From RNA to the Genetic Code and Back

Tidjani Négadi[1*]

Abstract

In this paper, which is an extension of a very recent one, we show, in a first part, that starting from RNA, more exactly, from the detailed atomic composition of its basic components, the four ribonucleotides UMP, CMP, AMP and GMP, full of Fibonacci numbers, we can derive the "One-Number Model" of the genetic code which was introduced by us some years ago, in 2007. In a second new and original part, we consider the opposite way, that is, starting from the "One-Number Model", we can recover, in the detail, the atomic composition of the four ribonucleotides, from which we started, and the "circle is complete".

Key Words: RNA, Fibonacci Numbers, Genetic Code, One-Number Model.

DOI Number: doi: 10.14704/nq.2020.18.1.NQ19103 **NeuroQuantology 2020; 18(1):01-07**

Introduction

It is an interesting fact that the present situation in fundamental biology resembles the one in fundamental physics. In the latter, only about 4% of all we know concerns visible matter (atoms, molecules, planets, stars, galaxies, etc., and ourselves) and the rest, that is, 96%, is attributed to what is now known as dark matter and dark energy. In biology, considering for example the human genome, less than 2-3% is comprised of coding sequences, that is sequences coding for proteins and the great remaining part (>97-98%), what was earlier called "junk DNA" and today more appropriately replaced by "dark matter of biology", is almost entirely comprised of what is now called non-coding RNAs. The RNA molecule is an awesome object. Recent research has shown that it has passive (information carrier or structural) as well as active (functional) roles in the innumerable biochemical processes in the living organisms. The number of kinds of RNA is growing fast. In addition to the well known three types mRNA (messenger RNA), tRNA (transfer RNA) and rRNA (ribosomal RNA), a countless number of RNAs, with an active role, have been discovered in the last decades.

Some few examples of them are: snoRNAs, miRNAs, siRNAs, shRNAs), tasiRNAs, rasiRNAs, eRNAs, etc., the list is still long and certainly not closed. In a recent paper (Chech and Steitz, 2014) more than fifty studied kinds of non-coding RNAs were listed.

The growing importance of RNA over DNA has led scientists to even challenge the time-honored Central Dogma of Molecular Biology, symbolized by the sequence DNA→RNA→proteins. For example, Brosius (Brosius, 2003a) writes "On a evolutionary level one might rather suggest a restatement as RNA→proteins→DNA". In another paper (Brosius, 2003b), he says that DNA and proteins, as *latecomers*, could have been subsequent "inventions" of RNA. Said otherwise, this would mean that the storage of genetic information and the enzymatic function were, at a certain time of the early history of life, devolved to DNA and proteins, respectively (Mercer et al. 2009). RNA is also of great importance in neurosciences. It seems therefore that the ribonucleic acid (RNA) is a major player, compared to DNA. Non-coding RNAs have been shown recently to play a major role in neurosciences. As an example, among others, it has been shown the gene Arc plays an essential role in the storage of information in the mammalian brain (Pastuzin et al., 2018).

Corresponding author: T. Négadi
Address: Physics Department, Faculty of Science, University of Oran1, Es-Sénia, Oran, Algeria.
[1*]E-mail: tnegadi@gmail.com
Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
Received: 16 August 2019 **Accepted:** 22 October 2019

All this has been nicely summarized by John Mattick (Mattick, 2009) in a seminar at the University of Sydney, Australia, in June 2009: "In conclusion, what was rejected as undesirable because it was not understood probably holds the key to understanding human development and human intelligence. RNA is not just a passive and ephemeral intermediate between the gene and the protein, but the computational engine of cell biology, development, brain, and possibly evolution itself." A natural question one would ask, in view of all what has been said above about the centrality of RNA and the wonderful things it is behind in the living world, is whether this awesome molecule has some other new kind of informational abilities, hidden in it, to be disclosed. We shall see below, in this paper, that, possibly, it is the case.

