

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 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 genomes were 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 unusual 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 the large magnitude of mutations, which threw the ancestor organisms out of the stable disarmed state and entered an unstable armed state. In armed state, the process genotype reshaping 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 reshaping 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 lasted 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.

Table of Contents

- 1. Introduction**
- 2. Life Timeline on Earth**
- 3. Origin of Life – Randomness Brings Life to the Nascent Earth**
- 4. Slow Evolution – the Quiet 3.5 Billion Years before Cambrian Explosion**
- 5. Protein Variants and Evolution**
- 6. Fast Evolution – New Species Arise in Evolution Cycle**
- 7. Fast Evolution – A Closer View**
- 8. Rethinking of Natural Selection**
- 9. Learning and Evolution Cycle versus Natural Selection**
- 10. The Zero Sum Rule and Evolution Cycle**
- 11. Life Incubator and Nascent Seas**
- 12. Genotype Configuration and Genotype Reconfiguration**
- 13. Genotype Reconfiguration, Evolution Cycle, and Zero Sum Rule**
- 14. Genotype Potential Energy and Zero Sum Rule**
- 15. Summary and Discussion**

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 a 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 on the earth today is so rich in forms ranging from simple bacteria and archaea to highly sophisticated 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 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.

Evolution from simple to advanced is the inherent property of life from the very beginning because the DNA genome displays dual distinguished characteristics that are fundamental to life – stability and mutability. The genome is the longest lived biological molecules, passed down from their 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 randomly and constantly, especial during geological and climate changes. Mutability is the foundation of evolution, a process that is responsible for the proliferation 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 most common and completely random, referring to single base deletion, insertion, or substitution. Point mutations are a type of replication errors. Mutations also include deletions or insertions of short pieces of DNA sequences. The large magnitude of mutations 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 rearrangements that change the orientation or location of a segment of DNA in the genome. Deletions of large chromosomal regions can lead to the loss of 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 account for most of the mutations introduced by DNA polymerases during germline division and become 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 thus the stable biotic world. Gene duplication is a process to make a new copy of DNA fragment that contains one or more genes, a special type of DNA rearrangement. Gene duplication is a major mechanism that the genome generates new genetic materials 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 in the past 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 what evolution of life is all about and how evolution of life has gone through from the history point of view. This paper also challenges the verity of the theory of natural selection as a cornerstone of evolution.

2. Life Timeline on Earth

Primitive life on earth can be dated back to 4 billion years ago, about 500 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 dedicates 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 500 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 begin 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.

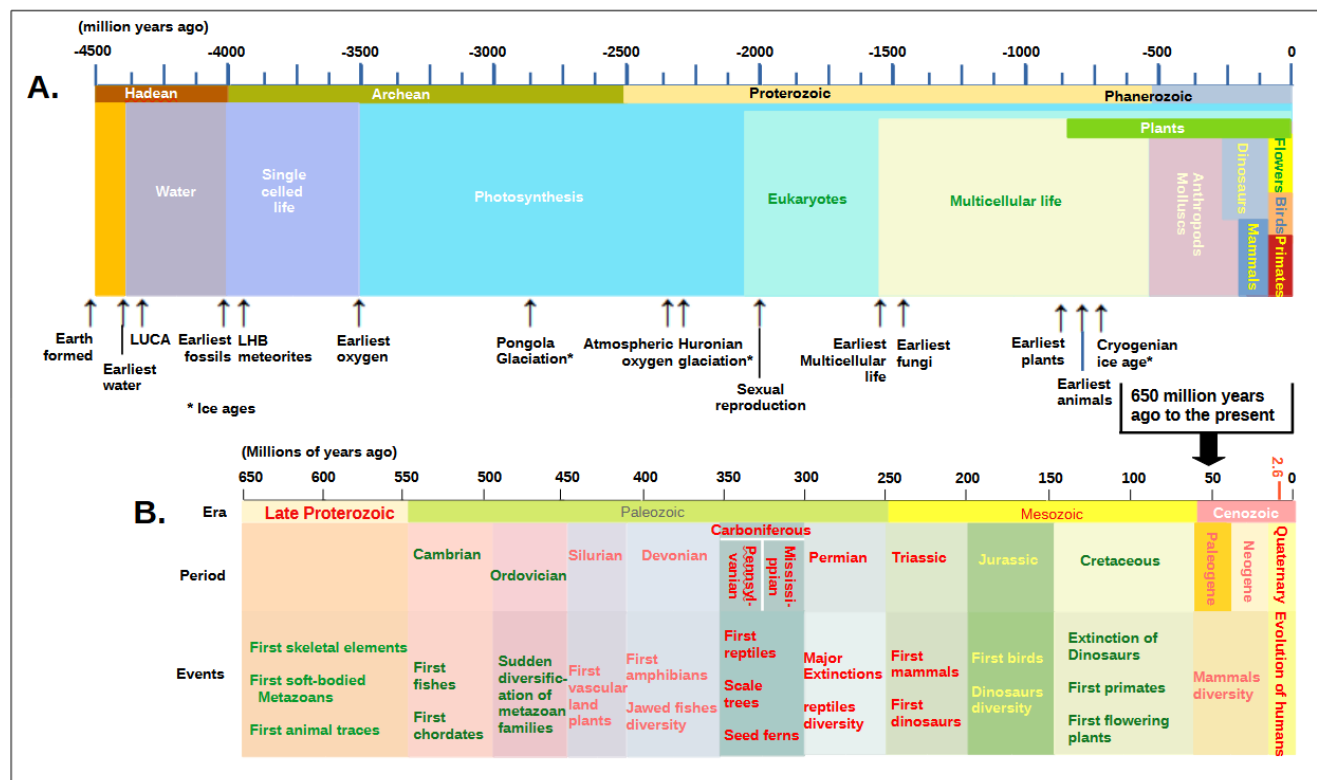


Figure 1. Timeline of the evolution of life on Earth (Adopted from Evolution on Wikipedia and Britannica). Geologic period Phanerozoic comprises the Paleozoic, Mesozoic, and Cenozoic periods

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 appeared in Cambrian explosion period.

In the Devonian period from 419 to 359 million years ago, arthropods insects, spiders, centipedes, etc. became part of the land ecosystem, and more vertebrates moved to the land as well. In Cretaceous period from 145 to 66 million years ago, numerous species of mammals, birds, and flowering plants appeared. In this period, first primates emerged and all dinosaurs went extinction. The last 66 million years were marked by the dominance of mammals, birds, and flowering plants. More insects, moths, butterflies, 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 were 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 the primitive life system from the basic chemical components over a period of 500 million years. This stage is not considered as evolution per se, but origin of life. The evolution process commenced only after life had formed. 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 has occurred. The entire evolution process is the adventure of life that has lasted 4 billion years. In this unthinkable long period, about 85% of the time has been devoted to the slow evolution, while only 15% to the fast evolution.

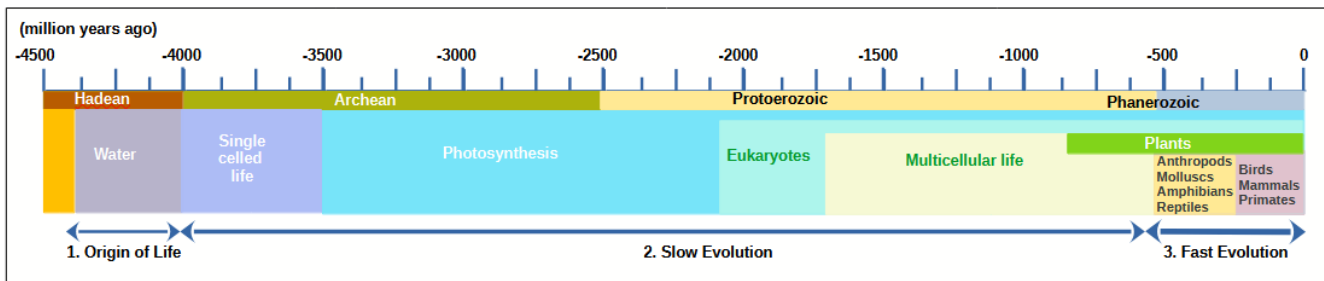


Figure 2. The evolution is a three-stage adventure, each later stage relies on the earlier stage.

Uneven distribution of evolutionary events over the evolutionary timeline is intriguing. Why must the slow evolution stage take about 3.5 billions of years to move single celled life only to low and simple multi-cellular aquatic plants and animals, while the fast evolution stage took only 500 millions of years,

especially the last 100 millions of years, to flourish life into millions of species of all complexities and forms? What is hidden behind this evolutionary timeline?

3. Origin of Life – Randomness Brings 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 in all forms of life. 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 that produce proteins, RNA and DNA, possibly aided with unknown non-enzymatic 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, the 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, higher the possibilities of enzymes with a wider variety of catalytic activities. Enzymes would do catalytic work to accelerate chemical reactions whenever their substrates were available, igniting all possibilities for life. The debut of enzymes and structural proteins must have had profound impact on the production of all types of simple and complex molecules in the very early stage in the incubator, bringing up the idea that enzymes had played decisive roles in the origin of life.

Ribonucleotides are composed of three totally different small molecules – a nitrogenous base, a pentose sugar ribose, and a phosphate group. Therefore, from pure chemistry point of view, ribonucleotides are more complex than amino acids in terms of chemical composition and structure. The ribose molecule exists in various configurations in solution, and only its β -D-ribofuranose form is found in RNA, which is relatively low in abundance. When a ribose molecule accepts a base at 1' position, it becomes a ribonucleoside, and the ribonucleoside reacts with a phosphate group at the 5' position of the ribose, it becomes a ribonucleotide. Riboses that carry bases and phosphates at other hydroxyl groups are not ribonucleotides for RNA. An implication is that the amount of ribonucleotide in the incubator would be insufficient to warrant RNA synthesis in any way. When polypeptides were produced in large quantity early in the nascent incubator and some of them folded into unique three-dimensional structures with catalytic activity for ribonucleotide synthesis, RNA production could become possible at least in terms of the available amount of ribonucleotide. A likely scenario would be that the emergence of a large random polypeptide pool was a prelude to the emergence of nucleic acids.

The chemical reaction that links ribonucleotides into a polymer in the strict order of 3'–5' orientation isn't thermodynamically favored in the absence of enzymes, if not impossible. One possibility is that ribonucleotides could adhere to some special surfaces in the incubator. If ribonucleotide molecules that laid on the surface were close enough, adjacent ribonucleotides could form 3'–5' phosphodiester linkage. The reactions could continue infinitely and produce RNA molecules of various lengths and base compositions. RNA molecules could replicate in similar fashion except that the complementary bases might be snapped into positions on the RNA molecule serving as a template through hydrogen bonding, the result of which is a complementary chain that forms double stranded RNA with the template. Like other random polymerization reactions, RNA production of this type was low in efficiency. The situation changed when the peptide pool happened to generate enzymes that could catalyze the formation of 3'–5' linkage, making RNA synthesis more efficient. These peptide pool based enzymes were rudimentary in catalytic activities and short lived, but were critical for the life to begin from the ground zero. Comparing with modern RNA polymerases, they were merely RNA synthase-like enzymes that incorporated random ribonucleotides into a polyribonucleotide chain in a random sequential order, producing RNA molecules of infinite lengths and base compositions, with or without templates.

The deoxy form of ribonucleotides – deoxyribonucleotides – is more stable and better suited to serve as the genetic materials. In all modern living organisms, production of deoxyribonucleotides from ribonucleotides requires an extra reaction catalyzed by the ribonucleotide reductases, in which the 2' hydroxyl group of the ribose is reduced into a hydrogen. Reduction of ribonucleotides in the ancient time could be possible without enzymes, for example 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 the random enzyme pool, enabling production of deoxyribonucleotides almost as early as ribonucleotides. Similar to RNA, DNA synthases might be lucky ones in the peptide pool as well, generating DNA molecules of random length and random base compositions in appreciable amount. It was even likely that the same enzyme served as the synthase for RNA and DNA generation as the enzyme wasn't good enough to distinguish deoxyribonucleotides from ribonucleotides. It's pure speculation, but any possibilities are possible in the face of a magic life incubator full of randomness on the mysterious nascent earth.

The early appearances of proteins, RNA and DNA could be independent events, but it's far more likely that the proteins came into existence first to form a large random peptide pool, in which some random peptides transformed into early enzymes of various activities that 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 join the rank of life due to the extra reduction step, DNA was more than likely to be the 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 be the greatest source of an extraordinary variety of useful molecules with biochemical significance if the random pool is large enough. 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 many crucial protein molecules, including a variety of enzymes with different specificities, among which were the rudimentary RNA polymerases, DNA polymerases, aminoacyl tRNA synthetases, ribonucleotide reductases, and many more. Even if these early forms of polymerases were largely lacking high specificities and only able to add substrates to the 3' or 5' ends in a totally random fashion, they enabled polymer chains to grow faster, thus greatly accelerating the

expansion of randomness of RNA and DNA sequence populations. In addition, other enzymes made basic life components like amino acids, ribonucleosides, ribonucleotides, deoxyribonucleotides more readily available and in larger quantity via biosynthesis from more basic chemical components present in the incubator. Over 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 is ready 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. The RNA without signature secondary structures could be the earliest forms of 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 meaningful complex formed in this way would be most likely the protoribosomes, a precursor to ribosomes 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 synthases 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 it made synthesis of RNA and DNA no longer completely random, but catalyzed by enzymes on a crude platform. The incubator so far had established itself as the common home for proteins, RNA and DNA, a scenario of life in its early embryonic stage.

Biosynthesis of macromolecules must have occurred spontaneously as necessary components appeared in the incubator. A tRNA molecule would be armed with an amino acid when its hydroxyl group of the 3' end formed an ester bond with the carboxyl group of any amino acid, a reaction 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 charged with random amino acids to align themselves along the mRNA template without much specificity. The complex so assembled would be the most primitive form of peptide synthesis platform, the rudimentary precursor of modern ribosomes, but it was a giant step forward in the origin of life.

The heterogeneity in RNA population produced initially without templates was enormous, and it would be augmented immensely further later by DNA template-dependent RNA production even if DNA wasn't a contemporary fellow of RNA at first. The presence of template-independent DNA synthases allowed the DNA populations to grow more rapidly through synthesis of new chains and elongation of the existing chains via randomly incorporating random deoxyribonucleotides at the 3' or 5' ends. On the other hand, DNA could replicate itself similar to RNA replication. The replication processes, regardless of their mechanisms, were awfully egregious. Assume there was a particular DNA molecule in the incubator. Every replication process would introduce a considerable amount of mismatches into the DNA template, quickly turning this grand parent DNA molecule into a heterogeneous DNA population, which resulted in an even far larger heterogeneous RNA population after being transcribed into RNA molecules. Since template dependent RNA population continuously mixed into the templateless RNA

population, the random RNA pool increased significantly in size, raising the possibility of producing more varieties of potentially useful proteins.

Life is not a random existence per se, but an unusually ordered and consistent living entity. Nascent life must move out of randomness and establish 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 had been made possible when ribonucleotide reductases appeared in the pool to produce deoxyribonucleotides in quantity. The DNA based randomness served as a firm ground on which randomness diminished as it was gradually replaced with ordered operations.

When we talk about modern aminoacyl tRNA synthetases, RNA and DNA polymerases, ribosomes, they are not simple protein molecules, but complexes formed from different protein components that are aggregated into special structures. Nevertheless, early forms of those complexes must be much simpler. DNA replication, RNA transcription, and protein translation, which were among the most basic, but also the most complicated biochemical processes in all forms of life, relied on those early simple complexes to accomplish their roles. From the evolution standpoint, those processes must be among the earliest processes to establish before life could develop further. The coexistence of those protein complexes, together with few accessory protein factors, enabled them to form the earliest super protein synthesis complex capable of DNA replication, RNA transcription, and protein translation. From early on this complex had performed the grand-old function that generated RNA intermediates from DNA templates and converted these intermediates into protein molecules. However, it must be functionally rudimentary, barely capable of linking amino acids or nucleotides into polymers and nothing more.

The appearance of the protein synthesis complex allowed the same polypeptides to be produced from the random DNA templates in a much more dependable way. As the transition of randomness moved on, the polypeptide pool boasted a wider spectrum of biological functionalities, including enzymes of more varieties and superior quality, structural components, regulatory protein factors, and so on. As a result, some components of the complex were replaced by proteins that outperformed and also novel components were added to make the complex more adequate in functionalities. It was the infinite time that allowed the simple, rudimentary protein synthesis complex to develop slowly into sophisticated and quite efficient protein synthesis machine that has been operating the most critical part of life since the origin of life. An improved protein synthesis machine had made protein production even more stable and dependable. On the other hand, proteins produced through old random polymerization played only diminishing roles until they disappeared from the processes. DNA independent randomness thus faded away altogether in the path to single-celled life.

As the protein synthesis machine evolved over time, it became increasingly capable and complex in structure and functionalities. Assembled from a large number of enzymes, structural proteins, and special protein factors, it was a super system responsible for DNA replication, RNA transcription, and protein translation, albeit still much immature. Such a system was the minimum fulfillment of the basic genetic information flow from DNA to RNA to proteins, a prerequisite for life. 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, template-independent and subsequent template-dependent, that complemented each other early on to generate enormous pools of random polypeptides in the life incubator that contained a large number of functionally active proteins, allowing early feeble life activities to appear. Without the initial random polymerization of amino acids and ribonucleotides, it's impossible to successfully build up a DNA template based genetic system for

stable and reliable protein production to keep the same biochemical processes go on indefinitely, and so that it's impossible for life to come into being.

Numerous chemical reactions in living organisms, including reactions that 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 was made possible before the appearance of adenosine triphosphates or ATP. One possibility would be that there were unknown non-enzymatic catalysts in the incubator that could move these reactions forward yet slowly with some form of energy input. A more likely scenario would be that enzymes for ATP production appeared early in the incubator. ATP as a metabolic 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 utilize basic chemicals in the environment for the biosynthesis of amino acids, lipids, bases, and sugars.

As DNA molecules grew infinitely in length and replicated 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 minute characteristics of their modern counterparts, they had established themselves as dependable genetic materials, the true DNA based genetic machine in its primitive forms. Most importantly, in such a genetic machine, proteins, RNA and DNA existed not as independent ordinary large molecules, but as inseparable parts for life. They interlinked in a single system in which one's production became impossible in the absence of other two. As the genetic machine evolved over time in form, accuracy, efficiency, and complexity, the mutational rates during DNA replication were greatly reduced, thus smaller chances of randomness in the genetic information flow. All this obviously accelerated the emergence of life as a self organizing living system.

Enormous heterogeneity of the DNA populations in the incubator implied that many functional protein molecules must be produced from different DNA templates, a huge problem for life as an integral entity. The appearances of DNA ligases and enzymes for DNA recombination allowed multiple DNA molecules to be linked or combined into larger ones. Some DNA molecules emerged as all-potent ones after a number of DNA sequences, each of which harbored a rich set of enzymes and structural proteins, were linked into single ones. They were much larger in length and showed some more structural characteristics analogous to small genomes. These genome analogues slowly grew into longer size by accepting more DNA sequences at their two ends, and their gene-like loci developed into an array of genes with regulatory features that could perform the very basic functions vital to the primitive life. The genome analogues finally transitioned into minimal genomes when they became self-sufficient to sustain themselves. A minimal genome must satisfy the minimal requirements to contain genes that would encode all proteins and RNA elements necessary to support the complete genetic information flow from DNA to RNA to proteins.

As the minimal genomes continued to expand, they hosted a growing list of functional proteins, including enzymes, ion transporters, proteins for cell division, structural protein filaments for the cytoskeletons of cells, etc. Nascent metabolism pathways emerged to start energy generation from carbohydrates and produce key chemical compounds for building basic cellular structures, especially for cell membranes and cell walls. The self organizing nature of proteins allowed some special protein molecules to work as a complex and perform the same functions in an open as well as in a closed system. For example, if a complex consisting of five proteins could pull some super protein structures

apart into two parts in a cell-free system, it could divide the same super structure apart into two parts inside a cell. If a protein complex embedded in the lipid bilayer could transport sugar molecules across the lipid bilayer, it would transport sugars across the cell membrane as well when embedded in the cell membrane. The magic moment finally came as the minimal genomes were engulfed into lipid bilayer membranes, forming the earliest primitive cells – the single celled life. The cell membranes established a closed micro-environment in which the genetic machines would perform their functions much free from the interference of random chemicals in the pool, and the metabolism pathways would produce energy and biochemical compounds largely shielded from free diffusion.

Not all primitive cells were born equal because of the heterogeneity of the minimal genomes, and some primitive cells flourished better than the others. The primitive cells with superior genetic machines gradually dominated the cell populations, and one cell became the ancestor of the most common cell population among the single celled life at the time. Through faster division and wide spread, this primitive cell population finally monopolized the early world of life. Life that descended from it shared the same set of amino acids and genetic codons for protein synthesis, and the same set of bases for the genetic materials. It further diverged into all types of cell populations through random mutations while keeping genetic codons constant. Today all forms of modern life are proud of the descendants of this grand old single celled ancestor.

Primitive single celled life was far from complete from the standpoint of a species. Randomness that occurred to the genomes remained to play major roles in generating more novel functional proteins, not only for the genetic machines and metabolic pathways, but also for cell structures. As more random proteins emerged over time, some of them would carry functions that made single celled life more complex and self-sustainable, allowing early life to develop and diverge continuously, until it finally transformed into the real single celled life – stable and sophisticated enough to be called species. These species bore basic capacity to survive and prosper in the face of various environmental changes and damaging attacks from other species.

