Rethinking of Evolution

Hangjiong Chen, PhD
hjchen1@yahoo.com

Abstract
Evolution of species can be divided into three stages, origin of life, slow evolution, and fast evolution. These three stages have accomplished the same thing – evolution of life, but in different ways. The nascent earth hosted a special place that served as a life incubator, where basic chemical components were abundant, and conditions were right for random polymerization reactions to occur, forming a pool of polypeptides and possibly random RNA. In origin of life, randomness was the only source of all kinds of macromolecules essential for forming life, and life per se rose from such a pool of randomness. Some of the random polypeptides displayed catalytic activities when folding into three dimensional structures. They were the earliest enzymes and the catalysts of the origin of life. As the pool increased in size, more enzymes of different specificities became available to produce basic small biochemical molecules, protein, RNA, and DNA. Consequently, the earliest tRNA and rRNA emerged from random RNA in the pool. When proteins, RNA, and DNA started to self associate and assemble into special complexes, the precursors to modern ribosomes, replication and transcription complexes emerged. When sequences in the random DNA acted as templates to produce random macromolecules, they slowly developed into genes. When many DNA molecules were linked into single ones, an all-potent DNA molecule – the minimal genome – emerged. The minimal genome was then enveloped in a lipid bilayer membrane, forming the earliest primitive cell – single celled life. This nascent form of life was far from mature and robust, but vulnerable and defenseless against natural elements. Furthermore, their genomes were too small to support evolution. What followed was the slow evolution that lasted 3.5 billion years. In this period life matured first into single celled eukaryotes and then into the simplest multicellular organisms. The genomes underwent dramatic enlargement and the coding gene count increased notably, both of which were based largely on random point mutations and DNA duplication. All of this marked the profound changes that occurred to the organisms in this unprecedentedly long period of time in the history of evolution, implying the unthinkable difficulties for the genetic system to create novel genes de novo and assimilate gene products to become the integral part of life. At the end of slow evolution, organisms were prepared well to enter the fast evolution track. Cambrian explosion marked the beginning of fast evolution, in which species evolved via evolution cycles. Protein variants and gene duplications had played critical roles in the emergence of new species. An evolution cycle was a series of genetic events, and all the organisms that appeared in the cycle were intermediates of the cycle. It started when the ancestor organisms were struck by mutations of large magnitude, which threw the ancestor organisms out of the stable disarmed state and entered an unstable armed state. In armed state, the process genotype reshape brought numerous mutations to the genomes at rates greater than the normal mutational rates, resulting in significant changes in morphology and physiology to the intermediates. The reshape process slowly diminished in magnitude and decayed into the process genotype healing, during which the survived intermediates gradually regained stable disarmed states, signifying the emergence of new species and the end of an evolution cycle. Thus evolution of species occurred only in evolution cycles. Once new species came into being, their genomes exhibited remarkable stability as the result of zero sum rule, which determined that all the mutations will result in net gain of zero. Most mutations were deleterious and undermined the dedicate balance maintained among all the biochemical and cellular components of the mutation carriers, causing the carriers to disappeared from the population. Therefore, mutational changes in a species is always short lasting negative sum changes. The zero sum rule maintains the stability and thus the continuity of species throughout the evolutionary timeline, as manifested by the extraordinary modern biodiversity.

1. Introduction

Life began on earth about 4 billion years ago, but how the primitive form of life came into being for the very first time is a forever mystery. Chemistry tells us that a chemical reaction will occur if conditions are right, regardless of whether it occurs in a laboratory or in natural environment. When all the basic chemical components and conditions for life existed, chemical processes took place spontaneously and led to forming early life: lipids for cell membranes, ribonucleotides for genetic materials, amino acids for proteins, and carbohydrates for energy and structures. The early earth must be such a lucky planet in the universe. It was shielded under the nourishing atmosphere and boasted an environment where biochemical reactions seen in a living organism could occur. This environment formed a cozy incubator for life, from which the emergence of the primeval form of life was just a matter of time.

Life today is so rich in forms ranging from simple bacteria and archaea to highly developed eukaryotes. Despite all this, all modern living organisms use the same set of amino acids, same set of genetic codons, same set of nucleobases, and same set of lipids, suggesting that life as we see today originates from a single ancestor in a single place on the prehistoric earth. Then a long journey of evolution brought the early life to such an extraordinary modern diversity. From the very beginning, life has striven on itself for existence, renewal and flourishing and orchestrated its own entire life cycle from inception, embryonic development, birth, maturation, reproduction, and finally to death without input of external instructions. All this occurs thanks to the genome enclosed in the nucleus of the cells. The genome is the most glorious wonder in the universe.

The DNA genome displays dual distinguished characteristics that are fundamental to evolution of life – stability and mutability. The genome is the longest lived biological molecules, passed down from their very ancestors that emerged million, even billion years ago. Genome stability ascertains the continuity of the species. Meanwhile, the genome is highly mutable as mutations occur to its nucleotide sequences all the time, especial during geological and climate changes. Mutability is the foundation of evolution for the divergence of a myriad of new species since the origin of life.

In genetics, any sequence alterations in the genome of an organism are mutations, or genetic mutations. Point mutations are completely random and refer to single base substitution, deletion or insertion. Point mutations are a type of replication errors. Mutations also include deletions or insertions of short pieces of DNA sequences. Mutations of large magnitude refer to changes that alter chromosomal structure in a considerable degree. Gene duplications are a type of sequence amplifications, while chromosomal translocations and chromosomal inversions are types of DNA rearrangement that change the orientation or location of a segment of DNA in the genome. Deletions of large chromosomal regions can lead to loss of the genes within those regions. Deletions or insertions of a segment of DNA sequence can bring together separate genes to produce functionally distinct hybrid genes. All genetic mutations can be lethal if they disrupt genes that are vital to the organisms.

Mutations that are more relevant to evolution are point mutations and gene duplications. Point mutations accounted for most of the mutations introduced by DNA polymerases during germline division and became more frequent when the fidelity of DNA polymerases is reduced. Point mutations can be lethal if they shift or disrupt the reading frames for protein translation. Normally the DNA polymerases replicate DNA with high fidelity, resulting in low mutational rates and stable biotic world. Gene duplication is a process to make a new copy of DNA fragment that contains a gene, a special type of DNA rearrangement. Gene duplication is a major mechanism that the genome generates new genetic material for the evolution of new species. Gene duplications remain common in most species today, but its biological significance is unknown.

How did such a wonder arise in the ancient earth is not only intriguing, but also awe-inspiring, worth every effort to ponder and explore. It has been firmly established that organisms evolve from simple to complex and from low to advanced over billions of years. However, how does evolution occur isn't certain, and a general consensus is that natural selection is a key mechanism of evolution. In this paper, I have presented my random thoughts about evolution of life and criticized the theory of natural selection as a misinterpretation of evolution.

2. Life Timeline on Earth

Primitive life on earth can be dated back to 4 billion years ago, about 5 million years after earth was formed. Figure 1 shows the timeline of life evolving from the most primitive forms to simplest single celled forms to modern mammals, although it is approximate only.

A striking characteristic of the timeline is that it dedicated a stunning long period of 2 billion years (from 4 to 2 billion years ago) to the development of the simplest forms of life, prokaryotes, including bacteria and archaea. This is followed by 5 million years (from 2 billion years to 1.5 billion years) for the single celled eukaryotes. This signifies the difficulty of life arising and surviving in the primeval time. The next 1 billion years witness the rise of multicellular eukaryotic life like fungi and slime molds. Until around 500 million years ago, an eon of accelerated evolution, living organisms began to diverge into all forms and complexities, resulting in the appearance of abundant new species of plants and animals that dominate the earth thereafter. This eon is divided into a few geological periods.

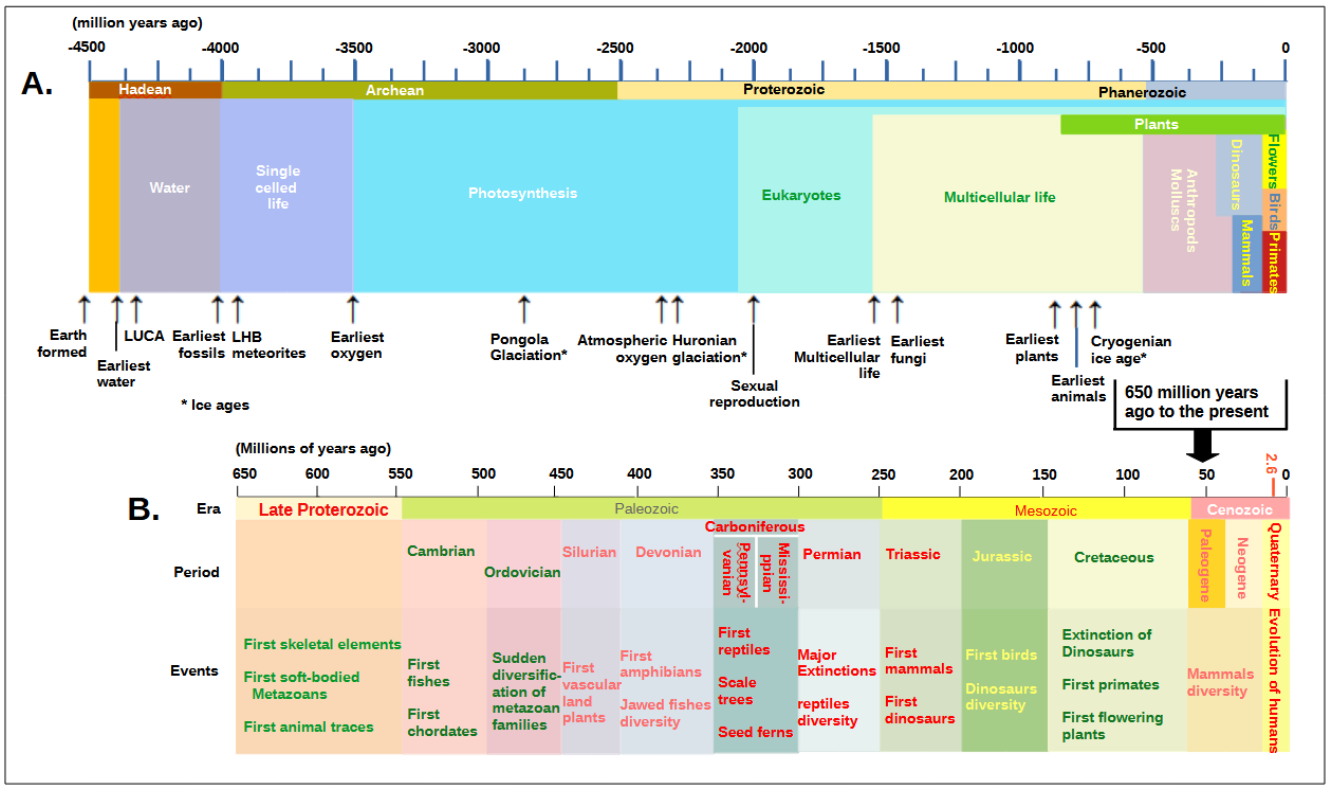

Figure 1. Timeline of the evolution of life on Earth (Adopted from Evolution on Wikipedia and Britannica). Geologic period Phanerozoic comprises the Paleozoic, Mesozoic, and Cenozoic periods

In the Cambrian period (about 539 to 485 million years ago) the earth endured large changes from the preceding geological period in climate, earth's biosphere, and geography that impacted life of that time

with the greatest significance. The changes caused the destruction of natural environments and mass extinction of species, but more importantly the changes led to the emergence of many new species, some of which started to move from ocean to land. This is a time of rapid evolution and diversification of life on earth and known as Cambrian explosion. The beginning of Cambrian explosion heralded the acceleration in biotic diversity, though the species were still as low and simple as comb jellies, sponges, corals, etc. The earliest known vertebrates also originate from Cambrian explosion.

In the Devonian period from 419 to 359 million years ago, arthropods insects, spiders, centipedes, etc. became part of the land ecosystem, and vertebrates started to move to the land as well. In Cretaceous period from 145 to 66 million years ago, more species of mammals, birds, and flowering plants appeared. In this period, first primates emerged and all dinosaurs went extinction. The last 66 million years of earth's history is marked by the dominance of mammals, birds, and flowering plants. More insects, butterflies, moths, fishes, amphibians, and reptiles with modern forms took over the earth long after mammals and birds emerged. Later appearance rewarded these low species with morphology and cellular and biochemical processes that are more advanced and sophisticated than their earlier cousins.

Diversification of primates occurred around 50 million years ago, while the apes, which were evolved from primates and gave rise to the early humans, emerged some 15–20 million years ago. Early humans called hominins diverged from the apes from 14 to 2 million years ago, a time span that is very short on the evolutionary timeline, giving the large morphological changes from apes to hominins. True modern humans are now generally believed to emerge in Africa approximately 300,000 years ago, and then migrate to other continents some 100,000 to 50,000 years ago.

From the timeline of the evolutionary process, the time taken for living organisms to evolve from the very beginning to present day can be divided into three stages (Figure 2). The first stage was dedicated to the origin of life, the buildup of primitive life system from basic chemical components over a period of 500 million years. This stage is not considered as evolution per se, but origin of life. Only when life has formed, the evolution process can commence. Evolution occurs in the two later stages, referred to as slow evolution stage and fast evolution stage, respectively. Dividing evolution into slow and fast stages has profound implications about how evolution really occurs. The entire evolution process is the adventure that has lasted 4 billion years. In this unthinkably long period, about 85% of the time was devoted to the slow evolution, while only 15% to the fast evolution.

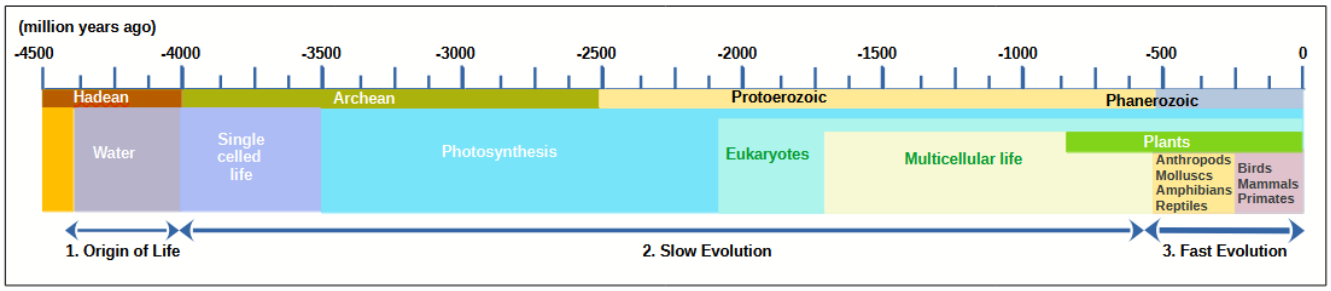

Figure 2. The evolution is a three-stage adventure, each later stage relies on the earlier stage.

Uneven distribution of evolutionary events on the evolutionary timeline is intriguing. Why did it take more than 3.5 billions of years for life to evolve to aquatic plants and animals that are very low and simple in terms of tissue differentiation when comparing with modern day life, while it took only 500 millions of years, especially the last 100 millions of years, for life to flourish with millions of species of all complexities and forms? What is hidden behind this timeline of the evolutionary history of life?

3. Origin of Life – Randomness Brought Life to the Nascent Earth

Proteins, RNA and DNA are not ordinary molecules, they are independent chemical entities that are life in its simplest forms. These molecules are so tightly interlinked that one can't be produced without the other two. A pressing question is how the initial proteins, RNA and DNA could be produced in the incubator? And how could these independent chemical entities become interlinked and assembled into the earliest form of life? What must be true is that the life incubator was an environment in which the conditions favored the chemical reactions to produce proteins, RNA and DNA, possibly facilitated with unknown non-emzymatic catalysts.

Basic chemicals for life are simple organic molecules, amino acids for proteins, and nitrogenous bases combined with pentose sugar riboses for RNA and DNA. The initial sources of bases, sugar, and amino acids could be either randomly produced in the incubator or traced to comets and meteorites traveling through the earth or both. However the extraterrestrial origin was less likely unless the earth was hit regularly with those outer space objects at that time. Regardless of their origins, chemistry of these small organic molecules play far more important roles in the origin of life.

A dipeptide is produced when the carboxylic acid group of one amino acid reacts with the amine group of another to form a covalent chemical bond called peptide bond. The peptide bond is relatively stable under physiological conditions. Dipeptides could elongate at both sides by accepting more amino acids through the same peptide bonding, resulting in polypeptides. The polypeptides so produced would be linear and random but infinite in sequence and length, forming a pool of polypeptides in the primeval incubator. Polypeptides in the pool could transform spontaneously from unstable random coils into more ordered three-dimensional structures, allowing some of them to become biologically functional, including catalytic activities or structural capabilities. If one random polypeptide molecule out of 100 millions could gain a specific three-dimensional structure to become an enzyme, 100 different enzymes could emerge when the size of peptide population reached, say, 10 billions. Larger the polypeptide pool, more enzymes with a wider variety of catalytic activities. Enzymes will do catalytic work to accelerate chemical reactions whenever their substrates are available, igniting all possibilities for life. The debut of enzymes and structural proteins must have had profound impact on what could happen in the very early stage in the incubator, bringing up the idea that enzymes had played decisive roles in the origin of life.

From pure chemistry point of view, ribonucleotides are more complex than amino acids in terms of chemical composition and structure, since they are composed of three totally different small molecules – a nitrogenous base, a pentose sugar ribose, and a phosphate group. The ribose molecule can conform to various configurations in solution, and only β-D-ribofuranose is the right form for ribonucleotide. Moreover, the relative abundance of β-D-ribofuranose in all ribose forms are quite low. A ribose must accept a base at 1′ position to produce a ribonucleoside molecule, which then reacts with a phosphate group at the 5′ position to becomes a ribonucleotide. All products with base and phosphate attached to other hydroxyl groups are not the building blocks for RNA and DNA. In all likelihood, the amount of ribonucleotide in the incubator couldn't be sufficient to warrant RNA synthesis in anyway unless special enzymes from the polypeptide pool made ribonucleotide production biologically significant. It could be concluded that polypeptides were produced earlier than ribonucleic acids in the nascent incubator and some of them folded into unique three-dimensional structures, among which were enzymes with specificity for ribonucleotide synthesis. The most likely scenario could be that the emergence of a large random polypeptide pool was a prelude to the emergence of nucleic acids.

Linking ribonucleotides into a polymer in the strict order of 3′–5′ orientation isn't a straight-forward chemical reaction in the absence of enzymes even with ribonucleotides available in sufficient amount, if not impossible. One possibility is that some unknown special surfaces in the incubator could attract ribonucleotides to adhere. If ribonucleotide molecules that laid on the surface were close enough, adjacent ribonucleotides could form 3′–5′ phosphodiester linkage. The reactions could continue infinitely, producing RNA molecules of various lengths. Replication of RNA molecules could be achieved in similar fashion except that the complementary bases might be snapped into positions on the RNA molecule serving as a template through hydrogen bonding. This step could continue as a process, producing a complementary chain in the form of double stranded RNA. Like all other polymerization reactions in similar scenarios, this type of RNA production was very low in efficiency. Nevertheless, when the peptide pool happened to contain enzymes that could catalyze the formation of 3′–5′ linkage, synthesis of RNA suddenly became not only possible, but also in quantities sufficient for RNA to do what RNA could do. This enzyme was only rudimentary relative to modern RNA polymerases, only serving as a RNA synthase to incorporate random ribonucleotides into a polyribonucleotide chain in a random sequential order. RNA synthases, like hypothesized special surfaces, produced RNA molecules of infinite lengths and base compositions, with or without templates.

The deoxy form of ribonucleotides – deoxyribonucleotides – is more stable and fits better to be genetic materials than ribonucleotides. In living organisms today production of deoxyribonucleotides from ribonucleotides requires an extra reaction, in which the 2′ hydroxyl group of the ribose is replaced with a hydrogen atom catalyzed by ribonucleotide reductases. Reduction of ribonucleotides in the ancient time could be different, but still possible if the incubator contained some non-enzyme catalysts to make this reaction happen. A more likely scenario was that ribonucleotide reductases happened to be part of random enzymes in the pool, thus production of deoxyribonucleotides was as easy as ribonucleotides. Similar to RNA generation and replication, DNA synthases might be lucky ones in the peptide pool, generating DNA molecules of random length and random base compositions. It was likely that RNA synthases and DNA synthases were the same enzyme that were not specific enough to distinguish ribonucleotides from deoxyribonucleotides. It's pure speculation, but any possibilities were possible when facing a magic life incubator in that unknown and mysterious world.

Proteins, RNA and DNA could appear independently of each other, but it's far more likely than any other possibilities that the proteins came into existence first and random enzymes from the peptide pool catalyzed the syntheses of proteins, RNA, DNA, and other small biochemical molecules essential for a process called life. Despite current general consensus that DNA is the last component to be part of life due to the extra reduction step, DNA was more than likely to be a contemporary fellow of RNA.

Life in its earliest moment could be conceived simply as a random existence. Polymerization of ribonucleotides, deoxyribonucleotides, and amino acids was all merely a type of random reactions, and the products were all random in terms of sequence and length. The beauty of random production in the dark and chaotic age is that randomness could have led to the availability of a tiny number of useful molecules with biochemical significance out of a vast amount of random stuff. The number of useful molecules would build up as the random pool continued to build up. In some point in time, the life incubator had accumulated a sufficient number of useful molecules, including a variety of enzymes with different specificities, among which were the rudimentary RNA polymerases, DNA polymerases, aminoacyl tRNA synthetases, ribonucleotide reductases, etc. Even if these early forms of polymerases might be only able to add substrates to the 3' or 5' ends in a totally random fashion, they had greatly accelerated the expansion of randomness of RNA and DNA sequence populations. In addition, amino acids, even ribonucleosides, ribonucleotides, deoxyribonucleotides could be synthesized in large quantities from more basic chemical components present in the incubator with specific enzymes. Over

a long period of time, the incubator had massed a variety of molecules large or small, such as nucleic acids, peptides, lipids, carbohydrates, and molecules of unknown identities and functions. And consequently, life started to form and develop.

Early synthesis of RNA molecules was template independent and totally random, resulting in a large and ever-increasingly heterogeneous RNA population in the incubator. Among the population were sequences that could fold on itself to assume double-stranded secondary structures characteristic of modern tRNA and rRNA. If one tRNA or rRNA like molecule showed up out of 100 millions, about 100 tRNA or rRNA like molecules would emerge when the RNA population increased faster than the rates of natural degradation and reached, say, 10 billions. These tRNA and rRNA like molecules could have played important roles in the early phase of life development and they were the early predecessors of modern tRNA and rRNA. Those RNA without secondary structures in the population would be the predecessors of modern mRNA.

When a myriad of random little things were moving around aimlessly in the dark, the chances for right components to come across and interact were high, resulting in the formation of special structural complexes. The first complex formed in this way would be most likely the precursor to ribosomes or protoribosomes for protein translation. It would form when rRNA-like RNA bumped into proteins with affinity for it. Such a complex would evolve slowly in size and complexity as more components joined in once they became available. Furthermore, there were random peptides that could aggregate with RNA or DNA polymerases to form masses that could act as the primitive platforms for the transcription of RNA and replication of DNA. Such platforms must be poor in performing its functions in terms of output and accuracy, but at least synthesis of RNA and DNA was no longer a type of random polymerization, but catalyzed by enzymes on a crude platform. The incubator had established itself as the common home for proteins, RNA and DNA. This is the likely scenario of the early phase of life in the making.

In the above early phase of life, enzyme catalyzed synthesis of large polymers was largely lacking specificity, but making chain linkages faster and more efficient. The hydroxyl group of the 3' end of a tRNA molecule would be able to form an ester bond with the carboxyl group of any amino acid, and the reaction was greatly facilitated by the early form of aminoacyl tRNA synthetase like enzyme. A rRNA containing ribosomal like complex held a mRNA-like template, allowing many tRNA molecules that were charged with random amino acids to align themselves along the mRNA template without much specificity. The complex so assembled would be the most basal form of peptide synthesis platform, the rudimentary ribosomes, but it was a giant step forward in the origin of life.

