‘Glass Half Full or Empty’: Illuminating the Human Transcriptome

Lecture 1

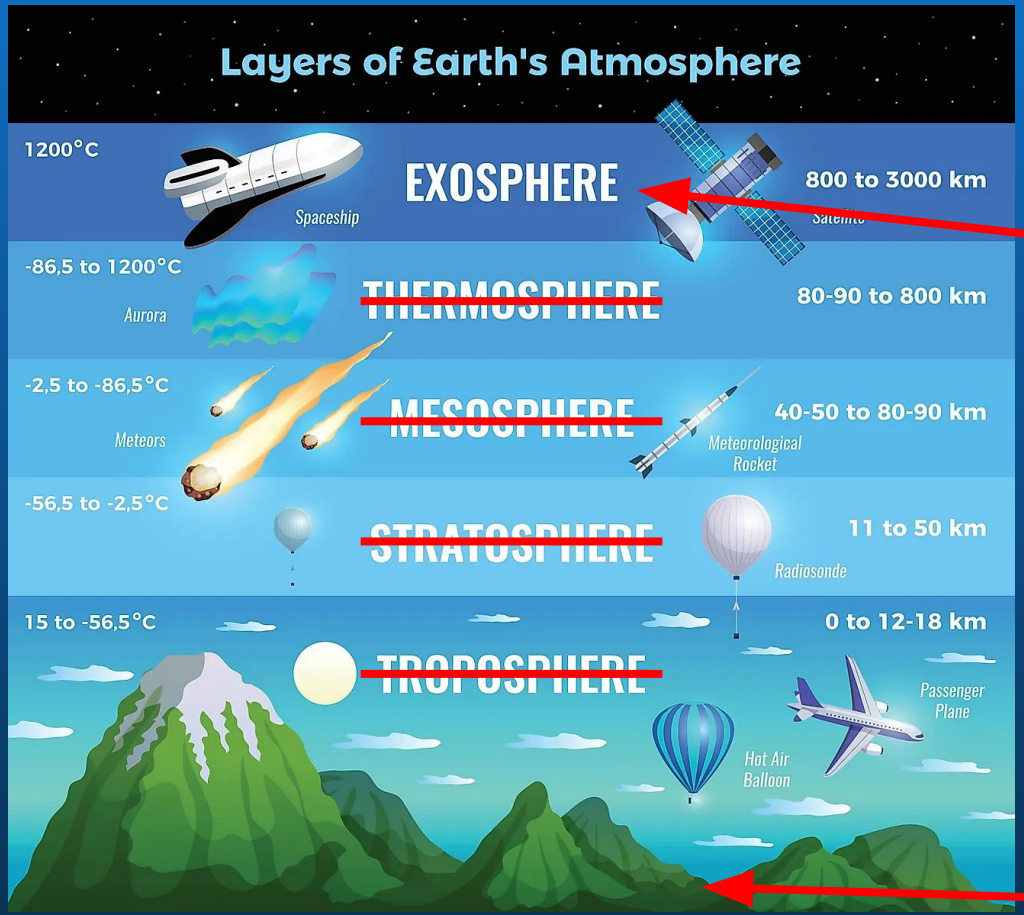
What is the world?

What is the world?

What is the world?

What is the world?

What is the world?



This course focuses on two specific layers within biology

Why Biology?

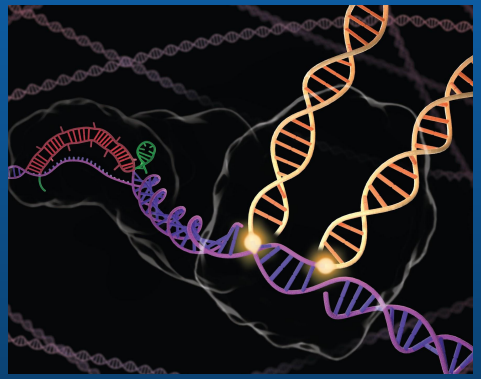
Biology = Life

Biology + Medicine = More Life

Why Biology?

Biology = Life

Biology + Medicine = More Life



Rewriting the code of life with CRISPR

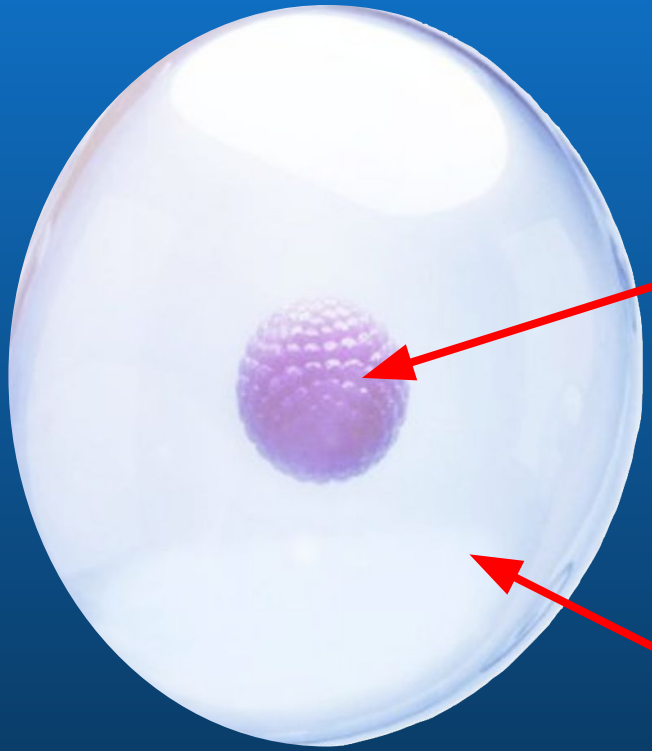
Experimental Gene Therapy Reverses Sickle Cell Disease for Years

