

A New Model of Permutation the Pieces of Nucleotides in DNA Sequences Using the Action of Dihedral Group and Graph Theory

Husein Hadi Abbass¹, Hussein Salman Mohammed Hussein²

Department of Mathematics, Faculty of Education for Girls, University of Kufa, Iraq

Abstract: In this paper, we give a new model of genetic algorithm using the action of largest subgroup Hof dihedral Group D_{3^m} , $n = 3^m$, $m \in \mathbb{N}$, $m \geq 2$, bipartite graph, and a markov basis for $\frac{n^2-3n}{3} \times 3 \times \frac{n}{3}$ - contingency tables with fixed two dimensional marginals such that Bis H-invariant.

Keywords: Computational algebraic statistics, sufficient statistics, linear transformation, connected graph, bipartite graph, dihedral group, algebra statistic, Markov basis.

1. Introduction

Let I be a finite set $n = |I|$ elements, we call an element of I a cell and denoted by $i \in I$. i is often multi-index $i = i_1 \dots i_m$. A non-negative integer $x_i \in \mathbb{N}$ denotes the frequency of a cell i . The set of frequencies is called a **contingency table** and denoted as $x = \{x_i\}_{i \in I}$, with an appropriate ordering of the cell, we treat a contingency table $x = \{x_i\}_{i \in I} \in \mathbb{N}^n$ as a n -dimensional column vector of non-negative integers. Not that a contingency table can also be considered as a function from I to \mathbb{N} defined as $i \mapsto x_i$. The L_1 -norm of $x \in \mathbb{N}^n$ is called the sample size and denoted as $|x| = \sum_{i \in I} x_i$. We will denote \mathbb{Z} be the set of integer numbers, also we denote to the $a_j \in \mathbb{Z}^v, j = 1, \dots, v$, as fixed column vectors consisting of integers. A v -dimensional column vector $t = (t_1, \dots, t_v)' \in \mathbb{Z}^v$ as $t_j = a_j' x, j = 1, \dots, v$. Here $'$ denotes the transpose of a vector or matrix. We also define a $v \times p$ matrix A , with its j -row being a_j' given by $A = \begin{bmatrix} a_1' \\ \vdots \\ a_v' \end{bmatrix}$, and if $t = Ax$ is a v -

dimensional column vector, we define the set $T = \{t: t = Ax, x \in \mathbb{N}^n\} = AN^n \subset \mathbb{Z}^v$. In typical situations of a statistical theory, t is **sufficient statistic** for the nuisance parameter. The set of x 's for a given t , $A^{-1}[t] = \{x \in \mathbb{N}^n: Ax = t\}$ (t -fibers), is considered for performing **similar tests**, for the case of the independence model of two-way contingency tables, for example, t is the row sums and column sums of x , and $A^{-1}[t]$ is the set of x 's with the same row sums and column sums to t . The set of t -fibers gives a decomposition of \mathbb{N}^n . An important observation is that t -fiber depends on given only through its kernel, $\ker(A)$. For different A 's with the same kernel, the set of t -fibers are the same. In fact, if we define $x_1 \sim x_2 \Leftrightarrow x_1 - x_2 \in \ker(A)$, this relation is an equivalence relation and \mathbb{N}^n is partitioned into disjoint equivalence classes. The set of t -fibers is simply the set of these

equivalence classes. Furthermore, t may be considered as labels of these equivalence classes. A n -dimensional column vector of integers $z = \{z_i\}_{i \in I} \in \mathbb{Z}^n$ is called a move if it is in the kernel of A , i.e. $Az = 0$ [10].

For a move z , the positive part $z^+ = \{z^+_i\}_{i \in I}$ and the negative part $z^- = \{z^-_i\}_{i \in I}$ are defined by $z^+_i = \max(z_i, 0), z^-_i = \max(-z_i, 0)$, respectively, Then $z = z^+ - z^-$ and $z^+, z^- \in \mathbb{N}^n, z^+, z^- \in \mathbb{N}^n$. Moreover, z^+ and z^- are in the same t -fiber, i.e., $z^+, z^- \in A^{-1}[t]$ for $t = Az^+ = Az^-$. We define the **degree** of z as the **sample size** of z^+ or (z^-) and denote it by $\deg(z) = |z^+| = |z^-|$. In the following we denote the set of moves (for a given A) by $M = M_A = \mathbb{Z}^n \cap \ker(A)$ [1].

Let G be a group and W be a set. A left action of G in W is a function from $G \times W$ into W , usually denote by $(g, w) \rightarrow gw \in W$ such that $g(hw) = (gh)w$ and $ew = w$ for all $g, h \in G$ and $w \in W$ where e is the identity element of G . We also say that G acts on W on the left.

Let a group G act on a set W , and $U \subseteq W$, $G_{(U)} = \{g: gu = u, \forall u \in U\}$ is called the pointwise stabilizer of U . Let a group G acts on a set W , $U \subseteq W$, and $GU = \{gu: u \in U, g \in G\}$. We call U invariant under G (or G -invariant) if $GU = U$ [9].

Let $A: \mathbb{Z}^n \rightarrow \mathbb{Z}^v$ be a linear transformation, $t \in \mathbb{Z}^v$, and $A^{-1}[t]$ be the set of t -fibers, and let $B \subset \ker_2(A)$, then we define $A^{-1}[t]_B$ be the graph with vertex set $A^{-1}[t]$ and $u - v$ an edge if and only if $u - v \in \pm B$ [12]. Let $A^{-1}[t] = \{x \in \mathbb{N}^n: Ax = t\}$. A set of finite moves B is called **Markov basis** if for all t , $A^{-1}[t]$ constitutes one B equivalence class, if $B \subseteq \ker_2(A)$ is a set such that $A^{-1}[t]_B$ is connected for all t , then B is a Markov basis for A [1].

If a group G act on $A^{-1}[t]$ on the left, B is a Markov basis, and $G(B) = \{gz : z \in B, g \in G\}$, B is called invariant under G (or G -invariant) if $G(B) = B$. We will denote to the polynomials in the p indeterminates (polynomial variables) p_1, p_2, \dots, p_p over the complex field \mathbb{C} by either $\mathbb{C}[p_1, p_2, \dots, p_p]$ or $\mathbb{C}[P]$, $P = (p_1, p_2, \dots, p_p)$. Let $A: \mathbb{Z}^n \rightarrow \mathbb{Z}^d$ be a linear transformation, the toric ideal I_A is the ideal $\langle P^u - P^v : u, v \in \mathbb{N}^n, A(u) = A(v) \rangle \subseteq \mathbb{C}[P_1, \dots, P_p]$ where $P^u = P_1^{u_1} P_2^{u_2} \dots P_p^{u_p}$ [11].

In [4] H. H. Abbass and H. S. Mohammed Hussein found a Markov basis B and toric ideals for $\frac{n^2-3n}{3} \times 3 \times \frac{n}{3}$ -contingency tables with fixed two dimensional marginals, n is a multiple of 3 greater than or equal 6, also they [5] found the largest subgroup H of dihedral Group D_{3^m} , $m \in \mathbb{N}$, such that B is H -invariant.

In this paper, we use the Markov basis B and action of the subgroup H of dihedral Group D_n on these contingency tables to give a new model of permutation the pieces of nucleotides in DNA sequences.

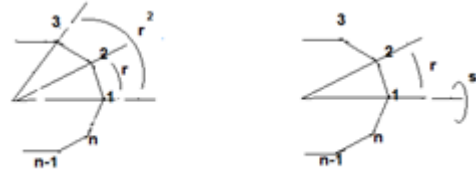
2. Preliminaries

In this section, we review some basic definitions and notations of dihedral group, connected graph, bipartite graph, moves, Markov basis, and toric ideals that we need in our work.

Definition 1 (see [12]). Let n be a positive integer greater than or equal 3. The group of all symmetries of the regular polygon with n sides, including both rotations and reflections, is called **dihedral group** and denoted by D_n . If we center the regular polygon at origin then the elements of the dihedral group acts as linear transformation of the plane. Let us represent the elements of D_n as matrix, with composition multiplication. Dihedral groups are among simplest examples of finite groups and they play an important role in group theory, geometry, and chemistry. The set of rotations is generated by r -counterclockwise rotation with angle $2\pi/n$ of order n , and the set of reflections is of order 2 and every element sr^j generates $\{e, sr^j\}$, where e is the identity element in D_n . The $2n$ elements in D_n can be written as: $\{e, r, r^2, \dots, r^{n-1}, s, sr, sr^2, \dots, sr^{n-1}\}$. In general, we can write D_n as: $D_n = \{s^j r^k : 0 \leq k \leq n-1, 0 \leq j \leq 1\}$ which has the following properties:
 $r^n = 1, sr^k s = r^{-k}$,
 $(sr^k)^2 = 1$, for all $0 \leq k \leq n-1$. The composition of two elements of the D_n is given by $r^i r^j = r^{i+j}$, $r^i s r^j = sr^{j-i}$, $sr^i r^j = sr^{i+j}$, $sr^i s r^j = r^{i-j}$.

Remark 2 (see [12]). If we label the vertices (of the regular n -gon) 1 to n in a counterclockwise direction around n -gon then the elements of D_n can be written as permutations of

vertices, let r be a counterclockwise rotation, and let s be the reflection of the n -gon about an axis through the center and vertex 1, as indicated in below. The element r generates the cyclic group of order n C_n which is a normal cyclic subgroup of D_n . In all cases, addition and subtraction should be performed using modular arithmetic with modulus n .



Elements of C_n

Elements of D_n

Any symmetry will fix the origin and is determined by the image of two adjacent vertices, say 1 and 2. The vertex 1 can be taken to any of n vertices and then the vertex 2 must be taken to one of the two vertices adjacent to the image of 1. Hence, D_n is a non abelian group of order $2n$ generated by r and s .

Now, we give some concepts about the action of a group on a set that we use later.

Definition 3 (see [3]). A graph G is connected if for every pair of distinct vertices $u, v \in V(G)$, where $V(G)$ be the set of vertices of the graph G , the graph G has a u, v -path. Otherwise, we say the graph is disconnected.

Definition 4 (see [3]). A graph G is a bipartite graph if there are $X, Y \subseteq V(G)$ meeting the following conditions:

1. $V(G) = X \cup Y$,
2. $X \cap Y = \emptyset$,
3. $G[X]$ and $G[Y]$ are both null graphs, where $G[X]$ and $G[Y]$ are subgraphs of the graph G induced by the set of vertices $X, Y \subseteq V(G)$ respectively.

Theorem 5 (see [3]).

