

The Human genome: structure and function of genes and chromosomes

Ved 5'-ende er der en **promotor region**, der er ansvarlig for ordentlig **initiering** af transskriptionen. Transskriptionen starter fra 5'-ende og kan fortsætte flere millioner af nukleotider, både gennem introner og exoner. Efter begge ender er blevet modificeret (**capping** og **polyA-tale**, der stabiliserer den primær RNA) og intronerne er splejset fra og exonerne er sat sammen, er den dannet mRNA som bliver transporteret til cytoplasmaet, hvor den bliver translateret til polypeptidkæden.

Ved **transskriptionen** bliver nukleotidsekvensen i genets ene DNA-streng kopieret til et komplementært RNA-molekyle vha. en DNA-afhængig RNA-polymerase II. Ved den første kodende sekvens initieres transskriptionen, hvor DNA'ets dobbelthelix åbnes og den DNA-streng der er orienteret i 3'-5'-retningen fungerer som template for transkriptionen. Syntesen af RNA foregår i 5'-3'-retningen.

Translationen sker i cytoplasmaet defineret af startcodon, **AUG**, der koder for aminosyren **methionin** på det færdig mRNA, dog den fjernes før syntesen af proteinet er færdigt. Den etablerer læseramme af mRNA. Ved translationen oversættes mRNA'ets basesekvens i grupper af 3 baser, som definerer polypeptidets aminosyre-sekvens. Aminosyrer føres til ribosomerne, opbygget af ribosomal RNA (kodet af 18S og 28S) vha. **tRNA**. Hver aminosyre har sit eget tRNA, opbygget af ca. 80 nukleotider, som har en såkaldt **anticodon**, der er komplementær til den codon i mRNA som tRNA'et bindes til, når det afleverer aminosyren. Da der er 4 nukleotider, en **codon** består af 3 nukleotider, kan man sige, at der kan kombineres 64 forskellige codoner, der udgør den genetiske kode. Vi ved at der kun er 20 aminosyrer, derfor kan en enkel aminosyre have flere codoner.

Selve translationsprocessen kan inddeles i 3 trin og foregår fra 5'-enden til 3'-enden: 1) **initiering**, hvor der dannes et initieringskompleks bestående af mRNA, et ribosom og tRNA sv.t. codon. 2) **elongering** er det næste trin, og består af codon-genkendelse (binding af tRNA), etablering af peptid-binding og flytning af ribosomet, vha. en **translokase**, 3 baser frem i 3'-retningen på mRNA, således der er en ny codon synlig, som tRNA med aminosyre kan genkende. Dette kræver elongeringsfaktorer sådan at de successive aminosyrer påsættes det voksende polypeptid. 3) **termineringen** er det sidste trin og translationen stopper når en af de 3 stopcodons **UAA**, **UGA** eller **UAG** nås. Det dannede polypeptid forlader ribosomet, som dissocierer til subunits (lille og store) og mRNA.

De posttranslationelle modifikationer typisk består i oxidation af cysteiner for at danne disulfidbroer samt i glykosylering, fosforylering m.m.. Mange polypeptider bliver splejset til korte polypeptider og er sat sammen på tværs, f.eks. insulin er sammensat af en 21 og 30 lange aminosyrekæder, hvor proinsulin er 80 aminosyre langt posttranslateret produkt.

Den DNA-streng der ikke fungerer som template, vil være kodende, forstået på den måde, den vil være identisk med RNA bortset fra der er U stedet for T.

Beta-globin har en promotor, som alle andre har nogle små funktionelle elementer, der interaktioner med transskriptions faktorer, der regulerer transskription. En vigtig promotor sekvens er **TATA-box**, en region

fyldt med Adenin og thymin, der er ca. 50 opstrøm fra startstedet i beta-globins tilfælde. Der er endnu en region længere op end TATA-box, CAAT-box. Hvis der opstår mutationer i disse boxe, vil man f.eks. udvikle **beta-thalassemia**.

Tandemrepeteret DNA er almindeligt forekommende i eukaryote genomer og kaldes for **satellit-DNA**. Disse familier består af blokke af tandemt repeterede DNA-sekvenser. De enkelte blokke kan forekomme på få eller flere kromosomale lokaliseringer (omkring centromer eller telomer = ca. 10 % af DNA) og kan have betydning i forbindelse med eksempelvis den rekombination der foregår under meiosen samt dannelsen af duplikationer.

Spredte sekvenser findes spredt ud over genomets totale længde. Udgør ca. 15 % af genomet.