December 13, 2021

The World of Biology

The Cell

The World of Biology



Nucleus => DNA

RNA

Protein

Cytoplasm

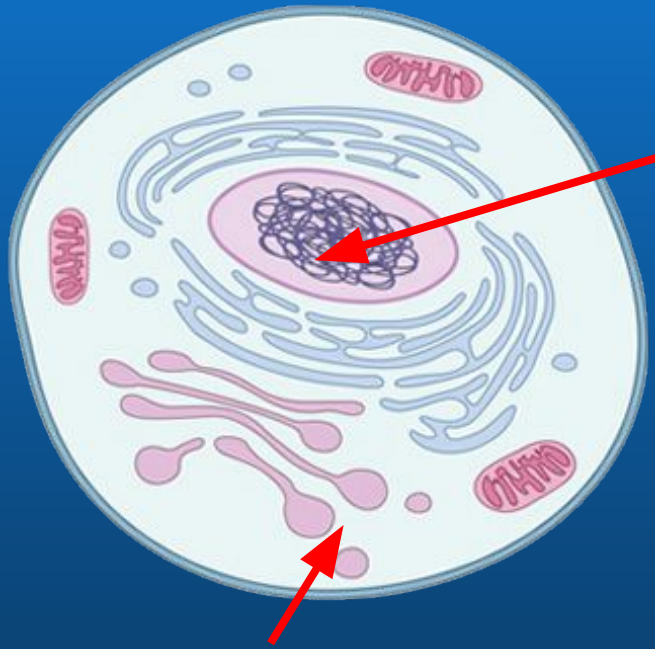
The Cell

The World of Biology

The Cell

The World of Biology

The Cell



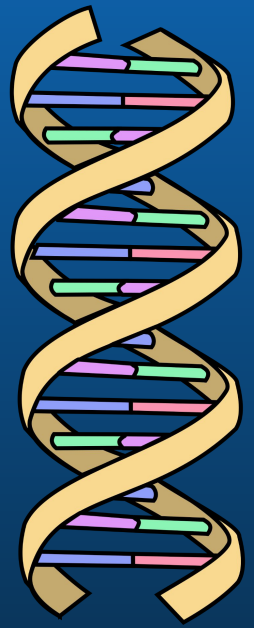
Nucleus

Cytoplasm

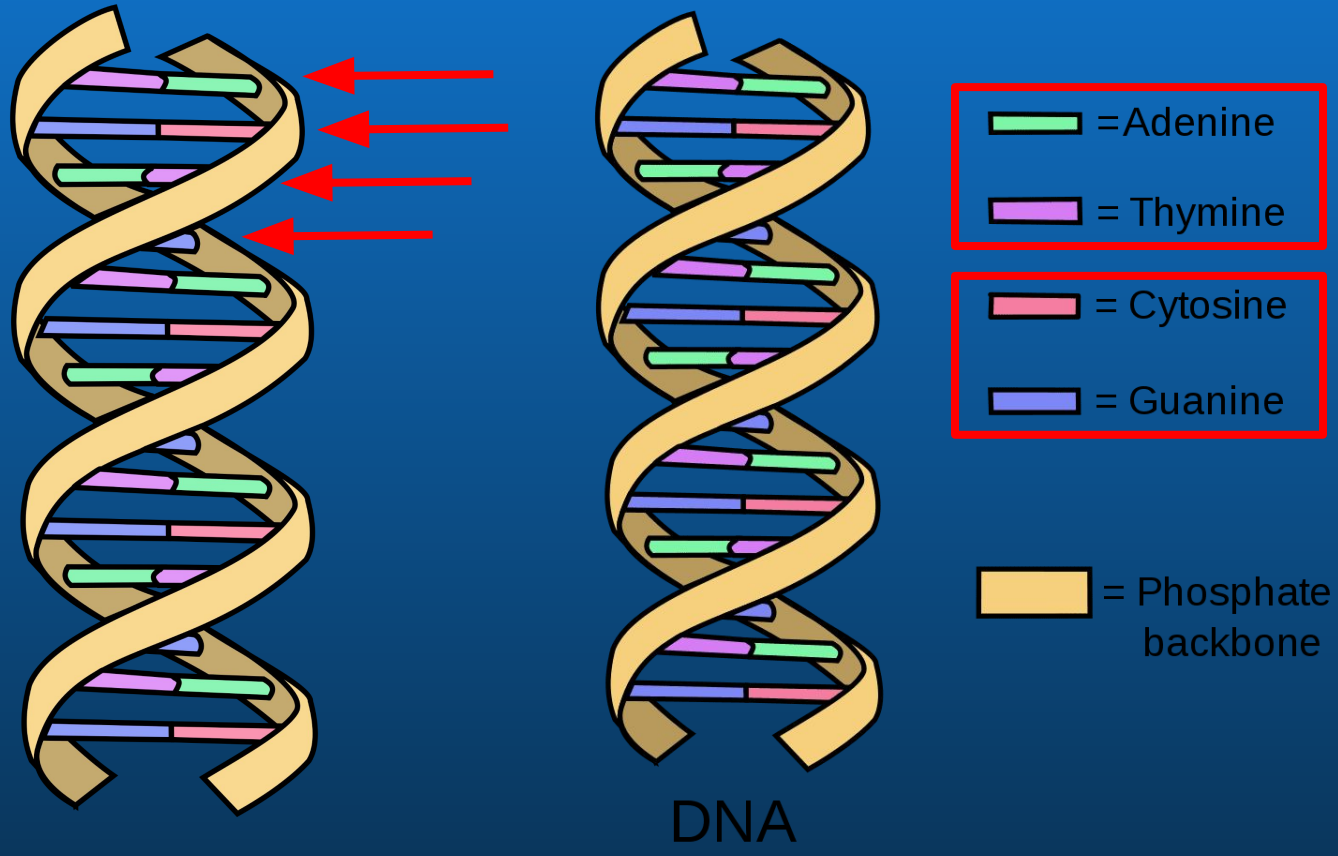
DNA in the Nucleus

The Cell

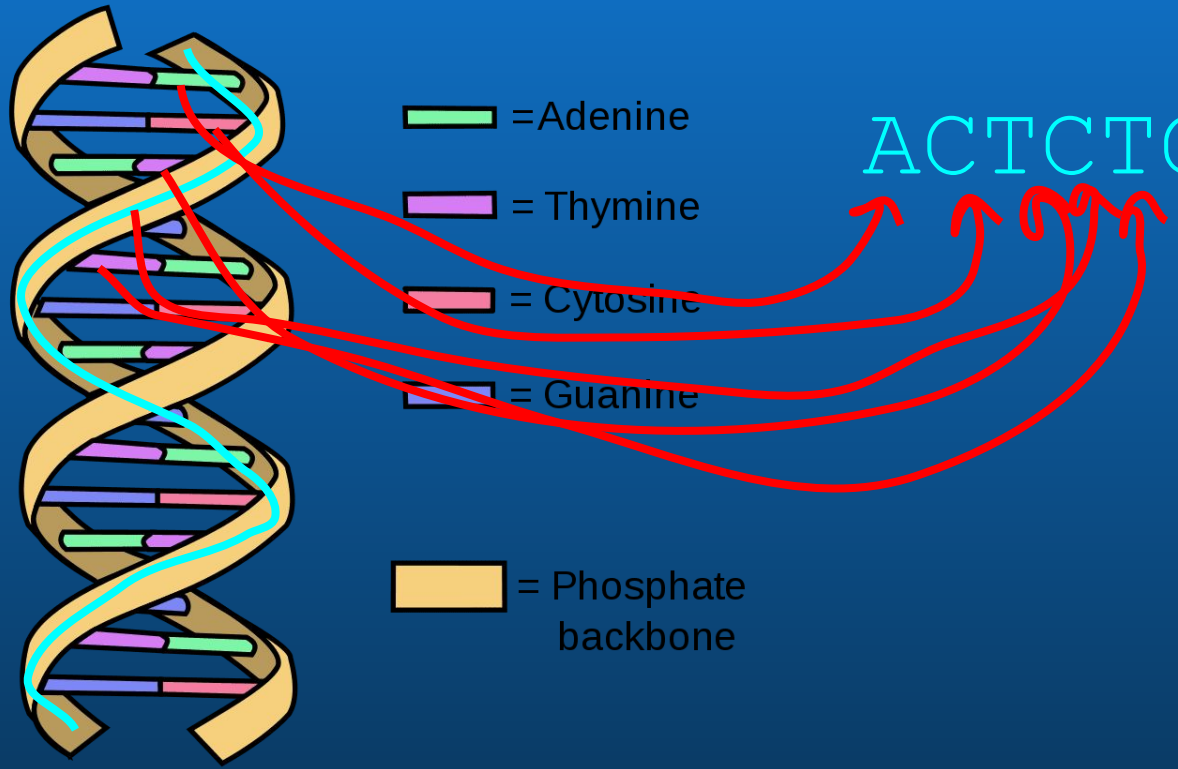
Nucleus => DNA



DNA in the Nucleus

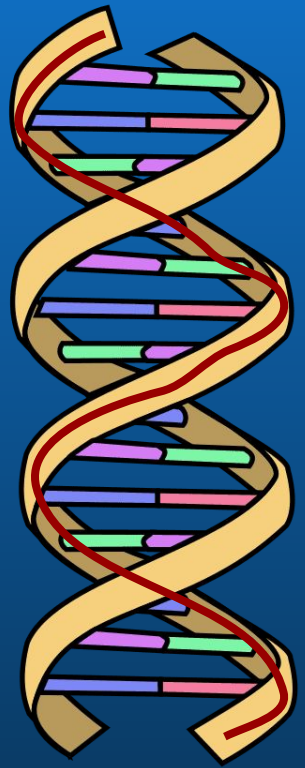


DNA in the Nucleus



ACTCTGAGACTCTG

Your Turn



DNA

=Adenine

= Thymine

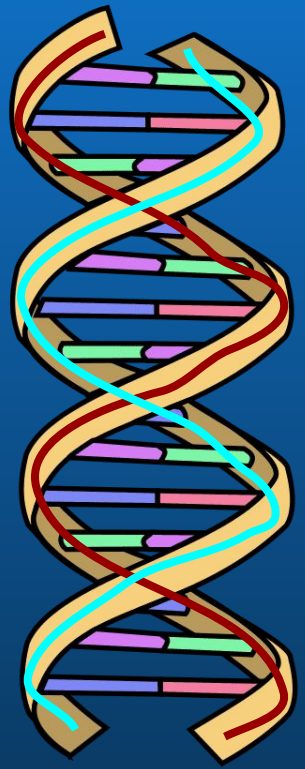
