Integration of Standardized Cloning Methodologies and Sequence Handling to Support Synthetic Biology Studies

Maurice HT Ling
Life Technologies, Inc.
33 Marsiling Ind. Est.Rd. 3
Singapore 739256
(65) 63629415
Maurice.Ling@lifetech.com

Angela Jean
Life Technologies, Inc.
33 Marsiling Ind. Est.Rd. 3
Singapore 739256
(65) 63629416
Angela.Jean@lifetech.com

Dunqiang Liao
Life Technologies, Inc.
33 Marsiling Ind. Est.Rd. 3
Singapore 739256
(65) 63629585
Dunqiang.Liao@lifetech.com

Ben BY Tew
Life Technologies, Inc.
33 Marsiling Ind. Est.Rd. 3
Singapore 739256
(65) 63629555
BengYong.Tew@lifetech.com

Shanice Ho
Life Technologies, Inc.
33 Marsiling Ind. Est.Rd. 3
Singapore 739256
(65) 63629307
Shanice.Ho@lifetech.com

Kevin Clancy
Life Technologies, Inc.
5791 Van Allen Way
CA 92008, USA
(01) 7602688356
Kevin.Clancy@lifetech.com

ABSTRACT

The assembly and downstream transformation of genetic constructs has been a fundamental scientific technology for the last thirty years. Synthetic biology is an engineering based approach to molecular biology as emphasizing the standardized assembly of characterized DNA fragments. The standards promoted by the BioBricks™ Foundation have enabled novel constructs to be developed based upon the expected function of these constructs. However scientists need a software environment that enables them to curate large collections of parts and assemblies, combined with appropriate tools to facilitate quick creation of constructs and identification of potential design issues *in silico*. In this paper, we present the implementation of BioBrick™ and GENEART® Assembly tools, coupled with an enhanced database to manage and develop such parts collections. Integration of these tools and data into the VectorNTI® software suite is a step towards implementation of BioCAD™, a computer based design approach to facilitate development of complex circuit based perturbation of cellular systems.

Categories and Subject Descriptors

J.3 [**Computer Applications**]: Life and Medical Sciences – biology and genetics

General Terms

Algorithms, Design, Experimentation

Keywords

Bioinformatics, Synthetic biology, BioBrick, GENEART Assembly, Vector NTI, Cloning, Parts, Devices

IWBDA'11, June 6-7, 2011, San Diego, CA, USA.

1. INTRODUCTION

The use of constructs has evolved from simple molecular cloning experiments to gene reporting [1], switch testing [2] and promoter characterization [3]. Assembly of constructs typically involves a backbone vector as well as the set of DNA fragments of interest. Depending on the DNA sequence and the assembly system in place, various combinations of restriction sites can be used for the extraction and recovery of the final target fragment [4].

Through the understanding of underlying gene expression mechanisms, novel and optimized biological systems can be created [5]. Although there are many ways to assemble constructs [6], typically biologists have not developed their materials with an eye to reuse or elaboration from simple to complex systems. This large scale assembly or reuse of constructs in series of experiments are often restricted and limited to the capabilities of the selected protocol [7]; leaving little room for interoperability or extensibility for future experiments.

A set of open cloning standards promoted through the BioBricks™ Foundation [8] can be used for the typing of constructs and the standardization of assembly protocols. Availability of various assembly standards [9] using combinations of restriction sites, amongst other criteria, also provides the option of selecting an appropriate assembly protocol for a given assembly. The resulting construct will have a restriction site scar between the 2 fragments which can have functional consequences, particularly when combining protein domains [10].

The GENEART® Assembly System is a homologous recombination based cloning system. This experimental design tool takes advantage of recombination in vitro or in yeast to join pre-existing DNA fragments or chemically synthesized fragments into a single recombinant molecule [11]. As the system relies on homologous recombination, the adjacent DNA fragments must share end-terminal homology [12] or they must be "stitched" together by means of primers [13], known as stitching primers. This system is thus reliant upon identification of possible issues that would disrupt homologous recombination and accurate design of the different types of primers needed in the experiment.

