YALE NATIONAL INITIATIVE
to strengthen teaching in public schools®

Curriculum Units by Fellows of the National Initiative
2017 Volume VI: Engineering of Global Health

Gene Therapy and Muscular Dystrophy: Structure, Function, and Dysfunction of the Muscular System

Curriculum Unit 17.06.01, published September 2017
by Kwame Adu-Wusu

Introduction

The human body is wonderfully elegant in its composition and operation, at once a complex assemblage of structures and an aggregation of simple, often intuitive, functions. The content of my Anatomy and Physiology course is organized around a central concept: *Mr. Adu-Wusu's 3 P's of Anatomy and Physiology* (Figure 1). The 3 P's are as follows: Parts, Purpose, Peril. Alternatively, they can be considered in the framework of anatomy-physiology-dysfunction or structure-function-dysfunction. The basic principle holds that the body is comprised of structures whose compositions dictate the functions they can perform and thus, changes in structure can lead to dysfunction.

Mr. Adu-Wusu's 3 P's of Anatomy & Physiology		
Parts	**P**urposes	**P**erils
anatomy	physiology	disease / disorder
structure	function	dysfunction
What is it made of? How is it put together?	*What does it do?*	*What can go wrong?*

As an example, consider the skin. The outermost layer consists of tightly-packed, flattened cells filled with the protein keratin. The arrangement of these epidermal components allows the integumentary system to serve a barrier function, rebuffing transport of water, germs, and other unwanted foreign substances. Any event that breaks the continuity of this layer (e.g. penetration by a sharp object) diminishes the ability of the skin to function as a protective barrier. Necessarily then, remedies to address such dysfunctions (e.g. adhesive

bandages, surgical stitches, scab formation, etc.) seek to restore the physical integrity of the epidermis that enables its appropriate function.

When considered in this way, the workings of the body are much less a set of enigmatic phenomena, and much more so a known or knowable system that can be understood through experimentation and data analysis: research. As my students come to internalize this fundamental construction, they can view various health ailments – disease, injury, aging, etc. – as just changes in functional capability owing to changes in structure. It follows then that solutions (actual and potential) for bodily infirmities result from restoration of normal structure and/or mitigating the effects of abnormal structure. Simply put, my goal is for my students to embrace the idea that we explore the parts of the body, so we can know how those parts work, because it is by understanding how the parts work, that we might be able to fix any problems that arise. In this structure-function-dysfunction framework, students can explore and better understand anatomical and physiological topics ranging from relatively simple skin lacerations and bone fractures to more complex gastrointestinal disorders and autoimmune diseases.

Throughout the course, the paradigm of the 3 P's is applied to the various organ systems of the body. The curriculum unit presented here is to be a short segment taught within the portion of the course that covers the structure, function, and dysfunction of the muscular system. Prior to the curriculum unit, students have come to understand that the structures of the muscular system cooperate to perform several critical functions; muscles generate movement of body parts and help to maintain joint stability; muscles help to provide the shape/structure for the body and help to maintain posture; muscles produce heat to help maintain body temperature (homeostasis); muscles help to protect internal organs; muscle contractions enable movement of blood and food/nutrients through the body.

The curriculum unit establishes Duchenne muscular dystrophy (DMD) as one of the representative disorders that students can explore in order to gain a better understanding of normal and abnormal functioning of the muscular system. The curriculum unit also presents a simplified framing device through which students can grasp the molecular basis of gene therapy. Advances in gene therapy for the treatment of DMD are presented as concrete examples to reinforce the central idea that even complex remedies fit the simple paradigm of regaining function by restoring structure.

Demographics

My school is a high-poverty (100% qualify for free and reduced-price meals at school) neighborhood (non-application) high school in Washington, DC. The student body is predominantly black (62%) and Hispanic/Latino (36%). There is a significant international presence at the school. Several west and east African countries including Cameroon and Ethiopia are represented and many of the Hispanic/Latino students emigrated from Mexico and Central America. The school is in the midst of implementing an International Academy (starting with the lower grades), so the newcomer population is set to increase in the years to come. 30 percent of the student population are classified as English Language Learners (ELL). There are several of my students whose primary language is not English. In school, they generally communicate with peers in Spanish or Amharic and only use English to converse with administrators and teachers. Most of the ELL students in my classes have sufficient command of English and require only limited support, but there are a handful who may require substantial support (e.g. dual-language academic resources, translators, tutors,

etc.).

Recent standardized test scores suggest significant deficits in English Language Arts, math, and science performance. Student surveys and anecdotal observations reflect a disinclination toward the sciences among the student population. Often the students' lack of interest in the sciences is a substantial factor in lack of classroom engagement. An expectation for this curriculum unit is to increase classroom engagement by emphasizing the direct connection between the textbook/academic material that students might find abstract or even arbitrary ("*Why do we have to learn this?*") and the contemporary real-world applications of those concepts. The hope is that students will embrace the idea that the foundational principles being introduced in class are the basis for new, life-altering medical advances impacting the lives of real people. This is a version of a notion I introduce on the first day of school and underscore throughout the year with each new unit; in academia in general, but in science more particularly, the point of learning is not simply to come to know information that others have previously discovered or described, but rather to acquire skills and insight that will lead you to develop new knowledge to be shared with and advance the world.

