

## Theoretical Analysis of DNA Informatics, Bioindicators and Implications of Origins of Life

Andrew Magdy Kamal

## Author Note:

DNA provides indication for the origin blocks of life. The process on how it stores information is algorithmic. This process, plus other correlations of natural phenomena leads us to a new outlook on the very foundation of life, and one of the most profound questions in all of science.

## Abstract

The usage of Quantum Similarity through the equation  $Z = \{ \forall \theta \in Z \rightarrow \exists s \in S \wedge \exists t \in T: \theta = (s,t) \}$ , represents a way to analyze the way communication works in our DNA. Being able to create the object set reference for z being (s,t) in our DNA strands, we are able to set logical tags and representations of our DNA in a completely computational form. This will allow us to have a better understanding of the sequences that happen in our DNA. With this approach, we can also utilize mathematical formulas such as the Euler–Mascheroni constant, regressional analysis, and computational proofs to answer important questions on Quantum biology, Quantum similarity, and Theoretical Physics.

*Keywords: [DNA, Molecular Biology, Bioinformatics, Computational Physics, Computational Models, Mathematics, Probability, Statistics, Comparative Modeling, Biostatics]*

### Introduction

DNA is the building block of life, and to some extent acts like an algorithm. The way our DNA replicates information as some sort of communicational ladder, isn't a coincidence. Everything, runs in sequences, including the strands of our own DNA. In some extent when cells are reproducing itself, it utilizes instructions that are placed in our DNA. The way this works is quite algorithmic and can be computationally modeled. Once you can encode and decode DNA, you can computationally compare it to other naturally occurring phenomena or even synthetically created bioindicators. This leads us to create a probabilistic distribution on how DNA creates processes of data that gives cells specific types of instructions. This analysis will then allow us to have a comparative model for the origin of life and even the statistical probability of us existing through intelligent means. The theoretical analysis and this proposal can accelerate our own understanding of life, the universe, and computational sciences as we know it.

Mathematical Overview of our DNA	3
----------------------------------	---

## Contents

Theoretical Analysis of DNA Informatics, Bioindicators and Implications of Origins of Life	1
[What is Quantum Similarity?]	4
[Quantum Similarity in our DNA]1	5
DNA Storage	8
Tunneling Algorithm	8
Elements of DNA	9
Logic of DNA's Computational Data	12
DNA's Metacodes	13
Probability of Intelligent Means for Design	14
Constant of Information of our DNA	16
Decomposed Analysis of DNA Design	17
Dynamic Phenotypes and Correctional Networks	17
Regressionally Analyzing our DNA	18
Quantum Origin Boolean Tags	19
Computational Proof Theory	19
Concluding Note	
References	20

[Theoretical Analysis of DNA Informatics, Bioindicators and Implications of Origins of Life]

This paper is an introduction to the object set references and recognizable mathematical patterns of DNA in terms of computational proofs, algorithmic nature, and mathematical representations or formations. This look on Quantum similarity, and approach to DNA informatics or DNA storage will provide a framework on the potential of the subjective nature of related research.

### [What is Quantum Similarity?]

Quantum similarity is an extremely important concept in computational chemistry. Utilizing Quantum similarity allows us to look at different molecular response variances and the relationships between molecular sets<sup>1</sup>. Being able to observe different response variances, and coming up with recognizable patterns and descriptions, allow us to better understand the molecular composition of what we are observing. Some of these patterns can hence be mathematically analyzed and figure out missing links or response sequences that are

difficult to see without using the concept of Quantum similarity. This utilizes advanced Quantum mechanical models that uses theoretical implications of Quantum mechanics and Quantum information to look at response variances, molecular position, states of polarity, chemical bonds, etc. We can then apply a geometric approach<sup>2</sup> to the observation of certain molecular structures and compounds. Utilizing this approach, we can recognize patterns mathematically or create some sort of computational algorithm that utilizes regressional analysis or computer vision to do so for us.

