Acta Scientific Medical Sciences (ISSN: 2582-0931)

Volume 9 Issue 9 September 2025

Hypothesis

QED (Quadruplet Expanded DNA) Eukaryote Genetic Code Developmental Journey and Incurable Diseases

Rama Shankar Singh*

Research and Development, GEN-MED R&D, LLC, Orlando, FL 32817, USA

***Corresponding Author:** Rama Shankar Singh, Research and Development, GEN-MED R&D, LLC, Orlando, FL 32817, USA. Email: gen-med@comcast.net

Received: July 15, 2025

Published: August 06, 2025

DOI: 10.31080/ASMS.2025.09.2139

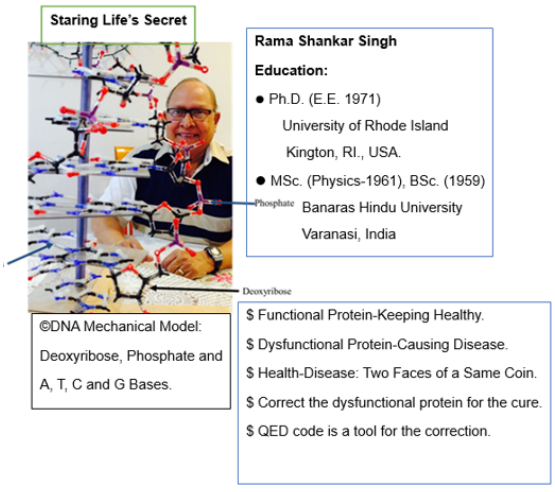
Abstract

After a decade of pioneering research, I have successfully developed the first Quadruplet Expanded DNA (QED) code for eukaryotic cells. This groundbreaking genetic system builds on the natural DNA bases (T, C, A, and G) to establish twenty independent, non-degenerate protein-coding and thirty-five distinct regulatory codons. These regulatory elements enable precise control of gene transcription, post transaction RNA splicing, and translating mRNA for protein synthesis across eukaryotes, viruses, and prokaryotes.

QED overcomes the limitations of the long-standing prokaryotic triplet genetic code and the orthogonal expanded quadruplet genetic code designed for unnatural amino acids. By facilitating the creation of functional proteins that maintain cellular homeostasis—and correcting dysfunctional proteins that cause disease—QED opens new frontiers in gene therapy.

This transformative technology holds the potential to revolutionize therapeutic strategies by enabling the repair of defective proteins and their underlying genetic sequences at both the DNA and protein levels, marking a paradigm shift in the treatment of genetic disorders and ushering in a new era of molecular medicine.

Keywords: QED (Quadruplet Expanded DNA); Human Diseases



QED eukaryote code and human diseases

Gene-based disease model

Human eukaryote cells synthesize functional proteins to maintain cellular homeostasis. Mutations in genes and errors in synthesis yield a dysfunctional protein that contributes to the disease. Correcting dysfunctions at both the DNA and protein levels is crucial for therapeutic advancements.

Dysfunctional protein correction

Dysfunctional proteins that cause diseases could be corrected is sketched at the protein level, Figure 1, or through gene therapy at the DNA level using QED code, Figure 2.

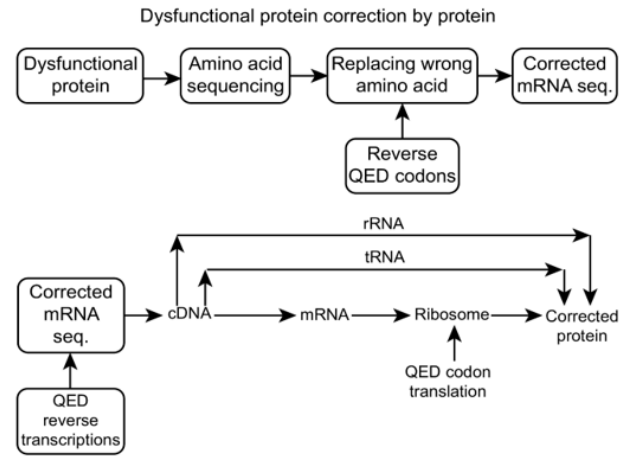


Figure 1: Dysfunctional protein correction at protein level.

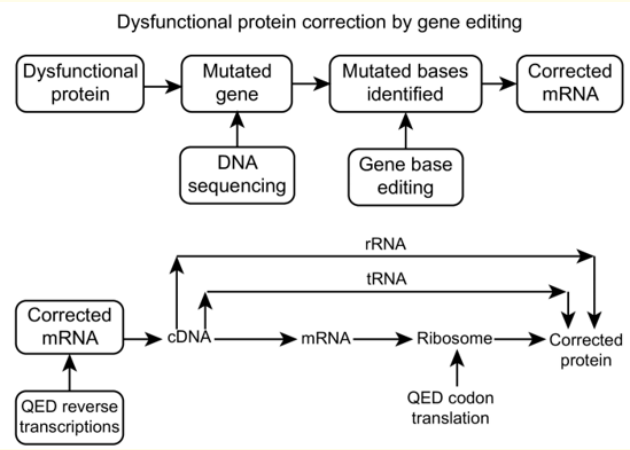


Figure 2: Dysfunctional protein correction at DNA level, Gene Therapy.

Timeline: 1968 to present

1. **May 1968**: Birth of a child diagnosed with a rare, incurable disease. An attending physician hypothesized a gene-disease causality—the origin of the scientific inquiry.
2. **1971-2005**: Pursued a career in physics and engineering, laying the groundwork in analytical and computational sciences.
3. **2005-2010**: Transitioned into molecular biology and genetics, acquiring the necessary skills to investigate gene-disease relationships.
4. **2010-2025**: Development of the QED eukaryotic genetic code model:
5. Developed the QED protein-coding and non-coding code (Dec 2023).
6. Established the correlation between QED non-coding codons and cis-regulatory elements (Sept 2024).
7. Illustrated the principle of information flow in biological protein synthesis (June 2025).
8. Predicted the QED codon table (June 2025).

- **2025-Present**:

1. Investigating the QED code's role in elucidating gene-disease causality.
2. Exploring the application of QED in developing cures for monogenic and multigenic rare diseases, cancers, and neurodegenerative disorders.

