The Worst Idea in the History of Science

Listening Carefully to the Curious Noise of Silent Mutations

By Mark White, MD

"Things should be made as simple as possible, but not simpler."

Albert Einstein

Introduction to History

Ideas have power. The notion is made quite literal in exploding an atomic bomb as a physical demonstration of the idea that even tiny amounts of mass contain enormous amounts of energy. But more so, ideas are powerful in how they shape other ideas, and even simple ones can explode into enormous power and complexity. The power of a single idea can be measured in terms of its clarity, pervasiveness, and centrality to other ideas. All of these things contribute to the longevity of an idea in the minds of mankind. Really good ideas live a long time, but so too can bad ones.

Of course, the concept of good versus bad introduces a large bit of additional subjectivity to the formal measurement of ideas. Judging ideas as the best or the worst is like arguing about the best golfer, the best painter, or the best leader of all time. In truth, it can never be an untainted process of pure objectivity. Plus, many ideas that might actually deserve the title of "the worst idea ever" rarely gain enough traction to merit any attention. There must be something "good" about a bad idea to become a candidate for the worst idea ever. Despite these subjective difficulties, some ideas clearly stand out. They are so powerful that, in time, they separate themselves from the field in terms of being the best or being the worst idea of all time. From these few exceptional ideas we can confidently say that the most powerful ideas tend to be simple ones, or at least they deal with fundamental relationships. They provide simple clarity to otherwise opaque relationships. Clarity derived from simplification is the foundation upon which more ideas and more complex ideas are built within

science. This is the basic pattern of science in its inexorable march toward evolving its collection of ideas. The best ideas march us quickly toward more good ideas and the worst ones lead us into dead ends, or worse still, they lead us in the wrong direction entirely.

Ironically, bad ideas can sometimes be more powerful than good ones, and a good idea need not be entirely correct to become powerful. For instance, one of the best ideas of all time surely is the one Newton used to explain motion in terms of force and mass. After all, this fundamental relationship serves to explain the motion of the earth around the sun. Once the proper relationships between the earth and sun are appreciated, the force of gravity can explain this motion. However, Einstein discovered that Newton was not entirely correct, so he proposed that an additional symmetry is required between the earth and sun. Neither body is really stationary, so their motions must always be seen as relative to each other. Einstein also realized that motion is merely a function of time and space, and symmetry must exist between time and space if we are to understand motion. Both men were true geniuses with great ideas, obviously, yet Newton was not entirely correct after all. Still, his ideas are deemed "good" and no doubt continue to be incredibly powerful. They are the foundation for the ideas of Einstein, who actually corrected Newton, or you might say that Einstein merely extended the ideas of Newton.

Regardless, my vote for the best idea in the history of science will always go to Charles Darwin and his idea of natural selection. This was a simplifying idea that provides an unexpected clarity in biology, a field dominated by overwhelming complexity. Darwin described a mechanism that can bridge the gap between organic and inorganic matter in the minds of many men. The idea explains with surprising clarity the logical relationship between random molecular activities and highly organized molecular information, even though Darwin perhaps never used the term "molecule." Natural selection is, therefore, the best ideological framework for understanding all of nature's many organizational schemes. It is a simple idea that makes sense of evolution in spanning the extremes of micro and macro in time, space and number. Darwin, like Einstein, couched his great idea within an general ideological framework of symmetry and time. He appreciated that natural selection was a function of time as well as a function of complex symmetry in the many parts making up any whole.

Leading candidates for the worst idea in the history of science surely must include the two ideas of a flat earth and a geocentric universe. I reject the flat earth as a viable candidate here because it is merely an example of a simple choice made wrongly. The geocentric universe, on the other hand, is a good example of a simple relationship wrongly inverted. The sun does not travel around the earth, yet the simple inversion can make perfect sense at some level. So perhaps we can expect to find the hallmarks of this bad idea masquerading as a good one for a long period of time. And so we do. These unfortunate situations involving simple inversions mean that bad ideas become troublingly powerful, show puzzling longevity, and are extremely difficult to eradicate. Worst of all, the basic inversion winds up inverting practically everything it touches, like cause and effect. This means that bad ideas ultimately require "add-on" ideas to

take care of the numerous inconsistencies that always arise in an inverted model. For instance, a geocentric universe must also be filled with complicated epicycles to account for otherwise simple motions in heavenly bodies.

