Research Journal of Applied Sciences 6 (4): 258-260, 2011
ISSN: 1815-932X

Modeling the Distribution of Primes in a Biomolecular System

O. Okunoye Babatunde
Department of Pure and Applied Biology, Ladoke Akintola University of Technology,
Ogbomoso, Nigeria

Abstract: A Prime number is a natural number larger than one which cannot be expressed as the product of two smaller natural numbers. Prime numbers have been shown to have significant applications in nature as shown by the life cycle of cicadas of the genus *Magicicada*. Determining the exact distances between prime numbers is an important unsolved problem in mathematics. This study investigated the distances between prime numbers generated from nucleotides in a segment of Bacteriophage T4 genome. The digits of Euler's number e and Pi π, two mathematical constants were found to be encoded spatially between the Viral DNA helical segments comprised of prime numbers.

Key words: Prime number, DNA, Euler's number, pi, segments, Nigeria

INTRODUCTION

A positive integer is said to be prime in case it has exactly two positive divisors (Eynden, 2001). The two divisors are one and itself. The prime numbers <20 are 2, 3, 5, 7, 11, 13, 17 and 19. Prime numbers, once thought to be of interest only to pure mathematicians have found application in applied science, mainly in the protocols of information security and cryptography with the development of Public Key Cryptography whose security lies in the difficulty of factoring large numbers into their prime factors.

In Biological science, the emergence of prime numbers as life cycles of cicada of the genus *Magicicada* (Hoppensteadt and Keller, 1976; May, 1979; Barnett, 1997; Goles *et al*., 2001; Campos *et al*., 2004) has been found to be an evolutionary strategy to escape predators. Advances in DNA computing (Stojanovic and Stefanovic, 2003; Okamoto *et al*., 2004; Margolin and Stojanovic, 2005) have brought biological molecules into the field of computer science and information technology. Deoxyribonucleic acid and other cellular organelles are now employed as input and output in performing routine calculations and building molecular computing devices.

Determination of the exact distribution of prime numbers is an important unsolved problem in mathematics. The question is closely linked with the Reimann hypothesis (Tao, 2007) which predicts a precise formula for the probability distribution of primes. The distances between 20, 000 prime numbers generated from Bacteriophage T4 DNA was investigated with the objective of modeling their distribution. Euler's number (e) and Pi (π) are two of the most important mathematical constants and a part of Euler's identity (Table 1):

$$e^{i\pi}+1 = 0$$

Table 1: Showing frequency distribution of distances between prime numbers In T4 Phage DNA

Distances between primes (x)	Frequency (f)
5	215
-4	90
-3	1455
-2	1051
-1	3662
0	7018
1	3696
2	1059
3	1446
4	119
5	189

The value of e and π to ten significant figures, respectively are e = 2.718281828 and π = 3.141592653.

MATERIALS AND METHODS

Bacteriophage T4 genome (Miller *et al*., 2003) was sourced from GenBank, the institutional genome depository with accession number AF158101. Bacteriophage T4 represents the most understood model for modern genomics and proteomics and its study has revealed many insights and paradigms in molecular biology. The numbers of each nucleotide base per DNA helical segment, comprising of ten bases (Weaver, 2005) were counted and recorded in successive helical segments of T4 phage DNA (complement 1'-99, 180' in the 3'-5' direction). The prime numbers present as numbers of nucleotides were 2, 3, 5 and 7 and the distances between them in 20, 000 samples was investigated (Fig. 1). T4 Phage DNA helical segments comprising only of prime numbers consists of the following combinations of nucleotide bases: 55, 37, 235 and 2233. The spatial distances between these helical segments was counted and recorded (Table 2).

Table 2: The spatial distances of digits of e as encoded between the helical chains of T4 DNA comprised of prime numbers

e	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}	e_{11}	e_{12}
2	26	170	270	410	620	702	780	896	1014	1173	1401	1571
7	35	175	280	428	630	714	789	915	1032	1174	1419	1575
1	36	183	281	431	633	723	814	921	1039	1178	1420	1579
8	69	196	293	571	635	731	865	929	1047	1191	1427	1589
2	90	202	300	572	650	735	866	931	1049	1204	1466	1594
8	103	222	308	585	661	750	867	955	1080	1205	1479	1621
1	105	225	320	593	662	753	879	959	1082	1208	1484	1623
8	122	258	348	595	686	758	881	995	1135	1338	1522	1635
2	130	261	351	613	687	760	885	1003	1136	1354	1526	1636
8	168	265	378	619	700	776	890	1007	1161	1398	1559	1642

Table 3: The spatial distances of digits of e as encoded between the helical chains of T4 DNA comprised of prime numbers