Key Peer-reviewed publications

1. Singh, Rama Shankar. "Quadruplet Expanded DNA (QED) Genetic Code for Eukaryotic Cells." Acta Scientific Medical Sciences, 7.12 (2023): 70-82. DOI: 10.31080/ASMS.2023.07.1720.

2. Singh, Rama Shankar. "Correlation between Eukaryotic Noncoding QED Genetic Codes and Cis-Regulatory Elements". Acta Scientific Medical Sciences, 8.10 (2024): 89-96. DOI: 10.31080/ASMS.2024.08.1929.

3. Singh, Rama Shankar. "QED Eukaryote Genetic Code and Principle of Information Flow and Biological Protein Synthesis". Acta Scientific Medical Sciences, 9.6 (2025): 100-114. DOI: 10.31080/ASMS.2025.09.2097.

Encoding assignment Table prediction.

Amino Acids	QED Codons	HB Bonds	QED Codons	Amino Acids
Arg	(GA)(GA)	10	(CU)(CU)	Leu
Asn	(AA)(CC)	10	(UU)(GG)	*Met
Cys	(UG)UU	9	(CA)AA	Gln
Glu	(GA)AA	9	(CU)UU	Ser
Gly	GGGG	12	CCCC	Pro
His	(CA)CC	11	(UG)GG	Trp
Lys	AAAA	8	UUUU	Phe
Thr	(AC)(CA)	10	(GU)(GU)	Val
Tyr	(UU)(CC)	10	(GG)(AA)	Asp
lle	(UC)CC	11	(GA)GG	Ala

Table 1: The QED code encoding assignment table prediction (2023 Ref . 1) (like triplet coding hypothesis 1963, verified 1968) needs verification.

QED codons:

- Twenty protein encoding codons.
- Thirty-five noncoding control codons.
- Nondegenerate independent codons.

The predicted QED encoding assignment table1 has some unique and interesting properties. The encoding codon of one amino acid is the anticodon of the other amino acid: (GA)(GA) encodes Arg, and its anticodon (CU)(CU) encodes Leu. Similar attributes are noted for the others. Also, the anticodon of tRNA (CU)(CU) Leu will decode the (GA)(GA) Arg codon, maintaining proper HB bonds and assuring a secure protein synthesis. Consequently, efficient use of tRNA is anticipated.

Start of the journey

My journey from a physicist and engineer to a pioneer in eukaryotic genetic code has been both inspiring and significant. The development of the "Quadruplet Expanded DNA (QED) Genetic Code for Eukaryotic Cells" marks a monumental achievement in molecular biology and overcomes the limitations of the 50 years accepted Triplet genetic code [4] and the orthogonal expanded quadruplet genetic code for unnatural amino acids [5-7].

The relentless desire to understand gene-disease causality and seek potential cures for rare diseases fueled my work. Human eukaryotic cells carry hereditary DNA information and the molecular tools necessary to synthesize proteins, maintaining homeostasis. However, genetic variants, transcription errors, and splicing defects can lead to dysfunctional proteins, ultimately causing disease. Thus, correcting dysfunctional protein ought to be the first developmental step in finding cures of human diseases.

In eukaryotes, DNA resides in the nucleus, where information flows through transcription and splicing, producing mRNA that is transported to the cytoplasm. There, the mRNA undergoes translation, where codons encode proteins. QED code is the first and only code to control the eukaryote transcription process. Also, the QED eukaryotic genetic code is designed to correct dysfunctional protein synthesis, offering a promising approach for treating monogenic incurable rare diseases, as well as complex conditions like multigenic cancers and neurodegenerative diseases [3].

Through decades of research, perseverance, and a profound personal motivation, the QED genetic code has emerged as a potential key to unlocking treatments for some of the most challenging human diseases.

Motivation

The development of the QED genetic code is deeply personal. In 1968, while pursuing my Ph.D. in Electrical Engineering, the birth of our child with a rare, incurable disease catalyzed a profound shift in our life's direction. The attending physicians suggested that the condition might be linked to our genes. At the time, the concept of genes was completely foreign to me, and the idea that they could be responsible for our child's illness was both

bewildering and haunting. Genetic testing in 1968 was purely science fiction. Despite the absence of genetic testing at the time, my determination to understand the underlying causes of genetic diseases set the foundation for my future endeavors, ultimately leading to the development of the QED genetic code.

Overcoming challenges and the driving force

Changing the field of expertise from one to the other requires time and resources. Resources can be raised, but time is fixed. When a physicist wants to be a biologist [8] or a biologist wants to be a physicist [9], they spend time and resources developing skills. I had no such luxury.

Balancing a full-time career as a physicist and engineer with the pursuit of genetic knowledge was no small feat. Raising a family while caring for a child with an incurable disease presented significant financial and emotional challenges. To navigate these obstacles, I pursued a dual professional path to navigate the challenges of time and resource constraints effectively: Ten years in academia followed by thirty years in the industry which paid the bills and part time developing the Molecular Biology and genetics skill.

- Physicist - Ten years academic. Post Doctoral fellowship for a year at NBS followed by teaching Physics and guiding research at the Physics Department, University of Puerto Rico for seven years. I had a two years sabbatical leave at MIT Lincoln Lab., Lexington, MA.
- Engineer - Thirty years: Sr Process Engineer (3yrs), GE-AESD, Utica, NY; Sr Professional Staff (10yrs) Martin Marietta, Orlando, FL; and Sr. Staff (17yrs), Lockheed Martin, Sunnyvale, CA.
- Molecular Biologist and Geneticist - A decade dedicated to developing the skills for understanding gene disease causality relation and be prepared to develop the QED eukaryote genetic code.

My dedication to self-educating in genetics and molecular biology was by drawing inspiration from Michael Faraday, who, despite limited formal education, became a renowned scientist through self-study and perseverance, I had embarked on a similar path. My approach mirrors Faraday's in many ways: leveraging available resources, committing to continuous learning, and applying acquired knowledge to address complex challenges. This mindset has undoubtedly been instrumental in my developing the Quadruplet Expanded DNA (QED) Genetic Code for Eukaryotic Cells.

