

Long Noncoding RNA

The Dark Matter of the Genome

Megan McSweeney

BMS 265

Long Noncoding RNA

LncRNA is defined as all transcribed RNA molecules greater than 200 nt in length.

The 200 nt cutoff is based on RNA isolation protocols.

Very broad category, so lncRNAs are a very large and functionally heterogeneous group.

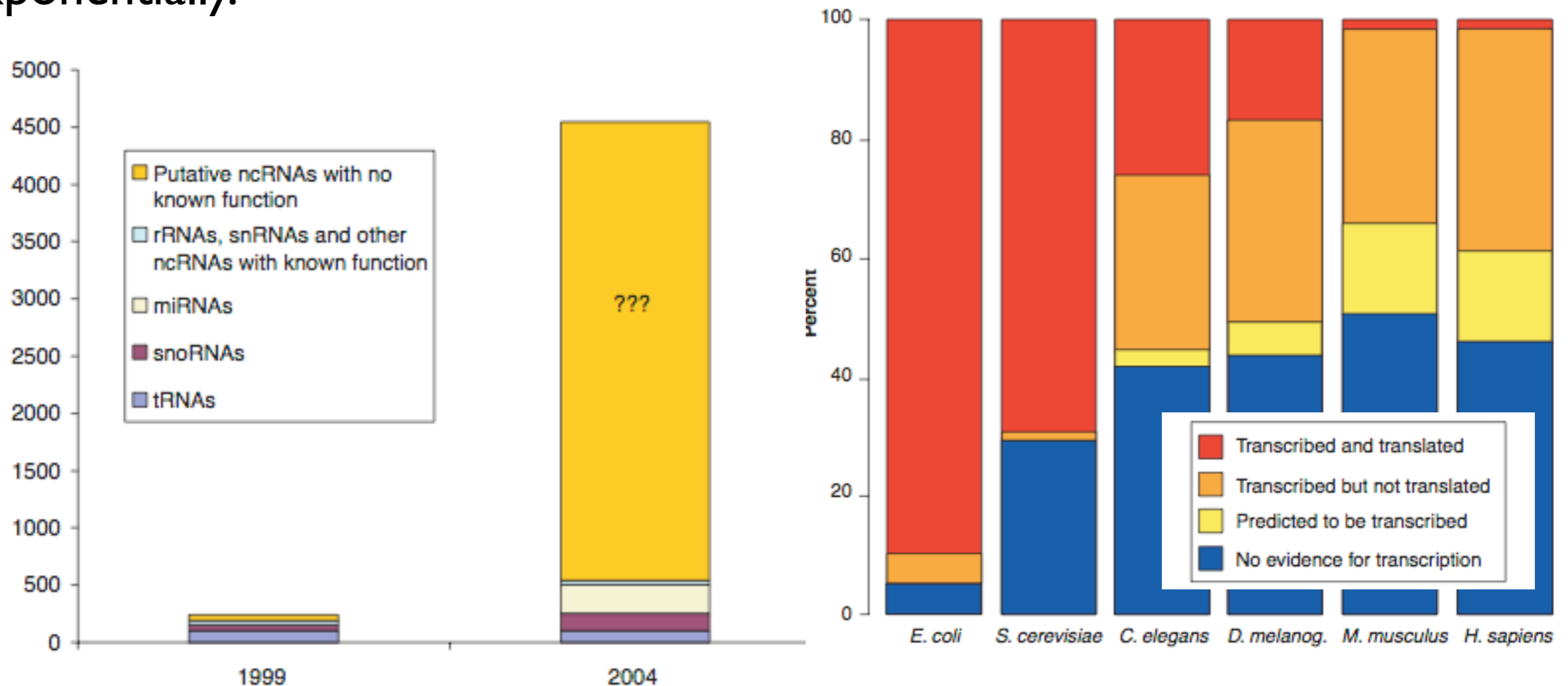
Mouse transcriptome: ~180,000 total transcripts, ~20,000 protein coding genes. The majority of the remaining 160,000 transcripts are lncRNAs

Discovery

1980s-1990s: Individual lncRNAs are discovered through traditional gene mapping approaches, *Xist* and *H19*.

Early 2000s: Development of large scale cDNA sequencing leads to the discovery of a surprising number of lncRNA transcripts.

Mid 2000s: The number of predicted genes in the mammalian genome goes down and the number of detected lncRNA transcripts increases exponentially.



Evolution and Conservation: Are lncRNAs Functional?

Transcriptional Noise

- Low affinity binding of RNA polymerase to randomly generated promoter sequences.
- More efficient to allow random transcripts than to downregulate nonspecific transcription.
- lncRNAs are generally expressed at low levels
- lncRNA sequences are not well conserved between species.

