**Paper Details : 1S00144 - S.Y.B.Sc. (Sem. IV) (Choice Base) / 78873 –**

**Life Sciences : Paper III. Date :03-05-2019 Time :02:30 pm - 05:30 pm**

**QPCode :65006**

**Q. 1 Do As Directed: 20M**

**A. Define / Explain 7M**

1. Critical value in hypothesis testing—The value of the sample statistic that defines the regions of acceptance and rejection, is called the critical value.

2. Null hypothesis-- The statistical hypothesis that is set up for testing a hypothesis is known as null hypothesis. It is set up to decide whether to accept or reject the null hypothesis. It asserts that there is no difference between the sample statistic and population parameter. It is denoted by $H_0$.

3. Left tailed test—A test of statistical hypothesis where either alternative is one sided is called as one tailed test. In the left tailed test the critical region or rejection region lies entirely on the left tail of the normal curve, that is $\mu < \mu_0$

4. Level of significance--The level of significance is the maximum probability of making a type 1 error denoted by (alpha), i.e, P (Rejecting). The best value for fixing the level of significance depends on the seriousness of the results of the type of error. The commonly used level of significance in practice are 5% (alpha=0.05) and 1% (alpha=0.01). if we use 5% level of significance (alpha=0.05) we shall mean that the probability of making type 1 error is 0.05 or 5%.

5. Alternative hypothesis-- Alternative hypothesis—the negation of null hypothesis is called 'alternative hypothesis. When Null hypothesis is rejected, automatically alternative hypothesis is accepted. It is denoted by $H_1$.

6. Confidence interval—The two limits within which the estimate for the parameter lies are known as confidence limits and the interval bounded by these two limits is called confidence interval.

7. test statistic—After setting up the null and alternative hypothesis, the test statistic is computed. The test statistic is statistic based on appropriate probability distribution.

**Q. 1 B. 7M**

**a**) Paralogs - iii) Genes within the same species with different function

b) Forward Reading Frame -- vii) 5' -- 3'

c) Evolutionary Clock - v) Fossil records

d) ACC – i) Threonine

e) Putative Genes ii) Unknown Protein function

f) Neighbor joining methods - iv) Phylogenetic tree building

g) UAA - vi) Stop Codon

**Q. 1 C) State Whether True or False:** **6M**

1.True 2.True 3.True 4. 5.True 6.True

**Q. 2 A. Answer Any One of the following:** **10M**

**1.Discuss the influence of reproductive techniques in evolution.**

Reproductive technologies have had a significant impact to the lives of many infertile and sub-fertile couples around the world. However, due to the high financial costs of these procedures, the access to these technologies is largely limited to Western society; particularly middle to high income earners. Consequentially, developing countries whom have the highest rates of infertility, have limited access to these technologies.

The use of these technologies is surrounded with controversy over the social implications involved. In the case of developing countries, some fear allowing access to these societies would lead to increased population growth in already overpopulated environments. A potential consequence of this would include further inequality to resource access, increased risk for the spread of disease, and subsequent extrapolation of financial costs. However this ignites further controversy, as denying the access of these services is considered to violate a basic human right, established in the UN Declaration of Human Rights Article 16.1: Xvi, which states "men and women of full age, without any limitation due to race, nationality or religion, have the right to marry and to found a family.

<u>In-Vitro Fertilisation</u>

In-Vitro Fertilisation (IVF) is an assisted reproductive technology that has been used since the 1950s in animal breeding, and successfully produced its first human child in 1978 with the birth of Louise Brown.The technique requires ovarian hyper stimulation in order to extract a number of developed ova from the ovaries. These are then fertilised external to the body, and the resulting embryo is replaced in the uterus several days later for implantation.IVF is considered to have a notable impact on society, mainly due to its risks and social-evils. The risks of IVF have been well documented, and include multiple pregnancy, ectopic pregnancy, and ovarian hyperstimulation syndrome (OHSS).

The major outcome of IVF is that it has provided a means for many infertile couples/individuals to have children. However in doing so, there are concerns regarding the fertilisation of oocytes outside of the body. Not only is this viewed as unnatural, but it also requires extensive laboratory work in order to retrieve, fertilize and replace the resulting embryo.Additionally, as with many assisted reproductive procedures, success entails an increased risk of having a multiple pregnancy, which has considerable increased health risks for the mother and fetuses. This is because more than one oocyte is often transferred into the fallopian tubes, with the potential for fertilization. This procedure also increases the risk of an ectopic pregnancy, miscarriage, premature birth and other complications. Therefore, it has the potential to lead to significant emotional and financial costs for the family and wider society. 'It has been reported that average, hospital charges for a twin delivery were four times higher than for a singleton, whereas charges for a triplet delivery were eleven times higher. Additionally, there are long term costs associated with complications; including mental retardation, cerebral palsy, chronic problems with lung development and learning disabilities, which increase in frequency with pre-maturity.'Another controversial issue is associated with age. There is debate over what age is too old for a person to undergo IVF in order to have a child, with reports of women utilizing its services after the onset of menopause. This raises concern for the mothers' health in surviving the pregnancy, as well as their ability to survive long enough to raise the child.

**2. Discuss the evolution of Homo sapiens w.r.t development of speech and language for communication.**

Speech is the most physical aspect of language, and as such is the most promising aspect to study in the context of the (biological) evolution of language. However, even for speech, individual facts tend to be equivocal, such that there is no consensus about the interpretation of the evidence that exists. Moreover, some researchers consider the physical instantiation of language to be unimportant for the study of its evolution, because the physical dynamics are seen as a separate, ancillary process that is only used for externalization (Chomsky, 2007).

Perhaps it is useful to start by making a case that the vocal tract did indeed undergo selection related to vocal communication. First of all, the human vocal tract is different from that of other primates: not just is the larynx lower, but humans also have a much bigger gap between the larynx and the velum than do other primates, and even than do other mammals with permanently lowered larynges (increasing the risk of choking on one's food; e.g., Heimlich, 1975). Moreover, the human vocal tract may be optimized for vocalization (de Boer, 2010, 2012; but see Badin, Boë, Sawallis, & Schwartz, 2014). Furthermore, as compared to other apes, we lack air sacs (Fitch, 2000) and have better breathing control (MacLarnon & Hewitt, 2004). In summary, humans have a higher risk of choking while eating but are better at producing carefully controlled vocalizations. Therefore, following Parker and Maynard Smith (1990), who observed that optimization for a function often indicates selective pressure related to that function, it seems likely that humans have undergone selective pressure related to speech.

The first question is important, because it not only addresses the issue of how long ago hominins started to speak, but also indirectly addresses the wider issue of how much time has been available for the evolution of language.

Finally, many conclusions have been based on anatomical evidence for or against the ability to produce complex vocalizations. However, this assumes that (modern) language somehow depends on modern anatomy, and the question is whether this is true.