Regardless of a large heterogeneous RNA population produced initially without templates in the incubator, the heterogeneity would be augmented sooner or later by template-dependent production even if DNA wasn’t a contemporary fellow of RNA at first. A DNA population in the incubator could be synthesized through elongation by incorporating random deoxyribonucleotides at the 3′ or 5′ ends under the catalysis of the template-independent DNA polymerases. Assuming a particular DNA molecule in the incubator. In the absence of modern DNA replication complex and RNA transcription machine, the replication and transcription processes were awfully egregious. Each replication process introduced a considerable amount of mismatches into the sequence, quickly turning this single DNA sequence into a heterogeneous DNA population. Each of these single DNA sequences could be transcribed into RNA molecules, forming a large heterogeneous RNA population. Since template dependent RNA population continuously mixed into the templateless RNA population, the random RNA population increased significantly, becoming a great treasure for potential useful proteins.

Life is not a random existence per se, but an unusually ordered and consistent living entity. Nascent life successfully moved out of randomness by establishing consistency through precisely controlling all the reactions vital to life with protein catalysts – enzymes. In this remarkable transition, gradual shift to DNA based randomness from total randomness is the turning point in the origin of life. This shift was made possible only when enzymes similar to ribonucleotide reductases appeared in the pool to produce deoxyribonucleotides. The DNA based randomness served as a firm ground on which randomness was gradually replaced with ordered operations.

DNA replication, RNA transcription, and protein translation are among the most basic, but the most complicated biochemical processes in all living organisms. All these processes are built upon their own unique macromolecular machines to proceed. When we talk about aminoacyl tRNA synthetases, RNA and DNA polymerases, ribosomes, each of them is not a simple collection of certain protein molecules, but a super complex of them packaged into unique structure. From the standpoint of evolution, these processes must be among the earliest processes to establish before life could develop further. When these super protein complexes coexisted in the incubator, they together were able to perform a job that started from DNA template, generated RNA as intermediates, and ended as protein molecules, the earliest mechanism of protein biosynthesis. At this stage the machine was preliminary, starting from functionally simple components barely capable of linking amino acids or nucleotides into polymers and nothing more, but eventually it grew slowly into modern macromolecular machine that has operated the most critical part of all forms of life since the origin of life.

The appearance of the protein synthesis machine allowed the same polypeptides to be produced from the random DNA templates in a more dependable way. As the randomness transition moved forward, the polypeptide pool bore a wider spectrum of biological functionalities, including enzymes of more varieties and better quality, regulatory protein factors, structural components, etc. As a result, the machine would have some components replaced by proteins that outperformed them and also have novel components added to become more adequate in functionalities. An improved protein synthesis machine would make protein production even more stable and dependable. On the other hand, proteins from old random polymerization played only complementary and diminishing roles until they all disappeared from the processes. DNA independent randomness thus faded away altogether.

As the protein synthesis machine evolved over time, it became increasingly dependable in functions. At the same time, it wasn’t simply a mix of some components to make random polypeptides from random DNA sequences any more, but was a sophisticated, functionally differentiated super structure that was assembled from many novel enzymes and protein factors in the pool, becoming a system composed of DNA replication machine, RNA transcription machine, and protein translation machine, albeit still very much immature. Such a system was the basic fulfillment of the prerequisite for the initial life to appear as an information flow from DNA to RNA to proteins. This information flow enabled all the future biochemical processes in early life to proceed with extraordinary consistency and regularity. From evolution point of view, it’s the two types of randomness that complemented each other early on to generate a large pool of random polypeptides containing a variety of quality enzymes to start early faint life activities. Without the absolute randomness to start with, it’s impossible to successfully establish DNA template based system for protein production, so it’s impossible to bring life into being.

Numerous chemical reactions in living organism, including reactions to form peptide bonds and ester bonds, are unfavorable in the absence of energy input and can’t take place spontaneously. It was a mystery how synthesis of peptides and nucleic acids in the incubator was made possible before the appearance of adenosine triphosphates or ATP. It was possible that there were unknown non-enzymatic catalysts to move these reactions forward slowly with some form of energy input. However, it must be

true that as the system developed over time, many more functional enzymes emerged, among which were enzymes that catalyzed energy generation in the form of ATP. ATP as a metabolism product had a far reaching impact on the evolution of life. The advent of ATP greatly increased the rates of energy-consuming biochemical reactions through chemical coupling, making it possible to metabolize chemicals available from the environment to produce necessary basic chemical components for the biosynthesis of amino acids, lipids, bases, and sugars. A more likely scenario was that ATP production enzymes appeared early in the incubator.

As DNA molecules grew infinitely in length and continued to replicate through error-prone DNA polymerases, they harbored more sequences that served as templates for all the heritable functional proteins in the pool and the templates for all types of the early heritable RNA, the predecessors to modern tRNA, rRNA, and mRNA. These DNA molecules were essentially the early forms of DNA genomes that accommodated sequence loci that were more or less the early forms of genes. Although these early forms of genomes and genes bore only the faint resemblance to their modern counterparts, they made up a true genetic machine in its primitive forms and established DNA as dependable genetic materials. Most importantly, in such a genetic machine, proteins, RNA and DNA were no longer the independent ordinary molecules, they had become interlinked in a single system in which one's production became difficult without the other two. As the genetic machine evolved over time in form, accuracy, efficiency, and complexity, it reduced the mutational rates in DNA replication, thus reducing the chances of randomness in the information flow. All this greatly accelerated the emergence of life as a self organizing living system.

The great heterogeneity of DNA populations in the incubator implied that many functional protein molecules were produced from different DNA templates, a huge problem for life as an integral entity. The DNA ligases and enzymes for DNA recombination provided an effective solution to this problem. Combining multiple DNA molecules into single ones occurred at the time DNA ligases appeared in the incubator. An all-potent DNA molecule emerged when a number of DNA sequences, each of which harbored a rich set of enzymes and structural proteins, were ligated into a single one. This was the earliest DNA molecule that showed structural characteristics analogous to a genome. This genome analogue slowly grew in length by appending more DNA sequences to its two ends and its gene-like sequence loci matured into an array of genes with regulatory features that could perform the very basic functions vital to the primitive life. The genome analogue transitioned into the minimal genome when it was sustainable on itself with complete information flow from DNA to RNA to proteins and encoded all proteins and RNA elements required to become the most basic, yet independent, self sufficient form of life. Quickly this minimal genome dominated DNA population because of its superior replication capability and became the most common ancestor of many life and finally the common ancestor of all life at the time. Life that descended from it shared the same set of amino acids, same set of codons and anti-codons, same set of bases, same set of metabolism pathways, etc. This DNA molecule is eventually the common ancestor of all modern life as well.

The minimal genome continued to expand and hosted a growing list of enzymes of more varieties, ion transporters, proteins for cell division, structural protein filaments to make cytoskeleton of the cell, etc. Nascent metabolism pathways could have started to generate energy from carbohydrates and produce key chemical compounds for building basic cellular structures, especially cell membranes and cell walls. The self organizing nature of proteins allowed individual proteins, when coexisting, to perform whatever functions the individual proteins could perform in their own space. For example, if some proteins could divide a cell apart into two when mixed together, they would divide the cell apart into two when coexisting in the same cell. If a protein could transport sugar molecules across the cell membrane, it would transport sugars across the cell membrane when embedded in the lipid bilayer. The

time was finally ripe for the minimal genome to be enveloped in a lipid bilayer membrane, forming the earliest primitive cell – single celled life. This single celled life relied on a single set of genetic codons corresponding to a single set of amino acids for protein synthesis. The biological significance of the cell membrane was that it shielded the genetic machines and metabolism pathways from interference of other random peptides and chemicals in the pool. As a result, this single celled life quickly became dominant in the incubator through replication and division, and diverged into all variants through genomic mutations while maintaining the consistency of the basic genetic machine. Today all forms of life are proud of the descendants of this grand single celled ancestor.

As early life continued to develop, diverge, and evolve, it finally transformed into the real single celled life – mature and complex enough to be called species. These species bore some basic capacity to survive and prosper in the face of environmental changes and attacks from other species. In the meantime the genetic information flow continued to improve and preserve to present days.

Randomness and consistency are incompatible and paradoxical with each other when life was at its inception. From a chronological point of view on the origin of life, life starts in randomness and reaches consistency in maturity. It's the randomness that results in consistency, and it's the consistency that reduces randomness. Reduced randomness in turn lowers the chances to generate new functional proteins, which slows down the system to develop and advance. In other words, reduced randomness slows down evolution of life. Nevertheless, randomness is adverse to the evolution of life because it disrupts consistency, yet it has led to the endless possibilities for anything that would start up life. Therefore, randomness is the initial chemical basis for the most effective trial and error approach, albeit time consuming and wasteful, to establish life on the earth. It was those small number of biochemically active proteins among vast polypeptides randomly produced in the absence of DNA templates that initiated the process to form life. Life arises from the total randomness at the very beginning. It is unbelievable, but it is extraordinarily clever.

When we think of the origin of life, it's essential to think of the environment on the nascent earth in which life arises. An environment or system that is dedicated purely to RNA synthesis or protein synthesis could be created only in the laboratory. It was utterly unthinkable for the nascent earth to host an environment that was rich in chemicals only for RNA production or only for protein production. A system in which amino acids, bases and ribose coexisted was the most likely situation before life forming activities started to appear, considering the similar chemistry of amino acids and nucleotides. In a mixed system, nothing could prevent amino acids from linking into peptides when ribonucleotides polymerized into RNA. In terms of pure chemistry, peptides were the things to be produced more easily prior to RNA synthesis in such a system as discussed earlier. If peptides were produced in the system and they happened to contain enzymes for RNA synthesis, their activities couldn't be excluded in the process of RNA replication even though RNA with special secondary structures had catalytic activity and could self-replicate. If proteins, RNA, and DNA were present in their own worlds that were separated from each other in space, then when and how could these three independent worlds finally come together to form primitive life? The emergence of life happened spontaneously in a pure random fashion in an open environment on the young and turbulent earth. It's totally different from making dish with each ingredient available to you on the kitchen table, allowing you to add them to the cooking pan any time you liked.

4. More about Life Incubator

The premise for the development of early life is that all required chemical reactions occurred randomly and spontaneously on the nascent turbulent earth, albeit at low, even insignificant rates. A period of about half a billion years for the development manifested a process that is driven by chances, lucks, and

coincidences, all of which are characteristic of randomness-based processes to build up an enormously complicated, but fully consistent state. It attested the utmost difficulties to establish a single celled life from ground zero purely through random events. However, when we think of the origin of life, the most fascinating part isn't how protein, RNA and DNA are produced for the first time, but is the kind of environments on the nascent earth where protein, RNA, and DNA have been produced for the first time and continuously thereafter. Although life incubator seems to be a plausible idea to make a point that once upon a time on the nascent earth there was such a place where life originated. However, it was still hard to envision such a place that could have ever possibly existed on the earth to serve the role of life incubator. Was it as big as a small pond in a neighborhood, or as a large pond near a highway, or even as large as a lake full of nutrients for life? Where could it be located if such a place did exist?

Coal and oil are the largest carbon deposits on the earth today. Coal is a type of fossil fuel, originating from dead plant matter buried deep into the ground where they decayed under the heat and pressure without oxygen and slowly converted into coal over millions of years. Oil is a fossil fuel as well. It is derived from fossilized microorganisms. As a vast number of dead microorganism layers settled into the sea or lake bed, they were covered by mud and silt before they could decompose in the absence of oxygen. Like coal formation, dead microorganisms gradually were converted into oil under the heat and pressure without oxygen over millions of years. If something that happened millions of years ago can be explained with one theory, it usually can be explained with another theory equally well.

The coal deposits are discovered around the earth. They vary in size, type, deposit amount, depth under the ground, and geologic location. Some deposits contain enormous amounts of coal that seem to require amounts of plants that were way beyond the capacity that the areas could grow, even after multiple events that resulted in the burial of all the plants in the same areas that grew over millions of years. If coal is converted from buried plants, then why are some coals stone like, leaving hard remains after burning, and why are coal deposits present only in selected areas? Did it mean that only plants grown in certain geologic locations were buried and converted into coal, while the majority of dense forests weren't?

Like coal, the oil deposits are discovered around the earth. They vary in size, type, deposit amount, depth under the ground, and geologic location. If oil was derived from decaying organic matter plankton in the anaerobic conditions, first it must be formed before numerous different plankton eating organisms appeared in the ocean environment at least four hundreds of millions of years ago. Second, oil deposits should be found more or less in all areas that were once covered with water in the earth history. But it doesn't seem to be the case. Moreover, the majority of oil is contained in oil shale, any sedimentary rocks about 900 meters below the surface. It seems unlikely that all oil shale once was the ocean beds for long periods of time and had accumulated enough remains of microorganisms to form oil and penetrate into the rocks.

If coal and oil aren't formed as a result of geologic processes that acted on the remains of organic matter plankton or plants, then how did they form many years ago? If we think of the origin of life and the origin of coal and oil together, we can reach a conclusion that will explain both of them well. In the nascent earth, a large portion of earth's surface was covered by thick layers of the carbon rich liquid matter or carbon soup. The carbon contents of the soup are sufficient to account for all carbon deposits, including all living things, coal, oil, natural gas, and other forms of carbon containing materials, like minerals and carbon dioxide in the atmosphere. The soup contains all other kinds of elements, oxygen, nitrogen, sulfur, metal ions, and so on, all essential to chemical and biochemical reactions and thus life. The turbulent nature of the nascent earth and direct, unattenuated light from the sun raise the

temperature of the soup to certain levels that make many chemical and biochemical reactions possible and widespread. As protein, DNA and RNA polymerization reactions occur in certain locations on the earth, early life development processes start. Therefore, the concept of life incubator doesn't have a size or specific location, but anywhere on the earth, where there exist environmental conditions that allow life-forming polymerization reactions to occur. This implies that the randomness that has lead to the single celled life could be exponentially larger than the randomness that was possible in a limited life incubator. From the standpoint of a single set of amino acid codons, the early form life emerges from a single place in the incubator out of the entire area covered with carbon soup.

Because of plate tectonics, the earth had been constantly experiencing the split of old continental crust to form new continental crust. As a result over the period of hundreds of million, even billion years, tectonic forces have caused old continental crust to rearrange into new continents, thus causing the carbon soup to redistribute as well by sucking various portions of it into any spaces created during crust rearrangements. Tectonic forces also have generated tremendous amounts of heat and pressure that act on the carbon soup buried in the spaces in the absence of oxygen, forcing them to form all forms of oil, all forms of coal, all forms of natural gas, and all forms of other forms of carbon materials according to specific environmental conditions in which carbon soup is trapped in geologic locations.

5. Slow Evolution – Quiet 3.5 Billion Years before Cambrian Explosion

The development of early life from ground zero and the subsequent evolution reveal an important fact that when conditions are right, life building blocks, amino acids and nucleotides, are capable of self-assembling into interlinked and inter-dependent large molecules proteins and nucleic acids. These large molecules, together with lipids and carbohydrates, are capable of self-organizing into larger structures called organelles. The organelles are capable of self-organizing into a base form of life called cells. Cells are capable of self-organizing into the high form of existence called animals or plants, which are capable of self-evolving into more complex and more advanced forms such as birds and mammals over time. The entire process of this is known today as evolution. The fruition of evolution is that flowering plants, amphibians, reptiles, birds, and mammals have emerged at an astonishing speed, giving rise to the mind-boggling biodiversity on earth today.

Slow evolution after origin of life can be subdivided into prokaryotic era and eukaryotic era. Transition from prokaryotes to eukaryotes was completed in about 2 billion years, while preparing eukaryotes for fast evolution cost another 1.5 billion years. Both eras were painstakingly slow and lengthy, and the organisms evolved in this period were simple in every aspect. Nevertheless, slow evolution is the prelude to the fast evolution that started from Cambrian explosion and was responsible for a myriad of new advanced species that populated the earth.

The single celled organisms from Stage 1 were far from mature and robust. They are vulnerable and defenseless against natural elements and their genomes were too small to support evolution. In the prokaryotic era, evolution brought up major changes to the single celled organisms by increasing their genome size and protein coding gene counts. Randomness from sporadic point mutations was still one of the factors that drove prokaryotes to diversify and expand, resulting in numerous species. All the different species carried a certain number of genes that made them distinct from each other. All these species specific genes formed a huge gene pool from the standpoint of the prokaryotic kingdom.

In the dynamic prokaryotic era, life didn't always exist inside the cell walls, but was influenced greatly by some forms of extra-chromosomal genetic material that could co-exist inside the cell walls as an independent entity. Plasmids are extra-chromosomal genetic material that transmitted from one bacterium to another (even of another species) mostly through conjugation. Plasmids transferred

genetic material from one bacterium to another. Bacteriophages are another extra-chromosomal genetic material in the forms of viruses, which played roles similar to plasmids. Bacteriophages transfer genetic material from one bacterium to another through infecting various prokaryotic species. Extra genetic materials transferred from other species not only enriched the genes coding for more functional proteins, but also increased the genome sizes. By taking the rapid cell division rates and a period of 2 billion years into consideration, even at very low rates of point mutations, single celled organisms could have accumulated an amount of mutations in intergenic DNA sequences that were sufficient to transform some of the DNA sequences into functional genes. A larger genome tended to generate more novel genes through random point mutations. Whenever a random sequence became gene-like, the transcription machine would bind to it and make RNA and protein from it. Regardless, prokaryotic genomes were enormously heterogeneous among prokaryotic population.

Genotype differences led to differentiation of cells into various sizes and robustness. Small cells could be engulfed by large cells, and their DNA could be integrated into the host genomes. This is a random merger process called symbiogenesis. The merged cell populations were heterogeneous immensely, in which individual organisms differed greatly in genome sizes, gene sets, metabolism pathways, and more importantly in physiology. All of these qualities could entitle them to exist as distinct species. Therefore, it was randomness and time again that drove the evolution of single celled organisms to become species that were more adequate in function and better formed and developed in morphology. Above all, new species were more robust in defending against adverse factors.

When the number of proteins in single celled organisms was relatively small, it was much easier for the organisms to diverge into numerous different lineages through mutations and symbiogenesis. More rounds of divergence made lineages different more from their ancestors. It’s not surprising that there are a billion of distinct species of prokaryotes today, including bacteria and archaea. Archaea may have a better evolutionary relationship with eukaryotes as they share certain similarities in cell structure and function. Some genes and metabolic pathways found in eukaryotes are more closely related to those of archaea, especially the enzymes involved in transcription and translation.

When some of the merged prokaryotes boasted of large genome sizes, large gene counts, and rich proteins with a variety of biological functions, they became the predecessors of eukaryotes, in which biochemical processes and cellular structures started to compartmentalize into organelles. One of the organelles is the nucleus, which conferred a designated space for genomes to replicate and transcribe, and more importantly to elude interference from other cellular activities. The appearance of histone like proteins further transformed naked genomes into tightly packed chromosomes. Appearance of chromosomes and confinement of the chromosomes in the nucleus means that life has entered the eukaryotic age, the landmark in the history of life.

Eukaryotic cells are full-fledged organisms at this time, showing off a genome size over 5 millions of base pairs and a variety of metabolism pathways. Emergence of chloroplasts allows the organisms to capture and store unlimited energy from sunlight through photosynthesis, thus resolving food problems. Having mitochondria as an energy generator, the organisms are supplied with ample chemical energy to power a variety of biochemical processes and cellular activities. Meantime, safekeeping of the genetic machine in nucleus guarantees the high fidelity of DNA replication, further reducing the randomness in DNA replication and slowing down the pace of evolution. An implication is that evolution must find another mechanism to move forward. Indeed, the rest of the slow evolution period is to make ready for the rapid evolution.

It’s apparent that the appearance of eukaryotic cells didn’t mean that evolution entered the fast track. In all likelihood, the genomes of nascent eukaryotes were still small and the protein coding gene counts were far from adequate. As a result, the eukaryotes must continue to enlarge the genome sizes and increase the protein coding gene counts. Confinement of genome in the nucleus and acquisition of more enzymes useful for DNA manipulation allowed the organisms to have more freedom to bring about genetic changes to the genomes.

Genes in prokaryotic organisms are continuous without intragenic sequences, suggesting that DNA insertion is likely to be lethal and can’t serve as a general mechanism to increase the genome size. However, only very few genes in modern eukaryotic organisms aren’t disrupted by large amounts of intragenic DNA, called introns. This suggests that DNA insertions were random but common in nascent eukaryotes. Foreign DNA fragments could come from internalized cells or viruses through endocytosis. Meantime duplication of DNA fragments was another major means to increase the genome size, while random point mutations accumulated constantly on the chromosomes, generating a variety of possible genetic loci with potential biological significance. All the random genetic changes diversified the population into different species and prompted organisms to differentiate into different cell types, a prelude to the rise of multicellular life.

Slime molds are amoeba-like, typically single-celled organisms, and some of them can aggregate into loosely associated colonies. Such colonies are the infant form of multicellular organisms, and the cells in the colony were just about to differentiate into cell types. The genetic basis of cell differentiation is the differential expression of genes in different cell types. In other words, gene expression must be regulated stringently according to the roles of individual genes in cell types. Therefore, it’s imperative to establish rigorous regulatory mechanisms to control gene expression to maintain the cell types.

Immediate questions were that how to guarantee that particular genes were expressed only in cell types in which they were intended to express? Were the regulatory elements in the promoters and any other regions sufficient to confine the expression of particular genes into particular cell types? The answer seemed to be a no. Leak expression in the wrong cell types seemed to be common occurrences for all genes, which would ruin cell differentiation, thus ruin evolution of life.

The presence of introns in the genes requires genetic machine to remove them from the newly transcribed RNA molecules before exporting them out of nuclei, a process called RNA splicing. RNA from leak expression might not be able to survive the RNA splicing process due to insufficient amount, thus eliminating the possibility of protein synthesis in the wrong cell types. On the other hand, house keeping genes are not specific to cell types, but common to all cell types. Splicing their RNA is a waste of resources, and many of these genes are indeed intron-free. Adding non-coding DNA sequences inside genes increased the genome size considerably, and as a result, the cells would consume more energy and material to operate and maintain large genomes. Therefore, introduction of introns isn’t a cost effective, but is a viable way to guarantee the integrity of cell types. On the other hand, the purpose of introns didn’t seem to make sure that gene expression was leak-proof, but it just happened randomly and was preserved in all eukaryotes. A likely possibility is that cells started to differentiate just because of the appearance of introns and splicing. What other strategy could replace introns to serve the same purpose, if not better?

Evolution of eukaryotes from the moment of their appearance to the moment before Cambrian explosion spanned a period of staggering 1.5 billion years, roughly two third of which was dedicated to organisms of multicellular nature. Prokaryotic era and eukaryotic era share a common, but significant and indicative, characteristic. Both era endured a period of about 1.5 billion years to conclude the

evolution triumph. Limited by the single cell nature, prokaryotic organisms didn't change much in morphology, indicating that all changes in the period were confined to their genomes. On the other hand, the multi-cellular nature permits the morphology of eukaryotic organisms to vary infinitely. Nevertheless, comparing with the organisms that emerged in the post Cambrian era, pre-Cambrian organisms just gained quite limited changes in phenotype that seemed too meager to worth 1.5 billion years of evolution. All this indicates that all changes in the period were confined to their genomes as well. This was a mystery until we can divulge into it with the availability of huge amounts of genome sequence data. The whopping 1.5 billion years for each of these two eras reveal the unthinkable difficulties for the genetic system to create novel proteins and then assimilate them into the existing biochemical processes and cellular structures to become the integral part of life. It also shed light on how the evolution itself is evolved over the entire course.

So far hundreds of genomes from species covering almost all levels of evolution have been sequenced and annotated, and data are available from several research institutions for public research. Table 1 lists minimum genome information, including genome size and the number of protein coding genes, from several selected species ranging from archaea to bacteria to organisms emerged from Cambrian explosion. What we can draw from the data in the table will help us understand evolution and find out what is the bottleneck of evolution.

