DOI: https://doi.org/10.24297/ijct.v24i.9635

Redundant Residue Number System (RRNS) Di-Base Table for SOLiD Sequencing

[1]Joshua Apigagua Akanbasiam, [2]Kwame Osei Boateng, and [3]Matthew Glover Addo,

[1]Department of Electrical/Electronics Engineering, Dr. Hilla Limann Technical University Wa, Ghana

ja.akanbasiam@dhltu.edu.gh or ja.akanbasiam@gmail.com

[2]Department of Computer Engineering, Kwame Nkrumah University of Science and Technology Kumasi, Ghana

boat.coe@knust.edu.gh

[3]Department of Theoretical and Applied Biology, Kwame Nkrumah University of Science and Technology Kumasi, Ghana

maddo.sci@knust.edu.gh

Abstract.

The next-generation sequencing (NGS) methodology, sequencing by oligonucleotide ligation and detection (SOLiD) uses a di-base table, a somewhat unusual method, to decode sequences. Its coding scheme is based on the binary number system. The di-base table is not connected to the genetic code, nor is the coding scheme structured in the space of an entire number system. Gamow also revealed the hidden attribute of a 4×4 code for the di-base table, supporting his proposal of a $4 \times 4 \times 4$ codons for the genetic code. Consideration for digital applications has focused more on the Residue Number System (RNS) and Redundant Residue Number System (RRNS) lately. Consequently, an RRNS di-base table based on the number tree concept is designed. The designed RRNS di-base table deviates from the canonical di-base table but retains every attribute necessary for effective SOLiD decoding. It shares a close relationship with the RNS-Genetic code and this presents a compelling argument for creating a single instrument that possesses the capabilities of both the genetic code and the di-base table.

Keywords: Residue Number System, Redundant Residue Number System, di-base table, RRNS di-base table, Sanger sequencing, Next Generation Sequencing, SOLiD Sequencing

INTRODUCTİON

The demand for DNA sequencing is growing exponentially and with the help of innovative tools and algorithms, the sequencing process has become faster, more accurate and less expensive. These processes can be enhanced with a well-structured number system and innovative algorithms. Basic research exploring biological processes as well as practical fields like diagnostics and forensics have become essential. Sanger and Maxam and Gilbert's methods are the two first-generation sequencing techniques that were developed. Sanger's chain termination technique gained popularity due to its expediency and practicality. The need to convey sequencing analysis and diagnosis to hospitals and clinical laboratories (Ansorge 2009) has ushered in new methods. Higher sequencing throughput techniques have become attractive due to the Human Genome Project (HGP). These methods, also known as Next Generation Sequencing (NGS), can produce millions or billions of bases per run and work in tandem with the sequencing process. This is a benefit that these NGS techniques have over previous sequencing techniques. Each NGS technology differs from the other and the advantages over each other are mainly based on read length, accuracy, run time, throughput etc. (Kulski 2016) (Li et al. 2011). SOLiD has the highest accuracy among NGS technologies, with an approximate accuracy of 99.99% [7]. It is often called 2-base encoding since it is predicated on 2-nucleotide sequencing via ligation. The colour space or sixteen (16) di-base matrix is crucial to the decoding process. The encoding matrix's architecture permits the integration of error-checking features. Two-base encoding, which double-examines each base during sequencing, allows for the extremely high certainty detection of the rare Single Nucleotide Polymorphism (SNP). With this approach, researchers can focus on the biological significance of the data instead of the poor-quality data. The major drawbacks are the long run times and the short read length but these are usurped by the lowest error rates. In SOLiD sequencing, four dyes; blue, green, yellow and red are used to encode for the sixteen (16) possible two-base combinations, thus ($4^2 = 16$). Just as the Rosetta stone was the key to deciphering Egyptian hieroglyphs (Stone 2007) the di-base matrix is vital to SOLiD sequencing. One of the highly researched areas in number systems for computer applications in the past decade is RNS. A notable mention is its application in molecular biology and bioinformatics, Smith Waterman Algorithm. An RNS di-base table for SOLiD sequencing is developed using the concept of number trees. Sorting the numbers into the four codon blocks and designating one of the moduli redundant resulting in the di-base table can be made simpler by presenting RNS as a number tree. The possibility of generating the genetic code as RNS and the di-base table as RRNS offer the opportunity to design tools with

the features of both the genetic code and the di-base table. The RNS di-base table generated mimics the canonical di-base table and also reserves all the properties for SOLiD sequencing.

