ONLINE INQUIRY & INVESTIGATION

Using Playing Cards To Simulate a Molecular Clock

KARIN E. WESTERLING

Changes in DNA base-pair order may serve as an indicator of the time elapsed since divergence from a common ancestor. Before DNA could be easily sequenced, comparisons of amino acid sequences in proteins were analyzed to determine divergence. Kimura (1983) compared changes in the amino acid sequence of the a- and b-globin chains of hemoglobin with the estimated time of divergence as determined by the fossil record. The result was a remarkable constancy in the rate of amino acid change regardless of generation time.

DNA sequences can now be analyzed. Some changes in DNA sequence are silent; they do not change the amino acid sequence of a protein. Evolution of these silent, and probably neutral, changes may be faster in organisms with short generation times. DNA mutations that result in a change in amino acid sequence appear to be constant over time regardless of generation time (Ridley, 1996). Some controversy exists as to whether neutral drift or natural selection drives the molecular clock. Possibly both phenomena play a role.

Nevertheless, the rate of change can be calibrated using multiple independent lines of evidence including the fossil record (Rambaut & Bromham, 1998). This calibrated rate of change is the molecular clock. A molecular clock may be constructed using either DNA nucleotide or protein amino acid sequences. It is not an exact time piece such as a Swiss chronometer or metronome; rather it is a relative time-keeper akin to a medieval clock with no minute hand. Caution must be used when analyzing different types of DNA and organisms. For example, the mitochondrial DNA of mammals may change two to four percent in 1,000,000 years while shark mitochondrial DNA changes at a significantly slower rate, about one percent in 6,000,000 years (Martin, 2003). At a minimum the molecular clock is useful for creating phylogenetic trees showing the order of divergence from a common ancestor. This is analogous to keeping time at school where second period comes between first and third regardless of the exact clock time.

KARIN E. WESTERLING is a science teacher at Matthew Gage Middle School, Riverside, CA 92506; e-mail: kwesterling@rusd.k12.ca.us.

Purpose

The purpose of this simulation is to allow students to observe the accumulation of changes in a randomly mutating sequence of playing cards. The cards are analogous to DNA nucleotide or protein amino acid sequences. In this simulation, the accumulation of mutations is solely the result of random drift; there is no natural selection. Although the simulation does not address all aspects of, nor complications in interpretation of, a molecular clock, students learn the general concept. Subsequent to this activity, they will be able to analyze actual DNA base-pair or protein amino acid sequence data with understanding. They will also see that molecular clocks are useful for limited periods of time. Eventually, the clock may "run down" and will no longer be able to provide useful data for the construction of a cladogram.

National standards addressed in this lab include Life Science Content Standard C (Grades 5-8: Diversity and Adaptations of Organisms, and Grades 9-12: Biological Evolution) as well as History and Nature of Science Content Standard G (Grades 5-8: Nature of Science). This lab may be used with a broad spectrum of student ages and abilities depending on the depth explored and the level of teacher assistance.

Concepts To Be Explored Prior to This Lesson

Students should understand that DNA is a template for the transcription of mRNA and that the information from mRNA is translated to synthesize protein. They should know that genetic differences between organisms can be observed directly by comparing the nucleotide order of the DNA or indirectly by inferring DNA base-pair order using the amino acid order of proteins. Comparison of amino acid order of proteins may reveal fewer changes than direct observation of DNA because genes may be broken up along the DNA (introns and exons), many amino acids have redundant codons, and proteins may undergo post-translational modification. This simulation ignores such complications because both DNA and protein comparisons can be useful in determining phylogeny.

Figure 1. Glossary

Cladogram: Is a branching tree diagram used in cladistics to show the evolutionary relationships between species. Cladistics groups organisms which share derived characteristics to determine phylogeny.

Cytochrome-c: Is a protein found in the membranes of the mitochondria. Cytochrome-c is involved with electron transfer in the electron transport chain. Since the sequence of amino acids in cytochrome-c is highly conserved, it is useful for determining phylogeny of a broad spectrum of eukaryotes.