Notably, the longevity of the geocentric universe was independent of the sheer number of epicycle-like ideas that were also required to sustain it. Humans are willing to accept untold complexity as long as they can hold onto an initial premise of simplicity. However, the remarkable longevity of geocentrism can be attributed to several other factors as well, not the least of which is that the initial inversion between earth and sun hardly seems to matter. In fact, this inversion still seems to find many practical and impractical uses today. A starting assumption of a stationary earth is "good enough" for practically everything. Our acceptance or rejection of it depends on the required level of detail and the practical need for anyone to have a "better" understanding of the system in question. Plus, defenders of this ruggedly simple paradigm had plenty invested in it before it began to falter. It really is intuitively simpler for humans to suppose a stationary earth at the center of a rotating universe. Consequently, total indoctrination to the erroneous paradigm was both simple and fierce. This makes sense of the puzzling fact that this one powerful yet demonstrably bad idea died very slowly, and was allowed to do immeasurable damage along its death march. Could it happen again?

Modern science is now in the process of lifting the veil on what may eventually be seen as the worst idea in the history of science. It is hard to imagine that 21st century science could either spawn or safely harbor such a thing, but in reality it is more likely than not that it actually does. It is modern hubris to expect otherwise. Science is so broad and complex these days that ideas are more difficult to understand and evaluate – not less - yet simple ones still send ripples throughout the entire pond of scientific ideas. Science has proven itself to constantly be full of both good and bad ideas, so today can be no different than yesterday or tomorrow. The magic of science lies in its ability to eventually tell the good from the bad. An honest scientist will tell you that there is really more faith required in most scientific ideas today than ever before, and certainly more faith involved than most modern scientists ever care to admit. When it comes to molecular biology, a huge percentage of the ideas simply must be taken on faith, because most of them are inherently hard to define, and they become even more difficult to conceptualize and conclusively test.

With this context in mind, consider the important case of an extremely powerful idea that widely exists today. It is a simple idea, so it is a good illustration of the relationship between simplicity and power with respect to scientific ideas. The idea to which I refer is the idea that molecular sequence determines molecular structure. It is hard to overstate the power that this one idea has had in shaping our current understanding of molecular biology. However, it is now the pivot around which we rotate our ideas of genetic translation and still more abstract general notions of molecular information. In other words, it is the initial point of reference for most of a much larger narrative. This same idea has been stated, reformulated, built upon, combined and

employed in too many different ways to mention here, so regardless of whether we ultimately deem it good or bad, it is undeniably powerful. The idea derives its basic power from its unique ability to simplify and clarify what would otherwise be an opaque relationship: the basic relationship between sequence and structure in the universe of molecular information. The idea gets this ability by merely inverting the fundamental relationship between sequence and structure. This idea is not believed to be a generally true, but it does seem to be true for the special case of large biopolymers. Let us review:

Every molecule can be viewed in two different ways: as a composition or a structure. The composition of any molecule is an information subset of its structure. In other words, knowing the structure is the same as knowing the composition, but knowing the composition is not the same as knowing the structure. This is a fact of basic chemistry, yet in the world of large, complex, "linear" biopolymers, where an amazing consistency of subunits is found, the sequence of subunits is indeed valid shorthand for molecular composition. Of course, sequence is not really the same as composition but they can easily be seen as equivalent because – purely in terms of molecular information - every sequence is only slightly above every composition. In other words, sequence adds just a little bit of organizing structure to composition, and that bit of structure in biopolymers is also remarkably consistent. This wee bit of equivalency simplifies the notion of composition in complex molecules; however, to simplify the situation entirely we must also know that sequence determines structure.

The natural hierarchy of total molecular information content is: structure – sequence – composition, yet this single, powerful, new idea collapses all molecular information to sequence by merely inverting the place of sequence and structure in the hierarchy. In the very special case of biomolecular information, composition, sequence and structure can now be seen as equivalent on an important level of understanding. This is exactly what is meant to say that sequence determines structure. If sequence determines structure then it logically also follows that molecular information equals sequence. This is, in fact, the standard way to define molecular information in the world of complex biopolymers. This is our first point of reference when we begin to discuss and contemplate the complex world of biopolymers, their fundamental relationships, and the translations of one into another.

