  (filamento ritardato o lagging strand) sarà sintetizzato a partire da piccole sequenze di DNA dette frammenti di Okazaki successivamente unite tra loro da un ulteriore enzima.

- **Geni**

- **Introni**

Un introne è la sequenza che resta fisicamente separata dopo l'escissione nello splicing dell'RNA: può codificare solo per snoRNA o

miRNA. Gli introni hanno numerose funzioni tra cui la regolazione dello splicing alternativo, determinazione dell'aumento della diversità genica, possono contenere enhancer che promuovono la velocità di polimerizzazione della RNA polimerasi oltre che contenere esoni di altri geni (in caso di geni sovrapposti). Una sequenza intronica presenta generalmente due nucleotidi GT in 5' e due nucleotidi AG in 3' e, ad una determinata distanza dall'estremità 3' anche 15pb di polipirimidine oltre che la presenza di una branch site.

#### ○ **Istoni**

Gli istoni costituiscono il core interno del nucleosoma. Sono proteine basiche (perlopiù arginina e lisina) che conferiscono stabilità al nucleosoma e sono suddivisibili in cinque differenti tipi: H2A, H2B, H3, H4 che costituiscono il vero e proprio core istonico e H1 che si collega alla regione linker. H2A, H2B, H3 e H4 sebbene differenti per peso molecolare, mantengono caratteristiche comuni tra cui una coda N-terminale, una regione carbossi-terminale (variabile in lunghezza) ed il dominio histone-fold, ovvero la porzione che forma il core istonico (3 domini a  $\alpha$ -elica conservati in tutti e quattro gli istoni). Gli istoni H2A e H2B si associano a dare dimeri mentre una coppia di istoni H3 ed una coppia di istoni H4 si associano a dare un tetramero.

#### ○ **Maturazione dell'mRNA**

Le fasi di maturazione dell'mRNA prevedono diverse reazioni chimiche. La prima, la fase del capping, avviene durante l'allungamento del filamento: prevede che al filamento di mRNA in trascrizione venga aggiunto una metilguanina in posizione 5' al filamento (trifosfatasi defosforila il fosfato in  $\gamma$  in 5', guaniltransferasi lega una guanina al fosfato in  $\beta$ , metiltransferasi metila la guanina in 7). Ulteriore fase di maturazione dell'mRNA è costituita dalla poliadenilazione in 3' della catena. Le proteine CTSF e CPSF giocano ruolo forza nel processo di poliadenilazione, esse si trovano in coda alla polimerasi e si legano al filamento di mRNA: CPSF taglia l'mRNA mentre CTSF è un fattore di taglio e stimolatorio. Una volta avvenuto

il taglio, viene richiamata una PoliA polimerasi che sintetizza numerose adenine e le lega alla coda 3' del filamento. Teoricamente potrebbero essere inserite infinite adenine, ma un meccanismo a feedback ne permette un massimo di 200 inserzioni.

#### ○ **Mismatch**

Una regione di mismatch è una zona del DNA che presenta un errore nell'appaiamento delle basi: se la polimerasi riconosce l'errore e riesce a correggerlo si ritorna alla situazione di partenza, altrimenti l'errore verrà corretto nel secondo ciclo replicativo portando ad una mutazione. I meccanismi che portano alla riparazione del mismatch sono molteplici, dal più semplice, come l'attività di proofreading della polimerasi, a sistemi più complessi, come il complesso proteico (Muts, Mutl e Muth). La proteina Muts ha una struttura a forma di chela che scorre sul filamento di DNA, andando a chiudersi sulle regioni di mismatch (processo ATP-dipendente): il legame va a piegare il DNA modificandone la conformazione che favorisce l'ancoraggio delle altre due proteine. Esse si assemblano a monte della regione di mismatch prendendo contatto con Muts formando un complesso: Muth ha attività endonucleasica che va a tagliare il DNA. Una esonucleasi riconoscerà il taglio e degraderà i nucleotidi fino all'estremo della regione di mismatch: la DNA polimerasi prima e la ligasi dopo andranno a copiare la sequenza corretta dal filamento stampo. Poiché solo il DNA del filamento stampo è metilato, il complesso Mut riconoscerà solo quello come stampo per la correzione e non l'altro (che non risulta metilato). Poiché Muth può portare avanti la degradazione in entrambe le direzioni, posso avere degradazione sia a monte che a valle della regione di mismatch. Il sistema di riparazione del mismatch funziona anche in caso di scivolamenti sulle sequenze nucleotidiche: in caso di ansa di nucleotidi non appaiati, Muts riconosce le anse e attiva il sistema di mismatch richiamando le altre due proteine.

