Hindawi Publishing Corporation
Computational Biology Journal
Volume 2013, Article ID 795418, 10 pages
http://dx.doi.org/10.1155/2013/795418

Research Article

Transition and Transversion on the Common Trinucleotide Circular Code

Emmanuel Benard and Christian J. Michel

Equipe de Bioinformatique Théorique, ICube, Université de Strasbourg, CNRS, 300 boulevard Sébastien Brant, 67400 Illkirch, France

Correspondence should be addressed to Christian J. Michel; c.michel@unistra.fr

Received 15 February 2013; Accepted 22 April 2013

Academic Editor: Alessandra Lumini

In 1996, a trinucleotide circular code which is maximum, self-complementary, and C^3, called X_0, was identified statistically on a large gene population of eukaryotes and prokaryotes (Arquès and Michel (1996)). Transition and transversions I and II are classical molecular evolution processes. A comprehensive computer analysis of these three evolution processes in the code X_0 shows some new results; in particular (i) transversion I on the 2nd position of any subset of trinucleotides of X_0 generates trinucleotide circular codes which are always C^3 and (ii) transversion II on the three positions of any subset of trinucleotides of X_0 yields no trinucleotide circular codes. These new results extend our theory of circular code in genes to its evolution under transition and transversion.

1. Introduction

We continue our study of properties of trinucleotide circular codes [1–5], trinucleotide comma-free codes [1, 6], strong trinucleotide circular codes [7], and the common trinucleotide circular code X_0 identified in genes [8] (see also the recent statistical analysis by [9]) which could be a translation code [10]. 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 (called genetic code), more precisely a uniform code but not a circular code (see Remark 2). In the past 50 years, codes, comma-free codes, and circular codes have been mathematical objects studied in theoretical biology, mainly to understand the structure and the origin of the genetic code as well as the reading frame (construction) of genes, for example, [11–13]. In order to have an intuitive meaning of these notions, codes are written on a straight line while comma-free codes and circular codes are written on a circle, but in both cases, unique decipherability is required. Circular codes only belong to some subsets of the 64 trinucleotide set while comma-free codes are even more constrained subsets of circular codes [1].

Before the discovery of the genetic code, Crick et al. [11] proposed a maximum comma-free code of 20 trinucleotides for coding the 20 amino acids. This comma-free code turned out to be invalid (see, e.g., [14]). In 1996, a maximum circular code X_0 of 20 trinucleotides was identified statistically on a large gene population of eukaryotes and also on a large gene population of prokaryotes [8]

$$X_0 = \{AAC, AAT, ACC, ATC, ATT, CAG, CTC, CTG, GAA, GAC, GAG, GAT, GCC, GGC, GGT, GTA, GTC, GTT, TAC, TTC\}. \quad (1)$$

This code X_0 has remarkable mathematical properties as it is a C^3 self-complementary maximum circular code (see the following). Since 1996, its properties have been studied in detail by different authors, for example, [9, 15–21]. Transition and transversions I and II are classical molecular evolution processes, for example, [22]. By using an algorithm based on the necklace, we perform here a comprehensive computer analysis of these three evolution processes in the code X_0. Some new results are identified with the code X_0 by computer analysis; in particular (i) transversion I on the 2nd position of any subset of trinucleotides of X_0 generates trinucleotide circular codes which are always C^3 and (ii) transversion II on the three positions of any subset of trinucleotides of X_0 yields no trinucleotide circular codes.

2. Preliminaries

The classical notions of language theory and codes can be found in [23, 24]. Let $\mathcal{A}_4 = \{A, C, G, T\}$ denote the genetic alphabet, lexicographically ordered by $A < C < G < T$. The set of words (nonempty words, resp.) on $\mathcal{A}_4$ is denoted by $\mathcal{A}_4^*$ ($\mathcal{A}_4^+$, resp.). The set of the 16 words of length 2 (dinucleotides or diletters) on $\mathcal{A}_4$ is denoted by $\mathcal{A}_4^2 = \{AA, AC, \ldots, TT\}$. The set of the 64 words of length 3 (trinucleotides or triletters) on $\mathcal{A}_4$ is denoted by $\mathcal{A}_4^3 = \{AAA, AAC, \ldots, TTT\}$.

Definition 1. A subset $X \subset \mathcal{A}_4^+$ is a code on $\mathcal{A}_4$ if for each $x_1, \ldots, x_n, x_1', \ldots, x_m' \in X$, $n, m \geq 1$, the condition $x_1 \cdots x_n = x_1' \cdots x_m'$ implies $n = m$ and $x_i = x_i'$ for $i = 1, \ldots, n$.

Remark 2. $\mathcal{A}_4^3$ is a code.

Any nonempty subset of $\mathcal{A}_4^3$ is a code called here trinucleotide code.

Definition 3. A trinucleotide code $X \subset \mathcal{A}_4^3$ is circular if, for each $x_1, \ldots, x_n, x_1', \ldots, x_m' \in X$, $n, m \geq 1$, $p \in \mathcal{A}_4^*$, $s \in \mathcal{A}_4^+$, the conditions $sx_2 \cdots x_n p = x_1' \cdots x_m'$ and $x_1 = ps$ imply $n = m$, $p = \varepsilon$ (empty word) and $x_i = x_i'$ for $i = 1, \ldots, n$.

Notation 1. A trinucleotide circular code is noted C.

Remark 4. $\mathcal{A}_4^3$ is not a trinucleotide circular code.

Let $l_1, l_2, \ldots, l_{n-1}, l_n$ be letters in $\mathcal{A}_4$, $d_1, d_2, \ldots, d_{n-1}, d_n$ diletters in $\mathcal{A}_4^2$, and n an integer satisfying $n \geq 2$.

Definition 5. We say that the ordered sequence $l_1, d_1, l_2, d_2, \ldots, d_{n-1}, l_n, d_n, l_{n+1}$ is an $(n+1)$LDCN (Letter Diletter Continued Necklace) for a subset $X \subset \mathcal{A}_4^3$ if

$$\begin{gathered} l_1 d_1, l_2 d_2, \ldots, l_n d_n \in X, \\ d_1 l_2, d_2 l_3, \ldots, d_{n-1} l_n, d_n l_{n+1} \in X. \end{gathered} \tag{2}$$

Only a few trinucleotide codes are circular. Two propositions based on the necklace concept allow to determine if a trinucleotide code is circular or not [2, 18].

Proposition 6 (see [18]). *Let X be a trinucleotide code. The following conditions are equivalent:*

(i) *X is a trinucleotide circular code;*

(ii) *X has no* 5 LDCN .

Definition 7. We say that the ordered sequence $l_1, d_1, l_2, d_2, \ldots, d_{n-1}, l_n, d_n, l_{n+1}$ is an $(n+1)$LDCCN (Letter Diletter Continued Closed Necklace) for a subset $X \subset \mathcal{A}_4^3$ if

$$\begin{gathered} l_1 d_1, l_2 d_2, \ldots, l_n d_n \in X, \\ d_1 l_2, d_2 l_3, \ldots, d_{n-1} l_n, d_n l_1 \in X. \end{gathered} \tag{3}$$

Proposition 8 (see [2]). *Let X be a trinucleotide code. The following conditions are equivalent:*

(i) *X is a trinucleotide circular code;*

(ii) *X has no n*LDCCN *for any integer $n \in \{2, 3, 4, 5\}$.*

Definition 9. A trinucleotide circular code $X \subset \mathcal{A}_4^3$ is maximal if, for each $x \in \mathcal{A}_4^3$, $x \notin X$, $X \cup \{x\}$ is not a trinucleotide circular code.

Definition 10. A trinucleotide circular code containing exactly k elements is called a k-trinucleotide circular code.

Definition 11. A 20-trinucleotide circular code is maximum as no trinucleotide circular code can contain more than 20 words.

Notation 2. A maximum trinucleotide circular code is noted MC.

Remark 12. A 20-trinucleotide circular code is both maximal and maximum.

We recall two classical genetic maps: complementary and circular permutation.

