# Supplementary Information

## Abstraction hierarchy to define biofoundry workflows and operations for interoperable synthetic biology research and applications

Haseong Kim[1,2,3,4], Nathan J. Hillson[5,6], Byung-Kwan Cho[4,7], Bong Hyun Sung[1,2,3,4], Dae-Hee Lee[1,2,3,4,8], Dong-Myung Kim[9], Min-Kyu Oh[10], Matthew Wook Chang[11,12,13,14], Yong-Su Jin[15,16], Susan J Rosser[17,18], Peter Vegh[17,18], Rennos Fragkoudis[17,18], Rosalind Le Feuvre[19], Nigel S. Scrutton[19], Marko Storch[20,21], Wonjae Seong[1,2], Paul S Freemont*[20,21,22], Seung-Goo Lee*[1,2,3,4]

[1]Korea Biofoundry, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon 34141, Republic of Korea

[2]Synthetic Biology Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon 34141, Republic of Korea

[3]Department of Biosystems and Bioengineering, KRIBB School of Biotechnology, University of Science and Technology (UST), Daejeon 34113, Republic of Korea

[4]Graduate School of Engineering Biology, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea

[5]DOE Agile BioFoundry, Emeryville, CA 94608 U.S.A.

[6]Biological Systems and Engineering Division, Lawrence Berkeley National Lab, Berkeley, CA 94720 U.S.A

[7]Department of Biological Sciences, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea

[8]Department of Integrative Biotechnology, College of Biotechnology and Bioengineering, Sungkyunkwan University, Suwon, Gyeonggi-do 16419, Republic of Korea

[9]Department of Chemical Engineering and Applied Chemistry, Chungnam National University, 99 Daehak-ro, Daejeon, 34134, Republic of Korea

[10]Department of Chemical and Biological Engineering, Korea University, Seongbuk-gu, Seoul 02841, Republic of Korea

[11]NUS Synthetic Biology for Clinical and Technological Innovation (SynCTI), National University of Singapore, Singapore

[12]Synthetic Biology Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

[13]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

[14]National Centre for Engineering Biology (NCEB), Singapore, Singapore

[15]Department of Food Science and Human Nutrition, University of Illinois at Urbana-Champaign, Urbana, IL 61801, U.S.A.

[16]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, U.S.A.

[17]School of Biological Sciences, University of Edinburgh, Edinburgh, United Kingdom

[18]Edinburgh Genome Foundry, Centre for Engineering Biology, University of Edinburgh, Edinburgh, United Kingdom

[19]Manchester Institute of Biotechnology, University of Manchester (UoM), United Kingdom

[20]Department of Infectious Disease, Imperial College London, London, United Kingdom

[21]London Biofoundry, Innovation Hub, White City Campus, Imperial College London, United Kingdom

[22]Care Research, and Technology Centre, UK Dementia Research Institute, London, United Kingdom

Table S1. Summary of 58 biofoundry workflows.

| ID | Stage | Workflows | Description |
|---|---|---|---|
| WD010 | Design | General Design of Experiment | This is a general-purpose experimental design workflow that utilizes the Design of Experiment approach. Variables can be independently designed from domains. |
| WD020 | Design | Adaptive Laboratory Evolution Design | Top-down design utilizing random mutations and artificial evolution. Designs may involve UV radiation or environmental stress conditions for inducing artificial sequence mutation. |
| WD030 | Design | Growth Media Design | Growth media design for strain culture, creating optimal growth conditions based on data-driven experimental design to optimize strain growth and productivity. A growth media composition database is established for different organisms, strains, and experiments. |
| WD040 | Design | Parallel Cell Culture/Fermentation Design | Designing conditions for large-scale culturing of proteins, enzymes, or strain activity tests. In fermentation, the Design of Experiment is used to explore initial scale-up conditions for the selected strains. Parameters such as media, temperature, pH expression levels, aeration, and more are optimized. |
| WD050 | Design | DNA Oligomer Pool Design | Designing an oligomer pool for assembling target DNA (Gene, Pathway). The oligomers used here are typically 50–300 bp in length. |
| WD060 | Design | Genetic Circuit Design | Designing genetic circuits for a specific purpose such as biosensors detecting metabolites or logic gate-based cell control. Part sequences stored in a part registry are utilized. This design process includes predictive modeling using quantified DNA parts (e.g., Cello) or predictive tools like RBS calculator for prokaryotic systems. |
| WD070 | Design | Vector Design | The design process for constructing DNA in the form of a plasmid vector, BAC, YAC, HAC, etc. |

| | | | |
|---|---|---|---|
| WD080 | Design | Artificial Genome Design | A bottom-up or middle-level design for creating a novel genome (e.g. Decompression, codon redesign). Data such as host metabolism, essential gene information, and gene/metabolic pathway databases can be used. |
| WD090 | Design | Genome Editing Design | CRISPR-based genome editing design. Factors such as types of CRISPRs and host species, as well as off-target effects, and thermodynamics and folding of RNA tertiary structure are considered when designing target gRNAs, using dedicated software. |
| WD100 | Design | Protein Library Design | Library design for optimizing protein activity, specificity, expression, etc. Various tools (e.g., Alphafold, ProteinMPNN) are used to facilitate protein design. This includes designing random mutation libraries. |
| WD110 | Design | *De novo* Protein/Enzyme Design | *De novo* protein or enzyme design using deep learning tools like Rfdiffusion, ProteinMPNN, and Alphafold. It includes *de novo* design of biocatalysts. |
| WD120 | Design | Retrosynthetic Pathway Design | Use of design tools to find pathways to a desired metabolic reaction to the product chemical (e.g., RetroPath). It supports *de novo* design of optimized pathways and allows the discovery of novel pathways through retrosynthesis. Design tools for database search for part selection of enzymes to complete the desired reaction pathway (e.g., Selenzyme). |
| WD130 | Design | Pathway Library Design | Designing libraries for optimizing protein function or metabolic pathways. Components stored in a pre-built DNA parts bank are used. Libraries for pathway expression optimization can be generated by incorporating promoters, RBS, terminators, and other regulatory elements. |
| WB010 | Build | DNA Oligomer Assembly | Workflow for assembling DNA oligomers into sequences several kb in length. This procedure can begins with an oligomer pool. |

| | | | |
|---|---|---|---|
| WB020 | Build | DNA Library Construction | Workflow for creating designed DNA mutant libraries, metagenomic libraries, or pathway libraries. It includes methods such as combinatorial and random mutagenesis libraries. |
| WB030 | Build | DNA Assembly | Workflow for assembling double-stranded DNA fragments into several kb sequences or larger. This includes assembling several DNA fragments such as parts or operon-level sequences, in a designed order. |
| WB040 | Build | DNA Purification | This workflow refines crude DNA extracts to achieve high purity suitable for downstream applications. It typically involves methods like column chromatography, magnetic beads, or precipitation to remove contaminants such as proteins, RNA, and salts. |
| WB045 | Build | DNA Extraction | This workflow focuses on releasing DNA from biological samples (e.g., cultured cells, tissues) through lysis and initial separation from major cellular components like proteins and lipids. |
| WB050 | Build | RNA Extraction | The isolation of RNA from biological samples, such as cells or tissues, to enable downstream applications like gene expression analysis, reverse transcription PCR (RT-PCR), or next-generation sequencing (NGS). |
| WB060 | Build | DNA Multiplexing | Workflow for selecting cells grown on solid or liquid media and assigning arbitrary barcodes for identification. Amplified DNA with barcode primers can be obtained from cultured cells, and the barcoded DNA is pooled for the next step of NGS sequencing. |
| WB070 | Build | Cell-free Mixture Preparation | Workflow for preparing master solutions for cell-free reactions. This involves collecting and processing large-scale cultures of specific strains to create the cell extract for cell-free systems, including stages such as cell lysis, purification, and separation. |