The idea that sequence determines structure, therefore, is an idea that is nothing short of a magic wand when it comes to simplifying the complex world of molecular information. To say that sequence equals molecular information is our way of first knowing the rate of exchange during molecular translations. We now hardly need be concerned with compositions and structures, simple or complex, because life, it turns out, is just like a computer: It is all about sequences! This is what molecular information is and this is what molecular information does in life's grand organizational scheme. The information of a biopolymer is contained wholly in its linear sequence of subunits. This is why this model for understanding things is also referred to as "one-dimensional" because it models only one-dimension in the universe of molecular information. Information can be reliably reduced to, thought of as, and translated by the single dimension of

information contained in any sequence of polymeric subunits. Naturally, this greatly accelerates learning and also happily expedites our use of computers in mastering the complex universe of bioinformation. It is an incredibly powerful idea that logically leads us to many more ideas that are clearly useful in countless ways.

Specifically, this simple idea serves as the basis for our understanding of the genetic code, which then serves as our paradigm for molecular translation. It is, after all, the supreme paradigm of molecular translation to know that a sequence of nucleotides is reliably translated in triplets – codons - by the genetic code into a sequence of amino acids. The product of translation then conveniently determines the structure of folding in the translated protein. The logic of translation is made complete merely by correctly knowing the relationship between sequences of codons and sequences of amino acids. It makes perfect sense that every protein can consistently have only one stable structure simply because only one dimension of information can be translated into its sequence of amino acids. Proteins absolutely must have consistent structures - no doubt - and the genetic code is the only way to achieve them. This also logically makes every sequence dwelling in the simple helical structure of every genome the correct initial point of reference from which all molecular information naturally emanates. So, the idea also paints for us the all-important main arrow in our current diagrams of information logic, or more precisely it illustrates to us the inevitable character and "flow" of information as described by the central dogma of molecular biology. In other words, sequences of DNA can legitimately be seen to form the central body around which all other bodies must revolve in the organic universe of molecular information. They are the gatekeepers of sequence and sequence is everything.

This kind of grand and truly inspiring narrative is familiar to anyone who has ever studied biology at any level. It derives its explanatory power at first from the idea that sequence determines structure and molecular information is, therefore, sequence. We know that sequences of nucleotides determine the double helix of DNA and sequences of amino acids determine the structure of proteins. The two are solidly welded together by the inescapable logic of the genetic code. Of course, these primary explanations must lead to countless other explanations because science is not in the business of giving answers but of asking questions. We must now explain how this basic system works, why it works the way it does, and how it ever came to work this way in the first place. This is the nature of science, and it all fits neatly and completely together - as it should. Therefore, the idea that sequence determines structure is an incredibly powerful idea in modern science, one that is clear, central and pervasive; one that instantly turns the opacity of nature into something that can be immediately understood in its simplest form possible. It is a powerful idea of science that has only been around for half a century, but one that will be around for a long, long time to come. This is an obvious candidate for the best idea in the history of science, which is probably why it, and its many close descendents, have garnered so many accolades in the annals of science. However, the sheer simplicity and raw power of this one idea now also make it a potential candidate

for the worst idea in the history of science… if only it were wrong. Clearly, it is completely wrong.

All one really needs to know about the veracity of this simple idea is that "silent" mutations actually do cause changes in protein structure. It is not a matter of whether all silent mutations cause structural changes but whether any silent mutation ever causes a structural change by any means whatsoever. And they do. However, if the basic idea were right in the first place, this should never happen; otherwise, it would not surprise us to learn now that it actually does. If silent mutations can change protein structures then sequence cannot determine structure. It's just that simple. The very real structural impact of silent mutations was demonstrated four years ago by Cortazzo, and recently verified in convincing fashion by Gottesman. These empiric demonstrations completely destroy the now utter myth that sequence determines structure, but the supporting evidence for this basic knowledge has, in fact, been abundant for decades. The very idea was completely illogical to begin with. There are now many ways to say it and many more ways to prove it, but one can no longer even recognize the orbits for all the epicycles.