Standards

The curriculum unit is intended for 11th and 12th grade Anatomy and Physiology students who may or may not have previously completed a high school biology course, and still lack mastery of some foundational concepts from biology. The curriculum unit incorporates Next Generation Science Standards (NGSS) for High School Life Sciences (HS-LS1: From Molecules to Organisms: Structures and Processes and HS-LS3: Heredity: Inheritance and Variation of Traits) and specifically emphasize the associated Disciplinary Core Ideas (DCI) (LS1.A: Structure and Function, LS1.B: Growth and Development of Organisms, LS3.A: Inheritance of Traits, and LS3.B: Variation of Traits).

The performance expectations in HS-LS1 are intended to help students formulate an answer to the question, *"How do organisms live and grow?"* The content of the curriculum unit addresses whole-body effects of the muscular system and the roles it plays in sustaining the human organism (i.e. the primary functions of the muscular system presented previously) as well as the requirements needed to maintain life at the level of the cell. Students are able to draw a direct line from a protein (dystrophin) and the specific role it plays (maintaining the structural integrity of individual muscle cells) through to the organ system (skeletal muscles) and the variety of functions they perform (e.g. generating movement).

The performance expectations in HS-LS3 are intended to help students formulate answers to the questions, *"How are characteristics of one generation passed to the next? How can individuals of the same species and even siblings have different characteristics?"* The fact that DMD results from gene mutations provides the avenue to address genetic variation in the population and the location of the dystrophin gene on the X-chromosome allows for the opportunity to address differential modes of inheritance and gene expression.

The curriculum unit necessarily also incorporates NGSS Science & Engineering Practices (SEP); 1 - asking questions (science) and defining problems (engineering), 2 - developing and using models, 3 - planning and carrying out investigations, 4 - analyzing and interpreting data, 5 - using mathematics and computational thinking, 6 - constructing explanations (science) and designing solutions (engineering), 7 - engaging in argument from evidence, 8 - obtaining, evaluating, and communicating information.

The Curriculum Unit

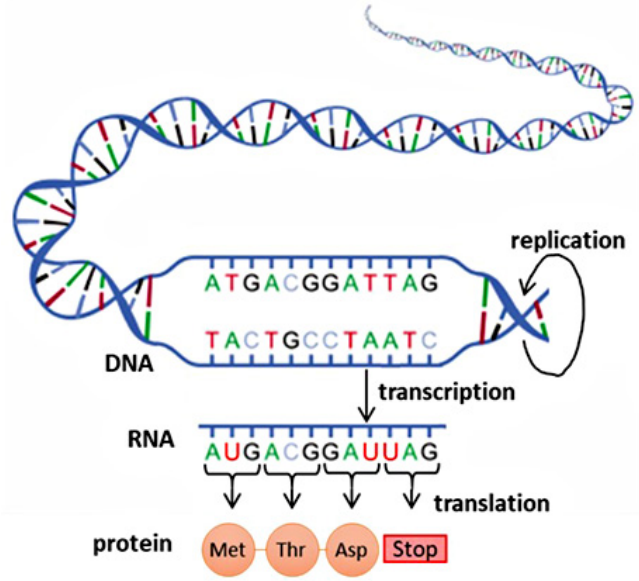
The curriculum unit is arranged in three parts. Part 1 presents a review/overview of molecular genetics. There is a specific focus on how information encoded in deoxyribonucleic acid (DNA) results in protein products and how the mutations of various kinds can interfere with those processes. Part 2 introduces gene therapy and the mechanisms thereof. A general overview is presented, but there is also a focus on specific methods being developed and employed by researchers as potential treatments for DMD. Part 3 asks students to employ their scientific learning to engage with some of the ethical and moral topics that are raised by gene therapy research and the advances thereof.

Part 1 - Overview of Molecular Genetics

It is widely understood that DNA is a biomolecule that contains the blueprint for any living organism. Within the base pairs of the double-helical structures housed in a single nucleus is all the information needed to build the entire organism or any part therein. In short, DNA is the molecular basis for heredity. As such, DNA must both preserve the genetic information and facilitate the expression of the genetic information.

The information stored in DNA is the result of the arrangements of its four chemical bases: adenine (A), guanine (G), cytosine (C), and thymine (T). Akin to the manner in the which letters of the alphabet can be arranged to form words and sentences with an endless variety of meanings, so too can these A, G, C, and T bases of DNA be ordered to hold an array of genetic information. This framing device of using letters and words to construct complex sentences with clear meanings is the through line of the unit. In the end, and at points throughout the unit, students construct examples and models to show foundational concepts underpinning molecular biology and gene therapy.

According to the central dogma of genetics, the pattern of information occurring most frequently in cells is as follows: information within DNA is preserved when the molecule is duplicated through the process of replication and in order for the genetic information to be expressed, DNA is transcribed into RNA, then RNA is translated into protein. DNA molecules are arranged in a double helix structure in which bases pair complementarily with each other, A with T and C with G. In this way, each strand of the double helix can serve as a template to reproduce the entire molecule (replication). Replication is essential because in most cell divisions each new cell must have a complete copy of the DNA contained in the original cell. To facilitate gene expression, a DNA sequence on one strand is used as a template to produce an RNA sequence (transcription), which in turn is to be used as guide for assembling protein (translation) (Figure 2).



