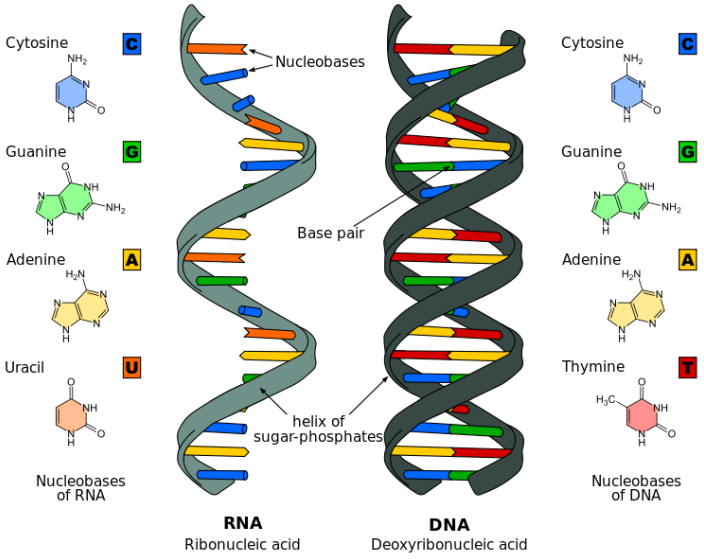
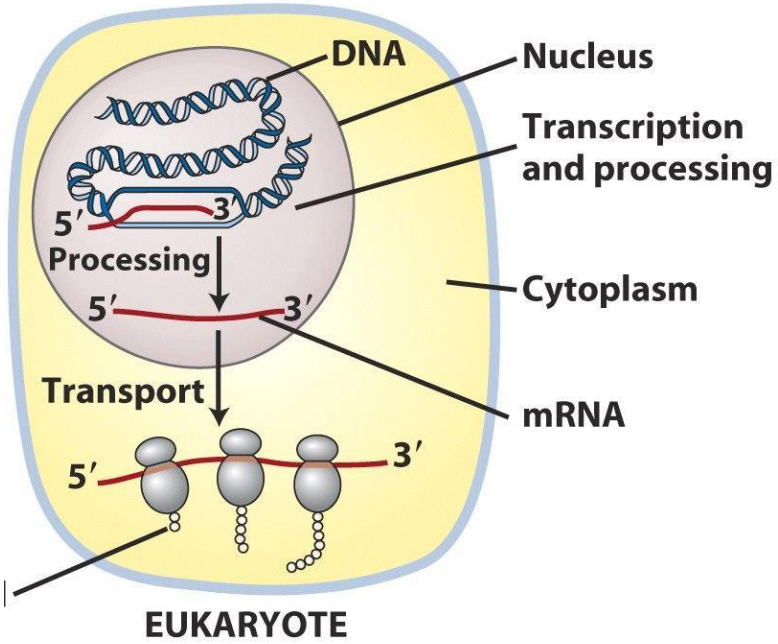
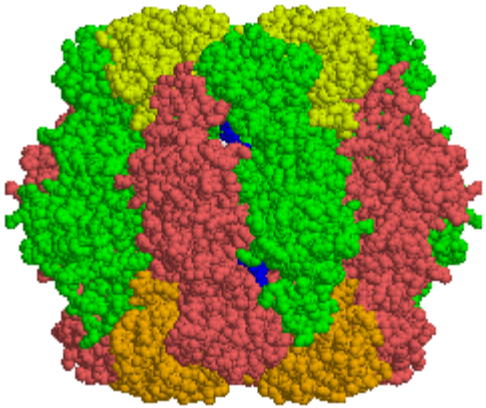
DNA to RNA to protein







Gene to Phenotype: The BAD2 gene and fragrance in rice

Plant Biotechnology Journal (2005) **3**, pp. 363–370 doi: 10.1111/j.1467-7652.2005.00131.x

The gene for fragrance in rice

Louis M. T. Bradbury[1,2], Timothy L. Fitzgerald[1,2], Robert J. Henry[1,2,*], Qingsheng Jin[1,2,3] and Daniel L. E. Waters[1]

[1]*Centre for Plant Conservation Genetics, Southern Cross University, Lismore, NSW 2480, Australia*
[2]*Grain Foods CRC, Southern Cross University, Lismore, NSW 2480, Australia*
[3]*Crop Research Institute, Zhejiang Academy of Agricultural Sciences, Hangzhou, China*

Received 29 October 2004;
revised 21 December 2004;
accepted 22 December 2004.
**Correspondence* (fax +61 2 66203010;
e-mail rhenry@scu.edu.au)

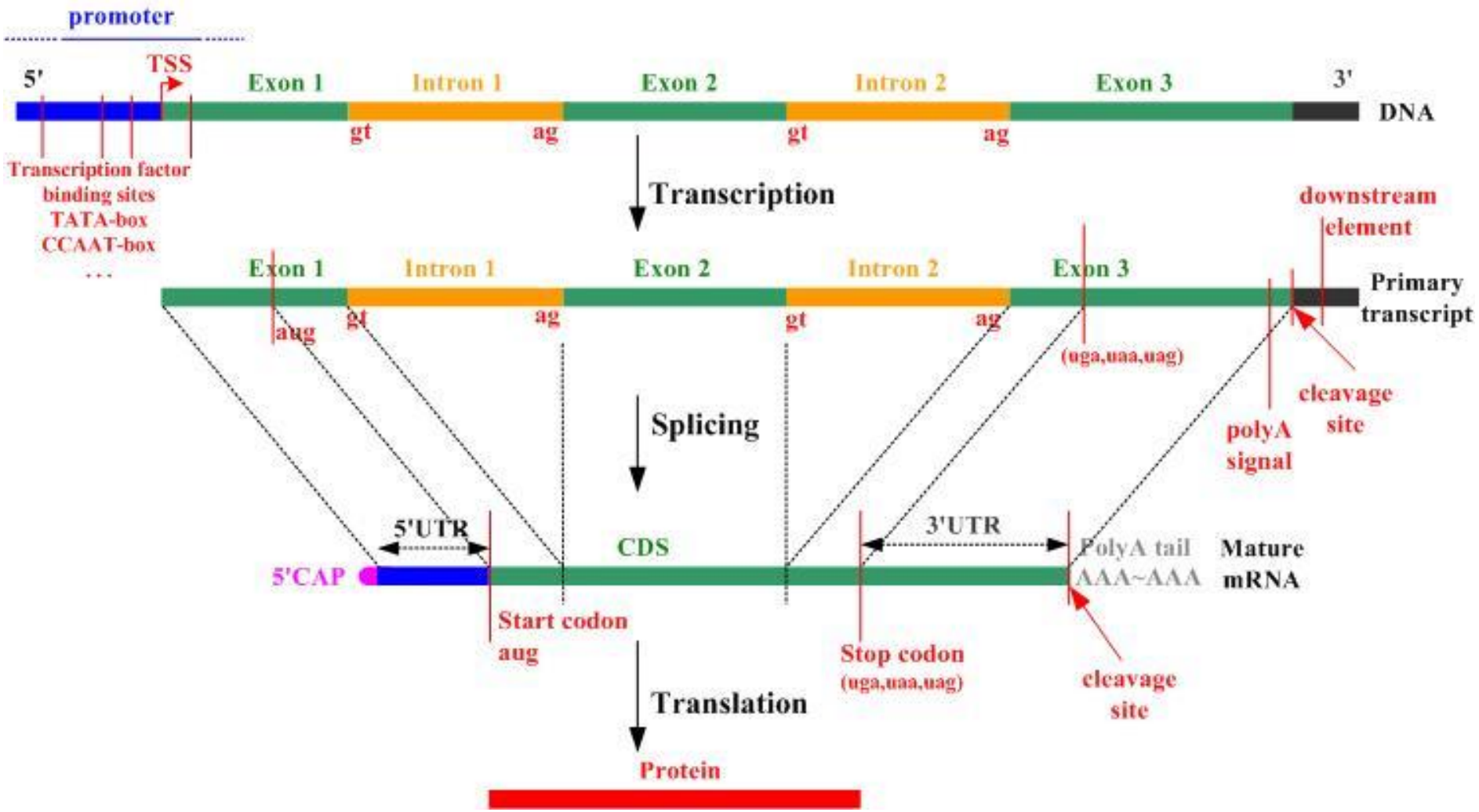
Summary

The flavour or fragrance of basmati and jasmine rice is associated with the presence of 2-acetyl-1-pyrroline. A recessive gene (*fgr*) on chromosome 8 of rice has been linked to this important trait. Here, we show that a gene with homology to the gene that encodes betaine aldehyde dehydrogenase (BAD) has significant polymorphisms in the coding region of fragrant genotypes relative to non-fragrant genotypes. The accumulation of 2-acetyl-1-pyrroline in fragrant rice genotypes may be explained by the presence of mutations resulting in a loss of function of the *fgr* gene product. The allele in fragrant genotypes has a mutation introducing a stop codon upstream of key amino acid sequences conserved in other BADs. The *fgr* gene corresponds to the gene encoding BAD2 in rice, while BAD1 is encoded by a gene on chromosome 4. BAD has been linked to stress tolerance in plants. However, the apparent loss of function of BAD2 does not seem to limit the growth of fragrant rice genotypes. Fragrance in domesticated rice has apparently originated from a common ancestor and may have evolved in a genetically isolated population, or may be the outcome of a separate domestication event. This is an example of effective human selection for a recessive trait during domestication.

Keywords: 2-acetyl-1-pyrroline, betaine aldehyde dehydrogenase, aroma, basmati, jasmine.

Eukaryotic Gene Structure

DNA sequence specifying a protein 200 – 2,000,000 nt (bp)



RNA

Ribonucleic acid (RNA) is a key nucleic acid in transcription and translation. RNA is like DNA except that:

1. Usually single rather than double stranded
2. Pentose sugar is ribose rather than deoxyribose
3. It contains the pyrimidine base uracil (U) rather than thymine (T)

Classes of RNA

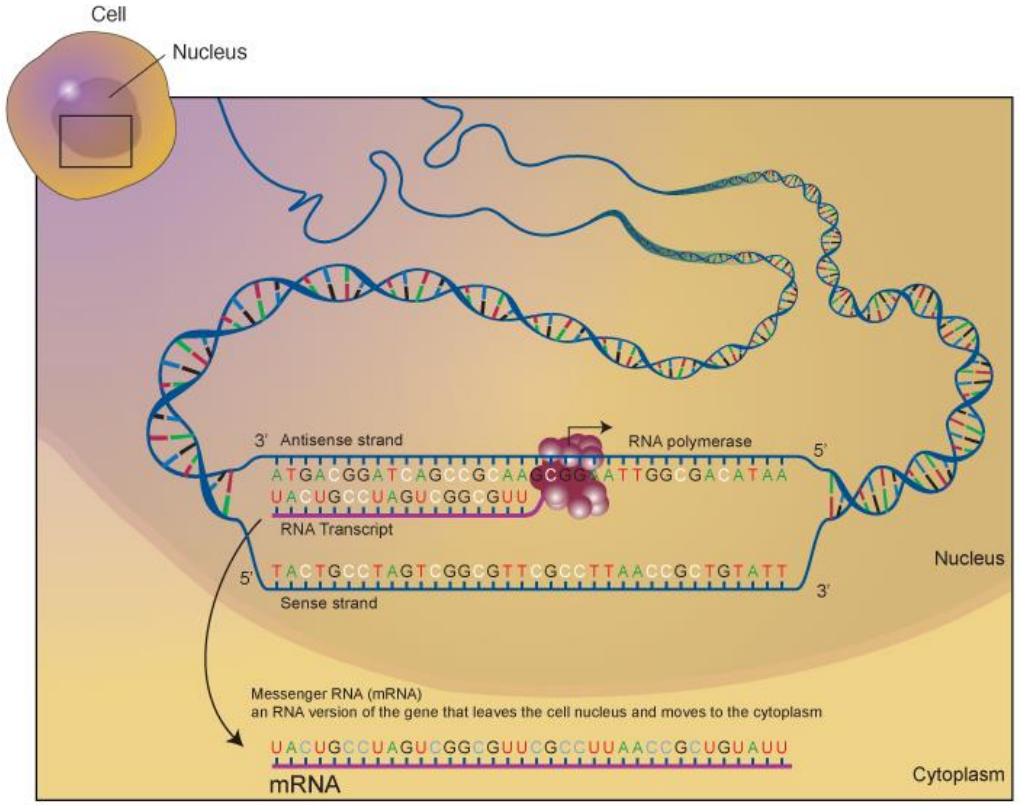
1. Informational (messenger); mRNA

2. Functional (transfer, ribosomal RNA)
   - tRNA
   - rRNA

3. Regulatory: (RNAi)

Informational (messenger) - mRNA

- single-stranded RNA molecule that is complementary to one of the DNA strands of a gene
- an RNA transcript of the gene that leaves the nucleus and moves to the cytoplasm, where it is translated into protein

