Research Article   MATHEMATICS

Applying Rough Set Theory on DNA Recombination

M. M. Elsharkasy*, M. Sh. Badr**and Wafaa M. Fouda***

* Department of Mathematics, Faculty of Science, Tanta University.
** Department of Mathematics, Faculty of Science, Assuit University, New Valley.
*** Faculty of Pharmacy, Tanta University.

Abstract: The generalized structure of DNA recombination is based on the rules of rough Set. In this paper we are discussing the relationship between rough Set and DNA recombination. We construct a new recombination operator using the properties of DNA and RNA. Using the process of cutting and sticking of a sequence of genes, new types of topological structures are constructed and some of their properties and characterization are investigated. We studied recombination operators in the statement" The restriction enzyme cuts both molecules at the same sequence ("Sticky" End)".

Key words: Recombination, DNA, Topology, Rough Set, Sticky ends.

1. Introduction and Preliminaries

The biological application of topology to the study of DNA structure and understanding protein–DNA interactions that involve alterations of DNA topology is an essential aspect of the existence of every living cell because of the extraordinary degree to which the genomes of free-living organisms are confined. Moreover, changes in DNA topology accompany a wide range of enzyme-mediated processes on DNA such as replication, recombination, and repair [31].

DNA is often like a recipe or a code, since it contains the instructions needed to construct other components of cells, such as proteins and RNA molecules. The DNA segments that carry this genetic information are called genes. Within cells, DNA is organized into long structures called chromosomes [27].

DNA is a long polymer made from repeating units called nucleotides [3,17,29].Although each individual repeating unit is very small, DNA polymers can be very large molecules containing millions of nucleotides [20].

Recombinant DNA are molecules constructed outside of living cells by joining natural or synthetic DNA segments to DNA molecules that can replicate in a living cell, or molecules that result from their replication. This technique involved in creating and purifying desired genes.

The process of recombination makes it possible to cut different strands of DNA, in with a restriction enzyme and join the DNA molecules together via complementary base pairing [4].

In this study paper, we consider some methods for generating topologies by using rough set theory via one of biological applications (DNA recombination processes) and getting new results, finally exploring the

extent of matching between mathematical and biological results.

1.2 DNA Recombination

The following is a summary of the process of making recombinant DNA (see Figure 1):

1. Treat the DNA taken from both sources with the same restriction end nuclease.
2. The restriction enzyme cuts both molecules at the same site, a sticky Ends for example (EcoRI)) (GAATTC) and at blunt ends (Hpa 1) for example(GAATTC).
3. The ends of the cut have an overhanging piece of single-stranded DNA called "sticky Ends."
4. These sticky ends are able to base pair with any DNA molecule that contains the Complementary sticky end.
5. Complementary sticky ends can pair with each other when mixed.
6. DNA ligase is used to covalently link the two strands into a molecule of recombinant DNA.
7. In order to be useful, the recombinant DNA needs to be replicated many times (i.e. cloned). Cloning can be done in vitro, via the Polymerase Chain Reaction (PCR), or in Vivo (inside the cell) using unicellular prokaryotes (e.g. E. coli), unicellular eukaryotes (e.g. yeast), or mammalian tissue culture cells. Some examples of the therapeutic products made by recombinant DNA techniques include:
a. Blood proteins: Erythropoietin; Factors VII, VIII, IX; Tissue plasminogen activator; Urokinase.
b. Human Hormones: Epidermal growth factor; Follicle stimulating hormone; Insulin; Nerve growth factor; Relaxing; Somatotropin

Figure (1) A pictorial representation of the recombinant DNA process

A new mathematical representation is proposed for the configuration space structure induced by recombination which is called " P-structure ". It consists of mapping of pairs of objects to power set of all objects in the search space. The mapping assigns to each pair of parental "genotypes" the set of all recombinant genotypes obtainable from the parental ones. P.F.stadler conclude spaces that the algebraic approach to fitness landscape analysis can be extended to recombination spaces and provides an effective way to analyze the relative

hardness of a landscape for a given recombination operator.
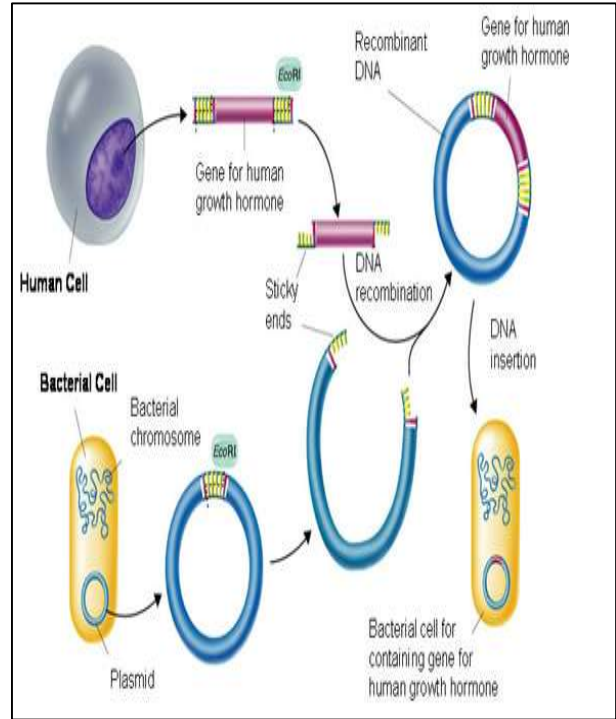
1.3 The DNA Recombination Operator

