Applied Mathematics, 2012, 3, 107-112
doi:10.4236/am.2012.31017 Published Online January 2012 (http://www.SciRP.org/journal/am)

On Conjugation Partitions of Sets of Trinucleotides

Lorenzo Bussoli[1], Christian J. Michel[2], Giuseppe Pirillo[1,3]
[1]Dipartimento di Matematica U.Dini, Firenze, Italia
[2]Equipe de Bioinformatique Théorique, Université de Strasbourg, Strasbourg, France
[3]Université de Marne-la-Vallée, Marne-la-Vallée, France
Email: {bussoli, pirillo}@math.unifi.it, michel@dpt-info.u-strasbg.fr

Received October 2, 2011; revised December 5, 2011; accepted December 13, 2011

ABSTRACT

We prove that a trinucleotide circular code is self-complementary if and only if its two conjugated classes are complement of each other. Using only this proposition, we prove that if a circular code is self-complementary then either both its two conjugated classes are circular codes or none is a circular code.

Keywords: Trinucleotide; Conjugated Trinucleotides; Code; Circular Code; Self-Complementary Circular Code; Complementary Circular Codes.

1. Introduction

We continue our study of the combinatorial properties of trinucleotide circular codes. A trinucleotide is a word of three letters (triletter) on the genetic alphabet $\{A,C,G,T\}$. The set of 64 trinucleotides is a code in the sense of language theory, more precisely a uniform code, but not a circular code [1,2]. In order to have an intuitive meaning of these notions, codes are written on a straight line while circular codes are written on a circle, but, in both cases, unique decipherability is required.

Comma free codes, a very particular case of circular codes, have been studied for a long time, e.g. [3-5]. After the discovery of a circular code in genes with important properties [6], circular codes are mathematical objects studied in combinatorics, theoretical computer science and theoretical biology, e.g. [7-23].

There are 528 self-complementary circular codes of 20 trinucleotides [6,24,25] and, as proved here, they are naturally partitioned into two quite symmetric classes.

Let $\mathcal{T}=\{AAA,CCC,GGG,TTT\}$ be the four trinucleotides with identical nucleotides. In this paper, we study some particular partitions of $\mathcal{A}_4^3\setminus\mathcal{T}$. Indeed, each circular code X_0 can be associated with two other subsets X_1 and X_2 of $\mathcal{A}_4^3\setminus\mathcal{T}$ simply by operating two circular permutations of one letter and two letters on the trinucleotides of X_0. Then, we prove our main result, *i.e.* a circular code is self-complementary if and only if the remaining two classes are complement of each other. Furthermore, we also show that a subset of $\mathcal{A}_4^3\setminus\mathcal{T}$ is a circular code if and only if the set consisting of all its complements is a circular code.

As a consequence of these results, we also prove that if a circular code is self-complementary then either both its two conjugated classes are circular codes or none is a circular code.

In Section 2, we give the necessary definitions and a characterization for a set of trinucleotides to be a circular code. In Section 3, we give the results, mainly expressed by Proposition 7 and Proposition 8.

2. Definitions

The classical notions of alphabet, empty word, length, factor, proper factor, prefix, proper prefix, suffix, proper suffix, lexicographical order, etc. are those of [1]. Let $\mathcal{A}_4=\{A,C,G,T\}$ denote the genetic alphabet, lexicographically ordered with $A<C<G<T$. We use the following notation:

- $\mathcal{A}_4^*$ (respectively $\mathcal{A}_4^+$) is the set of words (respectively non-empty words) over $\mathcal{A}_4$;
- $\mathcal{A}_4^2$ is the set of the 16 words of length 2 (diletters or dinucleotides);
- $\mathcal{A}_4^3$ is the set of the 64 words of length 3 (triletters or trinucleotides).

We now recall two important genetic maps, the definitions of code and circular code, and the property of C^3-self-complementarity for a circular code, in particular [1,6,17,24,25].