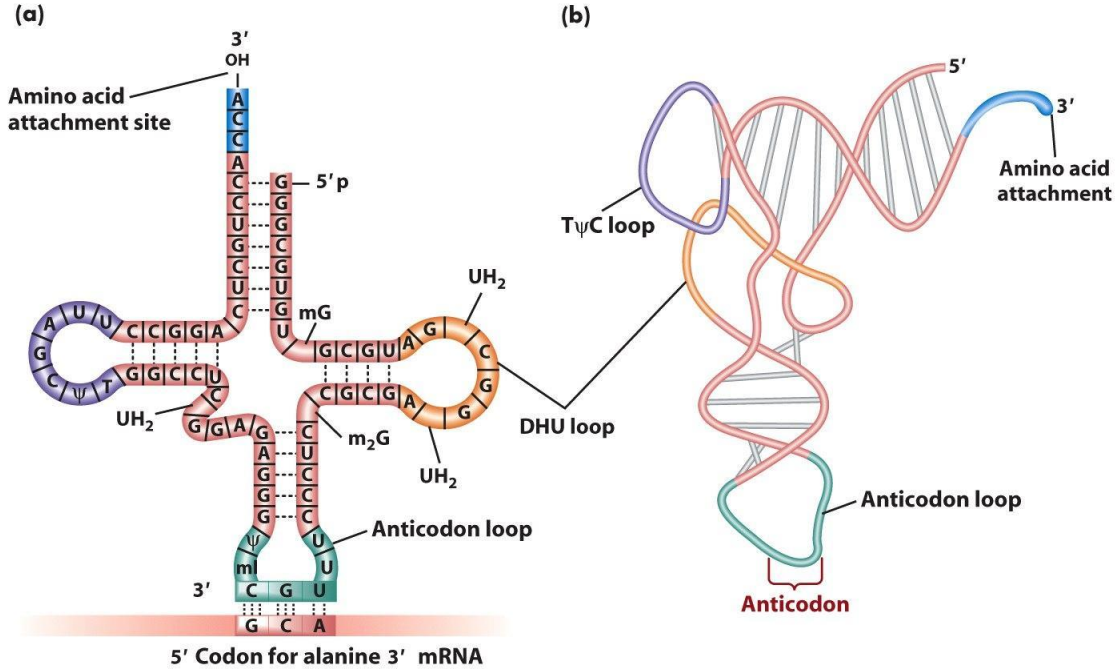


http://www.genome.gov/glossary

Functional (transfer) - tRNA

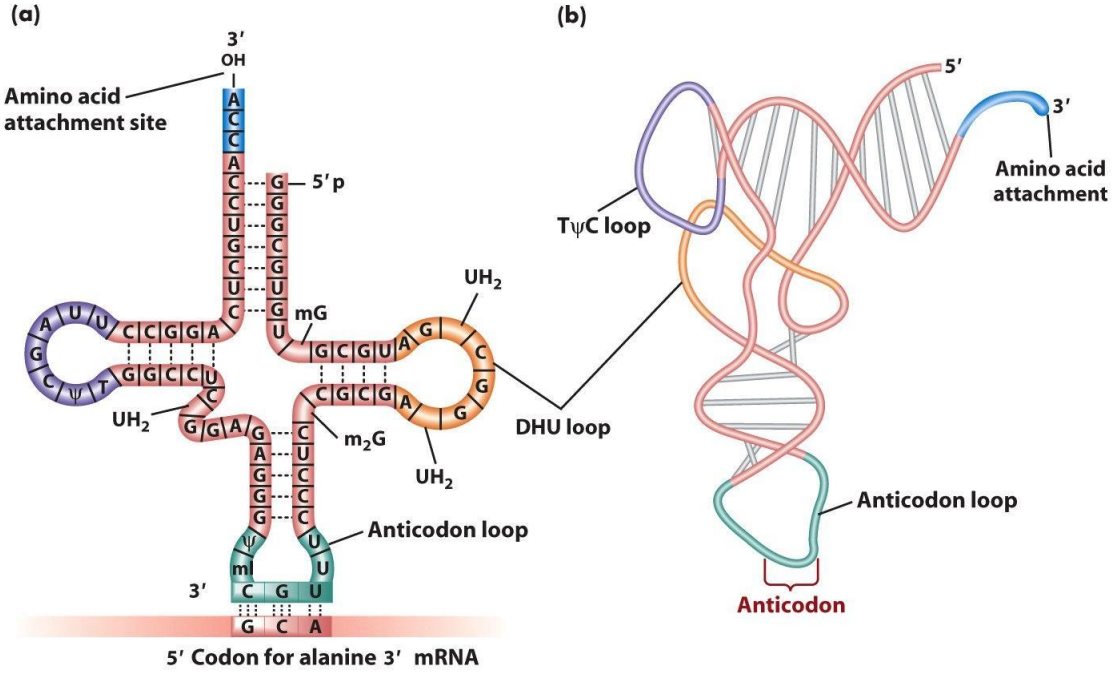
Molecules that carry amino acids to the growing polypeptide:
~ 32 different kinds of tRNA in a typical eukaryotic cell

- Each is the product of a separate gene.
- They are small containing ~ 80 nucleotides.
- Double and single stranded regions
- The unpaired regions form 3 loops



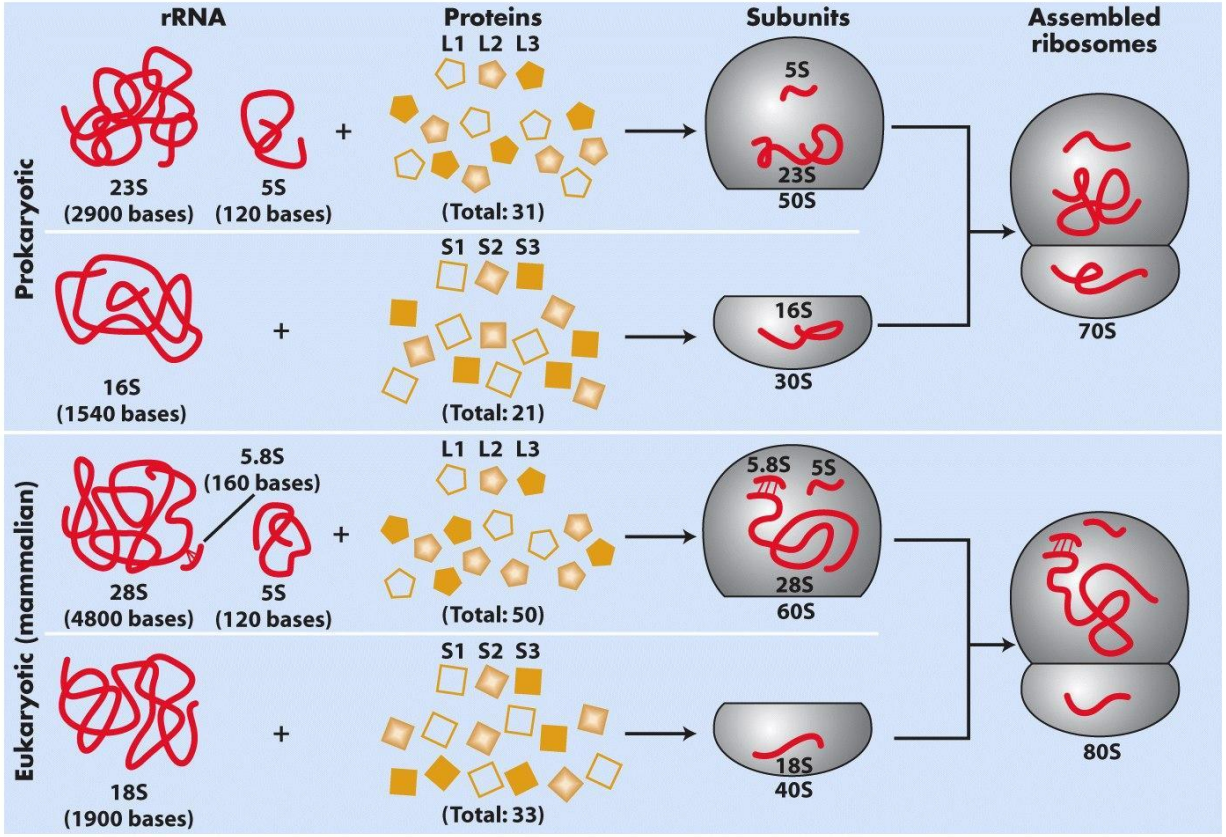
Functional (transfer) - tRNA

- Each kind of tRNA carries (at its 3′ end) one of the 20 **amino acids**
- At one loop, 3 unpaired bases form an **anticodon**.
- Base pairing between the anticodon and the complementary codon on a mRNA molecule brings the correct amino acid into the growing polypeptide chain



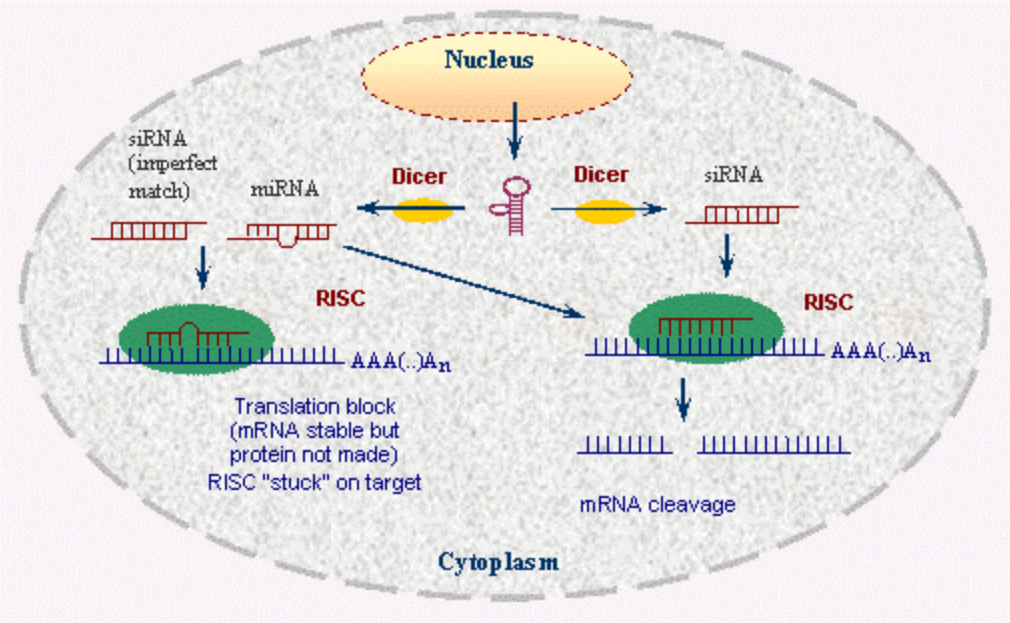
Functional (ribosomal) - rRNA

- The ribosome consists of RNA and protein
- Site of protein synthesis
- Ribosome reads the mRNA sequence
- Uses the genetic code to translate it into a sequence of amino acids



Regulatory (silent) - RNAi

Regulatory RNA: So special it deserves a section all its own.



www.ncbi.nlm.nih.gov

Transcription

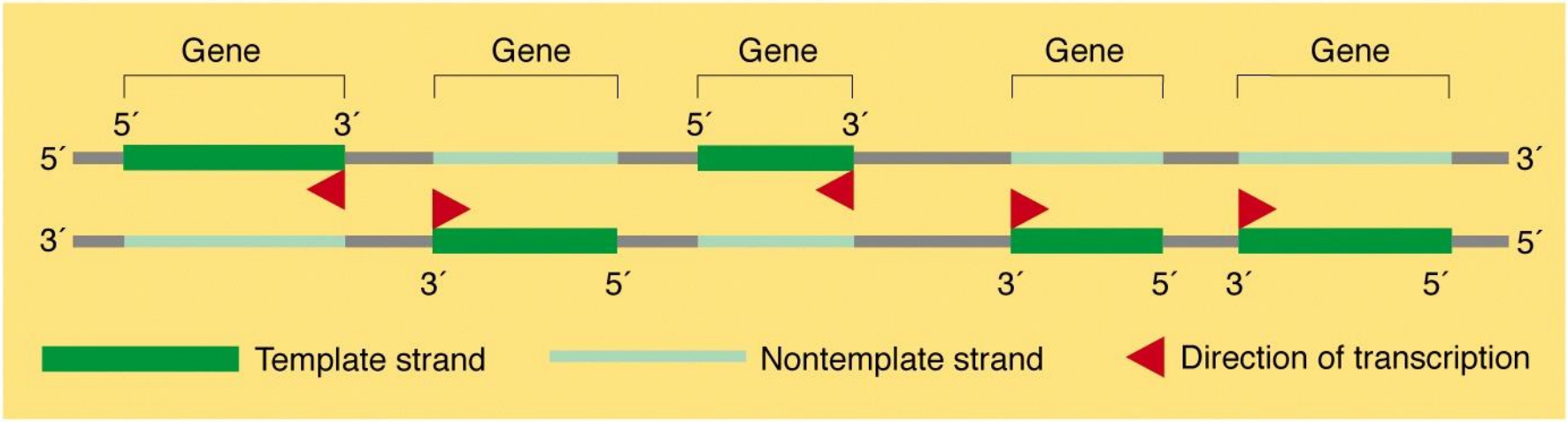
- Messenger RNA (**mRNA**) is an intermediate in the transcription process
  - Transmits the information in the DNA to the next step: translation

- Three transcription steps: **initiation, elongation, and termination**.

- Either DNA strand may be the template for RNA synthesis for a given gene.
  - For any given gene, the *template strand* is also referred to as the antisense (or non-coding) strand
  - Non-template strand is the sense (or coding) strand
  - The same *DNA strand* is not necessarily transcribed throughout the entire length of the chromosome or throughout the life of the organism.

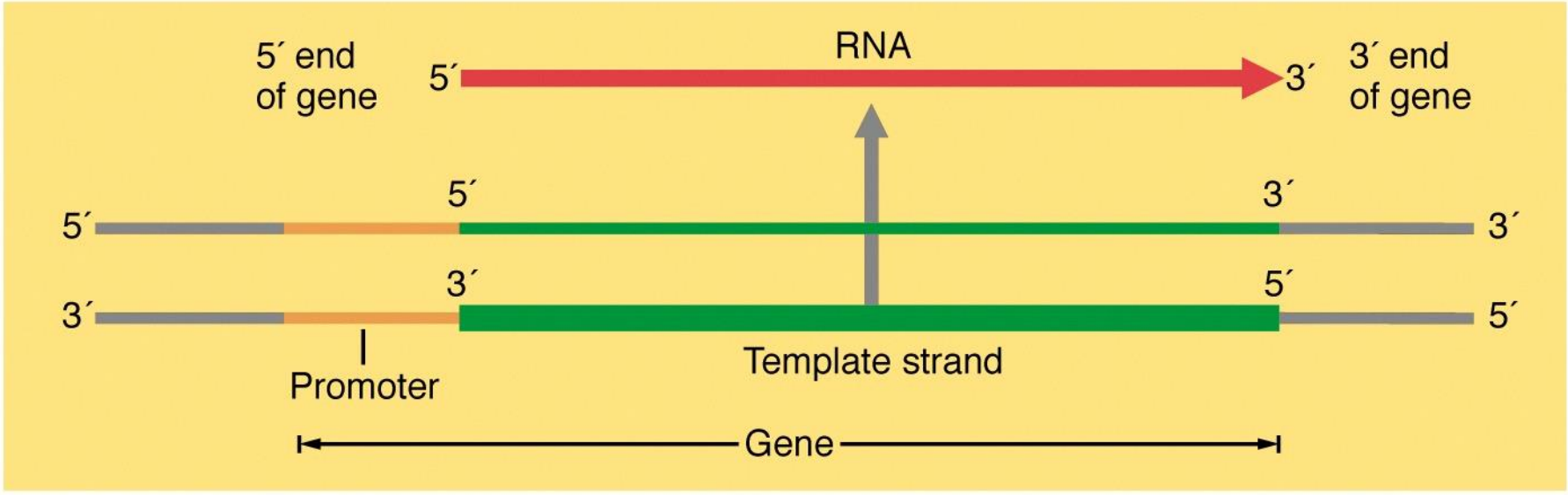
Transcription

Either strand of the DNA may be the template strand for RNA synthesis for a given gene.



Transcription

The *template strand* is also referred to as the antisense (or non-coding) strand and the non-template strand as the sense (or coding) strand. The same *DNA strand* is not necessarily transcribed throughout the entire length of the chromosome or throughout the life of the organism.



