Introduction
to the
TRICHROMATIC THEORY
OF EQUILIBRIUM
OF SYSTEMS

applied to
DNA or RNA Sequence
Analysis and Modification

Author: ***Nunzio Bonaventura***
Computer Technician: ***Vincenzo Viggiano***

13 January 2020

I INTRODUCTORY
INDEX

I.1 Principal Concepts of the Trichromatic Theory of Equilibrium of Systems (T.T.E.S.)

The aim of this introduction is to briefly present the **TRICHROMATIC THEORY OF EQUILIBRIUM OF SYSTEMS** (**T.T.E.S.**) and its important application to the **DNA or RNA sequence analysis and modification** (http://www.ttesystems.eu/application.php).

The **T.T.E.S.** is a theory of systems through which it is possible to observe, analyse, control and modify the state of every system (http://www.ttesystems.eu/index_en.php).

With the **T.T.E.S.** and its software, all scientists, from different disciplines, can share an universal view of reality. They can also use, at different levels of analysis and for different systems, the same theory of systems and the methodology for data analysis so to communicate more easily and effectively on an interdisciplinary level.

Essential condition for the application of the **T.T.E.S.** is the identification of *three specific parameters representative* of the general functioning of the system (or the subsystem) that needs to be analysed.

The task of scientists is to identify only this three parameters (**X**, **Y** e **Z**), to acquire directly or from historical archives the experimental data of the selected parameters and to process them through the **T.T.E.S.** .

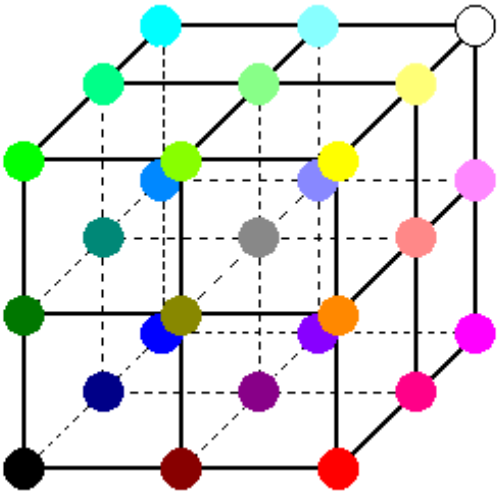
The **T.T.E.S.** was used for the first time to analyse the Vegetative Nervous System through the help of the peripheral biofeedback (The Peripheral Biofeedback and the Trichromatic Theory of Equilibrium of the Vegetative Nervous System;
The Future of Peripheral Biofeedback: The Trichromatic Theory of Equilibrium of the Vegetative Nervous System;
Hyperventilation: a privileged model for the quantitative and qualitative evaluation of the psychophysiological activation with the Trichromatic Theory of Equilibrium of the Vegetative Nervous System).

Many other applications of the **T.T.E.S.** are expected (http://www.ttesystems.eu/application.php) and some of them are being experimented and will be published in the future.

To calculate and visually represent every system variation, the **T.T.E.S.** has used the colour model RGB.

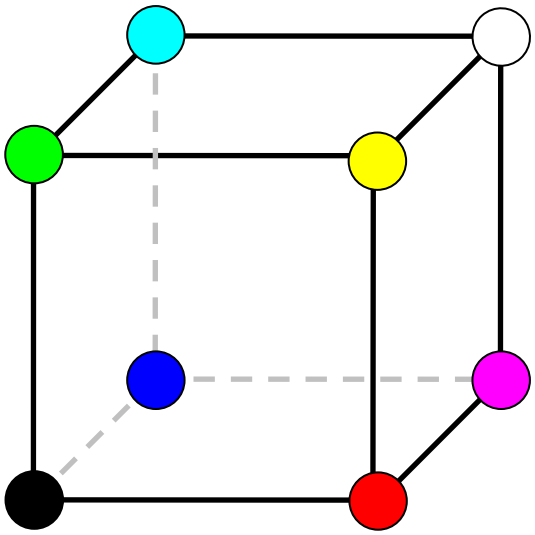
The RGB model is a method to define colours based on *THREE Primary colours* (Red, Green and Blue).

The **CUBE** is the solid used to visually represent all the possible variations of the state of a system (Pct. 1).



Pct. 1

In order to synthetically describe every possible functional state of a system, **8 PRINCIPAL CODES** (Pct. 2) of the **64 TOTAL CODES** have been used**.**



Pct. 2 *

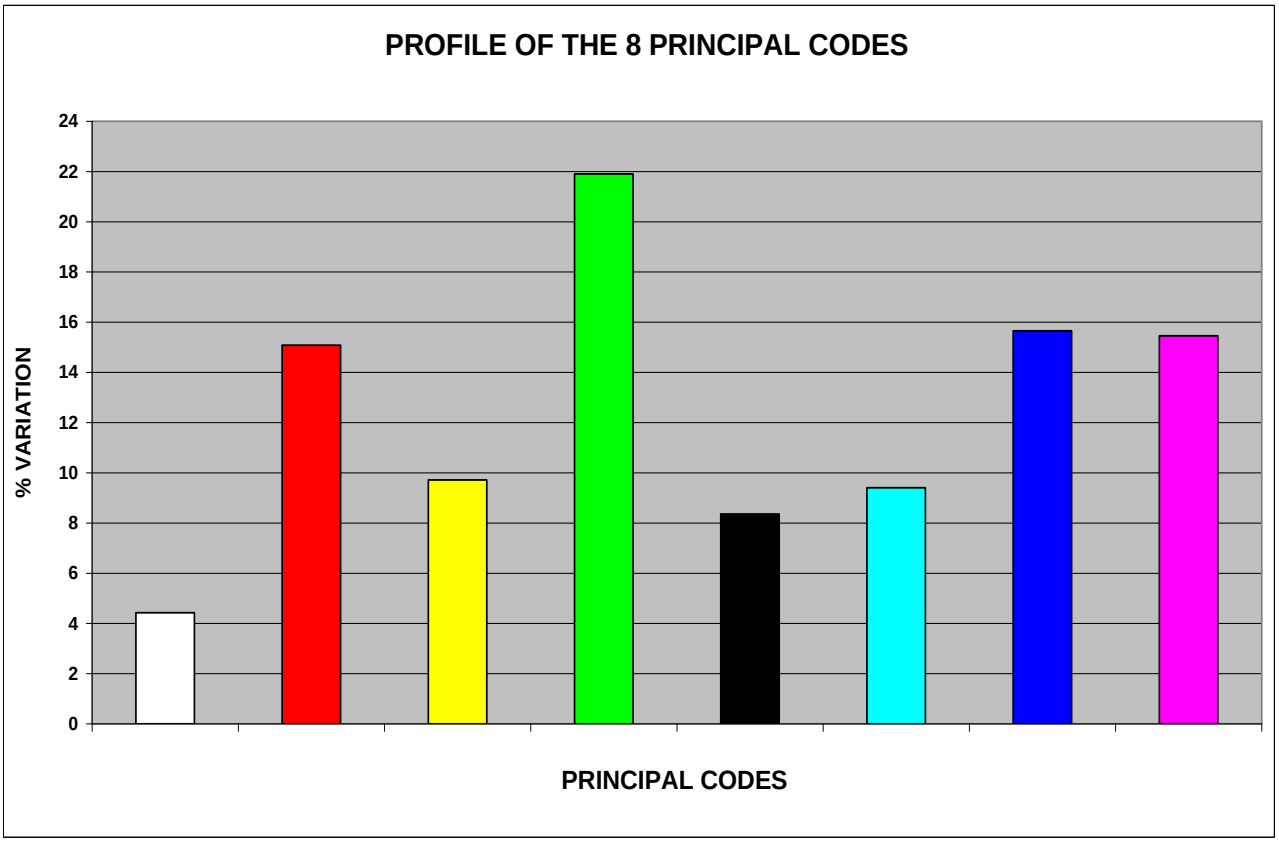
With <u>appropriate adaptations</u> of the basic **T.T.E.S.** software, it is possible to analyse and modify a DNA (or RNA) sequence in an *innovative way*.

* Excerpted and modified from: https://commons.wikimedia.org/wiki/File:Avl3119color4a.jpg

I.2 First Phase: Analysis of the Original Sequence

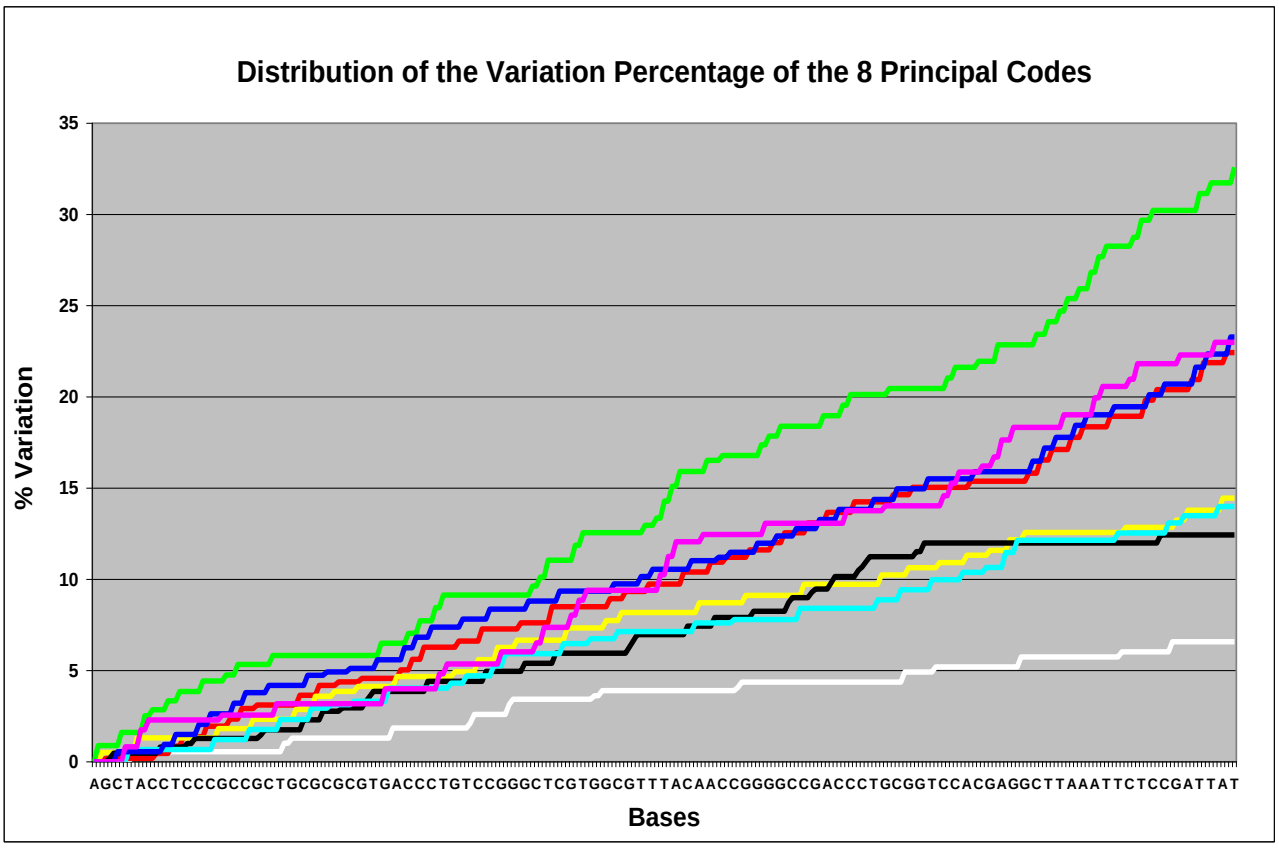
The **acquisition of the sequence** and its **graphic representation** constitute the **FIRST PHASE** of the study, that of the **ANALYSIS OF THE ORIGINAL SEQUENCE**.