Definition 1. *The complementarity map* $\mathcal{C}:\ \mathcal{A}_4^+\to\mathcal{A}_4^+$ *is defined by* $\mathcal{C}(A)=T$, $\mathcal{C}(T)=A$, $\mathcal{C}(C)=G$ *and* $\mathcal{C}(G)=C$, *and by* $\mathcal{C}(uv)=\mathcal{C}(v)\mathcal{C}(u)$ *for all* $u,v\in\mathcal{A}_4^+$, e.g., $\mathcal{C}(AAC)=GTT$.

The map $\mathcal{C}$ on words is naturally extended to a word set X: its complementary trinucleotide set $\mathcal{C}(X)$ is obtained by applying the complementarity map $\mathcal{C}$ to all the trinucleotides of X.

Definition 2. *The circular permutation map* $\mathcal{P}$: $\mathcal{A}_4^3 \to \mathcal{A}_4^3$ *permutes circularly each trinucleotide* $l_1l_2l_3$ *as follows* $\mathcal{P}(l_1l_2l_3) = l_2l_3l_1$.

The map $\mathcal{P}$ on words is also naturally extended to a word set X: its permuted trinucleotide set $\mathcal{P}(X)$ is obtained by applying the circular permutation map $\mathcal{P}$ to all the trinucleotides of X. We shortly write $\mathcal{P}^2(X)$ for $\mathcal{P}(\mathcal{P}(X))$.

Definition 3. *A set* X *of words is a code if, for each* $x_1,\cdots,x_n,x_1',\cdots,x_m' \in X$, $n,m \geq 1$, *the condition* $x_1\cdots x_n = x_1'\cdots x_m'$ *implies* $n = m$ *and* $x_i = x_i'$ *for* $i = 1,\cdots,n$.

Definition 4. *A trinucleotide code* X *is circular if, for each* $x_1,\cdots,x_n,x_1',\cdots,x_m' \in X$, $n,m \geq 1$, $p \in \mathcal{A}_4^*$, $s \in \mathcal{A}_4^+$, *the conditions* $sx_2\cdots x_np = x_1'\cdots x_m'$ *and* $x_1 = ps$ *imply* $n = m$, $p = \varepsilon$ *(empty word) and* $x_i = x_i'$ *for* $i = 1,\cdots,n$.

Definition 5. *A trinucleotide code* X *is self-complementary if, for each* $x \in X$, $\mathcal{C}(x) \in X$.

Definition 6. *If* X_0 *is a subset of* $\mathcal{A}_4^3 \setminus \mathcal{T}$, *we denote by* X_1 *the permuted trinucleotide set* $\mathcal{P}(X_0)$ *and by* X_2 *the permuted trinucleotide set* $\mathcal{P}^2(X_0)$ *and we call* X_1 *and* X_2 *the conjugated classes of* X_0.

Definition 7. *A trinucleotide circular code* X_0 *is* C^3-*self-complementary if* X_0, X_1 *and* X_2 *are circular codes satisfying the following properties*: $X_0 = \mathcal{C}(X_0)$ (*self-complementary*), $\mathcal{C}(X_1) = X_2$ (*and* $\mathcal{C}(X_2) = X_1$).

We have proved that there are exactly 528 self-complementary trinucleotide circular codes having 20 elements [6,24,25].

The concept of necklace was introduced by Pirillo [17] in order to characterize the circular codes for an efficient algorithm development. Let $l_1,l_2,\cdots,l_{n-1},l_n,\cdots$ be letters in $\mathcal{A}_4$, $d_1,d_2,\cdots,d_{n-1},d_n,\cdots$ diletters in $\mathcal{A}_4^2$ and $n \geq 2$ an integer.