| | | | |
|---|---|---|---|
| WB080 | Build | Cell-free Protein/Enzyme Expression | Workflow for mixing target DNA with cell-free reaction reagents and producing quantities of proteins or enzymes under specific conditions suitable for high-throughput assays. |
| WB090 | Build | Protein Purification | Workflow for purifying target proteins or enzymes to high purity. This can be conducted by automated equipment capable of handling 96-well plates for high-throughput purification. |
| WB100 | Build | Growth Media Preparation and Sterilization | Workflow for large-scale production, sterilization, and aseptic storage of designed solid and liquid media, including sterilizing media and fermentation equipment. |
| WB110 | Build | Competent Cell Construction | Workflow for creating competent cells for transformation. This can be done manually in bulk or using automated equipment to make plate-based competent cells. |
| WB120 | Build | Biology-mediated DNA Transfers | Workflow for transforming designed vector plasmid into cells. This includes 96/384-well plate-based automated or semi-automated transformation procedures. It includes conjugation or other DNA transfer protocols (e.g. phage-mediated). |
| WB130 | Build | Solid Media Cell Culture | Workflow for culturing cells on solid media. It includes post-transformation growth, activity screening, or single-cell/colony isolation from solid medium. |
| WB140 | Build | Liquid Media Cell Culture | Workflow for growing cells in liquid media. It includes inoculum culture and subsequent batch culture processes in liquid medium. |

| | | | |
|---|---|---|---|
| WB150 | Build | PCR-based Target Amplification | This workflow utilizes designed primers and PCR to specifically amplify a target gene sequence from complex templates such as genomic DNA or metagenomic samples, enabling gene screening and retrieval. |
| WT010 | Test | Nucleotide Sequencing | Workflow for sample preparation through native barcode multiplexing and library prep, using NGS equipment to sequence the target DNA and produce fastq files. It includes transcriptomics like RNA. |
| WT020 | Test | Protein Expression Measurement | Quantifying the expression levels of target proteins or enzymes. It includes measurements using techniques such as gel electrophoresis or automated capillary electrophoresis systems. High-throughput proteomics approaches like LC-MS can be integrated. It allows detailed identification and quantification of protein expression and post-translational modifications |
| WT030 | Test | Protein/Enzyme Activity Measurement | Workflow for measuring the activity of purified proteins/enzymes using general or specific methods (e.g., biosensors, chromatography, pNP). The method used is largely dependent on the specific protein/enzyme activity. |
| WT040 | Test | Parallel Cell-free Protein/Enzyme Reaction | Workflow for expressing and simultaneously measuring the activity of target proteins/enzymes in a cell-free reaction system under specific conditions. |
| WT050 | Test | Sample Pretreatment | Workflow for separating and preprocessing metabolites from cultured media using centrifugation, cell lysis, and cell removal steps before purification and analysis. It can be used in the process of proteomics, lipidomics, transcriptomics, and etc. |
| WT060 | Test | Metabolite Measurement | Workflow for quantifying metabolites using GC-MS, LC-MS, and spectroscopy after high-throughput pretreatment. Includes fast measurement of single or complex components and unknown compound analysis. |

| | | | |
|---|---|---|---|
| WT070 | Test | High-throughput Single Metabolite Measurement | Workflow for analyzing and measuring a single type metabolite in a well-plate using techniques like biosensors or other biochemical assays such as high-throughput LC-MS. |
| WT080 | Test | Image Analysis | Workflow for analyzing cell growth, morphology, chromatin structure, organelle, and sub-cellular protein localization using a high-throughput optical device, such as a microscope. This workflow includes sample preparation steps for imaging analysis. |
| WT090 | Test | High-speed Cell Sorting | Workflow for sorting cells based on target metabolite or cell activity using genetic circuits that convert the activity into a detectable signal. |
| WT100 | Test | Micro-scale Parallel Cell Culture | Workflow for culturing cells in 0.2 ml–1.2 ml 96 deep well plates. The process includes treatments to induce protein or cell activity. |
| WT110 | Test | Micro-scale Parallel Cell Fermentation | Workflow for performing fermentation in 0.8 ml–2.5 ml volumes while monitoring key parameters (OD, pH, temperature, DO). |
| WT120 | Test | Parallel Cell Fermentation | Workflow for performing fermentation in 15 ml–250 ml volumes while monitoring key parameters in real-time (OD, pH, temperature, DO). |
| WT130 | Test | Parallel Mammalian Cell Fermentation | Workflow for culturing animal cells in 15 ml volumes to explore conditions for maximizing protein production. Monitoring of key parameters such as OD, pH, temperature, DO is performed in real-time. |
| WT140 | Test | Lab-scale Fermentation | Workflow for performing fermentations of less than 10L while monitoring key parameters such as pH, temperature, dissolved oxygen (DO), etc. |

| | | | |
|---|---|---|---|
| WT150 | Test | Pilot-scale Fermentation | Workflow for performing fermentations between 10L and 500L while monitoring key parameters such as pH, temperature, dissolved oxygen (DO), etc. |
| WT160 | Test | Industrial-scale Fermentation | Workflow for performing fermentations of more than 500L while monitoring key parameters such as pH, temperature, dissolved oxygen (DO), etc. |
| WL010 | Learn | Sequence Variant Analysis | Workflow for verifying the sequence of template DNA (target genes, pathways, plasmids). This workflow is required for activities such as gene cloning and assembly. This workflow includes the comparison analysis of sequence variants. |
| WL020 | Learn | Genome Resequencing Analysis | Workflow for analyzing SNPs and other genome variations in organisms with reference genomes. |
| WL030 | Learn | *De novo* Genome Analysis | Workflow for analyzing the genome of new organisms without reference genomes. It includes *de novo* genome assembly from NGS data |
| WL040 | Learn | Metagenomic Analysis | Workflow for analyzing large volumes of metagenomic sequence data. It includes raw data collection, gene/strain identification, and functional predictions. Machine learning or AI algorithms can be used to identify candidate enzymes from metagenomes. |
| WL050 | Learn | Transcriptome Analysis | Workflow for analyzing transcriptomes (mRNA) from target organisms under different conditions. It includes mRNA sequence analysis and DEG. |
| WL055 | Learn | Single Cell Analysis | This workflow focuses on analyzing individual cells to understand cellular heterogeneity and functional characteristics. It includes techniques such as single-cell RNA sequencing, single-cell ATAC-seq, and other omics approaches. |

| | | | |
|---|---|---|---|
| WL060 | Learn | Metabolic Pathway Optimization Model Development | Workflow for analyzing measured metabolite data, including preprocessing and flux analysis. Specialized software can be used. Machine learning and AI models can be developed using labeled data from metabolic pathway gene sequences and the corresponding metabolite products. |
| WL070 | Learn | Phenotypic Data Analysis | Processing and analyzing phenotypic data, including growth rates, morphological traits, metabolic activity, and image-based phenotypic data. The workflow integrates statistical analysis, image processing, and machine learning to extract quantitative features, identify patterns, and establish phenotype-genotype relationships. |
| WL080 | Learn | Protein/Enzyme Activity Optimization Model Development | Workflow for developing models to optimize the activity of target proteins by utilizing phenotypic and sequence data obtained from protein/enzyme expression and activity measurements. It includes protein structure and function analysis, leveraging pre-trained models or publicly available models (e.g., Alphafold2, Rosettafold, MPNN). This can be used for designing libraries of new functional proteins. |
| WL090 | Learn | Fermentation Optimization Model Development | Workflow for exploring optimal conditions for target compound production based on fermentation data for a given strain. This includes simulation models based on reaction formulas from the process. |
| WL100 | Learn | Foundation Model Development | Workflow for training foundation models using large sequence datasets such as protein DBs or metagenomic CDS. |

The modularized workflows can be arranged sequentially to perform arbitrary services. Figure S1 represents an example of the construction of a protein library as a service (D, B, T, L). It starts from protein library design with tools such as ProteinMPNN[1] (WD110). The designed sequences can be constructed by assembling DNA oligomers so oligomer pools are required to be designed by a specific software tool (WD050). After designing an appropriate vector plasmid (WD070), we can conduct a series of build workflows, Oligomer assembly (WB010), Transformation (WB120), Solid Cell Culture (WB130), Liquid Media Cell Culture (WB140), DNA Multiplexing (WB060), and DNA Purification (WB040). Then, we can perform Nucleotide Sequencing (WT010) followed by Sequence Variant Analysis (WL010). The workflows listed in Table S1 can be expanded and updated to account for general usability and modular properties.