---

<sup>1</sup> Carbo, R., & Calabuig, B. (1992, November). Quantum similarity measures, molecular cloud description, and structure-properties relationships. Retrieved October 22, 2018, from <https://pubs.acs.org/doi/abs/10.1021/ci00010a005>

---

<sup>2</sup> Pothos, E. M., Busemeyer, J. R., & Trueblood, J. S. (n.d.). A quantum geometric model of similarity. Retrieved October 22, 2018, from <http://www.psy.vanderbilt.edu/jdmlab/papers/QuantumSimilarity.pdf>

**[Quantum Similarity in our DNA]<sup>1</sup>**

Quantum Similarity is an important aspect in the genomic analysis of our DNA. We utilize the principles of Quantum similarity to analyze the molecular composition of our DNA, its metacodes and the communication happening back and forth in our RNA systems. The decoding and encoding of expressions in genes can hence be computationally modeled and analyzed utilizing these said principals. The first thing to analyze computationally is structural analysis of the molecular composition, then we look at the geometrical modeling through Quantum similarity. After this, a preset optical flow model is formed. Having an optical flow model will then allow us to use advanced computational principles in data science to look at how our DNA mathematically performs and functions.

For modeling we set up the tagged

set:

$$Z = S * T$$

Where Z being the object you are computationally analyzing, I.e. (The DNA strand)

S is the object set and T is the tagged set formation<sup>3</sup>

When looking at the Boolean tags, we set:

$$Z = \{\forall \theta \in Z \rightarrow \exists s \in S \wedge \exists t \in T : \theta = (s, t)\}.$$

This creates a matrix set for what we are mathematically modeling. You end up having a matrix tree where you start out with the analysis of Z (The DNA) strand and end up with the tagged and object set formation, which is the analysis of the DNA metacodes. Utilizing this method allows you to prepare

---

<sup>3</sup> CARBÓ-DORCA, R., & GIRONÉS, X. (2014, August 13). Foundation of Quantum Similarity Measures and Their ... Retrieved October 22, 2018, from <http://iqc.udg.es/articles/pdf/iqc417.pdf>

for the implementation of QSPR for your computational analysis.

**[Molecular Biology].**



When looking at Quantum Similarity, we are utilizing a computational model for how molecular response variances “communicate” in our DNA systems. This is the origin point of our biological systems.

This will then allow us to create this representation of QSPR. Z represents the entire DNA strand. The tagged reference set T is the base pairs, S is the actual helix/chain holding the base pairs together. Now that you have the object set reference for DNA,

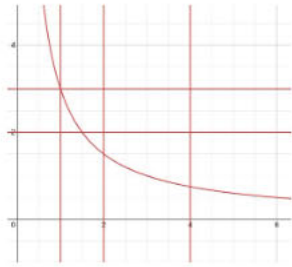
you have the numerical representation of DNA with the set Boolean tags to perform numerical analysis.

$$Z_{AB}(\Omega; \Theta) = \iint \rho_A(\mathbf{r}_1) \Omega(\mathbf{r}_1, \mathbf{r}_2) \rho_B(\mathbf{r}_2; \Theta) d\mathbf{r}_1 d\mathbf{r}_2, \quad 4$$

The equation above signifies QSPR when looking at the relationship of molecular descriptors. Our DNA has similar descriptors where we process information and biological instructions performing almost as a molecular data set. The mathematical and numerical characterization<sup>5</sup> of our DNA allows us to utilize these equations and have a different perspective on how our DNA performs. It performs like sets of sequential data.

<sup>4</sup> Robert, D., Besalu, E., Amat, L., Carbó, R., & Girones, X. (2000). *Molecular quantum similarity in QSAR and drug design*. Berlin: Springer.

<sup>5</sup> Yu, H., & Huang, D. (2013). Descriptors for DNA sequences based on joint diagonalization of their feature matrices from dinucleotide physicochemical properties. *Tsinghua Science and Technology*, 18(5), 446-453. doi:10.1109/tst.2013.6616518



