Understanding & Teaching Genetics Using Analogies

● SCOTT WOODY, ED HIMELBLAU

Abstract

We present a collection of analogies that are intended to help students better understand the foreign and often nuanced vocabulary of the genetics curriculum. Why is it called the "wild type"? What is the difference between a locus, a gene, and an allele? What is the functional (versus a rule-based) distinction between dominant and recessive alleles? It is our hope that by using these analogies, teachers at all levels of the K–16 curriculum can appeal to the common experience and common sense of their students, to lay a solid foundation for mastery of genetics and, thereby, to enhance understanding of evolutionary principles.

Key Words: *Genetics; analogies; locus; allele; dominant; recessive.*

As famously written in the pages of this very journal some 60 years ago, "Nothing in biology makes sense except in light of evolution." Given that the science and language of genetics are key to understanding evolution, and evolution is the key to understanding biology, we will argue that building a strong foundation based on a clear understanding of terms used in genetics education is essential for effective student education in biology.

There are two significant challenges that face both teachers and students in the course of the genetics "unit" in the biology curriculum. First, many of the terms used repeatedly in genetics have no common usage outside of that realm. There is no "real world" situation in which terms like *heterozygote* or *genotype* – or for that matter *allele* and *locus* – are used. It really is as if students are being asked to learn a foreign language, with all the attendant challenges that may entail. Second, many genetics terms are defined in opposition to similar-sounding terms whose meanings are quite distinct: *genotype* versus *phenotype*, *heterozygous* versus *homozygous*, *dominant* versus *recessive*, and *wild type* versus *mutant*. It seems nearly certain that if students do not understand one of those paired terms, they surely cannot understand the other.

In our teaching experience in both laboratory and classroom settings, we have witnessed the power of an apt analogy to help

A strong foundation based on a clear understanding of terms used in genetics education is essential for effective student education in biology.

students understand the vocabulary of genetics. In what follows, we will describe several that we believe are particularly effective in helping students gain a lasting comprehension of this fundamental biological science. We readily admit at the outset that our belief in their utility rests on intuition that has been bolstered by interactions with students and teachers alike. We will confess, further, that the analogies on offer here have no grounding in cognitive sciences research; they are simply approaches that we have developed and believe to be useful. They are offered "as is" and in what we hope is a conversational and accessible style that is of use to readers of *ABT*.

○ Locus versus Gene

The term *locus* (Latin for "place"; plural *loci*) refers to the physical site of a particular genetic element along a chromosome. *Gene* can be synonymous with *locus*. The distinction is subtle but essentially reflects what is known about the "particular genetic element," a phrase in the definition that is deliberately vague. Specifically, if all we know is that a mutation exists that causes a perceptible change in the phenotype of an organism upon transmission to progeny generations, it is best to use *locus*, because that usage does not imply that we know much or anything about the molecular mechanism(s) by which the phenotypic effects come about. Use of the term *gene* becomes appropriate when we have investigated the locus to understand the functional role of associated DNA sequences. Thus, *gene* refers to all protein-coding sequences as well as regulatory elements that govern expression of the encoded function.

