Assimilation: Indonesian Journal of Biology Education
ISSN 2621-7260 (Online)
Journal homepage: https://ejournal.upi.edu/index.php/asimilasi

Incusion of mathematics in biological concept at the senior secondary level in Indian education system

Jaspreet Kaur
Maitreyi College, University of Delhi, India
Corresponding author: jkaur@maitreyi.du.ac.in

OPEN ACCESS

ARTICLE HISTORY
Received: 2 September 2023
First Recised: 11 September 2023
Accepted: 30 September 2023
First Available Online: 30 September 2023
Publication Date: 30 September 2023

KEYWORDS
Quantification
Measurement count
Mathematical modeling

ABSTRACT

The students who study biology often develop fear from mathematical concepts and tend to avoid the pages of biology textbooks where mathematical details are included. The students often perceive Mathematics and Biology as two separate disciplines of study. One of the main reasons for this perception is that biology textbooks do not include relevant mathematical details in biological concepts. Moreover, biology teachers have little or no mathematical academic background while mathematics is taught by teachers with least or no interest in biology. This often leads to the lack of understanding among students that both subjects are actually linked to each other. The need to study mathematical details in biology concept becomes important in early stage during schooling years that both Mathematics and Biology are related to each other. In this article, illustrative examples have been presented with detailed explanation wherein the biology students can appreciate mathematical concepts efficiently in a teaching-learning environment at the senior secondary level.

INTRODUCTION

Textbooks are a rich source of information about specific disciplines. They help in the development of understanding about a specific subject as well as appreciate it in general. Biology textbooks are well-equipped with examples from the real biological world but often miss out on quantitative details. For instance, statement such as 'ligand binds with the membrane receptor' found in a standard senior secondary biology textbook can be made more quantitative to include the average value (in numbers) of the ligands or cell signalling molecules present in a cell, which further bind with receptor followed by the downstream processes. This will lead to integration of mathematics in a biological statement or a principle in a simple yet effective manner. Also, biology students will not perceive the discipline of biology as simply a subject of learning facts but will encourage them to understand it much better from the frontiers of mathematical skills.

The idea of linking biology with mathematics is not a new one. In fact, many of the 10 important equations stated by Jungck (1997) have already been incorporated in the biology curricula, but mostly at the undergraduate or postgraduate level, including logistic equation of Lotka-Volterra interspecific competition, Michaelis–Menten equation of enzyme kinetics, Species-Area relationship (MacArthur Wilson theory of Island Biogeography) or application of the Poisson distribution, for better understanding about the mode of action of bacteriophages (Grammer, 2017). However, many biological concepts are still waiting to be linked with mathematical equations or principles at the senior secondary level. If the educators and academicians think that the inclusion of such mathematical derivations at the school level will increase the difficulty level of 'Biology students', then this will also lead to an understanding of biological concepts with certain lacunae which become very difficult to fill up at the later stages of their education. This incomplete understanding may remain with students even after they enter professional heights of their career. Thus, there is an emergent need to revise the biology curriculum which is being followed at the senior secondary level across Indian schools. In this article, I try to bring out a few simple yet elegant mathematical details which can be easily applied by the educators in their classrooms. The inclusion of such mathematical ideas will enhance the understanding of biology students and will lead to an appreciation for mathematics as a discipline which will help them to *make some sense* of the biological data. In this article aim to understand a few important biological concepts with mathematics as an integral component in the following sections.

METHODS

The National Council of Educational Research and Training (NCERT) Biology textbooks of Senior Secondary level (Class XI and XII) were screened for selected topics (discussed in this review) which lacked mathematical details. The articles chosen based on the selected biology concept. The articles were analysed then augmented with quantitative details retrieved from different research articles.

RESULTS AND DISCUSSION

The following sections discuss some important biological concepts with mathematics as an integral component based on the selected research articles.

Fibonacci sequence

Leonardo Pisano, who was later known as Fibonacci presented the Fibonacci sequence, in which every number (after the first two) is the sum of the two preceding numbers, for example: 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, and so on. Fibonacci numbers appear in various biological forms

DOI: https://doi.org/10.17509/aijbe.v6i2.62077
e-ISSN 2621-7260

like the shape of sunflower head, shape of mollusc shell, human anatomy (cochlea, phalanges of hand), leaves and cones of plants, ciliary rows in eukaryotic protozoans, body plans in arthropods as well as microscopic structures like the DNA double helix, and many more (Figure 1) (Bormashenko, 2022; Sinha, 2019; Persaud-Sharma & Leary, 2015). The students at the senior secondary level study about this important sequence *only* in Mathematics (this concept has been included under chapter 'sequences and series' in National Council of Educational Research and Training (NCERT) Mathematics Textbook of Class XI). So, when Biology students are studying about various examples as stated above, they can easily appreciate the concept of 'sequence and series' via the Fibonacci sequence. Also, the non-biology students can discuss the presence of this simple pattern of numbers in the real-world examples.

Figure 1. Fibonacci Sequence in sunflower head, mollusk's shell and pine cone. Image credits: L. Shyamal, CC BY-SA 2.5 <https://creativecommons.org/licenses/by-sa/2.5>, via Wikimedia Commons, Suyothami, CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0>, via Wikimedia Commons, Richard Flink from USA, CC BY-SA 2.0 <https://creativecommons.org/licenses/by-sa/2.0>, via Wikimedia Commons

The Fibonacci sequence can be calculated mathematically, when the numbers are arranged in a sequence. Each number in the sequence is considered a term, denoted by Fn, where 'n' is the position of the number in the sequence, starting with 0 and 1 as the first 2 numbers of the sequence. The rest of the numbers in the sequence can be derived using a simple formula:

$$Fn = Fn\text{-}1 + Fn\text{-}2$$

So, 3rd digit in sequence (after 0, 1) will be

$F3 = F3\text{-}1 + F3\text{-}2$

$= F2 + F1$

$= 1+0$ (F1 = 0, F2 =1)

So, F3 = 1

Likewise, F4 = 2, F5 = 3 and so on.