Exons and Introns: Exons are the regions of DNA that are transcribed into RNA and are retained in the messenger RNA (mRNA). Exons are separated by introns. Introns are the regions of DNA which are transcribed but are then removed prior to the mRNA being translated to synthesize protein. The fragmenting of genes into exons and introns is common among eukaryotes but is usually not found in prokaryotes. This simulation does not deal with exons or introns.

Generation time: Is the amount of time needed for members of a species to reproduce the next generation. Fruit flies have a generation time of a few weeks while humans have a generation time of a few decades.

Hemoglobin: Is an oxygen-carrying protein. All tetrapods have both a-globin and b-globin. In adult mammals, hemoglobin is a tetrameric protein complex composed largely of two pairs of globular proteins: two a-chains and two b-chains. A heme group is bound to each chain. Hemoglobin amino acid sequences are sufficiently conserved to provide phylogenetic information within some tetrapod families.

Hemoglobin genes: The genes coding for the a- and b-globin proteins appear in gene clusters. Each gene cluster contains a set of similar globin genes which are expressed at different stages of development: embryo, fetus, and after birth. The gene cluster on human chromosome 16 includes z1, z2, a1, and a2 genes as well as two nonfunctional pseudogenes. The gene cluster on human chromosome 11 contains the e, two g, d, and b-globin genes and a nonfunctional pseudogene.

Hot spot: Is a region of DNA with a high rate of mutation.

Molecular clock: Is a method used to date the divergence of two species. The elapsed evolutionary time is deduced by counting the number of amino acid substitutions (a protein clock) or nucleotide substitutions (a DNA clock). The fewer the differences in the amino acid or nucleotide sequence, the more recent the common ancestor. This simulation models the molecular clock.

Myoglobin: Is another oxygen-binding protein with an amino acid sequence similar to hemoglobin. Myoglobin genes are found on human chromosome 22. Hemoglobin and myoglobin appear to have diverged about 450 million years ago.

Phylogeny: Refers to the evolutionary history of organisms and the relationships between organisms.

Pseudogenes: Have a nucleotide sequence similar to active genes, but the sequence does not yield a functioning gene product. Pseudogenes and active genes often appear in clusters and are derived from an ancestral gene that was duplicated.

Prior to this activity, students should be introduced to branching tree diagrams. They have practiced making pedigrees. They noticed that siblings share a parent and are located on a short branch while cousins share a more distant common ancestor, a grandparent. This lesson may be conducted during, or shortly after, the introduction of cladograms (branching trees which show the phylogeny of species over evolutionary time).

Students should also have observed that as organisms diverge from a common ancestor they become more different over time. For example, siblings are often more similar in appearance than cousins, while cousins may be more similar than complete strangers. Especially in classes of younger students, the teacher expects a student to mention a family where two cousins look almost identical while two siblings appear quite different. At this point it should be emphasized that phenotype is a function of one or more genes and their products interacting with the environment. Phenotype is often what we observe when looking at whole organisms, and natural selection acts upon phenotype. However, it is the DNA sequence, the genotype, which is actually inherited. Observations of DNA sequences are the most likely to reveal phylogenetic relationships. Because the mRNA code is redundant (many amino acids have multiple codons), observations of amino acid sequences may omit some changes in the DNA, but are still often useful. The teacher may want to bring up additional reasons for proteins not reflecting the exact DNA sequence in advanced high school classes, but such information is not necessary for this activity.

Materials

Each lab group needs two decks of 52 playing cards. It does not matter if the Jokers are used or not. Students also use six-sided dice. Four dice per group is ideal, but one or two is sufficient. Graph paper is used to analyze the data. Additionally, the teacher may prepare examples of card sequences to explain the simulation game. See the Sample Data Examples accompanying the method. The activity requires 45 to 90 minutes of class time.

