PU. M. A. Vol. **23** (2012), No.2, pp. 103 – 109

A note on self-complementary trinucleotide circular codes

Lorenzo Bussoli
Dipartimento di Matematica "U. Dini"
viale Morgagni 67/A
50134 Firenze, Italy
email: bussoli@math.unifi.it

and
Giuseppe Pirillo
Consiglio Nazionale delle Ricerche
Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti"
Unità di Firenze, Dipartimento di Matematica "U. Dini"
viale Morgagni 67/A
50134 Firenze, Italy
and
Université de Marne-la-Vallée
5 boulevard Descartes
77454 Marne-la-Vallée Cedex 2, France
email: pirillo@math.unifi.it

(Received: December 1, 2011 and in revised form May 31, 2012)

Abstract. For any positive integer n we introduce the property μ_n that holds for a set X of n self-complementary couples of trinucleotides (called in this paper "forbidden configuration") when, for all $m < n$, all the sets of m self-complementary couples of trinucleotides contained in X are circular codes but X itself is not a circular code. We prove that each forbidden configuration contains at most 4 self-complementary couples of trinucleotides, i.e. the maximum integer such that the property μ_n holds is 4.

Mathematics Subject Classification(2010). 94A45.

Keywords: trinucleotide circular codes, self-complementary trinucleotide circular codes, necklace.

1 Introduction

The "forbidden configurations"often plays an important role in solving combinatorial problems. This happens, for example, for sturmian words, for symmetric group and several other combinatorial stuctures.

Now, it is well known that a subset of a circular code is a circular code. Conversely a set X that is not a circular code can have several proper subsets that are circular codes.

We point out that in this paper we consider only sets of self-complementary couples of trinucleotides. In the sequel the set of all the self-complementary couples of trinucleotides will be denoted by S (see Table 1 for a complete list of couple of S). We call *n-subset* of S any set X of n self-complementary couples of trinucleotides. Note that $\binom{30}{n}$ is the number of all the n-subsets of S.

We introduce a new property for an n-subset X of S. We say that X has property μ_n if:

i) X is not a circular code;

ii) for each m-subset Y of S, if Y is strictly contained in X, then Y is a circular code.

It is well known that $\{ATA, TAT\}$ and $\{CGC, GCG\}$, 1-subsets of S, cannot be contained in a circular code.

In the section *Results* we show that:

-the number of 2-subsets of S having property μ_2 is 44;

-the number of 3-subsets of S having property μ_3 is 64;

-the number of 4-subsets of S having property μ_4 is 153.

In this paper we prove that if X is an n-subset of S and has property μ_n then $n \leq 4$; this means that a set X of self-complementary couples of trinucleotides having property μ_n can contain at most 4 self-complementary couples of trinucleotides. Consequently, for $5 \leq n \leq 10$, no n-subset of S can have property μ_n.

2 Definitions

For the classical notions of *alphabet, empty word, length, factor, proper factor, prefix, proper prefix, suffix, proper suffix, alphabetical order*, we refer to [3]. Let $\mathcal{A}_4 = \{A, C, G, T\}$ denote the genetic alphabet, lexicographically ordered with $A < C < G < T$. We use the following notation:

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

We now recall two important genetic maps, the definitions of code and circular code [3].

DEFINITION 2.1 The complementary map $\mathcal{C}$: $\mathcal{A}_4^+ \rightarrow \mathcal{A}_4^+$ 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. $\mathcal{C}(AAC) = GTT$. This map on words is naturally extended to word sets: a complementary trinucleotide set is obtained by applying the complementary map $\mathcal{C}$ to all its trinucleotides.

DEFINITION 2.2 Code: A set X of words is a code 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 $i = 1, \ldots, n$.

DEFINITION 2.3 Circular code: A set X is a circular code 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 $i = 1, \ldots, n$.

We consider in this paper only circular codes consisting of self-complementary couples of trinucleotides, i.e. only n-subsets of S that are circular codes.