The **acquisition of the DNA (or RNA) sequence** to be analysed and modified can be carried out directly from the *organism* studied, from the website of the **NCBI** [National Center for Biotechnology Information (1)] or from any other source.
After acquiring any DNA or RNA sequence (called **original sequence**), the **T.T.E.S.** software offers the possibility to obtain **4 principal graphic representations of the original sequence** [as case study see Pct. 3, Pct. 4, Pct. 5 and Pct. 6. These graphs refer to the Sequence XM_011721319.1 - *PREDICTED: Macaca nemestrina insulin (INS), transcript variant X4, mRNA*, obtained from significant alignments of the **Insulin A Chain**. The complete analysis of this sequence was published in the following document: DNA or RNA Sequence Analysis and Modification through the T.T.E.S. (Chapter II°)].

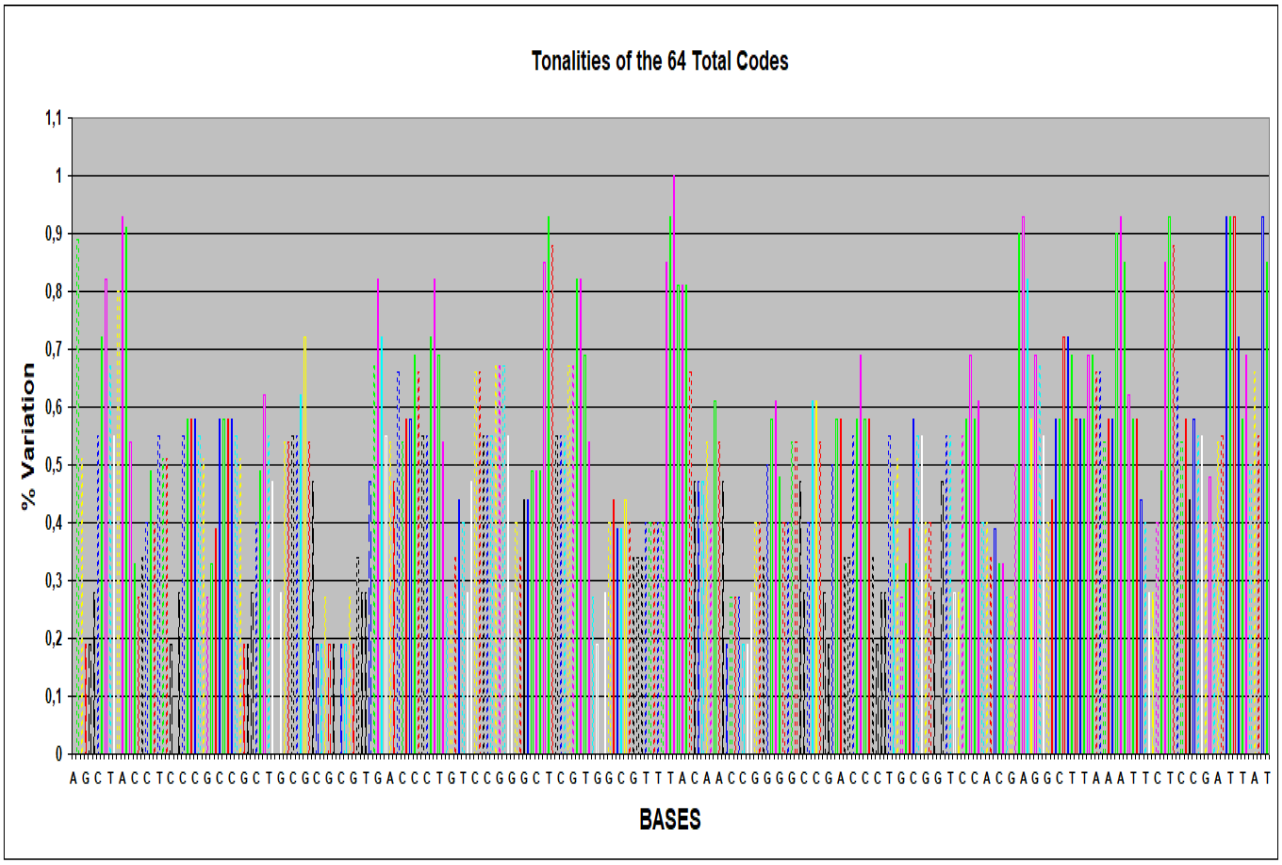


Pct. 3 (This graph constitutes a *very general synthesis* of the entire sequence)

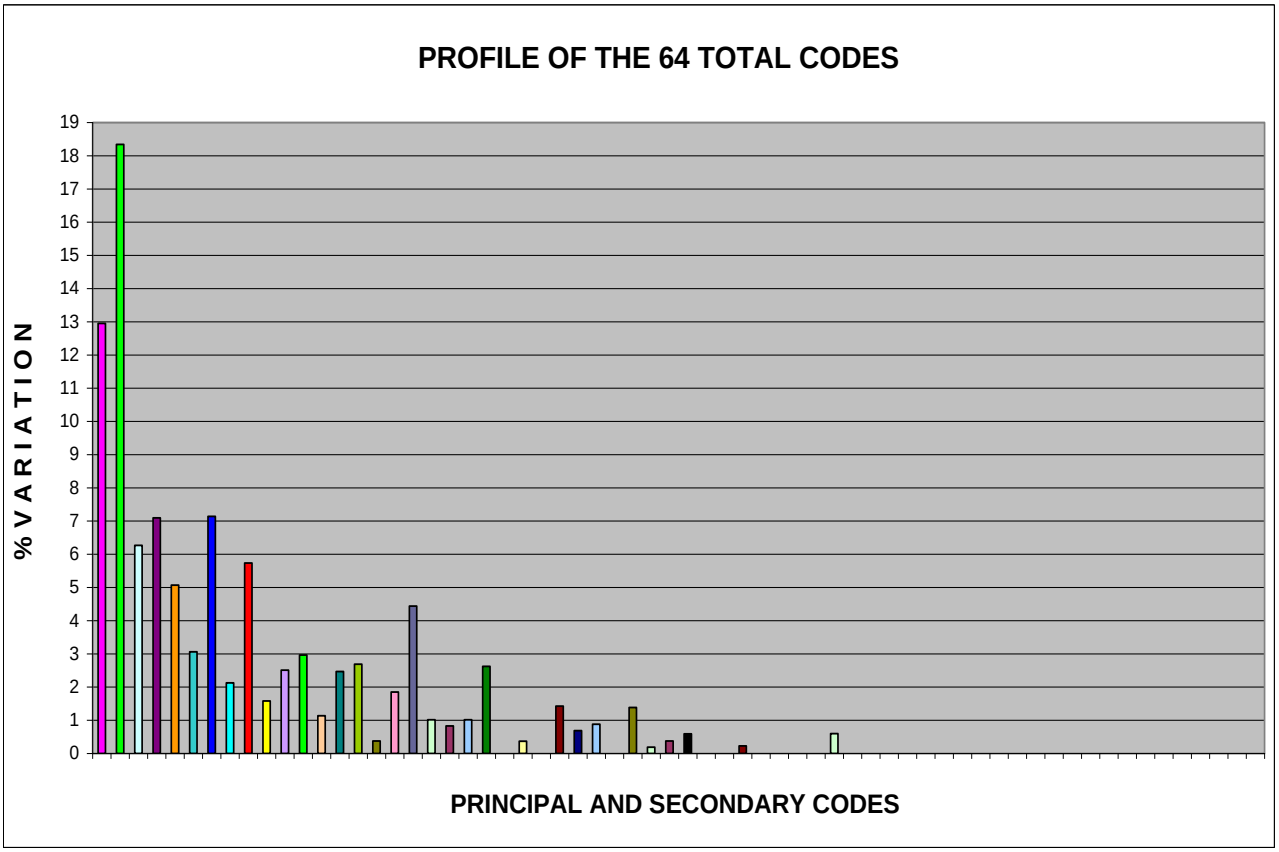
(1) National Center for Biotechnology Information (**NCBI**)[Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; [1988]. Available from: https://www.ncbi.nlm.nih.gov/



Pct. 4 (This graph highlights *precise aspects* of the "*Trend* "of the entire sequence)



Pct. 5 (This graph highlights the "*quality*" of the *single bases* of the entire sequence)



Pct. 6 (This graph constitutes a *very specific synthesis* of the entire sequence)

Up to this point, the evident novelty is the *completely original and innovative way* of graphically representing a DNA or RNA sequence.