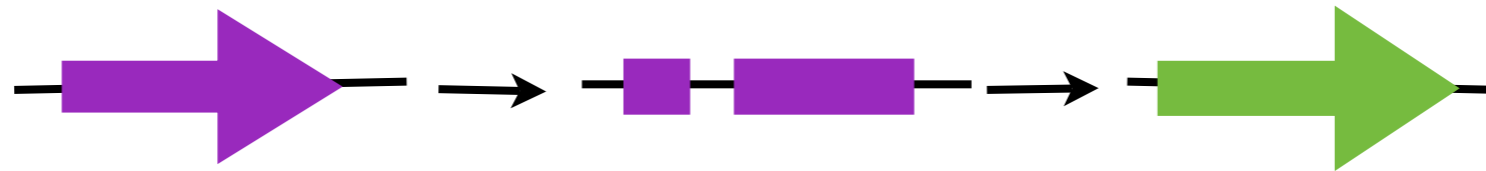
lncRNAs are Functional

- lncRNAs do not have the strict sequence conservation restraints that protein-coding genes do.
- lncRNAs may be more plastic than protein coding genes and thus can evolve rapidly.
- lncRNA promoter sequences are very well conserved.

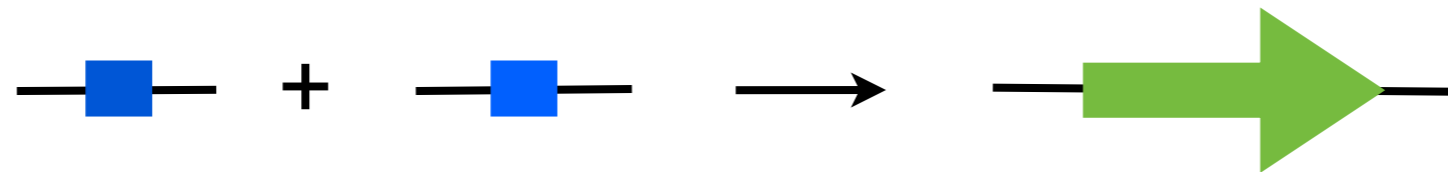
Origins of lncRNA Sequences

lncRNAs can be categorized based on their proximity to protein coding genes

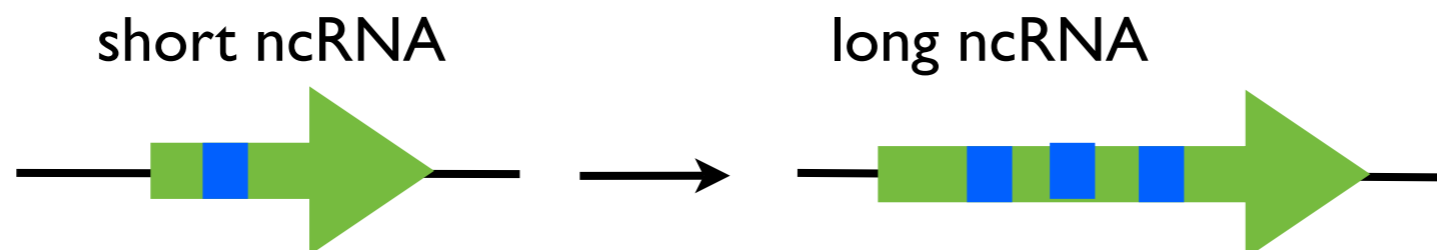
Mutation of a Protein Coding Gene



Chromosomal Rearrangement



Duplications



Transposon Insertion



Legend:

Green = lncRNA,
Purple = Protein
Coding Gene

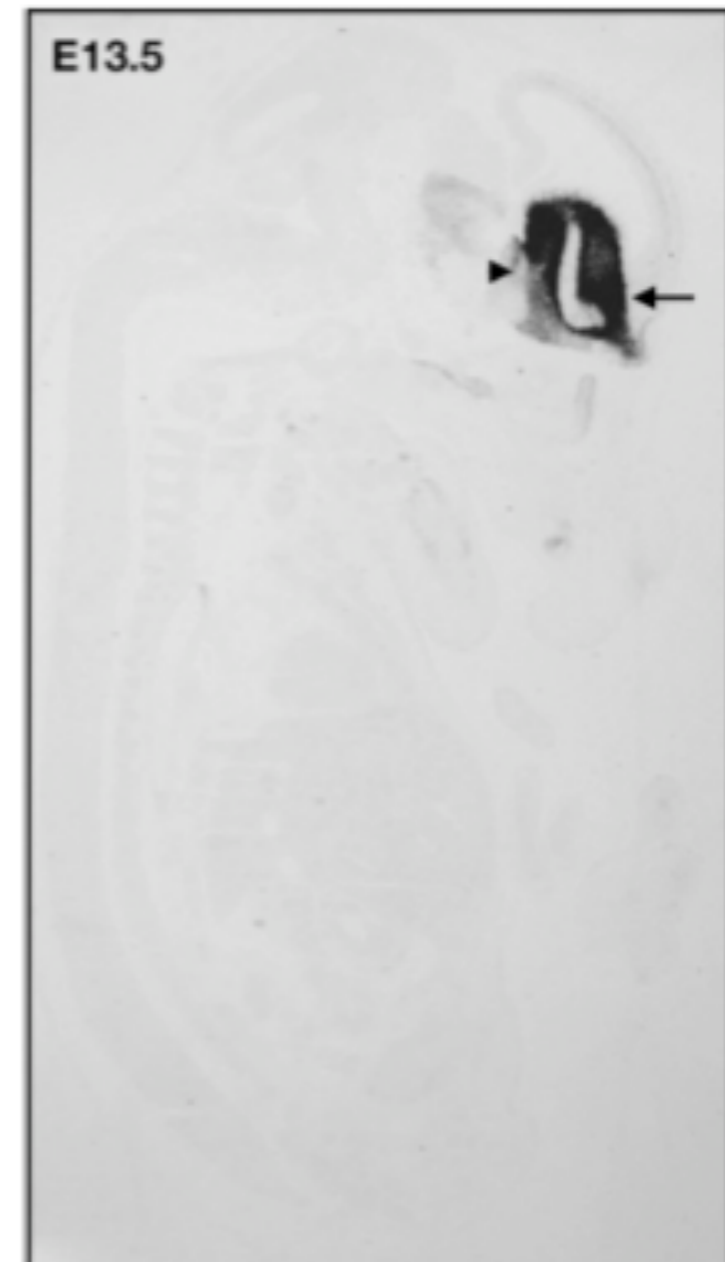
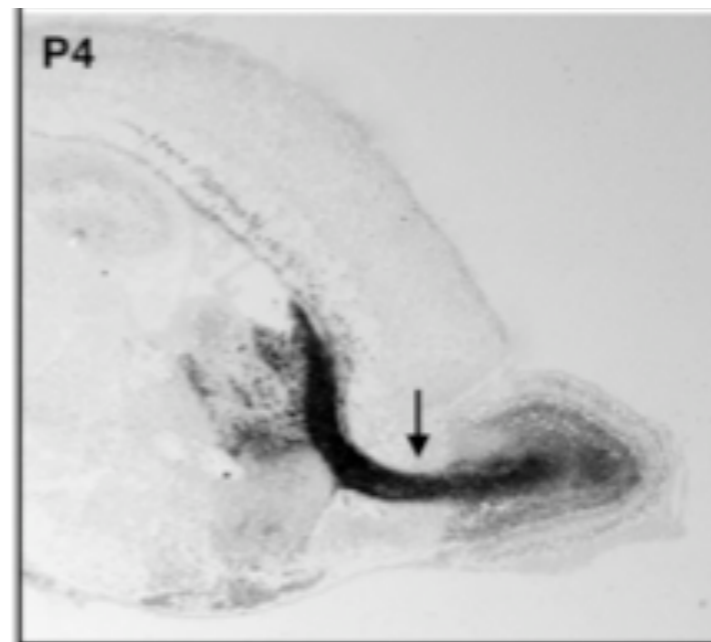
Biogenesis: Regulation of lncRNA Expression

The regulation of lncRNA expression has also been closely studied to determine if lncRNAs are functional

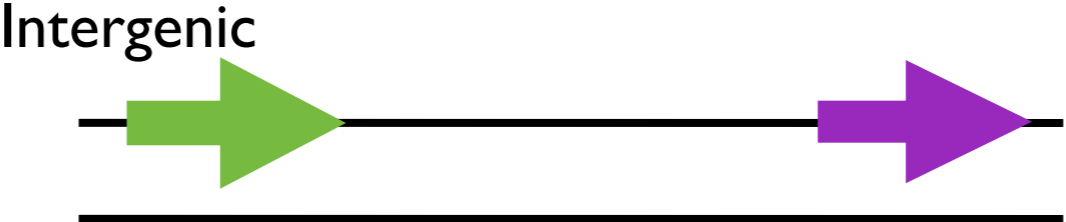
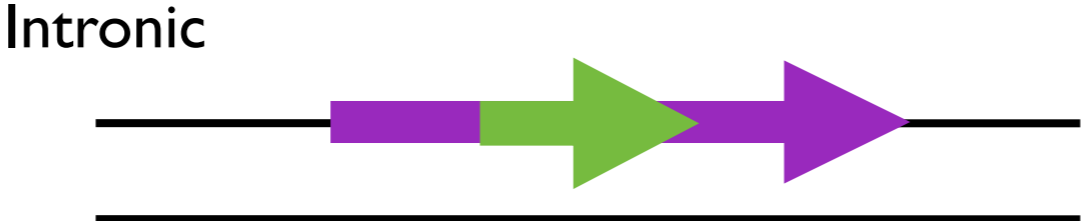
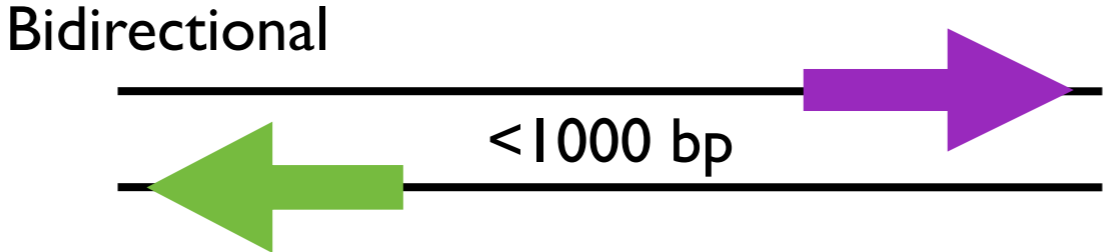
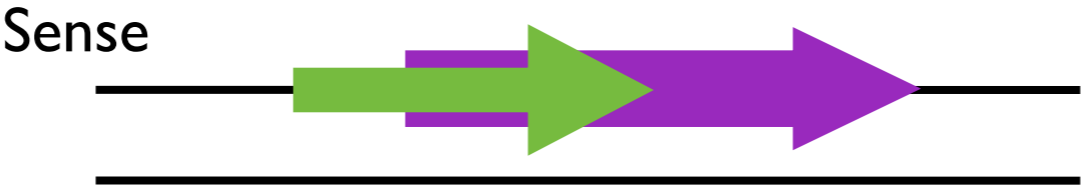
Evidence of Function: lncRNA expression levels vary with location, time, and physiological stimuli.

