# From genotype to phenotypes: *Red blood cell or White blood cell?*
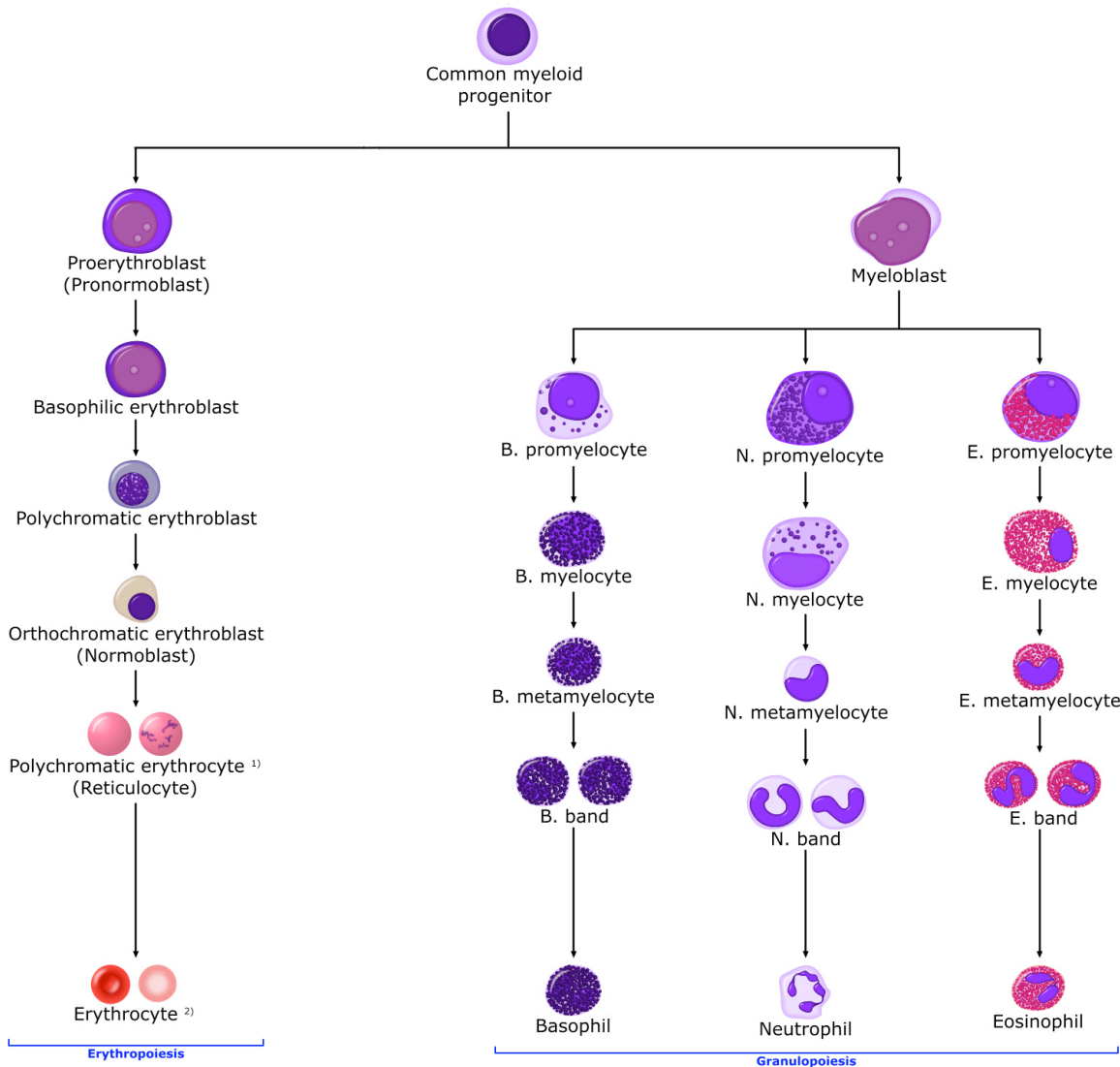
## *Introduction*

In this exercise, you will use a computer program to simulate the production of different types of mature blood cells from a common blood progenitor cell. The choice to proliferate then mature into red blood cells (erythrocytes) or a type of white blood cell (granulocyte) depends on which signaling molecule from the microenvironment acts on the progenitor cell. In reality, this choice depends on many outside signals that in combination change the phenotype.

Everyone starts out with identical copies of a short portion of the genome (raw DNA transcript) of a Common Myeloid Progenitor (CMP) cell. The figure below shows some differentiation pathways of this cell.