#### ○ **Mitosi**

La mitosi è l'ultima fase del ciclo cellulare ed è suddivisa in quattro

fasi più una. All'inizio si ha la profase dove il DNA è condensato e si perde la membrana nucleare, a questa segue la metafase dove si forma il fuso mitotico e ciascuna coppia di cromatidi si attacca agli opposti del fuso. Successivamente, nella anafase, si ha perdita di coesione e separazione dei cromatidi fratelli che vengono attratti ai poli del fuso; con la telofase si ha la diminuzione di condensazione dei cromosomi e formazione della membrana nucleare attorno ai cromosomi segregati. Con la citochinesi si ha la separazione finale delle cellule figlie.

- **Nucleosomi**

Un nucleosoma è un elemento strutturale ripetuto dei cromosomi eucariotici e permette di impacchettare il DNA di circa sei volte rispetto alla forma lineare. Un nucleosoma è costituito da un core composto da diversi tipi di istoni in cui si avvolge il DNA (circa 147pb per nucleosoma); due nucleosomi continui sono separati da un segmento di circa 20-60pb detto DNA linker.

- **Operoni (operone lac)**

Un operone è un insieme di geni strettamente correlati e regolati dal medesimo promotore. Un operone estremamente famoso è l'operone lac (o operone lattosio): questo operone codifica enzimi che permettono alla cellula batterica di metabolizzare il lattosio. In caso di assenza di lattosio (induttore), la proteina repressore inibisce la trascrizione dell'mRNA per dare le tre proteine responsabili della metabolizzazione del lattosio. In caso di presenza di lattosio invece, l'induttore si lega al repressore inibendone l'attività di impedimento alla RNA polimerasi che può a questo trascrivere i geni per gli enzimi permeasi, transacetilasi,  $\beta$ -galattosidasi: quando non c'è più l'induttore il sistema torna ad essere inibito dal repressore. L'operone lac è un sistema inducibile, di fatto il sistema rimane inattivo in assenza dell'induttore, il lattosio.

- **Pirosequenziamento**

Il pirosequenziamento o piattaforma Roche 454 è una tecnica di sequenziamento del DNA basata sulla sintesi di un filamento

complementare e sulla quantizzazione della liberazione di pirofosfato. Il metodo di base è il medesimo di quello di Sanger ma, invece di interrompere la catena, viene continuata e si valuta la quantità di energia emessa dalla liberazione di pirofosfato: si inserisce un nucleotide alla volta e se si ha legame sul complementare si osserva l'emissione di luce e si quantizza. È un processo estremamente più rapido rispetto agli altri ma è economicamente svantaggioso.

- **Primasi**

La primasi è un enzima (una RNA polimerasi) coinvolto nel processo di replicazione del DNA. Affinché la DNA polimerasi possa iniziare la sequenza di copiatura del DNA, essa deve attaccarsi ad un innesco ibrido costituito da RNA. La primasi sintetizza il primer usando come stampo ssDNA; richiede l'assemblaggio con elicasi ed agisce sulle sequenze contenenti il trimero G-T-A: se l'elicasi apre l'elica, la primasi potrà sintetizzare il primer. Una volta che l'opera di copiatura è completa, occorre rimuovere il primer costituito da RNA e sostituirlo con una sequenza a DNA: l'RNasi H taglia i ribonucleotidi adiacenti, un'esonucleasi rimuoverà l'ultimo nucleotide del primer e la polimerasi provvederà a sintetizzare i nucleotidi mancanti e la ligasi salderà le due estremità.

- **Promotori**

Un promotore è una regione di DNA costituita da specifiche sequenze dette consenso, alla quale si lega la RNA polimerasi per iniziare la trascrizione di un gene, o di più geni (operone).

- **Proprietà del DNA: idrolisi, denaturazione e rinaturazione, curva di melting, effetto ipercromico**

Sia il DNA che l'RNA possono andare incontro a fenomeni di idrolisi dovuti a digestione enzimatica o variazioni di pH con conseguenza della rottura dei legami fosfodiesterici. La denaturazione coinvolge solamente il DNA e permette la separazione delle due catene per rottura dei legami a idrogeno e la conservazione di quelli fosfodiesterici. La rinaturazione è il processo inverso della

denaturazione: mentre si può denaturare il DNA a 90°, il raffreddamento permette la riassociazione delle catene. Mentre il DNA nativo (doppia elica) ha un determinato picco di assorbimento a 260nm, si nota che il DNA denaturato ha un picco molto più alto alla stessa lunghezza d'onda, il cosiddetto effetto ipercromico: poiché le basi non sono più impegnate a formare legami a idrogeno sono più libere di assorbire la luce. Mettendo in relazione l'assorbanza a 260nm in funzione della temperatura si ottiene un grafico dove il punto di flesso indica il melting point, la temperatura in cui il DNA è denaturato al 50% (il grafico ricorda una S estremamente allargata e schiacciata con un punto di flesso); variazioni nella composizione in termini di basi possono spostare la curva a destra (se presenti più coppie G-C) o a sinistra (se presenti più coppie A-T).