*[Molecular Geometry].*

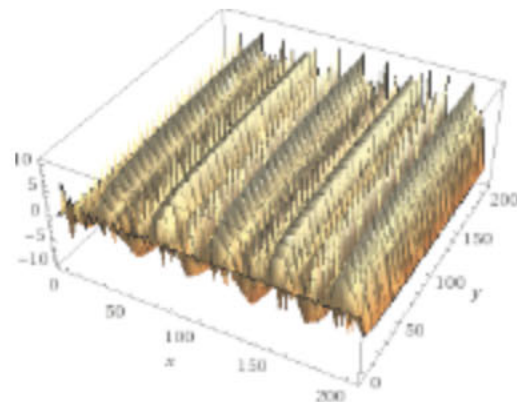
$$y = a + b[x]$$

$$b1 = [\sum(x-\bar{x}) \cdot (y - \bar{y})] / [\sum(x-\bar{x})^2]$$

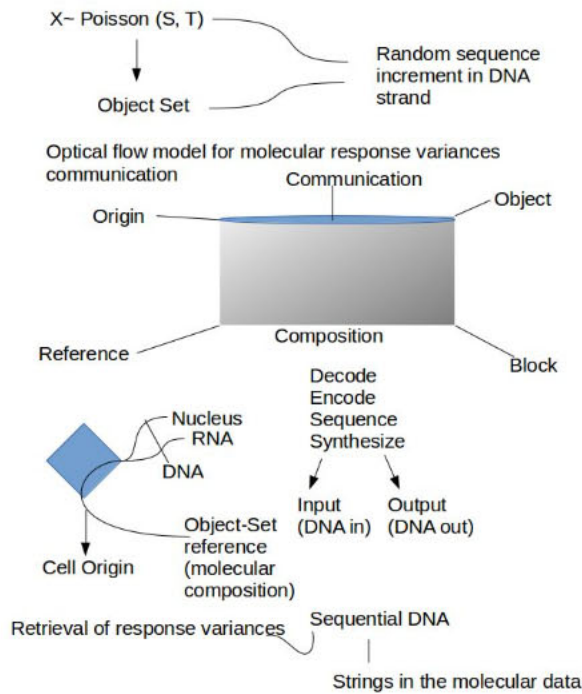
For molecular geometry, we are utilize a regressional model for analyzing object set references in dimensional data. This utilize the above sequence. b1 objectively being the first replicator set for transferring information is analyzed regressionally.

You have the summation of the first strand of molecular data minus its mean. This, it is being subtracted from the secondary strand of data and its mean. You then divide the equation for the first strand to the second power. This small numerical a small piece of data for the reoccurring computational model for processed information.

*The sample below is a harmonic pattern illustration using Wolfram Alpha's computational engine:*







### DNA Storage

This is an example of QSPR being diagrammed for the response variances that happen in our DNA. Here, you have the origin and block for the beginning to end of the DNA strand. In our model, a reference to the response variances and molecular composition of our DNA is also being included. Sequentially, the DNA has these different response variances that communicate information to our cells and genetic makeup.

These are the algorithmic sequences happening inside our DNA.

### Tunneling Algorithm

$$\Gamma = Z \sim (S, T)$$

This sample proof has the Euler–Mascheroni constant which sets the distributions for Z being the object set reference.

This equation can be simplified as the following in relation to Quantum Tunneling and QSPR, and its implications:

$$\Gamma = \lim_{s \rightarrow z} (-\ln s + \sum_{t=1}^s 1/t)$$

The limit is set for the strand and set sequence. The patterns are represented by the continuing base pairs. So for a tunneling algorithm,

$$\Gamma = \lim_{s \rightarrow z} (-\ln s + \sum_{t=1}^s 1/t) \cap \lim_{s \rightarrow z} \rightarrow \infty$$

theoretically you can have this sequential pattern and combined set of elements formulating over and over again. This will be until no chemical information is being



processed. The mathematical representation therefore represent the theoretical processing of information through DNA storage. Just as a computer, one doesn't have a set increment on the set time limit for the

biochemical processing of information. Predictive analysis isn't set for chemical response variances. These harmonic patterns therefore represent continuation of processing of biochemical information.

### Elements of DNA

DNA has a double helix pair and nucleotide bases that store information for biological functions. What is fascinating in the field of computational sciences, is what DNA is theoretically capable of doing. According to *Science Magazine*,<sup>6</sup> one can store 215 petabytes worth of data into a single gram of DNA. The fact that DNA has these base pairs to form these computational like functions that transfer data from one pair to another seem quite applicable on

what understanding DNA can mean for the future of computing. The idea that one can have this synthetic algorithmic network that functions in any systematically similar capacity, changes the way modern algorithms are designed in general.