Transcription & Gene Expression

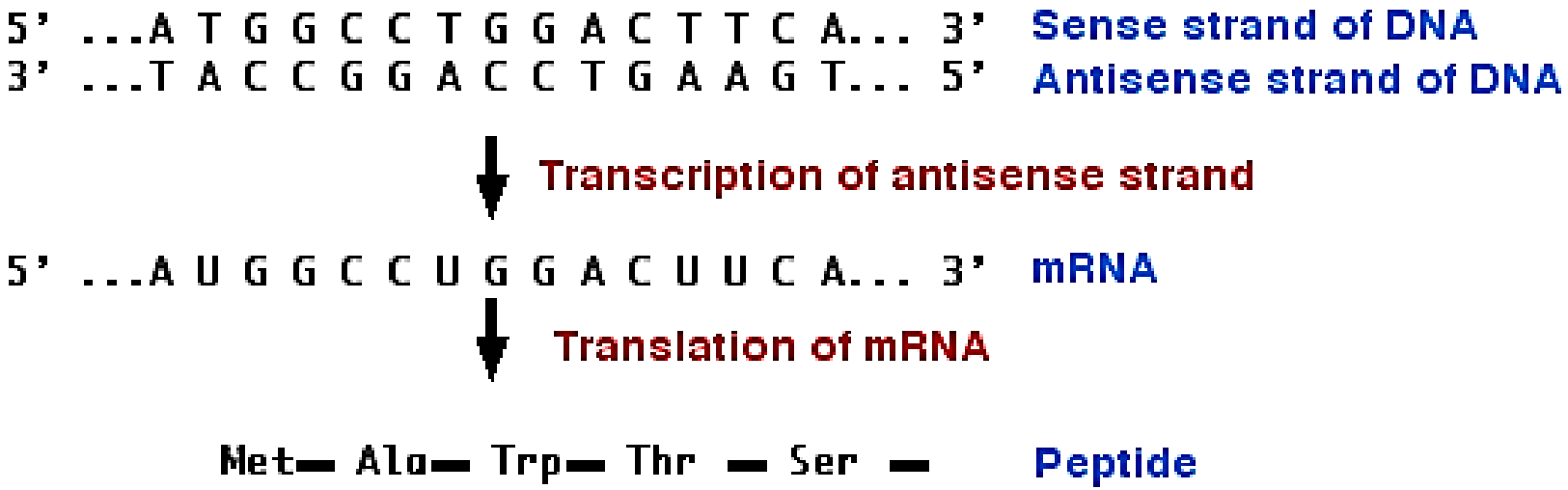
The majority of genes are expressed as the proteins they encode. The process occurs in two steps:

- **Transcription = DNA → RNA**
- **Translation = RNA → protein**

Taken together, they make up the "central dogma" of biology: DNA → RNA → protein.

DNA to RNA to protein

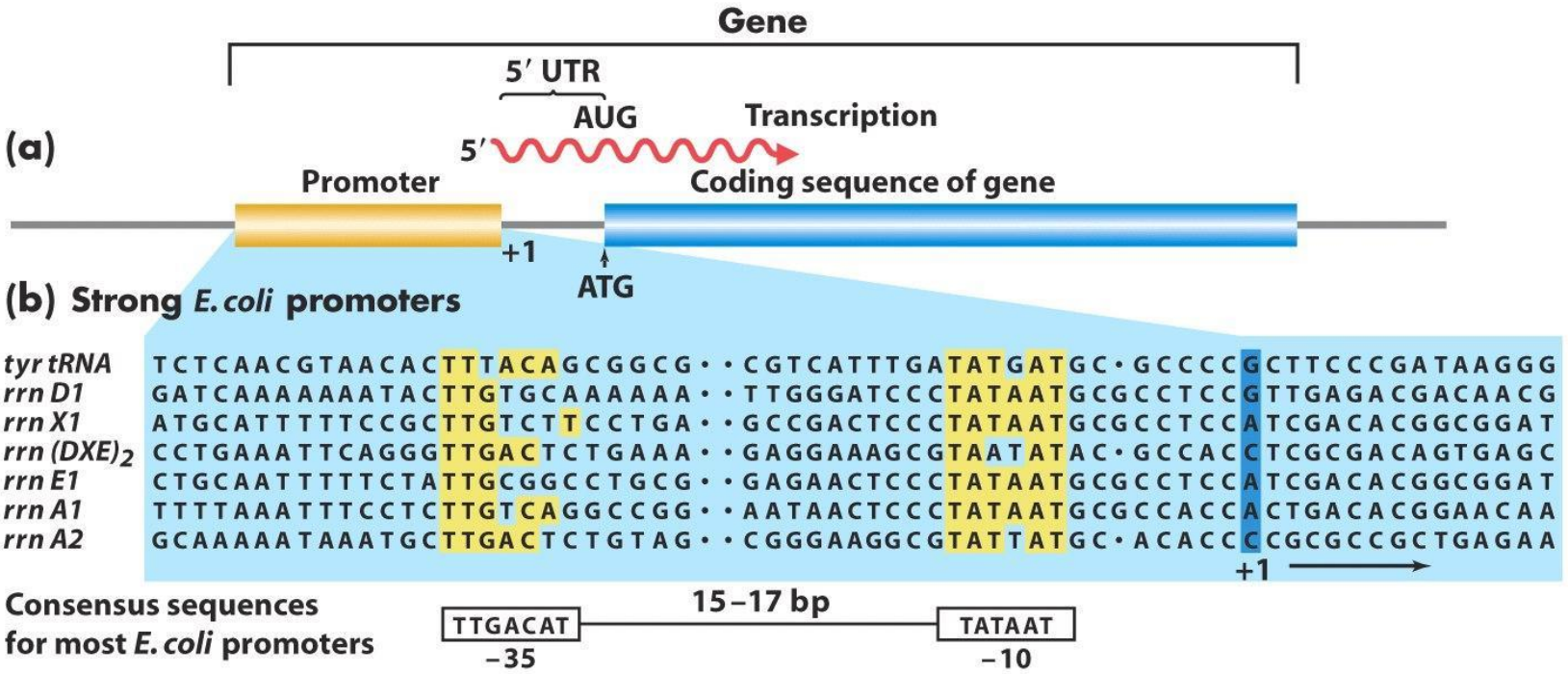
Here is an overview.



http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/T/Transcription.html

Transcription: Initiation

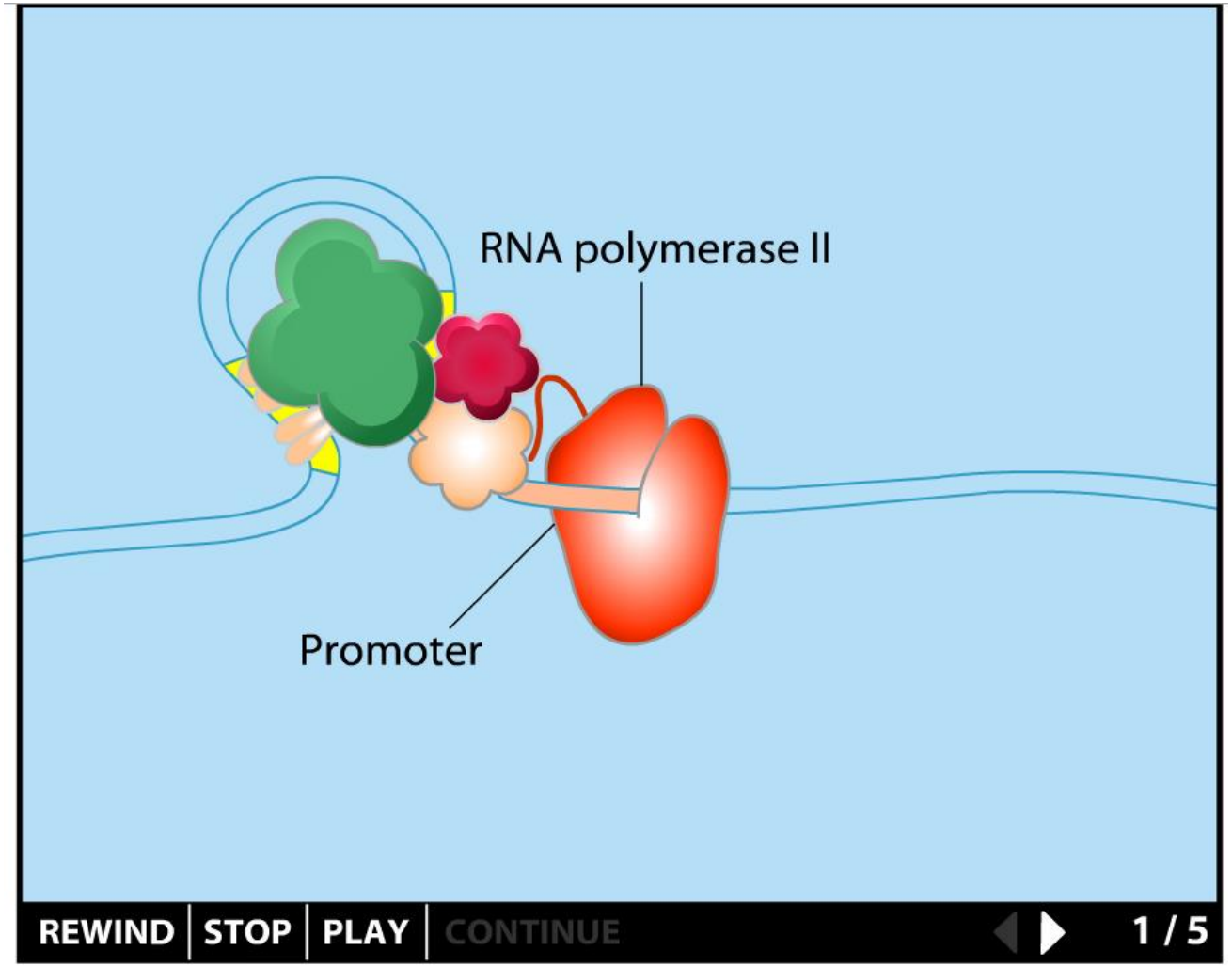
1. *Initiation:* Transcription is initiated at the promoter.



The promoter is a key feature for control of gene expression. Promoters have defined attributes, in terms of their sequence organization.

Transcription: Elongation

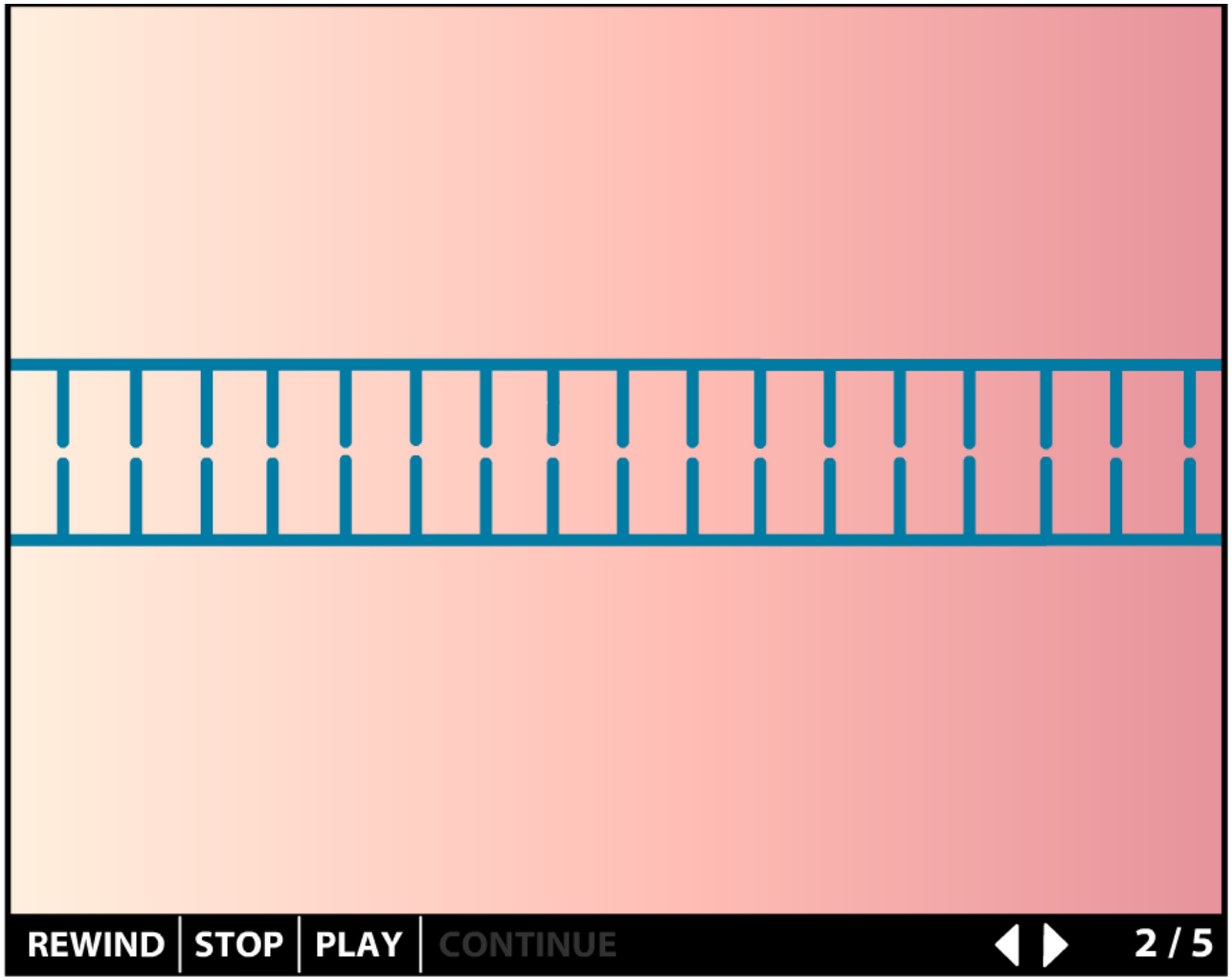
2. Elongation*:*



RNA polymerase is positioned at a promoter DNA sequence by various transcription factors.