For a graph G the following statements are equivalent:

1. G is bipartite.
2. Every cycle in G has an even length.

Definition 6 (see [3]). Let $B \subset M_A$ be the set of moves and let $x_1, x_2 \in A^{-1}[t]$. We say that x_2 accessible from x_1 by B if there exists a sequence of moves $z_1, \dots, z_k \in B$ and $\varepsilon_k \in \{-1, 1\}$, $k = 1, \dots, K$, such that

$$x_2 = x_1 + \sum_{k=1}^K \varepsilon_k z_k,$$

$$x_1 + \sum_{k=1}^K \varepsilon_k z_k \in A^{-1}[t] \text{ for } 1 \leq k \leq K.$$

Remark 7 (see [4]). Let n be a multiple of 3 such that $n \geq 6$, and let $x_j \in A^{-1}[t]$, $j = 1, \dots, k$ be the representative elements of the set of $3 \times \frac{n}{3}$ -contingency tables and $B = \{z_1, z_2, \dots, z_k\}$ such that each z_j

$j = 1, 2, \dots, k$, is a matrix of dimension $3 \times \frac{n}{3}$ either has two columns $(1, -1, 0)'$, $(-1, 1, 0)'$ $(1, 0, -1)'$, $(-1, 0, 1)'$ or either $(0, 1, -1)'$, $(0, -1, 1)'$ and the other columns are zero denoted by $+z_j$, or it two columns $(1, -1, 0)'$, $(-1, 1, 0)'$ $(1, 0, -1)'$, $(-1, 0, 1)'$ or either $(0, 1, -1)'$, $(0, -1, 1)'$ and the other columns are zero denoted by $+z_j$, or it has two columns $(-1, 0, 1)'$, $(1, 0, -1)'$ $(-1, 0, 1)'$, $(1, 0, -1)'$ or $(0, -1, 1)'$, $(0, 1, -1)'$ and the other columns are zero denoted by $-z_j$, like

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ -1 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ -1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix}$$

Also, we can write all elements of **B** as one-dimensional column vectors as follows:

$z_j = (z_1, \dots, z_n)'$, $j = 1, \dots, k$ and $z_t = 1$ or -1 or 0 such that

If $t = 1, 2, \dots, \frac{n}{3}$

$$z_t = \begin{cases} 1 & \text{if } z_{t+\frac{n}{3}} + z_{t+\frac{2n}{3}} = -1 \text{ and } \sum_{i=1, i \neq t}^{\frac{n}{3}} z_i = -1 \\ -1 & \text{if } z_{t+\frac{n}{3}} + z_{t+\frac{2n}{3}} = 1 \text{ and } \sum_{i=1, i \neq t}^{\frac{n}{3}} z_i = 1 \\ 0 & \text{if } z_{t+\frac{n}{3}} + z_{t+\frac{2n}{3}} = 0 \text{ and } \sum_{i=1, i \neq t}^{\frac{n}{3}} z_i = 0 \end{cases}$$

$$\sum_{i=1}^{\frac{n}{3}} x_i, \sum_{i=\frac{n}{3}+1}^{\frac{2n}{3}} x_i, \sum_{i=\frac{2n}{3}+1}^n x_i, x_1 + x_{\frac{n}{3}+1} + x_{\frac{2n}{3}+1}, x_2 + x_{\frac{n}{3}+2} + x_{\frac{2n}{3}+2}, \dots, x_n + x_{\frac{2n}{3}} + x_n$$

respectively. The entry in the column indexed by x_i in the matrix **A** will be equal to one, if x_i appears in the index of its row, and otherwise it will be zero. Then

$$A = \begin{bmatrix} 1 & 1 & \dots & 1 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & 1 & \dots & 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 1 & \dots & 1 & 1 \\ 1 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}_{\frac{n+9}{3} \times n}$$

Theorem 10 (see [4]). The set $\mathbf{B} = \{z_1, \dots, z_{\frac{n^2-3n}{3}}\}$ is a set of moves.

Corollary 11 (see [4]). The set **B** of moves in theorem 10 is a Markov basis.

If $t = \frac{n}{3} + 1, \frac{n}{3} + 2, \dots, \frac{2n}{3}$

$$z_t = \begin{cases} 1 & \text{if } z_{t-\frac{n}{3}} + z_{t+\frac{n}{3}} = -1 \text{ and } \sum_{i=\frac{n}{3}+1, i \neq t}^{\frac{2n}{3}} z_i = -1 \\ -1 & \text{if } z_{t-\frac{n}{3}} + z_{t+\frac{n}{3}} = 1 \text{ and } \sum_{i=\frac{n}{3}+1, i \neq t}^{\frac{2n}{3}} z_i = 1 \\ 0 & \text{if } z_{t-\frac{n}{3}} + z_{t+\frac{n}{3}} = 0 \text{ and } \sum_{i=\frac{n}{3}+1, i \neq t}^{\frac{2n}{3}} z_i = 0 \end{cases}$$

If $t = \frac{2n}{3} + 1, \frac{2n}{3} + 2, \dots, n$

$$z_t = \begin{cases} 1 & \text{if } z_{t-\frac{2n}{3}} + z_{t-\frac{n}{3}} = -1 \text{ and } \sum_{i=\frac{2n}{3}+1, i \neq t}^n z_i = -1 \\ -1 & \text{if } z_{t-\frac{2n}{3}} + z_{t-\frac{n}{3}} = 1 \text{ and } \sum_{i=\frac{2n}{3}+1, i \neq t}^n z_i = 1 \\ 0 & \text{if } z_{t-\frac{2n}{3}} + z_{t-\frac{n}{3}} = 0 \text{ and } \sum_{i=\frac{2n}{3}+1, i \neq t}^n z_i = 0 \end{cases}$$

Theorem 8 (see [4]). The number of elements in **B** equal to $\frac{n^2-3n}{3}$.

Remark 9 (see [4]). Given a contingency table $= (x_1, x_2, \dots, x_n)'$, the entry of the matrix **A** in the column indexed by x_1, x_2, \dots, x_n respectively and its rows indexed by

Corollary 12 (see [4]). The toric ideal I_A for $\frac{n^2-3n}{3} \times 3 \times \frac{n}{3}$ -contingency tables are $I_A = \langle P_{i+l}P_{j+k} - P_{j+l}P_{i+k} : i, j = 1, 2, \dots, \frac{n}{3} \text{ and } l, k = 0, \frac{n}{3}, \frac{2n}{3} \rangle$, such that $i < j$ and $l < k \rangle \subset \mathbb{C}[P_1, P_2, \dots, P_n]$.

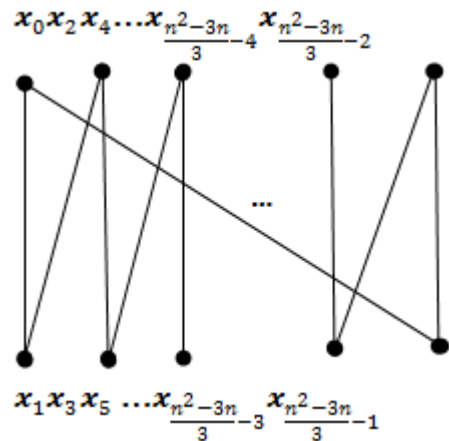
Remark 13 (see [4]). Now, we will construct a connected graph by using the elements of **B**. Let z_m be an element of

B such that $z_m = x_m - x_{m-1}$, $m = 1, 2, \dots, \frac{n^2-3n}{3}$ is an edge connect x_m and x_{m-1} , ..., and $z_{\frac{n^2-3n}{3}} = x_0 - x_{\frac{n^2-3n}{3}-1}$ is an edge connect x_0 and $x_{\frac{n^2-3n}{3}-1}$, where $x_i \in A^{-1}[t]$, $i = 0, 2, \dots, \frac{n^2-3n}{3} - 1$. Then

we can connect all $\frac{n^2-3n}{3} \times 3 \times \frac{n}{3}$ -contingency tables with fixed two dimensional marginals by $\frac{n^2-3n}{3}$ edges by applying moves from **B** one by one and go from x_0 to

$x_{\frac{n^2-3n}{3}-1}$ without causing negative cell frequencies on the way, and also from $x_{\frac{n^2-3n}{3}-1}$ to x_0 of this type, by forming undirected graph $G = (R, W, \mathbf{B}) = A^{-1}[\mathbf{t}]_{\mathbf{B}}$, where the contingency tables interpreted as vertices and connecting moves are interpreted as edges of a graph, $R = \{x_0, x_2, \dots, x_{\frac{n^2-3n}{3}-2}\}$ and $W = \{x_1, x_3, \dots, x_{\frac{n^2-3n}{3}-1}\}$ as shown in figure 1

Figure1. The graph $G = (R, W, \mathbf{B}) = A^{-1}[\mathbf{t}]_{\mathbf{B}}$.



In [5] H. H. Abbass and H. S. Mohammed Hussein assumed $n = 3^m, m \in \mathbb{N}$, and $m \geq 2$, and H is the subgroup

$$I_A = \langle P_{g(i+l)} P_{g(j+k)} - P_{g(j+l)} P_{g(i+k)} : i, j = 1, 2, \dots, \frac{n}{3} \text{ and } l, k = 0, \frac{n}{3}, \frac{2n}{3}, \text{ such that } i < j \text{ and } l < k \rangle \subset C[P_1, P_2, \dots, P_n],$$

for all $g \in H$.

3. The Main Results

Let $n = 3^m, m \in \mathbb{N}$ and $m \geq 2$, $g x_j \in A^{-1}[g\mathbf{t}], j = 0, \dots, \frac{n^2-3n}{3} - 1$ and $g \in H$ be representative elements of the set of $3 \times \frac{n}{3}$ -contingency tables. Then we write g as $n \times n$ permutation matrix $T_g = \{P_{ij}\} = \{\delta_{ij}, g(i)\}$, where δ is the Kronecker's delta such that $T_{g_1 g_2} = T_{g_1} T_{g_2}$ for $g_1, g_2 \in H$, and $T_{g^{-1}} = T_g^{-1}$. The identity matrix of the order n denoted by E_n for the unit element e .

Now, we consider a left action of dihedral group $D_n, n = |I|$, on $A^{-1}[\mathbf{t}]$ the set of $\frac{n^2-3n}{3} \times 3 \times \frac{n}{3}$ -

$$x_i = x_j + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k z_k, x_j + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k z_k \in A^{-1}[\mathbf{t}] \text{ for } 1 \leq k \leq \frac{n^2-3n}{3}. \text{ (Definition 6)}$$

Let $g \in H$

$$\text{If } g = e \Rightarrow E_n x_i = E_n x_j + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k E_n z_k, E_n x_j + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k E_n z_k \in H(A^{-1}[\mathbf{t}])$$

$$\text{for } 1 \leq k \leq \frac{n^2-3n}{3}.$$

If $g = r^{\frac{n}{3}}$ then we write $r^{\frac{n}{3}}$ as a $n \times n$ permutation matrix, i.e. $r^{\frac{n}{3}} = T_{(1 \frac{n}{3}+1 \frac{2n}{3}+1)(2 \frac{n}{3}+2 \frac{2n}{3}+2)\dots(\frac{2n}{3} n)}$, then

$\{e, r^{\frac{n}{3}}, r^{\frac{2n}{3}}, sr, sr^{\frac{n}{3}+1}, sr^{\frac{2n}{3}+1}\}$ of dihedral group D_n , where $r = (1 \ 2 \ 3 \ \dots \ n)$ and $s = (2 \ n)(3 \ n-1) \dots (\frac{n+1}{2} \ \frac{n+3}{2})$.

Theorem 14 (see [4]). The graph $G = (R, W, \mathbf{B})$ is a connected bipartite graph (up to graph isomorphism).

Theorem 15 (see [5]). The Markov basis \mathbf{B} is H -invariant.

Corollary 16 (see [5]). The subgroup H is the largest Subgroup of the group D_{3^m} such that the Markov basis \mathbf{B} is H -invariant.