Even Quantum Computing in essence will have to be put in an entirely different scale. When combining this with other theoretical technologies, the future seems almost limitless. Technologies such as interstellar communications, discovering nuclear fusion, or building space crafts traveling in light years would be considered

---

<sup>6</sup> Service Mar, R. F., Blaustein Oct, R., Mervis Oct, J., Voosen Oct, P., Waldman, S., & E&E News Oct. (2017, December 08). DNA could store all of the world's data in one room. Retrieved October 22, 2018, from <http://www.sciencemag.org/news/2017/03/dna-could-store-all-worlds-data-one-room>

child's play compared to understanding the elements of our DNA from a computational standpoint. Even simulating Super String theory or M theory would not be as

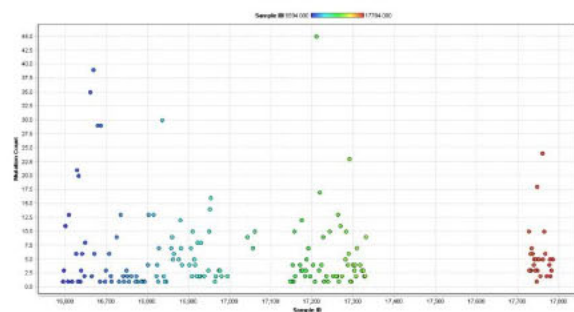
sophisticated as what this research can provide. Using the same type of sequential logic DNA uses, one can create custom algorithms for solving the unknown.

### Implications in terms of Genomics

In terms of genetics, it can take about a day for some of the most advanced computational infrastructure to analyze the human genome. The Microsoft Azure cloud has made significant advances in lowering the cost of genomic analysis.<sup>7</sup> Even advanced in mathematics and mathematical proof allow us to create algorithms that more advancedly analysis genomic sequencing. The ability to recursively analyze processes in computing and the trivial mathematical theorems one can

provide for forming large algorithmic proofs expands the possibility of this research. Quantum similarity and many other advanced concepts are also taking a large role. Advancements in modern medicine will be possible once people map out genomic data and computationally analyze it against case studies or open bio portals. An example of this is immunology or researching into cancer genomics.

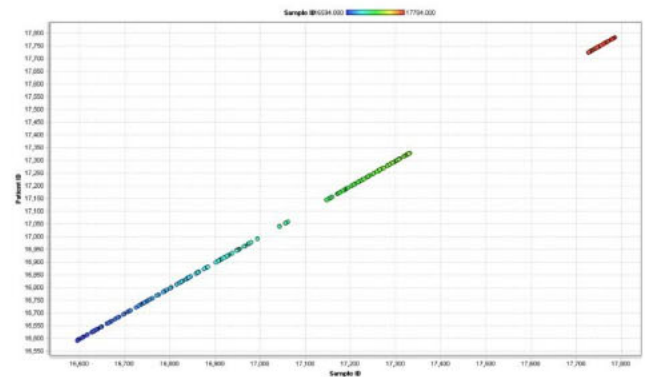
<sup>7</sup> Bergeson, T. (2018, June 15). Power the democratization of genomics with the cloud. Retrieved October 22, 2018, from <https://cloudblogs.microsoft.com/industry-blog/industry/health/power-the-democratization-of-genomics-with-the-cloud/>



The shrunk image above is take from an open source project I done<sup>8</sup> utilizing RapidMiner studio and open portals for genomics data. This is just a small example of what is possible with genomic sequencing. Creating a small algorithm for regressional analysis, one is able to look at genetic mutational views or complexities not easily viewed on open data. This is just a small fraction compared to the advanced processes that can be done computationally. Utilizing more advanced computational proofs or constraints, will allow us to solve complexities that can modernize medicine. The usage of Quantum similarity and logic along with the usage of genomic sequencing, allow for us to form more accurate analysis. This is because once you know how DNA communicates, how it

<sup>8</sup> Kamal, A. M. (2016). Data Mining Cancer Genomics Case Studies. Retrieved October 22, 2018, from <https://devpost.com/software/data-mining-cancer-genomics-case-studies>

stores information, and have a generalized idea of the sequential processes of DNA, the accuracy of your models will be much more in depth.



Also this shrunk image above from my same project, utilize the same predictive analysis for regressionally analyzing these sample ids, genetic mutations, and case studies in a linear data visualization.