There are examples of all of these types of regulation.

lncRNA Evf2 in
developing mouse
brain
in situ hybridization



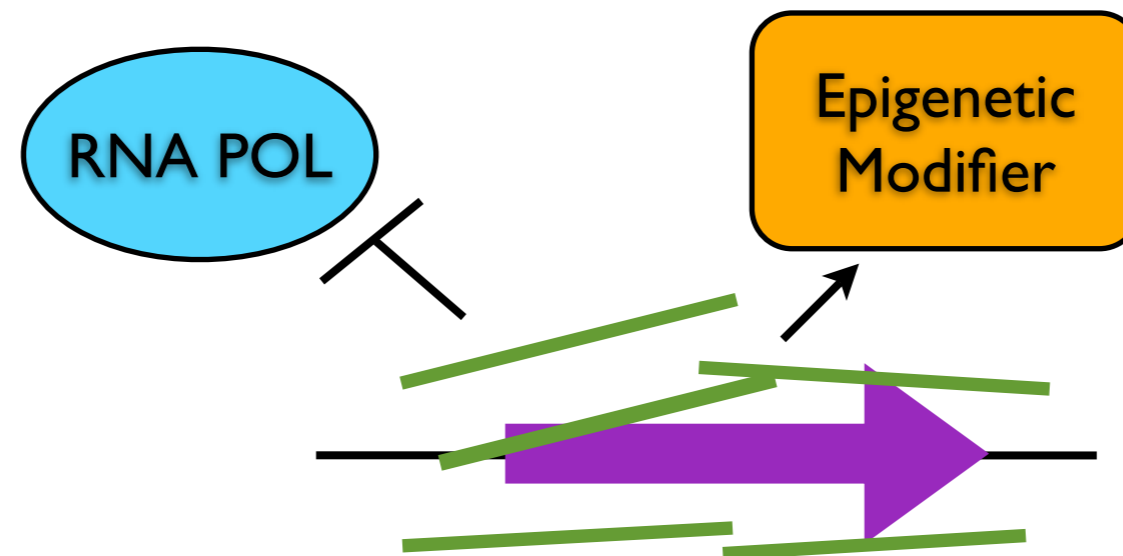
Biogenesis: Where are lncRNAs?



Legend:
Green = lncRNA,
Purple = Protein
Coding Gene

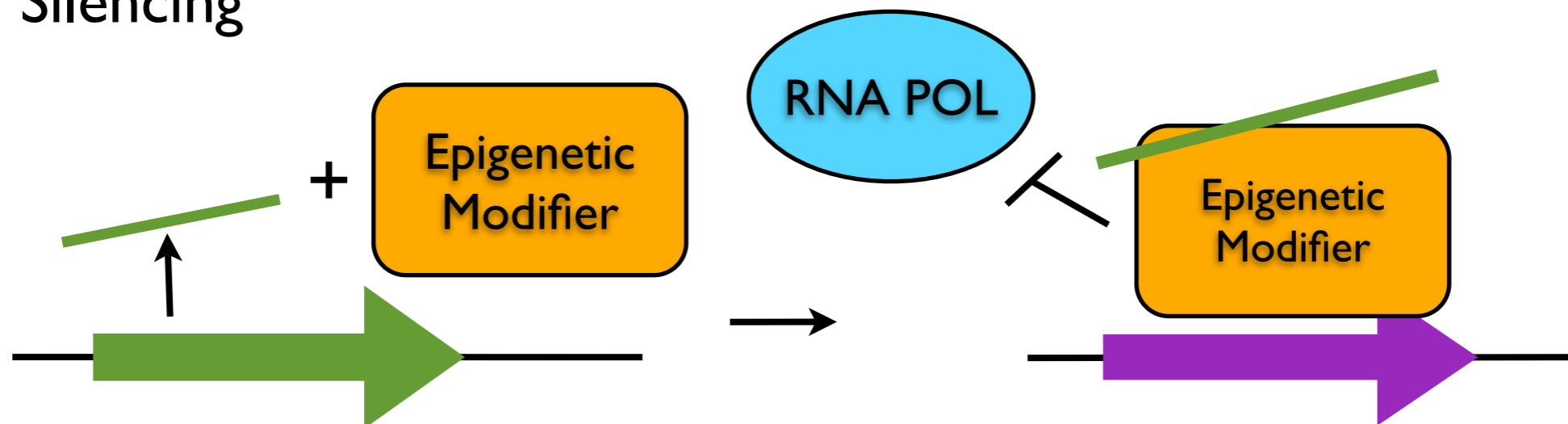
Mechanism of Action: Regulation of Chromatin Structure