Remark 17 (see [5]). Let $\mathbf{t} = (t_1, t_2, t_3, \dots, t_{\frac{n}{3}+3})'$, $x_i \in A^{-1}[\mathbf{t}]$ and $g \in H$.

Then $g x_i \in A^{-1}[g\mathbf{t}]$ where

$$g\mathbf{t} = (gt_1, gt_2, gt_3, \dots, gt_{\frac{n}{3}+3})', A^{-1}[g\mathbf{t}] = \{x \in \mathbb{N}^n : Ax = g\mathbf{t}\}$$

. So, we have six types of $g\mathbf{t}$ -fibers $A^{-1}[\mathbf{t}], A^{-1}[r^{\frac{n}{3}}\mathbf{t}], A^{-1}[r^{\frac{2n}{3}}\mathbf{t}], A^{-1}[sr\mathbf{t}], A^{-1}[sr^{\frac{n}{3}+1}\mathbf{t}]$ and $A^{-1}[sr^{\frac{2n}{3}+1}\mathbf{t}]$.

Theorem 18 (see [5]). If $g \in H$, then \mathbf{B} is a Markov basis for $\frac{n^2-3n}{3}$ contingency tables $g x_0, g x_1, \dots, g x_{\frac{n^2-3n}{3}-1}$ in $A^{-1}[g\mathbf{t}]$.

Corollary 19 (see [5]). The toric ideal for $\frac{n^2-3n}{3} \times 3 \times \frac{n}{3}$ - contingency table in $A^{-1}[g\mathbf{t}]$ is

contingency tables, and the action of dihedral group D_n on the set of Markov basis \mathbf{B} .

Theorem 20. Let $x_i, x_j \in A^{-1}[\mathbf{t}]$, if $g \in H$. Then x_i accessible from x_j by \mathbf{B} if and only if $T_g x_i$ accessible from $T_g x_j$ by \mathbf{B} , for all $i, j = 1, 2, \dots, \frac{n^2-3n}{3}$.

Proof: If x_i accessible from x_j by \mathbf{B} , then there exists a sequence of moves $z_1, \dots, z_k \in \mathbf{B}$ and $\varepsilon_k \in \{-1, 1\}, k = 1, \dots, \frac{n^2-3n}{3}$, such that

$$\begin{aligned}
 &T\left(1 \frac{n}{3}+1 \frac{2n}{3}+1\right)\left(2 \frac{n}{3}+2 \frac{2n}{3}+2\right) \dots\left(\frac{n 2 n}{3 3} n\right) \mathbf{x}_i \\
 &= T\left(1 \frac{n}{3}+1 \frac{2n}{3}+1\right)\left(2 \frac{n}{3}+2 \frac{2n}{3}+2\right) \dots\left(\frac{n 2 n}{3 3} n\right) \mathbf{x}_j \\
 &\quad + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k T\left(1 \frac{n}{3}+1 \frac{2n}{3}+1\right)\left(2 \frac{n}{3}+2 \frac{2n}{3}+2\right) \dots\left(\frac{n 2 n}{3 3} n\right) \mathbf{z}_k, T\left(1 \frac{n}{3}+1 \frac{2n}{3}+1\right)\left(2 \frac{n}{3}+2 \frac{2n}{3}+2\right) \dots\left(\frac{n 2 n}{3 3} n\right) \mathbf{x}_j \\
 &\quad + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k T\left(1 \frac{n}{3}+1 \frac{2n}{3}+1\right)\left(2 \frac{n}{3}+2 \frac{2n}{3}+2\right) \dots\left(\frac{n 2 n}{3 3} n\right) \mathbf{z}_k \\
 &\in H\left(A^{-1}[\mathbf{t}]\right) \text{ for } 1 \leq k \leq \frac{n^2-3n}{3}.
 \end{aligned}$$

If $g = r^{\frac{2n}{3}}$, then $r^{\frac{2n}{3}} = T\left(1 \frac{2n}{3}+1 \frac{n}{3}+1\right)\left(2 \frac{2n}{3}+2 \frac{n}{3}+2\right) \dots\left(\frac{n}{3} n \frac{2n}{3}\right)$,

$$\begin{aligned}
 &T\left(1 \frac{2n}{3}+1 \frac{n}{3}+1\right)\left(2 \frac{2n}{3}+2 \frac{n}{3}+2\right) \dots\left(\frac{n}{3} n \frac{2n}{3}\right) \mathbf{x}_i \\
 &= T\left(1 \frac{2n}{3}+1 \frac{n}{3}+1\right)\left(2 \frac{2n}{3}+2 \frac{n}{3}+2\right) \dots\left(\frac{n}{3} n \frac{2n}{3}\right) \mathbf{x}_j \\
 &\quad + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k T\left(1 \frac{2n}{3}+1 \frac{n}{3}+1\right)\left(2 \frac{2n}{3}+2 \frac{n}{3}+2\right) \dots\left(\frac{n}{3} n \frac{2n}{3}\right) \mathbf{z}_k, T\left(1 \frac{2n}{3}+1 \frac{n}{3}+1\right)\left(2 \frac{2n}{3}+2 \frac{n}{3}+2\right) \dots\left(\frac{n}{3} n \frac{2n}{3}\right) \mathbf{x}_j \\
 &\quad + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k T\left(1 \frac{2n}{3}+1 \frac{n}{3}+1\right)\left(2 \frac{2n}{3}+2 \frac{n}{3}+2\right) \dots\left(\frac{n}{3} n \frac{2n}{3}\right) \mathbf{z}_k \\
 &\in H\left(A^{-1}[\mathbf{t}]\right) \text{ for } 1 \leq k \leq \frac{n^2-3n}{3}.
 \end{aligned}$$

If $g = sr$, then $sr = T\left(1 n\right)\left(2 n-1\right) \dots\left(\frac{n-1 n+3}{2}\right)$,

$$\begin{aligned}
 &T\left(1 n\right)\left(2 n-1\right) \dots\left(\frac{n-1 n+3}{2}\right) \mathbf{x}_i = T\left(1 n\right)\left(2 n-1\right) \dots\left(\frac{n-1 n+3}{2}\right) \mathbf{x}_j + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k T\left(1 n\right)\left(2 n-1\right) \dots\left(\frac{n-1 n+3}{2}\right) \mathbf{z}_k \\
 &T\left(1 n\right)\left(2 n-1\right) \dots\left(\frac{n-1 n+3}{2}\right) \mathbf{x}_j + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k T\left(1 n\right)\left(2 n-1\right) \dots\left(\frac{n-1 n+3}{2}\right) \mathbf{z}_k \in H\left(A^{-1}[\mathbf{t}]\right) \text{ for } 1 \leq k \leq \frac{n^2-3n}{3}.
 \end{aligned}$$

If $g = sr^{\frac{n}{3}+1}$, then $sr^{\frac{n}{3}+1} = T\left(1 \frac{2n}{3}\right)\left(2 \frac{2n}{3}-1\right) \dots\left(\frac{n n}{3 3}+1\right)\left(\frac{2n}{3}+1 n\right)\left(\frac{2n}{3}+2 n-1\right) \dots\left(\frac{5n-35n+9}{6}\right)$,

$$\begin{aligned}
 &T\left(1 \frac{2n}{3}\right)\left(2 \frac{2n}{3}-1\right) \dots\left(\frac{n n}{3 3}+1\right)\left(\frac{2n}{3}+1 n\right)\left(\frac{2n}{3}+2 n-1\right) \dots\left(\frac{5n-35n+9}{6}\right) \mathbf{x}_i = \\
 &T\left(1 \frac{2n}{3}\right)\left(2 \frac{2n}{3}-1\right) \dots\left(\frac{n n}{3 3}+1\right)\left(\frac{2n}{3}+1 n\right)\left(\frac{2n}{3}+2 n-1\right) \dots\left(\frac{5n-35n+9}{6}\right) \mathbf{x}_j + \\
 &\quad + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k T\left(1 \frac{2n}{3}\right)\left(2 \frac{2n}{3}-1\right) \dots\left(\frac{n n}{3 3}+1\right)\left(\frac{2n}{3}+1 n\right)\left(\frac{2n}{3}+2 n-1\right) \dots\left(\frac{5n-35n+9}{6}\right) \mathbf{z}_k. \\
 &T\left(1 \frac{2n}{3}\right)\left(2 \frac{2n}{3}-1\right) \dots\left(\frac{n n}{3 3}+1\right)\left(\frac{2n}{3}+1 n\right)\left(\frac{2n}{3}+2 n-1\right) \dots\left(\frac{5n-35n+9}{6}\right) \mathbf{x}_j + \\
 &\quad + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k T\left(1 \frac{2n}{3}\right)\left(2 \frac{2n}{3}-1\right) \dots\left(\frac{n n}{3 3}+1\right)\left(\frac{2n}{3}+1 n\right)\left(\frac{2n}{3}+2 n-1\right) \dots\left(\frac{5n-35n+9}{6}\right) \mathbf{z}_k \in H\left(A^{-1}[\mathbf{t}]\right) \text{ for } 1 \leq k \leq \frac{n^2-3n}{3}.
 \end{aligned}$$

If $g = sr^{\frac{2n}{3}+1}$, then $sr^{\frac{2n}{3}+1} = T(1 \frac{n}{3})(2 \frac{n}{3}-1) \dots (\frac{n}{3}+1 n)(\frac{n}{3}+2 n-1) \dots (\frac{2n-2n}{3}+1)$
 $T(1 \frac{n}{3})(2 \frac{n}{3}-1) \dots (\frac{n}{3}+1 n)(\frac{n}{3}+2 n-1) \dots (\frac{2n-2n}{3}+1) x_i$
 $= T(1 \frac{n}{3})(2 \frac{n}{3}-1) \dots (\frac{n}{3}+1 n)(\frac{n}{3}+2 n-1) \dots (\frac{2n-2n}{3}+1) x_j + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k T(1 \frac{n}{3})(2 \frac{n}{3}-1) \dots (\frac{n}{3}+1 n)(\frac{n}{3}+2 n-1)$
 $\dots (\frac{2n-2n}{3}+1) z_k,$
 $T(1 \frac{n}{3})(2 \frac{n}{3}-1) \dots (\frac{n}{3}+1 n)(\frac{n}{3}+2 n-1) \dots (\frac{2n-2n}{3}+1) x_j +$
 $\sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k T(1 \frac{n}{3})(2 \frac{n}{3}-1) \dots (\frac{n}{3}+1 n)(\frac{n}{3}+2 n-1)$
 $\dots (\frac{2n-2n}{3}+1) z_k \in H(A^{-1}[t])$ for $1 \leq k \leq \frac{n^2-3n}{3}$.

Therefore $T_g x_i$ accessible from $T_g x_j$ by $H(\mathbf{B})$.