In conjunction with the BioBrick™ standards, the Repository of Biological Parts has developed to track DNA parts that have been generated by scientists and participants in the iGEM competition over the last several years [14]. The current collection has more than 15,000 entries, however many of these entries are poorly characterized and the lack of defined, parsable data records renders much of this data difficult to discover, use or validate in another laboratory. However these parts can be assembled using either traditional BioBrick or GENEART® Assembly approaches.

Ideally synthetic biologists would use software that can assist them in development of their projects. This would require the development of software that could 1. handle collections of parts and devices during various stages of development, managing the provenance of development of novel parts, characterization of existing parts and assembly of parts into more complex devices; 2. provide expert tools that could facilitate development of projects by automating or semi automating design decisions, and; 3. provide a means to identify design issues during in silico development. Enabling such a BioCAD™ environment in Vector NTI®, a widely used commercial software package, would greatly speed both development among existing synthetic biologists as well as introducing such practices to the larger population of existing molecular biologists.

2. ALGORITHMIC APPROACHES

The Vector NTI® Suite is a comprehensive desktop bioinformatics suite. The software uses a serialized, indexed flat file database for sequence record import, storage and retrieval. The database was modified from an ASCII format to a binary format and the indexing scheme was supplemented with an improved key value index. These changes both increased the numbers of records that can be stored from 10,000 to ~100,000 as well as decreased search and retrieval times. The improved database was used to import the entirety of the Standard Repository of Biological parts. These molecules were then functionally characterized based upon their BioBrick IDs, information encoded in the header lines of the files and sequence similarities. Subsets containing functionally characterized parts were used in subsequent *in silico* development projects.

The Vector NTI Molecule Viewer® was modified to identify sets of restriction enzymes important to the BioBrick Assembly standards. These sets were used to construct custom view and search sets for part characterization. This allowed us to quickly scan new and existing parts for the presence of these sites and helped us identify mutagenesis strategies to remove such sites from these affected parts.

Construct assembly using BioBrick™ standards was implemented as described in the BioBricks standards (Life Technologies, Inc.). To reflect the current Biobrick assembly approach, paired sets of molecules are submitted for assembly. Sequential runs using progressively assembled parts gives rise to a multi-fragment construct representing complex devices.

The implementation of construct assembly using BioBrick™ standards allows the final construct to be typed and assembled according to these standards. At the same time, the provision of chassis options and strand direction of assembly also allows various options to be visualized prior to the in vitro assembly of the construct. This facilitates rapid design, prototyping and debugging of potential design errors as in silico designs are developed.

The GENEART® Assembly System is a homologous cloning based tool which also supports synthetic biology cloning projects. There are 4 modes to the system – seamless cloning, high-ordered genetic assembly resulting in construct below 60 kilobases, high-ordered genetic assembly resulting in construct above 60 kilobases, and high-ordered stitching assembly. For 3 former modes required homologous ends between adjacent fragments and allow primers to be designed for use to amplify the original fragments. Seamless cloning allows up to 4 DNA fragments plus a vector totaling up to 13 kilobases in length while high-ordered genetic assembly allows up to 10 DNA fragments plus a vector totaling up to 110 kilobases in length. The stitching assembly does not require homologous ends between adjacent ends as up to 3 sets of stitching primers can be designed, implying a maximum of 5 fragments, including the vector, to be stitched together. The resulting clone can be displayed in a circular or linear form.

Both tools were used to design and subsequently construct constructs conforming to the two cloning methodologies. BioBrick™ constructs were validated to previously reported constructs in the Parts Repository. GENEART® Assembly constructs were verified by size, restriction patterns and through end and junction sequencing studies.