- **Proteine chaperone**

Le proteine chaperone o chaperonine sono strutture proteiche che permettono il ripiegamento di determinate sezioni di RNA solo nella conformazione attiva: l'RNA può assumere tutte le conformazioni energeticamente favorite, ma la chaperonina renderà stabile e duratura solo la conformazione attiva.

- **Pseudogeni**

Uno pseudogene è un frammento di DNA molto simile ad un gene normale ma che è stato leggermente alterato e quindi non risulta più funzionale.

- **Real time PCR (meccanismo e metodi di rilevazione dell'amplicone)**

La tecnica di analisi PCR permette di ottenere grandi quantità di un determinato frammento genico grazie a cicli successivi di duplicazione. Si parla di tre fasi distinte caratterizzate da tre differenti temperature: denaturazione dell'elica a doppio filamento per ottenere un filamento singolo (92-95°C), annealing (3-5°C sotto il punto di melting) per allineare i vari primer alle regioni dello stampo complementari, allungamento (68-72°C) si permette alla polimerasi termostabile di sintetizzare il filamento complementare allo stampo partendo dall'innesco oligonucleotidico. Una volta terminato il primo

ciclo, agendo sulla temperatura, si riavvia un nuovo ciclo: si può ottenere un numero massimo di filamenti (con una resa di reazione del 100%) di  $2^n$  dopo  $n$  cicli di PCR. In realtà dopo un certo numero di cicli si raggiunge dapprima la fase lineare e poi il plateau della reazione, cioè non si ottengono ulteriori filamenti. Occorre mantenere una soluzione equimolare (20-200  $\mu$ M per non sequestrare Mg) di deossinucleotidi in soluzione tampone di  $MgCl_2$ , essenziale per il funzionamento dell'enzima (generalmente la Taq polimerasi). In soluzione si aggiungono poi anche i primer, specificamente strutturati per una determinata amplificazione. La rtPCR permette di misurare direttamente l'amplificazione durante la fase esponenziale della stessa, dato che in questa fase le variabili di reazione possono influire minimamente sull'efficienza permettono di ottenere risultati estremamente più accurati rispetto alla tradizionale "end point". In questo caso si utilizzano probe fluorescenti che permettono l'analisi diretta via computer e la costruzione di un grafico detto curva di amplificazione.

#### ○ **Replicazione DNA**

La replicazione del DNA è quel processo chimico che porta alla formazione di una copia del DNA cellulare. Il processo è detto semiconservativo poiché ogni copia del DNA sarà composta da un filamento parentale ed un filamento copia. Gran parte del processo di replicazione del DNA è affidato alla DNA polimerasi, ma numerosi altri enzimi e proteine sono coinvolti nel processo. Durante le fasi di replicazione, vengono prima di tutto rimossi gli istoni per permettere al macchinario replicativo di accedere al DNA stampo. Sul sito di origine si lega il complesso ORC che recluta le proteine Cdc6 e Cdt1 che, cooperando, richiamano e legano il complesso elicastico Mcm-2-7 per completare la formazione del complesso di prereplicazione (pre-RC) che, una volta attivato, permetterà la singola replicazione del DNA per ciclo cellulare. Al livello del duplex inizia quindi lo svolgimento e il rilassamento dei superavvolgimenti, Cdc6 e Cdt1 sono rilasciati per reclutare altri fattori: Mcm-2-7 provvede all'apertura del duplex

mentre le topoisomerasi si occupano di risolvere i superavvolgimenti positivi. La DNA pol $\alpha$ /primasi a questo punto sintetizzerà i primer ed inizierà la sintesi del filamento ritardato: la giunzione primer-stampo è riconosciuta dal caricatore della pinza scorrevole (RFC) che assembla una pinza scorrevole su questi siti (PCNA). La polimerasi è sostituita da pol  $\epsilon$  (filamento ritardato) o pol  $\delta$  (filamento leader) che possono iniziare l'allungamento dei filamenti. FEN-1, con la sua attività endonucleasica, va a degradare i primer a RNA e le interruzioni vengono riempite da una polimerasi: la ligasi provvederà ad unire i nucleotidi. A questo punto vengono nuovamente depositati gli istoni per interazione del DNA nascente con il fattore I di assemblaggio della cromatina (CAF-1) e PCNA.

- **Repressori**

Un repressore è una particolare proteina che, legandosi all'operatore, inibisce l'espressione di uno o più geni. Un repressore può legare il DNA blocca l'attacco della RNA polimerasi al promotore impedendo la trascrizione dei geni nell'RNA messaggero. Un repressore legato all'RNA invece si lega all'mRNA ed impedisce la sua traduzione in proteine.

