DNASCRIPT

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

FIMM

Enzymatic DNA Synthesis (EDS) Enables Rapid and Broad-Based Access to Synthetic Oligos Needed for the Genetic Analysis and Functional Characterization of the SARS-CoV-2 Virus

Benoit Derrien[1], Carita Savolainen-Kopra[2], Teemu Smura[3], Tiina Hannunen[4], Arto Orpana[5], Pablo Guardado Calvo[6], Marija Backovic[6], Nelly Lourenço[1], Nadège Tardieu[1], Henri Lachaize[1], Maëllys Kevin[1], Maryke Appel[1], Christine Peponnet[1], Pekka Ellonen[4], Felix Rey[6] and Xavier Godron[1]

[1]DNA Script, Le Kremlin-Bicêtre, France; [2]THL, Helsinki, Finland; [3]University of Helsinki, Department of Virology, Helsinki, Finland; [4]University of Helsinki, FIMM, Helsinki, Finland; [5]HUSLAB, Helsinki University Hospital, Helsinki, Finland and [6]UMR 3569, CNRS Virology Lab, Paris, France

INTRODUCTION

The COVID-19 pandemic sparked a global scientific effort to study the epidemiology, genetics, biochemistry, and evolution of the SARS-CoV-2 virus. One unforeseen impact of this work was a worldwide bottleneck in the supply of synthetic DNA (primers, probes, assay controls, and gene fragments), which currently relies on highly centralized phosphoramidite-based production and third-party logistics.

We have developed a novel Enzymatic DNA Synthesis (EDS) technology, which utilizes a highly engineered TdT enzyme, reversibly terminated nucleotides, and a solid support (Figure 1). This innovation enables same-day, on-demand, in-house DNA production with a benchtop printer in a standard laboratory environment, requiring no specialized technical skills. The SYNTAX System (Figure 2) produces molecular biology-ready oligos, which can be rapidly deployed in new or existing workflows.

Here we report the use of EDS primers in amplicon-based SARS-CoV-2 sequencing as well as in a site-directed mutagenesis strategy to produce and confirm variants of the SARS-CoV-2 spike protein gene for functional studies.

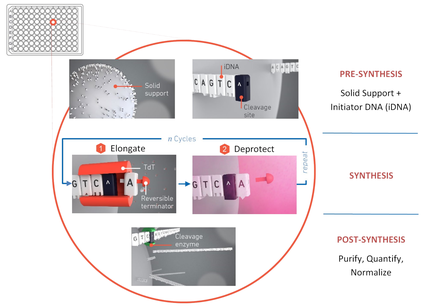

Figure 1. Overview of EDS process. The cyclic 2-step synthesis process is performed on a solid support and is repeated until the longest oligo on the plate has been completed. After the last cycle, all bases downstream from the cleavage site are released. Unlike conventional chemistry, oligos are synthesized in the "natural" 5'⟶ 3' direction and are delivered by default with a 5'-phosphate.

DOWNLOAD TECH NOTE TO LEARN MORE ABOUT EDS

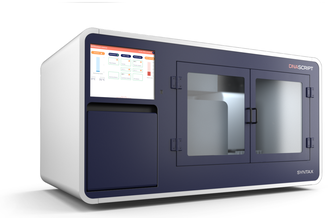

Figure 2. SYNTAX System: the first DNA printer powered by EDS. This fully automated system with touchscreen control requires only 15 minutes of setup time per run. Two runs of up to 96 desalted, quantified, and normalized ready-to-use oligos may be completed in a 24-hr period with a single instrument.

System Capabilities

- Fully automated, walk-away synthesis
- Plug-and-play integration
- Parallel synthesis in 96-well plates
- Up to 96 oligos per run
- 15-minute setup time per run
- Same-day synthesis of 15 – 30 nt oligos enables two runs in 24 hours
- Synthesize longer oligos overnight for next-day use

Oligo Specifications

- 15 – 60 nt de novo oligo synthesis (A, C, G and T)
- Custom iDNA length: 15 – 40 nt
- Ready-to-use oligos for molecular biology and genomics applications
- Default 5'-phosphate
- 200 – 300 pmol per well