Randomness and consistency are incompatible and paradoxical with each other along the evolutionary timeline. From a chronological point of view on the origin of life, life starts in pure randomness and reaches maturity in consistency. It's the randomness that has generated the endless possibilities for life to start and it's the randomness as well that has established the consistency of life. But consistency reduces the randomness and reduced randomness in turn lowers the chances to generate new functional proteins and slows down the system to develop and advance. Reduced randomness thus slows down evolution of life. On the other hand, randomness is adverse to life as it disrupts consistency and destroy the stability of species. During the origin of life, total randomness has been the sole and most effective trial and error approach to establish life on the earth, albeit time consuming and extremely wasteful. In this period, the focus of randomness is mostly on random chemical reactions that produce vast pools of random polypeptides and nucleic acids in the absence of DNA templates. After the origin of life, the randomness must be kept to certain levels that aren't too large to destroy the consistency of life, but aren't too small to halt the evolution of life. In this period, the randomness is almost exclusively on mutations that occur to DNA genomes. Indeed, life has been evolving and proliferating from the beginning by balancing evolution and stability of biodiversity through balancing the mutational rates on the genomes. It's safe to say that life arises from the total randomness at the very beginning and evolve on subtle random mutations thereafter. It is unbelievable, but it is extraordinarily clever.

When we think of the origin of life, it's essential to think of the environments 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 that the nascent earth had

hosted an environment in which only RNA or proteins, not both, could be produced. Nascent earth must be in a mess, on which amino acids, bases, ribose, and many other chemicals were randomly present and constituted the life incubator. Considering similar chemistry of nucleotides and amino acids, nothing could prevent amino acids from linking into peptides while ribonucleotides polymerized into RNA, and vice versa. In terms of pure chemistry, peptides were the things to be produced prior to RNA synthesis in such a system as discussed earlier. If peptide pools happened to contain enzymes for RNA synthesis, they couldn't be excluded in the process of RNA replication even if certain RNA that bore special secondary structures had catalytic activity and could self-replicate. If proteins, RNA, and DNA were present in their own worlds separated from each other in space, then when and how could these three independent worlds come together to form primitive life? Life emerges spontaneously in a pure random fashion in open environments on the young and turbulent earth, and it's not the same as making dish with each ingredient on the kitchen table, allowing you to get spoonfuls and add to the cooking pan any time you liked.

4. Slow Evolution – the Quiet 3.5 Billion Years before Cambrian Explosion

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

The single celled organisms from Stage 1 are far from mature and robust. They are vulnerable and defenseless against natural elements and their genomes are too small to support evolution. In the prokaryotic era, evolution must bring up essential changes to the single celled organisms to increase their genome size and protein coding gene counts. Over a period of 2 billion years, single celled organisms could have generated many new functional genes as a result of numerous spontaneous random point mutations on their genomes. A larger genome tended to generate more novel genes via random point mutations. These novel functional genes differentiated one species from others. From evolution point of view, randomness was still the single major factor that drove prokaryotes to diversify and expand, resulting in numerous species. When each species carried a certain set of unique genes, all these species-specific genes formed a huge gene pool in the prokaryotic kingdom.

In the dynamic prokaryotic era, life was influenced greatly by some forms of extra-chromosomal genetic material that could co-exist inside the cell walls as an independent genetic entity. Plasmids are extra-chromosomal genetic material that can transmit themselves from one bacterium to another (even of different species) mostly through conjugation. Plasmids can carry host genetic material via DNA recombination and transfer host genetic material from one bacterium to another. Bacteriophages are another extra-chromosomal genetic material in the forms of viruses. They transfer genetic material from one bacterium to another through infecting various prokaryotic species. Gene transfer among species through extra genetic materials have played significant roles in disseminating unique genes from one species to another, which not only enriched the genes coding for more functional proteins for that species, but also increased its genome size. However, because of rapid mutations, prokaryotic genomes remained enormously heterogeneous among prokaryotic population.

Genotype differences led to differentiation of cells into various sizes and robustness. Large cells could engulf small cells and integrate small cell DNA into their own genomes. This is a random merger process called symbiogenesis. Single celled organisms diverged into numerous different lineages that were different from their ancestors through rounds of mutations and symbiogenesis. The diverged 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 classify these species into different genera, even phyla. Therefore, it was randomness and time again that drove the evolution of single celled organisms to become vastly different species that were more adequate in function, better formed and developed in morphology, and more robust in defending against adverse factors.

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 structures and functions with eukaryotes. For example, 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 is 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 higher fidelity of DNA replication, further reducing the occurrence of randomness in the process 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 fast 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 genome sizes and increase protein coding gene counts. Confinement of genome in the nucleus and acquisition of more enzymes useful for DNA manipulation allowed eukaryotes 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, while foreign DNA for insertion could come from internalized cells or viruses through endocytosis. Duplication of DNA fragments was another major means to increase the genome size. Over time constant accumulation of random point mutations on the chromosomes could have generated a variety of possible genetic loci with potential biological significance. All the random genetic changes diversified the population into numerous 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 genes in cell types. Therefore, it's imperative to establish rigorous regulatory mechanisms to control gene expression in order 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 that genetic machine remove all introns 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 and RNA. Introduction of introns isn't cost effective, but is a viable way to guarantee the integrity of cell types. Splicing wasn't a purpose to make sure that gene expression was leak-proof, but it just happened randomly and solved the leak problem. It has been preserved in all eukaryotes since then. Therefore, it was the appearance of introns and splicing that saved life from a dead evolutionary end. Is there other strategy that could replace introns to serve the same purpose, if not better? In addition to the above roles, intron splicing has introduced an unexpected, but powerful mechanism to generate protein variants through alternative splicing as discussed in the next section.

Evolution of eukaryotes from the moment of their appearance to the moment right before Cambrian explosion spanned a period of staggering 1.5 billion years, roughly two third of which were 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 cellular 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 changes in the period were confined to their genomes as well. This is 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 the eras reveal daunting difficulties for the genetic system to create novel proteins and then assimilate them into the existing biochemical processes and cellular structures of the organisms. It also shed light on how the evolution itself is evolved over time.

So far genomes of hundreds of 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

selected species ranging from archaea to bacteria to organisms emerged during Cambrian explosion. The data in the table will help us understand evolution and find out the bottleneck of evolution.

Organism	Type	Genome Size	Coding Genes
<i>Pyrobaculum aerophilum</i>	Archaea	~2,200,000	2,562
<i>Haloferax volcanii</i>	Archaea	~4,000,000	3,895
<i>Chlamydomonas reinhardtii</i>	Bacteria	1,225,935	1,113
<i>Chlorobium chlorochromatii</i>	Bacteria	2,572,079	2,002
<i>Gloeobacter violaceus</i>	Bacteria	4,659,019	4,430
<i>Bacteroides fragilis</i>	Bacteria	5,205,140	4,260
<i>Micromonas pusilla</i>	Algae	~20,900,000	10,056
<i>Guillardia theta</i>	Algae	~87,200,000	24,840
<i>Dunaliella salina</i>	Algae	~343,700,000	16,697
<i>Breviolum minutum</i>	Algae	~1,500,000,000	47,014
<i>Tieghemostelium lacteum</i>	slime molds	~23,400,000	10,225
<i>Hortaea werneckii</i>	Yeast	~49,900,000	15,974
<i>Aspergillus nidulans</i>	Fungi	29,828,291	10,534
<i>Armillaria ostoyae</i>	Fungi	60,106,801	22,299
<i>Blumeria graminis</i>	Fungi	124,489,486	7,118
<i>Puccinia graminis</i>	Fungi	88,724,376	15,800
<i>Crambe crambe</i>	sponges	~370,500,000	
<i>Chrysaora quinquecirrha</i>	jellyfish	~337,400,000	
<i>Bolinopsis microptera</i>	comb jellies	~265,400,000	14,118
<i>Lytechinus pictus</i>	urchin	~811,700,000	20,054
<i>Cydia splendana</i>	Moths	~630,600,000	14,646
<i>Penaeus indicus</i>	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 in table 1 is obtained using genome analysis software, so it doesn't necessarily mirror their true expression in the organisms. However it shows that these genetic loci exhibit gene structures and can be considered as genes. An implication is that at least they can serve as genetic materials for new genes via point mutations and gene duplication.

Genomes of prokaryotes are much smaller and contain many fewer protein coding genes comparing with the genomes of eukaryotes. Furthermore, the sizes and gene counts varies greatly from species to species. On average bacterial genomes carry a size of 3 million bps and contain about 3000 genes. Each gene has an average length of about 1000 bps, encoding 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 single celled life must be much smaller with many fewer genes than modern 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 their 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 are 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. Protein coding gene counts usually fluctuate around 15,000 to 22,000 regardless of genome sizes and positions 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 early on the evolutionary timeline in some 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 proceeded.

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 that hundreds of millions of random events converged on the locus base by base over an inestimably long period of time. For example, a random piece of DNA in the genome would be first converted into a semi-gene locus coding for a random polypeptide of good length and then refined into a gene coding for a protein with some catalytic activity. Finally the locus would undergo further mutations to become a gene that would encode a biochemically active enzyme useful to the organism. In this process any early meaningful mutations could be canceled out by later ones, slowing down the progress. Therefore, evolution of a gene with biological significance must be an endless and repetitive trial and error process, bearing possible fruit only after random trials for tens, even hundreds of millions of years.

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 modern organisms to generate a limited amount of energy in the absence of oxygen. It's a mystery when glycolysis pathway first appeared in prokaryotic era, but from the evolution standpoint, all of these enzymes must be developed de novo slowly from random DNA sequences in different prokaryotic organisms, and then merged into one organism to form the final pathway via plasmids or bacteriophages mediated transfer of gene-containing DNA fragments among organisms. Merge could play very important roles as a general mechanism behind the evolution of multisubunit proteins and many cellular structures and biochemical processes that rely on multiple protein molecules to work as single units. Gene merge must have been an effective approach to greatly accelerate the evolution of prokaryotes, but the difficulties of obtaining functional proteins from random DNA sequences remained, always being a monotonous and lengthy trial and error process.

Consider a hypothesized project. A group of distinguished experts in protein engineering 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 any technology available today to design enzymes for the pathway D. A common approach would be to explore the vast protein databases to identify any enzymes with potential for modifications to obtain desired activities and specificities. Modifications wouldn't be random, but carefully engineered 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?

If about 2000 genes were new additions to the genomes of nascent cells, a period of whopping 2 billion years just for 2000 genes seemed to well speak for the extraordinary difficulty and complexity in the evolution of prokaryotic organisms. On one aspect, it could be understood purely by considering that many new genes had to be created first from random DNA sequences, and then integrated into the existing system to function smooth to make prokaryotic life more robust and diverse. On the other aspect, it could be understood from a different angle. Prokaryotic organisms, as a simple form of life, had come to a dead end and reached the evolution limit with its genome organizations and cellular structures. In other words, their potentials to evolve further into more complex and advanced forms had been exhausted. It is a forever mystery how much time it took for the eukaryotes to emerge from the prokaryotes as the clear time divisions shown in the evolutionary timeline are for illustration only.

On top of the average gene count of 3000 in prokaryotic cells, eukaryotic organisms added about 7 times more new genes to their genomes, reaching 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 pathway in all forms of life. There exists clear evidence of sequence homology between hexokinases from yeast, plants and vertebrates, but not between hexokinases from prokaryotes and eukaryotes. Function-only homology between protein counterparts from prokaryotes and eukaryotes provided 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. Right after departure from prokaryotes, eukaryotes must carry genomes that were closer to prokaryotes than to modern single celled eukaryotic species in terms of size, gene count, and sequence, and their genetic apparatus must be simple and basic in terms of genetic operations involving large pieces of DNA, especially DNA duplication. Moreover, as most genes functioned as groups, it was unlikely that all genes in a group were created at the same time. If some of the genes hadn't be attended by other genes in the group for some time, they could disappear after more mutations rendered them useless. Therefore, evolution of a functional gene was not only determined by mutations on its own locus, but also depended on other genes as well, if it belonged to a functional group. However, generation of almost 18,000 new genes in about 1.5 billion years could be considered a lightning speed in contrast to 2,000 genes in the previous 2 billion years, even though point mutation rates in eukaryotes were lower due to higher fidelity of DNA polymerases. An implication is that eukaryotes had developed mechanisms for DNA duplication early in the period. Despite continued de novo creation of genes from random DNA sequences, DNA duplications should take the credit for rapid increases in gene counts. If the low levels of cell differentiation were taken into consideration, the pre-Cambrian species apparently didn't warrant such high gene counts, suggesting that not all genes were expressed to serve cellular activities and biochemical functions. It's more than likely that many of the genes served as potential gene templates for future new genes, laying down the foundation for fast evolution.

In the later prokaryotic era, random mutation rates diminished notably due to the maturity of DNA replication apparatus, slowing down the evolution of genes to a great extent. In eukaryotes, the

genomes were protected in the nuclei, which increased the stability of genomes significantly. In addition, increased fidelity of DNA replication hindered evolution further, leaving time and rapid replication cycles as the major factors in augmentation of gene repertoires in the pre-Cambrian period, in which the chances to create new standalone enzymes, not to mention complex pathways like hypothesized pathway D, were extremely small. Therefore, establishing new biochemical and cellular functions from random DNA loci would be the most painstaking and enduring processes of trial-and-error, and it was possible only when great chances, lucks, and coincidences all had occurred and converged fortuitously in some individual organisms. The time span of 3.5 billion years for the slow evolution are the strong indication of the enormous difficulty, setback, frustration, and uncertainty of random mutation based evolution of the early living organisms.

5. 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 like mammals, albeit undersized genomes. Meanwhile the genetic machine has become better developed and more powerful. An implication is that evolution has switched to a new mode, in which gene variations and gene duplication are the major mechanisms behind the appearances of numerous biochemical and cellular processes that make new species more complex and advanced.

Assume that there was a new small chemical named X that was able to regulate body temperature in the extremely cold environment. To generate a receptor for X, an existing receptor gene for a different small molecule Y happened to have become duplicated. Turning 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 sequence changes must enable the receptor to bind X by creating a three-dimensional structure with an internal space that could 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 was another three-dimensional structure that could 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 would be expressed only in selected tissues and time. And many more. Having accumulated a large number of point mutations in the duplicated DNA locus over numerous generations, a Y-receptor-based functional X receptor emerged. Nevertheless, there wasn't a guarantee for such an outcome.

In the above hypothesized scenario, the gene for X receptor was initially duplicated from an existing gene. By taking advantage of the duplicated gene as a fully structured DNA sequence, it would be unnecessary to build a genetic locus with common gene structures in a random DNA sequence, but to focus on forging a protein coding sequence that could bind specifically the molecule X and transform it into an active form of the receptor. Although it remained to be an extremely lengthy trial and error process, without doubt, it would 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 in organisms living in the extremely cold area, would the organisms 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 didn't need to respond to cold temperature? It must be possible. Neither the need for X would trigger or accelerate the development of a functional receptor for X, nor the lack of need would prevent the development of a functional receptor for X. Most biological functions and structures didn't emerge on necessity or usefulness, but rather they emerged as the consequences of random

mutations. They would be preserved if they happened to enhance or complement some processes or structures for better functionalities, or if they could contribute as standalone factors to increase the well-being and survivability of the organisms.

The large magnitude of genome sequencing has revealed that species in the same genus, even in the same family, share extremely high percentage of identical sequences. Its implication is that large differences in morphologies don't mean similar large differences in genotypes. In fact as species move up the evolutionary ladder, the differences between genotypes have diminished greatly. For species that are classified into the same family, differences in morphology, biochemistry, and cellular structures can be generally attributed to variants or isoforms of the same proteins expressed in different species. The advent of protein variants is a giant evolutionary step forward for more complex and advanced species.

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 working 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 of genes that have existed in lower organisms, a strategy of derivation and reuse.

Protein variants in the evolution of species also reveal an important biochemical property of proteins. The sequence dependent three dimensional structures are not always rigid, but show great elasticity in the cells. In other words, the three dimensional structures of some proteins are elastic enough to withstand certain sequence changes and remain compatible with the existing biochemical and cellular processes, allowing variants to perform the same or similar functions. However, elasticity in structure can have more or less impact on the biological processes or activities they serve in subtle or indirect way. This is very useful from evolution standpoint. For example, a neurotransmitter receptor variant could have its affinity for the same ligand increased or decreased relative to the original receptor, thus changing the behavior of organisms accordingly. Signal transduction pathway could be altered because of changed physical interaction between altered protein components, thus eliciting more changes in the downstream events. The impact of structural variants is often visible. An organism could assume a different morphology when tissue orientation in normal development was skewed to a certain degree due to the substitution of a tissue growth factor with a variant. The numerous occurrences of such changes in a species indicate the emergence of a new species. As more and more such changes occur in species, new species emerge as a result, but new species are more complex and advanced as well.

From biochemistry standpoint, protein isoforms are different forms of a protein that are encoded by the same gene and perform the same or similar biological roles in the same species. They can arise from alternative splicings or variable promoter usage, which can be attributed to the result of base insertion, deletion, or substitutions, in the promoter or splicing sites of the gene. Protein variants often mean protein isoforms, but also refer to proteins that originate from gene duplication. The duplicated genes, if both are active, can carry their own unique sets of random mutations over time, thus, encoding proteins that differ more or less from each other in their amino acid sequences. In this case, the proteins encoded by duplicated genes are variants of each other if they perform the same or similar biological roles in the same species. A large portion of protein coding genes in eukaryotes has been found to have isoforms or variants, and each of them is usually expressed in different cell types and/or in different developmental stages.

Protein variants or isoforms have a broader meanings from the evolution standpoint. Protein variants have played critical roles in the evolution of species over hundreds of million years. In general most of the genes in a later species can be traced back to have homologous genes in the early species, and the proteins they encode are variants of the proteins encoded by those homologous genes if they play same or similar roles in both species. This greatly expands the concept of protein variants to evolution. This type of protein variants constitutes the foundation of evolution, ensuring the continuum of genetic materials and the common set of biochemical processes and cellular structures that have been the backbone of all later eukaryotic life. By comparing gene and protein sequence changes between protein variants across different species and across evolutionary tree, it would be possible to shed some light on the evolution of species at the molecular level.

Evolution is a continuous and gradual process to move from low forms to higher and more complex forms. The low forms serve as the base on which more complex and advanced phenotypes can be built. However, evolution wouldn't happen if genetic events were limited in scope and degree. For example, random mutations could increase the diversity of the base forms, but would make base forms no more complex and advanced. Evolution of species to a higher level required a lot more to be created on top of the base forms. In this quest creation of new protein variants, especially via gene duplication, must be the easiest and quickest way to do so. As a result, the higher species inherit the backbone of life from the base species, and display numerous new phenotypes formed on numerous protein variants. For example, fish and amphibians all have brain, eye, heart, limbs/fins, and so on, but these organs in the amphibians have undergone profound changes in order to adapt to the terrestrial life, and they are functionally and structurally more advanced and sophisticated than fish. It's the difference aspect of the two protein repertoires, including numerous new protein variants absent in the lower species, that is responsible partly for the higher species to be higher on the evolutionary tree. In all likelihood, random variations of protein molecules is one of the major facts that has dramatically shortened the time for new species to appear and contributed to the greater complexity and diversity of biochemical and cellular processes in the higher species.

The genetic origins of protein isoforms or variants are various, but alternative splicing site creation and gene duplication are responsible for most variants found in an organism. Gene duplication is especially important to our understanding of evolution as it gives rise to protein variants encoded by separate genes, which allows individual genes to diverge further in their own paths. Genes coding for protein variants could be traced back to some ancestor genes or master genes if these genes hadn't diverged too far away from their master genes. A master gene could bring about many child genes via successive gene duplication and random mutations over hundreds of millions of generations, but they were always their master gene, even after great divergence over time had hidid such a relationship. Parent-child gene relationship is more appropriate to describe the relationship between protein variant genes from species that are close on the evolutionary tree. In general, after enduring random mutations, the child genes-encoded proteins remained to be the variants of their parent proteins if functional similarities were not lost. The differences in amino acid sequences would widen as species diverged further apart until identifiable relationship between them disappeared, so the biological roles they served.

Opsins are a family of proteins that function as photon receptors in the eyes of animals when coupled with light sensitive chromophore 11-cis retinal. S-opsin absorbs short wavelength light, 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, respectively. By contrast, most mammals lack L-opsin and L-opsin-containing photoreceptors and are not sensitive to long wavelength light. When primates split from most mammals, duplication of M-opsin gene

occurred, which gave rise 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. M-opsin gene is the parent of L-opsin gene, and L-opsin is a variant of M-opsin. There is no further divergence of the L-opsin gene.

In general, a variant will assume the same biochemical roles of its parent protein, but with subtle changes in biochemical properties. It's these subtle changes that empower variants to better fit old or new occasions or fill the void that the parent proteins can't fill, or complement the action of parent 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 great levels. Cell differentiation became a major evolutionary event in the later stage of slow evolution. Multicellular life began to take more defined and more sophisticated morphologies 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 molecules, such as new protein factors to control the development of the morphology and copious protein components to perform the underlying biochemical processes and build up the cellular structures. The genes that encoded these early proteins must have served as master genes, from which protein variants were derived to enable new species to develop and assume more complex morphology with extensive differentiation of cell and tissue types, thus impacting the evolution of species in a fundamental way throughout the evolutionary history.

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. Assuming 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 parent 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 parent 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 expressed in the wrong part of the embryo, it could induce the growth of a fin at a wrong location.

The consequences of divergence of master genes into variant families are remarkable. The rise of myriad forms of phenotypes along the evolutionary timeline could be attributed to these variants as critical influencing factors. The early variants of fin inducing factor descending from the master gene gave rise to various forms of fin morphologies on different fishes. Further divergence gave rise to many more distant variants, which controlled the development of limbs unique to amphibians, reptiles, birds, mammals, primates, and finally humans. The degree of divergence from the master gene seemed to mirror 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 development of fins and later on the limbs has been amplified to the utmost degree in the past 500 million years, although many of these variants might no longer have identifiable sequence homology.

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 active 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 sequence of each of the keratin variants has been preserved over evolution in different organisms because they forms unique three dimensional structures that are particularly suited 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 to 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 DNA 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, and more. 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 to 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

one of the molecular bases for the evolution of closely related species like gorilla and chimpanzee and for the staggering biodiversity on the earth today.