Conversely,

if $T_g x_i$ accessible from $T_g x_j$ by $H(\mathbf{B})$, then there exists a sequence of moves $g z_1, \dots, g z_s \in H(\mathbf{B}), g \in H$, and $\varepsilon_k \in \{-1, 1\}, k = 1, \dots, \frac{n^2-3n}{3}$, such that

$$T_g x_i = T_g x_j + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k T_g z_k,$$

$$T_g x_j + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k T_g z_k \in H(A^{-1}[t]) \text{ for } 1 \leq k \leq \frac{n^2-3n}{3}. \text{ (Definition 6)}$$

By multiplying the previous equations by $T'_g, g \in H$, we have

$$T'_g(T_g x_i) = T'_g(T_g x_j) + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k T'_g(T_g z_k),$$

$$T'_g(T_g x_j) + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k T'_g(T_g z_k) \in H(A^{-1}[t]) \text{ for } 1 \leq k \leq \frac{n^2-3n}{3}.$$

This implies, $(T'_g T_g) x_i = (T'_g T_g) x_j + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k (T'_g T_g) z_k,$

$$(T'_g T_g) x_j + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k (T'_g T_g) z_k \in H(A^{-1}[t]) \text{ for } 1 \leq k \leq \frac{n^2-3n}{3}.$$

Hence $E_n x_i = E_n x_j + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k E_n z_k,$

$$E_n x_j + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k E_n z_k \in H(A^{-1}[t]) \text{ for } 1 \leq k \leq \frac{n^2-3n}{3}.$$

Therefore, $x_i = x_j + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k z_k,$

$$x_j + \sum_{k=1}^{\frac{n^2-3n}{3}} \varepsilon_k z_k \in A^{-1}[t] \text{ for } 1 \leq k \leq \frac{n^2-3n}{3}. \text{ [since } E_n x_j = x_j \text{ and } E_n z_k = z_k]$$

Then x_i accessible from x_j by \mathbf{B} . □

Remark 21. Now, we will construct a connected graph by using the elements of $H(\mathbf{B})$. Let $g z_k = T_g z_k$ be an element of $H(\mathbf{B})$ for all $g \in H$ such that $g z_k = g x_k - g x_{k-1} = T_g x_k - T_g x_{k-1}$

, $k = 1, 2, \dots, \frac{n^2-3n}{3}$ is an edge connected $g x_k = T_g x_k$ and $g x_{k-1} = T_g x_{k-1}$ and, \dots , and $g z_{\frac{n^2-3n}{3}} = g x_0 - g x_{\frac{n^2-3n}{3}-1} = T_g x_0 - T_g x_{\frac{n^2-3n}{3}-1}$ be an

edge connected $T_g x_0$ and $T_g x_{\frac{n^2-3n}{3}-1}$, where $g x_i \in A^{-1}[gt], g \in H, i = 0, 1, 2, \dots, \frac{n^2-3n}{3} - 1$. Then we can connect all $\frac{n^2-3n}{3} \times 3 \times \frac{n}{3}$ - contingency tables with fixed two dimensional marginals by $\frac{n^2-3n}{3}$ edges by applying moves from $H(\mathbf{B})$ to $T_g x_0$ one by one and go from $T_g x_0$ to $T_g x_{\frac{n^2-3n}{3}-1}$ without causing negative cell frequencies on the way, and also from $T_g x_{\frac{n^2-3n}{3}-1}$ to $T_g x_0$. This forms undirected graph as shown in figure 2

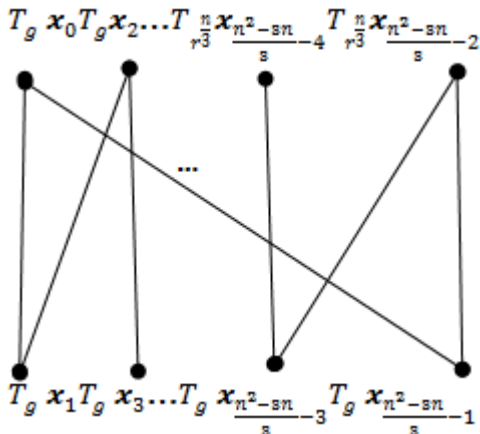


Figure 2. The graph $T_g G = (T_g R, T_g W, H(\mathbf{B})) = A^{-1}[gt]_{H(\mathbf{B})} = A^{-1}[gt]_{\mathbf{B}}$, where the contingency tables interpreted as vertices and connecting moves are interpreted as edges of a graph, $T_g R = \{T_g x_0, T_g x_2, \dots, T_g x_{\frac{n^2-3n}{3}-2}\}$ and $T_g W = \{T_g x_1, T_g x_3, \dots, T_g x_{\frac{n^2-3n}{3}-1}\}$.

Theorem 22. The graphs $T_g G = (T_g R, T_g W, H(\mathbf{B}))$ are connected bipartite graphs (up to graph isomorphism).

Proof : Let $T_g x_i, T_g x_j \in A^{-1}[gt]$, if $0 \leq i \leq j \leq \frac{n^2-3n}{3} - 1, i \neq j$, by remark 21 there exists a path $< T_g x_i, T_g x_{i+1}, T_g x_{i+2}, T_g x_{i+3}, \dots, T_g x_{j-1}, T_g x_j >$, and if $0 \leq j \leq i \leq \frac{n^2-3n}{3} - 1, i \neq j$, by remark 21 there exists a path $< T_g x_j, T_g x_{j+1}, T_g x_{j+2}, T_g x_{j+3}, \dots, T_g x_{i-1}, T_g x_i >$, and that implies there exists a path between every pair of distinct vertices $T_g x_i, T_g x_j \in A^{-1}[gt]$ of the graph, by (definition 3), G is a connected graph.

Now, we prove the graph $G = (T_g R, T_g W, H(\mathbf{B}))$ is a bipartite graph.

Let $T_g x_i, T_g x_{i+1}, \dots, T_g x_{j-1}, T_g x_j, T_g x_{j+1} = T_g x_i$ be a cycle in G . Suppose $T_g x_i \in T_g R$. Then $T_g x_{i+1} \in T_g W$, since the edge $T_g x_{i+1} = T_g x_{i+1} - T_g x_i \in H(\mathbf{B})$, then $T_g x_{i+2} \in T_g R$, since the edge $T_g x_{i+2} = T_g x_{i+2} - T_g x_{i+1} \in H(\mathbf{B})$. Continuing in this way,

we see that if k is an odd, then $T_g x_k \in T_g W$, and if k is even, then $T_g x_k \in T_g R$. Since $T_g x_{j+1} = T_g x_i \in T_g R$, it implies that $j+1$ is even and thus the cycle is of even length. By theorem 5, then the graph $G = (T_g R, T_g W, H(\mathbf{B}))$ is a bipartite graph. \square

4. Genomics and Phylogenetic

In this section, we describe some of the basic biological facts needed to understand phylogenetic models and then delve into the practical side of the algebraic statistics of these models. The basic genetic information of an organism is (almost always) carried in the form of DNA, a double helix consisting of two complementary B polymers bound together. The DNA molecules in a genome are typically represented as a number of frequencies of letters from the four letters alphabet = {A, C, G, T}. These letters correspond to the bases in the double helix that is the nucleotides Adenine, Cytosine, guanine and Thymine. The four nucleotides that form DNA come in two types: the purines (A and G) and the pyrimidine's (C and T). The two strands of the double helix are joined together via the base pairings A to T (via 2 hydrogen bonds) and C to G (via 3 hydrogen bonds). Since each cell typically contains a copy of the DNA of the organism, DNA copying occurs frequently. Several types of errors are possible during the replication of DNA. Single bases can mutate, or large pieces of DNA can separate and become reattached, possibly at another position, possibly in the opposite direction, these are just some of the events that occur over the course of evolution (C. Semple, M [2] and J. Felsenstein[6]).

5. A New Model of Genetic Algorithm Using the Action of Largest Subgroup of D_n for Invariance Markov Basis

In this section, we construct a new model of genetic algorithm that permutes the pieces of nucleotides in aligned DNA sequences using the actions of largest subgroup H of D_n for invariance Markov basis and toric ideals. Now, we describe our model in the following steps.

Step (1): Suppose we have l -taxons of DNA sequences each taxation of length L such as
 Taxon1: A G C T A A C G G T A T ...
 Taxon2: C G A T C T G A C C T T ...
 Taxon l : A C G T C A C G T A G C ...

Now, we define a pattern $i = i_1, i_2, \dots, i_m$ to be the sequence of characters. We look at a single site (column) of our sequence data. In the sequences above, we can look at the first site in the sequences and see the pattern "AC...A". A pattern frequency x_i is that i appears in our set of sequence data, and we denote to the number of frequencies by n where $n = 3^m, m \in \mathbb{N}, m \geq 2$.

Step (2): We can input pattern frequency x_i of above sequences in $3 \times \frac{n}{3}$ -contingency table as follows :

x_1	x_2	...	$x_{\frac{n}{s}}$	$\sum_{i=1}^{\frac{n}{s}} x_i$
$x_{\frac{n}{s}+1}$	$x_{\frac{n}{s}+2}$...	$x_{2\frac{n}{s}}$	$\sum_{i=\frac{n}{s}+1}^{2\frac{n}{s}} x_i$
$x_{2\frac{n}{s}+1}$	$x_{2\frac{n}{s}+2}$...	x_n	$\sum_{i=2\frac{n}{s}+1}^n x_i$
$x_1 + x_{\frac{n}{s}+1} + x_{2\frac{n}{s}+1}$	$x_2 + x_{\frac{n}{s}+2} + x_{2\frac{n}{s}+2}$...	$x_{\frac{n}{s}} + x_{2\frac{n}{s}} + x_n$	$ \mathbf{x} = \sum_{i \in I} x_i$

Where $|\mathbf{x}| = \sum_{i \in I} x_i = L$ is the length of sequences (the sample size), and x_1 is the frequency of the first pattern .

Step(4): From remark 9, A is $\frac{n+9}{3} \times n$ matrix and

$$A = \begin{bmatrix} 1 & 1 & \dots & 1 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & 1 & \dots & 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 1 & \dots & 1 & 1 \\ 1 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}_{\frac{n+9}{3} \times n}$$

$Ax = t$ i.e

$$\begin{bmatrix} 1 & 1 & \dots & 1 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & 1 & \dots & 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 1 & \dots & 1 & 1 \\ 1 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}_{\frac{n+9}{3} \times n} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ \vdots \\ t_{\frac{n}{3}+3} \end{bmatrix}$$

Where the columns of the matrix A index by the elements of the column vector \mathbf{x} .

Step (5): We can find the Markov basis from remark 7.

Step (6): $A: Z^{2 \times (\frac{n}{3})} \rightarrow Z^{2 \times (\frac{n}{3})}$ is a linear transformation, $t \in Z^{2 \times (\frac{n}{3})}$, and $A^{-1}[t]$ be the set of t -fibers, and $B \subset \ker_Z(A)$. Use remark 13 to find the bipartite graph $G = (R, W, B) = A^{-1}[t]_B$.

Step (7): We can find the toric ideals by using corollary 12 for each contingency tables.

Step (8): Find $T_g B = \{T_g z_1, T_g z_2, \dots, T_g z_{\frac{n^2-3n}{3}}\} = B$ for all $g \in H$, where T_g is a permutation matrix of g .

Step (9): Find the set $\{T_g x_0, T_g x_1, \dots, T_g x_{\frac{n^2-3n}{3}-1}\} \subseteq A^{-1}[gt]$ (gt -fibers) for all

x_2 is the frequency of the second pattern.

\vdots

$x_{\frac{n}{3}}$ is the frequency of the $\frac{n}{3}$ pattern.