π	π_1	π_2	π_3	π_4	π_5	π_6	π_7	π_8	π_9	π_{10}	π_{11}	π_{12}	π_{13}	π_{14}	π_{15}
3	23	161	238	429	525	665	786	982	1072	1146	1207	1358	1443	1498	1596
1	27	163	240	431	539	675	814	984	1078	1153	1208	1365	1448	1519	1601
4	37	179	288	433	545	688	818	987	1086	1155	1216	1369	1449	1530	1619
1	40	183	295	437	548	689	834	1002	1090	1157	1219	1376	1453	1532	1623
5	57	190	297	447	559	694	839	1008	1092	1158	1234	1379	1469	1535	1627
9	109	195	337	455	601	742	883	1040	1102	1177	1258	1411	1473	1561	1683
2	119	202	351	458	613	744	885	1049	1113	1190	1282	1412	1478	1571	1689
6	123	207	385	463	616	763	965	1057	1116	1195	1350	1413	1483	1576	1690
5	135	224	401	512	618	768	976	1061	1137	1196	1352	1437	1490	1580	1691
3	147	227	402	513	654	782	979	1070	1144	1198	1355	1441	1491	1583	1706

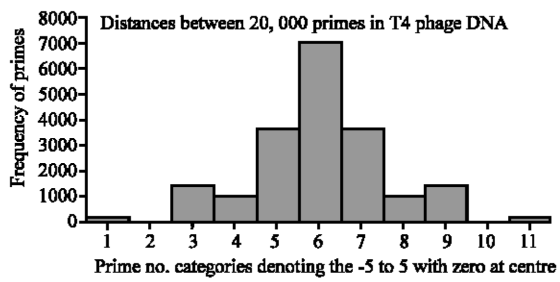

Fig. 1: Histogram showing the frequency distribution of the distances between prime numbers in T4 Phage DNA

RESULTS AND DISCUSSION

The distances between the prime numbers generated from Bacteriophage T4 nucleotide bases followed a binomial distribution.

The mean and standard deviation of a binomial random variable is expressed as: $\mu = np$ and $\sqrt{np\ (1-p)}$ (Weiss, 2002). The distances between T4 Phage DNA helices comprised of prime numbers encoded 12 values of e and 15 values of π (to ten significant figures) (Table 3).

Prime numbers are regarded as the atomic elements of natural number multiplication (Tao, 2007). Prime number research is an expansive field in mathematical number theory, often churning out interesting results such as the discovery of the 45th known Mersenne prime: $2^{43,112,609}$-1 which is the largest discovered prime number till date. This research work attempts to bridge the fields of mathematical number theory with molecular biology.

CONCLUSION

In this study the occurrence of the mathematical constants e and π encoded within DNA molecules also an interesting experimental result, perhaps shows that mathematics is the foundation of all there is even molecular biology.

REFERENCES

Barnett, A., 1997. Odd timing keeps cicadas in the prime. New Sci., 153: 18-18.

Campos, P.R.A., V.M. de Oliveira, R. Giro and D.S. Galvao, 2004. Emergence of prime numbers as the result of evolutionary strategy. Phys. Rev. Lett., 93: 098107-098110.

Eynden, C.V., 2001. Elementary Number Theory. 2nd Edn., McGraw-Hill, New York, ISBN-13: 9780072325713, pp: 278.

Goles, E., O. Schulz and M. Markus, 2001. Prime number selection of cycles in a predator-prey model. Complexity, 6: 33-38.

Hoppensteadt, F.C. and J.B. Keller, 1976. Synchronization of periodical cicada emergences. Science, 194: 335-337.

Margolin, A.A. and M.N. Stojanovic, 2005. Boolean calculations made easy (for ribozymes). Nat. Biotechnol., 23: 1374-1376.

May, R.M., 1979. Periodical cicadas. Nature, 227: 347-349.

Miller, E.S., E. Kutter, G. Mosig, F. Arisaka, T. Kunisawa and W. Ruger, 2003. Bacteriophage T4 genome. Microbiol. Mol. Biol. Rev., 67: 86-156.

Okamoto, A., K. Tanaka and I. Saito, 2004. DNA logic gates. J. Am. Chem. Soc., 126: 9458-9463.

Stojanovic, M.N. and D. Stefanovic, 2003. A deoxyribozyme-based molecular automaton. Nat. Biotechnol., 21: 1069-1074.

Tao, T., 2007. Structure and randomness in the prime numbers: A small collection of results in number theory. UCLA Science Colloquium, January 17, 2007. http://www.math.ucla.edu/~tao/preprints/Slides/primes.pdf.

Weaver, R.F., 2005. Molecular Biology. 3rd Edn., McGraw-Hill, New York, pp: 27.

Weiss, N.A., 2002. Elementary Statistics. 5th Edn., Addison-Wesley, Massachusetts, ISBN-13: 9780201710588, pp: 768.