Definition 8. *Letter Diletter Continued Necklace* (*LDCN*): *We say that the ordered sequence* $l_1,d_1,l_2,d_2,\cdots,d_{n-1},l_n,d_n,l_{n+1}$ *is an* $(n+1)$*LDCN for a subset* $X \subset \mathcal{A}_4^3$ *if*

$$l_1d_1,l_2d_2,\cdots,l_nd_n \in X$$

and

$$d_1l_2,d_2l_3,\cdots,d_{n-1}l_n,d_nl_{n+1} \in X .$$

Any trinucleotide set is a code (more precisely, a uniform code [1]) but only few of them are circular codes. We have the following proposition.

Proposition 1 [17]. *Let* X *be a trinucleotide code. The following conditions are equivalent*:

1) X *is a circular code*;
2) X *has no* 5*LDCN.*

The figure below explains the notion of 5*LDCN*.

l_1	d_1	l_2	d_2	l_3	d_3	l_4	d_4	l_5

3. Results

Proposition 2. *If* X_0 *is a trinucleotide circular code having* 20 *elements and* X_1 *and* X_2 *are its two conjugated classes then* X_0, X_1 *and* X_2 *constitute a partition of* $\mathcal{A}_4^3 \setminus \mathcal{T}$.

Proof. It is enough to prove that $X_0 \cap X_1 = X_0 \cap X_2 = X_1 \cap X_2 = \varnothing$. Suppose that the trinucleotide $l_1l_2l_3$ belongs both to the classes X_0 and X_1. Then $l_1l_2l_3$ and $l_3l_1l_2$ are both in class X_0. As no two conjugated trinucleotides can belong to a circular code, we are in contradiction. Suppose that the trinucleotide $l_1l_2l_3$ belongs both to the classes X_0 and X_2. Then $l_1l_2l_3$ and $l_2l_3l_1$ are both in class X_0. As no two conjugated trinucleotides can belong to a circular code, we are in contradiction. Suppose that the trinucleotide $l_1l_2l_3$ belongs both to the classes X_1 and X_2. Then $l_3l_1l_2$ and $l_2l_3l_1$ are both in class X_0. As no two conjugated trinucleotides can belong to a circular code, we are in contradiction. So, $X_0 \cap X_1 = X_0 \cap X_2 = X_1 \cap X_2 = \varnothing$. ☐

Proposition 3. *The class of self-complementary circular codes* X_0 *with both* X_1 *and* X_2 *in the class of circular codes is non-empty.*

Proof. Consider, for example, the following set X_0 of 20 trinucleotides

$$X_0 = \{AAC, AAG, AAT, ACC, ACG, ACT, AGC, AGG, AGT, ATC, ATT, CCT, CGT, CTT, GAT, GCC, GCT, GGC, GGT, GTT\}.$$

It is enough to prove that X_0 is a self-complementary circular code and that its two conjugated classes X_1 and X_2 are also circular codes.

X_0 **is a self-complementary circular code**.

X_0 is self-complementary. Obvious by inspection.

X_0 is a circular code. We use Proposition 1 [17]. By way of contradiction, suppose that X_0 admits a 5*LDCN*. As l_2 can be A, C, G or T, it is enough to prove that each choice leads to a contradiction.

1) If $l_2 = A$ then there is no possible d_1 as A is not a suffix of any trinucleotide of X_0, contradiction.

2) If $l_2 = C$, there are three possible d_2:

- if $d_2 = CT$ (a) or $d_2 = GT$ (b) then $l_3 = T$ (c) but there is no possible d_3 as T is not a prefix of any trinucleotide of X_0, contradiction,
- if $d_2 = TT$ (d), there is a contradiction as no trinucleotide of X_0 has a prefix TT.

3) If $l_2 = G$, there are six possible d_2:

- if $d_2 = CT$ or $d_2 = GT$, contradiction (a) and (b),
- if $d_2 = CC$ then $l_3 = T$, contradiction (c),
- if $d_2 = GC$ or $d_2 = AT$ then $l_3 = C$ or $l_3 = T$:

- ◆ if $l_3 = C$, there are three possible d_3: if $d_3 = CT$ or $d_3 = GT$ then $l_4 = T$, similarly to (c), contradiction, and if $d_3 = TT$, similarly to (d), contradiction,
- ◆ if $l_3 = T$, contradiction (c),
- if $d_2 = TT$, contradiction (d).

