Lecture 25

- Signs and observations that led to the discovery regulatory RNAs
  - Phenotypes associated with mutations in genes without an ORF
  - Extensive sequencing of RNA after depletion of ribosomal associated RNAs
  - CLIP cross-linking immunoprecipitation
  - Led to the discovery of piwi interacting RNA and circular RNA
- Anti-sense inhibition in bacterial gene regulation, can be utilized to knock-out (specific) gene expression in organisms lacking homologous recombination by simply introducing a transgene with an inverted, anti-sense sequence of the targeted endogenous gene (sense) → sense and anti-sense → RNA duplex, unstable and not-translatable
- Another way of introducing the anti-sense RNA into model organisms such as nematodes would be through injections of ss anti-sense RNAs, but this experiment needs a few controls to test for the procedure itself:
  - Water
  - Anti-sense for another gene from another unrelated organism
- Weird observation was analyzed when injecting ss sense RNA, which had the same phenotype as the above
- Fire/ Mello experiments
  - Ss RNA (sense and anti) versus → ds 100x more efficient
  - Test for three different dsRNA targeting different genes → different phenotypes → gene specific targeting
  - At the level of DNA, transcription or RNA, translation → dsRNA for promoter and intronic regions not found in the processed transcript → no associated phenotype → conclude that it is post-transcriptional regulatory mechanism leading to the loss of target mRNA
  - It is called RNAi
- Seen also in plants, example, CHS gene encoding an enzyme that lead to pigment accumulation in petals. A transgene of the CHS gene with a highly expressed promoter was introduced → observed white segments that indicate silencing gene expression from both the transgene and the endogenous gene → co-suppression due to:
  - Random integration of the transgene
  - Convergent transcription leading to the formation of a dsRNA
  - PTGS
- 16: populations of siRNA are also found in worms, insects and protozoan → indicating gene regulation by RNAi

John: This slide shows that RNAi is associated with the presence of small RNAs. RNAi is induced by siRNAs, not miRNAs, although in the latter case, small RNAs are also associated with gene silencing.

- The experiment was done to isolate these 25nt siRNA through Northern Blot:
  - Isolate RNA
  - Run through poly-acrylamide gel electrophoresis
  - Transfer to a membrane
  - Use sense and anti-sense probe
- The presence of the dsRNA is the trigger for RNAi, and the effector molecules would be these ssRNA called siRNA molecules
- RNAi involves transcript degradation
- Another class of RNA molecules involved in RNAi is called microRNA or miRNA found in C. elegans
- C. elegans have internal fertilization and the early stages of embryogenesis occur in the adult body before laying
- Lin mutants → affect the pattern/ timing of cell division and differentiation during development:
  - Lin4 mutant: delay in developmental progression, the adult does not lay eggs due to the lack of the egg laying apparatus vulva defects

  - Lin14 mutant: precocious, early pre-mature developmental progression, adult body structures at larval stages
- this type of mutants is known as heterochronic: mutation affecting the developmental timing/ temporal M
- in normal vulva development, the stability of high levels of lin-14 mRNA/ transcripts, while the decay in its protein levels after L1 stage indicates that there is a post-transcriptional regulatory mechanism blocking its translation, and guess who is that one that is blocking its translation, its L4!
  - Lin14 → protein required in the L1 stage of the vulva formation
  - Lin4 → encodes a small RNA that block the translation of lin14 gene after L1 stage → leading to reduction in the lin-14 protein
  - More about that mechanism: Lin4 encodes 2 transcripts (long 60nt and a short 22nt derived from the stem-loop transcript). The 22nt lin-4 transcript is partially complementary to multiple site in the 3' UTR of the lin4 transcript
  - This lin4 is transcribed from L2-L4 stages (so it's a small temporal regulatory RNA)
  - Conclude that it is a post-transcriptional and trans regulation process, involving the inhibition of protein synthesis
- Two examples, but the majority of the small RNAs are not temporally regulated (rare) called microRNAs
  - Present in all multicellular eukaryotes
  - Regulate more that 70% of the protein coding genes
  - Conserved, orthologs or miRNA are found across human, fish, nematodes, and flies
  - However, they are very different in plants suggesting independent origin of evolution
- Primary miRNA (pri-miRNA) endogenous and can be found in different regions in the genome:
  - Most lack an ORF
  - Most arise from splicing from non-protein coding region
  - Less commonly, they can be located in protein coding mRNA (in introns)

<u>miRNA biogenesis includes:</u>

- transcription, nuclear processing, nuclear export and cytoplasmic processing
- this is one-one end product processing

Transcription:

- Transcribed by RNA pol II into a ssRNA (less than 70nt) → processed to have a 5' cap and a 3' poly-A tail → special fold back regions → intramolecular association/ interaction due to reverse sequence complementarity → **<u>imperfect</u>** hairpin stem-loop structure

Nuclear processing:

- The hairpin stem-loop structure is recognized by Dorsha (type !! RNA endonuclease in the nucleus) , which bind to the <u>primary miRNA</u> at the stem and snips it → <u>precursor miRNA</u> (with 3' overhang and 5' phosphate)

Nuclear transport

- Precursor miRNA is actively transported into the cytoplasm

Cytoplasmic processing

- Dicer recognizes the transported precursor miRNA and also binds to the stem → removes the loop and trims stem to form miRNA duplex (we have a villa now: p), one cleavage → 21-22nt with 2nt 3' overhang