Being able to have set computational models or visualizations utilizing sequential processes and mapping allows you to get much more accurate data sets while consuming less storage. The computational data can sequentially be analyzed

recursively in a similar manner to the logic of QSPR.

### Logic of DNA's Computational Data

DNA works logically with its object set references. One looks at the patterns that our DNA creates as fairly algorithmic. You have biological information being communicated through the chemical synthesis that happens in your DNA. As mentioned, this shows that DNA acts almost as its own computer. One can view a computer as having these halting sequences, or mathematical complexities such as Quasi-Polynomial time. There are numerous examples or reoccurring logical sequences that a computer takes its time to solve. The way our DNA transcribes information, functions in a similar but more advanced manner.

You have a set limit for these base pairs and in time complexity algorithms, exponential time represented is as  $2^{\text{poly}(n)}$ , you have some

biochemical processes formulating slower than others. For Quasi-Polynomial proofs such as  $2^{\text{poly}(\log n)}$ , one can look at naturally occurring time patterns through approximation. Although genetic algorithms or representations are represented through more linear processes, DNA transmits data almost through polynomial set references. You have these bases having object set references that follow recursive patterns and biologically transcribe instructions over and over again. Each one is sort of tagged and has this reference point for an advanced biochemical process. Some biochemical processes have constraints and others don't. They all follow a set reference that is still naturally algorithmic. Quantum Similarity represents these set reference happening in

our molecular response variances and an overall metric that leans towards the mathematical representation of our DNA.

### DNA's Metacodes

Metadata describes other data or information about information.<sup>9</sup> If you look at our DNA in general, the chemical synthesis that happens transcribes instructions. The metacodes or data for that information is the chemical processes, combinations, and response variances that describe a certain biochemical process. Theoretically, one can say that every biochemical process follows a certain type of representative pattern at a representative time. Looking at QSPR and time complexity, one represents this process going on. Certain metacodes or descriptors of information are more unique than others.

This is why sequencing allows us to look at certain complexities or unique descriptors. Just as advanced computers are analyzing certain complexities such as unknown mathematical proofs, the complexities in DNA are much more advanced. One can theoretically represent a found complexity and utilize algorithmic designs or programs for formulating highly computational tasks. The knowledge unlocked in certain molecular descriptors are in essence what formulates certain building blocks or biochemical processes happening at certain times. Even micro evolution, heredity, or genetic mutations follow a set object set

---

<sup>9</sup> Metacode. (n.d.). Retrieved October 22, 2018, from <https://www.thefreedictionary.com/Metacode>

reference pattern of transcribed biochemical code.

### Probability of Intelligent Means for Design

Next we look at what this means in terms of the theoretical origin of things. As questioned in the “Evolution 2.0 Prize”,<sup>10</sup> the information in knowing the algorithmic processes of our DNA changes the way we look at science in general. One can see that these objective set of mathematical laws formulate and govern the universe.<sup>11</sup> People question the objectivity of whether or not math in essence can prove the existence of intellectualism involved in the creation of the universe. However, the mathematical implications and patterns that go into our DNA can provide some usage of practical

implications. Even human consciousness follows some sort of indicative pattern. As humans have free will, one can indicate that the patterns happening deeply in their neural thinking processing formulate at different times and time complexities.

The late Prof. Hawking<sup>12</sup> even argued that since the universe has fundamental laws of sciences, then a creator must be subject to these laws. However, the reverse logic of this argument is that it is theoretically feasible that before the existence of time itself something or someone fundamentally created set laws of origin. This is not an argument of pseudoscience because

mathematics can universally exists

<sup>10</sup> Marshall, P. (n.d.). Artificial Intelligence Origin of Life Prize, \$5 Million USD. Retrieved October 22, 2018, from <https://www.herox.com/evolution2.0>

<sup>11</sup> Nelson, R. H. (2017, May 30). An economist has studied the data and concluded God exists. Retrieved October 22, 2018, from <https://www.independent.co.uk/life-style/existence-of-god-rational-arguments-mathematics-human-consciousness-a7739841.html>

<sup>12</sup> Gabbatt, A. (2010, September 02). Stephen Hawking says universe not created by God. Retrieved October 22, 2018, from <https://www.theguardian.com/science/2010/sep/02/stephen-hawking-big-bang-creator>



independent of our own physical realities.<sup>13</sup> In comparison to how our DNA processes information and Quantum similarity in general, even biological processes are subject to the fundamental laws of mathematics. In theory one can argue that random genetic mutations are also indicative of a mathematical pattern. One can argue that the laws of mathematics existed since the beginning, however, the fundamental question lies that if math can lie outside our physical realities, it is indicative of some sort of intellectualism.