First, recognize that this is an idea - like being pregnant - that cannot be partially right. Either molecular information is contained and translated in one dimension or it is not. This is now not a question of degree but of absolute relationship. Specifically, is the relationship between sequence and structure normal or inverted? Either sequence determines structure or it does not. After all, we already knew that sequence was a subset of structure, so to say that sequence is but one determinant of structure is to say the obvious. To say that it is a very important determinant of structure barely merits a yawn. The clarity and sheer power of this idea is completely derived from the clever inversion of the normal relationship between sequence and structure by making the silly mistake of definitively saying that sequence determines structure. This is also precisely why we now say that molecular information is sequence. If there is no real inversion, there is no added clarity but merely added confusion. In fact, there no longer is a viable definition of molecular information. Consequently, we now have no working definition of the genetic code. What is the genetic code? Second, note that all subsequent arrows of time, cause and effect, as well as presumed evolutionary progress through time, teleology, virtually all of our key explanatory narratives depend on this one central relationship between organic sequence and structure. If the relationship is inverted then so too are all the dependent arrows. The power of this idea lies in its simplicity and its ability to clarify, but so too is its risk to obscure reality and mislead our thoughts. In other words, it is either taking us in the right direction in many and various ways or it is leading us to dead ends and wrong turns everywhere we look. The latter indeed appears to be the case.

The original man-bites-dog scientific story of sequence determining structure was fantastic and, therefore, widely reported and touted for its supreme significance. But the dull truth is once again just another example of dog-bites-man. Nothing is inverted, except for our current understanding of reality. The correction, as always, has not gotten nearly the attention of the sensational false

headlines. The continued absurdity now merely reflects our complete indifference to the tremendous power and influence that the initial idea has obtained. Unfortunately, one can look nowhere today and fail to see the fabulous number of required epicycles nested in an elaborate structure of ideas to support this one, really bad idea. It boggles the mind that this could go on for so long and yet now be fully expected to continue indefinitely. The sheer simplicity and intuitive appeal of the original false paradigm, coupled with the immeasurable investment and extreme indoctrination over decades mean that this bad idea will surely haunt us for a long time to come. However, the realization that sequence ***does not*** determine structure should rightly be seen as analogous to finally realizing that the earth actually does move within our solar system. It changes everything in a very fundamental way. Yet, think of all the simple flow and logic diagrams and all the captions for all the codon tables in the world today. Think of how many more will be printed in the future with this catastrophic idea embedded in not merely the captions but in the fabric of an all-too-clear picture being painted by the table itself. It is a picture of a molecular code completely devoid of structural organization and structural information.

Surely, some grizzled veterans of biochemistry will never concede the fundamental importance or even the obvious incorrectness of this one bad idea. They simply are incapable of doing this, so they will find half-truths within it, and attempt to rehabilitate the apparent utility of the idea right on 'til the embittered end of their scientific days. The important narratives will change slowly because this is the nature of bad scientific ideas – they obtain power, cause confusion, and eventually lead to ideological wars. When all else is lost, the defenders of the flawed paradigm will undoubtedly argue that perhaps we never would have gotten where we are today without it. In other words, it was an incomplete but necessary step in discovery. Even this weak argument is pure hogwash. The idea served us no good; nothing done of benefit would not have been done otherwise, and we simply would have gone farther and faster without it. It is a red herring that doomed the inevitable advance of science to become slower and more difficult not easier and faster. Plus, some of the damage it has already caused might never be reversed.

This is perhaps the key concept of this entire book: We have a binary decision to make. Does sequence determine structure or does structure determine sequence? Does the universe of molecular information revolve around sequences or does it revolve around structures? Choosing the former produces a line and choosing the later produces an intricate curve in our thinking. These are diametrically opposed ideas and they are mutually exclusive. One or the other is correct but both of them cannot be correct. Our entire model of molecular information and genetic translation are built upon this first decision in the model building process. Our entire explanatory language depends upon it. If sequence determines structure then the model becomes flat and we loose all sense of time and scale. The details evaporate and the explanations become necessarily ad hoc and tautological. If structure determines sequence then the details explode. Rather than studying a simple line we find ourselves looking at a curve that rivals the Mandelbrot set in its beauty and complexity. We must then

introduce many different dimensions of information, including time and scale to our understanding of molecular information. The shape of the curve, the number of details and the patterns we see in them all become dependent on the scale we choose to investigate this beautifully complex natural phenomenon.