Cis Silencing



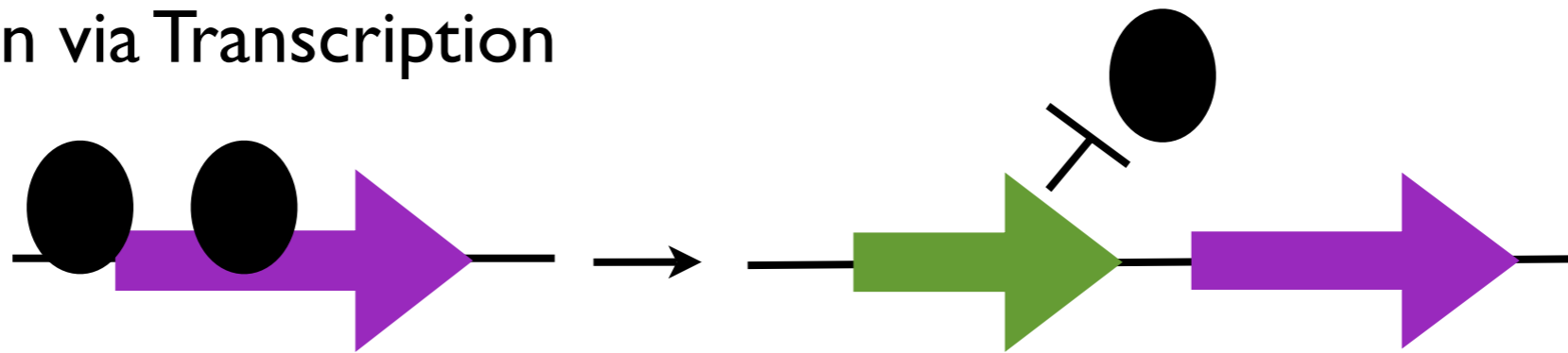
Legend:
Green = lncRNA,
Purple = Protein
Coding Gene

