

Analisi e Modificazioni di Sequenze di DNA o di RNA con la

TEORIA TRICROMATICA DELL'EQUILIBRIO DEI SISTEMI



Capitolo II°

***Confronto delle Analisi Complete di 4 Sequenze
dell'Organismo Macaca Nemestrina ottenute
dagli Allineamenti Significativi
della Catena A dell'Insulina***

**Autore: Nunzio Bonaventura
Tecnico informatico: Vincenzo Viggiano**

12 Agosto 2019

INDICE DEL CAPITOLO SECONDO

Introduzione	pag. 3
 CAPITOLO II°: Confronto delle Analisi Complete di 4 Sequenze dell'Organismo Macaca Nemestrina ottenute dagli allineamenti significativi della Catena A dell'Insulina	
 Analisi Completa della Sequenza XM_011721319.1	
PREDICTED: Macaca nemestrina insulin (INS), transcript variant X4, mRNA	
2.1 Caratteristiche della Sequenza XM_011721319.1	
PREDICTED: Macaca nemestrina insulin (INS), transcript variant X4, mRNA	pag. 7
2.2 Risultati grafici dell'analisi della Sequenza XM_011721319.1	
PREDICTED: Macaca nemestrina insulin (INS), transcript variant X4, mRNA	pag. 9
 Analisi Completa della Sequenza XM_011721318.1	
PREDICTED: Macaca nemestrina insulin (INS), transcript variant X3, mRNA	
2.3 Caratteristiche della Sequenza XM_011721318.1	
PREDICTED: Macaca nemestrina insulin (INS), transcript variant X3, mRNA	pag. 14
2.4 Risultati grafici dell'analisi della Sequenza XM_011721318.1	
PREDICTED: Macaca nemestrina insulin (INS), transcript variant X3, mRNA	pag. 16
 Analisi Completa della Sequenza XM_011721317.1	
PREDICTED: Macaca nemestrina insulin (INS), transcript variant X2, mRNA	
2.5 Caratteristiche della Sequenza XM_011721317.1	
PREDICTED: Macaca nemestrina insulin (INS), transcript variant X2, mRNA	pag. 21
2.6 Risultati grafici dell'analisi della Sequenza XM_011721317.1	
PREDICTED: Macaca nemestrina insulin (INS), transcript variant X2, mRNA	pag. 23
 Analisi Completa della Sequenza XM_011721316.1	
PREDICTED: Macaca nemestrina insulin (INS), transcript variant X1, mRNA	
2.7 Caratteristiche della Sequenza XM_011721316.1	
PREDICTED: Macaca nemestrina insulin (INS), transcript variant X1, mRNA	pag. 28
2.8 Risultati grafici dell'analisi della Sequenza XM_011721316.1	
PREDICTED: Macaca nemestrina insulin (INS), transcript variant X1, mRNA	pag. 30
 Confronto e Commento dei Risultati delle Analisi Complete delle Sequenze XM_011721319.1, XM_011721316.1, XM_011721317.1 e XM_011721318.1	
2.9 Confronto delle basi delle Sequenze XM_011721319.1, XM_011721316.1, XM_011721317.1 e XM_011721318.1.....	pag. 35
2.10 Confronto e Commento dei risultati grafici delle analisi complete delle Sequenze XM_011721319.1, XM_011721316.1, XM_011721317.1 e XM_011721318.1.....	pag. 37
 Discussione Generale dei Risultati e Conclusioni	
2.12 Discussione generale dei risultati e conclusioni	pag. 49

INTRODUZIONE

In questo Capitolo sono discussi i risultati del **Confronto delle Analisi Complete di 4 Sequenze dell'organismo Macaca Nemestrina**.

Le **Analisi Complete delle 4 Sequenze** sono state effettuate attraverso l'applicazione della **TEORIA TRICROMATICA DELL'EQUILIBRIO DEI SISTEMI** o **T.T.E.S.** (per approfondimenti si veda il sito web www.ttesystems.eu).

Prima di iniziare la lettura di questo Capitolo è indispensabile almeno la lettura propedeutica dell'Introduzione e dei Paragrafi 1.1 e 1.2 del Capitolo I° - Parte Prima [([Analisi e modificazioni di Sequenze di DNA o di RNA con la T T E S \(Capitolo I° - Parte Prima\)](#))] e dell'Introduzione e dei Paragrafi 1.1, 1.2, 1.3, 1.41, 1.42 e 1.43 del Capitolo I°- Parte Seconda [([Analisi e modificazioni di Sequenze di DNA o di RNA con la T T E S \(Capitolo I° - Parte Seconda\)](#))]].

Le **Analisi Complete** delle **4 Sequenze** sono state ottenute dagli *allineamenti significativi* della **Catena A dell'Insulina** (si veda il Paragrafo 1.1 del Capitolo I°- Parte Seconda):

Sequences producing significant alignments:							
Selected seq	Description	Max score	Total score	Query cover	E value	Ident	Accession
XM_011721319.1 2/1 3/1 8/1 17/1	PREDICTED: Macaca nemestrina insulin (INS), transcript variant X4, mRNA	110	110	100%	6e-21	98%	XM_011721319.1
XM_011721318.1 2/1 3/1 8/1 17/1	PREDICTED: Macaca nemestrina insulin (INS), transcript variant X3, mRNA	110	110	100%	6e-21	98%	XM_011721318.1
XM_011721317.1 2/1 3/1 8/1 17/1	PREDICTED: Macaca nemestrina insulin (INS), transcript variant X2, mRNA	110	110	100%	6e-21	98%	XM_011721317.1
XM_011721316.1 2/1 3/1 8/1 17/1	PREDICTED: Macaca nemestrina insulin (INS), transcript variant X1, mRNA	110	110	100%	6e-21	98%	XM_011721316.1

ATTENZIONE: La ricerca BLAST relativa a questi *allineamenti significativi* è stata effettuata in data **9/04/2019**. Come ci è stato comunicato dagli stessi responsabili del BLAST DATABASES, eventuali successive variazioni degli *allineamenti significativi* devono essere attribuite ai regolari aggiornamenti e/o alle modifiche dovute alle parti mobili di questo complesso DATABASES).

Sono state scelte queste **4 sequenze di mRNA** in quanto i loro **quattro** prodotti sono “**insulin isoform**”:

- 1) [XM_011721319.1](#) Product = "insulin isoform **X2**";
- 2) [XM_011721318.1](#) Product = "insulin isoform X1";
- 3) [XM_011721317.1](#) Product = "insulin isoform X1";
- 4) [XM_011721316.1](#) Product = "insulin isoform X1".

Queste **4 sequenze di mRNA** condividono tra di loro molte basi azotate.

In questo Capitolo sono esplorate le **potenzialità dei grafici** elaborati con i dati ottenuti dal software della **T.T.E.S.** e con il programma Excel.

L'obiettivo è quello di evidenziare la **validità**, l'**attendibilità** e la **sensibilità** dei calcoli effettuati per la generazione e la rappresentazione grafica dei dati.

La **validità** di uno *strumento* è il grado in cui lo strumento misura ciò che si prefigge di misurare. Nel caso in esame, la **validità** è provata se le relazioni tra le sequenze di Dna (o Rna), i calcoli effettuati per la generazione di tutti i grafici e le loro rappresentazioni sono *effettive, non dovute al caso*.

La **attendibilità** di uno *strumento* è il grado in cui lo strumento misura ciò che si prefigge di misurare, nonostante la presenza di fattori estranei e di disturbo.

Nel caso in esame, la **attendibilità** è provata se le relazioni tra le sequenze di Dna (o Rna), i calcoli effettuati per la generazione di tutti i grafici e le loro rappresentazioni restano **coerenti** nonostante la *lunghezza delle sequenze*, i *diversi tipi di “trend”*, la “*qualità*” delle singole basi azotate (per questi due ultimi aspetti si veda il paragrafo 2.2) il fatto che le *sequenze appartengano allo stesso organismo o a organismi diversi o a specie differenti*, ecc.

La **sensibilità** di uno *strumento* è la capacità dello strumento di discriminare fenomeni diversi e di evidenziare gli aspetti che differiscono tra questi fenomeni.

A tal proposito, dal **Confronto delle Analisi Complete delle 4 Sequenze dell'organismo Macaca Nemestrina** considerate in questo capitolo, ci aspettiamo che:

- 1) la sequenza XM_01172131**9**.1 (Product = "insulin isoform **X2**") venga rappresentata graficamente in maniera **leggermente diversa** dalle altre tre sequenze (Product = "insulin isoform X1");
- 2) le sequenze XM_01172131**6**.1, XM_01172131**7**.1 e XM_01172131**8**.1 siano rappresentate graficamente in maniera **abbastanza simile**, nonostante la *quantità* e “*qualità*” delle singole basi azotate siano in parte differenti in ciascuna di queste sequenze.

CAPITOLO II°

**Analisi
Completa della
Sequenza XM_011721319.1**

**PREDICTED: Macaca nemestrina insulin (INS),
transcript variant X4, mRNA**

2.1 CARATTERISTICHE DELLA SEQUENZA XM_011721319.1

PREDICTED: *Macaca nemestrina* insulin (INS), transcript variant X4, mRNA

Dei numerosi risultati ottenuti dalla ricerca **BLAST** (*Basic Local Alignment Search Tool* (1)) eseguita sulla **Catena A dell'Insulina** (si veda il Paragrafo 1.1 del Capitolo I° - Parte Seconda, [[“Analisi e modificazioni di Sequenze di DNA o di RNA con la T T E S \(Capitolo I° - Parte Seconda\)”](#)]), qui evidenziamo l'allineamento significativo con lo mRNA dell'organismo **Macaca nemestrina** (SEQUENZA XM_011721319.1).

PREDICTED: *Macaca nemestrina* insulin (INS), transcript variant X4, mRNA

Sequence ID: XM_011721319.1 2/1 3/1 8/1 17/1

Product="insulin isoform X2"

Length=297 Number of Matches: 1

Range 1: 232 to 294 [GenBank](#) [Graphics](#) [FASTA](#)

Score	Expect	Identities	Gaps	Strand
110 bits(121)	7e-21	62/63(98%)	0/63(0%)	Plus/Plus
Query 1	GGCATCGTGGAGCAGTGCTGCACCAGCATCTGTTCCCTCTACCAGCTGGAGAACTACTGC	60		
Sbjct 232	GGCATCGTGGAGCAGTGCTGCACCAGCATCTGCTCCCTCTACCAGCTGGAGAACTACTGC	291		
Query 61	AAC	63		
Sbjct 292	AAC	294		

PREDICTED: *Macaca nemestrina* insulin (INS), transcript variant X4, mRNA (NCBI Reference Sequence: XM_011721319.1)

LOCUS XM_011721319 297 bp mRNA linear PRI 24-APR-2018
DEFINITION PREDICTED: *Macaca nemestrina* insulin (INS), transcript variant X4, mRNA.
ACCESSION XM_011721319
VERSION XM_011721319.1
DBLINK BioProject: [PRJNA279145](#)
KEYWORDS RefSeq.
SOURCE *Macaca nemestrina* (pig-tailed macaque)
ORGANISM *Macaca nemestrina*
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Cercopithecidae; Cercopithecinae; *Macaca*.
COMMENT MODEL REFSEQ: This record is predicted by automated computational analysis. This record is derived from a genomic sequence ([NW_012013911.1](#)) annotated using gene prediction method: Gnomon, supported by mRNA evidence.
Also see:
[Documentation](#) of NCBI's Annotation Process

##Genome-Annotation-Data-START##
Annotation Provider :: NCBI
Annotation Status :: Full annotation
Annotation Name :: [Macaca nemestrina Annotation Release 101](#)
Annotation Version :: 101
Annotation Pipeline :: NCBI eukaryotic genome annotation pipeline
Annotation Software Version :: [8.0](#)
Annotation Method :: Best-placed RefSeq; Gnomon
Features Annotated :: Gene; mRNA; CDS; ncRNA
##Genome-Annotation-Data-END##

```

FEATURES             Location/Qualifiers
     source            1..297
                        /organism="Macaca nemestrina"
                        /mol_type="mRNA"
                        /isolate="M95218"
                        /db_xref="taxon:9545"
                        /chromosome="Unknown"
                        /sex="female"
                        /tissue_type="blood"
     gene              1..297
                        /gene="INS"
                        /note="Derived by automated computational analysis using
gene prediction method: Gnomon. Supporting evidence
includes similarity to: 1 mRNA, 1 Protein, and 17%
coverage of the annotated genomic feature by RNAseq
alignments"
                        /db_xref="GeneID:105469786"
     CDS              1..297
                        /gene="INS"
                        /codon_start=1
                        /product="insulin isoform X2"
                        /protein_id="XP_011719621.1"
                        /db_xref="GeneID:105469786"
                        /translation="MALWMRLRLPLALLALLWGPDPAPAFVNQHLGSHLVEALYLVC
ERGFFYTPKTRREAEDPQGSLLQPLALEGSLQKRGIVEQCCTSIICSLYQLENYCN"
     STS              1..296
                        /gene="INS"
                        /standard_name="PMC123023P3"
                        /db_xref="UniSTS:270424"
     STS              114..295
                        /gene="INS"
                        /standard_name="Ins1"
                        /db_xref="UniSTS:267003"
     STS              139..290
                        /gene="INS"
                        /standard_name="PMC24644P6"
                        /db_xref="UniSTS:265494"
ORIGIN
1   atggccctgt  ggatgcgcct  cttgcccttg  ctggcgctgc  tggccctctg  gggacctgac
61  ccggcccccg  cctttgtgaa  ccagcacctg  tgcggctccc  acctggtgga  agctctctac
121 ctggtgtgcg  gggagcgagg  cttcttctac  acaccaaga  cccgccggga  ggcagaggac
181 cctcagggca  gcctgcagcc  cttggcgctg  gaggggtccc  tgcagaagcg  cggcacgtcg
241 gagcagtgtc  gcaccagcat  ctgctccctc  taccagctgg  agaactactg  caactag
//

```

Le informazioni sulle caratteristiche della **SEQUENZA XM_011721319.1** sopra riportata sono state acquisite direttamente dal sito del **NCBI** [*National Center for Biotechnology Information* (2)].

- (1) 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/>
- (2) 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/>

2.2 RISULTATI GRAFICI DELL'ANALISI DELLA SEQUENZA XM_011721319.1 PREDICTED: Macaca nemestrina insulin (INS), transcript variant X4, mRNA

Tutti i grafici presentati in questo paragrafo si riferiscono all'intera sequenza di basi analizzata.

Il **Profilo degli 8 Codici Principali** (si veda il grafico in Fig. 1) costituisce una sintesi molto generale dell'intera sequenza.

Esso *SINTETIZZA IN MANIERA NON SPECIFICA* l'apporto congiunto delle *singole basi* (in minor misura) e del *"Trend"* (in maggior misura) nel *caratterizzare* la sequenza.

Questo **Profilo** non è limitato in alcun modo dalla lunghezza della sequenza; tuttavia, più lunga è la sequenza, minore è la capacità di evidenziare *aspetti peculiari della sintesi dell'intera sequenza*.

Questo grafico è molto utile per evidenziare le *caratteristiche generali* della sequenza (soprattutto, riguardo al ruolo svolto dal "Trend").

Attenzione: **A GRAFICI SIMILI potrebbero corrispondere SEQUENZE DIFFERENTI, MA CON CARATTERISTICHE GENERALI SIMILI.**

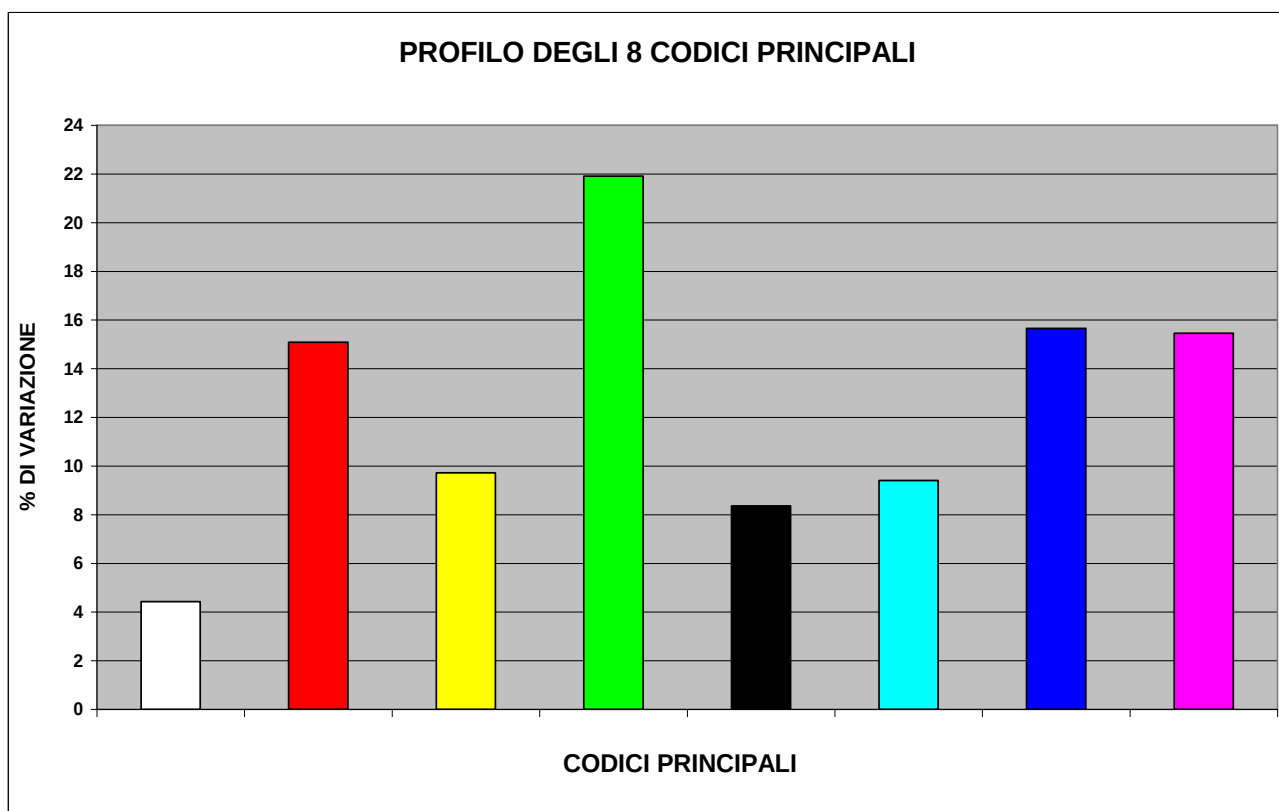


Fig. 1

La **Distribuzione della Percentuale di Variazione degli 8 Codici Principali** (si veda il grafico in Fig. 2) è un grafico che evidenzia, in maniera molto chiara, precisi aspetti del “Trend” dell’intera sequenza.

Questo **Profilo** non è limitato dalla lunghezza della sequenza; ; tuttavia, più lunga è la sequenza, minore è la capacità del grafico di rendere percettivamente discriminabili gli *aspetti peculiari del “Trend”*.