SANGER SEQUENCING

The process of figuring out the nucleotide bases' order within a DNA molecule is known as DNA sequencing (NCBI) n.d.). This area of study is now crucial for application domains including forensics and diagnostics research as well as basic research exploring biological systems. Understanding our DNA provides explanations for differences and similarities between organisms of the same species and those with others - related or unrelated species. Sanger's method and that of Maxam and Gilbert set the tone for further consideration in DNA sequencing. Sanger's method became the de facto standard because of its practicality (França, Carrilho, and Kist 2002). The Sanger sequencing method served the research community for some decades until the quest for high throughput, low cost and relatively accurate sequencing techniques ushered in the massively parallel sequencing technologies, NGS. Though still in use today (Sanger's method), the future of DNA sequencing, thus finding cures for cancer, forensics, criminal justice, and other clinical applications is shifting the attention of the research community greatly towards higher throughput NGS technologies.

NEXT GENERATION SEQUENCING (NGS)

The demand for low-cost, high throughput and timely sequencing necessitated the development of NGS technologies (Berglund, Kiialainen, and Syvänen 2011). The phrase "next-generation sequencing" was created to refer to a variety of contemporary sequencing technologies, including SOLiD sequencing. Different DNA sequencing methods have different trade-offs in speed, read length, cost, and accuracy. The next-generation sequencing method known as Sequencing by Oligonucleotide Ligation and Detection (SOLiD) is based on 2-nucleotide sequencing by ligation and was introduced by Applied Biosystems (ABI) (ABI Solid Sequencing - Wikipedia n.d.). The key to successful SOLiD sequencing is the di-base table (Płoski 2016). With each base being interrogated twice, this method's precision is by far its biggest benefit. The backbone of digital computing is number systems. A highly researched number system in the past decades that promises to offer solutions to the challenges of the binary number system is RNS. There have been many arguments made to support the use of number systems in molecular biological considerations (Mohanty et al. 2023) and data storage (Bhat 2018). RNS has seen some applications in molecular biology (Boateng and Baagyere 2012; Kehinde Bello and Alagbe Gbolagade 2018). A merger of next-generation sequencing technology (SOLiD) and Residue Number System (RNS) is presenting an alternative approach to SOLiD sequencing technological design. The proposed RRNS di-base table though a variant of the canonical di-base table, preserves all the properties set out by Applied Biosystems Instruments (ABI) for successful SOLiD sequencing. The tools that support molecular biological research and bioinformatics are taking on some exciting new dimensions thanks to the benefits promised by RNS in digital computing(E.Y. Baagyere 2011).

SOLiD SEQUENCING TECHNOLOGY

Sequential ligation with dye-labelled oligonucleotides is the foundation of the SOLiD sequencing technique(Płoski 2016). There are several, concurrent, and aggressively vying fluorescently tagged oligonucleotide probes. In Sanger sequencing, each colour designating a single nucleotide is immediately translated to A, C, G, or T. In SOLiD sequencing, however, each colour designating a potential pair of 2-bases represents four different possibilities. Utilising a 2-base encoding technique, the sequencing process double-checks each base for errors and can distinguish between true polymorphisms and system noise. The 2-base encoding also enables reading each base twice without performing twice the work (File:Two-base encoding scheme.pdf - Wikipedia n.d.).

THE CANONICAL DI-BASE TABLE

All four (4) bases are in a grid form to generate all sixteen (16) di-base combinations. The order of bases for the canonical di-base table are A, C, G and T for each row and column. An intersection of any two (2) bases forms a cell which makes up a particular colour for decoding purposes. Some key features make up the canonical di-base code and these are presented as follows. The available colours are four; Blue, Green, Yellow and Red and each cell results in one of these colours. The same colour is produced by a mono- and di-base probe and this can be seen in the leading diagonal as AA, CC, GG, and TT. Di-base probes have the same colour as their opposites, for example, (AC) = (CA), (CG) = (GC), and (GT) = (TG). Furthermore, two independent di-base probes with the same first base produce different colours, giving each cell in each row a unique colour. AA, AC, AG, and AT. Each column and cell in the pattern AC, CC, GC, and TC would have a different colour since two distinct di-base probes with the same second base have different colours. Ultimately, a di-base probe's and its complement's colours are the same. Cytosine (C) complements Guanine (G), while Adenine (A) complements Thymine (T). The di-base table is the key to SOLiD sequence decoding. Figure 1 illustrates the canonical di-base table developed by Applied Biosystems.