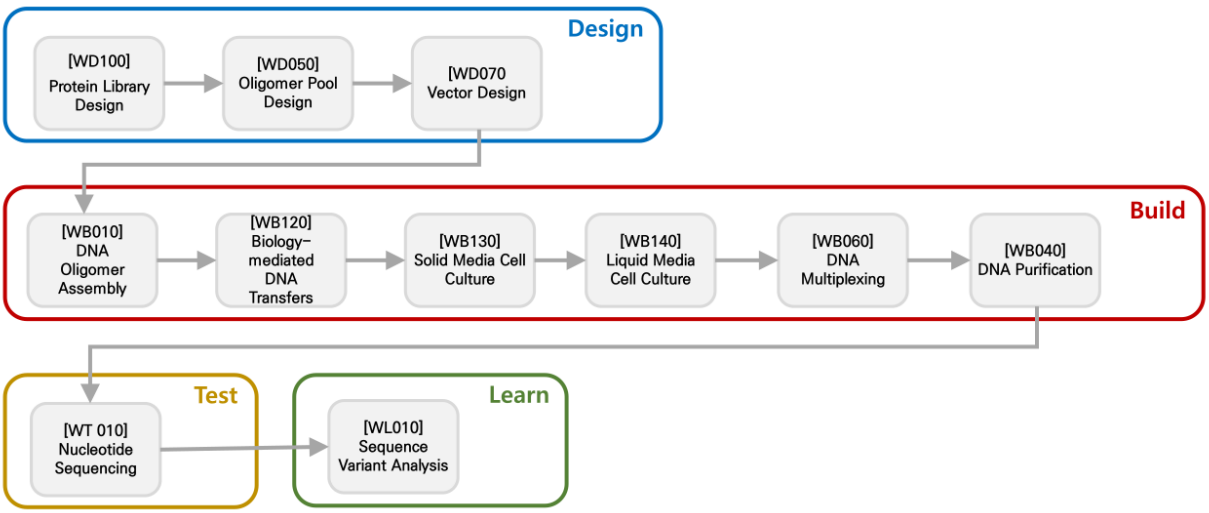


Figure S1. Example workflows for constructing a protein library as a service. This service consists of three Design workflows (blue), six Build workflows (red), and both the Test (yellow) and Learn (green) components, each consisting of a single workflow.

Table S2. List of unit-operations for hardware

| ID | Unit Operation | Equipment | Description |
|---|---|---|---|
| UH010 | Liquid Handling | Multiple dispenser system | Basic and fundamental liquid sample operations in laboratory processes such as reagent preparation, sample distribution, dilution, mixing, and washing. |
| UH015 | Bulk Liquid Dispenser | Bulk reagent dispenser | Rapid and accurate dispensing of larger volumes of a single liquid reagent such as growth media and buffers into multi-well plates. Optimized for speed and efficiency in plate preparation tasks. |
| UH020 | 96 Channel Liquid Handling | NGS library preparation system | High-throughput, simultaneous dispensing or transferring of liquids across a 96-well plate format. It includes NGS library preparation and magnetic bead-based purification. |
| UH030 | Nanoliter Liquid Dispenser | Acoustic liquid dispenser | Specialized for high-precision dispensing of extremely small liquid volumes, typically in the nanoliter range. It reduces reagent usage, minimizes waste, and allows for scalable, cost-effective workflows. |
| UH040 | Desktop Liquid Handling | Desktop liquid handling system | Small desktop liquid handling system for small-scale automated experiments or educational use. |
| UH050 | Single Cell Sequencing Preparation | Droplet-based high-throughput single-cell platforms | Preparation of single cells for single-cell analysis. It includes processes such as encapsulating single cells with reagents for simultaneous isolation and library preparation. |
| UH060 | Colony Picking | Colony picker | The process of isolating individual bacterial or yeast colonies from an agar plate and transferring them to a liquid culture or multi-well plate for downstream applications. |
| UH070 | Cell Sorting | High-speed flow cytometry | A process used to isolate and categorize cells based on their physical or biological properties, such as size, shape, fluorescence, or other markers. It is used for single-cell analysis, cell line development, or functional studies. |
| UH080 | Cell Lysis | Bead mill or ultrasonic homogenizers | The process of breaking open cells to release their internal components, such as DNA, RNA, proteins, or metabolites, for downstream processing. |
| UH090 | Electroporation | Electroporator | Technique for introducing foreign molecules, such as DNA and RNA, into cells by applying an electrical field. |

| | | | |
|---|---|---|---|
| UH100 | Thermocycling | Thermal cycler | The process of repeatedly heating and cooling a sample through defined temperature cycles to facilitate reactions. It is used in the workflows such as DNA assembly and DNA/RNA amplification. |
| UH110 | Real-time PCR | qPCR | Amplifying and simultaneously quantifying a specific DNA or RNA sequence. |
| UH120 | Plate Handling | Plate handling robot arm | Connecting workflows/unit operations among automated equipment through plate transportation. |
| UH130 | Sealing | Microplate sealer | Plate sealing for PCR, culturing, storing, and etc. |
| UH140 | Peeling | Microplate peeler | Plate cover removal after PCR, culturing, storing, and etc. |
| UH150 | Capping Decapping | Tube capper and decapper | Automatic opening and closing of sample tube caps from storage. |
| UH160 | Sample Storage | Sample storage system | Storing DNA or cell samples, with automated sample retrieval and storage system. |
| UH170 | Plate Storage | Plate hotel | Storing or retrieving plates with an automated system for high-throughput experiments. It improves workflow throughput with integration of automated robotic arms. |
| UH180 | Incubation | Incubator | Maintaining specific conditions for cells or chemical reactions to promote growth or activity. |
| UH190 | HT Aerobic Fermentation | High throughput microbial fermentation system | Process of cultivating microbial or cell cultures under oxygen-rich conditions in a parallelized, automated system. It is generally designed for small-scale (10~250mL) fermentations. |
| UH200 | HT Anaerobic Fermentation | High throughput anaerobic fermentation system | Process of cultivating microbial or cell cultures under oxygen-free conditions in a parallelized, automated system. It is generally designed for small-scale (10~250mL) fermentations. |
| UH210 | Microbioreactor Fermentation | Microbioreactor fermentation system | Process of cultivating microbial or cell cultures covered in micro-scale bioreactors (<10 mL) equipped with advanced monitoring and control capabilities. It is mainly used for high-throughput screening of strains, media, and process parameters. |
| UH220 | Bioreactor Fermentation | Bioreactor fermentation system | Process of cultivating microbial or cell cultures in bioreactors (Liter scale). Bioreactors can operate in batch, fed-batch, or continuous modes and are commonly used in both research and industrial settings. |