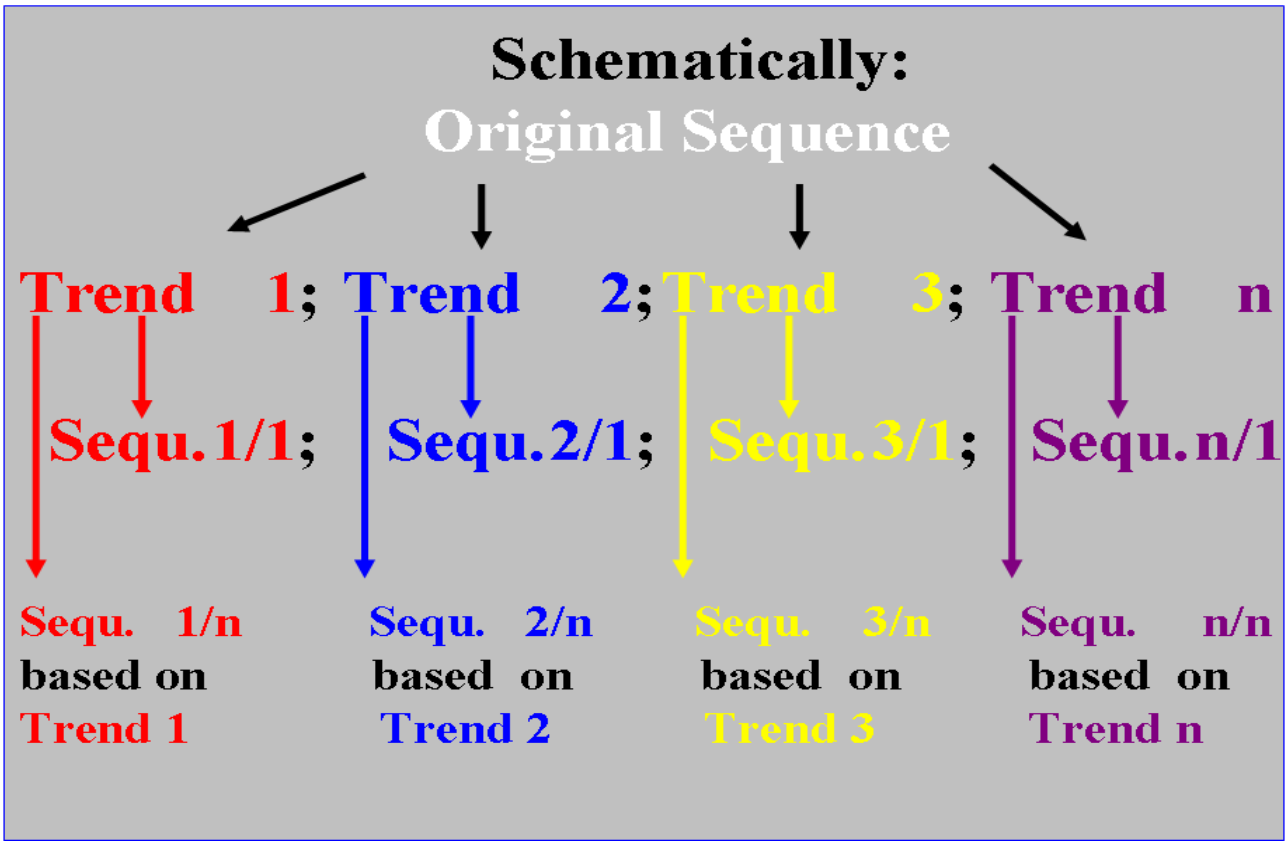
I.3 Second Phase: Modification of the Original Sequence

Subsequently (and at scientists' discretion), the **SECOND PHASE,** that is to say the **MODIFICATION OF THE ORIGINAL SEQUENCE (**and possibly the next ones) begins.

After performing the **graphic analysis** of the **original** DNA or RNA **sequence**, the **T.T.E.S.** software offers also the possibility to modify the **original sequence** and to generate numerous and different **new** DNA or RNA **sequences** that faithfully respect the numerous and different "*non-obvious trends*" of the **original sequence**.
The "**non-obvious trends**" are *trends*, not clearly identifiable, present in all DNA and RNA sequences.

Generating many different **new** DNA or RNA **sequences** is possible for **two reasons**:

1) Firstly, because every DNA or RNA specific sequence (*original sequence*) can be «transformed» into many different *new sequences*, following the many different "non-obvious trends" of the specific *original sequence* (Pct.7);
2) Secondly, because every "non-obvious trend" of the *original sequence* can generate many different new sequences (Pct.7).



Pct. 7

Starting from **each** of the **possible "Non-Obvious Trend"** of an **original sequence**, it is possible to generate an indeterminate number of "new sequences" (as schematized in Pct. 7). Therefore, from **Trend n°1** it is possible to generate sequences n°1/1, 1/2, 1/3, etc. Similarly, from **Trend n°2** it is possible to generate sequences n°2/1, 2/2, 2/3, etc. and so on.
The identification of "**non-obvious trends**" is the fundamental prerequisite for DNA and RNA sequences modifications performed through the **T.T.E.S.** software.
For the **COMPLETE MODIFICATION OF AN ORIGINAL SEQUENCE** and the generation of **ALL NEW SEQUENCES** it is necessary to **identify all possible NON-OBVIOUS TRENDS** (beyond the **Trend n°1** already identified, analyzed and graphically represented in the **first phase**) of the **original sequence**.

It is fundamental to underline that, usually, an **original sequence** differs from **all new sequences generated** by at least **70% of bases** (it is therefore a **very different biological material**).

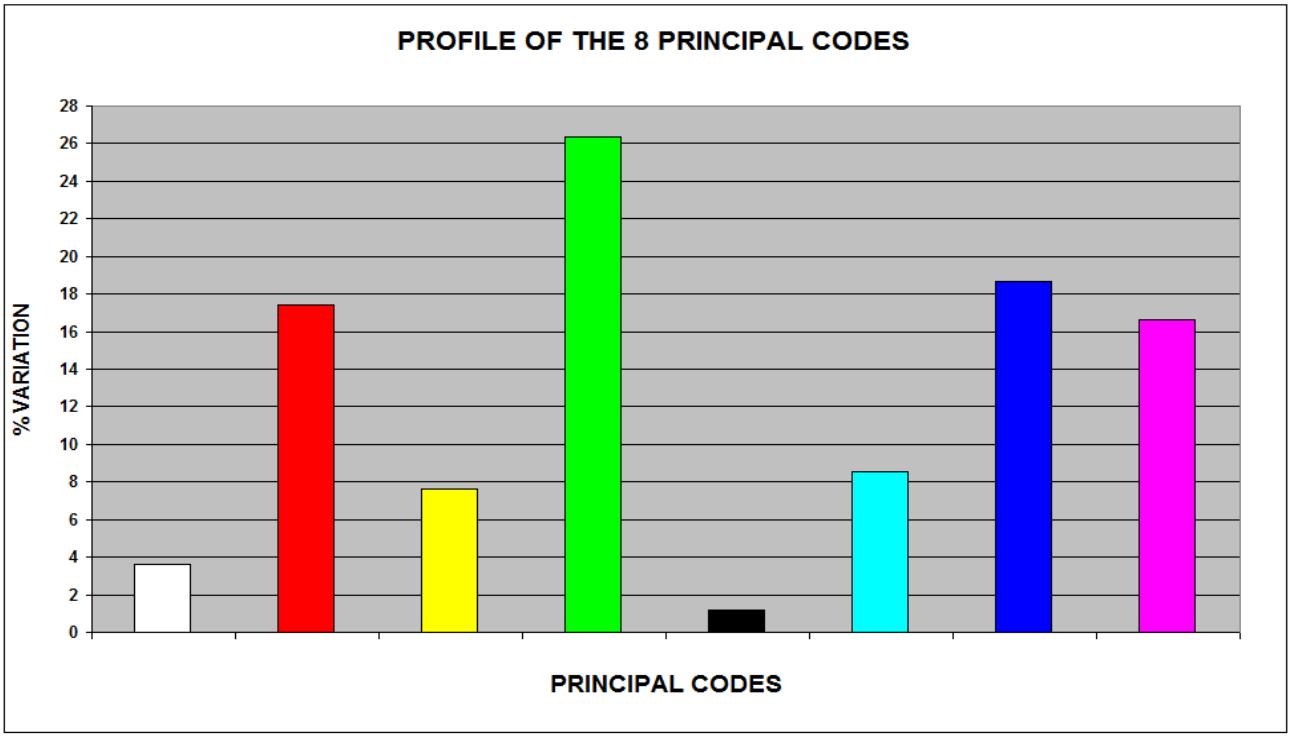
I.4 Third Phase: Comparison of Graphic Representations

The afore-mentioned account is followed by a **THIRD PHASE**: the **COMPARISON OF GRAPHIC REPRESENTATIONS**.
After *analysing and representing graphically* the **original sequence**, *identifying* all its possible **non-obvious trends**, *analysing* and *representing graphically* all the **new sequences** (or only those to which the scientist is interested in), it is possible to *compare the graphic representations* of an **original sequence** with those of the **new sequences**.

The images presented from page 11 to page 15 are the thorough example of a *comparison between the four main graphic representations* of the **original sequence** (**Insulin A Chain**) and those of one of its **new sequences** (**Sequence n°1/1**) generated from one of its **non-obvious trends** (the **Trend n°1**). The following example is taken from the two documents linked right below:
DNA or RNA Sequence Analysis and Modification through T_T_E_S_ (Chapter I° - First Part) and DNA or RNA Sequence Analysis and Modification through T_T_E_S_ (Chapter I° - Second Part).

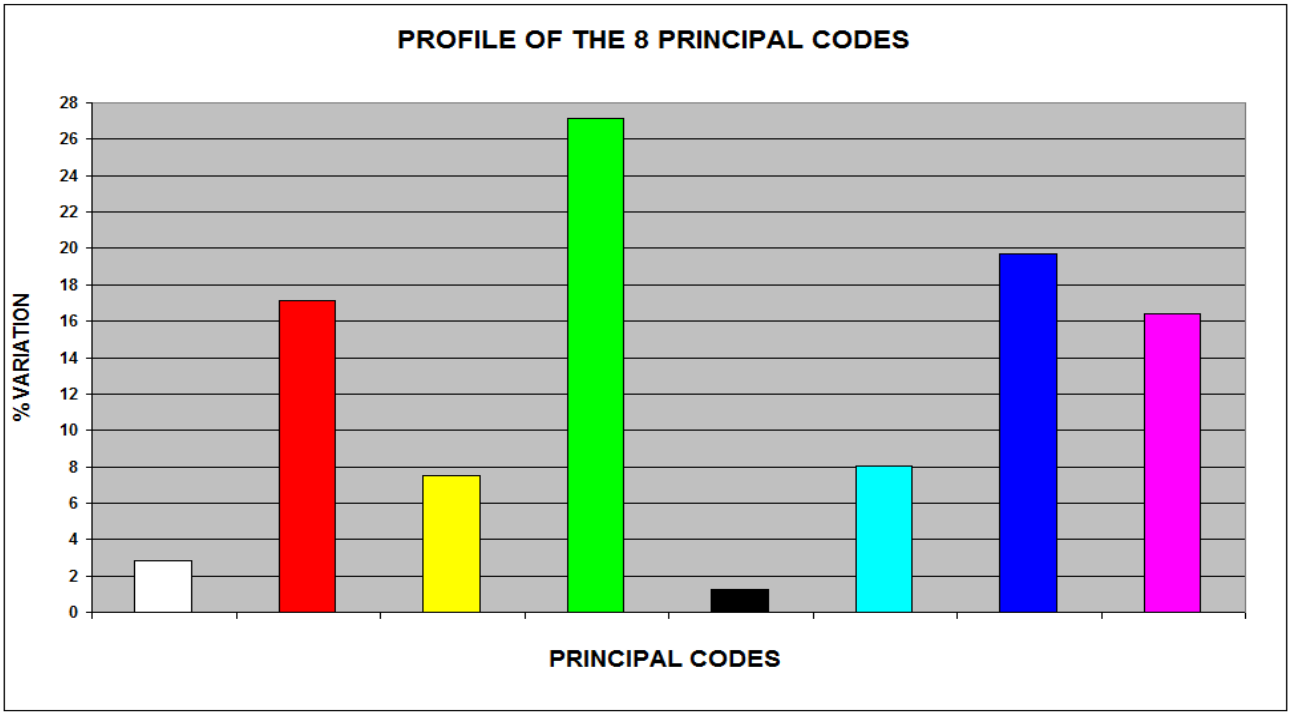
In the Pct. 8 (A e B) two **8 Principal Code Profiles** are compared.

