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?

<table>
<thead>
<tr>
<th>Organism</th>
<th>Base Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flowering plants</td>
<td>10^6-10^11</td>
</tr>
<tr>
<td>Birds</td>
<td>10^6-10^7</td>
</tr>
<tr>
<td>Mammals</td>
<td>10^7-10^8</td>
</tr>
<tr>
<td>Reptiles</td>
<td>10^7-10^8</td>
</tr>
<tr>
<td>Amphibians</td>
<td>10^7-10^9</td>
</tr>
<tr>
<td>Bony fish</td>
<td>10^7-10^9</td>
</tr>
<tr>
<td>Cartilaginous fish</td>
<td>10^7-10^9</td>
</tr>
<tr>
<td>Echinoderms</td>
<td>10^7-10^9</td>
</tr>
<tr>
<td>Crustaceans</td>
<td>10^7-10^9</td>
</tr>
<tr>
<td>Insects</td>
<td>10^7-10^9</td>
</tr>
<tr>
<td>Mollusks</td>
<td>10^8-10^10</td>
</tr>
<tr>
<td>Worms</td>
<td>10^8-10^10</td>
</tr>
<tr>
<td>Molds</td>
<td>10^9-10^10</td>
</tr>
<tr>
<td>Algae</td>
<td>10^9-10^10</td>
</tr>
<tr>
<td>Fungi</td>
<td>10^9-10^10</td>
</tr>
<tr>
<td>Gram(+) bacteria</td>
<td>10^9-10^10</td>
</tr>
<tr>
<td>Gram(-) bacteria</td>
<td>10^9-10^10</td>
</tr>
<tr>
<td>Mycoplasma</td>
<td>10^9-10^10</td>
</tr>
</tbody>
</table>

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

blogs.ajc.com
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
DNA synthesis (replication)

nucleotides

RNA synthesis (transcription)

mRNA

protein synthesis (translation)

PROTEIN

amino acids

“transcripts”
Eukaryotic mRNA is modified, processed, and transported

Transcription

Modification of ends

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

Splicing

Exons are joined

A₂₀₀

Transport

Nucleus

Cytoplasm

Translation

A₂₀₀

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.
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.
rRNAs

tRNAs

snRNAs

snoRNAs

microRNAs

gRNAs

lincRNAs

casiRNAs

natsiRNAs

pasRNAs

tRNAs

endo-siRNAs

microRNAs

gRNAs

snRNAs

sRNAs

tasRNAs

casiRNAs
microRNAs are **FUNCTIONAL** gene products 21 – 22 nucleotides long.

From: discovermagazine.com
FUNCTIONAL microRNAs are cut out of longer RNA molecules by an enzyme called DICER.
• 1 microRNA can control the expression of many target mRNAs.

• microRNAs are important for development and cell- and tissue-specific gene expression.

![Diagram showing temporal and spatial switches controlled by microRNAs](image-url)
**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>X</td>
<td></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>X</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>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bladder cancer</td>
<td></td>
<td></td>
<td>X</td>
<td></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.

microRNAs are involved in cancer metastasis.

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

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?