| | | | |
|---|---|---|---|
| UH230 | Nucleic Acid Fragment Analysis | Nucleic acid fragment analyzer | To separate, identify, and measure fragments of nucleic acids based on their size, concentration, or sequence characteristics. |
| UH240 | Protein Fragment Analysis | Protein fragment analyzer | To separate, identify, and characterize fragments of proteins for the study of structure, size, modifications, interactions, or expressions. |
| UH250 | Nucleic Acid Purification | Nucleic acid purification system | The process of isolating DNA or RNA from biological samples using an automated device. |
| UH255 | Centrifuge | Centrifuge | Separating components of different densities within a liquid sample by applying centrifugal force. Used for pelleting cells, clarifying solutions, separating precipitates, and sample preparation steps. |
| UH260 | Short-read Sequence Analysis | Short read sequence analyzer | Short-read-based sequencing using NGS technologies. It is essential for high-throughput genomic studies and variant detection. |
| UH270 | Long-read Sequence Analysis | Long read sequence analyzer | Long-read-based sequencing using platforms such as Nanopore or PacBio. It provides comprehensive insights into complex genomic regions and structural variations. |
| UH280 | Sequence Quality Control | Sequence quality control system | Evaluating single-cell quality for single-cell analysis. |
| UH290 | LC-MS-MS | LC-MS/MS | High-performance liquid chromatography coupled with tandem mass spectrometry. Provides highly sensitive and selective quantification and structural information by separating components with LC and analyzing specific ions through two stages of mass analysis. Ideal for targeted analysis in complex matrices like biological fluids (e.g., quantifying drug metabolites, analyzing specific peptides). |
| UH300 | LC-MS | LC-MS | Liquid chromatography coupled with mass spectrometry. Separates components using LC and identifies them based on their mass-to-charge ratio. Offers better identification than LC alone but less structural info and selectivity than LC-MS/MS. Used for general profiling (e.g., metabolic profiling, identifying major components in natural extracts). |
| UH310 | HPLC | HPLC | High-performance liquid chromatography. It separates, identifies, and quantifies components in a liquid mixture using high pressure and a detector such as UV-Vis. Common for routine analysis, purity checks, and QC (e.g., |

| | | | |
|---|---|---|---|
| | | | analyzing active pharmaceutical ingredient purity, quantifying vitamins in supplements). |
| UH320 | UPLC | UPLC | Ultra-high-performance liquid chromatography. Similar to HPLC but uses smaller particle columns and higher pressures for significantly faster analysis times and higher resolution. Advantageous for high-throughput screening or separating complex mixtures that are difficult for HPLC. |
| UH330 | GC | GC | Gas chromatography. It separates volatile or semi-volatile compounds in their gaseous phase based on their interaction with a stationary phase. Typically uses detectors like FID or TCD. Used for analyzing volatile organic compounds, solvent purity, and flavor/fragrance profiling (e.g., analyzing aroma compounds, measuring ethanol content). |
| UH340 | GC-MS | GC-MS | Gas chromatography coupled with mass spectrometry. It separates volatile components using GC and identifies them based on their mass spectra, often matched against libraries. It provides much better identification capabilities than GC alone. It is widely used in environmental analysis, forensics, and food safety. |
| UH350 | GC-MS-MS | GC-MS/MS | Gas chromatography coupled with tandem mass spectrometry. Offers enhanced sensitivity and selectivity compared to GC-MS by using two stages of mass analysis after GC separation. Ideal for trace analysis in complex matrices. |
| UH355 | SPE-MS-MS | SPE-MS/MS | Solid-Phase Extraction (SPE) system coupled with Tandem Mass Spectrometer. Combines automated SPE for sample cleanup, concentration, and matrix removal with Tandem Mass Spectrometry for highly sensitive and selective quantification. It often used for rapid screening or targeted quantification where extensive chromatography is not needed (e.g., quantifying specific pesticide residues in food extracts after SPE cleanup). |
| UH360 | FPLC | FPLC | Fast protein liquid chromatography. A type of liquid chromatography optimized for the purification of biomolecules like proteins and nucleic acids, typically using lower pressures and biocompatible materials compared to HPLC. Used for preparative scale separation. |
| UH370 | Oligomer Synthesis | Oligomer synthesizer | Parallel synthesis of oligomers using chemical methods or Tdt enzyme-based commercial technologies. It supports |

| | | | |
|---|---|---|---|
| | | | the production of custom DNA or RNA sequences for research and industrial applications. |
| UH380 | Microplate Reading | Multi-functional microplate reader | Quantification of protein/cell activity by measuring fluorescence, OD, etc. on 96 or 384-well plates. It is essential for high-throughput screening and assay development. |
| UH390 | Microscopy Imaging | Microscopy | Imaging for measuring activity in cells, such as animal cells. It provides detailed visualization for cell biology and histology studies. |
| UH400 | Manual | Manual operations | A general process of experiment including preparation of reagents, labwares, and any manual steps. It encompasses traditional laboratory techniques and procedures. |

Table S3. List of unit-operations for software

| ID | Unit Operations | Software | Description |
|---|---|---|---|
| US010 | DNA Oligomer Pool Design | Dsembler, DNAWorks | A software package that designs DNA oligomers which are pooled in a single tube. It optimizes the selection and combination of oligomers for efficient DNA assembly and synthesis. |
| US020 | Primer Design | SnapGene, Primer3, OligoAnalyzer | Designing primers regarding melting temperature and structure. The primers are used for mutant generation, PCR primers, and other molecular biology applications, ensuring specificity and efficiency in amplification. |
| US030 | Vector Design | VectorNTI, SnapGene, Geneious | Designing vector maps regarding inserts and a plasmid backbone. This might be possible to include primer design and DNA assembly processes, facilitating the construction of functional genetic vectors for cloning and expression. |
| US040 | Sequence Optimization | GeneOptimizer, JCat | Modifying codon usage of a DNA sequence to maximize protein expression in a specific host. This software ensures optimal translation efficiency and protein yield by adapting sequences to host-specific codon preferences. |
| US050 | Synthesis Screening | UltraSEQ, Common Mechanism, FAST-NA | Tools for screening of potentially dangerous DNA sequences. These tools help ensure biosafety by identifying sequences that may pose risks in synthetic biology applications. |
| US060 | Structure based Sequence Generation | ProteinMPNN | Generating sequences based on protein structures using AI models. This software aids in designing novel proteins with desired structural and functional properties. |
| US070 | Protein Structure Prediction | Alphafold, Rosettafold | Predicting protein structures using AI models. These tools provide insights into protein folding and stability, supporting protein engineering and drug discovery efforts. |
| US080 | Protein Structure Generation | Rfdiffusion | Generating protein structures using AI models. This software facilitates the design of new proteins and enzymes with specific catalytic or binding functions. |
| US090 | Retrosynthetic Pathway Design | RetroPath2.0, ECREACT, BioNavi-NP | To predict biosynthetic pathways using tools for checking reaction feasibilities and for novel pathway discovery. These tools support metabolic engineering by identifying efficient routes for chemical synthesis. |