Questo grafico è molto utile per evidenziare le *caratteristiche del “Trend” della sequenza.*

Attenzione: **A GRAFICI SIMILI potrebbero corrispondere SEQUENZE DIFFERENTI, MA CON ASPETTI DEL “TREND” SIMILI.**

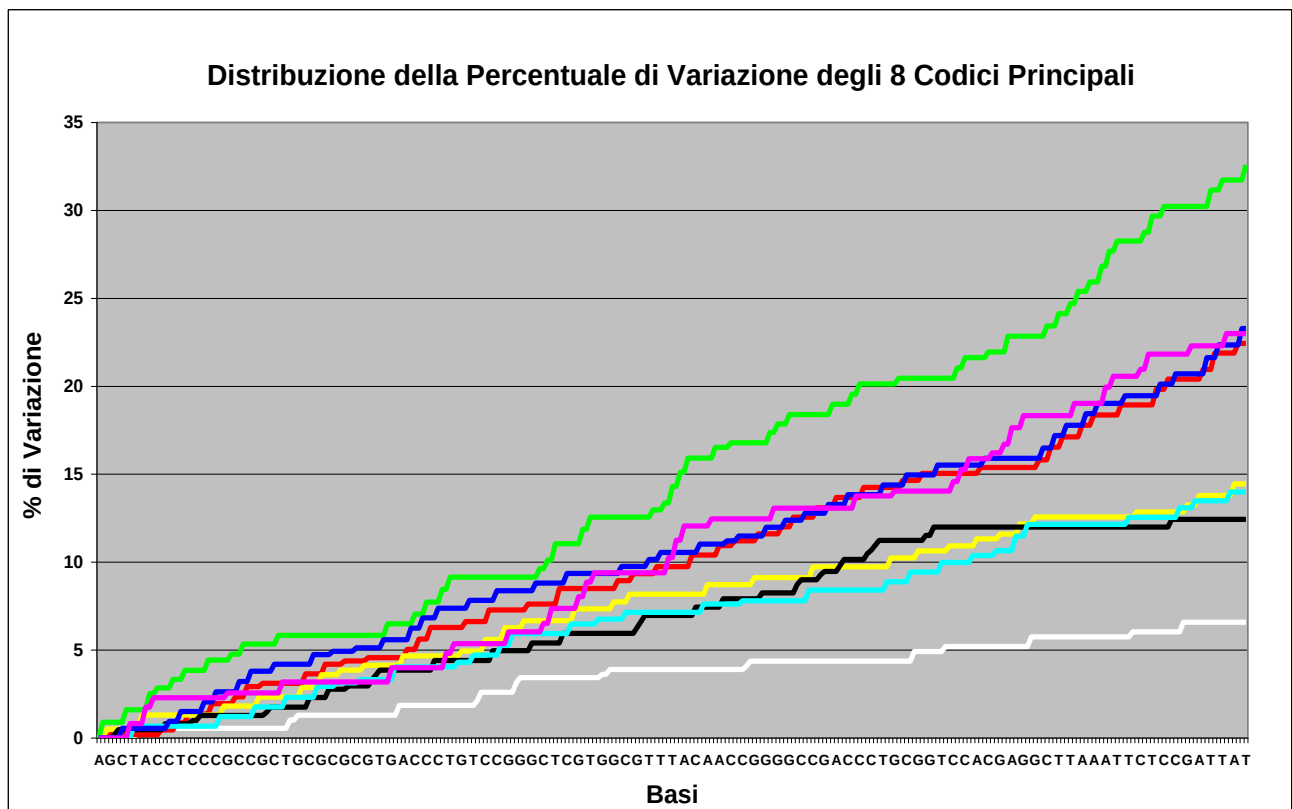


Fig. 2

Il grafico delle **Tonalità dei 64 Codici Totali** (si veda il grafico in Fig. 3) è quello che evidenzia le singole basi dell'intera sequenza, mentre offre informazioni molto meno chiare sul suo "Trend".

Questo **Profilo** non è limitato dalla lunghezza della sequenza; tuttavia, più lunga è la sequenza, minore è la capacità del grafico di rendere le *singole basi* percettivamente discriminabili.

Questo grafico è molto utile per evidenziare le "qualità" (Tonalità e % di Variazione) delle *singole basi* della sequenza.

Attenzione: **A GRAFICI SIMILI corrispondono SEQUENZE SIMILI.**

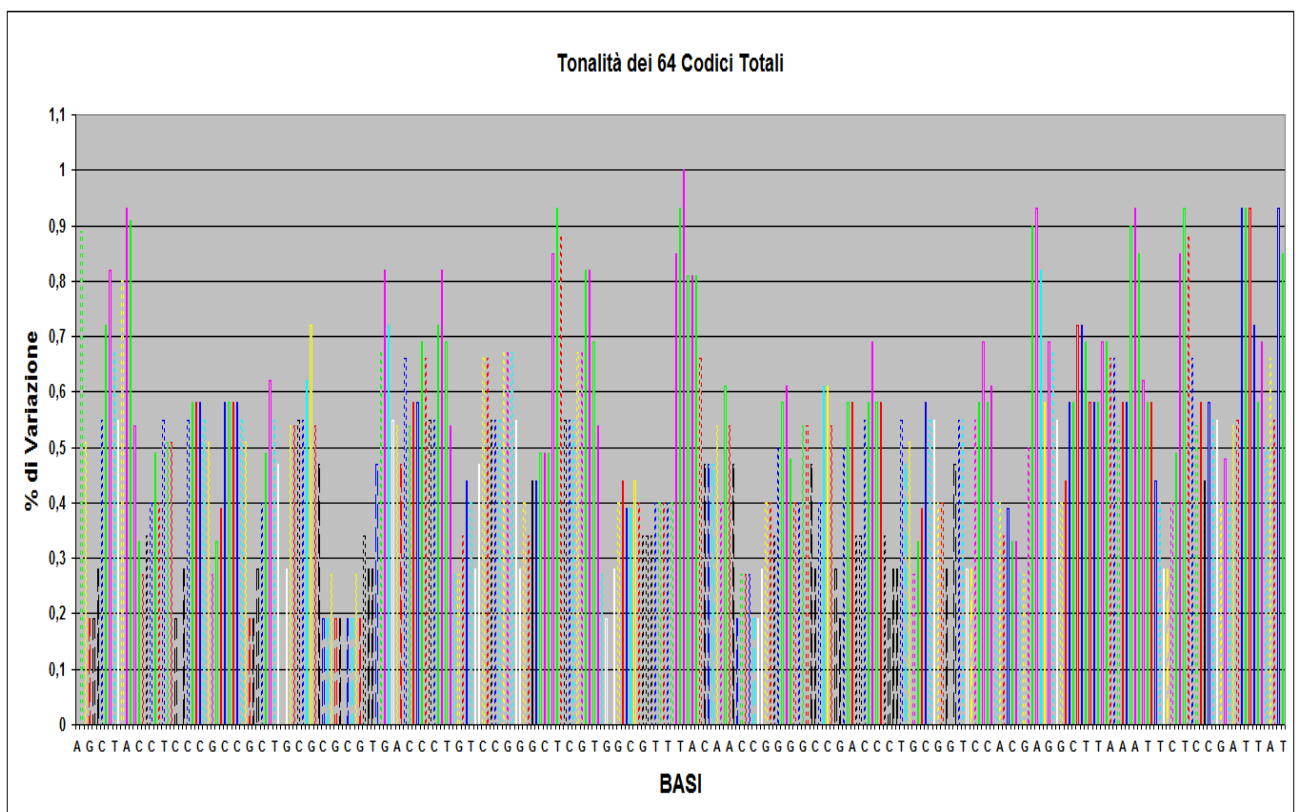


Fig. 3

Il **Profilo dei 64 Codici Totali** mostrato in Fig. 4 costituisce una sintesi molto specifica dell'intera sequenza.

Esso *SINTETIZZA IN MANIERA SPECIFICA* e nel modo migliore l'apporto congiunto delle *singole basi* (in minor misura) e del "*Trend*" (in maggior misura) nel *caratterizzare* la sequenza.

Questo **Profilo** non è limitato in alcun modo dalla lunghezza della sequenza; tuttavia, più lunga è la sequenza, minore è la capacità di evidenziare *aspetti altamente peculiari della sintesi dell'intera sequenza*.

Questo grafico è molto utile per evidenziare rapidamente le *caratteristiche indentificative di una sequenza*.

Attenzione: A GRAFICI SIMILI potrebbero corrispondere SEQUENZE DIFFERENTI, MA CON CARATTERISTICHE SPECIFICHE SIMILI.

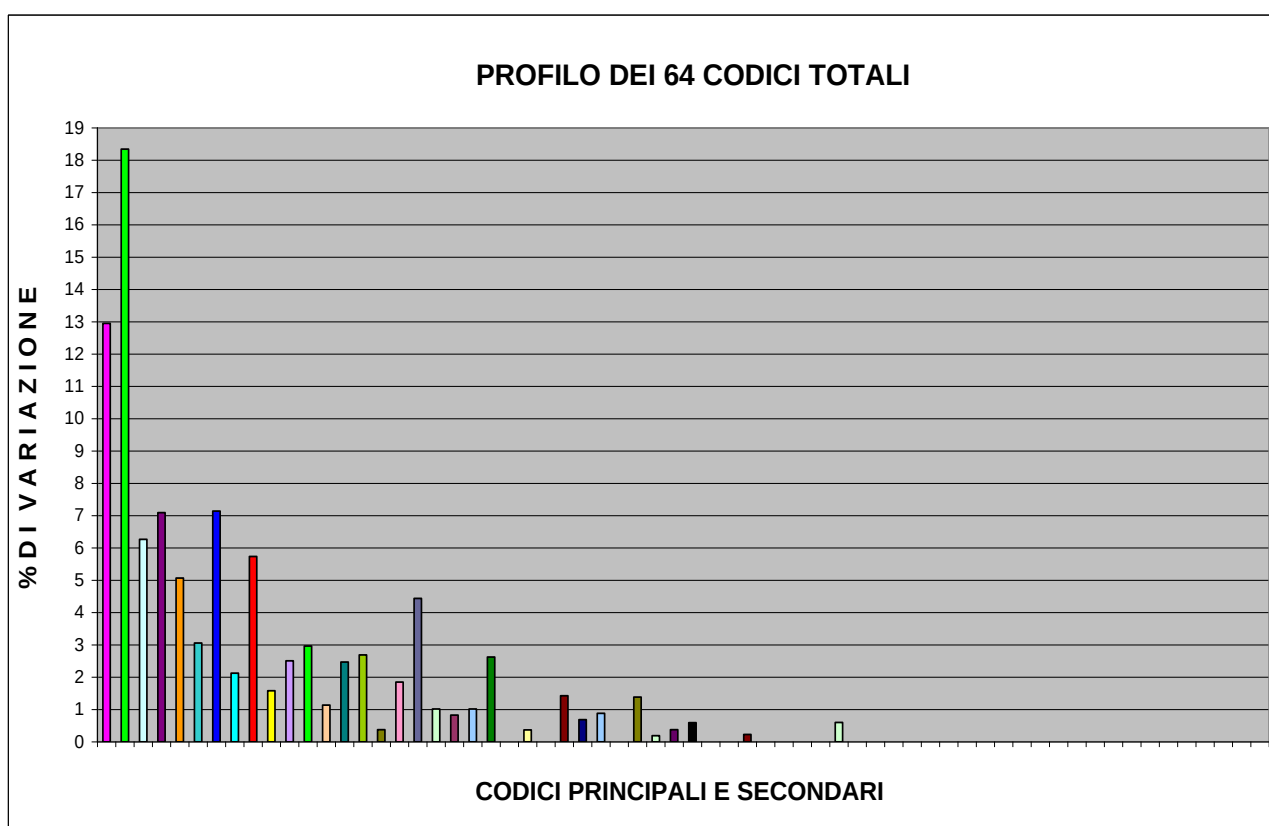


Fig. 4

**Analisi
Completa della
Sequenza XM_011721318.1**

**PREDICTED: Macaca nemestrina insulin (INS),
transcript variant X3, mRNA**

2.3 CARATTERISTICHE DELLA SEQUENZA XM_011721318.1

PREDICTED: Macaca nemestrina insulin (INS), transcript variant X3, mRNA

Dei numerosi risultati ottenuti dalla ricerca **BLAST** (*Basic Local Alignment Search Tool* (1)) eseguita sulla **Catena A dell'Insulina** (si veda il Paragrafo 1.1 del Capitolo I° - Parte Seconda, [[Analisi e modificazioni di Sequenze di DNA o di RNA con la T T E S \(Capitolo I° - Parte Seconda\)](#)]), qui noi evidenziamo l'*allineamento significativo* con lo **mRNA (SEQUENZA XM_011721318.1)** dell'*organismo Macaca nemestrina*.

PREDICTED: Macaca nemestrina insulin (INS), transcript variant X3, mRNA

Sequence ID: **XM_011721318.1** 2/1 3/1 8/1 17/1

Product="insulin isoform X1"

Length=532 Number of Matches: 1

Range 1: 395 to 457 [GenBank](#) [Graphics](#) [FASTA](#)

Score	Expect	Identities	Gaps	Strand
110 bits(121)	7e-21	62/63(98%)	0/63(0%)	Plus/Plus
Query 1	GGCATCGTGGAGCAGTGCTGCACCAGCATCTGTTCCCTCTACCAGCTGGAGAACTACTGC	60		
Sbjct 395	GGCATCGTGGAGCAGTGCTGCACCAGCATCTGCTCCCTCTACCAGCTGGAGAACTACTGC	454		
Query 61	AAC 63			
Sbjct 455	AAC 457			

PREDICTED: Macaca nemestrina insulin (INS), transcript variant X3, mRNA (NCBI Reference Sequence: XM_011721318.1)

LOCUS XM_011721318 532 bp mRNA linear PRI 24-APR-2018
DEFINITION PREDICTED: Macaca nemestrina insulin (INS), transcript variant X3, mRNA.
ACCESSION XM_011721318
VERSION XM_011721318.1
DBLINK BioProject: [PRJNA279145](#)
KEYWORDS RefSeq.
SOURCE Macaca nemestrina (pig-tailed macaque)
ORGANISM [Macaca nemestrina](#)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Cercopithecidae; Cercopithecinae; Macaca.
COMMENT MODEL [REFSEQ](#): This record is predicted by automated computational analysis. This record is derived from a genomic sequence ([NW_012013911.1](#)) annotated using gene prediction method: Gnomon, supported by mRNA evidence.
Also see:
[Documentation](#) of NCBI's Annotation Process

##Genome-Annotation-Data-START##
Annotation Provider :: NCBI
Annotation Status :: Full annotation
Annotation Name :: [Macaca nemestrina Annotation Release 101](#)
Annotation Version :: 101
Annotation Pipeline :: NCBI eukaryotic genome annotation pipeline
Annotation Software Version :: [8.0](#)
Annotation Method :: Best-placed RefSeq; Gnomon
Features Annotated :: Gene; mRNA; CDS; ncRNA
##Genome-Annotation-Data-END##
FEATURES Location/Qualifiers
source 1..532

```

/organism="Macaca nemestrina"
/mol_type="mRNA"
/isolate="M95218"
/db_xref="taxon:9545"
/chromosome="Unknown"
/sex="female"
/tissue_type="blood"
gene
1..532
/feature="INS"
/note="Derived by automated computational analysis using
gene prediction method: Gnomon. Supporting evidence
includes similarity to: 2 mRNAs, 10 Proteins, and 9%
coverage of the annotated genomic feature by RNAseq
alignments"
/feature="GeneID:105469786"
STS
8..254
/feature="INS"
/standard_name="GDB:181496"
/feature="UniSTS:155248"
CDS
128..460
/feature="INS"
/codon_start=1
/product="insulin isoform X1"
/protein_id="XP_011719620.1"
/feature="GeneID:105469786"
/translation="MALWMRLPLALLALWGPDPAFAFVNQHLCSHLVEALYLVCG
ERGFFYTPKTRREAEDPQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSL
YQLENYCN"
STS
128..459
/feature="INS"
/standard_name="PMC123023P3"
/feature="UniSTS:270424"
STS
241..458
/feature="INS"
/standard_name="Ins1"
/feature="UniSTS:267003"
STS
266..453
/feature="INS"
/standard_name="PMC24644P6"
/feature="UniSTS:265494"
STS
315..532
/feature="INS"
/standard_name="GDB:179433"
/feature="UniSTS:155046"

```

ORIGIN

```

1 gggacaggct gcatcagaag aggccagcaa gcaggtctgt tccaagggcc ttcacgtcag
61 gtgggctcag ggctgcccc cttggggggt ccaggggtggc tggacccag gtcactgtcc
121 ttccgccatg gccctgtgga tgcgcctctt gcccctgctg gcgctgctgg ccctctgggg
181 acctgaccgc gcccggcct ttgtgaacca gcacctgtgc ggctccacc tgggtggaagc
241 tctctacctg gtgtgcgggg agcgaggctt cttctacaca cccaagacc gccgggaggc
301 agaggaccct caggtggggc aggtggagct gggcgggggc cctggcgag gcagcctgca
361 gcccttggcg ctggaggggt ccctgcagaa gcgcggcatc gtggagcagt gctgcaccag
421 catctgtctc ctctaccagc tggagaacta ctgcaactag atgcggccc caggcgggcc
481 acaccctcca cctctgcac caagagagat cgaataaagc ccttgaacca gc

```

//

Le informazioni sulle caratteristiche della **SEQUENZA XM_011721318.1** sopra riportata sono state acquisite direttamente dal sito del **NCBI** [*National Center for Biotechnology Information* (2)].

- (1) 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/>
- (2) 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/>

2.4 RISULTATI GRAFICI DELL'ANALISI DELLA SEQUENZA XM_011721318.1 PREDICTED: Macaca nemestrina insulin (INS), transcript variant X3, mRNA

Tutti i grafici presentati in questo paragrafo si riferiscono all'intera sequenza di basi analizzata.

Per non appesantire inutilmente la lettura di questo lavoro, riguardo l'interpretazione di questi grafici rimando alle spiegazioni esposte nel **paragrafo 1.1**.

Il **Profilo degli 8 Codici Principali** (si veda il grafico in Fig. 5) costituisce una sintesi molto generale dell'intera sequenza.

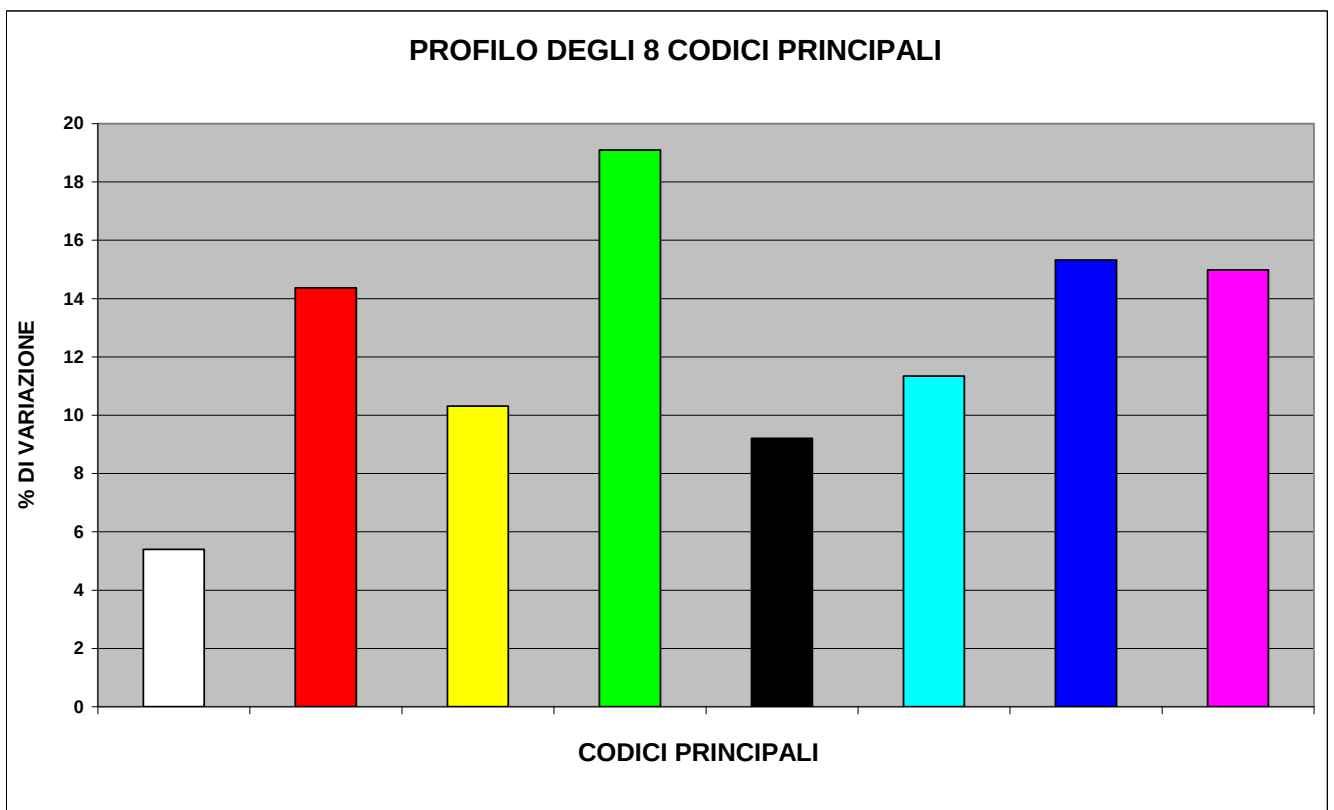


Fig. 5

La **Distribuzione della Percentuale di Variazione degli 8 Codici Principali** (si veda il grafico in Fig. 6) è un grafico che evidenzia, in maniera molto chiara, precisi aspetti del “Trend” dell’intera sequenza.

