

The Sound of DNA:

Musical Gene Expression

by

Janet A. De Souza-Hart and Joseph DeMasi

School of Arts and Sciences

Massachusetts College of Pharmacy and Health Sciences, Boston, MA

Introduction

They say that music is a universal language...and that every human being, regardless of their native tongue, can appreciate notes and chords, songs and symphonies. The notes on a page are translated into sounds and songs by the musical instruments we use to play them.

There is another universal language: DNA, hidden deep within our cells, is the language of life. No matter how different two human beings are on the outside, their cells contain the exact same DNA building blocks: nucleotides containing the bases cytosine (C), thymine (T), adenine (A), or guanine (G). Human beings have about three billion of these building blocks across 23 pairs of chromosomes. Chromosomes are just long strands of nucleotides packaged up with protein. We can think of these nucleotides as notes, and the order in which they are arranged helps create the symphony that makes us human.

A unit of DNA information (i.e., “a gene”) is used as a template to build mRNA which, in turn, is used as a set of directions to put together amino acids and synthesize a polypeptide. Transcription is often described as consisting of three phases: initiation, elongation and termination. Transcription initiation involves an enzyme called RNA polymerase that binds to a region of DNA (the “promotor” region) just upstream of a particular gene. Elongation consists of RNA polymerase separating the two strands of DNA followed by the creation of an mRNA molecule. This “messenger RNA” is made when RNA polymerase covalently attaches RNA nucleotides together using one of the separated DNA strands as a template. Finally, RNA polymerase reaches the end of the gene and, with the help of various accessory proteins, it falls off of the gene and the free mRNA is released.

In eukaryotic cells, once the mRNA is produced it is processed in a few special ways. “Caps” and “tails” are added to the ends, and little bits of extra (non-coding) RNA called introns are spliced out so that a continuous (mature) mRNA molecule is made from segments of coding DNA called exons. Then this mature mRNA leaves the nucleus and special protein-making machines called ribosomes jump on the mRNA floating around in the cytoplasm and the process of translation begins.

Translation also has three steps with the same names as the steps involved in transcription: initiation, elongation and termination. While the purpose of transcription is to produce RNA by “reading” the DNA, the purpose of translation is to produce protein by “reading” the RNA; a molecular machine called the ribosome performs this task. Initiation of translation begins when the small ribosomal subunit searches for a special mRNA start site (the nucleotides AUG). It waits for a complementary transfer RNA molecule to bring in the amino acid methionine. Elongation of translation occurs as the ribosome “reads” an mRNA molecule in three nucleotide “words” called codons (AUG is one of them) and waits for the appropriate tRNA to bring in the right amino acid. Each codon will specify the addition of one amino acid and only that amino acid will be allowed into a growing protein chain. Finally, termination happens when a ribosome reaches a stop codon in its reading frame. Once it reaches a stop codon, a release factor will help the whole complex dissociate, freeing our newly created protein so that it can go off and do its job in the cell.

In transcription, we make different forms (RNA from DNA) of the same chemical language: nucleic acids. But the next step, translation, involves a much more dramatic change. Getting back to our music analogy, think about how different it is to write notes on sheet music compared with actually playing that music. The notes are the instructions but the song we hear is the real functional outcome, what actually moves our emotions.

Just as written notes on a page help us enjoy music, or as the Rosetta Stone helped us to translate between hieroglyphics and classic Greek, the genetic code helps us to translate from the nucleic acid language to the amino acid language. Using this code, we can read each mRNA codon and know exactly which amino acid should be inserted into a protein, the functional outcome of gene expression.