Method

Preparing the Simulation

1. Listen to your teacher introduce, and explain the rules of, the molecular clock simulation game. Carefully observe the sample pairs of rows of playing cards. Each of these rows represents a strand of DNA. Then obtain playing cards and dice for your group.
2. Lay out a tableau of cards consisting of two matching rows of cards as shown (Figure 2). These two rows of cards represent a sequence of DNA bases in siblings. The rows are identical because the siblings inherited the same strand of DNA from the parent. Note that the card's value (e.g., Ace, 2, 3, King) is important. Suit (e.g., heart or spade) does not matter.

Figure 2. Step 2 Example. Note that the card's value (e.g., Ace, 2, 3, King) is important. Suit (e.g., heart or spade) does not matter. Each row of cards represents a sequence of DNA bases (or a sequence of amino acids for which the DNA codes).

Position:	1	2	3	4	5	6	7	8	9	10	11	12	13
Row 1 Cards:	A	2	3	4	5	6	7	8	9	10	J	Q	K
Row 2 Cards:	A	2	3	4	5	6	7	8	9	10	J	Q	K

3. Shuffle the remaining cards and place them face down as a stockpile.

Modeling the Molecular Clock

4. Divide your lab group into two teams. Team #1 will mutate Row #1 while Team 2 mutates Row #2. The two teams work simultaneously; do not rush ahead of the other team.
5. Your team rolls two dice. Then your team removes the card from your row at the position indicated by your dice roll and places that card into a discard pile. Note that the dice determine the position of the card you removed, not its value. The other team simultaneously rolls two dice and removes one card from its row.

Figure 3. Step 5 Example. In this example, the Row 1 team rolled a seven while the Row 2 team rolled a four. The cards in these spaces have been removed. Note: Your data may differ!

Position:	1	2	3	4	5	6	7	8	9	10	11	12	13
Row 1 Cards:	A	2	3	4	5	6	_	8	9	10	J	Q	K
Row 2 Cards:	A	2	3	_	5	6	7	8	9	10	J	Q	K

6. Replace the missing card in your row using a card drawn from the stock-pile. At the same time, the other team does the same thing to its row.
7. Count and record the number of differences between your lab group's two rows of cards.

Figure 4. Step 7 Example. The Row 1 team drew a Queen while the Row 2 team drew a 2 card. There are now two differences between the two simulated genes.

Position:	1	2	3	4	5	6	7	8	9	10	11	12	13
Row 1 Cards:	A	2	3	4	5	6	Q	8	9	10	J	Q	K
Row 2 Cards:	A	2	3	2	5	6	7	8	9	10	J	Q	K

8. Repeat the process beginning at Step #4. Remember that your dice roll determines the position of the card that you remove.

Figure 5. Step 8 Examples. In this example, the Row 1 team rolled another seven. The team discarded the Queen that was located at the 7 position. Then an Ace was drawn. Meanwhile, the Row 2 team rolled a nine and drew a 3 card. There are now three differences between the two simulated genes at positions 4, 7, and 9.

Position:	1	2	3	4	5	6	7	8	9	10	11	12	13
Row 1 Cards:	A	2	3	4	5	6	A	8	9	10	J	Q	K
Row 2 Cards:	A	2	3	2	5	6	7	8	3	10	J	Q	K

In the next round, the Row 1 team rolled a nine and drew a 3 card while the Row 2 team rolled a ten and drew a Jack. There are still three differences at positions 4, 7, and 10. There is no difference at position 9 since both cards have the same value.

Position:	1	2	3	4	5	6	7	8	9	10	11	12	13
Row 1 Cards:	A	2	3	4	5	6	A	8	3	10	J	Q	K
Row 2 Cards:	A	2	3	2	5	6	7	8	3	J	J	Q	K

Reconstructing the Phylogenetic Tree

9. After four to six rounds, clone one (or both, depending on your teacher's instructions) row using a second deck of cards. For example, this sample data shows Row 2 being cloned to form Row 3.

Figure 6. Step 9 Example. In this example, Row 3 is a "clone" of Row 2. It will begin to diverge from Row 2 in the next round.

