INNOVATION SPOTLIGHT:

Synbio Technologies
Genes for Life

DNA Data Storage

BCC Research

Report Code: BIO213A

DNA Data Storage

We're creating more data than we can possibly handle with our current storage technologies. To manage it, we are going back to the universe's original data filing system.

In 2012, IBM famously posted a statistic: Approximately 90% of all existing digital data in the world had been generated in the prior 2 years. That was true then, and statistically remains true in 2021—every 2 years humans produce 10 times more data than they did in the previous 2 years.

This rate of data generation is constantly accelerating due to the growth of "big" computing (governmental, military, industrial, transportation and academic) and civilian computing, namely the meteoric rise of search engines, social media and connected devices of all kinds. Every day Google receives 3.5 billion search requests. WhatsApp users exchange 65 billion messages. This explosion in data creates demand for efficient, high-density storage systems.

Currently, data is stored on magnetic tapes, hard drives, flash drives and optical discs. All of these technologies are nearing their specific channel limits, all can be easily damaged, and all physically degrade over time. The looming global data crunch demands new ways of storing information, and the answer may just be in our own genetic information.

DNA data storage is the process of encoding and decoding binary data to and from synthesized strands of DNA or RNA. It has the potential to exponentially increase data storage capabilities, and can also theoretically be preserved indefinitely. As of this writing, high cost and slow read and write times are keeping this from the mainstream. But, like all other information storage hurdles in the past, DNA data storage will get smarter, cheaper and faster.

Dr. Hanlee Ji, Associate Professor of Medicine, Stanford University, sums up the reasons for DNA as a storage medium: "It is estimated that 44 trillion gigabytes of data were generated in 2020. In the era of big data, more is considered better and what that practically means is that there are huge amounts of data that need to be stored, managed and potentially analyzed. If you look at the supply of current microchips, this amount of data clearly exceeds the supply that we can currently generate based on microchip-grade silicon. This represents a potential dilemma for maintaining humanity's history.

DNA is extraordinarily stable. DNA can retain its information through the course of decades if not a lot longer. And we can manipulate DNA with increasing finesse and increasing scale. So, it's kind of been a logical point and topic to consider in terms of new types of data storage technology."

Summary Table:
Global Market for DNA Data Storage, by End Use Industry, Through 2025
($ Millions)

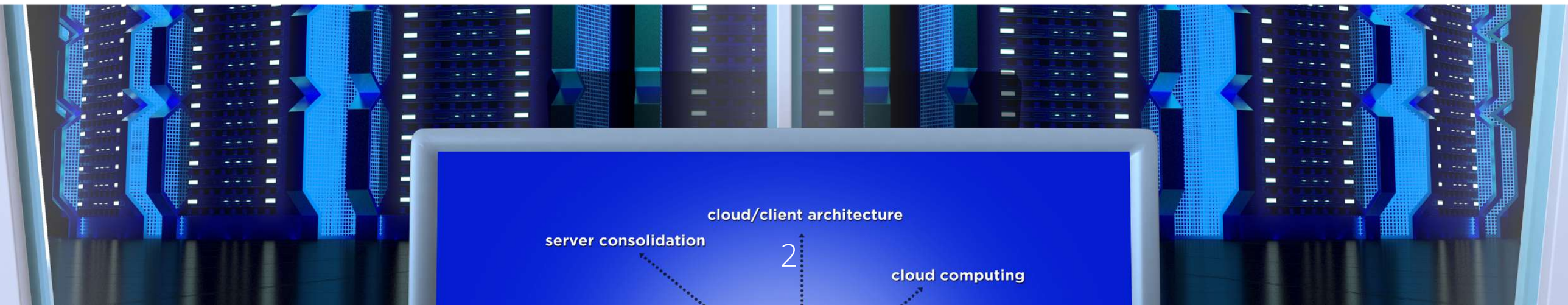
End Use Industry	2019	2020	2025	CAGR% 2020–2025
Government/defense	25.2	36.4	153.9	33.4
Non-government	---	---	371.4	---
Total	25.2	36.4	525.3	70.6

Source: BCC Research

BCC sat down with Synbio Technologies, Inc. Synbio develops and markets products and services for DNA reading, writing and editing. Products and services include: NGS, Sanger sequencing and immunosequencing (DNA reading); gene synthesis, cloning, pathway/genome synthesis, synthetic DNA libraries (DNA writing); and CRISPR-Cas9 sgRNA design, library construction and genome editing (DNA editing).

Synbio markets its Synotype platform, which is an integrated system for translation or reverse translation between genotype and phenotype. The company has expertise in synthetic biology technologies, including DNA engineering; DNA synthesis; genome synthesis; pathway synthesis; pharmacogenomics; microbiology; and translational biology.

