Nano Today 33 (2020) 100871

Contents lists available at ScienceDirect

Nano Today

journal homepage: www.elsevier.com/locate/nanotoday

Information stored in nanoscale: Encoding data in a single DNA strand with Base64

Yi Zhang[a,c,1], Linlin Kong[b,1], Fei Wang[d,e], Bo Li[a], Chao Ma[a], Dong Chen[b,**], Kai Liu[a,c,f,*], Chunhai Fan[d], Hongjie Zhang[a,c,f]

[a] State Key Laboratory of Rare Earth Resource Utilization, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun, 130022, China
[b] Institute of Process Equipment, College of Energy Engineering and State Key Laboratory of Fluid Power and Mechatronic Systems, Zhejiang University, Hangzhou, 310027, China
[c] University of Science and Technology of China, Hefei, 230026, China
[d] Frontiers Science Center for Transformative Molecules, School of Chemistry and Chemical Engineering, and Institute of Molecular Medicine, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, 200240, China
[e] Joint Research Center for Precision Medicine, Shanghai Jiao Tong University Affiliated Sixth People's Hospital South Campus, Southern Medical University Affiliated Fengxian Hospital, Shanghai, 201499, China
[f] Department of Chemistry, Tsinghua University, Beijing, 100084, China

ARTICLE INFO

Article history:
Received 19 February 2020
Received in revised form 16 March 2020
Accepted 1 April 2020
Available online 4 May 2020

Keywords:
Information storage
Coding DNA
Base64
Logical storage density
Double direction sequencing

ABSTRACT

DNA as a storage medium has enormous potential because of its high storage density, but the produced redundancy limits this potential. The introduction of less error corrections to fully increase the storage density in DNA remains a major challenge. To address this, an optimized Base64 method is developed and accordingly we realized a high specific storage density of 1.77 bits/nucleotide in a DNA single strand. In this strategy, by Base64 encoding, code reshaping and balancing, and data mapping, some random text information was encoded into a DNA sequence and the corresponding DNA molecule was synthesized. It was then inserted into a circular plasmid for long-term information storage. This is also particularly suitable for information replication at an exponential rate when it is transformed in a bacterium. The introduction of balance codes during the transcoding process effectively controlled the GC content and continuous base repeat, which is important to reduce the error rates in the encoded DNA synthesis and sequencing. Moreover, the circular plasmid platform enhanced the storage stability and sequencing accuracy. Therefore, our approach achieved a robust and high efficient storage and an accurate readout of digital data.

Introduction

Information explosion nowadays produces digital data at exponential rates, resulting in a significant challenge for high-density and long-term data storage. The current mainstream storage media, such as magnetic, optical, and solid-state media, only possess limited storage density due to their binary coding of only "1" or "0" on each bit, and their stored data become unrecoverable after a century or even less. To address the challenge, DNA, a robust genetic information carrier has evolved in nature for billions of years, appears as a promising storage media [1–4]. There are several advantages of DNA as a digital storage carrier, such as ultrahigh storage density, robust stability, low energy cost, easy replicability, and operability. There are 4 basic nucleotides, A, T, C and G, and each nucleotide can store two bits information, which is twice of the storage density of conventional storage media. Once data are stored in DNA, abundant transcripts could be created at high speed and low energy cost, and the DNA sequences remain recoverable and accurate after thousands of years [5]. Therefore, it is becoming advantageous and urgent for information storage by DNA molecules.

In the last few years, DNA has been regarded as a promising candidate for next-generation storage. Clelland [6] developed an approach to store information in a DNA microdot and this also offered a possible way for information encryption. Carter and coworkers [5] utilized polymerase chain reaction (PCR) and DNA

* Corresponding author at: State Key Laboratory of Rare Earth Resource Utilization, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun, 130022, China.
** Corresponding author.
E-mail addresses: chen.dong@zju.edu.cn (D. Chen), kai.liu@ciac.ac.cn (K. Liu).
[1] Equal contribution.