The last example of protein variants are homeodomains. A homeobox gene is a piece of DNA sequence of about 180 base pairs long, encoding a 60-amino acid long protein motif called homeodomain. The homeodomain recognizes and binds to specific DNA sequences with its helix-turn-helix structure composed of three alpha helices. Homeoboxes are found within many genes coding for transcription factors that are involved in the regulation of gene expression in embryonic morphogenesis. For this reason, homeodomains have become one of the most studied DNA-binding motifs. Homeoboxes comprise a large family of DNA binding motif variants in a species as the result of extensive gene duplication and divergence since pre-Cambrian period. For instance, 103 in *D. melanogaster* and *C. elegans*, 121 in sea snail, 111 in polychaete worm, 96 in sea urchin, 181 in leech, around 250 in most vertebrates (242 in humans and 289 in mouse). Homeodomain motifs are functionally conserved. Members in the family can share low amino acid sequence identity and recognize dramatically different DNA sequences, but they interact with the DNA nearly identical. Mutations in homeobox genes can cause developmental disorders and produce easily visible changes in body morphology. It's obvious that evolution of the homeobox gene family is partially responsible for the changing morphologies of species as they move up the evolutionary tree.

Assume there is a protein variant X1 that deviates from the master protein X via gene duplication and random changes in its gene 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, while alternative splicings contributed a lot more to the protein variant repertoire. Many genes in higher organisms are variations of the non-duplicated genes of lower organisms. It seems likely that all of the novel genes in a species are just a small portion of the protein coding gene counts. The large gene counts of multicellular species relative to their simple morphology 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.

6. Fast Evolution – New Species Arise in Evolution Cycle

Evolution from simple to advanced is the inherent quality of life from the very beginning because of the mutability of the genomes. However, 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 algae, slime molds and fungi. Then slightly more complex multicellular organisms like sponges, jellyfish, sea anemones, corals gradually emerged in the later phase 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, sea cucumbers, and sand dollars all appeared in this period. First plants and fishes rose at the later stage of the period as well. From evolution point of view, all these species remain very low on the evolutionary ladder despite stunning varieties in complexities and forms.

In the following 500 million years since Cambrian explosion, evolution greatly sped up, and numerous new species of higher complexities emerged in short periods. 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, organisms rely more on protein variants to build unique morphologies, cellular structures and biochemical processes as they evolve to higher levels. Protein sequence comparisons reveal a lot about those proteins that play similar biological roles in different species. A large number of the proteins can be classified into groups based on the overall similarities of amino acid sequences or sharing of certain short amino acid sequences called motifs. By taking advantage of protein variants and motifs, it was much easier and faster to create proteins with desired functions and properties, thus alleviating the challenge to assimilate new components into the existing biochemical processes and cellular structures. On the contrary, it would incur a formidable array of problems to create and integrate new protein components even in relatively low and simple species.

Evolution is a process of constant changing, but only changes that go beyond the threshold of evolution can bring about meaningful consequences to the organisms. Modern organisms, archaea, bacteria, animals and plants, don't seem to be on the path to evolution. The 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 in the past billion years, we wouldn't see organisms as low as archaea, bacteria, algae, fungi, jelly fish, sea urchins, etc. implying that not all low 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 an overall 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 genomes of ancestor organisms should be more prone to environmental changes and quite elastic and blessed with a genetic machine that could perform the large magnitude of genetic changes – changes that would generate new genes or gene variants at much faster rates in a short period. The result of it was the establishment of new phenotypes that could define the organisms as new species. In doing so, ancestor organisms must be able to accommodate new components created and added to the system at different times until all the new necessary components were in place to complete the new phenotype.

What exact events could trigger the large magnitude of genome changes 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, on conditions that we were not part of the mass extinction. However, it's still worth to think about it and envision something even 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 common DNA recombinational events during meiosis. When sudden geological and climate changes broke out, ancestor organisms in a population suffered from more mutations, resulting in a decrease in the replication fidelity of DNA polymerases, thus accelerating the genome wide accumulation of point mutations. These genetic changes acted as perturbations that drove the genome from its disarmed state to enter an inconsistent and unstable state, an armed state. In the armed state the genetic machine of an individual ancestor was activated to perform genetic changes that would reshape the genome to start an evolutionary event – a reshaping process for the genotype. As the reshaping process continued to reshape the genome by bringing further changes, it gradually diminished in magnitude into a healing process, in which the genome was gradually returning to a consistent and stable state, a new disarmed state. When the genomes of all individuals descending from the common ancestors regained disarmed states, it completed one evolution cycle (Figure 3). All the individuals that carried mutations in the cycle, including those dead at the embryonic stage, were intermediates of the cycle. The boundary between processes reshaping and healing is quite blurry and continuous, but the two processes are two separate concepts useful to reveal what was happening in an evolution cycle.

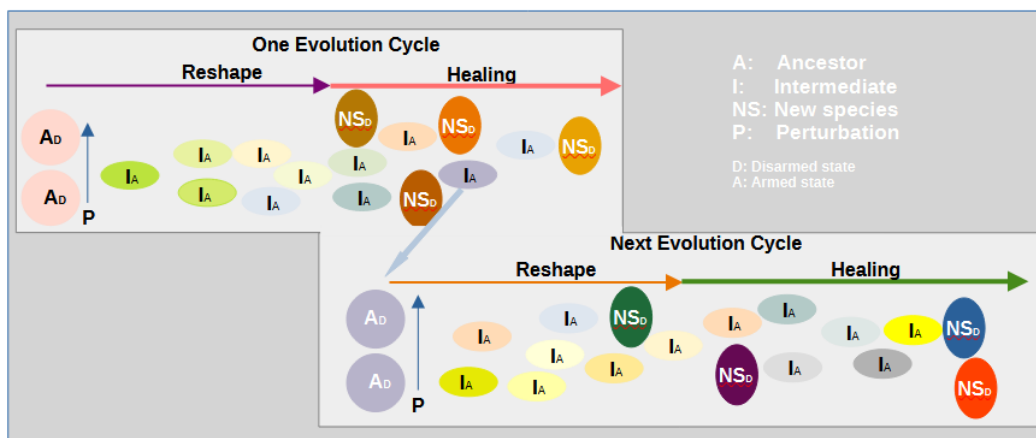


Figure 3. Evolution cycle from ancestor organisms in disarmed states to new species in new disarmed states, including processes reshaping and healing, and numerous intermediates in armed states.

Assume there was a population of a single ancestor species. Higher frequencies of random mutations brought bumpy starts to the calm genetic machines and aroused them to initiate reshaping processes and enter armed states. An evolution cycle began as such. The genetic perturbation must result in 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, the differences in genotypes grew, leading to increasingly widening heterogeneity of the population. Meanwhile, process reshaping was slowly reduced into process healing after it had incurred numerous changes to the intermediates. New components, including protein variants and novel proteins if any, gradually reached an equilibrium with the existing biochemical machines, bringing process healing to an end. All survived intermediates entered new disarmed states, in which the enduring instability and incompatibility among new and old components had been erased, and all the biochemical and cellular processes restored to balanced states after being agitated by genome wide changes. The difficulties to re-establish such balanced or disarmed states would be unthinkable if numerous protein variants were not employed in the cycle, especially for species as advanced as fishes, amphibians, etc. Regaining disarmed states permitted life to return to work in harmony and stability, but at higher levels. What's visible in the cycle was the morphological transformation of each intermediate after their genotypes were changed. At the end of the cycle, the genotypes of the survived intermediates were so different from their ancestor and from one another that they were no longer the same species, but distinct new species, the happy ending of an evolution cycle.

The nature of a genetic change is determined in lieu of survivability of the 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 on quietly from failure to survive seemingly endless mutations. They formed the dead ends of the cycle. Only those heritable changes were passed down, resulting in variations in phenotypes among the next generations. The concepts of beneficial and deleterious mutations were not applicable to the evolution cycle as the final effects of any non-lethal mutations produced in one generation must be shown and viewed from the whole cycle standpoint. Generally, the fates of all intermediates in the armed states were random and unpredictable.

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, reshaping 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 diversity that arises after each evolution cycle. All genetic changes that occurred in a cycle are deemed as genome-wide, but the changes to one generation must be limited and barely large enough so that some of the intermediates would 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 solid 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, it draws 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, which is 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. Since the lengths of those productive trails differ tremendously in a cycle, the end of the cycle, like processes reshaping and healing, is a concept that is difficult to determine but useful to indicate how new species arise along the long life timeline.

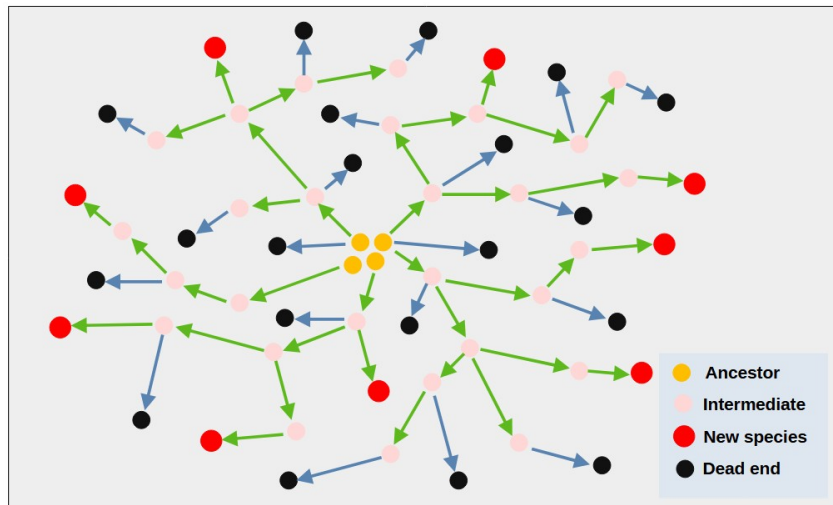


Figure 4. A simplified diagram to illustrate how new species arise in explosive mode in one evolution cycle. Distribution of new species is random relative to the ancestor organism. All new species can be classified into a single class.

Assume in a fish evolution cycle that one of the fin inducing factors 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 led to five new species, and this variant didn't diverge further, then 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 the variants biochemically and structurally, more similar the final fins on the 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 which 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 lucky species emerged from this intermediate after X8 generations.

In biology, there is a classification system that classifies living kingdom into eight levels based on shared characteristics. Class, order, family, genus, and species comprise the last five levels. All primates fall in Mammalia class and Primate order. 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 seemed to be the longest, hence referred to as human trail. It would be obvious that gorilla, chimpanzee and human shared the same ancestor and a series of common intermediates until reaching a particular intermediate, from which 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.

It's hard to determine which species reached their disarmed states earlier than others in a cycle. Fossil records are useful in estimating the approximate time species appear on the earth, but not the time an ancestor organism begins to evolve or starts an evolution cycle. Worse was few intermediates left fossils behind to record their evolutionary past. Scarcity of human related fossil records have hindered progress in tracing human evolution in the past 3 million years. Only the genome sequence homology between species seems correlated to their evolutionary relationship. Therefore, genome sequence comparison would be the viable resort to determine how close species are. Regardless of lack of details, life has been evolving endlessly through an unknown number of evolution cycles since the beginning of Cambrian period. What happened in human evolution isn't different too much from what happened in Cambrian explosion, 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 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 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.

7. Fast Evolution – A Closer View

Randomness has been changing its roles since life-like activity appeared in the incubator on the nascent earth. Randomness drove the origin of life. In this period, a large number of basic proteins, RNA, and DNA emerged from the random pools of extraordinary size to start the assembly of life via a trial and error approach at the cost of time. In the slow evolution, randomness incurred vast genome sequence heterogeneity among prokaryotic organisms as well as in the early eukaryotic organisms, resulting in gradual increases in protein coding gene counts. On the macro level, randomness was greatest in the origin of life and in the early phase of slow evolution, in which randomness was augmented by the short cell cycle, relatively error prone DNA replication system, and single cellularity. Consequently,

single celled organisms formed massive populations, in which individuals carried their own unique genomes, and many of them could be considered unique species, leading to vast genome heterogeneity in the single celled populations. If each of these species diverged into a few more new species, new species would arise in an exponential mode. Indeed, the number of modern day single celled species, including prokaryotes and eukaryotes, is too large to count. The greatest randomness in this period independently brought about numerous unique, novel genes in different species, enriching gene counts tremendously upon cell fusions, endocytosis, and plasmid and phage mediated DNA transfer.

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 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 became more complex and advanced, the integration processes became more challenging and risky, resulting in high failure rates, thus lowering the number of new species from an evolution cycle. As evolution proceeded forwards, randomness gradually changed its action mode, in which generating novel genes de novo out of huge random pools was replaced with generating protein variants through modifying existing or duplicated genes. On the macro level, randomness no longer referred to the genome uniqueness of individual organisms, but was limited greatly to the genome uniqueness of the population of a particular species, indicating that the number of species with high levels of tissue differentiation was limited, and became smaller as the levels increased. Transition of the randomness roles on the macro level seems to correlate well with the transition of slow evolution to fast evolution.

Organism	Genome Size	Coding Genes	Non-Coding Genes	Pseudogenes
<i>S. cerevisiae</i>	12,157,105	6,600	424	12
<i>C.elegans</i>	100,286,401	19,985	24,813	2,128
<i>Ciona intestinalis</i>	115,227,500	16,671	455	27
<i>D. melanogaster</i>	143,726,002	13,986	4,054	340
<i>Ciona savignyi</i>	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.

Genome sizes and protein coding gene counts shown in Table 1 hint what has happened in the slow evolution stage. Table 2 above shows similar information for the 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, but these organisms differ enormously in every aspect. An important implication is that evolution cycles, for example from fishes to amphibians or from amphibians to mammals, didn't seem to require many more novel proteins to produce new species, instead, protein variants must have played more significant roles than we thought. De novo generation of new genes wasn't always necessary, thus in fewer number even in higher species.

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. Some small changes in their amino acid sequences could have considerable impact on the organism's morphology, development, or physiology. They might have very few or even no variants in the same species, but share high sequence homology with proteins from species of the same class, even other classes. In genetics terms, these genes are very evolutionarily conserved. Some of the tissue or organ inducing factors might have this type of protein variants. They are not resulted from gene duplication, but from mutations that occur directly on these genes. If these genes were duplicated by chance, only one copy might survive, as more variants would be detrimental to the survival of the species, unless one copy found a good use elsewhere.

In contrast to those conserved proteins, there are proteins that display different degrees of tolerance to sequence changes and exhibit a high tendency to form protein variants. These protein variants perform similar jobs with their own characteristics in different cells and different species, and contribute to cell differentiation and speciation. GPCRs and keratins, as described earlier, are the extreme examples of protein variants of this kind. The presence of so many GPCRs or keratins variants in a single species attests their gene duplication dependent origin. Gene duplication is a random genetic occurrence, but is the 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 as implied by the observation that gene duplication occurs in modern day organisms. 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.

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 reshaping phase of the cycle and then slowed down gradually as the process healing progressed. In addition to the genome-wide random point mutations, the other known genetic changes, including insertion or deletion of functional motifs, exon shuffling, creation of alternative splicing sites, etc, might have all contributed to the conversion of various genetic loci into new functional genes or variant genes. It's the overall consequences of all those types of random mutations that slowly but steadily transformed the grand old biochemical and cellular landscape of the 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 ended up with 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 1×10^{-8} per base pair per generation. If one cycle spanned 10 million years, the final species could have accumulated about 100 million mutations, or 10% of bases have undergone mutations after 10 million generations. If each gene contained 900 bases to encode 300 amino acids on average and 200 bases for the regulatory sequences to control gene expression, the gene had a chance of a single point mutation 1.1×10^{-5} per generation. If 1,000 genes were house keeping genes with no 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 was 10%, which translated into average 110 ($1,100 \times 1 \times 10^{-2}$) point mutations per gene, and average 3.19×10^6 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, indirectly indicating the critical roles of gene variants in the fast evolution stage.

The above simple scenario was to get an estimate how random point mutations would accumulate in a gene in a cycle under the given assumption. To become a little more realistic with other assumptions unchanged, if the average number of offspring per one pair of parent organisms was 5 from a single birth in a life span of one year. All trails ended when the cycle ended, and all offspring survived mutations and produced their own offspring. In this scenario, the maximum number of offspring at the end of 10 millionth generation would be 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 number $5^{10,000,000}$ was an infinity, it couldn't represent the total number of new species emerged from one cycle. We could estimate the maximum number of new species from one cycle under a few more simple 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 had 110 random 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 due to lethality or infertility of the mutations, the number of survived offspring or new

species at the end of the cycle would be only an infinitesimal fraction of $110^{29,000}$. On the other hand, the functional effects of random point mutations was highly unpredictable since many of them could be synonymous or the number was too few to incur phenotype changes. Moreover, they could cancel each other over the period of cycle. For this reason, chances to bring about new species would be further reduced to a new infinitesimally small number. Such estimates would get worse when organisms climbed up the evolutionary tree. The number was infinitesimally small, but it was utmost significant since it truly represented the number of new species emerging out of an evolution cycle.

Sequence analysis has demonstrated that good functional homology but poor sequence homology among genes or proteins are common in species across different classes. This clearly indicates that the remarkable elasticity in the three dimensional structure is at work, allowing sequences to diverge through point mutations, while preserving their basic biological roles, the molecular basis of protein variants in the evolution of species. In an evolution cycle, protein variants displayed good homology among intermediates and usually differed more or less purely by some random mutations. However, all the differences from the whole protein repertoire would be added up sufficiently to result in distinct new species that fell in the same genus or family. However, during transition to higher class from low class, the large magnitude of sequence changes must have accumulated in the new species, reducing the overall functional and structural homology of some of its proteins with their fellow variants in the ancestor species. These proteins could have played essential roles as novel proteins to break the barrier of the class transition of the new species. Any such magnitude of genetic changes was the result of a genome reshaping process that was initiated by geological and climate changes and made possible after the genome had reached the level of maturity on 3.5 billion years of slow development.

The reproduction rates of amphibians 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 populations are taken into consideration. However, research shows that mutation rates of about 1×10^{-8} per base pair per generation are common almost in all modern species simple or advanced, revealing the two aspects of evolution. First the mutation rates must be higher than 1×10^{-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 to 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 encode, 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 the mutated proteins will change their functions in the cells. Only a relatively small number of non-lethal mutations are preserved, accumulated and finally give rise to the emergence of new species at the end of a cycle.

Division of evolutionary timeline into slow and fast stages is scientific as two fundamentally distinct mechanisms are working behind each stage. In slow evolution for both prokaryotes and eukaryotes, thousands of all-new genes are created de novo from random genomic sequences and their protein products are well integrated and assimilated into the cellular machine. At the end of this stage, the protein coding gene counts have reached the levels of higher species, while the morphologies remain 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 little usefulness of many genes created

and accumulated at this stage. It's even doubtful that many genes actually are active to produce proteins and contribute to any biological activities that occur in the organisms. Regardless of their true utility in the pre-Cambrian organisms, they comprise the abundant, ready-to-use genetic fodder, from which new proteins can be derived much more quickly to bring about species with the increasing sophistication of forms and phenotypes, which explains well what the evolution is all about, and which also reveals the genetic foundation of evolution. Clearly the slow evolution is a well-done preparation for the fast evolution to rapidly proliferate 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 and the task of creating genes from random DNA has been largely completed in the slow evolution stage. 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. It suggests that the apparent main task of the fast evolution is first to derive new properties from the existing properties and then to assimilate them into the existing system, the result of which is new species with distinct morphology and physiology in a much shorter period of time. This is 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 to reuse and recombine through mutation-based derivations. Because of this, post-Cambrian evolution proceeds in cycle. In each cycle, random mutations on protein coding genes result in distinct properties that drift away more or less from the parent genes. As the cycle progresses, more useful and handy properties appear and change the organisms overwhelmingly in every way, resulting in the emergence of new species as the cycle ends. Any new distinct properties that appear in a cycle are 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 to have new things added in and existing things removed or replaced, consequently many fewer new species emerged from a cycle. This is indeed what has happened 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 and reptile are known to have around 8,000 and 12,000 species, respectively. About 6,400 extant species of mammals have been recorded. The relative small number of amphibians implied that they were possibly the 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 long time ago. 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 were morphologically similar, but developed in different locations and classified into different suborders. 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 structures and functions. As a result evolution of passerine in separate subcontinents gave rise to 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, forming an evolution cycle. 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. Homologous traits, on the other hand, have similar forms, but can have varying functions such as limbs and wings. Homologous traits are originated from a common intermediate ancestor.

If two evolution trails were split early 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 longer track of 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 had shared with each other. In the evolution of birds, many evolution trails could lead to passerines that were morphologically similar, but genetically unrelated, although these evolution trails occurred in separate evolution cycles, different geographic locations, and timeline. 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 split 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 led to humans lasted 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 regions 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 in humans 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 human genes code for proteins that share the identical amino acid sequences with chimpanzees.

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. These differences must have encompassed all the changes needed to refine and hone every bit of human 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 have optimized most of the genes to achieve the best overall phenotype of a living organism.

In general for species that share a particularly long tract of an evolution trail, their genomes differ slightly due to limited point mutations, DNA deletion or insertion, etc. However it's these small differences that enable them to boast their own forms, expertise, and unique survival strategy and peculiarity, all of which together distinguish one species from the others. In other words, it's the sum of all the small differences scattered over the entire genome that confers each species with 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.

8. 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 have accumulated to a substantial amount over numerous generations, the individuals that carry those variations become a distinctly different species – new species.

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?

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 founded to have any relevance. If natural selection led to new species, an inference would be that all species throughout the timeline were in the process of evolution, and the appearance of new species meant the disappearance of old species or ancestor species. All this is clearly a glaring inconsistency with the extraordinary biodiversity today. More importantly, species that can be classified into one class emerge approximately in the same geologic period. It would be impossible for many new species to appear at the similar time simply due to natural selection. This section will focus on why natural selection theory can't be a cornerstone of modern biology.

Before diving into natural selection theory, it would be helpful to understand the essence of mutation based evolution. Once upon a time in the commercial TV market the competition between different TV makers were intense. Maker A introduced a new feature – picture-in-picture – to its TV set lines and gained market share considerably. Maker B followed suit and added to its TV set lines a similar picture-in-picture feature in attempt to regain its lost market share. Maker A immediately improved the picture-in-picture feature to allow users to watch another channel in the little picture-in-picture window. Maker A returned to the leadership position in TV industry because of this improvement.