In this paper, which is an extension of a very recent one (Négadi, 2019), we show in a first part (sections 3 and 4), that starting from RNA, more exactly, from the detailed atomic composition of its basic components the four ribonucleotides UMP, CMP, AMP and GMP, full of Fibonacci numbers, we can derive, as an outcome, the "*one-number model*" of the genetic code which was introduced by us some years ago (Négadi, 2007). In a second part (section 5), which is the truly original part of this paper, we consider the *opposite way*, that is, starting from the "*one-number model*", we can recover, as an outcome, in the detail, the atomic composition of the four ribonucleotides from which we started (in section 4), and "*the circle is complete*"! The mathematical tools we shall use for our computations are only elementary and well known arithmetic functions; they will be defined in the appropriate places. Finally, the last section (section 6) is devoted to summarize this paper and also to derive some interesting and meaningful results, as by-products of, and related to, the ones of sections 3, 4 and 5. Let us recall, here, some basic facts about the standard genetic code but we shall say more below. The genetic code is the "dictionary" of life that translate, in the ribosome(s), the 64 RNA-codons into 20 amino acids. More exactly, 61 codons, called meaningful codons, are translated into 20 amino acids and 3 codons serve as stops or terminations signals. As there are only 20 amino acids coded, 41 codons (61-20) are said degenerate. Also, 18 amino acids are coded by more than one codon and only 2 amino acids are coded by one codon (Tryptophane and Methionine). This means that 59 codons (61-2) correspond to the 18 amino acids mentioned above.

The Fibonacci Numbers and Some of Their Properties

The famous Fibonacci sequence, so much represented in the natural world (biology, botany, physics and so on) is defined by the formula

$$F_n := F_{n-2}+F_{n-1}\ ,\ n=2,\ 3,\ 4,\ \ldots \qquad (1)$$

with the initial conditions, or "seeds", $F_0=0$, $F_1=1$. There exist also a sequence, known as the Lucas sequence, L_n, defined similarly but with the initial conditions $L_0=2$ and $L_1=1$. We give in Table 1, for the two sequences, the first few terms

Table 1. The Fibonacci and Lucas sequences

n	0	1	2	3	4	5	6	7	8	9	10	11	12
F_n	0	1	1	2	3	5	8	13	21	34	55	89	144
L_n	2	1	3	4	7	11	18	29	47	76	123	199	322

For later use, we include the following useful identities involving both sequences

$$F_n+2F_{n+1}=F_{n+3} \qquad (2)$$
$$F_n+F_{n+3}=2F_{n+2} \qquad (3)$$
$$L_n+5F_{n+2}=L_{n+4} \qquad (4)$$
$$L_{n-1}+L_{n+1}=5F_n \qquad (5)$$
$$F_{n+2}-F_{n-2}=L_n \qquad (5)'$$
$$L_n+F_{n-2}=F_{n+2} \qquad (6)$$
$$F_{n-1}+F_{n+1}=L_n \qquad (7)$$

02

The four RNA Ribonucleotides are Full of Fibonacci Numbers

Consider the four basic units of RNA, the ribonucleotides *uridine monophosphate* (UMP), *cytidine monophosphate* (CMP), *adenosine monophosphate* (AMP) and *guanosine monophosphate* (GMP). It is well known that RNA may have played a key role in the early evolution of life (Gilbert's RNA-World hypothesis, Gilbert, 1986). Moreover, being the "computational engine" in biology, as explained above, it should have, and plays, a major role. The four basic molecules, mentioned above, have the following molecular brut formulae for their atomic composition and, in parenthesis, those for the four nucleobases (uracil U, cytidine C, adenine A and guanine G)

$$\text{UMP: } C_9H_{13}N_2O_9P \ \ (\text{U: } C_4H_4N_2O_2)$$
$$\text{CMP: } C_9H_{14}N_3O_8P \ \ (\text{C: } C_4H_5N_3O_1)$$
$$\text{AMP: } C_{10}H_{14}N_5O_7P \ \ (\text{A: } C_5H_5N_5)$$
$$\text{GMP: } C_{10}H_{14}N_5O_8P \ \ (\text{G: } C_5H_5N_5O_1) \qquad (8)$$

In these formulae H_k, C_k, O_k, N_k and P_k refer to hydrogen, carbon (C), oxygen (O), nitrogen (N) and phosphorus (P) atoms and their numbers (as indices) indicated in the corresponding molecule.