The chart in Pct. 8 (A) refers to the **original sequence** (**Insulin A Chain**).



Pct. 8 (A)

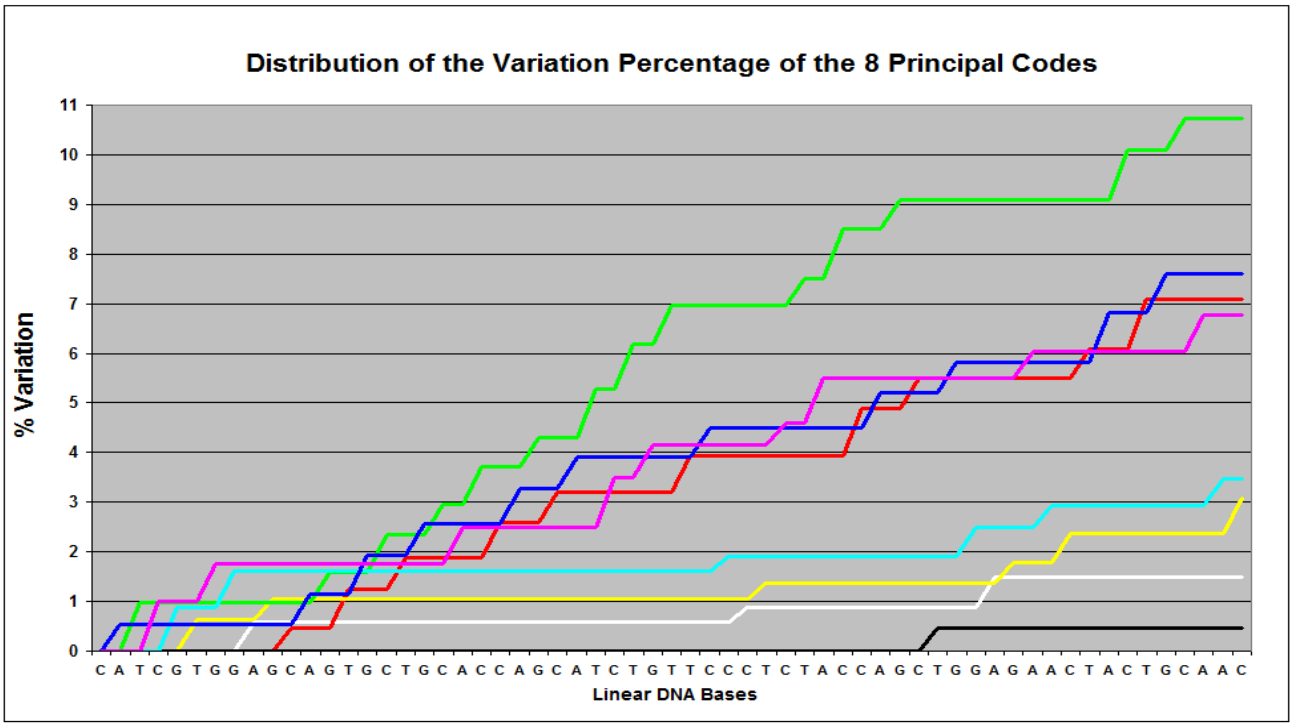
The chart in Pct. 8 (B) refers to the **new sequence** (**Sequence n° 1/1**).



Pct. 8 (B)

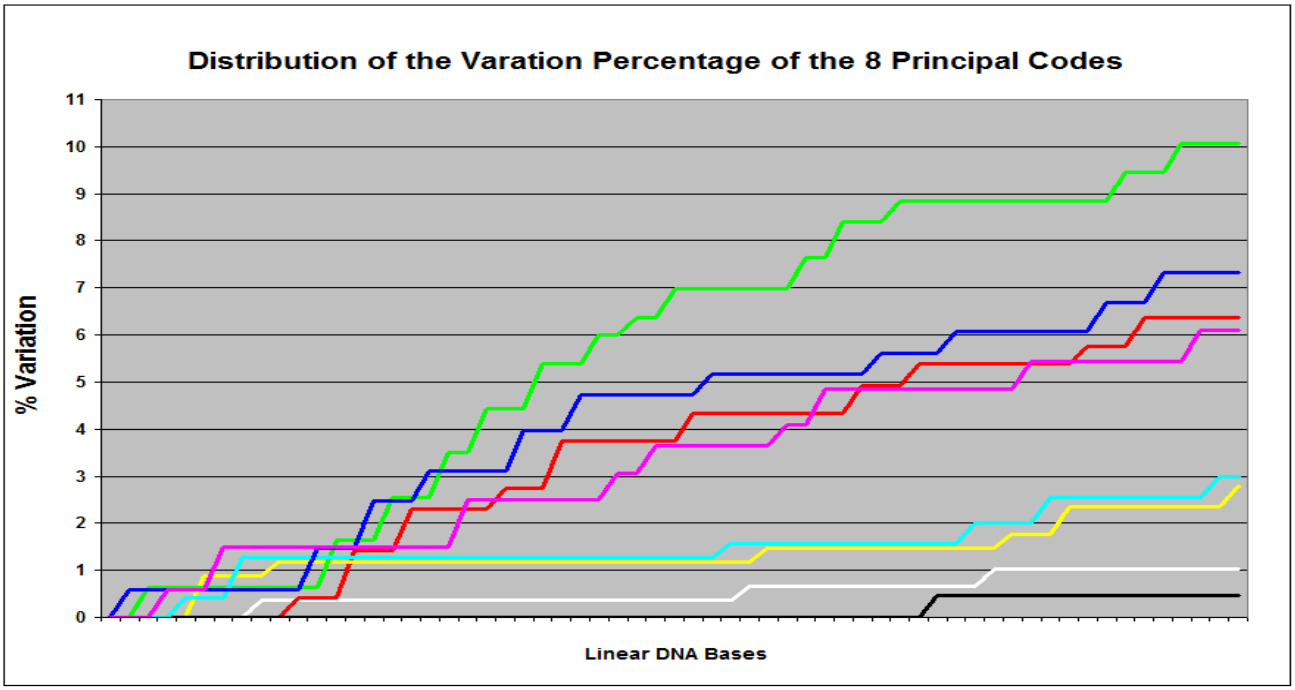
In Pct. 9 (A and B) two charts illustrating the **Distribution of the Variation Percentage of the 8 Main Codes** are compared.

The charts in Pct. 9 (A) refers to the **original sequence** (**Insulin A Chain**).



Pct. 9 (A)

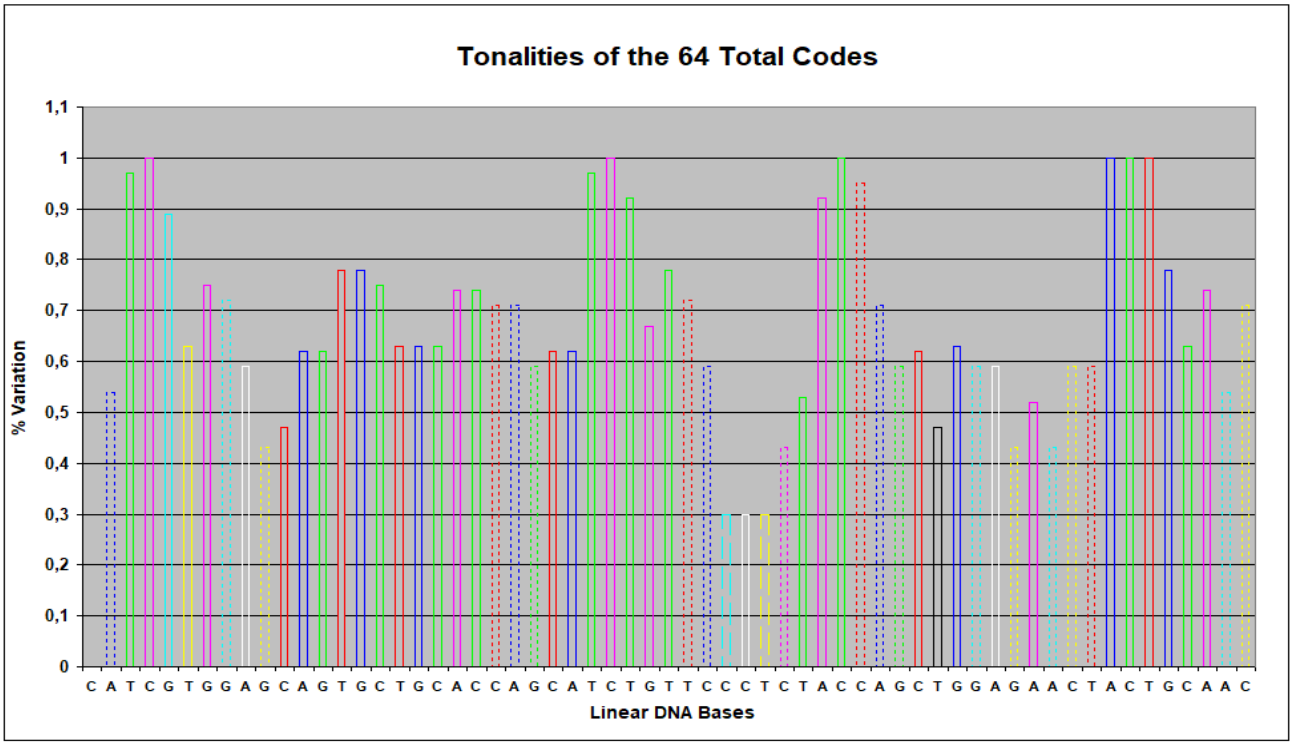
The chart in Pct. 9 (B) refers to the **new sequence** (**Sequence n° 1/1**).