Position:	1	2	3	4	5	6	7	8	9	10	11	12	13
Row 1 Cards:	A	2	7	4	5	6	A	8	3	10	3	Q	K
Row 2 Cards:	A	2	3	2	5	6	7	A	3	J	4	Q	K
Row 3 Cards:	A	2	3	2	5	6	7	A	3	J	4	Q	K

10. Reorganize your lab group so that a person or team is in charge of each row. You will now continue to play. During each round, each team rolls two dice, removes the indicated card, and replaces it from the stockpile. At the end of each round, count and record the number of differences between each possible pair of rows. Compare Row 1 with Row 2, Row 1 with Row 3, and Row 2 with Row 3. If you have four rows, compare Row 4 with the other three rows.

Figure 7. Step 10 Example. Later there are 10 differences between the card sequences in Rows 1 and 2, 11 differences between Rows 1 and 3, and four differences between Rows 2 and 3.

Position:	1	2	3	4	5	6	7	8	9	10	11	12	13
Row 1 Cards:	A	2	7	K	8	5	Q	8	10	9	3	K	K
Row 2 Cards:	A	6	3	2	8	6	7	A	J	10	4	5	K
Row 3 Cards:	A	6	3	4	2	6	K	A	J	10	7	5	K

11. After playing 10 to 12 total rounds, record the order (that is the value) of the cards for each row on a separate strip of paper.
12. Exchange your row sequence data with another lab group.
13. (**Note:** If you are doing this lab over two periods, you may do Step 13 after Step 15. Follow your teacher's specific instructions.) Using the sequence data from the other lab group, construct a cladogram (a branching-tree diagram) showing how the rows of cards are related. Show which rows of cards are descended from the row that was cloned mid-game. Show which rows of cards started to diverge at the beginning of the game. In real life, complete DNA sequences of species that have been extinct for millions of years are not available today. Therefore you may only use the sequence data. You may not ask the other lab group for the answer; nor should you tell them the phylogenetic history of your own rows of playing cards.

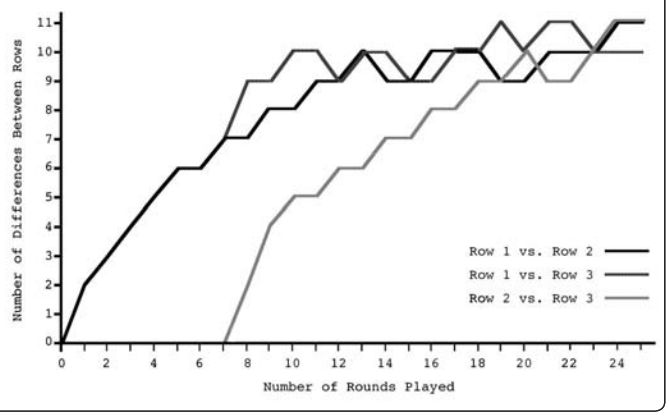
Observing the Limits of the Molecular Clock

14. Continue to play as many rounds as possible. You should be able to mutate your card sequence at least 15 to 20 rounds. If you run out of stock cards, shuffle the discards and place them face down as a new stockpile.

Figure 8. Step 14 Example. At the end of this sample game, there are 10 differences between each pair of rows. Phylogeny can no longer be determined using this sequence alone.

Position:	1	2	3	4	5	6	7	8	9	10	11	12	13
Row 1 Cards:	A	3	7	K	2	5	Q	8	10	9	3	A	K
Row 2 Cards:	A	6	3	2	8	6	7	A	4	5	4	9	K
Row 3 Cards:	A	2	3	4	2	Q	K	7	J	10	7	5	K

Figure 9. Accumulation of differences in playing card sequences over time.
This graph shows the results of a typical game. As the simulation progresses, the number of differences in the sequences of playing cards increases up to a maximum of 10 or 11 differences. Row 2 was duplicated to make Row 3 after the sixth round of play. By the end of the simulation, Row 3 lost its similarity to its parent.



15. Graph the number of differences between the card sequences for each generation. Place the time (the number of rounds played) on the *x* axis and the number of differences between rows of cards on the *y* axis. Discuss your results.

Discussion

This card game simulates a molecular clock. The row of cards with 13 different possible values is analogous to a protein which may contain 20 different amino acids or a DNA segment with its four nitrogenous bases. Initially, the number of card (or base or amino acid) differences increases steadily and indicates the approximate number of rounds (or generations) played.