Definition 13. The complementary genetic map $\mathcal{C}: \mathcal{A}_4^+ \to \mathcal{A}_4^+$ is defined by

$$\begin{aligned} \mathcal{C}(A) &= T, \qquad \mathcal{C}(C) = G, \\ \mathcal{C}(G) &= C, \qquad \mathcal{C}(T) = A \end{aligned} \tag{4}$$

and for all $u, v \in \mathcal{A}_4^+$ by

$$\mathcal{C}(uv) = \mathcal{C}(v)\,\mathcal{C}(u). \tag{5}$$

Example 14. $\mathcal{C}(ACG) = CGT$. This map $\mathcal{C}$ is associated with the property of the complementary and antiparallel double helix (one DNA strand chemically oriented in a 5′-3′ direction and the other DNA strand in the opposite 3′-5′ direction).

Definition 15. The complementary map $\mathcal{C}$ on a trinucleotide x is naturally extended to a trinucleotide code X as follows:

$$\mathcal{C}(X) = \left\{ y \mid x, y \in \mathcal{A}_4^3,\ x \in X,\ y = \mathcal{C}(x) \right\}. \tag{6}$$

Definition 16. The circular permutation genetic map $\mathcal{P}: \mathcal{A}_4^3 \to \mathcal{A}_4^3$ permutes circularly a trinucleotide $l_1 l_2 l_3$, $l_1, l_2, l_3 \in \mathcal{A}_4$, as follows:

$$\mathcal{P}(l_1 l_2 l_3) = l_2 l_3 l_1. \tag{7}$$

Example 17. $\mathcal{P}(ACG) = CGA$.

Definition 18. The circular permutation map $\mathcal{P}$ on a trinucleotide x is naturally extended to a trinucleotide code X as follows:

$$\mathcal{P}(X) = \left\{ y \mid x, y \in \mathcal{A}_4^3,\ x \in X,\ y = \mathcal{P}(x) \right\}. \tag{8}$$

Notation 3. The kth iterate of $\mathscr{P}$ is denoted by $\mathscr{P}^k$.

Remark 19. The trinucleotide codes $\mathscr{P}(X)$ and $\mathscr{P}^2(X)$ are the conjugated classes of the trinucleotide code X.

Definition 20. A trinucleotide circular code X is self-complementary if, for each $x \in X$, $\mathscr{C}(x) \in X$.

Notation 4. A self-complementary trinucleotide circular code is noted SC.

Remark 21. A k-trinucleotide circular code for k odd cannot be self-complementary.

Definition 22. A trinucleotide circular code X is C^3 if X, $\mathscr{P}(X)$, and $\mathscr{P}^2(X)$ are trinucleotide circular codes.

Notation 5. A C^3 trinucleotide circular code X is noted C^3.

Definition 23. A trinucleotide circular code X is C^3 self-complementary maximum if X is maximum, $X = \mathscr{C}(X)$ (self-complementary), and $\mathscr{P}(X)$ and $\mathscr{P}^2(X)$ are trinucleotide circular codes satisfying $\mathscr{C}(\mathscr{P}(X)) = \mathscr{P}^2(X)$.

Notation 6. A C^3 self-complementary maximum circular code is noted MSC^3.

The set X_0 of 20 trinucleotides identified in the gene populations of both eukaryotes and prokaryotes is a C^3 self-complementary maximum circular code MSC^3 [8]; that is, X_0 is maximum, $X_0 = \mathscr{C}(X_0)$, $\mathscr{P}(X_0) = X_1$, and $\mathscr{P}^2(X_0) = X_2$ are trinucleotide circular codes, and $\mathscr{C}(X_1) = X_2$.

We recall three classical evolution genetic maps: transition and transversions I and II, for example, [22] and extend their definitions to the positions of a trinucleotide.

Definition 24. The transition evolution genetic map $\mathscr{T}$: $\mathscr{A}_4^+ \to \mathscr{A}_4^+$ is defined by

$$\begin{aligned} \mathscr{T}(A) &= G, \qquad \mathscr{T}(C) = T, \\ \mathscr{T}(G) &= A, \qquad \mathscr{T}(T) = C. \end{aligned} \tag{9}$$

Definition 25. The transition map $\mathscr{T}$ on a letter l can be applied in different positions of a trinucleotide $x = l_1 l_2 l_3$: $\mathscr{T}^i$, $i \in \{1, 2, 3\}$, is the transition on the position i of x, $\mathscr{T}^{i,j}$, $i, j \in \{1, 2, 3\}$ with $i < j$, is the transition on the two positions i and j of x, and $\mathscr{T}^{1,2,3}$ is the transition on the three positions of x.

Example 26. $\mathscr{T}^1(ACG) = GCG$, $\mathscr{T}^2(ACG) = ATG$, $\mathscr{T}^3(ACG) = ACA$, $\mathscr{T}^{1,2}(ACG) = GTG$, $\mathscr{T}^{1,3}(ACG) = GCA$, $\mathscr{T}^{2,3}(ACG) = ATA$, and $\mathscr{T}^{1,2,3}(ACG) = GTA$.

Definition 27. The transition maps $\mathscr{T}^i$, $\mathscr{T}^{i,j}$, $\mathscr{T}^{1,2,3}$ on a trinucleotide x are also extended to a trinucleotide code X, in a similar way to the genetic maps $\mathscr{C}$ and $\mathscr{P}$.

Definition 28. The transversion I evolution genetic map $\mathscr{V}_{\mathrm{I}}$: $\mathscr{A}_4^+ \to \mathscr{A}_4^+$ is defined by

$$\begin{aligned} \mathscr{V}_{\mathrm{I}}(A) &= T, \qquad \mathscr{V}_{\mathrm{I}}(C) = G, \\ \mathscr{V}_{\mathrm{I}}(G) &= C, \qquad \mathscr{V}_{\mathrm{I}}(T) = A. \end{aligned} \tag{10}$$

Definition 29. The transversion I map $\mathscr{V}_{\mathrm{I}}$ on a letter l can also be applied in different positions of a trinucleotide $x = l_1 l_2 l_3$: $\mathscr{V}_{\mathrm{I}}^i$, $i \in \{1, 2, 3\}$, is the transversion I on the position i of x, $\mathscr{V}_{\mathrm{I}}^{i,j}$, $i, j \in \{1, 2, 3\}$ with $i < j$, is the transversion I on the two positions i and j of x, and $\mathscr{V}_{\mathrm{I}}^{1,2,3}$ is the transversion I on the three positions of x.

Example 30. $\mathscr{V}_{\mathrm{I}}^1(ACG) = TCG$, $\mathscr{V}_{\mathrm{I}}^2(ACG) = AGG$, $\mathscr{V}_{\mathrm{I}}^3(ACG) = ACC$, $\mathscr{V}_{\mathrm{I}}^{1,2}(ACG) = TGG$, $\mathscr{V}_{\mathrm{I}}^{1,3}(ACG) = TCC$, $\mathscr{V}_{\mathrm{I}}^{2,3}(ACG) = AGC$, and $\mathscr{V}_{\mathrm{I}}^{1,2,3}(ACG) = TGC$.

Definition 31. The transversion I maps $\mathscr{V}_{\mathrm{I}}^i$, $\mathscr{V}_{\mathrm{I}}^{i,j}$, $\mathscr{V}_{\mathrm{I}}^{1,2,3}$ on a trinucleotide x are also extended to a trinucleotide code X, in a similar way to the genetic maps $\mathscr{C}$ and $\mathscr{P}$.

Definition 32. The transversion II evolution genetic map $\mathscr{V}_{\mathrm{II}}$: $\mathscr{A}_4^+ \to \mathscr{A}_4^+$ is defined by

$$\begin{aligned} \mathscr{V}_{\mathrm{II}}(A) &= C, \qquad \mathscr{V}_{\mathrm{II}}(C) = A, \\ \mathscr{V}_{\mathrm{II}}(G) &= T, \qquad \mathscr{V}_{\mathrm{II}}(T) = G. \end{aligned} \tag{11}$$

Definition 33. The transversion II map $\mathscr{V}_{\mathrm{II}}$ on a letter l can also be applied in different positions of a trinucleotide $x = l_1 l_2 l_3$: $\mathscr{V}_{\mathrm{II}}^i$, $i \in \{1, 2, 3\}$, is the transversion II on the position i of x, $\mathscr{V}_{\mathrm{II}}^{i,j}$, $i, j \in \{1, 2, 3\}$ with $i < j$, is the transversion II on the two positions i and j of x, and $\mathscr{V}_{\mathrm{II}}^{1,2,3}$ is the transversion II on the three positions of x.