**Paleontological evidence for speech**

The oldest and most controversial evidence is about the position of the larynx. Lieberman and Crelin (1971) reconstructed the Neanderthal vocal tract on the basis of observations of fossils and of the infant modern human vocal tract. They then used a computer model to calculate the acoustic properties of articulations that could be made with the reconstructed vocal tract. They concluded that Neanderthals were not able to produce the complete range of sounds that modern humans can make. Although this work was groundbreaking in its method, the results have not been universally accepted, in part because their reconstruction was considered wrong (Arensburg, Schepartz, Tillier, Vandermeersch, & Rak, 1990; Schepartz, 1993), but also because the computer model explored only a small set of articulations (Boë, Heim, Honda, & Maeda, 2002; but see also de Boer & Fitch, 2010). More recent reconstructions and computer models have tended to lead to the conclusion that Neanderthals had articulatory abilities similar to those of modern humans (Boë et al. 2002). The main obstacle preventing consensus is that there is no universally accepted reconstruction of the Neanderthal vocal tract on the basis of the fossils we currently have. Nevertheless, more recent review articles have tended to attribute modern human-like articulatory abilities to Neanderthals (Barney et al. 2012; Dediu & Levinson, 2013).

A recent study (D'Anastasio et al. 2013) circumvented the problems of a direct reconstruction of the vocal tract by looking only at the internal composition of a Neanderthal hyoid bone

(Arensburg et al. 1989). Although it was already known that the shape of the Neanderthal hyoid bone falls within the range of modern human possibilities (Arensburg et al. 1990), D'Anastasio et al. found that in addition, the ways in which muscles attached to it and exerted stresses on it are indistinguishable between modern humans and Neanderthals.

Fossil hyoid bones are important for another reason, because they are associated with the presence of air sacs (Fitch, 2000; Steele, Clegg, & Martelli, 2013, for a recent review). All great apes except modern humans have air sacs, and both acoustic analysis (de Boer, 2009; Riede, Tokuda, Munger, & Thomson, 2008) and experiments (de Boer, 2012) have indicated that their presence would decrease the understandability of vocalizations. Fossil hyoid bones of Neanderthals (Arensburg et al. 1989), *Homo heidelbergensis* (Martínez et al. 2008), and contemporaries (Capasso, Michetti, & D'Anastasio, 2008) indicate that they most likely did not have air sacs, whereas a hyoid fragment (a bulla) of *Australopithecus afarensis* that is shaped like that of a chimpanzee (Alemseged et al. 2006) indicates that Australopithecines most likely did have air sacs. Still, this is only indirect evidence, since air sacs may have disappeared in human evolution for other reasons, such as prevention of hyperventilation during long calls (Hewitt, MacLarnon, & Jones, 2002).

MacLarnon and Hewitt (1999, 2004) have also proposed a different fossil indication of the evolution of speech: the thoracic vertebral canal. They argue that a larger thoracic vertebral canal indicates a larger number of neural fibers, and that this can be used to investigate how well the intercostal muscles are controlled. Better control is needed for the extremely long and accurate outbreaths that are used in speech. These researchers found that the thoracic vertebral canal is relatively larger in modern humans and Neanderthals than in other apes, whereas *Homo ergaster* had a thoracic vertebral canal of the size that would be expected in an ape their size.

Finally, Martínez et al. (2004) proposed that the hearing of *Homo heidelbergensis* was similar to that of modern humans, whereas that of chimpanzees and earlier hominins was different. They interpreted this as an indication that modern human (and Neanderthal) hearing is specialized for speech, because it is more sensitive to frequencies in the range of 2–4 kHz. However, it has to be noted that this is rather higher than what is usually considered essential for speech (300–3300 Hz was considered sufficient for analog telephone speech).

Although none of the individual pieces of evidence are unequivocal, when they are taken together a relatively consistent picture emerges: Both *Homo sapiens* and Neanderthals—and by extension their latest common ancestor, who lived about 400,000 years ago—have the same anatomical adaptations to speech, whereas earlier species were probably different. This indicates that the ability to produce complex vocalizations was already present 400,000 years ago. Moreover, the observed changes are very diverse, but they are all explained by the evolution of complex vocalization.

Even if one accepts that adaptations for complex vocalizations evolved around 400,000 years ago, this does not necessarily mean one accepts that language evolved at that time, as well. Two alternatives are possible: that language (in the sense of a learned symbolic and open-ended communication system) came before speech, or that language came after. The first position requires that language was initially conveyed through some other medium, and the gestural-origins hypothesis proposes just that. The second position requires that complex vocalizations were originally used for a different purpose, and this is what is advocated by the musical-origins hypothesis. Both positions, in my opinion, are unlikely.

It is unlikely that language first evolved in the gestural modality and later shifted to the acoustic modality, because even if the acoustic modality has certain advantages (e.g., one can use it in the dark), evolution tends to stick with solutions that work well but are not necessarily optimal. There is certainly a very good case for the multimodal origins of language (and for language still being multimodal today, because other apes (and by homology, our latest common ancestor) are more flexible in using gestures than in using vocalizations and because their communication is generally multimodal anyway. Nevertheless, apes do show some flexibility in vocalization. It therefore seems plausible that vocalizations have played a role in language from the beginning, and that this has caused selective pressure on the vocal tract and other systems involved in speech production and processing.

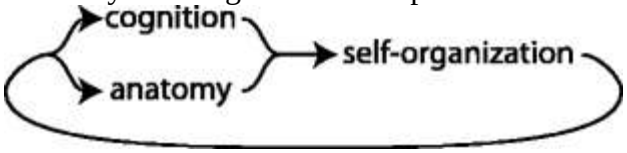
The hypothesis that vocalizations were first used for music (in the form of song) is also problematic, not because music cannot be old or important, but because it seems unlikely that the extra cognitive abilities needed for music would not also directly have given our ancestors an ability for (simple) language. The prelinguistic cognitive abilities of the latest common ancestor with apes, derived from comparisons of ape abilities, consist of the ability to learn a sizable lexicon of form–meaning mappings and the ability to do basic semantics. The innovations needed for music are vocal imitation and vocal control. In addition, if music is to be a social activity (as it appears to be in humans), an ability for music also includes the ability to cooperate, and to do so vocally. It therefore seems more parsimonious to propose that language and music evolved together and are most likely two sides of the same coin than to propose that music evolved much earlier than language. This does not mean that there can be no cognitive specializations for either. The increasing complexity of language and song (under the influence of coevolution; see the next section) could, for instance, lead to linguistic cognitive abilities to deal with complex combinatorial meaning and syntactic structure, as well as musical cognitive abilities to deal with complex rhythm and harmony. Moreover, it should be noted that proponents of musical protolanguage tend to situate the starts of their evolutionary scenarios earlier than 400,000 years ago, usually around the period of *Homo erectus,* so the theory deals with the earliest precursors of language.

These two points indicate that speech and language most likely evolved together, and that we can therefore assume that the beginnings of language are at least 400,000 years old as well.