Recently-diverged lineages are more similar to each other than to lineages with an earlier common ancestor. However, after enough time has elapsed, the "clock" runs down. Round 15 looks pretty much like Round 20. After the clock runs down, only a minimum amount of time elapsed can be determined.

The data for the b-globin chain of hemoglobin show a similar phenomenon. The common ancestor of the great apes lived recently enough that the b-globin amino acid sequence may be a useful tool to determine phylogeny. But the b-globin sequence cannot differentiate between the common ancestor of great apes and lemurs from the common ancestor of all primates and equines.

Models and simulations are useful ways for both scientists and students to examine natural phenomena. Appropriate models are clear and compelling. A classroom simulation is focused, rapid, and inexpensive. Simulation lessons also have a high rate of success. Since models simplify the real world, they may reinforce misconceptions. Such misconceptions should be overtly addressed. For example, as organisms evolve, mutations gradually accumulate. Accumulation of mutations is clearly seen in this simulation. However, each round of play resulted in exactly one mutation. Although the rate of mutation for a segment of DNA is often consistent, real DNA does not sustain a single-point mutation with each generation.

Since two dice never roll a one or a 13, the cards in the first and 13th position never change. Although the cause is different, the result is analogous to what is seen in actual proteins. Studies of proteins such as b-globin and cytochrome-c show some amino acids are constant. Presumably, if a mutation changes the DNA coding for these amino acids, the resulting protein is nonfunctional, causing the organism not to survive or reproduce.

Like the rolling of dice, variation is a random evolutionary event. An organism can neither predict nor choose a mutation and students do not get to pick which card they change. However, the distribution of dice rolls is not even. There is only one way to roll a two, but there are two ways to roll a three and six ways to roll a seven. Advanced students may note that they were more likely to mutate a middle card. This is analogous to the hot spots (sites of frequent mutation) that appear along the DNA. The mechanism for the hot spots in the simulation and in real DNA are, of course, completely different.

There is no selection in this lab. Any card can be placed in any position (2 through 12) without changing viability. This may not be realistic. However, limiting which cards may be placed where (e.g., only face cards can go into the 11th and

Table 1. Number of amino acid differences in the b-globin chain of hemoglobin between species. The number of amino acid differences suggests a very recent common ancestor for humans and chimpanzees and a slightly more distant common ancestor of humans, chimps, and gorillas. Old-world rhesus monkeys appear more closely related to the great apes than new-world squirrel monkeys. But there are fewer differences in b-globin sequence between horses and apes than between lemurs and apes! Lemurs diverged from other primates so long ago that the b-globin molecular clock is no longer useful. Fortunately, other sequences (such as cytochrome-c) may be used to differentiate between the ancestry of apes, lemurs, and horses.

	MAMMALS / PRIMATES / ANTHROPOIDS/OLD & NEW WORLD / CATARRHINES/OLD WORLD / HOMINIDAE					
	Chimp	Gorilla	Rhesus Monkey	Squirrel Monkey	Lemur	Horse
Human	0	1	8	11	30	26
Chimp		1	8	11	30	26
Gorilla			8	12	30	27
Rhesus Monkey				13	28	28
Squirrel Monkey					30	25

12th position) tends to complicate the simulation to the point that students fret over the game rules instead of thinking about the meaning of the model.

After the Lesson

Students observed that mutations accumulate and why amino acid sequence differences may reveal phylogeny. They are ready to analyze actual amino acid sequence data in order to construct a phylogenetic tree. Amino acid sequences of proteins such as b-globin and cytochrome-c are widely available (see Additional Resources). Since amino acid sequences are shorter than DNA sequences, it is more likely that children, who may lack computer access, will be able to successfully analyze protein data.

Differentiating Instruction

Some students are confused between the position of the card and its value. When they roll a three, they want to change any and all cards with the value of three instead of whatever card sits in the third position. A strip of paper with spaces for the cards labeled by position will help reduce confusion and is recommended when teaching middle school or challenged high school students.

