

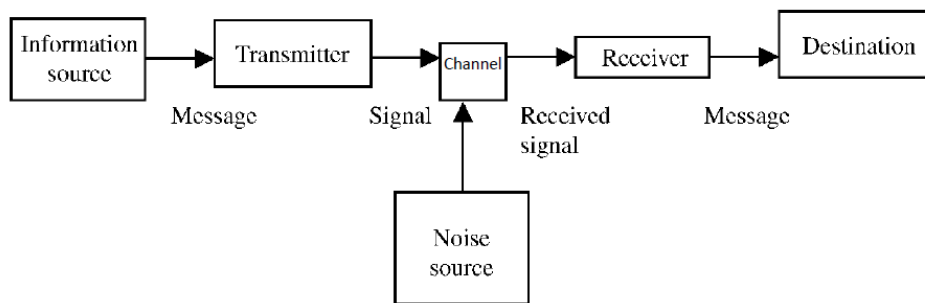
**ISTI Researcher's work brings new evidence for a primordial genetic code composed of two nucleotides:**

**Is the Genetic Code Optimal?**

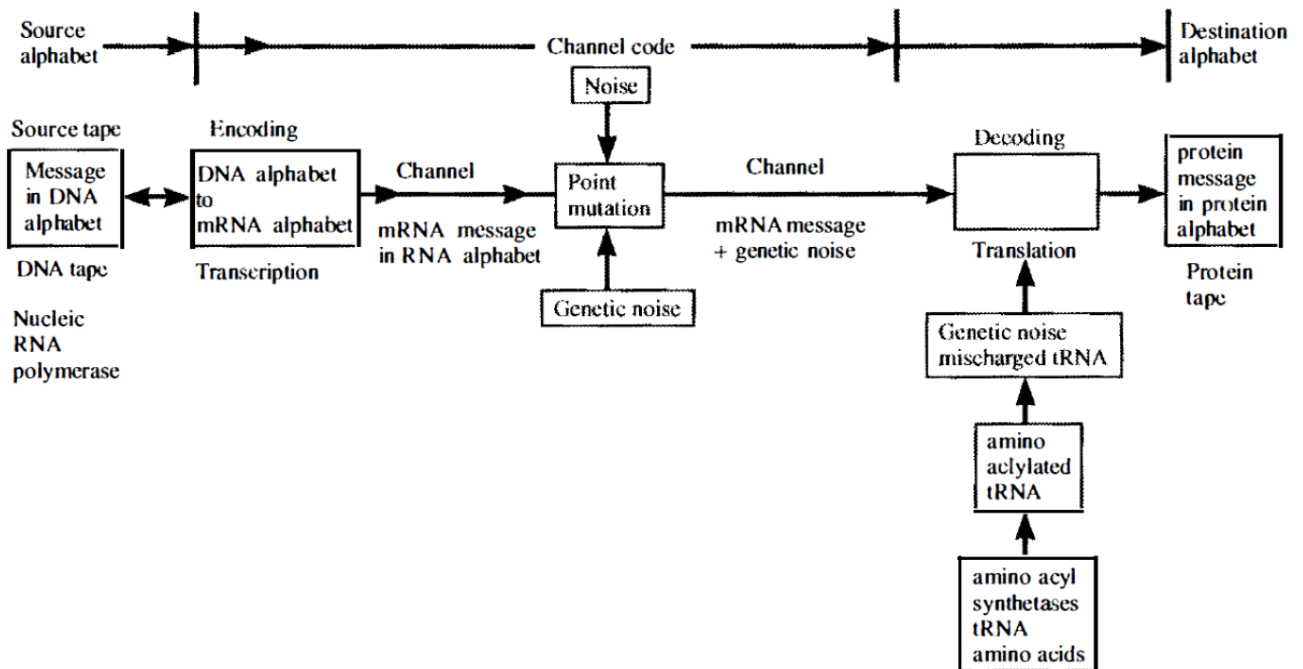
Ercan E Kuruoglu (Senior Researcher) of ISTI who has been a visiting scientist in the Max Planck Institute as an Alexander von Humboldt Experienced Research fellow and Peter Arndt of Max Planck Institute of Molecular Genetics in Berlin propose a new way of looking at genomic processes: they envision biological processes in the cell as information transfer processes and propose a mathematical framework for studying these processes using telecommunication theory due to Claude Shannon.

This vision leads naturally to looking at the basic genomic processes, such as translation, transcription and replication as a communication system. Visualising the protein synthesis process in this telecommunication analogy, the RNA corresponds to the encoder/encoded message, the mutations to the noisy channel, the ribosome to the decoder and aminoacid sequence making up the protein the decoded message. Kuruoglu and Arndt calculated for this channel, the information capacity hence providing us with the limits of reliable information transfer. Simple, it may seem, this calculation allows one to visualise the biological evolution in the framework of preservation/change in the information transferred to the new generations.