$x_{\frac{n}{3}+1}$ is the frequency of the $\frac{n}{3} + 1$ pattern.

$x_{\frac{n}{3}+2}$ is the frequency of the $\frac{n}{3} + 2$ pattern.

\vdots

x_n is the frequency of the n pattern.

Step (3): Represent the contingency table $\mathbf{x} = \{x_i\}_{i \in I} \in \mathbb{N}^n$ as a n - dimensional column vector of non-negative integers $\mathbf{x} = (x_1, x_2, \dots, x_n)$, Where ' denotes the transpose of a vector or matrix , as in remark (1.4.2) then \mathbf{x} is a t -fiber

(i.e) $\mathbf{x} \in A^{-1}[t]$ where $A^{-1}[t] = \{\mathbf{x} \in \mathbb{N}^n : A\mathbf{x} = t\}$.

$g \in H$, where

$$r^{\frac{n}{3}} t = (t_3, t_1, t_2, t_4, \dots, t_{\frac{n}{3}+3})', r^{\frac{2n}{3}} t = (t_2, t_3, t_1, t_4, \dots, t_{\frac{n}{3}+3})$$

$$srt = (t_3, t_2, t_1, t_{\frac{n}{3}+3}, t_{\frac{n}{3}+2}, \dots, t_4),$$

$$sr^{\frac{n}{3}+1} t = (t_2, t_1, t_3, t_{\frac{n}{3}+3}, t_{\frac{n}{3}+2}, \dots, t_4) \quad \text{and}$$

$$sr^{\frac{2n}{3}+1} t = (t_1, t_3, t_2, t_{\frac{n}{3}+3}, t_{\frac{n}{3}+2}, \dots, t_4).$$

Step (10): Use remark 21 and theorem 22 to find the graphs $T_g G = A^{-1}[gt]_B$ for all $g \in H$.

Step (11): Use corollary 19 to find the toric ideal $I_A = \langle P_{g(i+l)} P_{g(j+k)} - P_{g(j+l)} P_{g(i+k)} : i, j = 1, 2, \dots, \frac{n}{3} \text{ and } l, k = 0, \frac{n}{3}, \frac{2n}{3} \rangle$, such that $i < j$ and $l < k$ for all $g \in H$.

Step (12): Use the $\frac{n^2-3n}{3} \times 3 \times \frac{n}{3}$ -contingency tables (t -fibers) in step (6), and $\frac{n^2-3n}{3} \times 3 \times \frac{n}{3}$ -contingency tables (gt -fibers) in step (9) for all $g \in H$ to find the permutation of nucleotides in aligned DNA sequences.

Example (23): Suppose we have the following three aligned DNA sequences

Taxon 1: A G C T G A T T G G C C C G T T T
 T T G
 Taxon 2: A G A T C T A T C G A A A C A A A
 T T C
 Taxon 3: A T T C A G A T C T T T T A A A A
 T T C

Step (1): There are three taxons of above DNA sequences with $|x| = \sum_{i \in I} x_i = L = 20$ and nine patterns A A A, G G T, C A T, T T C, G C A, A T G, T A A, T T T, G C C with frequencies 1, 2, 4, 1, 2, 1, 4, 3 and 2 respectively where $n = 9$.

Step (2): Now, we input the patterns frequency x_i of above sequences in 3×3 -contingency table as follows :

1	2	4	7
1	2	1	4
4	3	2	9
6	7	7	20

Then the table of marginal and conditional probability is:

0.05	0.10	0.20	0.35
0.05	0.10	0.05	0.20
0.20	0.15	0.10	0.45

Step (3): Represent the contingency table $x = \{x_i\}_{i \in I} \in \mathbb{N}^9$ as a 9- dimensional column vector of non- negative integers $x = (1, 2, 4, 1, 2, 1, 4, 3, 2)'$, as in remark (1.5.2) then x is a t - fiber

(i.e. $x \in A^{-1}[t]$, where

$$A^{-1}[t] = \{x \in \mathbb{N}^9 : A \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ 9 \\ 6 \\ 7 \\ 7 \end{bmatrix}\}.$$

Step(4): A is 6×9 matrix and

$$z_1 = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, z_2 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ -1 & 1 & 0 \end{bmatrix}, z_3 = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}_{6 \times 9}, \text{ and}$$

$$Ax = t \text{ i.e. } \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}_{6 \times 9} \begin{bmatrix} 1 \\ 2 \\ 4 \\ 1 \\ 1 \\ 2 \\ 1 \\ 4 \\ 3 \\ 2 \end{bmatrix}_{9 \times 1}$$

$$= \begin{bmatrix} 7 \\ 4 \\ 9 \\ 6 \\ 7 \\ 7 \end{bmatrix}_{6 \times 1} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \end{bmatrix}_{6 \times 1}$$

Where the columns of the matrix A index by the elements of the column vector x , and $t_1 = \sum_{i=1}^3 x_i = 7, t_2 = \sum_{i=1}^3 x_{i+3} = 4, t_3 = \sum_{i=1}^3 x_{i+6} = 9, t_4 = x_1 + x_4 + x_7 = 6, t_5 = x_2 + x_5 + x_8 = 7, t_6 = x_3 + x_6 + x_9 = 7.$

Step (5): We can find the Markov basis from remark 7. Then the number of moves is $\frac{n^2-3n}{3} = \frac{9^2-3 \times 9}{3} = 18$ elements in the set

$$\begin{aligned}
 z_4 &= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}, z_5 = \begin{bmatrix} 0 & 1 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, z_6 = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix} \\
 z_7 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}, z_8 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ -1 & 1 & 0 \end{bmatrix}, z_9 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \\
 z_{10} &= \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, z_{11} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix}, z_{12} = \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \\
 z_{13} &= \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix}, z_{14} = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}, z_{15} = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \\
 z_{16} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix}, z_{17} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}, z_{18} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}
 \end{aligned}$$

Step (6): The connected graph $A^{-1}[t]_B = G = (R, W, B)$ with $\frac{n^2-3n}{3} = 18$ (t - fibres) 3×3 - contingency tables as vertices of it.

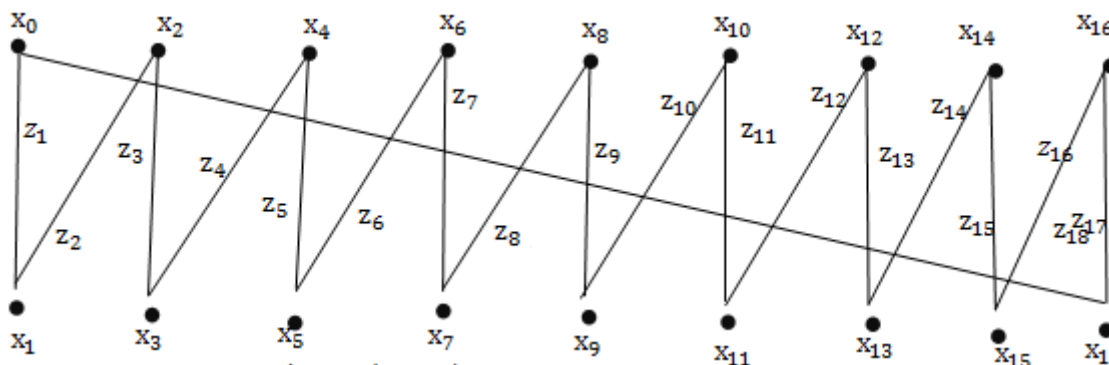
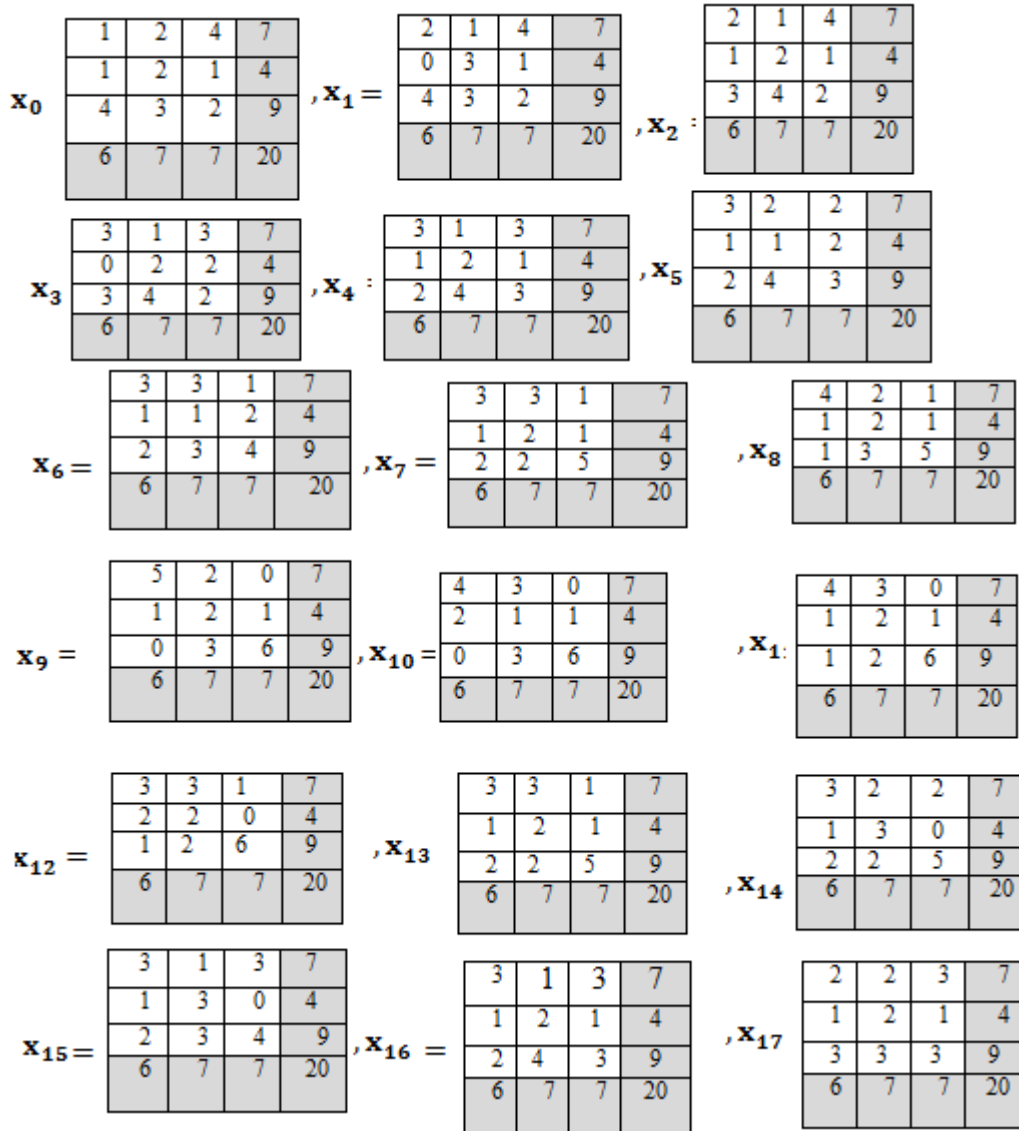