- **Ricombinazione omologa (ruolo di Rec e Ruv)**

La ricombinazione omologa è un processo di riparazione delle rotture a doppio filamento che recupera le informazioni necessarie alla riparazione dal cromosoma fratello (cromosoma omologo): se si ha un danno al doppio filamento non si ha più il filamento stampo utile alla polimerasi per sintetizzare i nuovi nucleotidi. Questo è possibile solo se è già avvenuta la segregazione e quindi si ha disponibilità del cromosoma omologo in caso contrario si opera con il meccanismo NHEJ (Non-Homologous End-Joining, giunzione delle estremità non omologhe). Il meccanismo NHEJ porta inevitabilmente a mutazioni per spirito di sopravvivenza. La ricombinazione omologa non ha la sola funzione di riparazione genetica, ma di generale riarrangiamento tra cromosomi omologhi che si scambiano pezzi di materiale genetico. In particolare si ricordano le funzioni di mantenimento

della variabilità genetica, recupero delle zone di DNA danneggiate e regolazione dell'espressione di geni (i cromosomi possono spostare geni espressi in una zona silente dove non vengono espressi ad una zona dove invece è funzionalmente attivo) ed è suddivisa in cinque fasi distinte. ALLINEAMENTO DELLE MOLECOLE DI DNA OMOLOGHE: se due porzioni di DNA devono scambiarsi i cromosomi omologhi è necessario che abbiano un'omologia una rispetto all'altra, quindi le molecole devono allinearsi (omologia totale o almeno 100pb per lo scambio). INTRODUZIONE DI ROTTURE AL DNA: perché avvenga lo scambio devo avere DNA a singolo filamento che viene quindi tagliato e processato, andando a "mangiucchiare" una delle estremità. INVASIONE DEL FILAMENTO: una volta formato il DNA a singolo filamento esso costituirà il filamento invasore ed entrerà nel cromosoma omologo, si legherà quindi alle basi complementari andando a costituire un eteroduplex. FORMAZIONE GIUNZIONI DI HOLLIDAY: in seguito all'invasione le due molecole di DNA sono connesse tra loro da filamenti che si incrociano. Questa struttura assume il nome di giunzione di Holliday e può scorrere lungo il DNA (migrazione del chiasma). RISOLUZIONE DELLE GIUNZIONI: l'incrocio viene risolto e si ricostituiscono i cromosomi mescolati geneticamente. Nella realtà è possibile avere anche un taglio dei due filamenti del medesimo cromosoma: si otterrà quindi una doppia invasione ed un duplice chiasmo.

- **Riparazione DNA**

I meccanismi che portano alla riparazione del DNA sono di diverso tipo e prevedono meccanismi ben precisi. La riparazione diretta porta alla cancellazione del danno, si ricorda la fotoriattivazione (un danno causato da radiazione UV porta alla formazione di un dimero di timina e alla contemporanea attivazione della DNA fotoliasi che, tramite luce visibile, porta alla riparazione diretta del danno) e la demetilazione (in caso di presenza di metilguanina che può accoppiarsi con l'adenina, la metiltransferasi trasferirà il metile dalla base al residuo -SH dell'enzima ripristinando la sequenza corretta. In

caso di riparazione per escissione si rimuove il nucleotide errato ed il filamento stampo viene usato come base per la riparazione. L'enzima glicosilasi attacca il legame glicosidico tra base e zucchero e lo rimuove: si crea quindi un sito abasico. Interviene a questo punto una nucleasi che taglia il legame fosfodiesterico tra zucchero e zucchero andandolo a rimuovere: polimerasi e ligasi andranno a riparare il filamento. È possibile anche rimuovere più nucleotidi, in questo caso si utilizza un complesso proteico di proteine del gruppo Uvr che riconoscono la regione di errore sul DNA. Le proteine UvrA e UvrB si uniscono a formare un tetramero e si vanno ad assemblare attorno alla regione di distorsione grazie all'idrolisi di ATP. Una volta agganciate, le UvrA si staccano lasciando che UvrB crei una bolla di apertura della doppia elica: la bolla richiama UvrC, una nucleasi che, ponendosi a ponte sopra UvrB, va a tagliare il DNA ai due lati (8 da una parte e 4-5 dall'altra). Il nick creato richiama UvrD che rimuove il frammento ed attiva la polimerasi e la ligasi per la riparazione. In caso di trascrizione del DNA da parte della RNA polimerasi, in caso di errore nel filamento, la RNA polimerasi si blocca, recluta il complesso Uvr per la riparazione del danno e poi riprende il lavoro di trascrizione.