In Table 2, below, we give the detailed distribution of the various atoms computed from Eqs.(8)

Table 2: The distribution of atoms in the ribonucleotides

Ribonucleotide(s)	Hydrogen	CNOP	total (HCNO)
UMP	13	21	34
CMP+AMP+GMP	42	68	110
All four ribonucleotides	55	89	144

First, we have for the total number of atoms in all four ribonucleotides 144, the twelfth Fibonacci number (see Table 1). This number, which is known to be, with the trivial unity 1, the only Fibonacci number with the property $F_i=i^2$ (with i=12), will play in the following a crucial role.This number splits between 55 hydrogen atoms and 89 CNOP atoms satisfying in this way the identity in Eq.(1) with n=12. Now uridine, *alone*, has 34 atoms, 13 hydrogen atoms and 21 CNOP atoms. (In the rest of this paper, we use uridine, in place of uridine monophosphate, for short.) These three last numbers are also Fibonacci numbers satisfying the identity in Eq.(1) with n=9. *This particular pattern shows itself only for uridine*. Note however that the *hydrogen* and CNOP atom numbers for uridine, on the one hand, and the remaining three ribonucleotides (CMP+AMP+GMP), on the other, satisfy the identity equation (2) with respectively n=7 and 8, that is 13+2×21=55 and 21+2×34=89. We see therefore that uridine, compared to the other three ribonucleotides, plays in some way a singular role. As an additional support for the oddity of uridine, we have that its number of hydrogen atoms, 13, splits into 5+8 (=13), also satisfying the identity (1), where 5 is the Hydrogen Donor Count (HDC) and 8 is the Hydrogen Acceptor Count (HAC)[1]. This hydrogen HAC/HDC-pattern does not exist for the other three ribonucleotides. Summarizing, we have *eight* Fibonacci numbers, 5, 8, 13, 21, 34, 55, 89 and 144, of which *five* concern uridine, only, and *three* concern all four ribonucleotides, in also a nice Fibonacci pattern (5+3=8). By the way, let us note that uridine plays a crucial role in the pyrimidine metabolism of the brain. It supplies nervous tissue with the pyrimidine ring and, in turn, participates in a number of important metabolic pathways. Uridine and its nucleotide derivatives may also have an additional role in the function of the central nervous system as signaling molecules[2].

[1]http://www.hmdb.ca/metabolites/HMDB0000288; https://www.drugbank.ca/drugs/DB03685
[2]http://proteom.elte.hu/_publications/PDFs/Arpi_uridin_Curr_Top_11.pdf

The One-Number Model of the Genetic Code, from RNA Fibonacci Numbers

In view of what has been said in the last section about the singular role played by uridine, let us consider the following partition of the total number of atoms in the four ribonucleotides 144=34+110 (see Table 2), where 34 and 110 are respectively the number of atoms in uridine and in the rest, that is in CMP, AMP and GMP. This partition satisfies also the identity (2): 34+2×55=144. For applications in the following, let us introduce the function A_0 of an integer n $A_0(n)=a_0(n)+SPI(n)+\Omega(n)$, where $a_0(n)$ is the sum of the prime factors of the prime factorization of n, including multiplicities, SPI(n) is the sum of their respective prime indices and Ω(n) is the number of these prime factors. As an example, consider the n=12. Its prime factorization, from the Fundamental Theorem of Arithmetic is $2^2\times3$ and we have $a_0(12)=7$, $SPI(12)=4$, $\Omega(12)=3$, so that $A_0(12)=7+4+3=14$. Now considering the above partition, we have the factorizations $34=2\times17$ and $110=2\times5\times11$. This leads to $a_0(34)=2+17=19$, $SPI(34)=1+7=8$, $\Omega(34)=2$, $a_0(110)=2+5+11=18$, $SPI(110)=1+3+5=9$, $\Omega(110)=3$ and therefore to 03