Transcription: Elongation

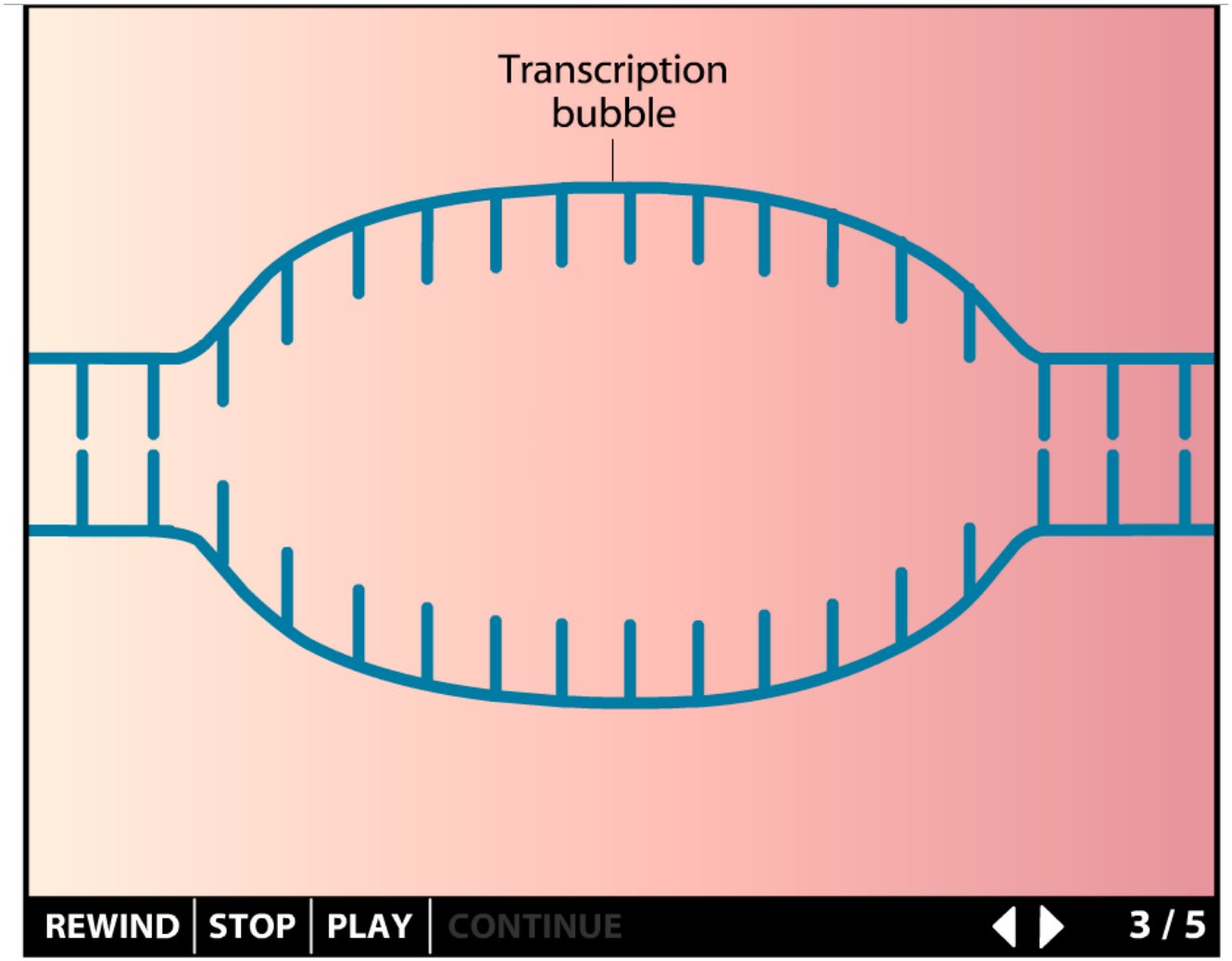
2. Elongation:



The polymerase separates the base pairs of DNA, forming a transcription bubble.

Transcription: Elongation

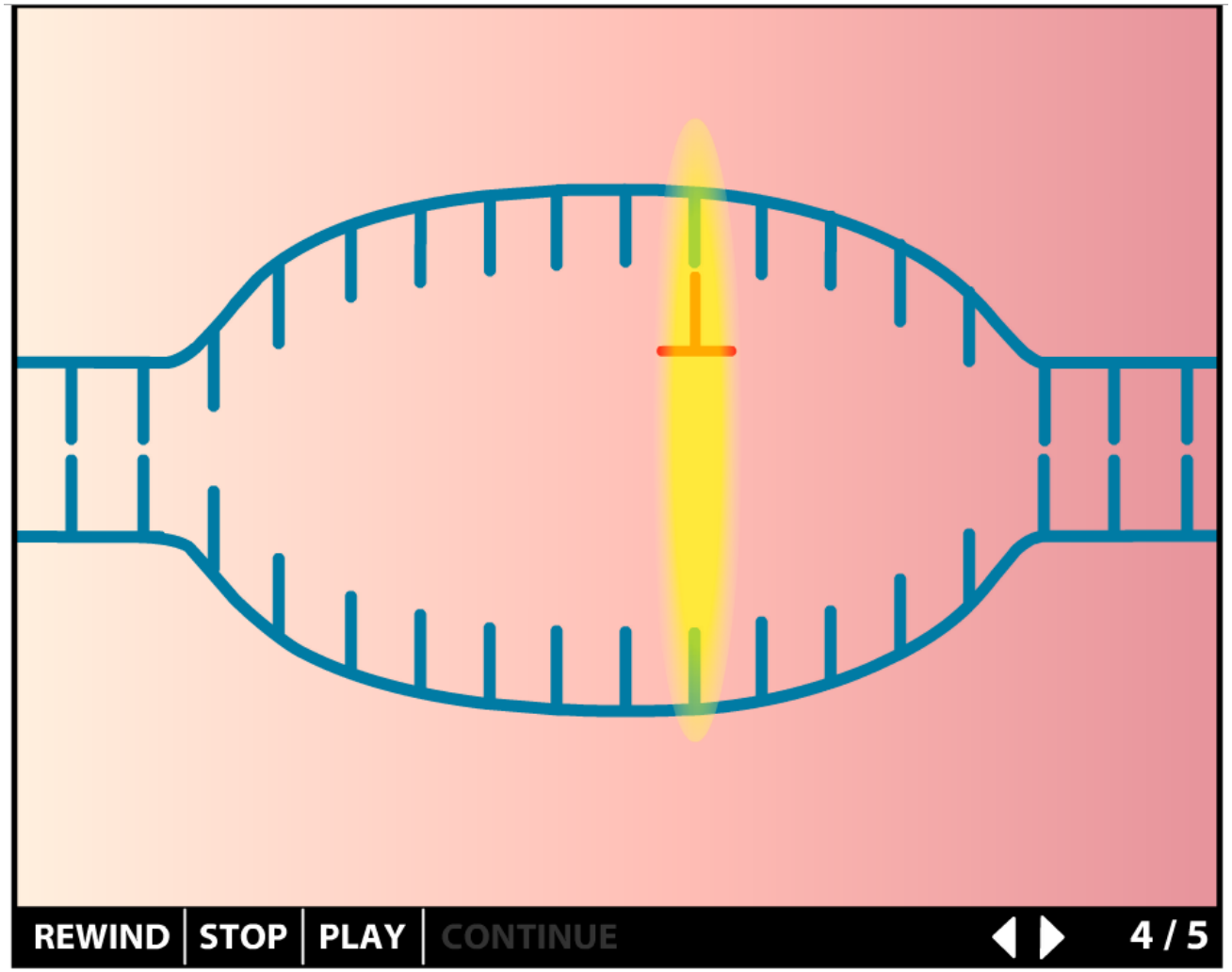
2. Elongation*:*



A ribonucleoside triphosphate diffuses into the active site and base pairs with the complementary base of the template.

Transcription: Elongation

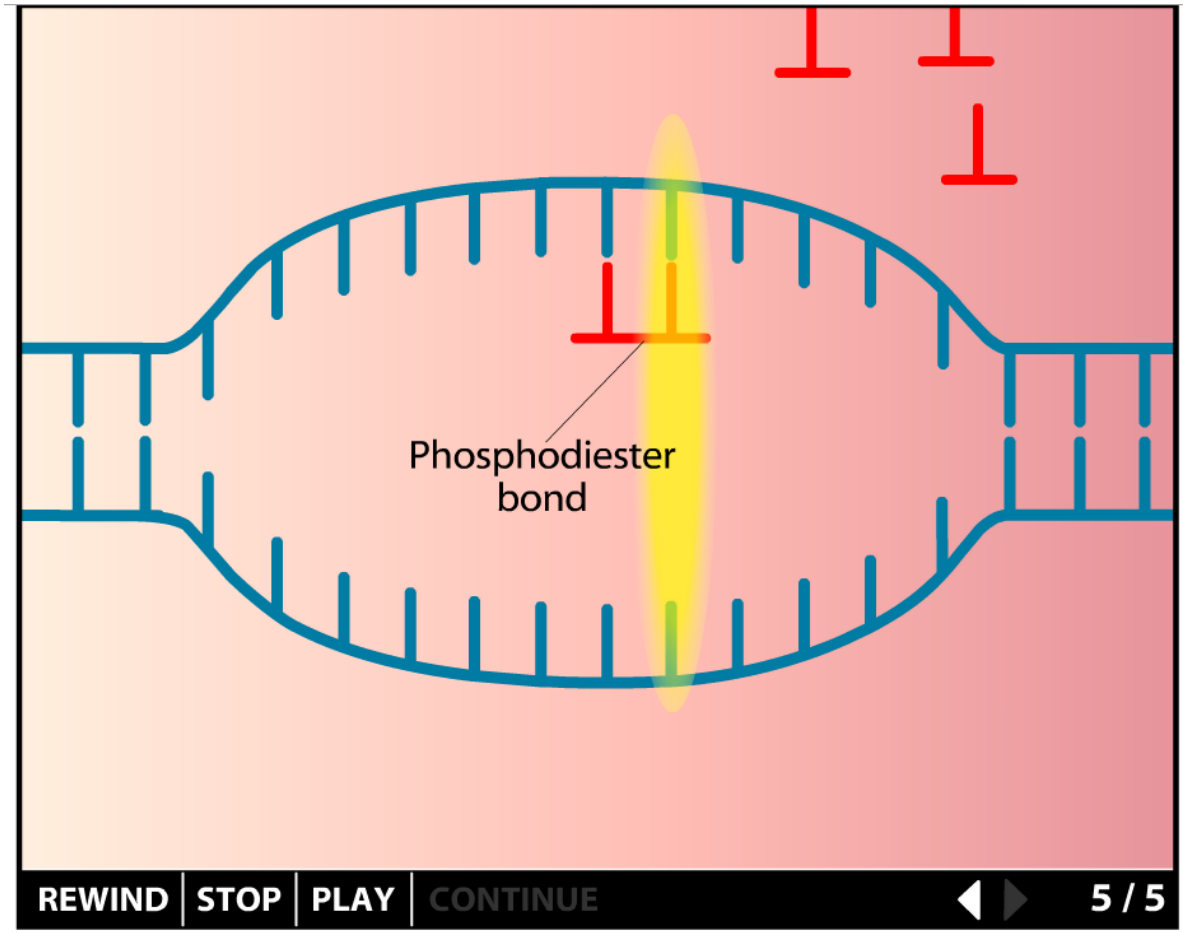
2. Elongation*:*



A second ribonucleotide pairs with the template and is joined to the first base by a phosphodiester bond.

Transcription: Elongation

2. Elongation*:*

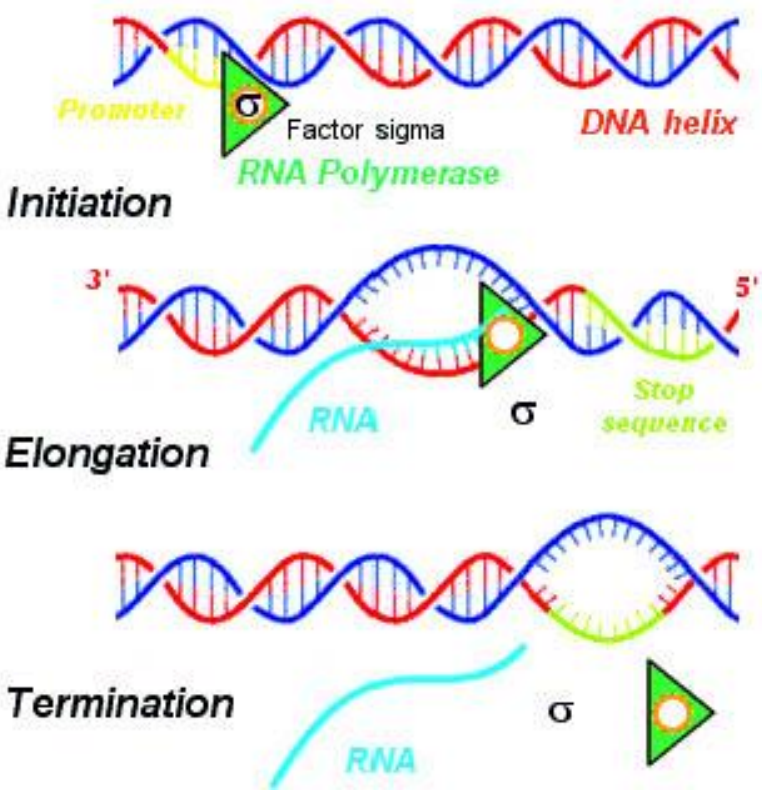


The polymerase then copies the DNA in a stepwise fashion, separating the DNA base pairs ahead and forming base pairs behind. The RNA is elongated as the polymerase transcribes the gene.