○ ***Risoluzione delle giunzioni di Holliday***

Per risolvere le giunzioni di Holliday occorre dapprima immaginare di ruotare il chiasma nello spazio per renderlo planare e successivamente effettuare il taglio. Se opero un taglio lungo i filamenti coinvolti nello scambio otterrò un prodotto patch (minima ricombinazione) mentre se opero un taglio lungo i filamenti non coinvolti nello scambio otterrò invece un prodotto di crossing over. Nella risoluzione delle giunzioni di Holliday, giocoforza è definito dal ruolo dei complessi proteici Rec e Ruv. Il sistema RecBCD (con funzione nucleasica ed elicastica) permette di preparare il filamento invasore: RecB e RecD degradano entrambe i filamenti, mentre RecC prende contatto con entrambe. Poiché RecB è più lenta di RecD, si formerà un'ansa. Il processo a doppia velocità continuerà fino a che

RecC non incontra e lega la sequenza  $\chi$ : il processo si blocca e si invertano le velocità, ottenendo il frammento libero voluto. A questo punto la proteina RecA lega il filamento protruso ed inizia l'invasione del filamento omologo. Il complesso proteico Ruv è invece coinvolto nella migrazione del chiasma. La proteina RuvA si lega all'altezza del chiasma prendendo contatto con tutti i filamenti: questo richiama RuvB che, ponendosi ai lati di RuvA, con un'azione ATP-dipendente, ruota e fa scorrere il chiasma sul DNA. Il chiasma prosegue la sua propagazione fino a che non incorre in sequenze specifiche che porta al distacco del complesso RuvAB e all'attacco di RuvC, un'endonucleasi che taglia il chiasma ed induce la risoluzione dello stesso.

- ***RNA polimerasi batteriche ed eucariotiche***

La RNA polimerasi è un enzima costituito da due subunità  $\alpha$ , due subunità  $\beta$  ed una  $\omega$ : solo le subunità  $\beta$  entrano in gioco nella polimerizzazione, in quanto centro catalitico dell'enzima. L'enzima ha la forma di una chela di granchio: il sito attivo è la cavità dove passa il DNA e la polimerasi sintetizza l'RNA, qui è contenuto magnesio bivalente. A differenza della DNA polimerasi non ha bisogno di un primer per poter iniziare il processo di copiature, ma riconosce specifici promotori per l'inizio della trascrizione di sequenze di DNA e sequenze terminatrici per la cessazione della trascrizione. Al contrario della DNA polimerasi, la RNA polimerasi ha anche azione elicasi, questo significa che il DNA entra a elica, viene processato ed esce nuovamente sottoforma di elica. La direzione di trascrizione è  $5' \rightarrow 3'$  per cui verrà operata la trascrizione del filamento stampo a dare una copia esatta del filamento codificante.

- ***Sequenze ripetute in tandem: satelliti, minisatelliti, microsatelliti***

Una sequenza ripetuta in tandem è una sequenza ripetuta esattamente uguale un certo numero di volte: si parla di sequenze satellite, minisatellite, microsatellite. Una sequenza satellite ha una localizzazione centromerica, è un prodotto della duplicazione di piccole sequenze e loro successiva mutazione; sono sequenze grandi

fino a 200pb con una densità differente rispetto al DNA (se centrifugo in gradiente di concentrazione otterrò una separazione in bande). Le sequenze minisatellite prevedono una grandezza massima di 5pb, le microsatellite sono invece inferiori a 4pb.

○ ***Sequenze ripetute interdisperse: trasposoni e retrotrasposoni, LTR, SINE e LINE***

Le sequenze interdisperse possono essere suddivise in trasposoni e retrotrasposoni. I trasposoni, o trasposoni a DNA, sono elementi mobili all'interno del genoma che hanno inserito nella propria sequenza un gene che codifica per la trasposasi, capace di tagliare il DNA e spostarlo da un'altra parte. I retrotrasposoni, o trasposoni a RNA, devono essere trascritti in RNA dalla RNA polimerasi e retrotrascritti dalla trascrittasi inversa in DNA copia (cDNA) che a questo punto è capace di inserirsi in un nuovo sito bersaglio. I retrotrasposoni sono suddivisi in tre tipologie: LTR, SINE e LINE). I retrotrasposoni di tipo LTR (Long Terminal Repeats) sono molto simili ai retrovirus e si tratta di sequenze terminali lunghe; come i retrovirus presenta i geni gag (codifica per il capsid) e pol (codifica per la trascrittasi inversa) ma non presenta env (codifica per le proteine di rivestimento). Le sequenze LINE (Long Interspersed Nuclear Elements) sono autonome e in grado di autotrasporsi, presentano i geni ORF1 e ORF2 ed una coda poliA terminale. Contrariamente, le sequenze SINE (Short Interspersed Nuclear Elements) non presentano i due geni e non sono quindi autonome: spetta alle LINE fungere da help alle SINE per la trasposizione.

○ ***Sequenziamento DNA***

Il sequenziamento del DNA vede nel suo apice il Progetto Genoma Umano che prevede l'analisi dell'intero genoma umano. Sequenziare il DNA significa riconoscere ed ordinare tutti i nucleotidi che compongono il patrimonio genetico così come posizionati nel genoma. Numerose sono le tecniche per ottenere questo risultato. Le fasi di sequenziamento sono comuni e comprendono: digestione del DNA in frammenti, inserzione in plasmidi e amplificazione in

batteri, sequenziamento dei frammenti, allineamento dei frammenti, sequenziamento.