= Cytosine

= Guanine

= Phosphate backbone

TGAGACTCTGAGAC

Think about it a different way



DNA

=Adenine

= Thymine

= Cytosine

= Guanine

= Phosphate backbone

ACTCTGAGACTCTG

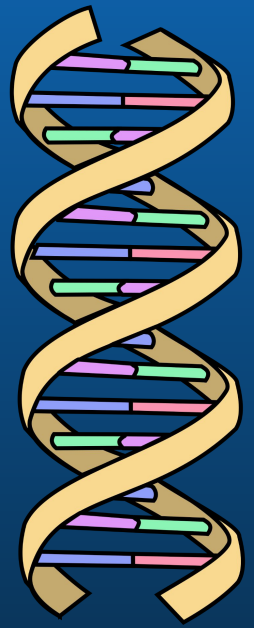
TGAGACTCTGAGAC

Worksheet Time!

The Genome

The Cell

Nucleus => DNA



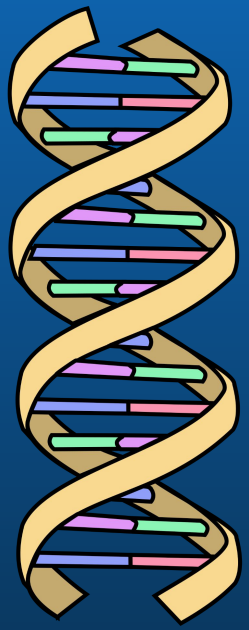
The Genome

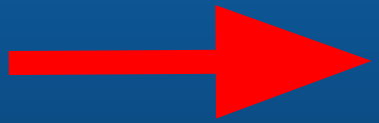
The Cell

Nucleus => DNA

The Genome

Nucleus => DNA





```
AGCCTTCTGGGTCCGAGGCTCCCAC
CTGCTCTAAGCGCTTGACACCCTTT
AAAAAAATGTATTTAAAGAG
GCTGGTTCCTATCCATCCGACTGGA
GGCATCTCAGTGCAAGAGCAAAGCT
AAGTCCTGCACACGCTCCTC...
```