Coevolution of speech and cognition

A final issue is the relation between cognition and anatomy. It is true that anatomical innovations allow for a larger and more fluent set of utterances, and that this puts pressure on the cognitive systems dealing with speech. Also, it has been shown that self-organization under functional constraints can explain, without recourse to specialist cognitive mechanisms, aspects of language such as the structure of vowel and consonant systems.On the other hand, it has been pointed out that essential ingredients for speech, such as vocal control and the ability to imitate, are lacking in other apes, and by homology in our latest common ancestor

Anatomy and cognition are expected to coevolve under the influence of self-organization:



When a small anatomical or cognitive innovation occurs, self-organization will cause this to be reflected in the language. This will then change the selective pressure on either the cognitive system or the anatomy, thus creating the potential for further adaptations.,instance, self-organization causes the vowel space to be used maximally, but this causes a selective advantage to speakers with a slightly better articulatory ability. Without the effect of self-organization, such modifications would not have an advantage. Thus, self-organization also helps overcome the problem of the frequency dependence of the selective advantage of adaptations to language.Self-organization causes phenomena to emerge in a language that the language users can then adapt to.

The evolution of cognitive and anatomical adaptations is therefore inherently a process of coevolution, and it is misleading to speak about one being more important than the other. Nevertheless, there are arguments for cognitive adaptations being the triggering factor. First, self-organization, in the sense investigated in the articles referred to above, would only be possible if certain cognitive innovations, such as flexibility in vocalization and the ability to do vocal imitation, were already present. In that respect, it is clear that historically, certain cognitive adaptations must have occurred before selective pressure for vocal communication became an issue.

Second, even fully complex modern language is possible without a very large set of speech sounds.
Convergent evidence for adaptations to complex vocalizations in Neanderthals and *Homo heidelbergensis* indicates that adaptations to producing complex vocalizations were already present 400,000 years ago. In combination with what we know about the prelinguistic abilities of other apes (and thus, of our latest common ancestor), it seems likely that some form of language must have been present as well. Given that even with a monkey-like vocal tract it is probably already possible to produce a range of articulations that is sufficient to be usable for language, it seems that cognitive adaptations must have triggered the emergence of language. Rather, the evidence reviewed here indicates a much more gradual process of coevolution between cognition and anatomy; between vocalizations, gesture, and communicative abilities; and between culture and biology, linked through self-organization. It is true that this coevolutionary account does not propose a clear causal factor that may potentially tell us when language emerged precisely, but the complexity of multiple coevolving systems is something that the field of language evolution has been coming to terms with over the last few decades.A coevolutionary account is perhaps less spectacular, but it is much more plausible biologically.

**Q.2.B.Answer any two of the following:**
**1. Comment on "Evolution brings about a rapid change in population".**

Ecological factors exert a range of effects on the dynamics of the evolutionary process. A particularly marked effect comes from population structure, which can affect the probability that new mutations reach fixation. Our interest is in population structures, such as those depicted by 'star graphs', that amplify the effects of selection by further increasing the fixation probability of advantageous mutants and decreasing the fixation probability of disadvantageous mutants. The fact that star graphs increase the fixation probability of beneficial mutations has lead to the conclusion that evolution proceeds more rapidly in star-structured populations, compared with mixed (unstructured) populations. Here, we show that the effects of population structure on the rate of evolution are more complex and subtle than previously recognized and draw attention to the importance of fixation time. By comparing

population structures that amplify selection with other population structures, both analytically and numerically, we show that evolution can slow down substantially even in populations where selection is amplified.The rate of evolution measures how quickly new traits can be established in a population. Typically, this is a function of three factors: mutation rate, population size and the fixation probability of new mutations but there is increasing recognition that ecological influences, such as population structure and the number of competing beneficial mutations contribute additional layers of complexity.The effect of population structure can be particularly strong. This is evident from theoretical studies of evolution on graphs, which show that fixation probability can be enhanced in certain spatially structured populations, compared with unstructured (i.e. well-mixed) populations. Within the category of structured populations, some spatial arrangements, for example, those referred to as 'stars', can significantly amplify the effects of selection. This means that in populations with such structures, the fixation probability of beneficial mutations is greater than in unstructured populations of the same size, whereas the fixation probability of deleterious mutations is smaller than in equally sized unstructured populations. Given that a beneficial mutation has a higher likelihood of fixation in a star-structured population, it follows that the rate of evolution will be more rapid in populations with star structures.Only spatially structured populations with specific structures show amplifying effects. Indeed, for the large class of population structures in which every individual has the same probability of replacement by an offspring from a neighbouring node, fixation probability is unaffected by population structure, as proved in the isothermal theorem. This implies that there are many population structures where the rate of evolution ought not to differ between structured and unstructured populations.

The effect of population structure on the rate of evolution is of more than just academic interest. Population structures that amplify selection therefore have the potential to open new avenues of research. There are four primary factors are required in order to predict the rate of evolution: mutation rate, population size, the fixation probability of new mutations and time to fixation. Only under conditions where the mutation rate is vanishingly small can the influence of the fixation time, mediated by population structure, be neglected.

**2. Explain any one type of post zygotic isolating mechanism with suitable example.**

The following are the types of post zygotic isolating mechanisms,elaborate anyone in detail.

a) Gametic incompatibility. Sperm transfer takes place, but egg is not fertilized.

b) Zygotic mortality. Egg is fertilized, but zygote does not develop.

c) Hybrid inviability. Hybrid embryo forms, but of reduced viability.

d) Hybrid sterility. Hybrid is viable, but resulting adult is sterile.

e) Hybrid breakdown. First generation (F1) hybrids are viable and fertile, but further hybrid generations (F2 and backcrosses) may be inviable or sterile.

**3. Write a note on Gradualism.**

Gradualism is an evolutionary model which theorizes that most speciation is slow, uniform and gradual. When evolution occurs in this mode, it is usually by the steady transformation of a whole species into a new one (through a process called anagenesis). In this view no clear line of demarcation exists between an ancestral species and a descendant species, unless splitting occurs

It also refers to the tiny variations in an organism or in society that happen over time to make a better fit for animals and humans in their environment. These variations allow them to

survive and thrive, resulting in the slow and consistent process of change in the whole population.

Examples:

- Species of butterfly is yellow and black in color. However, a butterfly is born that happens to be orange and yellow in color which makes it difficult to see. Over a long period of time, the yellow and black butterflies die out, because the orange and yellow color combination makes the butterflies less visible to predators.
- Small variations occur over time in a population of wolves - larger ears, longer teeth, and a heightened sense of smell. Wolves with these helpful traits tend to survive better than those without; as time progresses slowly, the traits gradually become the norm among the population.
- Over the period of many, many years, a population of elephants develops larger ears to help protect the elephants from the sun and keep them cool. This larger ear eventually becomes a physical feature of the entire population of elephants.
- A group of finches begins to exhibit different features from those in the rest of the population, including shorter beaks and black stripes on their wings. Eventually, the two distinct groups of birds are considered entirely separate species from one another.
- Changes occur among a group of tree frogs which include color patterns that help them hide from predators and poisonous skin secretions that deter predators from seeking them out as a food source. As time progresses, these traits become standard in the population of tree frogs.
- A bird is born with a longer beak than others in its population, which allows it to more easily forage for insects inside trees, offering it an advantage over its shorter-beaked relatives. As a great deal of time progresses, more and more birds in the population are born with this advantageous feature, eventually resulting in the entire population of birds with longer beaks.
- Over a long period of time, tigers develop the combination of orange and black stripes, which allow them to hide in tall grasses as they stalk their prey. This trait eventually becomes one of the distinguishing features of tigers, as the coloring on all future tigers changes to adopt the orange and black stripe coloring.
- Small variations appear over a long period of time among a population of penguins, including thicker coats of down that allow them to survive in very cold climates. Penguins with these traits survive much better than those with thinner down coats, leading to the changes becoming standard among all penguins.
- A species of flowers adapts over time to attract more bees, a morphological change that allows the flowers to spread their pollen further and increase the size of the population. The adaptation is advantageous to the flowers and, after a long period passes, the flowers all gradually adopt the adaptation.
- Certain moths in one population gradually adopt changes in color and wing shape. After a great deal of time, the two groups of moths develop into entirely distinct species from one another.