In July 2017 Synbio Technologies partnered with Twist Bioscience covering the supply of long-length genes up to 70 kilobases. Twist Bioscience will manufacture synthetic DNA up to 3.2 kilobases in length, and supply to Synbio Technologies, who will create genes up to 70 kilobases in length.



BCC's Interview With Synbio Technologies, Inc.

BCC: How and why did Synbio Technologies come into existence?

Synbio: In the recent decade, many unprecedented breakthroughs including the de novo synthesis of working microorganism genomes; translation of CRISPR into precision diagnostics & gene therapy; mega base pairs scale metabolic engineering; digital information DNA storage; COVID-19 diagnostics and vaccines; etc. started the era of synthetic biology.

Synbio Technologies was incorporated in 2013 by a group of passionate biologists with a single mission in mind: to empower scientific discovery and innovation by providing advanced DNA manufacturing and design. We are committed to consistently provide effective DNA solutions for design and synthesis that serve as the fundamental platforms of synthetic biology.

BCC: What is Synbio's role (so far) in the larger ecosystem of DNA data storage?

Synbio: Synbio Technologies' role is to help in manufacturing (i.e. synthesizing the digitalized DNA that will be used in storage) and working to develop efficient algorithms that will help expedite the translation from digital information into DNA.

BCC: What do you supply, and to what kind of clients/partners?

Synbio: As a DNA engineer and manufacturer, we provide DNA synthesis, sequencing, editing and various combinational DNA solutions to a variety of challenges, such as de novo antibody discovery, vaccine design and discovery, etc. We serve customers from bio-pharma, biotech and research universities.

BCC: What, in simple terms, is the Synotype platform, and what does it bring to the market?

Synbio: We know that a genotype is an organism's complete set of genetic material derived from nature. But, what about a set of artificial DNA that is de novo synthesized? Essentially, this is what Synotype represents. A Synotype contains unnatural DNA sequences with a novel phenotype/feature. It is the footprint and result of the synthetic biology cycle of Design-Build-Test-Learn (DBTL). Synthetic biology products with improved implication are obtained from this loop.

BCC: What is your organization/lab environment like, in terms of size and productivity, but also of culture?

Synbio: While our physical laboratory mirrors many other biopharma/bio-tech companies in terms of equipment, instruments, etc., what makes Synbio Technologies unique is our talented set of employees. New Jersey provides a huge pool of talent of people within the life sciences and biotech industry. As a company that prioritizes synthetic biology production, we are very lucky and grateful to see so many scientists bringing in a variety of novel ideas and products within this growing field. Our company's core values consist of integrity, profession, transparency and action. We are united not only under our love and passion for genes, but under our shared belief of the importance of each of these values in fostering a successful work culture.

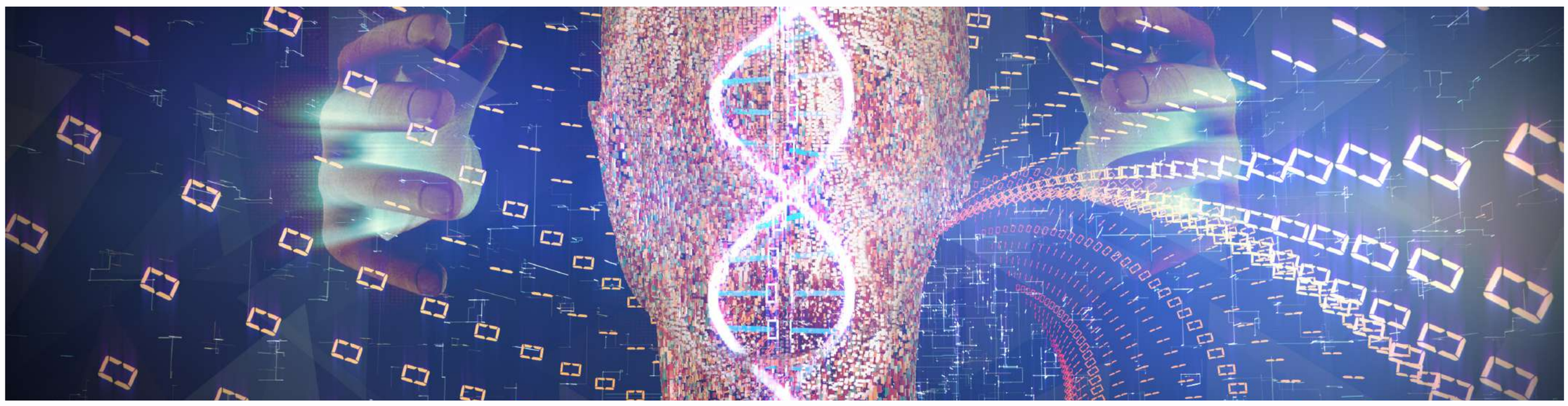
BCC: Based on all you've experienced to now, where do you see DNA data storage in 10 to 15 years? Will it be the norm?