Definition1.3.1[7,8,9,11,12] Let X be any set which represents the set of *types* that it may be strings of bites, vectors, DNA, RNA sequence .etc. A definition of a recombination operator on X defined by.

$$T: X \times X \to P(X)$$

The recombination operator maps every pair of X to a subset of X (i.e. .an element of the power set $P(X)$ such that $\forall s, t \in X$, the following condition holds:

$(i) T(s,t) = T(t,s)$

$(ii) T(s,s) = \{s\}$

$(iii) \{s,t\} \subseteq T(s,t)$

$(iv) \|T(s,u)\| \leq \|T(s,t)\|\ for \forall u \in T(s,t).$

Definition 1.3.2[17]

Let X be a set of possible genotypes and A is a set of realized genotypes, a fixed collection of genetic operators such as mutation, recombination, gene-rearrangement and crossover, then k(A) is the genotypes accessible from A which satisfies the

following properties.

(i). No spontaneous creation (i.e.) $k(\varnothing) = \varnothing$.

(ii). A more diverse population produces more diverse off springs: (i.e.) $A \subset B$ implies $k(A) \subset k(B)$; $A, B \subset X$.

(iii). All parental genotypes are also accessible in the next time step. (i.e.) $A \subset k(A)$.

(iv). Diversity of offspring depends only on the parent. (i.e.) $k(A) = \bigcup_{x,y \in A} K(A)$ s.t. $A \subseteq X$.

1.4 Rough set theory

Rough set theory [38] is a recent approach for reasoning about data. It has achieved a large amount of applications in various real-life fields, like medicine, pharmacology, banking, market research, engineering, speech recognition, material science, information analysis, data analysis, data mining, control and linguistics (see the bibliography of [39]) and The main idea of rough sets corresponds to the lower and upper set approximations. These two approximations are exactly the interior and the closure of the set with respect to a certain topology τ on a collection U of imprecise data acquired from any real-life field. The base of the topology τ is formed by equivalence classes of an equivalence relation E defined on U using the available information about data. Following the connection between rough set concepts and topological notions, we investigate new definitions of the lower and upper approximation operators for similarity relation R. The equivalence class may be replaced by an element of the base $\{< p > R | p \;\square\; U\}$ of the topology τ. It generalizes Pawlak's approach and other extensions [10, 13, 30, 33, 34, 35, 36, 37]. It can be also compared with other similarity-based generalizations of rough sets.

The notion of approximation spaces is one of the fundamental concepts in the theory of rough sets. This section presents a review of the Pawlak approximation space constructed from an equivalence relation and its generalization using any binary relations.

Suppose U is a finite and nonempty set called the universe. Let $E \subset U \times U$ be an equivalence relation on U. The pair (U, E) is called an approximation space [38, 39]. Let $[x]_E$ denote the class of x such that $[x]_E = \{y \in U : x\,E\,y\}$. Then the lower and upper approximation of a subset X of U are defined as

$\underline{E}(X) = \{x \in U: [x]_E \subset X\}$

$\overline{E}(X) = \{x \in U: [x]_E \cap X \neq \emptyset\}$

A rough set is the pair $(\underline{E}(X), \overline{E}(X))$. Obviously, we have $\underline{E}(X) \subset X \subset \overline{E}(X)$. The lower approximation of X contains the elements x such that all the elements that are indistinguishable from x are in X. The upper approximation of X contains the elements x such that at least one element that is indistinguishable from x belongs to X. This definition can be extended to any relation R, leading to the notion of generalized approximate space [32]. let xR be the right neighborhoods defined as

$xR = \{y \in U : xRy\}$

The lower and upper approximations of X according to R are then defined as

$\underline{R}(X) = \{x \in U: xR \subset X\}$

$\overline{R}(X) = \{x \in U: xR \cap X \neq \emptyset\}$

Obviously, if R is an equivalence relation, $xR = [x]_R$ and these definitions are equivalent to the original Pawlak's definitions. We list the properties that are of interest in the theory of rough sets.

L1. $\underline{R}(X) = [\underline{R}(X^c)]^c$, where X^c denotes the complementation of X in U.

L2. $\underline{R}(U) = U$.

L3. $\underline{R}(X \cap Y) = \underline{R}(X) \cap \underline{R}(Y)$.

L4. $\underline{R}(X \cup Y) \supset \underline{R}(X) \cup \underline{R}(Y)$.

L5. $X \subset Y \Rightarrow \underline{R}(X) \subset \underline{R}(Y)$.

L6. $\underline{R}(\emptyset) = \emptyset$.

L7. $\underline{R}(X) \subset X$.

L8. $X \subset \underline{R}(\overline{R}(X))$.

L9. $\overline{R}(X) \subset \underline{R}(\overline{R}(X))$.

U1 $\overline{R}(X) = [\underline{R}(X^c)]^c$.

U2. $\overline{R}(\emptyset) = \emptyset$.

U3. $\overline{R}(X \cup Y) = \overline{R}(X) \cup \overline{R}(Y)$.

U4. $\overline{R}(X \cap Y) \subset \overline{R}(X) \cap \overline{R}(Y)$.

U5. $X \subset Y \Rightarrow \overline{R}(X) \subset \overline{R}(Y)$.

U6. $\overline{R}(U) = U$.

U7. $X \subset \overline{R}(X)$.

U8. $\overline{R}(\overline{R}(X)) \subset X$.