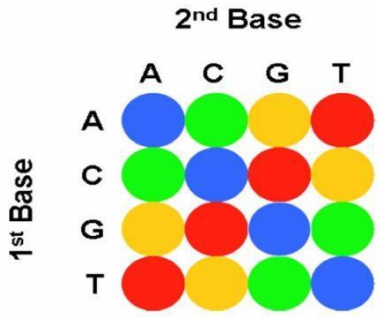

Fig.1 The canonical di-base table - Applied BioSystems(ABI Solid Sequencing - Wikipedia n.d.)

BINARY NUMBER SYSTEM DI-BASE

Binary number systems have proven to be the most effective notation for digital applications. The di-base table can be realised for computerised applications through a two-bit representation of each base; thus, the base order A, C, G, and T for the canonical di-base are assigned the binary digits, 00, 01, 10 and 11 respectively. This forms all 16 colour blocks as the intersection of any two digits. Each block's digits are concatenated to form the four (4) bit binary notation for each block, as shown in Fig. 2(a). This numbering is static and not flexible within the binary number space, as shown in Fig. 2 (c).

	A	C	G	T
A	AA	AC	AG	AT
C	CA	CC	CG	CT
G	GA	GC	GG	GT
T	TA	TC	TG	TT

(a)

	0	1	2	3
0	0000	0001	0010	0011
1	0100	0101	0110	0111
2	1000	1001	1010	1011
3	1100	1101	1110	1111

(b)

	0	1	2	3
0	0	1	2	3
1	4	5	6	7
2	8	9	10	11
3	12	13	14	15

(c)

Fig. 2 Binary notation of the canonical di-base table.

The traditional number system employed in contemporary computer arithmetic is the binary 2's complement. The carry propagation delay sets a cap on the processing performance needed today. This has ushered in other number systems with some comparative advantages. One such, which has been considered extensively by the research community in the past decade is the residue number system. This is therefore adopted for an RRNS di-base table for SOLiD sequencing.

RESIDUE NUMBER SYSTEM

RNS is a non-weighted number system with attractive features for modern-day digital systems design and computations (Daabo 2018). RNS is defined as modular structures (rings or fields) that have moduli which are pairwise relatively prime integers. If $[m_1, m_2, \ldots\ldots m_L]$ represent the set of moduli then the product $m_1 \times m_2 \times \ldots . m_l = M$, is the dynamic range of the RNS. Then the interval [0, M-1] is the set of all legitimate operational range of the number system. RNS can be presented as two (2), three (3), four (4) etc. moduli set depending on the application and dynamic range required (Amos and Premkumar 2007). RNS is non-weighted as such it is carry-free, and the lack of carry propagation means that errors in one digit cannot spill over to neighbouring digits. This physiognomy makes it inherently supportive of fault-tolerant computing – error control coding. One of the contemporary techniques for high-speed digital applications and computing is parallelism. RNS breaks operations into smaller units that can be executed in parallel and as such support high-speed computations. In the design of an RRNS di-base table (matrix) at least a two (2) moduli set RNS is required. The moduli set should be greater than or equal to four (4) to uniquely represent each of the four (4) known nitrogenous bases. The dynamic range should be greater than sixteen (16) since the operational range required for the generation of the di-base table is sixteen (16). The relationship between number systems and molecular biology is expanding (Mohanty et al. 2023) and RNS has seen some successful research in bioinformatics and molecular biology (Kehinde and Alagbe 2018) (Kehinde Bello and Alagbe Gbolagade 2018). An RRNS is an RNS with some moduli set declared redundant. This results in two computational ranges, the legitimate and the illegitimate. The redundant residue number system (RRNS), has the unique ability for error control coding. In this research, it is adopted for the di-base design. In the ensuing section, we look at the nature of the number tree and how it forms the ease and flexibility of generating the genetic code and the di-base table.