Transcription: Termination

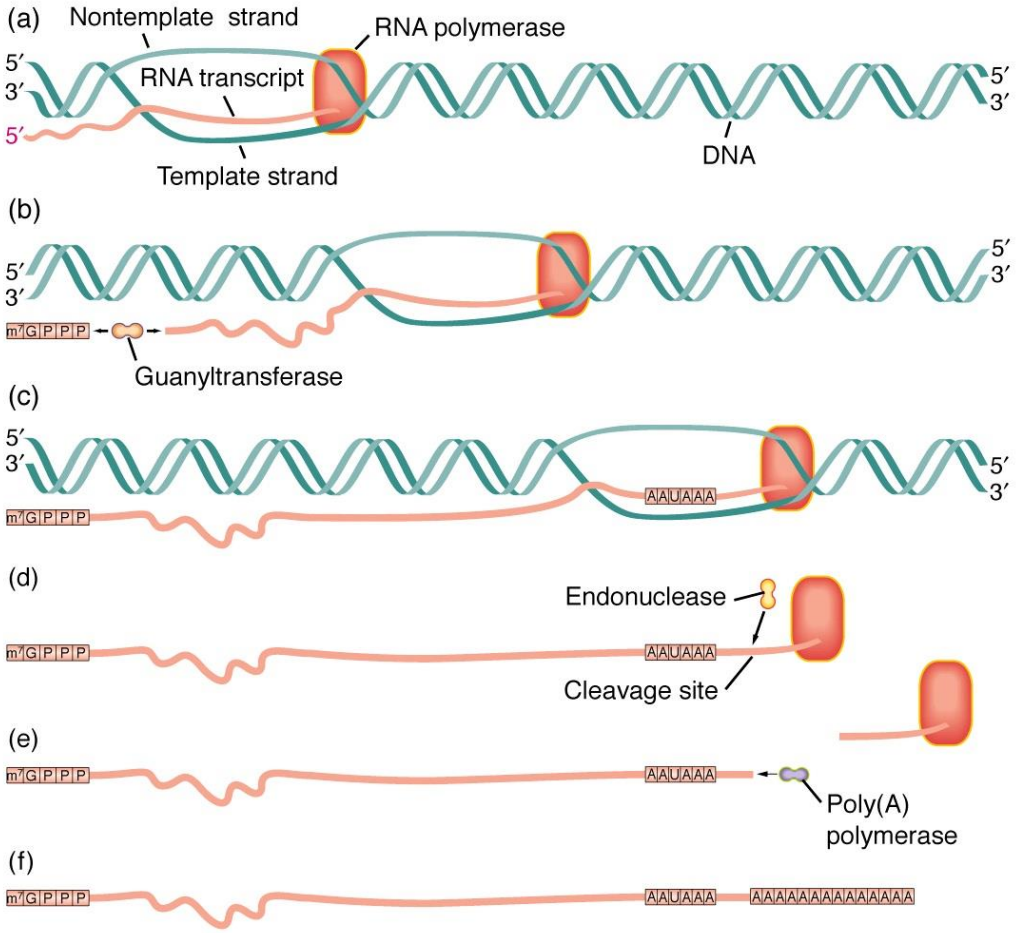
3. Termination

- Specific nucleotide sequence in the DNA signals end of transcription: the “terminator” sequence.
- The terminator sequences recognized by specific proteins associated with the RNA polymerase II.
- Signal end of transcription and additional of poly-A tails



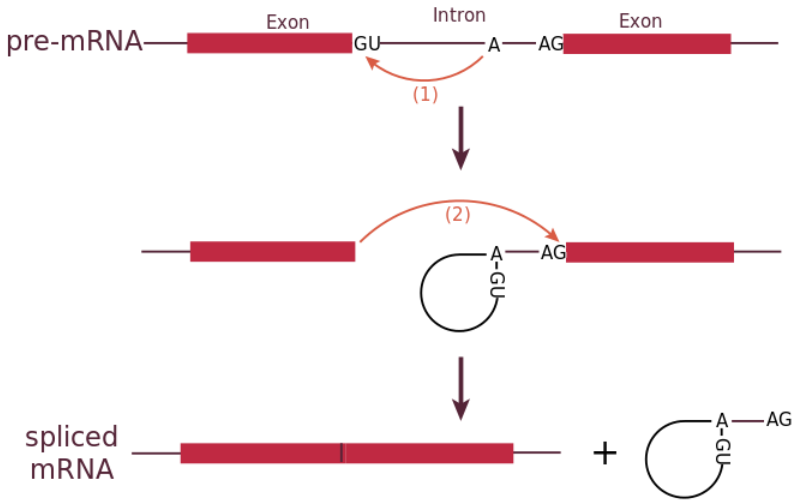
Transcript Processing

- Prokaryotes - mRNA is sent on to the ribosome for translation.
- Eukaryotes - primary RNA transcript is processed into a mature mRNA before exporting to the cytoplasm for translation.

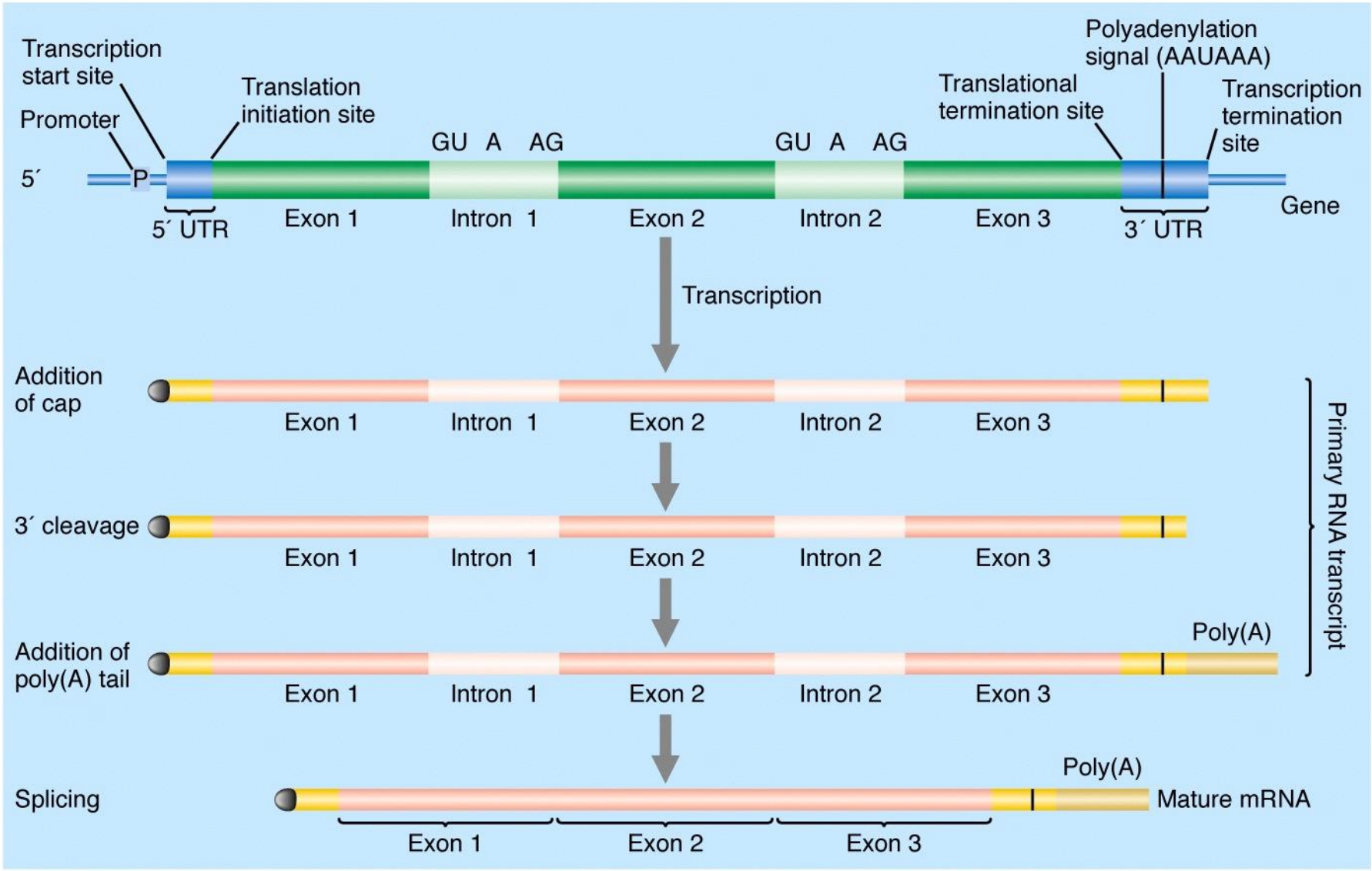


Transcript Processing

1. 5’ cap: 7-methylguanosine added to free phosphate at 5’ mRNA
   - *Prevents degradation and assists in ribosome assembly*
2. 3’poly(A tail): After pre-mRNA is cleaved, poly (A) polymerase adds ~200 A nucleotides
   - *Protects against degradation, aids export to cytoplasm, and involved in translation initiation*
3. Splicing: Removal internal portions of the pre-mRNA
   - Most eukaryotic genes have an intron/exon structure
   - Splicing removes introns and remaining exons are rejoined

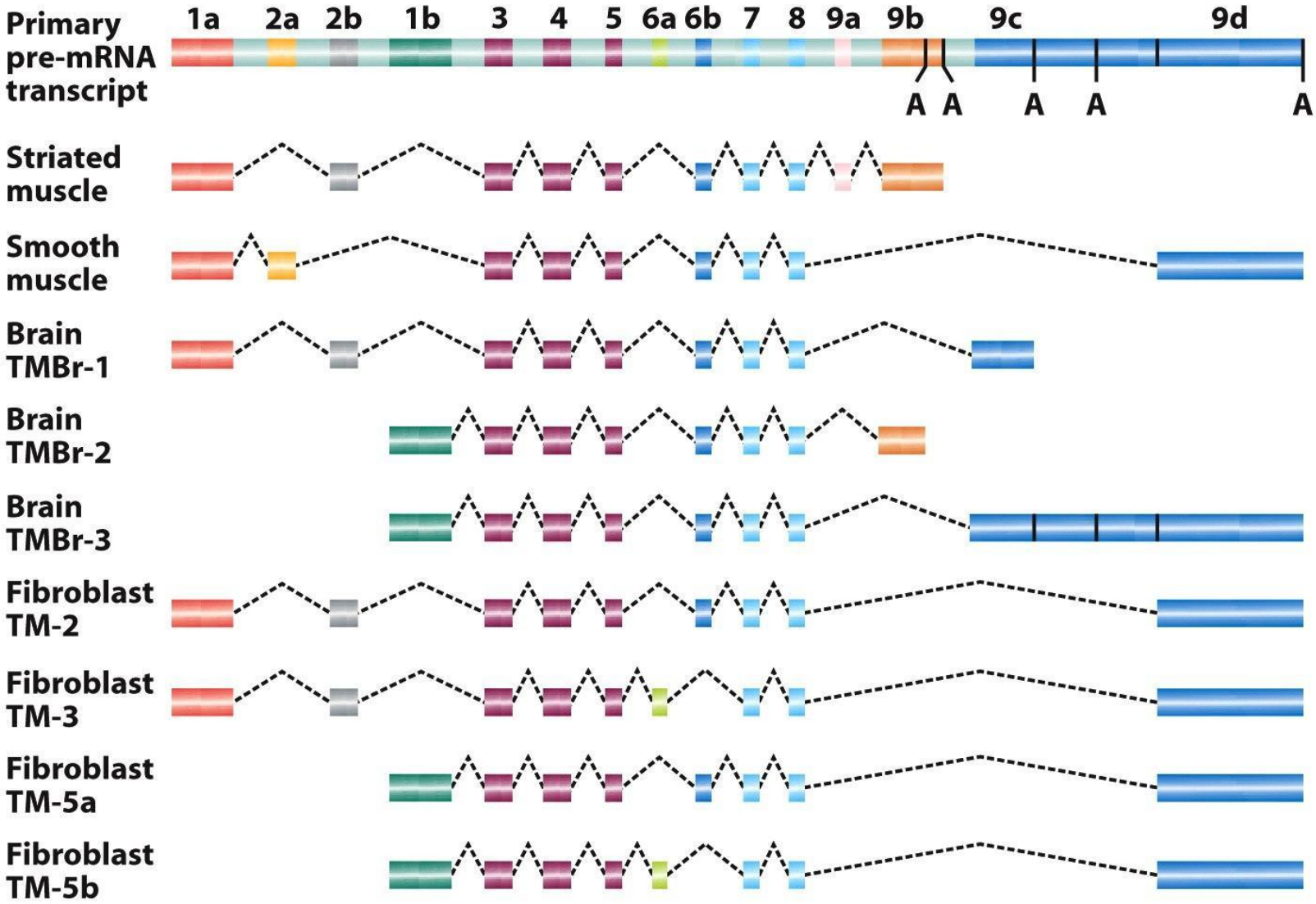


Transcript Processing



Transcript Processing

Changes in intron sequence *splicing* can affect what the gene encodes



The Genetic Code

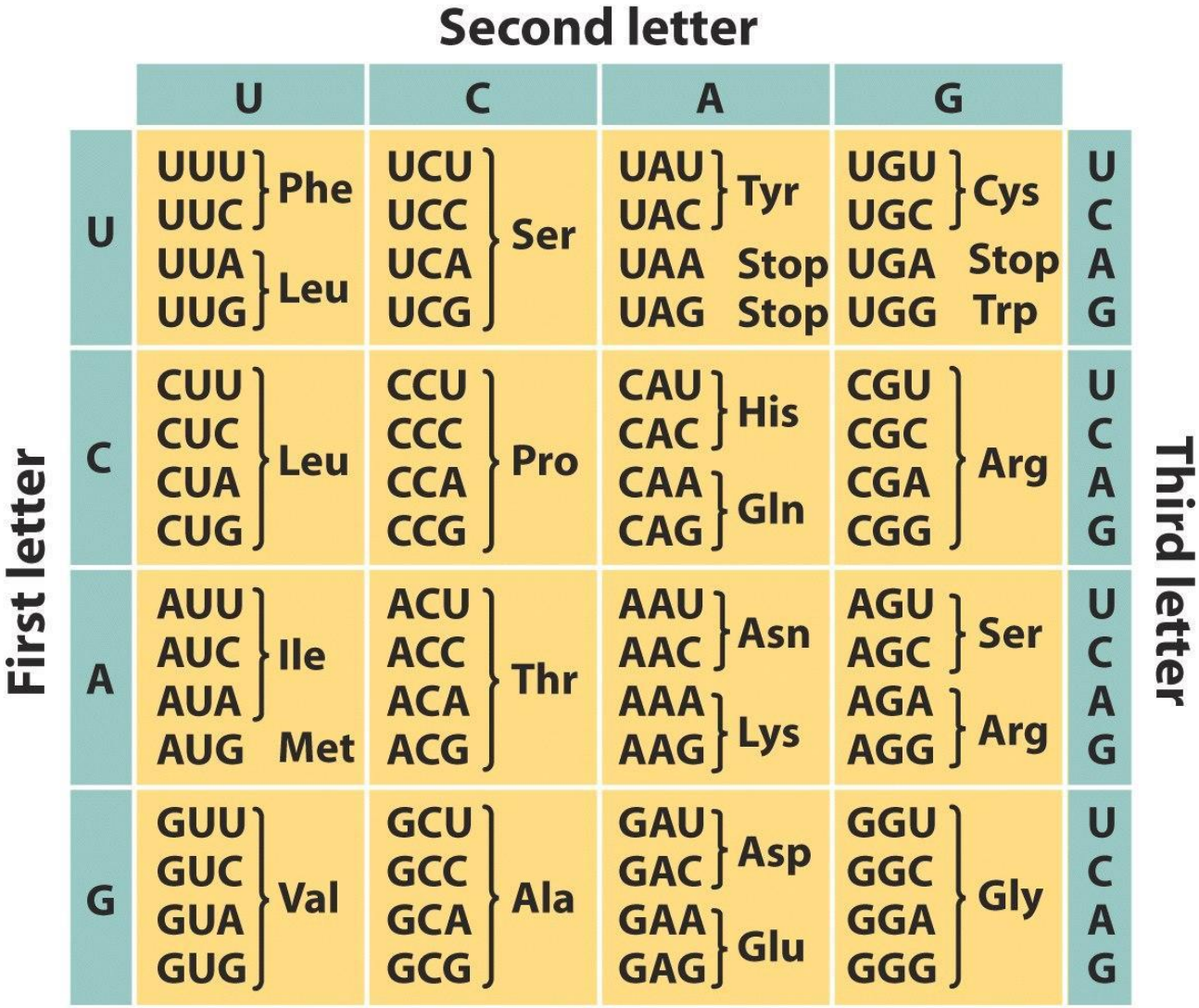
The sequence of a coding (sense, non-template) strand of DNA, read 5' – 3', specifies a sequence of amino acids (read N-terminus to C-terminus) via a triplet code. Each triplet is called a **codon** and 4 bases give **4^3 possible combinations**.