Pct. 9 (B)

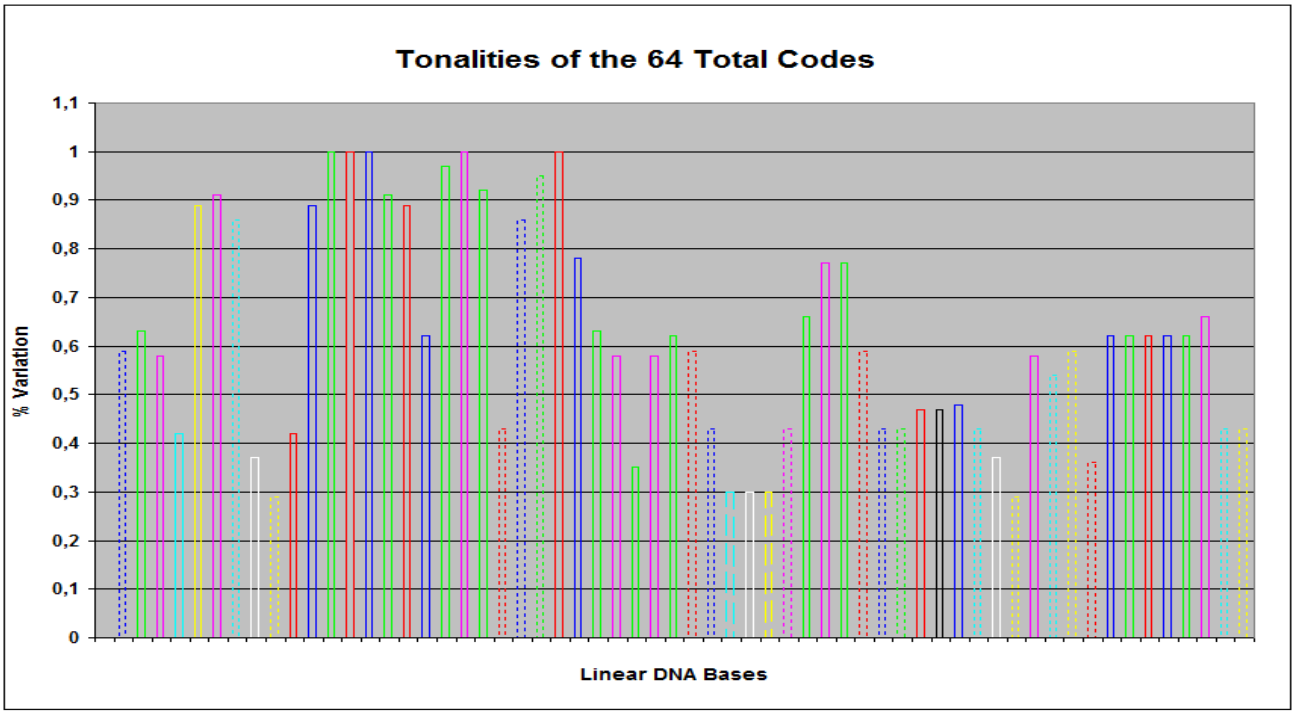
In Pct. 10 (A and B) two charts concerning the single **Tonalities of the 64 Total Codes** are compared.

The chart in Pct. 10 (A) refers to the **original sequence** (**Insulin A Chain**).



Pct. 10 (A)

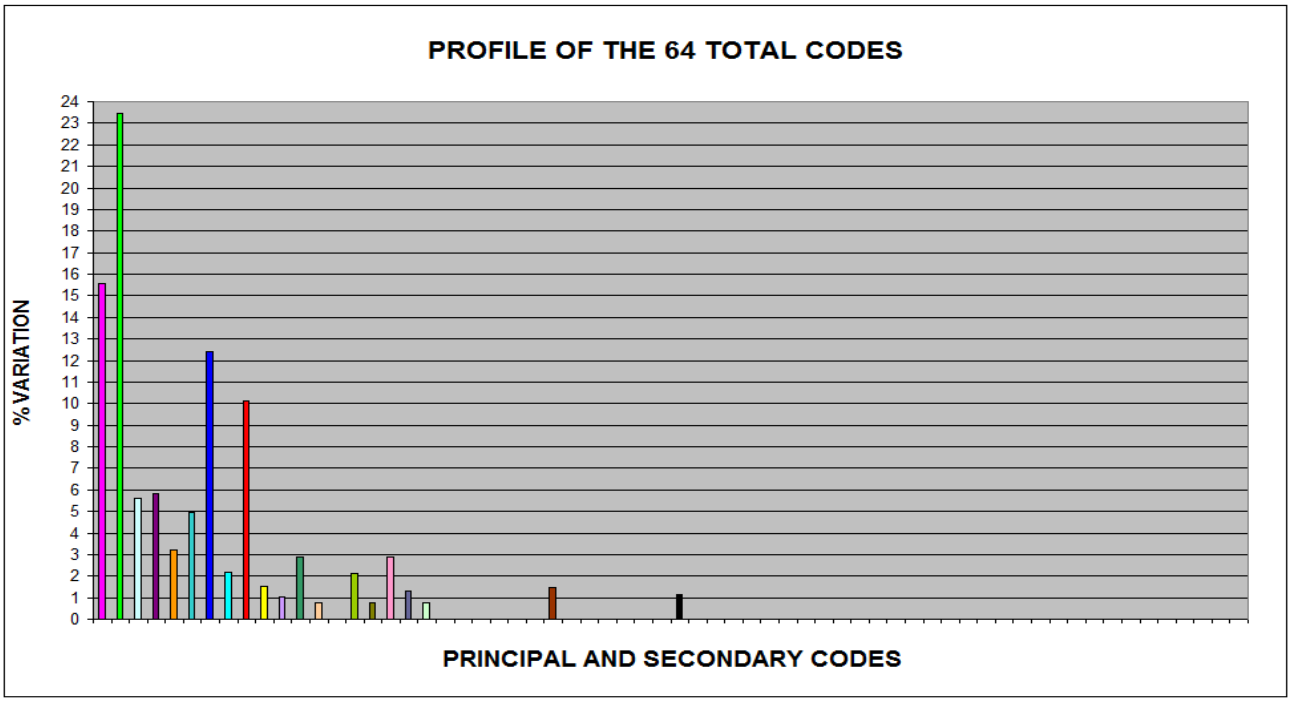
The chart in Pct. 10 (B) refers to the **new sequence** (**Sequence n° 1/1**).



Pct. 10 (B)

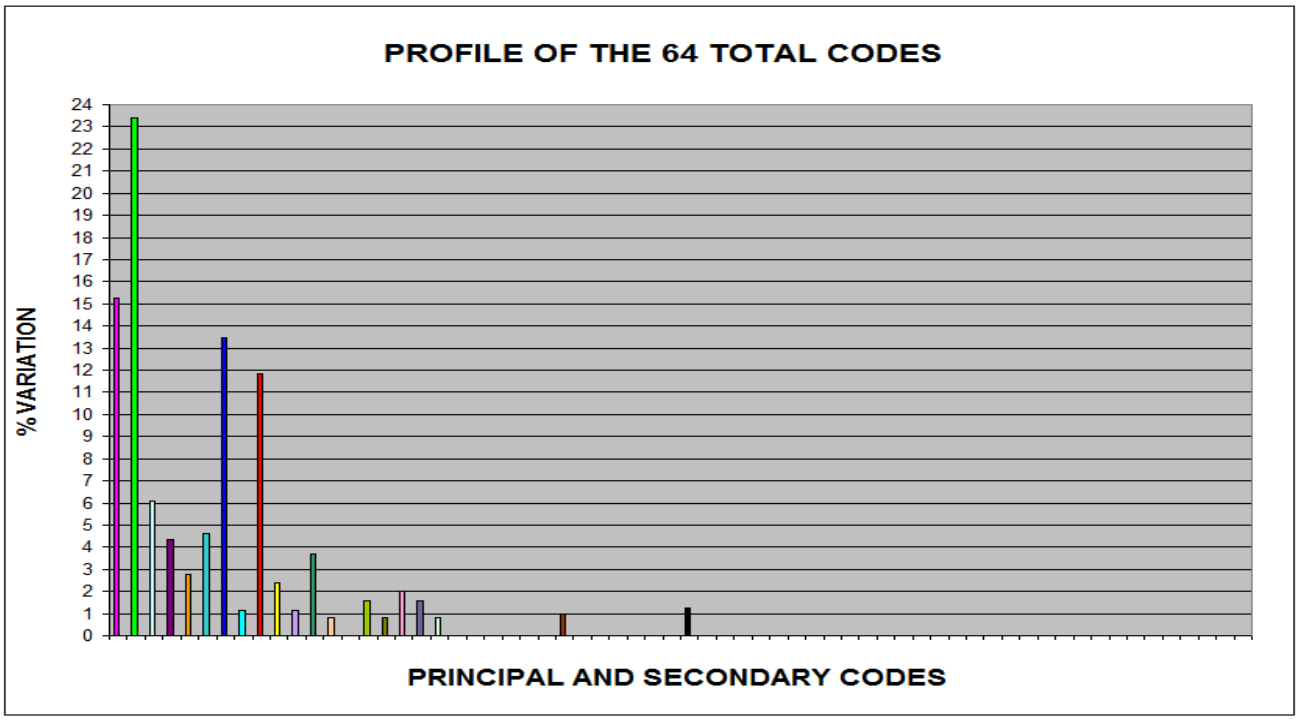
In Pct. 11 (A and B) two charts concerning the **Profile of the 64 Total Codes** are compared.

The chart in Pct. 11 (A) refers to the **original sequence** (**Insulin A Chain**).



Pct. 11 (A)

The chart in Pct. 11 (B) refers to the **new sequence** (**Sequence n° 1/1**).



Pct. 11 (B)

From the *comparison of the four principal graphic representations* of an **original sequence** with those of its **new sequences**, important *similarities* and *differences* can emerge.
The remarkable similarities between the "characteristics" of an **original sequences** and those of the **new sequences** (despite the fact that between the first and second there are bases which differ by at least 70%) [see all the results presented in: DNA or RNA Sequence Analysis and Modification through T_T_E_S_ (Chapter I° - Second Part)] invite us to reflect deeply on the importance and significance, to this date totally neglected, of **non-obvious trends** of DNA or RNA sequences.

I.5 Fourth Phase: Blast Research

The next phase, la **FOURTH PHASE**, is that of ***BLAST RESEARCH*** (*Basic Local Alignment Search Tool* (2)).

After comparing the four principal graphic representations of an **original sequence** with those of its **new sequences** and after analysing their *similarities* and *differences*, we can proceed by carrying out the **BLAST research**, respectively on the **original sequence** considered and on all its **new sequences** [see all results presented in: DNA or RNA Sequence Analysis and Modification through T_T_E_S_ (Chapter I° - Second Part)].

The purpose of this Phase is to identify, for the **original sequence** and for the new **sequences**, all *significant alignments* with the sequences and with the *organisms* stored in the database.