The set theories of mathematical laws is that even DNA has a higher proposed implication then what was once thought of. The theoretical analysis of how we look at our DNA could be quite intriguing to the fundamental unsolved

---

<sup>13</sup> Nelson, R. H. (2018, September 18). Arguments why God (very probably) exists. Retrieved October 22, 2018, from <http://theconversation.com/arguments-why-god-very-probably-exists-75451>

problems of theoretical physics. Even looking at Quantum gravitational theory of atomism, one sees that theoretical physics can be in harmony together. Questions such as the average live of a proton or things physics is yet to determine, could all be looked at as time complexities or recursive sequences. This in essence shows that DNA is subject to similar fundamental mathematical laws as other occurring phenomena. The main difference is in the mathematical representation of these laws as well as their proofs and formation.

According to Dr. Ali Binazir<sup>14</sup>, the odds of one existing is 1 in  $10^{2,685,000}$  and this number expands as the human population expands exponentially. Even looking at just a sperm meeting an egg, the probability is still 1 in 400 trillion. However, in

---

<sup>14</sup> Spector, D. (2012, June 11). The Odds Of You Being Alive Are Incredibly Small. Retrieved October 22, 2018, from <https://www.businessinsider.com/infographic-the-odds-of-being-alive-2012-6>



implications for genomics, the probability of humanity existing at all gets slimmer and slimmer. This is because the rarity of the computational patterns in our DNA happening the way that it is. These mathematical probability vs. possibilities have less to do on whether or not we look at the existence of the universe as unlikely, rather than the statistical nature of the laws of the universe being governed solely by math. Even randomization is impacted by the laws of logic. If math is in harmony with being outside of our physical realities, then the way we look at the fundamental building

blocks of life is quite profound as mathematically subjective. The theoretical implications of what science believes governs these mathematical laws has even bigger implications than what one might realize. The creation of dark matter itself or the universe's origin point can then be computationally representative of the algorithmic nature of the universe. As fundamental physics must have invoking algorithmic representations and proofs to lead to major breakthroughs in Quantum physics.

### **Constant of Information of our DNA**

Constants are non-altered values<sup>15</sup> or information being processed in computer programming. Regular chemical synthesis

and even genetic forms of mutation have constants in their values. Even when a mutation happens, a sequential flow also happens in the transfer of biochemical information. This sequential flow creates a trace of data in the nucleotides that store

---

<sup>15</sup> ". (2018). Constant. Retrieved October 22, 2018, from <https://www.encyclopedia.com/science-and-technology/computers-and-electrical-engineering/computers-and-computing/constant>

information about the chemical sub processing going on. The idea of Quantum similarity and a certain set of object references for these chemical processes show more of an algorithmic sequencing then what one might realize.

Chemical synthesis seems to always be taking place but the rate, complexity, and time of a set process is deterministic on the nature of the instructions being transcribed. These recurring processes happen in different forms and sequences.

### Decomposed Analysis of DNA Design

The way our DNA works seems almost computational in its natural form. You have these biological processes and mathematical thresholds demonstrating its chemical synthesis. Things such as the constraints showed previously, as well as the regressional transcribing of biochemical instructions formulate a base point on how

the information is transmitted. The metacodes and what is being described as well as the non-altered values represent a logical flow of information. Just as neural networks propagates output layers or recursively transmits information, the same can be said about our DNA's base pairs.

