The Emerging Field of Synthetic Biology: a Review

Narjeskhatoon Habibi
Soft Computing Research Group, Faculty of Computer Science and Information systems, Universiti Teknologi Malaysia, Malaysia.
Siti Zaiton Mohd Hashim
Soft Computing Research Group, Faculty of Computer Science and Information systems, Universiti Teknologi Malaysia, Malaysia.
Cesar A. Rodriguez
Genome Compiler Corporation, USA.
Mohd Saberi Bin Mohamad
Artificial Intelligence and. Bioinformatics Laboratory, Faculty of Computer Science and Information systems, Universiti Teknologi Malaysia, Malaysia.
Safaai Bin Deris
Artificial Intelligence and. Bioinformatics Laboratory, Faculty of Computer Science and Information systems, Universiti Teknologi Malaysia, Malaysia.

***Abstract*-Synthetic biology which is concerned with engineering organisms to perform novel functions and developing the ways by which organisms can be engineered easily and robustly, is a relatively young field with a large potential for growth. Much of the work in the field involves design and construction of genetic circuitry, the piecing together of biological "parts", such as promoters and ribosome binding sites, to form the basis for biological "devices" which are organisms that have an engineered, well-specified input-output behavior. The idea in synthetic biology is applying engineering principles, such as hierarchical design, modular reusable parts and standard interfaces to construct bio-systems. The results of these engineering efforts can be of great value to human interests such as medicine and industry. In this paper first some related terms and resources for synthetic biology are described. Then existing computational tools and methods for synthetic biology are explored. It is followed by a discussion on current challenges in the field and future directions and finally by conclusion.**

I. INTRODUCTION

Synthetic biology which is concerned with engineering organisms to perform novel functions and developing the ways by which organisms can be engineered easily and robustly, is a relatively young field with a large potential for growth. Much of the work in the field involves design and construction of genetic circuitry, the piecing together of biological "parts", such as promoters and ribosome binding sites, to form the basis for biological "devices" which are organisms that have an engineered, well-specified input-output an behavior.

Synthetic biology borrows wet-lab techniques from genetic engineering, modeling techniques from systems biology and design concepts from electrical and control engineering. The idea is applying engineering principles, such as hierarchical design, modular reusable parts and standard interfaces to construct bio-systems.

Oscillators, bistable switches, noise controlling networks and synchronizing circuits, are examples of bio-systems constructed by synthetic biologists.

The results of these engineering efforts can be of great value to human interests such as medicine and industry. In the future synthetic biology may become a key player in bioremediation, drug production and bio-fuel production [1].

In this chapter first some related terms and resources for synthetic biology are described. Then existing computational tools and methods for synthetic biology are explored. It is followed by a discussion on current challenges in the field and future directions and finally by conclusion.

II. BACKGROUND MATERIAL

In this section some related terms, concepts and resources for synthetic biology are described.

A. Abstraction Hierarchy

Abstraction hierarchies assist researchers to engineer very complex systems by ignoring

978-1-4577-1967-7/12/$26.00 ©2011 IEEE

unnecessary details. Abstraction hierarchy in synthetic biology allows the designer of a biological system to ignore some of the implementation details and focus only on the high-level design issues. The most common hierarchy in synthetic biology, which is depicted in Fig. 1, includes DNA, part, device and system [2]:

DNA

DNA forms the basic unit and everything else is built on it.

Part

Parts are non-reducible elements of genetic composition (e.g. promoters, ribosome binding sites and open reading frames).

Devices/circuit

A device includes one or more parts. It receives an input signal and responds to it by producing the appropriate output signal (e.g. invertor).

System

A system contains several devices (e.g. bio-sensors).

B. DNA Assembly

DNA assembly is building of larger DNA fragments from smaller ones which is one of the most fundamental techniques in synthetic biology.

C. Assembly Protocols/Standards

These are different proposed wet-lab methods for DNA assembly. Examples are BioBrick assembly, overlap methods, Gibson, Golden Gate, SLIC, SOEing and EIPCR.

