TRANSCRIPTOMICS

The study of the complete set of RNAs (transcriptome) encoded by the genome of a specific cell or organism at a specific time or under a specific set of conditions.
QUESTIONS

What is the relationship between transcriptome size and organismal complexity?

How do basic steps of gene expression contribute to transcriptome size?

What do new studies tell us about transcriptome size and complexity?

What are the roles of non-coding transcripts?
Is DNA content related to morphological complexity?

From: Essential Genes, Lewin
Minimum gene numbers range from 500 to 30,000

<table>
<thead>
<tr>
<th>Gene Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>500 genes</td>
<td>Intracellular (parasitic) bacterium</td>
</tr>
<tr>
<td>1,500 genes</td>
<td>Free-living bacterium</td>
</tr>
<tr>
<td>5,000 genes</td>
<td>Unicellular eukaryote</td>
</tr>
<tr>
<td>20,000 genes</td>
<td>Multicellular eukaryote</td>
</tr>
<tr>
<td>25,000 genes</td>
<td>Higher plants</td>
</tr>
<tr>
<td>22,000 genes</td>
<td>Mammals</td>
</tr>
</tbody>
</table>

From: *Essential Genes, Lewin*
Caenorhabditis elegans
959 somatic cells

Homo sapiens
100 trillion cells
*Caenorhabditis elegans*

- 959 somatic cells
- ~20,000 protein coding genes

*Homo sapiens*

- 100 trillion cells
- ~20,000 protein coding genes
BASIC STEPS OF GENE EXPRESSION

DNA synthesis (replication) → DNA → RNA synthesis (transcription) → mRNA (“transcripts”) → protein synthesis (translation) → PROTEIN → amino acids
Eukaryotic mRNA is modified, processed, and transported

1. **DNA**
2. **pre-mRNA**
3. **mRNA**
4. **protein**

**Transcription** → **Modification of ends** → **Splicing** → **Exons are joined** → **mRNA** → **Translation** → **protein**

- Cap at 5’ end
- Poly(A) at 3’ end
- Exon-intron junctions are broken

Copyright © 2006 Pearson Prentice Hall, Inc.
Alternative splicing adds to the size (complexity) of the transcriptome.

From: ncbi.nlm.nih.gov
RNA editing adds to the size (complexity) of the transcriptome.

*From: Maas Lab website*
The National Human Genome Research Institute (NHGRI) launched a public research consortium named ENCODE, the Encyclopedia Of DNA Elements, in September 2003, to carry out a project to identify all functional elements in the human genome sequence. The project started with two components - a pilot phase and a technology development phase.

The pilot phase tested and compared existing methods to rigorously analyze a defined portion of the human genome sequence. 1 percent (30 Mb) of the human genome was selected as the target for the pilot project. The full project was completed and published in 2012.
The human genome is pervasively transcribed. One stretch of DNA can encode multiple (m)RNAs.

The human genome is pervasively transcribed. DNA sequences far apart can be found on a single RNA ("fusion transcripts").

The human genome is pervasively transcribed. One stretch of DNA can encode an mRNA together with many other types of RNAs.
microRNAs are **FUNCTIONAL** gene products 21 – 22 nucleotides long.
• 1 microRNA can control the expression of many target mRNAs.

• microRNAs are important for development and cell- and tissue-specific gene expression.
**microRNAs are mis-regulated in human cancers.**

### Variations in microRNA genes linked to cancers

<table>
<thead>
<tr>
<th>Gene for:</th>
<th>miR-146a</th>
<th>miR-196a</th>
<th>miR-423</th>
<th>miR-27a</th>
<th>miR-492</th>
<th>miR-499</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung cancer</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Esophageal cancer</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast cancer</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Stomach cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liver cancer</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bladder cancer</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Papillary thyroid</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brain cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

This is important for cancer diagnosis.
microRNAs are involved in cancer metastasis.

High levels of microRNAs miR-103 and miR-107 are associated with poor cancer prognosis.

microRNAs have the potential to treat cancer (in mice).

microRNAs have the potential to **treat** cancer (in mice).

SOME ANSWERS

Transcriptome size scales with organismal complexity.

Alternative splicing and RNA editing contribute to transcriptome size.

The human genome is pervasively transcribed. One stretch of DNA can encode many different RNAs.

Many new functional species of RNAs have been identified (e.g. microRNAs).

microRNAs control the expression of many mRNAs and are involved in many cellular processes including development and disease, esp. cancer.
The need for a more RNA-centric viewpoint?
• Relationship of transcriptome size to morphological complexity

  • Transcriptome size scales with organismal complexity.

  • Alternative splicing and RNA editing contribute to transcriptome size.

  • The human genome is pervasively transcribed. One stretch of DNA can encode many different RNAs.

  • Many new species of RNAs have been identified.

  • microRNAs

    • They are 21 – 22 nucleotide long FUNCTIONAL GENE PRODUCTS.

    • They are processed from precursor RNAs by Drosha and Dicer.

    • They control the expression of MANY mRNAs.

    • They are involved in many cellular processes including development and disease.