U9. $\overline{R}(\overline{R}(X)) \subset \overline{R}(X)$.

U10. $\overline{R}(\underline{R}(X)) \subset \underline{R}(X)$.

K. $\underline{R}(X^c \cup Y) \subset \underline{R}(X)^c \cup \underline{R}(Y)$.

LU. $\underline{R}(X) \subset \overline{R}(X)$.

Definition 1.41. [2] Let R be any binary relation on U, a set $< p > R$ is the intersection of all right neighborhoods containing p, i.e., $< p > R = \bigcap_{p \in xR}(xR)$

Definition 1.4.2[1]. Let R be any binary relation on U, The lower and upper approximations of X according to R are then defined as

$\underline{R}(X) = \{x \in U :< x > R \subset X\}$

$\overline{R}(X) = \{x \in U :< x > R \cap X \neq \emptyset\}$

Proposition 1.4.1[1] For any binary relation R on a nonempty set U the following conditions hold for every $X \subset U$.

(i) $\underline{R}(X) = [\underline{R}(X^c)]^c$.

(ii) $\underline{R}(U) = U$.

(iii) $\underline{R}(X \cap Y) = \underline{R}(X) \cap \underline{R}(Y)$.

(iv) $X \subset Y \Rightarrow \underline{R}(X) \subset \underline{R}(Y)$.

(v) $\underline{R}(X \cup Y) \supset \underline{R}(X) \cup \underline{R}(Y)$.

(vi) $\underline{R}(X) \subset \underline{R}(\underline{R}(X))$

Definition 1.4.3[20] Let X is a non-empty set and a closure operator $cl : P(X) \rightarrow P(X)$ such that:

K_0: $CL(\emptyset) = \emptyset$

$K_1 : If \ A \subseteq B \Rightarrow CL(A) \subseteq CL(B)$ (Isotonic)

$K_2 : A \subseteq CL(A)$
(Expanding)

$K_3 : CL(A \cup B) = CL(A) \cup CL(B)$
(Sub-additive)

$K_4 : CL(CL(A)) = CL(A)$
(Idempotent)

$K_5 : \cup_{i \in I} CL(A_i) = CL(\cup_{i \in I}(A_i))$ (additive)

1.5 Generalized DNA Recombination Topological Space[19]

We aimed in this work is to use the topological concepts in constructing flexible mathematical models in the field of biomathematics. We constructed two new recombination operators using the properties of DNA and RNA. We study topological properties of the constructed operators and the associated topological spaces of DNA and RNA. Several examples are discussed to illustrate the new concepts. *Cut and Sticks * for sequence of genotypes new types of topological structures are constructed and some of their properties and characterization are investigated.

The definition of process recombination mathematically using a matrix can be   cut and the integration of two types. We have been made a definition of the process of recombination with two ways: first, we take the first sequence needed only and ignore the rest .The other way we take all parts and as a result of improved optimization of this trace reappeared in the form of the space and that he began the installation process is described in a more accurate.

Well the way recombination between genes and we have greater accuracy and better place, and how cutting separation from the rest of the injured part.

Definition 1.5. 1 general topological DNA recombination operator[19]

Let X set of "types", which may be strings of bits, vectors, DNA or RNA sequences etc. And $span \ of \ X$ contains all the linear combination elements of a X the

outputs as well as recombination.

We define the recombination function

$R_s : X \times X \rightarrow span \{X\}$.

Such that

$R_s(x,y) = \cup_{i=1}^{n} \left\{ c_i^* x_{3'}^{5'} + c_j^* x_{5'}^{3'}, c_{n-i}^* y_1 + c_{n-j}^* y_2 \right\}$, the mapping represents a general mapping.

Since $C_i^* = \begin{pmatrix} I & \cdots & O \\ \vdots & \ddots \ I & \vdots \\ O & \cdots & O \end{pmatrix}$ I.E.

$$C_i^* = \begin{pmatrix} 1\,0\,0\,0 & \cdots & 0\,0\,0 & \cdots \\ 0\,1\,0\,0 & \cdots & 0\,0\,0 & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ 0\,0\,0 & \cdots 1 & 0\,0\,0 & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ 0\,0\,0 & \cdots & 0\,0\,0 & \cdots \end{pmatrix}$$

Where the matrix represent the unity of level $i \times n$, O a zero matrix, c_i^* is called Matrix Hacking and $c_i^* \in M_n(f)$[Boolean matrix],for all $i \in \{2,4,6,8\}$.

We Will run the new Topological study on DNA recombination where recombination between types x, y since x first gene and y plasmid is the same between the x, y, this is called general topological recombination operator.

Proposition 1.5.1[19]

Let recombination operator

$R_s : X \times X \rightarrow span \{X\}$ Be an operator,

$R_s(x,y) = \cup_{i=1}^{n} \left\{ c_i^* x_{3'}^{5'} + c_j^* x_{5'}^{3'}, c_{n-i}^* y_1 + c_{n-j}^* y_2 \right\}$,

$R_s(x,y)$ Consists of all recombination products can be done or (offspring) can be obtained induced by x, y and satisfied:

$(i)\{x,y\} \subset R_s(x,y)$: That is recombination between "types", x and y forms a recombination set $R_s(x,y)$

$(ii) R_s(x,y) = R_s(y,x)$

$(iii) \forall z \in R_s(x,y) \Rightarrow \|R_s(x,z)\| \leq \|R_s(x,y)\|, \forall x,y,z \in B$.