https://doi.org/10.1016/j.nantod.2020.100871
1748-0132/

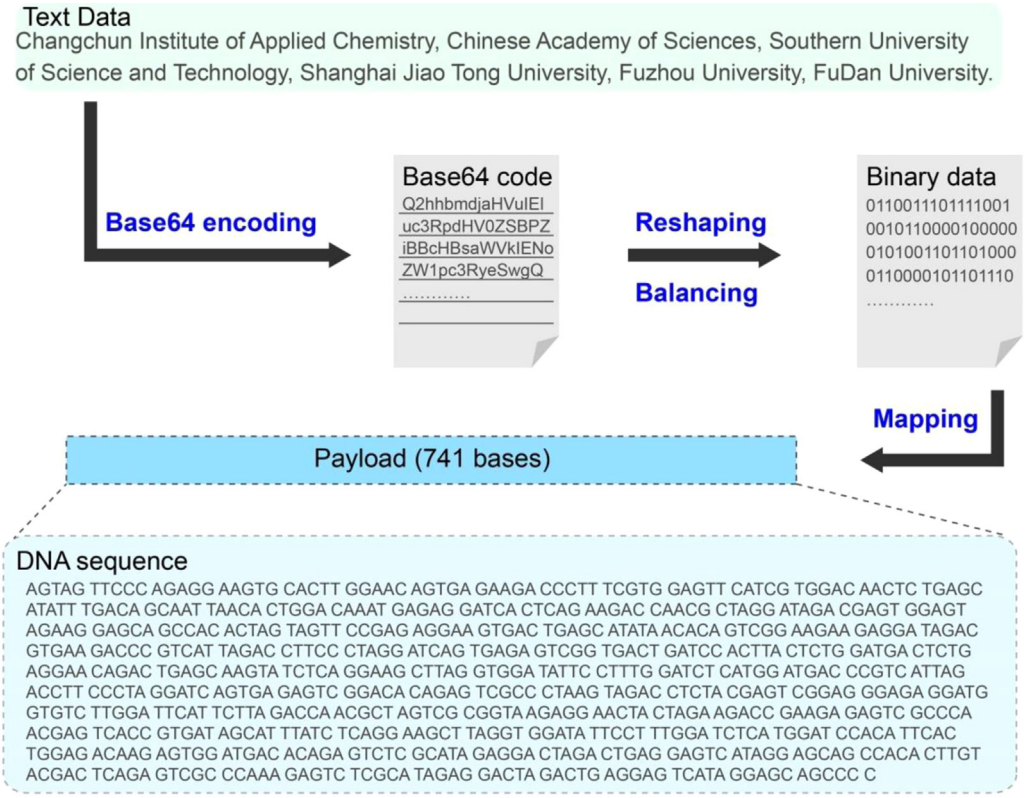


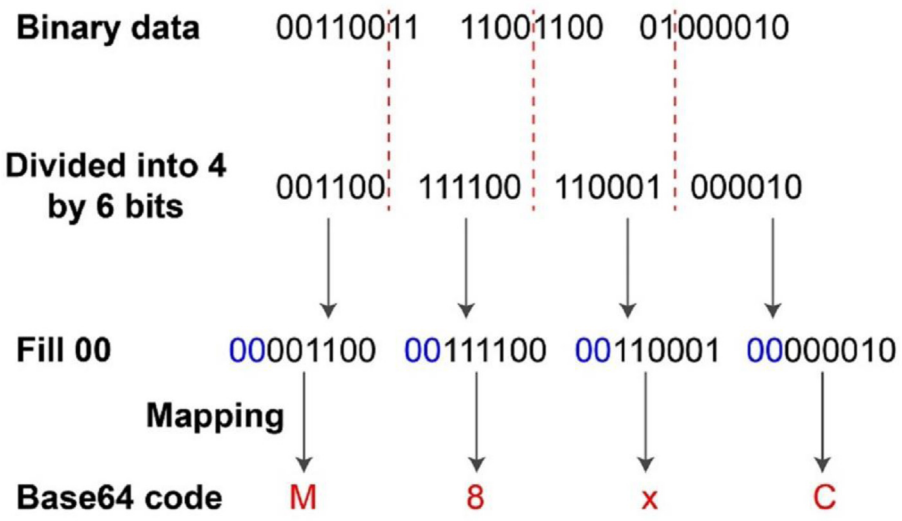
Fig. 1. Schematic of the text information encoding steps, including Base64 encoding, code reshaping and balancing, and data mapping. I) The names of five research institutions are converted into a specific Base64 code, which contains 64 different printable characters. II) The Base64 code is then reshaped and converted into two groups of 8-bit binary data, in which one group of the binary data is balanced by a specific code. III) The binary data and balance codes are finally mapped into a DNA sequence according to a customized mapping rule. The GC content in the DNA sequence is controlled about 50 % and continuous base repeat is also decreased by the balance algorithm, which would effectively reduce the error rate of sequencing.