Thus, a small section on the importance of Fibonacci sequence from biological point of view and few simple class outdoor exercises to study the sequence related to shape and pattern of leaves of plants while going on a walk to the nearby school park will help Biology students to understand the leaf pattern using a fascinating Mathematical principle in an interesting way.

The cell

'*Cell is a basic unit of life*'-this is a common biological statement and is found in almost every biology textbook. In fact, this statement describes the fundamental basis of life and helps the beginners in biological sciences to appreciate living organisms. The general approach to study a cell and its functions at a very basic level is to make a well labelled diagram of a cell and then distinguish between a unicellular and multicellular organism. This further helps the students to understand

DOI: https://doi.org/10.17509/aijbe.v6i2.62077
e-ISSN 2621-7260

about prokaryotes like bacterial cells as well as eukaryotes. The theoretical concept is reinforced with some basic laboratory exercises like observing an onion peel under a microscope or making a slide of dividing cells in an onion root tip (mitosis) or observation of blood cells from prepared blood film on a microscopic slide. But, the observations in these exercises are limited only towards biological features of cell like its *morphology*. Students seem to miss out the important details or questions like what is the size of cell which is being observed? What is its diameter? The answer to such questions will help in implementing Mathematics to understand the cell as a complete entity and a concept *in toto*. So, the first step in this direction can be the '*counting*' of red blood cells in a blood sample or bacterial cells in a culture medium using a haemocytometer. Then, a discussion on dimensions of a red blood cell includes its average diameter, which has been reported as 7.27±0.52µm in fresh cells (Silverman & Glick, 1969). Also, the concentration of erythrocytes in blood has been reported as 4,000,000 to 6,000,000 cells/mm^3 (Rosen, 1967). Further, human RBCs have a mean volume of ~97 μm^3 and a mean surface area of ~137 μm^2 (Fung, 1993). The study of these important dimensions of red blood cells will help the students to understand about physiological concepts like carrying capacity of blood and haemoglobin in human physiology topic.

Human physiology

Haemoglobin, a protein present in red blood cells carries oxygen to various body organs. At the basic level, the students can be encouraged to calculate the amount of Hb present in a red blood cell. Pierigè (2008) has reported a derived value of ~3×10^8 copies per cell which has been calculated using physiological level of Hb in erythrocyte, molecular mass of Hb, volume of erythrocyte and Avogadro's number. The calculation is as follows:

$$N(Hb)=\frac{C(Hb)\times V(RBC)\times NA}{MW(Hb)}$$

where:
N(Hb) =Number of Hb in erythrocyte,
C(Hb) =Concentration of Hb in erythrocyte, which is ~330 g/l
V(RBC) =volume of erythrocyte, 10^{-13} liter/cell
NA =Avogadro's number, 6×10^{23} Hb/mole,
MW(Hb)=molecular mass of Hb, 64,458 g/mole

After plugging in the values, we get

$$N(Hb) = \frac{330\ (g/l) \times 10^{-13}\ (liter/cell) \times 6\times10^{23}\ (Hb/mole)}{64,458\ (g/mole)}$$

$$= \frac{198 \times 10^{-12} \times 10^{23}}{64,458}$$

$$= \frac{198 \times 10^{11}}{64,458}$$

$$= 0.00307 \times 10^{11} \text{ or } 3.07\times10^8 \text{ (Hb/cell)}$$

Most of the oxygen in the blood is bound with haemoglobin except a minimal amount which dissolves in plasma water (dissolved oxygen). The average carrying capacity of Hb under normal conditions is around 1.39 ml oxygen per gram of Hb. This value of 1.39 was obtained through rigorous works of Dijkhuizen et al. (1977). They found that 1 mole of Hb binds to 4 moles of oxygen. Since mol the weight of Hb is 64,458.5g/mole, it means 1/64458.5 moles of Hb (i.e. 1 gram) will bind 4 /64458.5^{th} of a mole of oxygen. Now, one mole of oxygen typically occupies ~22.4L of volume at standard temperature and pressure; thus, the ideal oxygen-binding capacity of haemoglobin must be (4 × 22.4) ÷ 64458.5, or 0.00139 L/g or around 1.39 ml/g.

DOI: https://doi.org/10.17509/aijbe.v6i2.62077
e-ISSN 2621-7260

So, any kind of abnormality in Hb will change the carrying capacity of Hb. The total oxygen content in the blood is calculated by the equation:

$$CaO = 1.34 * [Hb] * (SaO) + 0.003 * PO \quad \text{(Equation 1)}$$

Where:
CaO = the concentration of oxygen
SaO = saturation of haemoglobin in %
PO = partial pressure of oxygen in blood in mmHg
0.003 = the content of dissolved oxygen in blood in ml/L/mmHg,
[Hb] = concentration of Hb in g/L

The oxygen bound with Hb corresponds to [1.34*[Hb] * (SaO)] and oxygen dissolved in plasma is given by 0.003 * PO.

If we want to calculate the concentration of oxygen in blood or the carrying capacity of Hb, given that Hb is 90% saturated with oxygen, and the alveolar PO corresponds to 60 mmHg and concentration of Hb is 14g/L. So, we can plug in these values in Equation 1 as

CaO = (1.34 x 14 x 0.90) + 0.003 x 60 (90% or 0.90)
= 16.884 + 0.18
= 17.064

In other words, the oxygen bound with Hb corresponds to 16.884 and oxygen dissolved in plasma is 0.18, giving the total concentration of oxygen in blood as 17.064 volume%. Thus, the simple cell size measurements of red blood cells can indicate the amount of Hb present in them which may further relate to the carrying capacity of oxygen in blood. So, the discussion of biological concepts should connect these simple mathematical details for their better understanding.