Second letter

First letter	U	C	A	G	Third letter
U	UUU, UUC Phe; UUA, UUG Leu	UCU, UCC, UCA, UCG Ser	UAU, UAC Tyr; **UAA Stop**; **UAG Stop**	UGU, UGC Cys; **UGA Stop**; UGG Trp	U C A G
C	CUU, CUC, CUA, CUG Leu	CCU, CCC, CCA, CCG Pro	CAU, CAC His; CAA, CAG Gln	CGU, CGC, CGA, CGG Arg	U C A G
A	AUU, AUC, AUA Ile; **AUG Met**	ACU, ACC, ACA, ACG Thr	AAU, AAC Asn; AAA, AAG Lys	AGU, AGC Ser; AGA, AGG Arg	U C A G
G	GUU, GUC, GUA, GUG Val	GCU, GCC, GCA, GCG Ala	GAU, GAC Asp; GAA, GAG Glu	GGU, GGC, GGA, GGG Gly	U C A G

DNA is the molecular basis for heredity because the order of its bases, the gene sequence, determines the protein products that will be produced. Genes are the specific segments of DNA, a specific set of instructions, that encode particular protein products. Genes can range in size from as small as a few hundred base pairs long to as large a million or more base pairs. According to the Human Genome Project, the human genome contains approximately 3 billion base pairs divided among 23 pairs of chromosomes. Within these base pairs is encoded information for approximately 25,000 genes (Griffiths, et al. 2005).

With billions of base pairs and tens of thousands of genes, it is necessary to have a comprehensive coding system to organize genetic information. To recall the earlier analogy of the alphabet and sentence construction, it is necessary to understand how letters are grouped to form words and how the words of a sentence are read in sequence across a page to convey complete ideas. In an analogous manner, it is necessary to know that the genetic code is a specific (unambiguous), universal, non-overlapping, degenerate (redundant), non-punctuated code in which groups of three bases (codons) code for the amino acids used to construct protein products (Griffiths, et al. 2005).

To begin, there are 20 different amino acids commonly found in cellular proteins. These in a way are the words that form sentences (proteins).

When a strand of messenger RNA is read end to end, there is only one base at each position. As there are only four different bases in RNA (A, U, G, and C), amino acids cannot be represented by a single letter (base); otherwise the size of the vocabulary, four, would be far too small to accommodate the 20 words (amino acids), needed to form fully coherent, complex, and meaningful sentences (functional proteins). If pairs of letters (AA, AU, GC, etc.) were the basis of word formation, the vocabulary would be larger ($4^2 = 16$), but still insufficient. With triplet codons (AAA, CGC, AUG, etc.), 64 (4^3) distinct word combinations are possible, more than enough for 20 amino acids. Experiments by Francis Crick and other researchers confirm that the codon is indeed exactly three bases long (Griffiths, et al. 2005). The genetic code is said to be degenerate (redundant) because some amino acids are in encoded by multiple codons, however it is also specific (unambiguous) in that no codon codes for multiple amino acids (Figure 2). It is also important to note that the code is non-

punctuated; once reading of the letters (bases) is initiated, each letter is read until a stop codon is reached – no commas, no spaces. Crucially, adjacent words (amino acids) do not share letters (bases); the code is said to have non-overlapping reading frames.

By applying these rules of genetic sentence construction, a DNA sequence such as ATGACGGATTAG, is transcribed into the messenger RNA transcript AUGACGGAUUAG and then translated into the polypeptide sequence Met-Thr-Asp (Figure 2).

Mutations

The framing device of sentence construction is helpful beyond just understanding how a DNA code reliably determines how proteins are constructed. It is also useful for understanding the consequences of errors in those processes, mutations in particular. Gene mutations are lasting changes in the DNA sequences that code for a protein. The causes of mutations are many and include random errors in the DNA replication process. The source of a mutation can involve a single base, a few bases, of even large segments of DNA including entire chromosomes. For the purposes of this curriculum unit, it is important to understand several types of mutations and even more so to grasp the potential consequences for the protein product of the associated gene.

Missense mutation – Missense mutations occur when a single base pair substitution occurs such that one amino acid in the peptide sequence is replaced with a different amino acid (Figure 3). The severity of the consequences will differ depending on the role of the particular protein, the location of the mutation, and the differences between the amino acids; a relatively benign missense mutation might slightly affect the shape and folding of protein without resulting in significant losses or gains of function. In other cases, a single amino acid change in a critical domain of the protein might result in a very diminished protein product and/or critical loss of function. The mutation that is associated with sickle cell anemia, for example, is caused by a single base pair substitution in which a GAG codon in the gene for hemoglobin is changed to GTG resulting in an amino acid valine where there should be a glutamic acid (Kumar, Fausto and Abbas 2004). The consequence of this minor change in structure is a significant shift functional capability.