Educational journey in genetics and molecular biology

Upon relocating to Silicon Valley in 1997 and working at Lockheed Martin, I seized various opportunities to develop my genetic skills:

- **Evening Classes**: I attended courses at esteemed institutions such as Stanford University, UC Santa Cruz, and UC San Francisco, covering topics like molecular biology, bioinformatics, gene expression, and genomic medicine.
- **Online Courses**: Through platforms like Coursera, I completed over 25 courses offered by institutions including Columbia University, Duke University, Johns Hopkins, Peking University, and UC San Diego. These courses spanned subjects like monogenic rare diseases, bioelectricity, epigenetics, recombinant DNA technology, precision medicine, virology, biostatistics, and genomics.
- **Seminars and Conferences**: I actively participated in weekly seminars at Stanford, UC Berkeley, UCSF, and local biotechnology companies, staying abreast of the latest developments in the field.

Nobel prize research

I closely followed Nobel Prizes in Chemistry and Medicine, recognizing them as valuable sources of breakthrough research and fundamental concepts. Notably, I studied key Nobel-winning discoveries related to DNA structure (1962), gene control (1965), triplet gene code verification (1968), split genes (1993), eukaryotic transcription (2006), and ribosome structure (2009).

A 1976, pre-QED code developmental incident - nonoptimal triplet coding

Around 1976, While teaching and guiding research in Physics, University of Puerto Rico, Mayaguez, I came across an article in Scientific American that DNA is a heredity material and has four bases: Adenine (A), Thymine (T), Cytosine (C) and Guanine (G).

The combination of three out of four bases, a coding problem, encodes twenty amino acids to synthesize proteins to keep us healthy. This triggered my memory of 1968 incidence of the birth of our daughter and the attending physicians' comment that our gene may be the cause. My curiosity peaked to explore further to find gene disease causality relationship.

Claude E. Shannon's coding theory [10] was taught to us in communication theory class. According to Shannon, the optimum bit required to code N object is $\log_2 N$. For coding twenty amino acids, the optimum bits would be $\log_2 20$ =4.325 bits. But triplet coding had sixty-four triplet codons (4X4X4=64) and their representation will require 6 (2^6) bits. Consequently, the triplet is not an optimum and degenerate coding where more than one codon encodes an amino acid. My physicist's investigative instinct peaked to explore it further. However, due to physics teaching, research and raring family responsibilities, I had to postpone for the future endeavor.

In 2010, after waiting for 30 years (1976-2005) and completing parallel Physics/Engineering and Genetic Skill development career, I retired as a Senior Staff from Lockheed to devote full time developing QED code. This was similar to a battlefield scene in epic Mahabharat- Bhagwat Geeta, where warrior Arjun bowing with folded hand before Lord Krishan, after listening his advice till the end of 18th chapter saying:

नष्टो मोहःस्मृतरि्लब्धा त्वप्रसादान्मयाच्युत ।

स्थतिोऽस्मि गतसंदेहःकरष्ये वचनं तव ॥18.73॥

naṣṭō mōhaḥ smṛtirlabdhā tvatprasādānmayācyuta.

sthitō.smi gatasandēhaḥ kariṣyē vacanaṅ tavall18.73ll

English translation: O Lord, by your grace, my delusion has ended and I have gained my memory. My reason has awakened and doubts ended. I will do whatever you command.

QED code development

Time was now ripped for developing QED code by critically reviewing triplet coding and its verification. The central idea of QED code development lies in the triplet coding verification. In 1968, Robert W. Holley, H. Gobind Khorana, and Marshall N. Nirenberg verified 64 triplet codons, 61 encodings, one START, and two STOPs triplet codons [11]. Holley established the tRNA structure. Khorana and Nirenberg synthesized polyribonucleotide. Khorana's synthesis [12] and his Nobel Prize lecture indicated that poly-rAU and - rCG containing two adjacent bases that naturally form base pairs are chain terminator or STOP.

Inspired by H. Gobind Khorana's findings that adjacent base pairs (AU and CG) are non-coding, the QED code extends this concept to a quadruplet coding system, using all four DNA bases (T, C, A, and G), base-position-independent, symmetric, and (AU) NN, (CG)NN, (N any base) noncoding. It defines 20 non-degenerate, independent coding codons and 35 non-degenerate, independent noncoding ones [1], aligning with known cis-regulatory elements [2] and offering a potential universal model [3] for gene regulation and protein synthesis in eukaryotes, prokaryotes, and viruses.

Similar to the triplet genetic code hypothesis introduced in 1963 (and experimentally validated by 1968), this model also awaits verification.

The proposed QED (quadruplet expanded DNA) eukaryote theoretical model is based on the following assumptions:

1. All four DNA (A, T, C, and G) bases are involved; in mRNA, T is replaced by U
2. The base positions are independent; i.e., for any A and B, AB and BA are equivalent.
3. The base positions are symmetric; i.e., for any A and B, (AB) and (BA) are synonymous
4. The self-complementarity forming adjacent base pairs with any two adjacent NN (N any A, T, C, or G) bases, (AT) NN and (CG) NN, is noncoding.

Following the assumption (3), (AT)(NN)) and (NN)(AT)) are synonymous; likewise, (CG)(NN) and (NN)(CG) are synonymous.

The independent QED codons are generated using the property of a square symmetric matrix. A N x N square symmetric matrix with N rows and N columns has N x (N+1)/2 independent elements. QED code elements are generated arranging four DNA bases (T, A, C, and G) in a square symmetric matrix in two steps. First, four bases are arranged in a 4X4 square symmetric matrix yielding 4 X (4+1)/2= 10 independent elements. Next ten elements

are arranged in a 10X10 square symmetric matrix leading to 10 X (10+1)/2=55 independent elements. Out of the fifty-five, the thirty-five independent noncoding elements are estimated following QED assumption (3):

Sixteen noncoding (AU)NN, N=any base; sixteen noncoding (CG) NN, N= any base; one noncoding (AU)(AU); one noncoding (CG)(CG); and one noncoding (AU)(CG).

The remaining (55-35=20) twenty are encoding. The independent twenty encodings and thirty-five noncoding codon bases were generated using four DNA (T, C, A, and G) bases [1], listed in Table 2 and 3.