Example 34. $\mathscr{V}_{\mathrm{II}}^1(ACG) = CCG$, $\mathscr{V}_{\mathrm{II}}^2(ACG) = AAG$, $\mathscr{V}_{\mathrm{II}}^3(ACG) = ACT$, $\mathscr{V}_{\mathrm{II}}^{1,2}(ACG) = CAG$, $\mathscr{V}_{\mathrm{II}}^{1,3}(ACG) = CCT$, $\mathscr{V}_{\mathrm{II}}^{2,3}(ACG) = AAT$, and $\mathscr{V}_{\mathrm{II}}^{1,2,3}(ACG) = CAT$.

Definition 35. The transversion II maps $\mathscr{V}_{\mathrm{II}}^i$, $\mathscr{V}_{\mathrm{II}}^{i,j}$, $\mathscr{V}_{\mathrm{II}}^{1,2,3}$ on a trinucleotide x are also extended to a trinucleotide code X, in a similar way to the genetic maps $\mathscr{C}$ and $\mathscr{P}$.

Definition 36. The evolution genetic maps in l trinucleotides of a trinucleotide circular code are defined by $\mathscr{T}(l)$ for transition, $\mathscr{V}_{\mathrm{I}}(l)$ for transversion I, and $\mathscr{V}_{\mathrm{II}}(l)$ for transversion II.

3. Results

An evolution genetic map, that is, $\mathscr{T}(l)$, $\mathscr{V}_{\mathrm{I}}(l)$, and $\mathscr{V}_{\mathrm{II}}(l)$, in l trinucleotides of the common trinucleotide circular code X_0 leads to $S(l) = \binom{20}{l}$ trinucleotide codes which are potentially circular. Table 1 gives these numbers $S(l)$.

Based on Proposition 6 allowing to test if a trinucleotide code is circular or not (algorithm not detailed, see, e.g., [2]), computer analyses of a great number of trinucleotide codes allow to identify here new properties with the common trinucleotide circular code X_0 observed in genes under evolution by transition and transversion.

Table 1: Number $S(l) = \binom{20}{l}$ of trinucleotide codes after an evolution genetic map (transition $\mathcal{T}(l)$, transversion I $\mathcal{V}_{\mathrm{I}}(l)$, and transversion II $\mathcal{V}_{\mathrm{II}}(l)$) in l trinucleotides of the common trinucleotide circular code X_0.

l	$S(l)$
{1, 19}	20
{2, 18}	190
{3, 17}	1140
{4, 16}	4845
{5, 15}	15504
{6, 14}	38760
{7, 13}	77520
{8, 12}	125970
{9, 11}	167960
10	184756
20	1

3.1. Transition Map

3.1.1. Transition Map $\mathcal{T}^i$

Result 1 (Table 2). For $l = 1, \ldots, 20$

$$\begin{aligned} &c\left(\mathcal{T}^1(l)\right) = c\left(\mathcal{T}^3(l)\right), \\ &\mathrm{mc}\left(\mathcal{T}^1(l)\right) = \mathrm{mc}\left(\mathcal{T}^3(l)\right), \\ &\mathrm{sc}\left(\mathcal{T}^1(l)\right) = \mathrm{sc}\left(\mathcal{T}^3(l)\right), \\ &c^3\left(\mathcal{T}^1(l)\right) = c^3\left(\mathcal{T}^3(l)\right), \\ &\mathrm{msc}^3\left(\mathcal{T}^1(l)\right) = \mathrm{msc}^3\left(\mathcal{T}^3(l)\right). \end{aligned} \tag{12}$$

As expected, the lists of trinucleotide circular codes C associated with $c(\mathcal{T}^1(l))$ and $c(\mathcal{T}^3(l))$ are different for $l = 1, \ldots, 13$ (not shown). No trinucleotide code is circular after a certain number of transitions $\mathcal{T}^i$ in the trinucleotides of the common trinucleotide circular code X_0. Precisely, for $l = 14, \ldots, 20$

$$c\left(\mathcal{T}^1(l)\right) = c\left(\mathcal{T}^3(l)\right) = 0 \tag{13}$$

and for $l = 10, \ldots, 20$

$$c\left(\mathcal{T}^2(l)\right) = 0. \tag{14}$$

The transition $\mathcal{T}^i$ generates a maximum number of trinucleotide circular codes C for

$$\begin{aligned} &\max\left\{c\left(\mathcal{T}^i(l)\right), i = 1, 2, 3, l = 1, \ldots, 20\right\} \\ &\quad = c\left(\mathcal{T}^1(7)\right) = c\left(\mathcal{T}^3(7)\right) = 1436 \end{aligned} \tag{15}$$

and a maximum number of C^3 self-complementary maximum circular codes MSC^3 for

$$\begin{aligned} &\max\left\{\mathrm{msc}^3\left(\mathcal{T}^i(l)\right), i = 1, 2, 3, l = 1, \ldots, 20\right\} \\ &\quad = \mathrm{msc}^3\left(\mathcal{T}^1(6)\right) = \mathrm{msc}^3\left(\mathcal{T}^3(6)\right) = 20. \end{aligned} \tag{16}$$

Table 2: Transition map $\mathcal{T}^i(l)$ in l trinucleotides of the common trinucleotide circular code X_0. Number $c(\mathcal{T}^i(l))$ of circular codes C, number $\mathrm{mc}(\mathcal{T}^i(l))$ of maximum circular codes MC, number $\mathrm{sc}(\mathcal{T}^i(l))$ of self-complementary circular codes SC, number $c^3(\mathcal{T}^i(l))$ of circular codes C^3, and number $\mathrm{msc}^3(\mathcal{T}^i(l))$ of C^3 self-complementary maximum circular codes MSC^3.

l	$\mathcal{T}^1(l)$ or $\mathcal{T}^3(l)$					$\mathcal{T}^2(l)$				
	c	mc	sc	c^3	msc^3	c	mc	sc	c^3	msc^3
1	12	0	0	12	0	8	0	0	8	0
2	68	6	10	68	6	28	4	8	28	4
3	240	0	0	240	0	60	0	0	56	0
4	587	15	33	586	15	92	6	18	74	6
5	1049	0	0	1040	0	108	0	0	62	0
6	1408	20	48	1372	20	96	4	8	32	4
7	1436	0	0	1352	0	60	0	0	8	0
8	1111	15	33	985	15	23	1	1	1	1
9	642	0	0	516	0	4	0	0	0	0
10	268	6	10	184	6	0	0	0	0	0
11	76	0	0	40	0	0	0	0	0	0
12	13	1	1	4	1	0	0	0	0	0
13	1	0	0	0	0	0	0	0	0	0
{14, ..., 20}	0	0	0	0	0	0	0	0	0	0

3.1.2. Transition Map $\mathcal{T}^{i,j}$

Result 2 (Table 3). For $l = 1, \ldots, 20$

$$\begin{aligned} &c\left(\mathcal{T}^{1,2}(l)\right) = c\left(\mathcal{T}^{2,3}(l)\right), \\ &\mathrm{mc}\left(\mathcal{T}^{1,2}(l)\right) = \mathrm{mc}\left(\mathcal{T}^{2,3}(l)\right), \\ &\mathrm{sc}\left(\mathcal{T}^{1,2}(l)\right) = \mathrm{sc}\left(\mathcal{T}^{2,3}(l)\right), \\ &c^3\left(\mathcal{T}^{1,2}(l)\right) = c^3\left(\mathcal{T}^{2,3}(l)\right), \\ &\mathrm{msc}^3\left(\mathcal{T}^{1,2}(l)\right) = \mathrm{msc}^3\left(\mathcal{T}^{2,3}(l)\right). \end{aligned} \tag{17}$$

The lists of trinucleotide circular codes C associated with $c(\mathcal{T}^{1,2}(l))$ and $c(\mathcal{T}^{2,3}(l))$ are different for $l = 1, \ldots, 14$ (not shown). No trinucleotide code is circular after a certain number of transitions $\mathcal{T}^{i,j}$ in the trinucleotides of the common trinucleotide circular code X_0. Precisely, for $l = 15, \ldots, 20$

$$c\left(\mathcal{T}^{1,2}(l)\right) = c\left(\mathcal{T}^{2,3}(l)\right) = 0 \tag{18}$$

and for $l = 12, \ldots, 20$

$$c\left(\mathcal{T}^{1,3}(l)\right) = 0. \tag{19}$$

The transition $\mathcal{T}^{i,j}$ generates a maximum number of trinucleotide circular codes C for

$$\begin{aligned} &\max\left\{c\left(\mathcal{T}^{i,j}(l)\right),\ i, j = 1, 2, 3,\ i < j,\ l = 1, \ldots, 20\right\} \\ &\quad = c\left(\mathcal{T}^{1,3}(6)\right) = 598 \end{aligned} \tag{20}$$