THE NUMBER TREE

A non-linear discrete data structure known as a tree depicts the hierarchical relationship between different components or nodes. It defines parent-child relationships by having a finite collection of directed edges and a finite set of nodes. The "beginning" node is called the root – which has no parent and every edge directly or indirectly originates from the root. The predecessor node is the parent and the successor nodes are the children, nodes without children are leaves. Trees are extensively applied in algebra and programming. A tree data structure is hierarchical because it uses a parent-child hierarchy to define the relationships between distinct nodes. Hierarchy data structures are most popular for simplifying and speeding up searching and sorting. The performance of a tree algorithm is directly related to the number of ascendants of a particular node and the number of descendants of that node. The sum of the overall number of ascendants is the path length, which is a quantity used to measure the complexity of the tree algorithm. This greatly suits this research consideration since a three-moduli system with two (2) ascendants for the genetic code is further reduced to two moduli with only one ascendant. This reduces the level of complexity required to generate the di-base table. The tree can be implemented using arrays, classes, and connected lists since it is a collection of interconnected nodes. When every level of a tree is full, there are no "gaps", it's termed a "perfect" tree. RNS-based trees are perfect, and the nodes on the last level are equal to the dynamic range of the moduli sets chosen. Also, the number of nodes at each level is determined by the moduli set at that level. Trees are adaptable and provide effective search and insertion operations, making it easy to move subtrees. As a result, they are implemented in a file directory, and classical inheritance tree in a single inheritance language. Next, the approach to achieving the objectives of this research work is discussed.

METHODOLOGY

The di-base table (matrix) is made of four dyes; blue, green, yellow and red which are used to encode all sixteen (16) possible di-base combinations. The idea of generating the di-base matrix for SOLiD sequencing has its roots in the early days of the genetic code where George Gamow formulated the mathematical thinking that a $4^3 = 64$ code was required to generate all 64 codons of the genetic code. Embedded in his thought is that a $4^2 = 16$ will suffice for the di-base matrix, which is required for SOLiD sequencing. The RNS di-base matrix is generated from a three residue number moduli - m_1, m_2, and m_3. The three moduli set RNS is designed into a tree or forest and with digit base substitution an appropriate truncation of the genetic code is produced. When one of the moduli is declared redundant with appropriate truncation and digit base substitution the RNS di-base table is formed. Figure 3 below gives a summary of the design of the di-base table.

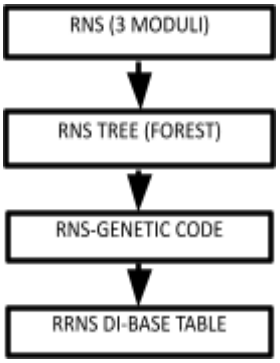

Fig.3 Design Flow RNS or RRNS di-base table.

A three moduli set [4, 5 and 7] RNS are generated using the concept of number tree, Fig. 4, and these digits have a unique characterization that is used to generate the di-base table. The number tree approach proposed is flexible since it allows for the generation of the genetic code and the di-base table in a single device. A three-moduli set is required to generate the genetic code, and when one of the moduli sets is declared redundant resulting in an RRNS, the di-base table is generated. In designing the di-base code from a tree residue number moduli set, some foundations are established. A three moduli sets are considered; m_1, m_2 and m_3. The moduli set chosen must be greater than or equal to four (4) to adequately represent the four (4) nitrogenous bases – T, C, A and G. The maximum achievable roots, nodes and leaves of the tree are m_1-1 and m_2-1 and m_3-1 respectively. The minimum achievable dynamic range for a moduli set of [4,5] where the third moduli is declared redundant is twenty (20) and thus [0 – 19]. This is illustrated in section 1.10. Any value of the moduli set can always be truncated to the 16-code matrix, allowing for flexibility and the code serving the entire RNS space. Figure 4 below depicts the RNS tree based on moduli sets, 4, 5 and 7.

