The RNA revolution: rewriting the fundamentals of genetics

Overview

1. Genetics 101
2. Recent lessons from genomics
3. The expanding world of noncoding RNAs
4. Long noncoding RNAs – an area of controversy
5. I’m a clinician - is any of this relevant to me?
Genetics 101

DNA: our genetic blueprint
Proteins: the building blocks of our body

- Hair and Nails: A protein called collagen forms your hair and fingernails and gives the skin its elasticity.
- Blood: Hemoglobin proteins carry oxygen in your blood to every part of your body.
- Muscles: Muscle proteins called myosin and actin enable all muscular movement—from smiling to throwing a baseball.
- Brain and Nerves: Neuron proteins control signals initiated by the small molecules life and death of nerve cells.
- Cellular Messengers: Receptor proteins allow the reception of your body and transmit signals to other proteins or the inside of the cells.
- Enzymes: Enzymes in your saliva, stomach, and small intestine are proteins that help you digest food.
- Antibodies: Antigens are proteins that help defend your body against foreign substances.
- Cellular Construction Workers: Major cellular molecules from molecular machines that do your work, such as working genes, making ribosomes and making new proteins.

Image from http://publications.nigms.nih.gov/

Our genes encode the instructions to make proteins

Gene A from Person 1
- Ala
- Arg
- Asp
- Asn
- Cys

Gene A from Person 2
- Ala
- Lys
- Asp
- Asa
- Cys

Codon change resulted in a different amino acid at position 2

Image adapted from http://www.jgi.doe.gov/
RNA: the middle-man of genetics

The central dogma in action
Recent lessons from genomics

The human genome has been sequenced
Our genes represent only \(\sim 2\%\) of the genome. That’s fewer than 20 DNA bases per 1000!

Is the rest all junk DNA?
Our genes are not much different from those of mice

Moreover, >99% of our protein-coding genes have homologs in mice

Data from http://www.ensembl.org

The transcriptome may unlock the genome’s secrets
There seems to be all this RNA that doesn’t code for protein.
Summary #1

- Genes function by encoding proteins
- Our genes represent only ~2% of the genome and are similar to those of other animals
- >90% of the RNA in our cells isn’t there to make proteins

The expanding world of noncoding RNAs
Noncoding RNAs (ncRNAs) everywhere!

RNA successfully combines the best features of both DNA and protein

- base-pairs with DNA/RNA in a highly specific fashion

- forms 3D structures allowing direct interaction with proteins and even catalysis
Discovery of hundreds of microRNAs
miRNAs regulate the production of thousands of our proteins

Short interfering RNAs turn off genes via RNA interference
But wait, there’s more! snoRNAs, piRNAs, PASRs, TASRs …
The Transcriptional Landscape of the Mammalian Genome

The FANTOM Consortium\textsuperscript{\textregistered} and RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group)\textsuperscript{\textregistered}

- 20,714 genes encoding proteins
- 22,839 long noncoding RNAs

What’s with all these long ncRNAs?

- Accidental by-products of transcription?
- A vast functional network of regulatory genes?
Tissue-specific expression of long ncRNAs


Long ncRNAs are co-expressed in developmentally-related tissues

A search for functional long ncRNAs

- Embryonic stem cell differentiation (0 - 16 days)
  - 2 cell embryo
  - 4 cell embryo
  - Blastocyst
  - Embryoid body

- T-cell differentiation and activation
  - Unactivated T-cells
  - Activated T-cells
  - Naive T-cells
  - Memory T-cells

- Male and female gonadal ridge development (10.5 - 13.5 dpc)
  - 10.5dpc embryo
  - 13.5dpc male/female gonad dissections

A search for functional long ncRNAs

- C2C12 Myoblast differentiation
  - Myoblasts (confluent)
  - Myotubes (5d)
  - Confluent 50% 24h 5d

- Macrophage activation
  - Unactivated (-LPS)
  - Activated (+LPS)

- Neural stem cell development
  - Neural Stem Cells (NSC)
  - Oligodendrocytes (OL)
  - Myelinating OL
  - GABAN
  - Myel OL
  - OLP
  - NOP
OK, so what did we find?

- >1500 long ncRNAs differentially expressed in at least one of the experimental systems
  - these represent potential candidates for further functional studies

Some long ncRNAs might be involved in early embryonic stem (ES) cell function
Some long ncRNAs overlap important genes

Some long ncRNAs might exert epigenetic effects
Summary #2

- There are lots of different types of noncoding RNAs in our cells, most of which have only recently been identified.

