1 Transcription II

1.1 E.coli promoters

The sequences which are transcribed are numbered with $\mathbb{Z} \setminus \{0\}$, with 1 at the start of transcription.

Positive ordinals are transcribed.

The higher the number, the more *downstream* the sequences are.

At -35 to -10 there are promoter consesus (most common, average) sequences.

E.coli has different σ-factors, which recognize different consensus sequences.

1.2 Simplified Model of Gene Expression

Transcription and translation are coupled.

In procaryotes, translation can begin before the transcription is finished.

A more realistic model would include polycistronic mRNA generating multiple proteins, which is usually a product of a single promoter to coordinate expression, as well as regulatory genes, which may stop a promoter.

Eukaryotic transcription is more complicated.

Transcription and translation are not coupled, since a nuclear export must happen before translation happens at the location of ribosomes. Note that primary RNA transcript is defferent from mature mRNA. The mature mRNA has a cap, with introns removed by splicing and polyadenylation end. Sometimes primary RNA might have a cap, with introns not removed and without polyadenylation end.

Observe that both eukaryotes and prokaryotes have mRNAs, rRNAs and tRNAs. However, eukaryotes have much more (siRNAs, snRNAs, snoRNAs,scaRNAs, miRNAs and other noncoding RNAs, lots of which are not involved in producing proteins).

Each of RNA polymerase I (larger rNAs), II (mRNAs), and III is a multi-subunit protein which are responsible for transcription of different RNAs. Animals and plants share thes RNAPs, but plants also have RNA polymerase IV and V.

1.3 Subunits of Eukaryotic RNA Polymerases

RNAPs are complex structures with many subunits, where some subunits are common to all three RNAPs and some subunits resembling the subunits of bacterial RNAPs. For example, in RNAP II there is a special CTD, carboxyterminal domain.

Eukaryotic RNA polymerases require proteins, called transcription factors, to help position them at the promoter. These factors fulfill a similar role to sigma subunits of the bacterial RNA ploymerases. Eukaryotic RNA polymerases also need to deal with chromosomal structures.

There is no splicing (no introns), no caps and no poly-a tail at 3' end in prokaryotic mRNA.

There are coding and noncoding sequences. For example, in mRNA the start of translation is marked by the AUG sequence.

Even if introns are removed, there are still coding and noncoding regions. Both in eukaryotes and prokaryotes, there are UTR untranslated regions at 3' and 5'. Translation of mRNA in eukaryotes results in one protein, while in prokaryotes there are several.

1.4 Transcription of Protein-Coding Genes

1.4.1 Eukaryotes

Eukaryotes have the TATA box, found 25-36 base pairs upstream from the start, which positions RNAP II and contains highly conserved and hyghly transcribed (not all) genes. The TATA box is one example of the highly conserved sequences called elements.

Eukaryotes can have any combination of $TATA$ or other elements (Inr, DPE).

TATA is ulually at $(-35) - (-25)$, while the Inr is at 1, with DPE is approximately at $+30$.

TBP binds along the rest of TFID, which includes TAFS, and mobilizes the binding of TFIIB complex adjacent to TATA box, so that RNAP can now bind in the correct orientation at the start site.

Thus, TBP, a subunit of TFIID, binds to the TATA box promoter in the minor mode, bending and distorting DNA. This attracts other transcription factors, which help to orient and bind RNAP II to the DNA. The helicase activity of TFIIH ??.

In RNAP II only the carboxy terminal domain of the largest subunit has a stretch of 7 amino acids that are repeated ??.

RNAPII proceeds by abortive transcription.

There addition factors required for transcription initiation. For example, enhancers control transcriptional activators, which recognize the sequence. There are also insulator sequences, which terminate the passage of transcription.

RNA polymerase II is activated by phosporylation of CTD S groups. There are about 100 subunits involved in initiating eukaryotic transcription.

1.5 RNA Factory

Phosphorylation of C-terminal tail of RNA Polymerase II results in the binding of RNA processing units. Phosphorylation of CTD depends on the phosphorylation of different protein patterns.

1.6 Capping

Capping helps to protect RNA from exonucleases. RNA triphoshatase removes ??. Guanyltransferase removes ??. RNA Methyltransferases then add a Methyl to the G base.

1.7 Introns

Key words: exons, introns, RNA processing increases the number of gene products.

1.8 Splicing reaction

Splicing is a two step process, with adenosine attacking the 5' splice site and 3' of one exon reacting with 5' of next exon to release intron, forming an intron lariat, exons ligate.

1.9 Catalytic Mechanism

2' OH group of the ribose sugar is not present in dexyribose. This group is necessary for the formation of the lariat structure in intron splicing.

Sequences required for the intron removel are almost invariant.

1.10 Termination

Transcription of the consensus sequences and recruitment of modifying proteins.

CPSF and CtsF move from the CTD to the specific sequences on the RNA.

PAP = poly-a polymerase

PABPN1 = Poly-a binding protein N1

mRNA translated in the cytoplasm, a site of protein synthesis.