Organism	Type	Genome Size	Coding Genes
Pyrobaculum aerophilum	Archaea	~2,200,000	2,562
Haloferax volcanii	Archaea	~4,000,000	3,895
Chlamydophila pneumoniae	Bacteria	1,225,935	1,113
Chlorobium chlorochromatii	Bacteria	2,572,079	2,002
Gloeobacter violaceus	Bacteria	4,659,019	4,430
Bacteroides fragilis	Bacteria	5,205,140	4,260
Micromonas pusilla	Algae	~20,900,000	10,056
Guillardia theta	Algae	~87,200,000	24,840
Dunaliella salina	Algae	~343,700,000	16,697
Breviolum minutum	Algae	~1,500,000,000	47,014
Tieghemostelium lacteum	slime molds	~23,400,000	10,225
Hortaea werneckii	Yeast	~49,900,000	15,974
Aspergillus nidulans	Fungi	29,828,291	10,534
Armillaria ostoyae	Fungi	60,106,801	22,299
Blumeria graminis	Fungi	124,489,486	7,118
Puccinia graminis	Fungi	88,724,376	15,800
Crambe crambe	sponges	~370.500,000	
Chrysaora quinquecirrha	jellyfish	~337,400,000	
Bolinopsis microptera	comb jellies	~265,400,000	14,118
Lytechinus pictus	urchin	~811,700,000	20,054
Cydia splendana	Moths	~630,600,000	14,646
Penaeus indicus	Shrimps	~1,900,000,000	21,824

Table 1. Genome sizes of various species on different levels of evolution. The cells that display the number of protein-coding genes are left blank if data are not available. Data are taken from NCBI, Ensembl Bacteria, and Ensembl Fungi.

The number of protein coding genes for each organism listed in the table is obtained using genome analysis software, so it doesn't necessarily mirror their true expression in the organisms. However it demonstrates that these genetic loci exhibit gene structures and can be considered as genes. It implicates that at least they can serve as genetic materials for new genes or gene duplication.

Genomes of prokaryotes are much smaller and contains much fewer protein coding genes comparing with the genomes of eukaryotes. Furthermore, the sizes and gene counts varies greatly from species to species. On average each prokaryotic gene takes up about 1000 bps and encodes a protein of about 250 amino acids. This shows that prokaryotic genomes contain sparse intergenic DNA sequences. What could be inferred from this is that the genomes of the earliest forms of life must be far smaller and encode far fewer proteins than present day prokaryotes.

Genomes of single celled eukaryotes can vary in size and gene counts even more greatly from species to species. They are usually 5 to 20 times larger than the genomes of prokaryotes, but protein coding gene counts are only 2 to 10 times larger. On average each eukaryotic gene takes up about 2000 bps. This clear disproportionality shows that eukaryotic genomes contain a large amount of intergenic and intragenic DNA sequences. What could be inferred from this is that the genome sizes and gene counts of the earliest forms of eukaryotes must be close to those of prokaryotes.

Genomes of multicellular eukaryotes vary in size and gene counts greatly from species to species as well. A general trend is that the genome sizes increase dramatically, but gene counts were relatively steady, as the organisms move up the evolutionary ladder. The average base pairs per gene are about 3000 bps for fungi, but dramatically increased to about 30,000 bps in pre-Cambrian organisms sponges, jellyfish and comb jellies, and to 45,000 bps in post-Cambrian organisms urchin and moths. This indicates that genome size increase is largely due to the increase in intergenic and intragenic sequences, not in protein coding sequences. Occurrence of DNA duplication could be frequent, and duplicated DNA sequences then diverged upon long time accumulation of point mutations. Protein coding gene counts usually fluctuate around 15,000 to 22,000 regardless of genome sizes and the position of the species on the evolutionary ladder. This indicates that there is a ceiling for protein-coding gene counts and this ceiling has been hit in some of the low species like sea urchin, fungi, and shrimps. It is a shock that in multicellular organisms protein coding gene counts are not well correlated to the complexity of the organisms. The gene count ceiling has a profound implication about how evolution of eukaryotes has occurred.

The earliest single celled organisms contained a minimum set of genes that could satisfy the working of the most basic living system called life. However, evolution from simple to advanced is the inherent quality of life from the very beginning. The initial enzymes involved in DNA manipulation have strong tendencies to expand the genetic repertoire of the genome via introducing random mutations as well as duplication and recombination. Organisms of all three domains of life today rely on similar metabolism networks to sustain life. However, it wouldn't be possible for the very early single celled life to enjoy modern metabolic systems, but all modern metabolic pathways must initially evolve and mature in prokaryotic organisms over 2 billion years before passing them down to eukaryotic organisms at the time they emerged from prokaryotic organisms.

As described earlier, life arises from total randomness at the very beginning and evolves on random genetic changes thereafter. If we reckoned with the difficulties in turning a random DNA locus into a functional gene through random mutations, we could envision hundreds of millions of random events converged on the locus base by base over an inestimably long period of time. This random DNA locus would be converted first into a semi-gene coding for a random polypeptide of good length and then

refined into a gene coding for a protein that could perform catalysis on some substrates. Finally the protein would undergo further mutations to achieve the catalytic activity useful to the organisms. In this process early good mutations could be canceled out by later ones, slowing down the progress further. Therefore, evolution of a gene of biological significance must be an endless and repetitive trial and error process, bearing fruit possibly after hundreds of millions of year random trials. If a group of protein molecules are required to perform a compound function, activity, structure, or pathway, the task of evolution of the group as a unity would be achieved after a far larger number of random events in a longer period. It is expected that it would be more difficult to generate functional protein molecules of multisubunit nature.

Glycolysis is the metabolic pathway that converts glucose into pyruvate and at the same time produces ATP and NADH(reduced nicotinamide adenine dinucleotide). It is a sequence of 10 reactions catalyzed by 10 enzymes and used by most organisms today to generate a limited amount of energy. From the evolution standpoint, the glycolysis pathway was likely to arise from the merge of multiple early prokaryotic species, each of which contributed a few enzymes to form the final pathway we see today. Obviously this would accelerate considerably the evolution of many biochemical processes and cellular structures in prokaryotes. However, the difficulties of generating any functional proteins from random DNA loci were not changed.

Consider a hypothesized situation. A group of distinguished experts in protein sequence and function was asked to design, implement, and test a drug production project. A small organic molecule D was identified as an effective drug against heart disease. In theory molecule D could be synthesized by a sequence of 5 reactions called metabolic pathway D similar to glycolysis. The effort was focused on creating 5 enzymes to catalyze 5 reactions sequentially. Experts could employ the latest biotechnology, protein databases, techniques, equipment, and the fastest supercomputers available to the cutting edge research to design enzymes to serve the project. A common approach would be to explore the vast protein databases to identify any enzymes with potential for modifications to obtain desired activities. Modifications wouldn’t be random, but carefully designed according to what we have learned about the relationship between protein sequences, 3-D structures and functions. Could this group of experts achieve their project goal by employing such learned and well equipped approaches?

It’s a mystery when glycolysis pathway appeared in the prokaryotic era, but it must not be part of the nascent single celled organisms. This means that most of the ten enzymes of the pathway rose and evolved slowly de novo from random DNA sequences before eukaryotic era began. On average, bacterial genomes carry about 4000 genes. It must be true that most of these genes rose and evolved slowly over the 2 billion year period. In addition, most of genes don’t act alone, but perform particular functions in sequential or interdependent fashion. Assuming about 2000 new genes added to the genomes of nascent cells, a period of whopping 2 billion years could be well accounted for, considering that so many additional new genes were to be created de novo, organized into functional groups, and incorporated into the existing system to act altogether to make life processes more complete, robust, and efficient. To this end, the potentials had been exhausted for the prokaryotic organisms to evolve further into more complex and advanced forms. The evolution of prokaryotic organisms reached the limit with its cellular structures and genome organization.

On top of the average gene count of 4000 in prokaryotic cells, eukaryotic organisms added about 5 times more new genes to their genomes to reach the neighborhood of 20,000. In general, prokaryotes and eukaryotes share many proteins only in functions, not in amino acid sequences. For example, enzyme hexokinase catalyzes the first reaction in glycolysis in all forms of life. There exists clear evidence of sequence homology between yeast, plants and vertebrates, but no relation in sequence

observed between prokaryotes and eukaryotes. Function-only homology between protein counterparts from prokaryotes and eukaryotes gave evidence of genomic upheaval in the evolution of eukaryotes. Without it, prokaryotes would be the only form of life on the earth.

If the complexity of post Cambrian species warrants gene counts around 20,000, it wasn't expected that the ceiling has been hit as early as in preCambrian organisms. After the departure from prokaryotes, the genome of eukaryotes must be closer to prokaryotes than to modern single celled eukaryotic species in terms of gene counts, size, and sequence structures, and the genetic apparatus must be simple and basic regarding genetic operations involving large pieces of DNA, especially DNA duplication. However, almost 16,000 new genes rose and evolved during 1.5 billion years in the eukaryotic era. This could be considered at a lightning speed comparing with around 2,000 genes in 2 billion years. In addition, there were only slim chances for multiple mutations to fall on the target DNA loci in one replication cycle, and most genes were members of functional groups and were unlikely to be created at the same time. If one functional component hadn't be attended by other components in the sequence for some time, they were likely to disappear after more mutations rendered them nonfunctional. It could be concluded that a small portion of the 16,000 genes rose and evolved de novo from random DNA sequences first, and then they became the targets for gene duplication to achieve a rapid increase in gene counts. Therefore, DNA duplications were likely already at work in the later period of the eukaryotic era. Evolution of a particular gene is not only dictated by mutations on its own locus, but also constrained by other genes as well, if it is part of a functional group. Low cell differentiation of pre-Cambrian organisms didn't warrant such high gene counts, but it suggested that not all genes were likely expressed to serve extant cellular activities and biochemical functions, but they laid the foundation for fast evolution.

In the prokaryotic era, random mutational rates seemed to have diminished due to the maturity of the DNA replication apparatus, thus slowing down the evolution of genes to a great extent. In eukaryotes the genomes were protected even better in the nuclei, which increased the fidelity of DNA replication and made genomes remarkably stable, thus hindering the speed of evolution to a new high. Time and rapid replication cycles had become the only factor in augmentation of gene repertoires of prokaryotes as well as pre-Cambrian eukaryotes. It could be well understood that establishing new biochemical and cellular functions from random DNA loci would be possible only when infinite chances, lucks, and coincidences all had occurred and converged fortuitously together through the most painstaking and enduring processes of trial-and-error. Even with time and rapid replication cycles, the possibilities to create new standalone enzymes, not to mention complex pathways like hypothesized pathway D, are extremely low. The time span of 3.5 billion years for the slow evolution are the strong indication of the enormous complexity, difficulty, frustration, and uncertainty of random mutation based evolution of living organisms.

6. Protein Variants and Evolution

At the end of slow evolution stage, species are still low and not sophisticated at all from any stand point of view, but the average protein coding gene counts are unexpectedly large, roughly on the par with higher animals, including mammals, albeit undersized genomes. Meanwhile the genetic apparatus became better developed and more powerful. An implication is that evolution has entered a new mode to move species forward. Gene variations and gene duplication are the mechanisms behind deviation of numerous biochemical and cellular processes that make new species more complex and advanced.

Assume that a new small chemical named X could regulate body temperature in the extremely cold environment. To generate a receptor for X, the X receptor, an existing receptor gene for a different small molecule Y happened to become duplicated. To turn Y-specific receptor to X-specific receptor was still a long evolutionary journey. We could imagine that many changes must be made to the Y

receptor so that it could be transformed into an X receptor. First the changes must enable the receptor to bind X by adjusting the overall three-dimensional structure to create an internal space that can specifically accommodate X. Second the X receptor, upon binding the X, must be able to undergo conformational changes into an active state. Third the active state of the X receptor is another three-dimensional structure that can interact with a downstream component involved in regulation of body temperature or act as an enzyme by itself. Fourth the X receptor gene must be subjected to regulatory control so that this receptor will be expressed only in selected tissues. After a large number of point mutations occurred and accumulated in the duplicated DNA locus over numerous generations, a Y-receptor based functional X receptor emerged. Nevertheless, such a success couldn't be guaranteed.

In the above hypothesized scenario, the gene for X receptor was duplicated from an existing gene. By taking advantage of the duplicated gene as a fully structured DNA sequence, no additional changes would be necessary to build gene structural features into a random DNA sequence before making it fit for X functionally. Though it remained to be an extremely lengthy trial and error process, undoubtedly it must be much faster than de novo creation of a protein receptor for the molecule X.

An interesting question is that if no functional receptor for X could be produced for organisms living in the extremely cold area, the organisms would die from the cold? It must be unlikely. If this small chemical was present in a warm climate as well, would the organisms there develop a receptor for X where organisms don't need to respond to cold temperature? It must be possible. A need for X would not trigger or accelerate the development of a functional receptor for X, while lack of need wouldn't necessarily prevent the development of a functional receptor for X either. Most biological functions and structures didn't emerge on necessity or usefulness, but rather they emerged randomly as consequences of random mutations. They would be preserved if they happened to be able to enhance or complement some processes or structures for better functionalities, or if they could function as standalone factors to increase the well-being and survivability of the organisms.

Genome sequencing of large magnitude indicates that species in the same genus, even in the same family, share extremely high percentage of identical sequences. The implication of this is that large morphological differences don't mean similar differences in genotypes. In fact as species move up the evolutionary ladder, the differences between genotypes diminish greatly. For species that are classified into the same family, differences in morphology, biochemistry and cell structures generally are attributed not to those novel proteins, but to variants or isoforms of the same proteins that are expressed in different species. The advent of protein variants in evolution is a giant step forward towards more complex and advanced species on a number of superior benefits since Cambrian explosion.

C. elegans is a free-living transparent nematode or worm, belonging to a type of metazoan organism with 959 cells. C. elegans genome is relatively small, consisting of 100,286,401 bps, and contains an estimated 19,985 protein-coding genes. 83% of proteins expressed in the worm were found to have human homologous genes. Only 11% or less genes are nematode specific. Some proteins can be exchanged between C. elegans and humans or mammals. This means that most genes present in much more complex organisms like mammals are already available in animals as low as C. elegans. An implication is that development of higher organisms doesn't depend on creation of a large number of animal specific genes, but primarily on re-utilization and re-localization of genes that have existed in lower organisms, a strategy of derivation and reuse.

All protein molecules must assume unique three dimensional structures to assume their biochemical functions, while the three dimensional structures are mostly determined by the amino acid sequences. Proteins will exhibit properties that are altered more or less in detail while playing the same roles in

biological processes after certain amino acids are substituted. They are the variants of the same protein molecule. The protein variants can have subtle impact on biological processes that rely on three dimensional structures. For example, a neurotransmitter receptor variant could have its affinity for the ligand increased or decreased relative to the original receptor, thus changing the behavior of organisms. Signal transduction pathway could be affected because of changed physical interaction between each protein component. The impact of structural variants is often visible morphologically. An organism could assume a different morphology if tissue orientation in normal development was skewed a certain degree due to a structural variation in the process. Limb could grow out of wrong place or into wrong forms if variants are engaged in embryonic development,

During the evolution process, the same protein could diverge into multiple variants due to randomness of the point mutations after numerous generations. This kind of divergence enables offspring to assume looks that are similar but not identical in size, morphology and behaviors, all of which are sufficient to qualify them as new species. Protein variants that are randomly produced in the evolution are one fact that has dramatically shortened the time for new species to appear.

Protein variants in the evolution of species reveal an important biochemical property of proteins. The sequence dependent three dimensional structure is not rigid, but shows great elasticity in the cells. In other words, the three dimensional structure of a protein is elastic enough to withstand certain sequence changes and remain compatible with the existing biochemical and cellular processes, allowing its variants to perform routine functions as if nothing has happened.

Gene duplication is the easiest and quickest way to produce protein variants for new species. For easy discussion, the ancestor gene from which all subsequent genes coding for the variants are derived is called master gene, and derived genes are child genes as a result of gene duplication. The child genes must be subjected to changes in DNA sequences so that they encode variants of the master protein. The differences in amino acid sequences will widen as species diverge further apart.

Opsins are protein molecules responsible for wavelength sensitivity when coupled with light sensitive chromophore 11-cis retinal. S-opsin absorbs short wavelength light, while M-opsin absorbs middle wavelength light, and L-opsin absorbs long wavelength light. Primate retina houses three types of photoreceptors, each of which contains S-opsin, or M-opsin, or L-opsin. By contrast, most mammals lack L-opsin-containing photoreceptor and are not sensitive to long wavelength light. When primates split from most mammals, M-opsin gene underwent a duplication event that led to an extra copy of the gene. Then random point mutations turned this extra gene into a functional gene coding for L-opsin. All this is evidenced by the fact that M-opsin and L-opsin are identical except 15 amino acids out of 364 total. This small difference in amino acid sequence confers L-opsin sensitivity to the long wavelength light. L-opsin is a variant of M-opsin.

Generally speaking, a variant assumes the same biochemical roles of its master protein, but with subtle changes in biochemical properties. It’s these subtle changes that empowers a variant to fit new occasion or fill the void that the master proteins can’t fill, or complement the action of master proteins in the new species. L-opsin is the brilliant example in this regard. Muscarinic acetylcholine receptor has about five subtypes in humans, all of which are variants of each other. These receptor variants show unique gene expression patterns, unique sensitivities to ligand acetylcholine and various drugs, and more importantly they elicit unique neurological effects in different target cells.

Most important impact of protein variants on evolution is not short-term, but long-term. Multicellular life emerged after cells started to differentiate into types when the genome size and protein coding gene

count increased to certain levels. Cell differentiation became a major change in the later stage of slow evolution, as multicellular life began to take a more defined and more sophisticated morphology like soft-bodied metazoans, some of which displayed a trace of skeletal elements. The appearance of more complex life forms must be supported by new protein factors, for example, new protein factors that control the development of a morphology, as well as proteins that carry out the underlying biochemical processes. The genes that encoded these early proteins would serve as master genes, from which protein variants would be derived to impact the evolution of species in a fundamental way, enabling new species to develop and assume more complex morphology with extensive differentiation of cell and tissue types.

Fin development begins in the morphogenetic fin field in the fish embryo. Some fin inducing factors act on mesenchymal cells in that field and cause the outer germ layer to proliferate and bulge out, forming a fin bud. A growth factor then guides further development of the fin bud into a fin. The fin inducing factors control the exact location and direction the fin bud bulges out in the morphogenetic fin field, which determines the final morphology and location of the mature fin on the body. Assumed that one genetic locus was duplicated from the gene encoding one of the fin inducing factors. Over time under random mutations, the sequence of this locus deviated gradually from its master gene and encoded a variant. This variant assumed a three dimensional structure that was slightly different from that of the master protein. Because of this slight difference it induced the fin bud to bulge out at slight different location and towards a slight different direction. The overall impact on the fin development was that the final fin was quite different from the fin seen on the ancestor organisms morphologically and in location. If a factor variant assumed a three dimensional structure that was too skewed to induce the normal fin development, the final fin could be in a deformed state. If a factor variant was produced in the wrong part of the embryo, it could induce the growth of a fin at a wrong location.

The consequence of divergence of a master gene into a variant family is remarkable. The rise of myriad forms of phenotypes along the evolutionary timeline can be attributed to these variants as critical influencing factors. Taking fin inducing factor as hypothesized example, the early variants descending from the master gene gave rise to various forms of fin morphology on different fishes. Further divergence in gene sequences gave rise to many more distant variants, each of which controls the development of limbs unique to amphibians, reptiles, birds, mammals, primates, and finally humans. The degree of divergence from the master gene mirrors well the positions of organisms on the evolutionary ladder. This is a vertical view of the roles the protein variants have played in the course of evolution. From a horizontal view, what we see is so many distinct fins on different fish species, so many distinct limbs on different species from amphibians, reptiles, birds, mammals, primates, finally one unique limb on humans. Through divergence into a large variant family, the overall impact of the fin inducing factor on the fin and later on the limb development has been amplified to the utmost degree in the past 500 million years.

Evolution of keratin genes occur in parallel with the evolution of organisms. Keratin consists of a large family of structural fibrous proteins called intermediate filaments. The master gene of keratin can be dated back to as early as in sea squirts before Cambrian explosion. Today numerous variants of keratin exist in almost all species, including vertebrates and invertebrates. They form the hair, outer layer of skin, horns, nails, claws, scales, shells, feathers, beaks and hooves for sea squirts, fishes, reptiles, birds, and mammals. It’s the sweeping divergence of keratin master gene in the past 600 million years that has made a large variety of tough structures of the same type possible. And it’s these variety forms of tough structures that confer animals to bear one of the structures with distinct capabilities to inhabit suitable environments. A particular keratin variant can function in many species, and a particular species can have many keratin variants to fulfill different functional requirements. In humans 54

functional keratin genes are located in two clusters on chromosomes 12 and 17 and expressed differentially in different types of cells and tissues to serve different roles. It's well established that the amino acid sequences of each keratin variant have been preserved over evolution in different organisms because each of them forms unique three dimensional structures that are particularly suitable to build beaks, or feathers, or hair, or nails, etc.

The most illustrious example of gene evolution via gene duplication goes to the largest superfamily of genes coding for a special group of proteins called G protein-coupled receptors (GPCRs), also known as seven-transmembrane domain receptors. GPCRs are cell surface membrane receptors that transduce myriad extracellular signals into the cells to regulate a variety of biochemical and cellular processes. Extracellular signals include but not limited to such as photons, lipids, hormones, peptides, proteins, odorants, neurotransmitters, and ions. The opsin molecules described earlier are GPCRs. The wide spectrum of ligand types and biochemical and cellular processes that they regulate are strong evidences that their vital roles in eukaryotic organisms are critical for the evolution of species.

Search of GPCRs in comprehensive protein sequence databases reveals a long history of evolution. The GPCRs superfamily can be dated back to the time of the multicellular origin. The receptor for cAMP and receptor for neurotransmitter glutamate have been shown to be the early GPCRs in Amoeba-like protozoa, which can be traced to the early time in eukaryotic evolution more than 1.4 billions of years ago. Main mammalian families of GPCRs are present in fungi, illustrating a long evolutionary link too. The size of the human GPCR superfamily is determined to be at least 800 different genes, accounting for about 4% of the entire protein-coding sequence. Classification of the GPCR superfamily was complicated and varies among researchers. The three main classes (A, B, and C) don't share detectable sequence homology, suggesting early divergence of the master gene along the evolutionary timeline.

GPCRs are involved almost in most, if not all, biochemical and cellular processes in humans, including senses, behavior and mood, immune system, nerve, homeostasis, growth, endocrine system, etc. Because of extreme versatility in structure to specifically bind many types of ligands outside of the cell membranes, and to specifically transduce signals to different types of protein factors inside of the cells for signal transduction, GPCR genes indisputably have become the easiest targets to derive variants that bear slightly different biological functions, which differentiate not only closely related species, but also various cell and tissue types in the same species. Without protein variants, it would be impossible to produce novel protein factors just to perform slightly different functions. Therefore, protein variants are the molecular basis for the evolution of closely related species like gorilla and chimpanzee, and for the staggering biodiversity on the earth today.

Assume there is a protein variant X1 that deviates from the master protein X via gene duplication and random changes in amino acid sequences. After some generations, protein variant X2 deviates from X1 in the same manner. Eventually a large family of variants X1 to Xn is established after divergence of numerous generations across different classes of animals. A particular variant may be expressed only in a particular species or in a wide range of species of different classes. This is a family of variants from evolution standpoint, but it is not necessarily the same family of variants from structure and function standpoint. In this large evolutionary family, some of the members have gradually become all-new protein molecules on their own in amino acid sequences, structures, and biological functions, and have lost the core functions of its master gene. Evolution is driven by random mutations, while random mutations place no constraints on duplicate genes. Therefore, every duplicated gene has the freedom to diverge, and the resultant variants, regardless of the extent of differences, will be preserved if they are good fits and not lethal to the organisms over time. When they lost qualifications to be the variants of other members per se, they couldn't be easily traced back to the master protein X. Evolution of all-new

proteins in this way is obviously far faster and economical than from random DNA loci. Derivation of new functional genes through deviation from gene variants has been greatly accelerated as the genetic machine develops more means to manipulate genes other than random point mutations. If this is what has happened in evolution, establishing DNA based evolution trees can be a daunting challenge.