### Dynamic Phenotypes and Correctional Networks

The way one analyzes neural networks are similar to types of computational networks being analyzed in our DNA. Binary Neural Networks in particular, have this deterministic threshold for their

functionality. A DBN or (Deterministic binary neuron) can hence be mathematically represented:  $DBN(x) = 1_x > 0$ <sup>16</sup>. This

---

<sup>16</sup> Dong, H., & Yang, Y. (2018). Binary Neurons. Retrieved October 22, 2018, from <https://salu133445.github.io/bmusegan/neurons.html>

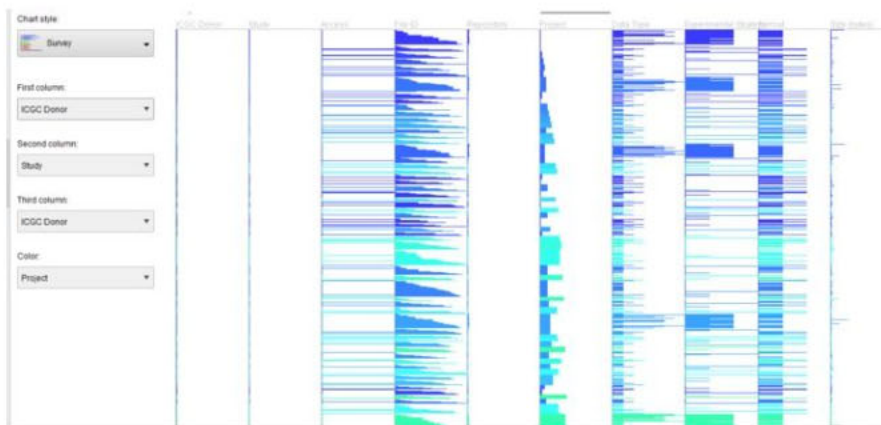
is similar, but not as advanced as the computational representation for the complexities of our DNA's processing of sequential and chemically synthetic information. Utilizing mathematical

representations of biological networks and formations allow us to create indicators or advanced mathematical models for our algorithms to systematically optimize functions to better analyze complexities.

### Regressionally Analyzing our DNA

Using QSPR, and genomic sequencing is important for the regressional analysis of our DNA. As proposed earlier, mapping out the

The figure above<sup>17</sup> is another sample from my RapidMiner project. This is a view of ICGC donor data from that same data set.



One is able to view different data by donor, sample ID, and see the complexities in certain genome types or mutations. Things such as combinations of helper cells, or gene therapy or immune boosters are more

feasible when you utilize genomic human genome and creating subsets of sequential patterns allow us to lead to further biomedical breakthroughs.

feasible when you utilize genomic

<sup>17</sup> GDGC Data View [Personal figure screenshotted from RapidMiner]. (2018).

sequencing with large data sets and algorithmic proofs. This is yet another small example of the theoretical potential that is possible with the mathematical representation of our DNA, and the overall

view of what genomic sequencing can provide. Hence, another reason on why mathematical formulations can be used to represent both biological or non-biological networks.

### Quantum Origin Boolean Tags

For object set references, one can represent Quantum similarity set references as well with Boolean tagging. The same Boolean tagging that can be used to represent Quantum information can be utilized in a representation for the states of sequential information being processed in our DNA. One set proof is the preliminary function  $f: \{0, 1\}^n \rightarrow \{0, 1\}$ <sup>18</sup> where you are representing  $n$  variables. The same tags can be applied to object set references.

### Computational Proof Theory

One equation I want to reference is the Ramanujan- $\Delta$ -function:  $\tau(n) = O(n^{2+\epsilon})$  for all  $\epsilon > 0$ <sup>19</sup>. You have an indicative reference to space and functional complexities for viewing regressional complexities. If you utilize the same constraint as  $\Gamma = Z \sim (S, T)$ , and a set its limit, you can apply a functional proof for representing recursive sequential data in chemical synthesis. So the reference is  $Z \sim (S, T) = O((S, T) + \epsilon)$ ,  $\epsilon$  utilizing the origin point of the strand's object set reference instead of zero.

---

<sup>18</sup> [arXiv:0810.2435](https://arxiv.org/abs/0810.2435) [quant-ph]

<sup>19</sup> Jannsen, U., Dr. (2016). Deligne's Proof of the Weil-conjecture. Retrieved October 22, 2018, from <http://www.mathematik.uni-regensburg.de/Jannsen/home/Weil-gesamt-eng.pdf>

### Concluding Note

The final response is a conclusion on these proofs. The view that DNA and biological life has an algorithmic design, among with the view of the mathematical representations regarding Quantum similarity, object set references, computational proof theory, and complexity algorithms leads to the introduction of Quantum Similarity Origin Point References or QSOPR Theorem in general.