sequence analysis and successfully demonstrated the storage and recovery of a ternary code. In 2012, Church [7] and Goldman [8] introduced addressing fountain block in DNA and realized the long bit stream to split into smaller information segments. This reduced the errors for DNA synthesis and sequencing, thus allowing for the storage of large scale digital data. Very recently, Erlich [9] and coworkers developed a new encoding strategy, called DNA Fountain. This enabled unlimited data retrieval and high physical density. These works have made great advances on the storage and recovery of digital data using DNA as a storage media and demonstrated its advantages, especially the fidelity and storage capacity. However, the applications of DNA storage also encounter several challenges, such as coding and decoding complicated algorithm [9–15], difficult information indexing and storage [16], the unpredicted error rates in DNA synthesis and sequencing [17–25], and so on. For example, the introduced error correction redundancy limits the utilization efficiency of DNA storage, leading to a final logical density less than 2 bits/nucleotide. In addition, to increase the error correction capacity, the algorithm gets more and more complicated, leading to the difficulties in DNA synthesizing and sequencing process. Innovations of new algorism, which reduces the recovery errors and increases the final logical density, remain a great challenge.

Here, we demonstrate a concise algorithm for DNA-mediated storage. Considering the expansion and complexity in digital data, Base64 instead of Base32 was chosen. Some random text information was selected as an example to be encoded in a long DNA sequence. Base64 encoding, code reshaping and balancing, and data mapping are involved in this process. Through a balance code, the developed algorithm led to the proper GC content and decreased continuous base repeat in the DNA strand. Thus the error rates were reduced in the process of DNA synthesis and sequencing. Interestingly, the encoded DNA information was stored in a circular plasmid platform. In the next step, decoding the DNA information was realized successfully via Illumina MiSeq sequencing. By this strategy, we achieved a high fidelity storage and recovery using a DNA macromolecule. A specific storage density of 1.77 bits/nucleotide has been realized, which is superior to other DNA-mediated storage systems. Therefore, the results offer new potentials for the development of DNA-mediated digital storage.

Experiments, results, and discussions

DNA digital data storage mainly involves four-steps: coding digital information into DNA sequences, synthesizing DNA sequences, storing DNA information, and recovering DNA information through sequencing. Limited by the DNA techniques, especially the DNA sequencing techniques, a proper coding algorithm is required to avoid continuous base repeat, that is homopolymers, e.g. TTTT (more than three same nucleotides in a row). The CG content should also be controlled properly. The Illumina MiSeq, the most accurate sequencing technique nowadays, could not give readout of DNA homopolymer sequences correctly. While high CG content may cause difficulty in the fracture of hydrogen bonding and denaturation, leading to the biased DNA sequencing coverage. Therefore, as demonstrated in Fig. 1, an encoding algorithm based on base64 code is developed to reduce the possibility of homopolymers and control the CG content about 50 %. The encoding algorithm includes three steps: I) converting the text information into Base64 code, which contains 64 different print-



A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Q	R	S	T	U	V	W	X	Y	Z	a	b	c	d	e	f
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v
32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47
w	x	y	z	0	1	2	3	4	5	6	7	8	9	+	/
48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63

Fig. 2. The pipeline of the Base64 encoding process. First, each 3 bytes of binary data are divided into 4 groups of 6 bits. Two "0"s are then filled in front of each group of 6 bits, forming a new 8-bit code. The 8-bit binary code is afterwards converted into Base64 code according to its decimal number and the coding table, which contains 64 printable characters. Therefore, the resulting Base64 data only contains 64 different characters, which is the prerequisite of code reshaping and balancing.

able characters; II) reshaping and converting the Base64 code into two groups of 8-bit binary data, in which one group is balanced by a specific code; III) the hybrid encoding is suitable to map the balance code and binary code into DNA sequences according to a customized mapping rule. The homopolymers and GC content in the DNA sequence are controlled by the balance code and the customized mapping rule, which effectively reduce the sequencing error rate.

The pipeline of the Base64 encoding process is shown in Fig. 2. The names of five research institutions are transformed into a stream of binary data using an online compiler (the website shown in Discussion S1). To encode the binary stream, each 3 bytes of binary data (total 24 bits) are divided into 4 groups, each group containing 6 bits. Two "0"s are then added at the front of each group, forming a new 8-bit code. The 8-bit binary codes are converted into individual characters according to its decimal number and the Base64 code table (Table S1). Therefore, each 3 bytes of data are encoded into 4 bytes of Base64 codes. The resulting Base64 data only contain 64 different characters and could reliably be transferred across different systems. In the encoding process, there is data amplification for filling "0" in the front of binary code. Base64 method generates less data amplification in comparison to base16 or base32. Moreover, the balanced sequences, which are constructed by permutating four "1"s and four "0"s is important to form a 8-bit balance binary code. In this context, only 70 arrangements could be built at the most, so we choose base64 method as the encoding method in this work.