- Many ncRNAs are small, and seem to have important regulatory roles in cells (e.g. by switching off genes).

- Many ncRNAs are long, and at least some of them are likely to be functional.

I’m a clinician - is any of this relevant to me?
Our genes are not much different from those of mice

Moreover, >99% of our protein-coding genes have homologs in mice

Data from http://www.ensembl.org

Same pieces, different configuration
Same materials, different configuration

Same proteins, different configuration
Junk DNA – is this where the architectural plans lie?

Noncoding RNAs and disease?
Mutations in the RNA Component of RNase MRP Cause a Pleiotropic Human Disease, Cartilage-Hair Hypoplasia

Maaret Ridiarpaa,1 Hans van Eenennaam,2 Katarina Pellet,1,2 Robert Chatr-Aria,1 Cheryl Johnson,4 Bo Yuan,4 Walther van Veenreij,3 Ger Pruij,3 Rika Salmela,1,2 Susanna Rockas,1,2 Ondrej Miktic,1,2 Ilkka Kaatila,1,2 and Albert de la Chapelle1,2

Prader-Willi phenotype caused by paternal deficiency for the H112-85 C/D box small nucleolar RNA cluster

Trilochan Saboo1, Daniela del Guercio1, Jennifer R German1, Marwan Shnawi1, Serika U Peters1, Richard E Person1, Adolfo Garnica1, Sau Wei Cheung1 & Arthur L Beaudet1

Prader-Willi syndrome (PWS) is caused by deficiency for one or more paternally expressed imprinted transcripts within chromosome 15q11-13, including SNURF-SNRPN and multiple small nucleolar RNAs (snRNAs). Balanced chromosomal translocations that preserve expression of SNURF-SNRPN and centromeric genes but separate the snRNA HBII-85 cluster from its promoter cause PWS. A microdeletion of the HBII-85 snRNAs in a child with PWS provides, in combination with previous data, effectively conclusive evidence that deficiency of HBII-85 snRNAs causes the key characteristics of the PWS phenotype, although some atypical features suggest that other genes in the region may make more subtle phenotypic contributions.
MicroRNA Expression Profiles Associated With Prognosis and Therapeutic Outcome in Colon Adenocarcinoma

Conclusions. Expression patterns of microRNAs are systematically altered in colon adenocarcinomas. High miR-21 expression is associated with poor survival and poor therapeutic outcome.

RNA interference – a new therapeutic frontier?

Modulation of HIV-1 replication by RNA interference

Jean-Marc Jacque, Karine Triques & Mario Stevenson

Program in Molecular Medicine, University of Massachusetts Medical School, 333 Plantation Street, Worcester, Massachusetts 01605, USA

RNA interference (RNAi) is the process by which double-stranded RNA (dsRNA) directs sequence-specific degradation of messenger RNA in animal and plant cells. In mammalian cells, RNAi can be triggered by 21-nucleotide duplexes of small interfering RNA (siRNA). Here we describe inhibition of early and late steps of HIV-1 replication in human cell lines and primary lymphocytes by siRNAs targeted to various regions of the HIV-1 genome. We demonstrate that synthetic siRNA duplexes or plasmid-derived siRNAs inhibit HIV-1 infection by specifically degrading genomic HIV-1 RNA, thereby preventing formation of viral complementary-DNA intermediates. These results demonstrate the utility of RNAi for modulating the HIV replication cycle and provide evidence that genomic HIV-1 RNA, as it exists within a nucleoprotein reverse-transcription complex, is amenable to siRNA-mediated degradation.
Summary #3

• ncRNAs might represent the architectural plans that tell our bodies how to configure our genes

• ncRNAs are increasingly being implicated in disease

• ncRNAs represent a novel strategy to help treat disease

Acknowledgements

Institute for Molecular Bioscience, Brisbane
  John Mattick
  Marcel Dinger
  Tim Mercer
  Paolo Amaral
  Andrew Perkins
  Sean Grimmond
  Mark Crowe

Ludwig Institute for Cancer Research, Melbourne
  Weisan Chen

Funding bodies
  NHMRC
  Australian-American Fulbright Commission
  Menzies Foundation
  RACP
  CSL
Questions?