Figure 3: The graph $G = (R, W, B) = A^{-1}[t]_B$, where the contingency tables interpreted as vertices and connecting Markov basis are interpreted as the edges of a graph, $R = \{x_0, x_2, x_4, x_6, x_8, x_{10}, x_{12}, x_{14}, x_{16}\}$ and $W = \{x_1, x_3, x_5, x_7, x_9, x_{11}, x_{13}, x_{15}, x_{17}\}$ Where



Step (7): We can find the toric ideals by using corollary 12 for each contingency tables. Then the toric ideals that correspond the Markov basis that shown in the a previous figure is:

$$I_A = \langle P_1 P_5 - P_2 P_4, P_4 P_8 - P_5 P_7, P_1 P_6 - P_3 P_4, P_4 P_9 - P_6 P_7, P_2 P_6 - P_3 P_5, P_2 P_9 - P_3 P_8, P_5 P_9 - P_6 P_8, P_1 P_8 - P_2 P_7, P_1 P_9 - P_3 P_7 \rangle$$

$$I_A = \langle 0.005 - 0.005, 0.0075 - 0.02, 0.0025 - 0.01, 0.005 - 0.01, 0.005 - 0.02, 0.01 - 0.03, 0.01 - 0.0075, 0.0075 - 0.02, 0.005 - 0.04 \rangle$$

$$\Rightarrow I_A = \langle 0.0025 \rangle$$

Step (8): we find the set $T_{r^3} \mathbf{B} = \{T_{r^3} \mathbf{z}_1, T_{r^3} \mathbf{z}_2, T_{r^3} \mathbf{z}_3, T_{r^3} \mathbf{z}_4, T_{r^3} \mathbf{z}_5, T_{r^3} \mathbf{z}_6, T_{r^3} \mathbf{z}_7, T_{r^3} \mathbf{z}_8, T_{r^3} \mathbf{z}_9, T_{r^3} \mathbf{z}_{10}, T_{r^3} \mathbf{z}_{11}, T_{r^3} \mathbf{z}_{12}, T_{r^3} \mathbf{z}_{13}, T_{r^3} \mathbf{z}_{14}, T_{r^3} \mathbf{z}_{15}, T_{r^3} \mathbf{z}_{16}, T_{r^3} \mathbf{z}_{17}, T_{r^3} \mathbf{z}_{18}\}$.

Such that

$$\begin{aligned}
 T_{r^2} z_1 &= \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix}, T_{r^2} z_2 = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}, T_{r^2} z_3 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix} \\
 T_{r^2} z_4 &= \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}, T_{r^2} z_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}, T_{r^2} z_6 = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \\
 T_{r^2} z_7 &= \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}, T_{r^2} z_8 = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, T_{r^2} z_9 = \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \\
 T_{r^2} z_{10} &= \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix}, T_{r^2} z_{11} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ -1 & 1 & 0 \end{bmatrix}, T_{r^2} z_{12} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix} \\
 T_{r^2} z_{13} &= \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, T_{r^2} z_{14} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix}, T_{r^2} z_{15} = \begin{bmatrix} 0 & 1 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \\
 T_{r^2} z_{16} &= \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}, T_{r^2} z_{17} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, T_{r^2} z_{18} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}
 \end{aligned}$$

Step (9): We find the set $\{T_{r^3} x_0, T_{r^3} x_1, T_{r^3} x_2, T_{r^3} x_3, T_{r^3} x_4, T_{r^3} x_5, T_{r^3} x_6, T_{r^3} x_7, T_{r^3} x_8, T_{r^3} x_9, T_{r^3} x_{10}, T_{r^3} x_{11}, T_{r^3} x_{12}, T_{r^3} x_{13}, T_{r^3} x_{14}, T_{r^3} x_{15}, T_{r^3} x_{16}, T_{r^3} x_{17}\} \subseteq A^{-1}[r^3 t]$ ($r^3 t$ -fibers) where $r^3 t = (9, 7, 4, 6, 7, 7)'$. We have.

$$\begin{aligned}
 T_{r^3} x_0 &= \begin{bmatrix} 4 & 3 & 2 & 9 \\ 1 & 2 & 4 & 7 \\ 1 & 2 & 1 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix}, T_{r^3} x_1 = \begin{bmatrix} 4 & 3 & 2 & 9 \\ 2 & 1 & 4 & 7 \\ 0 & 3 & 1 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix}, T_{r^3} x_2 = \begin{bmatrix} 3 & 4 & 2 & 9 \\ 2 & 1 & 4 & 7 \\ 1 & 2 & 1 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix} \\
 T_{r^3} x_3 &= \begin{bmatrix} 3 & 4 & 2 & 9 \\ 3 & 1 & 3 & 7 \\ 0 & 2 & 2 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix}, T_{r^3} x_4 = \begin{bmatrix} 2 & 4 & 3 & 9 \\ 3 & 1 & 3 & 7 \\ 1 & 2 & 1 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix}, T_{r^3} x_5 = \begin{bmatrix} 2 & 4 & 3 & 9 \\ 3 & 2 & 2 & 7 \\ 1 & 1 & 2 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix} \\
 T_{r^3} x_6 &= \begin{bmatrix} 2 & 3 & 4 & 9 \\ 3 & 3 & 1 & 7 \\ 1 & 1 & 2 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix}, T_{r^3} x_7 = \begin{bmatrix} 2 & 2 & 5 & 9 \\ 3 & 3 & 1 & 7 \\ 1 & 2 & 1 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix}, T_{r^3} x_8 = \begin{bmatrix} 1 & 3 & 5 & 9 \\ 4 & 2 & 1 & 7 \\ 1 & 2 & 1 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix}
 \end{aligned}$$

Step (10): Use remark 21 and theorem 22 to find the graph

$$\begin{aligned}
 T_{r^3} x_9 &= \begin{bmatrix} 0 & 3 & 6 & 9 \\ 5 & 2 & 0 & 7 \\ 1 & 2 & 1 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix}, T_{r^3} x_{10} = \begin{bmatrix} 0 & 3 & 6 & 9 \\ 4 & 3 & 0 & 7 \\ 2 & 1 & 1 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix}, T_{r^3} x_{11} = \begin{bmatrix} 1 & 2 & 6 & 9 \\ 4 & 3 & 0 & 7 \\ 1 & 2 & 1 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix} \\
 T_{r^3} x_{12} &= \begin{bmatrix} 1 & 2 & 6 & 9 \\ 3 & 3 & 1 & 7 \\ 2 & 2 & 0 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix}, T_{r^3} x_{13} = \begin{bmatrix} 2 & 2 & 5 & 9 \\ 3 & 3 & 1 & 7 \\ 1 & 2 & 1 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix}, T_{r^3} x_{14} = \begin{bmatrix} 2 & 2 & 5 & 9 \\ 3 & 2 & 2 & 7 \\ 1 & 3 & 0 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix} \\
 T_{r^3} x_{15} &= \begin{bmatrix} 2 & 3 & 4 & 9 \\ 3 & 1 & 3 & 7 \\ 1 & 3 & 0 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix}, T_{r^3} x_{16} = \begin{bmatrix} 2 & 4 & 3 & 9 \\ 3 & 1 & 3 & 7 \\ 1 & 2 & 1 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix}, T_{r^3} x_{17} = \begin{bmatrix} 3 & 3 & 3 & 9 \\ 2 & 2 & 3 & 7 \\ 1 & 2 & 1 & 4 \\ 6 & 7 & 7 & 20 \end{bmatrix}
 \end{aligned}$$

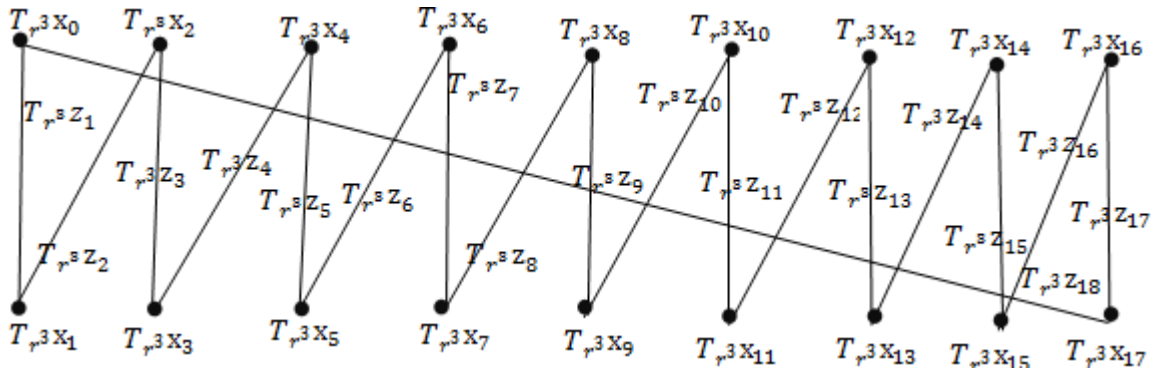


Figure 4: The graph $T_{r^3}G = (T_{r^3}R, T_{r^3}W, H(\mathbf{B})) = A^{-1}[r^3t]_{H(\mathbf{B})} = A^{-1}[r^3t]_{\mathbf{B}}$, where the contingency tables interpreted as vertices and connecting moves are interpreted as edges of a graph, $T_{r^3}R = \{r^3x_0, r^3x_2, r^3x_4, r^3x_6, r^3x_8, r^3x_{10}, r^3x_{12}, r^3x_{14}, r^3x_{16}\}$ and $T_{r^3}W = \{r^3x_1, r^3x_3, r^3x_5, r^3x_7, r^3x_9, r^3x_{11}, r^3x_{13}, r^3x_{15}, r^3x_{17}\}$.