Reading the DNA code: There are 64 codons; 61 represent amino acid codes and 3 cause the termination of protein synthesis (stop codons).

Degeneracy: Most amino acids represented by >1 triplet

Reading the Code

There are 64 codons; 61 represent amino acid codes and 3 cause the termination of protein synthesis (stop codons).



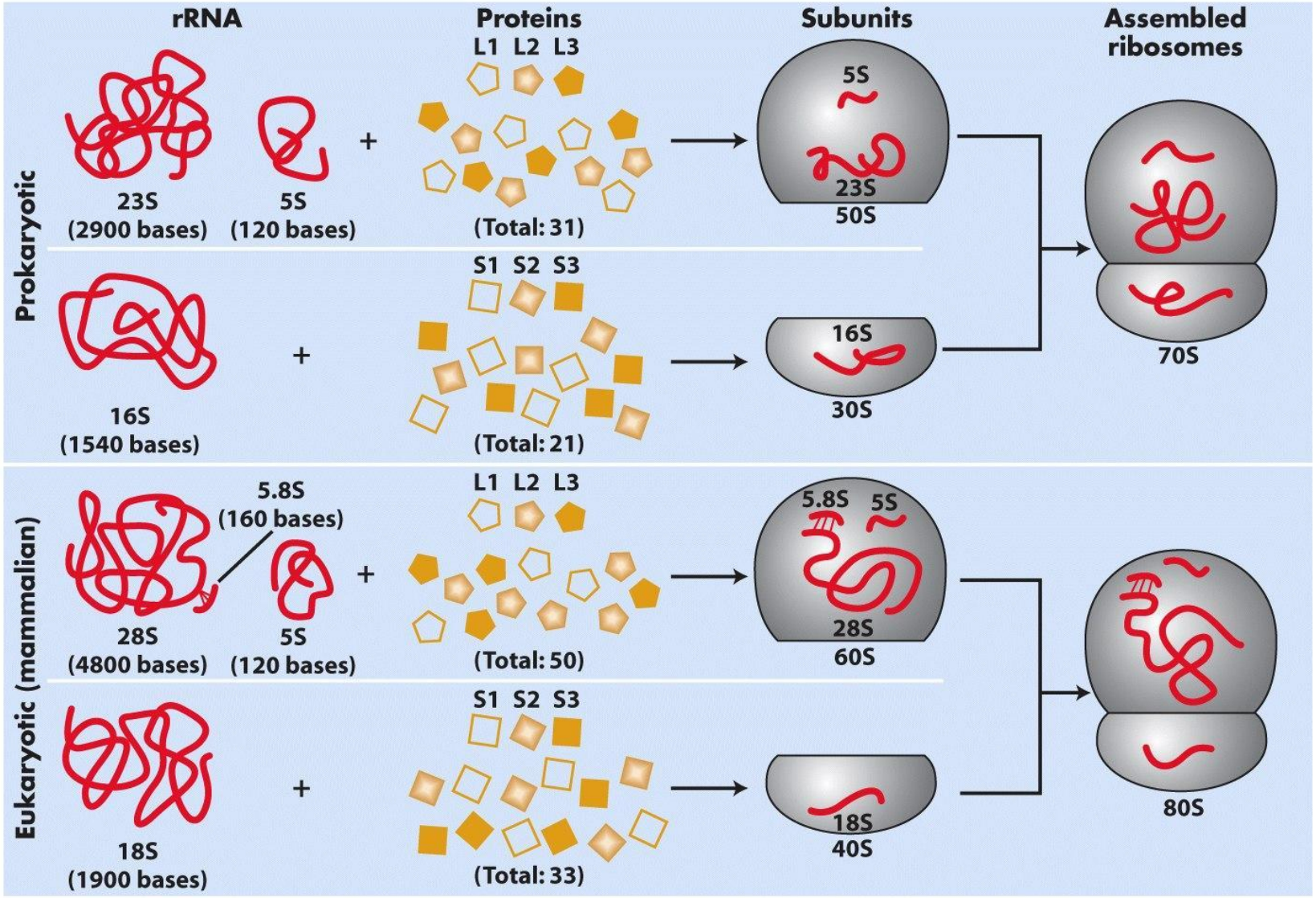
Translation

Overview: The process of *translation* takes the information that has been transcribed from the DNA to the mRNA and, via some more intermediates (ribosomes and transfer RNA), gives the sequence of amino acids that determine the polypeptide.

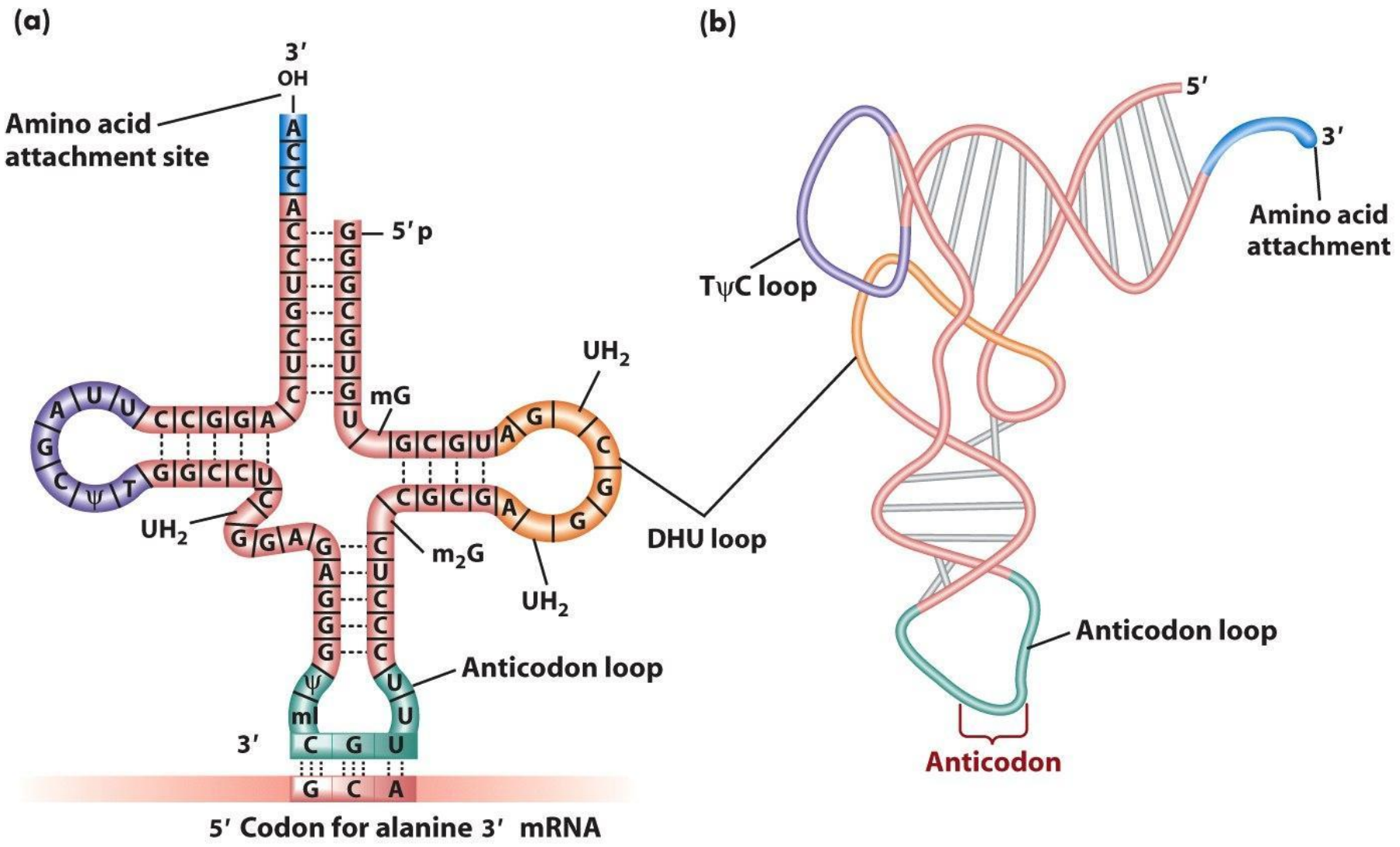
1. *Ribosomes:*

2. *Transfer RNA (tRNA).*

Ribosomes: Structure & Subunits



Transfer RNA (tRNA)



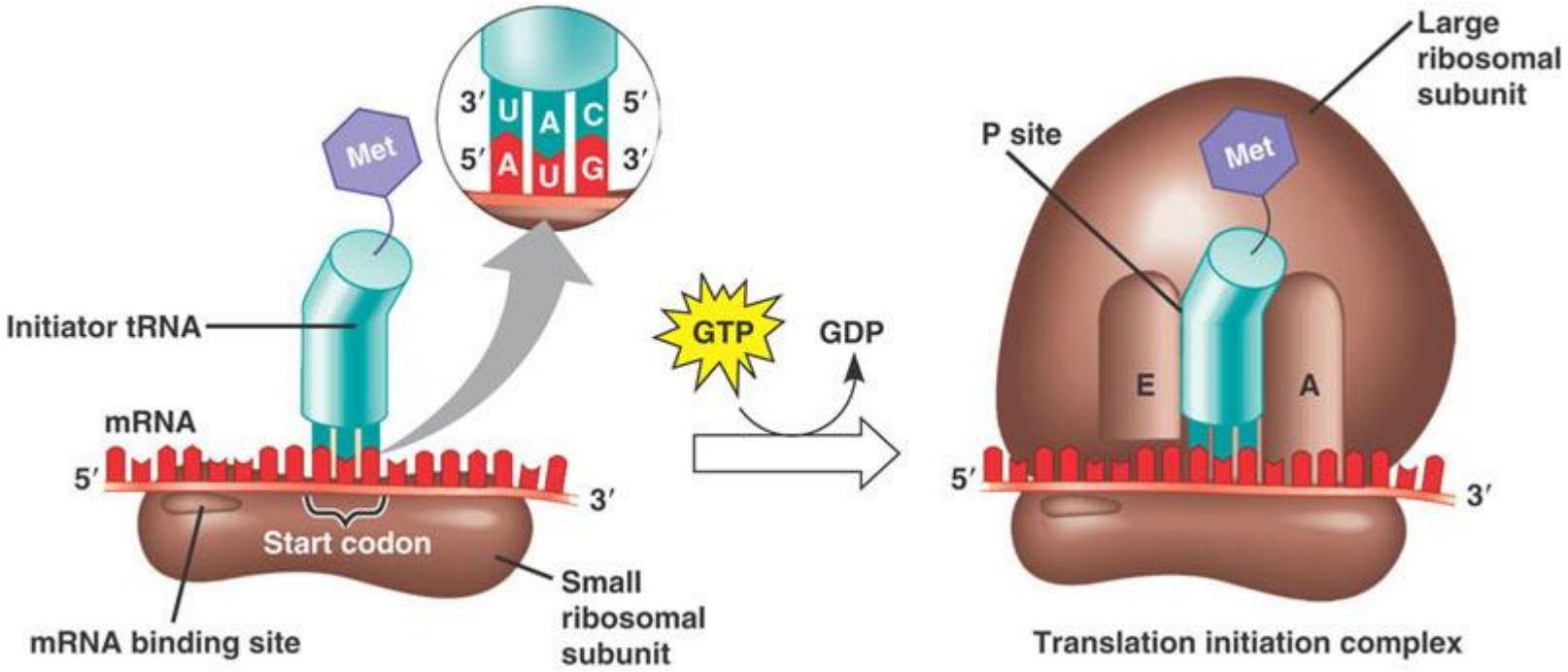
Translation – 3 Steps

1. *Initiation*: In addition to the mRNA, ribosomes, and tRNAs, initiation factors are required to start translation. The AUG codon specifies initiation, in the correct sequence context. It also specifies methionine (MET).

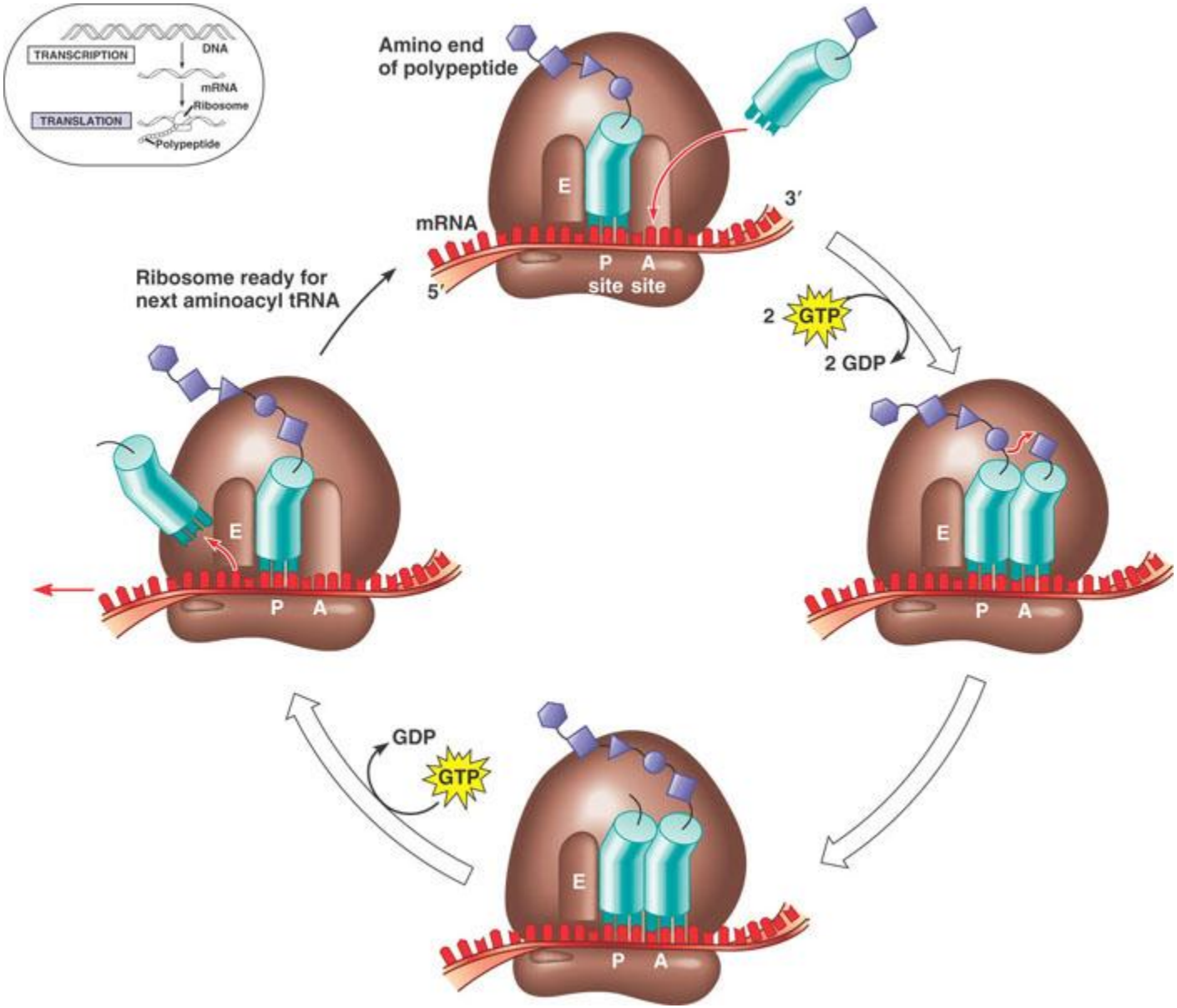
2. *Elongation*: Much as initiation factors were important in the first step, now elongation factors come into play. The reactions also require additional components and enzymes.