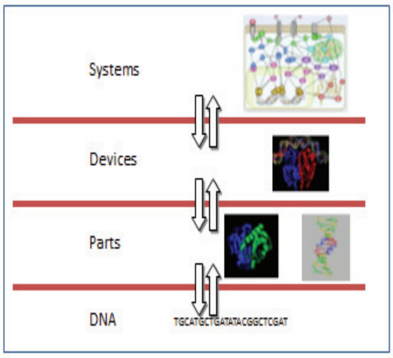

Figure 1. Abstraction hierarchy in synthetic biology

D. MIT Registry of Standard of Biological Parts

This database is a continuously growing set of genetic parts (BioBricks) that can be assembled to build synthetic biology devices and systems. It was founded in 2003 at MIT and provides genetic parts to iGEM teams and academic labs (http://partsregistry.org/).

E. BioBrick

BioBrick standard biological parts are DNA sequences with defined structure and function. Their special design allows assembling them to build larger and larger BioBrick devices and incorporating into living cells such as E. coli to construct new bio-systems. (http://biobricks.org/)

F. BIOFAB

The BIOFAB (International Open Facility Advancing Biotechnology) was founded in December 2009 as the world's first biological design-build facility. BIOFAB projects will be designed to produce useful collections of standard biological parts that can be made freely available to both academic and commercial users, while also enabling the rapid design and prototyping of genetic constructs needed to support specific needs (http://biofab.org/).

G. SBOL

Synthetic Biology Open Language (SBOL) is a language for the description and the exchange of synthetic biological parts, devices, and systems (http://www.sbolstandard.org/).

H. The International Genetically Engineered Machine (iGEM)

It is a worldwide synthetic biology competition for undergraduate university students. Student teams are given a kit of biological parts at the beginning of the summer from the Registry of Standard Biological Parts. They work at their own schools over the summer and use the kit and their own new parts to build biological systems and operate them in living cells (http://igem.org/).

I. OpenWetWare (OWW)

OWW provides a place (in Wiki format) for labs, individuals, and groups in biology, biological engineering and synthetic biology fields to organize their information and collaborate with others easily and efficiently (http://openwetware.org/).

III. COMPUTATIONAL TOOLS FOR SYNTHETIC BIOLOGY

Computational tools are necessary in any modern engineering discipline to assist the design process. They provide compositions of parts, mathematical models as well as simulation and analysis facilities to predict the design's behavior [3].

In synthetic biology it is necessary to have computer-aided design (CAD) tools in order to design, analyze and build biosystems. These tools provide a path from a design to a final product. Features of a good CAD tool for synthetic biology are [1]:

- Providing mathematical description for the models (parameters, equations of dynamics of the part, etc.).
- Providing several modeling methods (different ways of defining the dynamics of the model).
- Ability to interface with third-party programs.
- Ability to share tools.
- Ability to reuse existing smaller circuits in order to build larger ones.
- Considering the uncertainties in real system (ability to define parameters as a range or distribution of value and analyze the model by considering these uncertainties).

The typical design flow in bio-CADs is depicted in Fig. 2.

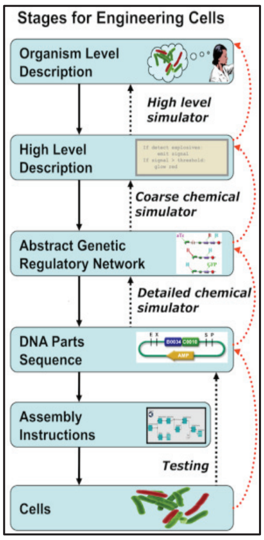

Figure 2. Typical design flow in synthetic biology

Different approaches have been taken for constructing CAD programs for synthetic biology. Existing computational applications are listed in tables I-IV [5]. Each of them is unique in its own ways and has its own respective focus areas.