	QED (Quadruplet expanded DNA) eukaryote codons				Hydrogen Bond
	Codons	Synonymous Isocodons, T(U)			
1	UUUU	UUUU			8
2	CCCC	CCCC			12
3	AAAA	AAAA			8
4	GGGG	GGGG			12
5	(AA)(CC)	(CC)(AA)			10
6	(UC)CC	(CU)CC	CC(UC)	CC(CU)	11
7	(UG)UU	(GU)UU	UU(UG)	UU(GU)	9
8	(UG)GG	(GU)GG	GG(UG)	GG(GU)	11
9	(CA)CC	(AC)CC	CC(CA)	CC(AC)	11
10	(UU)(GG)	(GG)(UU)			10
11	(AC)(CA)	(AC)(AC)	(CA)(CA)	(CA)(AC)	10
12	(GA)(GA)	(GA)(AG)	(AG)(GA)	(AG)(AG)	10
13	(GU)(GU)	(GU)(UG)	(UG)(UG)	(UG)(GU)	10
14	(GA)(GG	GG(GA)	GG(AG)	(AG)GG	11
15	(CA)AA	(AC)AA	AA(CA)	AA(AC)	9
16	UU(UC)	UU(CU)	(UC)UU	(CU)UU	9
17	(AG)AA	AA(GA)	AA(AG)	(GA)AA	9
18	(AA)(GG)	(GG)(AA)			10
19	(CU)(CU)	(CU)(UC)	(UC)(UC)	(UC)(CU)	10
20	(UU)(CC)	(CC)(UU)			

Table 2: QED twenty independent encoding codons, its synonymous and HB.

Number *	Noncoding	Noncoding Synonymous			H.B.	Cis- correlation
1	(TA)(TA)	(TA)(AT)	(AT)(TA)	(AT)(AT)	8	TATA -Trans. Start
2	(CG)(CG)	(CG)(GC)	(GC)(CG)	(GC)(GC)	12	(CG)(CG)-Intron
3	(AU)GG	GG(AU)	GG(UA)	(UA)GG	10	(AU)GG- START
5	(UG)(AG)	(GU)(AG)	(UG)(GA)	(GU)(AG)	10	(UG)(AG)-STOP
8	(UA)(GA)	(AG)UA)	(UA)(AG)	(GA)(AU)	9	(UA)(GA)-STOP

10	(UA)AA	AA(UA)	(AU)AA	AA(AU)	8	(UA)AA- STOP
6	(UG)AA	AA(UG)	(GU)AA	AA(GU)	9	AA (UG) Promotor
7	(TA)(GT)	(GT)(TA)	(TA)(TG)	(GT)(AT)	9	(TA)(TG) Promotor
11	(TA)(AC)	(AC)(TA)	(TA)(CA)	(AC)(AT)	9	CAAT Box, Promotor
15	TT(AC)	(AC)TT	(CA)TT	TT(CA)	9	TT(AC), Promotor
16	TT(AG)	(GA)TT	(AG)TT	TT(GA)	9	TT(AG), Promotor
22	AA(CT)	(CT)AA	(TC)AA	AA(TC)	9	AA(TC), Promotor
30	(CT)(TA)	(TC)(TA)	(CT)(AT)	(TC)(AT)	9	(TA)(CT), Promotor
12	(TT)(AA)	(AA)(TT)			8	TT AA, Promotor
14	TT(TA)	(TA)TT	(AT)TT	TT(AT)	8	TT(TA), Promotor
4	(UG)(AC)	(AC)(UG)	(UG)(CA)	(AC)(GU)	10	(UG)(CA), Promotor
9	(UA)(GC)	(UA)(CG)	(CG)(UA)	(CG)(AU)	10	(GC)(AU), Promotor
17	TT(CG)	(CG)TT	TT(GC)	(GC)TT	10	TT(CG), Promotor
18	CC(TA)	(TA)CC	(AT)CC	CC(AT)	10	(CC)(AT), Promotor
23	AA(CG)	(GC)AA	(CG)AA	AA(GC)	10	AA(CG), Promotor
28	(AC)(AG)	(AC)(GA)	(CA)(GA)	(CA)(AG)	10	(AC)(AG),Promotor
32	(CT)(AC)	(TC)(AC)	(CT)(CA)	(TC)(CA)	10	(TC)(AC), Promotor
33	(CT)(AG)	(TC)(AG)	(CT)(GA)	(TC)(GA)	10	(CT)(AG), Promotor
34	(CT)(TG)	(TC)(TG)	(CT)(GT)	(TC)(GT)	10	(CT)(AG),Promotor
19	CC(TG)	(TG)CC	(GT)CC	CC(GT)	11	CC(TG), Promotor
20	CC(AG)	(AG)CC	(GA)CC	CC(GA)	11	CC(AG), Promotor
24	GG(CT)	(CT)GG	(TC)GG	GG(TC)	11	GG(CT), Promotor
26	GG(AC)	(AC)GG	(CA)GG	GG(CA)	11	GG(AC), Promotor
27	(AC)(CG)	(CA)(CG)	(CA)(GC)	(AC)(GC)	11	(AC)(CG), Promotor
29	(AG)(CG)	(GA)(CG)	(AG)(GC)	(GA)(GC)	11	(AG)(CG), Promotor
31	(CT)(CG)	(TC)(CG)	(CT)(GC)	(TC)(GC)	11	(CG)(TC), Promotor
35	(GT)(CG)	(TG)(CG)	(GT)(GC)	(TG)(GC)	11	(CG)(TG), Promotor
13	(CC)(GG)	(GG)(CC)			12	(CC)(GG), Promotor
21	CC(CG)	(CG)CC	(GC)CC	CC(GC)	12	CC(CG), Promotor
25	GG(CG)	(CG)GG	(GC)GG	GG(GC)	12	GG(CG), Promotor

Table 3: The sequence number were rearranged while predicting their assignments.

QED encoding code assignment prediction

The QED protein-encoding codons encode proteins in both eukaryotes and prokaryotes as protein synthesis process is similar. The triplet coding encodes protein in prokaryote and has been verified and could be used as a guide to predict QED protein-encoding codon assignment

Nirenberg showed [13,14] that polyU, polyA and polyC encode the amino acids Phe, Lys and Pro, respectively. In the process, a direct link among mRNAs, tRNAs, amino acids, codons and anticodons was established in protein synthesis at ribosomes. Additionally, oligo chain lengths of 3 and 4: $(oU)_3$ and $(oU)_4$ showed nearly the same polypeptide formation. Therefore, it was reasonable to assume that if triplet UUU can encode Phe, quadruplet UUUU could

also encode Phe. Following this reasoning, LLLL-Lys and CCCC-Pro assignment prediction have been predicted. Since GGG encodes Gly, GGGG-Gly has also been assigned [1].