The Human Genome

Nucleus => DNA

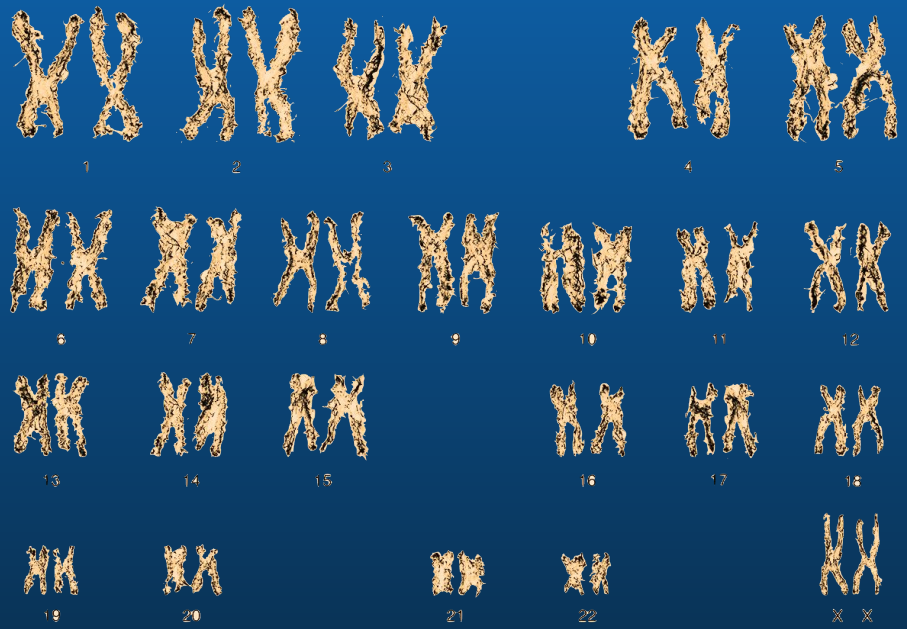
```
AGCCTTCTGGGTCCGAGGCTCCCAC
CTGCTCTAAGCGCTTGACACCCTTT
AAAAAAATGTATTTAAAGAG
GCTGGTTCCTATCCATCCGACTGGA
GGCATCTCAGTGCAAGAGCAAAGCT
AAGTCCTGCACACGCTCCTC…
```

+ 3 billion more

How do we Access a Location in the Genome?

The Cell

Nucleus => DNA



How do we Access a Location in the Genome?

How do we Access a Location in the Genome?

How do we Access a Location in the Genome?

UCSC Genome Browser Demo

https://genome.ucsc.edu/cgi-bin/hgGateway

The Human Genome

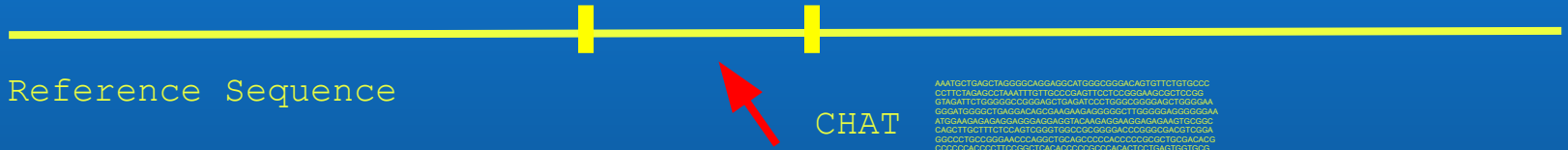
AGCCTTCTGGGTCCGAGGCTCCCACCTGCTCTAAGCGCTTGACACCCTTTAAAAAAATGTATTTAAAGAG
GCTGGTTCCTATCCATCCGACTGGAGGCATCTCAGTGCAAGAGCAAAGCTAAGTCCTGCACACGCTCCTC…