SARS-CoV-2 AMPLICON SEQUENCING: MODIFIED PROTOCOL ENABLES SINGLE-DAY SAMPLE PREP

- We have previously demonstrated that primers produced by EDS perform as well as commercial primers (phosphoramidite-based production) in the ARTIC network's SARS-CoV-2 amplicon sequencing protocol.
- The ARTIC protocol (v3 primer panel and library preparation protocol for Illumina® sequencing V.4) was adopted early in 2020—after emergence of the World Health Organization (WHO) Variant of Concern (VOC) Alpha—by the institutions in Finland collectively responsible for SARS-CoV-2 diagnostics and surveillance (THL, HUS, and the University of Helsinki's Department of Virology and FIMM). This broadly accessible protocol enabled the consortium to further coordinate efforts and improve throughput to >3,000 samples/month.
- Bottlenecks remained, primarily due to (i) specialized skills required to prepare ligation-based libraries, and (ii) the two-day turnaround from cDNA synthesis to sequencing-ready library—this despite streamlining library purification through the use of filter plates.
- To eliminate these constraints, the goal was to develop a true one-day method based on PCR only. This required modified ARTIC v3 primers with 5'-overhangs to enable the addition of barcodes and Illumina® adapter sequences during a second round of PCR (Figure 3). Such primers were not readily available from commercial suppliers, and thus were produced with the SYNTAX System. Custom iDNA plates were employed to incorporate the 5'-overhangs. Barcoded primers for the second PCR were produced using a similar approach.
- Modified primers were initially used in the 2-step PCR ("tailed") protocol published by the ARTIC network but yielded poor coverage uniformity and incomplete coverage of AT-rich amplicons 65 – 75. Subsequent introduction of the novel heat-pulse extension (HPE) PCR method (developed at HUS) has enabled complete coverage (Figure 4), rendering the method suitable for routine SARS-CoV-2 sequencing throughout the emergence of additional VOCs (Figure 5).

CONCLUSION: In-house production of primers with the SYNTAX System has enabled rapid development of a fast, high-throughput amplicon sequencing workflow to support Finland's SARS-CoV-2 surveillance program.

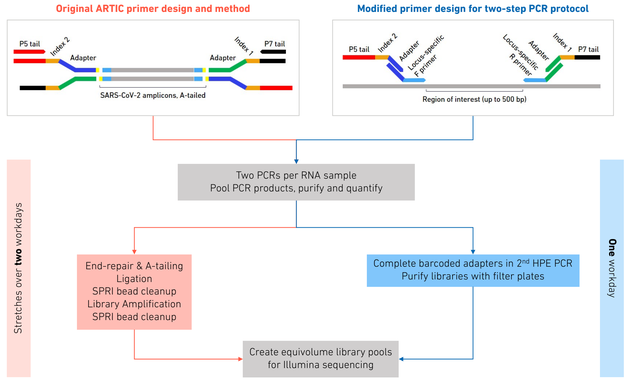

Figure 3. Original and modified ARTIC primer designs and library preparation workflows. The original v3 panel design (left) comprises virus-specific forward and reverse primers. Amplicons are converted to dual-indexed libraries for Illumina paired-end (2 x 250 bp) sequencing using a ligation-based library preparation method that requires two bead-based purification steps. For the 2-step PCR protocol (right), v3 primers with 5'-overhangs were synthesized using the SYNTAX System and SYNTAX 60-Cycle Custom iDNA Kits. Amplicon pools from the first round of SARS-CoV-2 genomic PCR are used directly (without purification) as templates in a second heat pulse extension (HPE) PCR. Full-length libraries are purified in pools of 96 using filter plates. An equivolume sequencing pool (1,000 – 1,500 libraries each) is prepared weekly for sequencing in a single NovaSeq™ 6000 sequencing run (SP flow cell).

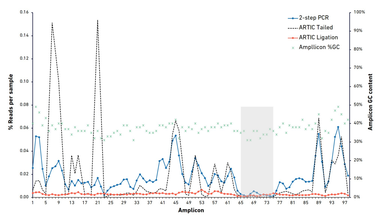

Figure 4. Amplicon coverage with different SARS-CoV-2 amplicon sequencing methods. The 98 overlapping amplicons generated with the v3 primer panel are shown on the x-axis, and the percentage of total reads from a representative NovaSeq 6000 run on the y-axis. The original ligation-based ARTIC library preparation protocol (red line) consistently produced complete and uniform coverage of the 30 kb SARS-CoV-2 genome. With the modified v3 primers (with 5'-overhangs), the ARTIC network's 2-step "tailed" PCR method (dotted black line) resulted in highly uneven coverage and dropout across AT-rich amplicons 65 – 75 (gray box). Complete coverage is achieved with the single-day protocol developed in this study (2-step PCR, filter plate purification; blue line). Further optimization is being done to improve coverage across amplicons 65 – 75.

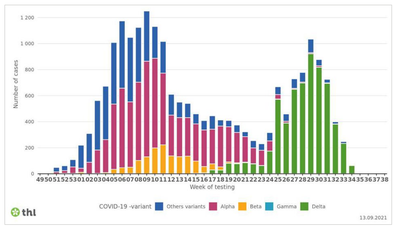

Figure 5. SARS-CoV-2 genetic lines recorded in Finland's national infectious diseases register between mid-December 2020 and end of August 2021. The Alpha variant appeared in Finland late in 2020, necessitating a higher throughput NGS pipeline to support more intensive surveillance efforts. Centralized bioinformatic analysis (focused on phylogeny and transmission) of all samples processed across the collaborating institutions in Helsinki demonstrated dominance of the Alpha variant, even after introduction of WHO VOC Beta early in 2021. The Delta variant was first detected at the end of April 2021 and has since become the dominant variant in Finland, as in many other countries. Data courtesy of THL.