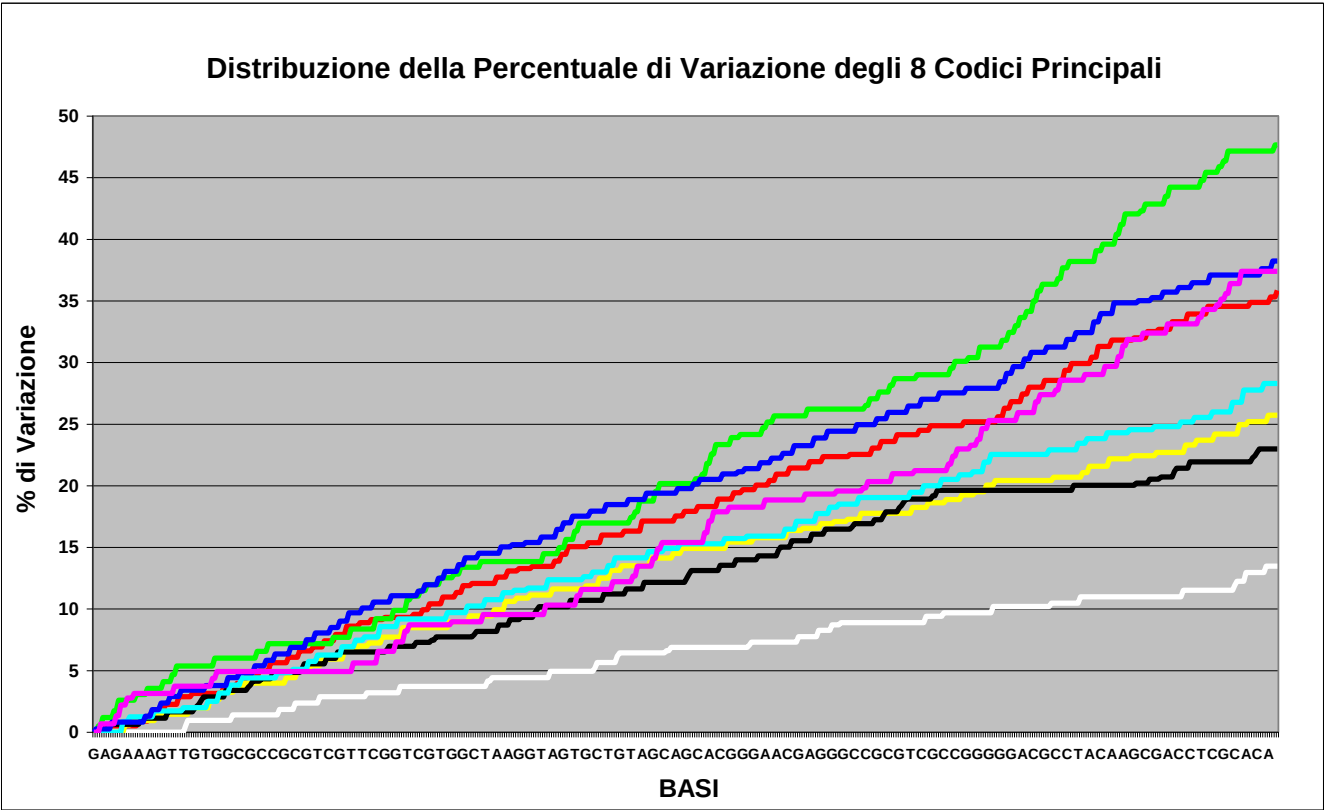


Fig. 6

Il grafico delle **Tonalità dei 64 Codici Totali** (si veda il grafico in Fig. 7) è quello che evidenzia le singole basi dell'intera sequenza, mentre offre informazioni molto meno chiare sul suo "Trend".

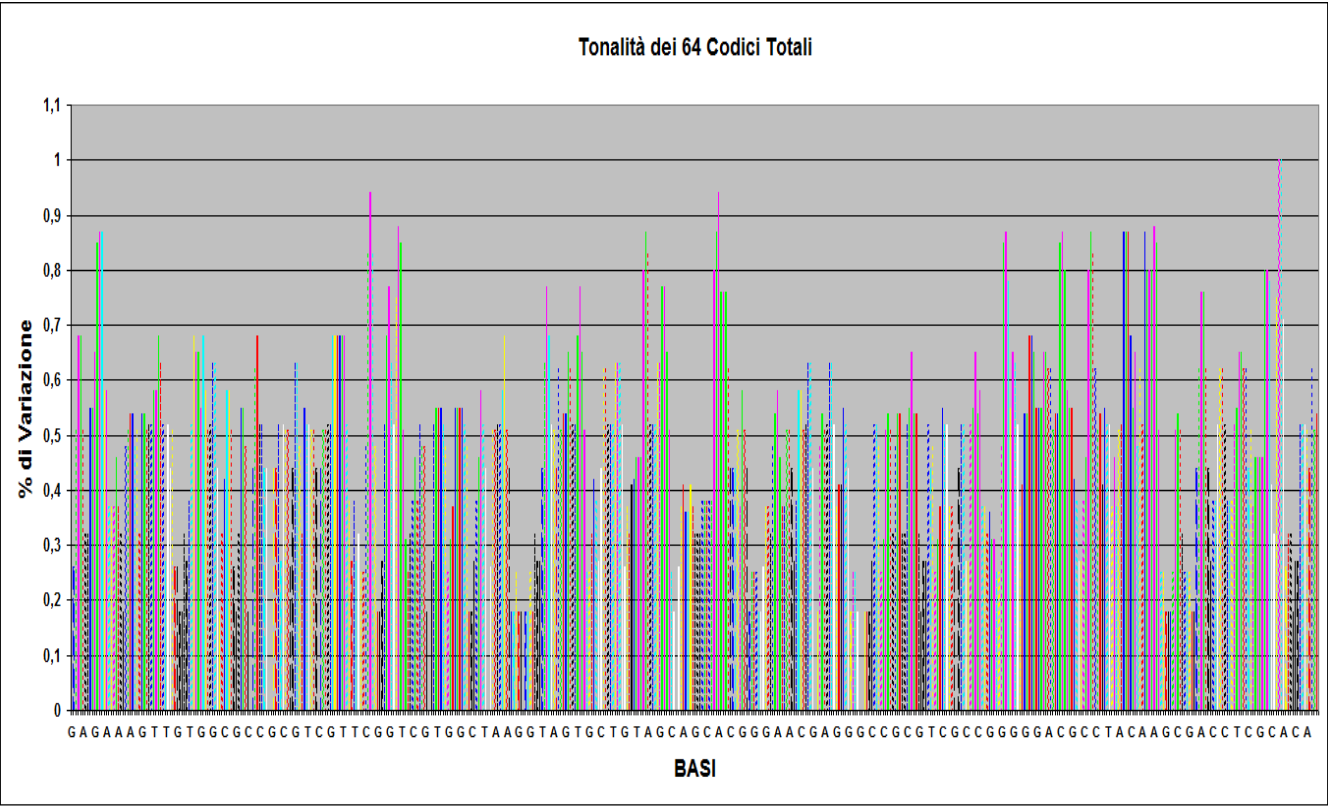


Fig. 7

Il **Profilo dei 64 Codici Totali** mostrato in Fig. 8 costituisce una sintesi molto specifica dell'intera sequenza.

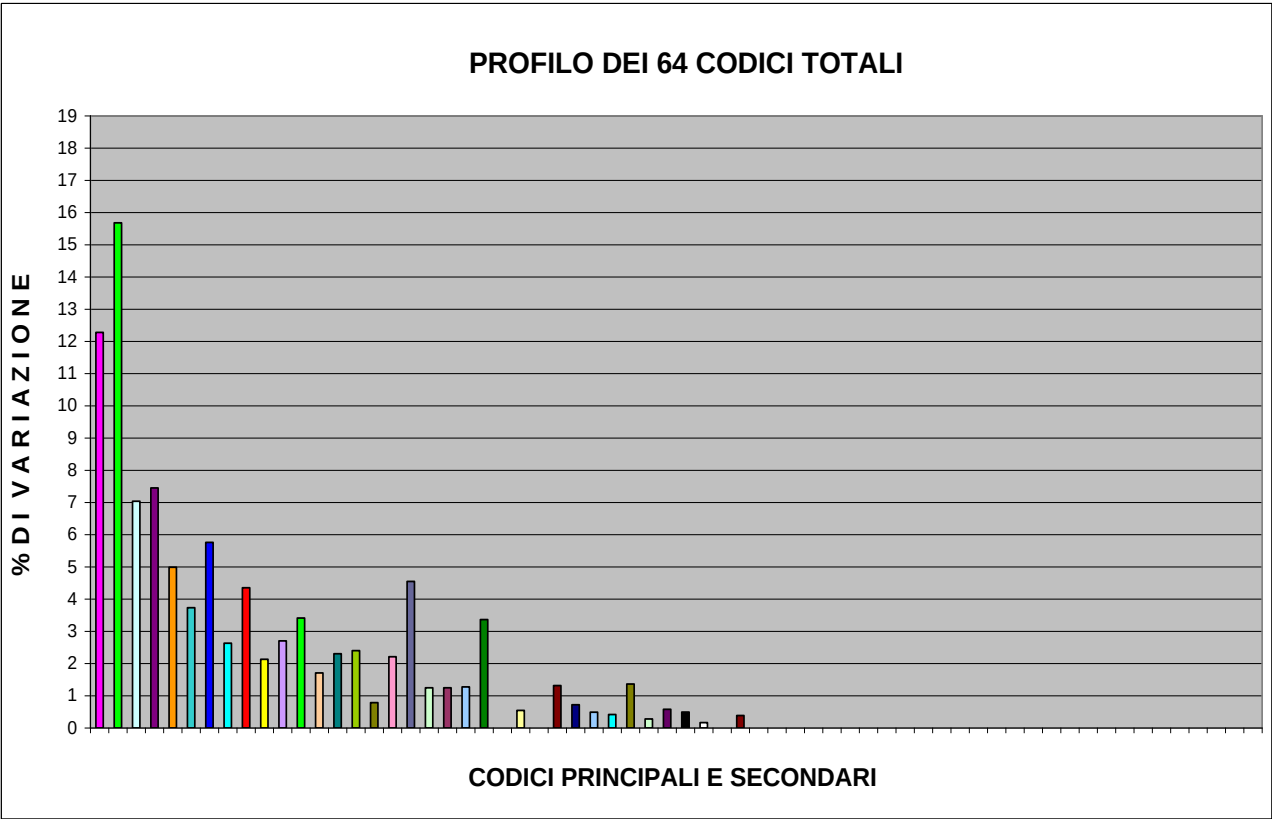


Fig. 8

**Analisi
Completa della
Sequenza XM_011721317.1**

**PREDICTED: Macaca nemestrina insulin (INS),
transcript variant X2, mRNA**

2.5 CARATTERISTICHE DELLA SEQUENZA XM_011721317.1

PREDICTED: Macaca nemestrina insulin (INS), transcript variant X2, mRNA

Dei numerosi risultati ottenuti dalla ricerca **BLAST** (*Basic Local Alignment Search Tool* (1)) eseguita sulla **Catena A dell'Insulina** (si veda il Paragrafo 1.1 del Capitolo I° - Parte Seconda, [[Analisi e modificazioni di Sequenze di DNA o di RNA con la T T E S \(Capitolo I° - Parte Seconda\)](#)]), qui noi evidenziamo l'*allineamento significativo* con lo **mRNA (SEQUENZA XM_011721317.1)** dell'*organismo Macaca nemestrina*.

PREDICTED: Macaca nemestrina insulin (INS), transcript variant X2, mRNA

Sequence ID: **XM_011721317.1** **2/1** **3/1** **8/1** **17/1**

Product="insulin isoform X1"

Length=482 Number of Matches: 1

Range 1: 345 to 407 [GenBank](#) [Graphics](#) [FASTA](#)

Score	Expect	Identities	Gaps	Strand
110 bits(121)	7e-21	62/63(98%)	0/63(0%)	Plus/Plus
Query 1	GGCATCGTGGAGCAGTGCTGCACCAGCATCTGTTCCCTCTACCAGCTGGAGAACTACTGC	60		
Sbjct 345	GGCATCGTGGAGCAGTGCTGCACCAGCATCTGCTCCCTCTACCAGCTGGAGAACTACTGC	404		
Query 61	AAC 63			
Sbjct 405	AAC 407			

PREDICTED: Macaca nemestrina insulin (INS), transcript variant X2, mRNA (NCBI Reference Sequence: XM_011721317.1)

LOCUS XM_011721317 482 bp mRNA linear PRI 24-APR-2018

DEFINITION PREDICTED: Macaca nemestrina insulin (INS), transcript variant X2, mRNA.

ACCESSION XM_011721317

VERSION XM_011721317.1

DBLINK BioProject: [PRJNA279145](#)

KEYWORDS RefSeq.

SOURCE Macaca nemestrina (pig-tailed macaque)

ORGANISM [Macaca nemestrina](#)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Cercopithecidae; Cercopithecinae; Macaca.

COMMENT MODEL [REFSEQ](#): This record is predicted by automated computational analysis. This record is derived from a genomic sequence ([NW_012013911.1](#)) annotated using gene prediction method: Gnomon, supported by mRNA evidence.

Also see:

[Documentation](#) of NCBI's Annotation Process

##Genome-Annotation-Data-START##

Annotation Provider :: NCBI

Annotation Status :: Full annotation

Annotation Name :: [Macaca nemestrina Annotation Release 101](#)

Annotation Version :: 101

Annotation Pipeline :: NCBI eukaryotic genome annotation pipeline

Annotation Software Version :: [8.0](#)

Annotation Method :: Best-placed RefSeq; Gnomon

Features Annotated :: Gene; mRNA; CDS; ncRNA

##Genome-Annotation-Data-END##

FEATURES Location/Qualifiers

source 1..482

/organism="Macaca nemestrina"

/mol_type="mRNA"

```

/isolate="M95218"
/db_xref="taxon:9545"
/chromosome="Unknown"
/sex="female"
/tissue_type="blood"
gene 1..482
/gene="INS"
/note="Derived by automated computational analysis using
gene prediction method: Gnomon. Supporting evidence
includes similarity to: 3 mRNAs, 10 Proteins, and 10%
coverage of the annotated genomic feature by RNAseq
alignments"
STS /db_xref="GeneID:105469786"
8..204
/gene="INS"
/standard_name="GDB:181496"
CDS /db_xref="UniSTS:155248"
78..410
/gene="INS"
/codon_start=1
/product="insulin isoform X1"
/protein_id="XP_011719619.1"
/db_xref="GeneID:105469786"
/translation="MALWMRLRLPLLALLALWGPDPAPAFVNQHLCGSHLVEALYLVC
ERGFFYTPKTRREAEDPQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSL
YQLENYCN"
STS 78..409
/gene="INS"
/standard_name="PMC123023P3"
STS /db_xref="UniSTS:270424"
191..408
/gene="INS"
/standard_name="Ins1"
STS /db_xref="UniSTS:267003"
216..403
/gene="INS"
/standard_name="PMC24644P6"
STS /db_xref="UniSTS:265494"
265..482
/gene="INS"
/standard_name="GDB:179433"
/db_xref="UniSTS:155046"
ORIGIN
1 gggacaggct gcatcagaag aggccagcaa gcaggctctgt tccaagggcc ttcacgtcag
61 gtcactgtcc ttccgccatg gccctgtgga tgcgcctctt gcccctgctg gcgctgctgg
121 ccctctgggg acctgacctg gcccggcct ttgtgaacca gcacctgtgc ggctcccacc
181 tgggtggaagc tctctacctg gtgtgcgggg agcgaggctt cttctacaca cccaagacct
241 gccgggaggg agaggacct caggtggggc aggtggagct gggcgggggc cctggcgagc
301 gcagcctgca gcccttggcg ctggaggggt ccctgcagaa gcgcgggcac gtggagcagt
361 gctgcaccag catctgtctc ctctaccagc tggagaacta ctgcaactag atgcggcccg
421 caggcgggcc acaccctcca cctctgcac caagagagat cgaataaagc ccttgaacca
481 gc
//

```

Le informazioni sulle caratteristiche della **SEQUENZA XM_011721317.1** sopra riportata sono state acquisite direttamente dal sito del **NCBI** [*National Center for Biotechnology Information* (2)].

- (1) 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/>
- (2) 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/>

2.6 RISULTATI GRAFICI DELL'ANALISI DELLA SEQUENZA XM_011721317.1 PREDICTED: Macaca nemestrina insulin (INS), transcript variant X2, mRNA

Tutti i grafici presentati in questo paragrafo si riferiscono all'intera sequenza di basi analizzata.

Per non appesantire inutilmente la lettura di questo lavoro, riguardo l'interpretazione di questi grafici rimando alle spiegazioni esposte nel *paragrafo 1.1*.

Il **Profilo degli 8 Codici Principali** (si veda il grafico in Fig. 9) costituisce una sintesi molto generale dell'intera sequenza.