Cell cycle

Cell cycle or cell-division refers to a series of events that take place in a cell as it grows and divides into daughter cells. The simplest model to study the cell-division cycle is a bacterial cell, which can be easily grown in a nutrient medium. For instance, the minimal generation time of *Escherichia coli* on Luria Bertani medium is 18.10 ± 0.52 min (Reshes et al., 2008) while the generation time on minimal gluconate medium is 60 minutes (Tao et al., 1999). Further, generation time falls to 26.4±7.2 min with amino acids and 46.8±17.0 min without amino acids in the medium (Ullman et al., 2012). On the other hand, the doubling time in *E. coli* shoots up to 67 min when grown in M9 minimal medium + glucose + thiamine. However, the time for a complete round of replication equals around 40 minutes (Helmstetter, 1968). To understand these *different values* of generation time in *E. coli*, it is important to understand about the bacterial cell cycle, which includes DNA duplication (S phase) and cell division (M phase). There happens to be an internal clock in this bacterium which determines the time of cell division which follows the initiation of DNA synthesis. In *E. coli*, it is around 60 minutes, out of which DNA replication takes ~40 minutes. This internal clock does not change even though the generation time changes due to availability of different nutrient conditions in the culture medium. In rich mediums, the generation time becomes shorter than 60 minutes, say 40 min or even 20 min, depending on the availability of nutrients rich in carbon, hydrogen, nitrogen, etc. In case of poor availability of growth nutrients, the generation time increases to 67 min (as discussed above).

The specific generation time (say 60 min), each new cell contains one copy of the genome. The origin of replication begins every 60 min, i.e., as soon as a new cell is formed, it initiates replication as part of the next cell cycle. This is followed by cell division to form two daughter cells, each one containing a copy of the genome. Each of these new cells initiate replication as part of

DOI: https://doi.org/10.17509/aijbe.v6i2.62077
e-ISSN 2621-7260

their cell cycle and after 60 min, again two new daughter cells arise, each with one copy of the parent's genome.

The bacterial cell cycle can be divided into three stages: A, B and C (Figure 2). The A phase is called the 'gap phase', chromosome replication takes place during B phase while cell division in the C phase. The A phase is highly variable in length, for instance, when the cells are growing fast, the A phase can be skipped totally and cells are born with replicating chromosomes (dotted line), while under slow growth conditions, cells spend much of their time in this phase, especially when the generation time is above 60 min. On the other hand, the length of B and C phases increases as the generation time increases. The length of B phase is generally in concordance with DNA replication (Dewachter et al., 2018).

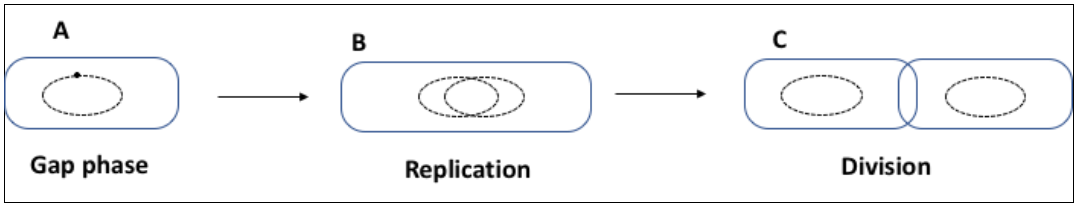

Figure 2. Stages of bacterial cell cycle. Image recreated from Dewachter et al. (2018)

However, DNA replication time is invariant to 40 min as long as the building blocks of DNA-nucleotides are available in the cell (Cooper et al., 1968; Zhang et al., 2020). It cannot go faster because the DNA polymerase incorporates ~200-1000 nucleotides/sec (DNA synthesis rate) by DNA polymerase III holoenzyme (Breier et al., 2005; Reyes-Lamothe et al., 2008). This DNA synthesis rate has been evolutionarily optimized. So, does the time required for DNA replication ever impose certain limitations on the growth rate of *E. coli*, given that the replication rate is several hundred nucleotides per second? The answer to this question can be obtained using a simple calculation as follows:

Genome size of *E. coli* is 4.6×10^6 bp;
Replication rate ~1000 bp per replies (taking the maximal value of the range reported for replication rate);
If we divide the genome size with the replication rate, we can estimate the time required for replication of complete genome, as:

$$\sim \frac{4.6 \times 10^6}{1000 \times 2}$$

(since DNA replication is bidirectional, so, 2 replisomes are expected to form during DNA replication);
~ 4.6×10^3 or 4600 sec or ~77 minutes, which is much longer than the doubling time of *E. coli* (20 minutes). So, this genome replication paradox can be explained by nested replication forks for maximum growth of bacterial cells within the stipulated time period.

So, the bacterial growth and the concept of cell cycle leads us to find out the answer of a pertinent question- How long will it take to get one billion cells from just one bacterial cell? The answer to this question requires *mathematical reasoning*. In order to understand this, let's take the following example (Figure 3).

DOI: https://doi.org/10.17509/aijbe.v6i2.62077
e-ISSN 2621-7260

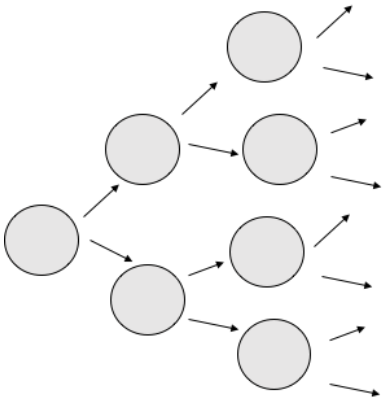

Figure 3. Pictorial representation of exponential doubling in bacterial cells

As we have discussed above, the generation time of *E. coli* is 60 min in a minimal medium. Since one bacterial cell divides to form 2 daughter cells or in other words, after one *generation or g*, 2 new cells are formed. So, after 2 generations, 4 cells will be formed, after 3 generations-8 cells, so on and so forth (Figure 3). Thus, bacterial growth is exponential in nature, more generally, starting with No cells:

$$N = N_0 \times 2^g \qquad \text{(Equation 2.1)}$$

Now, since we want to know how much time will it take to get 'N' no. of cells from one initial cell (N_0), so we can express generations (g) in terms of time interval (t). If we let
t_D = the generation time, or doubling time, then the number of generations that have elapsed during time interval (t) will be t/t_D, i.e.,

$$g = t/t_D \qquad \text{(Equation 2.2)}$$

or $t_D = t/g$

which means that t_D (generation time) = t (time, in minutes or hours)/g (number of generations)