Gene duplication based gene variations account only for a small portion of gene variations. More strictly speaking, most genes in higher organisms are variations of the non-duplicated genes in lower organisms, making all novel genes just an even smaller portion of the gene counts. The large gene counts of multicellular species in the pre-Cambrian period serve as an indicator that gene variations have started to become part of the key mechanism of evolution, and their roles and significance have manifested in the Cambrian explosion and all the evolutionary events thereafter.

7. Fast Evolution – New Species Arise in Explosive Mode and Evolution Cycle

It seems bewildering why all over a sudden evolution accelerated, bringing millions of complex and advanced new species into existence in a short period of time. In the fast evolution stage, a burst of new species in general accompanied certain dramatic climate and geological changes. For example the Cambrian world differed greatly from the preceding Proterozoic Eon in terms of geography and climate. During transition of the two periods, the earth experienced a gradual global warming, rising oxygen levels, and split of a single continent into two. Climate and geological changes could make mutations occur more frequently in all species. When mutations struck DNA polymerases, DNA polymerases replicated DNA at lower fidelity, causing organisms to suffer from accelerated genetic changes. Direct consequences of this are two folds, mass extinction of old species and proliferation of new species.

Prior to Cambrian explosion, most of the living organisms were classified under kingdom protists. They were small, unicellular or simple multicellular, including slime molds and fungi. Then a little more complex, multicellular organisms like sponges, jellyfish, sea anemones, corals gradually emerged in the later time of the slow evolution. The Cambrian period (about 539 to 485 million years ago) was particularly special in the evolutionary timeline, because it marks the start of fast evolution stage. Living organisms exploded into millions of forms and complexities in a period lasting only about 45 million years, commonly referred to as Cambrian explosion. Insects, flies, spiders, centipedes, ticks, mites, snails, scorpions, shrimps, shells, starfish, brittle stars, sea urchins, sand dollars and sea cucumbers all appeared in this period. First plants and fishes rose at the later part of the period. From evolution point of view, all these species remain very low on the evolutionary ladder despite of stunning varieties in complexities and forms.

In the following 540 million years, evolution greatly sped up, and emergence of new species of higher complexities occurred more frequently. In a nutshell, evolution of species is the evolution of genomes. The genomes became more advanced after each evolutionary event, laying down the foundation for more complex and advanced species to appear ahead.

As discussed earlier, when organisms evolve to higher levels, they rely more on protein variants to build unique morphology, cellular structures and biochemical processes. Protein sequence comparisons tell a lot about those proteins that play similar biological roles in different species. A large number of them can be classified into groups based on the similarities of overall amino acid sequences or sharing of certain short amino acid sequences called motifs. Proteins that shared amino acid or gene sequence homology are essentially variants or isoforms of each other. By taking advantage of protein variants and motifs, creation of proteins with desired functions and properties became easier, and at the same time it alleviated the challenge to assimilate new components into existing cellular structures and

biochemical processes, avoiding a formidable array of problems incurred at creation and integration of new proteins even when organisms were relatively low and simple.

Evolution is a process of constant changing, but only changes that go beyond the threshold of evolution will be able to bring about meaningful consequences to the organisms. All modern day organisms, archaea, bacteria, animals and plants, don't seem to be in the path of evolution. The planet earth is full of living organisms as simple as single celled life and as complex as mammals. If organisms were in a constant state of evolution over the past millions or billions of years, we wouldn't be able to see organisms as low as archaea, bacteria, algae, fungi, jelly fish, sea urchins, etc. This indicates that not all lower organisms have the quality to be the ancestors of higher organisms. Most species stay where they have been since they appeared long time ago.

The concept "ancestor" must be right because it agrees with the evolution of living things. Then what organisms can be ancestors from which higher forms of life arise? Formation of new species isn't a simple event or the sum of multiple simple events, but involves numerous changes in the genotype that generates a phenotype that is sufficiently different from the phenotypes of the old species. This level of genetic changes won't be possible in normal organisms even though they are under constant attacks of point mutations, suggesting that an ancestor organism must be special on its own. The genome of ancestor organisms should be more prone to environmental changes and quite elastic and blessed with a genetic machinery that could perform genetic changes of large magnitude – changes that would generate new genes or gene variants at much faster rates in shorter time frame, thus establishing new phenotypes that define organisms as new species. In doing so, ancestor organisms must be able to accommodate new components created and added to the system at different time points in the cycle before all the new components were created to complete the new phenotype.

What exact events could trigger genome changes of large magnitude is a forever mystery just as how life exactly started, unless humans could be fortunate enough to experience a new round of mass extinction and mass proliferation and we were not part of the mass extinction. However, it's still worth to think about it and envision something fictitious to get some idea.

In usual time, genomes of all organisms, including ancestor organisms, were in a disarmed state, in which the genome is consistent and stable except low random point mutations and normal DNA recombinational events from meiosis. When sudden geological and climate changes broke out, ancestor organisms in a population suffered from more mutations. When mutations lessened the replication fidelity of DNA polymerases, genome wide accumulation of point mutations accelerated. These genetic changes acted as a perturbation that drove the genome from its disarmed state to an inconsistent and unstable state, an armed state. In the armed state the genetic machinery of an individual ancestor was activated to perform genetic changes that would reshape the genome to start the evolutionary event, a phase called genotype reshape. As the reshape process continued to reshape the genome by bringing further changes, it gradually diminished in magnitude into a process called genotype healing. During the healing process, the genome gradually re-established a consistent and stable state, a new disarmed state. After genomes of all the individuals descending from the common ancestors underwent changes from the disarmed state to the armed state and then back to a new disarmed state, one evolution cycle completed (Figure 3). All the individuals that carried mutations in the cycle, including those dead at the embryonic stage, were intermediates of the evolution. Processes reshape and healing are quite blurry and continuous, but are two different concepts useful to reveal what is happening in an evolution cycle.

Assume there was a population of a single ancestor species. When sudden geological and climate changes struck, higher frequency of random mutations perturbed the normal biochemical machineries

and cellular structures to a greater extent, which acted as a bumpy start to bring the genomes of the population into armed states and start the reshape process. An evolution cycle began as such. The perturbation lead to the death of most of the early intermediates, but it would be necessary for an evolution cycle to proceed. The genotypes of individual intermediates differed soon after a few generations. As the cycle went on and the differences grew, the heterogeneity of the genotypes in the population widened. Meanwhile, process reshape was slowly reduced into process healing after it had given rise to numerous changes to the intermediates. When new components, including protein variants and novel proteins if any, gradually reached an equilibrium with the existing biochemical machine, process healing ended, and survived intermediates re-entered new disarmed states. In this state the enduring instability and incompatibility among components had been smoothed out, enabling all the cellular and biochemical processes to restore to balanced states again after they had been agitated by genome wide changes. The difficulty to achieve such balanced or disarmed states would be unthinkable if numerous protein variants were not employed in the cycle, especially in species as advanced as fishes, amphibians, etc. Regaining a disarmed state permitted life to return to work in harmony and stability at higher levels. What's visible in the cycle was the morphological transformation of each intermediate after their genotypes were changed. When the cycle ended at the end of healing process, the genotypes of the survived intermediates were so different from their ancestor and from one another as well that they were no longer the same species, but distinct new species. The happy ending of an evolution cycle was the emergence of new species in new disarmed states.

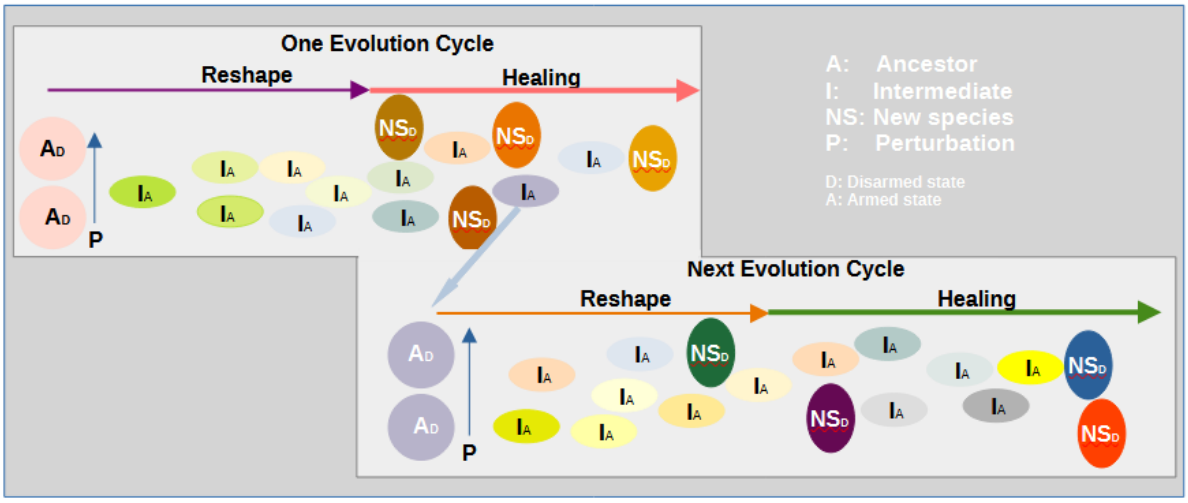

Figure 3. Evolution cycle from ancestor organisms in disarmed states to new species in new disarmed states, including processes reshape and healing, and numerous intermediates in armed states.

The nature of a change is determined in lieu of the survivability of intermediates. All lethal changes were eliminated from the population after causing carriers to die, leaving no impact in the evolution cycle. It was expected that the overwhelming majority of intermediates perished early and quietly upon failure to survive seemingly endless mutations. They are the dead ends of the cycle. Only heritable changes were passed down and caused variations of phenotypes among the next generations. The concepts of beneficial and deleterious mutations were not applicable to evolution cycle as the effects of any non-lethal mutations introduced in a generation must be decided from the whole cycle standpoint. Generally, the fates of intermediates in armed states were unpredictable. It's all up to their luck.

The genome evolution cycle is unique because one single cycle can take up to tens of million years or generations to complete, and its cycle path is composed of two disarmed states, a single armed state and two processes, reshape and healing. From genetic mutation standpoint, all changes that occur in the

path are random and irregular. Randomness and irregularity are the key to the fascinating biodiversity that arises after each evolution cycle. All genetic changes that occur in a cycle are deemed as genome-wide, but the changes to one generation of any intermediates must be on a granular level so that some of the intermediates will survive every change. Not all evolution cycles would lead to new species if no intermediates survived.

The random nature of genetic changes could result in a number of first generation intermediates, depending on the size of ancestor population. From the moment of birth, an intermediate would move along its own path and produce its own next generation intermediates in a manner independent of other intermediates. It wasn't possible to predict how many more generations were needed for a random intermediate to reach the final disarmed state if it was a lucky one. New species would resemble each other more strongly if they descended from the same intermediate fewer generations apart and differ with each other more strongly if they were more generations apart. Figure 4 illustrates an imaginary mini evolution cycle in its entirety starting from an ancestor population. Individual ancestors (orange solid circles) sit in the center and were surrounded by light pink sold circles that represent numerous intermediates, whose distance to the center represents the number of generations down from the ancestor. Intermediates that are dead ends are represented by outermost solid black circles. Some lucky intermediates that end up in disarmed states – the new species arising from the cycle – are indicated by outermost solid red circles. When a line with arrows is used to connect the center to one of the outermost circles through a series of intermediates in between, we could see the complete evolution trails. An evolution trail starts from an individual ancestor and passes through every intermediate that leads to the next intermediate, finally reaches the outermost circle. The picture clearly shows that new species arise in an explosive mode in an evolution cycle solely due to the random nature of genetic changes. Therefore, the size of new species descending from one common ancestor is determined by the number of intermediates that survive to the disarmed states. The trails that lead to new species are productive trails.

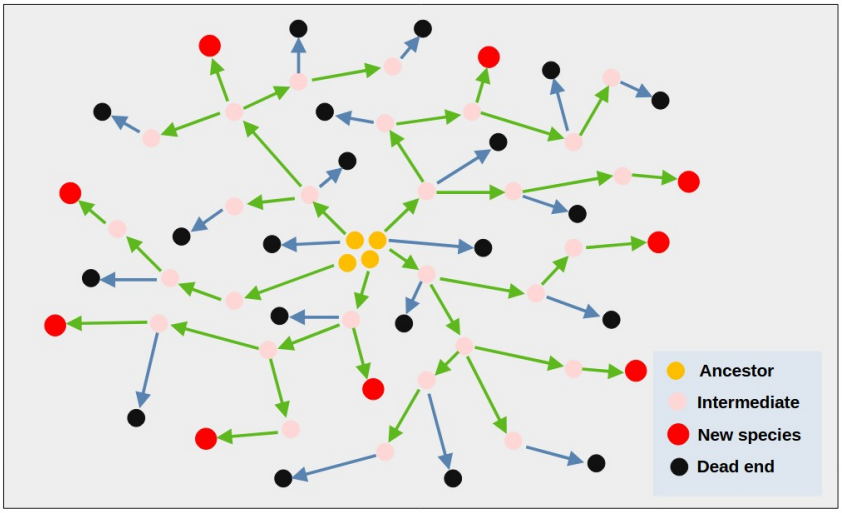

Figure 4. A simplified diagram to illustrate how new species arise in explosive mode in one evolution cycle. Distribution of new species is random relative to the ancestor organism. All new species can be classified into a single class.

Assume in a fish evolution cycle the fin inducing factor suffered from mutations in one intermediate to become a variant. The possible biological consequence was that it gave rise to a morphologically new and unique fin. If this intermediate ended up with five new species, while this variant didn't diverge

further, the fins on these new species would be very similar morphologically. Otherwise their fins would vary from identical to very different if the variant diverged further along the trails, depending on how far apart they were on the trails and the amount of mutations accumulated on each variant. Closer these variants biochemically and structurally, more similar the final fins on these new species. This isn't intended to explain the Cambrian explosion, but this illustrates a general principle behind the explosion of new species through evolution cycles.

Human evolution could help illustrate an evolution cycle at work, rough but a bit intuitive. It is more appropriate to say that all mammals arose not from a single ancestor, but from distinct ancestors that shared a lot of basic similarities. About 60 million years ago there was one ancestor X0. X0 could be an ancestor organism or an intermediate from another ancestor organism. Regardless of its origin, it diverged into a number of intermediates after X1 generations. One intermediate led to a variety of monkeys after X2 generations with many dead intermediates, while another one diverged into more intermediates of its own after X3 generations, among which one intermediate developed into early ape species after X4 generations, while another one moved on and diverged into more intermediates of its own. One of the intermediates became gorilla after X5 generations, and another one further split into more intermediates. One intermediate among them developed into different forms of chimpanzees after X6 generations, while another one led to more intermediates. One of these intermediates finally reached the earliest two-footed animal bipeda after X7 generations, establishing genus homo. Bipeda wasn't a dead end, but a lucky intermediate near the end of an evolution cycle. Its further divergence ended as humans, the only species emerged from this lucky intermediate after X8 generations.

In biology, there is a classification system that classifies living kingdom into eight levels based on shared characteristics. The last five levels are class, order, family, genus, and species. All primates fall in Mammalia class and Primate order. Then gorilla can be further classified into hominidae family, gorilla genus, and gorilla species. In similar way, chimpanzee as hominidae family, pan genus, and chimpanzee species. We humans belong to hominidae family, homo habilis genus, and homo sapiens species. In the above evolution cycle, the trail that reached humans may be the longest, hence referred to as human trail. It would be obvious that gorilla, chimpanzee and human share the same ancestor and a series of common intermediates until reaching a particular intermediate, from which their divergence occurred. Gorilla left the human trail and established its own genus gorilla. Chimpanzee and human continued to share some common intermediates before chimpanzees diverged from the human trail and established its own genus pan. However, the order of appearance can't be determined purely based on which species is more advanced physiologically and morphologically. In other words, the appearance of chimpanzees was unnecessarily earlier than humans.

Which species reached its disarmed state first in a cycle must be established by research only. In general, fossil records are useful in estimating the approximate time species appear on the earth, but not the time an ancestor organism began to evolve or started an evolution cycle. Moreover, not all intermediates left fossils behind to allow researchers to trace their evolutionary past. Scarcity of human related fossil records have hindered progress in tracing human evolution in the past 3 million years. On the other hand, the degree of the genome sequence homology between species is usually well correlated to their relationship on the evolutionary tree. Therefore, genome sequence comparison would be the viable resort to determine how close species are. Regardless of all this difficulty, life has been evolving through an unknown number of evolution cycles since the beginning of Cambrian period. What happened in Cambrian explosion was what happened in human evolution, albeit mammalian genomes being far more complex and richer in enzymes that can perform genetic operations.

New species are most likely to stay in a disarmed state indefinitely as long as they are not endangered by their natural habitats. Nevertheless, evolution has been a continuous process along the evolutionary timeline. While numerous new species emerged from intermediates in evolution cycles, some of them would transit into new generations of ancestors – daughter ancestors – to keep evolution going. When their DNA polymerases lost their high replication fidelity upon sudden geological and climate changes, they would start new evolution cycles. Therefore, evolution of life will occur when conditions strike ancestor organisms. It's an interesting and intriguing mystery if there are potential ancestors that are still crawling somewhere on the earth, waiting for a geologic event to rouse their evolutionary spirit.

8. Fast Evolution – A Closer View

Randomness has been changing its roles since life-like activity appeared in the incubator on the nascent earth. Origin of life is fully driven by randomness in the synthesis of proteins, RNA, and DNA. Slow evolution is driven by randomness that introduces vast sequence heterogeneity into the genomes. In both stages, a large number of basic proteins emerge from random pools to start and extend life via a trial and error approach at the cost of time. Therefore, on the macro level, randomness is the greatest during origin of life and early phase of slow evolution. In this period, randomness is the sum of short cell cycle, relatively error prone DNA replication system, and single cellularity. Consequently, there are numerous single celled organisms that carry their own unique genomes, in which many of them can be considered unique species, leading to vast heterogeneity of single celled populations. If each of these species diverges into a few more new species, new species will arise in an exponential mode. As a result, the number of modern day single celled species, including prokaryotes and eukaryotes, is too large to count. The greatest randomness during this period has independently brought about numerous unique, novel genes in different species, lifting gene counts upon cell fusions or endocytosis.

However, when multicellularity emerged, different cell types in an organism exerted constraints on the genomes to diverge freely, so limiting the genome heterogeneity. In addition, multicellular organisms are unable to grow as rapidly as single celled life, further reducing the randomness to a great extent. All this has limited the number of multicellular species.

On the micro level, emergence of a new species is the result of establishing a new balanced biological system, in which a series of changes brought up by newly generated functional genes have been successfully integrated into the existing system. As species become more complex and advanced, the integration processes become more challenging and risky, causing high failure rates and decreasing the number of new species from one evolution cycle. As evolution proceeds forwards, randomness has gradually changed its action mode from generating many novel genes de novo out of huge random pools to generating protein variants through modifying existing or duplicated genes. Randomness on the macro level no longer refers to the genome uniqueness of individual organisms in the living world, but is limited greatly to the genome uniqueness of the population of a particular species in the living world. This means that the transition of the roles of randomness on the macro level corresponds to the transition of evolution modes.

Genome sizes and protein coding gene counts shown in Table 1 give us a hint at what has happened in the slow evolution stage. Table 2 shows similar information for post-Cambrian species, from which we could infer what evolution is really about in the fast evolution stage.

Data are largely similar to what is shown in Table 1, but genome sizes have increased significantly as species move up the evolutionary ladder. Invertebrates ciona intestinalis and ciona savignyi are low species from Cambrian explosion era. Their genomes are relatively small, only 100 to 180 millions base pairs, but contain around 11,000 to 17,000 protein coding genes, accounting for about 50% to 90%

of the protein coding gene capacity of mammals. Tropical clawed frog genome contains about 22,000 protein-coding genes, which are comparable with the numbers from mammals while its genome is only about half the size of mammals. Exact protein coding gene number is virtually impossible to obtain just by analyzing the entire genome sequences, but it gives us a rough idea that gene counts and genome sizes are not proportional. The counts from fishes to mammals are largely similar, ranging from 15,000 to 25,000. An important implication of these numbers is that evolution cycles, for example from fishes to amphibians or amphibians to mammals, didn't seem to require additional proteins to become new species. Therefore, the gene counts are a relatively stable parameter of species across vertebrates, which suggests that protein variants have played more significant roles than we thought, while de novo generation of new genes was not only rare, but also unnecessary.

Some genes in an organism are so fundamental to life, or so unique in functions, or so stringent on protein sequences that they don't have much margin to tolerate sequence changes except on some noncritical positions. Mutations would occur to them as usual, but the consequences are either lethal or subtle in their biochemical or cellular properties and functions. Nevertheless, even very tiny changes could have considerable impact on the organism's morphology, even development process. They might have only very few or even no variants in the same species, but share high sequence homology with proteins from species of the same class, even species across classes. Some tissue or organ inducing factors are examples of this type of protein variants, which are independent of gene duplication, but the result of mutations that occur directly on the genes. If their genes are duplicated by chance, only one copy will survive, as more variants would be detrimental to the survival of the species. In genetics terms, these genes are very conserved.

Organism	Genome Size	Coding Genes	Non-Coding Genes	Pseudogenes
S. cerevisiae	12,157,105	6,600	424	12
C.elegans	100,286,401	19,985	24,813	2,128
Ciona intestinalis	115,227,500	16,671	455	27
D. melanogaster	143,726,002	13,986	4,054	340
Ciona savignyi	177,003,750	11,616	340	216
Sea lamprey	885,550,958	10,415	2,652	47
Zebrafish	1,373,471,384	25,545	6,599	315
Tropical clawed frog	1,451,301,209	22,107	2,023	484
Common carp	1,460,365,368	44,807	9,749	417
Common wall lizard	1,511,002,858	22,062	1,613	267
Softshell turtle	2,202,483,752	18,189	1042	97
Painted turtle	2,365,766,571	21,186	5,523	108
Coelacanth	2,860,591,921	19,569	2918	141
White-throated sparrow	1,052,600,561	13,862	4,340	50
Chicken	1,053,332,251	17,007	13,040	61
Small tree finch	1,064,016,346	16,191	883	93
Dog	2,396,858,295	20,567	9,944	610
Mouse	2,728,222,451	21,955	18,008	13,755
Black snub-nosed monkey	2,977,074,946	20,966	8,376	563
Gorilla	3,063,362,794	21,794	7,768	522
Human	3,099,750,718	19,830	26,462	15,222

Table 2. Genome sizes and coding gene counts of various species on different levels of evolution. Data are taken from Ensembl.

Sequence flexibility of some proteins allows their sequences to vary to various degrees while maintaining similar, but not identical structures or functions. It's this difference that enables protein variants to perform similar jobs with their own characteristics in different cells and different species, part of the molecular basis of cell differentiation and speciation. GPCRs and keratins are the best examples of protein variants of this kind. The presence of so many GPCRs or keratins variants in a single species attests the gene duplication dependent origin of these protein variants. Gene duplication is a fast and economical mechanism to derive protein variants for similar properties and functions on minimum sequence changes. Most duplicated genes were expected to end up as pieces of random DNA sequences or pseudogenes in the genome. Because of this, most of the genes didn't evolve into large families like GPCRs or keratins. This explains why there is a ceiling for the coding gene counts of about 25,000. Most of those "failed" duplicated genes were likely removed from the genome to keep genome size relatively steady after each evolution cycle. In this regard, GPCRs and keratins are clearly the exceptions, not norm.