- **Sequenziamento DNA – Maxam Gilbert**

A differenza del metodo di Sanger, il metodo Maxam-Gilbert prevede l'uso di composti chimici dannosi e non è automatizzabile. Questo metodo parte da DNA a doppia elica, marcato in 5' o 3' e successivamente denaturato. Il DNA viene suddiviso in quattro campioni e vengono utilizzati reagenti chimici che permettono la degradazione selettiva di ogni base ottenendo tanti frammenti quanti sono i nucleotidi. Viene effettuata una corsa elettroforetica e dalla lastra si ottiene, come per Sanger, una sequenza di basi (5' → 3', basso → alto) che rappresenta però il filamento originale e non lo stampo. La complessità di messa in opera, l'impossibilità di automatizzazione e l'interazione con sostanze chimiche dannose lo hanno portato in disuso in favore di altre tecniche di sequenziamento.

- **Sequenziamento DNA – Sanger**

Il sequenziamento secondo Sanger è la prima tecnica messa a punto per ottenere il sequenziamento del DNA. Riduco il DNA a singolo filamento e lo suddivido in quattro provette contenenti ognuna uno dei quattro dNTP, la polimerasi ed un ddNTP (e.g. provetta "A": ssDNA+dATP+polimerasi+ddATP): se la polimerasi in provetta invece di legare il dNTP lega il ddNTP impedirà il legame in 3' bloccando la catena. Si conduce a questo punto un'analisi elettroforetica posizionando ogni provetta in un diverso pozzetto: è consigliato aggiungere marcatura con il <sup>32</sup>P radioattivo, in tal modo l'elettroforesi potrà essere marcata direttamente su lastra fotografica. Leggendo la lastra dal basso verso l'alto le varie bande in corrispondenza dei diversi pozzetti, otterrò la sequenza complementare (5' → 3', basso → alto): per ottenere la sequenza originale è sufficiente disegnare il filamento complementare (3' → 5', basso → alto). In alternativa ai quattro pozzetti con marker radioattivi, si possono utilizzare marker fluorescenti in un singolo pozzetto:

l'analisi con il laser provvederà a separare le varie bande. A questo punto un sistema computerizzato, basandosi su database esistenti, proverà a sovrapporre ed ordinare i vari frammenti per ottenere la sequenza di riferimento.

### o **Splicing**

L'RNA viene trascritto come RNA immaturo che sarà utilizzabile solo dopo una serie di processi che si attivano già all'uscita della catena ribonucleotidica dalla polimerasi. Sebbene possa avvenire con più varianti, lo splicing è di fondamentale importanza perché dà variabilità genetica: dal medesimo trascritto immaturo posso ottenere tanti trascritti maturi e di conseguenza diverse proteine. Il DNA genomico (la porzione che viene codificata in mRNA) è costituito in parte da esoni, la parte codificante, e da introni, la parte non codificante: entrambi sono presenti nel pre-mRNA. Sebbene il numero di introni vari da cellula a cellula, essi vanno comunque eliminati poiché non hanno un ruolo funzionale: la reazione è detta di splicing. Esiste lo splicing catalizzato dallo spliceosoma (insieme di proteine e piccoli RNA nucleari che catalizzano la reazione; è lo splicing più tipico che si realizza sugli mRNA nucleari preposti alla codificazione di proteine) e lo splicing autocatalitico. Lo splicing autocatalitico viene distinto in classe I e classe II: lo splicing di classe II ha lo stesso meccanismo dello splicing catalizzato dallo spliceosoma ma è l'introne stesso che funge da catalizzatore. Anche lo splicing di classe I sfrutta l'introne come catalizzatore ma il meccanismo differisce leggermente. Su rRNA e RNA degli organelli si ha splicing alternativo, trans-splicing ed editing dell'RNA. Affinché si rimuovano gli introni, gli spliceosomi devono riconoscere determinate sequenze al confine tra introne ed esone: una regione in 5' con sequenze conservate GU, un sito accettore in 3' con sequenza AG, una sequenza PoliU vicina al sito di splicing in 3' ed un punto di ramificazione determinato dalla presenza di adenina. Il meccanismo di reazione è abbastanza semplice in quanto l'adenina sul punto di ramificazione lega con l'OH del ribosio in 2' il sito di

splicing in 5' andando a liberare l'esone: la struttura si richiude come un cappio tra l'estremità 5' con il punto di ramificazione. L'esone che si è liberato può a questo punto, con l'estremità 3', andare a legarsi alla prima base dell'esone successivo andando a sganciare l'introne dalla catena. Lo spliceosoma è un complesso di proteine e RNA nucleari detti snRNA indicati con U1, U2, U4, U5 ed U6: durante lo splicing si assemblano a delle proteine formando i snRNP.