SITE-DIRECTED MUTAGENESIS OF SARS-CoV-2 SPIKE PROTEIN FOR FUNCTIONAL STUDIES

- The emergence of SARS-CoV-2 variants during 2020 quickly created a demand for research materials to study the biology of mutant viruses. Plasmids expressing the spike protein (product of the SARS-CoV-2 *S* gene) of variants of concern (VOCs) are of particular interest, to support investigations focused on transmissibility and immune escape.
- Expression plasmids for functional studies may be obtained via a variety of strategies, ranging from the assembly and cloning of PCR products to the production of full-length synthetic constructs. While the latter appears to be the most advantageous approach, it is expensive and requires long turnaround times from a handful of commercial suppliers.
- In this study, we supported research at the CNRS Virology lab under the direction of Felix Rey by using the SYNTAX System to produce oligos for site-directed mutagenesis of a cloned copy of the 3.8-kb SARS-CoV-2 *S* gene. Oligos were designed to introduce the nine mutations associated with WHO VOC Alpha, as well as an additional mutation, using a multi-site approach (Figure 6).
- In addition to the oligos used for mutagenesis, the SYNTAX System was also used to produce sequencing primers for (i) the confirmation of individual mutations in picked clones, and (ii) tiled sequencing of the entire *S* gene. The experimental workflow and results are summarized in Figure 7, and examples of Sanger confirmation of individual mutations are shown in Figure 8.
- Through a combination of in-house oligo production and a pragmatic approach, expression studies could be initiated within a month from starting the first cloning step. The same approach has since been applied to generate spike protein expression plasmids for functional studies of the rapidly spreading WHO VOC Delta.

CONCLUSIONS: Primers synthesized with the SYNTAX System enabled high success rates in multi-site mutagenesis and were shown to produce high-quality Sanger sequencing data.

On-demand oligo synthesis for site-directed mutagenesis supports a fast and flexible approach for the generation of SARS-CoV-2 spike protein mutants as new VOCs emerge. Modification of existing plasmids is an iterative process that requires only a few days per cycle of mutagenesis and sequence confirmation, compared to weeks of waiting for full-length synthetic constructs.

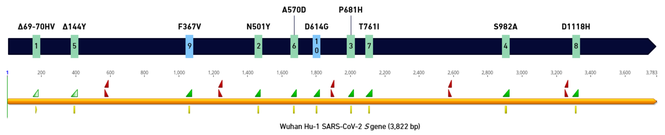

Figure 6. Primers produced by EDS for multi-site mutagenesis of the SARS-CoV-2 spike protein. Mutation sites highlighted in green correspond to the eight amino acid changes in WHO VOC Alpha (formerly B.1.1.7 lineage) known in January 2021. Sites are numbered based on the different multi-site mutagenesis strategies outlined in Figure 7. The F367V mutation (blue site 9) was introduced to revert the cloned *S* gene to the Wuhan Hu-1 reference sequence. The D614G mutation (blue site 10; present in all SARS-CoV-2 variants of concern) was not known when the mutagenesis strategy was first designed, and was added in the third round of mutagenesis. Mutations were introduced with the QuikChange Multi Site-Directed Mutagenesis Kit (Agilent Technologies), as outlined in Figure 7. Green triangles represent oligos used for mutagenesis, whereas red triangles correspond to sequencing primers.

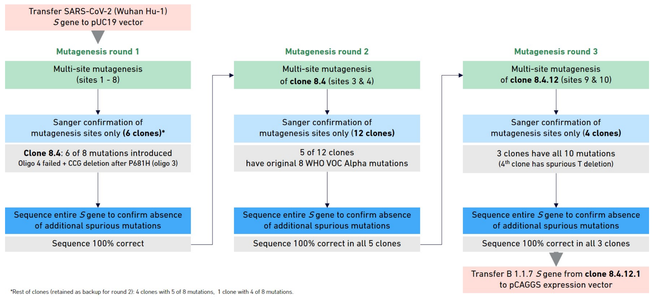

Figure 7: Mutagenesis and variant confirmation strategy and results. All primers used for site-directed mutagenesis (green) and sequencing (blue) were produced using the SYNTAX System. The original mutagenesis strategy (eight simultaneous mutations) was ambitious. Instead of sequencing more clones from round one, it was deemed more productive to perform a second, less complex round of mutagenesis to "fix" partially successful clones from the first round. The third round of mutagenesis was used to introduce additional mutations that were not included in the original design. Only a small number of clones had to be sequenced in each round in order to move ahead with the next step (pertinent results summarized in grey). The entire process, including the subcloning steps (red), took approximately one month to complete.

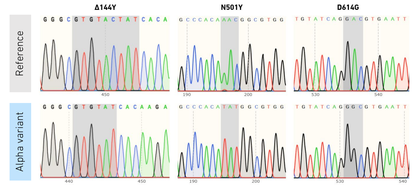

Figure 8. Representative Sanger sequencing results. Representative data from three of the ten sites targeted to generate a SARS-CoV-2 Alpha variant spike protein gene from the Wuhan Hu-1 sequence confirm successful site-directed mutagenesis. Results validate that EDS primers support high-quality Sanger sequencing.