The heuristic devastation wrought by our unquestioning faith in a model based on this single bad idea is incalculable. It has laid waste to the natural beauty of a proper model of a complex molecular biology. The idea smacks of pre-formation, or of a DNA homunculus that guides the growth of every protein. This is quite the same as believing that DNA would somehow store digital pictures, if it could, in an uncompressed format and then have the genetic code compress them for our viewing or printing. The idea that sequence determines structure has clearly led to a modern day version of alchemy in the derivative idea of protein folding. Millions of people and computers today are searching in vain for the elusive magic formula that converts sequence to structure like lead into gold. They are merely studying the structures of a decidedly biased group of proteins to ostensibly map the forces of nature that move these complex structures around the gravity of simple sequences. But a search for any single target of protein structure based wholly on amino acid sequences is doomed to failure. Noises are made by silent mutations, yet the blinding faith in this alchemy will surely persist as it has despite decades of abject failure.

Nowhere is the damage more evident than in our collapsed view of the genetic code. Only because it is collapsed is it widely seen as an all-too-disappointing kluge that is in turns described as simple, degenerate, unimpressive, arbitrary yet universal, frozen out of evolutionary competition by its banishment to one and only one dimension of information. In short, virtually every aspect of the genetic code is mistakenly perceived today, and all because of the idea that sequence determines structure. Our paradigm of a code of molecular translation is quickly and reliably burned into our brains today by looking at a spreadsheet and truly believing that it somehow depicts the actual code of protein synthesis. In fact, it's not even a code at all but a powerful ideological icon built from a simple arrangement of demonstrably incomplete data. Sure, there really is a consistent relationship between nucleotides and amino acids, but this is merely a subset of more complex molecular information and its ingenious logic of translation. What's more, the genetic code is not only a fabulous operating system for building proteins but a search engine to boot. It embodies the first principles of life itself. Time, complexity and symmetry are all playing major roles in these functions. Unfortunately, our entire definition of molecular information is now merely a subset of molecular information. Therefore, the codon table should never have defined information in full and stood as our paradigm of molecular translation. Unfortunately, that is exactly what it does today.

This absolute over-simplification has stood the entire concept of cause and effect completely on its head. In other words, we like to now say that a protein can fold only one way; therefore, the genetic code is this way. In truth, the genetic code is this way so that a protein can be made in any consistent fashion at all. It need not appear to us this way if only we can begin to imagine a

world without this one, powerfully bad idea. In the correct light, the genetic code becomes a stunningly complex and elegant system, optimized for function, yet able to evolve in various ways at a moments notice. Far from a simple example of sequences making structures it is part of a much larger algorithm where structures make sequences. After all, molecular sequences are nothing more than sequences of molecular structures. DNA stores these structures in sequences of remarkably simple overall structure, and they get translated through time into progressively more complex sequences of structure. One of the intermediate sequences of structure in translation - sequences of whole tRNA structures - are fabulously interesting and informative, but they are eliminated from the universe of molecular information whenever the genetic code is flattened and placed inside a codon table. Even sequences of amino acids are merely sequences of structure, but most of the structure in these sequences resides in the necessary peptide bonds that connect the amino acids. Proteins should be generally viewed as sequences of peptide bonds and not merely as sequences of amino acids simply because knowing the peptide bonds means knowing the amino acids but knowing the amino acids does not tell us the peptide bonds. Unfortunately, peptide bonds have also been completely eliminated from the codon table, and, therefore, no longer exist in the universe of molecular information either. Virtually nothing is left of molecular information when sequence determines structure.

Given the amazingly consistent set of amino acids in the genetic code, the next question should become: How many different sequences of peptide bonds can the genetic code make in both time and space? However, this code could never even be this simple because protein structures themselves are merely sequences of secondary structures, like helices, sheets, loops and turns. In fact, many of these sub-structures have been proven to contain no consistent structure at all. How could we ever hope to make a supra-structure consistently in only one way if its many required parts are allowed to somehow keep changing their own structures? Perhaps the genetic code is not at all concerned with simply making sequences of amino acids but is instead busy making peptide bonds, or secondary structures, or whole protein structures, or entire complex populations of protein structures. Regardless, we can now know for certain that it is not merely making sequences of amino acids; otherwise, all silent mutations would always be silent. It is obviously not a code that is only about simple sequences without inherent structure but will instead be far better understood as a complex algorithm that logically relates structures to other structures. It somehow is able to do this by relating many structures to each other sequentially in time and space. We will never understand this until we frame our questions and answers within the certainty that structure determines sequence. After all, it is logically and empirically true that structure determines sequence.

