Product Note

gSynth™

Synthesis and assembly of whole plasmids

Neil M. Bell, Sylwia A. Mankowska, Steven A. Harvey and Derek L. Stemple | Camena Bioscience Ltd

INTRODUCTION

Over the last 20 years, advancements in technologies used to sequence and manipulate DNA have driven a rapid expansion in synthetic biology. However, DNA synthesis technology, another core synthetic biology tool, has not advanced at a comparable pace.

For many years phosphoramidite synthesis has been the gold-standard DNA synthesis technology, however, over long stretches of nucleotides this method is error-prone[1]. By 200 nucleotides, with the best nucleotide coupling efficiency, only ~40% of the produced material is the correct and full-length DNA sequence. Consequently, phosphoramidite synthesis is usually only used to produce short single-stranded DNA molecules called oligonucleotides. However, the coding sequence of the average human gene is just over 1000 nucleotides long. Therefore, to meet synthetic biologist's needs, gene assembly methods are required to produce long, gene-length, DNA molecules.

Camena Bioscience has developed a novel, enzymatic, *de novo* synthesis and gene assembly technology called gSynth™. Previously we determined the accuracy of gSynth™ by producing a series of 300 nucleotide fragments and comparing them with the same sequences, produced by phosphoramidite synthesis[2]. Those benchmarking experiments highlighted the superior accuracy of gSynth™; on average 85.3% of the gSynth™ produced material was correct, whereas for phosphoramidite synthesis, just 22.7% was correct. However, despite this improved accuracy, to meet synthetic biologists needs any new DNA synthesis technology must be able to produce longer, gene-length, DNA molecules.

RESULTS AND METHODS

To demonstrate the utility of gSynth™, we produced a complete 2.7 kb plasmid. This length of DNA molecule provides a significant milestone, as it is 'gene-length' and enables a range of applications from protein production to the engineering of bacteria.

Using gSynth™ we reproduced the sequence of pUC19, a commonly used 2.7 kb plasmid[3]. pUC19 is a high-copy number plasmid, meaning that it efficiently replicates within bacteria. It also contains the ampicillin resistance gene and a multiple cloning site that spans the LacZ gene. This permits the screening of bacteria that contain pUC19 and blue-white screening to determine those plasmids that contain DNA within the multiple cloning site. As part of the synthesis and assembly we added the coding sequence for the amino acids CAMENA (Fig. 1A-B).