$(iv) R_s(x,x) = 2x$ (Duplication of genes)

$(v) R_s(x,y) \subseteq$ Span$\{x,y\}$,[with change the enzymes], where span

$A = \{z \in B: \forall i: \exists x \in A: z_i = x_i\}$ Is the linear span .

A related study on the structure of the genotype spaces that is introduced by so-called unequal crossover

(or no homologous recombination) show that isotonic spaces are adequate for mutation spaces but behave somewhat 'un nature' for certain recombination spaces[19].

In this section, we present a definition of closure recombination space by using recombination function

Definition 1.5.2 closure recombination operator[19]

Let a subset A of X. We take $T^*_{R_s}(A)$ represents the closure operator since

$T^*_{R_s}(A) = \cup_{x,y} R_s(x,y)$ s.t. $A \subseteq X$. We work in standard topological recombination operator. The resultant closure-space from the recombination operator R(x, y) is denoted by $(X, T^*_{R_s})$.

Theorem 1.5.1[19]

The closure space (c (n), $T^*_{R_s}(A)$) arising from recombination function satisfies

(i) $T^*_{R_s}(\emptyset) = \emptyset$

(ii) If A is a set of genes, then $A \subseteq T^*_{R_s}(A)$.

(iii) If $A \subseteq B \longrightarrow T^*_{R_s}(A) \subseteq T^*_{R_s}(B)$

(iv) $T^*_{R_s}(A) \cup T^*_{R_s}(B) \subseteq T^*_{R_s}(A \cup B)$ and

$\qquad T^*_{R_s}(A \cap B) \subseteq T^*_{R_s}(A) \cap T^*_{R_s}(B)$.

Resulting in a space $(X, T^*_{R_s})$, who will we call R_s DNA recombination achieve some topological properties which are useful for process a recombination of DNA. In the table a space is constructed Neighborhood space.

2. Upper and Lower Approximation on DNA Recombination

Following the connection between rough set concepts and DNA recombination, we investigate new definitions of the class of element depend on definitions of recombination set which resulting from the definition of recombination function which identifier as follows

$R_s : X \times X \to \text{span}\{X\}.$

Such that

$$R_s(x,y) = \cup_{i=1}^{n} \left\{ c_i^* x_{3'}^{5'} + c_j^* x_{5'}^{3'}, c_{n-i}^* y_1 + c_{n-j}^* y_2 \right\},$$

Definition Let $R_s(x,y)$ be the recombination set consists of all recombination products can be done or (offspring) can be obtained induced by x, y .

Definition Let $R_s(x,y)$ be a recombination set. Then the recombination class of x can be defined as $[x]_{R_s} = \{y \in U : y \in R(x,y)\}$.
Then the lower and upper approximation of a subset X of U are defined as
$\underline{R_s}(X) = \{x \in U: [x]_{R_s} \subset X\}$
$\overline{R_s}(X) = \{x \in U: [x]_{R_s} \cap X \neq \varphi\}$.

Example 1 Let the recombination process consists of three key pillars are: genes, plasmids and enzymes i.e. U = {g , e, p}, the recombination class of gene [G]ᵣ = {E},[P]ᵣ = {E,G} and [E]ᵣ = {∅} (see , Table , No 1)

A	$R_-(A)$	$R^-(A)$	$R_-(R^-(A))$	$R_-(R_-(A))$	A^c	$R^-(A^c)$	$[R^-(A^c)]^c$	$R_-(A^c)$	$[R_-(A^c)]^c$	$R^-R_-(A)$	$R^-R^-(A)$
{G}	{E}	{P}	{E}	{G,E}	{P,E}	{G,P}	{E}	{G,E}	{P}	{G,P}	∅
{P}	{E}	∅	∅	{G,E}	{G,E}	{G,P}	{E}	U	∅	{G,P}	∅
{E}	{G,E}	{G,P}	{E}	U	{G,P}	{P}	{G,E}	{E}	{G,P}	{G,P}	{P}
{G,P}	{E}	{P}	{E}	{G,E}	{E}	{G,P}	{E}	{G,E}	{P}	{G,P}	∅
{G,E}	U	{G,P}	{E}	U	{P}	∅	U	{E}	{G,P}	{G,P}	{P}
{P,E}	{G,E}	{G,P}	{E}	U	{G}	{P}	{G,E}	{E}	{G,P}	{G,P}	{P}
U	U	{G, P}	{E}	U	∅	∅	U	∅	U	{G,P}	{P}
∅	∅	∅	∅	∅	U	{G,P}	{E}	U	∅	∅	∅

(Table , No 1)

Proposition 2.1 For any recombination class $[x]_{R_s}$ on nonempty set U the following conditions hold for every X, Y $\square$ U.

(i) $\underline{R}(U) = U$, $\underline{R}(\emptyset) = \emptyset$, $\overline{R}(\emptyset) = \emptyset$, $\overline{R}(U) \subseteq U$

(ii) $\underline{R}(X \cap Y) = \underline{R}(X) \cap \underline{R}(Y)$.

(iii) $X \square Y \Rightarrow \underline{R}(X) \square \underline{R}(Y)$., $R^-(X) \square \overline{R}(Y)$

(iv) $\underline{R}(X \cup Y) \square \underline{R}(X) \cup \underline{R}(Y)$.

(v) $\underline{R}(X) \square [\overline{R}(X^c)]^c$.

(vi) $\underline{R}(\overline{R}(X)) \square \underline{R}X))$.