Table I
SYNTHETIC BIOLOGY: DESIGN TOOLS

Tool	Description
Antimony [6]	A modular model definition language where a system is defined as a set of reactions or basic DNA parts. The modules can be connected to each other, and the final system can be converted to a mathematical model for analysis.
BioJADE [7]	A visual design application where symbols from electrical engineering are used to represent biological networks. With the support of a parts database, BioJADE is able to associate DNA sequence information with the visual design.
GEC [8]	A language for describing biochemical reactions in terms of events and rules that govern interactions. DNA components such as promoters are used to specify the types of components in a model. GEC also supports modularity.
GenoCAD [9]	A web-based program where a sequence of parts are placed in accordance to a set of grammatical rules. These motifs are later 'mapped' to physical implementations. The graphical environment restricts the way in which the functional motifs are composed to encourage a functioning final product.
iBioSim [10]	iBioSim has been developed for the modeling, analysis, and design of genetic circuits. While iBioSim primarily targets models of genetic circuits, models representing metabolic networks, cell-signaling pathways, and other biological and chemical systems can also be analyzed.
ProMoT [11]	A modular model definition language and a visual tool where a system is defined as a set of reactions. The modules can be connected to each other, and the final system can be converted to a mathematical model for analysis.
SynBioSS [12]	A web-based and desktop-based tool that allows users to connect biological parts, such as promoters and coding regions, and build a consistent mathematical model.

TASBE [4]	A free and open source tool-chain to decompose the problems of design and assembly into automatable fragments. Organism design is possible using higher level descriptions, which are automatically transformed into genetic regulatory network designs, then assembled into DNA samples ready for in vivo execution.
TinkerCell [1]	A visual drawing tool for constructing modular and semantically structured diagrams. Due to semantic descriptions, the network diagrams can be automatically mapped to mathematical models or biological parts. TinkerCell has extensive support for plug-ins.

Table II
SYSTEMS BIOLOGY: MATHEMATICAL ANALYSIS TOOLS

Tool	Description
JDesigner	A visual model construction program where reaction networks are constructed using nodes and reactions. In addition to basic simulation, the SBW suite is used to perform various types of analyses.
Jarnac	A text-based model construction program where simple scripts are used to describe reactions and rate equations. Programming language features such as loops and matrix operations are available for analysis of models.
CellDesigner	A visual model construction program where reaction networks are constructed using molecules such as RNA, proteins, and genes. Simulation and other analyses are made available through simulation libraries and SBW.
COPASI	A graphical interface for analysis of models. Some of the types of analyses supported include simulations, optimization, and parameter scans.
PySCeS	A program that is built on Scientific Python and provides all the programming flexibility of Python. PySCeS uses a simple text-based language to describe reactions and nodes that comprise a model. Available analyses include simulations, metabolic control analysis, bifurcation analyses, and any function available in Scientific Python.

Table III
SYNTHETIC BIOLOGY: PART COMPOSITION AND MANAGEMENT TOOLS

Tool	Description
BioMotar	BioMortar is an application to facilitate the construction of BioBricks by consolidating information is an easy to access database and provides Standard Assembly-driven lab protocols.
Clotho [13]	Connects users to repositories of biological parts. Plugin tools then define the various functions that can be performed. Clotho makes it easier to share data and get it in a data model useful to synthetic biologists.
GenoCAD [9]	Web-based tool for the design of biological devices using an attribute grammar which defines the legal composition of parts.
j5 [14]	Designs assembly strategies for a variety of different assembly protocols given a list of initial parts. Works for SLIC, Gibson, CPEC, and Golden-Gate based assemblies.
MIT Registry of Standard Parts	A large repository of standard biological parts.
Viz-a-bricks	A visual environment for navigating registries of biological parts.

Table IV
GENETIC ENGINEERING: SEQUENCE REFINMENT TOOLS

Tool	Description
GeneDesign	An open-source web application that provides functions such as removal of unwanted restriction sites, codon optimization, and annotation.
GeneDesigner	A proprietary desktop application for primer-design *in silico* cloning, codon optimization, and visualization of DNA sequences.
Sequence Refiner	A package for modifying natural DNA sequences so that they conform to specific synthetic biology standard. Part of the BIOFAB code collection.
Vector NTI	A proprietary desktop application for primer-design *in silico* cloning, codon optimization, finding restriction sites, finding markers and key features, and visualization of DNA sequences.

IV. CHALLENGES AND FUTURE DIRECTIONS

Despite great potential of synthetic biology, important limitations exist. There are unresolved issues at several levels: wet-lab protocols, exchange of information and computational modeling.

Synthetic biology is based on the idea that biological parts can be assembled to build circuits with predictable behavior. Wet-lab protocols for assembling biological parts are time-consuming and difficult, especially to build large circuits.

Regarding information exchanging, standard protocols are required to exchange parts and description of their functions between research groups.