*Telecommunication system:*



*Genomic system:*



Kuruoglu and Arndt posed the question of whether the existing genetic code is optimal from the point of view of information capacity. They constructed artificial genetic codes which has higher information capacity hence demonstrating the sub-optimality of the natural genetic code. They designed, then, an intelligent machine learning algorithm to search for the optional genetic code. The optimal code they have discovered has significantly better information capacity and has redundancy in all three nucleotide positions in the codon providing error-correction coding potential to the code. The natural code on the other hand has redundancies almost exclusively at the third place. This provides a strong evidence to the hypothesis that the natural genetic code co-evolved from a primordial genetic code composed of 2-nucleotide codons.

*Near Optimal Genetic Code:*

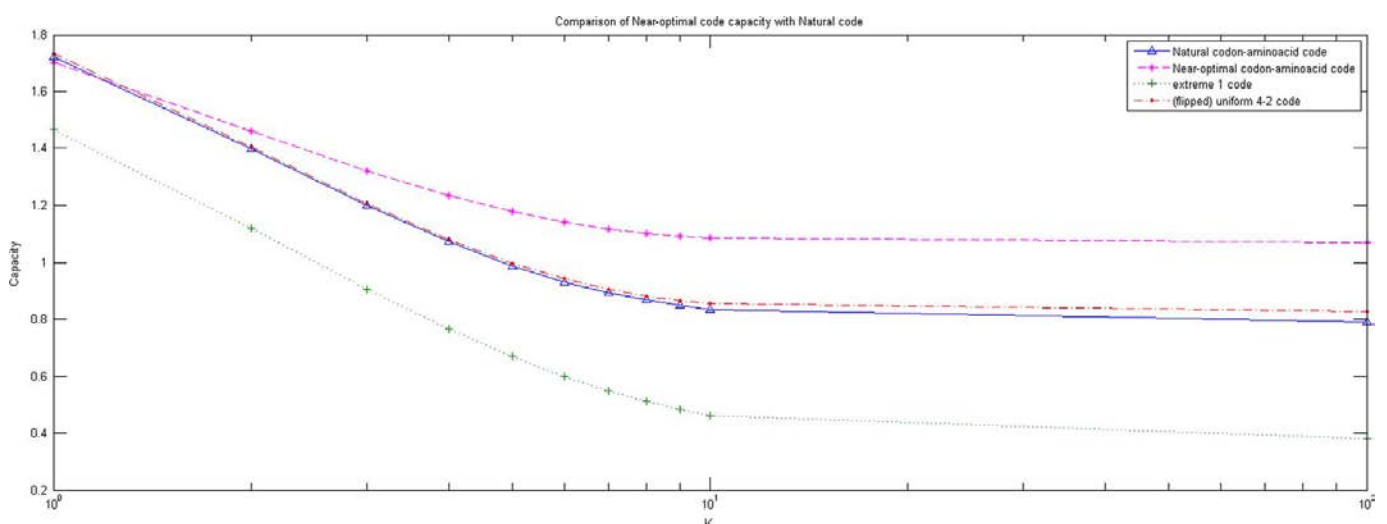
2	1	T	C	A	G	3
T		9	6	8	8	T
		9	6	15	15	C
		17	21	4	1	A
		17	21	4	1	G
C		9	6	8	8	T
		9	6	15	15	C
		17	21	4	1	A
		17	21	4	1	G
A		2	2	19	19	T
		14	14	13	13	C
		7	7	18	16	A
		7	7	18	16	G
G		10	10	5	5	T
		3	3	5	5	C
		11	20	12	12	A
		11	20	12	12	G

*Natural Genetic Code:*

2	1	T	C	A	G	3
T		14	11	10	20	T
		14	11	10	20	C
		11	11	13	20	A
		11	11	13	20	G
C		15	15	17	1	T
		15	15	17	1	C
		15	15	17	1	A
		15	15	17	1	G
A		19	9	3	4	T
		19	9	3	4	C
		21	6	12	7	A
		21	6	12	7	G
G		5	2	16	8	T
		5	2	16	8	C
		18	2	2	8	A
		18	2	2	8	G

Fig. 5. The natural genetic code (codon to amino acid map). 1: Alanine, 2: Arginine, 3: Asparagine, 4: Aspartate, 5: Cysteine, 6: Glutamate, 7: Glutamine, 8: Glycine, 9: Histidine, 10: Isoleucine, 11: Leucine, 12: Lysine, 13: Methionine, 14: Phenylalanine, 15: Proline, 16: Serine, 17: Threonine, 18: Tryptophan, 19: Tyrosine, 20: Valine, 21: STOP. We indicated the amino acids with numbers in the table to emphasize the fact that names are only labeling and should not affect our search for optimal codes in the sequel.

*Channel Capacity:*



Kuruoglu and Arndt now study the cancer dynamics again in the framework of information transfer using the framework provided by Shannon theory.

Kuruoglu is invited to give a plenary talk on this research at the conference Entropy 2018: From Physics to Information Sciences and Geometry, in Barcelona, Spain. <https://sciforum.net/conference/Entropy2018-1> . He has already given a plenary talk on the topic in IEEE SIU 2017 in Antalya, Turkey.