Synbio: It is very possible that DNA writing (synthesis) will start to follow Moore's law. In other words, the number of base pairs on a microchip will double every 18 months while the cost is halved. We hypothesize this because synthesis technology continues to rapidly advance as the main driving force behind DNA data storage. Products such as portable sequencers will be coming soon and will make data retrieving easier and faster, making using DNA as a form of digital information storage even more feasible and accessible.

BCC: How did COVID affect your operations and your company's mission?

Synbio: The COVID-19 pandemic left no one unscathed and proved to be a huge challenge for all of us. Fortunately, this challenge created an opportunity for the DNA industry. Specifically, Synbio Technologies has been working tirelessly with our partners to design and manufacture various products, including but not limited to DNA probes, various antibodies, DNA testing kits, etc. The necessity of DNA technology does not stop at COVID-19; it has and will continue to be used in exploring the old and new mysteries and questions of our world. As summarized by our motto: "Genes for life."

Technology advances in phosphoramidite-based synthesis are helping to propel the growth in the market. For example, Twist Bioscience combines phosphoramidite-based chemistry with a proprietary silicon support, then uses its expertise in miniaturization, engineering, fluid mechanics, robotics and laboratory information management systems to automate DNA synthesis with high precision and scale. Twist is a leading supplier of DNA for data storage applications.



What Does BCC See Ahead?

The DNA data storage industry is at the beginning stages of its growth, but there are tremendous opportunities in front of it. The industry is developing its own critical mass of enabling technologies (e.g., DNA sequencing, DNA synthesis, coding) and multidisciplinary initiatives (e.g., the Molecular Information Storage (MIST) program, the OligoArchive project, the DNA Data Storage Alliance, etc.). These forces are helping significant milestones to be reached and are driving commercial achievements in the industry, giving it a pathway to future adoption and growth.

Global Trends You Can Bank On

- The overall global market was $36.4 million in 2020 and is growing at a compound annual growth rate (CAGR) of 70.6% to reach a forecast market size of $525.3 million by 2025. The DNA data storage market can be segmented by the DNA synthesis technology employed: enzymatic or chemical (phosphoramidite) based.

- The phosphoramidite-based (chemical) synthesis market was $28.4 million in 2020 and is forecast to grow at a CAGR of 68% to reach $380.3 million by 2025.

- The enzymatic-based synthesis market was $8 million in 2020 and is growing at a CAGR of 78.5% to reach a forecast value of $145 million by 2025.

- The primary end users for DNA data storage systems in 2020 are government/defense, with multiple research and prototyping projects. This market was valued at $36.4 million in 2020 and is growing at a CAGR of 33.4% to reach $153.9 million by 2025. A key project in this market segment is the Intelligence Advanced Research Project Agency's (IARPA) and the MIST program, which is funding two teams of collaborators in an effort to develop and build a DNA data storage device that can store 1 terabyte of information in 24 hours at a low marginal cost.

- The non-government portion of the market includes data storage applications in multiple industries, including banking and financial services, healthcare and pharma, media and entertainment and nuclear, among others. This market segment is forecast to account for $371.4 million by 2025.

- Key applications for these two market segments include cold data storage archiving, where the data does not need to be accessed and read on a frequent basis. DNA-based data storage is able to compete favorably with magnetic tape-based storage in these applications because the read frequency is not so high and because DNA media, by virtue of its being central to all living systems, will not become obsolete, potentially lasting for hundreds if not thousands of years. This ensures that valuable archival data will be kept by humankind for future generations.

Related BCC Research Reports

DNA Data Storage: Global Markets and Technologies
Next-Generation Sequencing: Emerging Clinical Applications and Global Markets
Proteomics: Technologies and Global Market
Single-Cell Genomics and Proteomics: Emerging Technologies and Markets
Single Cell Omics: Emerging Technologies and Markets
Global DNA Sequencing: Research, Applied and Clinical Market
Global Biochip Markets: Microarrays and Lab-on-a-Chip
Biologic Imaging Reagents: Technologies and Global Markets
Synthetic Biology: Global Markets
Next Generation Cancer Diagnostics: Technologies and Global Markets
Liquid Biopsy Research Tools, Services and Diagnostics: Global Markets
Genetic Modification Therapies Clinical Applications
Global DNA Read, Write and Edit Market

50 years
BCC Research
Smart Decisions Start Here