Highly efficient and rapid generation of genetic variants

A new mutagenesis platform enables the fast, cost-efficient and automatable production of defined multi-site sequence variants for a wide range of applications. Demonstrations of this method included the generation of SARS-CoV-2 spike gene variants, DNA fragments for large-scale genome engineering, and adeno-associated virus 2 (AAV2) cap genes with improved packaging capacity.

This is a summary of:
Liu, L., Huang, Y. & Wang, H. H. Fast and efficient template-mediated synthesis of genetic variants. *Nat. Methods* https://doi.org/10.1038/s41592-023-01868-1 (2023)

Published online: 4 May 2023

The question

The production of kilobase-sized DNA building blocks is a key part of design–build–test–learn (DBTL) workflows used for applications in biotechnology such as synthetic biology, gene therapy, metabolic engineering and DNA data storage. Despite notable improvements in de novo DNA synthesis in recent years[1], current methods are still limited by the size of DNA they can produce and their fidelity, scale and cost. Beyond de novo DNA synthesis, available commercial mutagenesis kits can produce only a limited number of mutations and require cloning steps that are difficult to scale. Other alternative approaches to producing genetic variants, including recombineering, base editing or prime editing[2], involve the assembly of complex constructs and cannot be easily used in multiplex applications. As a result of the above shortcomings, rapid advances in the computational design of genetic variants[3] are poised to outpace our capacity to build and test DNA sequences large enough to encode proteins in the laboratory, thus highlighting a widening technology gap.

The discovery

DNA polymerases with proofreading activity, such as Q5 high-fidelity DNA polymerase, have a strong attraction to uracil nucleotides that leads to stoppage of DNA polymerization. However, the modified DNA polymerase Q5U, which contains mutations in its uracil-binding pocket, enables the amplification of DNA templates that contain uracil and inosine bases. Leveraging this property, we developed an in vitro gene variant synthesis platform called mutagenesis by template-guided amplicon assembly (MEGAA). In MEGAA, a desired variant sequence is first generated by annealing mutagenic oligonucleotides to uracil-containing template DNA, extending these oligonucleotides with Q5U, and ligating them with Taq DNA ligase, all in a single-pot reaction. The process allows 6–8 nucleotide mismatches and insertions, or deletions of 6–20 nucleotides, to be easily introduced per oligonucleotide. Next, Q5 DNA polymerase is used to amplify a desired variant amplicon from the MEGAA reaction by PCR, without interference from the original uracil-containing template DNA. The resulting product can then be used directly for downstream applications.

MEGAA enables the creation of many mutations in kilobases of DNA in a predictable manner at an efficiency of >90% per target site (Fig. 1a). Iterative cycling of MEGAA reactions by using the output from one round as the input for the next can tune the genotypic purity of the desired product, and additional combinatorial variants if desired (Fig. 1b). We developed an open-source lab automation workflow for MEGAA, dubbed MEGAAtron, to facilitate the desktop production and long-read sequence validation of variants. Using MEGAAtron, we demonstrated the successful construction of 31 natural SARS-CoV-2 spike gene variants, 10 recoded *Escherichia coli* genome fragments each with up to 150 mutations, and 125 AAV2 gene variants with combinatorial mutations at 6 defined sites. Some of the resulting AAV2 variants exhibited tenfold enhancement in viral packaging compared to wild-type AAV2.

The implications

MEGAA is lower cost and higher throughput for gene variant generation than traditional de novo DNA synthesis approaches and enables the routine production of DNA fragments of 4–5 kilobases. This new capacity opens up the possibility for synthetic genomics projects to build and modify genomes larger than those of bacteria and moves the field towards the synthesis of gigabase plant and mammalian cell genomes engineered to be genomically fortified and disease resistant[4].

There are some limitations to MEGAA. For some applications, completely pure and sequence-verified products are needed, which would require a clonal isolation step. Further, the efficiency of MEGAA can be influenced by factors such as local G+C content, the secondary structure of the template DNA, and the quality of the synthesized mutagenic oligonucleotides.

Future improvements to MEGAA might rely on orthogonally optimized DNA polymerases, heat-resistant DNA ligases with higher accuracy, desktop oligonucleotide synthesizers, droplet-based variant synthesis strategies, and more advanced predictive models for DNA folding, annealing and mutagenesis kinetics.

Liyuan Liu & Harris H. Wang
Columbia University, New York, NY, USA.

