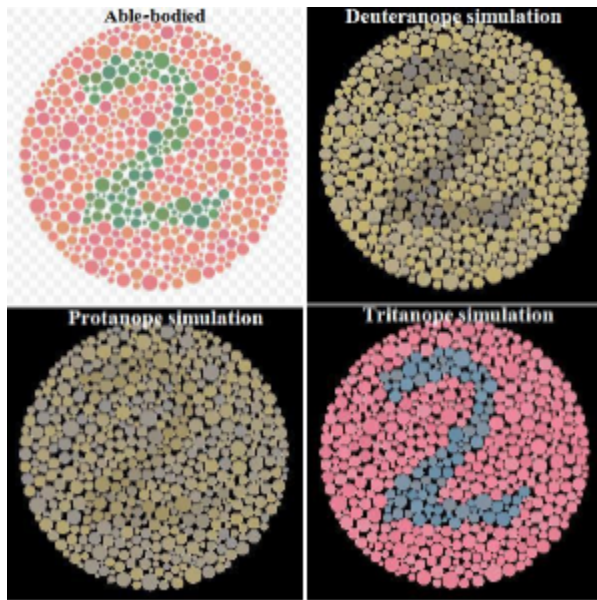


I: The Central Dogma

The Story

Those with protanopic color blindness (or red-green colorblindness) find it difficult to differentiate between red and green pigments, like in the image below.



OPN1LW is a protein coding gene which encodes for a light absorbing pigment called opsin 1 which specifically absorbs red light. It is found on the X chromosome of the human genome and defective copies of this gene are known to cause red-green colorblindness.

We have a DNA sequence that encodes for this gene. Can we identify from that, both the gene sequence and the protein sequence? Find the concepts and resources you need to do this below.

Molecular Biology Concepts

[DNA, RNA and Proteins: A molecular biology primer](#)

Programming concepts

[Variables, loops, conditionals: A Python Introduction](#)

Problem sets

1. Count how many nucleotides are in our DNA sequence
2. Convert the DNA sequence to RNA
3. Identify just the protein coding regions of the RNA
4. Translate RNA into protein

Files required:

[OPN1LW_nuclsequence.hg19.fasta](#)

[OPN1LW_annotation.hg19.gtf](#)

Additional Resources