STEP 1:

Generate Never Born Protein Sequences

B.K.N. Muthugobal[1], G. Ramesh[2], K. Muthuvel Prasath[3], S. Suvaithenamudhan[3] and S. Parthasarathy[3]*
[1]Department of Mathematics, Bharathidasan University Constitutent Arts & Science College, Nannilam, Thiruvarur 610 105, Tamil Nadu, India
[2]Ramanujan Research Centre, Department of Mathematics, Government Arts College (Autonomous), Kumbakonam 612 001, Tamil Nadu, India
[3]Department of Bioinformatics, School of Life Sciences, Bharathidasan University, Tiruchirappalli 620 024, Tamil Nadu, India
* For all correspondence bdupartha@gmail.com

HOME

The "GenNBPSeq" is an online web server to generate a Never Born Protein (NBP) sequence and its corresponding DNA sequence, using the Partitioned Gray code representations of Universal Genetic Code combined with a novel Toeplitz matrix approach, as outlined in Muthugobal et. al. (2018)[1]. We generate the three sample NBP sequences *NBP_SEQ_1, NBP_SEQ_2 and NBP_SEQ_3* using this "GenNBPSeq" online web server.

Choose Toeplitz Matrix Order

2x2 Matrix | 3x3 Matrix | 4x4 Matrix

Click Here

Click here for a brief Description of Generation of the three Sample NBP Sequences.

Click here to get Sample Input and Output *NBP_SEQ_1*, *NBP_SEQ_2* and *NBP_SEQ_3*.

Click here to get Partitioned Toeplitz Matrix types PTM_1, PTM_2, PTM_3 and PTM_4.

Click here for Universal Genetic Code.

Please Cite

[1] B.K.N. Muthugobal, G. Ramesh, K. Muthuvel Prasath, S. Suvaithenamudhan and S. Parthasarathy*(2018) Gray Code Representation of the Universal Genetic Code and Generation of Never Born Protein Sequences using Toeplitz Matrix Approach: Sequence and Structure Analysis. *To be communicated.*

Further Reference

He M, Petoukhov S and Ricci PE (2004) Genetic code, Hamming distance, and stochastic matrices. *Bulletin of Mathematical Biology*, **66**, 965-973.

Please, send your comments, suggestions to: bdubioserver@gmail.com

© 2018 - S. Parthasarathy, Bharathidasan University, India

STEP 2:

Generate Never Born Protein Sequences

INPUT

Choose PTM Matrix Type and Give 4 x 4 Toeplitz Matrix Values

PTM1 / PTM2 / PTM3 / PTM4	43 / 44 / 45 / 46	33 / 34 / 35 / 36	23 / 24 / 25 / 26	13 / 14 / 15 / 16
PTM1 / PTM2 / PTM3 / PTM4	13 / 14 / 15 / 16	43 / 44 / 45 / 46	33 / 34 / 35 / 36	23 / 24 / 25 / 26
PTM1 / PTM2 / PTM3 / PTM4	23 / 24 / 25 / 26	13 / 14 / 15 / 16	43 / 44 / 45 / 46	33 / 34 / 35 / 36
PTM1 / PTM2 / PTM3 / PTM4	33 / 34 / 35 / 36	23 / 24 / 25 / 26	13 / 14 / 15 / 16	43 / 44 / 45 / 46

Click Here to Generate NBP Sequence

Generate NBP Sequence | Reset

STEP 3 :



Generate Never Born Protein Sequences

RESULT

INPUT

```
Toeplitz Matrix Order : 4 x 4

Partitioned Toeplitz Matrix type for First Row  : M4

Partitioned Toeplitz Matrix type for Second Row : M3

Partitioned Toeplitz Matrix type for Third Row  : M2

Partitioned Toeplitz Matrix type for Fourth Row : M1

Toeplitz Matrix : 45  35  25  15

                  15  45  35  25

                  25  15  45  35

                  35  25  15  45
```

OUTPUT

```
Never Born Protein (NBP) and DNA Sequences

>NB_DNA_SEQ   3072 bp
ggggagaagaggaaaagaggagaagcggtgatgacgataacagcagtagccaccactgct
attgttgtcatccgccgtcatcactattactgctgtgtggcgacgatgacaatagtagca
gaggggaggaagagaaaagaaggatccccccctttctctttttttcctccgacggcagcaa
tagtaatgatggttgtcgccgctgccactattatcatagtggcggcagcgacaataatga
tcaccaccgtcgctgttgttactactactgccgccatcgtcattattgtggtagcagcgg
caacgatgataatcgttgctgccgctaccatcattagcaacaacggcgatggtggtaata
ctccttcctccctcttccttctttttttctcctcttcccctcttctcctattgtcgtcat
cgccactactgctaacaacagtagcggtggtgacgaataatgacgacagcggcagtagtg
tgttatcatcgtcaccgctgctactcttttcttcctctcccctccttcgaaaaaaaggag
agggggggaagaatcattactaccgctgccgtcgttggtgataatagtaacagcggcgac
gctgttattactatcaccgccgtcgacaacaatgatagtggtggcagcagcagtaataac
gatgacggcggtgttgctactattaccatcgtcgccgatggtagtaatagcaacgacggc
taccaccattatcgttgttgccgcagaaggaagaaagaggaaggagggcgccgtcatcac
tgctgttattacggggagaagaggggagaaaaaagagcggtgatgacggcagtaataaca
gccaccactgctgtcatcattgttcgacggcagcaatgatggtagtaagtggcgacgatg
gtagcaacaatagaggggaggaaggaaggaagaaaatccccccctttctttcctccttttt
ctactgccgccattattgtcgtcattgtcgccgctgttatcaccactatagtggcggcag
taatgacgacaatcaccaccgtcgttactactgttgctccttcctcccttctttttcttcc
tggtagcagcggtgataacaacgatcgttgctgccgtcattactaccagcaacaacggcg
gtaataatggtgagcagtaataacggcggtgatgacggtgataatagtggcgacaacagc
gctgttattactgccgtcatcaccgacaacaatgatggcagcagtggtagaaggaagaaa
ggaggggaggaagttgctactattgtcgccaccatcgatggtagtaatgacggcagcaac
taccaccattattgccgccgttgtataatgacgacagtagtggcggcattttctcctctt
ttctcccccctctattgtcgtcattactgccgccactaacaacagtagtgacgacggtgg
atcattactaccgtcgttgctgcctgttatcatcgttgctaccaccgctcttttcttcct
tccttcctccccgaaaaaaaggagggaagaaggggggccaccactgctatcattgttgtc
cgccgtcatcactgttattactgcggggagaagagggaaaaaagaggagcggtgatgacg
gtaataacagcatccccccctctctcctttttttccgacggcagcaatggtagtaatga
gtggcgacgatggcaacaatagtagaggggaggaagggaagaaaagaatcaccaccgtcg
ctactgttgttactactgccgccattgtcgtcattattgtcgccgctgtcaccactatta
tagtggcggcagtgacgacaataagcaacaacggcgataatggtggtactccttcctccc
ttttcttccttctggtagcagcggtaacaacgatgatcgttgctgccgttactaccatca
taccaccattatcgccgttgttgcagaaggaagaaaggggaggaaggagttgctactatt
gccaccatcgtcgatggtagtaatggcagcaacgactaacaacagtagcgacggtggtga
ataatgacgacagtggcggcagtattttctcctctttcccccctcttctattgtcgtcat
tgccgccactacgaaaaaaaggagagaaggggggggaatcattactaccgttgctgccgtc
```