EXPERT OPINION

"The authors constructed this new concept of high-throughput, in vitro variant generation based on an existing template, rather than de novo synthesis, to achieve genetic variants. This concept is fundamentally different to 'low-throughput' editing in vivo, such as base editing, prime editing, the principal investor's own previous work MAGE (which can only introduce a low number of variants) and the de novo synthesis of DNA (which is expensive and slow)." **Kaihang Wang, California Institute of Technology, Pasadena, CA, USA.**

FIGURE

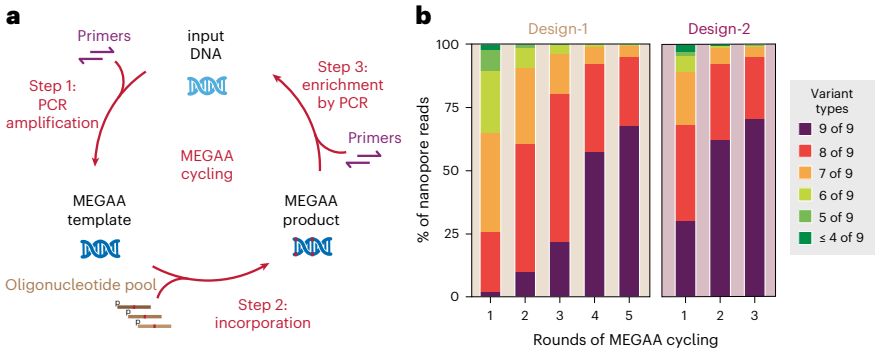

Fig. 1 | MEGAA cycling and optimization. a, An overview diagram illustrating the MEGAA process. Oligonucleotides and templates are introduced into the MEGAA workflow and assembled into variant products that can be further cycled for improved purity. **b**, Bar plots showing the purity of gene products generated over multiple rounds of MEGAA. Results from two mutagenic oligonucleotide design strategies are shown: random anneal (Design-1) and ordered anneal (Design-2), highlighting improved gene product generation through oligonucleotide optimization and MEGAA cycling. © 2023, Liu, L. et al.

BEHIND THE PAPER

I have been working on a synthetic genomics project for the past few years that involves generating sequence variants across many DNA constructs. With our high-throughput screening and protein structure prediction platform in place, I needed additional DNA constructs to functionally test variants. Unfortunately, the COVID-19 pandemic caused substantial delays and disruptions in the commercial DNA synthesis sector, which greatly impacted my work and led me to consider alternative options. I struggled with insomnia during this time, so I turned to regular meditation to reduce stress. The early MEGAA concept was conceived during a period of meditation while I was in the shower. It was noteworthy that we used MEGAA to quickly generate SARS-CoV-2 spike gene variants of the strains BA.1, BA.2 and XBB, which improved our understanding of these new variants of concern[5]. With invaluable input from my mentor Harris Wang and my co-author Yiming Huang, MEGAA became a highly efficient and automated technology. I owe much gratitude to my wife, Dr. Lingling Chen, for enabling me to finish this work. **L.L.**

REFERENCES

1. Venter, J. C., Glass, J. I., Hutchison, C. A. 3rd & Vashee, S. Synthetic chromosomes, genomes, viruses, and cells. *Cell* **185**, 2708–2724 (2022).
   This review highlights advances in DNA synthesis on the horizon that will open up new possibilities in medicine, industry, agriculture and research.
2. Wannier, T. M. et al. Recombineering and MAGE. *Nat. Rev. Methods Primers* **1**, 7 (2021).
   This Primer article covers recombineering, its derivative uses, diverse applications, and methods for combining it with other genetic editing tools in genetic engineering.
3. Yang, K. K., Wu, Z. & Arnold, F. H. Machine-learning-guided directed evolution for protein engineering. *Nat. Methods* **16**, 687–694 (2019).
   This article describes machine-learning approaches to predict sequence–function mapping without a detailed model, thus accelerating directed evolution by learning from characterized variants and selecting sequences likely to improve properties.
4. Boeke, J. D. et al. The Genome Project-Write. *Science* **353**, 126–127 (2016).
   In this article, leaders from Genome Project-Write advocate large-scale synthesis and engineering of gigabase genomes of higher-order eukaryotes.
5. Liu, L. et al. Striking antibody evasion manifested by the Omicron variant of SARS-CoV-2. *Nature* **602**, 676–681 (2022).
   This paper characterized SARS-CoV-2 Omicron variants made by on-demand variant gene synthesis.

FROM THE EDITOR

"MEGAA is a method allowing the high-efficiency construction of kilobase-sized DNA variants, with potential applications in the field of synthetic biology. It sidesteps the limitations of de novo synthesis and offers advances in multiplexability, tunability and speed." **Madhura Mukhopadhyay, Senior Editor, *Nature Methods*.**