(2) Altschul S. F., Madden T. L., Schaffer A. A., Zhang J., Zhang Z., Miller W. and D. J. Lipman. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res., 1997, 25 (17) :3389-3402. PMID: 9254694. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC146917/

I.6 Fifth Phase: Discovering and Highlighting Common Organisms

Upon conclusion of the BLAST RESEARCH PHASE, the **FIFTH PHASE** begins. This is dedicated to **DISCOVERING** and **HIGHLIGHTING** those "organisms" (and consequently the *Dna or Rna bases*) that are *in common* between the **original sequence** and its **new sequences**.

The following are the suggested **criteria** for performing the complex *highlighting* procedure in a relatively simple manner:

1) those **"organisms found to be in common**" between the results of the BLAST research carried out on an **original sequence** (e.g., **Insulin A Chain**) and the results of the BLAST research carried out on **one of its new sequences** (e.g., **Sequence n° 1/1**), are to be highlighted in **Red**;

2) those "**organisms found to be in common**" between the results of the BLAST research carried out on **an original sequence** (e.g., **Insulin A Chain**), the results of the BLAST research carried out on **one of its new sequences** (e.g., **Sequence n° 1/1**) and the results of the BLAST research carried out on **at least another** among all the *newly generated sequences*, are to be highlighted in **Green**;

3) those "**organisms found to be in common**" and the **"denomination of the sequences**" (1/1, 2/1, ... n/1) of the "**organism found to be in common**" between the results of the BLAST research carried out on **one of the new sequences** (e.g., the **Sequence n°1/1**) and the results of the BLAST research carried out on **at least another** among all *the newly generated sequences* (e.g., **Sequence n°2/1**), are to be highlighted in **Blue**;

4) those "**denomination of the sequences**" (1/1, 2/1, ... 19/1) of the "**organism found to be in common**" between the results of the BLAST research carried out on an **original sequence** (e.g., **Insulin A Chain**) and the results of the BLAST research carried out on **all** *the newly generated sequences*, are to be highlighted in **Yellow**.

The table presented on the next page is an example of a comparison between the alignments of the **Sequence 1/1** and the **Sequence of Insulin A Chain**. Such a comparison has been carried out following the criterion "**Species of Common Organisms**".
The same **comparison** is performed on **all** the *newly generated sequences* [see all comparisons presented in: DNA or RNA Sequence Analysis and Modification through T_T_E_S_ (Chapter I° - Second Part)].

Comparison between the alignments of Sequence 1/1 and Sequence of Insulin A Chain following the criterion "Species of Common Organisms":

Alignments Sequence 1/1	Description	Alignments Sequence Insulin A Chain	Description
1 Select seq CP010359.1	Pseudomonas plecoglossicida strain NyZ12, complete genome	Select seq CP026880.1 **18/1**	Pseudomonas sp. LH1G9 chromosome, complete genome
2 Select seq CP007620.1 **18/1**	Pseudomonas putida strain DLL-E4, complete genome	Select seq CP025263.1 **18/1**	Pseudomonas sp. S09G 359 chromosome
18 Select seq LT629788.1	Pseudomonas moraviensis strain BS3668 genome assembly, chromosome: I	Select seq CP018420.1 **18/1**	Pseudomonas veronii strain R02, complete genome
33 Select seq CP026674.1	Pseudomonas sp. SWI44 chromosome, complete genome	Select seq LT599583.1 **18/1**	Pseudomonas veronii 1YdBTEX2 genome assembly, chromosome: PVE_r1
34 Select seq CP026676.1	Pseudomonas sp. SWI6 chromosome, complete genome		
38 Select seq CP003961.1	Pseudomonas sp. VLB120, complete genome		
4 Select seq XM_027404801.1	PREDICTED: Cricetulus griseus pecanex 2 (Pcnx2), mRNA	Select seq XM_027409202.1	PREDICTED: Cricetulus griseus insulin (Ins), mRNA
5 Select seq XM_003496803.4	PREDICTED: Cricetulus griseus pecanex 2 (Pcnx2), mRNA	Select seq XM_003508080.2	PREDICTED: Cricetulus griseus insulin (Ins), mRNA
6 Select seq XM_026789982.1	PREDICTED: Microtus ochrogaster pecanex 2 (Pcnx2), transcript variant X2, mRNA	Select seq XM_005351571.2	PREDICTED: Microtus ochrogaster insulin (Ins), mRNA
7 Select seq XM_013345975.2	PREDICTED: Microtus ochrogaster pecanex 2 (Pcnx2), transcript variant X1, mRNA	Select seq DQ250572.1	Microtus kikuchii preproinsulin (Ins) gene, complete cds

Comparison between the alignments of Sequence 1/1 and Sequence of Insulin A Chain following the criterion "Species of Common Organisms":

Alignments Sequence 1/1	Description	Alignments Sequence Insulin A Chain	Description
8 Select seq XM_028095345.1	PREDICTED: Eumetopias jubatus TNFRSF1A associated via death domain (TRADD), mRNA	Select seq XM_028118258.1	PREDICTED: Eumetopias jubatus insulin (LOC114220406), mRNA
9 Select seq XM_027618249.1	PREDICTED: Zalophus californianus UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 9 (LOC113935781), transcript variant X5, mRNA	Select seq XM XM_027579931.1	PREDICTED: Zalophus californianus insulin (INS), mRNA
10 Select seq XM_027618248.1	PREDICTED: Zalophus californianus UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 9 (LOC113935781), transcript variant X4, mRNA		
11 Select seq XM_025888785.1	PREDICTED: Callorhinus ursinus TNFRSF1A associated via death domain (TRADD), mRNA	Select seq XM_025879485.1	PREDICTED: Callorhinus ursinus insulin (LOC112829807), mRNA
12 Select seq XM_021703964.1	PREDICTED: Neomonachus schauinslandi TNFRSF1A associated via death domain (TRADD), transcript variant X2, mRNA	Select seq XM XM_021685179.1	PREDICTED: Neomonachus schauinslandi insulin (INS), mRNA
13 Select seq XM_021703956.1	PREDICTED: Neomonachus schauinslandi TNFRSF1A associated via death domain (TRADD), transcript variant X1, mRNA		
14 Select seq XM_013122036.2 **15/1**	PREDICTED: Mesocricetus auratus pecanex homolog 2 (Drosophila) (Pcnx2), transcript variant X2, mRNA	Select seq XM_013112606.2	PREDICTED: Mesocricetus auratus insulin (Ins), mRNA

Comparison between the alignments of Sequence 1/1 and Sequence of Insulin A Chain following the criterion "Species of Common Organisms":

Alignments Sequence 1/1	Description	Alignments Sequence Insulin A Chain	Description
15 Select seq XM_005064691.3 **15/1**	PREDICTED: Mesocricetus auratus pecanex homolog 2 (Drosophila) (Pcnx2), transcript variant X1, mRNA		
16 Select seq XM_021170316.1 **15/1**	PREDICTED: Mus caroli pecanex homolog 2 (Drosophila) (Pcnx2), mRNA	Select seq XM_021152514.1 **6/1 8/1 10/1** **13/1 17/1 18/1**	PREDICTED: Mus caroli insulin-1 (LOC110286053), mRNA
17 Select seq XM_021220388.1 **15/1**	PREDICTED: Mus pahari pecanex homolog 2 (Drosophila) (Pcnx2), mRNA	Select seq DQ250565.1 **6/1 8/1 10/1** **13/1 17/1 18/1**	Mus caroli preproinsulin 1 (Ins1) gene, complete cds
20 Select seq XR_001778443.1	PREDICTED: Mus musculus pecanex homolog 2 (Pcnx2), transcript variant X3, misc_RNA	Select seq XM_021215010.1 **6/1 8/1 10/1** **13/1 17/1 18/1**	PREDICTED: Mus pahari insulin-1 (LOC110333420), mRNA
21 Select seq XM_011248396.2	PREDICTED: Mus musculus pecanex homolog 2 (Pcnx2), transcript variant X2, mRNA	Select seq NM_008386.4 **6/1 8/1 10/1** **13/1 17/1 18/1**	Mus musculus insulin I (Ins1), mRNA
22 Select seq XM_011248395.2	PREDICTED: Mus musculus pecanex homolog 2 (Pcnx2), transcript variant X1, mRNA	Select seq BC145868.1 **6/1 8/1 10/1** **13/1 17/1 18/1**	Mus musculus insulin I, mRNA (cDNA clone MGC:175755 IMAGE:40131171), complete cds
26 Select seq XM_006531060.1	PREDICTED: Mus musculus pecanex homolog 2 (Pcnx2), transcript variant X4, mRNA	Select seq DQ479923.1 **6/1 8/1 10/1** **13/1 17/1 18/1**	Mus musculus strain BTBR T+ tf/J insulin 1 precursor, gene, complete cds
27 Select seq NM_175561.4	Mus musculus pecanex homolog 2 (Pcnx2), mRNA	Select seq AC163452.12 **6/1 8/1 10/1** **13/1 17/1 18/1**	Mus musculus chromosome 19, clone RP23-405C7, complete sequence
29 Select seq BC068235.1 **15/1**	Mus musculus pecanex-like 2 (Drosophila), mRNA (cDNA clone IMAGE:30542978), containing frame-shift errors	Select seq AC136710.8 **6/1 8/1 10/1** **13/1 17/1 18/1**	Mus musculus chromosome 19, clone RP23-35B13, complete sequence

Comparison between the alignments of Sequence 1/1 and Sequence of Insulin A Chain following the criterion "Species of Common Organisms":

Alignments Sequence 1/1	Description	Alignments Sequence Insulin A Chain	Description
30 Select seq AK220342.1	Mus musculus mRNA for mKIAA0435 protein	Select seq AC140320.2 **6/1 8/1 10/1** **13/1 17/1 18/1**	Mus musculus BAC clone RP23-401C13 from chromosome 19, complete sequence
31 Select seq AK087907.1	Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:E330039K12 product:weakly similar to PECANEX 1 [Mus musculus], full insert sequence	Select seq BC098468.1 **6/1 8/1 10/1** **13/1 17/1 18/1**	Mus musculus insulin I, mRNA (cDNA clone MGC:107382 IMAGE:6432765), complete cds
32 Select seq AK030215.1	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4933424I21 product:hypothetical Homeodomain-like structure containing protein, full insert sequence	Select seq AK148541.1 **6/1 8/1 10/1** **13/1 17/1 18/1**	Mus musculus adult pancreas islet cells cDNA, RIKEN full-length enriched library, clone:C820020F18 product:insulin I, full insert sequence
		Select seq AK007345.1 **6/1 8/1 10/1** **13/1 17/1 18/1**	Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810005L03 product:INSULIN 1 PRECURSOR, full insert sequence
		Select seq XM_021168754.1 **6/1 8/1 10/1** **13/1 17/1 18/1**	PREDICTED: Mus caroli insulin-2 (LOC110299132), transcript variant X2, mRNA
		Select seq XM_021168753.1 **6/1 8/1 10/1** **13/1 17/1 18/1**	PREDICTED: Mus caroli insulin-2 (LOC110299132), transcript variant X1, mRNA
		Select seq NM_001185084.2 **6/1 8/1 10/1** **13/1 17/1 18/1**	Mus musculus insulin II (Ins2), transcript variant 3, mRNA
		Select seq NM_001185083.2 **6/1 8/1 10/1** **13/1 17/1 18/1**	Mus musculus insulin II (Ins2), transcript variant 1, mRNA
		Select seq NM_008387.5 **6/1 8/1 10/1** **13/1 17/1 18/1**	Mus musculus insulin II (Ins2), transcript variant 2, mRNA