(vii) $\overline{R}(X)) \square \overline{R}(\underline{R}(X))$

(viii) $\overline{R}(X) \square [\underline{R}(X^c)]^c$

(ix) $\overline{R}(X^c) \square \overline{R}(\underline{R}(X))$

Proof. (i) Since for every $x \in U$, $[x]_{R_s} \subset U$ hence, $x \in \underline{R}(U)$. Then $U \subset \underline{R}(U)$. Also since $\underline{R}(U) \subset U$. Thus, $\underline{R}(U) = U$., also there is no x $\in U$ such that $[x]_{R_s} \subset \varphi$ hence $\underline{R}(\varphi) = \varphi$.

(ii) $\underline{R}(X \cap Y) = \{x \in U : [x]_{R_s} \subset X \cap Y\}$
$= \{x \in U : [x]_{R_s} \subset X \wedge [x]_{R_s} \subset Y\}$
$= \{x \in U : [x]_{R_s} \subset X\} \cap \{x \in U : [x]_{R_s} \subset Y\}$
$= \underline{R}(X) \cap \underline{R}(Y)$.

(iii) Let $X \subset Y$ and $x \in \underline{R}(X)$, then $[x]_{R_s} \subset X$ and so $[x]_{R_s} \subset Y$, hence $x \in \underline{R}(Y)$. Thus $\underline{R}(X) \subset \underline{R}(Y)$.

(iv) Since $X \subset X \cup Y$ then $\underline{R}(X)$
$\subset \underline{R}(X \cup Y)$ also, Y
$\subset X \cup Y$ then

$\underline{R}(Y) \subset \underline{R}(X \cup Y)$, hence $\underline{R}(X \cup Y)$
$\supset \underline{R}(X) \cup \underline{R}(Y)$.

(v) $[R^-(X^c)]^c = \{x \in U : [x]_{R_s} \cap X^c$
$\neq \varphi\}^c$
$= \{x \in U : [x]_{R_s} \cap X^c = \varphi\} \supset \{x \in U : [x]_{R_s} \subset X\} \supset \underline{R}(X)$.

(vi) , (vii) , (viii) , (ix) ,and (x) The proof is the same as previous.

The set of all open recombination sets $= \{U, \emptyset, \}$ consist of a space called indiscrete recombination space satisfies the following:

Proposition 2.2.let (U,E) be an indiscrete recombination space, A$\subset$ U then the limit point of A

$$A' = \begin{cases} \emptyset & , \text{if } A = \emptyset \\ U - \{p\} & , \text{if } A = \{p\} \\ U & , \text{if A contains more then one element} \end{cases}$$

Proof : Obviously

These results means :

1- That there are always an end point of the process ,

2-there are an output of the process of recombination and this applicable with the biological concept.

Remark 2.1 . Every subset of an indiscrete recombination space is dense.

The sense that it produces a very large number of plasma carrying the gene

The aim of the following Preposition is described can be separated any item in the recombination process.

Preposition 2.3. Every indiscrete recombination space is regular space

Note: Since we use in recombination process {gene, plasmids, enzymes} i.e.

1. Gene +Enzyme=G-h since h The lump of gene.

2. Plasmid+ Enzyme=P-h.

3. 1+2=P+G.

Example 2.2.If the use of plasmid as an aid since U =$\{g, e, p\}$,hence the class of $[G]_R$ =$\{P\}$,$[P]_R = \{\emptyset\}$ and$[E]_R = \{G, P\}$,

$\underline{R_s}(X) = \{x \in U: [x]_{R_s} \subset X\}$

$\overline{R_s}(X) = \{x \in U: [x]_{R_s} \cap X \neq \varphi\}$. (see , Table , No 2)

A	$R\_(A)$	$R^-(A)$	$R\_(R^-(A))$	$R\_(R\_(A))$	A^c	$R^-(A^c)$	$[R^-(A^c)]^c$	$R^-(R\_(A))$	$R\_(A^c)$	$[R\_(A^c)]^c$	$R^-R^-(A)$
{G}	{P}	{E}	{P}	{G,P}	{P,E}	{G,E}	{P}	{G,E}	{G,P}	{E}	∅
{P}	{G,P}	{G,E}	{{P}}	U	{G,E}	{E}	{G,P}	{G,E}	{P}	{G,E}	{E}
{E}	{P}	∅	∅	{G,P}	{G,P}	{G,E}	{P}	{G,E}	U	∅	∅
{G,P}	U	{G,E}	{P}	U	{E}	∅	U	{G,E}	{P}	{G,E}	{E}
{G,E}	{P}	{E}	{P}	{G,P}	{P}	{G,E}	{P}	{G,E}	{G,P}	{E}	∅
{P,E}	{G,P}	{G,E}	{P}	{G,P}	{G}	{E}	{G,P}	{G,E}	{P}	{G,E}	{E}
U	U	{G,E}	{P}	U	∅	∅	U	{G,E}	U	∅	{E}
∅	∅	∅	∅	∅	U	{G,E}	{P}	∅	∅	U	∅

(Table , No, 2)

The set of all open recombination sets = {U, ∅, } consist of a space called indiscrete recombination space

Proposition 2.4. For any (U,R) be an indiscrete recombination space a nonempty set U the following conditions hold for every X, Y ⊂ U.