Now, if we plug in the value of g from equation 2 in equation 1, we get

$$N = N_0 \times 2^{(t/t_D)} \qquad \text{(Equation 2.3)}$$

or $N/N_0 = 2^{(t/t_D)}$

We can see that the bacterial growth is exponential with respect to time. Now, we can solve this equation for calculating the time t, required to get one billion cells (N) from one cell (N_0). To do this, we first take logarithm base 2 or base 10 or base e on both sides of equation 3,

$$\log_2(N/No) = t/t_D$$

or $$t = t_D * \log_2(N/N_0) \qquad \text{(Equation 2.4)}$$

After plugging the values of N_0 (1 cell), N (1 billion cells or 1000, 000, 000 cells) and t_D (60 min or 1 hr), we get

$$t = 1 * \log_2(10^9/1)$$

or $t = 29.9$ minutes

DOI: https://doi.org/10.17509/aijbe.v6i2.62077
e-ISSN 2621-7260

So, it means that it will take around 30 minutes to form 1 billion cells from one bacterial cell. The same equation 1 can also be solved by taking log on both sides, we get

$N = N_0 \times 2^g$
$\log N = \log N_0 + g \log 2$
$g = \frac{\log N - \log N_0}{\log 2}$
$g = \frac{\log N - \log N_0}{0.301}$ ($\log_2$ is 0.301)
This can be rearranged as
$g = 3.3 \log N/N_0$ (1/0.301 is 3.3)
Since $t_D = t/g$ (Equation 2), so if we solve for t_D
$t_D = \frac{t}{3.3 \log N/N_0}$

This will help us in calculating the generation time of a bacterial population that increases from N_0 cells to N cells in a particular time interval of growth.

Cell signaling

Cells in multicellular organisms communicate with each other via signalling molecules, which are often referred to as 'ligands', a general term for molecules that bind specifically to other molecules like receptors. So, when a signalling molecule binds to the specific receptor, it changes the shape or activity of receptor and leads to various other downstream changes along a signalling cascade, leading to changes in the cell such as the alteration of key proteins and changes in gene transcription (Figure 4).

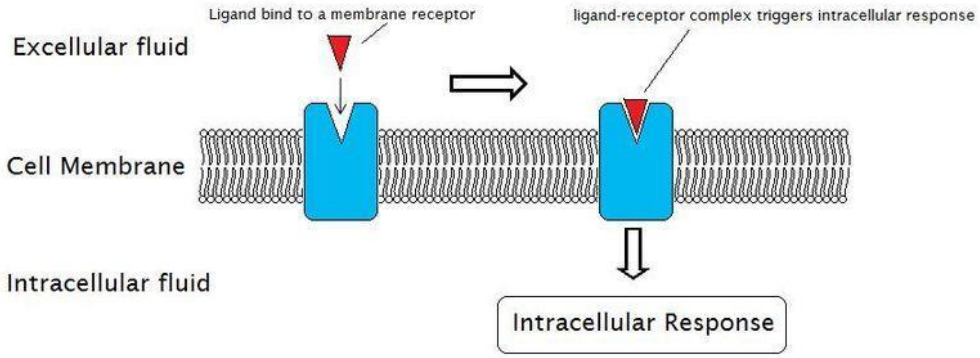

Figure 4. A simple representation of cell signalling cascade initiated by binding of a ligand with its receptor. Image taken under creative commons.

The most common receptors involved in the regulation of cell growth, differentiation, and survival are RTK or receptor tyrosine kinases (RTKs). These cell surface receptors bind to signalling molecules like growth factors and other proteins present in low concentration. The quantitative measurements of membrane tyrosine kinase receptors and their downstream intracellular signalling have been reported in different studies from non-transformed cell lines and are usually in the range of 10,000 – 1,000,000 molecules per cell (Legewie et al., 2008). But, the cellular protein concentrations of these membrane receptors can be even lower than 10,000 per cell. Further, some signalling proteins like caspases (cas3, 8 and 9), transcription factors (RelA, STAT2, Smads) and several phosphatases (e.g., CD45, PP1, PP2C, SHP-1) are also expressed at high levels (>10,000 per cell).

DOI: https://doi.org/10.17509/aijbe.v6i2.62077
e-ISSN 2621-7260

The measurement count of these important signalling molecules helps us to understand the cell energetics (Table 1). Since these signalling molecules constitute a significant portion of total cellular protein content, their production constitutes a central energy sink in the cell (Legewie et al., 2008).