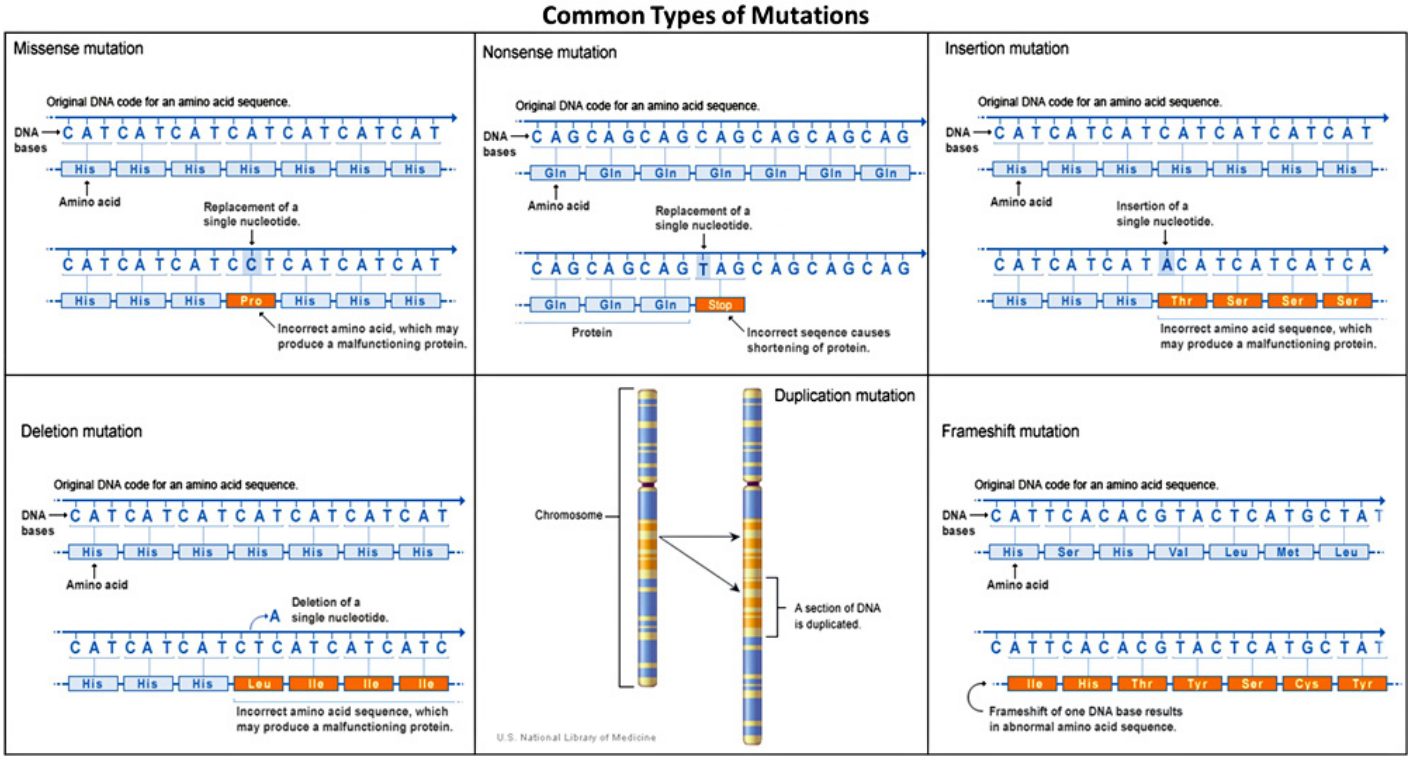
Nonsense mutation – Nonsense mutations occur when a single base pair substitution occurs such that one amino acid in the peptide sequence is replaced with a stop codon; a signal for the translation apparatus to stop building the protein (Figure 3). Like the single base pair missense mutations, the severity of the consequences of a nonsense mutation depends on the particular protein and the location of the mutation. In general, though, the closer a nonsense mutation is to the beginning of a gene, the more likely it is to have significant effects as it is more likely to disrupt critical domains downstream from the substitution. A sufficiently truncated protein might retain no function at all. Some of the most severe forms of DMD, for example, are the result of nonsense mutations early in the sequence of the dystrophin gene (Kumar, Fausto and Abbas 2004).

Insertion and Deletion Mutations– Insertion mutations add one or more bases somewhere in the sequence of the gene, whereas, deletion mutations remove one or more bases from the sequence of the gene (Figure 3). Very large deletions might even remove the whole gene. The potential consequences range depending on the nature of the particular sequences added or removed as well to/from where they are added or removed. These consequences are discussed further following the description of frameshift mutations (Kumar, Fausto and Abbas 2004).

Duplication Mutations – Duplication mutations are abnormal repetitions of one or more base pairs. As with

other mutations, severity of consequences varies. These consequences are discussed further following the description of frameshift mutations (Kumar, Fausto and Abbas 2004).

Frameshift Mutations – Frameshift mutations occur when the insertions, deletions, duplications, or similar mutagenic events disrupt the reading frame of the gene sequence (Figure 3). Recall that according to the sentence construction analogy established previously, the genetic code is non-overlapping and non-punctuated. Also, amino acids are constructed based on codons defined by base triplets. As such, for mutations that change the number of bases in the gene sequence, mutations that add or remove bases in multiples of three and occur between instead of within codons are likely to be less consequential than others. For frameshift mutations, the protein product is often nonfunctional because all downstream codons are likely to be altered (Kumar, Fausto and Abbas 2004).



Along with the previous descriptions and the images in Figure 3, the sentence construction framework is used to emphasize how changes in gene sequences can have a range of consequences on the protein product associated with a gene.

To illustrate the utility of this framework, consider a sentence comprising a sequence of three-letter words. As described previously, the letters in the sentences represent individual DNA bases in the gene and the words represent triplet codons coding for amino acids. The complete sentence represents the protein. The meaning conveyed by the sentence is also essential for the utility of the sentence construction framework in illustrating and understanding the consequences of mutation. The intended meaning of the sentence represents the function of the protein. In this construct, a variety of mutations can be simulated by adding, removing, or changing one or more letters to/from/in a sentence and assessing the extent to which the originally intended meaning of the sentence is affected. The extent to which the originally intended meaning of the sentence is distorted is a stand-in for how much of the function of a particular protein is changed after a mutagenic event.

Consider the sentence and the consequences of the example “mutations” shown in figure 4.