| | | | |
|---|---|---|---|
| US100 | Enzyme Identification | DeepEC, Selenzyme, SoluProt | To search for enzymes from database or to predict enzyme properties such as reactivities for selecting proper enzymes in pathways. This software aids in enzyme discovery and characterization for biocatalysis. |
| US110 | Sequence Alignment | BLAST, MUSCLE | Exploring and comparing sequence similarity using alignment algorithm. These tools are essential for identifying homologous sequences and understanding evolutionary relationships. |
| US120 | Sequence Trimming and Filtering | Trimmomatic, Cutadapt, Porechop, Filtlong | Preprocessing for removing low-quality long/short-read sequences. This step is crucial for ensuring data quality in sequencing projects. |
| US130 | Sequence Mapping and Alignment | BWA, Bowtie2, Minimap2, GraphMap | Mapping long/short-read sequences to reference sequences. These tools are used for genome assembly, variant calling, and transcriptomics. |
| US140 | Sequence Assembly | Velvet, SOAP, Quast, Canu, Flye | Assembling long/short-read sequences for complete gene, pathway, and chromosome. This software supports the reconstruction of genomes and metagenomes. |
| US145 | Metagenomic Assembly | MetaSPAdes, MEGAHIT | Assembling metagenomic data to reconstruct genomes from complex microbial communities. This software supports environmental and clinical metagenomics studies. |
| US150 | Sequence QC | FastQC, MultiQC, NanoPlot, pycoQC | Performing quality control (QC) on Long/short-read fast5 and fastq files. QC is essential for identifying and correcting errors in sequencing data. |
| US160 | Demultiplexing | bcl-convert, Guppy | Separating NGS reads based on native or user-defined barcodes. This process is critical for handling multiplexed sequencing data. |
| US170 | Variant Calling | GATK, bcftools, Sniffles, Longshot | Detecting variants based on read mapping data. These tools are used for identifying SNPs, indels, and structural variants in genomic data. |
| US180 | RNA-Seq Analysis | DESeq2/EdgeR (R), Galaxy, HISAt2 | Processing and analyzing transcriptomic data to quantify gene expression levels, identify splice variants, and detect differential gene expression. This software supports functional genomics studies. |
| US185 | Gene Set Enrichment Analysis | GSEA, DAVID | Analyzing gene expression data to identify enriched biological pathways. This software supports functional genomics and systems biology research. |

| | | | |
|---|---|---|---|
| US190 | Proteomics Data Analysis | MaxQuant, Perseus, Proteome Discoverer | Processing and interpreting data from mass spectrometry to identify and quantify proteins, understand modifications, and assess protein interactions. |
| US200 | Phylogenetic Analysis | MEGA, PhyML | Determining the evolutionary relationships among species or sequences by constructing phylogenetic trees based on sequence similarities and differences. |
| US210 | Metabolic Flux Analysis | COBRA Toolbox, FBA, CellNetAnalyzer | Analyzing the flow of metabolites through metabolic pathways, providing insights into cellular metabolism and pathway optimization. |
| US220 | Deep learning Data Preparation | pytorch::DataLoader, Huggingface::datasets | Preparing and batching datasets for AI model training and evaluation. This software supports machine learning workflows in bioinformatics. |
| US230 | Sequence Embedding | ProtT5, ProtBERT, ESM | Biological sequence embedding procedure. This process is used for transforming sequences into numerical representations for machine learning applications. |
| US240 | Deep Learning Model Training | CNN, LSTM, Transformer | Model training procedure using training data. This software supports the development of AI models for various bioinformatics tasks. |
| US250 | Model Evaluation | scikit-learn, TensorBoard | Utilizing model evaluation metrics (accuracy, precision, recall, F1 score, etc.). This software is used for assessing the performance of machine learning models. |
| US260 | Hyperparameter Tuning | Optuna, HyperOpt, Bayesian opt. | Efficiently exploring the search space using Bayesian optimization techniques. This software supports the optimization of machine learning and deep learning models. |
| US270 | Model Deployment | TorchScript, Flask, FastAPI | Deploying trained models as services. This software supports the integration of AI models into production environments. |
| US280 | Monitoring and Reporting | Prometheus, Grafana | Monitoring and visualizing performance and resource usage of AI models. |
| US290 | Phenotype Data Preprocessing | R, Python | A general preprocess of measured and collected phenotype data. It involves cleaning, organizing, and transforming raw phenotype datasets for downstream analysis. |

| | | | |
|---|---|---|---|
| US300 | XCMS Analysis | XCMS (R) | Analyzing and visualizing chromatographically separated and single-spectra mass spectral data. This software supports metabolomics research. |
| US310 | Flow Cytometry Analysis | flowcore, flowworkspace (R), flowJo | Analyzing and visualizing flow cytometry data. This software supports immunology and cell biology research. |
| US320 | DNA Assembly Simulator | pyDNA (python) | Simulating DNA assembly such as Golden Gate and Gibson for increasing assembly success rate. This software supports synthetic biology and genetic engineering. |
| US325 | Gene Editing Simulation | CRISPResso, CHOPCHOP | Simulating gene editing outcomes using CRISPR technology. This software helps predict off-target effects and optimize guide RNA design for precise genome editing. |
| US330 | Well Plate Mapping | Well plate mapping software | Software for mapping well positions from source plates to destination plates. |
| US340 | Computation | Computer | A general process of data collection, preprocessing, and analysis steps. |

**Tier 2 Example: (WB010) DNA Oligomer Assembly workflow**

The following table shows an example of unit operations for the DNA Oligomer Assembly workflow (WB010) (Table S4). We used the protocol from the genome synthesis study from Hutchison C. *et. al.*[2] The protocol describes the general process of assembling DNA oligomer fragments to create longer DNA sequences, starting with an oligomer pool containing the oligomer fragments as the initial starting material, and performing steps such as assembly, amplification, error correction, and recovery. Title and Description of each unit operation are more specialized and detailed within the experiment that a user performs. Note that in the case of fully automated biofoundries, additional unit-operations such as Plate Handling, Sealing, and Peeling will be required for this workflow.

Table S4. Details of unit-operations for the DNA Oligomer Assembly (WB010) workflow

| Unit operation ID | **Unit operation** | **Description** |
|---|---|---|
| UH010 | Liquid Handling | A DNA oligomer pool and primers resuspended in TE buffer pH8.0 with appropriate concentrations and dispense into a 384-well source plate. |
| UH010 | Liquid Handling | Mix polymerase master mix and DW in the source plate. |
| UH030 | Nanoliter Liquid Dispenser | Dispense oligomers from the source plate to the destination plate. |
| UH100 | Thermocycling | Perform PCA with the prepared PCR plate. The parameters are 98°C for 2 min, 30 cycles of 98°C for 30s and 65°C for 6 min (increasing 15 sec/cycle), followed by 72°C incubation for 5 min. |
| UH230 | Nucleic Acid Fragment Analysis | Analyze the PCR product along with a standard DNA ladder. |
| UH100 | Thermocycling | Perform PCA reaction with 98°C for 2 min, 2°C/s to 85°C, 85°C for 2 min, 0.1°C/s to 25°C, 25°C for 2 min for storage. |
| UH010 | Liquid Handling | Prepare samples for the storage |
| UH160 | Sample Storage | Store the sample plates in a storage at 4°C. |
| UH010 | Liquid Handling | Combine template DNA, error correction mix containing DW, endonuclease, and exonuclease into a 96-well PCR plate. |
| UH180 | Incubation | Incubate at 42°C for 1 hour. |
| UH010 | Liquid Handling | Dispense 1X Q5 or Phusion polymerase and PCR master mix to a destination plate. Prepare primers with appropriate concentration on a source plate. |
| UH030 | Nanoliter Liquid Dispenser | Dispense primers from the source plate to the destination plate. |
| UH100 | Thermocycling | Amplifying template. 98°C for 2 min, 30 cycles of 98°C for 30s and 65°C for 6 min (increasing 15 sec/cycle), followed by 72°C incubation for 5 min. |
| UH230 | Nucleic Acid Fragment Analysis | Analyze the PCR product along with a proper DNA ladder. |