The obtained Base64 codes are divided into groups of 7 characters, as shown in Fig. 3A. In each group, the 1 st, 3rd, and 5th characters are converted into 8-bit balance codes according to a customized balance code table (Table S2), in which each code contains four "1"s and four "0"s. The 2nd, 4th, 6th and 7th characters are converted back into 6-bit binary codes according to the Base64 code table, eliminating first two "0"s, and the 7th character are divided into 3 groups of 2 bits, each of which are added at the end of the 2nd, 4th and 6th characters, filling the 2nd, 4th, and 6th characters up to 8 bits, as demonstrated in Fig. 3B.

The binary code reshaped from the 2nd, 4th, 6th, and 7th characters and the balance code obtained from the 1 st, 3rd, and 5th characters are reorganized into two rows. Each column of the two rows is mapped into DNA sequences according to a customized mapping rule, i.e. converting {00, 10, 01, 11} to {A, T, C, G}, respectively, as shown in Fig. 3C. By this hybrid encoding, only when '1' appearing in the second row (balance code) it will make 'C' or 'G' be mapped, while the probability of '1' appearing in the balance code is 50 %. The introduction of the balance code makes the amplification of original files, but it leads to the waste of the nucleotide storage spaces. So, we just introduced one balance code in each mapping two rows, and the other was reshaped binary code. Comparing to the introduction of two balance code in each mapping group, one balance code and one binary code can save about 0.6 nucleotide/character.

To test the performance of the encoding algorithm, the input data of five research institutions names (185 bytes text information) is encoded into a long DNA sequence of 741 nucleotides using the above developed algorithm. The details of the coding program are shown in Discussion S1. Since the probabilities of A, T, G, and C are random, the probability of homopolymers is very low, which is directly proved by the obtained DNA sequence. EcoR I and Xho I restriction endonuclease sites are designed to be added at the