(i) $\underline{R}(U) = U$, $\underline{R}(\emptyset) = \emptyset$, $\overline{R}(\emptyset) = \emptyset$, $\overline{R}(U) \subseteq U$

(ii) $\underline{R}(X \cap Y) = \underline{R}(X) \cap \underline{R}(Y)$.

(iii) $X \subset Y \Rightarrow \underline{R}(X) \subset \underline{R}(Y)$., $\overline{R}(X) \subset \overline{R}(Y)$

(iv) $\underline{R}(X \cup Y) \supset \underline{R}(X) \cup \underline{R}(Y)$.

(vi) $\underline{R}(X) \subset [\overline{R}(X^c)]^c$.

(vii) $\underline{R}(\overline{R}(X)) \subset \underline{R}(X))$.

(viii) $\overline{R}(X)) \subset \overline{R}(\underline{R}(X))$

(ix) $\overline{R}(X)) \subset [\underline{R}(X^c)]^c$

(x) $\overline{R}(X^c) \subset \overline{R}(\underline{R}(X))$

Proof. The proof is the same as for proposition **2.1**

The set of all open sets = {U , ∅ }

The set of all closed sets = {U , ∅, {G},{G,P} }

3. Rough Set Theory for DNA recombination

In the following part we use the lower and upper approximation as**:**

$R_s(X) = X \cap \{x \in U: [x]_{R_s} \subset X\}$

$\overline{\overline{R}}_s(X) = X \cup \{x \in U: [x]_{R_s} \cap X \neq \varphi\}$, $\underline{R_s}(X) \equiv \underline{R}(X)$ and

The recombination process consists of three key pillars are: genes, plasmids and enzymes i.e.U = {g ,e, p},the recombination class of gene $[G]_R = \{E\}$,$[P]_R = \{E,G\}$ and$[E]_R = \{\emptyset\}$(see, Table , No 3)

A	$R\_(A)$	$R^-(A)$	$R\_(R^-(A))$	$R\_(R\_(A))$	A^c	$R^-(A^c)$	$[R^-(A^c)]^c$	$R\_(A^c)$	$[R\_(A^c)]^c$	$R^-R\_(A)$	$R^-R^-(A)$
{G}	∅	{G,P}	∅	∅	{P,E}	U	∅	{E}	{G,P}	∅	{G,P}
{P}	∅	{P}	∅	∅	{G,E}	U	∅	{G, E}	{P}	∅	{P}
{E}	{ E}	U	U	{ E}	{G,P}	U	∅	∅	U	U	U
{G,P}	∅	{G,P}	∅	∅	{E}	U	∅	{ E}	{G,P}	∅	{G,P}
{G,E}	{G,E}	U	U	{G,E}	{P}	{P}	{G,E}	∅	U	U	U
{P,E}	{E}	U	U	{ E}	{G}	{G,P}	{ E}	∅	U	U	U
U	U	U	U	U	∅	∅	U	∅	U	U	U
∅	∅	∅	∅	∅	U	U	∅	U	∅	∅	∅

(Table , No 3)

Proposition 3.1. For any $[x]_{R_s}$ is a recombination class on nonempty set U the following conditions hold for every X, Y ⊂ U.

L1. $\underline{R}$ (X ∩ Y) = $\underline{R}$ (X) ∩$\underline{R}$ (Y).

L2. $\underline{R}$ (X ∪ Y) ⊃$\underline{R}$ (X) ∪$\underline{R}$ (Y).

L3. X ⊂ Y ⇒$\underline{R}$ (X) ⊂$\underline{R}$ (Y).

L4. $\underline{R}$ (φ) = φ.

L5. $\underline{R}(X)$ ⊂ X.

U1$\overline{R}$ (X) = $[\underline{R}(X^c)]^c$.

U2. $\overline{R}$ (φ) = φ.

U3. $\overline{R}$ (X ∪ Y) = $\overline{R}(X)$ ∪$\overline{R}(Y$).

U4. $\overline{R}(X ∩ Y)$ ⊂$\overline{R}$ (X) ∩$\overline{R}(Y$).

U5. X ⊂ Y ⇒$\overline{R}(X)$ ⊂$\overline{R}$ (Y).

U6. $\overline{R}(U)$ = U.

U7. X ⊂$\overline{R}$ (X).

U8. $\overline{R}$ ($\overline{R}$ (X)) =$\overline{R}(X)$.

LU. $\underline{R}(X)$ ⊂$\overline{R}$ (X).

Proof. The proof is the same as for proposition 2.1

4. General rough Set theory for DNA recombination

The recombination process consists of three key pillars are: genes, plasmids and enzymes. In this way we infer mathematical relationships between the components of the process of recombination the process and the formation of the class of elements by $xR = [x]R = \{y \in U: xE\ y\}$

Example 4.1 Let U = {G ,E, P} and
R = {(G,G), (E,E), (P,P), (G,E), (P,E), (G,P) }.
Then: g R= {U} , p R ={P ,E} ,   e R = {E }
,$\overline{R}(A) = \{x: R_x ∩ A ≠ Ø \}$,$\underline{R}(A) = \{x: R_x ⊂ A\}$
(see ,Table , No 4)

A	$R^-(A)$	$R_-(A)$	$R_-(R_-(A))$	$R_-(R^-(A))$	$R^-(R_-(A))$
{G}	{G}	Ø	Ø	Ø	Ø
{P}	{G,P}	Ø	Ø	Ø	Ø
{E}	{E}	{E}	{E}	{E}	{E}
{G,P}	{G,P}	Ø	Ø	Ø	Ø
{G,E}	U	{E}	{E}	U	{E}
{P,E}	U	{P,E}	{P,E}	U	U
U	U	U	U	U	U
Ø	Ø	Ø	Ø	Ø	Ø

(Table, No 4)

Proposition 4.1 If a binary relation R on U is a reflexive relation, and then the following conditions hold.