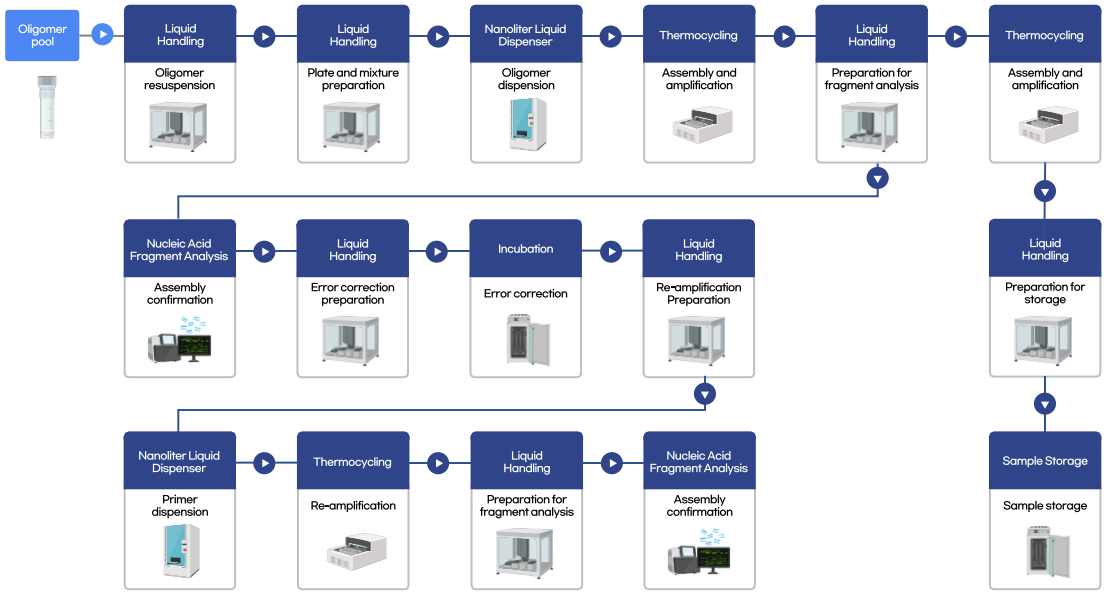


Figure S2. A set of unit operations for the DNA Oligomer Assembly workflow (WB010), illustrating the stepwise experimental process from oligomer resuspension to final sample storage. (Device icons from BioRender.)

### Tier 3 Example: Part DNA Assembly (Golden Gate Assembly) Workflows

We illustrate a tier 3 case-study for the application of abstraction hierarchy. This example is established based on the protocol described by Smith et al., Kanigowska, *et al.*, and Wang, Y. *et al.*[3–5]. It encompasses DNA Assembly (WB030), Biology-mediated DNA Transfers (WB120), Solid Media Cell Culture (WB130), Liquid Media Cell Culture (WB140), Nucleic Acid Size Verification (WB015, optional), DNA extraction (WB045), DNA multiplexing (WB060), and Nucleotide Sequencing (WT010). These workflows are utilizing DNA part stocks consisting of DNA part fragments, overhangs, and Type IIs restriction enzyme recognition sites (e.g. BsaI) to construct transcription units (TUs). Subsequently, plasmids were isolated by transforming and culturing the constructed TUs, and their sequences were analyzed to verify the assembly. Note that Plate Handling (UH120) is not represented for the simplicity. Relationships among the individual unit operations within each workflow were illustrated in Figure S3 (WB015 omitted).

Table S5. Details of workflows for Part DNA Assembly service/capability

| Workflow ID | Workflow | Description | Unit op. ID | Unit operation | Description |
|---|---|---|---|---|---|
| WB030 | DNA assembly | Preparation and conducting | UH160 | Sample Storage | Retrieve the DNA part stocks (including vector plasmid) |

| | | | | | |
|---|---|---|---|---|---|
| | | golden gate assembly | UH150 | Capping Decapping | DNA part stock decapping |
| | | | UH170 | Plate Storage | Retrieve source/ destination plate |
| | | | UH010 | Liquid Handling | Dispensing aliquots of reagent stocks and reaction mixtures into the source plate |
| | | | UH255 | Centrifugation | Spin down in the source plate |
| | | | UH030 | Nanoliter Liquid Dispenser | Golden Gate assembly mixture dispensing to the destination plate |
| | | | UH255 | Centrifugation | Mixture spin down in the destination plate |
| | | | UH130 | Sealing | Sealing the destination plate |
| | | | UH100 | Thermocycling | Golden Gate assembly reaction |
| | | | UH255 | Centrifugation | Reaction mixture spin down |
| | | | UH140 | Peeling | Plate peeling |
| WB120 | Biology-Mediated DNA Transfers | Heat shock transformation | UH160 | Sample Storage | Retrieve competent cell |
| | | | UH170 | Plate Storage | Retrieve well plates |
| | | | UH010 | Liquid Handling | Transformation reagent preparation |
| | | | UH010 | Liquid Handling | Heat-shock transformation |
| WB130 | Solid Media Cell Culture | Cell culture in solid media | UH180 | Incubation | Solid plate incubation |
| WB140 | Liquid Media Cell Culture | Transformant liquid culture | UH170 | Plate Storage | Retrieve a deep well plate for liquid incubation |
| | | | UH015 | Bulk Reagent Dispensing | Dispensing culture media into the deep well plate |
| | | | UH060 | Colony Picking | Inoculating colonies into liquid culture media in the deep well plate |
| | | | UH130 | Sealing | Sealing the deep well plate |
| | | | UH180 | Incubation | Liquid culture incubation |
| WB015 | Nucleic Acid Size Verification (Optional) | Part assembly confirmation through size confirmation | UH170 | Plate Storage | Retrieve a 96-well PCR plate |
| | | | UH140 | Peeling | Liquid culture plate peeling |
| | | | UH010 | Liquid Handling | Take a small aliquot of liquid culture from a colony-pooled liquid culture plate and inoculate it onto a PCR plate |
| | | | UH010 | Liquid Handling | Colony PCR mixture preparation and dispensing |
| | | | UH130 | Sealing | PCR plate sealing |
| | | | UH255 | Centrifugation | Mixture spin down in the PCR plate |
| | | | UH100 | Thermocycling | Colony PCR |
| | | | UH410 | Centrifugation | Plate spin down |

| | | | | | |
|---|---|---|---|---|---|
| | | | UH140 | Peeling | PCR plate peeling |
| | | | UH230 | Nucleic Acid Fragment Analysis | DNA size confirmation |
| WB045 | DNA Extraction | Plasmid mini prep and quality control | UH410 | Centrifugation | Cell down the cultured plate |
| | | | UH140 | Peeling | Plate peeling |
| | | | UH250 | Nucleic Acid Purification | Plasmid mini prep |
| | | | UH170 | Plate Storage | Retrieve a 96-well plate for measuring plasmid concentration |
| | | | UH010 | Liquid Handling | Transfer extracted DNA into empty well plate |
| | | | UH380 | Microplate Reading | Measuring DNA concentration |
| WB060 | DNA Multiplexing | Sample treatment & multiplexing | UH160 | Sample Storage | Retrieve DNA plasmids |
| | | | UH170 | Plate storage | Retrieve a 96-well plate |
| | | | UH020 | 96 Channel Liquid Handling | Adapter, size selection, barcoding |
| | | | UH280 | Sequence Quality Control | Measuring DNA concentration |
| | | | UH230 | Nucleic Acid Fragment Analysis | DNA size confirmation |
| WT010 | Nucleotide Sequencing | Sequencing | UH270 | Long-read Sequence Analysis | Long read sequencing |