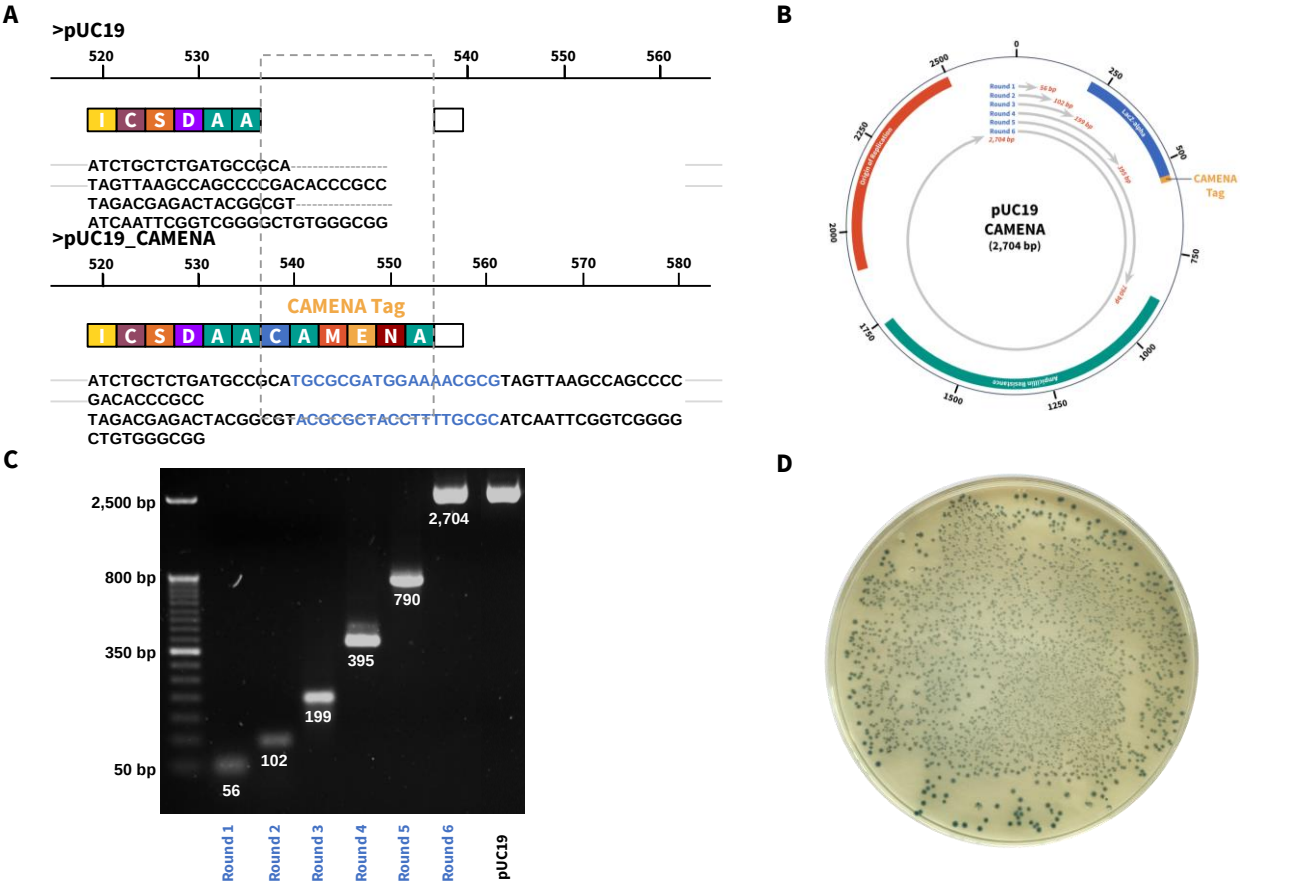

Figure 1. gSynth™ assembly of a whole plasmid. (A) Using pUC19 as a reference we designed gSynth™ experiments to reproduce pUC19, but also include the coding sequence for the amino acids CAMENA. (B) A map of the pUC19-Camena plasmid, which highlights the position of the tag and different rounds of synthesis and assembly. (C) An agarose gel highlighting different rounds of gSynth™, leading up to the full-length plasmid sequence of 2.7 kb. (D) After completion of synthesis and assembly the pUC19-Camena plasmid was transformed into DH5α bacteria cells and grown on LB-Amp plates.

Through sequential rounds of gSynth™ we produced exponentially longer double-stranded DNA fragments, leading up to a full length of 2.7 kb (Fig. 1C). If the correct plasmid sequence had been produced then it would be able to replicate in bacteria. Therefore, the gSynth™ produced pUC19-Camena plasmid was transformed into DH5α bacteria cells and grown overnight on LB-Amp plates at 37 °C (Fig. 1D). The formation of blue bacterial colonies indicated the correct formation of a plasmid sequence containing the LacZ and AmpR genes.

To confirm that gSynth™ produced the correct pUC19-Camena plasmid sequence, we picked individual bacterial colonies and, following overnight culture and DNA extraction, performed capillary sequencing analysis. Of 62 bacterial colonies sequenced, 100% contained the correct Camena tag sequence (Fig. 2).

DISCUSSION

Synthetic biology is a rapidly expanding field where biological pathways are engineered to produce new products in a range of markets from agriculture, textile production and therapeutics, the name a few. This revolution is being driven by the increasing accessibility to tools that permit the manipulation of those biological pathways. While our ability to sequence DNA has dramatically improved over the last 20 years, DNA synthesis technology has not progressed at a similar pace.

Recently there has been a drive to develop new DNA synthesis technologies[4]. By producing a series of 300 nucleotide