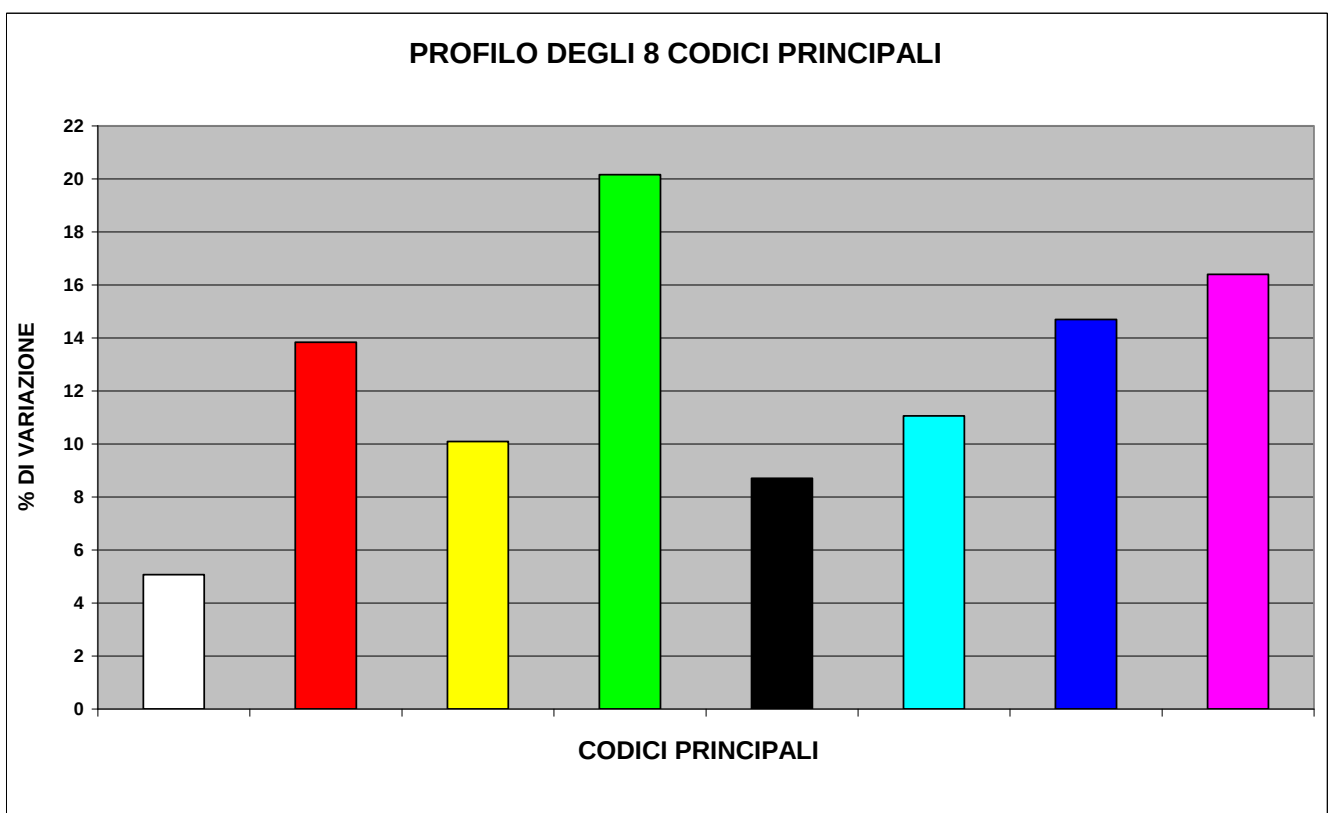


Fig. 9

La **Distribuzione della Percentuale di Variazione degli 8 Codici Principali** (si veda il grafico in Fig. 10) è un grafico che evidenzia, in maniera molto chiara, precisi aspetti del “Trend” dell’intera sequenza.

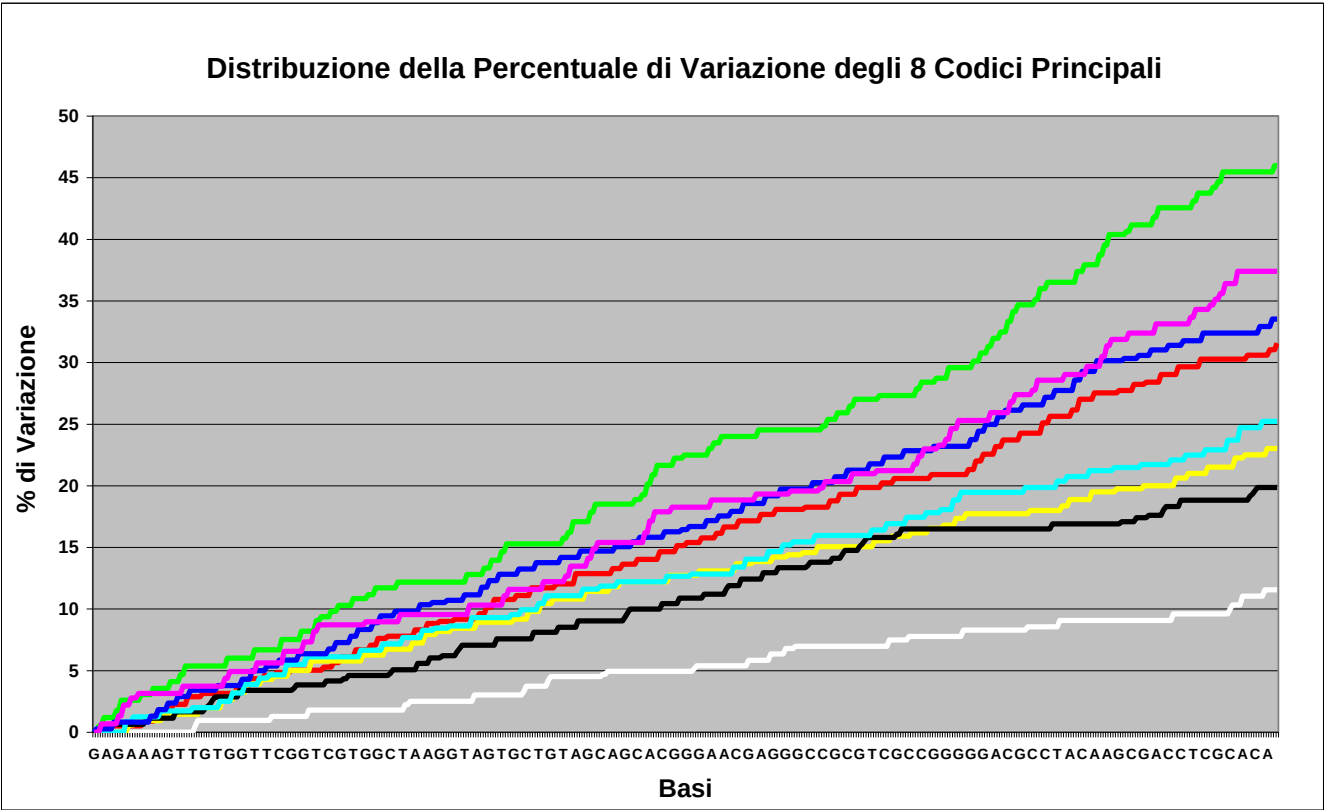


Fig. 10

Il grafico delle **Tonalità dei 64 Codici Totali** (si veda il grafico in Fig. 11) è quello che evidenzia le singole basi dell'intera sequenza, mentre offre informazioni molto meno chiare sul suo "Trend".

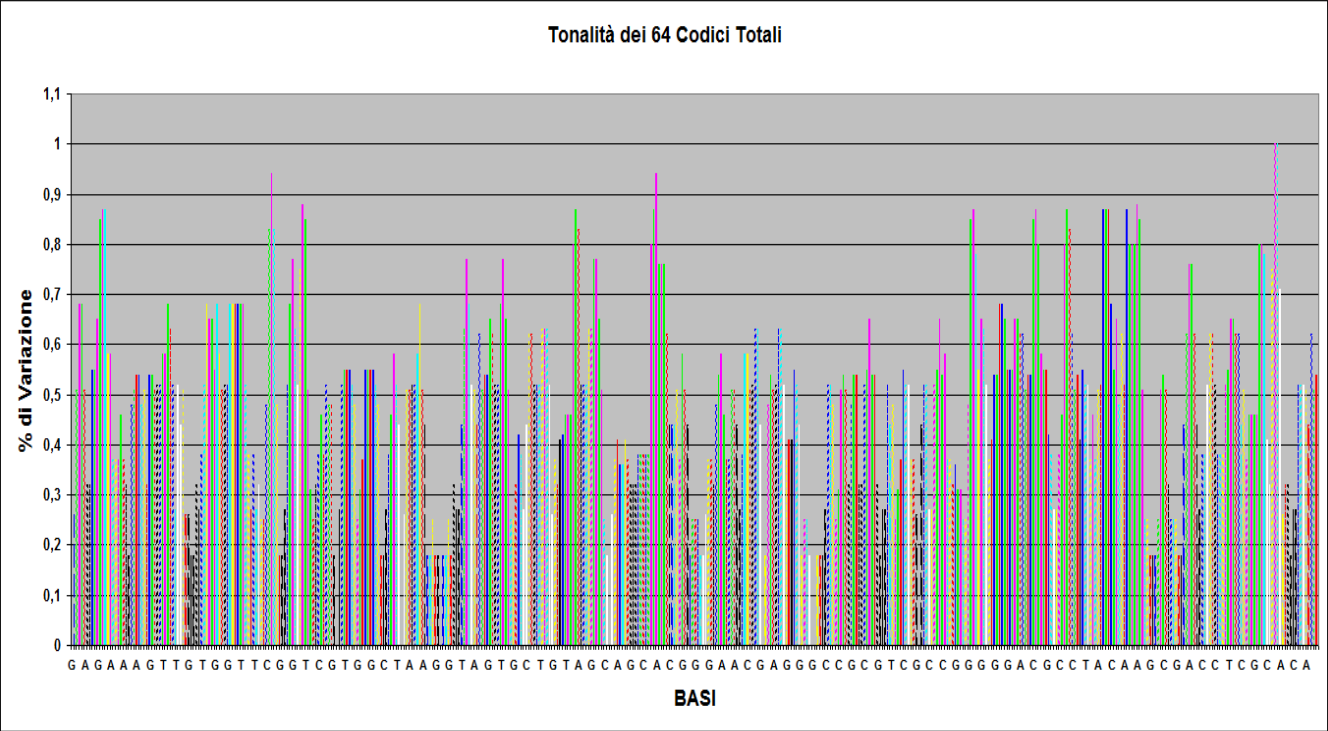


Fig. 11

Il **Profilo dei 64 Codici Totali** mostrato in Fig. 12 costituisce una sintesi molto specifica dell'intera sequenza.

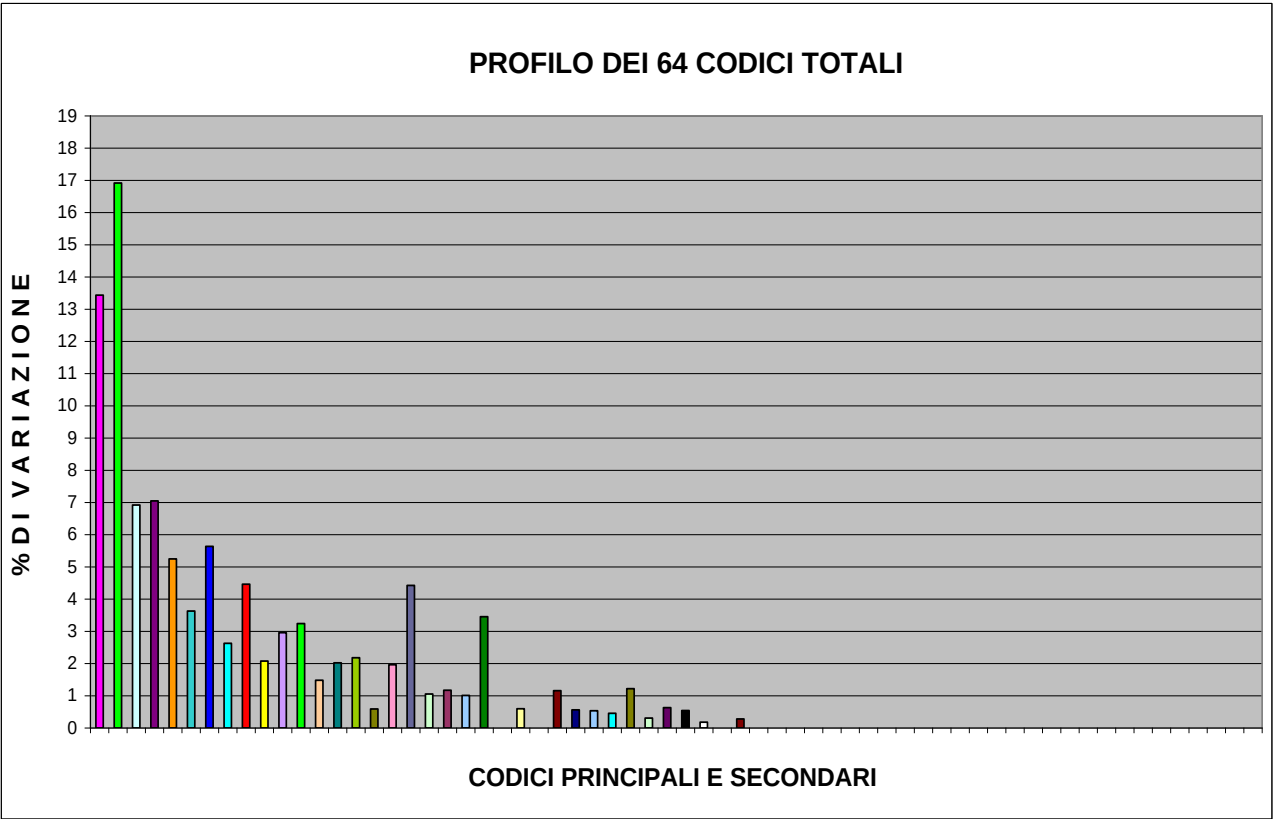


Fig. 12

**Analisi
Completa della
Sequenza XM_011721316.1**

**PREDICTED: Macaca nemestrina insulin (INS),
transcript variant X1, mRNA**

2.7 CARATTERISTICHE DELLA SEQUENZA XM_011721316.1

PREDICTED: *Macaca nemestrina* insulin (INS), transcript variant **X1**, mRNA

Dei numerosi risultati ottenuti dalla ricerca **BLAST** (*Basic Local Alignment Search Tool* (1)) eseguita sulla **Catena A dell'Insulina** (si veda il Paragrafo 1.1 del Capitolo I° - Parte Seconda, [[Analisi e modificazioni di Sequenze di DNA o di RNA con la T T E S \(Capitolo I° - Parte Seconda\)](#)]), qui noi evidenziamo l'allineamento significativo con lo mRNA (**SEQUENZA XM_011721316.1**) dell'organismo **Macaca nemestrina**.

PREDICTED: **Macaca nemestrina** insulin (INS), transcript variant X1, mRNA

Sequence ID: **XM_011721316.1** 2/1 3/1 8/1 17/1

Product="insulin isoform X1"

Length=454 Number of Matches: 1

Range 1: 319 to 381 [GenBank](#) [Graphics](#) [FASTA](#)

Score	Expect	Identities	Gaps	Strand
110 bits(121)	7e-21	62/63(98%)	0/63(0%)	Plus/Plus
Query 1	GGCATCGTGGAGCAGTGTGTCACCAAGCATCTGTTCCCTCTACCAAGCTGGAGAACTACTGC	60		
Sbjct 319	GGCATCGTGGAGCAGTGTGTCACCAAGCATCTGTTCCCTCTACCAAGCTGGAGAACTACTGC	378		
Query 61	AAC 63			
Sbjct 379	AAC 381			

PREDICTED: *Macaca nemestrina* insulin (INS), transcript variant **X1**, mRNA (NCBI Reference Sequence: XM_011721316.1)

LOCUS XM_011721316 454 bp mRNA linear PRI 24-APR-2018
DEFINITION PREDICTED: *Macaca nemestrina* insulin (INS), transcript variant X1, mRNA.
ACCESSION XM_011721316
VERSION XM_011721316.1
DBLINK BioProject: [PRJNA279145](#)
KEYWORDS RefSeq.
SOURCE *Macaca nemestrina* (pig-tailed macaque)
ORGANISM [Macaca nemestrina](#)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Cercopithecidae; Cercopithecinae; *Macaca*.
COMMENT MODEL [REFSEQ](#): This record is predicted by automated computational analysis. This record is derived from a genomic sequence ([NW_012013911.1](#)) annotated using gene prediction method: Gnomon, supported by mRNA evidence.
Also see:
[Documentation](#) of NCBI's Annotation Process

##Genome-Annotation-Data-START##
Annotation Provider :: NCBI
Annotation Status :: Full annotation
Annotation Name :: [Macaca nemestrina Annotation Release 101](#)
Annotation Version :: 101
Annotation Pipeline :: NCBI eukaryotic genome annotation pipeline
Annotation Software Version :: [8.0](#)
Annotation Method :: Best-placed RefSeq; Gnomon
Features Annotated :: Gene; mRNA; CDS; ncRNA
##Genome-Annotation-Data-END##

```

FEATURES             Location/Qualifiers
     source            1..454
                        /organism="Macaca nemestrina"
                        /mol_type="mRNA"
                        /isolate="M95218"
                        /db_xref="taxon:9545"
                        /chromosome="Unknown"
                        /sex="female"
                        /tissue_type="blood"
     gene             1..454
                        /gene="INS"
                        /note="Derived by automated computational analysis using
gene prediction method: Gnomon. Supporting evidence
includes similarity to: 5 mRNAs, 10 Proteins, and 11%
coverage of the annotated genomic feature by RNAseq
alignments"
                        /db_xref="GeneID:105469786"
     STS             8..178
                        /gene="INS"
                        /standard_name="GDB:181496"
                        /db_xref="UniSTS:155248"
     CDS             52..384
                        /gene="INS"
                        /codon_start=1
                        /product="insulin isoform X1"
                        /protein_id="XP_011719618.1"
                        /db_xref="GeneID:105469786"
                        /translation="MALWMRLLPLLALLLALWGPDPAFAFVNQHLGSHLVEALYLVCG
ERGFFYTPKTRREAEDPQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSL
YQLENYCN"
     STS             52..383
                        /gene="INS"
                        /standard_name="PMC123023P3"
                        /db_xref="UniSTS:270424"
     STS             165..382
                        /gene="INS"
                        /standard_name="Ins1"
                        /db_xref="UniSTS:267003"
     STS             190..377
                        /gene="INS"
                        /standard_name="PMC24644P6"
                        /db_xref="UniSTS:265494"
ORIGIN
1  gggacaggct  gcatcagaag  aggccagcaa  gcaggtcact  gtccttccgc  catggccctg
61  tggatgcgcc  tcttgcccct  gctggcgctg  ctggccctct  ggggacctga  cccggccccg
121 gcctttgtga  accagcacct  gtgcggtcc  cacctggtgg  aagctctcta  cctggtgtgc
181 ggggagcgag  gcttcttcta  cacaccaag  acccgccggg  aggcagagga  ccctcagggtg
241 gggcagggtg  agctgggcgg  gggccctggc  gcaggcagcc  tgcagccctt  ggcgctggag
301 gggtcacctg  agaagcgcg  catcgtggag  cagtgtgtga  ccagcatctg  ctccctctac
361 cagctggaga  actactgcaa  ctagatgcgg  cccgcaggcg  gcccacacc  tccacctct
421 gcaccaagag  agatcgaata  aagcccttga  acca
//

```

Le informazioni sulle caratteristiche della **SEQUENZA XM_011721316.1** sopra riportata sono state acquisite direttamente dal sito del **NCBI** [*National Center for Biotechnology Information* (2)].

- (1) 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/>
- (2) 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/>

2.8 RISULTATI GRAFICI DELL'ANALISI DELLA SEQUENZA XM_011721316.1 PREDICTED: Macaca nemestrina insulin (INS), transcript variant X1, mRNA

Tutti i grafici presentati in questo paragrafo si riferiscono all'intera sequenza di basi analizzata.

Per non appesantire inutilmente la lettura di questo lavoro, riguardo l'interpretazione di questi grafici rimando alle spiegazioni esposte nel *paragrafo 1.1* .

Il **Profilo degli 8 Codici Principali** (si veda il grafico in Fig. 13) costituisce una sintesi molto generale dell'intera sequenza.