Comparison between the alignments of Sequence 1/1 and Sequence of Insulin A Chain following the criterion "Species of Common Organisms":

Alignments Sequence 1/1	Description	Alignments Sequence Insulin A Chain	Description
		Select seq JN959239.1 **6/1 8/1 10/1 13/1 17/1 18/1**	Mus musculus targeted KO-first, conditional ready, lacZ-tagged mutant allele Ins2:tm1a(EUCOMM)Wtsi; transgenic
		Select seq JN951270.1 **6/1 8/1 10/1 13/1 17/1 18/1**	Mus musculus targeted non-conditional, lacZ-tagged mutant allele Ins2:tm1e(EUCOMM)Wtsi; transgenic
		Select seq BC145554.1 **6/1 8/1 10/1 13/1 17/1 18/1**	Mus musculus insulin II, mRNA (cDNA clone MGC:179126 IMAGE:9054118), complete cds
		Select seq BC099934.1 **6/1 8/1 10/1 13/1 17/1 18/1**	Mus musculus insulin II, mRNA (cDNA clone MGC:107381 IMAGE:6432976), complete cds
		Select seq BC132650.1 **6/1 8/1 10/1 13/1 17/1 18/1**	Mus musculus insulin II, mRNA (cDNA clone MGC:164281 IMAGE:40130927), complete cds
		Select seq DQ250569.1 **6/1 8/1 10/1 13/1 17/1 18/1**	Mus caroli preproinsulin 2 (Ins2) gene, complete cds
		Select seq AK007612.1 **6/1 8/1 10/1 13/1 17/1 18/1**	Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810027C14 product:INSULIN 2 PRECURSOR, full insert sequence
		Select seq AK007482.1 **6/1 8/1 10/1 13/1 17/1 18/1**	Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810013J24 product:INSULIN 1 PRECURSOR, full insert sequence
		Select seq BC066208.1 **6/1 8/1 10/1 13/1 17/1 18/1**	Mus musculus insulin II, mRNA (cDNA clone IMAGE:6436276)
		Select seq AC012382.14 **6/1 8/1 10/1 13/1 17/1 18/1**	Mus musculus chromosome 7, clone RP23-92L23, complete sequence

Comparison between the alignments of Sequence 1/1 and Sequence of Insulin A Chain following the criterion "Species of Common Organisms":

Alignments Sequence 1/1	Description	Alignments Sequence Insulin A Chain	Description
		Select seq AY899305.1 **6/1 8/1 10/1 13/1 17/1 18/1**	Mus musculus proinsulin mRNA, complete cds, alternatively spliced
		Select seq AC013548.13 **6/1 8/1 10/1 13/1 17/1 18/1**	Mus musculus chromosome 7, clone RP23-209O22, complete sequence
		Select seq AP003182.2 **6/1 8/1 10/1 13/1 17/1 18/1**	Mus musculus genomic DNA, chromosome 7 clone:B189M11, complete sequences
		Select seq GQ915612.1 **6/1 8/1 10/1 13/1 17/1 18/1**	Mus musculus insulin-2 precursor (Ins2) mRNA, partial cds, alternatively spliced
		Select seq XM_021204833.1 **6/1 8/1 10/1 13/1 17/1 18/1**	PREDICTED: Mus pahari insulin-2 (LOC110326410), transcript variant X2, mRNA
		Select seq XM_021204825.1 **6/1 8/1 10/1 13/1 17/1 18/1**	PREDICTED: Mus pahari insulin-2 (LOC110326410), transcript variant X1, mRNA
23 Select seq XM_004393402.2	PREDICTED: Odobenus rosmarus divergens TNFRSF1A-associated via death domain (TRADD), transcript variant X2, mRNA	Select seq XM_004403802.1	PREDICTED: Odobenus rosmarus divergens insulin (INS), mRNA
24 Select seq XM_012566139.1	PREDICTED: Odobenus rosmarus divergens TNFRSF1A-associated via death domain (TRADD), transcript variant X1, mRNA		
25 Select seq XM_006741502.1	PREDICTED: Leptonychotes weddellii TNFRSF1A-associated via death domain (TRADD), mRNA	Select seq XM_006750095.1	PREDICTED: Leptonychotes weddellii insulin (INS), mRNA

Comparison between the alignments of Sequence 1/1 and Sequence of Insulin A Chain following the criterion "Species of Common Organisms":

Alignments Sequence 1/1	Description	Alignments Sequence Insulin A Chain	Description
35 Select seq XM_022492727.1	PREDICTED: Enhydra lutris kenyoni trichohyalin-like (LOC111140481), partial mRNA	Select seq XM_022507720.1	PREDICTED: Enhydra lutris kenyoni insulin (LOC111150279), mRNA
37 Select seq XM_006880105.1	PREDICTED: Elephantulus edwardii putative scavenger receptor cysteine-rich domain-containing protein LOC619207-like (LOC102868011), mRNA	Select seq XM_006893212.1	PREDICTED: Elephantulus edwardii insulin (INS), mRNA

I.7 Sixth Phase: Specification of the Product of Significant Alignment

After **HIGHLIGHTING** the "organisms" that are *in common* between the **original sequence** and the **new sequences**, the **SIXTH PHASE SPECIFICIES**, for all "organisms" that are *in common*, the "**Product**" of the *significant alignment* and which *DNA* (*or Rna*) *modified bases* they have in common.

An example of this last procedure is presented in Chapter I – *First Part* [DNA or RNA Sequence Analysis and Modification through T T E S (Chapter I° - First Part)]. At page 33 of the above mentioned Chapter, the *partially complete* results (shown in the next page) of some "**Products**" of the *significant alignments* of the **Sequence n°1/1** are presented. These results refer to some species of **Pseudomonas** bacteria. These results are *partially complete* because they lack some fundamental indications to identify precisely which Dna bases, obtained through modification of the **Insulin A Chain**, are part of the DNA of the "**Product**" of the *organism* identified by the significant alignment of the **Sequence n°1/1**.

The information about the significant alignments produced by *BLAST research* concerning the **Sequence n°1/1** (and the other 19 sequences analysed in the *Second Part* of Chapter I) and all the other important information acquired from *GenBank* will be published in the **Appendix,** upon completion of the *General Conclusions* of the entire Book which is being written.

Significant Alignments of Pseudomonas

>CP010359.1
Pseudomonas plecoglossicida strain NyZ12, complete genome
Length=6233254
Product: cystathionine gamma-synthase
https://www.ncbi.nlm.nih.gov/nucleotide/752308899?report=gbwithparts&from=5560737&to=5562680&RID=1ZXSZEJC014

>CP007620.1
Pseudomonas putida strain DLL-E4, complete genome
Length=6484062
Product: cystathionine gamma-synthase
https://www.ncbi.nlm.nih.gov/nucleotide/635291785?report=gbwithparts&from=176815&to=178758&RID=1ZXSZEJC014

>LT629788.1
Pseudomonas moraviensis strain BS3668 genome assembly, chromosome:I
Length=6092541
Product: high-affinity iron transporter
https://www.ncbi.nlm.nih.gov/nucleotide/1086004611?report=gbwithparts&from=4649680&to=4651578&RID=1ZXSZEJC014

>CP003961.1
Pseudomonas sp. VLB120, complete genome
Length=5644569
Product: cytochrome c class I
https://www.ncbi.nlm.nih.gov/nucleotide/556072477?report=gbwithparts&from=5441006&to=5442901&RID=1ZXSZEJC014

The **two** significant alignments of *Pseudomonas* presented right below have been identified later by a new **Blast Research** carried out on the **Sequence n°1/1** (for further info, please see the data reported in the table at page 18).

>CP026674.1
Pseudomonas sp. SWI44 chromosome, complete genome
Length: 5919083
Product: cystathionine gamma-synthase
https://www.ncbi.nlm.nih.gov/nucleotide/CP026674.1?report=gbwithparts&from=2646332&to=2648260&RID=1KYS493H014

>CP026676.1
Pseudomonas sp. SWI6 chromosome, complete genome
Length: 5652054
Product: cystathionine gamma-synthase
https://www.ncbi.nlm.nih.gov/nucleotide/CP026676.1?report=gbwithparts&from=382732&to=384660&RID=1KYS493H014

Please note:

If of interests, the researcher may find useful **TO SPECIFY** the "**Product**" of the *significant alignment* of all "*organisms*" that are **NOT** *in common* between an **original sequence** and its **new sequences** and which *modified DNA* (*or RNA*) *bases* they have in common.

Thus, for example, the researcher might find interesting to identify the "products" of the **three** *organisms* presented in the diagram in this page [see also pages 37 - 40 in: DNA or RNA Sequence Analysis and Modification through T_T_E_S_ (Chapter I° - Second Part)] and which *modified DNA (or RNA) bases of* the **Insulin A Chain** *are part* of the DNA (or RNA) of these *organisms*, identified by the significant alignment of the **Sequence n°1/1** (**new sequence**) and which are **NOT** *in common* with the **Insulin A Chain** (**original sequence**).

Sequences producing significant alignments:

Selected seq	Description	Max score	Total score	Query cover	E value	Ident	Accession
3 **15/1** XM_017200197.1	PREDICTED: Drosophila ficusphila DNA topoisomerase 2-binding protein 1 (LOC108097709), mRNA	41.0	41.0	55%	2.7	89%	XM_017200197.1
19 XM_018455918.1	PREDICTED: Trachymyrmex zeteki uncharacterized LOC108727710 (LOC108727710), mRNA	40.1	40.1	38%	9.5	96%	XM_018455918.1
28 XM_002620491.1	Ajellomyces dermatitidis SLH14081 peroxisomal ABC transporter, mRNA	40.1	40.1	38%	9.5	96%	XM_002620491.1

Overall, the results of this phase are very important (along with the results of the SUCCESSIVE PHASES). It is indeed on them that will depend the eventual choice of intervening concretely in the modification of the **biological material** of the analysed sequences for the realization of the various and possible practical purposes.