If the total number of genes, on average, was assumed to be 30,000, including non-coding genes and pseudogenes, at the onset of an evolution cycle. The effective genetic changes, mostly point mutations and gene duplication, must happen to these 30,000 genetic loci. If one evolution cycle took 10 million years, all genetic changes that finally brought about new species must complete in this 10 million years. DNA recombination is largely independent of point mutations, and the occurrence of one would not interfere with the other. The frequencies of point mutations as well as recombinational events would be much greater in the early reshape phase of the cycle and then slowed down gradually as the process healing progressed. In addition to the genome-wide random point mutations, some of the other known genetic changes, including insertion or deletion of functional motifs, exon shuffling, creation of alternative splicing sites, etc, might have worked together to convert genetic loci into new functional genes or variant genes. It's the overall action of all those types of random mutations that slowly but steadily transformed the grand old biochemical and cellular landscape of an existing species into another grand new biochemical and cellular landscape – the birth of a new species.

In a hypothesized scenario for the purpose of illustration, suppose that there was a single individual ancestor organism with a genome size of 10^9 bps and one evolution cycle would result in a single new species. In other words, among numerous evolution trails, only one trail led to a new species. On average many animals produce offspring one year after birth, or one generation per year. If at the onset of an evolution cycle DNA polymerases incorporated 10 random point mutations in one meiosis per 1 billion bps, equal to $1X10^{-8}$ per base pair per generation. If one cycle spanned 10 million years, the final species could have accumulated about 100 million mutations. This means that 10% of the bases have undergone mutations after 10 million generations. Assume each gene contained 900 bases to encode 300 amino acids on average and 200 bases for regulatory sequences to control gene expression, the gene had a chance of a single point mutation $1.1X10^{-5}$ per generation. If 1,000 genes were house keeping genes that wouldn't result in phenotype changes on mutations, then the chance for the remaining 29,000 genes to be hit with a single point mutation was only 31.9% per generation, hence not a single gene would be subjected to one point mutation in one generation. Over 10 million generations, the chance for each base pair to receive 1 point mutation is 10%, which translated into average 110 ($1{,}100X1X10^{-2}$) point mutations per gene, and average $3.19X10^6$ point mutations for 29,000 genes. If mutation rate increased to 20 and the duration of one evolution cycle increased to 20 million years, then average mutations for one gene would be 440 in one cycle. Be noted that the weight of 110 mutations on a gene variant is much heavier than on a piece of random DNA sequence of the same size, as in the process of developing all-new genes during origin of life and slow evolution, indirectly indicating the critical rules of gene variants in the fast evolution stage.

Assume the average number of offspring per one pair of parent organisms was 5 from a single birth in a life span of one year, an evolution cycle lasted 10 millions of years, and each trail ended as the cycle ended. Assume again that all the offspring survived mutations and produced their own offspring, then the total maximum number of offspring produced at the end of the cycle or 10 millions of generations would have been an infinity number of $5^{10,000,000}$. This number would be much larger if the organisms could give multiple births in a life span of 2 or more years.

Although the total number of offspring produced in the 10 millionth years was an infinity, but it didn't represent the total number of new species from one cycle. We could estimate the maximum number of new species from one cycle under some assumptions. If all the intermediates survived to become new species, and the evolution trail for each new species had the same length, then every gene in the gene repertoire underwent 110 random point mutations per cycle, and diverged 110 times into 110 variations regardless when divergences occurred. In theory this would result in a combination of 110 variations of each gene for a total of 29,000 genes, an infinity number of $110^{29,000}$, the apparent maximum number of new species possible from this cycle. However, because the vast majority of evolution trails terminated randomly as dead ends at any time points due to lethality or infertility of the mutations, the number of survived offspring at the end of the cycle would be only an infinitesimal fraction of the maximum value of $5^{10,000,000}$, thus only an infinitesimal fraction of the maximum $110^{29,000}$ survived as new species. On the other hand, the functional impact of point mutations was highly unpredictable since they could be synonymous or too few to incur phenotype changes, even they could cancel each other over the period of cycle. For this reason, chances to bring about new species would be further reduced greatly to a new infinitesimally small number. Although the number was infinitesimally small, it was utmost significant since it truly represented the number of new species emerging out of an evolution cycle.

Sequence analysis has demonstrated that there exist good function homology but poor sequence homology among genes or proteins from species across different classes. This clearly indicates that protein functions and structures display remarkable elasticity on amino acid sequences, allowing sequences to diverge through point mutations, while preserving their basic biological roles, the molecular basis of protein variants in the evolution of species. However, after point mutations brought more sequence changes to proteins along the evolutionary timeline, the function, especially structure homology decreases. Within an evolution cycle, the homology is usually high, and protein variants differ more or less by pure chances. Nevertheless, these differences, when acting on the whole protein repertoire, are sufficient to result in new distinct species that fall in the same genus or family. When sequence changes of magnitude occur, protein products bear much reduced, but still identifiable overall function and structure homology. These new protein variants, when acting with any radically new proteins, can bring about changes in phenotypes that are profound enough to break the barrier of transition of species from lower classes to higher classes. Any genetic changes of such magnitude is the result of a genome reshape process ignited by geological and climate changes. The genome reshape process is the start of an evolution cycle and made possible only when the genome has reached the level of maturity after 3.5 billion years development. Gene and protein variants have played the critical roles in the fast evolution stage.

The reproduction rates of amphibian and organisms below are far larger by hundred or thousand folds per generation. Despite the fact that the survival rates of newborn life of low organisms are much lower, accumulation of random point mutations in low species must be much larger than the estimates for higher organisms when the whole population was taken into consideration. However, research shows that mutation rates of about $1X10^{-8}$ per base pair per generation occur almost in all modern species simple or advanced, revealing the two aspects of evolution. First the mutation rates must be

higher than $1X10^{-8}$ per base pair per generation in order to start an evolution cycle. Second, evolution doesn't seem related to reproduction rates too much. Higher reproduction rates aren't translated into higher accumulation of mutations that can trigger evolution, suggesting that normal random mutations are unable to drive evolution to occur.

Looking at evolution cycle from the point of whole genome, we could see a thread of events that went through the cycle. Higher random mutations change the discrete bases in genes, which results in discrete changes in amino acid sequences of the proteins they encoded, eventually affecting the biochemical properties of these proteins in a discrete manner. It can't be predicted that how the discrete changes in biochemical properties of these mutant proteins will change their functions in normal environments. Most mutations rendered protein products incompatible with the existing system by causing infertility, disruption of essential biochemical processes, collapse of cellular structures, all of which resulted in the death of intermediates. They were lethal or became lethal and eliminated from the populations immediately. Non-lethal mutations could become lethal as the cycle progressed. Only a relatively small number of non-lethal mutations was preserved, accumulated and finally gave rise to the emergence of new species at the end of a cycle.

In the evolution of more advanced species, some protein factors must have played critical roles in determining the final fate of an intermediate, while others such as fin inducing factors played deterministic roles in establishing the morphology of new species, for example, setting the general predisposition of organs in the body, deciding body shapes, brain size, development of wing, instead of limb, etc. The phenotype from these protein factors are all visible as the organisms go through their life cycles. Study of the evolution of these protein factors will reveal the process of how morphology of species has evolved from simple to complex to extremely complex.

Division of evolutionary timeline into slow and fast stages is scientific as two fundamentally distinct mechanisms are working behind each stage. The slow evolution stage took about 3.5 billion year to complete, reflecting the daunting difficulties and complexities of the tasks accomplished in this stage, during which many thousands of all-new genes are created de novo from random genomic sequences and their protein products assimilated into the existing small system, thus enlarging it into a vibrant and robust life system of high sophistication. It's interesting to see the protein coding gene counts lifted to the level of higher species, while the morphology remaining as simple as low multicellular life. The apparent lack of parallelism between the high coding gene counts and the simplicity of a morphology underlies the crudeness and immaturity of a number of genes created and accumulated in this stage. It's even doubtful that these genes actually coded for proteins that contributed to any biological activities that occurred in the organisms. Regardless of their true utility in the pre-Cambrian organisms, they are the abundant, ready-to-use genetic fodder, from which new proteins could be derived to build a more variety of features and more sophisticated forms, the essence of genetic basis of evolution. Clearly slow evolution is the preamble to the fast evolution by laying down the solid genetic foundations for rapid proliferation of new advanced species.

The relative stability of protein coding gene counts across the entire post-Cambrian living kingdom argues well for the conclusion that no more than 25,000 protein coding genes are required to build an organism as sophisticated as humans. As described earlier, various degrees of the protein sequence homology across the entire eukaryotic kingdom imply clear evolutionary links among most proteins with similar functions, suggesting what was the apparent main task of the fast evolution, first to derive new properties from the existing properties and then to assimilate them into the existing system, the result of which was new species with distinct morphology and physiology in a much shorter period of time. This was in stark contrast with the main task of the slow evolution – generating new protein

coding genes from random or semi-random DNA sequences to increase the coding gene counts, thus building up and enriching the genetic repertoire required for higher levels of life. If slow evolution is concentrated on creation, fast evolution takes advantage of what has been built in the slow evolution stage for reusing and recombining through mutation-based derivations. Because of this, post-Cambrian evolution has occurred in cycle. In each cycle, random mutations on functional coding genes resulted in distinct properties that drifted away more or less from the parent genes. As the cycle proceeded, more useful and handy properties appeared and changed the organisms overwhelmingly in every way, leading to the emergence of new species as the cycle ended. Any new distinct properties that appeared in a cycle were laid on top of the properties from the previous cycle, making the new species generally more sophisticated and more advanced than their ancestors virtually in all aspects of life.

It seemed apparent that higher complexity tended to have the capacity to generate more varieties, but as a matter of fact, the opposite was true. As species become more complex, the underlying biochemical and cellular machines become more delicate, constrained, and indivisible, requiring far greater balance among biochemical processes and cellular structures on the levels of cells, tissues, and organs. In other words, a simple life system was far more facile to have new components added and old components removed or replaced while developing into distinct new species. In sharp contrast, the larger and more intrinsically interlinked system was far less tolerant for having new things added in and existing things removed or replaced, consequently many fewer new species emerged from a cycle. This is indeed what happens in the modern animal kingdom. About one million of insects have been described and named from an estimated Insecta class size of around 5.5 millions. The described fish species is over 32,000, accounting for more than half of the vertebrate species. The amphibian species and species of reptiles are known to be around 8,000 and 12,000, respectively. About 6,400 extant species of mammals have been recorded. The relative small number of amphibians implied that they were possible living intermediates that survived the migration of organisms from water to land.

Birds are special in evolution. There are over 10,000 known bird species, and about 60 percent of them are passerine. Passerines are often small in size and grouped into families on the basis of morphological similarities. However, their morphological similarities aren’t the result of a close genetic relationship per se. All birds are evolved from common flying reptiles or small feathered dinosaurs around 160 million years ago, and many sequential evolution cycles led to the birth of ancient birds, many of which had disappeared from the earth today. The first passerines appeared 60 million years ago, and then diversified into three suborders, in coincidence with the separation of the southern continents into subcontinents. When the evolution cycles occurred in different subcontinents, genetic links among species in three suborders were totally broken. However, remarkably many species of passerines that were developed from different locations and classified into different suborders were morphologically similar. Many genes inherited from their common ancestors before geographic segregation showed no genetic sequence continuity in different species, but their protein products were similar in biochemical and cellular forms and functions. As a result evolution in separate subcontinents gave rise to passerine species that are genetically not close, but morphologically similar.

If morphological similarities aren’t the result of a close genetic relationship, then they are the result of convergent evolution. In convergent evolution new traits that appeared in different species had similar form or function but were developed independently. In contrast, the evolution discussed in this post is usually referred to as divergent evolution, in which new species evolve from a common ancestor by diverging from each other after developing their own specific new traits. Traits from convergent evolution is of analogous nature, while traits from divergent evolution is of homologous nature. Analogous traits are similar in forms or functions and aren’t present in their last common intermediate ancestor, so that they are independently developed. On the other hand, homologous traits have similar

forms, but can have dissimilar functions such as limbs and wings. Homologous traits are originated from a common intermediate ancestor.

If two evolution trails were split earlier in a cycle, the new species from the two trails would be more dissimilar in morphology. Therefore, in divergent evolution, morphological similarity was positively correlated to a larger common evolution trail that new species had shared with each other. However, in convergent evolution, morphological similarity showed no correlation to a common evolution trail that new species might have shared with each other. In the evolution of birds, many evolution trails, that occurred even in separate evolution cycles, different geographic locations, and timeline, all could lead to birds that were morphologically similar, but genetically unrelated called passerines. In the same time, new passerines could have emerged from the normal evolution cycle as well. Therefore, both divergent evolution and convergent evolution had contributed to the evolution of passerines, making it the largest order of birds in the bird world.

Human evolution is an interesting thing to look at. Modern humans appear just 300,000 to 80,000 years ago, while earliest primates appeared at least 90 million years ago. Monkeys that are closer to humans more than many other primates appeared about 40 million years ago, and the ancestors of the gorillas splits with the common ancestors of humans and chimpanzees about 10 million years. Chimpanzees, the closest relative of humans, split from early humans 8 million years ago. The exact time for these species to appear isn't important, but it's important to show clearly that the evolution cycle that leads to humans has been lasting more than 50 million years. DNA sequence comparisons show that genome sequences differ not as big as expected for humans and Chimpanzees. The two genomes are almost 99 percent identical in sequences that can be directly compared. The differences are attributed to single nucleotide substitutions, deletions and duplications of DNA fragments, insertion of transposable elements and chromosomal rearrangements. Human-specific single nucleotide substitutions accounts for 1.23% of human DNA, which seems to affect about 70% of proteins although the differences in amino acid sequences can be as small as only a couple of amino acids, the typical changes since chimpanzees and humans diverged from a common ancestor about 8 million years ago. More extended deletions and insertions cover about 3% of human genome. Therefore, when DNA insertions and deletions are taken into account, the sequence identity remains at 96 percent. The remaining 30 percent of genes code for the same amino acid sequences in chimpanzees and humans.

Humans differ from chimpanzees, gorillas, and other primates so extensively in every aspect from morphology to physiology to the brain size. Can 1.23% of genome differences mainly from point mutations account for all the differences between two species? There are no such genes that make humans humans or chimpanzees chimpanzees. There are a few classes of genes in humans that seem to be evolving more rapidly than in chimpanzees. These genes play key roles in human embryonic development, patterning of the nervous system, and more. Nevertheless, the vast majority of those limited genetic changes are widely scattered among the entire gene repertoire and they must have refined and honed every bit of genetic materials to form a perfect system, in which each gene expresses in such a precision manner in terms of cell type, timing, degree, coordination with others, and more. Derivation and utilization of protein variants in humans must have been so fine tuned that the changes in amino acid sequences and expression patterns can be small, even unnoticeable, but they are so subtle and to the point that have changed embryonic and post-birth development dramatically in morphology and physiology, especially the nervous system. It's the sum of all of these subtle changes that have made humans distinguish from all other primates.

In all likelihood, since splitting with chimpanzees, all the mutations that happened only on human genome optimized most of the genes to achieve the best overall phenotype of a living organism.

Generally speaking, in animal and plant kingdoms, for species that share a particularly long tract of an evolution trail, their genomes usually share a high degree of sequence homology with differences from limited point mutations, DNA deletion or insertion, etc. However it's these small differences that enable them to boast their own features, expertise, and unique survival strategy and peculiarity, all of which together distinguish one species from the others. In other words, it's again the sum of all the small differences scattered among the entire genome that depleted their identicalness, but conferred them distinct physiological and morphological characteristics.

All multicellular organisms start from a fertilized egg, while the egg provides only components to walk the first step in the entire life process. From life standpoint, it's the genome that directs the organism to complete its life cycle without input of external guidance or instructions. From evolution standpoint, the next generation of species always arises from the previous generation. As a result, evolution always moves species forward. From civil engineering standpoint, the genome is the greatest blueprint ever for making things from simple (in the eye of evolution) to unthinkably complicated. It plans and then executes every facet of a building process from design, layout, materials, organization, maintenance, and all the other aspects of engineering in precise and flawless manner, as well as in the greatest order, details, logic, and forms. A blueprint drawn from every genome in the living kingdom can be turned into a living marvel, coherently arranged, aesthetically pleasing, and economically efficient. Truly the genome is the finest thing in the universe.

9. Rethinking of Natural Selection

Natural selection is a fundamental element of the evolution theory. It illustrates how the remarkable biodiversity on earth has been driven and shaped by natural selection in the entire timeline of evolution in a simple and elegant way. Natural selection is a process through which some individuals in a population adapt and change better to suit the habitat than other individuals in the same population, and as a result survive better and reproduce more offspring. Differential survival and reproduction of individuals are due to genetic variations that produce some favorable traits to give them some surviving advantage. Upon passage of those favorable traits onto their offspring over generations, they become a better fit for the environment and more common in the population. Through natural selection, favorable genetic variations, thus favorable traits, are passed down through generations. After heritable genetic variations that underlie phenotypical changes in a population accumulated to a substantial amount over numerous generations, the individuals that carried those variations became a distinctly different species – new species.

Natural selection may be able to explain how species improve over time through heritable genetic variations, but its role in the origin of species is too ill grounded to have any relevance. If natural selection led to new species, an inference would be that all species throughout timeline were in the process of evolution, and the appearance of new species meant the disappearance of old species or ancestor species. This is clearly an inconsistency with the extraordinary biodiversity today. This section will focus on why natural selection theory can't be a cornerstone of modern biology.

On the evolutionary timeline eukaryotic organisms appeared to reproduce sexually at the single celled stage about 2 billion years ago. Since then sexual reproduction seemed parallel with the evolution of eukaryotic organisms. Almost all modern eukaryotic organisms produce offspring through sexual reproduction. Sexual reproduction is costly and of low efficiency, but it is universal for all multicellular organisms, indicating that it has advantage over asexual reproduction. Main advantage seems to increase genetic diversity in the population and mitigate accumulation of harmful genetic mutations.

Adoption of sexual reproduction system confers eukaryotic organisms two sets of genomes, germline genome and somatic genome. Information flow between the two genomes is unidirectional from germline genome to somatic genome. As a result, germline mutations will pass on to the somatic genome of next generation, while somatic mutations can't do the same to the germline genome, making somatic mutations short lived to the life span of the mutation carrier. When we talk about mutations, it always refers to heritable germline mutations.

Any mutations can exert some consequences to the organism regardless of being germline mutations or somatic mutations – deleterious, neutral, or beneficial. Only lethal type of deleterious mutations has a clear-cut meaning, while all the other consequences are quite variable in degree. Current understanding is that it's natural selection that determines if a mutation is deleterious, neutral, or beneficial. Beneficial mutations produce advantageous traits, which, under natural selection, allow the mutation carrier to survive better or reproduce more offspring, eventually becoming more common in the population. Only non-lethal mutations will be passed down to next generations.

So far there are many examples to demonstrate natural selection at work, and the origin of giraffes' long necks is the classic example. Giraffe's ancestor inhabited in dry savannahs of Africa with open plains and woodlands. There trees were tall and hard to reach for animals of normal necks like deer or antelope. Some day genetic mutations occurred in the ancestor's genome, making ancestor's necks grow longer. Long-necked individuals gained not only advantage to reach leaves on the high treetops, but also wider panorama view to maintain horizon vigilance that could allow giraffe to browse safely over wider areas, thus improving survival. As a result, long-necked giraffes were able to eat more and produce more offspring. As the genetic mutations passed down generations over time, their necks continued to grow longer until they reached present length. The long-neck seemed to be a favorable trait for giraffes as it made individuals adapt better to the dry savannahs. Thus, they became the most common in the population. Because the long necked animals were morphologically so different from their ancestors, they were called giraffe and qualified as a new species. It could have taken millions of years for the giraffe's ancestor to slowly develop into present-day giraffes. This explanation seems plausible and relatively easy to understand even for general public.

Nevertheless, the development of life through evolution over billions of years can't be as simple and straightforward as illustrated by natural selection at all. There are unsurmountable blocks when natural selection theory is used to explain evolution a little deeper and in more details. Let's take a closer look. Assume that giraffes shared a common normal necked ancestor with deer or antelope. One individual ancestor suffered from some mutations in the gene coding for the protein factor that guided neck muscle development. This new protein factor variant would now guide the neck muscle to grow longer. In other words, the appearance of giraffes as a new species was likely triggered initially by some random mutations of similar kinds.

The development of a long neck wasn't a single event isolated only to the necks, but affected giraffes in its entirety. To physically support a long neck, giraffes will have to pump more blood to the upper body and change its body shape in order to run at acceptable speed and keep body balanced. For this, giraffes must develop stronger cardiovascular, skeletal, and nervous systems and more. On the biochemical and molecular levels, a large number of protein molecules, new or variants, must be created to build up a phenotype – a long neck and everything else that must come to support the long neck. Meanwhile a corresponding gene regulatory mechanism must be established to make sure that each of those protein molecules would be produced in the right cells, tissues, and organs in the right time. The whole event of neck elongation was complicated and entangled with a large number of processes and activities that would be disturbed by those new or variant protein molecules, thus requiring the greatest coordination

and integration to guarantee that all the old and new would be working in harmony. The long neck could possibly become a viable outcome of evolution only if all the above conditions could be satisfied in similar time frames. However, changes of such large magnitude must be of far-reaching nature and well beyond what could be brought about by largely piecemeal genetic recombination and mutations that natural selection depended on regardless the length of time.

If giraffes' ancestor did have a compelling need for a long neck to survive better, it was still random mutations that started and drove giraffes' ancestor to evolve into long-necked giraffes, not because of compelling needs and natural selection. In a more likely scenario, most of the intermediates in the course of evolution must have died from the lethality of mutations, and some of the intermediates might have a neck longer or shorter than that of modern giraffes. The only few intermediates that had survived all the changes reached the end of the cycle and emerged as a new species called giraffes. If you truly would like to give natural selection some credit, the best role it could play was to check whether giraffes would survive the mutations in its natural habitat, not survive better at all during and after the cycle. Nature won't give organisms something just because they have a compelling need for it.

Attribution of evolution of giraffes' long necks to natural selection stood on flimsy ground. Giraffes' ancestor wouldn't be the only mammal living in such a habitat. Why did only giraffes develop such a long neck, while other mammals like deer or antelope remained normal necked and have survived just well? Does it mean that no other mammals have a need for long necks to gain survival advantage? Or does it mean no other mammals have a compelling need to eat leaves on tall treetops? From survival point of view, an excessively large body size gives animals more survival disadvantages rather than advantages. A large physical body easily hinders its movement and reproduction and requires extra food consumption to sustain normal life activity. All this seriously limits its population size and makes the animals succumb more easily to food shortage and natural disasters. Therefore, giraffes as a new species at the time of its appearance didn't gain any survival and reproduction advantage over normal-necked animals except the banal advantage to eat leaves on tall trees. Even this advantage might not exist if all the trees in the ancient habitats weren't as tall as today.

Giraffes and its closet relative short-necked okapi diverged from their common ancestor about 11.5 million years ago, yet giraffes and okapi shared only about 20% identical proteins, attesting the great magnitude of genome changes during the evolution of giraffes. Giraffes appear in the fossil record around 4 million years ago. A time span of 7 million years seemed too short to endure the grand genetic changes that gave rise to giraffes.

Bird vultures have a strange craving for dead animals. How such a behavior emerged couldn't be accounted for using natural selection theory either. No birds must rely on dead animals for food to survive. Feeding on dead animals required bird's many biological systems be reshaped in addition to its digestive system. First the bird must gain a strong appetite for putrid carcasses, involving olfactory cells and taste buds. Second the birds must develop a tough stomach to digest rotting flesh and kill infectious agents coming with the dead bodies. Third the birds must establish a good vision and nerve-muscle system that would allow them to look down for targets from high positions. Forth, the birds must develop some peculiar behavior to support their strange diet such as disgorging food from their crops to feed their young. All this couldn't be made possible without genome wide changes that must take place in sync. During the evolution, most bird intermediates must perish from the infectious agents when their stomach was weak for dead animals. Fortunately enough, some of the intermediates developed a stomach from random mutations that was so tough that this evolution trail survived and continued to an end with the emergence of more than 20 vulture species. In all likelihood, appearances of giraffes, vultures, and all other species were way beyond what natural selection could explain.