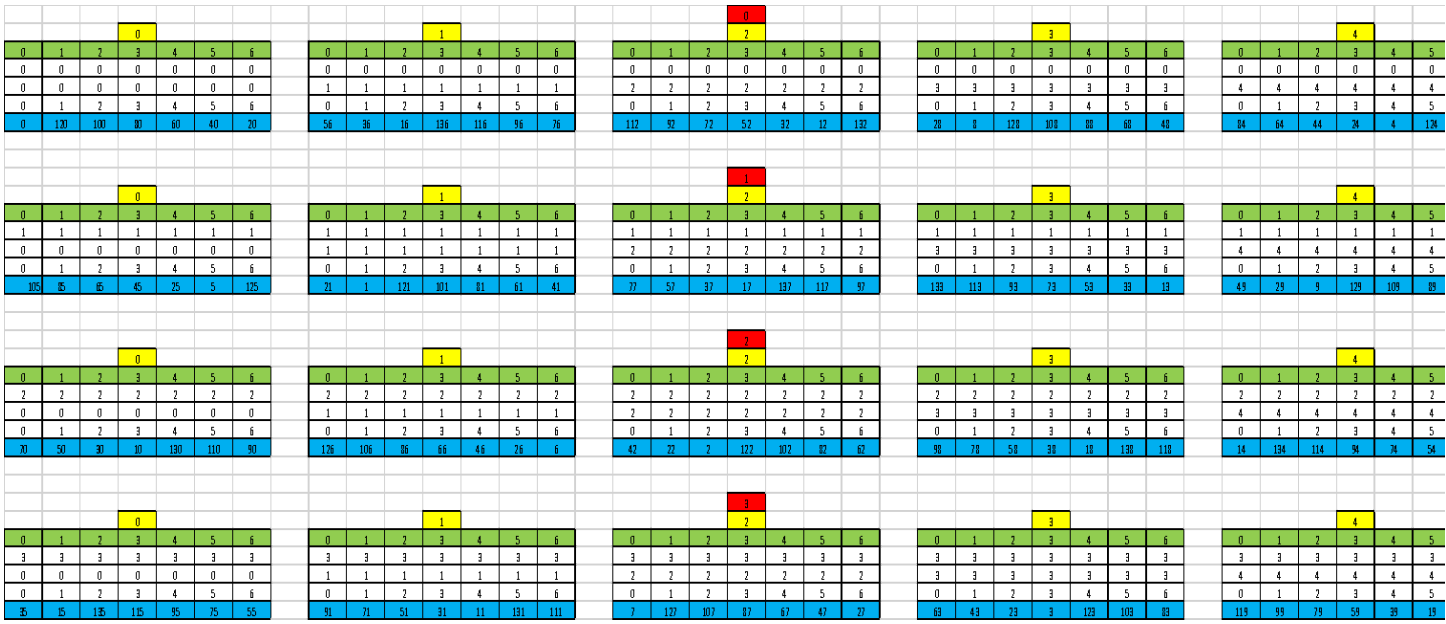

Fig.4 RNS tree with moduli 4, 5 and 7.

There are four (4) di-nucleotide sequences associated with each colour. The decoding matrix is generated using the concept of RNS and number trees. The di-nucleotide is viewed as the two moduli sets of RNS and the concept of number trees and the unique arrangement of digits sorts the RNS digits into blocks. A digit base substitution and appropriate truncation aid in converting the RNS digits to their respective codon family for generating the genetic code. By declaring one of the moduli redundant and most appropriately the third moduli set, which is responsible for wobbling, the di-base table is generated. The digit base arrangement for moduli digits 0, 1, 2, and 3 is T, C, A, and G. This is necessary for the success of both the genetic code and the di-base table as a single operation. This deviates from the canonical, which uses the digit bases A, C, G, and T for the numbers 0 through 3. This produces a non-canonical di-base table, a variant, which reserves all the properties for successful SOLiD sequencing, Fig. 5 illustrates the path of the RNS genetic code to the RRNS di-base table.

000	010	020	030
001	011	021	031
002	012	022	032
003	013	023	033
100	110	120	130
101	111	121	131
102	112	122	132
103	113	123	133
200	210	220	230
201	211	221	231
202	212	222	232
203	213	223	233
300	310	320	330
301	311	321	331
302	312	322	332
303	313	323	333

(a)

00	01	02	03
00	01	02	03
00	01	02	03
00	01	02	03
10	11	12	13
10	11	12	13
10	11	12	13
10	11	12	13
20	21	22	23
20	21	22	23
20	21	22	23
20	21	22	23
30	31	32	33
30	31	32	33
30	31	32	33
30	31	32	33

(b)

	U	C	A	G
U	00	01	02	03
C	10	11	12	13
A	20	21	22	23
G	30	31	32	33

(c)

Fig.5 The RNS genetic code and di-base table.

DESIGN FLOW – ALGORITHM

The Table below illustrates the algorithm required to generate the RNS dibase table for SOLiD sequencing.