(i) $\underline{R}(U)$ = U, $\underline{R}(Ø) = Ø$, $\overline{R}(Ø) = Ø$, $\overline{R}(U) = U$

(ii) $\underline{R}(X) ⊂ X ⊂ \overline{R}(X)$.

(iii)$\underline{R}\left(\overline{R}(X)\right) ⊂ \overline{R}(X)$,$\underline{R}(X) ⊂ \underline{R}(\overline{R}(A))$,

(iv) X⊂ $\underline{R}(\overline{R}(A))$.

(v) $\overline{R}$ (X ∪ Y) ⊃ $\overline{R}(X) ∪ ∪ \overline{R}$ (Y).

(vi) X □ Y ⇒ $\overline{R}(X) ⊂ \underline{R}(Y), \overline{R}(X) ⊂ \overline{R}(Y)$

(vii) $\underline{R}$ (X ∩ Y) = $\underline{R}$ (X) ∩$\underline{R}$ (Y).,, $\overline{R}$ (X ∩ Y) ⊂ $\overline{R}(X)$ ∩$\overline{R}(Y)$.

Proof. The proof is the same as for proposition **2.1**

The set of all open recombination sets = {U, Ø, {E},{P,E} },The set of all closed recombination sets = {U, Ø, {G},{G,P} }

Example 4.2. In this way we infer mathematical relationships between the components of the process of recombination the process ,change with working steps and the formation of the class of elements by xR=[x]R={y□U: xE y} ,since set U = {g ,e, p}

R = {(G,G), (E,E), (P,P), (E,G), (E,P), (G,P) }

g R= {G,P}   P R ={P }   e R = {U}

A	$R^-(A)$	$R_-(A)$
{G}	{G,E}	∅
{P}	U	{P}
{E}	{E}	∅
{G,P}	U	{G,P}
{G,E}	{G,E}	∅
{P,E}	U	{P}
U	U	U
∅	∅	∅

The set of all open recombination sets =
{U , ∅, {P},{G,P} }is T_0 space

The set of all closed recombination sets =
{U , ∅, {E},{G,E} }

Example 4.3. In this way we infer mathematical relationships between the components of the process of recombination the process ,change with working steps and the formation of the class of elements by [x]R= Rg= {x: xRg} since set U = {G ,E, P}

R = {(G,G), (E,E), (P,P), (E,G), (E,P), (G,P) }

Rg= {G,E}　　　Rp ={U }　　Re = {E}

A	$R^-(A)$	$R_-(A)$
{G}	{G,P}	∅
{P}	{P}	∅
{E}	U	{E}
{G,P}	{G,P}	∅
{G,E}	U	{G,E}
{P,E}	U	{E}
U	U	U
∅	∅	∅

The set of all open recombination sets =
{U , ∅, {e},{g,e} }

The set of all closed recombination sets =
{U, ∅, {p},{g,p} }

We can conclude that results of both example (4.2) and (4.3) are equivalent to that example (4.1). Where sets of all open

recombination sets in this group only two item of the three appear.

Conclusions

And the process of recombination is appear in the industrialization pharmaceutical and gene therapy, and we are working on a plan (draft) for a work program to help the biologists, the program consists of three phases (operations). The first process is the recombination, the second is determine mutation and third process of reform or Simulation, on the occasion the finished of the recombination process, this article talks about the process characterization and conformity mathematical results of biological results. Some of the suggestions to use rough set in future work.

References:

[1] A.A Allam, , M.Y Bakeir , E.A Abo-Tabl: New Approach for Basic Rough Set Concepts. Submitted.

[2]　Allam, A.A., Bakeir, M.Y., Abo-Tabl, E.A.: A relational view to some basic topological concepts. Submitted.

[3] Alberts, Bruce; Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts and Peter Walters (2002). Molecular Biology of the Cell; Fourth Edition. New York and London: Garland Science.

[4] Amy B. Vento and David R. Gillum Office of Environmental Health and Safety University of New Hampshire June 3, 2002.

[5] B.M.R Sadler, P.F. Sadler, M. Shapak and G.P .Wagner, Recombination space,metrics and pretopologies, Z.phy. 2001, in press, SFI preprint 01-02-001.

[6] B.M.R Sadler, P.F. Sadler, Generalized topological spaces in evolutionary theory and combinatorial chemistry, J. Chem. Inf. Computer science,Vol.42, No. 3, 2002.

[7] C.Flamm, I. L. Hofacker, P.F.Sadler, RNA in silico: The computational Biology of RNA secondary structures. Adv. Complex syst. 1999, 2, 65-90.

[8] Christoph flamm,Ivol.Hofacker,Barbel M.R.stadler,and peter F.stadler: Saddles and barrier in landscapes of generalized search operators.FOGA 2007, lens 4436, pp.194-212, 2007

[9]Colin adams and Robert franzosa:Introduction to topology pure and applied.2009.

[10] D. Kim: Data classification based on tolerant rough set. Pattern Recognition, 34(2001) 1613- 1624

[11]E.C. Webb, editor. Nature Publishing Group. Encyclopedia of life sciences. http://www.els.net [an Internet encyclopedia with up-to-date overview articles in every field of biochemistry and cell biology]. Enzyme nomenclature 1992. San Diego: International Union of Biochemistry and Molecular Biology/Academic Press, 1992.