**4. What is the importance of speech development in the course of evolution?**

In the course of evolution it has been observed that the anatomical structures may have adapted to speech over evolutionary time and how this can help estimate when speech evolved. It is also important to understand how cultural transmission shapes systems of speech sounds, and how this is important to understand the biological evolution of cognitive adaptations to learning and using speech. It discusses experimental techniques to investigate

cultural evolution of speech in a laboratory setting. From the evidence presented, it is likely that anatomical adaptations to complex vocal communication are at least as old as the latest common ancestor with Neanderthals (c 400 000 years ago), that cognitive adaptations are probably primary (and therefore even older than this), that cultural evolution is very important in shaping (systems of) speech sounds, and that therefore the evolution of speech was a complex co-evolution between anatomy, cognition, and culture.It has helped in better communication.

**Q.3.A Solve any one** **10M**

1. Null hypothesis- the colour of eyes of the son's is not associated with colour of father's eyes.

The observed frequencies are

| | Eyes colour in sons | | | |
|---|---|---|---|---|
| Eyes colour in fathers | | Not light | Light | total |
| | Not light | 230 | 148 | 378 |
| | Light | 151 | 471 | 622 |
| | total | 381 | 619 | 1000 |

Expected frequencies are

| | Not light | light |
|---|---|---|
| Not light | 378 X 381 /1000=144 | 378 x 619/1000=234 |
| light | 622 x 381/1000=237 | 622 x 619/1000=385 |

$\Psi^2 = \sum(O-E)^2 / E$

$=(230-144)^2 /144+ (151-237)^2 / 237+ (148-234)^2 / 234 + (471-385)^2 /385$

$= 133.39$

Given $\Psi^2$for 1 degree of freedom=3.84 < $\Psi^2$cal, therefore null hypothesis is rejected, that is there is an association between the colours of eyes of son and father.

2.

| Roll no. | Before x | AFTER Y | d=X-Y | $d^2$ |
|---|---|---|---|---|
| 1 | 12 | 15 | -3 | 9 |
| 2 | 14 | 16 | -2 | 4 |
| 3 | 11 | 10 | 1 | 1 |
| 4 | 8 | 7 | 1 | 1 |
| 5 | 7 | 5 | 2 | 4 |
| 6 | 10 | 12 | -2 | 4 |
| 7 | 3 | 10 | -7 | 49 |

| | | | | |
|---|---|---|---|---|
| 8 | 0 | 2 | -2 | 4 |
| 9 | 5 | 3 | 2 | 4 |
| 10 | 6 | 8 | -2 | 4 |
| | | | $\sum d=-12$ | $\sum d^2=84$ |
| | | | | |

Null hypothesis: there is no significant difference between memory capacity before and after nourishing food, ie $\mu_x=\mu_y$

Alternative hypothesis $\mu_x \neq \mu_y$ two tailed test

Mean d= -12/10=-1.2

$S^2$= 1/n-1 $[\sum d^2 - (\sum d)^2/n]$

= 1/9(84-144/10)

=7.7

Standard Error = $S/\sqrt{n}$=0.87

Student's t = Mean d / Standard error

=-1.2 /0.87

=-1.3

|t|= 1.3

Given ttab =2.26> tcal

Therefore null hypothesis is accepted that is there is no significant difference in memory capacity of students as an effect of nourishing food.

**Q.3.B. Solve any 2** **10M**

1. Given n=400, Mean X=1570, s= 150, μ=1600

Null hypothesis: the mean life time of the bulbs is 1600 hrs, i.e. μ=1600

Alternative hypothesis μ>1600 right tailed test

Standard error= $s/\sqrt{n}$

= 150/20=7.5

Z= Mean X- μ/ standard error

= 1570-1600 /7.5

= -4, |z|=4

Given Ztab=2.33 < zcal, Null hypothesis rejected, therefore the mean life time of the bulbs is greater than 1600.

**2.Compare and contrast between parametric and non-parametric tests**

| BASIS FOR COMPARISON | PARAMETRIC TEST | NONPARAMETRIC TEST |
|---|---|---|
| Meaning | A statistical test, in which specific assumptions are made about the population parameter is known as parametric test. | A statistical test used in the case of non-metric independent variables, is called non-parametric test. |
| Basis of test statistic | Distribution | Arbitrary |
| Measurement level | Interval or ratio | Nominal or ordinal |
| Measure of central tendency | Mean | Median |
| Information about population | Completely known | Unavailable |
| Applicability | Variables | Variables and Attributes |
| Correlation test | Pearson | Spearman |

3.

Given n1=9, s1= 2.1, n2=13, s2=1.8

$\sigma^2 = n1.s_1^2 / n-1 = 9 \times (2.1)^2 / 8 = 4.96$

$\sigma^2 = n2.\ S_2^2 / n-1 = 13 \times (1.8)^2 / 12 = 3.51$

Null hypothesis $H_0$: $\sigma_1^2 = \sigma_2^2$

Alternative hypothesis $\sigma_1^2 \neq \sigma_2^2$

F=4.96 / 3.51=1.41

F tab =2.85

Fcal < F tab Therefore accept Null hypothesis samples may be drawn from normal population with same standard deviation.

4. Given Mean X=15.8, $s^2$= 10.3, n=9, $t_{0.01}$=3.36

99% confidence limits are

mean$X \pm t_{0.01}\ s/\sqrt{n-1} = 15.8 \pm 3.36\ \sqrt{10.3/8}$

$= 15.8 \pm 3.36 \times 1.135$

= 11.98, 19.61

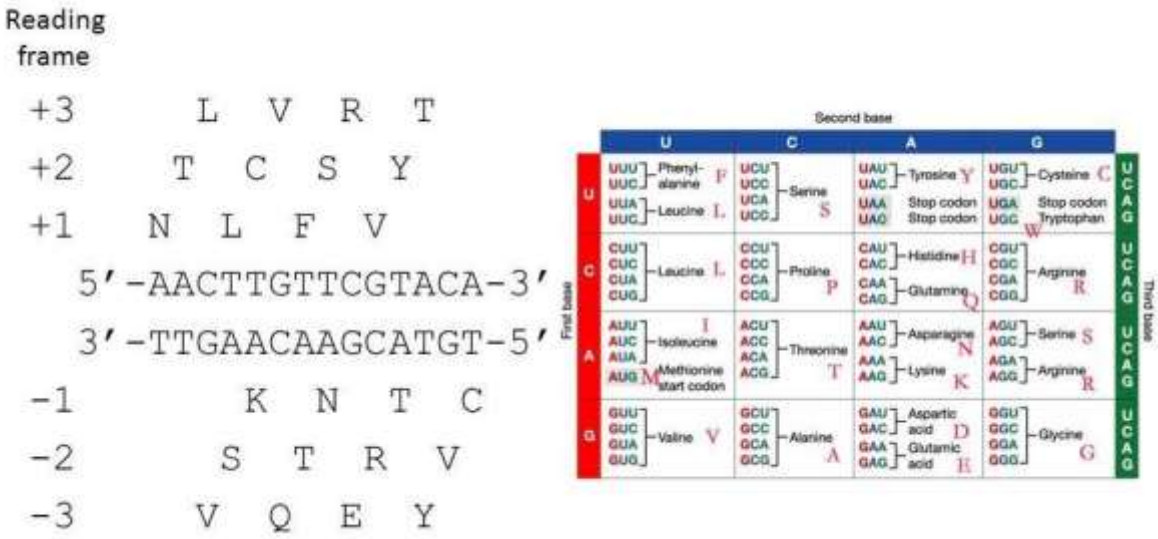
**Q.4. A) Describe any one of the following: (10)**