A straightforward analogy useful to explain the difference between loci and genes is street addresses. Imagine a fictional small town named Rustic (Figure 1, upper panel). We might say, "There is an occupied lot at 267 Main Street." This doesn't tell us anything, other than the fact that there is a building of some sort at that address that has some function (*locus*). However, if we know that it's a business that sells
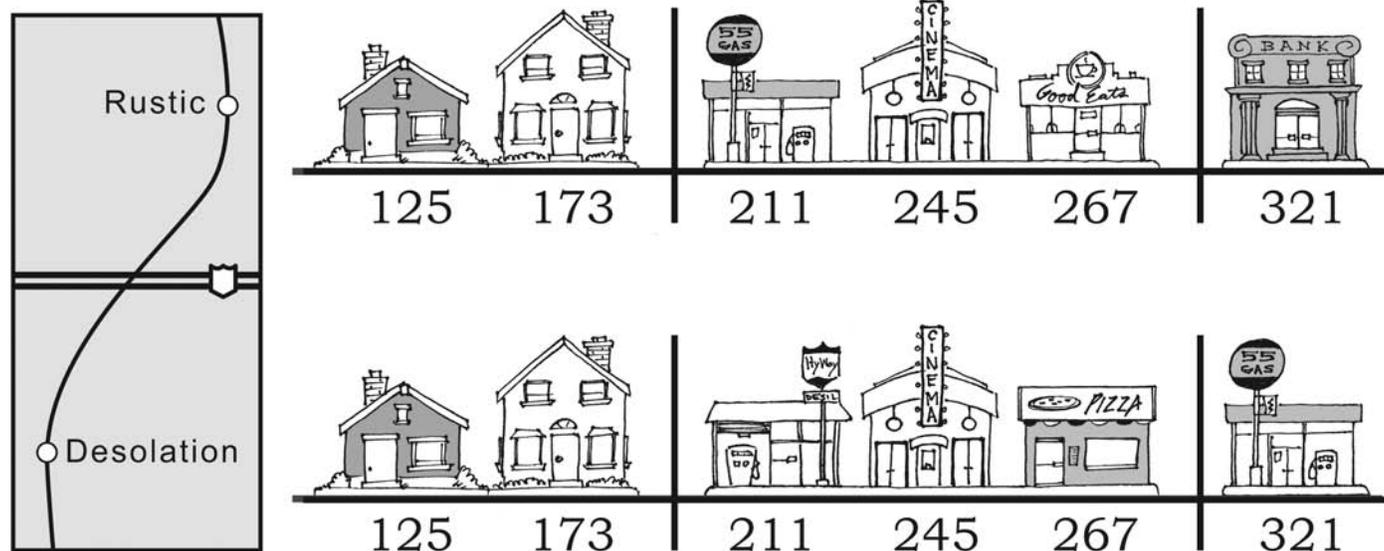
Figure 1. Towns of Rustic and Desolation.

prepared foods and is called "Good Eats," we explicitly assert that we have considerably more knowledge of many things regarding the structure and activities at 267 Main Street, perhaps including the quality of the food, the likelihood of getting a table during the dinner hour, the price of entrees, and so on. That is, we can be much more specific and authoritative in our choice of vocabulary (*gene*).

Locus versus Allele

Unlike the nearly synonymous terms *locus* and *gene*, *locus* and *allele* have decidedly different meanings. In genetics terminology, *allele* refers to a DNA sequence that differs from an alternative DNA sequence that might instead occupy the same *locus*. Please note that the particular DNA sequence difference that distinguishes between alternative alleles may or may not have functional consequences on the encoded product or genetic function and, thereby, affect the phenotype of an individual; the mere existence of such differences is sufficient to constitute alternative allelic forms – provided that they occupy the same chromosomal location (locus). Expanding on the small-town analogy introduced above, let's imagine that the town of Rustic lies just north of a very similar small town named Desolation (Figure 1, lower panel). Imagine that both Rustic and Desolation were originally cookie-cutter replicas of planned communities drawn from the same master plan by the county planning commission. Driving into either town on Main Street takes us through a residential area of modest homes and then into the business district with its shops, banks, gas stations, and restaurants – typical small-town fare.

Observant travelers might initially be struck by the similarity of Rustic and Desolation: homes in the 100 block of Main Street in both towns are of the same size, shape, and color, and even the business districts are similar, although not identical. That is, the 200 block in each town has a gas station, a movie theater, and a restaurant. However, although the movie theaters are the same, the gas stations and restaurants are of different kinds. A more striking difference is evident at 321 Main Street: in Rustic there is a bank, whereas at the same address in Desolation there is another gas station.