Reference Sequence

The Human Genome

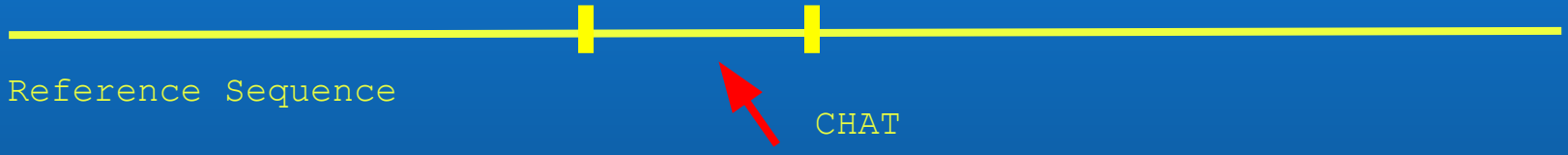
Reference Sequence

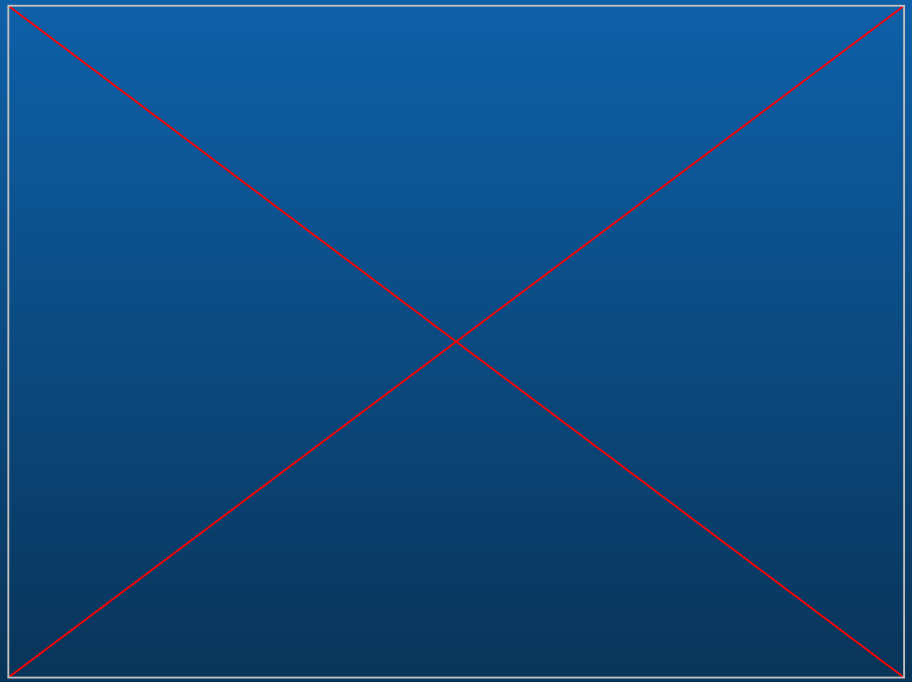
The Human Genome



AAATGCTGAGCTAGGGGCAGGAGGCATGGGCGGGACAGTGTTCTGTGCCC
CCTTCTAGAGCCTAAATTTGTTGCCGAGTTCCTCCGGGAAGCGCTCCGG
GTAGATTCTGGGGGCGGGAGCTGAGATCCCTGGGCGGGGAGCTGGGGAA
GGGATGGGCCTGAGGACAGCGAAGAAGAGGGGCTTGGGGCGAGGGGGAA
ATGGAAGAGAGAGGAGGGAGGAGGTACAAGAGGAAGGAGAGAAGTGCGGC
CAGCTTGCTTTCTCCAGTGGGGTGGCCGCGGGGACCCGGGCGCAGGTCGGA
GGCCCTGCCGGGAACCCAGGCTCCAGCCCCCACCCCCGCCGCTCGCACACG
CCCCCCACCCCTTCCGGCTCACACCCCCGCCCACACTCCTGAGTGGTGCG
GTGCAGCGTCGGCCGAGGCAGCAGAGCCGAGGAGAGCAGGTGAGAAGAAG
GGCTCGCGTGGCCTGCCGAGCCCGTCGCCCGGAGCAAGGCCGCCGCTCG
GGGGCTCTATCCAGGGACAGCACCGGGGCCGAGAGGCCCAGGACTCGTG
GAGTCCTCTGAGTCCTGCGCAGCAGCGGCCGTCGGGGGGTCAGCTAGGAAG
CTGCAGGATTGCGTTGCTTCCGTCCCTGGGCTTGGGGAGGCTTAACCATC
AGGGCTTCGGGACTCCAAGCGCATGTTCTCACTTCACTTGGCCCAGCATC
TGCTAAGTGACCAGGTCCGCGCTTTCTGTGTCTAGTCGGAGGGTCCTTGG
GTCTGACCCGACTCGCCTGAATTGCTGGTTGGCAACGCGTTCTCCCGCGA
GCAGGGGGTGGGGAAGGAGGAGACAGTGGCTGGTGGCAAGTCTGAGCCAGG
ACATAGCCGATGACAGCCTGTGAATGACCGGCCAGACTGGGGTTGGACCT
TCGGGGCAGAGCAGGCAGCCCTTGGAAAGCCAGGCTGGCAGGAGCATGGG
TGGCAGCCACAAACACAAACAGACATATATACACATACATCCTCTAGGC
ATCTGCCCAGGAGTGATGGAGGAAGAAGAGTTGGTAGGTGGCACCACAG
GAAACCATCCCACCCAGCTCTGCCTTCTTCAAGCTCAGGGTCCCTGAAGC
CTGACTAGTGTCTCCCGAATGGCGGGGGGAGGGGGGGCTCTGGCCCAGAGC
AAAAGGAGGCATTGGCTGCATGGTCTGTAGGGCCTTTACCAACCTGACCT
TGTCCTTGCTGTCCTCTCAAAAGACTGCCATGCTCCGGTGAATCCTCACC
AGGACGCCACCTCTTTCACTTTGCAGTTGGCGGCAGCCCCTACGAACCCA
CAGCACTGGGAGGCGTGGGAACAAGCCTGTGGAGAGTGTCCCTGCACTGT
GCCAGCAAGGGGTTTACATTCTGTGCCCAGCACTGCAAGGCCAGCCCCAGG
CTCCTCCATCCTTGGGCCTGGTTCATCCATCTTTGTGGACTCTCTGTTCCC
TGGCAACAACCCTGCCAGGTGCCCATACATGGGCGGTGCTTGCCTCTGGG
TAGGGGTGGTGTTGGGTGGGGAGAGGGGTGAGGAAATTAGCTAAACCTCA
GCTTTACACTTGAGACAATAGAGGGTGGGCTGCTGGAAAACATGAGTATG
TGCCTGTGGAGCCTTAGGTTCAGGGGGATCTCAGAGGGTAGCAATTGTGA
CCCACAGCCTAATAATAACAGTCTTTGCCTCTTGGCCTGGCTCACTTCT
ACTGTACCTTCAAAATTAATCTGCCTCTCCCCAGACTAAGGCAGGTGGCC
CCTAGGGCCCCTGGCTCCCCAGCCTCCCTGCATACCTCCCTGGGCACTCC
TATGGCATCTACACTGGCAGCCTGCAGACCCGAGACTGCCTAGGTCACAA
CGCCTCCAGCAGGCCCAGAGTCTCCCTGCCCTGGCCACAGGCCAGGCTCC
CAATTACCCCAGATGCATCCTGAGCCACAGGGCCTTAAAGGCCTCCATTC
ACCCACGCATGCATTCCTTTCCCTCCACCAACGGAGTGAGGGAATTCACC
CTACTCCACACCAGAGATGTGGCCGGAATGCAGAGATGAAGCACTGAGCA
CAGTAGGTATGGTCTACCCTTGACAACTAGTGGGAAAGAGAGCTTTG
ATTGGCAAAGCACTTATCTATTGGGTTCCAGGCAGGAACAAAGGCAGCAC
TCTGCCAGCAACCCCTGGTGGATTTGGATTGCCCCAGGGTCGTCAGCTGG
CCTTGGGTGACAACAGCTCATGAGACGAGACCCTCCACTTTAGGAACCAC
TCTCTCCCATATGGGGGCCCAGAGAGAACACTGTGGTCAGCACGTACAGG
TGGAAGTAGAGACTGGGGCCACGCCTGCAATGAGACCCCTATACACAGCT
GTCAGCTCTGGCCTCCTCATGGGCCTGGGGTGGGGGTCTGTTGCGCGGA
GGTGGAGGGTTTGTGACAGGCCTATGTGGCTGCTGTTGATGCTTCCCA
CTTCTTGGTCCCCAGGTCCACACCTCTGCATCCGTGCACCAGGACTCACC
AAGACGCCATCCTGGAAAAGGTCCCCCGTAAGATCGCAGCAAAAACTCC
CAGCAGTGAGGAGTCTGTGAGTGACTTTTGGAGCCCTCTCTCAACCCTGC
TTTCCCCACCTACATGCCCTTGCTTCTAGAACAGTCCTGAGTGGGACTGG
CCCCCAGCATTTGCCTGGCCCAGCCCTGCCCTCTACTGGCACAAGGCCC
AGGCTACAGCTCCACCCCCAGCTTGCTGAGGATCCCAGAAGGGCCTGGAG
GGGCTGAGACGGGCCTCAGGGTGCCAAGGAAATGCTAAGAGCTTAACTTC
CAAATCCCCTGGCCAGGCTAGGCCCTATCTTTCCTTCCAGGGCCTCTCTT
CATCCCAGCTTCCACTTCTGATAGCTTTGAGCCCAGTGATTTTGCCTTAG
GGTGGGGCTGAATTCCTACTCCAAGCCCTGCAAGTCCCAGCTCCAGCTGC
TATCACTAGTTCAGCCTTGGAGCCTTAGGCACCATGTTCTCCACTGGCCT
GAGACAATGCCATCCTCCCTGGGCAGGCATCTCCCTCCCGCCCCAGCCA
GCCTTCGGGTGGCAGAGCAACACCACCTGACTTCAGGTGGTTGCATGTGTT
TCCCCTCTCCCACACTTGTCTTTTTGCCACTCCTCTTCCCCCTCCCCCT
TCCCCTTTCCCCCAGCCCTCATTTGAAACCCTGAGCAGAATTTTCTGGCT
TCTCCCCAAGAAGCTGTTCACTCCATTGCCCTGAATTCTGCAGCTGACAA
ACATCTATTGTCACCTCTGGTCCACTGGCATGCTTCTCTACCCACCCCAT
TGCGCTGTGGCTTCCTTCAGAGCAGGCTCTGTTTCTCCATCGTCTCTGAG
TCCAGCCCAGGATCTGACTCACATGAACTGCCCAGGAAACATGGAGTTTT
GGAGCATTGATCTGAGACTCCCAGGTCCAAAGACACTCCTTAGCACAGGG
CCAGAAGTCCAGGAGAGTCTGGGGAGGATCCCTGGGGGGTTGAGATGTCT
GAGCTGCGATTTCCTGCTTTATTGCTTTTTACCCACTGCTGATTGCCTGT
TAGCCAGGCTAAGTCACAGCGCACTGTCAGAGCCTGGAAGAAGGGGAGCA
GTGTCTGGCAGAGGATCTAGCCAGACTTGTGGGCCCTTTTGGCCAGGTCA
CTAACTGGGCTGGGCTTTCTGTAGGTGCTCACCCACCCATCACACAGACC
ACCCCCCCCACCGCAAGAGATGGAAGAACAAGATCAGCTGGGTTGGGGGC
CCAGACCTGGCTCACTGGGGTTGGGCAGGGCAGTCTGGGATGCTGTCCAG
ATCCATCCTCGCATCATTCCCCAGGGCAGGAGCAGCCCGTAGTTGGCACT
G...