For the remaining sixteen amino acids, a two base matching procedure was used. Since only two bases could code only sixteen amino acids, a third base was added creating degeneracy and allowing the third base to form a dangling bond at the third base. Using the first two bases of verified triplet code, ignoring the degeneracy and third base wobble, was compared with the first two bases of QED code. If a match occurred, the corresponding twelve QED code assignment was predicted [1].

The remaining five: Ala (GCN), Asp (GA/U, C), lle (AUN), Met (AUG), and Tyr (UA/U, C) have (GCN) and (AUN), where N is any T(U), C, A and G correspond to noncoding QED code features. For these, additional condition to maintain the Hydrogen Bond needed for codon-anticodon pairing was imposed. The assignment predictions were done replacing (AU) by (AA) or (UU) to preserve quadruplet H.B. Similarly, (CG) replaced by (CC) or (GG) will preserve sextuplet H.B. The assignment predictions were made maintaining H.B. [3] and listed in Table 1.

QED noncoding code assignment prediction

The thirty-five noncoding QED codons are anticipated to control the transcription and splicing in eukaryote protein synthesis. TATA box with other *cis*-regulatory elements is considered the initiation of the eukaryote transcription process. A correlation between noncoding QED codon and cis- regulatory elements has recently been established [2], shown in Table 4, will facilitate in predicting the assignment. Based on the correlation result, the noncoding QED codon assignment prediction is listed in Table 3 and corresponding row number in Table 4.

The QED noncoding START and STOP codons are similar to the triplet three noncoding control START and three STOPs codons. Using verified triplet START and STOP codons, QED noncoding START and STOP assignment has been predicted and listed row# 3,5,8 and 10 in Table 3.

Since the QED eukaryote code was developed to find cures for incurable rare diseases, the role of QED was illustrated for three incurable rare diseases- Harlequin Ichthyosis (HI), Cystic Fibrosis (CFTR), Sickle cell disease, Cancer, and neurodegenerative diseases [3].

Cis-regulatory	Noncoding QED code	Table 3 row #
TATA Box	(TA)(TA)	1
CAAT Box	(CA)(TA)	11
CG/GC	(CG)(CG)	2
YCAY	(TC)(AT)	30
(Y-T (U)Or C)	CC(AT)	18
	(TC)(AC)	32
UAGG	(UA)GG	3
UGCAUG	(GC)(AU)	9
UGCAUG	(UG)(CA)	4
AT-Rich	AT-Rich	7,14
GC-Rich	CG or GC- Rich	17,21,23,25
		27,29,31,35

Table 4: Correlation between *cis*-regulatory elements and noncoding QED code bases.

Rare diseases

More than 7,000 rare diseases listed on the NIH site are autosomal, single-gene mutations occurring in eukaryotic cells. With all the resources available, NIH failed to develop a eukaryote genetic code to find a cure, and was left with no option but to manage the symptoms.

HI, CFTR, and Sickle Cell are incurable monogenic rare diseases.

The approach does not exhaust the scientific discussion of the diseases, but instead points to the role of QED coding in overcoming the triplet genetic coding hurdle in finding a cure. Three examples are selected to illustrate the concept. HI and CFTR are related to epithelial cells, while sickle is related to blood cells. The genetic mutations that substitute incorrect amino acids create dysfunctional proteins and lead to diseases.

HI - More than sixty incurable Ichthyosis skin diseases are listed on the Foundation for Ichthyosis and Related Skin Type

website [15], and its status [16]. HI is an autosomal recessive, life-threatening, severe skin disease. The normal functioning gene ABCA12 [17-19] secretes lipids from the Cell. When a mutation occurs, protein ABCA12 fails to secrete the lipids that accumulate in the epidermis, causing the disease. Gene testing provides the mutation and corresponding changes in the amino acids.

HI gene testing reveals [3], (Table 2) that the DupAA mutation causes Thr to be replaced by Arg, and another mutation, G, is replaced by A, resulting in Cys being replaced by Tyr.

Triplet coding

Four CAN, (N any 4 base) encode Thr. Sextuplet CGN, AGA, and AGG encode Arg.

QED coding

(AC)(CA) encode Thr, and (GA)(GA) encode Arg.

To correct the dysfunctional protein, Arg has to be replaced by the correct Thr. However, in triplet coding, each one of them is encoded by multiple codons. Selecting the proper codon out of a multicodon has been a hurdle that has prevented the discovery of cures for rare diseases.

The QED encoding of protein codons is nondegenerate and easier to select unambiguously to correct dysfunctional proteins.

Cystic fibrosis (CFTR)

CFTR is an autosomal recessive disease caused by a mutation [20] the ABCC7 gene [21], which fails to secrete the chloride ion (Cl^-) from the Cell, causing life-threatening accumulation of mucus in the epithelial cells of the lung. Mutations at codons 507 (ATC) and 508 (TTT) cause the disease. At 507, C is deleted, and the first two TTs at 508 leads to ATT (lie). Thus, Phe is lost.

Triplet coding

Normal, AUU, AUC, and AUA encode Ile, UUU and UUC encode Phe.

Mutation, only AUU is available to encode IIe and Phe is lost.

Since both UUU and UUC encode phenylalanine, which one should be select to regain it? A hurdle in correcting dysfunctional protein.

QED coding

Nondegenerate and independent UUUU encodes Phe, no ambiguity in selecting the code to correct a dysfunctional protein.

No cure has been found for CFTR so far. However, the appearance of molecular targeted therapies [22] is a ray of hope for its cure.

Sickle Cell

Sickle cell disease (SCD) is an autosomal recessive blood disease due to a monogenic mutation in the oxygen-carrying Hemoglobin protein [23]. Hemoglobin has four subglobin proteins. Two globin proteins at birth (HbF) produce normal blood and turn it off when adults. The other two adult globins (HbS) produce sickle-shaped blood cells, which are the source of the disease. A single pair mutation from G A G (glutamine) to G T G (Val) replaces Glutamine with Val, resulting in a dysfunctional protein causing the disease.