Let's abandon the travelogue and return to the analogy related to loci and alleles. Lot 125 Main Street as well as 173, 211, 245, and 267 Main Street (loci) are, in pairwise comparisons, allelic. The homes and the movie theater at 245 Main Street are identical and so do not constitute *alternative* alleles, but they can be considered allelic (put Rustic and Desolation together as a pair and you have the equivalent of the homozygous condition). The gas stations at 211 Main Street and the restaurants at 267 Main Street are not identical but they serve essentially the same function and so, too, can be considered as allelic forms (in heterozygous condition). What are we to make of the lot at 321 Main Street, at which Rustic has a bank and Desolation has another gas station? If we accept the premise that both towns were designed from the same master plan, this difference may seem a bit perplexing, but a reasonable evolutionary scenario can be imagined that plausibly explains the disparity. Just as the gas stations and restaurants in the 200 block have diverged since the towns were founded, other lots in Rustic and Desolation have experienced their own evolution over time. Perhaps the bank in Desolation failed at some point and, after a period in which the remains of the building decayed and ultimately collapsed, a new business formed on the same lot at 321 Main Street – in this case, a new gas station, the same as the old gas station at 211 Main Street in Rustic. Are the establishments at the 321 Main Street addresses allelic? The answer is no: there are no structural or functional similarities between a bank and a gas station, so even though they share the same address, the ultimate test of allelism is functional similarity; thus, the establishments at 321 Main Street do not fit the bill. Are the gas stations in Rustic and Desolation allelic? No, they reside at different addresses (loci). That conclusion begs one last question that we will leave to readers (and hopefully their students) to decide: Are the gas stations at 211 Main Street in Rustic and at 321 Main Street in Desolation allelic?

Wild Type versus Mutant Alleles

A common point of confusion for students is the term *wild type*, as often casually applied to both phenotypes and genotypes without clear definition. Let's cover *wild type* (*wt*) first and then turn

to an analogy to help students appreciate the distinction between *wt* and mutant alleles. As geneticists, we typically select a particular animal, plant, or other organism to serve as the prototype of the kind. "Model" organisms such as *Drosophila melanogaster* (insects, metazoans), *Arabidopsis thaliana* (flowering plants), and *Caenorhabditis elegans* (nematodes) have been selected as common objects of study in expectation that what we learn through research on those model organisms will be broadly applicable in explaining biological processes in related species. A secondary but very important consideration in selecting model organisms is determining which organisms are experimentally tractable: do we have the tools available to advance our research?

We've developed an analogy to help students appreciate why we study fruit flies, weeds that grow in sidewalk cracks, or small worms that grow in soil. We ask them to imagine being a space alien that just landed on planet earth. Looking around from the center of New York City, our alien selves notice that among the many things in this foreign environment are automobiles of varying shapes and sizes. Our innate curiosity compels us to try to understand how cars work, but given the wide range of alternative models available and our limited resources, studying all models is not possible. Thus, the first choice we face is "Which car model should we study?" One answer might be to select the Lamborghini, a very expensive and relatively rare model, but surely an exceptionally spectacular automobile; an alternative might be to select a Toyota Camry, a fine but common and somewhat unexceptional model.

Which should we choose? Clearly it is the Camry: its design and operating features are shared in common with many other cars that we see on the road, and both tools and repair parts are readily available. Most importantly, what we learn by studying a Camry seems likely to apply to Corollas, Chevrolets, and perhaps even Saabs; it is doubtful that studying a Lamborghini could yield similarly generalizable results.

Would study of a Camry necessarily imply that all of its parts and operating features are optimal forms in relation to all other cars? The obvious answer is no: What we learn from a focused study of the Camry merely establishes a model for how a common automobile and its subsystems operate. As our studies expand to include other models and brands, we will likely find both similarities and differences, and some of the latter may prove superior or inferior to those incorporated into the design of the Camry. Here, then, the Camry serves merely as a reference type, just as a "wild type" *Arabidopsis* plant of the ecotype Columbia has served as a well-studied model of flowering plants.