What is CHAT?





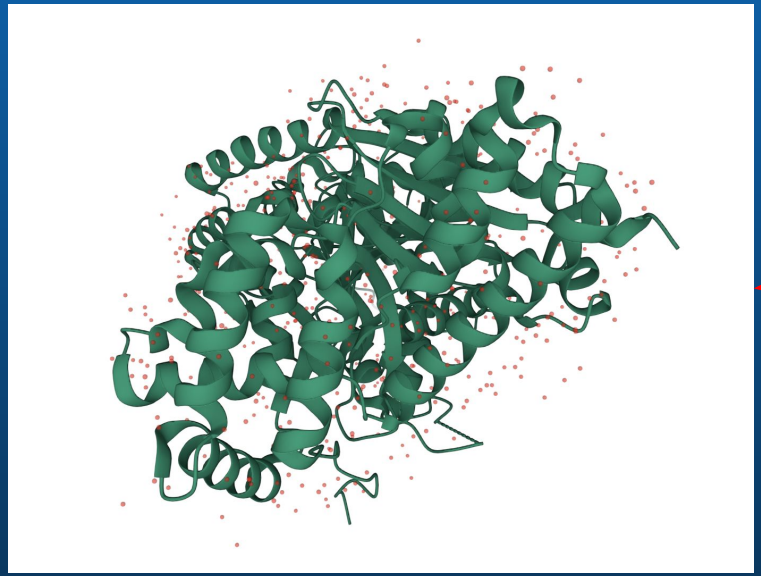
What is CHAT?

CHAT

Nucleus => DNA

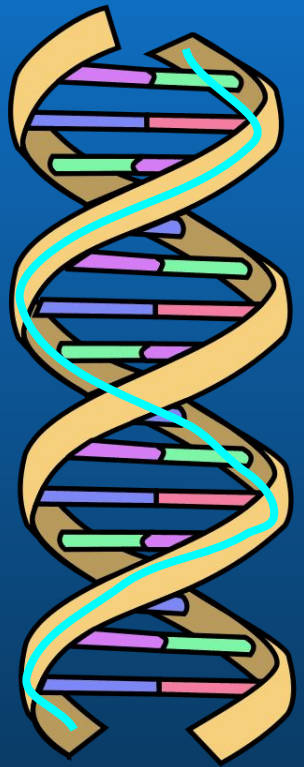
RNA

Protein



Cytoplasm

What is RNA?