Table 3: Transition map $\mathcal{T}^{i,j}(l)$ in l trinucleotides of the common trinucleotide circular code X_0. Number $c(\mathcal{T}^{i,j}(l))$ of circular codes C, number $\mathrm{mc}(\mathcal{T}^{i,j}(l))$ of maximum circular codes MC, number $\mathrm{sc}(\mathcal{T}^{i,j}(l))$ of self-complementary circular codes SC, number $c^3(\mathcal{T}^{i,j}(l))$ of circular codes C^3, and number $\mathrm{msc}^3(\mathcal{T}^{i,j}(l))$ of C^3 self-complementary maximum circular codes MSC^3.

l	$\mathcal{T}^{1,2}(l)$ or $\mathcal{T}^{2,3}(l)$					$\mathcal{T}^{1,3}(l)$				
	c	mc	sc	c^3	msc^3	c	mc	sc	c^3	msc^3
1	9	1	0	9	0	10	2	0	10	0
2	37	4	6	36	4	49	5	9	45	5
3	91	4	0	85	0	152	8	0	122	0
4	148	6	11	130	6	325	10	27	223	10
5	166	6	0	133	0	508	12	0	292	0
6	129	4	6	90	4	598	10	30	286	10
7	67	4	0	38	0	532	8	0	214	0
8	21	1	1	9	1	353	5	11	121	5
9	4	1	0	1	0	162	2	0	44	0
10	3	0	0	0	0	42	1	2	10	1
11	4	0	0	0	0	4	0	0	0	0
12	3	0	0	0	0	0	0	0	0	0
13	4	0	0	0	0	0	0	0	0	0
14	1	0	0	1	0	0	0	0	0	0
$\{15,\ldots,20\}$	0	0	0	0	0	0	0	0	0	0

Table 4: Transition map $\mathcal{T}^{1,2,3}(l)$ in l trinucleotides of the common trinucleotide circular code X_0. Number $c(\mathcal{T}^{1,2,3}(l))$ of circular codes C, number $\mathrm{mc}(\mathcal{T}^{1,2,3}(l))$ of maximum circular codes MC, number $\mathrm{sc}(\mathcal{T}^{1,2,3}(l))$ of self-complementary circular codes SC, number $c^3(\mathcal{T}^{1,2,3}(l))$ of circular codes C^3, and number $\mathrm{msc}^3(\mathcal{T}^{1,2,3}(l))$ of C^3 self-complementary maximum circular codes MSC^3.

l	$\mathcal{T}^{1,2,3}(l)$				
	c	mc	sc	c^3	msc^3
$\{1,19\}$	8	0	0	8	0
$\{2,18\}$	28	4	8	28	4
$\{3,17\}$	56	0	0	56	0
$\{4,16\}$	71	7	19	70	6
$\{5,15\}$	64	0	0	56	0
$\{6,14\}$	56	8	16	28	4
$\{7,13\}$	64	0	0	8	0
$\{8,12\}$	72	8	20	1	1
$\{9,11\}$	64	0	0	0	0
10	56	8	16	0	0
20	1	1	1	1	1

and a maximum number of C^3 self-complementary maximum circular codes MSC^3 for

$$\max\left\{\mathrm{msc}^3\left(\mathcal{T}^{i,j}(l)\right),\ i,j=1,2,3,\ i<j,\ l=1,\ldots,20\right\} = \mathrm{msc}^3\left(\mathcal{T}^{1,3}(4)\right) = \mathrm{msc}^3\left(\mathcal{T}^{1,3}(6)\right) = 10. \tag{21}$$

The numbers $c^3(\mathcal{T}^{1,2}(l)) = c^3(\mathcal{T}^{2,3}(l))$ of circular codes C^3 have a particular growth function

$$c^3\left(\mathcal{T}^{1,2}(14)\right) = c^3\left(\mathcal{T}^{2,3}(14)\right) = 1,$$

$$c^3\left(\mathcal{T}^{1,2}(l)\right) = c^3\left(\mathcal{T}^{2,3}(l)\right) = 0 \quad \text{for } l = 10,\ldots,13. \tag{22}$$

3.1.3. Transition Map $\mathcal{T}^{1,2,3}$

Result 3 (Table 4). The transition $\mathcal{T}^{1,2,3}$ always generates trinucleotide circular codes. Indeed, for $l=1,\ldots,20$

$$c\left(\mathcal{T}^{1,2,3}(l)\right) > 0. \tag{23}$$

The lists of trinucleotide circular codes C associated with $c(\mathcal{T}^{1,2,3}(l))$ and $c(\mathcal{T}^{1,2,3}(20-l))$ are different for $l=1,\ldots,9$ (not shown). The transition $\mathcal{T}^{1,2,3}$ generates a maximum number of trinucleotide circular codes C for

$$\max\left\{c\left(\mathcal{T}^{1,2,3}(l)\right),\ l=1,\ldots,20\right\} = c\left(\mathcal{T}^{1,2,3}(8)\right) = c\left(\mathcal{T}^{1,2,3}(12)\right) = 72 \tag{24}$$

and a maximum number of C^3 self-complementary maximum circular codes MSC^3 for

$$\max\left\{\mathrm{msc}^3\left(\mathcal{T}^{1,2,3}(l)\right),\ l=1,\ldots,20\right\} = \mathrm{msc}^3\left(\mathcal{T}^{1,2,3}(4)\right) = \mathrm{msc}^3\left(\mathcal{T}^{1,2,3}(16)\right) = 6. \tag{25}$$

3.2. Transversion I Map

3.2.1. Transversion I Map $\mathcal{V}_I^i$

Result 4 (Table 5). For $l=1,\ldots,20$

$$c\left(\mathcal{V}_{\mathrm{I}}^1(l)\right) = c\left(\mathcal{V}_{\mathrm{I}}^3(l)\right),$$

$$\mathrm{mc}\left(\mathcal{V}_{\mathrm{I}}^1(l)\right) = \mathrm{mc}\left(\mathcal{V}_{\mathrm{I}}^3(l)\right),$$

$$\mathrm{sc}\left(\mathcal{V}_{\mathrm{I}}^1(l)\right) = \mathrm{sc}\left(\mathcal{V}_{\mathrm{I}}^3(l)\right),$$

$$c^3\left(\mathscr{V}_{\mathrm{I}}^1(l)\right) = c^3\left(\mathscr{V}_{\mathrm{I}}^3(l)\right),$$
$$\mathrm{msc}^3\left(\mathscr{V}_{\mathrm{I}}^1(l)\right) = \mathrm{msc}^3\left(\mathscr{V}_{\mathrm{I}}^3(l)\right). \tag{26}$$

The lists of trinucleotide circular codes C associated with $c(\mathscr{V}_{\mathrm{I}}^1(l))$ and $c(\mathscr{V}_{\mathrm{I}}^3(l))$ are different for $l = 1, \ldots, 9$ (not shown). No trinucleotide code is circular after a certain number of transversions I $\mathscr{V}_{\mathrm{I}}^i$ in the trinucleotides of the common trinucleotide circular code X_0. Precisely, for $l = 10, \ldots, 20$

$$c\left(\mathscr{V}_{\mathrm{I}}^1(l)\right) = c\left(\mathscr{V}_{\mathrm{I}}^3(l)\right) = 0 \tag{27}$$

and for $l = 18, 19, 20$

$$c\left(\mathscr{V}_{\mathrm{I}}^2(l)\right) = 0. \tag{28}$$

The transversion I $\mathscr{V}_{\mathrm{I}}^i$ generates a maximum number of trinucleotide circular codes C for

$$\max\left\{c\left(\mathscr{V}_{\mathrm{I}}^i(l)\right),\ i = 1, 2, 3,\ l = 1, \ldots, 20\right\} = c\left(\mathscr{V}_{\mathrm{I}}^2(9)\right) = 24310 \tag{29}$$

and a maximum number of C^3 self-complementary maximum circular codes MSC^3 for

$$\max\left\{\mathrm{msc}^3\left(\mathscr{V}_{\mathrm{I}}^i(l)\right),\ i = 1, 2, 3,\ l = 1, \ldots, 20\right\} = \mathrm{msc}^3\left(\mathscr{V}_{\mathrm{I}}^2(8)\right) = 70. \tag{30}$$

A remarkable code property only found with transversion I $\mathscr{V}_{\mathrm{I}}^2$ is, for $l = 1, \ldots, 20$,

$$c\left(\mathscr{V}_{\mathrm{I}}^2(l)\right) = c^3\left(\mathscr{V}_{\mathrm{I}}^2(l)\right), \tag{31}$$

and furthermore, after a detailed computer analysis, the lists of trinucleotide circular codes C and C^3 associated with $c(\mathscr{V}_{\mathrm{I}}^2(l))$ and $c^3(\mathscr{V}_{\mathrm{I}}^2(l))$, respectively, are identical for $l = 1, \ldots, 17$.

