

ABSTRACT

Gene synthesis from commercial vendors is typically conducted by using overlap-extension PCR from oligos manufactured on microarrays. Here we apply Data-Optimized Assembly Design of Golden-Gate overhangs to rapidly construct genes from oligo pools in three steps and as a result create a workflow that can be employed in any standard molecular biology lab.

INTRODUCTION

Golden Gate Assembly has proven to be a valuable cloning method for the construction of high complexity assemblies. As opposed to homology-based cloning techniques utilizing 15-20 base pair complementary regions, Golden Gate uses short sticky-ends generated by restriction digest. In order to increase the fidelity of ligations using these overhangs, New England Biolabs has characterized the fidelity of T4 DNA ligase with all possible overhangs in a competition experiment¹. The resulting data set is then used to predict high-fidelity junction sets to enable complex multipart assemblies. By using Data-Optimized Assembly Design, Golden Gate has been used in a single-pot reaction to generate a 52-part assembly².

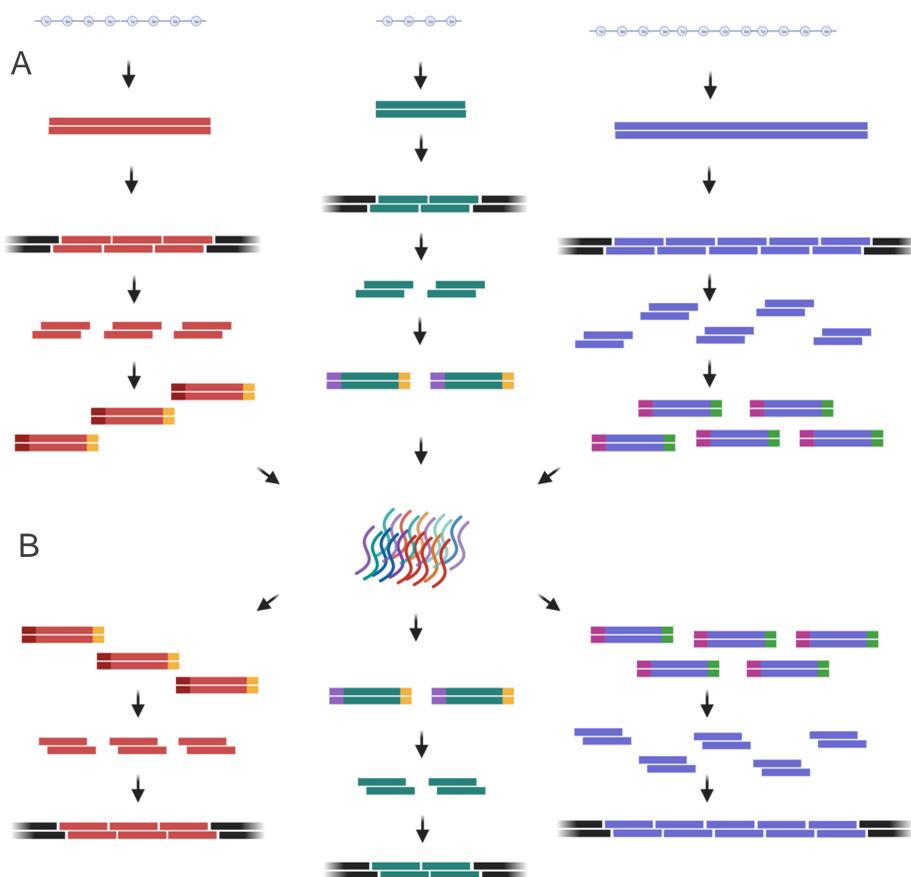
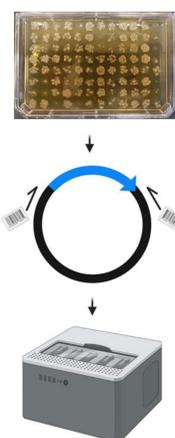


Figure 1: A. Design of oligo pool. Protein sequences are first codon-optimized for expression and domestication (exclusion of BsaI sites). Each design is broken into corresponding fragments using Data-Optimized Assembly Design. Each fragment is then barcoded and a pool containing multiple designs is gathered. **B. Construction of genes in parallel.** Parts are amplified using design-specific primer combinations. Once PCRs are purified, a single step assembly is conducted into a vector of choice. The mixture is then transformed and colonies are assessed by sequencing.