The best examples of evolution and natural selection came from Charles Darwin's observations of bird finches in Galápagos islands. Finches' bill sizes and shapes are attributed to each bird's adaptation to a specific type of food on the islands. For example, a thick beak adapted to feeding on crunchy seeds and arthropods, while a slender, pointy bill adapted to catching tasty insects hiding between the leaves. There are more examples to the list. Curlew's long bill can probe deep into the mud and shallow water to catch aquatic invertebrates. Great egret's long legs allow the birds to walk in relatively deep water to search for fishes. Woodpecker's long and strong bill with chisel-like tip is good for prying arthropods out of holes on tree trunks.

If we think a little more, it's not difficult to realize that appearances of these highly specialized bills or legs on different birds in the course of evolution and the formation of their life styles are actually the chicken or the egg problem, a causality dilemma if you are stubborn enough to put them in order. Birds developed a thick beak because there were abundant crunchy seeds and arthropods to eat. Similarly birds grew long legs in order to adapt to deep water habitat. This is clearly the answer from natural selection theory, in which the natural pressure seemed to have played dominant roles in deciding what types of bills or legs birds would develop. Nevertheless, the opposite explanation is much sounder and more agreeable to how evolution has occurred. Birds had developed specialized bills or legs first. Because of the thick beak, the birds became able to feed on crunchy seeds and arthropods. Similarly, because of the long legs, the birds gained the ability to enter the deep water to look for fish. This is an active adaptation of a habitat that well fitted the traits unique to each bird species. Therefore, the long legs enable egrets to enter the deep water for food, not food in the deep water that forced birds to develop long legs. Thick beaks allow birds to feed on crunchy seeds and arthropods, not crunchy seeds and arthropods that drove birds to develop thick beaks. However, both views were purely inferred from what types of beaks or legs birds possess at present. From the evolution standpoint, different types of beaks or legs we see today were all developed over tens of millions of years. It's meaningless to argue which comes first, the chicken or the egg.

Could natural forces make birds' legs long or beaks thick so that birds could survive better? The answer from the second explanation is that natural selection was no more than a type of adaptation to the natural environment. It's the birds that have played active roles upon having developed special traits that allow them to adventure into new proper natural environments. A bird can't choose what types of traits to have, but it can achieve the best use of whatever traits it has by actively finding a habitat that those traits fit well. It's the organisms that select a habitat, not the habitat that selects organisms.

The active adaptation is more likely to be what has happened during billions of years of evolution. There are more examples to support this view. Fish living in the dark cave is usually blind. It isn't the darkness that made the fish blind, instead blind fish could adapt to or survive in the dark cave due to lack of predators. The limbs of animal sloths are long, and their hands and feet are specialized to have long, curved claws. Their unusually low metabolism inhibits fast movement and intense activities. All this made sloths adapt to a stationary lifestyle by hanging effortlessly upside down from tree branches virtually whole life. Asian vine snakes are adapted to arboreal life, because their green color allows the snake to camouflage in dense green leaves to avoid predators and prey on lizards, frogs and other small animals. A mangrove is a shrub that adapts to grow in saline water along coastlines and tidal rivers, since they can take in extra oxygen and excrete salt, allowing them to live where most plants can't. They also produce offspring using a special mechanism to increase the survivability. In general, more specialized in morphology and physiology, higher tendency to adapt to habitats of narrower conditions.

An array of prominent phenotypes observed among animals in nature shouldn't be considered as traits to favor survival and reproduction after million years of natural selection. The enormous body sizes of many dinosaurs shouldn't be selected to fit certain natural environments. Excessively large antlers on some male deer could be detrimental to their survival when traced by predators in dense woodland. The birds Rhinoceros Hornbill possess a long, down-curved bill with a brightly colored, unusually enlarged bony structure on its top. It's hardly to think of any utility of such structures for reproduction and survival, but a burden to carry. A lot of frogs are in danger of extinction since they are extremely susceptible to environmental variations and barely able to exist outside their present special habitats. It seems that many million years of natural selection hasn't brought up traits that would favor their survival and reproduction to broaden their natural habitats. All traits, good or not good, fall on species in random fashion and have nothing to do with better survival and reproduction as claimed by natural selection. All species must live with whatever traits they have, so long as these traits would not cause serious morphological and physiological defects that would be lethal to their physical activities.

On the Galápagos islands again, a completely new finch species was created in the wild in just two generations by the mating of two different finch species. The importance of this observation was over exaggerated. Mating between different species is not often. First different species don't attract to each other for mating. Second fertilization couldn't occur due to recognition failure between an egg and a sperm. Third if fertilization succeeded, the hybrid offspring would carry two sets of proteins serving the same functions. These two sets of proteins, more often than not, were unlikely to be 100% comparable, thus disrupting normal biochemical processes and leading to the death of the hybrid offspring. Fourth if hybrid offspring did develop normally, it was often sterile or reproduced with difficulty. If the hybrid offspring was able to reproduce, then it indicated that the two parent species were close enough for mating, and nothing more. It would not be scientific to conclude that the hybrid offspring was a new species based on their appearance, food preferences, etc.

Almost all the species possess an innate ability – the tolerance that can be stretched to some degrees to sustain the environmental changes big or small. A lot of species would have gone extinction without such tolerance. Animals like rats are blessed with exceptionally strong and flexible tolerance, which enables them to adapt to a broad range of harsh and mild environments. In contrast, animals like some amphibian species are well known for their poor and rigid tolerance, forcing them to survive only in the niche space with ecologically strict conditions. Phenotypes or traits that determine the environmental tolerance of the species can't be shaped by natural selection over time. As usual they come out of the evolution process. In general, strong and flexible traits of a population are the most critical factors that drive the population to defy natural pressures and spread widely to a variety of environments.

Natural selection is often taught in the basic biology classes, and its importance as a theory in modern biology doesn't need to be emphasized more. Without giving it a second thought, natural selection is established in many minds as a mechanism by which populations adapt and evolve. It is an engine that moves evolution forward through natural pressures on the organisms. However, can natural selection really explain evolution of species as it has been claimed for many many years?

Every object, living or non-living, has inherent properties that determine its behavior on the macro level. For example, the freezing point of water is zero degree C, an inherent property of water. This property determines the behavior of water at zero degree C – freezing into a solid state. Temperature coexists with water and decides its state in the wild. However, the freezing point of water is the critical factor in turning water into a solid state, while zero degree C isn't an inherent property of temperature, but is one value in its range, at which water freezes. Therefore zero degree C is only an external condition to realize the freezing behavior of water at zero degree C. The temperature zero degree C is

applicable to any substances, while the freezing point of water is specific to water. When the night temperature drops to zero degree C, water freezes, but alcohol, gas, and many other liquids won't because their freezing points are far below zero degree C.

In the living world, organisms' behavior includes diet, life style, and the way they interact with other individuals of the same or different species. Every behavior is the outer manifestation of the collective morphological traits and the underlying physiological and biochemical processes, all of which are the inherent properties encoded in the genome. For example, some bats can emit ultrasonic sounds to produce echoes. By comparing the outgoing pulse with the returning echoes, bats are able to detect prey and navigate in the darkness, an ability termed echolocation. Echolocation is an inherent property of bats with cricothyroid muscle located inside the larynx. Cricothyroid muscle can generate and emit ultrasound through the open mouth. Bats' ears then measure the time delay and the relative sound intensity between its own sound emission and echoes returned from the object. This time delay and relative sound intensity information is sent to the auditory cortex in the brain where the distance and positions to the object are determined. Echolocation allows those bats to live in dark caves and go out to prey on flying insects at night – the general behavior of those bats. Ants produce an array of pheromones in different situations, and each type of pheromone elicits different biochemical and cellular responses on the receiving side, which then are translated into unique kinds of behavior. When all those behaviors combine together, ants become well known as sophisticated social insects. They form a variety of colonies with clear division of labor and unique pheromone-based methods of communication between individuals. As a result, ants operate as a well managed organization, working together to search for food, reproduce, defend and support their colonies, and much more.

Different species behave very differently. Some species are quite hostile, even belligerent in behavior towards others, showing very competitive nature. Some are weak and posing no risk to the others, but vulnerable to predators. Most of the species behave in between. A full display of the behavior of an organism is greatly influenced by the constraints from surroundings such as food availability, space occupancy, threat from predators. As a result, all the organisms of the same or different species are constantly facing challenges, competing with each other for food and space, enduring natural disaster, diseases, predators, etc. Only organisms that have the capacity to beat those challenges will be able to survive and reproduce. The geologic conditions of an environment are also a factor in determining what species can fare well in it. A given environment is a fixed territory, never changes except seasonal changes, and therefore is fair and unbiased to all the organisms that happen to arrive at it. Only those organisms, animals and plants, that can live with the geologic conditions can settle and survive. For example, plant cactus are succulents, and their thickened stems and highly modified leaves are an inherent property to store water and prevent water loss in very dry environments, thus giving rise to the behavior to live in dry desert. It isn't dry desert that selects cactus to be its resident, but because this special behavior enables cactus to adapt to the dryness of desert. Otherwise, cactus, like all the other plants, would succumb to dryness. This behavior isn't developed to suit the desert, but randomly acquired during evolution.

Looking closer at the savage wild, hostile behaviors would spur fierce fight for more food, mating rights, territory supremacy among individuals of the same or different species, inevitably deciding which individuals would prevail over the differential survival and reproduction and dominate the population. Whenever individuals of aggressive nature came together, fights broke out and produced winners and losers. Suppose that some individuals acquired a heritable variation of a particular trait that favored physical strength, thus increasing the chances to win the reproductive advantage. If winners' favorable trait spread to individuals in broader environment, they would become dominant in the population. This was an endless process and would go on and on for an unknown period of time until

the winners became the most common and the losers disappeared from the population. According to natural selection, most of the individuals in the population would comprise descendants of the winners, and their ability to survive and reproduce was superior to the individuals from which they descended. This is the process that drives the evolution of species.

However, this statement is deceiving and single-minded. First, what happened in the wild was more than likely to be a different story because of the complexity of animal behavior and the ability to adapt and survive even for the losers. The winners might never have the chance to become the most common in the population. Second the fight was driven spontaneously by the innate hostile behavior of the individuals, while the natural pressure or genetic variations that favored survival was not relevant at all, unless this kind of behavior itself was considered a type of natural selection. Take one step back, if every individual of the population was descended from the winners and carried those favored traits, the fierce fight continued and generated winners and losers even in the absence of new favorable genetic variations, truly an endless event without a possibility to conclude ultimate winners. Therefore, nothing could change such a population in any meaningful way and novel genetic changes for a new species would be a far remoter possibility even after a million or tens of million years. It would be the case that the grand old winner individuals would be the same species and remain susceptible to this and that as usual, because evolution didn't occur to them at all. Third the time factor was missing when drawing the statement, which will be left to the next section.

Different species reproduce offspring differently in radical way in the wild. The population density and the behavior of species could influence how animals reproduce in considerable degrees. Under either high or low population density individuals carrying favored traits may not have more chances to mate and reproduce offspring than normal individuals. For example, low organisms like fishes or frogs lay numerous eggs and sperms to increase the chances of fertilization and offset low survival rates of the new born from harsh elements in the aquatic environment. Larger the number of eggs and sperms, larger the number of new born organisms, more diluted the concentration of those that carry favored traits, and smaller the chances for offspring with favored traits to become the most common in the population. Many reptiles and mammals, especially those solitary species, spread over a large area, and the chances for them to meet and select the ones with favored traits are even smaller. On the other hand, traits that favor survival don't necessarily favor reproduction, especially when species become more complex and advanced, because too many factors are involved in reproduction and embryonic development. In all likelihood, the chances to pass favored traits down generations aren't that high in the wild. So long as there is no universal pattern for organisms to mate and reproduce across living kingdom, no general statement should be made as the one from natural selection.

Most of the species on earth today behave mildly, and they are usually located at the bottom of the food chains, thus extremely vulnerable to natural predators. Nevertheless, they have been in existence ever since they emerged at the some time points on the evolutionary timeline. Evolution must have come up with all types of morphological and physiological features for these species that will arm them with unique behaviors to defy and evade the danger from predators, thus survive and even strive in their own habitats. For example some areas with stringent environmental conditions have been the safe heavens for certain vulnerable species. It suggests that all weak species must have been equipped with features and behaviors, especially features and behaviors most critical for survival, no later than the moment they appeared in an evolution cycle. Otherwise they would have succumbed to all the dangers from the environment.

Some morphological and physiological features observed in certain species are minor and uncommon, but they are believed to be the result of evolution and important for survival and reproduction. These

species also display special behavior comparable with those features and believed to be important for survival and reproduction as well. This belief is based on a more general and broader belief that any features observed on an organism are important for survival and reproduction as evolution won't make thing without usefulness. Nevertheless, such a belief is based on a flimsy ground as well.

Colorful feathers of male birds like peacocks were considered to have a large influence on sexual selection in the mating season. This would be a valid explanation if only some male birds of the same species have colorful feathers. Are there male peacocks that have dull feathers? In addition, not all male birds in the avian world have colorful feathers and they mate with female birds just as well. More interestingly, many dull feathered birds like sparrows can grow in populations that are far larger than the populations of birds with colorful feathers, suggesting that colorful feathers don't provide any advantage for sexual reproduction, but a dangerous sign to attract predators. Some animals and plants can change colors, a phenomenon called camouflage. Camouflage was also believed to be useful in evading being eaten by predators, thus increasing the survivability. But a vast majority of animals and plants can't camouflage, and they have no problem to survive. Some animals or plants can inject venom into their predators as defense weapons, but it isn't a must since they can still survive without it. Other similar traits displayed on organisms can contribute to the survivability of the species more or less, but are not as important as believed by many. Generally speaking, it's the major physical and physiological features like bill types and leg length of an organism that determine the most critical behaviors, and these behaviors in turn determine the diet types, natural habitats, and the way the organisms co-exists with other individuals. All the traits, thus all the behavior, are a giving, not a choice to all the species. Those features like colorful feathers, camouflage, and venom are exceptions for some species, not the norm in the living world. If a male bird is suddenly given colorful feathers, then it could show off them to attract female bird, although it is a common behavior of male birds with colorful feathers. Similarly venom injection and camouflage were of auxiliary nature, and can be used against predators to some advantage only. As emphasized before, all the features, common, special, even weird, are the results of random mutation based evolution. There is no reason why some species own them and other don't. Is an explanation important in this regard? It's an utter waste of time to explain why and why not. The only true thing is that it's the diversity of traits and behaviors displayed on organisms of different species that constitutes the tremendous biodiversity on the earth today.

What could be said about traits, behavior, and natural selection in lieu of evolution cycle? When viable intermediates sustained prolonged genome-wide changes and succeeded in reaching new disarmed states, an evolution cycle ended, and new or changed phenotypes or traits dominated the species that survived the cycle and became new species. Major traits determined how new species would behave, and the new behavior further determined how the organisms would feed themselves and defend or shield themselves against predators or unfriendly environments, ultimately the survivability of the new species. Minor traits, including those rare traits, determined how the new species would behave in their own unique, even peculiar ways, enabling them to occupy special habitats and display funny or strange acts in their life cycle. If new traits put the new species into disadvantageous situations, for example, the new species faced food scarcity due to diet changes, or increased risk of predators, or intolerable to environmental conditions, the new species would migrate to areas where conditions would fit their behavior and life style if they were able to migrate. In this regards, new traits or variants of the existing traits wouldn't be always useful to benefit the new species in their native environment, rather be a force to disseminate species across lands and settle in safe territories they could reach. Clearly natural selection wasn't part of the entire process of evolution.

As a possible example, the ancestor of giant pandas might be a mammal indigenous to an area where bamboo was a rare plant species. During evolution, panda intermediates developed a quaint digestive

system for bamboo, including bamboo-loving taste buds, strong teeth and a tough stomach, all suited to eat bamboo as diet. Such an unusual digestive system prompted giant panda to migrate to places where bamboo was abundant. From genetic point of view, panda genome determines bamboo as its major diet, and bamboo diet in turn determines what kind of behavior panda will exhibit. It's this special behavior that inspired panda to migrate to bamboo rich terrain as its native habitat. However, the narrow appetite for bamboo put panda in a grave disadvantageous situation by being confined to a bamboo rich area only. Regardless of being disadvantageous, a trait can't be manipulated or discarded. The species either find a way to live with it or perish because of it.

10. The Zero Sum Rule and Evolution Cycle

If you look at the extraordinary biodiversity and ponder how it grows over evolutionary timeline, you would feel strongly that natural selection is merely an empty shell without substance. Its explanation of the origin of species is too simple-minded and skin-deep, particularly it can't explain how the traits expressed on ancestor organisms can become so radically different that they turn ancestor organisms into new species, why new species emerge in explosive mode in relatively short periods, and why millions of species of all complexity and variety coexist today.

A biological trait in general has a quite complicated genotype behind and is stable over time once it becomes part of a species at the certain point during its evolution. The stability of traits is the key foundation of biodiversity. Today's biodiversity contains a sweeping collection of species that could possibly form within earth's atmosphere and geology. In addition to millions of advanced species ranging from the early arthropods, early fish to modern insects, fishes, amphibians, reptiles, birds and mammals, modern biodiversity encompasses the simplest forms of prokaryotic life that formed 3 billion years ago on the nascent earth, the earliest forms of single celled eukaryotes that evolved from prokaryote archaea about 2 billion years ago, and millions of low species that appeared before, during, and right after Cambrian Explosion. It's the trait stability that has made it possible to preserve the continuity of species since they emerged hundreds of millions, even billions of years ago. Strictly speaking, evolution of a species has virtually stopped after emerging from an evolution cycle.

Natural selection is unable to explain why a genetic trait is stable over billions of years, but predicts that a genetic trait is under constant selection. It's true that random mutations occur spontaneously and sporadically to any genes in all organisms during DNA replication, causing genetic variations. In most cases, these genetic variations are not lethal, but have irregular effects on the phenotypic traits of the individuals in varying degrees. According to natural selection, only favored genetic variations will be inherited by and spread to more individuals and finally dominate the population. As a matter of fact, most of the mutations are believed to be slightly deleterious, which drifts the phenotypic traits away from the norm, resulting in weakened survival and reproduction. Those individuals, under natural selection, will be negatively selected for gradual elimination from the population. Furthermore, in the absence of natural selection, these weakened traits would become more variable and deteriorate over time, possibly ending as a vestigial evidence of their existence in the history of evolution. Natural selection always favors the heritable genetic variation that results in the fittest individuals, and leaves those deleterious variations to become vestigial. All this sounds so flawless and convincing, especially as easy as a piece of cake, but it neglects the most critical part of the evolution – time.

Evolution is a process of infinite nature in time, in which the impact of natural mutations on the phenotypes of individual organisms, thus the evolution of the species, is the overall result of extremely occasional mutations accumulated over periods of millions and even tens of millions of years. It must be very cautious to draw general conclusions largely based on observation of some modern day species, such as the shapes of finches' beaks or giraffe's long neck. Point mutations occur independently and

sporadically, indicating that every base of a gene mutates with similar probability during DNA replication and a mutated base can undergo one or even more mutations with the same probability so long as time goes on. A consequence is that the effects of mutations aren't lasting, and early mutations can be changed, even reversed by later mutations, regardless of the nature of earlier mutations on the phenotypes. Moreover, a deleterious mutation at time A can be the basis of beneficial mutations at a later time B, and vice versa. Therefore, the overall effects of non-lethal mutations on evolution seem more likely to maintain the status quo of species, not to be a key mechanism of evolution through natural selection.

In game theory the zero-sum game is a situation that involves two competing players, where player one's gain is equivalent to player two's loss. The name "zero-sum" is not used only in game theory, it is used to describe any situations involving two or more entities, where the sum of all winners' gains and the sum of all losers' losses cancel each other out, and the net result is that the final sum is zero. If zero sum theory is applied to mutations based evolution, there are differences. In game, the winner's gain is the instant loss of the loser. In evolution, gain and loss are not simultaneous, but separate in time by tens, even hundreds of thousands of years. A gain from one change can remain to be a gain indefinitely, until a later change results in a loss that eliminates the early gain, vice versa. If an early gain is never overthrown, then it isn't of zero sum nature.

In addition, the impact of any mutational events on the phenotype are varying and not clear cut largely because it can't be accurately measured in quality as well as quantity, indicating that the sum of advantageous events and disadvantageous events can't be exactly zero, but oscillates approximately along a sort of baseline phenotype. This isn't exactly a type of zero sum game, but neo-zero sum game. From an evolution standpoint, neo-zero sum theory is more correct to be a zero sum rule. Because of this zero sum rule, nematode C. elegans born hundreds of million years ago can pass on to today still as C. elegans, although the worm must have been struck with random mutations billions of time over the period. In a disagreement with natural selection, the long term effects of spontaneous and random mutations on species are of zero-sum nature. The immediate significance of the zero sum rule is that it has maintained the essential stability of biological traits and thus species over billions of years. Hence, the zero sum rule has maximized the biodiversity by preserving the continuity of existing species in parallel with the steady emergence of new species. Even if some mutations do change phenotypes permanently and positively, they can't be broad enough as the causes for new species to appear, rather they just tweak the traits in small degrees, making them perform better and improving the stability of genetic traits, thus the stability of species.

Differentiation of cell types into nerve cells, thus nervous system, has changed every aspect of life of multicellular organisms, particularly it has shaped the behavior of an organism and its responses to external stimuli in a fundamental way. The nervous system exerts its influence through learning, while learning is a process of acquiring knowledge or skill largely from experience as well as coaching by the previous generations. Aplysia, also called sea hare, is a pre-Cambrian organism classified into the Mollusca phylum. Aplysia is well known for its long term memory, associative and non-associative learning, in spite of its simple nervous system comprised of only about 20,000 neurons. The behavior of Aplysia is shaped by learning. For example, learning allows Aplysia to associate a shock with a touch on its siphon, and as a result, it retracts its gill, siphon and tail for protection. This is a quick neural response necessary for a speedy reaction to danger. The learning displayed by Aplysia doesn't depend on coaching or demonstration, but is a simple type of involuntary response to a stimulus, called reflex. In this aspect, Aplysia's nervous system is too simple to be able to demonstrate to other individuals. As the nervous system becomes more complex and advanced, the learning becomes more

complex activities as well, requiring coaching more than experience. Through learning, organisms acquire all the skills essential for survival and reproduction, not just the simple action through reflex.

Organisms of species sitting at the bottom of the evolutionary ladder learn largely from experience, the cost of which is the great mortality, which must be compensated for by producing an overwhelming number of offspring per generation. More advanced organisms learn skills by following the behavior of their parents or individuals of previous generation in the population in addition to experience. Fruit fly D. melanogaster has developed a nervous system that is quite complex and advanced comparing to that of Aplysia. This nervous system enables the flies to learn not only from experience but also through mimicking other individuals. If a naive female fly has observed other flies to copulate with a certain type of male, it tends to copulate more with this type of male. Male flies learn how to copulate with female flies more than female flies do. For example, naive males attempt to court and even copulate with female flies of other species, immature female flies, and even other male flies. This kind of nondiscriminatory behavior becomes much less likely after lessons learned from failed copulation. In addition, after male flies have experienced copulation, they change their courting behavior to finish courtship in less time. Nevertheless, the reproductive learning curve exhibited by fruit fly is still very rudimentary comparing with more advanced species. For social insects like ants, their behavior is far more sophisticated, requiring comprehensive elaborate coordination among various types of cells, tissues, and systems through a series of pheromones, sounds, and touch. To establish this type of behavior, young insects must learn by following adults to place each of themselves into specific position inside the colonies and fulfill their roles as queen, worker, and males, respectively.