DNA

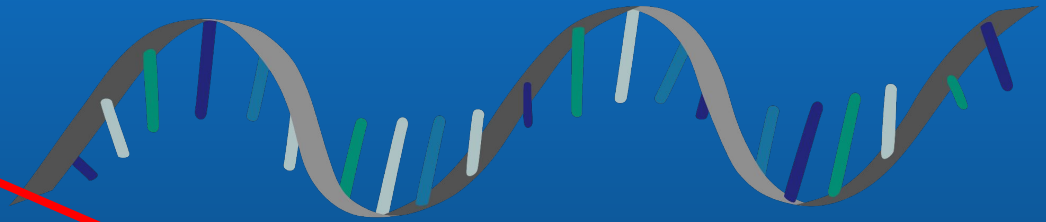
=Adenine

~~Thymine~~

= Cytosine

= Guanine

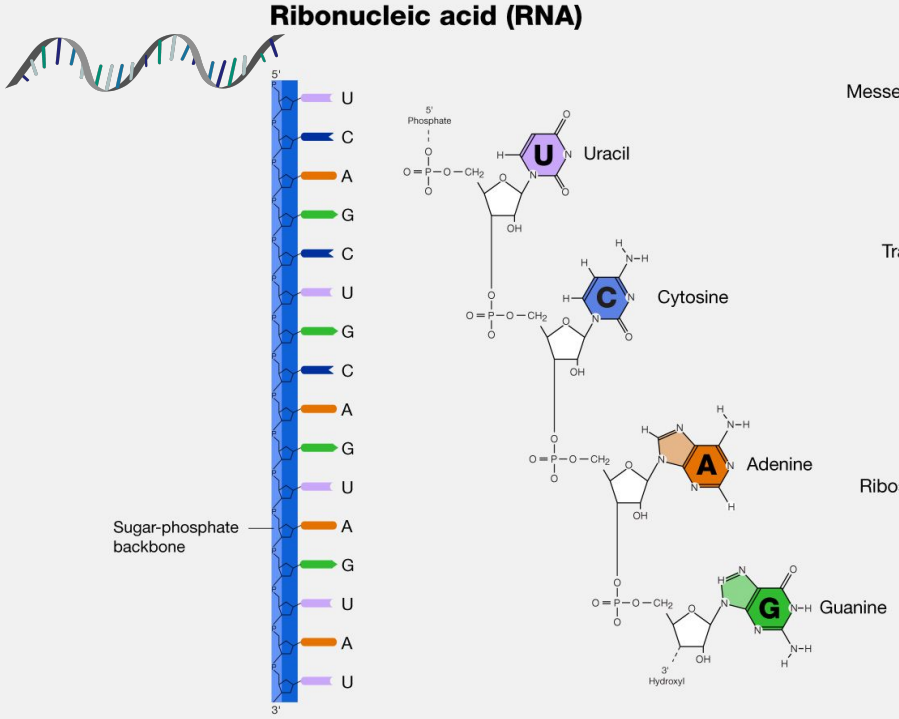
= Phosphate backbone

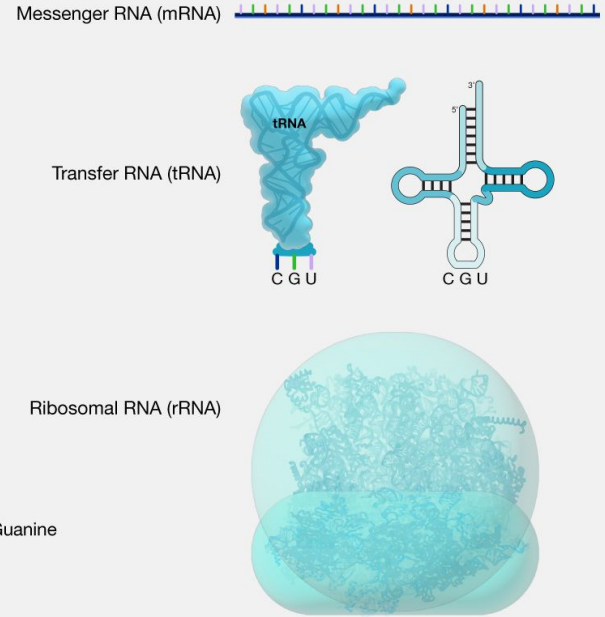


Uracil

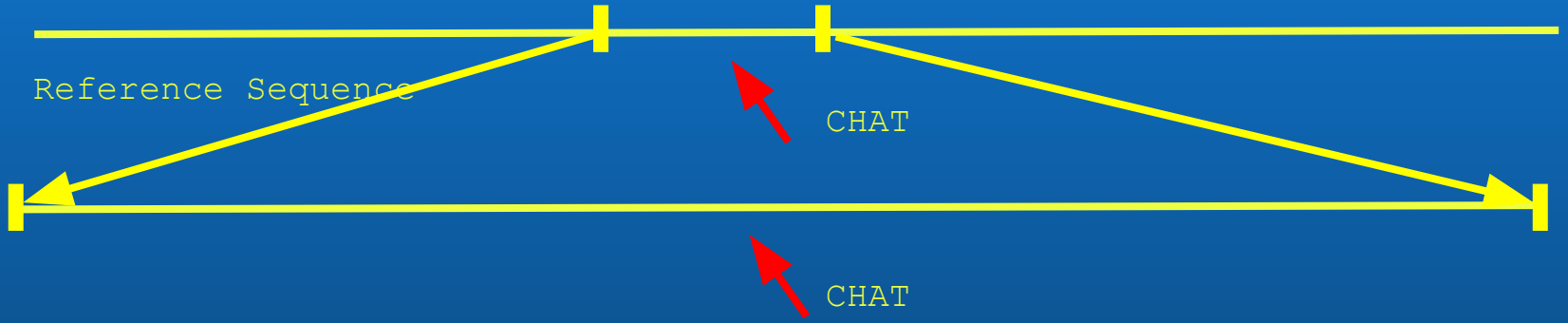
Theo Loves U

RNA Can Make Lots of Structures!