3.2.2. Transversion I Map $\mathscr{V}_I^{i,j}$

Result 5 (Table 6). For $l = 1, \ldots, 20$

$$c\left(\mathscr{V}_{\mathrm{I}}^{1,2}(l)\right) = c\left(\mathscr{V}_{\mathrm{I}}^{2,3}(l)\right),$$
$$\mathrm{mc}\left(\mathscr{V}_{\mathrm{I}}^{1,2}(l)\right) = \mathrm{mc}\left(\mathscr{V}_{\mathrm{I}}^{2,3}(l)\right),$$
$$\mathrm{sc}\left(\mathscr{V}_{\mathrm{I}}^{1,2}(l)\right) = \mathrm{sc}\left(\mathscr{V}_{\mathrm{I}}^{2,3}(l)\right), \tag{32}$$
$$c^3\left(\mathscr{V}_{\mathrm{I}}^{1,2}(l)\right) = c^3\left(\mathscr{V}_{\mathrm{I}}^{2,3}(l)\right),$$
$$\mathrm{msc}^3\left(\mathscr{V}_{\mathrm{I}}^{1,2}(l)\right) = \mathrm{msc}^3\left(\mathscr{V}_{\mathrm{I}}^{2,3}(l)\right).$$

The lists of trinucleotide circular codes C associated with $c(\mathscr{V}_{\mathrm{I}}^{1,2}(l))$ and $c(\mathscr{V}_{\mathrm{I}}^{2,3}(l))$ are different for $l = 1, \ldots, 12$ (not shown). No trinucleotide code is circular after a certain number of transversions I $\mathscr{V}_{\mathrm{I}}^{i,j}$ in the trinucleotides of the common trinucleotide circular code X_0. Precisely, for $l = 13, \ldots, 20$

$$c\left(\mathscr{V}_{\mathrm{I}}^{1,2}(l)\right) = c\left(\mathscr{V}_{\mathrm{I}}^{2,3}(l)\right) = 0 \tag{33}$$

and for $l = 19, 20$

$$c\left(\mathscr{V}_{\mathrm{I}}^{1,3}(l)\right) = 0. \tag{34}$$

The transversion I $\mathscr{V}_{\mathrm{I}}^{i,j}$ generates a maximum number of trinucleotide circular codes C for

$$\max\left\{c\left(\mathscr{V}_{\mathrm{I}}^{i,j}(l)\right),\ i, j = 1, 2, 3,\ i < j,\ l = 1, \ldots, 20\right\} = c\left(\mathscr{V}_{\mathrm{I}}^{1,2}(6)\right) = c\left(\mathscr{V}_{\mathrm{I}}^{2,3}(6)\right) = 630 \tag{35}$$

and a maximum number of C^3 self-complementary maximum circular codes MSC^3 for

$$\max\left\{\mathrm{msc}^3\left(\mathscr{V}_{\mathrm{I}}^{i,j}(l)\right),\ i, j = 1, 2, 3,\ i < j,\ l = 1, \ldots, 20\right\} = \mathrm{msc}^3\left(\mathscr{V}_{\mathrm{I}}^{1,2}(4)\right) = \mathrm{msc}^3\left(\mathscr{V}_{\mathrm{I}}^{2,3}(4)\right) = 6. \tag{36}$$

The numbers $\mathrm{sc}(\mathscr{V}_{\mathrm{I}}^{1,3}(l))$ of self-complementary circular codes SC have a particular growth function

$$\mathrm{sc}\left(\mathscr{V}_{\mathrm{I}}^{1,3}(l)\right) = 1 \quad \text{for } l = 12, 14, 16, 18,$$
$$\mathrm{sc}\left(\mathscr{V}_{\mathrm{I}}^{1,3}(l)\right) = 0 \quad \text{for } l = 8, 10. \tag{37}$$

The numbers $c^3(\mathscr{V}_{\mathrm{I}}^{1,3}(l))$ of circular codes C^3 have a particular growth function

$$c^3\left(\mathscr{V}_{\mathrm{I}}^{1,3}(l)\right) = 1 \quad \text{for } l = 16, 18,$$
$$c^3\left(\mathscr{V}_{\mathrm{I}}^{1,3}(17)\right) = 2, \tag{38}$$
$$c^3\left(\mathscr{V}_{\mathrm{I}}^{1,3}(l)\right) = 0 \quad \text{for } l = 7, \ldots, 15.$$

3.2.3. Transversion I Map $\mathscr{V}_I^{1,2,3}$

Result 6 (Table 7). The transversion I $\mathscr{V}_{\mathrm{I}}^{1,2,3}$ always generates trinucleotide circular codes. Indeed, for $l = 1, \ldots, 20$

$$c\left(\mathscr{V}_{\mathrm{I}}^{1,2,3}(l)\right) > 0. \tag{39}$$

The lists of trinucleotide circular codes C associated with $c(\mathscr{V}_{\mathrm{I}}^{1,2,3}(l))$ and $c(\mathscr{V}_{\mathrm{I}}^{1,2,3}(20-l))$ are different for $l = 1, \ldots, 9$ (not shown). The transversion I $\mathscr{V}_{\mathrm{I}}^{1,2,3}$ generates a maximum number of trinucleotide circular codes C for

$$\max\left\{c\left(\mathscr{V}_{\mathrm{I}}^{1,2,3}(l)\right),\ l = 1, \ldots, 20\right\} = c\left(\mathscr{V}_{\mathrm{I}}^{1,2,3}(10)\right) = 66 \tag{40}$$

and a maximum number of C^3 self-complementary maximum circular codes MSC^3 for

$$\max\left\{\mathrm{msc}^3\left(\mathscr{V}_{\mathrm{I}}^{1,2,3}(l)\right),\ l = 1, \ldots, 20\right\} = \mathrm{msc}^3\left(\mathscr{V}_{\mathrm{I}}^{1,2,3}(4)\right) = \mathrm{msc}^3\left(\mathscr{V}_{\mathrm{I}}^{1,2,3}(16)\right) = 9. \tag{41}$$

Table 5: Transversion I map $\mathcal{V}_{\mathrm{I}}^{i}(l)$ in l trinucleotides of the common trinucleotide circular code X_0. Number $c(\mathcal{V}_{\mathrm{I}}^{i}(l))$ of circular codes C, number $\mathrm{mc}(\mathcal{V}_{\mathrm{I}}^{i}(l))$ of maximum circular codes MC, number $\mathrm{sc}(\mathcal{V}_{\mathrm{I}}^{i}(l))$ of self-complementary circular codes SC, number $c^3(\mathcal{V}_{\mathrm{I}}^{i}(l))$ of circular codes C^3, and number $\mathrm{msc}^3(\mathcal{V}_{\mathrm{I}}^{i}(l))$ of C^3 self-complementary maximum circular codes MSC^3.