$$A_0(34)+A_0(110)=29+30=59 \quad (9)$$

This number, 59, is equal to the number of *meaningful* codons, see above in the introduction. Moreover, if we write the components of the two A_0-functions in Eq.(9) as follows

$$[A_0(34)+SPI(110)+\Omega(110)]+[a_0(110)]=41+18=59 \quad (10)$$

we have a first interesting result: 41 is the number of codons for the 18 amino acids coded by more than one codon. To reach the total number of codons, 64, including the two non-degenerate codons for the two singlets and the three stop codons, we have two simple possibilities for adding the number 5 as 2+3: either we take the length function λ which gives the number of digits of a number, as we have done in reference (Négadi, 2019), or take again the Ω-functions, defined above.

In either case, we have the same result $\lambda(34)=\Omega(34)=2$ and $\lambda(110)=\Omega(110)=3$. Let us choose one of them, λ for example, to get

$$[A_0(34)+ SPI(110)+ \Omega(110)]+[a_0(110)+ \lambda(34)]+ \lambda(110)=41+(18+2)+3=41+20+3=64 \quad (11)$$

This is the desired result: 20 amino acids, 18 amino acids coded by more than one codon and 2 "non-degenerate" amino acids (the two singlets), 41 degenerate codons and 3 stop codons. We shall return to the above two additions at the end of this section. Now, let us split Eq.(11) in two *sub-partitions*, as follows

$$[A_0(34)+ SPI(110)+ \Omega(110)]=41 \quad (12)$$

and

$$[a_0(110)+ \lambda(34)]+ \lambda(110)=(18+2)+3=20+3 \quad (13)$$

and consider only the first one, in Eq.(12). The functions in this equation, taken separately, are $a_0(34)=19$, $SPI(34)=8=1+7$, $\Omega(34)=2$, $SPI(110)=9$ and $\Omega(110)=3=1+1+1$. Before going further, let us mention that $A_0(34)=29$ which is the seventh Lucas number (see Table 1). Also, using the additive property of the A_0-function, we have $A_0(34)=A_0(2)+A_0(17)=4+25$ and we see that it satisfies the identity relation in Eq.(4) with n=3. Moreover, using the identity in Eq.(5), we have 25=7+18. Therefore we have several Lucas numbers 4, 7 (4+7=11), 18 and 29, hidden in the function A_0-function corresponding to uridine. From this observation, we can write the number 7 in the function SPI(34) above as 3+4, from the definition of Lucas numbers (see above). Summing up, we have the following set of nine numbers: [19, 9, 4, 3, 2, 1, 1, 1, 1], arranged in decreasing order. As our last step, we use a Gödel encoding procedure[3] (see Négadi, 2008, 2009a) and apply it to the above number sequence:

$$enc[19, 9, 4, 3, 2, 1, 1, 1, 1]=2^{19}\times3^{9}\times5^{4}\times7^{3}\times11^{2}\times13\times17\times19\times23 \quad (14)$$

Now, it appears that the number obtained this way in the right hand of Eq.(14) is nothing but the factorial of the number 23:

$$2^{19}\times3^{9}\times5^{4}\times7^{3}\times11^{2}\times13\times17\times19\times23=23!=25852016738884976640000 \quad (5)$$