**1. W.r.t. Amino acid codes, explain Six Frame Translation process with a suitable example.**

1. **Six Frame Translation and its significance:**

In molecular biology, a **reading frame** is a way of dividing the sequence of nucleotides in a nucleic acid (DNA or RNA) molecule into a set of consecutive, non-overlapping triplets. Where these triplets equate to amino acids or stop signals during translation, they are called codons.A single strand of a nucleic acid molecule has a phosphoryl end, called the 5′-end, and a hydroxyl or 3′-end. These define the 5'→3' direction. There are three reading frames that can be read in this 5'→3' direction, each beginning from a different nucleotide in a triplet. In a double stranded nucleic acid, an additional three reading frames may be read from the other, complementary strand in the 5'→3' direction along this strand. As the two strands of a double-stranded nucleic acid molecule are antiparallel, the 5'→3' direction on the second strand corresponds to the 3'→5' direction along the first strand.

One needs to consider *six* reading frames when considering the potential of DNA to encode protein (three frames for each strand). But only one strand is transcribed into RNA — the so-called *coding strand*. It would therefore seem to me that there are actually only *three* reading frames to consider. Explain with suitable example. Of +1,+2,+3 and -1,-2,-3 frames.




**Biological Significance:**

One common use of open reading frames (ORFs) is as one piece of evidence to assist in gene prediction. Long ORFs are often used, along with other evidence, to initially identify candidate protein-coding regions or functional RNA-coding regions in a DNA sequence.[3] The presence of an ORF does not necessarily mean that the region is always translated. For example, in a randomly generated DNA sequence with an equal percentage of each nucleotide, a stop-codon would be expected once every 21 codons. A simple gene prediction algorithm for prokaryotes might look for a start codon followed by an open reading frame that is long enough to encode a typical protein, where the codon usage of that region matches the frequency characteristic for the given organism's coding regions. By itself even a long open reading frame is not conclusive evidence for the presence of a gene. On the other hand, it has been proven that some short ORFs (sORFs) that lack the classical hallmarks of protein-coding genes (both from ncRNAs and mRNAs) can produce functional peptides. 5'NTR of about 50% of mammal mRNAs are known to contain one or several sORFs. 64–75% of experimentally found translation initiation sites of sORFs are conserved in the genomes of human and mouse and may indicate that these elements have function. However, sORFs can often be found only in the minor forms of mRNAs and avoid the selection; the high conservatism of initiation sites may be connected with their location inside promoters of the relevant genes. Such kind of situation is characteristic of SLAMF1 gene, for example.

**2. Construct a cladogram using the information in the table below and answer the following :**

Students should draw a dendrogram with Amoeba as an outgroup and whale and human most closely related species.

1) Identify the organism in the table that is least closely related to the others.

Amoeba

2. Which animal is more closely related to the Humans; the Owl or the Snake? Why?

Whales as it has most of its features common to Humans and it shares the most recent ancestor with Humans.

**Q. 4. B) Explain any two of the following: 10M**

**1. Gene annotation and its importance.**

**DNA annotation** or **genome annotation** is the process of identifying the locations of genes and all of the coding regions in a genome and determining what those genes do. An annotation (irrespective of the context) is a note added by way of explanation or commentary. Once a genome is sequenced, it needs to be annotated to make sense of it. For DNA annotation, a previously unknown sequence representation of genetic material is enriched with information relating genomic position to intron-exon boundaries, regulatory sequences, repeats, gene names and protein products. This annotation is stored in genomic databases such as Mouse Genome Informatics, FlyBase, and WormBase. Educational materials on some aspects of biological annotation from the 2006 Gene Ontology annotation camp and similar events are available at the Gene Ontology website. The National Center for Biomedical Ontology (www.bioontology.org) develops tools for automated annotation of database records based on the textual descriptions of those records.

As a general method, dcGO has an automated procedure for statistically inferring associations between ontology terms and protein domains or combinations of domains from the existing gene/protein-level annotations.

Genome annotation consists of three main steps:. identifying portions of the genome that do not code for proteins

2. Identifying elements on the genome, a process called gene prediction, and
3. attaching biological information to these elements.

Automatic annotation tools try to perform all this by computer analysis, as opposed to manual annotation (a.k.a. curation) which involves human expertise. Ideally, these approaches co-exist and complement each other in the same annotation pipeline.The simpliest way to perform gene annotation relies on homology based search tools, like BLAST, to search for homologous genes in specific databases, the resulting information is then used to annotate genes and genomes. However, nowadays more and more additional information is added to the annotation platform. The additional information allows manual annotators to deconvolute discrepancies between genes that are given the same annotation. Some databases use genome context information, similarity scores, experimental data, and integrations of other resources to provide genome annotations through their Subsystems approach. Other databases (e.g. Ensembl) rely on both curated data sources as well as a range of different software tools in their automated genome annotation pipeline.

*Structural annotation* consists of the identification of genomic elements.

- ORFs and their localization
- gene structure
- coding regions
- location of regulatory motifs

*Functional annotation* consists of attaching biological information to genomic elements.

- biochemical function
- biological function
- involved regulation and interactions
- expression

These steps may involve both biological experiments and *insilico* analysis. Proteogenomics based approaches utilize information from expressed proteins, often derived from mass spectrometry, to improve genomics annotations. A variety of software tools have been developed to permit scientists to view and share genome annotations. Genome annotation remains a major challenge for scientists investigating the human genome, now that the genome sequences of more than a thousand human individuals and several model organisms are largely complete. Identifying the locations of genes and other genetic control elements is often described as defining the biological "parts list" for the assembly and normal operation of an organism. Scientists are still at an early stage in the process of delineating this parts list and in understanding how all the parts "fit together".Genome annotation is an active area of investigation and involves a number of different organizations in the life science community which publish the results of their efforts in publicly available biological databases accessible via the web and other electronic means. Here is an alphabetical listing of on-going projects relevant to genome annotation:

- Encyclopedia of DNA elements (ENCODE)
- Entrez Gene
- Ensembl
- GENCODE
- Gene Ontology Consortium
- GeneRIF
- RefSeq
- Uniprot
- Vertebrate and Genome Annotation Project (Vega)