Table 1 RNS Di-base Algorithm

```
Array GeneticCode[4,4,4]
for i ← 0 to m1 – 1
    for j ← 0 to m2 – 1
        for k ← 0 to m3 – 1
            if ((i < 4) && (j < 4) && (k < 4))
                GeneticCode[i,j,k] ← ijk
        Od
    Od
Od

for x ← 0 to ma - 1 do
    for y ← 0 to Mb - 1 do
        if ((x < 4) && (y < 4))
            DibaseTable[i,j] ← ij
    od
od
return DibaseTable
```

RESULTS AND DISCUSSION OF MODEL

The leading and trailing diagonals represent the colours blue and red respectively. The residue digits 01 and 02 also represent the colours yellow and green respectively. Thus, the rules or properties set out in the canonical di-base table (Applied-Biosystems 2011) are highly conserved in the RRNS di-base table. The RNS di-base table generated is a copycat of the canonical di-base table (Cheng, Fei, and Xiao 2023) designed by applied Biosystem for any SOLiD sequencing. If the digit-base substitution is 0 –> T, 1–> C, 2–> A, 3–> G following from the genetic code, then a variant of the canonical di-base table is achieved. Thus, the canonical di-base table (Life Technologies 2010),(Byosystems 2008) has di-bases different from that of the RRNS di-base table, Fig. 6. Another unique feature of this method is that the combination of digits in a cell has decimal values. These decimal values are generated automatically compared with the canonical (SOLiD 2008) and can be exploited for further biological analysis, Fig. 7. These decimal values also change with changing moduli but the di-base digits remain the same since the digits that make each cell do not change. Thus, the residue digits are static while the decimal values are dynamic, Fig. 6 (a), (b) and (c) present the static residue digits for different moduli sets whereas Fig. 7 (a), (b) and (c) illustrate the difference in moduli values. Figures 8 (a), (b) and (c) present the final RRNS di-base table.

4,5	0	1	2	3	4
0	00	01	02	03	04
1	10	11	12	13	14
2	20	21	22	23	24
3	30	31	32	33	34

(a)

4,7	0	1	2	3
0	00	01	02	03
1	10	11	12	13
2	20	21	22	23
3	30	31	32	33
4	40	41	42	43
5	50	51	52	53
6	60	61	62	63

(b)

5,7	0	1	2	3	4
0	00	01	02	03	04
1	10	11	12	13	14
2	20	21	22	23	24
3	30	31	32	33	34
4	40	41	42	43	44
5	50	51	52	53	54
6	60	61	62	63	64

(c)

Fig.6 The residue digits of a 4,5; 4,7 and a 5,7 RNS di-base.

4,5	0	1	2	3
0	0	16	12	8
1	5	1	17	13
2	10	6	2	18
3	15	11	7	3

(a)

4,7	0	1	2	3
0	0	8	16	24
1	21	1	9	17
2	14	22	2	10
3	7	15	23	3

(b)

5,7	0	1	2	3
0	0	15	29	10
1	21	1	16	31
2	7	22	2	17
3	28	8	23	3

(c)

Fig.7 The decimal values. of a 4,5; 4,7 and 5,7 RNS di-base.

4,5	T	C	A	G
T				
C				
A				
G				

(a)

4,7	T	C	A	G
T				
C				
A				
G				

(b)

5,7	T	C	A	G
T				
C				
A				
G				

(c)

Fig.8 Final RRNS di-base Table.

CONCLUSION

This research presents a redundant residue number system di-base table using the concept of number trees. This design is flexible across the entire residue number system space and is a variation of the canonical di-base table presented by Applied Biosystems Instruments (ABI) but retains all the guidelines and characteristics necessary for successful decoding of a SOLiD sequence. Its close relationship to the genetic code permits the development of a single tool that combines the functionality of both the di-base table and the genetic code.

Abbreviations

A	Adenine
ABI	Applied Biosystems Instruments
C	Cytosine
DNA	Deoxyribonucleic Acid
G	Guanine
HGP	Human Genome Project
NGS	Next-Generation Sequencing
RNS	Residue Number System
RRNS	Redundant Residue Number System
SOLiD	Sequencing by Oligonucleotide Ligation and Detection
T	Thymine

Acknowledgements

This research is part of a thesis submitted to the Kwame Nkrumah University of Science and Technology, Kumasi, Ghana in 2024 by Ing. Dr. Joshua Apigagua Akanbasiam for the award of Doctor of Philosophy in Computer Engineering. The other authors Ing. Prof. Kwame Osei Boateng and Prof. Matthew Glover Addo supervised the Thesis.