tgttatcatcgttaccaccgctgctcttttcttcctttcctcccctccgacaacaatgat
agcagtggtggcagcagtaataacggtgatgacggcggtgataatagtgacaacagcggc
gctgttattactgtcatcaccgccccgtcactattgtcgccattactgccgctatcattg
ccattatcgctgccctctttcctttcctttctccctccgttgtcactaccactgtcgtta
actgccatcgttgctaccgtcattcatcgctactgtcactgctatcgtacagcggtaatg
gcagtgataacgaaggggaaagaaaaagggagggacattaccgctgttatcactgccgt
actatcgccgttaccgtcgctattaacgatggcagtgacggtagcaatactgttgccatc
accattgctgtcaaggaaagaggggagaaaggaaggaagagagaagggaaaggagagagg
caatagtgacggtaatggcgacagcattgttaccgccaccgttattgcacggcaatagtg
gcgacagtaatgcctctctcctttcccttctctcttaccgctgtcattgccgttatcact
ccttttccctccccctttctttcccttccctctttctccttcctaatagcgacggt
aacggcgatagtaaagagggaagggaagggagaaagcagtggtaacgacaacggtagtga
aatgacagcggtgataacggcagtacgatagcagtgacagtagcgatgcactattgccgt
tactgtcgccatacggtggcaataacaatggcggtacagtaacgatggtagcaatgacgg
cagcgataatggcaatgatagcggccatcgttactgtcattgctaccgaatggtgacagc
aacagtgatggc

>NBP_SEQ 976 aa
GEKRKRGEAVMTITAVATTAIVVIRRHHYYCCVATMTIVAEGRKRKEGSPPSLFFLRRQQ
WLSPLPLLSWRQRQSPPSLLLLLLPPSSLLWQRQRSLLPLPSLATTAMVVILLPPSSFFF
SPLPLFSYCRHRHYCQQRWRIMTTAAVVCYHRHRCYSFLPLPSFEKKERGGRIITTAAVV
GDNSNSGDAVITITAVDNNDSGGSSSNNDDGGVATITIVADGSNSNDGYHHYRCCRRRKK
EEGGRRHHCCYYGEKRGEKRAVMTAVITATTAVIIVRRQQWVATMVATIEGRKEGRKSPP
SFLLFLLPPLLSSLSPLLSPLWRQRQSPPSLLLLLLPPFFSSWQRQRSLLPSLLPATTAV
IMVSSNNGGDDGDNSGDNSAVITAVITDNNDGSSGRRKKGGEEVATIVATIDGSNDGSNY
HHYCRRCIMTTVVAAFSPLFSPLYCRHYCRHQQRRWIITTVVAACYHRCYHRSFLPSFLP
EKKEGRRGATTAIIVVRRHHCYYCGEKREKRGAVMTVITASPPSLLFFRRQQWVATMATI
VEGRKGRKESPPSLLLLLLPPLSSLLSPLSPLLWRQRQATTAIMVVLLPPFSSFWQRQRS
LLPLLPSYHHYRRCCRRKKGEEGVATIATIVDGSNGSNDQQRRWIMTTVAAVFSPLSPLF
YCRHCRHYEKKERRGGIITTVAAVCYHRYHRCSFLPFLPSDNNDSSGGSSNNGDDGGDNS
DNSGAVITVITAPSLLSPLLPLSLPLSLPSFLSFLPPLSLPLSLTAIVATVIHRYCHCYR
TAVMAVITKGERKREGHYRCYHCRTIAVTVAINDGSDGSNTVAITIAVKERGEKGRKREG
KGERQRWRQHCYRHRYCTAIVATVMPLSFPFSLTAVIAVITPFSLPLSFPSLFSPFLNSD
GNGDSKEGREGRKQWRQRNDSGDNGSTIAVTVAMHYCRYCRHTVAITMAVQRWQRQRWQR
PSLLSLLPNGDSNSDG
(Protein sequence after removing stop codons)