A note of caution: when we say that an n-subsets X of S is a circular code we mean that the set of the $2n$ trinucleotides belonging to the n couples of X is a circular code.

DEFINITION 2.4 A trinucleotide circular code X_0 is self-complementary if, for each $y \in X_0$, $\mathcal{C}(y) \in X_0$, i.e. if it consists only of self-complementary couples of trinucleotides.

REMARK 2.5 A self-complementary trinucleotide circular code is necessarily an n-subset of S but several n-subsets of S are not self-complementary trinucleotide circular codes.

The concept of necklace was introduced by Pirillo [9] for circular codes in order to have an algorithmic characterization of circular codes. Let $l_1, l_2, \ldots, l_{n-1}, l_n, \ldots$ be letters in $\mathcal{A}_4$, $d_1, d_2, \ldots, d_{n-1}, d_n, \ldots$ be diletters in $\mathcal{A}_4^2$ and $n \geq 2$ be an integer.

DEFINITION 2.6 Letter Diletter Continued Necklaces ($LDCN$): 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 $(n+1)LDCN$ for a subset $X \subset \mathcal{A}_4^3$ if $l_1d_1, l_2d_2, \ldots, l_nd_n \in X$ and $d_1l_2, d_2l_3, \ldots, d_{n-1}l_n, d_nl_{n+1} \in X$.

PROPOSITION 2.7 *[9]Let X be a trinucleotide circular code. The following conditions are equivalent.*
(i) X is a trinucleotide circular code.
(ii) X has no $5LDCN$.

DEFINITION 2.8 Letter Diletter Continued Closed Necklaces ($LDCCN$): 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 $(n+1)LDCCN$ for a subset $X \subset \mathcal{A}_4^3$ if $l_1d_1, l_2d_2, \ldots, l_nd_n \in X$ and $d_1l_2, d_2l_3, \ldots, d_{n-1}l_n, d_nl_{n+1} \in X$ and $l_1 = l_{n+1}$.

NOTATION 2.9 We denote a $2LDCCN$ by l_1, d_1, l_1, a $3LDCCN$ by l_1, d_1, l_2, d_2, l_1, a $4LDCCN$ by $l_1, d_1, l_2, d_2, l_3, d_3, l_1$, and a $5LDCCN$ by $l_1, d_1, l_2, d_2, l_3, d_3, l_4, d_4, l_1$.

PROPOSITION 2.10 *[8]Let X be a trinucleotide circular code. The following conditions are equivalent.*
(i) X is a trinucleotide circular code.
(ii) X has no $nLDCCN$ for any integer $n \in \{2, 3, 4, 5\}$.

Definition 2.8, Notation 2.9 and Proposition 2.10 are used in literature, in particular in [8].

The previous propositions and the following table are very useful for the computations presented in this paper.

$a = \{AAC, GTT\}$	$d = \{ACA, TGT\}$	$m = \{CAA, TTG\}$
$b = \{AAG, CTT\}$	$h = \{AGA, TCT\}$	$u = \{GAA, TTC\}$
$c = \{AAT, ATT\}$	$\{ATA, TAT\}$	$z' = \{TAA, TTA\}$
$f = \{ACG, CGT\}$	$r = \{CGA, TCG\}$	$v = \{GAC, GTC\}$
$g = \{ACT, AGT\}$	$s = \{CTA, TAG\}$	$z = \{GTA, TAC\}$
$p = \{CCA, TGG\}$	$n = \{CAC, GTG\}$	$e = \{ACC, GGT\}$
$j = \{AGG, CCT\}$	$t = \{CTC, GAG\}$	$y = \{GGA, TCC\}$
$q = \{CCG, CGG\}$	$\{CGC, GCG\}$	$x = \{GCC, GGC\}$
$l = \{ATG, CAT\}$	$k = \{ATC, GAT\}$	$z'' = \{TCA, TGA\}$
$o = \{CAG, CTG\}$	$i = \{AGC, GCT\}$	$w = \{GCA, TGC\}$

Table 1. The 30 self-complementary couples of trinucleotides divided in 10 conjugation classes.