Step (11):

Use corollary 19 to find the toric ideal $I_A = \langle P_{i+l}P_{j+k} - P_{j+l}P_{i+k} : i, j = 1, 2, \dots, \frac{n}{3} \text{ and } l, k = 0, \frac{n}{3}, \frac{2n}{3} \text{, such that } i < j \text{ and } l < k \rangle = \langle P_1P_5 - P_2P_4, P_4P_8 - P_5P_7, P_1P_6 - P_3P_4, P_4P_9 - P_6P_7, P_2P_6 - P_3P_5, P_2P_9 - P_3P_8, P_5P_9 - P_6P_8, P_1P_8 - P_2P_7, P_1P_9 - P_3P_7 \rangle$

if $g = sr$, then $T_{sr}\mathbf{B} = \{T_{sr}z_1, T_{sr}z_2, T_{sr}z_3, T_{sr}z_4, T_{sr}z_5, T_{sr}z_6, T_{sr}z_7, T_{sr}z_8, T_{sr}z_9, T_{sr}z_{10}, T_{sr}z_{11}, T_{sr}z_{12}, T_{sr}z_{13}, T_{sr}z_{14}, T_{sr}z_{15}, T_{sr}z_{16}, T_{sr}z_{17}, T_{sr}z_{18}\}$,

$$\begin{aligned}
 T_{sr}z_1 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}, T_{sr}z_2 = \begin{bmatrix} 0 & 1 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, T_{sr}z_3 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix} \\
 T_{sr}z_4 &= \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, T_{sr}z_5 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ -1 & 1 & 0 \end{bmatrix}, T_{sr}z_6 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ -1 & 1 & 0 \end{bmatrix} \\
 T_{sr}z_7 &= \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, T_{sr}z_8 = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}, T_{sr}z_9 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & -1 \end{bmatrix} \\
 T_{sr}z_{10} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix}, T_{sr}z_{11} = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}, T_{sr}z_{12} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix} \\
 T_{sr}z_{13} &= \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}, T_{sr}z_{14} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix}, T_{sr}z_{15} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix} \\
 T_{sr}z_{16} &= \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, T_{sr}z_{17} = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}, T_{sr}z_{18} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}
 \end{aligned}$$

Also, we find the set $\{T_{sr}x_0, T_{sr}x_1, T_{sr}x_2, T_{sr}x_3, T_{sr}x_4, T_{sr}x_5, T_{sr}x_6, T_{sr}x_7, T_{sr}x_8,$

$T_{sr}x_9, T_{sr}x_{10}, T_{sr}x_{11}, T_{sr}x_{12}, T_{sr}x_{13}, T_{sr}x_{14}, T_{sr}x_{15}, T_{sr}x_{16}, T_{sr}x_{17}\} \subseteq A^{-1}[srt](srt\text{-fibers})$ where, $srt = (9, 4, 7, 7, 7, 6)'$, such that

$T_{sr}x_0 = \begin{bmatrix} 2 & 3 & 4 & 9 \\ 1 & 2 & 1 & 4 \\ 4 & 2 & 1 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$	$T_{sr}x_1 = \begin{bmatrix} 2 & 3 & 4 & 9 \\ 1 & 3 & 0 & 4 \\ 4 & 1 & 2 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$	$T_{sr}x_2 = \begin{bmatrix} 2 & 4 & 3 & 9 \\ 1 & 2 & 1 & 4 \\ 4 & 1 & 2 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$
$T_{sr}x_3 = \begin{bmatrix} 2 & 4 & 3 & 9 \\ 2 & 2 & 0 & 4 \\ 3 & 1 & 3 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$	$T_{sr}x_4 = \begin{bmatrix} 3 & 4 & 2 & 9 \\ 1 & 2 & 1 & 4 \\ 3 & 1 & 3 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$	$T_{sr}x_5 = \begin{bmatrix} 3 & 4 & 2 & 9 \\ 2 & 1 & 1 & 4 \\ 2 & 2 & 3 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$
$T_{sr}x_6 = \begin{bmatrix} 4 & 3 & 2 & 9 \\ 2 & 1 & 1 & 4 \\ 1 & 3 & 3 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$	$T_{sr}x_7 = \begin{bmatrix} 5 & 2 & 2 & 9 \\ 1 & 2 & 1 & 4 \\ 1 & 3 & 3 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$	$T_{sr}x_8 = \begin{bmatrix} 5 & 3 & 1 & 9 \\ 1 & 2 & 1 & 4 \\ 1 & 2 & 4 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$
$T_{sr}x_9 = \begin{bmatrix} 6 & 3 & 0 & 9 \\ 1 & 2 & 1 & 4 \\ 0 & 2 & 5 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$	$T_{sr}x_{10} = \begin{bmatrix} 6 & 3 & 0 & 9 \\ 1 & 1 & 2 & 4 \\ 0 & 3 & 4 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$	$T_{sr}x_{11} = \begin{bmatrix} 6 & 2 & 1 & 9 \\ 1 & 2 & 1 & 4 \\ 0 & 3 & 4 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$
$T_{sr}x_{12} = \begin{bmatrix} 6 & 2 & 1 & 9 \\ 0 & 2 & 2 & 4 \\ 1 & 3 & 3 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$	$T_{sr}x_{13} = \begin{bmatrix} 5 & 2 & 2 & 9 \\ 1 & 2 & 1 & 4 \\ 1 & 3 & 3 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$	$T_{sr}x_{14} = \begin{bmatrix} 5 & 2 & 2 & 9 \\ 0 & 3 & 1 & 4 \\ 2 & 2 & 3 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$
$T_{sr}x_{15} = \begin{bmatrix} 4 & 3 & 2 & 9 \\ 0 & 3 & 1 & 4 \\ 3 & 1 & 3 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$	$T_{sr}x_{16} = \begin{bmatrix} 3 & 4 & 2 & 9 \\ 1 & 2 & 1 & 4 \\ 3 & 1 & 3 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$	$T_{sr}x_{17} = \begin{bmatrix} 3 & 3 & 3 & 9 \\ 1 & 2 & 1 & 4 \\ 3 & 2 & 2 & 7 \\ 7 & 7 & 6 & 20 \end{bmatrix}$

Now, use remark 21 and theorem 22 to find the graph $T_{sr}G = A^{-1}[srt]_B$.

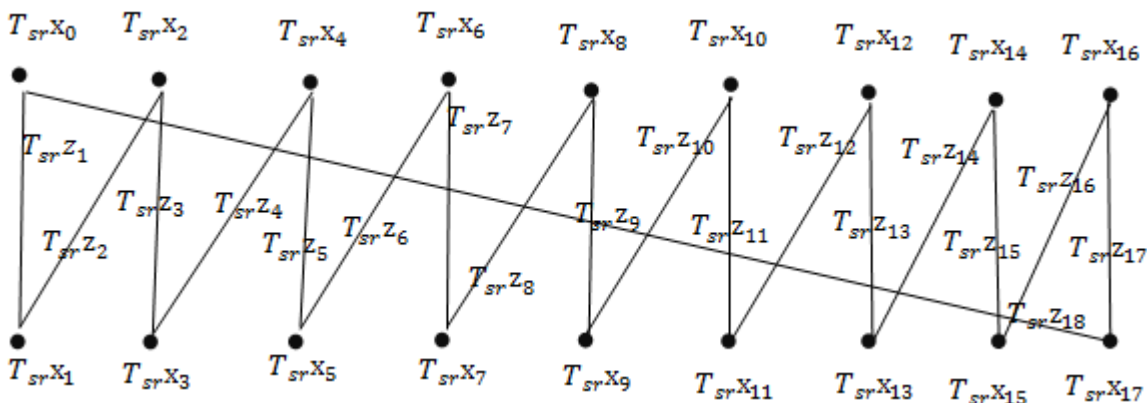


Figure 5: The graph $T_{sr}G = (T_{sr}R, T_{sr}W, H_2(B)) = A^{-1}[srt]_{H_2(B)} = A^{-1}[srt]_B$, where the contingency tables interpreted as vertices and connecting moves are interpreted as edges of a graph, where $T_{sr}R = \{T_{sr}x_0, T_{sr}x_2, T_{sr}x_4, T_{sr}x_6, T_{sr}x_8, T_{sr}x_{10}, T_{sr}x_{12}, T_{sr}x_{14}, T_{sr}x_{16}\}$ and $T_{sr}W = \{T_{sr}x_1, T_{sr}x_3, T_{sr}x_5, T_{sr}x_7, T_{sr}x_9, T_{sr}x_{11}, T_{sr}x_{13}, T_{sr}x_{15}, T_{sr}x_{17}\}$.

$x_0 = \begin{matrix} \textcircled{1} \textcircled{2} \textcircled{4} \textcircled{1} \textcircled{2} \textcircled{1} \textcircled{4} \textcircled{3} \textcircled{2} \\ \text{Taxon1: AGCTGA TTG G CCCC TTTTT G} \\ \text{Taxon2: AGA TC T AT C GA A AAAAA TT C} \\ \text{Taxon3: AT TCA GATCTTTTAAAA TTC} \end{matrix}$

$x_1 = \begin{matrix} \textcircled{2} \textcircled{1} \textcircled{4} \textcircled{0} \textcircled{3} \textcircled{1} \textcircled{4} \textcircled{3} \textcircled{2} \\ \text{Taxon1: AGCGA TTG ACCCG G TTTTTG} \\ \text{Taxon2: A GAC T AT C AAA A C CA AATTC} \\ \text{Taxon3: A T TA GATC AT TT AAA AATTC} \end{matrix}$

$x_2 = \begin{matrix} \textcircled{2} \textcircled{1} \textcircled{4} \textcircled{1} \textcircled{2} \textcircled{1} \textcircled{3} \textcircled{4} \textcircled{2} \\ \text{Taxon1: AGCTGA TTG ACCCG TTTTTG} \\ \text{Taxon2: AGA TC TAT C AA A ACAA TTTC} \\ \text{Taxon3: AT TCA G ATC ATTTAAAA TTC} \end{matrix}$

Step(12): Use the $18 \times 3 \times 3$ -contingency tables (t -fibers) in step (6), $18 \times 3 \times 3$ -contingency tables (gt -fibers) in step (9) to find the permutation of nucleotides in aligned DNA sequences. Then the change in the type of DNA sequences under the Markov basis. Be as Figure 3 where

$x_3 =$ ③①③①②②③④②
 Taxon1: AGCGATT G AACCG ATTTTT G
 Taxon2: AGAC T AT C AAA ACTA ATTT C
 Taxon3: AT TA GATC AATTGA ATTC

$x_4 =$ ③①③①②①②④③
 Taxon1: AGCTGA TTG AA CCGTTTT G G
 Taxon2: AGA TC T AT C AA AACA TTT C C
 Taxon3: AT TCA GATC AATT AA TTTC C

$x_5 =$ ③②②①①②②④③
 Taxon1: AGCTGA TTG A A G CATT TT G G
 Taxon2: AGA TC T AT CAA GATA TTT C C
 Taxon3: A T TCA GATCA A TTGA TTTC C

$x_6 =$ ③③①①①②②③④
 Taxon1: AGCTGA TTG A A GG ATTTG G G
 Taxon2: AGA TC T AT CAAGGTA TTC C C
 Taxon3: AT TCA GATCA ATTGATTC C C

$x_7 =$ ③③①①②①②②⑤
 Taxon1: AGCTGA TTG A A GG G TTGG G G
 Taxon2: AGA TC T AT CAAGGCA TCC C C
 Taxon3: A T TCA GATCA ATTAA TCC C C

$x_8 =$ ④②①①②①①③⑤
 Taxon1: AGCTGA TTG A A AG G TTGG G G
 Taxon2: AGA TC T AT CAAA GCTTCC C C
 Taxon3: AT TCA GATCA AATATTCC C C

$x_9 =$ ⑤②①②①①③⑥
 Taxon1: AGTGA TGA AAA G G TTGG G G G
 Taxon2: AG TC TT CAAA A GCTTCC C C C
 Taxon3: A T CA GTCA AA ATATTCC C C C

$x_{10} =$ ④③①②①①③⑥
 Taxon1: AGTGA TGA AA GG TTTGG G G G
 Taxon2: AG TC TT CAAA GGTTTCC C C C
 Taxon3: A T CA GTCA AA TTCTTCC C C C

$x_{11} =$ ④③①②①①②⑥
 Taxon1: AGTGATTG A A GG G TGG G G G
 Taxon2: AG TC TAT CAAA GGCTCC C C C
 Taxon3: A T CA GATCA AA TTATCC C C C

$x_{12} =$ ③③①②②①②⑥
 Taxon1: AGCTGTTG A A GGT G TGG G G G
 Taxon2: AG ATC AT CAAGGTCTCC C C C
 Taxon3: A TTCA ATCA ATTCATCC C C C