A lot of animals born and raised in captivity are unable to survive when released into the wild. Young migratory fishes produced in an artificial environment can hardly survive to adult age in the wild river, because they haven't learned how to migrate from fresh water to salt water where they will grow to full size. Tigers born and raised in the zoo don't have the skills to prey, and even shows great fear when seeing a chicken running around in front. Based on these observations, animals raised in captivity must undergo extensive training to regain skills for survival before being released into the wild. All this is to make one point that the behavior of an organism isn't formed over night but over millions of years living in the wild. The behavior must be bolstered and shaped throughout constant learning from surviving in their native habitats in addition to the genotypical traits. More properly, all behavior is preserved and passed down generations via learning in the wild under the mentorship of their parents or individuals of previous generation. Without the right environment and lessons to follow, behavior as native and fundamental for survival as hunting for prey can be lost in a single generation. When learning is a prerequisite for life to sustain in the wild by establishing proper surviving behavior, it becomes an inherent part of the life cycle of many species. Learning is especially a critical requirement for those organisms that exhibit unusually complicated and peculiar behavior. In more general terms, the behavior of organisms is founded on their genetic compositions, but it is not genetically heritable like body shape. It's largely determined and influenced by learning interactions with the environments. Changes in behavior will change the survivability of organisms.

The nervous system is essential for learning-based behavior and it is ubiquitous even in organisms from pre-Cambrian period. An observable variation of a trait could be the result of a change in the nervous system, even in the environment, rather than in the trait itself as it is commonly recognized. For example, an individual Aplysia lost the ability to retract its siphon upon touch. From a behavior point of view, this is an observable change in behavior and also an observable changes in heritable traits. At the gene level, this could be due to either deleterious mutations in muscle genes engaged in retraction or genes responsible for generating reflex impulse in the nerve cells elsewhere. It can be common for behavior or behavior changes to cancel out the genetic changes in traits incurred from mutations of any

nature. Together with behavior changes due to the environment factors, the concept heritable trait lacks genetic clarity and is too blurry to trace the root causes for the observable changes in heritable traits, making it meaningless to state that evolution occurs when changes in the heritable traits favor survival and reproduction. It appears that there are two types of biological traits. The first type isn't dependent on behavior, such as feather colors, bill shapes, sharp teeth, etc, while the second type depends on the behavior of the organisms like fierce fight and net building, etc. Obviously only the first type of traits are truly heritable, while the second type is the utilitarian embodiment of the first type in practical use. It's improper and misleading to interpret complicated biological phenomena with theories as simple as natural selection.

The scope of natural selection is generally limited to evolutionary changes within species, not between species. Since the food chains were formed in the animal kingdom, the greatest threat to any organisms are not from within the population, but from the predators that preyed on them as food. By watching numerous documentary films about animal life, all the organisms are in constant struggle for food, and all preys are watchful and on habitual alert, preparing to run or fly away to escape predators. In the dangerous wild, individual organisms will succumb more easily to predator attacks if they suffer from mutations that weaken their innate protective mechanism against predators, while individuals will survive better if mutations strengthen their ability to evade or outsmart predators. An enhanced nervous system obviously enhances the survival of organisms. In this regard, predators play an important role in eliminating individual organisms that are in disadvantageous positions for survival for any reasons. The phrase survival of the fittest seems more proper for survival of the fittest from predators.

If evolution by natural selection isn't a valid explanation of modern biodiversity, what mechanism has been working behind? Earth today has been the common home for millions of species, ranging from very old to old to middle aged to new to very new, and this strongly indicates as in the previous section that not all organisms are in a state of evolution even though random mutations occur equally to all of them. On the other hand, it is the random mutations that have remodeled existing organisms into new species. All this suggests that evolution must have occurred not by means of natural selection as generally accepted, and but through organisms with yet-to-be-discovered genetic mechanisms that enabled them to be ancestors of new species. These organisms possessed extraordinary capacity to withstand a gradual and lengthy buildup of a variety of new or protein variants from random mutations. Some of these proteins can have immediate visible effects if they are related to morphology, and other assimilated into more complicated and functional cellular machine over time, all of which turn the old species into new species. In other words, ancestor organisms are organisms that are predisposed to be ancestors of new species. Only few ancestor organisms should be found in a class, such as in fishes, amphibians, reptiles. Each ancestor organism in a class evolves into species classified into the class of next level. The evolution cycle hypothesis seems adequate to clear up the mystery of how new species arise from ancestor organisms, especially in an explosive mode and give stamp of approval to the present biological classification system that it is scientific and well-grounded.

As discussed in earlier section, an evolution cycle starts when right external conditions, including climate and geologic changes, strike the ancestor organisms. One cycle will bring about millions of evolution trails in a radiant fashion, each of which ends either as an intermediate carrying lethal mutations or an organism qualified as new species, resulting in explosive appearance of new species from a limited number of trails. It can be assumed that novel and variants of phenotypic traits have been generated and preserved continuously in some intermediates through random mutations over the cycle. When new traits accumulate to a sufficient number, they turn the intermediates into new species and the trails reach ends. These early phenotypic traits are heritable, stable, and relatively immune to genetic variation of large degree. In the meanwhile new species establishes unique behavior as they

learn from interacting with the environment by employing their newly evolved first type of traits. Over time, their unique behavior stabilizes into behavior-dependent traits that enable them to survive better. In a sense, evolution concerns only those organisms that are predisposed to evolution, and the evolution cycle adheres to the principle of zero sum rule as the destination of evolution.

All the mammal species have mammary glands that produce milk to nurse their young, while they differ in reproductive strategies and in a number of anatomical structures. They are divided into three groups, monotremes, marsupials and placentals. Monotremes are the oldest mammals, and lay eggs to produce young, rather than bear live young. Marsupials and placentals both carry their fetus in the uterus of its mother, but marsupials bear live young to a relatively undeveloped state and must nurture them within a pouch on mother's abdomen, while placentals bear live young to a relatively late stage of development.

Evolution of mammals can be dated back to the first fully terrestrial vertebrates amniotes, which descended from earlier amphibious tetrapods about 320 million years ago. Within a few million years, amniotes diverged into two lineages: the synapsids, the common ancestor organisms of the mammals, and the sauropsids, the common ancestor organisms of reptiles, and later birds. Synapsids then diverged into monotreme mammals and therian mammals about 275 million years ago, and the therian mammals further diverged into marsupial mammals and placental mammals about 125 to 160 million years ago. All the modern mammals are descended from these early mammal ancestors. According to Wikipedia Mammals, the monotremes, including platypus and echidnas, contain one order, 4 families, and 10 extant species. Marsupials, which include bandicoots, wombats, opossums, kangaroos, are classified into 7 orders, 19 families, and about 334 extant species. Placentals, which encompass the vast majority of extant mammals, are classified into 21 orders, 130 families, and about 5000 extant species, mostly rodents and bats.

When species evolved from amphibious tetrapods to amniotes to synapsids to monotremes, marsupials and placentals, intermediate species must endure a series of grand magnitude transformation in morphology, anatomical structures, development, and the underlying genome sequences in about 160 million years. The two lineages, synapsids and sauropsids, are derived from the ancestor organism belonging to an early species of amniotes during the Carboniferous period. In a simplified view, synapsids, the ancestor organisms for all mammals, is the origin of the mammal evolution cycle. The cycle first diverged into therian trail and monotremes trail, and then the former trail further diverged into marsupial trail and placental trail. All the species descending from continued divergence of these three trails are classified into Mammalia class. Depending on the distance of divergence from the origins of main trails, species that share the shortest common distances are classified into the same order, longer common distances into the same family, the longest common distances into the same genus. For example, if two intermediates diverge early on the trail to start their own trails, all species descending from one trail will be classified into one order, and all species descending from other trail will be classified into another order. The species in these two orders share the shortest common distances. As intermediates diverge closer to the end of trails, they share longer common distances and more common features, and species descending from each intermediate will be classified into the same family, same genus, until the intermediate itself becomes a new species.

The great diversity of mammal species in terms of morphology, physiology, diets, life style, and living habitats is a strong indication of multiple evolution cycles in the evolution of mammals over the last 300 millions of years. Each cycle started at different geologic periods or areas and ended up generating new orders of species, and at the same time leaving some ancestor organisms for future evolution cycles. Mammals that emerge in later cycles are more advanced, not necessarily more complex, in

many aspects than those from the earlier cycles. Available fossil records are limited, but show common, even massive extinction of mammalian species on the scope of entire genera or entire families. Mass extinction has some implications for evolution of species. In the history of evolution, the number of mammal species produced from each cycle are likely far more than the number of mammals living today. Nevertheless, due to the randomness of mutations, many mammals were too delicate and fragile to tolerate environmental changes of different kinds, and consequently perished from future adverse environmental changes. Some species were ill-formed morphologically or physiologically on critical mutations, for example having limbs growing out on the back, rendering them unfit for survival and being eliminated after few generations. Not all new species from an evolution cycle have the same survivability. Only those mammals that are well-formed and able to survive the adverse environmental changes continue to live to present days. Like selection of a liquid chemical that freezes at zero degree C, water is the only selection after a variety of liquids are tested.

Looking back at bats and ants, they are common names for numerous similar species in the biological classification system. They are distributed all over the world except two polar regions. It would be interesting to ask how bats attained their echolocation system to process ultrasonic sounds and how ants developed the pheromone-based capability to communicate with each other. Such capabilities are precise and efficient, but complicated and sophisticated, requiring many genes to work as a whole in an exact coordinated manner to achieve full potency. Furthermore, differences in capabilities among bat or ant species can vary greatly from subtle to large degrees. Any biological system that is as sophisticated, complicated, and diverse as these two is truly awe inspiring.

Bat echolocation calls and ant pheromone communication are claimed to be remarkable examples of good design through evolution by natural selection. Indeed the evolution of bats' echolocation system poses a challenge to natural selection theory. What natural pressures can drive the evolution of such a system? From development standpoint in lieu of evolution, it must be initiated and continued purely by random mutational events, a totally directionless trial-and error process. Random mutations led to a protein pool that could organize into the echolocation system. Some of the proteins in the pool must be novel and preserved specifically for the system, while others were the products of the reuse strategy. Regardless the protein pool couldn't be completed in a short time frame, but over a period of tens of millions of years. Only upon random, protein-wise, continuous appearances of proteins that happened to be needed to complete the pool, the system would develop and mature into what we see today in bats. This is a process solely depending on randomness-based chances. Numerous intermediates failed to be bats in part due to lack of luck to complete the pool. An implication is that the system wasn't clear at all in terms of functionalities at least in the early phase of evolution, thus its development would be utterly unrelated to the purpose of better survival and reproduction.

The genotype changes that natural selection is based on seem too inadequate to lead to such remarkable biological traits. Evolution in the context of evolution cycle obviously provides an elegant answer to explain the origins of these two systems. Assume 3 bat species. The echolocation systems of bat A and bat B were closely related, while bat C's seemed only distantly related. Looking at the evolution trails, an early intermediate from one trail carried mutations X that started the evolution of the echolocation system. After sharing generation X1, the intermediate diverged, from which one intermediate ended up as bat C after total generations X2, while another intermediate diverged again at generation X3. Starting from X3, one ended up as bat A at generation X4, and one ended up as bat B at generation X5. X4 and X5 were not too distant apart in terms of generations, therefore their echolocation systems were closely related as they shared many common mutations toward the final system, while bat C didn't share many mutations with bat A and bat B, resulting in a system that is more different from the other two. In terms of the positions in the cycle, bat A and bat B are close to each other both on the horizontal

dimension and the vertical dimension, while bat C is distant from bat A and bat B on both dimensions. Hence the evolutionary relationship of species can be well exhibited on the two dimensional system. Closer the two species on the two dimensional system, closer the two species in an evolution cycle, and greater their genetic similarity.

Bats' echolocation phenotype must be backed up by a super genotype that is composed of a large family of genes. An ancestor organism descending from placentals starts an evolution cycle that leads to the explosive growth of bats. In the early phase of reshape, one intermediate acquired one or more genes in the ear or in the larynx upon random mutations, whose protein products displayed unusual properties that enabled trachea to emit ultrasound if they were in the larynx, or enabled the ear to respond to ultrasound if they were in the ears. These genes or some other genes must be the earliest members of the gene family that played a pivotal role as the seed genes to initiate the formation of a nascent echolocation system, upon which all the later development was built.

As the reshape process progressed, the gene family grew in size and functionality, and its members scattered in throat, ears and brain, gradually assembling into a system capable of echolocating objects non-visually. In the process it's pivotal to coordinate and integrate various newly formed components into the single, existing system. The bat genomes seemed to contain DNA sequences with a tendency to encode protein components for the echolocation system upon mutations. In evolution it's likely that after some mutations started a process that could lead to a new trait or change an existing trait, this process couldn't be stopped any more, but continued until the trait became part of the new species.

Generally speaking, larger the system, larger the space for the system to grow to more variety in functions and details. Inherent elasticity of the three dimensional structures of protein molecules allows a large system to accommodate different protein variants of the same function, resulting in many different, but closely related species, each of which displays the same system with its own flavor. The echolocation system of bats is such a system with numerous different flavors, which seems to be responsible in part for the unusually large size of the bat orders.

The reshape process diminished and gradually became the healing process until the echolocation gene family could support a fully functional echolocation phenotype. Not only all the proteins necessary for the echolocation functions must be created and assimilated into an increasingly complicated system over time, their encoding genes must be regulated on the level of expression in various tissues in parallel to assure organizational and functional unity of the system. It's hard to imagine the difficulty encountered in the development of such a system, in which the functional operations of the system is guaranteed by a completely matching gene family and its regulatory network. If the development of a phenotype as complicated and sophisticated as echolocation is looked at from the standpoint of zero sum rule, it must have gone through a long period of evolution, in which the mutational effects on the phenotype transit gradually from positive sum changes to zero sum changes, equivalent to the transition of an armed state to a disarmed state in an evolution cycle. In the positive sum state, the net gain of beneficial mutations over deleterious mutations is positive to the system, thus improving the phenotype in functions and capabilities. When the mutational effects reach the zero sum state, the phenotype reaches the optimal state as well. At this time all aspects of the biochemical and cellular network are well balanced with no much space left for further improvement, an inevitable result of the evolution over tens of millions of years. The positive sum process is more like what natural selection refers to, but it progresses at faster pace and without selective pressure due to steady positive effects.

Though all species are experiencing non-stop random mutations throughout time, many codon changes are neutral to protein functions or structures on codon degeneracy. In inevitable cases amino acid

changes could exert more or less non-lethal negative effects on the functions of protein molecules. However as discussed earlier, their ultimate biological effects must be assessed in lieu of an infinite time scale. In the zero sum state, most changes will be neutralized in the long run on function level, leaving some changes that are too small to change the phenotype of species. In strict sense, the zero sum rule isn't referring to protein sequences, but more to protein functions and structures.

A zero sum balance established out of highly entangled processes and activities is vulnerable to mutational changes, tipping the affected biological system into agitation of various degrees. Mutational changes are mostly deleterious to a balanced zero sum state, resulting in a negative net gain. All the species are in the zero sum state at the end of evolution, but the zero sum state differs greatly from species to species in terms of stability and resilience, which is reflected on the tolerance for deleterious mutational changes, which is highly correlated to their ability to return to the zero sum state upon negative impact. If individuals are resilient to negative sum changes, they are the most common species found around the world, such as rodents and bats. Otherwise they are species that are quite susceptible to environmental changes and must be confined to certain niche habitats for survival such as many butterflies and frogs. A fragile zero sum state is too delicate to embrace new changes. Consequently, these fragile species generally would undergo mass distinction when climate or geological changes occurred. In theory, they had adequate space and time to evolve into species like rodent and bats, but they have remained vulnerable in the past several hundreds of million years, an evidence that indirectly disapproves natural selection as a universal mechanism of evolution. On the other hand, randomness of the deleterious mutations limits the changes only to some of the individuals. If they died from it, most other individuals remain normal in the zero sum state, thus keeping them as the common ones in the population, the basis of stable biodiversity. The balanced zero sum state can be tightly tied to the habitat and broken if a species migrates to elsewhere. In the new environments, mutations that are negative in the old habitat can be positive, thus causing genetic variations of certain heritable traits.

The evolution of phenotypes, simple or complicated, has been a trial and error endeavor of the genetic machine over the periods of hundreds of millions of years. It is made possible only when more functionally relevant genes and their protein products emerge and fit well into the system non-stop. Otherwise earlier genes will fade into functionally inactive pseudogenes or even random sequences after barraged by millions of random mutations. This implies that an evolution cycle requires the right mutation rates first to avoid damages by causing too much changes to the organisms and second to avoid lack of sufficient new proteins to sustain the cycle to the end. It could be predicted that the right mutational rates must be far faster than the mutational rates that occur in present organisms. In the same time, zero sum rule suggests that species living on the earth today haven't changed much since they achieved this ceiling state at certain periods in the history. Bats or ants today are no difference from bats or ants that lived on earth million years ago. Evolution cycles increase the biodiversity and the out of cycle zero sum rule maintains the stability of species and biodiversity. Figure 5 summarizes the achievement of the zero sum state in the process of evolution of species and the maintenance of the zero sum state afterwards.

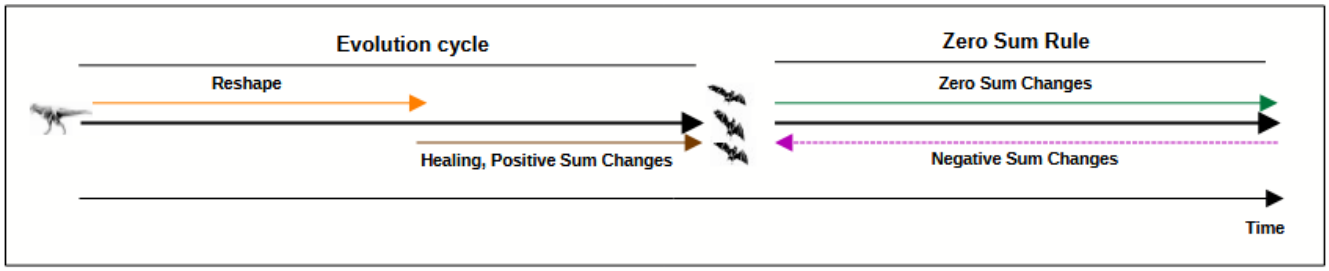

Figure 5. Evolution of species, evolution cycle and the zero sum rule. Any phenotypes in the healing period are not matured yet when measured with the zero sum rule. In the healing process, the net gain of prolonged mutational changes is positive. The healing process ends when the net gain becomes zero, bringing the species into the zero sum state, in which further changes will no longer result in further improvement, but possible deterioration of survival and reproduction, marked as negative sum changes. Therefore, mutational changes are characterized by positive sum during healing, and by either zero sum or negative sum under the zero sum rule. Individuals bearing negative sum changes will eventually die out if they failed to return back, preventing degeneracy of species and enabling the population stable indefinitely, the very foundation of the extraordinary biodiversity on the earth today. All organisms are in an armed state if in the cycle and in a disarmed state if under zero sum state. Mutational rates vary in different states. Far more mutations are required in the evolution cycles for new species to emerge, thus the mutational rates are much faster in the evolution cycles than in the zero sum states.

Everything on the planet earth is covered by sky, and similarly almost everything that is happening or existing on earth can be explained by natural selection as why they look like what they are today. River XX started from a place deep in the mountain range. In its early life, river XX was converged from many small branches, each of which started in different areas in the mountain and flew into river XX when they made way out. As time passed, many of the branches got blocked in their way out and emptied their water into branch X, making it wider. After a long time evolution, branch X continued to widen, accepting most water emptied from other branches. The main branch X flew out of mountain and merged with few remaining branches, forming river XX, the largest river in the area. This is so familiar with how an advantageous trait made the trait carrier more common in the population. Did the evolution of river XX result from natural selection as well? One more example will end this random thoughts. On a flat land stands a big stone, while its surrounding area is covered by small stones. It's evident that these small stones were left there after some big stones eroded by wind, rain or other natural elements. What can be drawn from these small stones is that the toll big stone has resisted the same natural elements that eroded its neighboring stones over years. Its survival seems fit well with natural selection theory. Is there anything that isn't the result of natural selection?

11. Summary and Discussion

Looking at life of all forms that appears along the timeline, evolution of species could be divided into three stages as shown in Figure 2. Stage 1 is dated back to about 4 billion years ago, in which primitive forms of life arise on the nascent earth. Stage 2 is an extraordinarily long period of slow evolution, lasting the next 3.5 billion years, in which life evolves into multicellular forms with slow but steady pace and the genomes of many species have reached moderate sizes at the end of this stage. Stage 3, by contrast, is an extraordinarily short period of fast evolution. Starting from Cambrian explosion, stage 3 has passed 600 millions of years so far, during which life of all forms and complexity has inhabited the earth, boasting an unprecedented biodiversity.

There is every reason to believe that the nascent earth was a life-welcoming planet, in which a place called incubator of life was furnished with a mix of amino acids, nucleobases, sugars, lipids, and other inorganic and organic chemicals. Random chemical reactions occurred spontaneously and constantly in the incubator, generating all kinds of possible chemical products, including polypeptides, ribonucleic acid RNA and DNA. All these polymers were of random nature, varying in length and sequence. As the amount of random polymers increased, some RNA happened to fold into structures similar to tRNA, rRNA and mRNA, and some polypeptides happened to fold into three dimensional structures with rudimentary biochemical properties. Among biochemical properties are preliminary enzymatic activities such as RNA and DNA polymerases, ribonucleotide reductases, and structural components such as ribosomal proteins and other primitive proteins engaged in DNA and RNA synthesis.

With the availability of the low grade protein complex for RNA and DNA synthesis and different kinds of early enzymes, DNA grew longer in a random fashion and even replicated to make numerous copies, while RNA was copied from DNA templates through transcription. Among RNA population, including those transcribed from DNA templates and those randomly polymerized, were RNA precursors to modern tRNA and rRNA. The rRNA precursors complexed with some ribosomal-like proteins and acted as simple ribosomes. When the tRNA precursors which carried an amino acid at 3′ end aligned on the mRNA molecule attached to the ribosomal platform, adjacent amino acids reacted to form peptide bonds with efficiency greater than random polymerization.

In the very early phase of the Stage 1, the life system was constantly changing in all its components. DNA sequences were quite random due to random elongation and error-prone replication, so the RNA and peptides derived from the DNA templates. When peptides were produced from DNA templates, their production became template based, more reliable and fixed in sequence. As more peptides were template based, they overtook random polypeptides in the incubator. More proteins displayed good properties or function as enzymes, ribosomal proteins, structural proteins, trans-membrane transporters, and so on, and furthermore they became slowly available in stable fashion. Gradual appearance of template based enzymes with an increased variety, better catalytic activities and higher specificities brought the early life into an enzyme era, making DNA replication, RNA transcription, and protein translation more reliable and consistent. Meantime, DNA sequences that served as templates for peptide synthesis were slowly transformed into gene-like structures, which further increased the reproducibility of protein molecules. All this indicated that the minimum genomes started to form. Slow but steady improvement and maturation of the basic biochemical machine marked the successful transition of early life away from randomness into consistent and disciplined operations. It’s the infinite randomness at the beginning of life that generated infinite amount of random peptides, RNA and DNA. Only infinite amount of random stuff could serve as a cache of great treasure and made de novo buildup of an all new self-sustaining system called life possible on the nascent earth. No question, life was born out of sheer randomness.

Early life arising from randomness must vary in forms, dependent on the amino acids and nucleic bases available in the environments. Different forms of life were ultimately attributed to the use of different set of codons for protein translation. The earliest primitive cell – single celled life – formed when one particular minimal genome was enveloped in a lipid bilayer membrane. This single celled life relied on a single set of genetic codons corresponding to a single set of amino acids for protein synthesis. It prevailed in the incubator and became the common ancestor of all modern living organisms.

The early single celled life was too simple and too flimsy to withstand any adverse impacts from the environments. It was in stage 2 that life fully developed its biochemical processes, cellular structures, and genetic machine, all of which greatly improved the overall efficiency, reliability, and survivability, successfully metamorphosing into full-fledged organisms.