3 Results

Immediately one realizes that each one of the 1-subsets $\{ATA, TAT\}$ and $\{CGC, GCG\}$ of S cannot be contained in a circular code; indeed

$$A, TA, T, AT, A$$

is a necklace for $\{ATA, TAT\}$ and similarly

$$C, GC, G, CG, C$$

is a necklace for $\{CGC, GCG\}$.

For this reason the names of $\{ATA, TAT\}$ and $\{CGC, GCG\}$ are omitted in Table 2.

We denote by M the set $\{\{ATA, TAT\}, \{CGC, GCG\}\}$. Note that M contains all the 1-subsets of S that cannot be contained in a circular code. So M has property μ_1 and is the first set of forbidden configurations that we consider in this paper. Note that $|M| = 2$.

We denote by:

- C the set of the 2-subsets of S that have property μ_2;
- T the set of the 3-subsets of S that have property μ_3;
- Q the set of the 4-subsets of S that have property μ_4.

PROPOSITION 3.1 *The cardinality of C is 44.*

This proposition was proved by inspection and after also verified by computer calculus. The complete list of the elements of C with the associated necklaces is given in the following Table 2. We implemented a function in environment Matlab that inspected all the 2-subsets of S that form a $5LDCN$ and consequently are not a circular code. Avoiding the 2-subsets of S that contain an element of M we obtain the set C. It consists of 2-subsets of S that have property μ_2 and it verifies $|C| = 44$.

PROPOSITION 3.2 *The cardinality of T is 64.*

This proposition was proved by computer calculus (again with a function in environment Matlab). The complete list of the elements of T with the associated necklaces is given in the following Table 3.

PROPOSITION 3.3 *The cardinality of Q is 153.*

Also this proposition was proved by computer calculus (and again with a function in environment Matlab). The complete list of the elements of Q with the associated necklaces is given in the following Table 4.

2-subsets	Minimal necklace	2-subsets	Minimal necklace
a,d	A,AC,A	i,s	A,GC,T,AG,C,TA,G,CT,A
a,m	C,AA,C	i,w	A,GC,A
b,h	A,AG,A	i,z'	A,GC,T,TA,A
b,u	C,TT,C	j,t	C,CT,C
c,r	A,AT,T,CG,A	j,y	A,GG,A
c,w	A,AT,T,GC,A	k,l	G,AT,G
c,z'	A,TT,A	k,q	C,CG,G,AT,C
d,m	A,CA,A	k,r	A,TC,G,AT,C,GA,T,CG,A
d,n	A,CA,C,AC,A	k,s	C,TA,G,AT,C
e,n	C,AC,C	k,z''	A,TC,A
e,p	A,CC,A	l,w	A,TG,C,AT,G,CA,T,GC,A
f,r	A,CG,A	l,x	C,AT,G,GC,C
f,v	C,GT,C	l,z	C,AT,G,TA,C
f,w	A,CG,T,GC,A	l,z''	A,TG,A
f,z	A,CG,T,AC,G,TA,C,GT,A	n,p	C,CA,C
f,z'	A,CG,T,TA,A	o,w	C,TG,C
g,s	A,CT,A	q,x	C,GG,C
g,z	A,GT,A	q,z	C,CG,G,TA,C
h,t	A,GA,G,AG,A	r,v	C,GA,C
h,u	A,GA,A	s,x	C,TA,G,GC,C
i,o	C,AG,C	s,z	C,TA,C
i,r	A,GC,T,CG,A	t,y	C,TC,C

Table 2. The 2-subsets of S which are not a circular code.