In essence this is a consumer selection process, similar to natural selection in the evolution of life, but there exist two fundamental differences between the two selections. First consumers' purchase of TV sets with picture-in-picture feature would feed back to TV makers in terms of sales. Second, TV maker knew exactly what consumers wanted from their feedback. If consumers liked the picture-in-picture feature, the maker would make this feature more powerful, otherwise simply remove it from the next generation of TV sets. In summary, consumer selection will be fed back to the makers, and makers make necessary changes to their products based on consumer feedback. There exist active and mutual influences in the consumer selection process. In the absence of mutual influences, more changes in products could nullify earlier gain in sales, and it would be possible for a good feature to turn worse, not better down the road. Getting better isn't always achievable if no positive feedback mechanism exists. A gain in the early stage can't guarantee that it's not a loss in the later stage in terms of sales for industry. In the natural selection process, any consequences of mutations, regardless of their nature, are not fed back to the genome, and as a result, the genomes are unable to make corresponding changes to deal with these mutations. Evolution is a totally random, unguided, wasteful, and prolonged process.

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 consequence, while all other mutations are varying greatly in consequences. It is generally understood that natural selection will determine 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, where trees were tall and hard to reach for animals of normal necks like deer or antelope. One day genetic mutations occurred in the ancestor's genome, which made ancestor's necks grow longer. The long-necked individuals gained not only advantage to reach leaves on the high treetops, but also have 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 develop slowly and gradually into present-day giraffes. This explanation seems plausible and relatively easy to understand even for general public.

Nevertheless, evolution of species 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. 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 a 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 would 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, digestive system, 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 could be as complicated and entangled as we could imagine. A large number of existing processes and activities would be disturbed, even disrupted by those new or variant protein molecules. Therefore, it required the greatest and careful coordination and integration to guarantee that all of the old and new would be working together in peace. Therefore, 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, such large magnitude of changes must be of far-reaching nature and well beyond what could be brought about by largely piecemeal genetic mutations and recombination that natural selection depended on regardless of 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. Only few intermediates had survived all changes, reached the end of the cycle, and emerged as a new species – giraffes. It is important to recognize that nature won't give organisms something just because they have a compelling need for it. Nature doesn't know what an organism needs in order to survive better.

Attribution of evolution of giraffes' long necks to natural selection stood on the 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 had 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 advantage in survival and reproduction over normal-necked animals except the banal advantage to eat leaves on tall trees. Even this advantage might not exist at all if all the trees in the ancient habitats weren't as tall as today.

Giraffes and its closest 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 through natural selection.

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 could rely on dead animals for food to survive without special expertise. 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 fitted well 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 and their respective diets. From the evolution standpoint, different types of beaks or legs and the life styles 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 fits those traits 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. 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 many million years of natural selection. The enormous body sizes of many dinosaurs were unlikely to be the result of natural selection to fit their natural

environments, but they obviously contributed to their sudden demise. Excessively large antlers on some male deer could be detrimental to their survival when they were 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 hard to think of any utility of such structures for reproduction and survival, but it is obviously 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 survive outside their present special habitats. It seems to be the case that hundreds of million years of natural selection hadn't brought up good traits that would allow them to survive and reproduce in broader 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 the traits they have, so long as these traits would not cause serious morphological and physiological defects that would be lethal to their life.

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% compatible, 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 a sort of 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 niche space with ecologically strict conditions. Phenotypes or traits that determine the environmental tolerance of species can't be shaped by natural selection over time, as they come out of the evolution process. In general, strong and flexible survival 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.

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. On the other hand, the darkness only provides an environment in which bats can fully exhibit their spectacular echolocation skill just like zero degree C at which only water freezes.

Ants' pheromone system and different situations provide another example of water and zero degree C relationship. Ants produce an array of pheromones in different situations and locations, and each type of pheromone elicits different biochemical and cellular responses on the receiving side, which then are translated into unique kinds of behavior. Because of all those well displayed behaviors, 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 the 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 developed the capacity to beat and deal with those challenges will be able to survive and reproduce.

The geologic conditions of an environment are a factor in determining what species can fare well in it. A given environment is a fixed territory, which don't changes often except seasonal changes. It 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. 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. An implication is that early genes had tendency to retain later genes if they are compatible in functions to build up phenotypes.

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 or areas. 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 their chances to gain a reproductive advantage as winners. 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 vast wild was more than likely to be a different story because of the complexity of animal behavior and the ability to adapt and survive for the losers, for example by migrating to elsewhere. It is in sharp contrast to the Galápagos islands where the space is closed, preventing loser species from migrating to other areas. The observations on Galápagos islands were too uncommon to make generalization about the evolution of species. As a result, the winners might never have the chance to become the most common in the population despite their superior survival ability. Second, the fight was driven spontaneously by the innate hostile behavior of the individuals, and 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 favored traits that made them winners, the fierce fight would continue as usual and generate winners and losers with or without new favorable genetic variations. This was truly an endless event without a possibility to conclude ultimate winners. Therefore, nothing could change such a population in any meaningful way, and the possibility for genetic changes for a new species would be too remote to be realistic even after a million or tens of million years. The grand old winner individuals would remain to be the same species that was good for this or bad to 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 later 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 points on the evolutionary timeline. Evolution must have come up with all types of morphological and physiological features for these species to survive. Because of this, these species have formed unique behaviors to defy and evade the danger from predators, thus surviving and even thriving 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 those most critical for their survival, at the

moment they appeared in an evolution cycle. Otherwise they would have succumbed to all the dangers from the environment.

It has been long believed that some morphological and physiological features observed in certain species are minor and uncommon, but they are the result of evolution and important for survival and reproduction. These species are also observed to display some special behavior that is compatible with those features and is 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 things 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 had 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. Many similar but uncommon 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 of an organism like bill types and leg length that determine the most critical behaviors. It's these behaviors that determine the diet types, natural habitats, and the way the organisms co-exists with other individuals. All the traits and behaviors are givings, not choices to all species. Those features like colorful feathers, camouflage, and venom are exceptions for some species, but not the norm in the living world. If a male bird was suddenly armed with colorful feathers, then it could show off 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 have 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? After a few intermediates sustained prolonged genome-wide changes and succeeded in reaching new disarmed states, the evolution cycle ended. These survived intermediates carried new and changed traits or phenotypes and became the new species of the cycle. Major traits determined how new species would behave, which determined how the organisms would feed themselves and defend or shield themselves against predators or unfriendly environments, ultimately their survivability. 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 and strange acts in their life cycle. Any disadvantageous situations that incurred from some new traits, for example, food scarcity, adverse environmental conditions, increased risk of predators, and etc, would become a force to disseminate

new species across lands to settle in safe homes they could reach. Such a passive adaptation must be common in the 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 diet prompted giant pandas to migrate to places where bamboo was plentiful. From genetic point of view, panda genome determined bamboo as its major diet, and bamboo diet determined that pandas would constantly search for bamboo. It's this special behavior that inspired pandas to settle in bamboo rich terrain as its native habitat. However, the narrow appetite for bamboo put pandas in a grave disadvantageous situation, being confined to a bamboo rich area. If pandas couldn't settle in the bamboo rich area, they would have gone extinction in the course of its evolution. Despite a disadvantage, this trait can't be manipulated or discarded. The pandas must either find a way to live with it or perish from it. More generally, there must be numerous species that perished throughout the entire evolutionary time, partly because their genetic phenotypes couldn't be supported by their living habitats.

9. Learning and Evolution Cycle versus Natural Selection

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 experiences as well as coaching by 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 have the capacity to learn from other individuals. As the nervous system becomes more complex and advanced, the learning becomes more complex activities as well, requiring coaching more than experiences. Through learning, organisms acquire all the skills essential for survival and reproduction, not just the simple action through reflex. Learning also incurs profound changes in the behavior of organisms.

The nervous system of the fruit fly *D. melanogaster* is quite complex and advanced comparing to that of Aplysia. This nervous system enables the flies to learn not only from experiences 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 that 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 sophisticated behaviors are the result of comprehensive elaborate coordination among various types of cells, tissues, and systems through a series of action involving pheromones, sounds, and touch. Such behaviors are established gradually over long period of time and passed down generation after generation through

learning. 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 migration seems to be a trait acquired over millions of years, and can't be acquired in the artificial environment. As a result, most of them don't migrate 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 show great fear when seeing a chicken running around in front. Based on these observations, animals raised in captivity must undergo extensive training to regain essential survival skills 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 phenotypical traits. More properly, all behaviors are 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, behaviors as native and fundamental for survival as hunting for prey can be lost in a single generation. When learning is imperative for life to sustain in the wild, learning proper surviving behavior 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 behaviors. In more general terms, the behavior of organisms is largely founded on their genetic disposition, but is not genetically heritable like body shape and anatomical structures. It's determined and influenced by coaching and interactions with the environments. Only reflex based learning seems genetically heritable. Changes in behaviors will change 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 resulted from a change in the nervous system or a change in the environment like loss of coaching, rather than in the trait itself because of mutations 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 as it is reflex-based. 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 nerve cells elsewhere, complicating the root causes. The concept heritable trait thus lacks genetic clarity and is too blurry to trace the true root causes for an observable change. Consequently, it doesn't make much sense to state that evolution occurs when changes in the heritable traits favor survival and reproduction.

The scope of natural selection is generally limited to evolutionary changes within, 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 coexisted in the same habitats. 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 in their habitats, preparing to run or fly away quickly to escape the danger. In the wildness, individual organisms would succumb more easily to predators if their innate protective mechanisms against predators were weakened by mutations, while individuals would survive better if mutations strengthened their ability to evade or outsmart predators. In this regard, predators played a far more important role in maintaining healthy populations of the preys by eliminating individuals that were in disadvantageous positions for any reasons. A strong nervous system obviously boosts the survivability of organisms via learning to acquire skills to catch preys or escape from predators.

Two types of biological traits seem to be at work. The first type isn't dependent on behavior, such as feather colors, bill shapes, sharp teeth, etc, but contributes to the behavior of the animals. The second type depends on the behavior like fierce fight, territory defense, and nest building, etc. Obviously only the first type traits are truly heritable, and the second type is the utilitarian embodiment of the first type in practical use, as learning plays a critical role in their establishment. All biological traits and many observable changes are complicated with multiple mechanisms working behind them, and it's improper and misleading to try to explain them all with theories as simple as natural selection.

Evolution by natural selection obviously isn't a convincing explanation of modern biodiversity. 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, strongly indicating that organisms aren't in a state of evolution, even though random mutations occur equally to all of them. However, it is the random mutations that have remodeled old organisms into new species in the past billions of years. The solution to this apparent contradiction seems to be in special genetic mechanisms that enabled some of the organisms to be ancestors of new species. The ancestor organisms would possess extraordinary capacity to withstand a gradual and lengthy buildup of a variety of new proteins or protein variants from random mutations. Some of these proteins could have immediate visible effects if they were related to morphology, and all others must be integrated into elaborate cellular machine over time for new phenotypes, as this would be necessary to transform old species into new species. Ancestor organisms seemed to be predisposed to be ancestors of new species and were part of organisms in small number in a class, such as in fishes, amphibians, reptiles. Each ancestor organism in a class would evolve into species that would be classified into the same order, the level below class via evolution cycle. The evolution cycle hypothesis seems adequate to clear up the mystery of how new species arise from ancestor organisms, especially in an explosive mode. It provides a strong evolutionary basis to the present biological classification system that is scientific and well-grounded from the evolution standpoint.

All the mammal species have mammary glands that produce milk to nurse their young, but 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 differ after birth. 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, were 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, was 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 were classified into Mammalia class. Depending on the distance of divergence from the origins of main trails, species that shared the shortest common distances were 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 diverged early to start their own trails, all species descending from one trail would be classified into one order, and all species descending from other trail would be classified into another order. The species in these two orders shared the shortest common distances. As intermediates diverged closer to the ends of trails, they shared longer common distances and more common features, and the species descending from each intermediate would be classified into the same family, same genus, until the intermediates themselves became 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 left some ancestor organisms for future evolution cycles. Mammals that emerged in later cycles were 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 families. Mass extinction has important implications for the evolution of species. In the history of mammalian evolution, the number of mammal species produced from each cycle were likely far more than the number of mammals living today. It could be expected that some of the mammals were eliminated slowly because they were ill-formed morphologically or deficient physiologically, while many of them were weak against adverse environmental changes, and perished when natural geologic or climate changes struck. Generally not all new species from an evolution cycle enjoyed the same survivability. Only well-formed species were able to survive adverse environmental changes and continue to live to present days. It's like selecting a liquid chemical that freezes at zero degree C, only water would be selected after broad testing,

Evolution of mammals has been a prolific process involving numerous genome wide genetic changes and producing millions of new phenotypes among thousands of mammalian species. This process, upon decomposition, is merely a number of evolution cycles occurring over few hundreds of millions of years, but it is indeed far beyond the possibility for natural selection theory to give an explanation.

Bats and ants 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 the echolocation system to process ultrasonic sounds and how ants developed the pheromone-based means to communicate with each other. Such systems 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, variations in capabilities among bat or ant species can range from subtle to large degrees. A biological system that is as sophisticated, complicated, and diverse as these two is truly awe inspiring. Because of this, bat echolocation calls and ant pheromone communication are claimed to be remarkable examples of good design by natural selection.

However, more correctly it's a challenge for natural selection theory to give a satisfactory account for the evolution of these two systems. What kind of natural pressures could drive the evolution of such complicated systems? The genetic changes that natural selection is based on seem too inadequate to explain such sophisticated and massive biological traits. On the other hand, evolution in the context of evolution cycles apparently provides elegant answers 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 was quite distantly related. From the evolution trails, an early intermediate from one trail got mutations X to start the construction of the echolocation system. After sharing generation X1, the intermediate diverged again, from which one intermediate ended up as bat C after generations X2, and another one diverged much later at generation X3. From X3, one intermediate ended up as bat A at generation X4, and another one ended up as bat B at generation X5. X4 and X5 were not too distant apart in the cycle and shared many mutations toward the final systems. Therefore, their echolocation systems were closely related, but quite different from bat C. From the standpoint of evolution trails, bat A and bat B were close to each other on both horizontal and vertical dimensions, 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 in an evolution cycle, greater their genetic similarity.

Bats' echolocation phenotype must be backed up by a super genotype composed of a large family of genes. An ancestor organism descending from placentals started an evolution cycle that led to the explosive growth of bats. In the early phase of reshaping, 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 a nascent echolocation system. As the reshaping process progressed in the next millions of years, 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 the early protein products would retain and integrate any new proteins that happened to be needed to complete the system, slowly and protein-wise, allowing the system to develop and mature into what we see in the bats today.

Numerous intermediates must have failed to become bats partly due to lack of luck to develop all the necessary proteins on random mutations. An implication is that the system must be the result of a random development in the early phase of evolution cycle, until more components were produced and integrated into the rudimentary system, a total trial-and error process. In general, the development of a trait big or small would be utterly unrelated to the purpose for better survival and reproduction. On the other hand, if some mutations started a process that could lead to a new trait or change an existing trait, this process couldn't be stopped, but continued until the trait became part of the new species or ended up in failure, leading to death. This kind of evolution is unimaginable with the natural selection theory.

Generally speaking, larger the system, greater the space for the system to grow in more flavors in functions and details. Inherent elasticity of the three dimensional structures of protein molecules allows a large system to accommodate protein variants of the same function, resulting in many different, but closely related species, each of which displays the same system with its own characteristics. 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 order.

Development of life through evolution over billions of years can't be accounted for far beyond by one theory that is as simple and straightforward as natural selection. There are numerous unsurmountable

blocks for natural selection to explain evolution a little deeper. Egg sorting machines are a more proper analogue for natural selection. The egg sorting machine sort eggs into small, medium, large, and jumbo, simply based on the size, regardless of where eggs come from and where the sorted eggs will go. All organisms live in their native habitats, which provide all necessary natural resources for the organisms to sustain life. The habitats also play the sole role in sorting mutations organisms carry based on lethality as the environmental factors are unchanged in the period. Organisms die from lethal mutations and move on from non-lethal mutations. Any consequences from the non-lethal mutations would neither be fed back to the genomes for further changes, nor be certain in their long term effects on organisms' survivability on the evolution time scale. Therefore, natural selection removes lethal mutation carriers from the population, while leaving non-lethal mutation carriers in the wild on their own regardless of living better, worse, or neutral. This is what natural selection is all about, and nothing else regarding to the evolution of species.

Natural selection is so broad that it can explain almost everything present on the earth. Why do these things look like what they look like today? Why do these things work like this or like that? River XX was the largest river in the area and 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, most of the branches got blocked and emptied their water into branch X, making branch X wider. After a long time, branch X continued to widen and became the main branch to accept most water emptied from other branches. When branch X flew out of mountain, it merged into river XX with few remaining branches. Did the evolution of river XX result from natural selection as well? One more example will end this discussion about natural selection. On a flat land stood a big stone, while its surrounding area was 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 could be drawn from these small stones was that the tall big stone had resisted the same natural elements that eroded its neighboring stones over years. The survival of the big stone seemed fit well with natural selection theory. It seems that anything under the sky is the result of natural selection.

10. The Zero Sum Rule and Evolution Cycle

If you look at the extraordinary biodiversity and ponder how it grows over evolutionary timeline, you will 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 equal 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 in 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* were born hundreds of million years ago, it 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.

A simple phenotype is the expression of a genotype consisting of multiple genes. If the phenotype is at the zero sum state, mutational changes to the genes would cause damage to the phenotype rather than improve it. For example, if mutations lowered the catalytic activity of any enzyme in the glycolysis pathway, the mutations would bring down the metabolic rates of the pathway entirely. On the other hand, if mutations increased the catalytic activity of one of its enzymes, the metabolic rates would not be affected noticeably unless they happened to the rate-limiting enzyme, which obviously would disrupt the regulatory mechanisms of the metabolism network, resulting in the waste of resources to the least of all. Therefore, it's much easier to ruin a well-established phenotype than to improve it further, regardless of time and external natural pressures. This is the most basic and plain illustration of the zero sum rule in the evolution of species, and also clears up why natural selection has no place in the evolution of species.

Cavefish is fish living in caves and other underground habitats. More than 200 obligate cavefish has been described around the world. Many cavefish is believed to be the cave forms of normal fish that entered the caves on occasions millions of years ago and has adapted to the dark underground habitats ever since. The adaptation has the largest impact on pigmentation and eyes as they are useless in the absence of light. The loss of eyes can be complete or partial in different species, resulting in no or incomplete and non-functional eyes. As a result, cavefish is usually pale colored and blind. Cavefish that entered the caves more recently show fewer adaptation signs than cavefish that had been living in the dark for longer time. As a result, some cavefish looks still like their surface fish, and others look just like cavefish. Some are pale but have eyes while others have no eyes but are fully pigmented. Cavefish usually has larger fins for more energy-efficient swimming, and has lost scales and swim bladder. Cavefish often show little fear of humans and can sometimes be caught with the bare hands.

The cave form of the Mexican fish tetra, in addition to being blind, displays another characteristic – asymmetry of their left and right sides of the body, which makes the cavefish to be left-leaning, swimming in a counterclockwise pattern along the contours of the cave. Tetra is genetically so close to its symmetrical and normal-sighted cousins that live nearby creeks and rivers, and they can interbreed and produce fertile young. What's interesting about the Mexican tetra is that the cavefish is born with eyes, but the eyes recede and disappear completely at the adult stage. Similarly, this cavefish is born with symmetrical body features like their river cousins. As the fish matures, their skull bones take shape towards a visibly skewed direction, resulting in an asymmetrical body pattern.

The Mexican cavefish provides a very robust model for understanding the zero sum rule in action. The early phase of the cavefish is largely normal, having eyes and displaying body symmetry comparing with their river cousins. However, some random mutations have resulted in the loss of genes required to complete the later stage of the development of eyes and symmetry. These random mutations occur randomly to any individuals of all fish species and incur negative sum changes to the mutation carriers, regardless of their living environments, but only fish that lives in the caves can survive because of the lack of natural predators in the darkness. This indicates that random mutations can act randomly to any genes at any time at any place, the results of which depend on what kinds of damages the mutations will incur to the organisms. The loss of vision or body symmetry is critical for survival on condition of natural predators foraging around for food, but isn't critical on condition of no natural predators in the

habitats. In other words, that the changes are of negative sum is conditional, and the biological basis of the zero sum rule is partly the presence of natural predators in the wildness. Lack of any major novel phenotypes despite millions of years living in the darkness re-enforces the early statement that it's much easier to ruin a well-established phenotype than to improve it further, not to say to create new phenotype.

As discussed in earlier section, an evolution cycle started when right external conditions, including climate and geologic changes, struck the ancestor organisms. One cycle would bring about millions of evolution trails in a radiant fashion, each of which ended either as an intermediate carrying lethal mutations or an organism qualified as new species, resulting in explosive appearance of new species. The phenotypic traits of new species were heritable, stable, and relatively immune to genetic variations of large degree. In the meanwhile new species would establish unique behavior over time as they learned from interacting with the environments and other organisms in the habitats by employing their first type of traits. Their unique behavior would be stabilized into behavior-dependent traits slowly, allowing them to survive better under adverse circumstances. Therefore, evolution concerns only organisms that are either predisposed to evolution or any species that adhere to the principle of zero sum rule – once a species, always the same species.

The complexity of the echolocation development is enormous. All the proteins necessary for the echolocation functions must be created and assimilated into an increasingly complicated system, and their encoding genes must be regulated on the level of expression in various tissues to guarantee the organizational and functional unity of the system. From the zero sum rule standpoint, the development of any phenotypical trait, simple or complex, must have been the result of an unthinkable protracted trial-and-error process, in which the mutational effects on the phenotype must transit gradually from negative sum changes to positive sum changes to zero sum changes, corresponding to the transition of reshaping process to healing process in the armed state to the disarmed state at the end of cycle. In the negative sum state, genome-wide changes must be a sort of disruptive to the entire cellular machine, albeit not lethal, but it was imperative for the higher species to appear. In the positive sum state, the net gain of beneficial mutations over deleterious mutations is positive to the system, thus heals the system by improving the phenotypes in functions and capabilities. For example, a signal transduction pathway, regardless of being new or modified, involved 10 proteins. At the end of reshaping process, a few of these proteins fitted with other members of the pathway sub-optimally, resulting in sub-optimal signal transduction. Further mutations shaped the unfit proteins positively or negatively. However, only positive mutations would be preserved as they increased the overall signal transduction in the absence of genome-wide changes that were incurred only in the reshaping process. This was what happened in the healing process. When the mutational effects reach the zero sum state, the phenotypes, the signal transduction pathway in the example, reach the optimal state, the cycle ends, and 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 happens in an evolution cycle and progresses at faster pace. Unlike perpetual natural selection, the healing process is finite.