**2. Any one type of technique employed in studying Molecular phylogeny.**

**There are many methods of studying Molecular Phylogeny listed as follows:**

**a. Immunological** **b. Amino acid comparison**

**b. DNA-DNA and DNA/RNA Hybridization** **d. Repetitive sequence comparison**

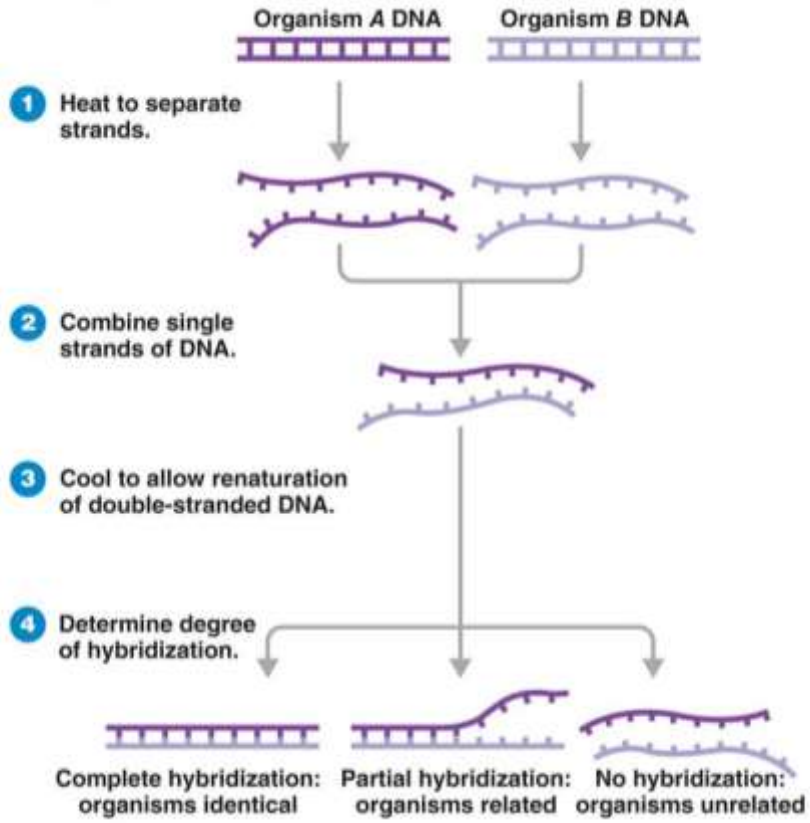
**e. Restriction Enzyme site comparison**

**DNA–DNA hybridization** generally refers to a molecular biology technique that measures the degree of genetic similarity between pools of DNA sequences. It is usually used to determine the genetic distance between two organisms. This has been used extensively in phylogeny and taxonomy.

The DNA of one organism is labelled, then mixed with the unlabelled DNA to be compared against. The mixture is incubated to allow DNA strands to dissociate and then cooled to form renewed hybrid double-stranded DNA. Hybridized sequences with a high degree of similarity will bind more firmly, and require more energy to separate them: i.e. they separate when heated at a higher temperature than dissimilar sequences, a process known as "DNA melting". To assess the melting profile of the hybridized DNA, the double-stranded DNA is bound to a column and the mixture is heated in small steps. At each step, the column is washed; sequences that melt become single-stranded and wash off the column. The temperatures at which labelled DNA comes off the column reflects the amount of similarity

between sequences (and the self-hybridization sample serves as a control). These results are combined to determine the degree of genetic similarity between organisms.

The modern approach is to carry out DNA–DNA hybridization *in silico* using completely or partially sequenced genomes. The GGDC developed at DSMZ is the most accurate known tool for calculating DDH-analogous values. Among other algorithmic improvements, it solves the problem with paralogous sequences by carefully filtering them from the matches between the two genome sequences.




**3. Describe any two types of Phylogenetic tree building methods.**

Algorithmic methods: defined only on the basis of the algorithm. Accomplish the goal of estimating a phylogeny by defining a specific sequence of steps that lead to the determination of a tree. Methods that use algorithms include: Cluster analysis (UPGMA) and Neighbor-joining. These methods are also fall under the category of phenetic methods because they rely on measures of overall similarity. You may recall that one major criticism of phenetic techniques is the inability to distinguish between homology and homoplasy, AND being able to even identify support for specific relationships. A second complaint of Cluster analysis (but not Neighbor-joining) refers to the ultrametric properties of the generated trees. Ultrametric distances are defined by satisfying the three point criteria.

**Neighbor-joining** (Saitou and Nei, 1987) is similar to cluster analysis but removes the assumption that the data are ultrametric. In this method the raw data are provided as a distance matrix, and the initial tree is a star tree (completely unresolved). A modified distance matrix is then created in which the separation between each pair of nodes is adjusted (normalized) on the basis of their average divergence from all other nodes. The tree is constructed by linking the least distant pair of nodes in the modified matrix. When two nodes

are linked, their common ancestral node is then added to the tree and the terminal nodes and their branches are removed. At each stage in the process two terminal nodes are replaced by one node until only two nodes remain, separated by a single branch.

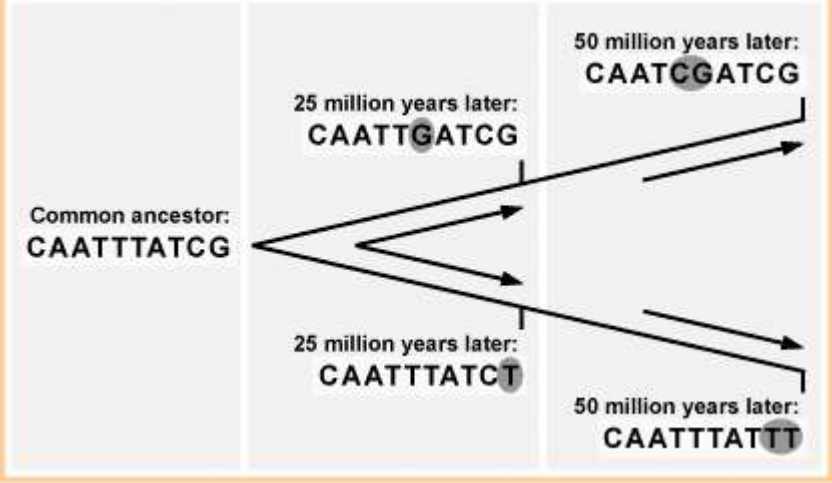
**Maximum Likelihood** methods (Cavalli-Sforza and Edwards, 1967; Felsenstein, 1981; 1993) evaluate hypotheses about evolutionary history in terms of the probability that a proposed model of the evolutionary process and the hypothesized history would give rise to the observed data. The hypothesis with the higher probability of giving rise to the observed data is preferred to one with a lower probability. Likelihood was defined by Edwards (1972) in the following way: "The likelihood, L(H|R), of the hypothesis H given the data R and a specific model, is proportional to P(R|H)...."

**Parsimony methods**
While likelihood methods require an a-priori model of evolution as part of tree construction parsimony methods are considered to be comparatively model free. Several methods based upon the parsimony optimality criterion have been developed and are summarized below.