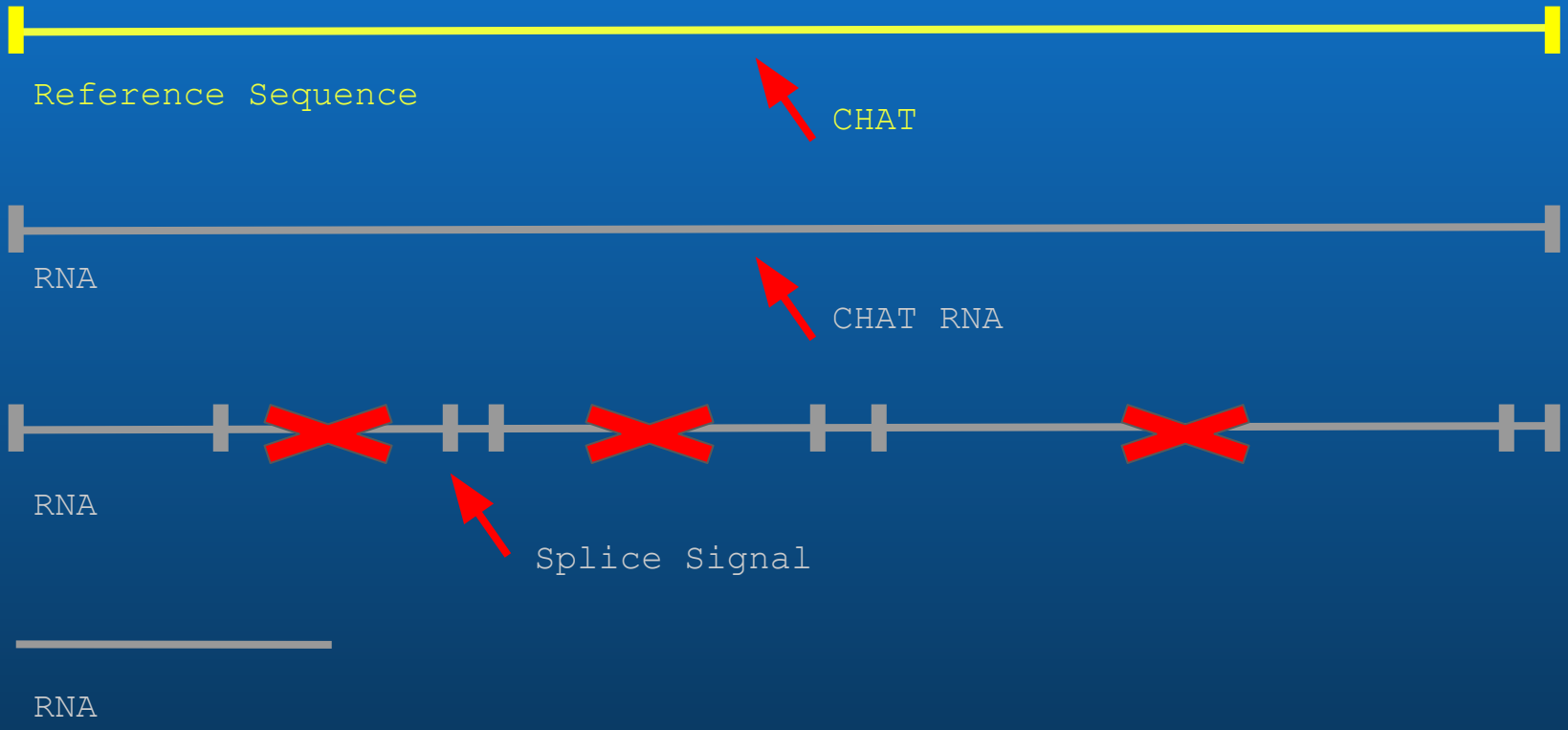




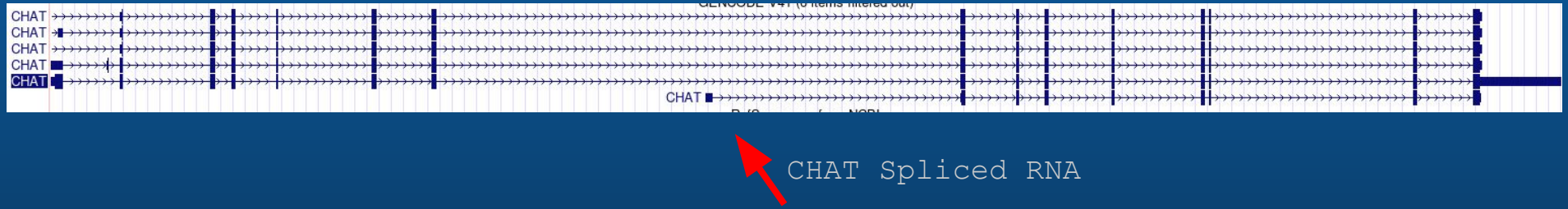
Alternative Splicing



Alternative Splicing



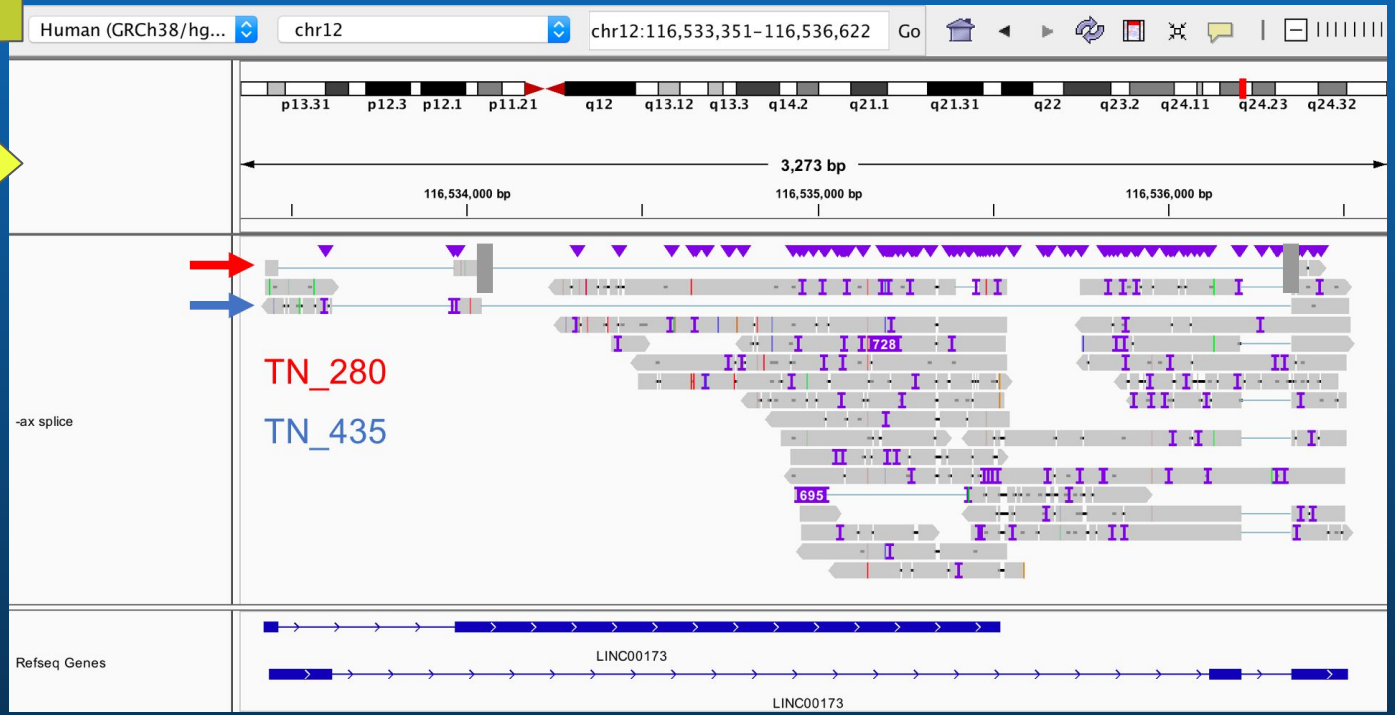
Alternative Splicing



There are many possible splice variants!

Interpreting IGV Charts

Reference Sequence



LINC00173
RNA

Worksheet Time!

What Will I Learn?

- Deoxyribonucleic Acid
- Ribonucleic Acid
- Sequencing
- *Alignment*

What Will I Learn?

- Theory
- Methods !
- The Great Unknown