All species are experiencing non-stop random mutations throughout time, but many codon changes are neutral to protein functions or structures on codon degeneracy. In inevitable cases, amino acid changes exert more or less non-lethal negative effects on the functions of protein molecules, but 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, which can tip 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 is not equal to all species, but differs greatly among species in terms of stability and resilience. Stability and resilience are the tolerance to deleterious mutational changes, the ability to return to the zero sum state after 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 species are quite susceptible to adverse environmental changes and must be confined to certain niche habitats for survival such as some frogs and butterflies. A fragile zero sum state is too delicate to embrace new changes. Consequently, these fragile species would undergo mass extinction when climate or geological changes occurred. In theory, they had adequate space and time to evolve into tough species like rodent and bats, but they have remained vulnerable since their appearance several hundreds of million years ago, an evidence that indirectly rejects natural selection as a universal mechanism of evolution. The zero sum state must be tightly tied to the habitat and will break if a species migrates to elsewhere. In the new environments, mutations that are negative in the old habitat might become positive, thus causing genetic variations of certain heritable traits.

The evolution of species 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 required functionally relevant genes and their protein products emerge and fit well into the system continuously to complete traits or processes. Otherwise, earlier genes can 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 mutational rates first to avoid damages to the organisms and second to avoid lack of timely supply of sufficient new proteins to sustain the cycle to the end. It would be reasonable to predict that the right mutational rates must be much faster than the mutational rates observed in modern organisms. The 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 history. Bats or ants today are no difference from bats or ants million years ago. Evolution cycles increase the biodiversity and the zero sum rule maintains the stability of species and biodiversity. Figure 5 summarizes the achievement of the zero sum rule both in the process of evolution of species and the maintenance of the zero sum state afterwards.

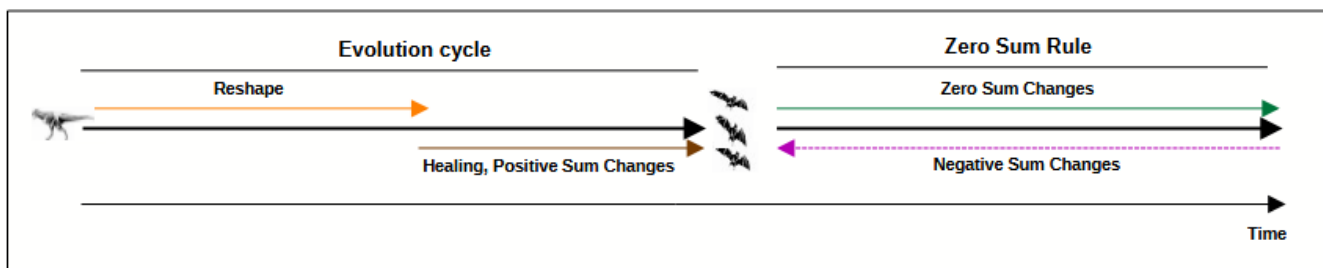


Figure 5. Evolution of species, evolution cycle and the zero sum rule. Any phenotypes in the healing period are not matured yet when measured with the zero sum rule. In the healing process, the net gain of prolonged mutational changes is positive. The healing process ends when the net gain becomes zero, 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.

11. Life Incubator and Nascent Seas

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 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 could be 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, 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 the place as big as a small pond in a neighborhood, or as large as a large pond near a highway, or even as large as a lake full of nutrients for life? Where could such a place be located if it did exist?

According to Wikipedia, the primordial earth formed about 4.54 billion years ago, and its oceans and atmosphere were formed by volcanic activity and outgassing. Water that filled the nascent oceans were partly condensed water vapor from volcanic activity, and partly water and ice from asteroids, comets, and protoplanets. Like the origin of life, the true nature of that primordial earth, especially the true nature of its oceans, is a forever mystery. Nevertheless, we could peep into the mystery on natural resources and life activity on the earth today and infer the bygone physical and chemical environments that were essential to the origin of life.

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. As the plant matter decayed under the heat and pressure without oxygen, they slowly converted into coal over millions of years. Oil is a fossil fuel as well. It is derived from fossilized microorganisms. A vast number of dead microorganism layers settled into the sea or lake bed, where 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.

The oil deposits are discovered around the earth. They vary in size, type, deposit amount, depth under the ground, and geologic location. The world oil reserves are too large to get a reasonable account. The microorganisms based oil formation has some implications for the origin of life, especially it sheds some light on the mysterious environments on the nascent earth called life incubator. A better approach would be to look at the origin of life by correlating the evolutionary path of life with the formation of oil back to the time when life rose from nowhere on the ancient earth.

The abundance of global oil reserves implied that the required amount of microorganisms must be far beyond what the last 500 millions of years could have produced. A large number of advanced marine species appeared in Cambrian period, indicating that the nutrients in the marine water would not be consumed exclusively to grow oil-forming microorganisms like algae and zooplankton. In addition, these advanced marine species feed on smaller and lower algae and zooplankton, which further reduced

the amount of organic materials to form oil. A likely scenario was that oil formation had started far before Cambrian explosion.

Photosynthesis is the landmark on the evolutionary timeline. It allows organisms to capture sunlight to fix atmospheric CO₂ into complex organic compounds such as carbohydrates. Unlimited supply of sunlight as energy source and CO₂ as the main carbon source allowed life to grow and renew in full swing. Photosynthetic organisms emerged around 3.5 billions of years ago. These early organisms used hydrogen sulfide as electron donors to fix CO₂ and release elemental sulfur. Cyanobacteria are the prokaryotic bacteria known to be the first photosynthetic organisms that used hydrogen from water for carbon fixation and released the oxygen as a byproduct. Cyanobacteria can be dated to 2.5 billions of years ago, and over billions of years since their appearances, they are thought to have directly changed the earth atmosphere from being devoid of gaseous oxygen to full of gaseous oxygen. In the absence of photosynthesis, the only organisms able to exist under such anoxic conditions would be the anaerobic bacteria which extracted the chemical energy from inorganic compounds or organic compounds, which were not biochemically produced, but present in the incubator. The organic compounds would be exhausted if they couldn't be replenished fast enough and if bacteria were settled into the soil. With sulfur-producing photosynthesis, organisms achieved limited capacity to utilize sunlight as energy input to fix CO₂. With oxygen-producing photosynthesis, cyanobacteria fundamentally transformed the carbon cycles and the nature of atmosphere, thus allowing life to evolve to more complex and diverse forms.

The emergence of the eukaryotic organisms is a milestone in the evolution of life. Early eukaryotic organisms are single-celled life and like prokaryotes they diverged rapidly into numerous species. Eukaryotic cells are typically much larger than prokaryotes, allowing them to acquire bacteria via endosymbiosis to be part of their own cellular organelles. Mitochondria were energy producing organelles in eukaryotic organisms. They were originally aerobic prokaryotic cells and its aerobic respiration enables the host organisms to yield more energy than anaerobic respiration, providing adequate energy for cells to grow and reproduce. Specialized photoautotrophic organelles chloroplasts were originally cyanobacteria, allowing photosynthesis to occur inside eukaryotic cells, thus producing unlimited amounts of organic compounds for cells to grow and reproduce. Thanks to photosynthesis and aerobic respiration, photoautotrophic eukaryotes evolved into a large and diverse group of photosynthetic eukaryotes phytoplankton, including algae and diatoms. Eukaryotes without organelles chloroplasts evolved into another large and diverse group of organisms zooplankton of varying sizes. Zooplankton must acquire nutrients by feeding on other organisms such as phytoplankton. Familiar organisms that fall in zooplankton include protozoans, metazoans, dinoflagellates, and amoeba.

Single cellularity of the organisms means short life cycle and exponential multiplication if supplies of energy and basic chemicals are ample. More importantly, single cellularity allows mutations to occur and accumulate readily, producing superior individuals from numerous species in a short period that would grow and reproduce more rapidly. At the time organisms capable of aerobic respiration and oxygen-producing photosynthesis were wide spread, marine organisms posed for massive expansion. Although cyanobacteria were abundant enough on the young earth to change the nature of atmosphere, they seemed too small in volume and too low in carbon content to form large oil reserves. It was the arrival of eukaryotic organisms phytoplankton and zooplankton that made it possible to produce organic masses that were large enough to form oil deposits. From evolution point of view, zooplankton are at the top of marine food chain in pre-Cambrian period, while they are also the organic materials for oil formation. Therefore, almost all marine organisms produced before Cambrian explosion could have the potential to be fossilized to form oil.

The vast oil reserves depict a general picture about the nascent earth. The early CO₂ in atmosphere would be far more concentrated than today, and the amount of water in the seas was far more abundant than today as well. Most importantly marine water must contain large quantities of organic and inorganic chemicals and elements necessary for life to form. With oxygen-producing photosynthesis, water cycle and carbon cycles in marine water and atmosphere remained relatively constant before significant oil formation started, as synthesis and breakdown canceled each other. When vast amounts of marine life remains were trapped to sea or lake bottoms in speeds faster than they could decompose aerobically, they removed corresponding amounts of sea water and atmospheric CO₂ with it, and increased O₂ in the atmosphere. The growth of marine life was greatly influenced by nutrient, typically phosphorus and nitrogen, present in the water. Ample nutrient could give rise to massive blooms of phytoplankton, which then fueled a rapid increase or accumulation in the population of predatory zooplankton. The total biomass produced in the period that lasted about 1.5 billions of years until Cambrian explosion would be an astronomical number, part of which were preserved as oil reserves around the world. This period is the oil period on the evolutionary timeline.

Despite its critical importance in our modern society, oil is nothing but a byproduct of evolution of life in the slow evolution stage, an inadvertent, but lavish dividend of the extreme hardship in de novo gene creation left to the evolution end product. Oil formation removed a considerable amount of CO₂ from the atmosphere and profoundly changed the carbon cycle on the earth. In addition, oil formation cleared up marine water by converting vast amounts of organic compounds into oil reserves. In all likelihood, in the oil period the earth underwent a gradual transition away from a quite suffocating one to the one with more fresh air to breathe. At the end of oil period, multicellular organisms had their genomes and protein coding gene counts increased to sizes large enough for derivation and reuse, and the earth environments were transformed to be more favorable for life to proliferate and evolve. Cambrian explosion followed.

Cyanobacteria and algae are microscopic unicellular organisms capable of photosynthesis. They are part of the diverse collection of organisms called plankton that are unable to actively propel themselves against currents but float or drift in water. Cyanobacteria and algae are early species to populate the fresh water and marine water during slow evolution stage. In modern days, these organisms can grow excessively to form algae bloom as the result of increases in nutrient, like nitrogen or phosphorus from industrial and agricultural populations. Algae bloom can be monitored from satellites to determine the location and organisms involved by detecting the kind of chlorophyll and its amount present in bloom. The higher the concentration of chlorophyll, the larger the bloom. Algae bloom occurs often along the coastline where sunlight is abundant and nitrate and phosphate are ample due to water runoff from the land. Such geographical locations and nutrient dependence for algae bloom to occur likely reflected the ancient environments where these grand aged organisms arose billions of years ago and remained to be their preferred areas for growth, old habits die hard. It seemed to provide possible answers to the pressing questions how the life incubator should look like, and where it was once located on the nascent earth.

On the nascent earth, a large portion of earth's surface was submerged under the vast water. The then water is not now water. The then water must contain abundant elements, carbon, oxygen, nitrogen, sulfur, metal ions, and so on. Right after formation, the earth was very hot and its surface was molten. It took a long period to cool down and form continental crust. During this period and under extreme environmental conditions, a variety of chemicals, including amino acids, bases, lipids, small organic compounds, large carbohydrate, and more could have been produced in mass amount and mixed in the nascent seas. All these chemicals were essential to form life. In the tropical or sub-tropical areas near the equator there were coastline locations that formed large bays or gulf covered with shallow water.

Such geographical locations seemed ideal to serve as life incubators. The turbulent nature of the nascent earth and greenhouse gases kept the temperature of the water to certain levels that made basic chemical reactions possible and widespread. Furthermore, the shores or shallow bottoms of the bays or gulf could provide solid surface supports, on which chemical reactions could be made more efficient with the possible catalysis from some active inorganic matter. As polymerization reactions brought about protein, RNA, and DNA, the early life development processes started. Therefore, the life incubator didn't have to have a size or specific geographical location, but anywhere on the nascent earth, where there existed environmental conditions that would allow basic chemical reactions and then life-forming polymerization reactions to occur. If this was the case, the randomness that had led to the single celled life could be exponentially larger than the randomness that could be provided by a single limited life incubator. Because of the universality of a single set of genetic codons, the early form of life emerged likely from a single incubator out of unknown possible ones.

Because of plate tectonics, the earth has been constantly experiencing split of old continental crust to form new continental crust. As a result tectonic forces have caused old continental crust to rearrange into new continents in the past billion years, We simply can't extrapolate the geographical locations of the life incubators from the earth geology today.

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 regrowth and reburial of plants 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 that grew in certain geologic locations were buried and converted into coal, while majority of the dense forests weren't? Therefore, coal formation is harder to think of and seems not related to life evolution.

12. Genotype Configuration and Genotype Reconfiguration

According to Cambridge Dictionary, word "configuration" means the particular arrangement of the parts of something or of a group of things, the way in which all the equipment that makes up a computer system is set to operate. A situation in which small changes are made to something, especially a computer system or software. The meaning of "configuration" sounds so applicable to the life system, albeit far more complicated, entangled, and not obvious. In this regard, the most important and relevant part of the meaning is the particular, not random, arrangement of the parts of something.

From development standpoint, life building blocks amino acids and nucleotides self-assemble into interlinked and inter-dependent large molecules proteins and nucleic acids, and proteins and nucleic acids self-assemble into chromosomes. Proteins, chromosomes, and numerous other chemicals, large or small, self-organize into larger structures – organelles. All the organelles self-organize into the basic form of life – cells. Cells finally self-organize into the highest forms of existence – animals and plants. From the beginning, animals and plants self-evolve into more and more complex and advanced forms, such as birds and mammals over time, leaving trails of evolution, from which all species can find their own positions to determine the time of their origins. Such a grand organization is self-governing and subject only to its own laws since its inception. The laws are written in the genome in the form of genes, including gene regulatory elements.

From configuration standpoint from top down, life, especially life of higher forms, is a grand massive configuration consisting of numerous sub-configurations, and each sub-configuration consists of many of its own sub-configurations, which consists of many of its own sub-configurations, until subs to the

levels of the basic molecules, amino acids, bases, sugars, lipids, all essential chemicals, and more. However, configurations to the molecular levels seem excessive. The genotype configuration is the life configuration at the gene level and is most proper for us to understand the origin, maturation and evolution of life. In simple terms, a genotype configuration refers to a particular functional arrangement of all the genes in the genome, and a particular organism is merely a live instance of a genotype configuration. Therefore, evolution of life is the evolution of genotype configurations per se, and the three stages of evolution of life parallel the three stages, through which genotype configurations have gone through from establishment to maturation to reconfiguration. Genotype configurations or configurations are often referred to as the same thing for the sake of simplicity.

Life arises from total randomness, so does its configurations. During the origin of life, vast randomness generated massive, unlimited amounts of proteins and DNA templates that could encode all of the proteins essential to life in the life incubator. Because of self-assembling nature of proteins and DNA, their complexes and all the necessary components that were randomly produced and present in the boundless incubator were enveloped into membrane bounds, forming singled celled life, which was sustained by minimum genotype configurations consisting of a limited number of genes organized into nascent genomes. They were the most basic genotype configurations that made continuity and evolution of life possible. The greatest significance of these basic genotype configurations is that they serve as the first baseline genotype configurations, from which long term evolution of prokaryotic life started. However, the membrane bounds of the single celled organisms severely reduced once unlimited availability of randomness, which the incubator had relied on to generate a variety of useful genes. As a consequence, it took next 2 billion years for the first baseline genotype configurations to evolve into the second baseline genotype configurations – the blueprints of the first single-celled eukaryotic organisms. At the end of slow evolution, the second baseline genotype configurations accumulated a large number of genes and gene-like sequences, becoming the final baseline genotype configurations. This final baseline was sufficient to serve as the starting point for the fast evolution. Because evolution of life relies on very different fundamentals in the slow and fast evolution stages, the evolution of genotype configurations aren't the same as well in the two stages.

Why did evolution of genotype configurations in the slow stage take about 3.5 billions of years to reach the point of maturity? We can get some thought to this question from Covid-19 pandemic. Covid-19 pandemic lasted about three years and the virus infected majority of the global population. The virus is a positive-sense single stranded RNA genome with a size of about 30,000 bases. It's more proper to say that the RNA genome is organized into a few read frames, not a few coding genes. The first two open reading frames ORF1a and ORF1b are overlapping and account for the first two-thirds of the genome. ORF1a and ORF1b are translated into two large overlapping polyproteins, pp1a and pp1ab. The larger pp1ab polyprotein is resulted from a ribosomal frameshift that allows for the continuous translation of ORF1a followed by ORF1b. The polyproteins are cleaved by the virus's own proteases at different specific sites into a set of small nonstructural protein products, including various replication proteins such as RNA-dependent RNA polymerase, RNA helicase, and exoribonuclease. The remaining genome contains a few more later reading frames that code for the four major structural proteins: spike, envelope, membrane, and nucleocapsid. Some minor reading frames are interspersed between these reading frames and code for the accessory proteins. When so many proteins are packed in such a small genome, Covid-19 RNA genome must be obviously the result of long time evolution.

The RNA-dependent RNA polymerase is responsible for replicating and spreading the RNA virus around the world. It is error-prone and incorporates mutations during RNA replication. As a result, zillions of mutant virus have been produced during the pandemic years. Millions of mutant virus have been sequenced and are available for analysis. There are no standard sequences but reference sequences

for the virus. Many major variants and their sub-variants emerged during the pandemic years and had been identified. These variants and sub-variants displayed increased or decreased infectivity and pathology, but caused symptoms that almost always conformed to the symptoms of Covid-19. In other words, despite mutations on such a large scale, the mutated viruses are still Covid-19 virus.

Covid-19 virus is a strain of coronavirus, while coronaviruses are a group of related RNA viruses that cause various diseases in mammals and birds. Covid-19 virus was believed to be animal origin and became a new coronavirus strain after gaining the capacity to enter human cells through genome recombination with other coronavirus strains. The result of it was that humans became the host for its replication, causing global outbreak of the highly infectious respiratory illness. Detailed genome comparisons among different strains of coronaviruses showed that many recombination events had occurred in multiple regions of Covid-19 virus through co-infection and genetic recombination with both closely related and distantly related coronaviruses that used other animals as hosts.

What can Covid-19 pandemic tell us about evolution? It played a live evolution of a virus on the global scale. If all the open reading frames of the virus is seen as a virus configuration, this configuration is stable and resilient to point mutations of any scales, a brilliant display of the zero sum rule in a small virus in a modern theater with highly developed biomedical technology. In a stable virus configuration, all the elements coexist and co-function in such a harmonious state that they resist all changes that can break such a harmonious state. Lethality of mutations is one of the effective weapons to prevent a virus strain to vary in an unconstrained way. Acquisition of foreign but similar viral fragments via recombination is the most viable way to re-establish a new virus configuration, which isn't too dramatic to change the nature of the virus, but is enough to enable the virus to infect new hosts and change some symptoms it causes. In addition, the Covid-19 RNA genome is compact and terse, and highly successful as a virus in infectivity and life cycle, indicating that except maintaining its viral activities, there is almost no space left in the virus configuration for further changes. This is another major attribute that limits virus's capacity to mutate beyond recognition.

Covid-19 virus configuration is one of the simplest in the living kingdom, but it defines a virus with clear virology and pathology in terms of virtually immutable biochemical and cellular properties. It restricts its infection host to humans and limits variations in genome sequence within the acceptable bounds to maintain the status quo as Covid-19 viral configuration, implying that most of the mutational events are lethal to the virus. From a broader sense, all viruses, regardless of RNA or DNA, share similar life characteristics with Covid-19 virus. They are all defined by their own unique virus configurations to complete their life cycles. If the zero sum rule applies to a virus, this virus is long lived on evolution sense. Otherwise, a virus must emerge and spread for a short period of time only.

As the early singled celled organisms developed into bacteria, their baseline genotype configurations evolved to become bacterial genotype configurations. A major difference from Covid-19 RNA virus is that the bacteria are an independent life entity, and their genomes can grow in theory to any size with no limitation. Furthermore, any new genes that emerge from random DNA fragments via random mutations will be preserved in genotype configurations as long as they aren't lethal to the bacteria, resulting in a configuration in which not all the genes are functional and active. From a single cell population standpoint, it's unlikely that such a population could generate all the proteins necessary to form a complete functional pathway or structure. Taking glycolysis as an example, some enzymes for the pathway were likely to rise independently in vastly distinct bacterial strains. These genes were being transferred between different bacterial strains via bacteriophages, or plasmids, or some unknown mechanisms over time. When all the ten enzymes converged in one bacterial strain in this way, a complete glycolysis pathway was established. Such a pathway is a stable ten-gene sub-configuration

under the grand bacterial genotype configurations. From evolution point of view, creation of such a sub-configuration must be a prohibitively prolonged trial-and-error process. Expansion of bacterial genotype configurations resembles the Covid-19 virus configuration to some extent, partly acquiring new genes from foreign sources via extrachromosome mediated gene transfer, a process of convergence of genes originated from various sources.

The genotype configurations of eukaryotic organisms at the beginning of Cambrian explosion have contained a large number of protein-coding genes that are close to modern mammals even though their real biological roles in the organisms are quite dubious. Multi-cellularity imposes even severer constraints on growth of genotype configurations as any new addition of genes must be coordinated not only in a single cell scope, but between cell types, and later tissue and organ types. Evolution of the eukaryotic genotype configurations in the fast evolution stage must have abandoned the old strategy developed in the slow evolution stage – acquisition of foreign genetic materials via transposon-mediated gene transfer, cell fusions, and endocytosis.