3-subsets	Minimal necklace	3-subsets	Minimal necklace
a,g,r	A,AC,T,CG,A	f,n,x	C,AC,GGC,C
a,g,w	A,AC,T,GC,A	f,n,z”	A,CG,T,CA,C,GT,G,TG,A
a,g,z’	A,AC,T,TA,A	f,p,y	A,CG,T,CC,A
a,l,u	C,AT,G,AA,C	f,u,z”	A,CG,T,TC,A
a,q,u	C,CG,G,AA,C	g,n,w	A,GT,G,CA,C,AC,T,GC,A
a,s,u	C,TA,G,AA,C	g,n,z”	A,GT,G,TG,A
b,g,r	A,AG,T,CG,A	g,r,t	A,CT,C,GA,G,AG,T,CG,A
b,g,w	A,AG,T,GC,A	g,t,z”	A,CT,C,,TC,A
b,g,z’	A,AG,T,TA,A	h,k,o	A,GA,T,CT,G,AT,C,AG,A
b,k,m	C,AA,G,AT,C	h,k,w	A,GA,T,GC,A
b,m,x	C,AA,G,GC,C	h,k,z’	A,GA,T,TA,A
b,m,z	C,AA,G,TA,C	h,o,v	A,GA,C,AG,A
c,d,z	A,AT,T,AC,A	h,s,v	A,GA,C,TA,G,TC,T,AG,A
c,h,s	A,AT,T,AG,A	i,l,t	C,AT,G,AG,C
c,m,z”	A,AT,T,CA,A	i,m,z”	A,GC,T,TG,A
c,p,y	A,AT,T,CC,A	i,p,y	A,GC,T,CC,A
c,u,z”	A,AT,T,GA,A	i,q,t	C,CG,G,AG,C
d,i,z	A,GC,T,GT,A	i,t,z”	A,GC,T,GA,G,CT,C,TC,A
d,l,r	A,CA,T,CG,A	i,u,z”	A,GC,T,CA,A
d,l,v	A,CA,T,GT,C,AT,G,AC,A	j,k,o	C,AG,G,AT,C
d,l,z’	A,CA,T,TA,A	j,o,x	C,AG,G,GC,C
d,o,v	A,CA,G,AC,A	j,o,z	C,AG,G,TA,C
d,o,z	A,CA,G,TA,C,TG,T,AC,A	k,o,p	C,TG,G,AT,C
e,j,r	A,CC,T,CG,A	l,v,y	C,AT,G,TC,C
e,j,w	A,CC,T,GC,A	n,q,w	C,CG,G,CA,C
e,j,z’	A,CC,T,TA,A	n,s,w	C,TA,G,CA,C
e,l,v	C,AT,G,AC,C	o,p,x	C,CA,G,GC,C
e,q,v	C,CG,G,AC,C	o,p,z	C,CA,G,TA,C
e,s,v	C,TA,G,AC,C	q,v,y	C,CG,G,GA,C
f,h,s	A,CG,T,AG,A	r,t,x	C,TC,G,GC,C
f,k,n	C,AC,G,AT,C	r,t,z	C,GA,G,TA,C
f,m,z”	A,CG,T,TG,A	s,v,y	C,TA,G,GA,C

Table 3. The 3-subsets of S which are not a circular code.

REMARK 3.4 The number 263 in the statement of the next theorem is the sum of the cardinalities of the sets of forbidden configurations:

- $M = \{\{ATA, TAT\}, \{CGC, GCG\}\}$;
- C the set of 44 elements (Proposition 3.1);
- T the set of 64 elements (Proposition 3.2);
- Q the set of 153 elements (Proposition 3.3).