- **Struttura del DNA (variazioni strutturali nella conformazione)**
- **Struttura primaria DNA (basi, zuccheri, legame fosfodiesterico)**

Il DNA è una lunga sequenza di basi organiche dette nucleotidi organizzate lungo un doppio filamento. I nucleotidi presenti nel DNA sono composti da basi azotate, gruppi fosfato e deossiribosio (presenza di -OH solo in 3'). Le basi azotate sono quattro, due purine (Adenina e Guanina) e due pirimidine (Timina e Citosina). Ogni base può interagire e legarsi con una sola base nel filamento opposto secondo la regola A-T C-G. Se terminale, lo zucchero presenta in posizione 3' il gruppo -OH e in 5' il gruppo fosfato; se in posizione interna legheranno entrambi il gruppo fosfato della base precedente e successiva. Il legame tra nucleotide e nucleotide è detto fosfodiesterico e coinvolge il fosfato in 5' e l'OH in 3'.

- **Struttura primaria RNA (basi, zuccheri, legame fosfodiesterico)**

L'RNA è una lunga sequenza di basi organiche dette nucleotidi organizzate lungo un singolo filamento continuo. I nucleotidi presenti nel RNA sono composti da basi azotate, gruppi fosfato e ribosio (presenza di -OH in 2' e 3'). Le basi azotate sono quattro, due purine (Adenina e Guanina) e due pirimidine (Uracile e Citosina). Ogni base può interagire e legarsi con una sola base nel filamento opposto secondo la regola A-U C-G. Se terminale, lo zucchero presenta in posizione 3' il gruppo -OH e in 5' il gruppo fosfato; se in posizione interna legheranno entrambi il gruppo fosfato della base precedente e successiva. Il legame tra nucleotide e nucleotide è detto fosfodiesterico.

- **Struttura secondaria DNA (conformazione A-B-Z, tripla elica, forcina, curvatura)**

Il DNA assume la caratteristica conformazione a doppia elica grazie ai legami a idrogeno che si vanno a creare tra base e base tra i filamenti opposti. In soluzione il DNA si dispone a forma di elica con le basi rivolte verso l'interno ed i gruppi fosfato all'interno: questo perché le basi sono idrofobiche. La catena è caratterizzata dalla presenza di due solchi, uno minore ed uno maggiore. Le basi sono soggette a movimenti di roll, tilt e twist lungo i tre assi x, y, z. Il DNA può assumere una delle tre conformazioni B Z o A. La conformazione B è la più comune e si trova principalmente nella cellula. La conformazione A (tozza e tarchiata) è presente perlopiù nell'RNA duplex o negli eteroduplex RNA-DNA. La conformazione Z è reperibile solo in vitro ed assume una forma a zig-zag. Esistono anche conformazioni alternative quali la tripla elica (triplex) dovuta a legami a idrogeno non canonici per presenza di una sequenza di 20-30pb di sole purine o pirimidine. La presenza di sequenze invertite (e.g. 5'-TGCGAT|ATCGCA-3') o invertite ripetute (e.g. 5'-TGCGAT|ACTC|ATCGCA-3') in un solo filamento permette la formazione di strutture a forcina (con un'ansa di sequenza non ripetuta ed uno stelo formato da sequenze ripetute), se invece la sequenza è su due filamenti si forma una struttura cruciforme (con due anse e due steli).

- **Struttura secondaria RNA (tetraloop, pseudonodo)**

A differenza del DNA, l'RNA si presenta in forma di singolo filamento per ragioni steriche e repulsione elettroniche per la presenza di -OH in 2'. Oltre alle strutture già viste nel DNA, l'RNA può presentare anche anse interne agli steli, regioni di mismatch o gemme, dove si ha lo spostamento della base verso l'esterno. Le strutture stelo-ansa possono presentare anche sequenze "tetraloop" con sequenza UUUU: questo particolare tipo di struttura (un semi ricciolo) risulta estremamente stabile per le interazioni di impilamento all'interno della struttura. Un'ansa che presenti basi esterne e complementari al

filamento tali da portare alla formazione di una struttura tridimensionale per impilamento coassiale è detto pseudonodo.

- **Struttura terziaria DNA (superavvolgimento)**

Il DNA tende ad assumere una conformazione estremamente compatta ed avvolta in una struttura detta cromatina. Il superavvolgimento può essere paragonato agli arrotolamenti del filo del telefono, già avvolto. Il DNA assume questa conformazione per compensare le tensioni indotte da modifiche strutturali quali duplicazione o replicazione. Un superavvolgimento è definito dal numero di legame (Lk), definito dalla somma dei twist (avvolgimenti) e dei writhe (arrotolamenti):  $Lk = Tw + Wr$ . Se la differenza tra Lk superavvolto e rilassato è minore di 0, allora si ha un superavvolgimento negativo, viceversa positivo. Se effettuo una elettroforesi su gel d'agarosio si vede che più il DNA è superavvolto e più velocemente si muove (→ è più distante dai pozzetti di deposito). Endonucleasi come le topoisomerasi possono mutare il numero di legame.