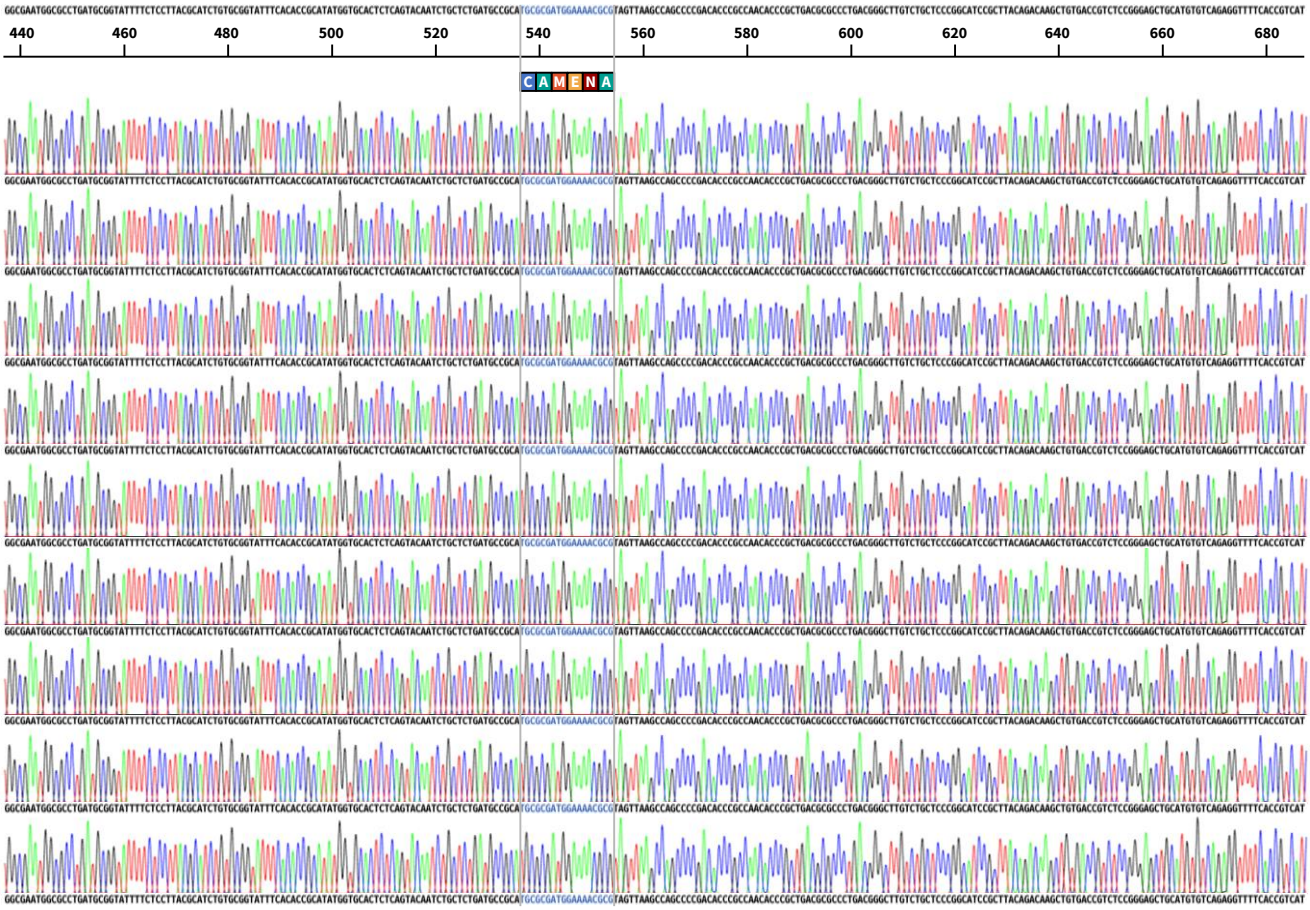

Figure 2. Capillary sequencing results of pUC19-Camena colonies. Following transformation of pUC19-Camena into DH5α bacteria cells and subsequent growth on LB-Amp plates (Fig. 1), individual colonies were picked and analysed with capillary sequencing. Each trace spans 250 bp around the Camena tag (highlighted) and represents an individual colony.

DNA fragments, we previously demonstrated that gSynth™ is a dramatically more accurate synthesis technology. However, the coding of the average human gene is just over 1000 nucleotides and therefore, any new DNA synthesis technology must be able to produce longer DNA fragments.

Our production of a 2.7 kb plasmid represents a significant enablement of gSynth™. By producing gene-length DNA molecules we are able to use gSynth™ to produce proteins and engineer bacteria, making this a critical step forward in supporting the needs of the synthetic biology community.

REFERENCES

1. Lubock, NB. *et al*., "A systematic comparison of error correction enzymes by next-generation sequencing". *Nucleic Acid Research,* 2017, 45(15), 9206-9217. doi: 10.1093/nar/gkx691

2. gSynth. A highly accurate, enzymatic, *de novo* synthesis and gene assembly technology https://www.camenabio.com/assets/media/2019-11-17-application-note.pdf

3. Norrander, J. *et al*., "Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis". *Gene*, 1983, 26(1), 101-106. doi: 10.1016/0378-1119(83)90040-9

4. Perkely, JM. *et al*., "The race for enzymatic DNA synthesis heats up". *Nature*, 2019, 566(7745), 565. doi: 10.1038/d41586-019-00682-0