Consequences of Mutation: A Sentence Model

THE BIG BAD RED FOX RAN FOR THE PEN AND ATE THE HEN BUT DID NOT EAT THE FAT PIG
Example 1 - A missense mutation in which the 17th letter in the sequence is changed to U THE BIG BAD RED FOX RUN FOR THE PEN AND ATE THE HEN BUT DID NOT EAT THE FAT PIG Consequence – The sentence sounds a little awkward, but the intended meaning of the original sentence is largely intact. This represents a mutation with little effect on the function of the protein.
Example 2 - A missense mutation in which the 47th letter in the sequence is changed to S THE BIG BAD RED FOX RAN FOR THE PEN AND ATE THE HEN BUT DID NST EAT THE FAT PIG Consequence – The non-word "NST" is not clearly understood to be NOT and with the change of a single letter, the meaning of the sentence's last clause is exactly the opposite of what it was originally intended to be. This missense mutation, though quantitatively no more than that in the first example, has a more significant effect on the function of the protein because it occurs in a more critical domain of the gene sequence.
Example 3 - A nonsense mutation that introduces a premature stop after the 51st letter of the sequence THE BIG BAD RED FOX RAN FOR THE PEN AND ATE THE HEN BUT DID NOT EAT ~~THE FAT PIG~~ Consequence – The truncated sentence retains some of its intended meaning; some is lost. Some function is lost, but the protein is not completely useless.
Example 4 - A nonsense mutation that introduces a premature stop after the 9th letter of the sequence THE BIG BAD ~~RED FOX RAN FOR THE PEN AND ATE THE HEN BUT DID NOT EAT THE FAT PIG~~ Consequence – Stopped so early, the sentence loses too many critical domains after the mutation and retains essentially none of its intended meaning. This severely truncated protein is useless; retains none of it functions.
Example 5 - A single letter, R, is inserted after 8th letter THE BIG BAR DRE DFO XRA NFO RTH EPE NAN DAT ETH EHE NBU TDI DNO TEA TTH EFA TPI G Consequence – The sentence retains none of its original intended meaning. The insertion of a single base shifted the reading frame resulting in a non-functional protein product.
Example 6 – Three letters, RAD, are inserted after the 9th letter THE BIG BAD RAD RED FOX RAN FOR THE PEN AND ATE THE HEN BUT DID NOT EAT THE FAT PIG Consequence – The sentence retains its original meaning and in fact, has added another descriptor for the fox. The is an in-frame insertion mutation; it can even be seen to represent a gain of function mutation.

Muscular Dystrophy and the DMD Gene

Duchenne muscular dystrophy (DMD) is a neuromuscular disorder affecting roughly one in every 3500-5000 male births worldwide. DMD is characterized by progressive skeletal muscle degeneration and weakness. Early symptoms begin with the legs and pelvis which are associated with the large muscle groups of the legs and gluteal region, but also occur less severely in the arms, neck, and other areas of the body. The first symptoms are usually observed between ages 3 and 5 and often include difficulty standing upright from a laying position and frequent falls. These boys also tend to have trouble climbing stairs. Over time the condition progresses and the motor skill deficits (running, hopping, jumping, etc.) increase. In time, even

walking becomes too challenging and many of these boys are confined to wheelchairs by age 12. Beyond this point sleeping and breathing aids also become necessary. By age 20 most DMD patients are characterized by severe breathing difficulties and cardiac disease. Most do not survive beyond the third decade of life due to respiratory or cardiac complications (Blake, et al. 2002).

In this section of the unit, students' comprehension of the primary functions associated with the muscular system (described previously) is reinforced through analysis of still images and videos made of DMD patients from around the world. Students analyze patient histories of people who lived or are living of with DMD. Students also watch and read, interviews and accounts by DMD patients and their loved ones and caregivers. Where the opportunity is possible, students conduct interviews with DMD patients, their families, and physicians who treat DMD patients. Gaining familiarity with a condition that highlights, in an almost visceral manor, the functional consequences of the living with less muscle, underscores for students the roles that their own muscles play, even when they are not thinking about them.

According to the 3 P's, the functional consequences associated with DMD can be sourced to changes in structure. At the level of muscle tissue, students emphasize the connection between the small boys' small frames and their diminished abilities to move. Students zoom in further to the level of the cell and draw connections between the mutated dystrophin gene and the reduced ability of muscle cells to remain stable during contraction-relaxation cycles.

DMD is caused by an absence of normal quantities of dystrophin, a cytoskeletal protein that stabilizes the plasma membrane of muscle fibers helping to keep muscle cells intact. Loss-of-function mutations in the *DMD* gene coding for dystrophin diminish the muscles' ability to withstand physical stresses owing to lengthening and shortening associated with repeated muscle contractions (Kendall, et al. 2012).

DMD largely affects boys because it is inherited in an X-linked recessive fashion due to the location of the *DMD* gene on the short arm of the X-chromosome. At 2.4 Mb, the *DMD* gene is among the largest in the human genome. Even after introns (nucleotide sequences that are not included in the final RNA product to be translated) are removed from the precursor messenger RNA (pre-mRNA) transcript, the mature *DMD* mRNA transcript produced by splicing its 79 exons (regions coding the final protein product) is still 14 kb in length. The sheer size of this gene and its transcript make the *DMD* gene more susceptible to mutations, but research also shows that certain regions of the *DMD* are mutation hotspots (nucleotide sequences with a higher concentration of mutations (Rodino-Klapac, et al. 2007). These mutations exist in the varieties previously described and are associated with the ranges of consequences modeled by the previous sentence construction analogies.