3. *Termination*: There are three "stop" codons.

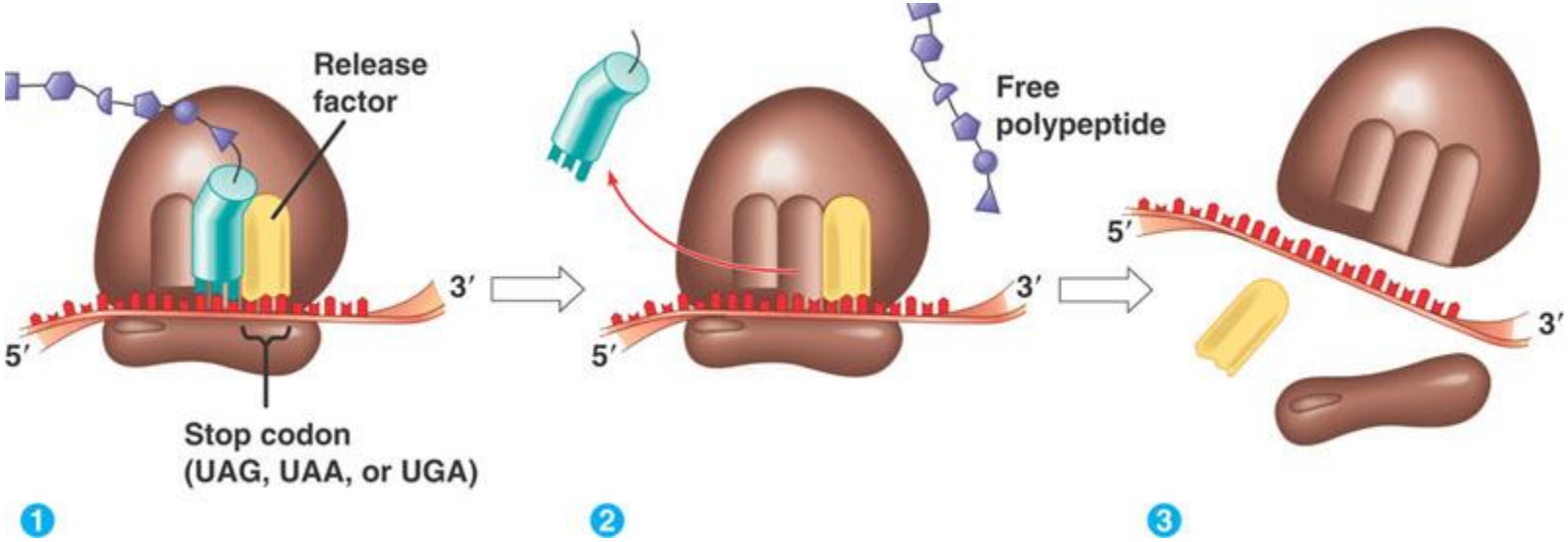
Translation – Initiation



Translation – Elongation



Translation – Termination



DNA to RNA to Protein

The following data from GenBank (accession No.AY785841) illustrate several points made in the preceding sections on transcription, the DNA code, and translation.

Reading Sequence Databases

```
gene            <1..>77
                /gene="CBF2A"
mRNA            <1..>772
                /gene="CBF2A"
                /product="HvCBF2A"
5'UTR           <1..12
                /gene="CBF2A"
CDS             13..678
                /gene="CBF2A"
                /note="HvCBF2A-Dt; AP2 domain CBF protein; putative CRT
                binding factor; monocot HvCBF4-subgroup member
                /codon_start=1
                /product="HvCBF2A"
                /protein_id="AAX23688.1"
                /db_xref="GI:60547429"
                /translation="MDTVAAWPQFEEQDYMTVWPEEQEYRTVWSEPPKRRAGRIKLQE
                TRHPVYRGVRRRGKVGQWVCELRVPVSRGYSRLWLGTFANPEMAARAHDSAALALSGH
                DACLNFADSAWRMMPVHATGSFRLAPAQEIKDAVAVALEVFQGQHPADACTAEESTTP
                ITSSDLSGLDDEHWIGGMDAGSYYASLAQGMLMEPPAAGGWREDDGEHDDGFNTSASL
                WSY"
3'UTR           679..>772
                /gene="CBF2A"
```

HvCBF2A DNA Code

ORIGIN

```
1       tagctgcgag ccatggacac agttgccgcc tggccgcagt ttgaggagca agactacatg
61      acggtgtggc cggaggagca ggagtaccgg acggtttggt cggagccgcc gaagcggcgg
121     gccggccgga tcaagttgca ggagacgcgc cacccggtgt accgcggcgt gcgacgccgt
181     ggcaaggtcg ggcagtgggt gtgcgagctg cgcgtccccg taagccgggg ttactccagg
241     ctctggctcg gcaccttcgc caaccccgag atggcggcgc gcgcgcacga ctccgccgcg
301     ctcgccctct ccggccatga tgcgtgcctc aacttcgccg actccgcctg gcggatgatg
361     cccgtccacg cgactgggtc gttcaggctc gcccccgcgc aagagatcaa ggacgccgtc
421     gccgtcgccc tcgaggtgtt ccaggggcag cacccagccg acgcgtgcac ggccgaggag
481     agcacgaccc ccatcacctc aagcgaccta tcggggctgg acgacgagca ctggatcggc
541     ggcatggacg ccgggtccta ctacgcgagc ttggcgcagg ggatgctcat ggagccgccg
601     gccgccggag ggtggcggga ggacgacggc gaacacgacg acggcttcaa cacgtccgcg
661     tcgctgtgga gctactagtt cgactgatca agcagtgtaa attattagag ttgtagtatc
721     agtagctagt actactagct gtgttcttcc accaggcgtc aggcctggca ag
```

5’ Untranslated Region (UTR)

Start Site (Methionine Codon

Stop Site Codon

3’ Untranslated Region (UTR)

HvCBF2A DNA Code Details

1. This sequence of 772 nucleotides encodes the gene HvCBF2A is from gDNA (genomic DNA) from the barley cultivar Dicktoo. Start reading the codons at nucleotide 1; the coding sequence starts at nucleotide 13 (codon = AUG = Met) and ends with nucleotide 678 (codon UAG = Stop).

2. When DNA base sequences are cited, by convention it is the sequence of the non-template (*sense, coding*) strand that is given, even though the RNA is transcribed from the template strand. The following Table shows highlighted sequences from the HvCBF2A gene and their interpretation.

More Code Details

Sequence	***Type***
5' atg gac aca.....tag 3’	Non-template DNA (decode replacing T with U)
3' tac ctg tgt.....atc 5'	Template DNA
5'aug gac aca.....uag3'	RNA (decode)
M D T Stop	Amino acid code (See Table)
Methionine, Aspartic acid, Threonine	Amino acid code (See Table)

Amino Acid Abbreviations

TABLE 3-1 The 20 Amino Acids Common in Living Organisms

Amino Acid	ABBREVIATION 3-Letter	ABBREVIATION 1-Letter	Amino Acid	ABBREVIATION 3-Letter	ABBREVIATION 1-Letter
Alanine	Ala	A	Leucine	Leu	L
Arginine	Arg	R	Lysine	Lys	K
Asparagine	Asn	N	Methionine	Met	M
Aspartic acid	Asp	D	Phenylalanine	Phe	F
Cysteine	Cys	C	Proline	Pro	P
Glutamic acid	Glu	E	Serine	Ser	S
Glutamine	Gln	Q	Threonine	Thr	T
Glycine	Gly	G	Tryptophan	Trp	W
Histidine	His	H	Tyrosine	Tyr	Y
Isoleucine	Ile	I	Valine	Val	V

Transcription, Translation, Phenotype

A. Allelic variation at the DNA sequence level: the fragrance in rice example

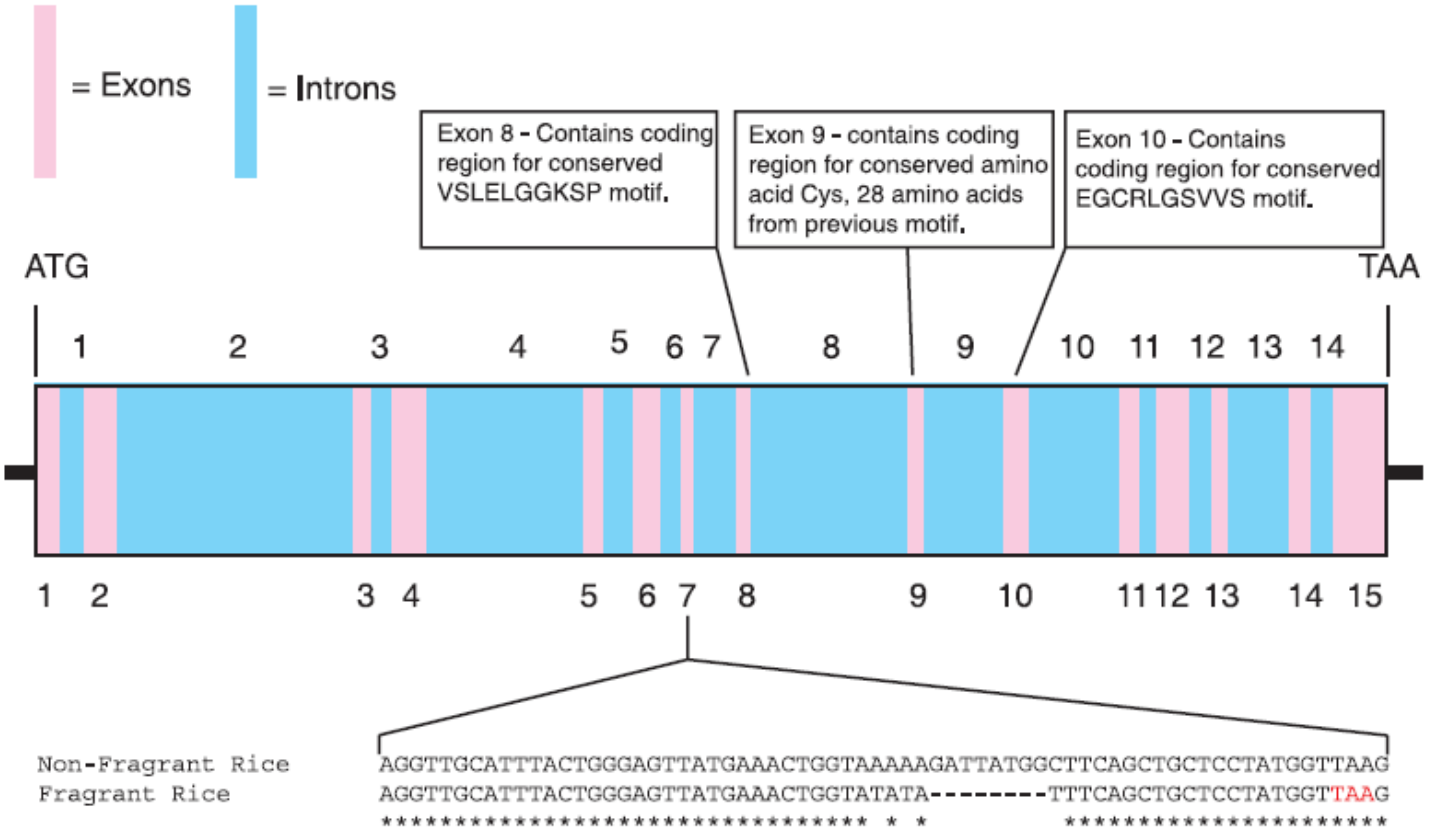


Figure 2 Structure of the fragrance gene (*fgr*) [Knowledge-based Oryza Molecular Biological Encyclopedia (KOME) ID: J023088C02] showing initiation codon (ATG), 15 exons, 14 introns and the ATT termination site. The nucleotide sequence of exon 7 is shown for both non-fragrant and fragrant rice varieties. The fragrant variety shows a large deletion and three single nucleotide polymorphisms (SNPs), and then terminates prematurely (stop codon in red), within this exon. The truncated protein encoded in fragrant rice varieties would therefore lack the highly conserved sequences encoded by exons 8, 9 and 10, which are believed to be critical for protein function.