The nucleus provides a safe heaven for the genome to work as a largely independent entity free of interference from numerous cellular and biochemical activities in the cytoplasm. The genomes organize into a set of chromosomes, removing the difficulty of clumsiness in DNA replication and separation during cell division. In addition to the universal point mutations, the genetic machine has evolved to operate in more modes, including gene duplication, alternative splicing, motif insertion, and more. All this has made it possible for eukaryotic organisms to evolve on a self-reliance mode independent of external input in any forms. One consequence is that new genes added to the configurations are largely not of de novo creation, but from derivation and reuse of existing genes via point mutations, gene duplication, alternative splicing, and so on. The strategy of derivation and reuse is effective and powerful to enrich and expand the existing genotype configurations into new species.

Reconfiguration is the process that makes a new or different arrangement or pattern of a group of related things. In evolution, genotype reconfiguration is dynamic and continuous, and a new or different arrangement or pattern is attributed to changes in the groups of related things encoded in the genome. The changes are mostly the results of modification via addition and deletion. In reconfiguration, new genes could be generated through different genetic operations, including, but not limited to preservation by inheritance, variation by duplication, alternative splicing, drifting in sequences, motif, creation from pseudogenes and random sequences, and deletion by lethal mutations, and degeneration to become pseudogenes. During reconfiguration, any genes that are produced by variations, reuse, and creation from pseudogenes are instant genes to distinguish from tardy genes that are generated by creation from random sequences.

To make genotype reconfiguration clearer, genes in a configuration can be divided roughly into three groups. Majority of the genes fall in group one. They are generally conserved in sequence and function to maintain the basic biologic processes, activities, and structures of the eukaryotic organisms, for example, enzymes catalyzing basic metabolic pathways and genetic operations. Genes responsible for the morphology of a species can be classified into group two. They subject to changes in large degrees between classes, even between orders in the same class. The changes are very noticeable in phenotype. The group two genes largely determine where a species will be placed in the biology classification system. For example, deletion or degeneration of the limb inducing factor genes gives rise to limbless animal snakes, which are classified into Reptilia class and Serpentes suborder because of their limbless morphology. All other genes, including all the instant or existing variant genes and those tardy genes, are group three genes. Group three genes can vary greatly in sequence and function from small to large, accounting for the majority of sequence changes during the reconfiguration not only between ancestor

organisms and new species, also among new species descending from the same ancestor. It's the group two and group three genes that support a particular morphology and furnish each species with its own unique characteristics and behavior. Therefore, genotype reconfiguration is largely about genes in group two and group three, which together set one species apart from all other species in the living world.

Like viral genotype configurations, one genotype configuration defines one unique species, and one unique species is determined only by one genotype configuration. What's different is that animal genotype configurations are far more complex in terms of the number of genes, genome size, and variations in gene structures. Comparing with small and compact configurations, large genotype configurations have the advantage of having a much larger capacity to develop into more complex and advanced configurations. On the other hand, the increasing complexity generates greater constraints on reconfigurability, thus diminishing the potential for further genotype reconfiguration, and at last bringing evolution to a stop.

Genome data shown in Table 2 seem to suggest that genome sizes and protein coding gene counts vary more broadly in low and simple species like *C. elegans*, *D. melanogaster*, *Ciona intestinalis*, and Sea lamprey than in higher and more advanced species like fishes and amphibians. An implication is that many genes in these low species are included to calculate the protein coding gene counts, but in reality they shouldn't be included when their biological roles are concerned. The listed counts are more like some nominal numbers, and the true counts are similarly smaller in number across different, but evolutionarily comparable species. In other words, any genotype configurations contain varying numbers of genes, which look like functional, but are actually non-functional. These genes constitute latent gene reserves for reconfiguration once an evolution cycle begins.

The fast evolution started from the final baseline genotype configurations at Cambrian period and progressed in successive evolution cycles thereafter. Any genotype configuration that started an evolution cycle served as the basal genotype configuration for that cycle. In the early phase of Cambrian explosion the genotype configurations were almost all translated into invertebrate species until some 518 million years ago, at which time genotype reconfiguration had reaped the greatest dividends in the evolution of life, resulting in numerous species of both invertebrates and vertebrates.

Organisms are placed in different groups based on shared common characteristics. These groups form the biological classification system. If a group sits high in the hierarchy of the system, organisms under it share fewer common characteristics, while a group sits low in the hierarchy of the system, organisms under it share more common characteristics. From evolution standpoint, if some characteristics are shared by all the species that fall into a group of higher hierarchy, they must appear early in evolution, which allows us to peep into how genotypes are reconfigured over time.

As high as 97% of animal species are invertebrates. However, it is hard to estimate the number of present-day invertebrates that emerged before Cambrian period, as most of them appeared during or post Cambrian period. Therefore, it's hard to estimate the number of baseline genotype configurations that started the Cambrian explosion. Nevertheless, the baseline configurations in theory must be potent to undergo a variety of reconfigurations into species with numerous different morphologies and body plans, on the premise that being baseline often means versatile in expansion. Without doubt, it didn't disappoint. Phylum Arthropoda is the largest group of up to ten million species, including insects, spiders, scorpions, ticks, mites, shrimps, prawns, crabs, lobsters and crayfish. Phylum Mollusca is the second largest animal phylum after Arthropoda. Numerous slugs, snails, clams, oysters, cockles, squid, mussels, scallops, cuttlefish, and octopuses all belong to this phylum. Phylum Onychophora, by

contrast, is small with approximately 200 species known as velvet worms so far. It's a mystery how many evolution cycles have gone through from the final baseline configurations to arrive at these phyla. The remarkable inequality in the size of different phyla indicates that not all basal genotype configurations are equal in reconfigurability. Larger the phylum size, greater the reconfigurability of the basal genotypes for the phylum.

Insects represent more than half of all animal species, and more than a million have been described. Their extraordinary diversity makes it difficult to classify. Because of this, insects are simply divided into two groups: wingless insects and winged insects. Insects in both groups have a three-part body (head, thorax and abdomen), a chitinous exoskeleton, three pairs of jointed legs, compound eyes, and a pair of antennae. It is truly remarkable for these characteristics to be shared by millions of species. From the evolution cycle standpoint, these features must have appeared very early on in the cycles and made species evolutionarily prolific if bearing them. Thus they are the foundations upon which all future diversity of insects are built. From the reconfiguration standpoint, after the basal configurations acquired unknown numbers of new randomly generated genes, both instant and tardy, with useful functions, they would differentiate into tremendously different configurations early on. The dominant configurations included gene clusters that encoded the above common characteristics for further evolution. The reconfiguration process was explosive, leading to diverse orders of early batches of insects on the evolutionary timeline, although the number of species in each of the early orders was relatively small yet.

Evolution of insects didn't stop there. Remarkable success in the appearances of more advanced insects occurred much later when the early insects served as basal genotype configurations. Beetles appear about 300 million years ago, flies about 250 million years ago, and moths, wasps, bees, and ants about 150 to 66 million years ago. The number of species in each of these later orders is much larger than early insects. For example, about 400,000 species for beetles, 150,000 for butterflies and moths, 150,000 for flies, and 117,000 for sawflies, wasps, bees, and ants. These large numbers indicate that the morphological characteristics common to all insect species in an order are highly effective for churning out a large number of species, reflecting high reconfigurability of their basal genotypes.

The exact time the early insects appeared couldn't be determined, nor their ancestors. Insects are classified into subphylum hexapods, and share morphological similarities with their closest relatives Diplura, an order under class Entognatha, which is at the same level as class Insecta. Under such a classification scheme, insects and species in the order Diplura shared common ancestors, but differ significantly in morphology, anatomy, physiology, internal system, and more, suggesting considerable divergence of the two groups since departing from their common ancestors. From an evolution standpoint, evolution of early insects and Diplura must be more challenging than evolution down the path, since in general, invertebrates at the early phase of Cambrian period were developmentally not close to insects and Diplura at all. Therefore, early genotype reconfiguration was to develop a number of tissues and organs, which required more new genes, both instant and tardy, to be generated and reconfigured, a trial-and-error process that was far lengthier and harder than simply advancing the functionalities and structures of existing tissues and organs. Even though insects and Diplura might share some common evolutionary trails, the outcomes at the ends of their respective cycles were totally poles apart. Insects account for more than half of all animal species, while class Entognatha contains only about 20,000 species total, a clear stunning failure in the evolutionary history. It also indicates that the genotype configurations of early insects are extraordinarily reconfigurable almost to the end of the cycle trails, leading to a large number of species that differ only slightly in morphology and physiology. The genotype configurations of the insects are an undisputed winner of evolution, revealing the pivotal roles of the reconfigurability of a genotype configuration in the evolution of species.

According to zero sum rule, the genotype configurations of the present-day invertebrates are stable with little chances of undergoing reconfiguration. It was certain baseline genotype configurations in the Cambrian period that led to the genotype configurations of the earliest vertebrates, one of which was jawless fish from 520 million years ago. The configurations of jawless fish weren't versatile in terms of reconfigurability, as indicated by about 120 living jawless fish species in total today. However, some fishlike vertebrates that appeared in the Cambrian period formed the basal configurations for jawed fish, which were far more reconfigurable than jawless fish, making jawed fishes the largest group of vertebrates, accounting for more than half of extant vertebrate species.

Comparing with the extraordinary reconfigurability of invertebrate genotypes, jawed fish genotypes exhibited greatly reduced reconfigurability because of their much more complex nature. However, it's remarkable relative to amphibian species, suggesting that their configurations still had ample room for advance. Jawed fishes are divided mainly into two types, ray-finned fish and lobe-finned fish, but lobe-finned fish was blessed only with limited evolution potential, accounting for only 4% of total present fish species. Over 26,000 fish species are ray-finned fish grouped in about 40 orders and 448 families, indicating unusual reconfigurability of their genotypes. Significant inequality in the genotype reconfigurability of the two groups suggests that after divergence from the basal configurations, the lobe-finned fish configurations could be considered an awful failure in reconfigurability. Some new genes arising in the lobe-finned fish during reconfiguration might have resulted in some morphological and physiological changes that made survival rates of the intermediates low or made new fish species ill-fated for survival in changing environments after the cycle. By contrast, ray-finned fish genotypes were much more reconfigurable, allowing intermediates to diverge more freely and acquire diverse morphologies and physiology that provided them with a common capability – agility of movement in the water, thus partly increasing their survivability.

Lobe-finned fish seemed to be a dumb end for fish, but one of its species might have become the basal genotype configurations for amphibians. The genes responsible for the fins underwent reconfigurations to become limb-like fins first and then limbs, which enabled them to crawl and move onto the dry land as early as around 370 million years ago. The transition from water species to partial land species is a long and costly endeavor from the evolution standpoint, requiring extensive genotype reconfigurations to endure an enormous amount of changes in morphology, anatomy, and physiology necessary for the terrestrial environment. Comparing with reptiles, birds, and mammals, amphibians are still low on the evolutionary tree, but class Amphibia for amphibians is relatively small with 8,000 species, of which nearly 90% are frogs and toads, indirectly indicating that what happened in the overall reconfigurations from jawed fish to amphibians are more than what happened from amphibians to reptiles or mammals.

Because of the complexity and diversity of the vertebrate animals and lack of useful fossil records, their true origins, thus their basal genotype configurations, are hard to trace. In an evolution cycle, all intermediates were free to diverge via random mutations, and the basal genotype configurations diverged quickly as instant genes emerged constantly throughout the cycle in different configurations. Enormously varying reconfigurability of individual genotypes resulted in great inequality in the size of an order or family. For example, lizards and snakes account for about 96% of 9546 reptiles total, and passerine birds account for half of 11,000 bird species, while rodents account for 40% of about 6,600 mammal species, followed by about 22% bats. The genotype reconfiguration seems to show strong bias towards some configurations, indicating that certain particular morphology and physiology that randomly appear early in an evolution cycle are commonly favored by evolution and have determined their future outcome down the path.

13. Genotype Reconfiguration, Evolution Cycle, and Zero Sum Rule

Modern evolution theory seems to suggest that evolution and biodiversity are two apparent conflicting things. Biodiversity is the result of evolution, but evolution is ruining biodiversity by placing species under constant evolutionary pressure to become new species. However, it isn't what we are seeing in our time because the spectrum of modern life covers species of all complexity. Evolution is a long term project, and its outcome has been enriching the biodiversity periodically since Cambrian time. Any visible decline in biodiversity isn't owing to evolution, but the result of environmental destruction by geological or climate changes, recently by human activities. Constant evolution and highly conserved biodiversity have been made possible through genotype reconfiguration, evolution cycle, and zero sum rule, the three foundations of all life on the earth at least since Cambrian period.

Genotype reconfiguration focuses on the gene level to look at evolution of species, while evolution cycle focuses on the species level to look at evolution of species. And together they provide a more tenable account of how species evolve in the fast evolution stage. Genotype reconfiguration parallels the evolution cycle from the start to end, resulting in new species with new genotype configurations. It's the zero sum rule that provides a genetic basis to bring the process of genotype reconfiguration to a stop. All of the above sheds light on some operational details about an evolution cycle – disarmed state, armed state, process genotype reshaping, and process genotype healing.

Evolution cycle starts when the genome transits from a disarmed state to an armed state upon genetic perturbations caused by environmental changes. Genotype reshaping process dominates the early phase of the armed state, resulting in large scale changes to the basal genotype configurations. As the reshaping process lessens, it transits gradually to the healing process. The healing process is the continuation of the reshaping process, and conceptually, it is to heal the “damages” the genome has incurred in the reshaping process. Therefore, any changes to the configurations in the healing process are mild in nature to smooth out those reshaped biological processes, activities, and structures, a characteristic of the later stage of the cycle. From the zero sum rule point of view, the reshaping process has incurred net losses from prolonged mutational changes, while the healing process gives rise to net gains from further mutational changes. If any genetic changes to the configurations result in neither gains nor losses, the configurations have reached a zero sum state, the end of an evolution cycle. Intermediates that have reached the end of an evolution cycle are new species of the cycle, which are in a new disarmed state. The zero sum rule ensures that the new species will be the same species forever by maintaining the stability of their configurations.

Evolution cycle, on the high level, adequately explains why evolution of species occurs in an explosive mode and provides the evolutionary basis of feature commonalities shared by species, the backbone of modern biological classification system. Genotype reconfiguration, on the other hand, reveals the bottom line of evolution. Evolution of species is nothing but a continuous process that makes different arrangements of genes contained in the configurations, a process called reconfiguration. The results of the arrangements are new genotype configurations when viewed from the end products. Each distinct arrangement, a result of composition of both old and new instant and tardy genes, was a unique genotype configuration, an instance of which is a unique new species.

A genotype configuration not only includes all protein coding genes, pseudogenes, regulatory genes in the genome, but also refers to a particular arrangement of all these genes in terms of their biological functions and roles in a life cycle, as a general flat genotype configuration doesn't really have much utility for our purpose. A better approach is to group genes in a configuration into different sub-configurations according to their functions and roles or according to specific uses for the study. The word “config” is used widely in computer systems, and is borrowed here to mean a sub-configuration.

Therefore, a config is simply a sub-configuration. A config can have its own sub-config, and a sub-config has its own sub-config, and so on. There is no necessity to put any constraints to limit sub-grouping of a config to a particular level. The purpose of grouping and sub-grouping is to establish a detailed map, which shows the relationship of genes in a biological system to untangle the biological mess into clearly defined processes, structures, and activities in the language of genotype configuration. Therefore, a genotype configuration can be compiled or sub-grouped in any forms to suit studies.

Genes that form glycolysis pathway are a great example of a config. And similarly, enzymes that form citric cycle are another example of a config. Glycolysis config and citric cycle config are not related, but can be grouped as sub-configs into a larger or parent config, sugar metabolism config or energy production config, whatever you want to call it. However, not all genes can be grouped or sub-grouped with the clarity of the glycolysis pathway. Compilation of a genotype configuration with details about all of its genes isn't an easy task. Taking echolocation config in bats as an example. Because this capability is overwhelmingly complicated, involving an unknown large number of genes, it requires great effort to compile all genes involved in echolocation from vast genome databases and research literature to make the echolocation config as complete and accurate as possible with sub-configs and even sub-sub-configs.

A well compiled or constructed genotype configuration of a species is like a well designed data structure in computer science, it holds a vast amount of data in a structured fashion. Any piece of data in the configuration represents a group of genes that work together to carry out a particular biological function or role. When such a piece of data are sub-grouped into a few smaller groups, it clears up more concrete roles of individual genes in the group at the lower levels of granularity. Therefore, if a genotype configuration is well compiled, it offers great advantage for evolution study. Many configs, like glycolysis config and citric cycle config, can be excluded from evolution study, because they don't change much between species across different classes or orders. The emphasis can be focused on configs that separate one species from others across the evolutionary tree.

A genotype configuration of an ancestor species was the basal configuration, from which a genotype reconfiguration process started at the onset of an evolution cycle. Random mutations of any kinds would result in random appearances of numerous instant genes and some tardy genes over time, driving evolution forward throughout the post-Cambrian history. Random releases of new protein products into cytoplasm would reshape the biochemical processes and cellular structures and activities with unpredictable consequences. If an instant gene was produced from one of the glycolysis genes, its impact would not be noticeable as long as it wasn't disrupting to the glycolysis path. Majority of instant genes, especially those derived from group one genes, didn't have much effects on the evolution of the intermediates. It seems that instant genes that were derived from group two or group three would change the morphology and physiology of the intermediates more or less.

It could be postulated that in evolution cycles from fishes to amphibians, a gene in the fin config became an instant gene that coded for a variant of the fin inducing factor. This instant gene would play a critical role by changing the morphology of the fins into limb-like fins, and could be considered a pivotal gene for the evolution of amphibians. Therefore, the emergence of the fin variant must be among the first sign of evolution towards amphibian species. As more genes emerged to make limb-like fins to develop into hands and feet with five or more digits, a new, evolutionary landmark config – limb config was established, completing the unusually complicated morphogenesis to bring animals from water to land. At the end of the cycle, the intermediates that survived the cycle formed a new class Amphibia. Amphibians are mostly semiaquatic and their limbs and feet allow them to inhabit a wide variety of habitats, able to live in freshwater, wetland or terrestrial ecosystems.

When fishes moved from water to land, their genotype configurations must have undergone numerous genome wide changes in addition to limbs. Other essential massive changes included reconfigurations of the configs for respiratory system, morphology, circulatory system, digestive system, nervous and sensory system, and many more. The scope for changes was so broad and basic in its evolution cycle that the number of trails that would reach the ends of the cycle was expected to be quite limited, thus the number of orders. This had been indirectly demonstrated by the fact that amphibians are classified into 3 extant orders and species in different orders display quite different morphologies. Furthermore, their life cycle still retains aquatic phase, in which the larvae must first complete their development in water before terrestrial phase. It would be the case that the amphibian species would not exist if any one of the above crucial configs failed to develop into one suited for land life.

The evolutionary significance of the limb-like fins is far-reaching. It formed an architectural platform on which all new features observed on amphibians had been built. Limb-like fins allowed restless intermediates to move out of water and expose to air directly, which could have occurred at the early stage of development, thus preserving new early genes if they could be part of the primitive lung config. Otherwise those new genes would be useless for the fishes and lost anytime without being noticed. Instant and tardy genes must appear randomly in vastly different time periods throughout the cycle. If any of new genes could play roles in moving the intermediates to the dry terrain, they would be blended into some configs and preserved as new gene members of those configs in amphibians. Transformation of fish to amphibians is the result of accumulation of numerous such genes over tens of millions of years. It's the very nature of genotype reconfiguration, a random and prolonged trial and error process. The pivotal roles of limb-like fins have been manifested by the fact that except caecilians, which are limbless, worm- or snake-shaped, all amphibian species bear hands and legs, indicating their limb-like fin origin from the ancestor fish.

Bats form the second largest order of mammals and are sub-divided into insect eating bats or microbats and fruit eating bats or megabats. Insect eating bats are known for their echolocation system for navigating and hunting for prey. Echolocation system isn't unique to microbats, and it works for some whale species. Simpler forms of echolocation can be found in other animals, including cave-dwelling birds cave swiftlets and oilbirds, and terrestrial mammals shrews and some rodents. Genes involved in echolocation could be grouped into an echolocation config. In bats the echolocation config is complex with possible multi-layer config structure, responsible for a sophisticated, full-fledged echolocating capability. There are evidences indicating that convergent evolution is involved in other animals.

From the evolution cycle standpoint, microbats and megabats were bats because they shared a common evolution trail before they diverged. The appearances of some genes with ultrasonic responsiveness in one of the microbat intermediates formed the base echolocation config, from which the complex echolocation system was developed, resulting in the microbat lineage capable of echolocating. These genes must be the earliest anchor members of the echolocation config required to start the development of the system. It was expected that a brand new biological system as complex as echolocation would allow a considerable degree of divergence during evolution, which partly explained why bats could become the second largest order of mammals.

Although all microbats enjoy the similar echolocation function, evolution of the echolocation configs must vary considerably among microbat species. Genes in the final configs appeared randomly in time, implying that each echolocation config among intermediates contained quite different sets of genes at any moment, and the differences were the greatest in the early phase of reshaping process. As the cycle approached the end, echolocation configs shared more common genes among intermediates which

shared longer common trails, and defined echolocating systems that were more similar in structures, functions, and physiology. Because of this, degrees of sequence homology of each gene in the configs varies considerably among microbat species, ranging from high to medium to low, a result of their evolutionary origin, not mutations afterwards. In addition, formation of echolocation also heavily depended on the reconfiguration of other configs, such as taste bud config, vision config, ear config, toe config, nervous config, etc. Until all of these dependent configs were in place and compatible with each other, the echolocation had achieved a full capability. When any changes to the echolocation config resulted in the net loss of the capability, the config had entered a zero sum state, and evolution of the system completed.