Given these underlying genetic elements of DMD, the availability of locus-specific databases of DMD mutations is invaluable for research, diagnosis, and clinical care. Advances in gene sequencing technologies have helped in developing just such databases that can aggregate data from groups of DMD patients. The unit introduces students to such data from the United Dystrophinopathy Project (a multicenter research consortium) and the TREAT-NMD DMD Global database. Analyses of these kinds of data of known *DMD* gene mutations find that deletions comprise the majority, but duplications and point mutations are also known (Bladen, et al. 2015).

As modeled in the sentence construction analogy, the characteristics and locations of specific mutations affect the severity of functional consequences. The range of consequences associated with mutations in the *DMD* gene includes but is not limited patients who endure occasional bouts of cramping and myalgia, patients who live with a related but milder form of muscular dystrophy called Becker's muscular dystrophy (BMD), and

patients with rapidly progressing DMD (Bladen, et al. 2015). The availability of mutation databases, gene maps, and the DystrophySNPs resource developed by the University of Utah Genome Depot allows students to make direct connections between changes in the *DMD* gene sequences and consequences in the dystrophin protein. Students' understanding of types of mutations allows them to explore the information in these data sets, then predict and describe consequences likely to occur in the protein. Students are also able to provide reasonable explanations to support the assertions they make. The sentence construction frame work helps the students develop their own understanding and then provides a useful way to communicate meaningful scientific information to science and non-science colleagues alike.

The example of DMD is used in the unit because it so clearly delineates the connection between a change in structure and a corresponding dysfunction. The exploration of gene therapy later in the unit is to emphasize the idea that even complex, cutting edge technological advances are rooted in the central idea that solutions for dysfunctions mollify the effects of abnormal structure.

Part 2 - Mechanisms of Gene Therapy

Whereas part 1 of the curriculum unit was concerned with elucidating the nature of the structural problems that result in the dysfunctions associated with DMD, the focus of part 2 is on how a clear understanding of the abnormal structures can lead to the development of remedies that regain function by restoring structure. This is the purpose of the 3 P's construction.

The molecular basis for DMD has been known for more than two decades and consequential advancements in the management of the disease have been achieved in the that time. Indeed, DMD patients are diagnosed more readily and have available to them systems of interventions that allow them to live comparatively more robust lives for longer than in years past. Still, the treatment options for DMD patients remain limited. The interventions available are largely palliative, seeking to mitigate the symptoms of the dysfunction rather being able to address the structural root causes of the disease. Corticosteroids, cardiac and pulmonary medications, orthopedic supports and procedures, and several other factors are among the critical components of a thorough interdisciplinary approach to managing DMD. Improved muscular strength, prolonged ambulation, and improved respiratory function can be expected for most DMD patients presently (Bushby, et al. 2017).

Strategies attempted in pursuit of curative rather than merely palliative therapies have included both the transplantation of healthy muscle progenitor cells into DMD patients as well gene-therapy mediated delivery of functional copies of the *DMD* gene into patients. Success with the former strategy has been limited by "issues of immune rejection and poor systemic delivery and viability of transplanted cells." That is, the administered cells don't survive after transplantation, and are not able to contribute to muscle function. Success of the latter strategy has been limited by issues associated with the large, complex structure of the gene (Lim, Maruyama and Yokota 2017).

Gene therapy is an approach to address genetic disorders by targeting the source of the disorder; that is, gene therapy refers to treating a disease condition resulting from an abnormal version of a gene by introducing modified DNA products with the correct gene into the cells of the patient. The most intuitive application of this broad concept is to introduce a functioning gene to correct the effects of a disease-causing mutation (e.g., introducing a functional dystrophin protein where none is present or it is present in insufficient quantities). However, gene therapy might also be implemented to inhibit a cell from producing a damaging product or even to completely destroy a problematic cell. In this sense, the basic concept of gene therapy is simple, but several challenges make implementation on a large scale more difficult; these include but are not limited to developing appropriate vectors to reliably deliver the genetic material to the desired sites, avoiding

immune responses, and not disrupting the normal function of non-related genes (Rodino-Klapac, et al. 2007).

As suggested earlier, previous curative strategies for DMD involving gene therapy were unsuccessful at reliably delivering the entire *DMD* gene into all the necessary tissues in significant part due to its large size. Recent advances in molecular biology techniques have given rise to a gene therapy strategy that does not necessitate that the whole *DMD* gene be transferred into cells: exon skipping.

In this exon skipping approach, the translational reading frame of a gene is restored using synthetic nucleic acid analogs called antisense oligonucleotides (AOs). These AOs can bind to pre-mRNA sequences and interfere with splicing. Specific AO sequences can be designed to target complementary sequences in pre-mRNA. The cell's splicing machinery does not incorporate the targeted sequence in the final transcript. In this way, whole exons or even multiple exons can be targeted for exclusion without disrupting the reading frame (Aartsma-Rus and van Ommen 2017).

In the parlance of the sentence construction analogy, rather than having to write a new, error-free version of a long complex sentence, exon skipping simply allows the reader to jump over the portion of the sentence containing the mistake. So long as the reading frame is kept intact, the resulting sentence, though somewhat shorter, may still be able to convey most of the intended meaning of the original sentence.

Figure 5 recalls the previously discussed example 5 (insertion of a single base resulted in a nonsensical protein product) and uses the sentence construction framework to illustrate how exon skipping can fix the effects of certain mutations.

