AUTOMATED DATA STORAGE SYSTEM

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefits of priority to U.S. Provisional Patent Application No. 63/492,593, filed March 28, 2023, and U.S. Provisional Patent Application No. 63/514,227, filed July 18, 2023, the entireties of which are incorporated herein by reference. All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BACKGROUND

[0002] DNA is a compelling data storage medium given its superior density, stability, energy-efficiency, and longevity compared to current storage methods. However, automated systems may contain a number of different components execute steps such as encoding, writing/synthesizing, storing, and reading/sequencing. Therefore, there is a need to develop systems which can perform these functions.

SUMMARY

[0002a] In an aspect, the present disclosure refers to a system for data storage comprising:

a computing system comprising at least one processor and instructions executable by the at least one processor to perform one or more operations comprising:

receiving digital information;

encoding digital information in one or more nucleic acid sequences; and

synthesizing a library of polynucleotides corresponding to the nucleic acid sequences; and

a modular rack-mount synthesis unit, wherein the modular rack-mount synthesis unit comprises:

a computer controller;

one or more reservoirs;

at least one flow cell block comprising one or more flow cells;

an extraction stage;

a post-processing unit;

a storage unit;

a robotic interference; and/or

a rack interface.

[0003] Provided herein are systems for data storage comprising a computing system comprising at least one processor and instructions executable by the at least one processor to perform one or more operations, and a modular rack-mount synthesis. In some examples, the one or more operations comprises: receiving digital information; encoding digital information in one or more nucleic acid

sequences; and synthesizing a library of polynucleotides corresponding to the nucleic acid sequences. In some examples, the modular rack-mount synthesis unit comprises: a computer controller; one or more reservoirs; at least one flow cell block comprising one or more flow cells; an extraction stage; a post-processing unit; a storage unit; a robotic interference; and/or a rack interface. In some examples, the at least one flow cell block comprises at least 12 flow cells. In some examples, the one or more flow cells comprise: one or more solid supports configured for synthesizing the library of polynucleotides; and one or more ports for exchanging gases, synthesis reagents, and/or extracted polynucleotides.

[0004] In some examples, the modular rack-mount synthesis unit comprises the extraction stage and/or the post-processing unit, and the extraction stage and/or the post-processing unit comprises: an energy supply chain; and one or more ports for exchanging synthesis reagents and/or extracted polynucleotides. In some examples, the modular rack-mount synthesis unit comprises the storage unit, and the storage unit comprises: a storage plate; a robotic interface configured to move storage plates; and an access port. In some examples, the modular rack-mount synthesis unit comprises the rack interface, and the rack interface comprises: a power supply; a rack reagent bulkhead; a synthesis unit reagent bulkhead; one or more reservoirs; an energy supply chain; and a pilot valve bank. In some examples, the storage unit provides the library of polynucleotides to a sequencing unit configured to sequence the library of polynucleotides.

[0005] In some examples, the systems for data storage further comprise one or more sensors configured to: detect a status, wherein the status comprises a temperature, a pressure, a humidity, a change in voltage, a

change in current, a capacitance, a conductivity, a storage plate position, a liquid volume, a flow rate, and/or a presence of a liquid; and report the status to the computing system. In some examples, the modular rack-mount synthesis unit comprises the at least one flow cell block, and the one or more sensors are further configured to pause synthesis at the one or more flow cells or the at least one flow cell block. In some examples, the system is a first system for data storage, and wherein the computing system is configured to divert encoded data from (i) a paused flow cell or (ii) a paused flow cell block to an active flow cell, an active flow cell block, or to a second system for data storage.

[0006] In some examples, the computing system further comprises a cache, wherein digital information is retained in the cache until polynucleotides encoding the digital information are stored.

[0007] Further provided herein are assemblies for storing information comprising: a plurality of compartments comprising a first material, wherein the plurality of compartments is configured to receive a plurality of polynucleotides encoding information; a base plate comprising a second material; and a cover plate comprising a third material, wherein a coefficient of thermal expansion (CTE) of the first material, the second material, or both, is less than the third material. In some examples, the plurality of compartments are positioned in an array on the base plate. In some examples, the base plate, the cover plate, or both, comprise a plurality of recessed features, and wherein each compartment is located at least partially within a recessed feature. In some examples, the first material comprises borosilicate. In some examples, the second material has a specific heat capacity of about 0.5 J/k-°C to 2.5 J/k-°C. In some examples, the third material includes: a CTE of about 15 μm/m-°C to 20 μm/m-°C; a thermal conductivity of about 15 W/m-K to about 20 W/m-K; and/or a specific heat capacity of about 0.5 J/g-°C. In some examples, the third material comprises stainless steel.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] A better understanding of the features and advantages of the present subject matter will be obtained by reference to the following detailed description that sets forth illustrative embodiments and the accompanying drawings of which:

[0009] **FIG. 1A** is a flowchart depicting a workflow for an exemplary automated data storage system using nucleic acids, including synthesis modules, according to aspects of the present disclosure.

[0010] **FIG. 1B** is a flowchart depicting a workflow for automated data storage system using nucleic acids, including storage modules. In some instances, these steps are performed by a robotics stage, according to aspects of the present disclosure.

[0011] **FIG. 1C** is a flowchart depicting a workflow of an exemplary system architecture at the level of a server rack (as described herein) for automated data storage system using nucleic acids including firmware, control systems (such as a network interface controller) and virtual data path, according to aspects of the present disclosure. In some aspects, data may move to the server rack for storage and validation, and the server rack may send data to an external interface for database management.

[0012] **FIG. 2A** is an illustration depicting an exemplary standardized DNA synthesis unit (DSU) with data/addressing, power, and reagent/DNA input/removal (a flow cell, flow cell block, or entire data storage unit), including a table describing dimensions and chip density for exemplary 6U module,