l	$\mathcal{V}_{\mathrm{I}}^{1}(l)$ or $\mathcal{V}_{\mathrm{I}}^{3}(l)$					$\mathcal{V}_{\mathrm{I}}^{2}(l)$				
	c	mc	sc	c^3	msc^3	c	mc	sc	c^3	msc^3
1	8	0	0	8	0	16	0	0	16	0
2	31	4	5	30	4	122	8	12	122	8
3	76	0	0	71	0	590	0	0	590	0
4	130	6	8	114	6	2030	28	56	2030	28
5	160	0	0	125	0	5278	0	0	5278	0
6	142	4	5	92	4	10738	56	132	10738	56
7	87	0	0	42	0	17446	0	0	17446	0
8	32	1	1	10	1	22880	70	174	22880	70
9	5	0	0	1	0	24310	0	0	24310	0
10	0	0	0	0	0	20878	56	132	20878	56
11	0	0	0	0	0	14378	0	0	14378	0
12	0	0	0	0	0	7826	28	56	7826	28
13	0	0	0	0	0	3290	0	0	3290	0
14	0	0	0	0	0	1030	8	12	1030	8
15	0	0	0	0	0	226	0	0	226	0
16	0	0	0	0	0	31	1	1	31	1
17	0	0	0	0	0	2	0	0	2	0
$\{18, \ldots, 20\}$	0	0	0	0	0	0	0	0	0	0

Table 6: Transversion I map $\mathcal{V}_{\mathrm{I}}^{i,j}(l)$ in l trinucleotides of the common trinucleotide circular code X_0. Number $c(\mathcal{V}_{\mathrm{I}}^{i,j}(l))$ of circular codes C, number $\mathrm{mc}(\mathcal{V}_{\mathrm{I}}^{i,j}(l))$ of maximum circular codes MC, number $\mathrm{sc}(\mathcal{V}_{\mathrm{I}}^{i,j}(l))$ of self-complementary circular codes SC, number $c^3(\mathcal{V}_{\mathrm{I}}^{i,j}(l))$ of circular codes C^3, and number $\mathrm{msc}^3(\mathcal{V}_{\mathrm{I}}^{i,j}(l))$ of C^3 self-complementary maximum circular codes MSC^3.

l	$\mathcal{V}_{\mathrm{I}}^{1,2}(l)$ or $\mathcal{V}_{\mathrm{I}}^{2,3}(l)$					$\mathcal{V}_{\mathrm{I}}^{1,3}(l)$				
	c	mc	sc	c^3	msc^3	c	mc	sc	c^3	msc^3
1	10	2	0	9	0	6	2	0	4	0
2	48	5	5	38	4	15	3	5	7	3
3	147	8	0	102	0	20	4	0	8	0
4	319	10	8	191	6	15	3	5	7	3
5	514	12	0	265	0	8	4	0	4	0
6	630	10	5	282	4	11	3	1	1	1
7	595	8	0	240	0	22	4	0	0	0
8	435	5	1	162	1	32	4	0	0	0
9	245	2	0	88	0	38	2	0	0	0
10	103	1	0	34	0	38	2	0	0	0
11	29	0	0	8	0	34	0	0	0	0
12	4	0	0	1	0	39	0	1	0	0
13	0	0	0	0	0	56	0	0	0	0
14	0	0	0	0	0	65	0	1	0	0
15	0	0	0	0	0	52	0	0	0	0
16	0	0	0	0	0	27	0	1	1	0
17	0	0	0	0	0	8	0	0	2	0
18	0	0	0	0	0	1	0	1	1	0
$\{19, 20\}$	0	0	0	0	0	0	0	0	0	0

TABLE 7: Transversion I map $\mathscr{V}_{\mathrm{I}}^{1,2,3}(l)$ in l trinucleotides of the common trinucleotide circular code X_0. Number $c(\mathscr{V}_{\mathrm{I}}^{1,2,3}(l))$ of circular codes C, number $\mathrm{mc}(\mathscr{V}_{\mathrm{I}}^{1,2,3}(l))$ of maximum circular codes MC, number $\mathrm{sc}(\mathscr{V}_{\mathrm{I}}^{1,2,3}(l))$ of self-complementary circular codes SC, number $c^3(\mathscr{V}_{\mathrm{I}}^{1,2,3}(l))$ of circular codes C^3, and number $\mathrm{msc}^3(\mathscr{V}_{\mathrm{I}}^{1,2,3}(l))$ of C^3 self-complementary maximum circular codes MSC^3.

l	$\mathscr{V}_{\mathrm{I}}^{1,2,3}(l)$				
	c	mc	sc	c^3	msc^3
{1, 19}	6	0	0	6	0
{2, 18}	17	5	7	17	5
{3, 17}	32	0	0	32	0
{4, 16}	47	10	17	45	9
{5, 15}	54	0	0	46	0
{6, 14}	51	12	19	31	7
{7, 13}	48	0	0	12	0
{8, 12}	53	13	11	2	2
{9, 11}	62	0	0	0	0
10	66	14	6	0	0
20	1	1	1	1	1

TABLE 8: Transversion II map $\mathscr{V}_{\mathrm{II}}^{i}(l)$ in l trinucleotides of the common trinucleotide circular code X_0. Number $c(\mathscr{V}_{\mathrm{II}}^{i}(l))$ of circular codes C, number $\mathrm{mc}(\mathscr{V}_{\mathrm{II}}^{i}(l))$ of maximum circular codes MC, number $\mathrm{sc}(\mathscr{V}_{\mathrm{II}}^{i}(l))$ of self-complementary circular codes SC, number $c^3(\mathscr{V}_{\mathrm{II}}^{i}(l))$ of circular codes C^3, and number $\mathrm{msc}^3(\mathscr{V}_{\mathrm{II}}^{i}(l))$ of C^3 self-complementary maximum circular codes MSC^3.

l	$\mathscr{V}_{\mathrm{II}}^{1}(l)$ or $\mathscr{V}_{\mathrm{II}}^{3}(l)$					$\mathscr{V}_{\mathrm{II}}^{2}(l)$				
	c	mc	sc	c^3	msc^3	c	mc	sc	c^3	msc^3
1	6	0	0	6	0	8	0	0	8	0
2	17	3	4	17	3	30	4	8	28	4
3	30	0	0	29	0	74	0	0	58	0
4	35	3	4	31	3	132	6	18	82	6
5	28	0	0	20	0	176	0	0	76	0
6	16	1	1	7	1	170	4	8	44	4
7	6	0	0	1	0	120	0	0	12	0
8	1	0	0	0	0	61	1	1	1	1
9	0	0	0	0	0	26	0	0	0	0
10	0	0	0	0	0	10	0	0	0	0
11	0	0	0	0	0	2	0	0	0	0
{12, ..., 20}	0	0	0	0	0	0	0	0	0	0

3.3. Transversion II Map

3.3.1. Transversion II Map $\mathscr{V}_{II}^{i}$

Result 7 (Table 8). For $l = 1, \ldots, 20$

$$\begin{aligned}
c\left(\mathscr{V}_{\mathrm{II}}^{1}(l)\right) &= c\left(\mathscr{V}_{\mathrm{II}}^{3}(l)\right), \\
\mathrm{mc}\left(\mathscr{V}_{\mathrm{II}}^{1}(l)\right) &= \mathrm{mc}\left(\mathscr{V}_{\mathrm{II}}^{3}(l)\right), \\
\mathrm{sc}\left(\mathscr{V}_{\mathrm{II}}^{1}(l)\right) &= \mathrm{sc}\left(\mathscr{V}_{\mathrm{II}}^{3}(l)\right), \\
c^3\left(\mathscr{V}_{\mathrm{II}}^{1}(l)\right) &= c^3\left(\mathscr{V}_{\mathrm{II}}^{3}(l)\right), \\
\mathrm{msc}^3\left(\mathscr{V}_{\mathrm{II}}^{1}(l)\right) &= \mathrm{msc}^3\left(\mathscr{V}_{\mathrm{II}}^{3}(l)\right).
\end{aligned} \tag{42}$$

The lists of trinucleotide circular codes C associated with $c(\mathscr{V}_{\mathrm{II}}^{1}(l))$ and $c(\mathscr{V}_{\mathrm{II}}^{3}(l))$ are different for $l = 1, \ldots, 8$ (not shown). No trinucleotide code is circular after a certain number of transversions II $\mathscr{V}_{\mathrm{II}}^{i}$ in the trinucleotides of the common trinucleotide circular code X_0. Precisely, for $l = 9, \ldots, 20$

$$c\left(\mathscr{V}_{\mathrm{II}}^{1}(l)\right) = c\left(\mathscr{V}_{\mathrm{II}}^{3}(l)\right) = 0 \tag{43}$$

and for $l = 12, \ldots, 20$

$$c\left(\mathscr{V}_{\mathrm{II}}^{2}(l)\right) = 0. \tag{44}$$

The transversion II $\mathscr{V}_{\mathrm{II}}^{i}$ generates a maximum number of trinucleotide circular codes C for

$$\begin{aligned}
&\max\left\{c\left(\mathscr{V}_{\mathrm{II}}^{i}(l)\right), i = 1, 2, 3, l = 1, \ldots, 20\right\} \\
&\quad = c\left(\mathscr{V}_{\mathrm{II}}^{2}(5)\right) = 176
\end{aligned} \tag{45}$$

and a maximum number of C^3 self-complementary maximum circular codes MSC^3 for

$$\begin{aligned}
&\max\left\{\mathrm{msc}^3\left(\mathscr{V}_{\mathrm{II}}^{i}(l)\right), i = 1, 2, 3, l = 1, \ldots, 20\right\} \\
&\quad = \mathrm{msc}^3\left(\mathscr{V}_{\mathrm{II}}^{2}(4)\right) = 6.
\end{aligned} \tag{46}$$