I.8 Seventh Phase: Bibliographic Research focused on a single Organism

In the **SEVENTH PHASE** a **bibliographic research** focused on a single organism is carried out in order to confirm the existence of **important relationships** between the characteristics (e.g., the "products" highlighted in the previous phase) of one of the *organisms* [e.g., *Pseudomonas*. For further information, see pages 37 - 46 of DNA or RNA Sequence Analysis and Modification through T_T_E_S_ (Chapter I° - First Part)], that has been identified by the Blast research performed on one of its **new sequences** (e.g., **Sequence n °1/1**), and some functional characteristics of the **original sequence** (e.g., **Insulin A Chain** or, simply, **Insulin**).

Regarding the case of **Pseudomonas** bacteria, it is possible to find overall many and important relationships between different species of **Pseudomonas** bacteria (identified by the BLAST research concerning the significant alignments of the **Sequence n°1/1** and some of their DNA bases related to the following "products": *Cystathionine Gamma-Synthase*, *Cytochrome c class I* and *High-Affinity Iron Transporters*) and **insulin**, **diabetes mellitus**, **melioidosis**, **obesity**, **cystic fibrosis**, different types of **infections** (especially **pulmonary**), **malignant external otitis**, **endocarditis**, **immune system**, **apoptosis**, **cellular respiration** and **iron levels** and **its transport**.

I.9 Eight Phase: Bibliographic Research focused on two or more Organisms

The researcher, at their discretion, and in order to reach specific goals, can carry out an **EIGHTH PHASE**, that is to say, a **bibliographic research** focused on two or more organisms in order to confirm the existence of **important relationships** between the characteristics (e.g., the "products" highlighted in the previous phase) of **two** or **more** *organisms* [e.g., *Pseudomonas* and *Heligmosomoides polygyrus*. For further information, see DNA or RNA Sequence Analysis and Modification through T_T_E_S_ (Chapter I° - First Part)], identified by the Blast research performed on a **new sequence** (e.g., **Sequence n°1/1**), and some functional characteristics of the **original sequence** (e.g., **Insulin A Chain** or, simply, **Insulin**).

As concerns the nematode *Heligmosomoides polygyrus* (which significant alignments are reported below), the bibliographic research has pointed out many and important relationships between this hookworm (identified by BLAST research concerning the significant alignments of the **Sequence n°1/1** and some of its DNA bases) and **insulin**, **immune system**, **apoptosis, diabetes type 1, obesity** and **insulite.**
It was not possible to investigate into the relationships between *Pseudomonas* and *Heligmosomoides polygyrus*, because only one article was found on the web. Furthermore, it was dated and not available online.

Significant Alignments of Heligmosomoides polygyrus

>LL188962.1
Heligmosomoides polygyrus genome assembly H_bakeri_Edinburgh, scaffold HPBE_scaffold0000593
Length=94530
https://www.ncbi.nlm.nih.gov/nucleotide/688429340?report=genbank&log$=nuclalign&blast_rank=2&RID=27MWTXV3014

>LL194531.1
Heligmosomoides polygyrus genome assembly H_bakeri_Edinburgh, scaffold HPBE_contig0000102
Length=27221
https://www.ncbi.nlm.nih.gov/nucleotide/688443549?report=genbank&log$=nuclalign&blast_rank=23&RID=27MWTXV3014

I.10 Ninth Phase: Concluding Considerations and Pragmatic Choices

In this **LAST PHASE**, it is possible to draw the **Conclusions** of the work performed in the previous phases and to make the necessary **Pragmatic Choices**, such as, to use the *acquired knowledge and/or the new* (partial or integral) *DNA or RNA sequences* for scientific, industrial, food, etc. research purposes, or to intervene concretely in the modification of the **biological material** (DNA or RNA) of one or more *sequences analysed* for the realization of the different possible purposes.

To give understandable examples of **Conclusions** and proposals for **Pragmatic Choices**, we use the results of the studies reported in this introduction.

From the obtained results, it can be **concluded** that **Insulin** (and, therefore, also the **Insulin A Chain**, called the original sequence) is *very* likely to *implicated* be in various ways *with certain characteristics* of both the *Pseudomonas* bacteria and the nematode *Heligmosomoides polygyrus* (*organisms* both consisting of DNA bases of **Sequence n°1/1**, called the new sequence).

In particular, for *Pseudomonas* bacteria, it is necessary to highlight the importance of the bibliographic research on **cellular respiration**, **ROS** (*Reactive Oxygen Species*) and **iron levels** and **its transport**, that are (individually and non) an *essential convergence point* among different aspects of the phenomena and pathologies in which this species of bacteria is often involved [see page 48 in: DNA or RNA Sequence Analysis and Modification through T_T_E_S_ (Chapter I° - First Part)].
The **pragmatic choice** to expertly interfere in these **biological processes** in which *Pseudomonas* is extremely involved, directly or by modifying the **biological material** of one or more *analysed sequences*, is certainly a goal to be pursued.
It must be added that, in Chapter I First Part, the parameters set for the BLAST research carried out on the **Insulin A Chain** sequence have considered only the first **100** (Hitlist size 100) significant alignments. Differently, in the new BLAST research carried out in in Chapter I Second Part, significant alignments have been extended to **1000** (Hitlist size 1000). From new Blast research, it has emerged that **849 sequences** highlight *significant alignments* with the **Insulin A Chain**. Among these significant alignments, **four** of them, refer to the *Pseudomonas* bacteria.
At the time of the completion of the Chapter I First Part we were not aware of such information and we speculated on the possible relationships between *Pseudomonas* bacteria, the features of the **Sequence n° 1/1** (from which were identified the significant alignments with *Pseudomonas*) and **Insulin**.
Today we know that, besides the **Sequence n°1/1**, the **Sequence n°18/1**, a **new sequence** generated from one of the **non-obvious trends** (the **Trend n°18**) of the **Insulin A Chain**, also shows significant alignments with *Pseudomonas* [see page 272 in: DNA or RNA Sequence Analysis and Modification through T_T_E_S_ (Chapter I° - Second Part)].
Furthermore, **one** of the **two** significant alignments of the **Sequence n°18/1** (Select seq CP025263.1 Pseudomonas sp. S09G 359 chromosome) with *Pseudomonas* is **identical** to the same sequence (although the *bases* are largely different and related to a different "product") of **one** out of the **four** significant alignments of the **Insulin Chain A** with *Pseudomonas* (see table shown at page 18 in this introduction).

Of these **four** significant alignments of the **Insulin A Chain** with *Pseudomonas*, **two** of them (one of which is that in common with the **Sequence n°18/1**) are referred to *bioprojects* which have the goal to identify gene clusters among *Pseudomonads* whose products inhibit human pathogens [not only for the treatment of patients with **Cystic Fibrosis** but also for individuals infected with **MDR** (Select seq CP025263.1; BioProject: PRJNA419203)] and biosynthetic gene clusters within environmental bacteria whose products have been shown to inhibit the growth of these multi-drug resistant pathogens derived from **Cystic Fibrosis** (Select seq: CP026880.1; BioProject: PRJNA433821).
It is also worth noting that **two** out of the **six** significant alignments of the **Sequence n°1/1** with *Pseudomonas* (see table shown on page 18 in this introduction) are referred to *bioprojects* that have as objective that of using *Pseudomonads* to facilitate **antibiotic discovery** (Select seq: CP026674.1 - BioProject: PRJNA433544 ; Select seq: CP026676.1 - BioProject: PRJNA433544).
For detailed explanations of the above results, see pages 302 - 304 in: DNA or RNA Sequence Analysis and Modification through T_T_E_S_ (Chapter I° - Second Part).
Overall, the new results obtained from the significant alignments related to *Pseudomonas* would seem to reinforce the hypothesis elaborated in Chapter I° First Part (that we were capable of deriving also thanks to an articulated bibliographic research): that is to say, the hypothesis according to which **Insulin** (and, therefore, also the **Insulin Chain A**) is *involved in very different ways with some characteristics* of *Pseudomonas* bacteria (such data was at the time discovered through Blast research carried out on the **Sequence n°1/1** and it has later been corroborated by the Blast research carried out on the **Sequence n°18/1** and the **Sequence of Insulin Chain A**).

Regarding the *Heligmosomoides polygyrus*, for reasons that are not known, research related to its significant alignments were removed from the National Center for Biotechnology Information (NCBI) archive [see pages 40 and 191 in: DNA or RNA Sequence Analysis and Modification through T_T_E_S_ (Chapter I° - Second Part)]. We can point out and **conclude** that, beyond the **Sequence n°1/1**, the **Sequence n°12/1**, a **new sequence** generated from one of the **non-obvious trends** (**Trend n°12**) of the **Insulin A Chain**, underscores also significant alignments with *Heligmosomoides polygyrus* (see page 191 of the document mentioned above). This last data reinforces the hypothesis of a relationship between **Insulin** and *Heligmosomoides polygyrus*.

As regards the relationship between *Pseudomonas* and *Heligmosomoides polygyrus*, a **pragmatic choice** should consider future specific studies, especially because some features of *Pseudomonas* appear to be opposite to those of *Heligmosomoides polygyrus;* in facts, in case of infection of *Heligmosomoides polygyrus*, the seriousness of ***Diabetes type 1***, ***obesity***, ***insulite***, ***apoptosis*** seem to decrease (while in case of infection of *Pseudomonas* it increases). For more information, see pages 47 - 49 in: DNA or RNA Sequence Analysis and Modification through T_T_E_S_ (Chapter I° - First Part).

In **conclusion**, the analysis (through **T.T.E.S.**) of the *Original Sequence* (the **Insulin A Chain**) based on one of its "Non-Obvious Trends" (the ***Trend n°1***), the creation of a *New DNA Sequence* (the **Sequence n°1/1**) from ***Trend n°1*** of the *Original Sequence* and the *congruence* with the data obtained from the *bibliographic study* largely confirms the hypotheses tested and open up exciting new research perspectives.
More generally, all studies published so far, downloadable for free from the "applications" section of the http://www.ttesystems.eu/index_en.php website, confirm the hypothesis that **the New Sequences** (generated faithfully respecting each of the possible specific "Non-Obvious Trends" of the **Original Sequence**) **have strong relations with the characteristics of the Original Sequence**.

I would like to end this introduction by anticipating the future publication (within the next four months), on the website **http://www.ttesystems.eu/index_en.php** and on my **LinkedIn Profile**, of **Chapter III** of the book.
In the following chapter I will discuss the **RELATIONSHIPS** between **INSULIN A CHAIN**, **PCNX1** and **PCNX2** (*Pecanex-like protein 1 and 2*), **TRADD** (*Tumor Necrosis Factor Receptor Type 1 Associated DEATH Domain Protein*), **APOPTOSIS** and **TUMORS** (a part will be also reserved also for the implications of the ***Pseudomonas*** bacterium).
These **RELATIONSHIPS** have been identified by a careful comparative analysis of the results of ***all significant alignments*** of the now-known **Sequence n°1/1**.

END OF INTRODUCTION

www.ttesystems.eu

Correspondence: ***nunzio.bonaventura@libero.it***

13 January 2020