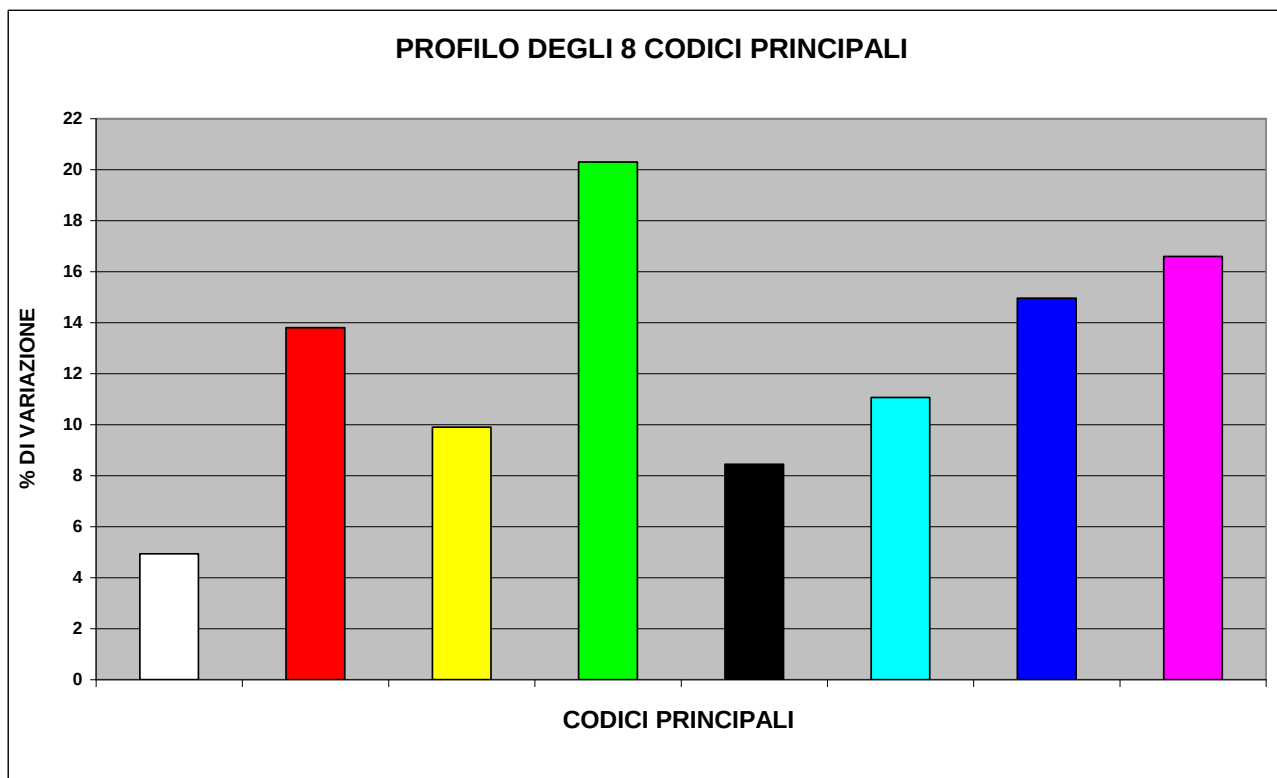


Fig. 13

La **Distribuzione della Percentuale di Variazione degli 8 Codici Principali** (si veda il grafico in Fig. 14) è un grafico che evidenzia, in maniera molto chiara, precisi aspetti del “Trend” dell’intera sequenza.

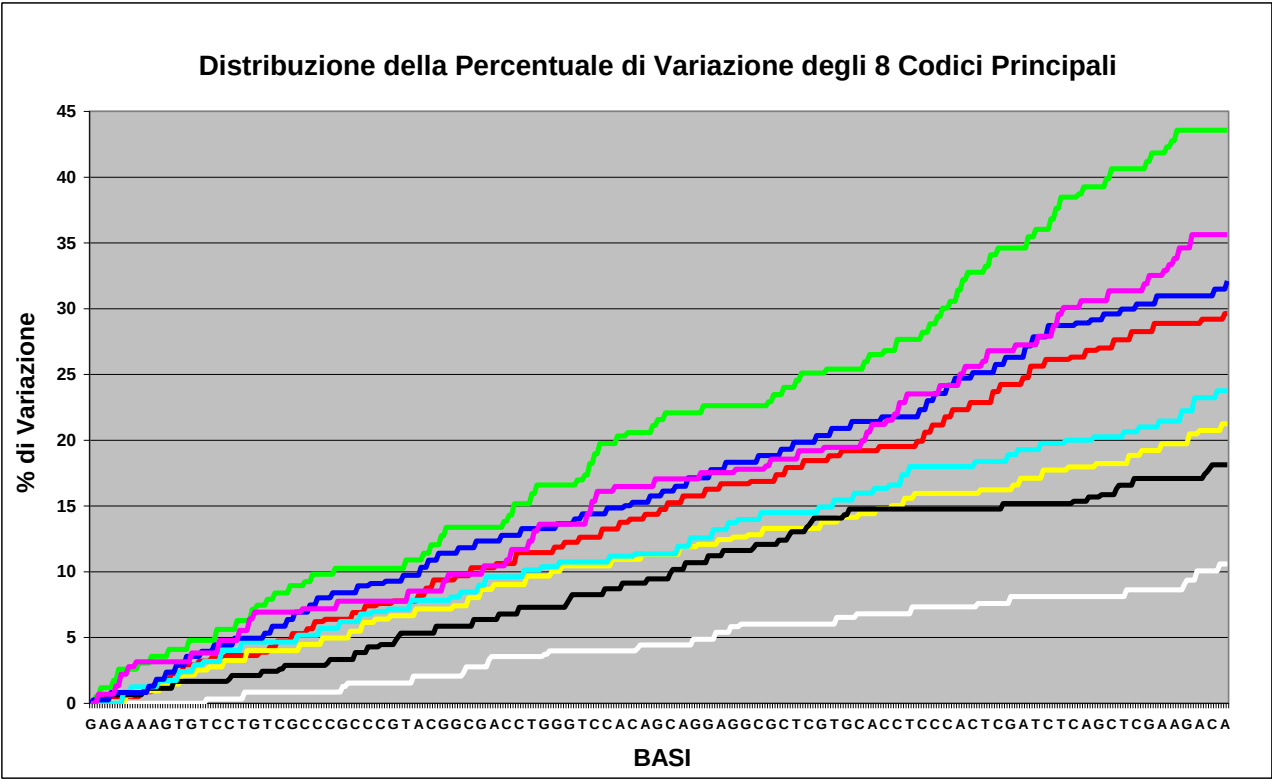


Fig. 14

Il grafico delle **Tonalità dei 64 Codici Totali** (si veda il grafico in Fig. 15) è quello che evidenzia le singole basi dell'intera sequenza, mentre offre informazioni molto meno chiare sul suo "Trend".

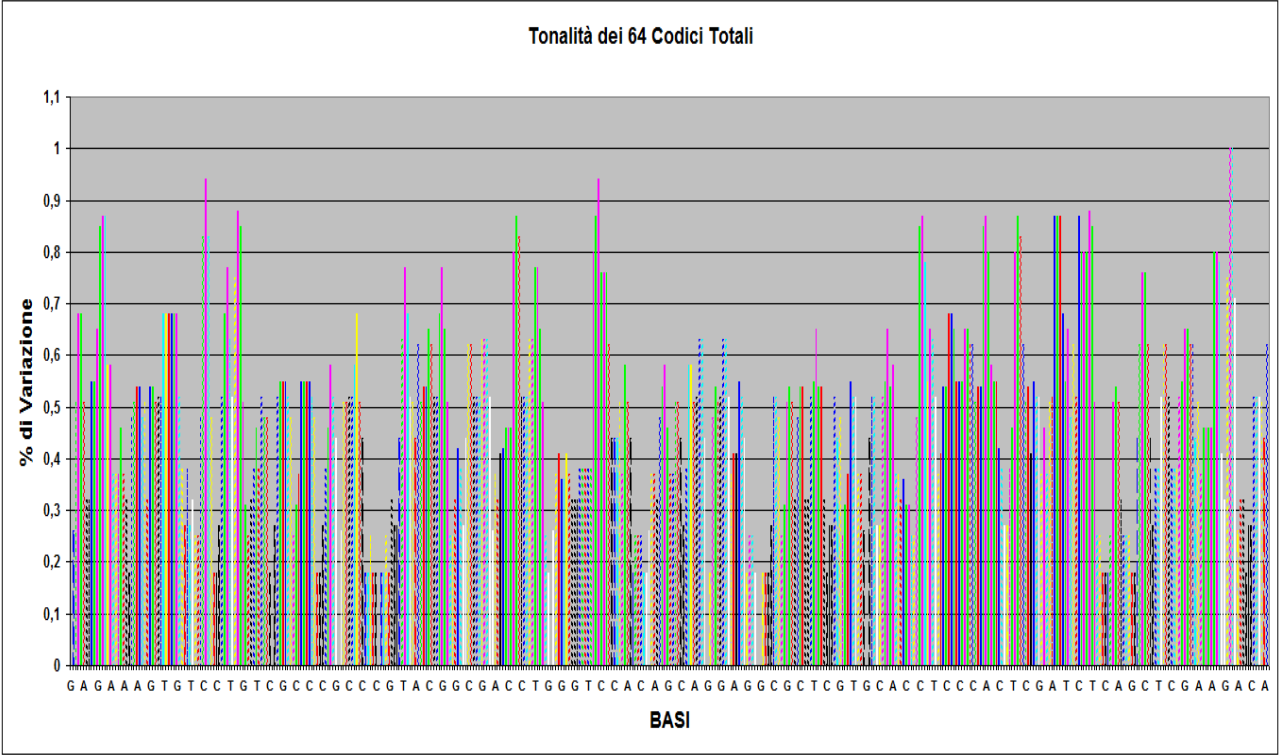


Fig. 15

Il **Profilo dei 64 Codici Totali** mostrato in Fig. 16 costituisce una sintesi molto specifica dell'intera sequenza.

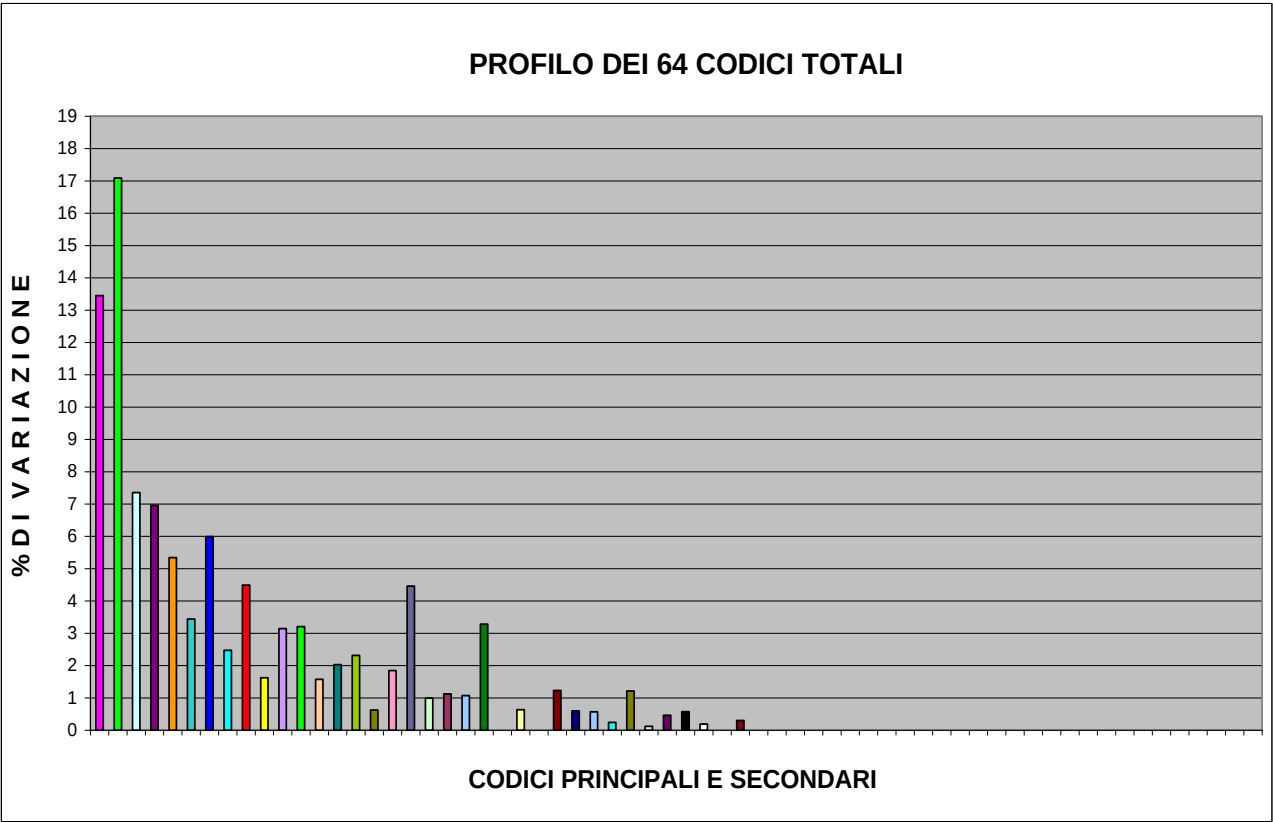


Fig. 16

Confronto e Commento dei Risultati delle Analisi Complete delle Sequenze

XM_011721319.1

XM_011721316.1

XM_011721317.1

XM_011721318.1

2.9 CONFRONTO DELLE BASI DELLE SEQUENZE

XM_011721319.1, XM_011721316.1, XM_011721317.1 e XM_011721318.1

Nella pagina seguente sono confrontate le **basi** delle *quattro sequenze di mRNA* oggetto di studio di questo capitolo.

Le basi della sequenza **XM_011721319.1** sono **presenti in tutte** le altre sequenze e sono riportate in **NERO** su fondo BIANCO.

Diversamente, sono evidenziate a **COLORI** le basi progressivamente aggiunte alla sequenza **XM_011721319.1** e a quelle che ad essa seguono.

Sono evidenziate in **GRIGIO** e in **MAGENTA** le basi aggiunte alla sequenza **XM_011721319.1** e presenti nella sequenza **XM_011721316.1**.

Sono evidenziate in **ROSSO** le basi aggiunte alla sequenza **XM_011721316.1** e presenti nella sequenza **XM_011721317.1**.

Infine, sono evidenziate in **GIALLO** le basi aggiunte alla sequenza **XM_011721317.1** e **presenti solo** nella sequenza **XM_011721318.1**.

XM_011721319.1 PREDICTED: Macaca nemestrina insulin (INS), transcript variant X4, mRNA

Product="**insulin isoform X2**" Length=**297**

ATGGCCCTGTGGATGCGCCTCTTGCCCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCCGGCCCCGG
CCTTTGTGAACCAGCACCTGTGCGGCTCCACCTGGTGGAAAGCTCTCTACCTGGTGTGCGGGGAGCGAGG
CTTCTTCTACACACCCAAGACCCGCCGGGAGGCAGAGGACCCTCAGGGCAGCCTGCAGCCCTTGGCGCTG
GAGGGGTCCCTGCAGAAGCGCGGCATCGTGGAGCAGTGCTGCACCAGCATCTGCTCCCTCTACCAGCTGG
AGAACTACTGCAACTAG

XM_011721316.1 PREDICTED: Macaca nemestrina insulin (INS), transcript variant X1, mRNA

Product="**insulin isoform X1**" Length=**454**

GGGACAGGCTGCATCAGAAGAGGCCAGCAAGCAGGT**CACTGTCTCTCCGCC**ATGGCCCTGTGGATGCGCC
TCTTGCCCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCCGGCCCCGGCCTTTGTGAACCAGCACCT
GTGCGGCTCCACCTGGTGGAAAGCTCTCTACCTGGTGTGCGGGGAGCGAGGCTTCTTCTACACACCCAAG
ACCCGCCGGGAGGCAGAGGACCCTCAGGTGGGGCAGGTGGAGCTGGGCGGGGGCCCTGGCGCAGGCAGCC
TGCAGCCCTTGGCGCTGGAGGGGTCCCTGCAGAAGCGCGGCATCGTGGAGCAGTGCTGCACCAGCATCTG
CTCCCTCTACCAGCTGGAGAACTACTGCAACTAGATGCGGCCCGCAGGCGGCCACACCCTCCACCTCCT
GCACCAAGAGAGATCGAATAAAGCCCTTGAACCA

XM_011721317.1 PREDICTED: Macaca nemestrina insulin (INS), transcript variant X2, mRNA

Product="**insulin isoform X1**" Length=**482**

GGGACAGGCTGCATCAGAAGAGGCCAGCAAGCAGGT**CTGTTCCAAGGGCCTTCACGTCAGGT**
CACTGTCTCTCCGCCATGGCCCTGTGGATGCGCCTCTTGCCCCCTGCTGGCGCTGCTGGCCCTCTGGGGAC
CTGACCCGGCCCCGGCCTTTGTGAACCAGCACCTGTGCGGCTCCACCTGGTGGAAAGCTCTCTACCTGGT
GTGCGGGGAGCGAGGCTTCTTCTACACACCCAAGACCCGCCGGGAGGCAGAGGACCCTCAGGTGGGGCAG
GTGGAGCTGGGCGGGGGCCCTGGCGCAGGCAGCCTGCAGCCCTTGGCGCTGGAGGGGTCCCTGCAGAAGC
GCGGCATCGTGGAGCAGTGCTGCACCAGCATCTGCTCCCTCTACCAGCTGGAGAACTACTGCAACTAGAT
GCGGCCCGCAGGCGGCCACACCCTCCACCTCCTGCACCAAGAGAGATCGAATAAAGCCCTTGAACCA**GC**

XM_011721318.1 PREDICTED: Macaca nemestrina insulin (INS), transcript variant X3, mRNA

Product="**insulin isoform X1**" Length=**532**

GGGACAGGCTGCATCAGAAGAGGCCAGCAAGCAGGT**CTGTTCCAAGGGCCTTCACGTCAGGT**
GGGCTCAGGGCTGCCCCACTTGGGGGTTCAGGGTGGCTGGACCCAGGT**CACTGTCTCTCCGCC**
ATGGCCCTGTGGATGCGCCTCTTGCCCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCCGGCCCCGG
CCTTTGTGAACCAGCACCTGTGCGGCTCCACCTGGTGGAAAGCTCTCTACCTGGTGTGCGGGGAGCGAGG
CTTCTTCTACACACCCAAGACCCGCCGGGAGGCAGAGGACCCTCAGGTGGGGCAGGTGGAGCTGGGCGGG
GGCCCTGGCGCAGGCAGCCTGCAGCCCTTGGCGCTGGAGGGGTCCCTGCAGAAGCGCGGCATCGTGGAGC
AGTGCTGCACCAGCATCTGCTCCCTCTACCAGCTGGAGAACTACTGCAACTAGATGCGGCCCGCAGGCGG
CCCACACCCTCCACCTCCTGCACCAAGAGAGATCGAATAAAGCCCTTGAACCA**GC**

2.10 CONFRONTO E COMMENTO DEI RISULTATI

GRAFICI DELLE ANALISI DELLE SEQUENZE

XM_011721319.1, XM_011721316.1, XM_011721317.1 e XM_011721318.1

Nella pagina seguente (si vedano le Figure 17, 18, 19 e 20) sono confrontati i **Profili degli 8 Codici Principali** delle quattro sequenze oggetto di analisi in questo capitolo. Come è stato già detto nel paragrafo 2.2, il **Profilo degli 8 Codici Principali** *SINTETIZZA IN MANIERA NON SPECIFICA* l'apporto congiunto delle *singole basi* (in minor misura) e del *“Trend”* (in maggior misura) nel *caratterizzare* la sequenza. Questo grafico è molto utile per confrontare le *caratteristiche generali* delle quattro sequenze analizzate (soprattutto, riguardo al ruolo svolto dal *“Trend”*).

A pagina 42, in un **unico grafico** (si veda la Fig. 21), sono confrontati i quattro **Profili degli 8 Codici Principali** (quelli relativi alle Figure 17, 18, 19 e 20).