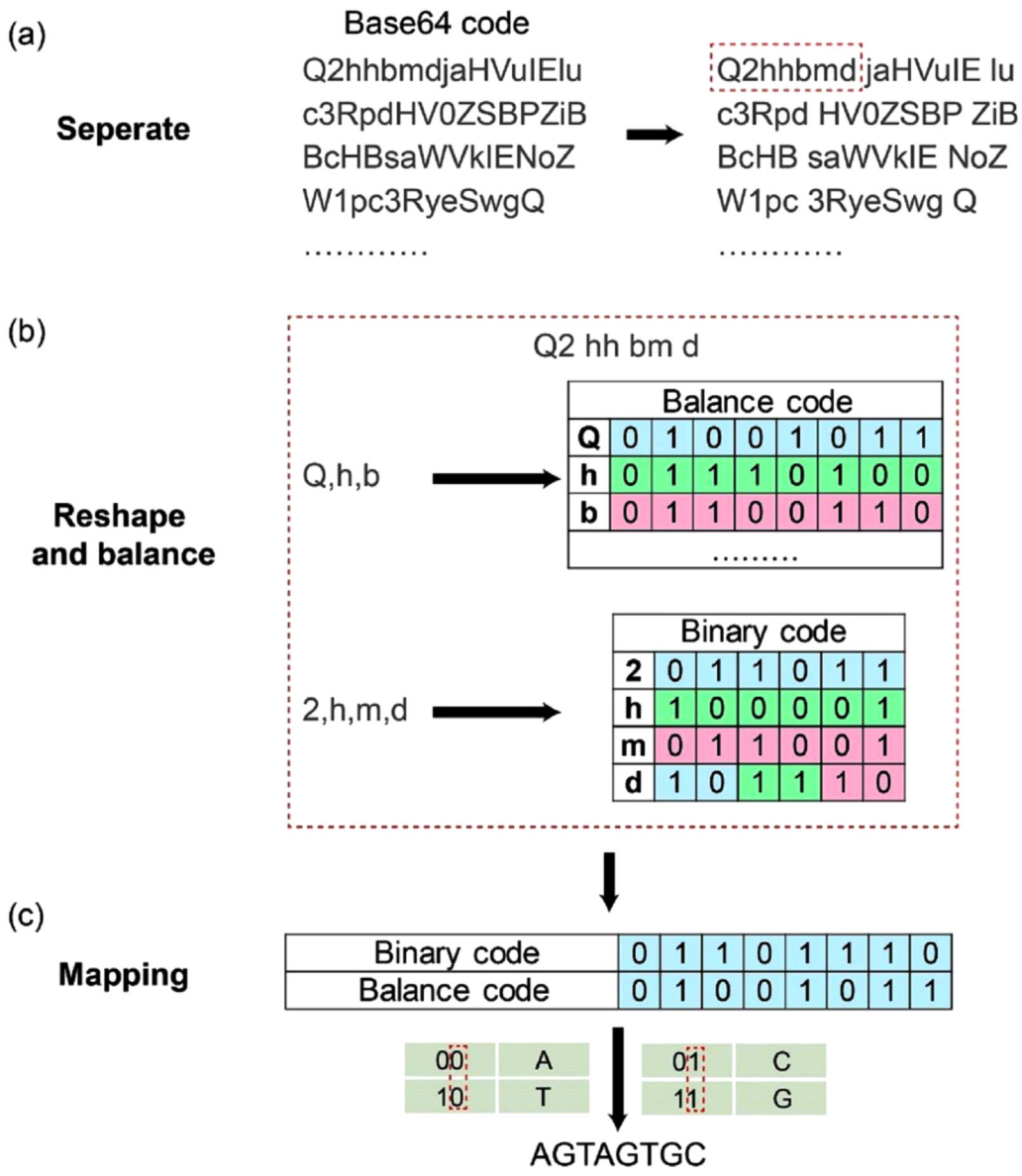


Fig. 3. Code reshaping and CG content balancing. (a) Each 7 characters in the Base64 code are divided into one group. (b) In each group, the 1 st, 3rd, and 5th characters are converted into 8-bit binary codes according to a customized balance coding table (Table S2), in which each code contains only four "1"s. The 2nd, 4th, 6th and 7th characters are converted into 6-bit binary codes according to Base64 code and the 7th character are divided into 3 groups of 2 bits, each of which are added at the end of the 2nd, 4th, and 6th characters, as suggested by their same color. Therefore, the 2nd, 4th, 6th, and 7th characters are reshaped into 3 groups of 8-bit data. (c) The binary code of the 2nd, 4th, 6th, and 7th characters and the balance code of the 1 st, 3rd, and 5th characters are then reorganized into two rows and each column of the two rows are mapped into DNA sequences by converting {00, 10, 01, 11} to {A, T, C, G}, respectively. According to the mapping rule, only when '1' appears in the balance code will 'C' or 'G' be mapped, while the probability of '1' appearing in the balance code is about 50 %.

two ends of the target DNA sequence, respectively, forming an inserting element of 753 nucleotides (Fig. 4). The DNA sequence is then prepared by solid state method and enzyme connection. After synthesis, the 753 nucleotides are inserted into a circular pGH plasmid by EcoRI and XhoI complementation sites, which serves as a storage platform and enhances the stability of the DNA sequence during storage. To read out the information from the DNA sequence, the whole plasmid is sequenced by Illumina MiSeq twice, one from M13(+) primer (forward sequencing) and the other from M13(-) primer (backward sequencing), which ensure the sequencing accuracy. We get total accurate 825 nucleotides from forward sequencing and 834 nucleotides from backward sequencing, both of which contain the whole payload section, as shown in Fig. S1. In comparison of the forward and backward sequencings, the two sequencing results show 100 % consistent DNA sequences and a specific density of 185*8/834 = 1.77 bits/nucleotide. The final GC content is 49.3 %, which locates in the ideal range of 45%–55%. Moreover, the introduction of plasmid as a platform for DNA digital data storage is also particularly suitable for highly efficient information replication in a form of exponential rate when it is transformed in a bacterium. To investigate the robustness of our strategy, we transformed this plasmid into E. coli and monitored the optical density of bacterial for 9 h (Fig. 4b). The S-shape curve demonstrated the robust plasmid in the system and the exponential growth replication in the period of 7 h. In comparison to other DNA mediated digital storage methods (Table S3), the present Base64 strategy exhibits attractive storage characteristics, especially the logical density is higher than other reported methods. We developed a specific method to prevent error instead of error correction, and it is suitable to release the storage space occupied by error correc-