[3]https://en.wikipedia.org/wiki/Gödel_numbering

which was the starting point of our 2007 paper entitled "*The genetic code multiplet structure, in one number*". Note, importantly, that 23! is the total number of permutations of the symmetric group S_{23}. We have therefore a first important result which was published very recently (Négadi, 2019) and encapsulated in the title "*Fibonacci number in RNA imply the one-number model of the genetic code*". Before continuing, let us recall some other facts concerning the multiplet structure of the genetic code. It is well known that the 64 triplet-codons of the Standard Genetic Code are organized into five *multiplets* of amino acids: 5 *quartets* (total 20 codons), Proline (P), Alanine (A), Threonine (T), Valine (V), Glycine (G); 3 *sextets* (total 18 codons) Serine (S), Leucine (L), Arginine (R); 9 *doublets* (total 18 codons) Phenylalanine (F), Tyrosine (Y), Cysteine (C), Histidine (H), Glutamine (Q), Asparagine (N), Lysine (K), Glutamic Acid (E), Aspartic Acid (D); 1 *triplet* (3 codons) Isoleucine and finally 2 *singlets* (1 codon each) Methionine (M) and Tryptophane (W). There are 64 codons in all, among them 61 code for 20 amino acids coded so 41 codons are said "degenerate" (61-20). Now, the 18 digits and the 5 symbols "0" in Eq.(15) could be sorted to match, exactly, the above described multiplet structure. We have (Négadi, 2007, 2009a)

$$\begin{aligned}&5\ \text{“}quartets\text{”}: \{3, 5, 5, 7, 7\}\rightarrow \{G, A, P, V, T\}\\ &3\ \text{“}sextets\text{”}: \{1, 2, 9\}\rightarrow \{S, L, R\}\\ &9\ \text{“}doublets\text{”}: \{4, 4, 6, 6, 6, 8, 8, 8, 8\}\rightarrow\{C, N, D, K, Q, E, H, F, Y\}\\ &1\ \text{“}triplet\text{”}: \{2\}\rightarrow \{I\}\\ &2\ \text{“}singlets\text{”}: \{0, 0\}\rightarrow \{M, W\}\\ &3\ \text{“}stops\text{”}: \{0, 0, 0\}\rightarrow \{UAA, UAG, UGA\}\end{aligned} \quad (16)$$

where we have indicated the one-letter code for the amino acids and the three stop codons are written explicitly. Here, we recall, briefly, how the detailed degeneracies are computed from the digits in Eq.(16), see (Négadi, 2007, 2009a). Let the number of digits in each multiplet be the number of amino acids of the multiplet, as explicitly shown in Eq.(16). The total degeneracy of a given multiplet, that is the number of degenerate codons in that multiplet, is computed as the sum of the a_0-functions of the digits the multiplet is composed of (*the digits taken without repetition, that is, counted only once*). We have

$$\begin{aligned}&5\ \text{“}quartets\text{”}: a_0(3)+a_0(5)+a_0(7)=3+5+7=15\\ &3\ \text{“}sextets\text{”}\ : 1+a_0(2)+a_0(9)=1+2+(3+3)=9\\ &9\ \text{“}doublets\text{”}:\\ &\quad a_0(4)+a_0(6)+a_0(8)=(2+2)+(2+3)+[2+2+2]=9+6=15\\ &1\ \text{“}triplet\text{”}:\ a_0(2)=2\\ &2\ \text{“}singlets\text{”}: \{0, 0\}\rightarrow 0\\ &3\ \text{“}stops\text{”}: \{0, 0, 0\}\rightarrow 0\end{aligned} \quad (17)$$

This is well known Rumer's division of the genetic

code table into two equal sets M_1 and M_2 (Rumer, 1966): M_1 comprises 5 quartets with 15 degenerate codons (total 20=5+15 codons) and the three quartet-part of the 3 sextets with 9 degenerate codons (total 12=3+9 codons) and M_2 comprises the 9 doublets with 9 degenerate codons (total 18=9+9 codons) *together* with the 6 doublet-part of the 3 sextets (6=2+2+2 codons), the triplet with 2 degenerate codons (total 3=1+2 codons), 2 singlets (2 codons) with zero degeneracy and finally 3 stop codons (3 codons) also with zero degeneracy. M_1 has 32 codons and M_2 has also 32 codons. There are also two other and different interesting ways to compute the degeneracies, but we shall not insist on this in the present paper. Also, many interesting results could be deduced from the above model based on the *unique* number 23! in Eq.(15) as, for example, the chemical structure of the 20 amino acids (number of hydrogen atoms, number of atoms, etc.), see (Négadi, 2009b).