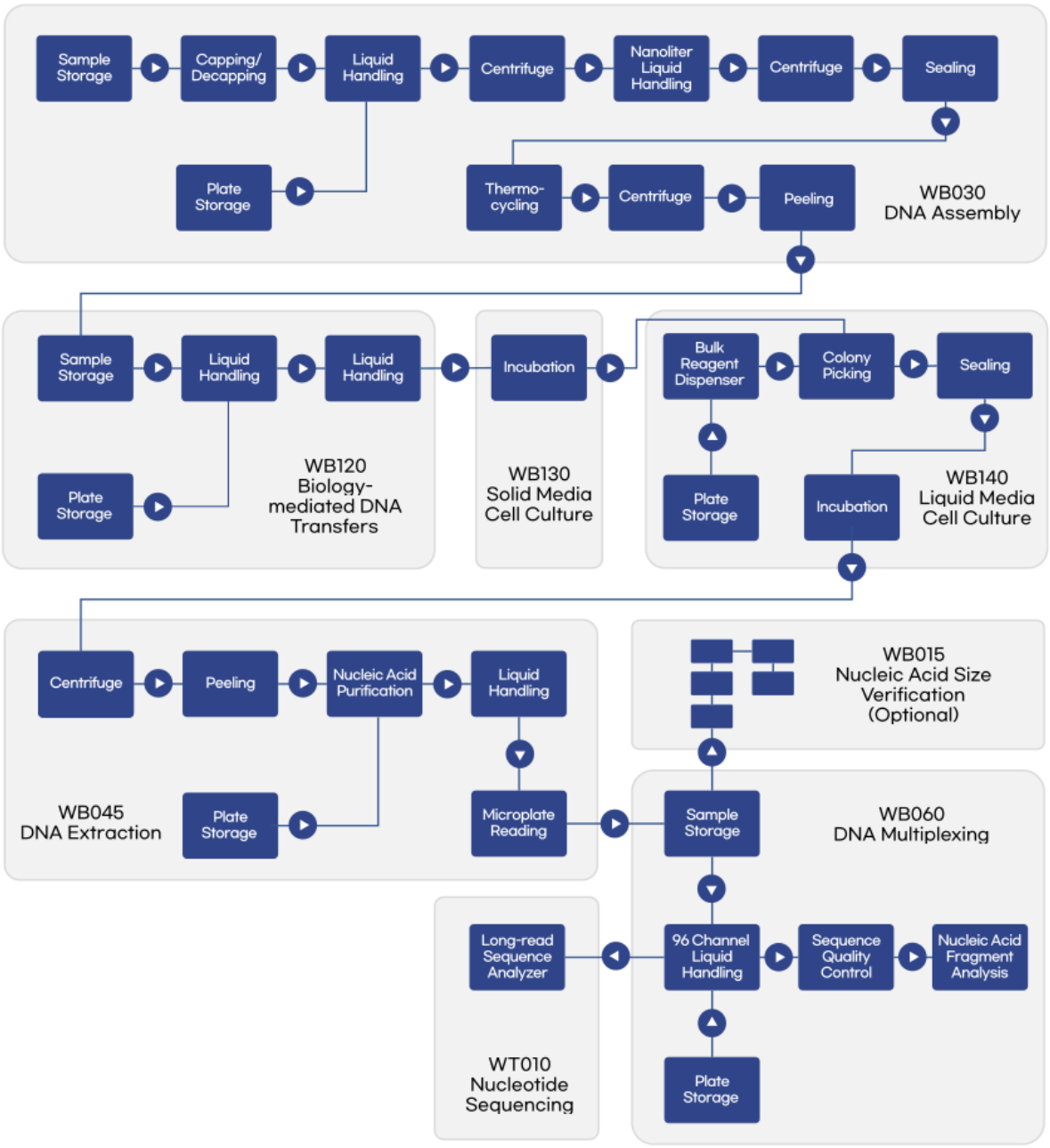


Figure S3. A set of unit-operations comprising the workflows from DNA Assembly to Long Read Plasmid DNA Sequencing (Workflow IDs: WB030, WB120, WB140, WB040, WB060, and WT010). It illustrates how sequential unit operations are orchestrated to automate complex synthetic biology tasks. Each workflow segment can be flexibly reconfigured, highlighting the practical application of the proposed abstraction hierarchy in constructing scalable and interoperable biofoundry workflows.

The three workflows and their unit operations constituting the Tier 3 capability of Part DNA Assembly can be electronically documented in markdown format as shown in Table S6. The first three workflows, DNA Assembly (WB030), Biology-mediated DNA Transfers (WB120), and Solid Media Cell Culture (WB130) are represented using Heading 2 ("##") level markdown headers, and the individual unit operations within each workflow are described using Heading 3 ("###") level headers. Each unit operation is systematically described with the necessary records. The unit operation name is enclosed in brackets, followed by a specific description of the experiment in the title. Metadata related to the specific operation including operator information, date, and time is recorded. Detailed methodologies outlining the reagents, consumables, and equipment required for each operation are also included. Input and Output records are specified to allow tracking of samples throughout the workflows, supporting traceability and reproducibility. In addition, Python code snippets can be included to dynamically display sample names and IDs based on mapping files. When using interactive editors like Jupyter Notebook, the executed code will display actual sample information, whereas standard text editors will show the code itself. This structure not only organizes the experimental records systematically but also enhances workflow modularity and reusability, aligning with the proposed abstraction hierarchy framework for biofoundry operations.