○ Dominant versus Recessive Alleles

Now let us turn to another bugaboo of genetics vocabulary: the relationship between *dominant* and *recessive* alleles. Many students have a rather superficial, rule-based understanding of the distinction: Dominant alleles are those whose effect on phenotype is evident when present in a single copy; conversely, recessive alleles must be present in homozygous condition in order to produce a detectable change in organismal phenotype as compared with that of the (phenotypically) *wt* individual. Those rules certainly apply, but they fail to convey any sense of the molecular basis for why those rules exist in the first place. To lay the groundwork for the analogy that follows, recognize that mutant alleles that arise spontaneously through errors in DNA replication or damage-repair processes, or mutant alleles that result from chemical mutagenesis of the reference *wt* genome, typically yield what geneticists refer to as "loss of function" alleles, where the coding potential of the locus has been altered in such a way that no functional gene product can be produced from the altered DNA template – say, introduction of a stop codon in a protein-coding gene. In most cases, recessive alleles have little or no effect on phenotype because the remaining *wt* allele in heterozygotes can suffice to provide the required function. But in the homozygous state, the near or total lack of functioning gene product may impose phenotypic consequences upon individuals that lack the *wt* or other functional allele. The FPsc *albino* mutant allele is a good example of a simple recessive allele (Figure 2). In a segregating F2 population as shown in the figure, a quarter of individuals in the population are unable to properly elaborate the photosynthetic machinery, and so they grow as weakly pigmented (yellow/white) seedlings. Although we do not yet know the precise molecular basis of the *albino* phenotype, the recessive nature of the mutant allele as compared with the *wt* alternative makes it nearly certain that the *albino* mutation has disrupted a locus that is absolutely essential in order for plants to "be" green.
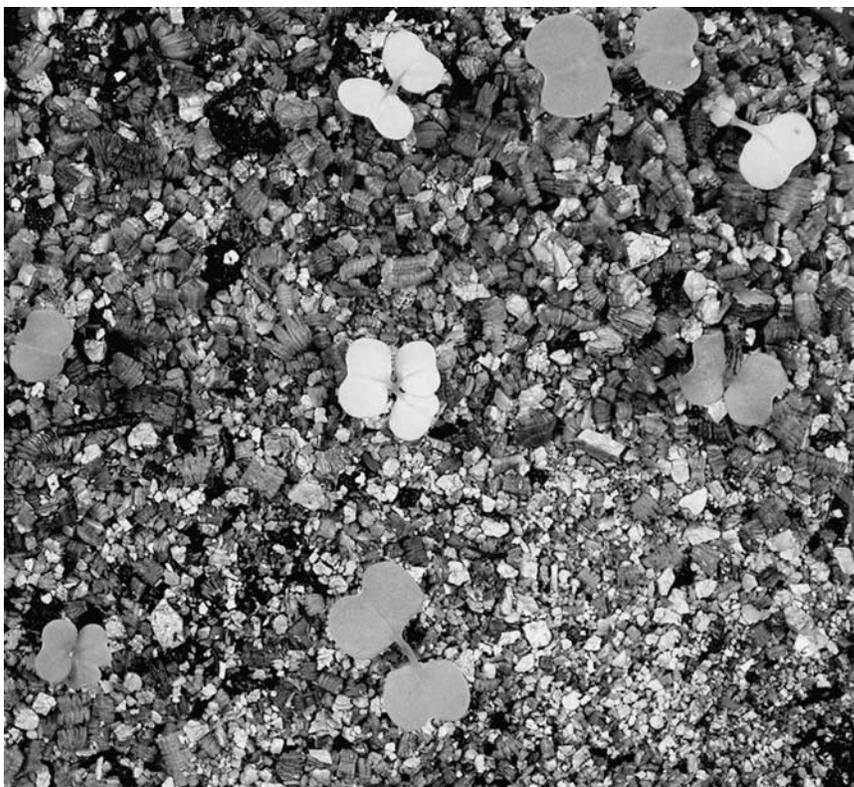

Figure 2. The recessive FPsc *albino* mutant allele prevents elaboration of the photosynthetic machinery (white segregants among the F2s shown above) if present in homozygous condition.

Figure 3. The dominant FPsc *ale* mutant allele causes pronounced leaf curling.

On the other hand, dominant alleles often result from "gain of function" mutations whose effects are evident even if present in just a single copy. As a hypothetical example, imagine a mutation that alters the substrate specificity of a key metabolic enzyme such that the encoded gene product enables the organism to harvest energy from an additional carbohydrate source in its environment – for example, let's say an organism carrying the mutant allele, expressed in precisely the same tissues and at the same level of expression as its *wt* counterpart can digest both glucose *and* sucrose, whereas individuals that carry only the *wt* allele must live on glucose alone. If sucrose is plentiful in the environment and carbon sources are growth-rate limiting, the heterozygote will likely enjoy a fitness advantage over the *wt* homozygote.