Table 1. Genetic Code

		2 nd base				
		U	C	A	G	
1 st base (5')	U	UUU (Phe/F)	UCU (Ser/S)	UAU (Tyr/Y)	UGU (Cys/C)	U
		UUC (Phe/F)	UCC (Ser/S)	UAC (Tyr/Y)	UGC (Cys/C)	C
		UUA (Leu/L)	UCA (Ser/S)	UAA Stop	UGA Stop	A
		UUG (Leu/L)	UCG (Ser/S)	UAG Stop	UGG (Trp/W)	G
	C	CUU (Leu/L)	CCU (Pro/P)	CAU (His/H)	CGU (Arg/R)	U
		CUC (Leu/L)	CCC (Pro/P)	CAC (His/H)	CGC (Arg/R)	C
		CUA (Leu/L)	CCA (Pro/P)	CAA (Gln/Q)	CGA (Arg/R)	A
		CUG (Leu/L)	CCG (Pro/P)	CAG (Gln/Q)	CGG (Arg/R)	G
	A	AUU (Ile/I)	ACU (Thr/T)	AAU (Asn/N)	AGU (Ser/S)	U
		AUC (Ile/I)	ACC (Thr/T)	AAC (Asn/N)	AGC (Ser/S)	C
		AUA (Ile/I)	ACA (Thr/T)	AAA (Lys/K)	AGA (Arg/R)	A
		AUG (Met/M)	ACG (Thr/T)	AAG (Lys/K)	AGG (Arg/R)	G
	G	GUU (Val/V)	GCU (Ala/A)	GAU (Asp/D)	GGU (Gly/G)	U
		GUC (Val/V)	GCC (Ala/A)	GAC (Asp/D)	GGC (Gly/G)	C
		GUA (Val/V)	GCA (Ala/A)	GAA (Glu/E)	GGA (Gly/G)	A
		GUG (Val/V)	GCG (Ala/A)	GAG (Glu/E)	GGG (Gly/G)	G

Part I – The Huntingtin Song

What would happen if you combined our two universal languages: music and DNA? A group of researchers explored this question in a fascinating way. Rie Takahashi, Frank Pettit, and Jeffrey Miller at UCLA created a program that created music from a DNA sequence (a gene) based on the protein that that gene encodes. The program transcribed and then translated a gene, producing a protein sequence. Each of the 20 main amino acids that you might find in a human protein was assigned to a specific three-note chord, so the last step of the program read the amino acid sequence of that protein to create a musical composition.

The huntingtin gene is the DNA sequence associated with Huntington disease, a very serious, inherited neurological disorder. Listen to the **audio file** of the huntingtin gene that was created with this program.

Questions

1. What is unusual about the huntingtin “song”?
2. Everyone has this gene but some people unfortunately carry a variant (also known as an “allele”) that causes them to develop the symptoms of this disorder. Use the Genetics Home Reference (<<https://ghr.nlm.nih.gov/condition/huntington-disease#genes>>) to make a connection between the music and the mutant form of this gene. If you played the songs from huntingtin variants, could you tell which ones were functional and which ones caused the disorder? How?

Part II – Song Titles

Let's continue with our music analogy by reading a short gene, turning it into mRNA sequence (transcription), and then turning the mRNA sequence into a protein (translation), which will give us a song title. This will be a small protein, which is often referred to as a "peptide." There are many peptides that are important in biological systems. Some peptides are even used therapeutically as drugs or dietary substitutes. The artificial sweetener, aspartame, is a di-peptide consisting of aspartic acid (abbreviated "D" for short) and phenylalanine (abbreviated "F").

Here's the template strand of the gene: 3' TACCTCGCAACAATGATC 5'

Questions

- Each "letter" represents a DNA nucleotide. How many nucleotides are in this gene? _____
- Perform the process of transcription on this gene, using the DNA sequence to build the complementary mRNA sequence (remember that RNA uses the nucleotide uracil paired with adenine instead of thymine):

5' _____ 3'

- Each three-letter group is one codon. Starting with the AUG, how many codons are in this gene? _____
- Now, perform the process of translation on this mRNA, using the mRNA sequence to construct the appropriate protein sequence using the genetic code (see Table 1, p. 2 above). This table shows amino acids using their 3- and 1-letter abbreviations, and which codons they are associated with. For this step, use the 1-letter abbreviation. If you do this correctly, you should get the title of a 2015 song from the British band Muse: _____.

Next, you and a partner will choose different song titles (representing peptides), to create a complementary mRNA and then a DNA sequence—going in the reverse direction as above. You will then swap your DNA sequences, transcribe and translate them, and double-check that you get back the correctly spelled song title that your partner chose.