Now we have the following

THEOREM 3.5 *Let X be an n-subset of S with $n \geq 5$. If X is not a circular code then X contains one of the* 263 *forbidden configurations.*

Proof. We implemented a function in environment Matlab that inspected all the possible 5-subsets, 6-subsets, 7-subsets and 8-subsets of S. We point out that for 9-subsets and 10-subsets of S the computer calculus is unuseful because, as to form a $5LDCN$, we need eight trinucleotides then at most an 8-subset of S. □

4-subsets	Minimal necklace	4-subsets	Minimal necklace	4-subsets	Minimal necklace
a,b,e,z'	A,AC,C,TT,A	b,m,v,y	C,AA,G,TC,C	f,m,p,x	A,CG,T,TG,G,CC,A
a,b,j,z'	A,AG,G,TT,A	c,d,q,v	A,AT,T,GT,C,CG,G,AC,A	f,n,u,y	A,CG,T,TC,C,AC,G,GA,A
a,e,h,o	A,AC,C,AG,A	c,d,s,v	A,AT,T,GT,C,TA,G,AC,A	f,o,x,z”	A,CG,T,CA,G,GC,C,TG,A
a,e,h,s	A,AC,C,TA,G,GT,T,CT,A	c,e,m,z	A,TT,G,GT,A	f,q,u,y	A,CG,G,GA,A
a,e,j,s	A,AC,C,CT,A	c,h,o,x	A,AT,T,CT,G,GC,C,AG,A	f,s,u,y	A,CG,T,TC,C,TA,G,GA,A
a,e,l,r	A,AC,C,AT,G,GT,T,CG,A	c,h,o,z	A,AT,T,CT,G,TA,C,AG,A	f,t,x,z”	A,CG,T,GA,G,GC,C,TC,A
a,e,l,z'	A,AC,C,AT,G,TT,A	c,j,s,u	A,TT,C,CT,A	g,h,o,x	A,CT,G,GC,C,AG,A
a,e,o,z”	A,AC,C,TG,A	c,k,m,p	A,AT,C,CA,A	g,k,n,z'	A,GT,G,AT,C,AC,T,TA,A
a,e,q,r	A,AC,C,CG,A	c,l,u,y	A,AT,G,GA,A	g,k,o,z'	A,CT,G,AT,C,AG,T,TA,A
a,e,q,w	A,CC,G,GT,T,GC,A	c,m,p,u	A,TT,C,CA,A	g,l,t,z'	A,CT,C,AT,G,AG,T,TA,A
a,e,q,y	A,AC,C,GG,A	c,m,p,x	A,TT,G,CC,A	g,l,v,z'	A,GT,C,AT,G,AC,T,TA,A
a,e,q,z'	A,CC,G,TT,A	c,m,p,z	A,TT,G,TA,C,CA,A	g,m,p,t	A,CT,C,CA,A
a,e,r,s	A,AC,TA,G,GT,T,CG,A	c,m,u,y	A,TT,G,GA,A	g,m,p,v	A,GT,C,CA,A
a,e,r,t	A,AC,C,GA,G,GT,T,CG,A	c,n,o,z”	A,AT,T,CA,C,TG,A	g,n,r,x	A,GT,G,GC,C,AC,T,CG,A
a,e,r,u	A,AC,C,CG,A	c,n,q,z”	A,AT,T,CA,C,CG,G,TG,A	g,n,t,z'	A,CT,C,AC,T,TA,A
a,e,s,w	A,AC,C,TA,G,GT,T,GC,A	c,n,s,z”	A,AT,T,CA,C,TA,G,TG,A	g,n,u,y	A,GT,G,GA,A
a,e,s,z'	A,AC,C,TA,A	c,n,t,z”	A,AT,T,GA,G,TG,A	g,n,v,z'	A,GT,C,AC,T,TA,A
a,e,t,z”	A,AC,C,TC,A	c,o,v,z”	A,AT,T,CA,GTC,A	g,n,x,z'	A,GT,G,GC,C,AC,T,TA,A
a,f,n,u	C,AC,G,AA,C	c,o,x,z”	A,AT,T,CA,G,GC,C,TG,A	g,o,r,x	A,CT,G,GC,C,AG,T,CG,A
a,f,p,x	A,AC,G,CC,A	c,o,z,z”	A,AT,T,CA,G,TA,C,TG,A	g,o,t,z'	A,CT,G,AG,T,TA,A
a,f,u,y	A,AC,G,GA,A	c,q,u,y	A,TT,C,GG,A	g,o,u,y	A,CT,G,GA,A