Our failed geocentric model of the molecular information universe today depends on the idea that we should accept a simple inversion of reality. This idea lends required mass to DNA and places it in a stationary central position so that all other information can revolve around it. This idea not only ignores Darwin completely, it reverses many of his most basic concepts, and it inverts most of

the key arrows he lays down in time. In fact, it virtually eliminates the concept of time from the entire model. Worst of all, the many beautiful symmetries that make the system work are also destroyed in this single erroneous process of simplification. The symmetries between sequence and structure, DNA and protein, time and complexity, random and organized, are completely flattened. So, not only is this model obviously geocentric but it includes a flat earth component as well - with a huge measure of spontaneous generation thrown in to boot.

The truth is that DNA does not make DNA and DNA does not make RNA and DNA does not make protein. In fact, DNA makes nothing. Protein makes DNA and protein makes RNA and protein makes protein by using DNA, RNA and protein. Only by seeing things first in this way can we understand how insentient molecules can efficiently turn an otherwise unimaginable molecular trick of complexity in performing a code of structural information processing. It turns out that protein is far closer to the center of the real molecular information universe than is DNA because structure makes sequence. In reality, nothing is stationary in this universe but everything constantly moves relative to everything else through space and time.

Making complex structures from simple sequences is a combinatoric cakewalk. The real challenge to the crystallographer and to the cryptographer is to figure out how life makes simple sequences from complex structures. In any complex molecular world there are too many different things that could possibly happen to be certain that the "right" thing will ever happen at all, so the question becomes: How does any molecular system make certain things always happen? Another way to ask the question is: how can complex structural information be captured, stored and reliably utilized? A way to ask this same question such that we might catch a glimmer of the answer today is: How did a molecular code of molecular structures make itself appear to us to contain no structure? The answer is that it uses perfect structures. In other words, if all of the structures and the components of structure are somehow the same, then the amount of structural information required by the system can be minimized. It is purely a question of information efficiency. When things in an otherwise random system become consistent their information content goes way down. Think of the case of tossing 100 coins and the resulting 100 bits of information that it generates. Now think of another case that involves 99 of 100 two-headed coins, and imagine the tremendous decrease in information it will provide. So, by strongly biasing components in a structural system you can greatly reduce the need for structural information. This trick can be done by biopolymers by selecting complex yet highly consistent structural symmetry, in huge numbers over lots of time. By making structures extremely consistent and highly symmetrical, information can be stored and translated via much simpler structural sequences. After all, the structural information can never completely disappear; it just gets hidden by the monotony of its participating structures. It is a case of steganography, where the true message got hidden in the physical form of the message itself. It never was about a simple relationship between sequence and structure. It is a complex symmetry between the two, but only if that relationship is properly understood.

It is useless to assume that the genetic code spontaneously appeared on earth with a magical inversion of reality that instantly simplified everything, but it is quite useful to note that through time it has almost achieved one. At first blush, sequence really does appear to determine structure, but this is merely an informative illusion, one that can teach us something useful. Proteins have figured out a way to compress huge amounts of complex structural information to be stored in simpler structures built of simpler sequences. There logically must be a huge amount of structural symmetry in this system that must operate through many scales in time. Therefore, the system ultimately selects only parts that enhance that symmetry. This means that the structures and their sequences are hugely biased toward this system's many symmetries over enormous periods of time. This is the only way a system like this could ever work or ever even get itself going in the first place. This also can explain the apparent lack of diversity in its structures today.

We do not currently have a workable definition for molecular information, but we logically know that all molecular information is a function of structural events in time. It would seem then that a good method for controlling events in time is to create molecular sequences. However, complex molecular structures must first be well-suited for this task before it can be reliably done. Even now it takes time for simple molecular sequences to translate into complex molecular structures, but it clearly takes more time for complex structures to make the appropriate molecular sequences. DNA can be uncompressed into enormous amounts of complex protein information relatively quickly, but protein information gets stuffed back into new DNA to make new proteins only about once per generation. The required symmetry between sequence and structure also means that the same structures that make sequences must also be able to recognize those sequences in order to make more structures. The greatest irony is that the current system of storing and translating molecular information appears to us now to contain no structural information. However, the very first system to perform these functions must logically have been purely structural. Molecular structure was the only molecular information for molecules to work with before any system began. Furthemore, molecular structure is the only information any molecule in any system today could ever understand in operating the system we see. Therefore, the original genetic code had to be a primitve language of molecules that related one group of structures to another group of structures, and only through time and intense evolutionary pressure did it become so fabulously sequential that it could travel the long path toward consistent structures that became so well behaved to appear before us now as sequences. In other words, the first genetic codes would have been diverse, erratic and with few truly silent mutations; whereas, today's genetic codes are consistent to the point of appearing "universal" and well behaved to the point where silent mutations seem like they should be expected events. Our current narrative to define, describe and debate this process has been squashed flat and turned around backwards. The entire explanatory language has become so perverted that many of the necessary words are completely missing from the language, and many more have absurd, backward meanings. The supposed first