In Figura 21, per necessità di sintesi, la **Sequenza XM_011721319.1** è stata denominata **9**; la **Sequenza XM_011721316.1** è stata denominata **6**; la **Sequenza XM_011721317.1** è stata denominata **7**; e, infine, la **Sequenza XM_011721318.1** è stata denominata **8**.

Come è possibile constatare osservando la Fig. 21, *le sequenze XM_011721316.1 e XM_011721317.1 sono molto simili rispetto a quanto lo siano le altre sequenze tra loro confrontate*.

Nonostante questa maggiore somiglianza delle *sequenze XM_011721316.1 e XM_011721317.1*, le differenze con le altre sequenze sono minime, come del resto ci si deve aspettare da un grafico che *sintetizza in maniera non specifica le caratteristiche delle sequenze*.

In conclusione, i confronti dei **Profili degli 8 Codici Principali** delle quattro sequenze analizzate, evidenziano che le *sequenze XM_011721316.1* (Fig. 18) e *XM_011721317.1* (Fig. 19) sono quelle che hanno *caratteristiche generali* (soprattutto, in riferimento al ruolo svolto dal *“Trend”*) **molto più simili** rispetto a quanto le abbiano tutte le altre sequenze tra loro confrontate. Viceversa, i confronti dei **Profili degli 8 Codici Principali** delle quattro sequenze analizzate, evidenziano che le *sequenze XM_011721319.1* (Fig. 17) e *XM_011721318.1* (Fig. 20), sono quelle che hanno *caratteristiche generali* (soprattutto, in riferimento al ruolo svolto dal *“Trend”*) **molto meno simili** rispetto a quanto le abbiano tutte le altre sequenze tra loro confrontate.

A pagina 43 (si vedano le Figure 22, 23, 24 e 25) sono confrontati i **Profili della Distribuzione della Percentuale di Variazione degli 8 Codici Principali** delle quattro sequenze analizzate.

Questo grafico è molto utile per confrontare le *caratteristiche dei “Trend”* delle quattro sequenze analizzate.

Come è possibile constatare osservando le Figure 22, 23, 24 e 25, *le sequenze XM_011721316.1* (Fig. 23) e *XM_011721317.1* (Fig. 24) sono molto simili rispetto a quanto lo siano le altre sequenze tra loro confrontate.

Mentre questo risultato era abbastanza prevedibile, sorprendono invece le differenze riscontrate tra la sequenza ***XM_011721318.1*** (Fig. 25) e rispettivamente le sequenze ***XM_011721316.1*** (Fig. 23) e ***XM_011721317.1*** (Fig. 24).

Il motivo di queste differenze non può che essere imputato al “peso” assunto (nell’influencare il “Trend”) dalle basi evidenziate in **GIALLO** a pagina 36 e **presenti solo** nella sequenza ***XM_011721318.1***.

Al contrario, la stessa importanza non sembrano assumere, nell’influencare il “Trend” e nel generare differenze tra le sequenze, le basi evidenziate in **ROSSO** a pagina 36, **presenti** nella sequenza ***XM_011721317.1*** e **assenti** nella sequenza ***XM_011721316.1***.

In conclusione, i confronti dei **Profili della Distribuzione della Percentuale di Variazione degli 8 Codici Principali** delle quattro sequenze analizzate, evidenziano che le sequenze ***XM_011721316.1*** (Fig. 23) e ***XM_011721317.1*** (Fig.24) sono quelle che hanno *caratteristiche dei “Trend” molto più simili* rispetto a quanto le abbiano tutte le altre sequenze tra loro confrontate. Viceversa, i confronti dei **Profili della Distribuzione della Percentuale di Variazione degli 8 Codici Principali** delle quattro sequenze analizzate, evidenziano che le sequenze ***XM_011721319.1*** (Fig.22) e ***XM_011721318.1*** (Fig.25) sono quelle che hanno *caratteristiche dei “Trend” molto meno simili* rispetto a quanto le abbiano tutte le altre sequenze tra loro confrontate.

A pagina 44 (si vedano le Figure 26, 27, 28 e 29) sono confrontati i **Profili delle Tonalità dei 64 Codici Totali** delle quattro sequenze analizzate.

Questo grafico è molto utile per confrontare le “qualità” (Tonalità e % di Variazione) delle *singole basi* delle quattro sequenze analizzate.

La sequenza ***XM_011721318.1***, essendo quella più lunga tra le quattro, contiene in sé **tutte le basi** presenti nelle altre tre sequenze.

Per questo motivo il grafico in Fig. 29, rispetto ai **Profili delle Tonalità dei 64 Codici Totali**, rappresenta il grafico di riferimento per il confronto con tutti gli altri.

Come è possibile constatare (anche se con molta più difficoltà di discriminazione visiva rispetto agli altri tipi di profili grafici) e, come prevedibile, osservando le Figure 26, 27, 28 e 29, le *maggiori differenze* si riscontrano tra la sequenza ***XM_011721319.1*** (Fig. 26) e la sequenza ***XM_011721318.1*** (Fig. 29). Questo risultato è comprensibile, essendo la sequenza ***XM_011721319.1*** quella con il minor numero di basi. Al contrario, la differenza tra le sequenze ***XM_011721316.1*** (Fig.27) e ***XM_011721317.1*** (Fig.28) è molto minore.

In conclusione, essendo la differenza del numero di basi minima tra le sequenze ***XM_011721316.1*** (Fig.27) e ***XM_011721317.1*** (Fig.28), i confronti di tutti i **Profili delle Tonalità dei 64 Codici Totali**, evidenziano che le sequenze ***XM_011721316.1*** e ***XM_011721317.1*** sono quelle che hanno “qualità” (Tonalità e % di Variazione) **molto più simili** rispetto a quanto le abbiano tutte le altre sequenze tra loro confrontate.

Viceversa, essendo la differenza del numero di basi massima tra le *sequenze XM_011721319.1* (Fig.26) e *XM_011721318.1* (Fig.29), i confronti di tutti i **Profili delle Tonalità dei 64 Codici Totali**, evidenziano che le *sequenze XM_011721319.1* e *XM_011721318.1* sono quelle che hanno “*qualità*” (Tonalità e % di Variazione) **molto meno simili** rispetto a quanto le abbiano tutte le altre sequenze tra loro confrontate.

A pagina 45 (si vedano le Figure 30, 31, 32 e 33) sono confrontati i **Profili dei 64 Codici Totali**.

Questo grafico *SINTETIZZA IN MANIERA SPECIFICA* e nel modo migliore l’apporto congiunto delle *singole basi* (in minor misura) e del “*Trend*” (in maggior misura) nel *caratterizzare* la sequenza.

Questo grafico è molto utile per confrontare le *caratteristiche indentificative* delle quattro sequenze analizzate.

A pagina 46, in un **unico grafico** (si veda la Fig. 34), sono confrontati i quattro **Profili dei 64 Codici Totali** (quelli relativi alle Figure 30, 31, 32 e 33).

In Figura 34, per rendere discriminabili visivamente gli istogrammi del grafico di ogni sequenza, alla **Sequenza XM_011721319.1** (già denominata **9**) è stato assegnato il **NERO**; alla **Sequenza XM_011721316.1** (già denominata **6**) è stato assegnato il colore **GIALLO**; alla **Sequenza XM_011721317.1** (già denominata **7**) è stato assegnato il colore **ROSSO**; infine, alla **Sequenza XM_011721318.1** (già denominata **8**) è stato assegnato il **BIANCO**.

Con una certa sorpresa, come è possibile constatare osservando la Fig. 34, le differenze tra le quattro sequenze *sembrano* veramente **minime**.

Questo risultato non era molto prevedibile perché i confronti tra i **Profili della Distribuzione della Percentuale di Variazione degli 8 Codici Principali** (si vedano le Figure 22, 23, 24 e 25) e, soprattutto, quelli tra i **Profili delle Tonalità dei 64 Codici Totali** (si vedano le Figure 26, 27, 28 e 29) avevano evidenziato importanti differenze tra le sequenze. Per chiarire meglio il risultato dei confronti grafici dei quattro **Profili dei 64 Codici Totali** (quelli relativi alle Figure 30, 31, 32 e 33), sono stati elaborati *altri due grafici*, quelli mostrati nelle Figure 35 e 36.

In Fig. 35 sono confrontati i profili delle sequenze **XM_011721316.1** (Fig. 31) e **XM_011721317.1** (Fig. 32). Come è possibile constatare osservando la Fig. 35, le differenze tra queste due le sequenze sono **sostanzialmente insignificanti**. Questo dato conferma i risultati già evidenziati dagli altri tipi grafici, ovvero che *le sequenze XM_011721316.1 e XM_011721317.1 sono molto più simili* (anche riguardo i 64 Codici Totali) *rispetto a quanto lo siano le altre sequenze tra loro confrontate*.

In Fig. 36 sono invece confrontati i profili delle sequenze **XM_011721319.1** (Fig. 30) e **XM_011721318.1** (Fig. 33). Come è possibile constatare osservando la Fig. 36, le differenze tra queste due sequenze sono **più evidenti**, anche se contenute rispetto a quanto ci saremmo aspettati. Anche questo dato conferma in parte i risultati già evidenziati dagli altri tipi grafici.

In conclusione, i confronti dei **Profili dei 64 Codici Totali** delle quattro sequenze analizzate, evidenziano che:

- 1) le sequenze ***XM_011721316.1*** e ***XM_011721317.1*** sono quelle che hanno *caratteristiche identificative molto più simili* rispetto a quanto le abbiano tutte le altre sequenze tra loro confrontate;
- 2) le sequenze ***XM_011721319.1*** e ***XM_011721318.1*** sono quelle che hanno *caratteristiche identificative molto meno simili* rispetto a quanto le abbiano tutte le altre sequenze tra loro confrontate;
- 3) il **Profilo dei 64 Codici Totali** sembra essere capace, più di quanto previsto, di **SINTETIZZARE** e insieme cogliere molto bene le **CARATTERISTICHE SPECIFICHE SIMILI** (in questo caso il fatto che i *prodotti* sono sempre delle “*insulin isoform*”) di sequenze che possono avere **quantità** e “**qualità**” di basi anche molto diverse tra loro. Questa conclusione riguardante il **Profilo dei 64 Codici Totali** è comunque stata verificata anche dalle analisi dei risultati grafici ottenuti nel Capitolo Primo (Parte Prima e Seconda), anche se non è stata particolarmente enfatizzata. Infatti, le 19 nuove sequenze prodotte a partire dalla Catena A dell’Insulina (sequenza originaria) hanno più del 73 % di basi diverse (per tipologia e per posizione assunta nella sequenza) dalla sequenza originaria, ma i **Profili dei 64 Codici Totali** delle *nuove sequenze prodotte* **non differiscono molto** da quelli della Catena A dell’Insulina.

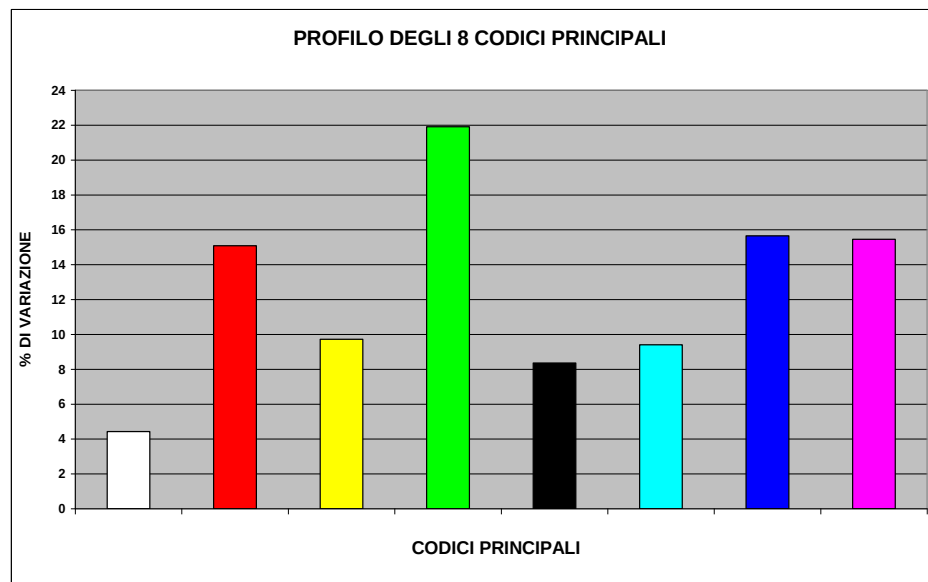


Fig. 17 (Sequenza XM_011721319.1)

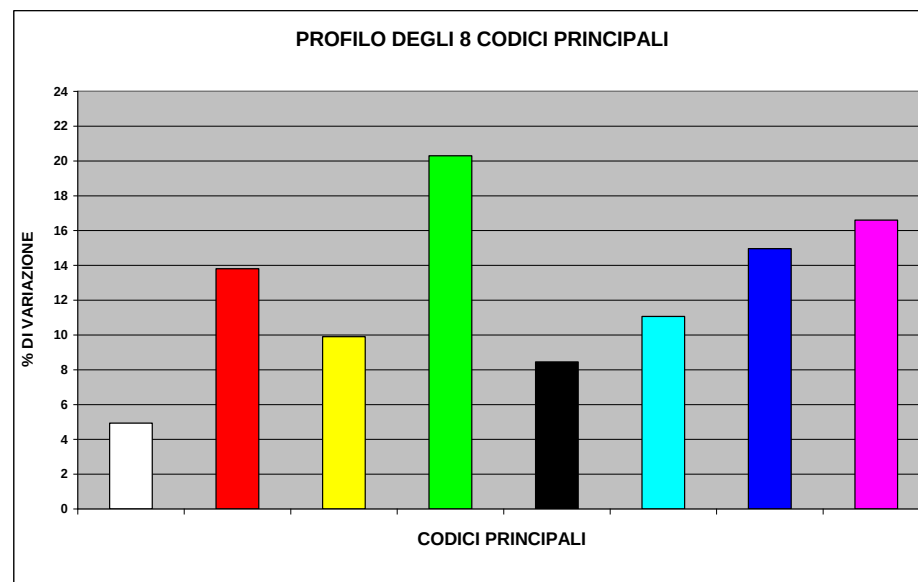


Fig. 18 (Sequenza XM_011721316.1)

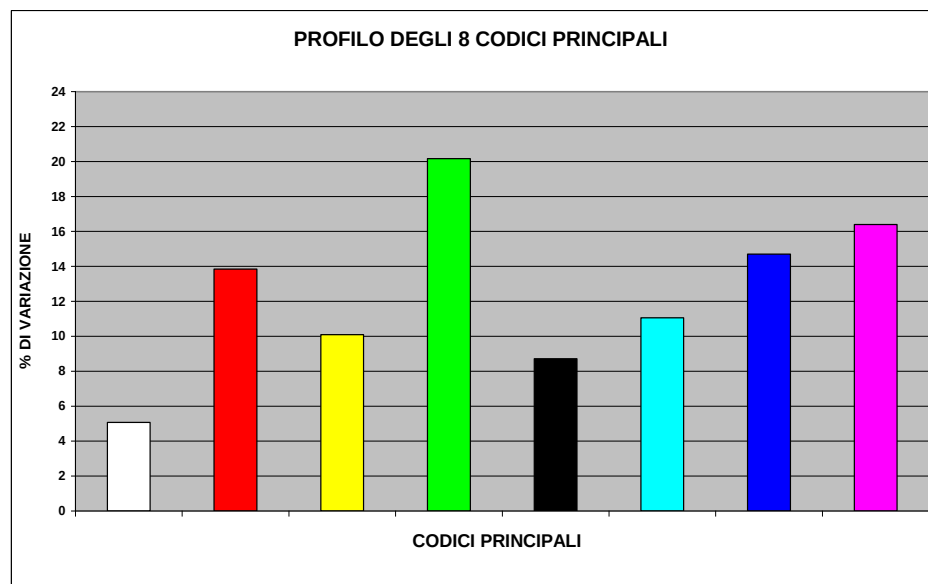


Fig. 19 (Sequenza XM_011721317.1)

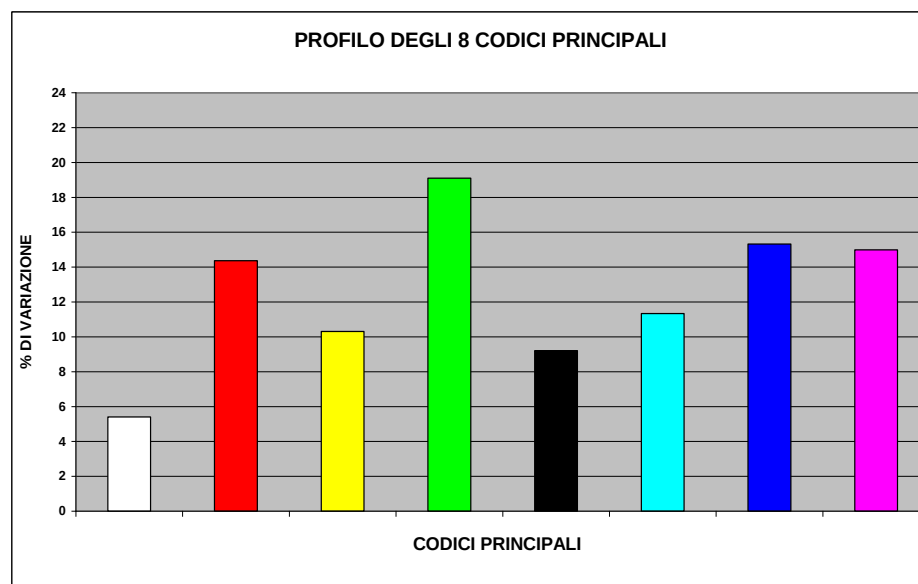


Fig. 20 (Sequenza XM_011721318.1)