Trans Silencing

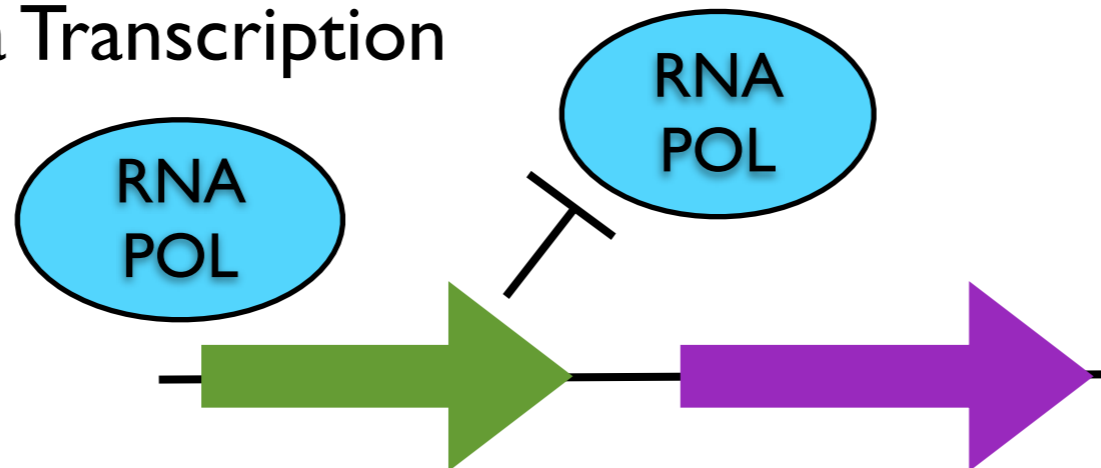


Mechanism of Action: Gene Regulation by lncRNA Transcription

Activation via Transcription

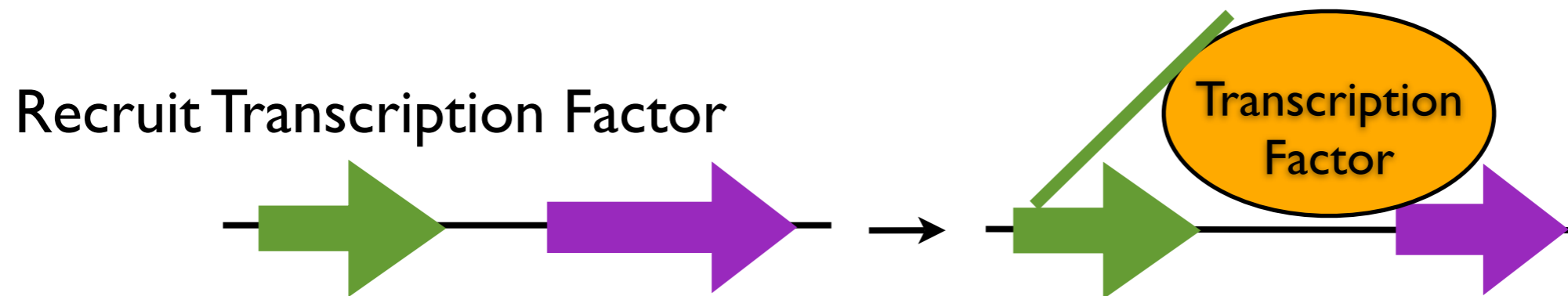
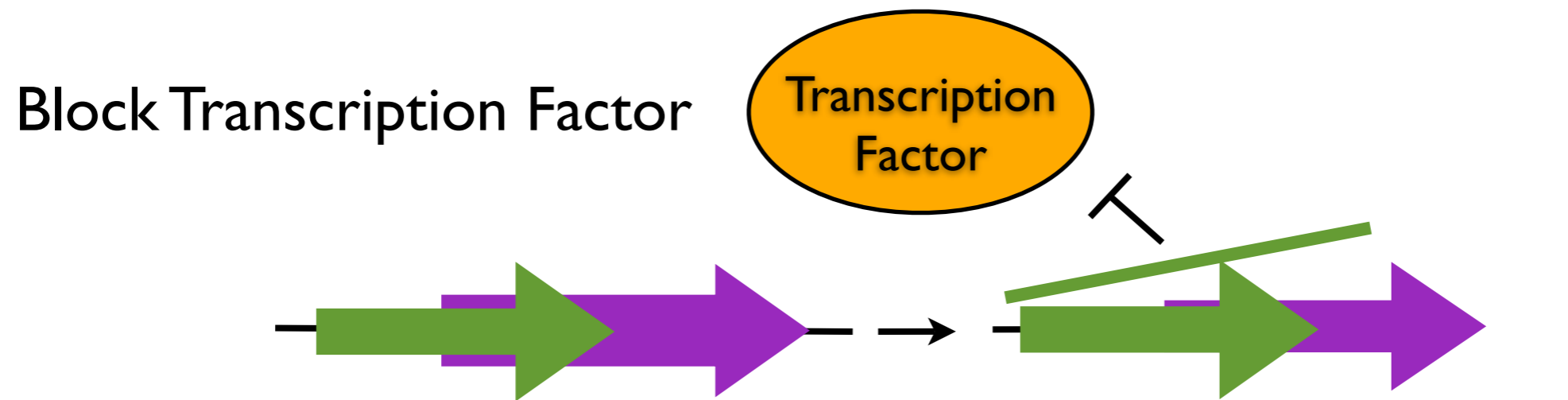


Silencing via Transcription



Legend:
Green = lncRNA,
Purple = Protein
Coding Gene

Mechanism of Action: Regulation of Transcription in cis



Legend:
Green = lncRNA,
Purple = Protein
Coding Gene

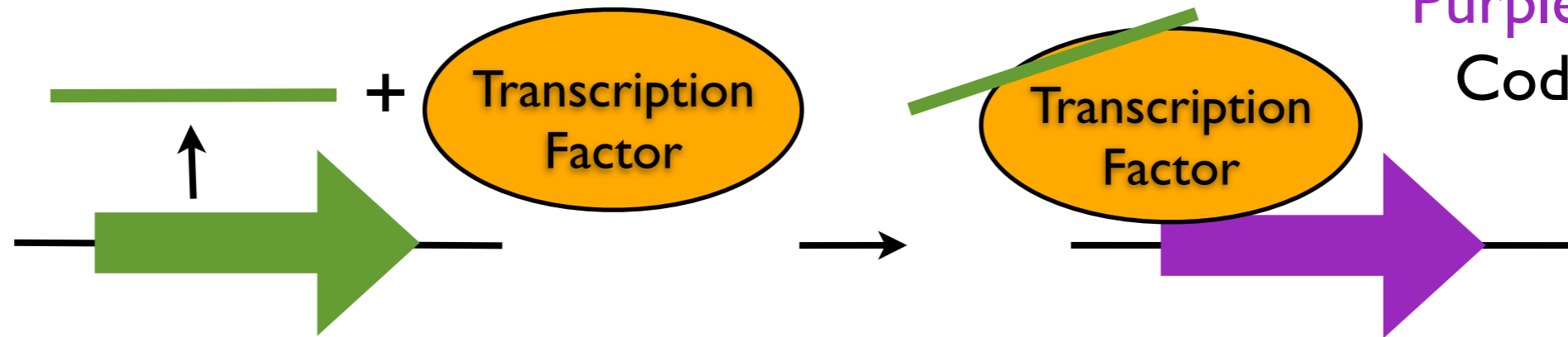
Mechanism of Action: Regulation of Transcription in trans