4) If $l_2 = T$, similarly to (c), contradiction.

As, for each letter, we cannot complete the assumed 5*LDCN* for X_0, we are in contradiction. Hence, X_0 is a circular code.

$X_1 = \mathcal{P}^1(X_0)$ **is a circular code**. We have to prove that

$$X_1 = \{ACA, AGA, ATA, ATG, CCA, CCG, CGA, CTA, CTC, CTG, GCA, GCG, GGA, GTA, GTC, GTG, TCA, TTA, TTC, TTG\}$$

is a circular code. By way of contradiction, assume that X_1 admits a 5*LDCN*.

1) If $l_2 = A$, there are four possible d_2: CA, GA, TA and TG, but no possible l_3, contradiction.

2) If $l_2 = C$, there are three possible d_1: CT, GT and TT, but no possible l_1, contradiction.

3) If $l_2 = G$, there are six possible d_1: AT, CC and GC, and the cases CT, GT and TT already seen, but no possible l_1, contradiction.

4) If $l_2 = T$, there is no possible d_1, contradiction.

Hence, X_1 is also a circular code.

$X_2 = \mathcal{P}^2(X_0)$ **is a circular code**. Finally, we have to prove that

$$X_2 = \{CAA, CAC, CAG, CAT, CGC, CGG, GAA, GAC, GAG, TAA, TAC, TAG, TAT, TCC, TCG, TCT, TGA, TGC, TGG, TGT\}$$

is a circular code. By way of contradiction, assume that X_2 admits a 5*LDCN*.

1) If $l_2 = A$, there is no possible d_2, contradiction.

2) If $l_2 = C$, there are six possible d_2: AA, AC, AG, AT, GC and GG, but no possible l_3, contradiction.

3) If $l_2 = G$, there are three possible d_2: AA, AC and AG which are cases already seen, contradiction.

4) If $l_2 = T$, there are four possible d_1: CA, TA, TC and TG, but no possible l_1, contradiction.

Hence, as X_0 and X_1, X_2 is also a circular code. □

Proposition 4. *The class of self-complementary circular codes* X_0 *having* 20 *elements with neither* X_1 *nor* X_2 *in the class of circular codes is non-empty.*

Proof. Consider, for example, the following set X_0 of 20 trinucleotides

$$X_0 = \{AAC, AAG, AAT, ACC, ACG, ACT, AGC, AGT, ATC, ATT, CGT, CTT, GAT, GCC, GCT, GGA, GGC, GGT, GTT, TCC\}.$$

It is enough to prove that X_0 is a self-complementary circular code and that neither its conjugated class X_1 nor its conjugated class X_2 are circular codes.

X_0 **is a self-complementary circular code**.

X_0 is self-complementary. Obvious by inspection.

X_0 is a circular code. We use Proposition 1 [17]. By way of contradiction, assume that X_0 admits a 5*LDCN*.

1) If $l_2 = A$ then there is one possible $d_1 = GG$ but no possible l_1, contradiction.

2) If $l_2 = C$, there are two possible d_2:

- if $d_2 = GT$ then $l_3 = T$ (a) and $d_3 = CC$ (b) but there is no possible l_4, contradiction,
- if $d_2 = TT$ (c) then there is no possible l_3, contradiction.

3) If $l_2 = G$ we have seven possible d_2:

- if $d_2 = AT$ then $l_3 = C$ or $l_3 = T$:
  - ◆ if $l_3 = C$ (d) then $d_3 = GT$ or $d_3 = TT$:
    - if $d_3 = GT$ then $l_4 = T$ and $d_4 = CC$ but there is no possible l_5, contradiction,
    - if $d_3 = TT$ then there is no possible l_4, contradiction,
  - ◆ if $l_3 = T$, contradiction (a),
- if $d_2 = CC$, similarly to (b), contradiction,
- if $d_2 = CT$, $d_2 = GA$ or $d_2 = GT$ then $l_3 = T$, contradiction (a),
- if $d_2 = GC$ then $l_3 = C$ or $l_3 = T$, contradiction (a) and (d),
- if $d_2 = TT$, contradiction (c).