Exon Skipping: A Sentence Model

THE BIG BAD RED FOX RAN FOR THE PEN AND ATE THE HEN BUT DID NOT EAT THE FAT PIG

Example 5 - A single letter, R, is inserted after 8th letter

THE BIG BAR DRE DFO XRA NFO RTH EPE NAN DAT ETH EHE NBU TDI DNO TEA TTH EFA TPI G

Consequence – The sentence retains none of its original intended meaning. The insertion of a single base shifted the reading frame resulting in a non-functional protein product.

Exon skipping approach to remove the damaged section of code and restore the reading frame

skipped

THE BIG BAR DRE DFO XRA NFO RTH EPE NAN DAT ETH EHE NBU TDI DNO TEA TTH EFA TPI G

Result of exon skipping

THE BIG RED FOX RAN FOR THE PEN AND ATE THE HEN BUT DID NOT EAT THE FAT PIG

Consequence – The new sentence is shorter than the original, but still conveys the intent of the original message. The protein is functional

Eteplirsen

In recent months, the US Food and Drug Administration (FDA) granted approval for the first gene therapy drug, using this novel exon skipping strategy, Eteplirsen. The drug uses an AO designed to bind in a complementary way to—and thus exclude—exon 51 from the final *DMD* gene transcript. Skipping exon 51 addresses structural problems associated with the mutated gene in approximately 14% of patients with *DMD* mutations (Lim, Maruyama and Yokota 2017).

In healthy individuals (no mutations), the 2.5-Mb *DMD* gene is transcribed to produce a 21-kb pre-mRNA transcript: 79 exons separated by introns. The cell's splicing machinery then removes the intron sequences from the pre-mRNA and combines the remaining exon regions into a 14-kb mature mRNA transcript. This mRNA transcript is used as template to translate a normal 427 kDa protein of 3685 amino acids.

Patients who may be helped by Eteplirsen have a *DMD* gene with a consequential mutation in or near exon 51 (e.g. a deletion that introduces a premature stop codon or a deletion that shifts the reading frame for the downstream sequence). The RNA produced from these do not result in a functional dystrophin protein.

In Eteplirsen-treated DMD patients the action of the AO binding to a specific portion of the pre-mRNA transcript results in a mature mRNA in which exon 48 is followed immediately by exon 52 and the reading frame is intact. The end result of translation is a shorter, but still functional dystrophin protein (Lim, Maruyama and Yokota 2017).

The FDA approval of Eteplirsen is not without various controversies. This provides opportunities for students to engage with these and other issues from a more informed perspective. In one such issue, some scientists tout the available data from clinical trials and foresee an effective treatment option for certain classes of DMD patients. Other scientists raise questions about the methodologies employed by the researchers and cast a skeptical eye on the reliability of the clinical trial data and on the utility of the drug at this time (Kesselheim and Avorn 2016). Now armed with a clearer, more accessible understanding of the molecular mechanisms associated with the drug's actions, students can now more meaningfully examine the published data from clinical trials and draw conclusions about the efficacy of the drug and reliability of claims made by the manufacturer. Students construct and communicate informed, evidence-based arguments related to contemporary issues in the scientific community. There is a direct and immediate connection between the students' learning in class, and significant real-world implications and applications of that learning. In considering the effectiveness and viability of gene therapy as a treatment for DMD, students engage with the question "*For whom and under what conditions would you recommend gene therapy as a treatment for DMD?*"

Part 3 – Ethical and Moral Issues Raised by Gene Therapy

Although gene therapy seems to be an increasingly-promising treatment option for several conditions, it remains a risky and often controversial technique. In general, gene therapy is only being explored for conditions with no other curative remedies. In the last part of the unit students combine their learning about the mechanisms of gene therapy, their analysis of clinical research and regulatory data along with their capacities both critical and abstract thinking, to engage with ethical and moral issues.

Beyond questions of efficacy and safety, gene therapy is often considered controversial due to associated moral and ethical questions. As with other scientific developments, as the technology improves, questions shift from *'can we'*, to '*should we'*. Among other issues, students consider the following questions presented on the NIH Genetics Home Reference website: *How can "good" and "bad" uses of gene therapy be distinguished? Who decides which traits are normal and which constitute a disability or disorder? Will the high costs of gene therapy make it available only to the wealthy? Could the widespread use of gene therapy make society less accepting of people who are different? Should people be allowed to use gene therapy to enhance basic human traits such as height, intelligence, or athletic ability?*

Classroom Activities

Among the classroom activities to be incorporated in this unit are the following.

Molecular Genetics Sentence Construction - Students construct sentences comprising three letter words. The letters in the sentences are to represent individual DNA bases in the gene. The words are to represent triplet codons coding for amino acids. The sentence, as a whole, represents the protein. Most critically, the intended meaning of the sentence represents the function of protein. The individual letters of the sentences are to be transferred on to separate pieces of paper large enough to be easily manually manipulated (e.g. index cards). Students can simulate various mutations by adding, removing, or changing one or more letters to/from/in their sentences and assessing the extent to which the originally intended meaning of the sentences are affected. The extent to which the originally intended meaning of the sentence is distorted is a stand-in for how much of the function of a particular protein is changed after a mutagenic event. Students can work in pairs or small groups in which they exchange sentences, identify the types of mutations being represented and compare the effects of these mutations. Depending on the availability of technological resources in the classroom, the digital "Mutations" model developed by The Concord Consortium (https://concord.org/stem-resources/mutations) can also be incorporated in this activity. The "Mutations" model allows students to create and edit DNA sequences and observe the effects changes in DNA sequence on the peptide sequence as well as the interactions among the peptides themselves.