[12]E.Cech topological space, interscience publishers, John Wiley and sons, New York (1966)

[13] J.A. Pomykala: Approximation operations in approximation space. Bulletin of the Polish Academy of Sciences, Mathematics, 35 (1987) 653-662

[14] J. Cupal, P. Schuster and P.F. Saddler, "Topology in Phenotype Space" Computer Scienc Biology, GCB'99 Proceedings, University Bielefeld, Hannover, 1999, 9-15.

[15] J. Cupal, S. Kopp, and P.F. Saddler, "RNA Shape Space Topology" Alife 6 (2000), 3-23.

[16] J. Kruger and F. Vogel. Population genetics of unequal crossing over. J. Mol. Evol., 4:201-247, 1975.

[17]K.ChandrasekharaRao,R.Gowri,V.Swamin athan and S.Stalin. Advanced studies in biology, Vol. 1, no. 2, 95-104, and 2009.

[18] M. B .Smyth: Semi- metreic, Closure space and digital topology, Theor. Comput. Sci. 151,(1995)   275-276

[19]M. M. Elsharkasy, M. Sh. Badr, Wafaa M. Fouda: Topological overview of DNA recombination via closure space under submitted.

[20] M. R. Bremner: DNA computing, insertion of words and left-symmetric algebras .Proc.of the Maple Conf. 2005. Maple Inc. Waterloo. 2005.

[21] M. Shpak and G. p .Wagner. Asymmetry of configuration space induced unequal crossover: implications for a mathematical theory of evolutionary innovation. Artificial life, 6:25-43, 2000.

[22] P. Gitchoff and G. P. Wagner. Recombination induced hyper graphs: A new approach to mutation recombination isomorphism. (1996) 1-14.

[23] P. C. Hammer.; Extended topology, set-valued , set function . nieuw Arch. III,10, (1962), 55-77.

[24] P.F.Sadler,G.P.Wagner,The algebaric theory of recombination. Spaces Evol.Comp.1998,5,241- 275

[25]P.F. Sadler, R. Seitz, G. P. Wagner: Evolvability of complex characters, population dependent Fourier decomposition of fitness landscapes over recombination space, bull. Math.,boil,62, pp399-428, (2000)

[26] R. Slowinski, Vanderpooten, D.: A Generalized Definition of Rough Approximations Based on Similarity. IEEE Transactions on Knowledge and Data Engineering,12(2)(2000)331-336

[27] Russell, Peter . Genetics. New York: Benjamin Cummings. (2001)

[28] S .Gregory, K.F .Barlow, K.E .McLay, R .Kaul; D .Swarbreck, A. Dunham, CE. Scott, KL. Howe, et al. (2006). "The DNA sequence and biological annotation of human chromosome 1".Nature 441 (7091): 315--21..

[29] Saenger, Wolfram (1984). Principles of Nucleic Acid Structure. New York: Springer-Verlag.

[30] U. Wybraniec-Skardowska: On a generalization of approximation space. Bulletin

of the Polish Academy of Sciences, Mathematics, 37 (1989) 51-61

[31] V.A. Bloomfield, D.M. Crothers, and I.J. Tinoco, Nucleic AcidsStructures, Properties and Functions (University Science Books, Herndon, VA, 2000

[32] Y.Y. Yao: Two views of the theory of rough sets in finite universes. International

 Journal of Approximation Reasoning, 15 (1996) 291-317

[33] Y.Y. Yao: Constructive and algebraic methods of the theory of rough sets. Information Sciences, 109 (1998) 21-47

[34] Y.Y. Yao: Generalized rough set models. In: Rough Sets in Knowledge Discovery, Polkowski, L. and Skowron, A. (Eds.). Physica-Verlag, Heidelberg (1998) 286-318

[35] Y.Y. Yao: Relational interpretations of neighborhood operators and rough set approximation operators. Information Sciences, 111 (1998) 239-259

[36] Y.Y. Yao, T.Y Lin.: Generalized of rough sets using model logic. Intelligent Automation

 and Soft Computing, 2 (1996) 103-120

[37] Y.Y. Yao, S.K.M.Wong, T.Y Lin: A review of rough set models. In: Rough Sets

 and Data Mining: Analysis for Imprecise Data , T.Y. Lin and N. Cercone, (Eds).Kluwer Academic Publishers, Boston, (1997) 47-75

[38] Z. Pawlak: Rough sets. International Journal of Information and Computer Sciences,

 11 (1982) 341-356

[49] Z.Pawlak: Rough sets. Theoretical Aspects of Reasoning about Data. Kluwer

 Academic Publishers (1991)

الملخص العربي

الشكل المعمم لعملية إعادة التركيب (اعادة تركيب الحمض النووي ناقص الاوكسجين "DNA recombination ") تعتبر نواة لدراسة المجموعات الاستقرابية (rough set).

في هذه البحث نناقش العلاقة بين المجموعة الاستقرابية وعملية إعادة التركيب . واستخدام المجموعة الاستقرابية في التاكد من صحة داله عملية اعادة التركيب (بيولوجيا) .

وباستخدام مفهوم عملية التركيب والمجموعة الاستقرابية امكن ايجاد خواص رياضية جديدة ومطابقة للمعني البيولوجي لعملية

اعادة التركيب .