4) If $l_2 = T$, similarly to (a), contradiction.

Hence, X_0 is a circular code.

$X_1 = \mathcal{P}^1(X_0)$ **is not a circular code**. We have

$$X_1 = \{ACA, AGA, ATA, ATG, CCA, CCG, CCT, CGA, CTA, CTG, GAG, GCA, GCG, GTA, GTC, GTG, TCA, TTA, TTC, TTG\}.$$

We use a technique developed in [23]. Observe that X_1 contains $\{AGA, CCT, GAG, TTC\}$. So,

$$(l_1, d_1, l_2, d_2, l_3, d_3, l_4, d_4, l_5) = (A, GA, G, AG, A, GA, G, AG, A)$$

is a 5*LDCN* for this 4-element subset of X_1 and, a fortiori, for X_1 itself which, consequently, is not a circular code.

$X_2 = \mathcal{P}^2(X_0)$ **is not a circular code**. We have

$$X_2 = \{AGG, CAA, CAC, CAG, CAT, CGC, CGG, CTC, GAA, GAC, TAA, TAC, TAG, TAT, TCG, TCT, TGA, TGC, TGG, TGT\}.$$

We again use a technique developed in [23]. Remark that X_2 contains $\{GAA, CTC, AGG, TCT\}$. So,

$$\begin{aligned}&(l_1,d_1,l_2,d_2,l_3,d_3,l_4,d_4,l_5)\\&=(T,CT,C,TC,T,CT,C,TC,T)\end{aligned}$$

is a 5*LDCN* for this 4-element subset of X_2 and, a fortiori, for X_2 itself which, consequently, is not a circular code. □

We need the propositions hereafter and, in particular the following one which states a general property of the involutional antiisomorphisms such as the complementary map $\mathcal{C}$.

Proposition 5. *A subset* X *of* $\mathcal{A}_4^3 \setminus \mathcal{T}$ *is a circular code if and only if* $\mathcal{C}(X)$ *is a circular code.*

Proof. Suppose, first, that X is not a circular code and that $\mathcal{C}(X)$ is a circular code. So X has a 5*LDCN*. This means that there are 13 nucleotides, say

$$b_1,b_2,b_3,b_4,b_5,b_6,b_7,b_8,b_9,b_{10},b_{11},b_{12},b_{13}$$

such that the trinucleotides

$$b_1b_2b_3, b_4b_5b_6, b_7b_8b_9, b_{10}b_{11}b_{12} \in X$$

and

$$b_2b_3b_4, b_5b_6b_7, b_8b_9b_{10}, b_{11}b_{12}b_{13} \in X.$$

Now, consider the sequence

$$\begin{aligned}&\mathcal{C}(b_{13}),\mathcal{C}(b_{12}),\mathcal{C}(b_{11}),\mathcal{C}(b_{10}),\mathcal{C}(b_9),\mathcal{C}(b_8),\mathcal{C}(b_7),\\&\mathcal{C}(b_6),\mathcal{C}(b_5),\mathcal{C}(b_4),\mathcal{C}(b_3),\mathcal{C}(b_2),\mathcal{C}(b_1).\end{aligned}$$