## 4. Evolutionary Clocks and their significance.

**Evolutionary clocks** are genetic sequences within genes that can help determine when in the past species diverged from a common ancestor. There are certain patterns of nucleotide sequences that are common among related species that seem to change at a regular time interval. Knowing when these sequences changed in relation to the Geologic Time Scale can help determine the age of the species' origin and when speciation occurred. Evolutionary clocks were discovered in 1962 by Linus Pauling and Emile Zuckerkandl. While studying the amino acid sequence in hemoglobin of various species. They noticed that there seemed to be a change in the hemoglobin sequence at regular time intervals throughout the fossil record. This led to the assertion that the evolutionary change of proteins was constant throughout geologic time. Using this knowledge, scientists can predict when two species diverged on the phylogenetic tree of life. The number of differences in the nucleotide sequence of the hemoglobin protein signifies a certain amount of time that has passed since the two species split from the common ancestor. Identifying these differences and calculating the time can help place organisms in the correct place on the phylogenetic tree in respect to closely related species and the common ancestor. There are also limits to how much information an evolutionary clock can give about any species. Most of the time, it cannot give an exact age or time when it was split off of the phylogenetic tree. It can only approximate the time relative to other species on the same tree. Often, the evolutionary clock is set according to concrete evidence from the fossil record. Radiometric dating of fossils can then be compared to the evolutionary clock to get a good estimation of the age of the divergence. A study in 1999 by FJ Ayala came up with five factors that combine to limit the functioning of the evolutionary clock. Those factors are as follows:

- Changing the amount of time between generations
- Population size
- Differences specific to a certain species only
- Change in the function of the protein
- Changes in the mechanism of natural selection




**Q.5. Write Short notes any 4 of the following: 20M**

**1. Social Evolution.**

Social evolution is actually the result of 'group selection', meaning the competition between groups organized according to different rules, that are selected on the basis of their functional adaptation. A social process in some way analogous to the process of biological evolution. Different thinkers have had different aspects of biological evolution in mind (and sometimes, different conceptions of the nature of the biological process). In its most minimal sense, social and cultural evolution can just be thought of as social and cultural change. Social evolution is a process of directional social change, and evolutionary theories attempt to describe and explain this process. Theories of social evolution go back to the second half of the nineteenth century to Spencer, Morgan, Tylor, and Marx and Engels. After a lapse, evolutionary theorizing revived in the 1930s and 1940s with the work of Childe, White, and Steward, and continued into the 1960s and 1970s with the work of Sahlins, Service, Carneiro, Lenski, and Harris. Important typologies of stages of evolutionary development have been developed by most of these thinkers. Although there is far from complete consensus regarding the most important dimensions of social evolution, virtually everyone recognizes the Neolithic Revolution and the rise of civilization and the state as two extremely important evolutionary transformations.

**2. Benefits of Bipedalism.**

Bipedalism allowed hominids to free their arms completely, enabling them to make and use tools efficiently, stretch for fruit in trees and use their hands for social display and communication. They could also see further over the savannah grass – but this also could have been a disadvantage since predators could probably spot them more easily.Bipedal hominids could spend more time foraging and scavenging out in the open savannah because their bodies would be exposed to less sunlight standing upright.Bipedalism allowed hominids to free their arms, allowing the use of tools.Walking on two limbs was also more energy efficient than walking on four – giving early hominids more energy to reproduce and therefore more chance of producing offspring bearing this unique trait.But even with these advantages, these transitional hominids probably spent time in the trees as well.At least some Australopithecus species, including the one represented by "Little Foot" at Sterkfontein, which is as yet unnamed, were at least partly arboreal between 4-million and 3-million years ago, when there was some forest in the Cradle of Humankind environment.Similarly, further north in Africa, the Australopithecus species of Ethiopia and Tanzania between 3-million and 2-million years ago would have been able to climb trees better than modern humans, but were

simultaneously adapting to more full-time upright walking. Australopithecus afarensis, which populated the Afar Depression in Ethiopia, would have lived in an environment typified by wetlands, woodland and forest. But the bipedal footprints of Australopithecus afarensis in Laetoli, Tanzania, are found in an area where the environment was probably drier and sparsely wooded 3.6-million years ago.“Little Foot”, which represents a species of Australopithecus more than 3.3-million years old, was most certainly not a knuckle-walker like some of the great apes. It probably could have walked and climbed effectively.“Little Foot” and other early australopithecines probably climbed trees to escape predators and maybe even to sleep in at night.

**3.**

A. TYPE 1 ERROR: It is the error of rejecting null hypothesis when it is true. When a null hypothesis is true, but the difference (of mean) is significant and the hypothesis is rejected then a Type 1 Error is made. The probability of making a type 1 error is denoted by (alpha), the level of significance. In order to control the type 1 error, the probability of type 1 error is fixed at a certain level of significance (alpha). The probability of making a correct decision is then (1-a).
B. TYPE 2 ERROR: It is the error of accepting the null hypothesis when it is false. In other words when null hypothesis is false, but the difference of means is insignificant and hypothesis is accepted, a type 2 error is made. The probability of making type 2 error is denoted by (beta).

**4. Steps of Mean Square within:**

Find out the deviations of the values of the sample items for all the samples from corresponding means of the samples. Then square of such deviations and finally total the values. This is named as sum of squares within the samples or

SS within = $\sum(X_1\text{-}X_1\text{bar})^2 + \sum(X_2\text{-}X_2\text{bar})^2 + \sum(X_3\text{-}X_3\text{bar})^2 + \ldots\ldots$

Divide the above answer by the degrees of freedom. This is named as Mean Square within the samples or MS within.

**5. Orthologues, Paralogues and their significance in evolutionary relationships between genes.**

**Homolog**

- A gene related to a second gene by descent from a common ancestral DNA sequence. The term, homolog, may apply to the relationship between genes separated by the event of speciation (see ortholog) or to the relationship between genes separated by the event of genetic duplication.

**Ortholog**

- Orthologs are genes in different species that evolved from a common ancestral gene by speciation. Normally, orthologs retain the same function in the course of evolution. Identification of orthologs is critical for reliable prediction of gene function in newly sequenced genomes.

**Speciation**

- Speciation is the origin of a new species capable of making a living in a new way from the species from which it arose. As part of this process it has also acquired some barrier to genetic exchange with the parent species.