The genetic system of early life was far from complete and robust, and its genome expansion was basically the continuation of Stage 1, largely random and of low efficiency. Increase in gene number allowed organisms to produce more enzymes of different kinds, which in turn allowed organisms to operate more metabolic pathways and perform genetic recombination with increased accuracy and efficiency. It was expected that numerous forms of life arose in the process due to randomness in the early phase of stage 2. Each form of life was likely to possess a unique set of proteins despite of using the same set of genetic codons. Existence of various life forms made it possible for multiple cells of different origins to merge into a single cell at the prokaryotic time. Merger accelerated the enlargement

of genetic materials, widening the coverage of metabolic pathways, and finally compartmentalizing cellular structures into organelles, all of which were characteristic of an eukaryotic cell. Confinement of genome inside the nucleus ushered in the era of eukaryotic life.

Transition from prokaryotic life to eukaryotic life must be supported by additional set of new proteins, so the same for transition of single celled eukaryotic life to multicellular eukaryotic life. Multicellular organisms are not simple aggregates of cells of the same type, but the aggregates of cells differentiated into different types packaged in a specific way. Data in Table 1 show dramatic increases not only in the genome size but also in protein coding gene counts in selected eukaryotic organisms over selected prokaryotic organisms. Similar data are unavailable for full-fledged prokaryotic organisms and early forms of life, because of no baseline available for comparison, but the data would be expected to be compatible with Table 1. Dramatic increases are expected not only in the genome size but also in protein coding gene counts in prokaryotic organisms over early forms of life. Considering the ultra long period of stage 2 and buildup of multicellular life from the very bare necessities of early life, creation and assimilation of a large number of new genes, including regulatory elements and protein products, seem to be the hardest bottleneck to break for evolution to move forward. Infinite randomness is the means to achieve the data shown in Table 1, albeit at the cost of time. The entire stage proceeded automatically without external intervention. Simply put, infinite randomness created new genes with novel functions, which then incorporated into the cellular machine to make it more complicated and more advanced. Observably the later organisms have more sophisticated morphology that their predecessors, which was exactly what evolution is all about. Infinite randomness is again the sole driving force that has moved evolution forwards from stage 2 to stage 3.

The end of stage 2 is the beginning of stage 3, which started from Cambrian explosion. Stage 3 is completely different from stage 2. Organisms at the end of stage 2 had so far amassed a large number of protein coding genes comparable even with mammals, and as a consequence, the mode of evolution had changed from total randomness to reuse via random mutation based re-polishing of existing genes or pseudogenes, advancing evolution in the form of cycles. Organisms emerged from each cycle could be classified into the same class, a subdivision of a phylum in the biology classification system or into the same order under the same class. Multiple cycles would be necessary for the evolution of species that are classified into a single class.

Not all early organisms were eligible for evolution. Only organisms with special genetic capacity would become ancestors of later organisms. When earlier amphibious tetrapods evolved into amniotes, which further evolved into the synapsids and the sauropsids, amphibious tetrapods remained, because only those amphibians of ancestor nature evolved into amniotes. Natural upheaval was served as perturbations to push ancestor organisms from a disarmed state into an armed state, in which the genetic machine became more error prone than at normal time. Higher frequency of mutations caused genome-wide changes, bringing the ancestor organisms into an evolution cycle. All offspring that born in the cycle were intermediates of the cycle and in the armed state. Upon entering the cycle, the genomes were constantly reshaped by random point mutations, gene duplications, and recombination. Some of the duplicated genes were genetic fodder for deriving new biochemical properties through the action of point mutations. Point mutations continued to refine mutated genes over time, resulting in the creation of new functional properties or phenotypical traits for the organisms. This process healed the genome back to a disarmed state – a healing process. Most of the intermediates died from mutations of lethal nature in the course of a cycle, while lucky ones not only survived but emerged as new species. Evolution cycles bring about new species in explosive mode.

Evolution cycle is an important concept. It delimits the time period from the time organisms begin to evolve to the time new species emerge in a consistent, stable, calm, and disarmed state. As a result when we talk about evolution, we just focus on the evolution cycle. What happens in the cycle is what happens during evolution at a specific time in the history of evolution. Another important thing to keep in mind is that evolution cycle can take up to tens of million years or generation to conclude, which, from a cumulative standpoint, implies that genome wide changes spread over millions or more generations and changes that happen in each generation must be limited in scope. On the other hand, mutation rates must be much higher than in a disarmed state, otherwise, early mutations could be reversed or canceled out by later mutations. The genetic machine of ancestor organisms must have the right mechanisms to maintain a delicate balance between mutational damages and creation of new properties for new species. In this way by keeping mutations at the right rates, it is ensured that there were always intermediates that survived intermittent type of changes and at the same time generating novel features at the right speed to allow new species to emerge.

Reuse based evolution is the most distinguished characteristic of evolution at stage 3. Reuse greatly accelerated the emergence of new species. The essence of reuse is the generation of protein variants. A large majority of protein variants are variants only in different species. For example, enzymes involved in glycolysis in a particular species have no variants, but they are variants among different species because they don't share identical amino acid sequences across species. A small portion of protein have variants not only across species, but also in the same species. For example, opsin molecules sensitive to different wavelengths are variants of each other in the same species, but each of them is also a variant of the same opsin molecule in other species. The later type of protein variants is encoded by genes of duplication origin, including active genes and pseudogenes. Gene duplication, as one type of DNA rearrangement, still occurs in modern day organisms. However, duplicated genes here refer only to those that led to functional protein variants after bombarded with random point mutations over tens of millions of years. Duplicated genes are free to diverge via gradual accumulation of random mutations so long as they are either not expressed or their protein products are not lethal or not fragmented in the intermediate organisms. Because a large majority of protein variants aren't derived from duplicated genes, the gene counts across species are stabilized around 22,000, regardless of the complexity of the species.

The essence of evolution since Cambrian explosion could be viewed from the reuse standpoint. Gene duplication combined with random point mutations is part of the reuse mechanism. In each evolution cycle, some genes are duplicated from master genes from the previous evolution cycle. Most existing active genes, together with certain duplicated genes, will undergo a series of random mutation based sequence changes over millions of years. Only sequence changes that result in functional or structural changes that won't disrupt the normal working of the cellular machine will be preserved and embrace more changes to come. When an intermediate has accumulated a sufficient amount of proteins with changed functional or structural properties, it becomes a new species. As a result, new species from this cycle are usually more advanced and sophisticated than species from the earlier cycles. Simply put, reuse based evolution cycle is to generate a series of protein variants from proteins that are actively engaged in the life processes in ancestor species. When new protein variants, complemented by a certain number of all new proteins, are assimilated into the life machine at different developmental stages, in different cells and tissues according to their expression control, they have reshaped the organisms in a fundamental way, and as a result, the organism has become entirely different and qualified to be a new species.

The changes brought up by protein variants are broad and their impact on the existing biochemical processes and cellular organization can be subtle or far reaching. For example, substitution of some

subunits in a multisubunit protein with subunit variants can change the biochemical properties of the protein in a subtle way, making the protein more fined tuned for the process or structures in that cell types. On the other extreme, variants of some developmental inducing factors can radically change the final morphology of the organisms by tweaking the embryonic development processes. Protein variants can fill the functional void in the old processes or improve old functions in the new species as in color vision of primates, stronger stomach of vultures, more sensitive olfactory buds in some organisms. Protein variants are generated and expressed at different stage of the cycle, and change the visible look of intermediates accordingly. In a nutshell, a large number of genes in the ancestor organisms have been substituted with their variant counterparts at different time points and tissue locations over an evolution cycle, resulting in the re-establishment of a balanced system, in which all components of new and old work together as a single unit just as in the ancestor organism, except the organism is no longer the same as the ancestor organisms morphologically and physiologically. Intermediates that fail to re-establish such a balanced system perish in the cycle.

Evolution cycle will continue until all survived intermediates emerge as new species. New species will reproduce, grow and die indefinitely due to the zero sum rule. The genetic effects of any mutational changes on species are either neutral or deleterious, but barely beneficial on the evolutionary timeline. In a sense, evolution of species is all about evolution cycle and zero sum rule. Evolution cycles proliferate species, while the zero sum rule maintains the stability of existing species.

Vision of the fruit fly comes from its compound eye, which is composed of 760 unit eyes or ommatidia, each of which is a tiny independent photoreception unit that consists of a cornea, lens, and eight photoreceptor cells (R1-R8). The R7 and R8 cells each comes in two subtypes R7p and R7y, and R8p and R8y, respectively. These subtypes form strict R7p and R8p pair and R7y and R8y pair. Comparing with the vision of animals high on the evolutionary ladder, fly vision is rudimentary but unique in its own way for mating, navigation, foraging, avoiding predators, etc. The fly genome encodes seven opsin molecules Rh1 to Rh7, each of which is sensitive to light of different wavelengths. Rh1 absorbs maximally blue light (~480 nm), Rh2 absorbs maximally violet light (~420 nm), Rh3 absorbs 345 nm light, Rh4 absorbs UV-light (375 nm), Rh5 absorbs light of 435 nm, Rh6 absorbs light of 508 nm, and Rh7 absorbs maximally 350 nm light. In vivo spectral sensitivities differ due to the presence of sensitizing pigments or screening pigments. For example, the opsin R6 shifts from 508nm to 600 nm in vivo. Each of the opsin genes is expressed only in one photoreceptor cell. The opsin Rh1 is expressed in photoreceptor cells R1-R6. Rh3 expressed in R7p cells, Rh4 expressed in R7y cells, Rh5 expressed in R8p cells, and Rh6 expressed in R8y cells. In addition to expression in photoreceptor cells, Rh2 is expressed in the extra small eyes called ocelli, and Rh7 is expressed in the central pacemaker neurons to regulate the circadian rhythm of the fly. Phototransduction pathway in fly photoreceptor cells is as complicated as in vertebrate vision via a G protein-coupled pathway. Light stimulation elicits a conformational change in the Rh molecule, turning it into an active form. Metarhodopsin activates Gq, the fly version of vertebrate G protein. The pathway starts.

Because of compound nature, fly eye absorbs light through each ommatidium in slightly different angles and produces an image that is a combination of numerous unit images from each ommatidium. When an object in the view is moving, or the fly itself is flying, the light that enters photoreceptor cells changes continuously in intensity, turning the light signals in the ommatidia on and off. This on and off signaling effect creates flickering in the brain, the frequency of which is the rate at which ommatidia are turned on and off. In this way the fly can detect and respond to movement in extremely fast fashion. About two-thirds of the fly brain is dedicated to visual processing, implying the importance of the compound eye for the survival of fly. Truly the image is a very wide angle view that enables the fly to detect fast movement in surroundings for protection despite poor resolution.

Assume Rh6 suffers from mutations that shifted its absorption peak from 600 nm to 700 nm. Because Rh6 expressing R8y is strictly paired with Rh4 expressing R7y cells, absorption change in Rh6 will disrupt the interaction between R7y/R8y pair. As a matter of fact, changes in light sensitivities of any of these opsin molecules could potentially tip the balance among photoreceptor cells and transmit altered visual signals that likely result in wrong interpretation by the central nervous system. Mutations to Gq proteins can impede, even break the phototransduction pathway. All the mutational changes can be considered as negative if not neutral, breaking the well established zero sum state of the fly visual system. What's more? Assume a new opsin protein with peak absorption at 700 nm emerges from a duplicated gene. For this newcomer to become a functional part of the compound eye, it must have its own expressing photoreceptor cells and its own interpreter neurons in the central nervous system. It must coordinate fully with the existing eye components without any incompatibility. Therefore, to simply accommodate a new additional opsin molecule, the compound eye must undergo quite large changes on all levels, even on morphological level. Under the zero sum state, the possibility to gain some feature upgrade seems to be prohibitively small even over periods of hundreds of millions of years, periods that might be far longer than the time for an evolution cycle. When a new opsin protein with peak absorption at 700 nm appears in a fly, the fly is no longer the fruit fly, but a different species with another set of compound eyes or completely different eye after a new evolution cycle.

The fly compound eye is an intricate system of multiple dimensions, but the citric acid cycle of the fly is a simple one dimensional metabolic pathway carried out by eight enzymes. If a mutation in one enzyme increases its catalytic activity, the overall performance of the pathway remains unchanged. Only when all the enzymes have improved their catalytic activities, the overall performance of the pathway will improve. However, it’s quite unlikely that this will ever occur regardless of time outside an evolution cycle, as the cancellation effects will maintain the status quo of the pathway. On the other hand, mutations that lower the catalytic activity of an enzyme can easily exert negative impact to slowdown the overall performance of the pathway, which is the basis of the rate limiting enzyme in regulation of a metabolic pathway. Therefore, it’s easy to break a state, but hard to improve a state. In a zero sum state, individuals that suffer negative changes will disappear from the population either through death or long term cancellation of the negative effects.

Above discussion indicates that all the components of the fly compound eye have been refined and honed painstakingly on molecular and cellular levels and organized into a precision biological machine, the visual system for the fly in an evolution cycle. Generally speaking, a biological machine is the assembly of all its components in a strict sequential order or spacial arrangement in such a way that the machine can execute its functions via exact and ordered interactions of these components. It is stable because only machine-wide changes can have the potential to advance its performance. It is fragile because most non-neutral changes will disrupt the balance among its components and put it into a disadvantageous state. The mutational effects on biological machine are severely constrained by its two sides – stability and fragility. The zero sum rule is to maintains the stability of the phenotype of species outside evolution cycles.

The fly genome contains 7 opsin genes, encoding 7 opsin molecules with light absorption peaks in close ranges, especially in the presence of sensitizing pigments. It is interesting to ask if the fly really needs all these 7 opsin molecules just to create low resolution images and produce flickering effects to detect movement. If we look at the development of vision phenotype from an evolution standpoint, we can conclude that visual development on the evolutionary timeline agrees well with the meaning of evolution. It improves continuously as organisms move up the evolutionary ladder. As it has been pointed out repeatedly, evolution means that the development of a phenotype is achieved through the

most time consuming and most wasteful trial and error approach. In the evolution of vision, it is never known in advance that how many opsin molecules would be enough to cover the light from the visible electromagnetic spectrum, and how complex a visual system should be in order to capture light and then convert it into neural signals for images. It's highly unlikely that all these 7 opsin molecules are needed for the fly to achieve such a primitive vision, but it's one of the numerous stages through which more advanced vision systems are developed by dropping off unnecessary components and accepting improved components for better biological performance.

The vision system is the most complicated biological system, reflecting the formidable difficulty in this quest. Complexity begets variety, thus visual system enjoys the most variety in morphology and the underlying molecular and cellular mechanisms in the invertebrate kingdom. Most invertebrates have some form of eyes, but different species has their own unique forms of eye to perceive light, color, distance, movement, preys, danger, etc. well fitting their specific living environments. Such a variety of visual system in invertebrates imply their unique and independent evolutionary origins, at least imply a remarkable divergence of the rudimentary visual functions at its very early stage of evolution. Despite simpleness in function, all these vision systems are obviously made possible only through evolution by cycle mechanism. Once a unique eye is formed for a species, it stays as is indefinitely. The species has to live with it without chances for it to evolve better, regardless of its extremely limited functionality. Moreover, most of these primitive forms of eye are simply the trial and error versions that appear in some stages of vision evolution, and their underlying mechanisms are not suited for further vision evolution. Because of this they all disappeared along the evolutionary timeline. Indeed they are the failure examples of trial-and-error strategy for the evolution. The eye conforms to a more common morphology since fishes while the underlying biochemical and cellular machine remains quite diverse.

The fly compound eye is amazing and fascinating and well suited for the fly's life style, but it is too simple to be comparable in every aspect with the vertebrate vision system. Nevertheless, it is far more complex and advanced than eyes of numerous lower organisms, such as garden snails, mosquitoes, mantis shrimp, worms, etc. which shows a clear evolutionary track. With the available genome data from numerous species ranging from low to high on the evolutionary ladder, a particular biological system can be compared in different species across animal kingdom to study how a particular phenotype evolves over time from the primitive forms to the advanced forms on the molecular and cellular levels. Obviously the visual system seems to be one of those systems for the study.

Life is the sum of many distinguishing phenomena that occur in an organism, especially metabolism, growth, reproduction, and adaptation to environment. The mechanisms to establish these distinguishing phenomena are different for organisms in three evolution stages although all are driven by randomness. A brief summary for each mechanism will be given below to end my random thoughts on evolution.

Life originates in the life incubator, an imagined environment on the nascent earth, a chemical reservoir full of components vital to life. A variety of chemical reactions, especially polymerization reactions, occur in the incubator under the seemingly fortuitous conditions and in random fashion. Among all the polymerization products are DNA, RNA, and proteins of random sequences and lengths, forming a vastly heterogeneous populations. A tiny portion of the random macromolecules happen to possess biochemical activities that can serve as enzymes, structural proteins, tRNA, rRNA, mRNA, and DNA. When DNA molecules contain sequences that encode some of these molecules, the template-based production of macromolecules starts to appear. It's the collection of these macromolecules that turn the earliest weak life-like activities into concrete and significant components of life in the incubator. The self-organizing nature of macromolecules allow them to assemble into superstructures like ribosomes, transcription and replication apparatus, which are then fully enveloped into cell members to become the

most primitive single-celled form of life. Randomness generates possibilities and possibilities increase with the increasing degree of randomness to generate endless possibilities – the chemical basis on which life arises.

The appearance of single-celled life signifies that the DNA genome contains all the information needed to support the continuous existence of life. Changes to the genome change the information contained in the genome, which changes the sum of the distinguishing phenomena of life, the molecular basis of both slow and fast evolution. All changes to the genome are of random nature and confined to the DNA sequences in the forms of base substitutions, deletions, insertions, and recombination. This is in sharp contrast to the unlimited production of random macromolecules on random movements of basic chemical components in the life incubator. How does the slow evolution differ from the fast evolution? The difference can be clearly illustrated in an instinctive way.

The wooden brain teaser puzzles are toys designed to amuse by presenting difficulties to fit different shaped pieces into a space defined by a boundary. The difficulty of a puzzle increases as the number of pieces and the similarity of shapes increase. Imaginably the makers of puzzles first design all pieces for a puzzle on a piece of paper by dividing the space into a pre-determined number of different shapes and then turn the design into wooden pieces accordingly using cutting tools and wood. The challenge to design such a teaser puzzle is obviously quite limited. On a sudden whim, one maker attempts to make a puzzle by making pieces of random shapes, hoping that some of these pieces can be assembled into a whole, which has a clear boundary like any puzzles on the market. Can this maker succeed? If he succeeded, how many pieces must he make in order for them to contain a few that could fit together to form a puzzle? A rough estimate would be at least a few millions. Such an attempt will become easier if the maker is experienced and the shapes for the puzzle are relatively simple. But it can become far more difficult if the shapes are a little more complicated and if the maker has to learn first how to prepare the materials and then how to make pieces. This provides the most direct analogy for what the evolution really is. In a sense, evolution is to make all kinds of puzzles, each of which consists of a set of different shapes to fulfill one specific task that supports one of the distinguishing phenomenon of life. How these puzzles are made is different in the slow and fast evolution.

In the early phase of slow evolution, the genome is small in size and contains very limited number of genes just sufficient to sustain the continuation of the most primitive life. As genome size increases by getting more and more DNA of random sequences, the task to make additional puzzles becomes possible. At this point in time, the task faces two difficulties. First, a random fragment of DNA must be turned into a gene, and second, the gene must encode a protein that has its unique functional place in the puzzle. Taking glycolysis pathways as an example, in which ten enzymes are the ten shapes for the puzzle. It can be imagined how onerous it will be to convert a random piece of DNA into a sequence called gene, which codes for a protein molecule of desired function by means of random mutations. It can be imagined that by means of random mutations alone it will be exponentially more onerous to develop ten genes from random DNA sequences to encode ten proteins, which, when getting together, will fulfill a complicated series of biological functions to turn glucose into pyruvates. Nevertheless, glycolysis pathway is relatively simple comparing with photosynthesis, citric acid cycles, aerobic respiration, etc. Taking this and that difficulties into consideration, it isn’t surprised that it takes 2 billions of years for the primitive life to be armed with many more biological puzzles and develop and mature into full-fledged prokaryotic life. Such a slowness continues into the time of early eukaryotic life, as far more puzzles are required to convert prokaryotic life into eukaryotic life. Although novel shapes for novel puzzles continue to emerge from random DNA sequences throughout the slow evolution, the slowness starts to ease after eukaryotic life has gained larger genomes and accumulated many more genes. At the end of the slow evolution, eukaryotic life has manufactured most of the

puzzles that more advanced organisms must have, including muscle, skeleton, nerve, digestive track in the forms of rudimentary tissues or organs.

More accurately, slow evolution lays the foundation for the reuse of existing puzzles to either derive more novel puzzles or improve to make them more advanced and sophisticated. Random mutations change the functional properties of proteins via their encoding genes, thus changes to the genes change the arrangement and interactions of all the protein components in the puzzles, resulting in changes in functions, structures, performance, and more. Addition of new protein molecules into the puzzles, despite far more difficulty, can expand the functions and complexity of the puzzles, and possibly transform some of them into novel puzzles. Although random mutations are the sole catalyst for all the changes in both slow and fast evolution, reuse has significantly reduced the effort to bring beneficial changes to the puzzles through modification and refinement of shapes that have been proved to work and fit well in the puzzles. Overall, reuse is a sound mechanism in the fast evolution to make function improvement and derivation easier and far more efficient.

It's quite clear now how puzzles are made in slow and fast evolution. In the slow evolution, all the materials that are used to make shapes are still raw. It must be prepared into materials that can be used to make shapes before manufacturing into different shapes. In the absence of knowledge about how to prepare and manufacture, these two steps must be completed via painfully long trial-and-error random processes, taking infinitely long time with high possibility of failure. Because of the trial-and-error nature, some good early shapes can be lost by re-shaping into unfit by the later trials. A more arduous problem is that it's in the absolute dark as what kind of shapes the evolution wants to make and what's the boundary into which the shapes will be fitted. Therefore, all these efforts are totally aimless and erratic, making the painfully long trial-and-error processes more painful and more infinite in time. To turn this seemingly unlikely puzzle making process into reality, most of the shapes must be shapes-in-waiting, and produced continuously in the factory. All of a sudden, as the number of shapes have accumulated to form a pool, a puzzle is born automatically when a few shapes in the pool fit together into a group with well defined boundary and market value. This puzzle is preserved and manufactured indefinitely thereafter. For example, when all the ten shapes that make up the glycolysis puzzle become available in the factory, the glycolysis puzzle is born without surprise and manufactured continuously and indefinitely from ten templates in the genome. Making puzzles in the slow evolution is the most exhausting, most time consuming, and most aimless trial-and-error processes in the universe. But there is no other feasible alternatives to replace it.

With the availability of a considerable number of puzzles in the factory, puzzle making business has become a different process, although it is still of trial-and-error nature. In the fast evolution, all the old puzzles are base puzzles from which new puzzles are derived through modification and replacement of old pieces and acceptance of some new additional pieces. If one piece assumes a slightly altered shape, other pieces in the puzzle can change shapes accordingly to allow it to refit into the boundary. If additional new pieces are to be added to the puzzle, again other pieces can reshape in a little more dramatic way to arrange room to accommodate them. In lieu of evolution, the derived puzzles are usually more refined, sophisticated and complicated in functions for tasks. This puzzle reuse strategy becomes increasingly difficult when puzzles are made from an increasing number of pieces. For this reason, there are fewer complicated puzzles on the market. Over tens and even hundreds of millions of years, old puzzles have been continuing to endure shape changes and at the same time used as templates to derive numerous next generation puzzles, many of which have been so different from the puzzles they are based on. Evolution of the puzzles parallels the evolution of puzzle manufacturing factory. This approach to make more advanced puzzles in more variety is still exhausting, time consuming, and aimless, but the randomness is only a small fraction of the slow evolution. All the

materials are ready for reuse, puzzle boundary is largely defined and relatively easy to extend and adapt, and manufacturing tools are better as well. What’s left is to refit the pieces into the existing puzzles or form new boundaries as new puzzles through random mutation based trial-and-error processes, On the other hand, once the puzzle is created, it is infinitely stable and resistant to all changes that are below the threshold of evolution. Are you surprised by the speed at which fast evolution has brought about new species into the world of life in a short period of time?