PROFILO DEGLI 8 CODICI PRINCIPALI

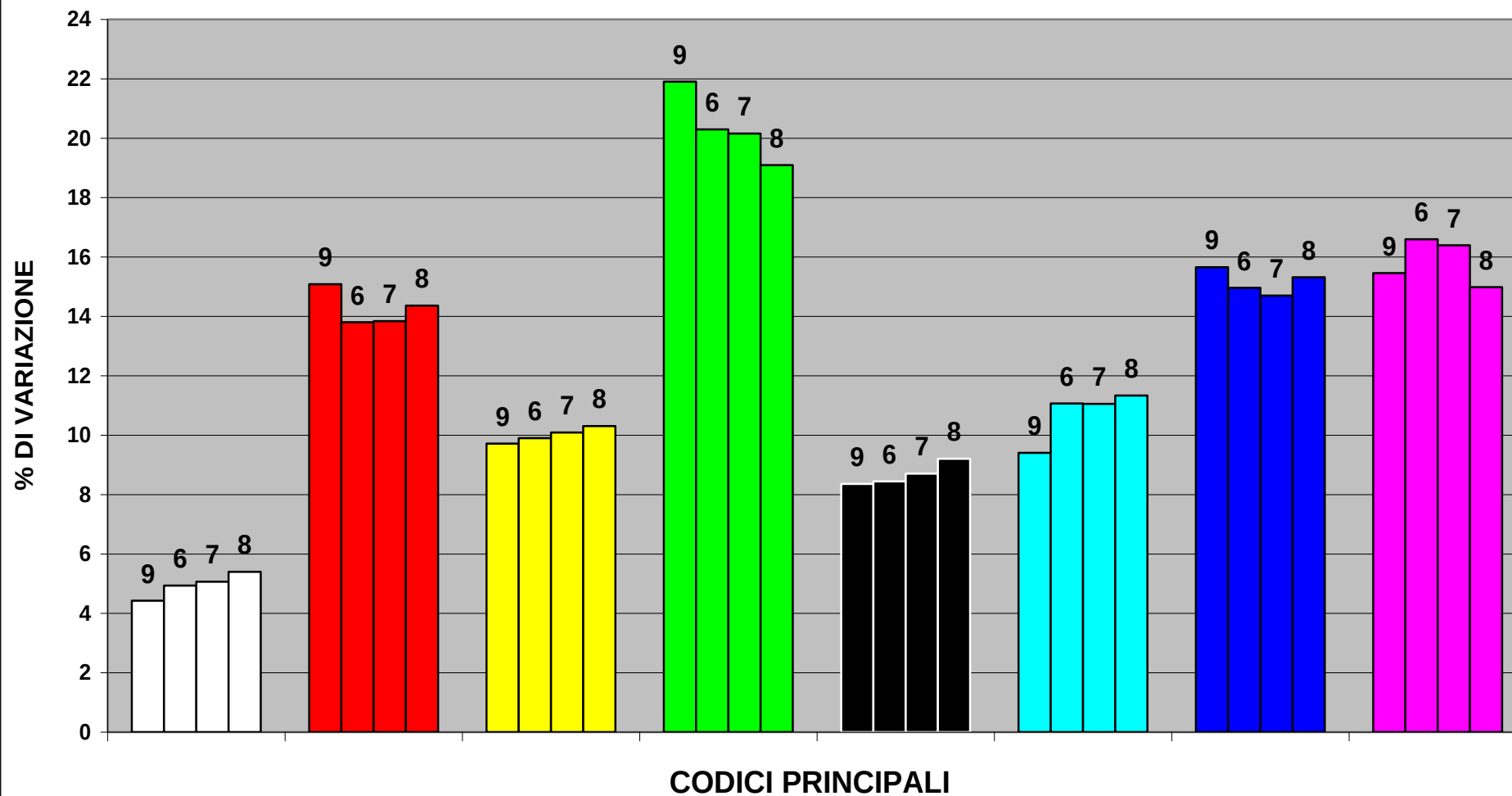


Fig. 21

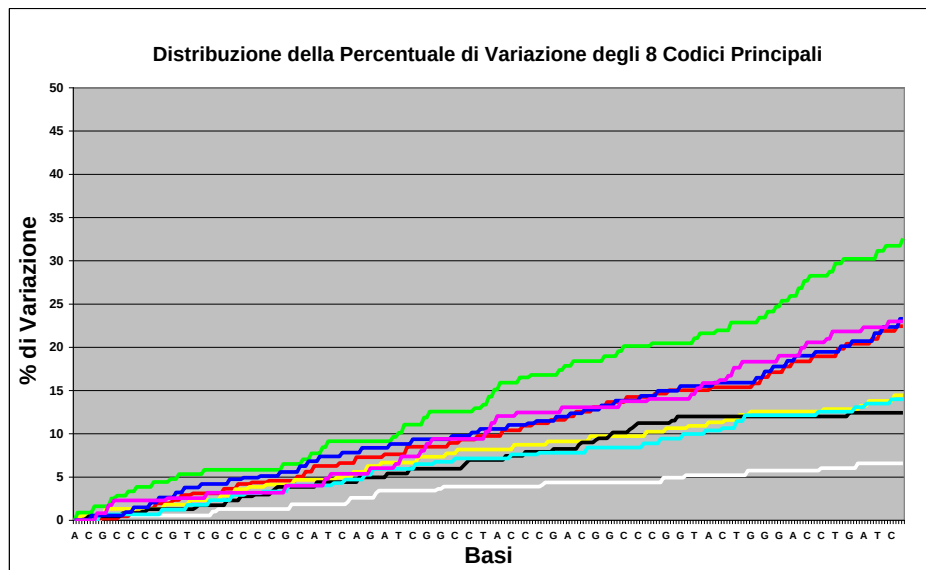


Fig. 22 (Sequenza XM_011721319.1)

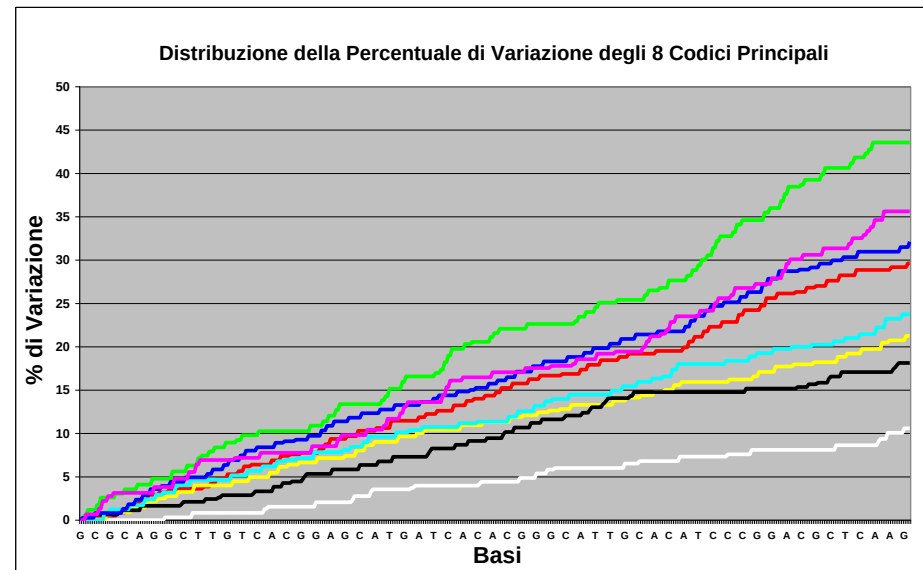


Fig. 23 (Sequenza XM_011721316.1)

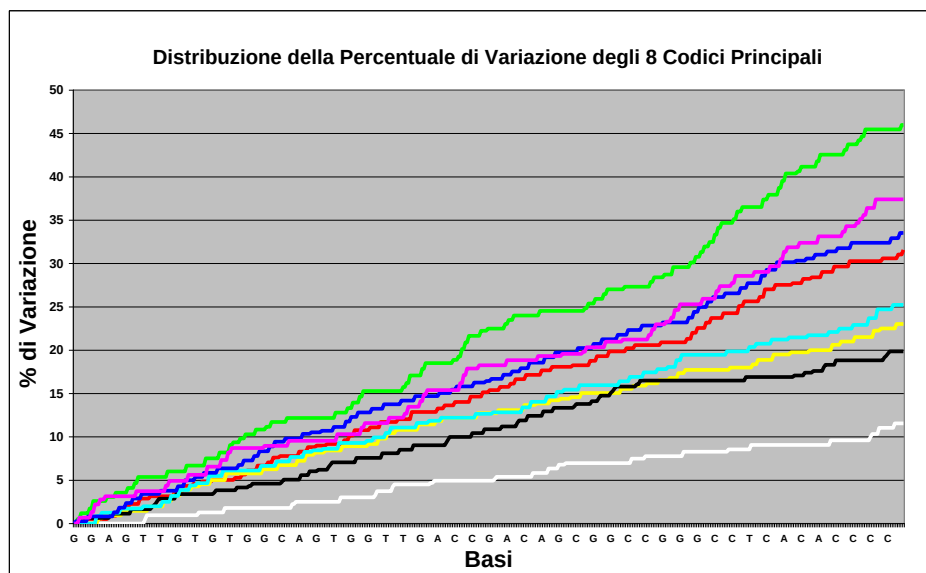


Fig. 24 (Sequenza XM_011721317.1)

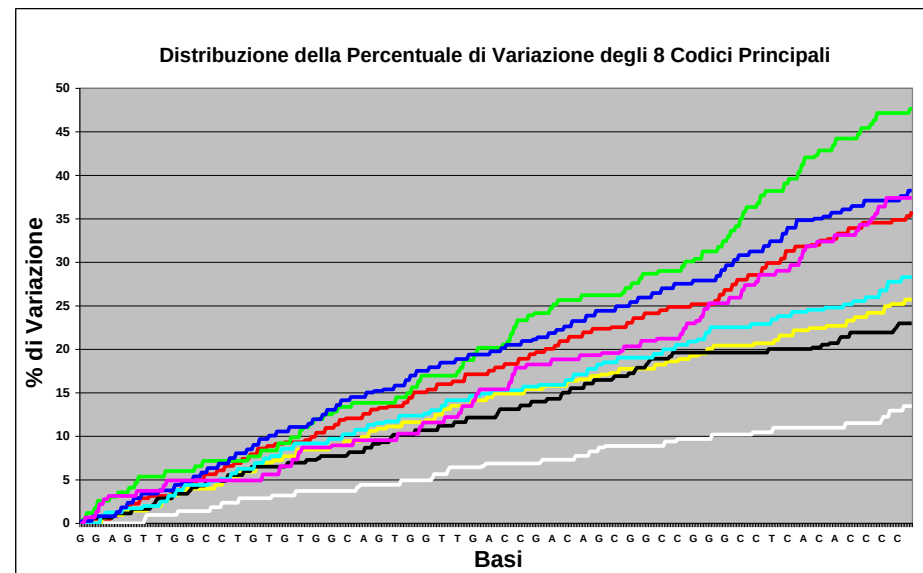


Fig. 25 (Sequenza XM_011721318.1)

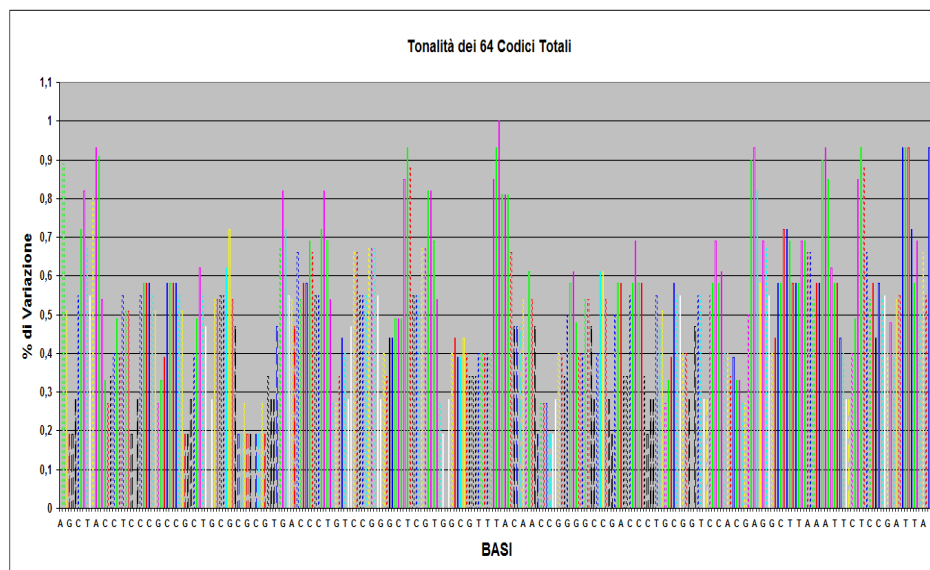


Fig. 26 (Sequenza XM_011721319.1)

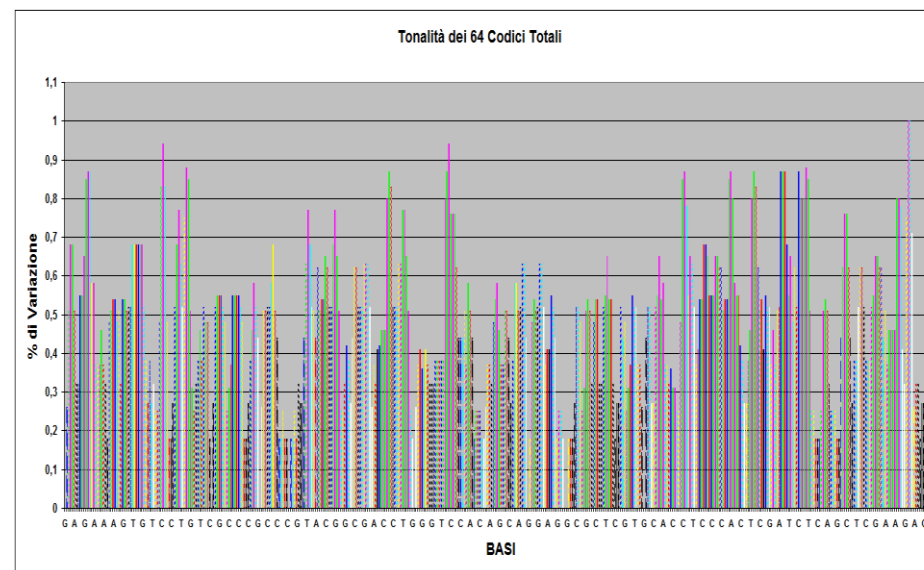


Fig. 27 (Sequenza XM_011721316.1)

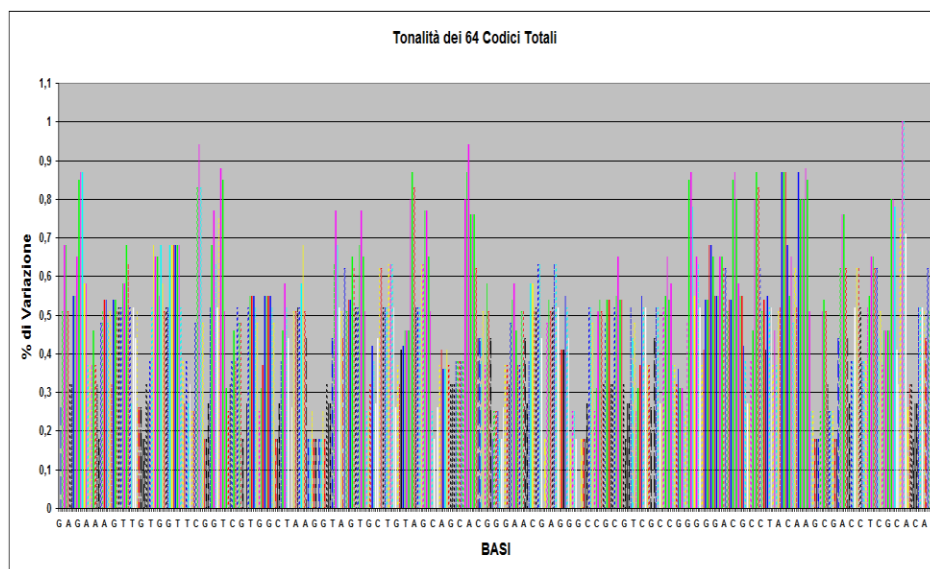


Fig. 28 (Sequenza XM_011721317.1)

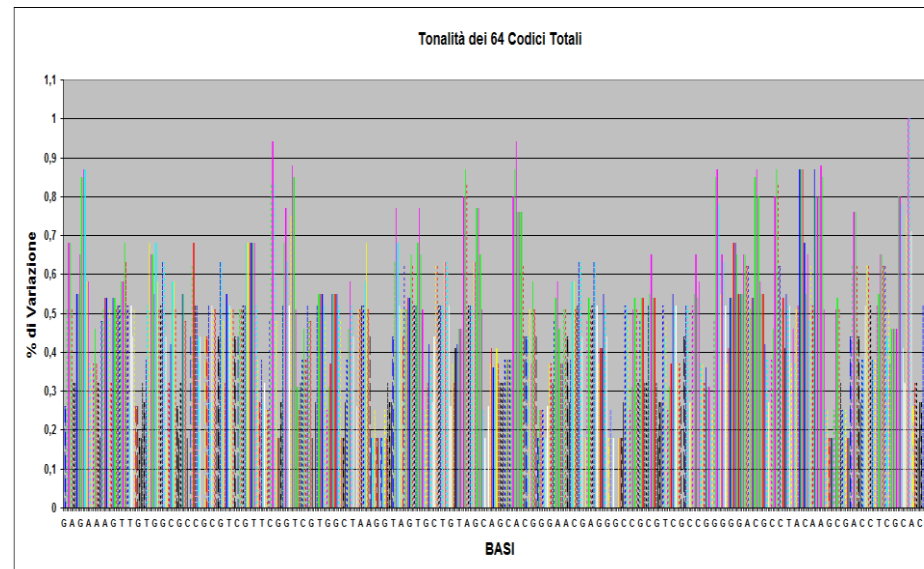


Fig. 29 (Sequenza XM_011721318.1)

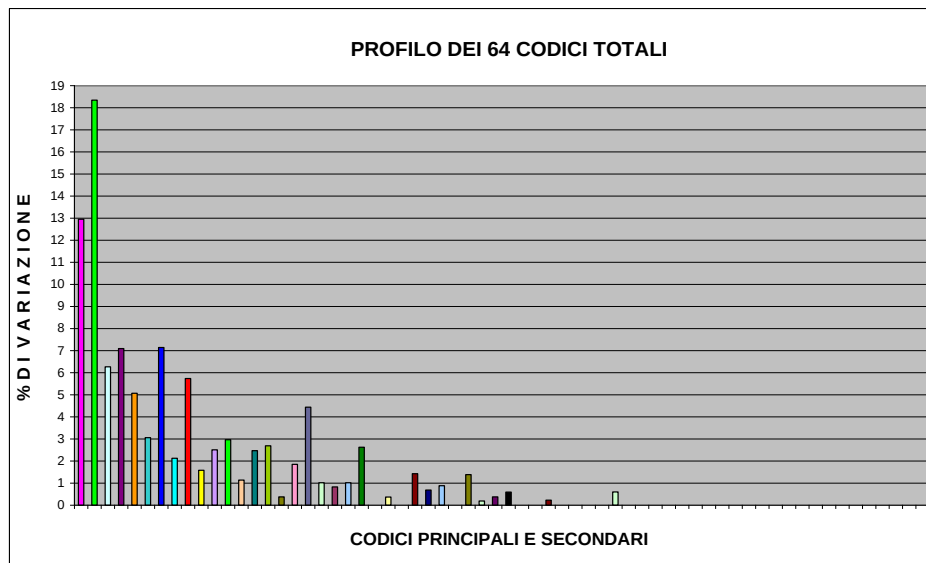


Fig. 30 (Sequenza XM_011721319.1)

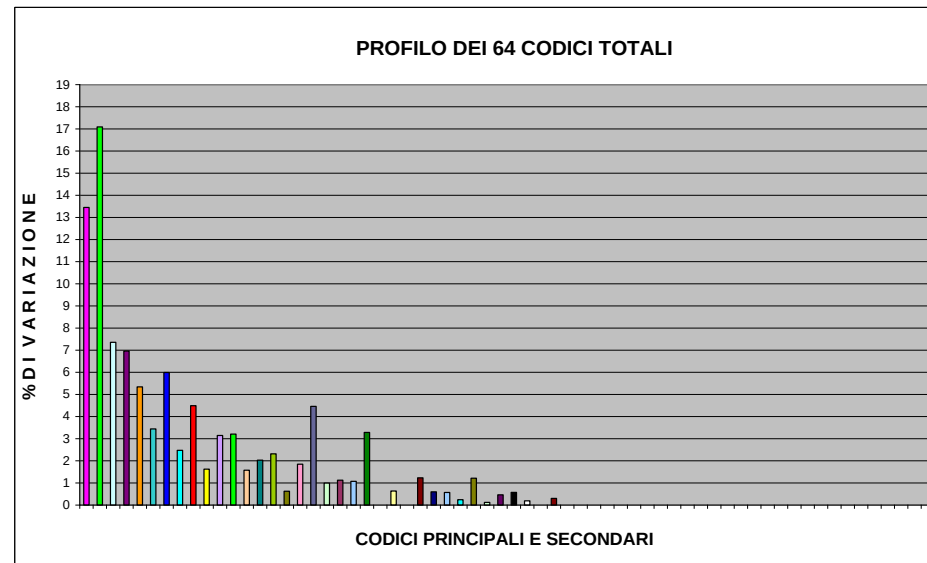


Fig. 31 (Sequenza XM_011721316.1)