a,g,p,y	A,AC,T,CC,A	c,q,v,z”	A,AT,T,GA,C,CG,G,TC,A	g,o,v,z'	A,CT,G,AC,T,TA,A
a,g,u,z”	A,AC,T,TC,A	c,s,u,y	A,TT,C,TA,G,GA,A	g,o,v,z”	A,CT,G,TC,A
a,j,o,u	C,AG,G,AA,C	c,s,v,z”	A,AT,T,GA,C,TA,G,TC,A	g,o,x,z'	A,CT,G,GC,C,AG,T,TA,A
a,l,p,z'	A,TG,G,TT,A	c,t,v,z”	A,AT,T,GA,G,TC,A	g,o,x,z”	A,CT,G,GC,C,TG,A
a,o,p,u	C,CA,G,AA,C	c,t,x,z”	A,AT,T,GA,G,GC,C,TC,A	g,q,t,w	A,CT,C,CG,G,AG,T,GC,A
a,r,t,u	C,GA,G,AA,C	c,t,z,z”	A,AT,T,GA,G,TA,C,TC,A	g,q,t,z'	A,CT,C,CG,G,AG,T,TA,A
b,d,j,v	A,AG,G,AC,A	d,e,j,z	A,CC,T,GT,A	g,q,v,w	A,GT,C,CG,G,AC,T,GC,A
b,d,j,z	A,AG,G,TA,C,CT,T,GT,A	d,g,q,v	A,GT,C,CG,G,AC,A	g,q,v,z'	A,GT,C,CG,G,AC,T,TA,A
b,e,j,z	A,AG,G,GT,A	d,h,k,z	A,GA,T,GT,A	g,q,v,z”	A,GT,C,CG,G,TC,A
b,e,m,v	C,AA,G,GT,C	d,h,l,s	A,CA,T,AG,A	h,k,m,p	A,GA,T,TG,G,AT,C,CA,A
b,g,m,z”	A,AG,T,TG,A	d,h,o,x	A,CA,G,GC,C,AG,A	h,k,p,y	A,GA,T,GG,A
b,g,p,y	A,AG,TGG,A	d,h,q,v	A,GA,C,CG,G,AC,A	h,l,v,z'	A,GA,C,AT,G,TC,T,TA,A
b,i,m,p	A,AG,C,CA,A	d,i,q,v	A,GC,T,GT,C,CG,G,AC,A	h,m,p,v	A,GA,C,CA,A
b,i,m,t	C,AA,G,AG,C	d,k,o,z'	A,CA,G,AT,C,TG,T,TA,A	h,q,v,w	A,GA,C,CG,G,TC,T,GC,A
b,i,q,y	A,AC,C,GG,A	d,l,p,y	A,CA,T,GG,A	h,q,v,z'	A,GA,C,CG,G,TC,T,TA,A
b,j,k,w	A,AG,G,AT,C,CT,T,GC	d,l,u,y	A,CA,T,TC,C,AT,G,GA,A	h,q,v,z”	A,GA,C,CG,G,TC,A
b,j,k,z'	A,AG,G,AT,C,TT,A	d,o,r,x	A,CA,G,GC,C,TG,T,CG,A	i,k,m,p	A,GC,T,TG,G,AT,C,CA,A
b,j,m,w	A,AG,G,CA,A	d,o,u,y	A,CA,G,GA,A	i,l,u,y	A,GC,T,TC,C,AT,G,GA,A
b,j,n,w	A,AG,G,CA,C,CT,T,GC,A	d,o,x,z'	A,CA,G,GC,C,TG,T,TA,A	i,m,p,t	A,GC,T,TG,G,AG,C,CA,A
b,j,n,z”	A,AG,G,TG,A	d,o,x,z”	A,CA,G,GC,C,TG,A	i,m,p,x	A,GC,C,CA,A
b,j,p,x	A,AG,G,CC,A	e,h,j,s	A,CC,T,CT,A	i,m,p,z	A,GC,T,TG,G,TA,C,CA,A
b,j,r,x	A,AG,G,CC,T,CG,A	e,j,m,z”	A,CC,T,TG,A	i,n,q,z”	A,GC,T,CA,C,CG,G,TG,A
b,j,r,z	A,AG,G,TA,C,CT,T,CG,A	e,j,o,v	C,AG,G,AC,C	i,q,u,y	A,GC,T,TC,C,GG,A
b,j,v,z”	A,AG,G,TC,A	e,j,u,z”	A,CC,T,TC,A	i,q,v,z”	A,GC,T,GA,C,CG,G,TC,A
b,j,w,x	A,AG,G,GC,A	e,m,q,w	A,CC,G,CA,A	j,m,p,x	A,GG,C,CA,A
b,j,w,z	A,AG,G,TA,C,CT,T,GC,A	e,q,u,y	A,CC,G,GA,A	j,r,u,x	A,GG,C,GA,A
b,j,x,z'	A,GG,C,TT,A	f,h,o,x	A,CG,T,CT,G,GC,C,AG,A	k,m,p,y	A,TC,CCA,A
b,j,z,z'	A,AG,G,TA,A	f,i,n,t	C,AC,G,CT,C	l,p,u,y	A,TG,G,GA,A
b,k,y,z'	A,TC,C,TT,A	f,k,m,p	A,CG,T,TG,G,AT,C,CA,A	n,r,t,w	C,GA,G,TG,C
b,m,n,w	C,AA,G,TG,C	f,l,u,y	A,CG,T,TC,C,AT,G,GA,A	o,p,v,y	C,CA,G,GA,C