Here is a list of song titles that begin with the letter "M." One of you should pick a title from the left column and the other should pick one from the right column.

MACARENA	MANHATTAN
MADNESS	MASTERPIECE
MAGNIFICENT	MIRACLES
MIGRAINE	MAGGIE
MERCENARY	MIDNIGHT
MACHETE	MACHINES
MAGNETIC	MAINSTREET
MANDINKA	MANEATER
MYSTIFY	MICHELLE

- Using the genetic code, find a three-letter codon for each letter of your peptide song title. Write down each of these codons below (in order) to create an mRNA for your peptide. Keep in mind that for some letters/amino acids, there are multiple codons you can choose. That is okay; just choose one. Also, don't forget a stop codon at the end! Write this mRNA for your peptide song title here:

5' _____ 3'

- Write the DNA sequence complementary to your mRNA here, and write it on the next page to give to your partner to complete. Notice that polarity (direction) of the mRNA is the opposite of the DNA sequence (they are antiparallel):

3' _____ 5'

Partner Sheet for Part II

7. Write the DNA sequence complementary to your mRNA here (the same DNA sequence written *carefully* from Question 6) and then give this page to your partner to complete.

3' _____ 5'

8. Transcribe and translate the DNA sequence/gene in the space above. Use the one-letter abbreviations for the protein sequence.
9. How many *codons* does this gene contain, including the stop codon? _____
10. How many *nucleotides* does this gene contain, including the stop codon? _____
11. Why did we choose song titles that begin with the letter “M”?
12. If you and your partner had chosen the same song title, could you have generated different DNA sequences? Why or why not?

Part III – Wrong Notes

Even the best musicians sometimes play a wrong note. And even the best enzymes sometimes make mistakes. We can think about mutations in the DNA as “wrong” notes, as we heard in the huntingtin example. This could be broadly applied to many other genes. If you play the music from a wild-type protein at the same time as a mutant protein, you would be able to hear the mutations—the places in the music where the notes don’t match. If you did this with the fully functional hemoglobin *A* allele and the sickle-cell anemia causing hemoglobin *S* allele, you would hear only one chord changed, but at a crucial spot.

Questions

1. Have your partner introduce a point mutation (a single nucleotide change) into the DNA sequence from your peptide song title. Transcribe and translate it below to see if you can determine what type of mutation (missense, frameshift, nonsense, or silent) and where it occurred. Does the song title stay the same or change?

References

- Carey, J. 2016. Science and culture: musical genes. *PNAS* 113(8): 1958–9.
<<https://doi.org/10.1073/pnas.1601004113>>
- Chu, J. 2007. A prelude in protein. *M.I.T. Technological Review* May 9, 2007.
<<https://www.technologyreview.com/s/407885/a-prelude-in-protein/>>
- Cold Spring Harbor Laboratory. n.d. Transcription and translation: RNA splicing. [Animation]. Running time: 1:37 min.
<<https://www.dnalc.org/resources/3d/>>
- Howard Hughes Medical Institute (HHMI) Biointeractive. n.d. DNA transcription (advanced detail). [Animation]. Running time: 1:55 min.
<<http://www.hhmi.org/biointeractive/dna-transcription-advanced-detail>>
- Howard Hughes Medical Institute (HHMI) Biointeractive. n.d. Translation (advanced detail). [Animation]. Running time: 3:04 min.
<<http://www.hhmi.org/biointeractive/translation-advanced-detail>>
- Reece, J. B., L.A. Urry, M.L. Cain, S.A. Wasserman, P.V. Minorsky, R.B. Jackson, and N.A. Campbell. 2014. *Campbell Biology*, 10th ed. Boston: Pearson.
- Takahashi, R. and J.H. Miller. 2007. Conversion of amino-acid sequence in proteins to classical music: search for auditory patterns. *Genome Biology*. 8: 405.
<<https://genomebiology.biomedcentral.com/articles/10.1186/gb-2007-8-5-405>>
- U.S. National Library of Medicine. 2018. Huntington disease. *Genetics Home Reference*.
<<https://ghr.nlm.nih.gov/condition/huntington-disease#genes>>