GENE STRUCTURE, ORGANISATION, GENOMES

Dr Neda M Bogari
GENOMES AND NUMBER OF GENE

In human

- Between 80,000 to 100,000 gene arranged on 23 chromosomes.
- About 3000 genes per chromosome.
- Unique single copy genes:

- Most human genes are coding for polypeptides which are involved in or carry out a variety of cellular functions. These include enzymes, hormones, receptor and structural and regulatory proteins.
THE GENE

5' Promoter Ex1 In1 Ex2 In2 Ex3 In-- Ex--- 3'
1. **Exons**: are the functional portions of gene sequences that code for proteins.

2. **Introns**: are the noncoding sequences which separate the coding sequence (exons).

3. **The open reading frame**: a sequence with variable length that does not contain stop codons and therefore can be translated. The sequence beginning with **ATG** which exist at the 5' end of genes.
4. **TATA boxes**: These regions are about 20-30 bases to the 5' end (left) of the open reading frame (ATG). TATA boxes direct important enzymes to the correct initiation site for transcription.

5. **Termination codon**: the end of translation is signified by a termination codon at the 3' end of genes. The termination codon could be TAA, TAG, or TGA.
OTHER GENOMES

* In mouse 75% are similar (16% the intron shorter than human) and the gene are 80% proteins seem
* In chimp > between 98% to 95% similar (24 chromosomes), in the gene 99% are same
* In dogs 85% similar
* In Lily 40 time large
MULTIGENE FAMILIES

Many genes have similar functions making up what are known as multigene families.

Multigene families can be split into two types:

1- **Classical gene families** which show a high degree of sequence **homology**.

2- **Gene superfamilies** which have limited sequence **homology** but are functionally related.
The genetic code describes how base sequences are converted into amino acid sequences during protein synthesis.

The DNA sequence of gene is divided into a series of units of three bases.

Each set of three bases is called a codon and specifies a particular amino acid.

The four bases in DNA and RNA can combine as a total of $4^3 = 64$ codons which specify the 20 amino acid found in proteins.
**The Genetic Code**

<table>
<thead>
<tr>
<th></th>
<th>U</th>
<th>C</th>
<th>A</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>UUU, UUC, UUA, UUG</td>
<td>Phe, Leu</td>
<td>Ser, Tyr, Stop</td>
<td>Cys, Stop, Trp</td>
</tr>
<tr>
<td>C</td>
<td>CUU, CUC, CUA, CUG</td>
<td>Leu</td>
<td>Pro, His, Arg</td>
<td>U</td>
</tr>
<tr>
<td>A</td>
<td>AUU, AUC, AUA, AUG</td>
<td>Ile, ACU, ACC, ACA, ACC</td>
<td>Thr, Asn, Lys, Arg</td>
<td>U</td>
</tr>
<tr>
<td>G</td>
<td>GUU, GUC, GUA, GUG</td>
<td>Val, GCU, GCC, GCA, GCG</td>
<td>Ala, Glu</td>
<td>U</td>
</tr>
</tbody>
</table>

**Amino acid names:**
- Ala = alanine
- Arg = arginine
- Asn = asparagine
- Asp = aspartate
- Cys = cysteine
- Glu = glutamate
- Gln = glutamine
- His = histidine
- Ile = isoleucine
- Leu = leucine
- Lys = lysine
- Met = methionine
- Phe = phenylalanine
- Pro = proline
- Ser = serine
- Thr = threonine
- Trp = tryptophan
- Tyr = tyrosine
- Val = valine
**MITOCHONDRIAL CHROMOSOME**

- It is a circular, double-stranded DNA molecule.
- All genes in mitochondrial DNA have been defined (37 genes).
- 93% of the DNA sequence represent coding sequence.
- They are responsible for genes necessary for mitochondrial protein synthesis, and for proteins essential to oxidative phosphorylation.
THANKS