As an extension to these defined assemblies, we asked whether parts amplified from oligo pools could be used to conduct single pot assemblies with high fidelity. Individual assemblies are broken into fragments that can be ordered in an oligo pool and appended with a unique barcode pair. A unique barcode pair is then used to retrieve parts associated with a design by PCR using the complex oligo pool as a template. While high fidelity assemblies have been shown to benefit from stoichiometrically balanced inputs, using oligo pools as template would not be expected to yield normalized parts.

METHODS



Each design is amplified in parallel using design specific primers from a 300 nt oligo pool. The amplifications are purified before being subjected to Golden Gate Assembly in the presence of an accepting vector. Once transformed, isolates are picked and subjected to a colony PCR appending barcodes indicating the location on a plate of a specific isolate³. The PCRs are pooled and sequenced using a MinION flow cell. Sequencing traces are demultiplexed using Minibar and assembled using Amplicon Sorter.

Figure 2: Low-cost high-throughput sequencing workflow³. Forward and reverse primers bind outside of the insert and to the vector. The forward primer has a 5' barcode indicating the well on a plate while the reverse primer's 5' barcode codes for the plate in which the isolate is found. Using this strategy, up to 16 plates of amplicons have been sequenced on a single MinION flow cell.

RESULTS

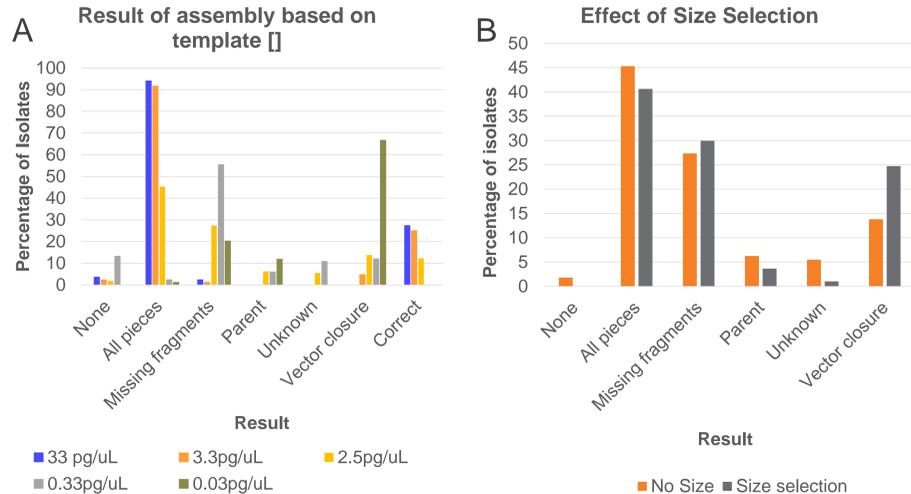


Figure 3: A. Optimizing template concentrations for PCR. 8 problematic assemblies of 9-pieces were selected for investigation of the effect of template concentration on success rate. A clear trend was observed in that assemblies from PCRs with high amounts of template tended to yield more isolates with all fragments present. **B. Investigating the use of size selection after PCR.** Size selection for larger PCR products in the initial amplification were selected for using SPRI Select beads. There appeared to be no effect of size selection on increasing the number of isolates with all fragments.

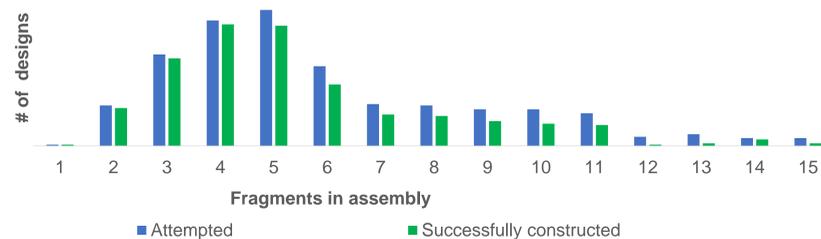


Figure 4: Design space attempted for synthesis and the result after checking 4 colonies. Each fragment corresponds to an average length of 220 bp. The sequence space attempted is indicated in blue while the space successfully constructed after checking 4 colonies from each design is in green.