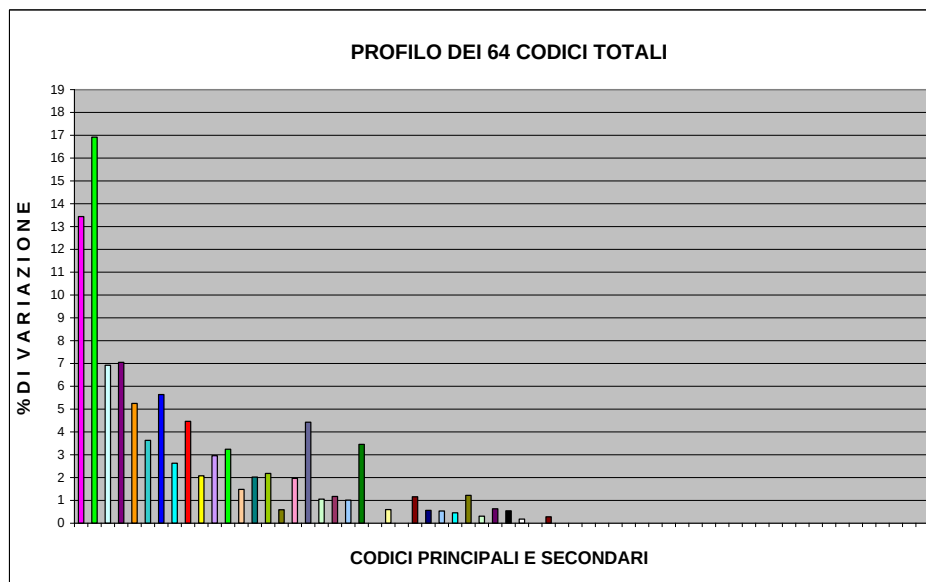


Fig. 32 (Sequenza XM_011721317.1)

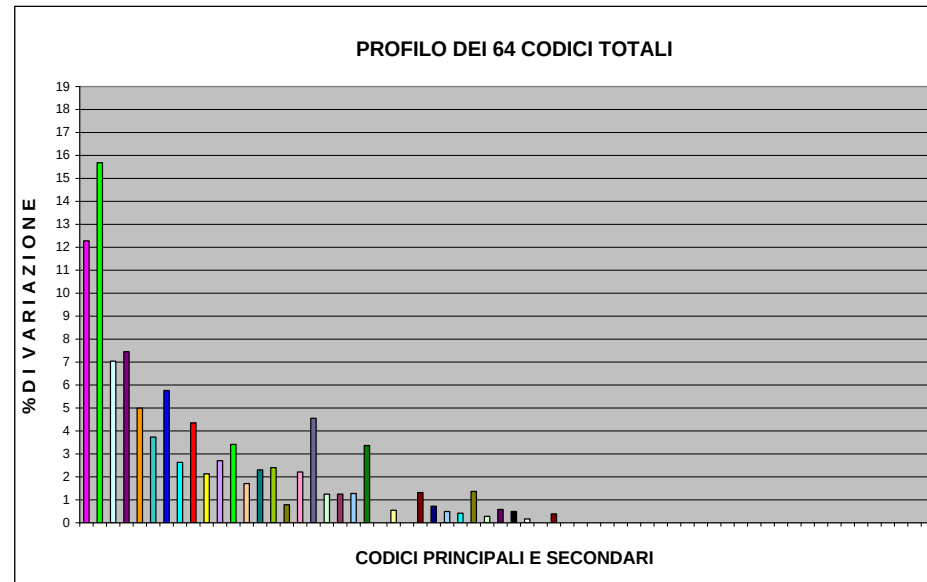


Fig. 33 (Sequenza XM_011721318.1)

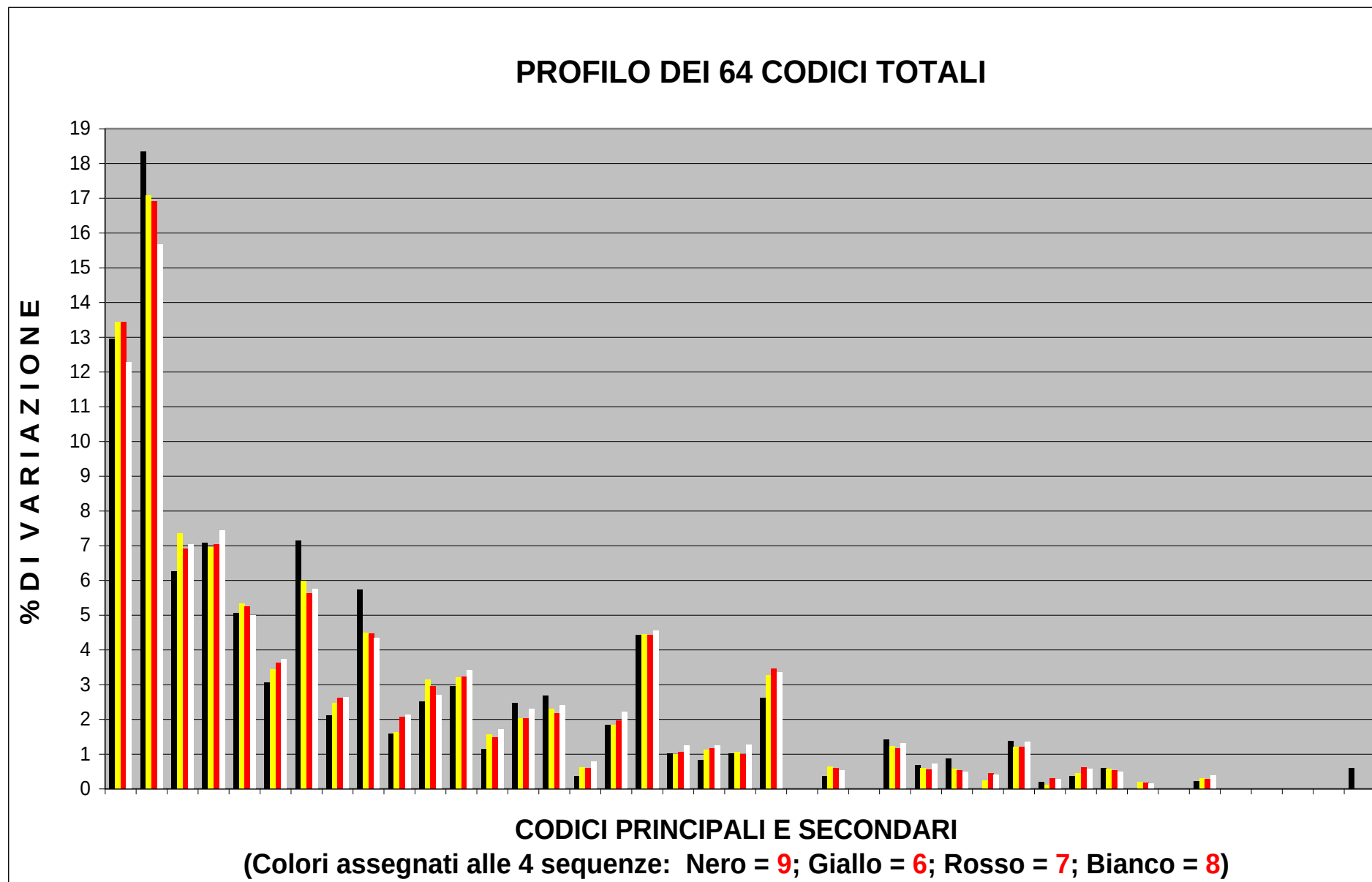


Fig. 34

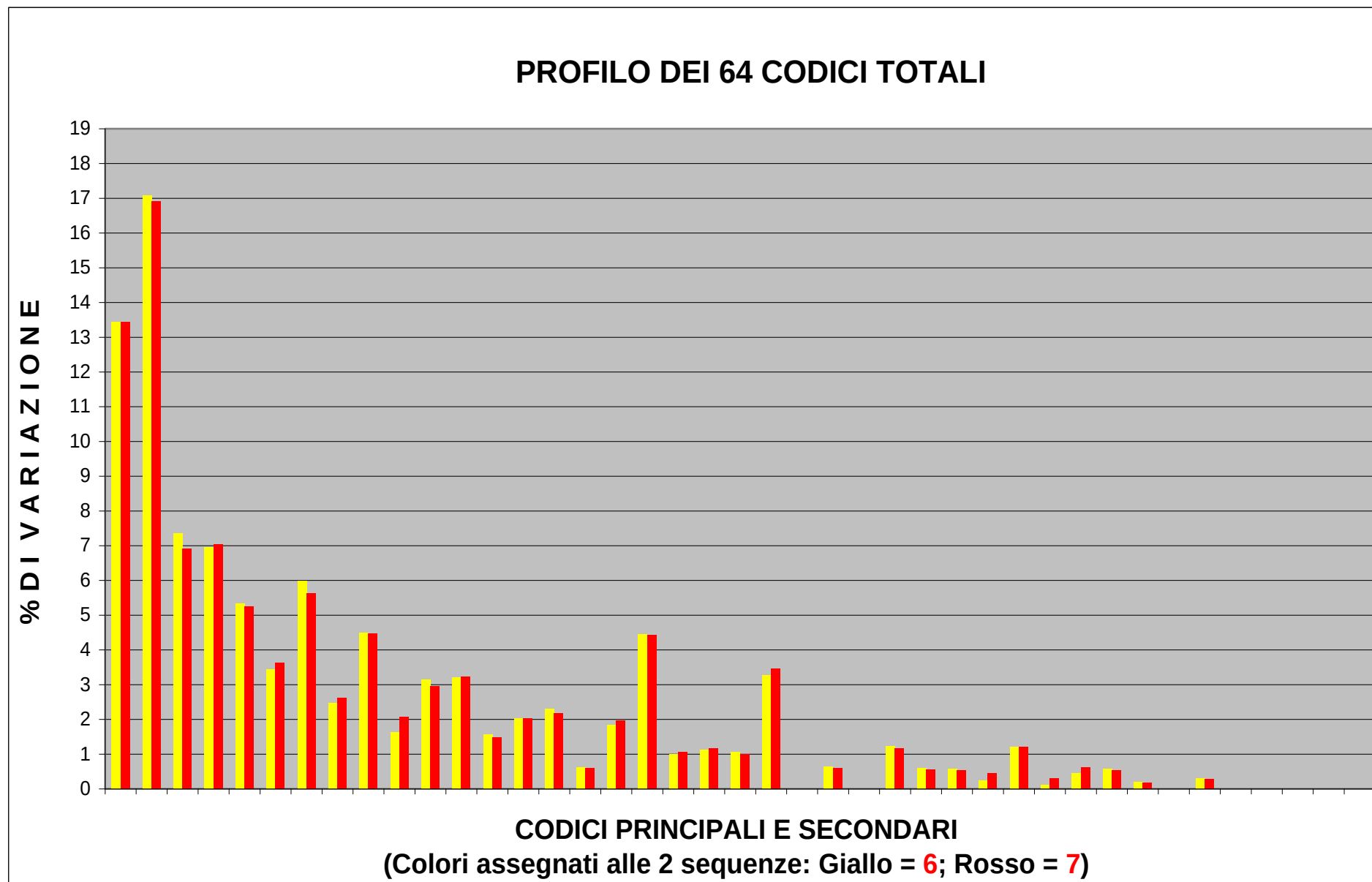


Fig. 35

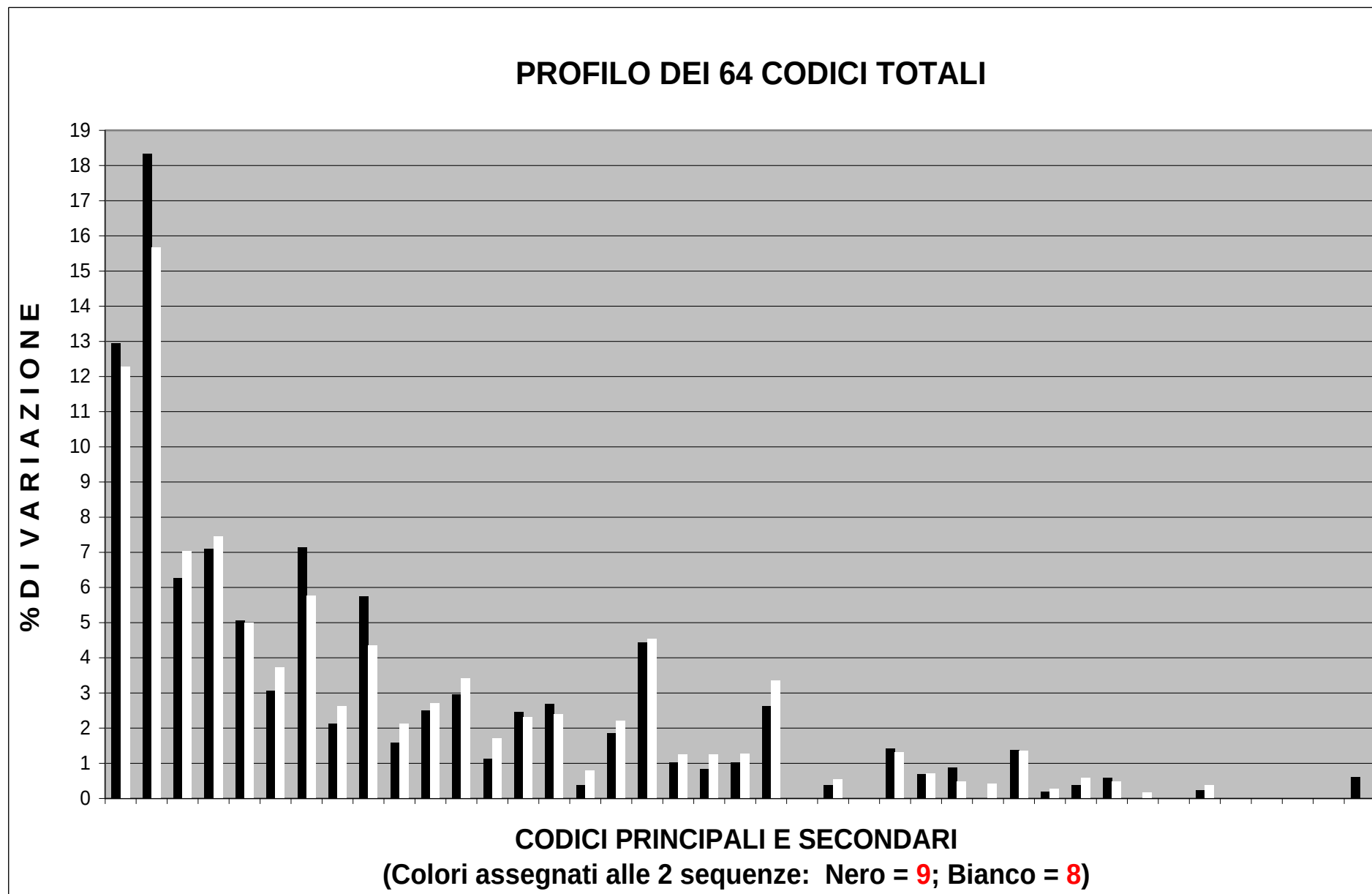


Fig. 36

Discussione Generale dei Risultati e Conclusioni

2.11 DISCUSSIONE GENERALE DEI RISULTATI E CONCLUSIONI

In questo Capitolo sono state esplorate le **potenzialità dei grafici** elaborati con i dati ottenuti dal software della **T.T.E.S.** e con il programma Excel.

L'obiettivo era quello di evidenziare la **validità**, la **attendibilità** e la **sensibilità** dei calcoli effettuati per la generazione e la rappresentazione grafica dei dati.

I risultati grafici di questo Capitolo, insieme a quelli del Capitolo precedente (Parte Prima e Seconda), sembrano confermare la **validità** dei calcoli effettuati.

Le relazioni tra le **sequenze di mRNA** (oggetto di studio in questo Capitolo), i calcoli effettuati per la generazione di tutti i grafici e le loro rappresentazioni si dimostrano *effettive e non dovute al caso*.

Anche la **attendibilità**, intesa come **coerenza interna**, sembra confermata; infatti, le relazioni tra le **sequenze di mRNA**, i calcoli effettuati per la generazione di tutti i grafici e le loro rappresentazioni restano **coerenti** nonostante la *lunghezza delle sequenze*, i *diversi tipi di "trend"* e la *"qualità" delle singole basi azotate*.

Infine, la **sensibilità** dei grafici, intesa come la loro capacità di discriminare le 4 **sequenze di mRNA** e di evidenziare gli aspetti che differiscono tra di esse, ovvero la diversa *lunghezza* delle sequenze analizzate, i loro *diversi tipi di "trend"*, la differente *"qualità" delle singole basi azotate* è stata ampiamente provata.

Piu nello specifico, dal **Confronto delle Analisi Complete delle 4 Sequenze dell'organismo Macaca Nemestrina** considerate in questo capitolo, abbiamo ampiamente provato che:

- 1) la sequenza XM_01172131**9**.1 (Product = "insulin isoform **X2**") è rappresentata graficamente in maniera **leggermente diversa** dalle altre tre sequenze (Product = "insulin isoform X1");
- 2) le sequenze XM_01172131**6**.1, XM_01172131**7**.1 e XM_01172131**8**.1 sono rappresentate graficamente in maniera **abbastanza simile**, nonostante la *quantità* e la *"qualità" delle singole basi azotate siano in parte differenti* in ciascuna di queste sequenze.

Per essere ancora più precisi, abbiamo accertato, senza alcun dubbio, che le sequenze XM_01172131**6**.1 e XM_01172131**7**.1 sono quelle che, tra tutte, hanno *caratteristiche generali*, dei *"Trend"*, *"qualità"* delle basi (Tonalità e % di Variazione) e *caratteristiche identificative molto più simili* rispetto a quanto le abbiano tutte le altre sequenze tra loro confrontate.

Infine, è stato riscontrato che il **Profilo dei 64 Codici Totali** sembra essere capace, più di quanto previsto, di **SINTETIZZARE** e insieme cogliere molto bene le **CARATTERISTICHE SPECIFICHE SIMILI** (in questo caso il fatto che i *prodotti* sono sempre delle “*insulin isoform*”) di sequenze che possono avere **quantità** e “**qualità**” di basi anche molto diverse tra loro.

Alla luce di questi risultati, possiamo tranquillamente concludere che i **grafici** elaborati con i dati ottenuti dal software della **T.T.E.S.** e con il programma Excel sono più che adeguati a **descrivere**, in maniera innovativa, chiara ed esaustiva, una sequenza di Dna (o di Rna), di qualsiasi lunghezza essa sia.

Più precisamente, i grafici elaborati sono capaci di:

- 1) sintetizzare ed evidenziare le *caratteristiche generali* di una *sequenza*, soprattutto riguardo al ruolo svolto dal “Trend” (**Profilo degli 8 Codici Principali**);
- 2) evidenziare, in maniera molto chiara, le *caratteristiche del “Trend”* di un’intera sequenza (**Distribuzione della Percentuale di Variazione degli 8 Codici Principali**);
- 3) evidenziare le “*qualità*” (Tonalità e % di Variazione) delle *singole basi* di una sequenza (**Tonalità dei 64 Codici Totali**);
- 4) sintetizzare ed evidenziare fedelmente le *caratteristiche indentificative e specifiche* di una sequenza (**Profilo dei 64 Codici Totali**);
- 5) discriminare ed evidenziare accuratamente gli aspetti simili e quelli differenti di sequenze parzialmente o molto diverse tra loro.

In conclusione, i **grafici** elaborati con i dati ottenuti dal software della **T.T.E.S.** e con il programma Excel possono essere considerati una nuova ed importante risorsa, sicuramente ancora perfezionabile, a disposizione dei ricercatori di biologia molecolare, genetica, genomica, ecc.

FINE CAPITOLO II °



www.ttesystems.eu

Corrispondenza: *nunzio.bonaventura@libero.it*

12 Agosto 2019