Legend:

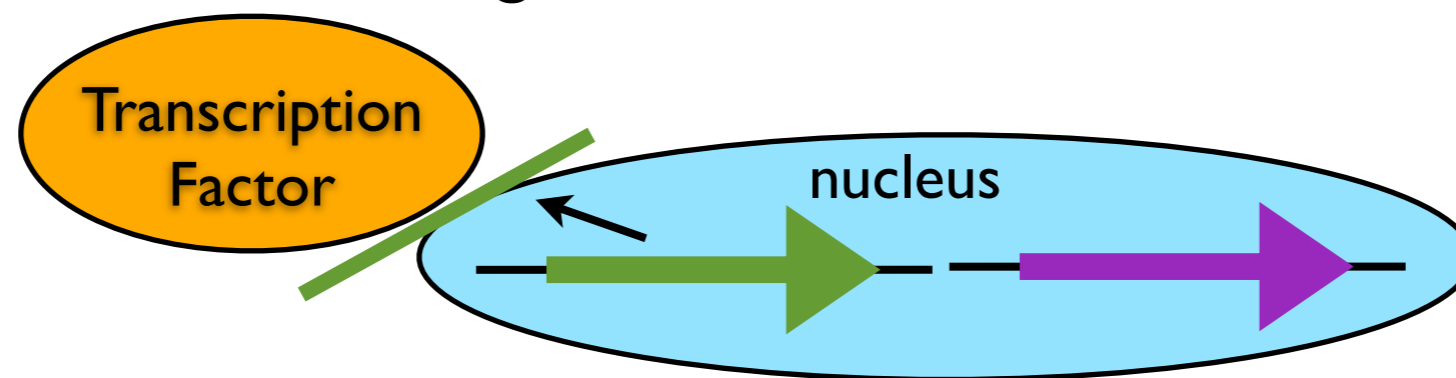
Green = lncRNA,

Purple = Protein
Coding Gene

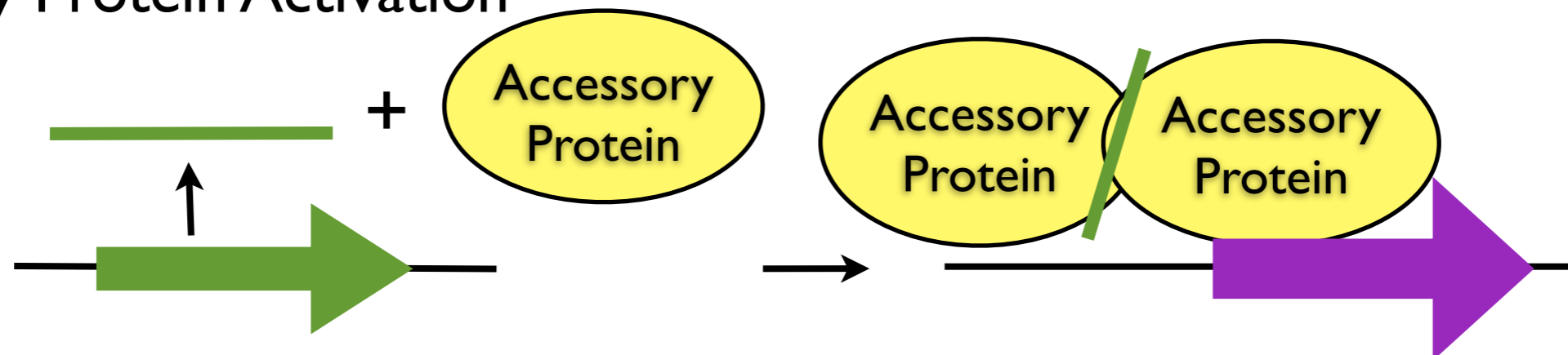
Transcription Factor Activation



Transcription Factor Trafficking

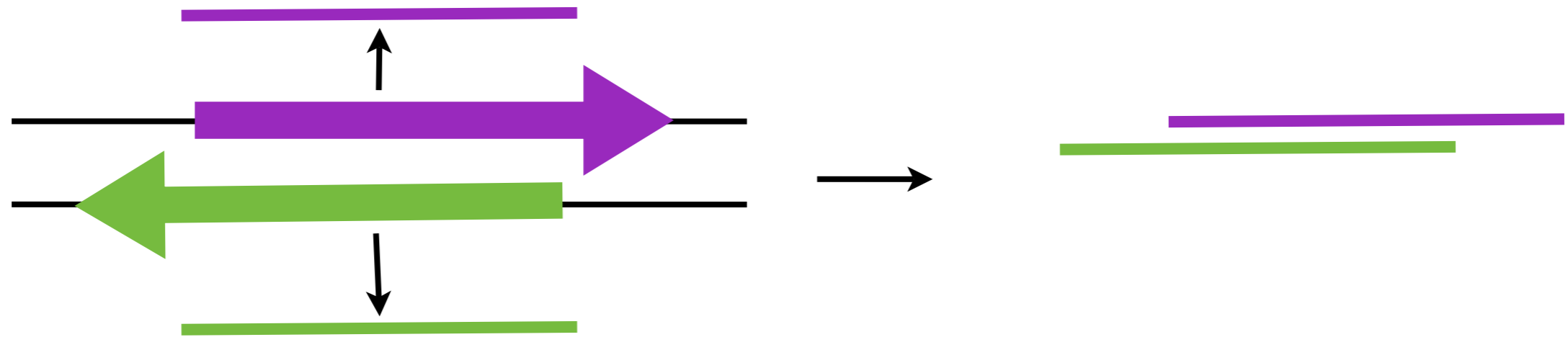


Accessory Protein Activation

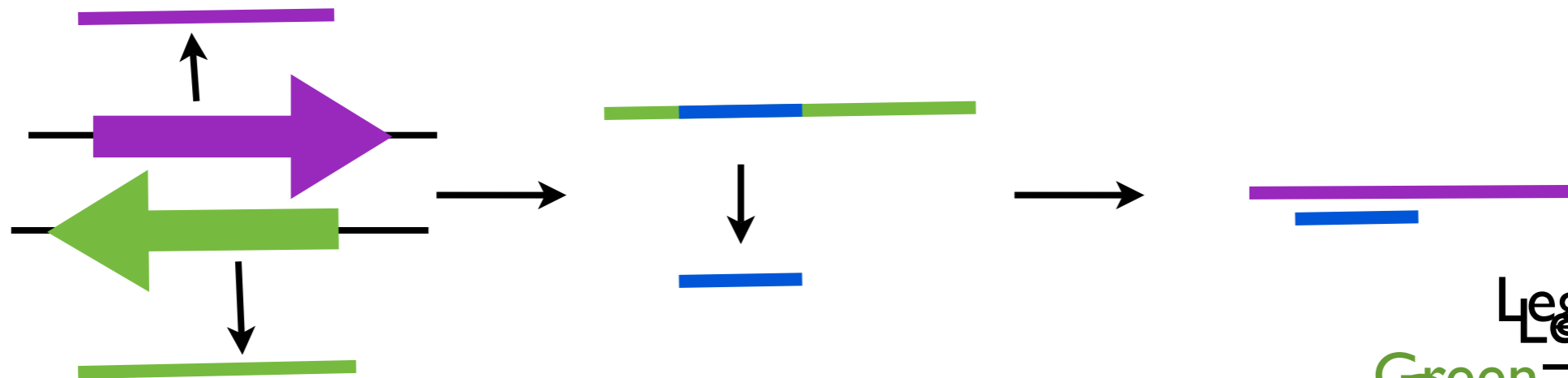


Mechanism of Action: Post Transcriptional Regulation

Regulation of mRNA Transcripts



Gene Silencing Pathway



Legend:
Green = lncRNA,
Purple = Protein
Coding Gene

Diseases

- **Cancer**

LncRNA expression profiles may be useful as biomarkers for cancer diagnosis.

- **Alpha-Thalassaemia**

An inherited form of this type of anemia is caused by the translocation of of an antisense lncRNA near the alpha-globin gene, resulting in the epigenetic silencing of the HBA2 gene and causing the disease.

- **Alzheimer's Disease**

The beta-secretase-1 (BACE1) protein is a crucial enzyme in the progression of Alzheimer's disease. The expression level of the antisense transcript of the BACE1 gene increases in response to cell stressors such as amyloid-beta 1-42.

- **Human Spinocerebellar Ataxia Type 8**

Patients with SCA8 have been shown to have a trinucleotide expansion in an lncRNA named ataxin 8 opposite strand (ATXN8OS) antisense to the KLHL1 gene. Transgenic mice with this repeat expansion show a similar progressive neurological phenotype to humans with SCA8.

- **Myotonic Dystrophies**

Repeat expansions in lncRNAs are involved in multiple organ system myotonic dystrophies. These mutations alter the lncRNA secondary structures and prevent splicing regulators from interacting normally with pre-mRNAs.