Alternatively, dominant alleles may arise via mutations that alter the temporal or spatial pattern of gene expression. The dominant FPsc *abnormal leaf* (*ale*) allele is a case in point (Figure 3). The *ale* locus encodes a "master" regulatory factor that governs allocation of cell types in the course of leaf development between those that serve evolved functions and reside on the upper (adaxial) leaf surface and those that differentiate to serve the functions of cells that occupy lower (abaxial) tissues in mature leaves. Proper allocation of adaxial and abaxial cell types is essential for formation of the flat solar panels we recognize as leaves. Misallocation or overproliferation of primordial cells into one class over the other during leaf development results in formation of curled leaves. Fascinating though they may be, we need not go further into the molecular details of the *ale* mutant phenotype here; it is sufficient to note that dominant mutant alleles can have effects even if present in a single copy, since it only takes "one bad apple" to perturb the developmental fate of cells during organ formation.

To help students better appreciate (and care about!) the molecular basis that distinguishes between loss-of-function recessive alleles and gain-of-function dominant alleles, we have used the following musical analogy to good effect with both students and teachers (Figure 4):

Imagine a symphony orchestra that intends to play a piece that includes a soaring violin solo that is the musical climax of the work. Imagine,

further, that each bit of sheet music provided to the assembled musicians calls for two players of each part: two violin soloists, two oboe parts, two timpani scores, etc. Now, during that violin "solo" passage, the orchestral score directs the timpani to observe a 60-measure rest while the twin first-chair violinists bring the listening audience to tears (Figure 4A).

What if the sheet music for one of those violinists had been garbled to the point of uselessness during the printing process, an error that had gone unnoticed until the evening of the performance? Unable to perform, the violinist with unreadable sheet music would effectively be silenced, but the violinist with the proper sheet music could continue nonetheless, playing a bit louder if necessary to achieve the desired and dramatic effect; the "phenotype" of the movement would likely be little changed (Figure 4B). However, if both violinists' sheets were similarly garbled and neither had memorized the passage, nearly all is lost: the musical piece fails to reach its emotional crescendo and the phenotype has changed decidedly for the worse (Figure 4C).

The preceding neatly illustrates the practical consequence of most recessive, loss-of-function mutant alleles: as long as a good copy of the locus remains available (as in the heterozygous condition), there may be little or no effect on phenotype, but if no functional copy of the locus is available, as would be the case in the homozygous condition, the phenotypic effects of that loss of function may be manifest.

Now imagine that the printing company charged with producing sheet music for the various musicians in the orchestra was, well, just really bad at the job. Specifically, imagine that of the two pieces of sheet music distributed to the timpanists, one properly indicated a 60-measure rest during the violin solo but the other had a typo that, instead, specified merely a 6-measure rest. Suddenly, and just as the violin solo begins to soar, the percussionist – dutifully following his or her errant instructions – breaks in, with the result that all the music that follows is arrhythmic and discordant (Figure 4D). Even though the second timpanist properly remained silent, the musical performance was doubtlessly altered, and probably for the worse; the phenotype had changed and all it took was one altered bit of encoded information.

This last bit nicely models the phenotypic consequences that affect leaf development in the FPsc *ale* mutant: misexpression of the locus whose activity is controlled, at least in part, by the regulatory sequence altered in the dominant *ale* mutant allele has profound downstream developmental consequences.