Let us return now to an important aspect of the main subject of this paper. The number in Eq.(15) has two representations, much like the "two faces of the same coin". First, the decimal place-value representation in the right hand of Eq.(15) and, second, the prime factorization from the Fundamental Theorem of Arithmetic in the left hand of Eq.(15). From the former, we have 18 digits, with possible multiplicities and 5 zeros in the form 2+3 (see above). Consider Eqs.(12) and (13) obtained, we recall, from the atomic composition of the RNA-ribonucleotides. We show, below, that they are in complete agreement with both representations of the number 23! in Eq.(15). Let us, first, look at Eq.(13), the second sub-partition, and compare what it says to the digits in Eq.(15). Besides $\lambda(34)=2$ which corresponds to the two singlets and $\lambda(110)=3$ which corresponds to the three stop codons, we have $a_0(110)=2+5+11=18$ and it corresponds, as we saw above, to the 18 amino acids coded by more than one codon.The prime factors of the number 110 are 2, 5 and 11. Comparing with the 18 digits in Eq.(15), we see that there are 7 prime numbers (2+5) and 11 are non-prime numbers; 2 (the only even prime) appears two times and there are 5 odd primes. This matching is quite amazing. Second, looking at the other sub-partition, Eq.(12), we have shown above, through Eq.(14), that, *alone*, it leads to the number 23!, which constitutes the basis of our one-number model of the genetic code. Note, importantly, that the two additions considered in Eq.(11), $\lambda(34)=2$ and $\lambda(110)=3$, were appropriate, and welcome, to include the two amino acids singlets (Methionine and Tryptophane) and the three stop codons, but absolutely not necessary to derive Eq.(15), and hence our model.

Back to RNA

In the previous section we have shown, starting from the (numerous) Fibonacci numbers found in the chemical structure of the four ribonucleotides, that the "one-number" model of the genetic code, introduced by us in 2007, naturally emerges. In this section, which constitutes the core and the main "*raison d'être*" of this paper, we take the *opposite way*, that is, starting from our "one-number" model, we show that the detailed chemical structure, and the appropriate "uridine-and-the-rest" partition of the four ribonucleotides is recovered, fully. To show this, let us begin by considering the 18 digits in Eq.(15). It is seen that *all* the individual decimal digits, from 1 to 9, occur and, except for 1, 3 and 9, all the others show multiplicities. Now, (i) the sum of the digits, taking into account the multiplicities is equal to 99 and (ii) ignoring the multiplicities the sum is equal to 45. Adding these two sums, we have 99+45=144. This is very interesting as this last number is nothing but the total number of atoms in the four ribonucleotides (see sections 3 and 4). In fact, we can do far better. As a matter of fact, let us split the sum 99, mentioned above, into two parts, the sum of the digits which occur only once 1+3+9=13, on the one hand, and the sum of the digits which occur with multiplicity 99-13=86, on the other. We have therefore 144=13+86. Now, for the number 45, we have tht $A_0(45)=21$ so that $45-A_0(45)=24$ or $45=21+24$. As a first conclusion we have $144=(13+86)+(21+24)$ or, by appropriately grouping the terms

$$144=(13+21)+(86+24)=34+110 \qquad (18)$$

This is just the partition of the total number of atoms into "uridine-and-the-rest" which was the starting point in section 4 and allowed us to make contact with the one-number model of the genetic code. For uridine, the matching is already perfect: 13 hydrogen atoms and 21 CNOP atoms (13+21=34, *i.e.*, Fibonacci numbers). For the other three ribonucleotides, a bit more work is necessary. Let us compute the A_0-functions of the numbers 21 (=3×7) and 24 (=2^3×3) whose sum is 45. We have $A_0(21)=A_0(24)=18$ or, written as an "identity", $A_0(21)-A_0(24)=0$. Introducing this identity in the second parenthesis of Eq.(18), we get

$$144=(13+21)+[(86-18)+(24+18)]=(13+21)+(68+42)=34+110$$
$$=(13+42)+(21+68)=55+89 \quad (19)$$