Triplet coding

Normal HbS, GAG encoding Glu, Mutated HbS, where GTG replaces GAG, encodes Val, resulting in a dysfunctional protein. Glu is encoded by GAA and GAG, Quadruplet GUN encodes Val, (N any four bases), six codons.

To replace Val, which quadruplet encoding should be used? Not an easy selection.

Quadruplet coding

Nondegenerate, independent (GU)(GU) encode Val, an easy selection to correct dysfunctional protein.

So far, no cure exists. Gene Therapy, applying molecularly targeted approaches to suppress HbS in the presence of HbF [24-27], has been reported to achieve normal blood cells. No side effects have been reported and are still under observation. The cost of treatment may be high and is unknown.

Cancer

Cancer is usually regarded as a multigenic disease. No cure exists, and extending life by five years has been the goal. After detection, the usual clinical steps are surgery, radiation, and chemotherapy. However, once metastasized, no successful cure exists. The biggest hurdle is that no biological technique exists to deliver chemotherapy directly to the cancerous Cell, as GPS is used

to reach a destination. Since its inception in 1970 and spending more than $ 200 billion, why has the NCI not yet developed a technique to deliver chemotherapy drugs to cancerous Cells to cure cancers selectively? Researchers working at the NCI had the resources, time, and expertise to collect data and develop alternative methods, but not a direct one for curing cancers.

Neurodegenerative diseases - Alzheimer's, Huntington's, ALS, and Parkinson's

Alzheimer's, Huntington's, ALS, and Parkinson's disease are among the most well-known. Similar to rare and cancer diseases, there is no cure. The diagnosis is primarily based on pathophysiological observations, and the gene-disease causality is not well understood. The biomarkers and diagnostic techniques are not universally accepted.

Alzheimer's disease

In 1906, Dr. Alois Alzheimer gave a talk to the 37th Congress of Psychiatrists of Southern Germany about his patient who had an unusual disease of the cerebral cortex, including memory loss, disorientation, and hallucinations. Later, the disease was named after him.

Alzheimer's disease hypothesis

Three hypotheses have been proposed for the gene-disease causality: Amyloid beta (Aß) [28]; Tau [29]; and Lysosomal dysfunction [30] hypothesis. The Aß and Tau hypotheses have been followed for the last twenty years. The lysosomal dysfunction causing the disease was recently introduced. Thus, it is essential to establish a clear gene-disease causality relationship before a cure can be found.

The amyloid hypothesis states that Alzheimer's disease (AD) is primarily caused by the deposition of Aß plaques in brain tissue. Amyloid precursor protein (APP) is normally cleaved by protease-β and γ-secretases, yielding Aβ. Aß proteins consist of multiple isoforms with 37-43 amino acid residues [31]. The Aβ 1-42 (Aβ42) and Aβ 1-40 (Aβ40) are the dominant isoforms, with Aβ42 having two additional amino acids Aβ42 (41- IIe, 42-Ala).

Aβ42/ Aβ40 Ratio test

Aβ42 tends to clump into plaques in the brain, causing its levels in blood or CSF to drop relative to Aβ40. A reduced Aβ42/40 ratio is therefore considered a strong marker of amyloid plaque accumulation, an early Alzheimer's pathology, often before cognitive symptoms are noticeable. To detect abnormal brain Aβ status in patients with early AD, the performance of plasma Aβ42/40 is measured [32,33]. Plasma β-amyloid 42/40 ratio of Alzheimer's Disease Risk: Lower Risk: > or = 0.170; Intermediate Risk: 0.150 - 0.169; and Higher Risk: <0.150.

Apolipoprotein E (apoE) modulates the aggregation, clearance, and toxicity of Aβ. The apoE gene encodes three isoform proteins, E2, E3, and E4, which differ in amino acids at two positions: 112 and 158 (Table 5).

Position	112	136	158
E2	Cys	Arg	Cys
E3	Cys	Arg	Arg
E4	Arg	Arg	Arg
R136SE3	Cys	Ser	Arg
R136SE4	Cys	Ser	Arg

Table 5: apoE isoform (E2, E3 and E4).

The E3 is the most common form, while E2 is neuroprotective, and E4 increases AD risk. E4 homozygotes (E4/E4) are 15-fold more susceptible to AD than E3 carriers [34], and disease begins several years earlier in E4 carriers than in those with E3 or E2.

RS136S mutations in E3 [35] and E4 [36] at position 136, replacing Arg with Ser, have shown promising results in controlling ADs.

Tau-p hypothesis (29)

In human brains, tau pathology appears about a decade before the formation of Aß plaques, especially targeting glutamate projection neurons in the association cortex. The tau proteins (tubulin-associated unit) form a group of six highly soluble protein isoforms_produced by alternative splicing_from the gene MAPT (microtubule-associated protein_tau) and maintain microtubule stability in the axons of neurons of the central nervous system (CNS). The hyperphosphorylated Tau (p-Tau) dysfunctional protein disrupts the microtubule, leading to the nonfunctioning of the axon and loss of memory. Additionally, Insoluble neurofibrillary

tangles are formed in hyperphosphorylated Tau (p-tau), which increases Aβ production. The increased A*β* production promotes hyperphosphorylation of Tau, driving vicious cycles leading to AD pathology over a long lifespan and dementia.

Plasma P-tau217 immunoassay accurately identifies biological ADs, comparable with cerebrospinal fluid (CSF) Aβ42/40 biomarkers [37-39]. The Alzheimer's Association Global Biomarker Standardization Consortium has recently reported on phosphorylated Tau (p-tau) assays that provide the largest fold-changes in Alzheimer's disease (AD) [40].

FDA-approved blood test

On May 16, 2025, the FDA approved the blood test Lumipulse G pTau217/ß-Amyloid 1-42 Plasma Ratio for the early detection of Alzheimer's disease. The Lumipulse G pTau217/ß-Amyloid 1-42 Plasma Ratio measures two proteins, pTau217 and β-amyloid 1-42, found in human plasma, a component of blood. It calculates the numerical ratio of the two protein levels. This ratio is correlated to the presence or absence of amyloid plaques in the patient's brain, reducing the need for a PET scan.