Table 1. Number of signaling molecules present per cell

Name	Pathway	Number present per cell*
POGFR	RTK	90000
EGFR1	RTK	200000
EGFR2	RTK	60000
InsulinR	RTK	250000
IGFR1	RTK	1200000
GO2	RTK	23000
Rs	MAPK	11000000
Raf	MAPK	10000
MeK	MAPK	20000000
ErK	MAPK	20000000
PI3K	PI3K	10000
Protein Kinase B	PI3K	600000
PP1	misc	500000
PP2A	misc	20000000
PP2B	misc	20000000
PP2C	misc	20000
PTP 1-B	misc	10000
SHP-1	misc	800000
CD45	TCR Signalling	100000
TGFR	TGF-beta	80000
Smad 2	TGF-beta	450000
Smad 3	TGF-beta	450000
Smad 4	TGF-beta	900000
Calmodulin	TGF-beta	10000000
Bcl-2	Apoptosis	50000
ApaT-1	Apoptosis	2000000
XIAP	Apoptosis	36000
Pro-Caspase9	Apoptosis	160000
Pro-Caspase8	Apoptosis	8000000
Pro Caspase 3	Apoptosis	1600000
c-abl	DNA damage	250000
CREB	cAMP	50000
Protein Kinase A	cAMP	720000
p300	cAMP	28000
IFN-gamma receptor	Cytokine	25000
Epo receptor	Cytokine	60000
STAT2	Cytokine	150000
STAT3	Cytokine	750000
RelA	NF—kB	125000
T cell receptor	Immune	30000
Lck	Immune	60000
ZAP7O	Immune	1200000
FC receptor	Immune	700000
TOTAL		**122747000**

*Data taken from Legewie et al. 2008; maximal values were taken here if a range of protein expression was given

DOI: https://doi.org/10.17509/aijbe.v6i2.62077
e-ISSN 2621-7260

If the number of all the signalling molecules (given in Table1) are summed up, it gives ~1.2 x 10^8 molecules per cell. Now, given the concentration of total protein content of a cell, which is ~300 pg (Volpe et al., 1990) with an average molecular weight of cell protein as 50 kDa (Hendil et al., 2002), we can calculate the number of total protein molecules in a cell (n) as follows:

$n \times 5x10^4 \times 1.66x10^{-27} = 300 \times 10^{-15}$ (1Da is 1.66×10^{-27} kg; 1pg = 10^{-15} kg)

$$n = \frac{3 \times 10^{-13}}{5 \times 1.66 \times 10^{-23}}$$

$$n = \frac{3 \times 10^{10}}{8.3}$$

$n = 0.361 \times 10^{10}$

$n = 3.61 \times 10^9$ molecules per cell

Thus, if we find out the molecular weight of the signalling proteins listed in Table 1 as:
$= 1.2 \times 10^8 \times 50 \times 10^3 \times 1.66 \times 10^{-27}/10^{-15}$ (1pg = 10^{-15} kg)
= 10 pg, which means that the signalling molecules constitute ~3.3% of the total protein content of the cell (10/300*100).

Thus, such discussions and simple calculations related to the number of signalling molecules present in a cell and their relatively low concentrations in the total protein content of a cell will give a better understanding about the energetics involved in cell signalling mechanisms. Given that the cellular protein turnover requires around 30 to 70% of the total cellular energy budget, the production of signalling molecules does not pose an energy burden on cells. This is because constitutively expressed signalling protein molecules are generally stable like other cellular proteins including housekeeping and structural proteins. So there seems to be a considerable evolutionary pressure in signal transduction networks, which keeps the turnover of uninduced signalling molecules low (Legewie et al., 2008).

Molecular basis of inheritance

Deoxyribonucleic acid or DNA is present as the genetic material in most of the living organisms including prokaryotes and eukaryotes, except certain RNA based viruses. The structure and function of this important biological molecule is critical to understand the molecular basis of inheritance. So, in order to understand this molecule in detail, we need to emphasize the *mathematical details* as well. It is known that DNA can either be present in A, B or C forms, depending on the type and amount of salt or the amount of hydration. Out of these, B-DNA is most commonly observed in different organisms. DNA is packaged into chromosomes so as to fit inside the small size of the nucleus of a cell. This simple concept discussed in the classroom can be further supplemented with the comparison of dimensions of cell, nucleus and the total length of the DNA molecule. Let's illustrate this concept using simple concepts/formulae related to DNA molecule:

The diameter of the DNA double helix is 20.4 Å. The distance between two adjacent base pair (bp) = 0.34×10^{-9} m. The number of base pairs is characteristic of every organism/species. e.g., bacteriophageϕ174 has 5386 bp, Lambda phage has 48502 bp, *E. coli* has 4.6 Mbp (4.6×10^6 bp) and human has 3286 Mbp (3.3×10^9 bp; haploid number). If we want to find out the length of a DNA molecule, we can simply multiply total number of base pairs with the distance between two consecutive base pairs or

DOI: https://doi.org/10.17509/aijbe.v6i2.62077
e-ISSN 2621-7260

Length of DNA = Total No. of base pairs x Distance between two consecutive base pairs

$= 6.6 \times 10^9 \times 0.34 \times 10^{-9}$ (diploid content is $2 \times 3.3 \times 10^9$)
$= 2.2$ m

So, to fit 2.2m DNA double helical molecules in the human nucleus with size of ~10 µm in diameter (Sun et al., 2000), the DNA needs to be packaged well. The packaging makes the molecule fit well in a small nuclear volume, e.g., nuclear volume of HeLa cell is 220 μm^3 and the cytoplasmic volume of HeLa cell is 940 μm^3 (Fujioka et al., 2006). Now, if we want to calculate the weight of human DNA, we need to find out the weight of a single base pair in grams. This value will then be multiplied by the number of bps in a single diploid cell to get the total weight of human DNA in the diploid cell. This is calculated as follows using Avogadro's number:

Since 1 mole of bp weighs 660 Da (1 Da = 1gram/mol)
or it also means that 6.023×10^{23} molecules of bp weigh 660 g/mol
So, 1 molecule of bp will weigh = $\frac{660}{6.023 \times 10^{23}}$ g
$= \frac{6600 \times 10^2}{6023 \times 10^{23}}$ g
$= 1.09 \times 10^{-23}$ or 1.1×10^{-21} g
So, weight of one base pair is 1.1×10^{-23} g

Now, total weight of human DNA $= 1.1 \times 10^{-21} \times 6 \times 10^9$ g
$= 6.6 \times 10^{-12}$ g
or $= 6.6 \times 10^{-12} / 10^{-12}$ (1pg = 10^{-12}g)
= 6.6 pg or 7 pg

Therefore, a single diploid human cell contains 7pg of genomic DNA.