Occasionally students may try to count all of the changes that have ever occurred in the evolutionary history of their team's cards! The teacher may need to remind students that only the current generation's genome can be observed; thus extinct sequences are not included in the current count. Demonstrating that only 11 differences are possible may alert students to count only the differences for the current generation. Younger and more challenged students may need to have prepared graph paper with the *x*-axis ending at 11.

Many middle school students may be confused modeling the molecular clock and reconstructing a phylogenetic tree in the same lab. The reconstruction stage, Steps 9 through 13, may be skipped entirely or done as a separate activity.

Instead of having all students begin with the cards in the same order, advanced students in high school may randomly deal the cards to form Row 1. Then they find the matching cards to form Row 2. After shuffling the remaining cards, they play as described above.

Acknowledgments

The author gratefully acknowledges the support of Larry Flammer, ENSI Web Master; and of Becky Saylor, science teacher, Dublin City Schools, Dublin, Ohio, who piloted and critiqued an early version of this lesson.

Additional Resources

Hemoglobin data:

http://web.indstate.edu/thcme/mwking/hemoglobinmyoglobin.html.

http://sickle.bwh.harvard.edu/hbsynthesis.html.

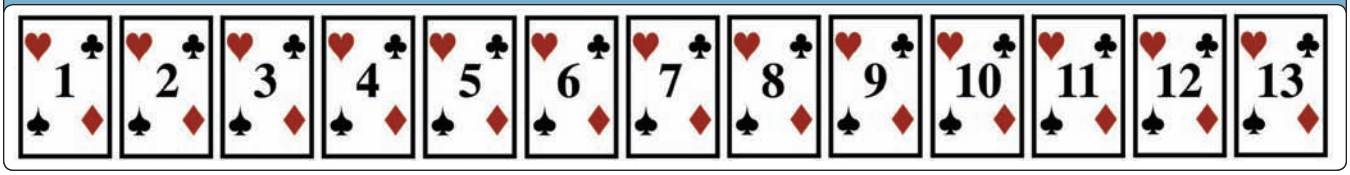
Kramer, B. & Westerling, K. (2001). Molecular biology and phylogeny: "The cytochrome-c" lab. (L. Flammer, Ed.). Available online at: http://www.indiana.edu/~ensiweb/lessons/mol.bio.html.

Myoglobin data:

http://oregonstate.edu/dept/biochem/hhmi/hhmiclasses/bb450/winter2002/ch07/c07emhp.html.

Nelson, C. & Nickels, M. (2005). Molecular sequences & primate evolution: Amino acid differences in beta

Figure 10. Card placement strip.
Place one or two enlarged copies of this strip on the students' work area to help them keep track of the position of their cards.



hemoglobins in primates. (L. Flammer, Ed.). Available online at: http://www.indiana.edu/~ensiweb/lessons/mol.prim.html.

Pealer, G. & Ferris, S. (1995). Using amino acid sequences to show evolutionary relationships. Available online at: http://www.woodrow.org/teachers/bi/1995/simulation_amino.html.

Puterbaugh, M. & Burleigh, J.G. (2001). Tutorial: Investigating evolutionary questions using online molecular databases: Bats/Birds, whale feet, and reptile feathers? (L. Flammer, Ed.). Available online at: http://www.indiana.edu/~ensiweb/lessons/p.tut.db.html.

References

Bunk, S. (2000). Hot paper: Molecular clock. In Analyzing sequence substitutions, more genes tell better evolutionary time. *The Scientist*, *14*(22), 17.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution.* Cambridge, U.K.: Cambridge University Press.

Martin, R.A. (2003). Molecular clocks and shark evolution. Available online at: http://www.elasmo-research.org/education/evolution/molecular_clocks.htm.

Ridley, M. (1996). *Evolution, 2nd Edition*. Cambridge, MA: Blackwell Science, Inc.

Rambaut, A. & Bromham, L. (1998). Estimating divergence dates from molecular sequences. *Molecular Biology and Evolution*, *15*(4), 442-8.