All the following trinucleotides belong to $\mathcal{C}(X)$:

$$\begin{aligned}&\mathcal{C}(b_{13})\mathcal{C}(b_{12})\mathcal{C}(b_{11}),\mathcal{C}(b_{10})\mathcal{C}(b_9)\mathcal{C}(b_8),\\&\mathcal{C}(b_7)\mathcal{C}(b_6)\mathcal{C}(b_5),\mathcal{C}(b_4)\mathcal{C}(b_3)\mathcal{C}(b_2)\in\mathcal{C}(X)\end{aligned}$$

and

$$\begin{aligned}&\mathcal{C}(b_{12})\mathcal{C}(b_{11})\mathcal{C}(b_{10}),\mathcal{C}(b_9)\mathcal{C}(b_8)\mathcal{C}(b_7),\\&\mathcal{C}(b_6)\mathcal{C}(b_5)\mathcal{C}(b_4),\mathcal{C}(b_3)\mathcal{C}(b_2)\mathcal{C}(b_1)\in\mathcal{C}(X)\end{aligned}$$

as they are the complement of trinucleotides in X. So, $\mathcal{C}(X)$ admits a 5*LDCN* and it cannot be a circular code. Contradiction.

The case X is a circular code and $\mathcal{C}(X)$ is not a circular code is similar. □

Proposition 6. *Let* S *be a self-complementary subset of* $\mathcal{A}_4^3 \setminus \mathcal{T}$. *If* S *is partitioned into three classes such that two of them are the complement of each other then necessarily the third one is self-complementary.*

Proof. Let X, Y and Z be the three classes of an arbitrary partition of S and suppose that Y and Z are complementary, *i.e.* Y and Z satisfy $\mathcal{C}(Y)=Z$. Let t be a trinucleotide of X. We claim that $\mathcal{C}(t)\notin Y$. Indeed, in the opposite case, Z should not be the complement of Y because $t\in X$. We also claim that $\mathcal{C}(t)\notin Z$. Indeed, in the opposite case, Y should not be the complement of Z because $t\in X$. It remains the case $\mathcal{C}(t)\in X$. So, X is self-complementary. □

Remark 1. *Clearly, if* X, Y *and* Z *constitute an arbitrary partition of* $\mathcal{A}_4^3 \setminus \mathcal{T}$ *then the self-complementarity of* X *is not enough to ensure that* Y *and* Z *are complementary of each other. This remark is again true if, in addition,* X *is a self-complementary circular code having* 20 *elements. Indeed in this case, it is easy to make a partition* $\mathcal{A}_4^3 \setminus \{X\cup\mathcal{T}\}$ *in two classes* Y *and* Z *that are not complementary of each other. Any case, if we consider the partition of* $\mathcal{A}_4^3 \setminus \mathcal{T}$ *in the three classes given by a self-complementary trinucleotide circular code* X_0 *having* 20 *elements and by its two conjugated classes* X_1 *and* X_2 *then the necessary and sufficient condition holds* (*Proposition* 7 *below*).

Proposition 7. *A trinucleotide circular code* X_0 *having* 20 *elements is self-complementary if and only if* X_1 *and* X_2 *are complement of each other.*

Proof if part. It is a trivial consequence of Proposition 6.

Only if part. Suppose that X_0 is self-complementary and consider the partition X_0, X_1 and X_2 of $\mathcal{A}_4^3 \setminus \mathcal{T}$. Suppose that the trinucleotide, say $l_1l_2l_3$, belongs to X_0. Then, also

$$\mathcal{C}(l_3)\mathcal{C}(l_2)\mathcal{C}(l_1)\in X_0.$$

We have

$$l_2l_3l_1, \mathcal{C}(l_2)\mathcal{C}(l_1)\mathcal{C}(l_3)\in X_1$$

and

$$l_3l_1l_2, \mathcal{C}(l_1)\mathcal{C}(l_3)\mathcal{C}(l_2)\in X_2.$$

As $l_1l_2l_3$ is a generic trinucleotide of X_0 and as

$l_2l_3l_1$ is the complement of $\mathcal{C}(l_1)\mathcal{C}(l_3)\mathcal{C}(l_2)$

and

$\mathcal{C}(l_2)\mathcal{C}(l_1)\mathcal{C}(l_3)$ is the complement of $l_3l_1l_2$

then X_1 is the complement of X_2. □

As a consequence, we have the following proposition.