Evolutionary biology

Students studying at the senior secondary level often perceive evolutionary biology as a study limited with bones, fossils or museums, wherein the theoretical concepts are discussed at length without much connection with the 'real' biological world. Some students also perceive this subject as simply 'rote-learning' of facts. Although, this notion has been revamped in recent years with the introduction of varied aspects of evolutionary biology including 'population genetics' wherein the students understand the changes in gene frequency of a population and appreciate the process of evolution of different species. However, the description related to population genetics is still confined within the boundaries of brief discussion with few examples only. In order to understand the principles of evolution, one needs to focus on the mathematical aspects. For instance, the factors affecting population equilibrium (Hardy-Weinberg equilibrium) like genetic drift, gene flow or migration, mutation, natural selection are discussed with some examples. These factors change the allele frequency of a population from one generation to another. But the real question lies- can we quantitate this change? How much the allele frequency changes from one generation to another?

The answer to these questions can be obtained using simple mathematical calculations. Let us illustrate a real-life example depicting change in gene frequency due to migration in a population over one generation. Gottelli et al. (1994) studied the molecular inheritance of Ethiopian wolf *Canis simensis* in 1994. The study involved populations of wolves from the Sanetti Plateau and Web Valley. These two areas were separated by rocky peaks and crossed by narrow corridors of suitable habitat. They found that the Ethiopian wolves are genetically distinct from

DOI: https://doi.org/10.17509/aijbe.v6i2.62077
e-ISSN 2621-7260

domestic dogs, but hybrids were also seen in Web Valley of Ethiopia where these two populations co-exist. On the other hand, the population from Sanetti Plateau was relatively pure. The extent of mixing of gene pools of populations of Ethiopian wolves and domestic dogs was estimated using allele frequencies at microsatellite locus 344. The dogs lacked 'J allele' while wolves were found to be homozygous for J allele. Frequencies of this allele were found as:

Sanetti Plateau	1.0
Web Valley	0.78
Domestic dogs	0.0

If we want to calculate the percentage of genetic composition of the Web valley population of Ethiopian wolves which was constituted from the domestic dog population, then we need to understand the basic concept of change in allele frequency of a population due to migration or gene flow. The concept can be understood using the continent-island model of Sewall Wright (1931), which describes that when a migrating population m (proportion of migrants) interbreeds with members of another population, then the descendants of the next generation will have 'm' genes from the migrants and the remaining (1 – m) genes from members of the original or native population. This statement can be expressed as-

$q_1 = mq_m + q_o(1\text{-}m)$, where (Equation 3.1)

q_1 = allele frequency of new population
q_m = allele frequency of migrating population
q_o = allele frequency of original/old/native population

So, the respective allele frequencies of original, migrating and new population will be:
$q_o = 1$
$q_m = 0$
$q_1 = 0.78$

Now, if we put in these values in Equation 1, we get

$$m = \frac{0.78 - 1}{0-1}$$

m = 0.22 or 22% of the genetic composition of the Web valley population of Ethiopian wolves was constituted from the domestic dog population. The same equation can be used to calculate the change in allele frequency due to migration. Since, change in allele frequency will be the difference between the allele frequency of new population (due to migrants) and the original population, or

$\Delta q = q_1 - q_o$ (Equation 3.2)

Now, if we put in the value of q_1 in equation 1, we get
$\Delta q = mq_m + q_o - mq_o - q_o$
$\Delta q = mq_m - mq_o$
$\Delta q = m(q_m - q_o)$

This equation clearly illustrates that the change in allele frequency of new population formed due to migration is dependent on the proportion of migrants (m). So, using such simple calculations, the '*actual change*' in allele frequency which happens in a population over a period of time can be *calculated*.

DOI: https://doi.org/10.17509/aijbe.v6i2.62077
e-ISSN 2621-7260

Mathematical modelling

Mathematical modelling is often used to understand different biological processes. For instance, modelling the population dynamics to study epidemiology using differential equations. This kind of modelling helps to model the transmission of a disease in a population, which will further help the communities and health care providers to combat a disease outbreak. In recent years, such type of mathematical modelling was witnessed to predict and study the COVID-19 disease (Khan et al., 2022). Mathematical modelling is the conversion or translation of a real-life or a physical situation into mathematical form, which helps in better representation and solution of certain problems. This topic is well explained and discussed in NCERT textbook of Mathematics, but the ironical part is that biology students should study this aspect in Biology textbooks as well to emphasize its importance in understanding biological aspects. So, while they are studying about different diseases and underlying mechanisms and/or diagnostics and therapeutics, they should also be exposed to the use or application of mathematical modelling in tackling such problems in a population. Here, I present an inclusion of a simple simulation, by which students can understand the concept of mathematical modelling for biological concepts.

In past two years, world has witnessed the wrath of COVID-19 disease, caused by SARS-CoV-2 virus. During this time, many scientists including mathematicians and biologists developed mathematical models and simulations to study the epidemiology of this disease. For biology students, it becomes imperative to study such models in a simplistic manner so that they can interpret the results of such simulations, even though they are not able to understand the intricacies of such models. One such model which was used to study the clinical progression of COVID-19 disease was the classic *SEIR model* (Figure 5), wherein,

S = Number of Susceptible individuals
E = Number of Exposed individuals
I = Number of Infected individuals
R = Number of Recovered individuals
D = Number of Dead individuals

The rate of progression from E (exposed) stage to I (infected) stage is alpha (α). I_1, I_2 and I_3 refer to the stages of infection from mild to severe to critical respectively at respective rates, i.e., from I_1 to I_2 at the rate p_1 and from I_2 to I_3 at rate p_2. The rate of recovery from stage I_1 is gamma (γ_1), while from stage I_2 is γ_2 and from stage I_3 is γ_3. Lastly, the individuals die at the rate μ (Zunyou et al., 2020).