$x_{13} =$ ③③①①②①②②⑤
 Taxon1: AGCTGA TTG A A GG G TTGG G G
 Taxon2: AGA TC T AT CAAGGCA TCC C C
 Taxon3: AT TCA GATCA ATTAA TCC C C

$x_{14} =$ ③②②①③①②②⑤
 Taxon1: AGCTGTTG A A GC G G TTGG G G
 Taxon2: AGA TC AT CAAGAC CA TCC C C
 Taxon3: AT TCA ATCA ATTAAA TCC C C

$x_{15} =$ ③①③①③①②③④
 Taxon1: AGCTGTTG A ACC G G TTTG G G
 Taxon2: AGA TC AT CAAA AC CATT C C C
 Taxon3: A T TCA ATCA ATTAAA TTC C C

$x_{16} =$ ③①③①②①②④③
 Taxon1: AGCTGA TTG AA CCGTTTT G G
 Taxon2: AGA TC T AT C AA AACA TTT C C
 Taxon3: AT TCA GATC AATT AA TTTC C

$x_{17} =$ ②②③①②①③③③
 Taxon1: AGCTGA TTG AGCCGTTTT G G
 Taxon2: AGA TC T AT C AG AACA ATT C C
 Taxon3: AT TCA GATC ATTT AA ATTC C

And the change in the type of DNA sequences under the action of r^3 on the set of Markov basis. Be as Figure 4 where

$T_{r^3}x_0 =$ ④③②①②④①②①
 Taxon1: AGCTGA TTG A A AGGCG A A AT
 Taxon2: AGA TC T AT C A A AGG A CTTT
 Taxon3: A T TCA GATC A A ATTTAGGGT

$T_{r^3}x_1 =$ ④③②②①④①③①
 Taxon1: AGCTGA TG A AA GGCTA AATT
 Taxon2: A GA TC T T C AAA GG ATTTTT
 Taxon3: A T TCA GTC A A ATTTCCGGTT

$T_{r^3}x_2 =$ ③④②②①④①②①
 Taxon1: AGCTGA TTG AA GGG CTA A AT
 Taxon2: AGA TC TAT C AA GGGATTTTT
 Taxon3: AT TCA G ATC A ATTTTCCGGT

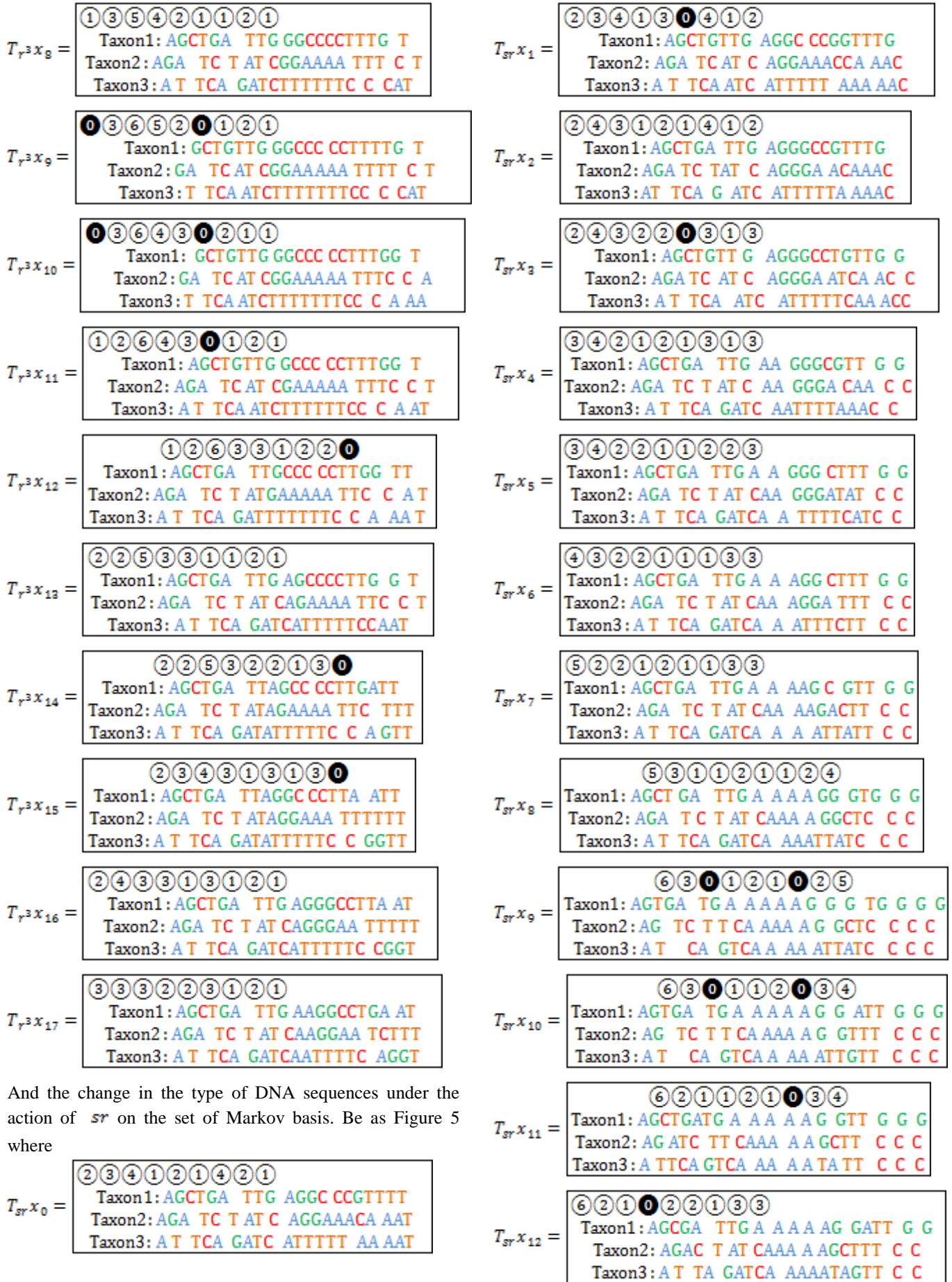
$T_{r^3}x_3 =$ ③④②③①③①②②
 Taxon1: AGCTGAT G AAGGG CTTAAT G
 Taxon2: AGA TC T T C AAGGGATTTTT C
 Taxon3: A T TCA GTC AATTTTC CGTC

$T_{r^3}x_4 =$ ②④③③①③①②①
 Taxon1: AGCTGA TTG AGGGCCTTA AT
 Taxon2: AGA TC T AT C AGGGAA TTTT
 Taxon3: A T TCA GATC ATTTTTC CGGT

$T_{r^3}x_5 =$ ②④③③②②①①②
 Taxon1: AGCTGA TTG AGGGCCTTGA G
 Taxon2: AGA TC T AT CAGGGAA TTCT C
 Taxon3: A T TCA GATCATTTTTC CAG C

$T_{r^3}x_6 =$ ②③④③③①①①②
 Taxon1: AGCTGA TTG AGGCCCTTGG G
 Taxon2: AGA TC T AT CAGGAAA T TCC C
 Taxon3: A T TCA GATCATTTTTC CAA C

$T_{r^3}x_7 =$ ②②⑤③③①①②①
 Taxon1: AGCTGA TTG AGCCCCCTTGG T
 Taxon2: AGA TC T AT CAGAAAA TTC C T
 Taxon3: A T TCA GATCATTTTTC CA AT



And the change in the type of DNA sequences under the action of sr on the set of Markov basis. Be as Figure 5 where

$T_{sr}^{x_{13}} =$

⑤②②①②①①③③
Taxon1: AGCTGA TTGAAAAG CGTTGG
Taxon2: AGA TC T AT CA AA AGA CTT CC
Taxon3: AT TCA GATCAAAATTATT CC

$T_{sr}^{x_{14}} =$

⑤②②①③①②②③
Taxon1: AGCGA TTGAAAAG CGGTTGG
Taxon2: AGAC T AT CA AA AGA C CA TCC
Taxon3: AT TA GATCAAAATTTAA TCC

$T_{sr}^{x_{15}} =$

④③②①③①③①③
Taxon1: AGCGA TTGAAAAGG CGGTTGG
Taxon2: AGAC T AT CAA A GGA C CA ACC
Taxon3: AT TA GATCAAAATTTAA ACC

$T_{sr}^{x_{16}} =$

③④②①②①③①③
Taxon1: AGCTGA TTG AA GGG CGTTGG
Taxon2: AGA TC T AT C AA GGA CAA CC
Taxon3: AT TCA GATC AATTTTAAAC C

$T_{sr}^{x_{17}} =$

③③③①②①③②②
Taxon1: AGCTGA TTG AA GGCCGTTT G
Taxon2: AGA TC T AT C AA GG AACA AT C
Taxon3: AT TCA GATC AATTTT AA AT C

Remark 24:

- i. We refer to ①, ②, ③, ④, ⑤ and ⑥ in example 23 to the frequencies of the patterns in DNA sequences.
- ii. We refer ① to the hidden in the pattern frequency of DNA sequences.

References

- [1] A. Takemura, and S. Aoki, "Some characterizations of minimal Markov basis for sampling from discrete conditional distributions", Annals of the Institute of Statistical Mathematics, 2003.
- [2] C. Semple, M. Steel, "Phylogenetics", volume 24 of Oxford Lecture Series in Mathematics and its Applications, Oxford University Press, Oxford, (2003).
- [3] G. Agnarsson, R. Greenlaw. "Graph theory, modeling, applications, and algorithms ", United States of America, (2007).
- [4] H. H. Abbass and H. S. Mohammed Hussein" On Toric Ideals for $3 \times \frac{n}{3}$ - Contingency Tables with Fixed Two Dimensional Marginals, n is a multiple of 3", European Journal of Scientific Research, Vol. 123, pp. 83-98, 2014.
- [5] H. H. Abbass and H. S. Mohammed Hussein"An Invariant Markov basis Under the Action of Largest Subgroup of Dihedral Group D_{3^m} ", European Journal of Scientific Research,(accept), 2014.
- [6] J. Felsenstein, "Inferring Phylogenies", Sinauer Associates, Sunderland, (2003).
- [7] P. Diaconis, and B. Sturmfels, "Algebraic algorithms for sampling from conditional distributions", The Annals of Statistics, 26, pp. 363-397, 1998.
- [8] P. Diaconis, D. Eisenbud, B. Sturmfels. "Lattice walks and primary decomposition", Mathematical Essays in Honor of Gian-Carlo Rota, eds. B. Sagan and R.

Stanley, Progress in Mathematics, Vol. 161, Birkhauser, Boston, (1998), pp. 173-193.

- [9] S. Aoki, and A. Takemura," Minimal Invariant Markov Basis for Sampling Contingency Tables with Fixed Marginals", Annals of the Institut of Statistical Mathematics, pp.230-256, 2006.
- [10] S. Aoki, A. Takemura, "The largest group of invariance for Markov bases and toric ideals", J. Symbolic Computation 43 (5) (2008) 342–358.
- [11] S. Sullivant, J. Rauh, T.Kahle. "Positive Margins and Primary Decomposition", North Carolina State, pp.1-22, 2012.
- [12] W. J. Gilbert," Modern Algebra With Application ",John Wiley, Inc., New Jersey, U.S.A, 2004.