Proposition 8. *If a trinucleotide circular code* X_0 *having* 20 *elements is self-complementary then either*

1) X_1 *and* X_2 *are both circular codes*

or

2) X_1 *and* X_2 *are not circular codes* (*both have a necklace*).

Proof. We have four possibilities:

X_1 is a circular code and X_2 is a circular code;

X_1 is a circular code and X_2 is not a circular code;

X_1 is not a circular code and X_2 is a circular code;

X_1 is not a circular code and X_2 is not a circular code.

Now, by applying Propositions 3 and 4, we have that the first and the last possibilities can be effectively realized.

Suppose that, by way of contradiction, the second possibility is realized. So, X_1 is a circular code. By Proposition 7, we have $\mathcal{C}(X_1) = X_2$. So, by Proposition 5, X_2 must also be a circular code. Contradiction.

Suppose that, by way of contradiction, the third possibility is realized. So, X_2 is a circular code. By Proposition 7, we have $\mathcal{C}(X_2) = X_1$. So, by Proposition 5, X_1 must also be a circular code. Contradiction.

So, only the first and the last possibilities can occur. □

Hence, our proposition holds.

Proposition 9. *The* 528 *self-complementary circular codes having* 20 *elements are partitioned into two classes*: *one class contains codes with the two permuted sets* X_1 *and* X_2 *which are both circular codes while the other class contains codes with the two permuted sets* X_1 *and* X_2 *which both are not circular codes.*

Proof. It is enough to apply Proposition 8 to each of the 528 trinucleotide circular codes having 20 elements. □

4. Acknowledgements

We thank Jacques Justin for his advices. The second author thanks the Dipartimento di matematica U. Dini for giving him a friendly hospitality.

REFERENCES

[1] J. Berstel and D. Perrin, "Theory of Codes, Vol. 117, (Pure and Applied Mathematics)," Academic Press, London, 1985.

[2] J.-L. Lassez, "Circular codes and Synchronization," *International Journal of Computer & Information Sciences*, Vol. 5, 1976, pp. 201-208.

[3] F. H. C. Crick, J. S. Griffith and L. E. Orgel, "Codes without commas," *Proceedings of the National Academy of Sciences of the USA*, Vol. 43, 1957, pp. 416-421. doi:10.1073/pnas.43.5.416

[4] S. W. Golomb, B. Gordon and L. R. Welch, "Comma-Free codes," *Canadian Journal of Mathematics*, Vol. 10, No. 2, 1958, pp. 202-209. doi:10.4153/CJM-1958-023-9

[5] S. W. Golomb, L. R. Welch and M. Delbrück, "Construction and Properties of Comma-Free Codes," *Biologiske Meddel Danske Vidensk Selsk*, Vol. 23, 1958, pp. 1-34.

[6] D. G. Arquès and C. J. Michel, "A Complementary Circular Code in the Protein Coding Genes," *Journal of Theoretical Biology*, Vol. 182, No. 1, 1996, pp. 45-58. doi:10.1006/jtbi.1996.0142

[7] A. J. Koch and J. Lehman, "About a Symmetry of the Genetic Code," *Journal of Theoretical Biology*, Vol. 189, No. 2, 1997, pp. 171-174. doi:10.1006/jtbi.1997.0503

[8] M.-P. Béal and J. Senellart, "On the Bound of the Synchronization Delay of a Local Automaton," *Theoretical Computer Science*, Vol. 205, No. 1-2, 1998, pp. 297-306. doi:10.1016/S0304-3975(98)80011-X

[9] F. Bassino, "Generating Function of Circular Codes," *Advances in Applied Mathematics*, Vol. 22, No. 1, 1999, pp. 1-24. doi:10.1006/aama.1998.0613

[10] N. Štambuk, "On Circular Coding Properties of Gene and Protein Sequences," *Croatica Chemica Acta*, Vol. 72, No. 4, 1999, pp. 999-1008.