3.3.2. Transversion II Map $\mathscr{V}_{II}^{i,j}$

Result 8 (Table 9). For $l = 1, \ldots, 20$

$$\begin{aligned}
c\left(\mathscr{V}_{\mathrm{II}}^{1,2}(l)\right) &= c\left(\mathscr{V}_{\mathrm{II}}^{2,3}(l)\right), \\
\mathrm{mc}\left(\mathscr{V}_{\mathrm{II}}^{1,2}(l)\right) &= \mathrm{mc}\left(\mathscr{V}_{\mathrm{II}}^{2,3}(l)\right), \\
\mathrm{sc}\left(\mathscr{V}_{\mathrm{II}}^{1,2}(l)\right) &= \mathrm{sc}\left(\mathscr{V}_{\mathrm{II}}^{2,3}(l)\right), \\
c^3\left(\mathscr{V}_{\mathrm{II}}^{1,2}(l)\right) &= c^3\left(\mathscr{V}_{\mathrm{II}}^{2,3}(l)\right), \\
\mathrm{msc}^3\left(\mathscr{V}_{\mathrm{II}}^{1,2}(l)\right) &= \mathrm{msc}^3\left(\mathscr{V}_{\mathrm{II}}^{2,3}(l)\right).
\end{aligned} \tag{47}$$

The lists of trinucleotide circular codes C associated with $c(\mathscr{V}_{\mathrm{II}}^{1,2}(l))$ and $c(\mathscr{V}_{\mathrm{II}}^{2,3}(l))$ are different for $l = 1, \ldots, 5, 13, 14$ (not shown). The distribution of trinucleotide codes which are not circular under transversions II $\mathscr{V}_{\mathrm{II}}^{i,j}$ in the trinucleotides of the common trinucleotide circular code X_0 is very unusual. Indeed, for $l = 6, \ldots, 12, 15, \ldots, 20$

$$c\left(\mathscr{V}_{\mathrm{II}}^{1,2}(l)\right) = c\left(\mathscr{V}_{\mathrm{II}}^{2,3}(l)\right) = 0 \tag{48}$$

and for $l = 15, \ldots, 20$

$$c\left(\mathscr{V}_{\mathrm{II}}^{1,3}(l)\right) = 0. \tag{49}$$

Table 9: Transversion II map $\mathcal{V}_{\mathrm{II}}^{i,j}(l)$ in l trinucleotide of the common trinucleotide circular code X_0. Number $c(\mathcal{V}_{\mathrm{II}}^{i,j}(l))$ of circular codes C, number $\mathrm{mc}(\mathcal{V}_{\mathrm{II}}^{i,j}(l))$ of maximum circular codes MC, number $\mathrm{sc}(\mathcal{V}_{\mathrm{II}}^{i,j}(l))$ of self-complementary circular codes SC, number $c^3(\mathcal{V}_{\mathrm{II}}^{i,j}(l))$ of circular codes C^3, and number $\mathrm{msc}^3(\mathcal{V}_{\mathrm{II}}^{i,j}(l))$ of C^3 self-complementary maximum circular codes MSC^3.

l	$\mathcal{V}_{\mathrm{II}}^{1,2}(l)$ or $\mathcal{V}_{\mathrm{II}}^{2,3}(l)$					$\mathcal{V}_{\mathrm{II}}^{1,3}(l)$				
	c	mc	sc	c^3	msc^3	c	mc	sc	c^3	msc^3
1	5	1	0	5	0	8	0	0	8	0
2	11	2	3	10	2	40	4	8	30	4
3	13	2	0	11	0	132	0	0	72	0
4	8	1	1	7	1	312	6	22	117	6
5	2	1	0	2	0	532	0	0	130	0
6	0	0	0	0	0	662	4	24	92	4
7	0	0	0	0	0	592	0	0	34	0
8	0	0	0	0	0	389	1	9	4	1
9	0	0	0	0	0	188	0	0	0	0
10	0	0	0	0	0	87	0	1	2	0
11	0	0	0	0	0	54	0	0	6	0
12	0	0	0	0	0	34	0	2	7	0
13	1	0	0	0	0	14	0	0	4	0
14	1	0	0	0	0	3	0	1	3	0
$\{15,\ldots,20\}$	0	0	0	0	0	0	0	0	0	0

The transversion II $\mathcal{V}_{\mathrm{II}}^{i,j}$ generates a maximum number of trinucleotide circular codes C for

$$\max\left\{c\left(\mathcal{V}_{\mathrm{II}}^{i,j}(l)\right),\ i,j=1,2,3,\ i<j,\ l=1,\ldots,20\right\} = c\left(\mathcal{V}_{\mathrm{II}}^{1,3}(6)\right) = 662 \tag{50}$$

and a maximum number of C^3 self-complementary maximum circular codes MSC^3 for

$$\max\left\{\mathrm{msc}^3\left(\mathcal{V}_{\mathrm{II}}^{i,j}(l)\right),\ i,j=1,2,3,\ i<j,\ l=1,\ldots,20\right\} = \mathrm{msc}^3\left(\mathcal{V}_{\mathrm{II}}^{1,3}(4)\right) = 6. \tag{51}$$

The numbers $c(\mathcal{V}_{\mathrm{II}}^{1,2}(l)) = c(\mathcal{V}_{\mathrm{II}}^{2,3}(l))$ of circular codes C have a particular growth function

$$\begin{aligned} c\left(\mathcal{V}_{\mathrm{II}}^{1,2}(l)\right) &= c\left(\mathcal{V}_{\mathrm{II}}^{2,3}(l)\right) = 1 \quad \text{for } l = 13, 14, \\ c\left(\mathcal{V}_{\mathrm{II}}^{1,2}(l)\right) &= c\left(\mathcal{V}_{\mathrm{II}}^{2,3}(l)\right) = 0 \quad \text{for } l = 6,\ldots,12. \end{aligned} \tag{52}$$

3.3.3. *Transversion II Map* $\mathcal{V}_{II}^{1,2,3}$

Proposition 37. *For* $l = 1,\ldots,19$

$$c\left(\mathcal{V}_{\mathrm{II}}^{1,2,3}(l)\right) = 0 \tag{53}$$

and obviously, by letter invariance, $c(\mathcal{V}_{\mathrm{II}}^{1,2,3}(20)) = 1$ *as in Tables 4 and 7.*

Proof. The common trinucleotide circular code X_0 can be partitioned according to the maps $\mathcal{V}_{\mathrm{II}}^{1,2,3}$, $\mathcal{P}$, and $\mathcal{P}^2$ as shown in Table 10.