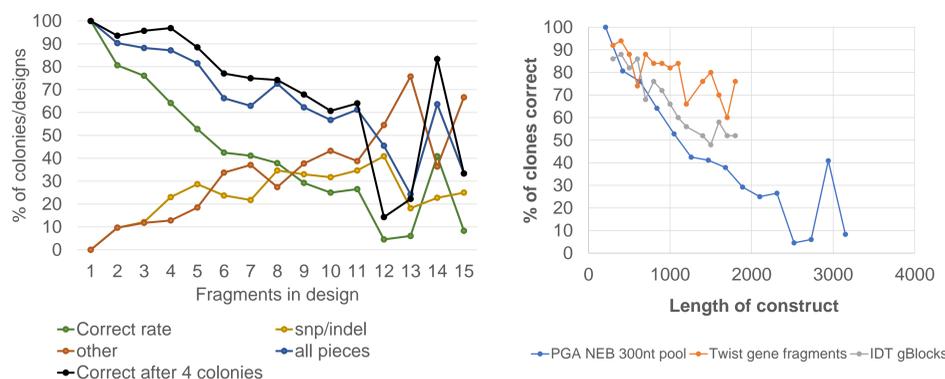


Figure 5: A. Summary of relationship between design complexity and assembly rates. For each design, 4 colonies were sequenced. As the length of the design increased, the number of SNPs and INDELS from the oligo array synthesis accordingly increased. As the number of parts goes up, the assembly success rate drops. **B. Comparison of Parallel Gene Assembly (PGA) compared to commercial DNA fragment suppliers.** Data from gBlocks and Twist gene fragments from Twist Biosciences available on website⁴.

SUMMARY AND FUTURE DIRECTIONS

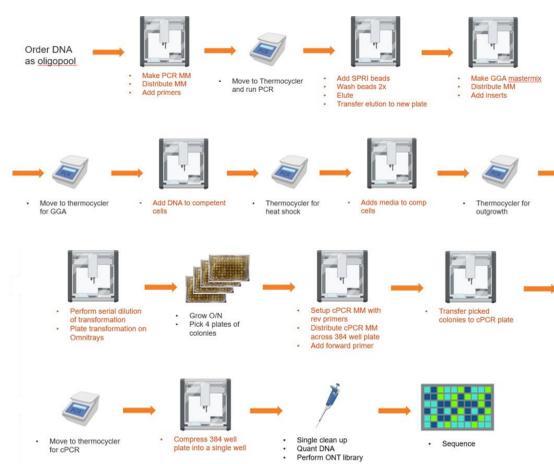


Figure 6: Automating gene construction with the Opentrons OT2. Future directions include automation of the entire gene construction process from oligos using an off-deck thermocycler, ONT sequencer, and Opentrons OT2 pipetting robot.

Parallel gene synthesis from oligo pools lags in success rates compared to commercially available gene fragments. However, assemblies as large as 17-piece and 3600 bp were found after checking 4 colonies. While less successful, DNA costs are more than a magnitude cheaper than the cost of gene fragments.

REFERENCES

- Potapov, Vladimir et al. "Comprehensive Profiling of Four Base Overhang Ligation Fidelity by T4 DNA Ligase and Application to DNA Assembly." *ACS synthetic biology* vol. 7,11 (2018): 2665-2674. doi:10.1021/acssynbio.8b00333
- Pryor JM, Potapov V, Bilotti K, Pokhrel N, Lohman GJS. Rapid 40 kb Genome Construction from 52 Parts through Data-optimized Assembly Design. *ACS Synth Biol*. 2022 Jun 17;11(6):2036-2042. doi: 10.1021/acssynbio.1c00525. Epub 2022 May 25. PMID: 35613368; PMCID: PMC9208013.
- Curran, Andrew et al. "Highly multiplexed, fast and accurate nanopore sequencing for verification of synthetic DNA constructs and sequence libraries." *Synthetic biology (Oxford, England)* vol. 4, 1 yz025. 29 Oct. 2019. doi:10.1093/synbio/yz025
- <https://www.twistbioscience.com/products/genes?tab=fragment>. Accessed May 22, 2023.