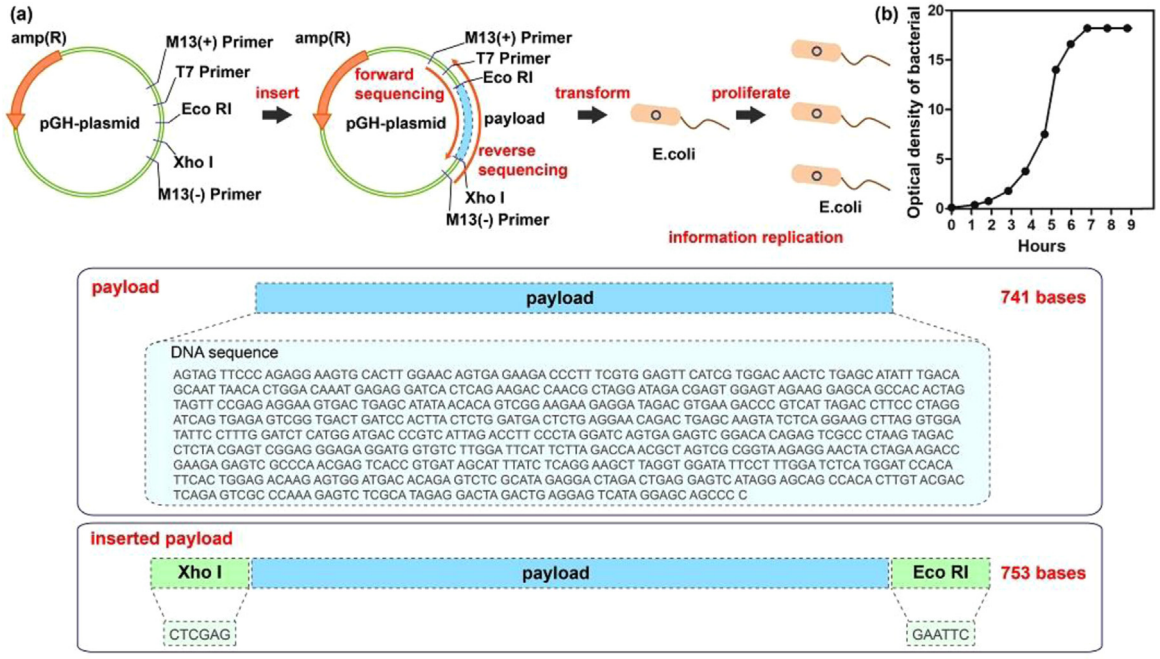


Fig. 4. The circular plasmid platform for the encoded DNA storage, sequencing and replication. (a) The encoding of five institute names produces a payload of 741 nucleotides. EcoRI and XhoI restriction endonuclease sites are added at the two ends of the 741 nucleotides payload, respectively, forming an inserting element of 753 nucleotides. The 753 nucleotides are inserted into pGH plasmid by EcoRI and XhoI complementation sites, enhancing the stability of the DNA coding during long-term storage. To read out the information from the DNA coding, the whole plasmid is sequenced by Illumina MiSeq twice, one from M13(+) primer and the other from M13(-) primer, to ensure sequencing accuracy. There are four "G" and "C" in every eight nucleotides. Comparing to the GC content in the whole strand, the proper content in every local part is also important to generate less errors. The even distribution of "G" and "C" can avoid biased DNA sequencing and prevent the local sequence error. The two sequencing results gave consistent coding sequences. The plasmid is transformed into E. coli and data replication by generational division is realized. (b) The plot of bacterial optical density verse time. The sample was measured at a wavelength of OD600 nm.

tion. The plasmid scaffold is important for the sequencing accuracy by the two direction sequencing, which is self-check. The length of information encoded DNA strand is extended, and meanwhile the addressing block is reduced in comparison to the short oligo-based strategy. In this context, every nucleotide in our approach can be used fully for information storage. Moreover, the plasmid-based scaffold is advantageous in sequencing accuracy, data stability, and data replication. Therefore, the Base64 algorithm is advantageous to ensure the high accuracy and high storage density for the DNA-mediated storage.

Conclusion

The paradigm of the Base64 algorithm was performed in the DNA molecules. Some random text information stored as binary streams in computer were mapped into DNA sequences using the Base64 encoding algorithm, which converted the binary data of "0" and "1" into the quaternary encoding nucleotides of A, T, C, and G. The hybrid encoding consisting of balance codes and reshaped binary codes, acted as an error-proofing strategy, which is suitable to control the CG content about 50 % and reduce the repeat of continuous base in the DNA sequence. A circular plasmid was introduced as a scaffold for the long-term storage of DNA-mediated digital data. The forward and backward sequencing results gave 100 % consistence of the encoded DNA sequence. A high specific logical density of 1.77 bits/nucleotide was realized, that is comparable or even higher than other reported methods. Therefore, a robust strategy for DNA-mediated digital data storage was constructed on the basis of Base64 algorithm, and was also demonstrated by the data-loaded bacterial proliferation. Equally important, the developed method is compatible for both text and graphic information storage, and it is also useful for video and audio files. Thus it offers great potentials for practical digital storage and other technical applications. Extension of the existed four nucleotides, A, T, C, and G, to six and eight nucleotides, including synthetic nucleotides [26–28], would further significantly increase the DNA-based storage density. A theoretical maximum logical storage density of 3 bits/nucleotide might be realized. Regarding the data readout in unnatural DNA molecules, advances in nanopore-based sequencing technique is also important [29,30]. In addition, the development of new indexing methods, such as the introduction of engineered framework of DNA molecules [31–36], will further facilitate the high-density storage in multiple dimensions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is supported by National Key R&D Program of China (Grant No. 2018YFA0902600), the National Natural Science Foundation of China (Grant Nos. 21704099, 21877104, 21834007, and 21878258), K. C. Wong Education Foundation (Grant No. GJTD-2018-09), and Dong Chen is also grateful for support by the Open Funds of the State Key Laboratory of Rare Earth Resource Utilization, Changchun Institute of Applied Chemistry, CAS (Grant No. RERU2019008).

