New Online Platform for Well-Characterized Meganucleases

June 18, 2012

ME Arnegard

Targeted gene modification, which can be accomplished by gene-specific double-stranded DNA breakage followed by mutagenic end-joining or homologous recombination, is becoming an increasingly powerful tool for genetically manipulating model organisms and treating certain human diseases. Just a decade ago, precise genome editing was only efficient in the yeast and mouse models. Due to the emergence of engineered nucleases, genome editing is now on its way to becoming standard procedure for a menagerie of other systems.

Three different protein scaffolds with sequence-specific DNA affinity and nuclease activity can achieve targeted gene modification: zinc-finger nucleases (ZFNs); TAL effector nucleases (TALENs); and LAGLIDADG homing endonucleases (LHEs; also known as 'meganucleases'). Although the modularity of ZFNs and TALENs makes it easier to engineer their target specificities, each of these scaffolds consists of large chimeras that only act as dimers and therefore must be expressed as separate proteins. While it is more difficult to re-engineer LHEs, these nucleases are better suited for gene targeting due to their compact, monomeric structures and high DNA cleavage specificity.

The laboratory of Dr. Barry Stoddard (Basic Sciences Division) recently discovered a wealth of naturally occurring LHEs residing in microbial genome databases (Takeuchi *et al.*, 2011). To help make the most of these untapped resources, joint first authors Gregory Taylor (Basic Sciences Division) and Lucas Petrucci (Department of Computer Science and Engineering at the University of Washington), along with Dr. Stoddard and others, recently developed a customized platform for organizing information about well-characterized LHEs. They call their new platform the LAGLIDADG Homing Endonuclease Database and Engineering Server, or 'LAHEDES.'

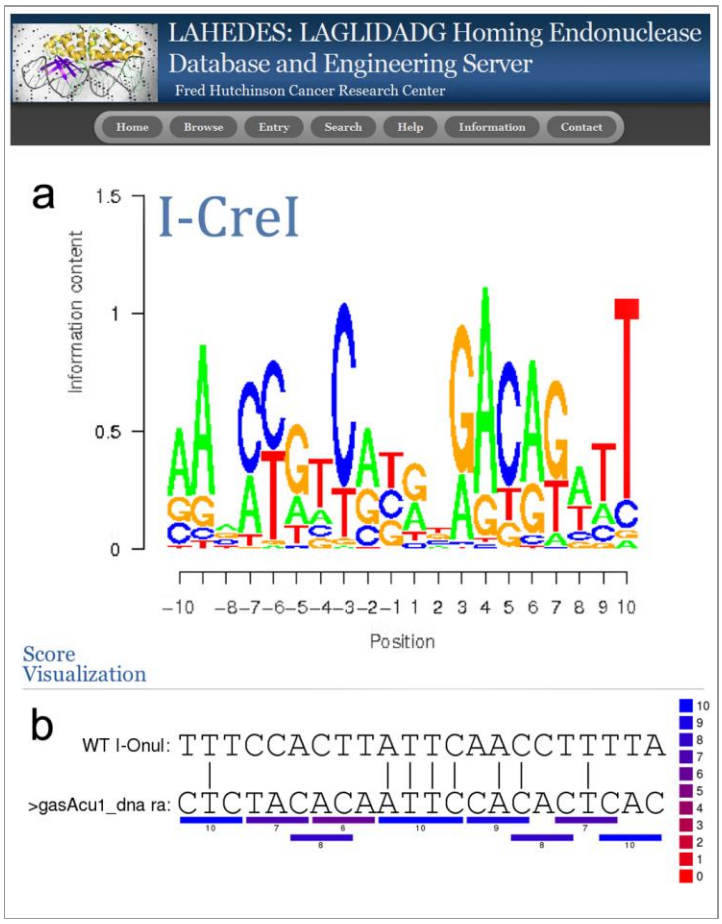
Large numbers of putative but ill-characterized homing endonuclease genes are already being cataloged in several other online databases. Thus, Taylor *et al.* focused their programing skills on the best-described LHEs, which meet strict criteria regarding well-defined and biochemically-characterized DNA target specificities. The authors reasoned that these are exactly the kinds of LHEs that will be most effective for future gene-targeting reagents. To facilitate the development of

such reagents, the authors designed several tools for LAHEDES. *Entry* tools allow for the cataloging of new endonucleases meeting the authors' rigorous selection criteria. An *Endonuclease Browser* summarizes the gene targeting information for wild-type and engineered LHEs alike, and *Genomic Search* tools help users identify DNA sequences that can be targeted by cataloged LHEs (see figure). Powerful search algorithms allow genomes to be queried using several position weight matrices (PWMs) and scoring approaches. These include: perfect identity searches across the central four base pairs of a given LHE; searches for the highest identity across the entire 22 base pair recognition site; searches for the best match against a PWM that accounts for variable recognition fidelity at individual nucleotides; and searches for combinations of three base pair modules that appear to be 'engineerable' based on cleavage activity screens in yeast.

Currently, LAHEDES contains 24 wild-type LHEs, three engineered LHEs, six chimeric LHEs and one pseudoendonuclease. These totals are sure to increase rapidly in the years to come, as the properties of many new LHEs are determined, and more and more researchers discover the utility of LAHEDES for facilitating targeted gene modification.

Taylor GK, Petrucci LH, Lambert AR, Baxter SK, Jarjour J, Stoddard BL. 2012. LAHEDES: the LAGLIDADG homing endonuclease database and engineering server. *Nucleic Acids Research,* Epub ahead of print, doi: 10.1093/nar/gks365.

Also see: Takeuchi R, Lambert AR, Mak AN, Jacoby K, Dickson RJ, Gloor GB, Scharenberg AM, Edgell DR, Stoddard BL. 2011. Tapping natural reservoirs of homing endonucleases for targeted gene modification. *Proc. Natl. Acad. Sci. USA* 108:13077-13082.



Images downloaded from http://www.homingendonuclease.net/

Appearance and functionality of the LAHEDES server. Functions in the top menu include: browsing basic properties of LHEs having validated activities and known specificities; entry of new LHEs meeting established criteria for protein engineering and selection experiments; and web-based searches of DNA sequences for potential target sites. For example, (a) shows a sequence logo plot summarizing the binding affinity of wild-type I-CreI. (b) Example output from a 'Module Search' performed on stickleback genome sequence in a 10-kb window around the gene 'eda.' This small part of the total output is the detailed view of the best candidate I-OnuI target (lower sequence), together with the natural I-OnuI target (upper sequence). 'Engineerable' DNA modules are represented by horizontal bars. The score of each module is indicated by the color scale at the right (favorably-scoring modules in blue, poorly-scoring modules in red).