- **Telomerasi**

La telomerasi è un complesso ribonucleoproteico che copia la parte di DNA terminale detta telomero: il cromosoma termina con estremità a singolo filamento con sequenze ripetute (TTAGGG) dette telomeri. La telomerasi riconosce per complementarità le sequenze del filamento singolo e, usando sé stessa come stampo, allunga le estremità telomeriche. Il processo di allungamento è regolato dalla presenza di alcune proteine legate ai telomeri (Rif e Rap) che inibiscono il processo di retrotrascrizione. Per evitare che i telomeri vengano "riparati" e rimanere quindi nella conformazione a singolo filamento, essi formano un "cappuccio" andandosi a ripiegare all'interno della singola elica.

- **Topoisomerasi (I e II)**

Le topoisomerasi sono enzimi della famiglia delle endonucleasi che permettono di aumentare il numero di legame e rilassare il superavvolgimento di DNA. La topoisomerasi II taglia entrambi i

filamenti, srotola l'attorcigliamento e permette Lk+2 tramite la riduzione dei writhe. La topoisomerasi I invece sfrutta la tirosina per rompere il legame fosfodiesterico di un solo filamento ed aumentare Lk+1.

- **Transelesione**

La sintesi transelesione è l'ultima speranza di riparazione del danno e sopravvivenza dopo che gli altri tentativi sono falliti: invece di riparare il danno, la cellula diventa tollerante ad esso. Il processo di replicazione è condotto da particolari polimerasi di translesione che differiscono tra procarioti (pol IV e pol V) ed eucarioti (polimerasi Y). Il processo è il medesimo della replicazione: la polimerasi III copia il filamento e ad un certo punto trova un danno dovuto ad un dimero di timina. A questo punto si stacca il complesso della polimerasi richiamando la polimerasi di translesione che copia il frammento danneggiato senza troppa specificità. Una volta copiato il frammento danneggiato, la polimerasi di translesione si stacca e la polimerasi III riprende il suo lavoro di copiatura

- **Trascrizione RNA (ruolo di  $\sigma 70$  come fattore di inizio)**

L'intero processo di trascrizione da DNA a RNA può essere riassunto in tre passaggi fondamentali: inizio, allungamento e terminazione. All'inizio la polimerasi riconosce il promotore: se il legame tra promotore e polimerasi, la sequenza di trascrizione può avviarsi. La polimerasi si aggancia al filamento (complesso chiuso) ed apre il doppio filamento di DNA formando la bolla di trascrizione (complesso aperto). In questa fase la polimerasi è in fase abortiva, cioè rimane adesa al promotore e sintetizza al più nucleotidi. Successivamente si stacca dal promotore e corre sul DNA; la polimerasi qui gioca diversi importanti ruoli: catalizza la sintesi dei nucleotidi, separa e poi riavvolge il DNA, stacca l'RNA dallo stampo e corregge le bozze. Quando la polimerasi raggiunge determinate sequenze di terminazione si stacca completamente dal DNA e termina la fase di trascrizione. Affinché la trascrizione parta è necessaria la presenza di un promotore, la cui forza è derivata dalla

capacità di legare la polimerasi e di esprimere trascritti nell'unità di tempo. I promotori sono costituiti da due sequenze di 6bp nelle regioni -35 e -10 intervallate da una sequenza consenso di circa 17-19pb: promotori in possesso di una sequenza discriminatore o l'elemento 10 esteso sono comunque funzionanti. Perché un promotore possa legare la polimerasi si rende necessario l'intervento del fattore  $\sigma$ , costituito da numerose subunità: la regione 4.2 prende contatto con la regione -35 del promotore, la regione 2.4 interagisce invece con la -10, la regione 2.3 è importante per l'apertura del DNA perché ricca di amminoacidi aromatici che stabilizzano l'elica in posizione aperta, il dominio 1.1. permette l'avvio dell'allungamento dell'mRNA. Il fattore  $\sigma$  si attiva solamente quando si lega alla polimerasi, altrimenti rimane in conformazione ripiegata atta a coprire i siti di legame al promotore.  $\sigma$  1.1 si trova al punto di ingresso del DNA nella polimerasi e, finché il complesso è chiuso, blocca l'ingresso del DNA alla polimerasi; dopo l'aggancio del fattore  $\sigma$  alle regioni -35 e -10 del promotore si ha un cambio conformazione dell'enzima e  $\sigma$  1.1 ruota di circa  $50^\circ$  determinando l'apertura del canale di entrata. Esistono numerosi fattori  $\sigma$ , con operatività dipendente dalle condizioni ambientali e di stress, il fattore che è attivo in condizioni normali per la trascrizione della maggior parte dei geni è il fattore  $\sigma^{70}$ .