One possible assessment method for this learning activity also emphasizes the random nature of many mutagenic mutations. The instructor has two containers with slips of paper, one container with "mutation instructions" (e.g. "remove letter", "add 3 letters", "repeat letter four times", etc.) and the other container with "location instructions" (e.g. 5th letter). Students chose at random one paper from each container, then must identify/describe the type of mutation resulting and assess the extent of the consequence on the *gene's function*.

Students can also by assessed by creating a large size poster display that show the sentence construction framework and explicitly makes the connections to gene mutations (with examples).

Data Analysis and Conclusions – One of the emphases of this unit is analyzing and interpreting data, one of important NGSS Science & Engineering Practices. In this learning activity students are provided data sets, tables, graphs, etc. from published research journal articles. Students then are asked to examine the data and use a graphic organizer to make sense of the information presented. The graphic organizer is simple and contains three boxes, each with a set of sentences intended to guide the student's exploration of the date.

Box 1 questions - Describe the data (*what kind of information, how is it organized, where did it come from, etc.*). Box 2 questions - What *patterns/trends/correlations* do you notice in these data? Which trends stand out the most? Do all the data fit these patterns (*are there any outliers?*)? Box 3 questions - What do the patterns/trends/correlations lead you to *conclude*? (*what might the trends and patterns mean?*)

In subsequent lessons and units, students will use the insights from data analysis done in this manner to construct scientific claims the supported by data and structured in a coherent way.

Letter to the Author/Editor – A goal of this unit, and indeed the course overall, is to develop students who are more critical consumers of scientific information. My students sometimes have a tendency to accept

information presented to them without critical examination. I remind students that when assessing the reliability of information, it is important to consider how an author/speaker presents an argument, but it is also important to keep in mind that even well-organized data is sometimes flawed and eloquent speakers can be wrong. At the very least, scientists often disagree in whole or in part with information that is presented to them. In these cases, in is important to be able to assess the information presented to you and construct a coherent argument of your own.

The controversy regarding the recent FDA approval of Eteplirsen provides a useful example to emphasize this point. One of the important texts used in this unit is from the opinion section of the December 13, 2016 issue of the Journal of the American Medical Association (JAMA) – Volume 316, number 22. In the piece doctors Aaron S. Kesselheim and Jerry Avorn, make the case that the approval of Eteplirsen is problematic because of certain methodological and procedural issues (e.g. small sample size of the study group). In this learning activity students compare these doctors' critiques with the data and arguments presented by the original researchers and then present their arguments as to the validity of the arguments of either party. Following this, students apply critical eyes to conclusions drawn from research data and write a "letter to the editor/author" in which the present critique the research and its conclusions. The emphasis here must be on the strengths and weakness of the research/data (*"What was done well? What can be improved? What other questions should have been asked?, etc..)* and crafting an evidence-based argument.

Bibliography

Aartsma-Rus, Annemieke, and Gert-Jan B van Ommen. 2017. "Less is more: therapeutic exon skipping for Duchenne muscular dystrophy." *The Lancet Neurology* (Elsevier) 8 (10): 873-875.

Bladen, Catherine L., David Salgado, Soledad Monges, Maria E. Foncuberta, Kyriaki Kekou, Konstantina Kosma, Hugh Dawkins, et al. 2015. "The TREAT-NMD DMD Global Database: Analysis of More than 7,000 Duchenne Muscular Dystrophy Mutations." *Human Mutation* 36 (4): 395-402.

Blake, Derek J., Andrew Weir, Sarah E. Newey, and Kay E. Davies. 2002. "Function and Genetics of Dystrophin and Dystrophin-Related Proteins in Muscle." *Physiological Reviews* 82: 291-329.

Bushby, Katharine, Richard Finkel, David J Birnkrant, Laura E Case, Paula R Clemens, Linda Cripe, Ajay Kaul, et al. 2017. "Diagnosis and management of Duchenne muscular dystrophy, part 2: implementation of multidisciplinary care." *The Lancet Neurology* (Elsevier) 9 (2): 177-189.

Griffiths, Anthony J.F., Richard C. Lewontin, Jeffrey H. Miller, David T. Suzuki, and William M. Gelbart. 2005. *Introduction to Genetic Analysis, Eighth Edition.* New York, NY: W H Freeman & Co.

Kendall, Genevieve C., Ekaterina I. Mokhonova, Miriana Moran, and Natalia E. Sejbuk. 2012. "Dantrolene Enhances Antisense-Mediated Exon Skipping in Human and Mouse Models of Duchenne Muscular Dystrophy." *Science Translational Medicine* 164ra160.

Kesselheim, A S, and J Avorn. 2016. "Approving a problematic muscular dystrophy drug: Implications for fda policy." *JAMA* 316 (22): 2357-2358.

Kumar, Vinay, Nelso Fausto, and Abul Abbas. 2004. *Robbins & Cotran Pathologic Basis of Disease, Seventh Edition.* Philadelphia, PA:

Elsevier Saunders.

Lim, Kenji Rowel Q., Rika Maruyama, and Toshifumi Yokota. 2017. "Eteplirsen in the treatment of Duchenne muscular dystrophy." *Drug Design, Development and Therapy.*

Rodino-Klapac, Louise R., Louis G. Chicoine, Brian K. Kaspar, and Jerry R. Mendell. 2007. "Gene Therapy for Duchenne Muscular Dystrophy: Expectations and Challenges." *JAMA Neurology* 1236-1241.