References

[1] L. Ceze, J. Nivala, K. Strauss, Nat. Rev. Genet. 20 (2019) 456–466, http://dx.doi.org/10.1038/s41576-019-0125-3.
[2] J.P.L. Cox, Trends Biotechnol. 19 (2001) 247–250, http://dx.doi.org/10.1016/s0167-7799(01)01671-7.

[3] V. Zhirnov, R.M. Zadegan, G.S. Sandhu, G.M. Church, W.L. Hughes, Nat. Mater. 15 (2016) 366–370, http://dx.doi.org/10.1038/nmat4594.
[4] R. Heckel, Nat. Biotechnol. 36 (2018) 236–237, http://dx.doi.org/10.1038/nbt.4093.
[5] C. Bancroft, T. Bowler, B. Bloom, C.T. Clelland, Science 293 (2001) 1763–1765.
[6] C.T. Clelland, V. Risca, C. Bancroft, Nature 399 (1999) 533–534, http://dx.doi.org/10.1038/21092.
[7] G.M. Church, Y. Gao, S. Kosuri, Science 337 (2012) 1628, http://dx.doi.org/10.1126/science.1226355.
[8] N. Goldman, P. Bertone, S. Chen, C. Dessimoz, E.M. LeProust, B. Sipos, E. Birney, Nature 494 (2013) 77–80, http://dx.doi.org/10.1038/nature11875.
[9] Y. Erlich, D. Zielinski, Science 355 (2017) 950–953, http://dx.doi.org/10.1126/science.aaj2038.
[10] L. Organick, S.D. Ang, Y.J. Chen, R. Lopez, S. Yekhanin, K. Makarychev, M.Z. Racz, G. Kamath, P. Gopalan, B. Nguyen, C.N. Takahashi, S. Newman, H.Y. Parker, C. Rashtchian, K. Stewart, G. Gupta, R. Carlson, J. Mulligan, D. Carmean, G. Seelig, L. Ceze, K. Strauss, Nat. Biotechnol. 36 (2018), http://dx.doi.org/10.1038/nbt.4079, 242-+.
[11] S.M.H.T. Yazdi, R. Gabrys, O. Milenkovic, Sci. Rep. 7 (2017), http://dx.doi.org/10.1038/s41598-017-05188-1.
[12] G.C. Smith, C.C. Fiddes, J.P. Hawkins, J.P.L. Cox, Biotechnol. Lett. 25 (2003) 1125–1130, http://dx.doi.org/10.1023/a:1024539608706.
[13] S. Yazdi, Y.B. Yuan, J. Ma, H.M. Zhao, O. Milenkovic, Sci. Rep. 5 (2015) 10, http://dx.doi.org/10.1038/srep14138.
[14] M. Blawat, K. Gaedke, I. Huetter, C. Xiao-Ming, B. Turczyk, S. Inverso, B. Pruitt, G. Church, Procedia Comput. Sci. (Netherlands) 80 (2016) 1011–1022, http://dx.doi.org/10.1016/j.procs.2016.05.398.
[15] M. Ailenberg, O.D. Rotstein, Biotechniques 47 (2009) 747–751, http://dx.doi.org/10.2144/000113218.
[16] R.N. Grass, R. Heckel, M. Puddu, D. Paunescu, W.J. Stark, Angewandte Chemie-International Edition 54 (2015) 2552–2555, http://dx.doi.org/10.1002/anie.201411378.
[17] L. Anavy, I. Vaknin, O. Atar, R. Amit, Z. Yakhini, Nat. Biotechnol. 37 (2019), http://dx.doi.org/10.1038/s41587-019-0240-x, 1229-+.
[18] S. Newman, A.P. Stephenson, M. Willsey, B.H. Nguyen, C.N. Takahashi, K. Strauss, L. Ceze, Nat. Commun. 10 (2019) 6, http://dx.doi.org/10.1038/s41467-019-09517-y.
[19] S. Kosuri, G.M. Church, Nat. Methods 11 (2014) 499–507, http://dx.doi.org/10.1038/nmeth.2918.