Let a partition $\mathbf{P}_i = \{x, x'\}$, $i \in \{1,\ldots,10\}$, composed of two trinucleotides $x, x' \in X_0$. For $l = 1$, any transversion II of a trinucleotide $x \in \mathbf{P}_i$ generates a trinucleotide y which is a permuted trinucleotide of the other trinucleotide $x' \in \mathbf{P}_i$. So, any transversion II of a trinucleotide $x \in X_0$ leads to a trinucleotide code which is not circular. For $2 \le l \le 19$, the proof needs a computer analysis of the necklace for the nontrivial cases when two transversions II occur with two trinucleotides in the same partitions. □

Remark 38. Very surprisingly, for the three maps of transition, transversions I and II, $\mathcal{T}^i(l)$, $\mathcal{V}_{\mathrm{I}}^i(l)$, and $\mathcal{V}_{\mathrm{II}}^i(l)$, $i \in \{1,2,3\}$, $\mathcal{T}^{i,j}(l)$, $\mathcal{V}_{\mathrm{I}}^{i,j}(l)$, and $\mathcal{V}_{\mathrm{II}}^{i,j}(l)$, $i,j \in \{1,2,3\}$ with $i < j$ and $\mathcal{T}^{1,2,3}(l)$ (not for $\mathcal{V}_{\mathrm{I}}^{1,2,3}(l)$ and $\mathcal{V}_{\mathrm{II}}^{1,2,3}(l)$), the numbers msc^3 of self-complementary maximum circular codes MSC^3 for the first even values of l follow a series of binomial coefficients. For $\mathcal{T}^i(l)$, $\mathcal{V}_{\mathrm{I}}^i(l)$, and $\mathcal{V}_{\mathrm{II}}^i(l)$, $i \in \{1,2,3\}$, $\mathcal{T}^{1,2}(l)$, $\mathcal{T}^{2,3}(l)$, and $\mathcal{V}_{\mathrm{II}}^{i,j}(l)$, $i,j \in \{1,2,3\}$ with $i < j$, the numbers mc of maximum circular codes MC for the first even values of l follow a series of binomial coefficients. For $\mathcal{T}^{1,2,3}(l)$, the numbers c^3 of circular codes C^3 for the values l and $(20-l)$ with $l = 1,\ldots,8$ follow a series of binomial coefficients. These binomial properties with some numbers of circular codes for the three maps of transition, transversions I and II have no combinatorial explanation so far.

4. Conclusion

A comprehensive computer analysis of transition and transversions I and II in the C^3 self-complementary maximum circular code X_0 shows some new results; in particular (i) transversion I $\mathcal{V}_{\mathrm{I}}^2(l)$ on the 2nd position of any subset of trinucleotides of X_0 generates trinucleotide circular codes

Table 10

Partition of X_0	X_0	$\mathscr{V}_{\text{II}}^{1,2,3}(X_0) = X_2$	$\mathscr{P}(X_2) = X_0'$	$\mathscr{V}_{\text{II}}^{1,2,3}(X_0') = X_1$	$\mathscr{P}^2(X_1) = X_0$
$\mathbf{P}_1 = \{AAC, ACC\}$	ACC	CAA	AAC	CCA	ACC
$\mathbf{P}_2 = \{AAT, GCC\}$	GCC	TAA	AAT	CCG	GCC
$\mathbf{P}_3 = \{ATC, GAC\}$	ATC	CGA	GAC	TCA	ATC
$\mathbf{P}_4 = \{ATT, GGC\}$	ATT	CGG	GGC	TTA	ATT
$\mathbf{P}_5 = \{CAG, TAC\}$	TAC	GCA	CAG	ACT	TAC
$\mathbf{P}_6 = \{CTC, GAA\}$	CTC	AGA	GAA	TCC	CTC
$\mathbf{P}_7 = \{CTG, GTA\}$	CTG	AGT	GTA	TGC	CTG
$\mathbf{P}_8 = \{GAG, TTC\}$	TTC	GGA	GAG	TCT	TTC
$\mathbf{P}_9 = \{GAT, GTC\}$	GTC	TGA	GAT	TCG	GTC
$\mathbf{P}_{10} = \{GGT, GTT\}$	GTT	TGG	GGT	TTG	GTT

which are always C^3 and (ii) transversion II $\mathscr{V}_{\text{II}}^{1,2,3}$ on the three positions of any subset of trinucleotides of X_0 yields no trinucleotide circular codes. In addition to the classical self-complementary (Definition 20) partition of X_0 known since 1996, a new partition of X_0 based on the transversion II map $\mathscr{V}_{\text{II}}^{1,2,3}$ (Definition 33) and the circular permutation maps $\mathscr{P}$ and $\mathscr{P}^2$ (Definition 18) is also identified here. These results here extend our theory of circular code in genes to its evolution under transition and transversion.

References

[1] 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, pp. 17–26, 2008.

[2] C. J. Michel and G. Pirillo, "Identification of all trinucleotide circular codes," *Computational Biology and Chemistry*, vol. 34, no. 2, pp. 122–125, 2010.

[3] L. Bussoli, C. J. Michel, and G. Pirillo, "On some forbidden configurations for self-complementary trinucleotide circular codes," *Journal for Algebra and Number Theory Academia*, vol. 2, pp. 223–232, 2011.

[4] L. Bussoli, C. J. Michel, and G. Pirillo, "On conjugation partitions of sets of trinucleotides," *Applied Mathematics*, vol. 3, pp. 107–112, 2012.

[5] C. J. Michel, G. Pirillo, and M. A. Pirillo, "A classiffication of 20-trinucleotide circular codes," *Information and Computation*, vol. 212, pp. 55–63, 2012.

[6] C. J. Michel, G. Pirillo, and M. A. Pirillo, "Varieties of comma-free codes," *Computers and Mathematics with Applications*, vol. 55, no. 5, pp. 989–996, 2008.

[7] C. J. Michel and G. Pirillo, "Strong trinucleotide circular codes," *International Journal of Combinatorics*, vol. 2011, Article ID 659567, 14 pages, 2011.

[8] D. G. Arquès and C. J. Michel, "A complementary circular code in the protein coding genes," *Journal of Theoretical Biology*, vol. 182, no. 1, pp. 45–58, 1996.

[9] D. L. Gonzalez, S. Giannerini, and R. Rosa, "Circular codes revisited: a statistical approach," *Journal of Theoretical Biology*, vol. 275, pp. 21–28, 2011.

[10] C. J. Michel, "Circular code motifs in transfer and 16S ribosomal RNAs: a possible translation code in genes," *Computational Biology and Chemistry*, vol. 37, pp. 24–37, 2012.

[11] F. H. C. Crick, J. S. Griffith, and L. E. Orgel, "Codes without commas," *Proceedings of the National Academy of Sciences*, vol. 43, pp. 416–421, 1957.

[12] S. W. Golomb, B. Gordon, and L. R. Welch, "Comma-free codes," *Canadian Journal of Mathematics*, vol. 10, pp. 202–209, 1958.

[13] S. W. Golomb, L. R. Welch, and M. Delbrück, "Construction and properties of comma-free codes," in *Biologiske Meddelelser*, vol. 23, Det Kongelige Danske Videnskabernes Selskab, 1958.

[14] B. Hayes, "The Invention of the Genetic Code," *American Scientist*, vol. 86, no. 1, pp. 8–14, 1998.

[15] A. J. Koch and J. Lehmann, "About a symmetry of the genetic code," *Journal of Theoretical Biology*, vol. 189, no. 2, pp. 171–174, 1997.

[16] R. Jolivet and F. Rothen, "Peculiar symmetry of DNA sequences and evidence suggesting its evolutionary origin in a primeval genetic code," in *Proceedings of the 1st European Workshop Exo-/Astrobiology*, P. Ehrenfreund, O. Angerer, and B. Battrick, Eds., ESA SP-496, pp. 173–176, May 2001.

[17] C. Nikolaou and Y. Almirantis, "Mutually symmetric and complementary triplets: differences in their use distinguish systematically between coding and non-coding genomic sequences," *Journal of Theoretical Biology*, vol. 223, no. 4, pp. 477–487, 2003.

[18] G. Pirillo, "A characterization for a set of trinucleotides to be a circular code," in *Determinism, Holism, and Complexity*, C. Pellegrini, P. Cerrai, P. Freguglia, V. Benci, and G. Israel, Eds., Kluwer, 2003.

[19] G. Pirillo and M. A. Pirillo, "Growth function of self-complementary circular codes," *Rivista di Biologia*, vol. 98, no. 1, pp. 97–110, 2005.

[20] J. L. Lassez, R. A. Rossi, and A. E. Bernal, "Crick's hypothesis revisited: the existence of a universal coding frame," in *Proceedings of the IEEE 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW '07)*, May 2007.

[21] G. Pirillo, "A hierarchy for circular codes," *RAIRO: Theoretical Informatics and Applications*, vol. 42, no. 4, pp. 717–728, 2008.

[22] M. Kimura, "Estimation of evolutionary distances between homologous nucleotide sequences," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 78, no. 1, pp. 454–458, 1981.

[23] J. Berstel and D. Perrin, *Theory of Codes*, Academic Press, London, UK, 1985.

[24] J. L. Lassez, "Circular codes and synchronization," *International Journal of Computer and Information Sciences*, vol. 5, no. 2, pp. 201–208, 1976.