### References

1. Carbo, R., & Calabuig, B. (1992, November). Quantum similarity measures, molecular cloud description, and structure-properties relationships. Retrieved October 22, 2018, from <https://pubs.acs.org/doi/abs/10.1021/ci00010a005>
2. Pothos, E. M., Busemeyer, J. R., & Trueblood, J. S. (n.d.). A quantum geometric model of similarity. Retrieved October 22, 2018, from <http://www.psv.vanderbilt.edu/jdmlab/papers/QuantumSimilarity.pdf>
3. CARBÓ-DORCA, R., & GIRONÉS, X. (2014, August 13). Foundation of Quantum Similarity Measures and Their ... Retrieved October 22, 2018, from <http://iqc.udg.es/articles/pdf/iqc417.pdf>
4. Robert, D., Besalu, E., Amat, L., Carbó, R., & Girones, X. (2000). Molecular quantum similarity in QSAR and drug design. Berlin: Springer.
5. Yu, H., & Huang, D. (2013). Descriptors for DNA sequences based on joint diagonalization of their feature matrices from dinucleotide physicochemical properties. *Tsinghua Science and Technology*, 18(5), 446-453. doi:10.1109/tst.2013.6616518
6. Service Mar, R. F., Blaustein Oct, R., Mervis Oct, J., Voosen Oct, P., Waldman, S., & E&E News Oct. (2017, December 08). DNA could store all of the world's data in one room. Retrieved October 22, 2018, from <http://www.sciencemag.org/news/2017/03/dna-could-store-all-worlds-data-one-room>
7. Bergeson, T. (2018, June 15). Power the democratization of genomics with the cloud. Retrieved October 22, 2018, from <https://cloudblogs.microsoft.com/industry-blog/industry/health/power-the-democratization-of-genomics-with-the-cloud/>
8. Kamal, A. M. (2016). Data Mining Cancer Genomics Case Studies. Retrieved October 22, 2018, from <https://devpost.com/software/data-mining-cancer-genomics-case-studies>
9. Metacode. (n.d.). Retrieved October 22, 2018, from <https://www.thefreedictionary.com/Metacode>
10. Marshall, P. (n.d.). Artificial Intelligence Origin of Life Prize, \$5 Million USD. Retrieved October 22, 2018, from <https://www.herox.com/evolution2.0>



11. Nelson, R. H. (2017, May 30). An economist has studied the data and concluded God exists. Retrieved October 22, 2018, from <https://www.independent.co.uk/life-style/existence-of-god-rational-arguments-mathematics-human-consciousness-a7739841.html>
12. Gabbatt, A. (2010, September 02). Stephen Hawking says universe not created by God. Retrieved October 22, 2018, from <https://www.theguardian.com/science/2010/sep/02/stephen-hawking-big-bang-creator>
13. Nelson, R. H. (2018, September 18). Arguments why God (very probably) exists. Retrieved October 22, 2018, from <http://theconversation.com/arguments-why-god-very-probably-exists-75451>
14. Spector, D. (2012, June 11). The Odds Of You Being Alive Are Incredibly Small. Retrieved October 22, 2018, from <https://www.businessinsider.com/infographic-the-odds-of-being-alive-2012-6>
15. ". (2018). Constant. Retrieved October 22, 2018, from <https://www.encyclopedia.com/science-and-technology/computers-and-electrical-engineering/computers-and-computing/constant>
16. Dong, H., & Yang, Y. (2018). Binary Neurons. Retrieved October 22, 2018, from <https://safu133445.github.io/bmusegan/neurons.html>
17. GDGC Data View [Personal figure screenshotted from RapidMiner]. (2018).
18. Montanaro, A., & Osborne, T. J. (2010, August 24). Quantum boolean functions. Retrieved October 22, 2018, from <https://arxiv.org/abs/0810.2435>
19. Jannsen, U., Dr. (2016). Deligne's Proof of the Weil-conjecture. Retrieved October 22, 2018, from <http://www.mathematik.uni-regensburg.de/Jannsen/home/Weil-gesamt-eng.pdf>

