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?

- Flowering plants
- Birds
- Mammals
- Reptiles
- Amphibians
- Bony fish
- Cartilaginous fish
- Echinoderms
- Crustaceans
- Insects
- Mollusks
- Worms
- Molds
- Algae
- Fungi
- Gram(+) bacteria
- Gram(−) bacteria
- Mycoplasma

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

- **500 genes**
  Intracellular (parasitic) bacterium

- **1,500 genes**
  Free-living bacterium

- **5,000 genes**
  Unicellular eukaryote

- **20,000 genes**
  Multicellular eukaryote

- **25,000 genes**
  Higher plants

- **22,000 genes**
  Mammals

From: *Essential Genes*, Lewin
Caenorhabditis elegans

Homo sapiens
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
Eukaryotic mRNA is modified, processed, and transported

Transcription

Modification of ends
- Cap at 5’ end
- Poly(A) at 3’ end

Splicing
- Exon-intron junctions are broken

Exons are joined

Transport
- Nucleus
- Cytoplasm

Translation

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

<table>
<thead>
<tr>
<th>Variations in microRNA genes linked to cancers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gene for:</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Lung cancer</td>
</tr>
<tr>
<td>Esophageal cancer</td>
</tr>
<tr>
<td>Breast cancer</td>
</tr>
<tr>
<td>Stomach cancer</td>
</tr>
<tr>
<td>Liver cancer</td>
</tr>
<tr>
<td>Bladder cancer</td>
</tr>
<tr>
<td>Papillary thyroid</td>
</tr>
<tr>
<td>Brain cancer</td>
</tr>
</tbody>
</table>

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

microRNAs are involved in cancer metastasis.

microRNAs can be used to treat liver cancer (in mice).

microRNAs can be used to treat liver cancer (in mice).

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?