*Note: Not all CMP differentiation pathways are shown. You do not need to know the names for each step.*
*Image license information: http://creativecommons.org/licenses/by-sa/3.0/*



**Red blood cells**          **White blood cells**

Whether daughter cells of CMPs become red blood cells or white blood cells depends on which genes are expressed (transcribed and translated into proteins that lead to different phenotypes.) This differentiation process can be dictated by factors in the extracellular environment, such as the introduction of *Erythropoietin* or *Granulocyte-Colony Stimulating Factor*.

**Key words:**

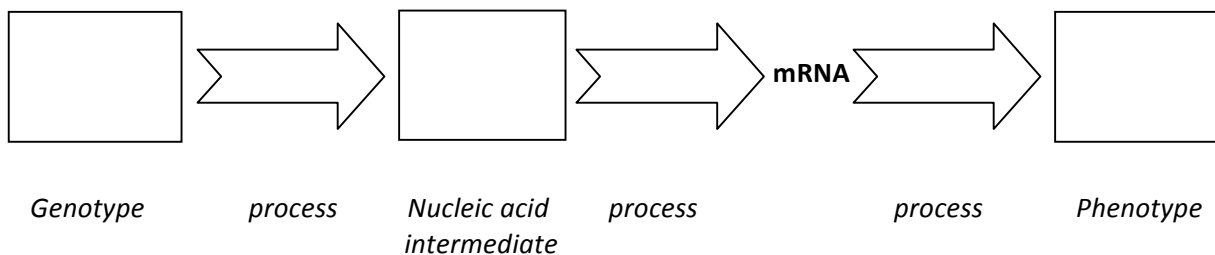*Define the following.*

Genotype: _____

Phenotype: _____

Proliferation: _____

Differentiation: _____

Myeloid progenitor cell: _____

*Fill in the steps with the following terms* (hints are *italicized* below each step):

Translation, pre-mRNA, protein, DNA, Transcription, splicing



| *Genotype* | *process* | *Nucleic acid intermediate* | *process* | **mRNA** | *process* | *Phenotype* |

**Directions**

Step 1: Open Another Plasmid Editor (ApE). You will be following the same steps as above.

Step 2: Copy and paste the **RAW DNA TRANSCRIPT** sequence on the next page into the ApE window.

Step 3: Based on whether you get Erythropoietin (Epo) OR Granulocyte-Colony Stimulating Factor (G-CSF), proceed from genotype to phenotype according to your corresponding set of directions on the next pages.

Circle the signaling molecule you got:          **Epo**          **G-CSF**

Step 4: Write down the mRNA sequence of your gene (after splicing).

_____

Step 5: Write your resulting protein's **amino acid sequence** _____

and **molecular weight** _____ in KiloDaltons.

What do M and * stand for? (Hint: Why does the amino acid sequence begin with M and end with *?)

Everyone copy this **RAW DNA TRANSCRIPT** and paste it into ApE:

AAGTTATACGCACTGCGCGTTTGTTTAGAAACTGGTGAGTCAGAGGCTAAATCCCACTCCAAATTCGGGCTCAATGCACAGTAACCTG
TGCATCGGGCCTTACCATCGACGGACGGACGTTGACTCTACGACTGGCACATTTTCAGACTACACGTCCTCGAGAGGGCATCATCATC
AGACGACTATGGGACAGTTTTTTGTAGCAATTACGACAACGATCCCCATACAAGATGCCATTGTCCCCCGGCCTCCTGCTGCTGCTGC
TCTCCGGGGCCACGGCCACCGCTGCCCTGCCCCTGGAGGGTGGCCCCACCGGCCGAGACAGCGAGCATATGCAGGAAGCGGCAGG
AATAAGGAAAAGCAGCCTCCTGACTTTCCTCGCTTGGTGGTTTGAGTGGACCTCCCAGGCCAGTGCCGGGCCCCTCATAGGAGAGGA
AGCTCGGGAGGTGGCCAGGCGGCAGGAAGGCGCACCCCCCCAGCAATCCGCGCGCCGGGACAGAATGCCCTGCAGGAACTTCTTCT
GGAAGACCTTCTCCTCCTGCAAATAAAACCTCACCCATGAATGCTCACGCAAGTTTAATTACAGACCTGAATATCGCACGTGATCTGTA
ACAGCCACTATATTAGGTTGTAATATGTAATCCTTGTAAATTAAAAAAAGTTAACCATATGTTATCTACTCAAAAAAGTTTATCGAATA
GTACATTAACATTAGCGAGTAAATCTCTGAGTACACAAAAAACGTACTTGAGCTTTCTATTACCTGGTACTTTTTAAGTGACTTAGATA
AATCAAAACCGTTTATACATACACTCGGGGACTTCACATTAACTAAATAGGCCGATACCAGGTACTTTGAATTAATCCATATATTATTA
TTAAGTGAACTGGAATAAAATAGTACTTATAAATAACGCACGAAAAAGTGTAAATAAACGAATCTTTATTCGGGGGTGGTGGTGGTG
GACCACAAGTTGTTTGTAGAATCAAATCAGTCGGTCTCAATGGGACAGATCTGAGTTGGGACACAAGTGTGATTGGGAAATCTCTAT
GGTTGACTACGGATCCACATTTCAGGTCTGTAATTAAACTTGCGTGAGCATTCATGGGTGAGGTTTTATTTGCAGGAGGAGAAGGTC
TTCCAGAAGAAGTTCCTGCAGGGCATTCTGTCCCGGCGCGCGGATTGCTGGGGGGGTGCGCCTTCCTGCCGCCTGGCCACCTCCCGA
GCTTCCTCTCCTATGAGGGGCCCGGCACTGGCCTGGGAGGTCCACTCAAACCACCAAGCGAGGAAAGTCAGGAGGCTGCTTTTCCTT
ATTCCTGCCGCTTCCTGCATATGCTCGCTGTCTCGGCCGGTGGGGCCACCCTCCAGGGGCAGGGCAGCGGTGGCCGTGGCCCCGGA
GAGCAGCAGCAGCAGGAGGCCGGGGGACAATGGCATCTTGTCCCTAAATCGCCCCCCACCAAGTAGCACACTCGACGAGCGCCGTT
TGCATTCACCTAGTGAGTCGCTTTCCAGACACTAAGATGCTAAACAACAGACTCGTTTTCTTTTCAACAGCTACC

**If you receive Granculocyte-Colony Stimulating Factor** (G-CSF):

Signaling cascade alpha, initiated by binding of G-CSF to its plasma membrane receptor, leads to activation of gene 1.

- Find **promoter region 1** by selecting **Edit → Find**, then copy and paste this sequence and click **Find Next**.
  TTCAGGTCTGTAATTAAACTTGCGTGAGCATTCATGGGTGAGGTTTTATTTGCAGGAGGAGAAGGTCTTCCAGAAGAAGTT
  CCTGCAGGGCATTCTGTCCCGGCGCGCGGATTGCTGGGGGGGTGCGCCTTCCTGCCGCCTGGCCACCTCCCGAGCTTCCTCT
  CCTATGAGGGGCCCGGCACTGGCCTGGGAGGTCCACTCAAACCACCAAGCGAGGAAAGTCAGGAGGCTGCTTTTCCTTATT
  CCTGCCGCTTCCTGCATATGCTCGCTGTCTCGGCCGGTGGGGCCACCCTCCAGGGGCAGGGCAGCGGTGGCCGTGGCCCCG
  GAGAGCAGCAGCAGCAGGAGGCCGGGGGACAATGGCATCTTGT

- Highlight this entire promoter region by clicking **Highlight All**, then Close. (If you make a mistake in highlighting, go to Edit → Clear Find Highlighting.)

Since we cannot "transcribe" a particular gene using ApE, we will have to do it by hand.
- With the promoter region highlighted, click anywhere in the sequence. Then click on **Edit → Reverse-Complement**. After doing this, delete the promoter region and all nucleotides that come before it, since these would not appear in the pre-mRNA transcript.

At the very beginning of your sequence, which was directly "downstream" of promoter region 1, is the **start sequence** (ATG) of gene 1. The **stop sequence** (in this case, TAA) marks the end of the gene.
- Find the stop sequence (make sure that it is **in frame**—that is, it can be read as "TAA" during translation, which involves reading the nucleotides in groups of three) by clicking on the **very beginning** of the sequence, selecting **Edit → Find**, and typing "**TAA**". Once you've found the first *in frame* TAA, delete everything that comes after it, since none of this would appear in the pre-mRNA transcript either.

Using this sequence (from start codon to stop codon) as a DNA template, transcriptional machinery (proteins and enzymes) converts gene 1 into **pre-messenger RNA, which includes introns and exons**.*

- Now we must "splice" out the **introns** within gene 1 by finding (same way as above) and deleting the following DNA sequences.