[20] Y. Erlich, Genome Res. 25 (2015) 1411–1416, http://dx.doi.org/10.1101/gr.191692.115.
[21] A.V. Bryksin, I. Matsumura, Biotechniques 48 (2010) 463–465, http://dx.doi.org/10.2144/000113418.
[22] H. Li, R. Durbin, Bioinformatics 25 (2009) 1754–1760, http://dx.doi.org/10.1093/bioinformatics/btp324.
[23] R.C. Edgar, Nucleic Acids Res. 32 (2004) 1792–1797, http://dx.doi.org/10.1093/nar/gkh340.
[24] H.H. Lee, R. Kalhor, N. Goela, J. Bolot, G.M. Church, Nat. Commun. 10 (2019) 12, http://dx.doi.org/10.1038/s41467-019-10258-1.
[25] C.N. Takahashi, B.H. Nguyen, K. Strauss, L. Ceze, Sci. Rep. 9 (2019) 5, http://dx.doi.org/10.1038/s41598-019-41228-8.
[26] Y.J. Seo, G.T. Hwang, P. Ordoukhanian, F.E. Romesberg, J. Am. Chem. Soc. 131 (2009) 3246–3252, http://dx.doi.org/10.1021/ja807853m.
[27] M.M. Georgiadis, I. Singh, W.F. Kellett, S. Hoshika, S.A. Benner, N.G.J. Richards, J. Am. Chem. Soc. 137 (2015) 6947–6955, http://dx.doi.org/10.1021/jacs.5b03482.
[28] S. Hoshika, N.A. Leal, M.J. Kim, M.S. Kim, N.B. Karalkar, H.J. Kim, A.M. Bates, N.E. Watkins, H.A. SantaLucia, A.J. Meyer, S. DasGupta, J.A. Piccirilli, A.D. Ellington, J. SantaLucia, M.M. Georgiadis, S.A. Benner, Science 363 (2019), http://dx.doi.org/10.1126/science.aat0971, 884-+.
[29] M. Jain, H.E. Olsen, B. Paten, M. Akeson, Genome Biol. 17 (2016) 11, http://dx.doi.org/10.1186/s13059-016-1103-0.
[30] K.K. Chen, J.L. Kong, J.B. Zhu, N. Ermann, P. Predki, U.F. Keyser, Nano Lett. 19 (2019) 1210–1215, http://dx.doi.org/10.1021/acs.nanolett.8b04715.
[31] W. Bae, S. Kocabey, T. Liedl, Nano Today 26 (2019) 98–107, http://dx.doi.org/10.1016/j.nantod.2019.03.001.
[32] X.G. Liu, F. Zhang, X.X. Jing, M.C. Pan, P. Liu, W. Li, B.W. Zhu, J. Li, H. Chen, L.H. Wang, J.P. Lin, Y. Liu, D.Y. Zhao, H. Yan, C.H. Fan, Nature 559 (2018) 593–598, http://dx.doi.org/10.1038/s41586-018-0332-7.
[33] H.L. Zhang, J. Chao, D. Pan, H.J. Liu, Y. Qiang, K. Liu, C.J. Cui, J.H. Chen, Q. Huang, J. Hu, L.H. Wang, W. Huang, Y.Y. Shi, C.H. Fan, Nat. Commun. 8 (2017) 7, http://dx.doi.org/10.1038/ncomms14738.
[34] J. Li, A.A. Green, H. Yan, C.H. Fan, Nat. Chem. 9 (2017) 1056–1067, http://dx.doi.org/10.1038/nchem.2852.
[35] Y.Y. Zhang, F. Li, M. Li, X.H. Mao, X.X. Jing, X.G. Liu, Q. Li, J. Li, L.H. Wang, C.H. Fan, X.L. Zuo, J. Am. Chem. Soc. 141 (2019) 17861–17866, http://dx.doi.org/10.1021/jacs.9b09116.
[36] Y. Zhao, X.P. Dai, F. Wang, X.L. Zhang, C.H. Fan, X.G. Liu, Nano Today 26 (2019) 123–148, http://dx.doi.org/10.1016/j.nantod.2019.03.004.