○ DNA as a Musical Score, *Not* a Blueprint

The preceding brings me to the last analogy we wish to present. Many biology teachers make use of analogies in many different contexts,
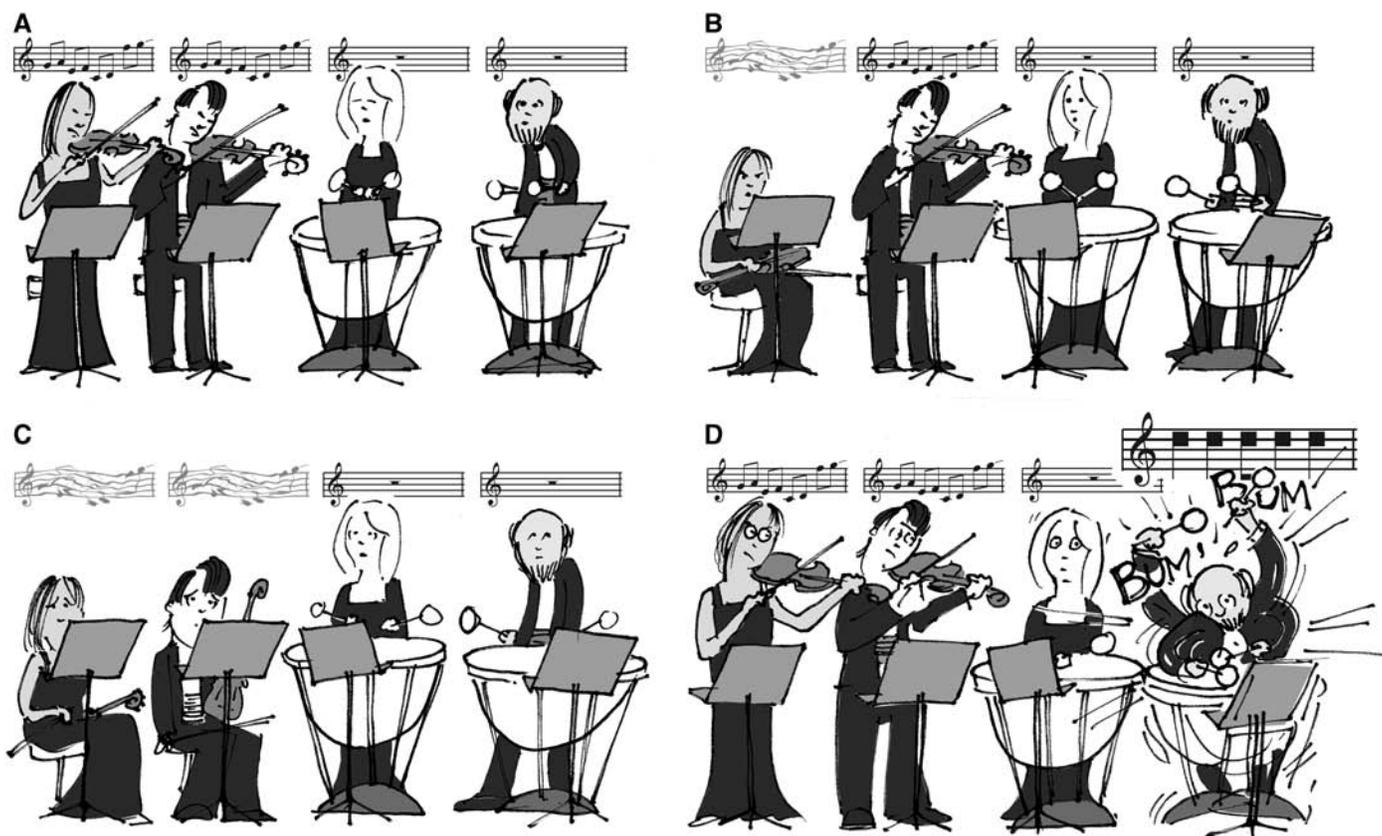
Figure 4. Musical mutations. (A) With all sheet music printed correctly, the violins play while the timpani rest. (B) Garbled sheet music prevents one violin from performing, but the second violin can perform and the piece proceeds. This represents a heterozygous genotype in which the functional (wild type) allele is dominant over a recessive, loss-of-function allele. (C) If the sheet music for both violins is garbled, the piece cannot be performed. This represents the phenotypic effects of a homozygous recessive genotype. (D) The sheet music for the violins is correct, but one timpani score is garbled, prompting the timpanist to play when he should rest. Even though one timpanist has the correct (wild type) score, a single mutant copy of the score is sufficient to disrupt the performance. This is analogous to a dominant, gain-of-function mutation.

but when it comes to genetics and the role of DNA as the mechanism by which evolved information is both stored and expressed, many turn to the ever popular "DNA as the blueprint of life." To do so is understandable, if for no other reason than the fact that it is the analogy frequently used in most print and visual media reports that appear in the press. However, the analogy fails spectacularly to convey to our students the role of DNA.