Evolution of giraffe is another example of genotype reconfiguration. As discussed earlier, the giraffes and the okapi belong to the same family, but different genera. Despite great morphological differences on first sight, they share a number of common features, including a long, dark-colored tongue, lobed canine teeth, and horns covered in skin. The okapi is the only extant species in its own genus, similarly giraffe genus is considered to have one species with nine subspecies. From the morphology standpoint, the okapi is closer to their deer-like ancestor, while giraffes departed from their deer-like ancestors far and away, suggesting that genotype reconfigurability of the giraffe is far higher than that of the okapi. There were a large number of species that showed intermediate morphologies between deer-like okapis and giraffes, for example, their necks were shorter than giraffes but longer than deer-like animals, but they had gone extinction for whatever reasons. All this indicates that genotype reconfiguration in the evolution of giraffes had suffered from awfully low success rates.

Okapis and giraffes share the same family, indicating that they shared only a short common evolution trail. There are two possibilities for the separation of two species. They diverged from the same intermediate that had developed early instant genes that induced long neck growth. However, an overwhelming majority of the later intermediates were unable to bring in all the instant and tardy genes necessary to support a transition to a long neck, resulting in only two species in the family, the giraffes with a full transition and the okapis with a short aborted transition. As usual it implied that although numerous instant genes appeared continuously along the evolution trail, randomly at vastly different time, genes required for the long neck development were rarities along the trail. Or one intermediate produced the early instant genes for long neck growth after it separated from the okapis, resulting in giraffes being the only evolution product thereafter.

If we assume that giraffe's neck config consisted of many genes that were organized into several sub-configs. An intuitive way to see the evolution of the neck config was that some instant genes had played deterministic roles in the initiation of such an unusual neck growth. In parallel with the reconfiguration of neck config, many other non-neck configs must be reconfigured as well to support an ever growing long neck, including cardiovascular config, skeletal muscle config, nervous system config, and so on. Due to random nature of the appearances of instant genes and tardy genes in the cycle, reconfigurations of the non-neck configs were not intended for the sake of the neck config, but occurred as normal part of the cycle. By extremely low chances, only when the results of non-neck reconfigurations happened to meet the demand for the neck that was growing longer and longer, a long neck would become possible. For example, the heart config, a sub-config of cardiovascular config, could undergo reconfiguration by random instant genes, resulting in an occasion in which the heart was larger than the size needed for a normal sized ancestor species. Did such an oversized heart do good or bad to the intermediates? A likely answer was that it did bad to the intermediates because the normal chest wasn't large enough to house the heart or it was an overkill for the species. However, if the neck went unexpectedly long, requiring a stronger and bigger heart to pump more blood to the head, the coming of an oversized heart together with a more spacious chest would be an opportune event for the

neck to grow long. Oversized hearts due to reconfiguration of heart sub-config must be a common occurrence in the evolution history in all vertebrate animals, but rarely there are species that carry oversized hearts, suggesting that oversized hearts must be detrimental to the intermediates, even to the species.

It is a fact that new instant genes and tardy genes that constantly appeared in the cycle were random, individual, unrelated with one another with huge time gaps up to millions of years, indicating that the order and timing of their appearances weren't always critical for reconfiguration, if their sudden presence didn't have lethal consequences to the survival of intermediates. If functional interdependence of individual configs in the configuration could be re-established with the continuous appearances of random genes, a trail in the evolution cycle would end up with a new genotype configuration, thus a new species. However, as stated earlier, this was an immensely prolonged trial-and-error process with low success rates. The rates would be even lower if the new configurations differed from the basal configuration in width and depth too extensively. Giraffes' neck configs in the great majority of intermediates never had chances to attain the configs that were able to support a fully functional and unusually long neck, as most of them failed the reconfiguration processes as needed, resulting in the giraffes being the lucky lone member in giraffe genus. It also showed that species with unusual and unfavorable physical morphologies are hard to come by from evolution point of view.

We could make a tentative conclusion that the final successful reconfiguration of a genotype was the sum of successful reconfigurations of all the configs, the result of which was a genotype configuration, in which all new genes had been incorporated into configs they belonged to, and performed functions as they were evolved to perform. The new species were made exactly as specified by their genotype configurations, such as long necked giraffes, echolocating bats, corpse eating vultures, etc. If reconfiguration of any one of the configs broke down the config interdependency, intermediates would never reach the end of the cycle.

The distinct differences in morphology between giraffes and okapis are rooted in their genotype configurations. As stated earlier, only about 20% proteins are identical in the two species, and most differences are genes engaged in metabolism, skeletal development and differentiation, nervous system, and cardiac muscle contraction, all of which support the view that the neck can't grow out of range on its own, but is a result of concerted reconfiguration of many configs achieved via body-wise mutational changes in a cycle. As the cycle was near the end, some small variations in the final genotype configurations took place among some of the intermediates, resulting in divergence into nine very similar sub-species. The genotype reconfiguration of deer-like ancestor was poor for okapis, and most likely the process was terminated early in the cycle, probably explaining why the number of species in the genus *Okapia* is limited to a single one.

The last example of genotype reconfiguration is from the primate color vision system. Primates are an order of mammals, accounting for about 8% of all mammalian species. Primates arose 74–63 million years ago from small terrestrial mammals. One of the characteristics of primates is color vision, which is the result of an extra type of opsin molecules L-opsin in the retinal. L-opsin is absent in the ancestor species and all other non-primate mammalian species. The opsin config is a sub config of the vision config. The reconfiguration of the opsin sub-config generated L-opsin gene, an instant gene via duplication of M-opsin gene. However, an extra L-opsin gene alone would not give rise to a full color vision in primates. A whole lot of events must take place during the overall genotype reconfiguration. At least, regulation of gene expression must ensure that there were photoreceptor cells that were dedicated only to L-opsin gene. Furthermore, L photoreceptors must not line up the retinal surface in cluster, but distribute evenly with M photoreceptors. Any genes involved in this regard must be part of

opsin sub-config. Sub config for photo signal transduction pathway must be reconfigured to accommodate and process vision signals generated from L-opsin, and work with signals from M-opsin to determine the intensity and wavelength of incoming light, as required for full color perception. It might involve some instant genes to produce variants of neurotransmitter receptors to manage extra nervous signals. In central visual cortex, the nerve network must undergo extensive changes to decipher the upgraded visual signals to perceive full colored images of the external objects. Unless all aspects of a color vision system have been reconfigured to function as a well balanced and coordinated unit, the vision wouldn't be fully colored despite L-opsin.

From the evolution cycle point of view, vision reconfiguration started in the reshaping process. The initial appearance of L-opsin instant gene was likely to interfere with the normal vision of the ancestor, which was not fully colored. From the zero sum rule point of view, the initial changes must incur a net loss to vision. This net loss was the sum of all changes to the vision config plus all its sub configs during the reshaping process. Only when all the necessary instant genes could appear randomly to complete all sub-configs involved in full color vision, the net changes to vision would turn to a net gain, symbolizing the start of the healing process. Genotype reconfiguration in the healing process was mild to polish the vision config to full color capable and reach the zero sum state.

The reshaping process of an evolution cycle is more like home building in progress, while the next healing process turns the completed main structure into a well finished home with fine details and pleasing views. The major difference is that in both the reshaping and healing processes, new materials are produced randomly and most of time either garbage or don't fit well with other components, new or old, thus requiring nonstop trials and errors. This is especially problematic in the work in progress phase. Therefore, genotype reconfiguration is a risky business with overwhelming chances of failure.

As species approached the top of the evolutionary tree, genotype reconfiguration seldom gave rise to vertebrate species with peculiar or unusual morphologies. Starting from amphibians, vertebrates can be grouped into two general types, legged and legless. In amphibians, legless species are grouped into one order, accounting for only 2.5% of the total. Legged animals after amphibians can have their front limbs transformed into wings, giving rise to winged dinosaurs, birds or bats. It implies that genotype configurations of the post-fish species have lost great morphological extensibility beyond this two general types. Reconfigurations could bring about morphological pattern changes only within the limits allowed by the basal configurations of amphibians. Furthermore, the genotype configurations of amphibians already had evolved a relatively complete set of configs for many organs common in higher species, including limb, skin, heart, liver, kidney, blood circulation system, colon, stomach, small intestine, lung, spleen, reproductive system, nervous and sensory systems, etc. After amphibians, reconfigurations that occurred in all evolution cycles reshaped all those configs to great extent without exception, bringing them more sophistication and advanced functions, structures, and physiology, although the basics were never changed across classes over time. Therefore, genotype configurations didn't have the freedom to undergo any reconfigurations that random mutations could incur on them. In other words, the genotype configurations of the ancestor organisms had constrained and determined what kind of new species could be, which in a broader sense was a type of inheritance. Any reconfiguration that could damage any one of the above organs and result in the loss of functions or structures would not be survivable.

14. Genotype Potential Energy and Zero Sum Rule

Genotype reconfiguration is a process driven purely by random genetic changes. Randomness has two faces. Randomness means immense possibilities to generate something good, and it can also destroys a well established system. Evolution has been advancing in such a paradox from the very beginning. The

zero sum rule indicates that the species today don't undergo reconfigurations, because reconfiguration will result in net loss of functionality, ensuring the lasting massive biodiversity. Then what kind of genotype configurations is prone to random genetic changes or reconfiguration?

In chemistry, chemical potential energy is a form of energy related to the structural arrangement or the configuration of the atoms in the molecule. Chemical reactions take place because it will naturally let molecules go from a higher chemical potential to a lower one and release free energy at the same time. Molecules therefore with higher chemical potential energy are less stable and chemically more reactive. From the standpoint of chemical potential energy, the vulnerability of species to external perturbation could also be considered an index of the instability of their genotype configurations. The concept genotype potential energy seems useful to measure the stability of a genotype configuration in a sort of quantitative way. However, unlike chemical potential energy for doing work, genotype potential energy isn't any form of energy per se, but an internal potential to fuel genome-wide changes in the face of geological or climate changes. If new configurations at the end of a cycle varied in their genotype potential energy, configurations with higher potential energy would be less stable and more vulnerable to external changes.

The genotype configuration in a nutshell is a collection of all the genes in a species, in which the genes are well organized into a hierarchy of configs possibly with multiple layers of sub-configs. Each config defines a unit of work or function at the highest level in a biological system, and each of its sub config defines a sub unit of work or function as a functional component of the parent config. Indeed, the way the configuration is organized is chosen for the purpose to clearly illustrate the intrinsic relationships between genes present in an organism to suit a particular research. The genotype configuration of a species is unique and stable against environmental pressures as it is governed by the zero sum rule. The configuration therefore is largely unchanging over time, and the instance of this configuration is simply a living organism of this species that has changed little since its inception.

The genotype reconfiguration refers to a genetic process that occurs in an evolution cycle to make numerous changes to the basal configuration, resulting in a different genotype configuration, which is characterized with new arrangements of old and new genes across all configs. The majority of active genes in the genome are housekeeping genes and they comprise the old gene population in the reconfiguration process. These old genes became gene variants of the ancestor organisms if they were changed by random mutations in the process. More importantly, random genetic changes would bring up a substantial number of new genes in all configs, and most of them were instant genes. These instant genes were largely variants of genes active in the same species, for example, L-opsin gene in primates. A small number of tardy genes accounted for the rest of new genes. Tardy genes and some of the instant genes could carry out functions that were absent in the basal configuration. A new GPCR variant, for example, could respond to a new ligand to start a new signal transduction pathway in new species, possibly forming a new sub config of its own.

All genes in the configuration are equally impacted by random mutations in the cycle, and most of the random mutations are deleterious by preventing genes or their protein products from functioning properly or completely. Therefore, the reconfiguration was a process that is highly uncertain for success from the standpoint of individual genes. The expression of housekeeping genes was obviously vital to survival, and any serious functional changes could have lethal consequences. Good thing was that all the intermediates that carried lethal mutations would not survive a few generations, and thus be eliminated naturally from the intermediate population, leaving only viable intermediates on the evolution trails. Bad thing was that the reconfiguration process would be abolished if no intermediates had survived. This is the natural elimination rule that has ensured that only random mutations of non-

lethal nature are retained by the evolution since evolution started billions of years ago if the genes were functional useful to the organisms.

It's mostly the instant and tardy genes that brought forth new morphologies and more advanced and complex biological processes, features, and characteristics to the survived intermediates, qualifying them to become new species. Because most of the instant genes were variants of some of the active genes and their functions and biochemical properties were largely comparable as well, making their integration into sub configs more readily. On the contrary, some instant genes and all the tardy genes formed a group of anonymous genes with unpredictable biochemical properties and functions. These genes would impose unknown consequences on any sub configs. While intermediates that carried lethal new genes would be eliminated from the population immediately, the biggest challenges were if there were any anonymous genes that could be integrated into some of the sub configs to bring to maturity complete pathways, structures, functions, etc., or if there were any anonymous genes whose products could provide the organisms with some advantage or capability for survival. There were many possibilities how the new genes could become new members of some sub configs or form their only sub configs. A critical point was that they emerged randomly over a period of tens of million years, and their functional status would subject to change as the cycle moved on. It is certain that the fates of any new genes would be largely dependent on the other new genes, emerged before or after. If they could comprise functional groups with old sub configs or form new sub configs with other new genes under proper parent configs at the certain time points in the cycle, they would be preserved as part of the configuration. In the meantime, intermediates had moved one step forward to become new species. The remaining anonymous genes would simply sink into oblivion. Therefore, only time would tell if any of the anonymous genes would find places in the configurations and survive an evolution cycle.

Given the redundant and interdependent nature of configs in organisms with organ differentiation, the life of an organism is the result of concerted action of numerous features, activities, and processes defined as configs in the genotype configuration. In other words, all configs in the configuration are compatible with one another in their roles to support a sustainable life. Therefore, merging into existing sub configs or forming their own sub configs under parent configs is not sufficient for new genes to be fitted into the configuration. All configs that were changed by any new genes must remain compatible with all other functionally related configs in order to achieve an absolute unity of functions and structures on all levels. This is a requirement that must be satisfied for the process of reconfiguration to be productive at the config level at least under a given natural environment.

Therefore, it could be concluded that any new genes could be preserved into some configs only if dependent changes occurred in other functionally related configs, and vice versa. Unsynchronized appearances of new genes with random functions and properties over the period of an evolution cycle had serious biological implications. All intermediates must have experienced prolonged mutational changes that could incur more or less severe pain and distress to their physical conditions, even though some of them had survived to become new species. In some sense, the reconfiguration process could have been one of the longest trial and error process that brought all configs to a compatible state in terms of functions and cellular activities, during which overwhelmingly most of the intermediates perished. However, it was unavoidable for the evolution to continue and diverge into numerous trails.

A cactus is a succulent plant with some 1,750 known species in the plant family Cactaceae. The evolution of cacti is complicated with possibility of convergent evolution in addition to divergent evolution. The morphology and physiology of cacti are evolved to conserve water for survival in very dry environments. Pereskias are species in a small genus of cacti, but they are not real cacti per se, because they still have leaves for photosynthesis and form bark early in plant life. However, all

evidence indicates that they are cactus species aborted at the beginnings of cactus evolution. All other cacti have transformed into an all-new morphology and physiology to conserve water by switching to stems for photosynthesis.

Assume that evolution of cacti started from an ancestor plant in which there was no succulent config to conserve water as in cacti. Random mutations generated a gene or few genes with potential to form a new config – succulent config, which became the starting point to transform stems into an organ that would store water and carry out photosynthesis. Like all other ancestor organisms of animals or plants, a plant was cactus ancestor capable, because by random genetic changes some genes in its genome happened to contain sequences that could be mutated readily into new genes to form the initial succulent config. Such ancestor plants for cacti lived in South America, where they evolved into cacti and moved northwards to Central and North America. However, the emergence of a succulent config didn't mean that the plants would evolve into cacti. Only when more genes that were essential for generating the basic cactus phenotype appeared continuously to enlarge the succulent config, albeit in random order, a core succulent config would be possible. The random nature of changes indicated that the possibilities for a core succulent config would be extremely low. Even worse, the initial config could be canceled out by new mutations over the period if there were no new genes for the config to emerge, resulting in the quiet disappearances of the configs from all intermediates. Fortunate enough, it didn't occur to the evolution of cacti, although some evolution trails terminated early, leaving four species of cacti behind. Pereskias are species of cacti, but look more like regular plants, having persistent leaves and non-succulent stems. From genotype reconfiguration point of view, some genes in the succulent config in Pereskias prevented the process of reconfiguration from continuing, resulting in species only with some characteristics of cacti.

Premature termination of genotype reconfiguration is only an isolated event in the evolution of cacti. The vast majority of the core succulent configs continued to enlarge and diverge into a large number of species. While cacti display a wide range of shapes, sizes, and growth habits, the succulent configs from all cactus species contain the basic sub configs to build up morphologies and physiology that guarantee the basic strategy to conserve water and carry out photosynthesis in the stems.

Throughout the evolution, random genetic changes could have brought up numerous genes that possibly borne all kinds of random functions or random structures, but only genes that could be assimilated into some configs had the tendency to be preserved into the configurations and change the phenotypes of the configs. Cacti differ from all other plants with their conspicuous morphologies in stems, leaves, spines, and areoles, while stems are the major part of evolution, determining other parts of the cacti. Stem sub config would have its own next layers of sub configs for water storage structures, photosynthesis, stem surfaces, and more. However, it was possible for the configurations of any plants to contain genes that had tendency to undergo mutations into genes involved in the development of succulent stems. And it was also possible to turn them into succulent genes via common mutational changes in a reconfiguration event. Nevertheless, the functions of these genes would be appreciated only when they could be part of the core succulent configs. Generally random genes, despite some occasional interesting functions, would be treated as trash if they were useless to any configs in the organisms. Because new genes in the sub configs appeared in random order in the cycle, the sub configs initially must be in an immature state, maturing as more gene components emerged and joined in. Upon completion as full functional sub-configs, their phenotypes determined the final looks and internal workings of the intermediates. Therefore, evolution of cacti was a process in which all sub configs in the succulent config must have undergone similar changes to become fully functional, each of which would build up spines, areoles, special roots and flowers, and etc..

In an evolution cycle, for a trail to go completion, it must be supported by continuous addition of functional genes that could fill the missing part in the configs. Therefore, to retain new genes, their functions must be worth a place in the configuration. This is the gene retention rule for evolution. All new genes, regardless of being instant or tardy, will either be dangling around looking for a home config to settle down or mutate into nowhere if no home config could be found in time.

The numbers of family and genus are far greater than the number of order in a class, indicating that it's far more challenging to diverge into the orders from the ancestor organisms than into the family and genus from the orders. It seems likely that setting the morphologies for new species was the priority of the early part of an evolution cycle with extraordinarily high failure rates. Vertebrates have only limited choices for morphologies, legged or legless, the former had choices of winged or wingless, because the presence of a backbone has severely constrained the possibilities for a variety of forms and shapes. Morphogenesis was a process very sensitive to genetic changes in morphology related configs, which could easily lead to deformation of the physical body, affecting survivability. For example, homeobox genes were first discovered in the fruit fly where legs grew from the head instead of the expected antennae. Therefore, when organisms evolved from low to high, the first and most critical changes were to establish a morphology that must be strong, well formed, and extensible above the basal configurations. Though this step was error prone, it decided the success or failure of the evolution cycle and must be settled early to avert new species with weak physical bodies.

A large number of organisms have gone extinction since Cambrian explosion, even though most of them were at a zero sum state, raising a question whether a zero sum configuration is really stable and resilient against environmental perturbation? The zero sum rule can't be the only rule to work behind evolution. The zero sum rule dictates that more genetic changes would result in the net loss of the phenotypes to species after they exited the evolution cycle. However, it doesn't mean that the genotype configurations of the new species are at the low potential energy. If the species inhabit in the same habitats and their environments don't experience geological or climate changes, the species are stable and remain the same indefinitely. It's the genotype potential energy, not the zero sum rule, that determines if the genotype configurations of species are resilient against external pressures.

Because of randomness and unequal reconfigurability, related configs could fall in a compatible state only conditionally or sub-optimally during reconfiguration, resulting in a configuration of higher potential energy at the zero sum state. For example, if the config for the heart was compromised in building contractile smooth muscle, the heart would pump less blood into circulation, a condition of higher potential energy in its configuration. The organism would live without defects in the norm time, but could die, even go extinction from heart failure in the face of geological or climate changes. Lampreys are a group of jawless fish and found in most temperate water because their larvae are sensitive to high water temperatures. The genotype potential energy in their configurations was low at temperate regions, but elevated in the tropics, a typical conditional low potential energy case, or a result of sub-optimal reconfiguration case. Another group of bony fish is commonly referred to as lobe-finned fish. These vertebrates are characterized by prominent muscular limb buds (lobes) within their fins, which were believed to be the ancestor organisms of amphibians. In this case, the fin config had harbored genes for limbs, but it stopped here. It could be assumed that the configuration of lobe-finned fish arrived at the zero sum state prematurely, a configuration that could be considered as sub-optimal in terms of genotype potential energy, which was much lower in ray-finned fish. This premature zero sum state limited lobe-finned fish from diverging into many species. Today lobe-finned fish species account for less than 4% of extant fish species total. Most of the lobe-finned fishes have undergone extinction. It's likely that high genotype potential energy of the lobe-finned fish is the genetic basis for the fish to become ancestor organisms of amphibians.

If the genotype configuration of a given species contained configs that were fragile against external pressures, its genotype potential energy would be high, and the species would be prone to extinction. In other words, a genotype configuration could be seemingly stable under the zero sum rule, but show otherwise in the face of geological or climate changes due to higher internal potential energy. A low potential energy manifests the extraordinary functional compatibility and stability among all configs to achieve the best concerted action. Under such a condition not only individual configs, but also the entire genotype configuration are resilient against adverse circumstances in a variety of natural habitats. By contrast, higher potential energy manifests sub-optimal functional compatibility and stability among genes in individual configs and in the entire configuration as well. Evolution of species seems to be triggered by higher potential energy in the genotype configurations of the ancestor organisms, while the lasting biodiversity is the result of low internal potential energy, even though the low can be conditional in the absence of geological or climate changes.