Author's Contributions

Each of the authors have made significant contributions to the manuscript. JAA the corresponding author was a PhD candidate under the supervision of KOB and MGA. All authors read and validated the final manuscript.

Funding

Not applicable

Availability of data and materials

Not applicable

Competing Interest

The authors declare that they have no competing interest

REFERENCES

"ABI Solid Sequencing - Wikipedia." https://en.wikipedia.org/wiki/ABI_Solid_Sequencing (May 24, 2021).

Amos, Omondi, and Benjamin Premkumar. 2007. Electronic Computers, IRE Transactions on Residue Number System Theory and Implementation.http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5219515. https://doi.org/10.1142/9781860948671

Ansorge, Wilhelm J. 2009. "Next-Generation DNA Sequencing Techniques." New Biotechnology 25(4): 195–203.

https://doi.org/10.1016/j.nbt.2008.12.009

Applied-Biosystems. 2011. "Principles of Di-Base Sequencing and the Advantages of Color Space Analysis in the SOLiD System." Application Note: 2–5.

Berglund, Eva C., Anna Kiialainen, and Ann Christine Syvänen. 2011. "Next-Generation Sequencing Technologies and Applications for Human Genetic History and Forensics." Investigative Genetics 2(1): 23. http://www.investigativegenetics.com/content/2/1/23. https://doi.org/10.1186/2041-2223-2-23

Bhat, Wasim Ahmad. 2018. "Bridging Data-Capacity Gap in Big Data Storage." Future Generation Computer Systems 87(2018): 538–48. https://doi.org/10.1016/j.future.2017.12.066.

Boateng, Kwame O, and Edward Y Baagyere. 2012. "A Smith-Waterman Algorithm Accelerator Based on Residue Number System." International Journal of Electronics and Communication Engineering 5(1): 99–112. http://www.irphouse.com.

Byosystems, Applied. 2008. "Analysis Tools (SAT)." Processing. www.appliedbiosystems.com.

Cheng, Chu, Zhongjie Fei, and Pengfeng Xiao. 2023. "Methods to Improve the Accuracy of Next-Generation Sequencing." Frontiers in Bioengineering and Biotechnology 11(January): 1–13. https://doi.org/10.3389/fbioe.2023.982111

Daabo, Mohammed I. 2018. "A Hybrid Residue to Binary Converster for The Moduli Set." International Research Journal of Engineering and Technology (IRJET): 458–64.

E.Y. Baagyere, K.O. Boateng and K.A. Gbolagade. 2011. "Bioinformatics: An Important Application Area of Residue Number System." https://doi.org/10.3923/jeasci.2011.174.179

"File:Two-Base Encoding Scheme.Pdf - Wikipedia." https://en.wikipedia.org/wiki/File:Two-base_encoding_scheme.pdf (June 4, 2021).

França, Lilian T.C., Emanuel Carrilho, and Tarso B.L. Kist. 2002. "A Review of DNA Sequencing Techniques." Quarterly Reviews of Biophysics 35(2): 169–200. https://doi.org/10.1017/S0033583502003797

Kehinde Bello, Hassan, and Kazeem Alagbe Gbolagade. 2018. "Acceleration of Biological Sequence Alignment Using Residue Number System." Asian Journal of Research in Computer Science 1(2): 1–10. https://doi.org/10.9734/ajrcos/2018/v1i224735

Kehinde, Hassan, and Kazeem Alagbe. 2018. "Residue Number System: An Important Application in Bioinformatics." International Journal of Computer Applications 179(10): 28–33. https://doi.org/10.5120/ijca2018916106

Kulski, Jerzy K. 2016. "Next-Generation Sequencing — An Overview of the History, Tools, and 'Omic' Applications." Next Generation Sequencing - Advances, Applications and Challenges: 3–60. https://research-repository.uwa.edu.au/en/publications/next-generation-sequencing-an-overview-of-the-history-tools-and-o-2 (December 20, 2021). https://doi.org/10.5772/61964

Li, Lijin et al. 2011. "Cancer Genome Sequencing and Its Implications for Personalized Cancer Vaccines." Cancers 3(4): 4191–4211. https://doi.org/10.3390/cancers3044191

Life Technologies. 2010. "SOLiD Data - 2 Base Encoding." : 24.