Here also, we have a perfect matching for the CNOP atoms in the remaining three ribonucleotides: 42 hydrogen atoms and 68 CNOP atoms (see Table 2). Of course, we recover also the total number of hydrogen atoms and CNOP atoms in all four ribonucleotides: 13+42=55 and 21+68=89 (see Table 2). (Observe by the way, the nice symmetry inversion in the above digits operated by the above identity: 86→68 and 24→42.)

Summary and Some Additional Relevant Results

This work is a continuation of a recent short paper (Négadi, 2019), where we considered the atomic composition of the RNA four ribonucleotides UMP, CMP, AMP and GMP full of Fibonacci numbers, as a starting point, to derive our 2007 "*one-number model of the genetic code*". Here, in this paper, we have first given more details on the computations of the work mentioned above and, next, have considered the opposite way, that is, starting from the one-number model we derived the explicitly detailed atomic composition of the four ribonucleotides, conspicuously partitioned as "uridine-and-the-rest", exactly as explained in section 4. This double approach has been encapsulated in the title of the present paper.

To end this paper, let us also include some interesting by-product results, completing those of section 5. In our paper (Négadi, 2014), we used the A_0-function of the number 144, the total number of atoms in the four ribonucleotides, to derive a mathematical representation of so-called "*condensation equation*" for the formation of a ribonucleotide from the three subunits, a nitrogenous base with the number of atoms N, a phosphate group (8 atoms) and a ribose (20 atoms) with the release of two water molecules (6 atoms) in the form

$$N+ribose+phosphate - 2H_2O = N+20+8-6 \quad (20)$$

This relation means that the number of atoms in a given ribonucleotide is the sum of the number of atoms in the corresponding nucleobase N and the number of atoms in the *common block* with the number of atoms equal to 22 (20+8-6), see equation (20). For the four ribonucleotides UMP, CMP, AMP and GMP, the number of atoms in the four nucleobases is equal to 56 (12+13+15+16), see Eq.(8), and the number of atoms in the four common blocks is equal to 88 (=4×22) and we have, like it should be 56+88=144. In what follows, we shall see that the number of atoms in uridine, 34, and the total number of atoms 144, appear to be mathematically "linked". As a matter of fact, using the well known arithmetic function π(n), which gives the number of prime numbers less or equal to a given integer n, we have π(144)=34 so that 144−π(144)=110 and therefore, rearranging

$$144 = 34+110 \quad (21)$$

Equation (21) is nothing but the partition "uridine-and-the-rest" considered in section 4. Other interesting results leading to the partition 144=56+88 could also be inferred using the numbers 144, 34 and 110 and two other arithmetic functions(i) the "totient function" φ of a positive integer n greater than 1 which is defined to be the number of positive integers less than n that are coprime to n (with φ(1)=1) and (ii) the σ-function which is simply the sum of the divisors function of an integer n. We have, using (i) φ(34)=16 and φ(110)=40, so that

$$144 - [\varphi(34)+\varphi(110)]= 144-56=88 \quad (22)$$

Rearranging the above result gives 56+88=144, which is nothing but the total number of atoms in the four nucleobases, on the one hand, and the numbers of atoms in the four blocks, on the other, see above. Using now (ii), we get σ(34)=54 so that