Intron 1. ATCCGTAGTCAACCATAGAGATTTCCCAATCA

Intron 2. TTGTGTCCCAACTCAGATCTGTCCC

Intron 3. GAGACCGACTGATTTGATTCT

Intron 4. AACAACTTGTGGTCCACCACCACCACCCCC

You now have a processed **mRNA** transcript ready for translation. Record the mRNA sequence.

To simulate translation, highlight the mRNA transcript from start codon through stop codon. Then press CTRL + T to translate (or go to ORFs → Translate...) then click OK.

Record the amino acid sequence and molecular weight of your new protein!

* For convenience, we have been working with the **reverse complement of the gene's DNA** sequence, and not the pre-mRNA sequence. This is a shortcut, since the pre-mRNA sequence nearly identical to the one we have been using, but with uracil bases (U's) instead of thymine bases (T's). Also, keep in mind that, in reality, splicing converts pre-mRNA to mRNA, so it does not change the gene itself—just the pre-mRNA transcript.

**If you receive Erythropoietin** (Epo):
Signaling cascade beta, initiated by binding of Epo to its plasma membrane receptor, leads to activation of gene 2.

- Find **promoter region 2** by selecting **Edit → Find**, then copy and paste this sequence and click **Find Next**.
ACAAGATGCCATTGTCCCCCGGCCTCCTGCTGCTGCTGCTCTCCGGGGCCACGGCCACCGCTGCCCTGCCCCTGGAGGGTGGCCCCAC
CGGCCGAGACAGCGAGCATATGCAGGAAGCGGCAGGAATAAGGAAAAGCAGCCTCCTGACTTTCCTCGCTTGGTGGTTTGAGTGGA
CCTCCCAGGCCAGTGCCGGGCCCCTCATAGGAGAGGAAGCTCGGGAGGTGGCCAGGCGGCAGGAAGGCGCACCCCCCCAGCAATC
CGCGCGCCGGGACAGAATGCCCTGCAGGAACTTCTTCTGGAAGACCTTCTCCTCCTGCAAATAAAACCTCACCCATGAATGCTCACGC
AAGTTTAATTACAGACCTGAA

- Highlight this entire promoter region by clicking **Highlight All**, then Close. (If you make a mistake in highlighting, go to Edit → Clear Find Highlighting.)

Since we cannot "transcribe" a particular gene using ApE, we will have to do it by hand.
- With the promoter region highlighted, click anywhere in the sequence. Then click on **Edit → Reverse-Complement**. After doing this, delete the promoter region and all nucleotides that come before it, since these would not appear in the pre-mRNA transcript.

At the very beginning of your sequence, which was directly "downstream" of promoter region 2, is the **start sequence** (ATG) of gene 2. The **stop sequence** (in this case, TAA) marks the end of the gene.
- Find the stop sequence (make sure that it is **in frame**—that is, it can be read as "TAA" during translation, which involves reading the nucleotides in groups of three) by clicking on the **very beginning** of the sequence, selecting **Edit → Find**, and typing "**TAA**". Once you've found the first *in frame* TAA, delete everything that comes after it, since none of this would appear in the pre-mRNA transcript either.

Using this sequence (from start codon to stop codon) as a DNA template, transcriptional machinery (proteins and enzymes) converts gene 2 into **pre-messenger RNA, which includes introns and exons**.*

- Now we must "splice" out the **introns** within gene 2 by finding (same way as above) and deleting the following DNA sequences.

Intron 1. GGGATCGTTGT

Intron 2. AATTGCTACAAAAAACTGTCCCATAGTCGTCTGATGATGATGCCCTCTC

Intron 3. GTGTAGTCTGAAAATGTGCCAGTCGTAGAGTCAACGTCCGTCCGTCGATGG

You now have a processed **mRNA** transcript ready for translation. Record the mRNA sequence.

- To simulate translation, highlight the mRNA transcript from start codon through stop codon. Then press CTRL + T to translate (or go to **ORFs → Translate**…) then click OK.

- Record the amino acid sequence and molecular weight of your new protein!


\* For convenience, we have been working with the **reverse complement of the gene's DNA** sequence, and not the pre-mRNA sequence. This is a shortcut, since the pre-mRNA sequence nearly identical to the one we have been using, but with uracil bases (U's) instead of thymine bases (T's). Also, keep in mind that, in reality, splicing converts pre-mRNA to mRNA, so it does not change the gene itself—just the pre-mRNA transcript.