[11] R. Jolivet and F. Rothen, "Peculiar Symmetry of DNA Sequences and Evidence Suggesting its Evolutionary Origin in a Primeval Genetic Code," *First European Workshop Exo-/Astro-Biology*, Noordwijk, 21-23 May 2001, pp. 173-176.

[12] G. Frey and C. J. Michel, "Circular Codes in Archaeal Genomes," *Journal of Theoretical Biology*, Vol. 223, No. 4, 2003, pp. 413-431. doi:10.1016/S0022-5193(03)00119-X

[13] C. Nikolaou and Y. Almirantis, "Mutually Symmetric and Complementary Triplets: Difference in their Use Distinguish Systematically between Coding and Non-Coding Genomic Sequences," *Journal of Theoretical Biology*, Vol. 223, No. 4, 2003, pp. 477-487. doi:10.1016/S0022-5193(03)00123-1

[14] E. E. May, M. A. Vouk, D. L. Bitzer and D. I. Rosnick, "An Error-Correcting Framework for Genetic Sequence Analysis," *Journal of the Franklin Institute*, Vol. 341, No. 1-2, 2004, pp. 89-109. doi:10.1016/j.jfranklin.2003.12.009

[15] G. Frey and C. J. Michel, "Identification of Circular Codes in Bacterial Genomes and their Use in a Factorization Method for Retrieving the Reading Frames of Genes," *Computational Biology and Chemistry*, Vol. 30, No. 2, 2006, pp. 87-101. doi:10.1016/j.compbiolchem.2005.11.001

[16] J.-L. Lassez, R. A. Rossi and A. E. Bernal, "Crick's Hypothesis Revisited: The Existence of a Universal Coding Frame," *Proceedings of the* 21*st International Conference on Advanced Information Networking and Applications Workshops/Symposia* (*AINA*'07), Niagara Falls, Vol. 2, 21-23 May 2007, pp. 745-751.

[17] G. Pirillo, "A characterization for a Set of Trinucleotides to be a Circular Code," In: C. Pellegrini, P. Cerrai, P. Freguglia, V. Benci and G. Israel, Eds., *Determinism, Holism, and Complexity*, Kluwer, Boston, 2003.

[18] C. J. Michel, G. Pirillo and M. A. Pirillo, "Varieties of Comma-Free Codes," *Computers & Mathematics with Applications*, Vol. 55, No. 5, 2008, pp. 989-996. doi:10.1016/j.camwa.2006.12.091

[19] G. Pirillo, "A Hierarchy for Circular Codes," *Theoretical Informatics and Applications*, Vol. 42, No. 4, 2008, pp. 717-728.

[20] M. V. José, T. Govezensky, J. A. García and J. R. Bobadilla, "On the Evolution of the Standard Genetic Code: Vestiges of Critical Scale Invariance from the RNA World in Current Prokaryote Genomes," *PLoS ONE*, Vol. 4, No. 2, 2009, p. e4340. doi:10.1371/journal.pone.0004340

[21] G. Pirillo, "Some Remarks on Prefix and Suffix Codes," *Pure Mathematics and Applications*, Vol. 19, No. 2-3, 2008, pp. 53-60.

[22] G. Pirillo, "Non Sharing Border Codes," *Advances in*

Applied Mathematics, Vol. 3, No. 2, 2010, pp. 215-223.

[23] L. Bussoli, C. J. Michel and G. Pirillo, "On some Forbidden Configurations for Self-Complementary Trinucleotide Circular Codes," *Journal for Algebra Number Theory Academia*, Vol. 2, 2011, pp. 223-232.

[24] G. Pirillo and M. A. Pirillo, "Growth Function of Self-Complementary Circular Codes," *Biology Forum*, Vol. 98, 2005, pp. 97-110.

[25] C. J. Michel, G. Pirillo and M. A. Pirillo, "A Relation between Trinucleotide Comma-Free Codes and Trinucleotide Circular Codes," *Theoretical Computer Science*, Vol. 401, No. 1-3, 2008, pp. 17-25. doi:10.1016/j.tcs.2008.02.049