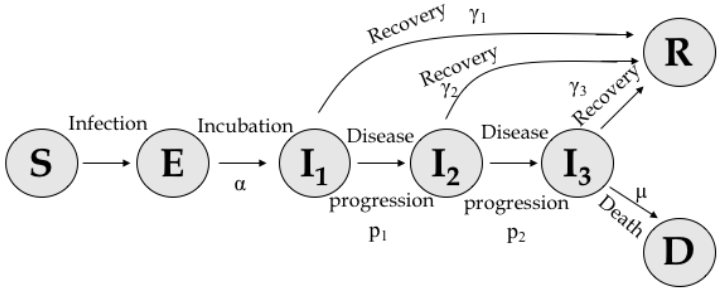

Figure 5. SEIR model (Figure recreated from Creative commons source)

So, the following exercise can be done in the class to study epidemiology of COVID-19 disease using mathematical modelling. The steps are given below:
Step 1. Open the online simulation using link:
https://alhill.shinyapps.io/COVID19seir/?fbclid=IwAR2aXJT79M2AmZxMdy8jsiEuSC4i7ijU8Av6oB4dmlZIeJ2VQgL7Tt3QGxA

DOI: https://doi.org/10.17509/aijbe.v6i2.62077
e-ISSN 2621-7260

Step 2. Click on 'Spread' tab

Step 3. Set the parameters given on the left side of the user interface-clinical parameters and transmission rates. Initially, the graph can be plotted using default parameters, which are as follows:

- Population size – 1000
- Infected individual- 1
- Duration of incubation period- 5 days
- Duration of mild infections- 6 days
- % of infections that are severe-15%
- Duration of severe infection (hospital stay)- 6 days
- % of infections that are critical- 6%
- Duration critical infection (ICU stay)- 8 days
- Death rate for critical infections- 40%
- Mild infections- 0.5/day
- Severe and critical infections- 0.1/day
- Not allowing seasonality in transmission, asymptomatic infections and pre-symptomatic transmissions

After setting the default parameters, we get the following graph (Figure 5).

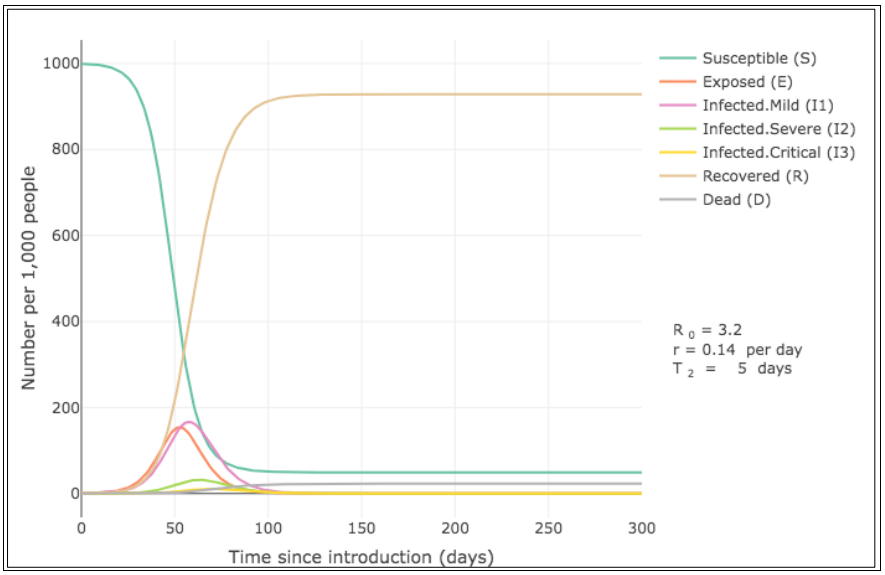

Figure 6. Number of individuals who are susceptible (S), infected (I), recovered (R) or dead (D) over a period of time.

This graph depicts the number of individuals who are susceptible (S), infected (I), recovered (R) or dead (D) over a period of time. As can be seen, the infected individuals (pink line) first pass go through an incubation period, where they do not show any signs of infection or are asymptomatic and then move into the symptomatic and infections stage, which may be mild, severe, or critical (I_1, I_2 or I_3). The value of R_0 denotes the average number of people that a single sick individual infects in an entirely susceptible population.

Step 4. Change the settings like incubation period or duration of mild infection and then also vary the transmission rates. Then observe how the graph changes over the time period and how the value of R_0 changes.

DOI: https://doi.org/10.17509/aijbe.v6i2.62077
e-ISSN 2621-7260

So, including such simple exercises, the biology students are able to visualize a biological aspect from mathematical point of view and the significance of using maths in biology is also emphasized in a simplistic manner. Thus, the examples illustrated in this review present with an alternative pedagogical approach of teaching and learning biology at the senior secondary level, wherein the educators can easily integrate simple mathematical details in explaining the intricacies of biological concepts. The educators can discuss the examples presented in this review in their classroom. The present study presents few examples which are just the tip of an iceberg and deliberate attempts must be done by educators to discuss biological concepts including quantitative and mathematical details so that school students can appreciate biology in its real sense.

CONCLUSION

The inclusion of simple mathematical details in the biological concepts can help the students at senior secondary level towards better understanding of those concepts. The quantification of at least some of the biological concepts which are discussed in this article will enrich the learning of biology as well as mathematics. The article is an attempt to spark and ignite the notion that biology needs to be understood from mathematical point of view as well and encourages the educators to bring out more such concepts and add the layers of quantification around those biological concepts.

REFERENCES

Bormashenko, E. (2022). Fibonacci sequences, symmetry and order in biological patterns, their sources, information origin and the landauer principle. *Biophysica, 2*(3), 292–307.

Breier, A. M., Weier, H. U., Cozzarelli, N. R. (2005). Independence of replisomes in *Escherichia coli* chromosomal replication. *Proceeding of the National Academy Science, 102*(11), 3942-7.