Lysosomal dysfunctional hypothesis [30]

Lysosomes are membrane-bound organelles in the cytosol with digestive enzymes for breaking down proteins, polysaccharides, and lipids at a balanced acidic ph. The dysfunctional lysosome-A*β* protein interaction leads to the early onset of all forms of Alzheimer's disease.

QED eukaryote dysfunctional protein-disease hypothesis

The dysfunctional proteins are the cause of human diseases, including but not limited to neurodegenerative diseases. Human eukaryotic cell maintains homeostasis through functional proteins—mutations in a gene result in a dysfunctional protein, leading to disease.

Since Alzheimer's disease typically sets in after the age of 50, the possibility of cell duplication error exists. The APP is cleaved by protease-β and γ-secretases to produce A*β,* leading to the accumulation of Aβ plaques. Out of the two Aβ isoforms, Aβ40 and Aβ42, Aβ42 has two additional amino acids and accumulates more plaques than Aβ40. The different functions may be related to the number of amino acids, resulting in distinct protein folding and their corresponding functions.

The other source of dysfunctional protein error in AD is errors in the apoE protein, with isoforms E2, E3, and E4 influencing the metabolism of Aβ. Even though E2, E3, and E4 differ only by amino acids at two positions, 112 and 158 (Table 5), E4 accelerates ADs more than E2. Since modifying the position at 136 in E3 and E4 by replacing Arg with Ser (R136S), shows a positive Ads result, is a clear indication of protein folding and its functions.

Tau protein maintains the stability and functioning of axons in the CNS. The hyperphosphorylated Tau (p-Tau) is a dysfunctional protein causing the axon destruction and degradation of the memory function.

Dysfunctional proteins occur due to tandem repeats (TR) in eukaryotic cells, causing several neurodegenerative diseases [41] that are located in different regions of the gene, including the 5'-UTR, exon, intron, and 3'-UTR. TR in the exon area may synthesize proteins, but the intron is a noncoding area. More than standard TR in the upstream of the gene will introduce errors in the start of the transcription, and at the exon/intron interfaces, causing splicing errors. In either case, dysfunctional protein will result causing the disease.

Huntington's, ALS, and Parkinson's

Tandem repeat (TR)

TR in eukaryote genes has been identified as the leading cause for Huntington's, ALS, and Parkinson's diseases using triplet coding without the eukaryote genetic code. Furthermore, gene distribution in eukaryotes is discrete, with <2% encoding (exon) parts separated by > 98% noncoding (intron). Thus, only 2% proteins produce functional proteins to maintain the well-being of the cell. Due to a gene mutation in the encoding region or an error at the exon-intron interface, dysfunctional proteins are synthesized, causing the disease.

The approach here is not a scientific review of the disease, but rather a study of the role of the QED code in resolving the occurrence of tandem repeat (TR) in Huntington's and other diseases. Most TRs are noncoding QED codons, which could aid in identifying the TR-disease relationship.

TR (CAG)n and (GCX)n occur in exon regions. In triplet coding, CAG encodes glutamine but causes HD and several other diseases. GCX encodes Alanine, which causes neurological, muscular, and other developmental diseases.

HD - Gln is encoded by CAA and CAG. In QED code, CAA only encodes Gln, and CAG is noncoding. CAA encoding yields polyglutamine but not by CAG. When TR <30, no HD is observed, but it is active when TR >36 or higher.

ALS - Ala is encoded by GCX (X, any T, C, A, and G). Under QED, GCX is noncoding.

The QED noncoding codons CGX, GCX, CXG, XCG, and CGXX are synonymous and include most TR. For example, (CGG)n in 5'-UTR, (GGGGCCC)n in an intron, and (CTG)n in 3'-UTR are all noncoding under QED code.

Acknowledgments

Over a decade of pioneering research, I was encouraged by some experts and others who criticized the work, which provided the opportunity to improve it. I am grateful to Dr**.** Nawin Mishra, Distinguished Professor Emeritus of Genetics and Genomics at The University of South Carolina, Columbia, SC 29208, for his encouragement and valuable comments regarding this work. I am thankful to Dr. Rakesh Dixit, Tiburon, CA, and Dr. Naresh Bhatt, Mountain Dora, FL, for the medical discussions. My daughter, Usha Singh's rare, incurable disease, was the catalyst to start the lifelong journey to find a gene-disease causal relationship. The outcome is the development of QED genetic code for eukaryotic cells and its correlation with cis-regulatory elements. This work would have never materialized without the unwavering support of my family: my grandchildren, Sonya Khan and Omar Khan, for their invaluable assistance in assembling the DNA mechanical model; my wife, Amrita Singh, and children, Asha Singh and Om Prakash Singh (also known as Tom Singh).

Author Contribution

Rama Shankar Singh - 100%.

Competing Interests

The authors have no competing interests.

Additional Information

Correspondence and requests for materials should be addressed

To: Rama Shankar Singh, email: gen-med @comcast.net

Bibliography

1. Rama Shankar Singh. "Quadruplet Expanded DNA (QED) Genetic Code for Eukaryotic Cells". *Acta Scientific Medical Sciences* 7.12 (2023): 70-82.

2. Rama Shankar Singh. "Correlation between Eukaryotic Noncoding QED Genetic Codes and Cis-Regulatory Elements". *Acta Scientific Medical Sciences* 8.10 (2024): 89-96.

3. Rama Shankar Singh. "QED Eukaryote Genetic Code and Principle of Information Flow and Biological Protein Synthesis". *Acta Scientific Medical Sciences* 9.6 (2025): 100-114.

4. Crick FH. "On the genetic code". *Science* 139 (1963): 461-464.

5. Liu CC and Schultz PG. "Adding new chemistries to the genetic code". *Annual Review of Biochemistry* 79 (2010): 413-444.

6. Malyshev DA., *et al*. "Efficient and sequence-independent replication of DNA containing a third base pair establishes a functional six-letter genetic alphabet". *Proceedings of the National Academy of Sciences of the United States of America*. (2012): 12005-12010.

7. Hoshika S., *et al*. "Hachimoji DNA and RNA: a genetic system with eight building blocks". *Science* 363 (2019): 884-887.