appearance of the code on earth now rivals any act of creation ever proposed, simply because the story of its beginning must necessarily begin in the middle. This is the inevitable consequence of biting so hard on a simple, powerful, bad idea.

The notion that the genetic code somehow reflects a property of our last common ancestor is also absurd simply because the notion of a last common ancestor is absurd. Nothing so fantastic as a potential candidate for "the last common ancestor of all life on earth" for tomorrow could arise today, nor could it have arisen yesterday. Life has never fundamentally changed its ways. This simple absurd notion of a last common ancestor merely results from a confusion of the concept of ancestry that is always involved with the concept of reproduction. After all, it is not the goal of life to reproduce but to merely produce. Life does not dwell on making copies but thrives on invention. More so than simple mutation, life most effectively invents things by making combinations of things as a useful way to perpetually generate new things. It can only do this when the things to be combined share components of a common system, what can easily be called symmetry. Nowhere is this more obvious than in the case of sexual reproduction, but it is also apparent in everything that life does. Therefore, life most strongly selects the things that make the most effective things for making more combinations of new things. The genetic code is the most effective thing on earth for making combinations of the most things. In this way we might say that it is not the last common ancestor reflected in the structure of the genetic code but rather the first common ancestor. All life today makes protein from nucleotides.

The idea that sequence determines structure has not been disproved in the exact same sense that Pasteur did not disprove spontaneous generation. They are but two examples of false questions. After all, we still basically believe in the correctness behind the general idea of spontaneous generation. However, it is more rightly understood now as a question regarding life on earth as something that can "spontaneously" generate via the process of natural selection – at least once, and perhaps many times. The question of how life arises on earth is one of degrees from the standpoint of requirements of time and circumstance. However, the correct question of the relationship between sequence and structure is truly one of degrees and never should have been one of absolutes. Unfortunately, the question was falsely turned into an absolute. The wrong answer was quickly assumed, yet nobody seemed to have seriously bothered to rigorously test or even question the original idea. The point is that bad ideas must be proven first; otherwise, we only give them power and cause the need to disprove them later. We have always known that sequence determines structure to some degree because it logically must, but to what degree? Does sequence really determine structure? Obviously not, but it was never really proven in the first place, so it hardly seems to be an issue to be disproved now. Unfortunately, our erroneous assumption and unquestioning faith that it was indeed somehow proven at any level has had an enormous impact on the ideas and narratives that such proof would quickly allow us. Now that we can clearly realize that no proof exists, we must carefully but quickly

dismantle the explanatory structures that are built up around this simple, bad idea. They are everywhere and they are insidiously controlling.

It is good to laud and contemplate good ideas. It is sometimes better to clearly recognize and completely reject bad ones. Science is really good at the former and horrible at the latter. It is far easier to convince somebody that they are right than that they are wrong. Science has fallen victim over the past half century to a really bad idea, perhaps the worst idea ever. It competes admirably on virtually every metric with all of history's worst ideas. We should recognize this quickly and work hard to correct it and limit its negative effects going forward. The good news is that bad answers in science yesterday are always the source of excellent questions in science tomorrow. The bad news is that this bad idea has had free reign for so long, so much indoctrination and so much investment has occurred that our entire model is now built in some way around it. Therefore, it will take a long, long time to get rid of and reverse the obvious damage done by this bad idea. This simply means that more patience than expected will now be required before we reach the promised land of bioinformation.

The universe of molecular information is far more complex and far more opaque than the universe of heavenly bodies. The job of science is to help us understand our place in the universe. Clearly, we are not the product of simple molecular sequences but of complex molecular structures operating on complex molecular information. However, the correct relationships between simplicity and complexity, sequence and structure can be clarifying principles in how that universe behaves only if we correctly understand them. The relationship can surely be made simpler, but it has heretofore been made far too simple.