Life is a dynamic system, and its stability and its evolution are governed by the two rules. The zero sum rule maintains the stability of the genotype configurations of species, while genotype potential energy provides species with an index that measures the possibilities of genotype reconfiguration, thus the possibilities of either becoming ancestor organisms for evolution or going extinction. The levels of genotype potential energy determine the rates and scope of genetic changes in the form of genotype reconfiguration that can occur under the given evolutionary pressures. Only when the potential energy is high to a degree, it can incur the magnitude of genetic changes that are sufficiently monumental to break the stability barrier of a genotype configuration maintained by the zero sum rule, pushing the organisms to enter an evolution cycle. Otherwise, high potential energy would result in the extinction of species in the face of geological or climate changes.

But high potential energy isn't the only factor for a configuration to be basal for evolution. It must harbor some genetic loci that contained special sequences and, upon random mutations on certain bases, could be converted into seed genes that encoded protein factors with properties to start new phenotypes, new morphologies, or even more fundamentally to initiate the development of the next generation class of species. Such genetic loci with special sequences are latent genes for evolution. Latent genes could be genes engaging in normal embryonic or tissue development, in regulation of gene expression, or even in cellular activities or biochemical processes. The latent genes essentially are the precursors of seed genes to start new configs or sub-configs. For example, the fin inducing factor gene of the lobe-finned fish could harbor such special sequences. When some bases in the sequences were changed by random mutations, the gene accidentally became a seed gene, encoding a protein factor that functioned as limb-like fin factor instead. This mutant protein molecule could induce the formation of limb-like fin buds to initiate the development of fins towards limbs. The limbs allowed the intermediates to crawl onto the land, bringing about the next generation class of animals – amphibians. Such fin inducing factor genes were likely present only in the configurations of some lob-finned fish, making them the ancestor organisms of amphibians. In the evolution of echolocation, there must be some special precursor genes as well. When mutated into seed genes upon random mutations, they could code for protein molecules with properties to respond to echo sound waves, forming the seed echolocation config. This seed config acted like a big selective sponge that retained into its possession all functionally related instant and tardy genes generated by random mutations over a period of tens of million years, forming a biological active sonar that conferred microbats full-fledged echolocating ability. Such precursor genes might have be mutated into the seed genes only in the intermediates that diverged into microbats. The common intermediates that diverged into humans and chimpanzees were likely to harbor some precursor genes. After these precursor genes were mutated into seed genes coding for growth factors, they induced the brain to develop into high sophistication. However, the mutations

for this conversion might have occurred only in the intermediates that led to humans, leaving chimpanzees forever with small brains.

It seemed plausible that it was the presence of precursor genes in the basal configuration that made an organism a potential ancestor organism. And it was the right mutations that converted the precursor genes into seed genes that made a potential ancestor organism into a real ancestor organism that was able to start an evolution cycle. And it was the subsequent successful reconfiguration of the entire basal configuration over tens of millions of years that made an ancestor organism become the grand ancestor organism of new species. The presence of precursor genes in the common ancestors prior to geographic separation could be the genetic basis of convergent evolution as well.

Assume there is a lone mountain with moderate slope on all sides. The mountain is surrounded by a flat expanse of grassland at its foot. The slope is a tricky terrain, rugged, treacherous, covered with copious little swamps, waiting for heavy things to sink in. On the mountain top there are piles of all shaped rocks. A sudden tremor pushes the rocks to roll off the top edges and down the slope because of potential energy stored in the rocks. Most of the rocks sink into the soft mud and disappear on their way down. One rock is magically stopped by some obstacles, sitting perilously at the edge of a little cliff and overlooking the distant flat grassland. Some rocks settle next to swamps at some distances from the grassland. Only very few rocks are fortunate enough to reach the grassland and lay securely on the surface. All rocks that have survived the journeys are at a zero sum state, stable and firm in their positions for the time being, but their potential energy is different. The cliff rock is stopped prematurely and possesses the highest potential energy. It will readily slip off the cliff, rolling down the slope again if another tremor hits the mountain. Those swamp rocks are more stable than cliff rock, but still vulnerable to another tremor and sink into the swamps. Only rocks that lay on the grassland are absolutely as solid as a rock and will be the lasting landmarks of the grassland regardless of tremors of degree. Physics tells us that more distant away from the mountain top, smaller the potential energy of the rocks, and consequently, less vulnerable to external impact.

Evolution of amphibians and amniotes from the lobe-finned fish is always intriguing, a fertile ground for rethinking of evolution on what has been presented in this paper. Genotype reconfigurations that underlie the emergence of new species from the Cambrian period to mammals and birds have been progressive and successive in the form of evolution cycles. The lobe-finned fish is similar to the modern coelacanth and lungfish. From morphology and anatomy standpoint, amphibians and all other terrestrial animals have shared a lot of commonalities, indicating that both groups of animals evolved from the same evolution cycle that started from those ancient lobe-finned fish. It was likely that one of the intermediates diverged into two distinct intermediates after sharing a substantial amount of common cycle trails. One of the intermediates descended into species grouped in class Amphibia, while the egg configs of the other intermediate gained new genes that provided egg cells with protective extra-embryonic membranes. This extra-embryonic membranes is an evolutionary landmark that permitted animals to lay eggs in non-aquatic environments. As a result, species that descended from this intermediate became fully terrestrial and known as amniotes. As described earlier, amniotes later diverged into two groups, namely the sauropsids, from which all reptiles and birds descended, and synapsids, from which all mammals descended.

According to Wikipedia, class Amphibia is divided into three subclasses, two of which are extinct. The extant subclass includes three extant orders and a few extinct orders. Simply based on differences in morphology, salamanders (order Caudata) and caecilians (order Gymnophiona) are more closely related to each other than they are to frogs (order Salientia). A likely evolutionary tree to depict relationships between the three extant orders of amphibians could be conceived as the following. Some intermediates

diverged productively again around 250 million years ago. One of the intermediates descended into order Salientia, representing 90% of all amphibian species, while another intermediate descended into the common ancestor of orders Caudata and Gymnophiona. This common ancestor diverged again, and as a result, one of the intermediates formed Caudata, and the another experienced the loss of limbs and descended into limbless species grouped into order Gymnophiona. In the entire process, there were numerous intermediates, even species, that went extinct shortly after they came into being.

Most amphibians maintain a life cycle that starts out as aquatic larvae with gills known as tadpoles. Tadpoles must undergo a metamorphic process to develop lungs and legs and to lose their tails to become adult frogs that can live both on land and in water. Nevertheless, amphibians are restricted to moist land habitats to keep their skin damp because their skins allow the oxygen to diffuse at a high rate and absorb water directly through it. Behind these typical and general characters, the differences between fish and amphibians and between amphibians in different orders are extensive and enormous on all aspects of morphology, physiology, and genetics. Furthermore, amphibians in different orders display a remarkable variety of life styles, life cycles, and terrestrial habitat requirements. All this shows that amphibians are species that are incomplete in their transition to become terrestrial animals, and incomplete in their own ways. Unusual diversity of the extant amphibian species, overwhelming extinction of entire subclasses and orders, and huge disparate distribution of species in extant orders are all observable characteristics that point to a bumpy evolutionary course, hard and costly, involving multiple evolution cycles. In a broader sense, changes in genotype configurations behind all class transitions would be enormous and far beyond what could be achieved in a single cycle.

Transition from lobe-finned fish to amphibians likely started with the appearance of limb-like buds and completed after multiple evolution cycles. If we could treat those evolution cycles as a single giant cycle, its reshaping phase would be the first evolution cycle that started amphibian transition, in which the entire genotype configuration had experienced unusually intense reconfigurations that laid the morphological and corresponding physiologic foundations for the subsequent prolonged evolution into amphibians and amniotes, the ancestor organisms of reptiles and mammals. In the healing phase of the giant cycle, though genotype reconfigurations in the amphibian lineage continued, but continued at much milder intensities. Those viable trails diverged in varying degrees into amphibian species that fell into different orders and then different families, all of which constituted the class Amphibia.

The reshaping phase of the giant cycle determined the morphologies of amphibians and amniotes, which in turn determined the underlying physiology required to support the emerging morphologies. It would be expected that a far larger number of new instant and tardy genes must be created to bring about new biochemical processes and cellular structures for land adaptation. At the same time, virtually configs across the entire genotype configuration must undergo mutational changes and mingle with all the new genes in such a way that all reconfigured configs would function as an interdependent and unified whole to enable the organisms to inhabit the land. However, randomness of the mutations in terms of biological properties and time of their appearances in the cycle and the changes in such a broad depth and extent made the success of the reshaping phase a grand challenge. It could be assumed that in this uncontrollable, and largely random reshaping phase, few intermediates stopped prematurely at the “mountain cliffs” and became species with high genotype potential energy, while overwhelming majority of the intermediates failed to survive and became dead ends. Species with high genotype potential energy usually carried highly unsettled genotype configurations. Some of these configurations were wide open for further reconfiguration, becoming the ancestor organisms for the next rounds of evolution cycles, and the others went extinct upon climate and geological changes. Therefore, the reshaping phase of the giant cycle might contain more than one horizontal or parallel daughter cycle that ended with few species. The genotype configurations of some of these species became the basal

configurations, from which all modern three extant amphibian orders could evolve in their own evolution cycles, comprising the healing phase of the giant evolution cycle.

Nevertheless, the reconfigurabilities of those basal configurations were not equal. The one from which order Salientia evolved claimed the highest reconfigurability, churning out about 90% of all amphibian species. The sheer size of order Salientia seemed to indicate that its evolution cycle was exceptionally large in the number of productive trails. Each genus would contain species that shared extraordinary similarities not only in their morphologies, but also in their individual configs, even in their genome sequences. Although the basal configuration for order Salientia was highly successful to generate new species, they were a dead end for further evolution. The large size of order Salientia also suggested that species in this order stored low genotype potential energy and should be less vulnerable to external impact. However, there has been a dramatic decline in the amphibian populations for many species globally. This is not necessarily because of high genotype potential energy per se, but because of reasons like permeable skins, which make them sensitive to environmental pollution, and complex reproductive needs, that make them vulnerable to climate and geological changes. However, these two and other phenotypes have been stable over hundreds of millions of years since amphibians emerged. At present, many amphibian species are found only in some special environments, indicating that their habitats were a sort of special, which generally had avoided large climate and geological changes. Confinement in such niche spaces for survival suggested that their low genotype potential energy was conditional. Large scale human activities have destroyed their natural habitats considerably, making many of them endangered species.

All class transitions could be treated as giant evolution cycles, composed of destructive reshaping phase and mild productive healing phase. And each phase of the giant cycle could contain more than one daughter cycle, both horizontal and parallel. Lucky survivors of the reshaping phase stored high levels of genotype potential energy, thus were prone to external changes. Majority of the early species of the class went extinct with a few starting new evolution cycles to generate new species which had lower levels of genotype potential energy. As giant cycles entered the healing phase, individual daughter cycles produced new species with low levels of genotype potential energy, forming modern day orders, families, and genera. If a class transition required a giant evolution cycle with more destructive reshaping phase, it would result in many fewer orders for the class, while a class transition required a giant evolution cycle with less destructive reshaping phase, it would result in many more orders for the class. Transitions from aquatic organisms to amphibians or reptiles required extremely vigorous genotype reconfigurations, and as a result, both of the classes comprise 3 orders for amphibians and 4 for reptiles. By contrast, transition from reptiles to birds was less dramatic, resulting in 44 orders out of a giant evolution cycle. Mammals were the only living organisms descending from Synapsida as early as 2 hundred millions of years ago, a hundred of millions of years after Synapsida was evolved from the early amniotes around 3 hundred millions of years ago. Fossil records indicated that there were exceptionally large number of non-mammalian species that descended from Synapsida before mammals appeared, and all of them had gone extinct in different geological periods. Many of those extinct non-mammalian species possessed reptile-like morphologies, but they were neither reptiles nor part of reptile lineage, implying that they only shared a remotely common ancestor amniotes. All this indicates that mammals were the end state of an unknown, but prolonged series of evolution cycles, each of which produced essentially “cliff” species, whose genotype potential energy was so high that they kept starting next cycles once external conditions were right. At the time of their appearance, like birds, the differences between early mammals and their direct ancestor organisms were no longer too dramatic, forming 27 orders with 6,640 extant species. It seemed that more than one giant evolution cycle would be involved for the mammals to appear.

When the genotype reconfiguration took place in response to geological or climate changes, random mutations incurred a series of changes to the configuration, resulting in something good and something bad. The bad was eliminated at the expense of life of numerous intermediates, while any neutral and beneficial changes were managed at single gene levels, sub config levels, and config levels to ensure that all of the configs remain compatible with each other in the entire configuration. Those assimilated changes conferred more advanced features and capabilities to new species. Evolution is simply an inevitable result of the plain paradox of randomness.

It could be concluded that if the ancestor organisms are viewed as rocks on the mountain top, then new configurations are rocks that have survived the swamps and reached the zero sum state, but their potential energy can be radically different from one another. The configurations that are the farthest away from the center of evolution cycle have the lowest potential energy, and consequently the organisms are more insensitive to geological or climate changes and able to survive in wide weather and conditions. On the other hand, the “cliff” configurations have the highest potential energy, making them unusually vulnerable to changes, a feature the basal configurations must have. In-between species can be full of life in their own native habitats in peaceful time, but prone to extinction when geological or climate changes strike. For example, animal rodents are everywhere. Their genotype configurations must store very low potential energy, which enables the animals to endure a variety of habitats and even geological or climate changes. On the other hand, the configurations of those endangered insects are in the zero sum state, but their potential energy is only sub-optimal and conditional. Consequently, they are confined in their native habitats and climate, a safe heaven for their survival. If genotype configurations collapse under geological or environmental pressures because of less optimal potential energy, the damages to the configuration would be irreparable, resulting in extinction of certain population, even the species in its entirety.

15. Summary and Discussion

Looking at life of all forms that appear along the timeline, evolution of species can 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 refers to as slow evolution, an extraordinarily long period of time, lasting the next 3.5 billion years, in which life evolves into multicellular forms at a slow but steady pace. The genomes of many species have reached moderate sizes at the end of this stage. Stage 3, by contrast, refers to as fast evolution. It is an unusually short period of time. Starting from Cambrian explosion, stage 3 has passed only 600 millions of years so far, during which life of all forms and complexity has inhabited the earth, boasting a prodigious biodiversity.

There is every reason to believe that the nascent earth was a life-welcoming planet, and many places in the vast seas could be called incubators of life, where existed a rich mix of nucleobases, amino acids, sugars, lipids, and other inorganic and organic chemicals. In the incubators, a variety of random chemical reactions occurred spontaneously and constantly, generating all kinds of possible chemical products, including polypeptides, ribonucleic acid RNA and DNA, all of which were random in length and sequence. As the amount of random polymers increased, some of the RNA happened to fold into structures similar to tRNA, rRNA and mRNA, and some of the polypeptides formed three dimensional structures with rudimentary biochemical properties, including 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 early forms of enzymes, DNA grew longer in a random fashion and multiplied by replicating. Meanwhile, the RNA population was a mixture of RNA molecules either transcribed from DNA templates or polymerized

randomly, many of which were RNA precursors to modern tRNA and rRNA. When rRNA precursors complexed with some ribosomal-like proteins, they became simple ribosomes. Similarly, the tRNA precursors could carry an amino acid at 3' end and align onto the mRNA molecule which was attached to the ribosomal platform, allowing 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 incubators. More proteins displayed good properties or function as enzymes, ribosomal proteins, structural proteins, trans-membrane transporters, and so on. More importantly they slowly became available in stable fashion. Gradual appearances 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, increasing the reproducibility of protein molecules. All this indicated that the minimum genomes started to emerge. Slow but steady improvement and maturation of the basic biochemical machines 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 an infinite amount of random peptides, RNA and DNA, a cache of great treasure that 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 and depend on 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 cells – single celled life – formed when minimal genomes were 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. The one that prevailed in the incubators 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 random mutations in the early phase of stage 2. Each form of life was likely to possess a unique set of proteins despite using the same set of genetic codons. Existence of numerous life forms made it possible for multiple cells of different origins to merge into single cells 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 eukaryotic cells. Confinement of the 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 were 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 for early life, creation and assimilation of a large number of new genes, including regulatory elements, 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 machines to make them more complicated and advanced. Observably the later organisms have more sophisticated morphology than 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. Evolution in Stage 3 is very different from stage 2. Organisms at the end of stage 2 had amassed a large number of protein coding genes comparable even with mammals. As a consequence, the mode of evolution had changed from total randomness to the reuse of existing genes or pseudogenes via random mutations and other mechanisms, 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 needed for the evolution of species that could be 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 amphibious tetrapods with ancestor nature evolved into amniotes. Natural upheaval 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. Once in the cycle, the genomes were constantly reshaped by random point mutations, gene duplications, and recombination. Some of the duplicated genes were genetic fodder to derive new biochemical properties through the action of point mutations and other genetic operations. 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 those lucky ones not only survived but emerged as new species. They were stable indefinitely under the zero sum rule. 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 and stable disarmed state. Therefore, when we talk about evolution, we can focus on an evolution cycle. What happens in a cycle is what has happened at a specific time in the history of evolution. Another important thing to keep in mind is that an evolution

cycle can take up to tens of million years or generation to settle, 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 to avoid early mutations being 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 phenotype for new species. By keeping mutations at the right rates, it is ensured that there were always intermediates that survived intermittent types of changes and at the same time generated novel features for new species to emerge.

Reuse based evolution is the most distinguished characteristic of evolution at stage 3. Reuse has greatly accelerated the emergence of new species. The essence of reuse is the generation of protein variants. Many proteins have variants not only in different species but also in the same species. For example, most of enzymes involved in glycolysis in human have more than one variant, some of which are produced by alternative splicing. Opsin molecules sensitive to different wavelengths are variants of each other in the different species and in the same species as well, and they are encoded by genes of duplication origin. Gene duplication, being one type of DNA rearrangements, still occurs in modern day organisms. Duplicated genes are free to diverge via gradual accumulation of random mutations so long as they don't cause lethal effects, but serve useful biological functions. Because a majority of protein variants don't seem to be derived from duplicated genes, the gene counts across species are stabilized around 22,000, regardless of the complexity of species.

In each evolution cycle, some genes were duplicated from master genes from the previous evolution cycle. Almost all genes, including duplicated genes, would be bombarded by random point mutations over millions of years. Many sequence changes resulted in functional or structural changes, but only changes that survived the normal working of the cellular machine would be preserved as functional variants and embrace more changes. The success rates of producing new functional proteins by reuse are much higher than de novo creation of novel proteins. When an intermediate accumulated a large number of protein variants and some novel proteins and survived, its cellular machine had been reshaped in a fundamental way, allowing the intermediates to emerge as new species. As a result, new species from later cycles are usually more advanced and sophisticated than species from early cycles.

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 its biochemical properties in a subtle way, making the protein more fine tuned for the process or structures in that cell types. On the other extreme, variants of some of the 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 in 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, the periods that might be far longer than the time for an evolution cycle. In an evolution cycle,

on the contrary, the genome was being reshaped globally, resulting in numerous changes, possibly including a new opsin molecule Rh8 with absorption peak at 700 nm. If all the components for a new type of photoreceptor cells for Rh8 appeared over the period of the cycle, a new photoreceptor R9 came into being in the fly compound eye. However, changes in an evolution cycle were far beyond the compound eyes, and this fly was no longer the fruit fly, but a different and more advanced fly species.

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 of the enzymes 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. It isn't an easy task to improve a phenotype as simple as citric acid cycle or glycolysis pathway, how can phenotypes as complex as the fly compound eye or bat echolocation system be improved overall by random mutations outside evolution cycles?

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 over a few evolution cycles. Generally speaking, a biological machine is the assembly of all of 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. Therefore, mutational effects on any biological machine are badly constrained by its two sides – stability and fragility. The zero sum rule 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 a particular phenotype is developed 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 required to cover the light from the visible 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 have been developed by dropping off unnecessary components, while preserving improved components for better biological performance. As species evolve, fewer opsin molecules are used in vision. Some reptiles have 5 opsins at the most, and majority of mammals have 3 opsins at the least. Primates, including humans, have 4, which includes rhodopsin for night vision. In other words, evolution itself is under constant evolution to learn how to create better and advanced species.

The vision system is the most complicated biological system, reflecting the formidable difficulties in this quest. Complexity begets variety, thus visual system enjoys the most varieties in morphologies and the underlying molecular and cellular mechanisms in 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, and danger, well fitting their specific living environments. Such a variety of visual system indicate their unique and largely independent evolutionary origins, at least implying a remarkable divergence of the rudimentary visual functions at the 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. Most of these primitive forms of eye are simply the trial and error versions that appear only in some stages of vision evolution, and their underlying mechanisms are not suited for further vision evolution. Because of this they all have disappeared along the evolutionary timeline. They are mere evidence of failed trial-and-error strategy in the course of evolution. The eyes have conformed 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 it is evolved over time from the primitive forms to the advanced forms on molecular and cellular levels. Obviously the visual system seems to be one of the best 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 have been 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 in the nascent seas – a giant 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 within a boundary that turns the earliest weak life-like activities into concrete and significant components of life. The self-organizing nature of macromolecules allows 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. Randomness from the nascent seas generates unlimited 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 will design all pieces for a puzzle on a piece of paper by dividing the space into a pre-determined number of pieces 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 puzzle maker attempted to make a puzzle by making pieces of random shapes, hoping that some of the pieces could be assembled into a whole, which had 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 would become easier if the maker was experienced and the shapes for the puzzle were relatively simple. But it could become far more difficult if the shapes were a little more complicated and if the maker had 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. 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 different phenomena 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 ten random pieces of DNA sequences to encode ten proteins, which, as a group, will fulfill a complicated series of biological activities that 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 about 2 billion 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 can have and extend, including muscle, skeleton, digestive track, nerve 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 all new puzzles. Although random mutations are the sole catalyst for all the changes in both slow and fast evolution, reuse strategy 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 it easier and far more efficient to improve and derive functions.

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 as a few shapes in the pool fit together into a group with a 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 the 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.

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 facility. The approach to make more advanced puzzles in variety is still exhausting, time consuming, and aimless, but the randomness accounts only for 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 to make new puzzles. 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. Nevertheless, this puzzle reuse strategy becomes increasingly difficult when puzzles are made from an increasing number of pieces. For this reason, there are fewer reused-based, complicated puzzles on the market. Regardless, 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?