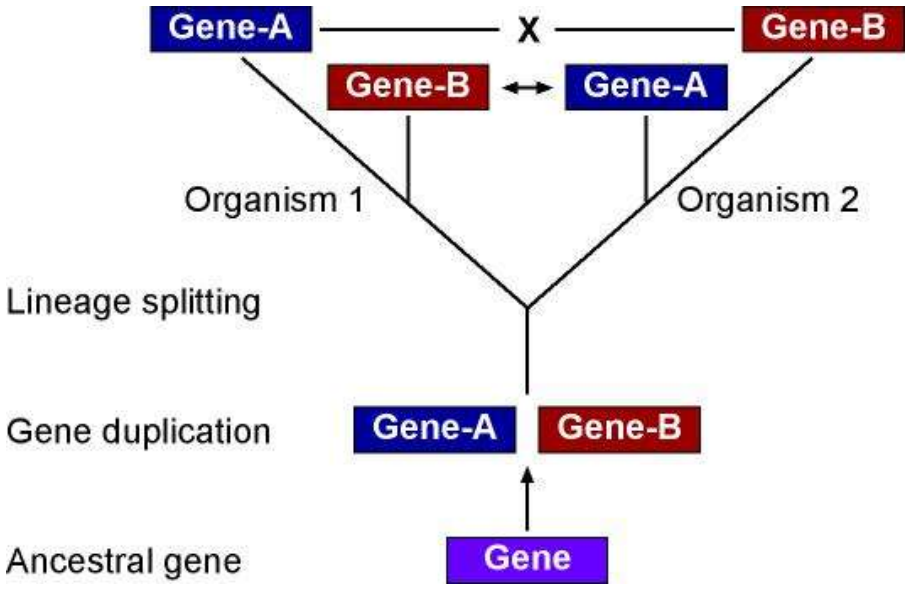
**Paralog**

- Paralogs are genes related by duplication within a genome. Orthologs retain the same function in the course of evolution, whereas paralogs evolve new functions, even if these are related to the original one.

**Why does this study matter?**

In the absence of biochemical assays, the best possible inference for gene function is that it is shared by orthologs, and that gene duplications allow one copy to diverge to take on a new function or to be otherwise specialized (e.g., in timing or location of expression).

- The figure below illustrates how the commonly used method of reciprocal best-BLAST matching leads to incorrect assignment of gene identities (and their correlate, gene function). In this example (found for many real world examples), the evolutionary split between the two organisms has occurred after a gene duplication that generated paralogs named "Gene-A" and "Gene-B". Genes do not all evolve at the same rate and, in this example, we're imagining that it is Gene-B in organism 1 and Gene-A in organism 2 that happen to have the slower rates. That being the case, the reciprocal best matches are between Gene-B of organism 1 and Gene-A of organism 2, so these paralogs are erroneously inferred to be orthologous and assigned the same function. The other two genes are assigned no function at all, since the best match to Gene-A of organism 1 is Gene-A of organism 2, but this is not reciprocal, and similarly for Gene-B of organism 2.
- Only a complete phylogenetic reconstruction using accurate methods - such as is done in the PHRINGE pipeline - can reconstruct this and make guide the proper inference of orthology and functional assignment.

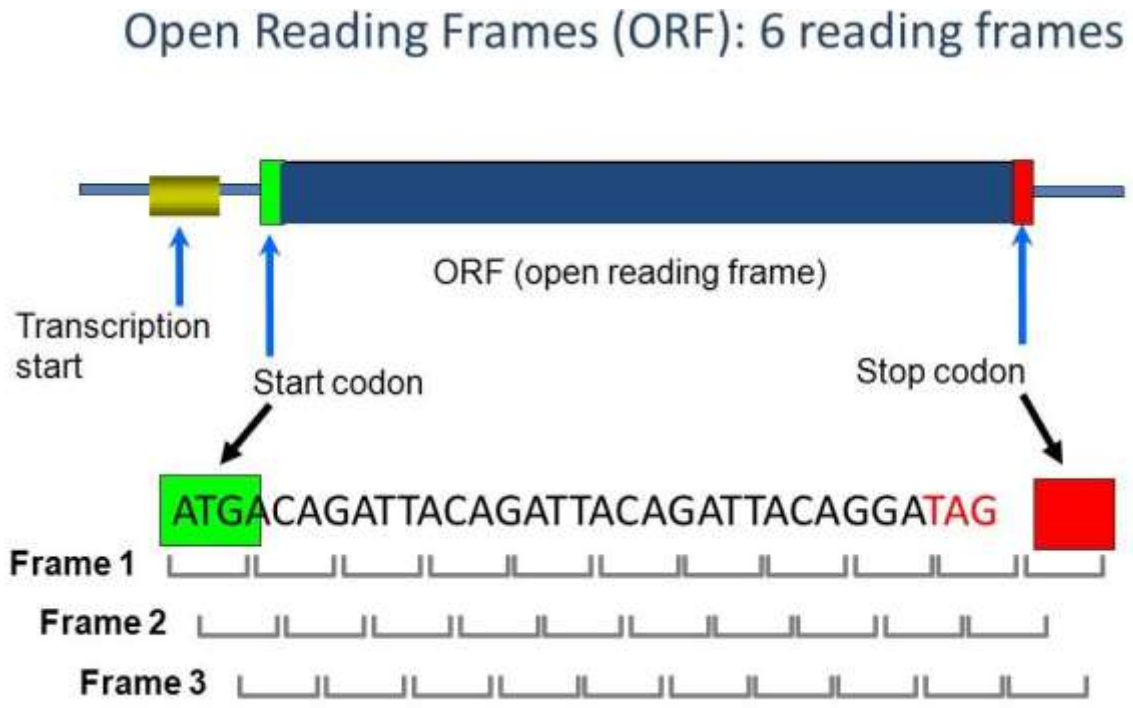



**6. ORF and its significance in protein translation.**

In molecular genetics, an **open reading frame** (**ORF**) is the part of a reading frame that has the ability to be translated. An ORF is a continuous stretch of codons that begins with a start codon (usually AUG) and ends at a stop codon(usually UAA, UAG or UGA). An ATG codon (AUG in terms of RNA) within the ORF (not necessarily the first) may indicate where translation starts. The transcription termination site is located after the ORF, beyond the translationstop codon. If transcription were to cease before the stop codon, an incomplete protein would be made during translation. In eukaryotic genes with multiple exons, introns are removed and exons are then joined together after transcription to yield the final mRNA for protein translation. In the context of gene finding, the start-stop definition of an ORF therefore only applies to spliced mRNAs, not genomic DNA, since introns may contain stop codons and/or cause shifts between reading frames. An alternative definition says that an ORF is a sequence that has a length divisible by three and is bounded by stop codons. This more general definition can also be useful in the context of transcriptomics and/or metagenomics, where start and/or stop codon may not be present in the obtained sequences. Such an ORF corresponds to parts of a gene rather than the complete gene.

One common use of open reading frames (ORFs) is as one piece of evidence to assist in gene prediction. Long ORFs are often used, along with other evidence, to initially identify candidate protein-coding regions or functional RNA-coding regions in a DNA sequence. The presence of an ORF does not necessarily mean that the region is always translated. For example, in a randomly generated DNA sequence with an equal percentage of each nucleotide, a stop-codon would be expected once every 21 codons. A simple gene prediction algorithm for prokaryotes might look for a start codon followed by an open reading frame that is long enough to encode a typical protein, where the codon usage of that region matches the frequency characteristic for the given organism's coding regions. Therefore, some authors say that an ORF should have a minimal length, e.g. 100 codons or 150 codons. By itself even a long open reading frame is not conclusive evidence

for the presence of a gene. On the other hand, it has been proven that some short ORFs (sORFs) that lack the classical hallmarks of protein-coding genes (both from ncRNAs and mRNAs) can produce functional peptides. 5'NTR of about 50% of mammal mRNAs are known to contain one or several sORFs. 64–75% of experimentally found translation initiation sites of sORFs are conserved in the genomes of human and mouse and may indicate that these elements have function.




**********