To wit, a blueprint is essentially a two-dimensional representation of three-dimensional objects that are intended to be produced. If one looks at the blueprint for a room or a house, it is almost immediately apparent what those entities will look like and what functions they will serve: an entryway here, a window there, and a bathroom in the corner, and so on. However, DNA sequences in the raw form of A's, C's, G's, and T's carry no such implications regarding either form or function of the encoded element. To imply otherwise by invoking a blueprint analogy suggests that genes on a chromosome somehow resemble in form and function the product or activity they encode, but nothing could be further from the truth.

DNA sequences that comprise the genome of an organism or species can much more usefully be analogized as the complete orchestral score carried by the maestro to the podium prior to the intended musical performance. Each note (nucleotide) and each piece of sheet music (gene) instructs the performers on how to produce the intended melodic outcome (the phenotype, as expressed from the genotype, as encoded in musical notations of half-notes, chords, rests, etc.). Excepting those who are preternaturally able to "hear" the music by inspection of the complete score as indicated on the sheet music, a bit of translation and elaboration is required before the performance and expression of the musical phenotype become apparent. Beyond the fact that it is considerably more apt than a blueprint analogy, the DNA-as-musical-score analogy has the additional virtue of injecting a dynamic sense to patterns of gene expression and the arrangement of biological processes that program growth and development. It is also quite flexible and may be extended as you prefer to include notions of harmonic interplay among different musical sections.

Finally, and assuming that the reader accepts the musical-score analogy, the analogy can be extended even further to incorporate epigenetic mechanisms: *adagio*, *pianissimo*, and *allegro* are all musical notations imposed over raw musical scores (nucleotide sequences) to indicate the pace or volume with which the encoded notes are to be expressed; it is by now well established that various modifications to DNA bases or to associated histone proteins can have significant influence over expression of the native DNA sequence. As a final embellishment to complete the analogy, one could point out

to students the feedback loops that exist between musicians and the audience (the environment). This is a much richer, more vivid, and more evocative analogy than the essentially static and erroneous view of the genome as portrayed by the blueprint analogy.

○ Conclusion

It is our hope that the analogies presented here might be useful to teachers in helping their students master the vocabulary and vernacular of genetic science. Just as mastery of a second language can be advanced by a "total immersion" experience gained by traveling abroad, so too can mastery of genetics be advanced by constant use of the associated vocabulary and attention to its nuances. It must be noted that all analogies, as abstractions from the primary subject of interest, will inevitably break down if taken to logical extremes. For example, and with particular reference to the car analogy, what would it mean if, say, the stereo of a BMW was deemed "superior" to that of the Camry? Does that translate to a dominant/recessive relationship? We will withhold judgment on this call, except to note that extending the analogy to consider the point could be useful in helping our students appreciate that dominance relationships between alleles are meaningful only in a pairwise comparison of function in a common circumstance. In any case, we would consider it "mission accomplished" if use of any of the preceding analogies prompted objection from your students, who might recognize the breaking point at which they become inapt.

A recurrent theme within the analogies we suggest here is an attempt to clarify the sometimes nuanced distinctions between alleles and loci, and between wild type, dominant, and recessive alleles. This emphasis is entirely intentional and also timely. A common theme of the Next Generation Science Standards and revised AP Biology curriculum is the centrality of evolution to all areas of biological study. Evolution may be broadly defined as the change in allele frequencies within a population over time. That definition usefully encompasses both of the major drivers of evolutionary change – natural selection and genetic drift – and quite properly emphasizes allelic differences as the primary agents through which evolutionary processes operate. We trust that it is self-evident that a clear understanding of what an "allele" is would be useful to help our students grasp the larger picture, and we hope that the approaches advocated here can advance our mutual goals for more effective genetics education, in particular, and biology education in general.

○ Acknowledgments

SCOTT WOODY is Director, FPsc Genetics Resource Development Program, Department of Biochemistry, University of Wisconsin-Madison, 207 Biochemistry, 433 Babcock Dr., Madison, WI 53706, USA; e-mail: swoody@wisc.edu. ED HIMELBLAU is Associate Professor in the Department of Biological Sciences, California Polytechnic State University, San Luis Obispo, CA 93407; e-mail: ehimelbl@calpoly.edu