For computational analysis, it is not clear yet what kinds of models are most appropriate to predict biological circuit behavior, especially in the presence of so many unknown parameters in an average synthetic circuit [1].

Because of complex nature of biological systems, each of the mentioned issues still needs lots of investigations and efforts.

V. CONCLUSIONS

Synthetic biology merges knowledge from various disciplines including molecular biology, engineering and mathematics to design and build new genetic circuits and metabolic networks. Early efforts concentrated on altering the behavior of individual elements but now synthetic biologists try to construction of complex networks in single-cell and multicellular systems.

The goal of a CAD tool in synthetic biology is to provide an efficient design flow for creating and analyzing new biological systems. There are some unresolved issues in building CAD tools due to lack of established methods in biology as well as relative lack of understanding regarding biological behavior which require continuous efforts to overcome [5].

ACKNOWLEDGMENT

This work was funded by the Malaysia's Ministry of Higher Education (MOHE).

REFERENCES

[1]. D. Chandran, F. T. Bergmann. And H. M. Suaro, "Computer-aided Design of Biological Circuits Using TinkerCell", Bioengineered Bugs, vol. 1(4), pp. 274-281, 2010.

[2]. L. Bilitchenko, A. Liu, S. Cheung, E. Weeding, B. Xia, M. Leguia, et al, "Eugene-A Domain Specific Language for Specifying and Constraining Synthetic Biological Parts, Devices, and Systems", PLoS ONE, vol. 6(4): e18882, 2011.

[3]. M. A. Marchisio and J. Stelling, "Computational Design Tools for Synthetic Biology", Current Opinion in Biotechnology, vol. 20, pp. 479–485, 2009.

[4]. J. Beal, T. Lu and R. Weiss, "Automatic Compilation from High-Level Biologically-Oriented Programming Language to Genetic Regulatory Networks", PLoS One. Vol. 6(8): e22490, 2011.

[5]. D. Chandran, F. T. Bergmann, H. M. Sauro and D. Densmore, Computer-Aided Design for Synthetic Biology (book chapter), Design and Analysis of Bio-molecular Circuits, Springer-Verlag, pp. 203-224, 2011.

[6]. L. P. Smith, T. F. Bergmann, D. Chandran, and H. M. Sauro. "Antimony: a Modular Model Definition Language", Bioinformatics, vol. 25(18), pp. 2452, 2009.

[7]. J. A. Goler, "BioJADE: A Design and Simulation Tool for Synthetic Biological Systems", Doctoral Philosophy Thesis, MIT, USA, 2004.

[8]. M. Pedersen and G. Plotkin, "A Language for Biochemical Systems: Design and Formal Specification", In Transactions on Computational Systems Biology XII, vol. 5945, pp. 77-145, 2010.

[9]. Y. Cai, "GenoCAD: Linguistic Approaches to Synthetic Biology", Doctoral Philosophy Thesis, Virginia Technology University, USA, 2010.

[10]. C. J. Myers, N. Barker N., K. Jones, H. Kuwahara, C. Madsen and N-P. D. Nguyen, "iBioSim: A Tool for the Analysis and Design of Genetic Circuits", Bioinformatics, vol. 25 (21), pp. 2848-2849, 2009.

[11]. S. Mirschel, K. Steinmetz, M. Rempel, M. Ginkel, E. D. Gilles, "PROMOT: Modular Modeling for Systems Biology", Bioinformatics, vol. 25(5), pp. 687, 2009.

[12]. A. D. Hill, J. R. Tomshine, E. Weeding, V. Sotiropoulos and Y. N. Kaznessis, "SynBioSS: The Synthetic Biology Modeling Suite", Bioinformatics, vol. 24(21), pp. 2551–2553, 2008.

[13]. D. Densmore, A. V. Devender, M. Johnson and N. Sritanyaratana, "A Platform-Based Design Environment for Synthetic Biological Systems", Proceeding of TAPIA '09, The Fifth Richard Tapia Celebration of Diversity in Computing Conference: Intellect, Initiatives, Insight, and Innovations, ACM, 2009.

[14]. J. H. Nathan, R. D. Rosengarten and J. D. Keasling, "j5 DNA Assembly Design Automation Software", *ACS Synth. Biol.*, vol. 1 (1), pp. 14–21, 2012.