$$144 - [34+\sigma(34)]= 144-88=56 \quad (23)$$

and the conclusion is the same as for Eq.(22) above. Finally, let us use the Fibonacci number 34, only, and compute the product τ(34)×σ(34)×φ(34), where the functions σ and φ were already defined above and τ is the *number of divisors function*. We have with τ(34)=4, σ(34)=54 and φ(34)=16 so that

$$\tau(34)\sigma(34)\varphi(34)=3456 \quad (24)$$

Now, a related function to the function φ is the function that gives *the sum of the coprimes*; it is defined for an integer n by (1/2)×[n×φ(n)]. Using this latter for n=34 and adding the above σ-function, we have

$$\frac{1}{2}[34\times\varphi(34)] + \sigma(34)=326 \quad (25)$$

The two numbers, 3456 and 326, are interesting; they have been already derived from the number 144, only (see Négadi, 2014). Here, as we see, they are derived from the (related) number 34, only. First, taking two times equation (24), gives 6912 and this number is equal to the *total number of atoms in the 64 RNA-codons* (using the ribonucleotides UMP, CMP, AMP and GMP), see (Négadi, 2014). On the other hand, 3456 is equal to the number of atoms in the 32 codons for the class-II Aminoacyl t-RNA Synthetases (AARS) amino acids and the difference 3456−326=3130 is equal to the number of atoms in the 29 codons for the class-I Aminoacyl t-RNA Synthetases amino acids (see Négadi, 2014). As for the number 326, it is equal to the number of atoms in the three stop codons (UAA, UAG, UGA) as computed using the ribonucleotides UMP, CMP, AMP and GMP, that is $3\times34+4\times37+2\times38=326$. A little manipulation gives

$$\begin{gathered}\text{Class-I AARS: } 3456 \quad (26)\\ \text{Class-II AARS: } (3456-326)+326=3130+326\end{gathered}$$

We reach therefore the same result as in (Négadi, 2014) but, here, using only the other interesting Fibonacci number 34.

References

Brosius J. The contribution of RNAs and retro position to evolutionary novelties. Genetica. 2003a; 118: 99-116.

Brosius J. Gene duplication and other evolutionary strategies: from the RNA world to the Future. Journal of Structuraland Functional Genomics. 2003b; 3: 1-17.

Chech TR, Steitz JA. The noncoding RNA revolution trashing old rules to forge new ones. Cell. 2014; 157(1): 77-94.

Gilbert W. Origin of Life: The RNA world. Nature. 1986; 319: 618.

Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nature Reviews, Genetics. 2009; 10(3): 155-159.

Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nature reviews genetics. 2009; 10(3): 155-159

Négadi T. The genetic code multiplet structure, in one number. Symmetry: Culture and Science. 2007; 18: 2-3: 149-160.

Négadi T. Fibonacci numbers in RNA imply the one-number model of the genetic code. Symmetry: Culture of Science. 2019; 30(2): 181-185.

Négadi T. The genetic code via Gödel encoding. The Open Physical Chemistry Journal. 2008; 2: 1-5.

Négadi T. A taylor-made arithmetic model of the genetic code and applications. Symmetry: Culture and Science. 2009a; 20(1-4): 51-76.

Négadi T. The genetic code degeneracy and the amino acids chemical composition are connected. NeuroQuantology. 2009b; 7-1: 181-187.

Négadi T. The genetic code invariance: when Euler and Fibonacci meet, Symmetry: Culture and Science. 2014; 25-3: 261-278.

PastuzinED, Day CE, Kearns RN, Kyrke-Smith M, Taibi AV, McCormick J, Yoder N, Belnap DM, Erlendsson S, Morado DR, Briggs JAG, Feshotte C, Shepherd JD. The Neuronal Gene Arc Encodesa Repurposed Retrotransposon Gag Protein that Mediates Intercellular RNA Transfer. Cell. 2018; 172(1-2): 275-288.

Rumer Y. About systematization of the genetic code. Dok. Akad Nauk SSSR. 1966; 167: 1393-1394.