3. CONCLUSION

Experimental design tools have facilitated final results to be visualized before *in vitro* experiments are performed; allowing potential issues to be identified and resolved through *in silico* means. At the same time, the results provide baselines against which *in vitro* results can be compared. The implementation of a construct assembly algorithm utilizing open standards prescribed by the BioBricks™ Foundation has not only enabled the visualization and inspection of the final construct, but allowed various hypotheses and scenarios to be carried out without significant costs. The GENEART® Assembly tool permits the construction of 100 bp to 120,000 bp constructs without the use of restriction enzymes. The improvement of the Vector NTI database permits the improved use and reuse of parts and constructs during design. This integrated system is an initial step to providing robust, scalable and easy to use access to Synthetic Biology tools.

4. REFERENCES

[1] Wesley, S.V., Helliwell, C.A., Smith, N.A., Wang, M., Rouse, D.T., Liu, Q., Gooding, P.S., Singh, S.P., Abbott, D., Stoutjesdijk, P.A., Robinson, S.P., Gleave, A.P., Green, A.G. and Waterhouse, P.M. 2001. Construct design for efficient, effective and high-throughput gene silencing in plants. The Plant Journal 27, 581–590.

[2] Gardner, T.S., Cantor C.R., Collins J.J. 2000. Construction of a genetic toggle switch in Escherichia coli. Nature 403, 339-342.

[3] Wick M., Zubov D., Hagen G.. 1999. Genomic organization and promoter characterization of the gene encoding the human telomerase reverse transcriptase (hTERT). Gene 232(1), 97-106.

[4] Karn, J. and Matthes, H.W.D. and Gait, M.J. and Brenner, S. 1984. A new selective phage cloning vector, [lambda] 2001, with sites for XbaI, BamHI, HindIII, EcoRI, SstI and XhoI. Gene 32(1-2). 217-224.

[5] Agapakis, C.M. and Silver, P.A. 2009. Synthetic biology: exploring and exploiting genetic modularity through the design of novel biological networks. Molecular BioSystems 5(7), 704-

713.

[6] Hebelstrup, K.H., Christainsen, M.W., Carciofi, M., Tauris, B., Brinch-Pederson, H. amd Holm, P.B. 2010. UCE: A uracil excision (USER™)-based toolbox for transformation of cereals. Plant Methods 6(1), 15.

[7] Harbers, M. 2008. The current status of cDNA cloning. Genomics 91(3), 232-242.

[8] Smolke, C.D. 2009. Buidling outside of the box: iGEM and the BioBricks Foundation. Nature Biotechnology 27(12), 1099-1102.

[9] Shetty, R.P., Endy, D. and Knight Jr, T.F. 2008. Engineering BioBrick vector from BioBrick parts. Journal of Biological Engineering 2, 5.

[10] Muyrers, J.P.P., Zhang, Y. and Stewart, A.F. 2001. Techniques: recombinogenic engineering new options for cloning and manipulating DNA. Trends in Biochemical Sciences 26(5), 325-333.

[11] Larionov, V., Kouprina, N., Graves, J., Chen, X.N., Korenberg, J.R. and Resnick, M.A. 1996. Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination. Proceedings of the National Academy of Science USA 93, 491-496.

[12] Gibson, D.G., Glass, J.I., Lartigue, C., Noskov, V.N., Chuang, R.Y., Algire, M.A., benders, G.A., Montague, M.G., Ma, L., Moodie, M.N., MerrymN, c., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E.A., Young, L., Qi, Z.Q., Segall-Shapiro, T.H., Calvey, C.H., Parmar, P.P., Hutchison 3rd, C.A., Smith, H.O. and Venter, J.C. 2010. Creation of a bacterial cell controlled by a chemically synthesized genome. Science 329, 52-56.

[13] Raymond, C.K., Sims, E.H. and Olson, M.V. 2002. Linker-mediated recombinational subcloning of large DNA fragments using yeast. Genome Research 12, 190-197.

[14] http://partsregistry.org/Main_Page