Table 4. The 4-subsets of S which are not a circular code.

Acknowledgement. The second author thanks the Dipartimento di Matematica “U. Dini” for giving him a friendly hospitality.

References

[1] D. G. Arquès and C. J. Michel, *A complementary circular code in the protein coding genes*, J. Theoret. Biol., 182 (1996) 45–58.

[2] D. G. Arquès and C. J. Michel, *A circular code in the protein coding genes of mitochondria*, J. Theoret. Biol., 189 (1997) 273–290.

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

[4] G. Frey and C. J. Michel, *Circular codes in archaeal genomes*, J. Theoret. Biol., 223 (2003) 413–431.

[5] 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*, Comput. Biol. Chem., 30 (2006) 87–101.

[6] C. J. Michel, G. Pirillo and M. A. Pirillo, *Varieties of comma-free codes*, Comput. Math. Appl., 55 (2008) 989–996.

[7] C. J. Michel, G. Pirillo and M. A. Pirillo, *A relation between trinucleotide comma-free codes and trinucleotide circular codes*, Theoret. Comput. Sci., 401 (2008) 17–26.

[8] C. J. Michel and G. Pirillo, *Identification of all trinucleotide circular codes*, Comput. Biol. Chem., 34 (2010) 122–125.

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

[10] G. Pirillo, *A hierarchy for circular codes*, RAIRO Theor. Inform. Appl., 42 (2008) 717–728.

[11] G. Pirillo, *Some remarks on prefix and suffix codes*, Pure Math. Appl. (PU.M.A.), 19 (2008) 53–59.

[12] G. Pirillo, *Non sharing border codes*, Adv. in Appl. Math., 3 (2010) 215–223.

[13] G. Pirillo and M. A. Pirillo, *Growth function of self-complementary circular codes*, Biology Forum, 98 (2005) 97–110.