Mohanty, Vaibhav et al. 2023. "Maximum Mutational Robustness in Genotype – Phenotype Maps Follows a Self-Similar Blancmange-like Curve." https://doi.org/10.1101/2023.03.11.532236

NCBI). The Principles of DNA Sequencing.

Płoski, Rafał. 2016. Clinical Applications for Next-Generation Sequencing Next Generation Sequencing-General Information about the Technology, Possibilities, and Limitations. Elsevier Inc. http://dx.doi.org/10.1016/B978-0-12-801739-5.00001-5.
https://doi.org/10.1016/B978-0-12-801739-5.00001-5

SOLiD. 2008. "SOLiD TM Data Format and File Definitions Guide." 2008.

Stone, Rosetta. 2007. "Rosetta Stone." 11(December): 1–8.

BIOGRAPHIES

Ing. Joshua Apigagua Akanbasiam (PhD) holds PhD, degree in Computer Engineering, a Masters in Telecommunications Engineering and BSc. in Computer Engineering, KNUST. He is a Senior Lecturer and currently the head of the Department of Electrical/Electronics Engineering, at Dr. Hilla Limann Technical University formerly Wa Polytechnic - Ghana. Before Lecturing he served as the Maintenance Manager of Mass Telecom Innovation (MTI) a company that maintained Telecom Power for MTN and Airtel cell sites in the three Northern Regions. He is currently the head of Department of Electrical/Electronics Engineering and a member of GhIE and IEEE. His research interest is in Broadband services, Computer Arithmetic and Applications, RNS and its application, Digital Signal Processing, Computational Biology and Bioinformatics.

Prof. Ing. Kwame Osei Boateng holds a Doctorate in Systems Engineering from the Graduate School of Engineering and Science of Ehime University, Japan, Masters in Computer Science from the same university and a BSc. (Hons.) degree in Electrical/Electronic Engineering from the University of Science and Technology (now Kwame Nkrumah University of Science and Technology), Kumasi Ghana. Since March 2003, Prof. Ing. Boateng has been working for Kwame Nkrumah University of Science and Technology and has risen through the ranks as a Senior Lecturer, an Associate Professor and currently a Professor. He has held several positions as head of Department of Computer Engineering, the dean of the Faculty of Electrical and Computer Engineering of the College of Engineering and director of the Institute of Distance Learning and ICT consultant. Prof. Ing. Boateng is a member of the GhIE, IEEE and IEEE Computer Society. He has served on the technical programme committee for the 11th IEEE (ATS'02) and IEEE VLSI Test Symposium (VTS), ICAST 2012, WTS 2013, and ESTE 2015. His present research interests are in the areas of design, test and diagnosis of logic and VLSI circuits, test of mixed-signal circuits, network security protocols, applications of residual number systems (RNS), image processing and smart metering.

Prof. Matthew Glover Addo; Mathew Glover Addo is a Professor of Microbiology/Molecular Biology at the Department of Theoretical and Applied Biology, Faculty of Biosciences at the Kwame Nkrumah University of Science and Technology (KNUST). Currently, he is the Director of the Institute of Distance Learning, KNUST. Prof Addo holds a BSc in Biological Sciences from KNUST-Ghana, an MSc in Biotechnology from the University of Bergen, Norway and Doctor of Science/Doctor of Philosophy degrees at the Université Paris Sud IX, Paris, France and KNUST, Kumasi. He completed the doctorate programme in May 2011. During the doctorate research programme, Prof. Addo designed an efficient screening method for the identification of genes involved in the mitochondrial genome stability using Caenorhabditis elegans as a model organism. He worked with the Functions and Dysfunction of Mitochondrial group of the Institute of Genetics and Microbiology where for the first time, he identified four (4) new nuclear genes (Y105E8A.23, dnj-10, atad-3, and phi-37) involved in mitochondrial stability. Prof. Addo has 20 years of teaching, collaborative research and consultancy experience. His specialisation is in Gene Expression, Clinical Infectious Microorganisms and Food and Water Microbiology. He has served the international and local communities in several capacities and has several peer-reviewed publications in reputable international journals. He has also assessed several MSc/MPhil and PhD theses. Prof. Addo is a visiting Lecturer/Assessor at the ISA-Lille University, Lille, France. Before he was appointed Director of IDL, Prof Addo was the Dean of the Faculty of Biosciences.