8. Kenneth S Kosik. "A Biologist talks to a Physicist". *Nature* 557 (2020): 281.

9. Sarah Bohndiek. "A Physicist talks to a Biologist". *Nature* 557 (2020): 283.

10. Claude E Shannon. "A Mathematical Theory of Communication". *The System Technical Journal* 27 (1948): 379-423, 623-656.

11. 1968 The Nobel Prize in Physiology or Medicine was awarded jointly to Robert W. Holley, Har Gobind Khorana and Marshall W. Nirenberg "for their interpretation of the genetic code and its function in protein synthesis".

12. AR Morgan., *et al*. "Studies on polynucleotides, LIX. Further codon assignments from amino Acid incorporations directed by ribopolynucleotides containing repeating trinucleotide sequences". *Proceedings of the National Academy of Sciences of the United States of America* 56 (1966): 1899-1906.

13. Nirenberg M and Leder P. "RNA codewords and protein synthesis. The effect of trinucleotides upon the binding of sRNA to ribosomes". *Science* 145 (1964): 1399-1407.

14. Jones O W and Nirenberg M W. "Qualitative survey of RNA codewords". *Proceedings of the National Academy of Sciences of the United States of America* 48 (1962): 2115-2123.

15. FIRST - Foundation for Ichthyosis & Related Skin Types, Inc.

16. Joosten MDW., *et al*. "New development in the molecular treatment of Ichthyosis_ Reve. of literature". *Orphanet Journal of Rare Diseases* (2022)17-269

17. Kelsell David P., *et al*. "Mutations in ABCA12 Underlie the Severe Congenital Skin Disease HI". *American Journal of Human Genetics* 76 (2005): 794-803.

18. Anna C Thomas., *et al*. "ABCA12 Is the Major Harlequin Ichthyosis Gene". J*ournal of Investigative Dermatology* 126 (2006): 2408-2413.

19. Akiyama Masashi., *et al*. "Mutations in lipid transporter ABCA12 in harlequin ichthyosis and functional recovery by corrective gene transfer". *Journal of Clinical Investigation* 15 (2005): 1777-1784.

20. Hunt John F., *et al*. "Cystic Fibrosis Transmembrane Conductance Regulator (ABCC7) Structure". *Cold Spring Harbor Perspectives in Medicine* 3 (2013): a009514.

21. Ratjen Felix., *et al*. "Cystic fibrosis". *Nature Reviews Disease Primers* 1 (2015): 15010.

22. Hartmut Grasemann and Felix Ratjen. "Cystic Fibrosis". *NEJM* 389 (2023): 1693 .

23. Inusa Baba PD., *et al*. "Sickle Cell Disease—Genetics, Pathophysiology, Clinical Presentation and Treatment". *International Journal of Neonatal Screening* 5 (2019): 20.

24. George Q Daley. "Welcoming the Era of Gene Editing in Medicine". *The New England Journal of Medicine* (2024).

25. Frangoul H., *et al*. "CRISPR-Cas9 Gene Editing for Sickle Cell disease and beta-Thalassemia". *The New England Journal of Medicine* 384 (2021): 252-260.

26. Doerfle Phillip A., *et al*. "Genetic therapies for the first molecular disease". *Journal of Clinical Investigation* 131.8 (2021): e146394.

27. Sharma A., *et al*. "CRISPR-Cas9 Editing of the HBG1 and HBG2 Promoters to Treat Sickle Cell Disease". *NEJM* 389 (2019): 820-832.

28. John Hardy and Dennis J Selkoe. "The Amyloid Hypothesis of Alzheimer's Disease_ progress and problems on the road to therapeutics". *Science* 297 (2021): 353-356.

29. Arnsten Amy FT., *et al*. "Hypothesis: Tau pathology is an initiating factor in sporadic Alzheimer's disease". *Alzheimer 's Dementia* 17 (2021): 115-124.

30. Im., *et al*. "Lysosomal dysfunction in Down syndrome and Alzheimer mouse models is caused by v- ATPase inhibition by Tyr682-phosphorylated APP βCTF". *Science Advances* 9 (2023): eadg1925.

31. Dangjoon Im and Tae Su Choi. "Distinctive contribution of two additional residues in protein aggregation of Aβ42 and Aβ40 isoforms". *BMB Report* 57 (2024): 263-272.

32. Lu Yifei., *et al*. "Changes in AD Blood Biomarkers and associations with Incident All-cause Dementia". *JAMA* (2024).

33. Janelidze Shorena., *et al*. "Head-to-Head Comparison of 8 Plasma Amyloid-β42/40 Assays in Alzheimer Disease". *JAMA Neurology* 78 (2021): 1375-1382.

34. Xia Zengjie., *et al*. "Co-aggregation with Apolipoprotein E modulates the function of Amyloid-β in AD's disease". *Nature Communication* 15 (2024): 4695.

35. Naguib Sarah., *et al*. "The R136S mutation in the APOE3 gene confers resilience against tau pathology via inhibition of the cGAS-STING-IFN pathway". *Immunity* (2025).

36. Nelson Maxine R., *et al*. "The APOE-R136S mutation protects against APOE4-driven Tau pathology, neurodegeneration and neuroinflammation". *Nature Neuroscience* 26 (2023): 2104-2121.

37. Thijssen Elisabeth H., *et al*. "Association of Plasma P-tau217 and P-tau181 with clinical phenotype, neuropathology and image markers in Alzheimer's disease and frontotemporal lobar degeneration: a retrospective". *Lancet Neurology* 20.9 (2021): 739-752.

38. Ashton Nicholas J., *et al*. "Diagnostic Accuracy of a Plasma Phosphorylated Tau217 Immunoassay for Alzheimer Disease Pathology". *NEJM* 81.3 (2024): 255-253.

39. Leuzy Antoine., *et al*. "Considerations in the clinical use of amyloid PET and CSF biomarkers for Alzheimer's disease". *Alzheimer 's Dementia* 21 (2025): e14528.

40. Ashton Nicholas J., *et al*. "The Alzheimer's Association Global Biomarker Standardization Consortium (GBSC) plasma phospho-tau Round Robin study". *Alzheimer's Dementia* 21 (2025): e14508.

41. Anthony J Hannan. "Tandem repeats mediating genetic plasticity in health and disease". *Nature Reviews Genetics* 19 (2018): 286-298.