---

siRNA biogenesis

- Can be made up of viruses, endogenous repetitive elements and some transgenes (external dsRNA)
- Long and perfectly paired dsRNA precursor
- Dorsha is not needed in here, as all the processing of the long dsRNA takes place in the cytoplasm by Dicer
- Also end up with multiple 21-25 nt with 2nt 3'overhang and 5'phosphate
- Then one stand of the ds duplex is incorporated into a cytoplasmic complex (with multi-component ribonucleoprotein complex) called RISC, always containing Argonaute protein (AGO)
  - Passenger RNA degraded or destructed
  - Guide miRNA bind to the target sequence through the RISK complex through base-pairing with complementary sequences → gene silencing
- Gene silencing through the RISC complex can be either by:

- Target sequence degradation → siRNA
- Translation/ protein synthesis blockage → micro-RNA/ miRNA

Lecture 26 Outline

- Gene silencing pathways:
  - miRNA: blockage in protein synthesis by binding to the target transcript, deadenylating and decay
    one miRNA effector/ precursor molecules from a single primary miRNA transcript
    hairpin stem-loop structure, imperfect pairing of the fold back regions
  - siRNA: degradation of the target transcript
    multiple siRNA effector/ precursor molecules from a single primary siRNA transcript
    long almost perfectly paired duplex and get chopped into smaller precursors
    (is the duplex here is a result of the pairing between to transcripts or folding within the same?)
- precursor molecules are processed in the cytoplasm by Dicer, which then binds to Argonaute and gets incorporated into one of the two RISC complex:

1) RISC that slice target transcript
   - Extensive complementarity and nearly perfect pairing between the si/miRISC and the target site
   - Target sites can be in 5', 3' or the coding sequence of the processed transcript
   - Seen in: most siRNA, some plants miRNA and few animal miRNA
   - Repression or gene silencing here is though endonucleolytic cleavage of the target mRNA
   - Slicing occurs at a site complementary to residues 10-11 of the guide RNA strand
   - Cleaved mRNA is then degraded

2) RISC that block protein synthesis
   - Complementarity to short repetitive sequences almost always in the 3'UTR
   - The number of sites and their efficiency/ complementarity → extent of repression
   - Imperfect pairing between the guide RNA strand/ miRISC and the target mRNA
   - Repression or gene silencing here is though blockage in protein synthesis
   - Often lead to de-adenylation of the tail → decay of the target mRNA
   - Seen most animal miRNA
   - When miRISC binds to its target mRNA it recruits:
     a) PABPC-interacting proteins
     b) De-adenylase complex → also recruits factors that prevent the formation of pre-initiation complex S43
        Both cause: shortening of the poly-A tail, de-capping and eventually degradation of the target mRNA

	RISC that slice	RISC that block protein synthesis
Complementarity	Extensive, nearly perfect pairing	Imperfect pairing
Target sites can be	3', 5' or coding sequence of the processed transcript	Almost always in the 3'UTR, multiple repetitive sequences
Repression/ gene silencing	endonucleolytic cleavage of the target mRNA → degradation	is though blockage in protein synthesis Or de-adenylation → target mRNA decay
Mechanism	Slicing at residues 10-11 of the complementary guide RNA strand	The number of sites and their efficiency/ complementarity → extent of repression
Seen	In most siRNA, most plant miRNA and only few animal miRNA	Most animal miRNA

- Note: the outcome of de-adenylation will be dependent on the developmental stage:

- Embryonic: affect protein synthesis
- Post-embryonic: affect mRNA stability

Gene silencing → miRNA RNAi→ siRNA

miRNA in plants vs. animals

	Plants miRNA	Animal miRNA
Complementarity	Near perfect 18-20nt	Strict complementarity in the Seed site/ region with 7nt match, central bulge and moderate complementary site (13-16nt)
Location of target sites in the processed transcript	Coding sequences (typically)	In the 3' UTR (typically), based in *lin-4 and let-7* targets
Target genes	TF involved in patterning, cell differentiation and nutrient homeostasis	TF involved in differentiation, physiological pathways and maintenance of tissue identity in zebrafish
Approaches for identification	1) Clone miRNA and invert 2) Align to genome sequence and look for sequence with 4 or less mismatch	1) Identify related genes (from genomic sequences) and align 3'UTR 2) Search within that for occurrence of conserved seed match sites 3) Search for 3'UTR with two or more seed math sites
Further comments		Bulge prevent RISC from cleaving its target transcript
Other General observations	– One miRNA target one gene or closely related members of a gene family	– One miRNA can have many unrelated targets – Single mRNA/ gene can have target sites for multiple miRNA with different seeds – 50% of human protein coding genes showed conservation of the miRNA sites in their 3'UTR → indicating an essential function

Final comments:

- Note: although the given example of lin-4 and let-7 are temporally regulated miRNA, most animal miRNA are spatially regulated → tend to be tissue specific and showing high degree of specificity and are remarkably diverse
- miRNA that are not involved in early development or embryogenesis are often have a role in differentiation and maintenance of tissue identity rather than developmental patterning

---

Regulatory roles of miRNA

➢ Genetic Switch: Completely suppress gene activity
- Temporal and spatial regulation of gene activity
- Target gene promoter is active in the 2 tissues of study
- miRNA is co-expressed with its target gene in the same tissue → preventing the target mRNA from being translated → shutting off target protein levels (spatial regulation)
- another example in hematopoiesis in zebrafish, miR451 → gata2, used in situ hybridization to analyse their expression pattern
- showed that the miRNA eliminates the target transcript through mRNA decay