Cooper, S., & Helmstetter, C. E. (1968). Chromosome replication and the division cycle of Escherichia coli B/r. *Journal of Molecular Biology*, 31(3), 519-540.

Dewachter, L., Verstraeten, N., Fauvart, M., Michiels, J. (2018). An integrative view of cell cycle control in *Escherichia coli*. *FEMS Microbiology Review*, *42*(2), 116–136.

Dijkhuizen, P., Buursma, A., Fongers, T. M. E., Gerding, A. M., Oeseburg, B., & Zijlstra, W. G. (1977). The oxygen binding capacity of human haemoglobin. *Pflugers Arch,* 369, 223–231.

E. coli genome project. University of Wisconsin. (2023, June 6). Growth of E. coli strain MG1655. https://www.genome.wisc.edu/resources/k12growth/mg1655growthcurve.htm

Fujioka, A., Terai, K., Itoh, R. E., Aoki, K., Nakamura, T., Kuroda, S., Nishida, E., Matsuda, M. (2006). Dynamics of the Ras/ERK MAPK cascade as monitored by fluorescent probes. *Journal of Biochemical Chemistry*, *281*(13), 8917-8926.

Fung, Y. C. (1993). *Biomechanics: Mechanical Properties of Living Tissues (2nd ed.).* Springer-Verlag.

Gottelli, D., Sillero-Zubiri, C., Applebaum, G. D., Roy, M. S., Girman, D. J., Garcia-Moreno, J., Ostrander, E. A., & Wayne, R. K. (1994). Molecular genetics of the most endangered canid: the Ethiopian wolf *Canis simensis*. *Molecular Ecology, 3*(4), 301–312.

Grammer, R. T. (2017). The poisson distribution in biology instruction. *The American Biology Teacher*, *79* (1), 10–13.

Helmstetter, C. E. (1968). DNA synthesis during the division cycle of rapidly growing Escherichia coli B/r. *Journal of Molecular Biology*, *31*(3), 507-518.

Hendil, K. B., Hartmann-Petersen, R., & Tanaka, K. (2002). 26 S proteasomes function as stable entities. *Journal of Molecular Biology*, *315*(4), 627-636.

Jungck, J.R. (1997). Ten equations that changed biology: Mathematics in problem-solving biology curricula. *Bioscene*, *23*(1), 11–36.

Khan, M.A., & Atangana, A. (2022). Mathematical modeling and analysis of COVID-19: A study of

DOI: https://doi.org/10.17509/aijbe.v6i2.62077
e-ISSN 2621-7260

new variant Omicron. *Physica A*, 599:127452.
Legewie, S., Herzel, H., Westerhoff, H. V., & Blüthgen, N. (2008). Recurrent design patterns in the feedback regulation of the mammalian signalling network. *Molecular Systems Biology*, *4*(190), 1-7.
Persaud-Sharma, D., & O'Leary, J. P. (2015). Fibonacci series, golden proportions, and the human biology. *Austin Journal of Surgery*, *2*(5), 1-6.
Pierigè, F., Serafini, S., Rossi, L., & Magnani, M. (2008). Cell-based drug delivery. *Advanced Drug Delivery Reviews*, 60(2), 286-295.
Reshes, G., Vanounou, S., Fishov, I., & Feingold, M. (2008). Timing the start of division in E. coli: A single-cell study. *Physical Biology*, 5(4), 1-9.
Reyes-Lamothe, R., Possoz, C., Danilova, O., & Sherratt, D. J. (2008). Independent positioning and action of *Escherichia coli* replisomes in live cells. *Cell*, 133(1), 90–102.
Rosen, R. (1967). *Optimality principles in biology.* Springer.
Silverman, L., & Glick, D. (1969). Measurement of protein concentration by quantitative electron microscopy. J*ournal of Cell Biology*, *40*(3), 773-778.
Sinha, S. (2019). The Fibonacci numbers and its amazing applications. *International Journal of Engineering Science Invention*, *6*(9), 7-14.
Sun, H. B., Shen, J., & Yokota, H. (2000). Size-dependent positioning of human chromosomes in interphase nuclei. *Biophysical Journal*, *79*(1), 184–190.
Tao, H., Bausch, C., Richmond, C., Blattner, F. R., & Conway, T. (1999). Functional genomics: Expression analysis of Escherichia coli growing on minimal and rich media. *Journal of Bacteriology*, *181*(20), 6425-6440.
Ullman, G., Wallden, M., Marklund, E. G., Mahmutovic, A., Razinkov, I., & Elf, J. (2013). High-throughput gene expression analysis at the level of single proteins using a microfluidic turbidostat and automated cell tracking. *Philosophical Transaction Royal Society B*, 368, 1-8.
Volpe, P., & Eremenko-Volpe, T. (1970). Quantitative studies on cell proteins in suspension cultures. *European Journal of Biochemistry, 12*(1), 195-200.
Zhang, Q., Zhang Z., & Shi, H. (2020). Cell size is coordinated with cell cycle by regulating initiator protein DnaA in *E. coli*. *Biophysical Journal*, *119*(12), 2537–2557.
Zunyou, W., & McGoogan, M. (2020). Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases from the Chinese Center for Disease Control and Prevention." JAMA, February.

Acknowledgment
I am highly thankful to the Principal, Maitreyi College, University of Delhi for constant support and encouragement.

Authors' Note
The authors declare that there is no conflict of interest regarding the publication of this article. Authors confirmed that the paper was free of plagiarism.

How to Cite this Article
Kaur, J. (2023). Incusion of mathematics in biological concept at the senior secondary level in Indian education system. *Assimilation: Indonesian Journal of Biology Education*, *6*(2), 117-132.

DOI: https://doi.org/10.17509/aijbe.v6i2.62077
e-ISSN 2621-7260