Transcription, Translation, Phenotype

Allelic variation at the DNA sequence level: the fragrance in rice example

- Mutations are changes in sequence from wild type
- Can affect transcription, translation, and phenotype
  - An insertion/deletion event can produce a frameshift
  - Premature stop codon in frame, as in the rice example

Frameshift

*** CTGGGAGATTATGGCTTTAAG***
*** CTGGGA - - - - - - - - - -TAAG*** 11 bp deletion, alignment

*** CTG GGA GAT TAT GGC TTT AAG
*** CTG GGA TAA G codon alignment

Leu Gly Asp Tyr Gly Phe Lys
Leu Gly STOP G translation

Sequence Changes & Translation

Silent

*** CTG GGA GAT TAT GGC TTT AAG***
*** CTG GGA GAT TAT GGC TTC AAG*** alignment

Leu Gly Asp Tyr Gly Phe Lys
Leu Gly Asp Tyr Gly Phe Lys translation

Missense

*** CTG GGA GAT TAT GGC TTT AAG***
*** CTG GGA GAT TAT GGC TAT AAG*** alignment

Leu Gly Asp Tyr Gly Phe Lys
Leu Gly Asp Tyr Gly Tyr Lys translation

Nonsense

*** CTG GGA GAT TAT GGC TTT AAG***
*** CTG GGA GAT TAG GGC TTT AAG*** alignment

Leu Gly Asp Tyr Gly Phe Lys
Leu Gly Asp STOP translation

Transcription, Translation, Phenotype

Allelic variation at the DNA sequence level: the fragrance in rice example

- Mutations are changes in sequence from wild type
- Can affect transcription, translation, and phenotype
  - An insertion/deletion event can produce a frameshift
  - Premature stop codon in frame, as in the rice example

- Rice fragrance gene patenting - Basmati
- Rice fragrance gene patenting - Thailand

Thai Scientists Patent Rice Genes Responsible for Aromatic Rice

Scientist at the National Centre for Genetic Engineering and Biotechnology (BIOTEC) and Kasetsart University announced recently that they have successfully isolated and patented a group of genes found in rice that are responsible for producing aromatic rice. The discovery was made last year and the patent registered earlier this year.

Patent registration was secured first with the US Patent and Trademark Office. This is being followed by patent applications in another seven countries – specifically Australia, France, China, Vietnam, Japan, India and the Philippines. The leader of the Thailand Rice Genome Project, Dr. Apichart Wannavijitr, said yesterday that registration marked the first time Thai researchers had received patent protection for genetic material and showed that Thailand's capacity in biotechnology was now world class.

Dr. Morakot Tanticharoen

BIOTEC's director, Dr. Morakot Tanticharoen, said the team had spent several years studying the rice genome to isolate the genes in Thai jasmine rice that gave it its unique fragrance. The sequence of genes giving a rice plant a fragrant aroma is now not only known but can also be used to create aromatic qualities in ordinary, non-aromatic rice varieties, she said.

The researchers found that fragrant rice was the result of a genetic mutation in that the plant actually was the result of abnormal genes. In the case of Thai jasmine rice, eight genetic "characters" are not working.

The rice genome has about 50,000 genes. This discovery led researchers to conclude that if the same eight characters were stopped in other rice varieties, they would develop aromatic qualities, she said.

To prove their theory, laboratory tests were carried out on non-fragrant Japanese rice to successfully change it into fragrant rice.

With this technique, researchers can improve other rice breeds to produce a better fragrance. Dr. Morakot said, the technique could also be used to improve other crops, including wheat, corn, soybeans and coconuts,

Apichart said it was important for Thailand to register the patent so as to protect Thailand's unique natural heritage in producing Thai aromatic rice, which is world renown for its scent and flavor.

Patenting Native Genes?

(12) **United States Patent**
Vanavichit et al.

(10) **Patent No.: US 7,319,181 B2**
(45) **Date of Patent: Jan. 15, 2008**

(54) **TRANSGENIC RICE PLANTS WITH REDUCED EXPRESSION OF OS2AP AND ELEVATED LEVELS OF 2-ACETYL-1-PYRROLINE**

(75) Inventors: **Apichart Vanavichit**, Nakornpathom (TH); **Somvong Tragoonrung**, Pathumthani (TH); **Theerayut Toojinda**, Samutsakorn (TH); **Samart Wanchana**, Nakornpathom (TH); **Wintai Kamolsukyunyong**, Nakornpathom (TH)

(73) Assignee: **National Science & Technology Development Agency**, Klong Luang, Phathumthani (TH)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **11/043,520**

(22) Filed: **Jan. 25, 2005**

(65) **Prior Publication Data**

US 2006/0168679 A1 Jul. 27, 2006

(51) **Int. Cl.**
C12N 15/82 (2006.01)

(52) **U.S. Cl.** **800/285**; 800/286; 435/468

(58) **Field of Classification Search** ..... None
See application file for complete search history.

(56) **References Cited**

Colliver SP, Morris P, and Robbins MP. Differential modification of flavonoid and isoflavonoid biosynthesis with an antisense chalcone synthase construct in transgenic *Lotus corniculatus*. (1997) Plant Molecular Biology, vol. 35, pp. 509-522.*

Bradbury, Louis M.T. et al. (2005) "The gene for fragrance in rice" Plant Biotechnology Journal 3: 363-370.

Christopher, Mandy et al. (Aug. 2004) "Marker assisted selection in rice improvement" Rural Industries Research and Development Corporation: 15 pages.

Elsley, Kevin (Nov. 18, 2004) "Fragrant gene found" The Land: 53.

Garland, Stephen and Robert Henry (May 2001) "Application of Molecular Markers to Rice Breeding in Australia" Rural Industries Research and Development Corporation: 21 pages.

Kamolsukyunyong, W. et al. (Jul. 18, 2003) "Isogenic Lines Carrying the 26.7-kb Genomic Region on Chromosome 8 of KDML105 are Characterized as Aromatic Rice," Proceeding of the Conference on Rice Biotechnology, Peach, Pattaya, Thailand, pp. 151-155.

Kamolsukyunyong, Wintai et al. (2003) "Aroma Gene of Thai Hom Mali Rice," Lab Today, 9:66-69.

Nagsuk, Anujaree et al. (2003) "Identification of 2-Acetyl-1-Pyrroline, the Principal Aromatic Rice Flavor Compound, in Fungus Cultures," Proceedings of the 2nd International Conference on Medicinal Mushrooms & International Conference on Biodiversity and Bioactive Compounds, pp. 395-400.

Vanavichit, A. et al. (Aug. 31-Sep. 3, 2004) "Discovering Genes for Rice Grain Aroma," Proceedings of the 1st International Conference on Rice for the Future, Bangkok, Thailand, pp. 71-80.

Wanchana, S. et al. (Jul. 18, 2003) "Sequence Variation in BADH is Associated with Synthesis of 2AP, a Potent Aroma Determination in Rice," Proceedings of the Conference on Rice Biotechnology 2003, Peach, Pattaya, Thailand, pp. 157-160.

Wanchana, Samart et al. (2001) "Physical Mapping of the Region Proximal to Genes Controlling Aroma in Rice," RGJ-Ph.D. Congress II, Chonburi, Thailand, p. 153.

Wanchana, Samart et al. (2003) "Sequence Variation in BADH is

The Protein Code

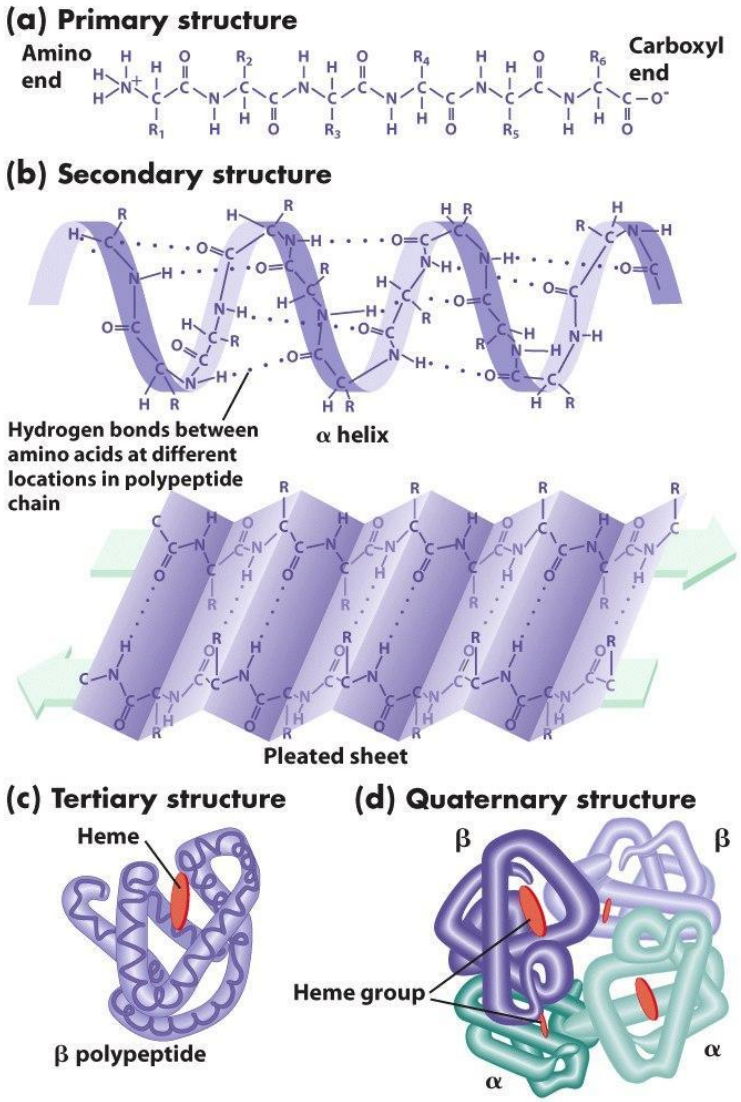
From gene to polypeptide: There are 20 common amino acids and these are abbreviated with three-letter and one-letter codes.

TABLE 3-1 The 20 Amino Acids Common in Living Organisms

Amino Acid	ABBREVIATION 3-Letter	ABBREVIATION 1-Letter	Amino Acid	ABBREVIATION 3-Letter	ABBREVIATION 1-Letter
Alanine	Ala	A	Leucine	Leu	L
Arginine	Arg	R	Lysine	Lys	K
Asparagine	Asn	N	Methionine	Met	M
Aspartic acid	Asp	D	Phenylalanine	Phe	F
Cysteine	Cys	C	Proline	Pro	P
Glutamic acid	Glu	E	Serine	Ser	S
Glutamine	Gln	Q	Threonine	Thr	T
Glycine	Gly	G	Tryptophan	Trp	W
Histidine	His	H	Tyrosine	Tyr	Y
Isoleucine	Ile	I	Valine	Val	V

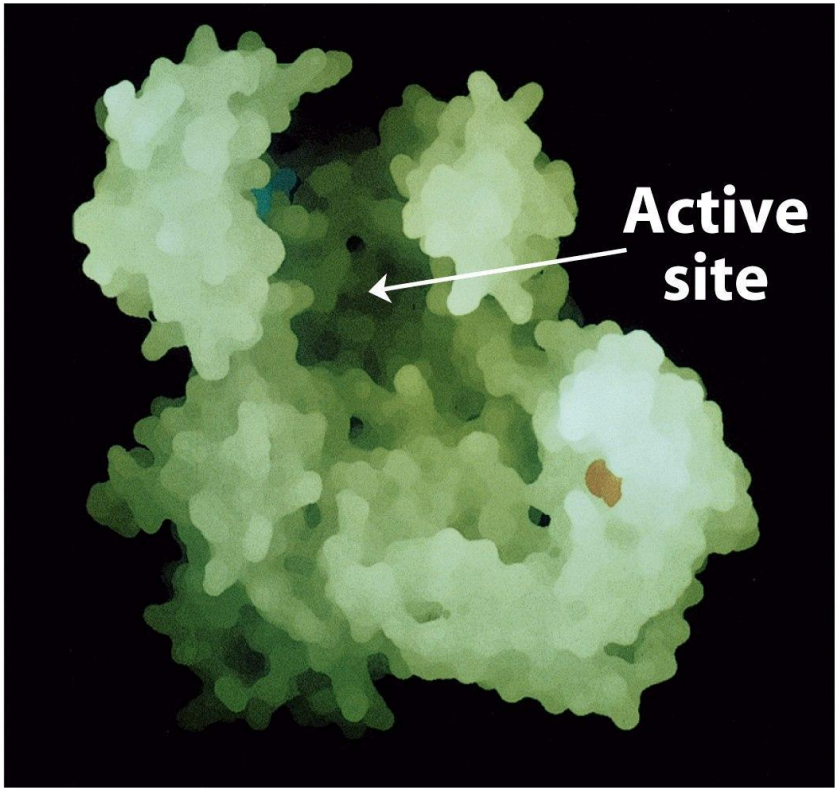
Protein Variation - Structure

Levels of protein structure: The primary, secondary, tertiary, and quaternary structures of protein.



Protein Variation - Function

- Functional - Enzymes (biological catalysts) have active sites
  - Change in site can give change in activity/function



Protein Variation - Structure

- Structural proteins can have tremendous economic and cultural value, e.g. wheat endosperm storage proteins. The same proteins can cause intense suffering in certain individuals - e.g. celiac disease

PROTEINS

Activation of Celiac Disease Immune System by Specific α-Gliadin Peptides

MARJA-LEENA LÄHDEAHO,[1,2] EEVA VAINIO,[3] MATTI LEHTINEN,[4] PÄIVI PARKKONEN,[1] JUKKA PARTANEN,[5] SAIJA KOSKIMIES,[5] and MARKKU MÄKI[6,7]

Cereal Chem. 72(5):475–479

ABSTRACT

Two different gliadin molecules (designated α-gliadin and α/β-gliadin) were synthesized as 52 and 58 ten amino acid (aa) long overlapping peptides for the determination of their B-cell epitopes. Monoclonal antibodies and human serum pools revealed two epitopes common for both gliadins (peptide 14 aa:s 66–75 and peptides 34α aa:s 166–175, 36α/β aa:s 176–185) and two unique epitopes (α-gliadin peptides 48 aa:s 236–245 and α/β-gliadin peptide 52 aa:s 256–275). In addition, peptide 9 (QPYPQPQPFP) aa:s 41–50 and peptide 42 (LGQGSFRPSQ) aa:s 206–215 were detectable by monoclonal antibodies and serum pools from patients with untreated celiac disease but not by serum pools from disease control patients who had antigliadin antibodies. Patients with celiac disease were also studied for their human leukocyte antigen (HLA) class II status (the presence of genetically determined proteins on antigen-presenting cells that are important for immunological recognition). Antigliadin antibody response to peptide QPYPQPQPFP was restricted by celiac disease (and HLA class II) because relative amounts of the antipeptide antibodies were significantly ($P < 0.05$) increased in celiac disease patients. The HLA alleles DQA1*0501 and DQB1*0201 are strongly associated with celiac disease. The difference between patients with celiac disease and healthy control subjects with regard to peptide QPYPQPQPFP suggest that this region in the gliadin molecule is of pathogenetic importance in celiac disease.

DNA to RNA to protein

Protein function and non-function: Changes in DNA coding sequence (mutations) can lead to changes in protein structure and function.

Proteomics: "If the genome represents the words in the dictionary, the proteome provides the definitions of those words".