Table S6. Details of workflows and unit operations in markdown format. Workflows are represented using Heading 2 ("##") level markdown headers, and the individual unit operations within each workflow are described using Heading 3 ("###") level headers.

```
## DNA Assembly
### [Sample Storage] DNA Part Stock Preparation

#### Meta
-   Operator: User
-   Date: DD.MM.YYYY
-   Time: 00:00

#### Reagents

-   DNA part stock (stock list: part_DB.xlsx)

#### Consumables
-  None

#### Equipment
-   Sample storage system

#### Input
- Assembly design file (Capital letter means part's own sequence)
      - P0010, Promoter,      BBa_K088007,
        ggtctcagcctTATAAGATCATACGCCGTTATACGTTGTTTACGCTTTGaaagagagacc
      - P0017, Promoter,      BBa_R1075,
        ggtctcagcctTTAAATTTCCTCTTTTCAGGCCGGAATAACTCCCTATAATGCGCCACCAaaagagagacc
      - P0008, Promoter,      BBa_J23116,
        ggtctcagcctTTGACAGCTAGCTCAGTCCTAGGGACTATGCTAGCaaagagagacc
      - P0007, Promoter,      BBa_J23115,
        ggtctcagcctTTTATAGCTAGCTCAGCCCTTGGTACAATGCTAGCaaagagagacc
```

```
      - R0023, RBS,    BBa_B0034,     ggtctcactttTCTAGAGAAAGAGGAGAAATActgcagagacc
      - R0021, RBS,    BBa_B0030,     ggtctcactttTCTAGAGATTAAAGAGGAGAAATActgcagagacc
      - R0022, RBS,    BBa_B0032,     ggtctcactttTCTAGAGTCACACAGGAAAGTActgcagagacc
      - C0002, CDS,    sfGFP, ggtctcagcag[HQ873313.1]ttagagagacc
      - T0031, Terminator,    L3S2P22,
        ggtctcactaaCTCGGTACCAAATTCCAGAAAAGAGGCCGCGAAAGCGGCCTTTTTTCGTTTTGGTCCGTGAagagacc
      - T0032, Terminator,    L3S1P51,
        ggtctcactaaAAAAAAAAAAAAGGCCTCCCAAATCGGGGGGCCTTTTTTATTGATAACAAAAGTGAagagacc
      - T0030, Terminator,    L3S1P52,
        ggtctcactaaTCTAACTAAAAAGGCCTCCCAAATCGGGGGGCCTTTTTTATTGATAACAAAAGTGAagagacc
      - T0033, Terminator,    L3S3P56,
        ggtctcactaaGCATTAGGTCTCACTAATTTTCGAAAAAACACCCTAACGGGTGTTTTTTTGTTTCTGGTCTCCCTCACAGA
GACCAAAGCAGTGAagagacc
      - V0001, Vector, pACBB-carrier, ggtctcatcac[CmR, p15A]aggcagagacc

#### Method

-   Check DNA samples for using Golden Gate assembly
-   Bring Capped DNA samples for Golden Gate assembly from storage system

#### Output
-   Capped DNA tubes from Sample storage system


### [Capping Decapping] DNA Part Stock Decapping
#### Meta
-   Operator: User
-   Date: DD.MM.YYYY
-   Time: 00:00

#### Reagents
-   None

#### Consumables
-   None

#### Equipment
-   Tube capper and decapper

#### Input
-   Capped DNA tubes from Sample storage system

#### Method
-   Decapping DNA stocks using Tube capper and decapper

#### Output
-   Decapped DNA tubes


### [Plate storage] Golden Gate Assembly Mixture Preparation
#### Meta
-   Operator: User
-   Date: DD.MM.YYYY
-   Time: 00:00

#### Reagents
-   None

#### Consumables
-   384 PP plate for source plate
```

```
-   96 well plate for destination plate

#### Equipment
-   Plate hotel

#### Input
-   Experiment design documents

#### Method
-   Retrieve plates for experiment as designed

#### Output
- WPL0001, a 96 well plate for Golden Gate assembly (destination plate)
- WPL0002, a 384 well plate for Golden Gate assembly (source plate)


### [Liquid handling] Dispensing DNA parts in Source Plate for Golden Gate Reaction
### Meta
-   Operator: User
-   Date: DD.MM.YYYY
-   Time: 00:00

#### Reagent
-   None

#### Consumables
-   None

#### Equipment
-  Automated liquid handler

#### Input
- Decapped DNA tubes
      - P0010, Promoter,      BBa_K088007,
        ggtctcagcctTATAAGATCATACGCCGTTATACGTTGTTTACGCTTTGaaagagagacc
      - P0017, Promoter,      BBa_R1075,
        ggtctcagcctTTAAATTTCCTCTTTTCAGGCCGGAATAACTCCCTATAATGCGCCACCAaaagagagacc
      - P0008, Promoter,      BBa_J23116,
        ggtctcagcctTTGACAGCTAGCTCAGTCCTAGGGACTATGCTAGCaaagagagacc
      - P0007, Promoter,      BBa_J23115,
        ggtctcagcctTTTATAGCTAGCTCAGCCCTTGGTACAATGCTAGCaaagagagacc
      - R0023, RBS,    BBa_B0034,     ggtctcactttTCTAGAGAAAGAGGAGAAATActgcagagacc
      - R0021, RBS,    BBa_B0030,     ggtctcactttTCTAGAGATTAAAGAGGAGAAATActgcagagacc
      - R0022, RBS,    BBa_B0032,     ggtctcactttTCTAGAGTCACACAGGAAAGTActgcagagacc
      - C0002, CDS,    sfGFP, ggtctcagcag[HQ873313.1]ttagagagacc
      - T0031, Terminator,    L3S2P22,
        ggtctcactaaCTCGGTACCAAATTCCAGAAAAGAGGCCGCGAAAGCGGCCTTTTTTCGTTTTGGTCCGTGAagagacc
      - T0032, Terminator,    L3S1P51,
        ggtctcactaaAAAAAAAAAAAAAGGCCTCCCAAATCGGGGGGCCTTTTTTATTGATAACAAAAGTGAagagacc
      - T0030, Terminator,    L3S1P52,
        ggtctcactaaTCTAACTAAAAAGGCCTCCCAAATCGGGGGGCCTTTTTTATTGATAACAAAAGTGAagagacc
      - T0033, Terminator,    L3S3P56,
        ggtctcactaaGCATTAGGTCTCACTAATTTTCGAAAAAACACCCTAACGGGTGTTTTTTTGTTTCTGGTCTCCCTCACAGA
GACCAAAGCAGTGAagagacc
      - V0001, Vector, pACBB-carrier, ggtctcatcac[CmR, p15A]aggcagagacc
-   WPL0002, a 384 PP plate from plate hotel

#### Method
-   DNA part preparation
    -   Thaw the decapped DNA parts at room temperature.
```

```
        - All of DNA parts should be duplex primer or PCR-product containing Type IIs
          restriction enzyme recognition site and overhang sites
    -   repeated pipetting the thawed DNA parts, then to collect the liquids.
    -   Place the tube containing the DNA part into the liquid handler
    -    Dispense 40 µL of DNA parts into the 384-well PP plate (WPL0002) using liquid
         handler.
        -   Refer to the JANUS G3 mapping file for precise dispensing
    -   Source_plate_information: \Source_for_Golden Gate-PCR.csv
    -   add at least 30 µL for refill

#### Output:

-   WPL0002, Source plate with DNA parts


### [Liquid handling] Golden Gate Reaction Enzyme Mixture Dispensing in Source Plate

### Meta

-   Operator: User
-   Date: DD.MM.YYYY
-   Time: 00:00

#### Reagent
-   Distilled Water
-   T4 DNA ligase
-   Type IIS restriction enzyme
-   10x T4 DNA ligase buffer

#### Consumables
-   Pipet tips
-   2 ml tube

#### Equipment
-  Automated liquid handler
-  Manual Pipettes
-  Freezer

#### Input
-   WPL0002, Source plate with DNA parts

#### Method
-   Enzyme mixture preparation

    -   Keep the Type IIs restriction enzyme and T4 DNA ligase on ice.
    -   Thaw 10x T4 DNA ligase buffer at room temperature.
    -   Vortex the thawed buffer, then perform a light spin-down to collect the liquids.
    -   Prepare the Golden Gate reaction mixture by combining the following components in
        a 2mL tube. The volume is depending on the number of samples.
        -   0.04 µL T4 DNA ligase
        -   0.04 µL BsaI-HF (restriction enzyme)
        -   0.1 µL 10x T4 DNA ligase buffer
        -    Fill up Distilled Water (DW) to 1 µL after subtracting the volume of the DNA
           parts (adjust based on total DNA part volume).
    -   Mix thoroughly by pipetting or vortexing, then perform a spin-down.
    -   Keep the prepared mixture on ice.
    -   add in 384-well plate (WPL0002)
    -   Source_plate_information: \Source_for_Golden Gate-PCR.csv

#### Output
-   WPL0002, Source plate for Golden Gate assembly
```

```
### [Centrifugation] Removing Bubbles from The Source Plate
#### Meta
-   Operator: User
-   Date: DD.MM.YYYY
-   Time: 00:00

#### Reagent
-   None

#### Consumables
-   None

#### Equipment
-   Micro plate centrifuge

#### Input
-   WPL0002, a Source plate for Golden Gate assembly

#### Method
-   Slightly spin down the source plate for removing bubbles

#### Output
-   WPL0002, a Bubbles removed source plate for Golden Gate assembly


### [Nanoliter liquid handling] Transfer DNA Parts Using Acoustic Liquid Handler
### Meta
-   Operator: User
-   Date: DD.MM.YYYY
-   Time: 00:00

#### Reagents
-   None

#### Consumables
-   None

#### Equipment
-   Acoustic liquid handler

#### Input
-   WPL0001, a 96 well plate for Golden Gate assembly (destination plate)
-   WPL0002, a Bubbles removed source plate for Golden Gate assembly

#### Method
-   Use acoustic liquid handler to transfer the DNA parts from the 384-well PP plate (source
    plate) to the 96-well skirted plate (destination plate)
-   Follow the mapping file for precise transfer and dispense 50 fmol of each DNA part
    based on the DNA concentration and size
-   Mapping file: mapping_file.csv

#### Output
-   WPL0001, a Part mixture prepared for Golden Gate assembly



### [Centrifugation] Removing Bubbles from Destination Plate
#### Meta
```

```
-   Operator: User
-   Date: DD.MM.YYYY
-   Time: 00:00

#### Reagent
-   None

#### Consumables
-   None

#### Equipment
-   Micro plate centrifuge

#### Input
-   WPL0001, Part mixture prepared for Golden Gate assembly

#### Method
-   Slightly spin down the destination plate for removing bubbles

#### Output
-   WPL0001, a Bubble-free destination plate prepared for Golden Gate assembly


### [Sealing] Golden Gate Assembly Mixture-Contained Plate Sealing
#### Meta
-   Operator: User
-   Date: DD.MM.YYYY
-   Time: 00:00

#### Reagent
-   None

#### Consumables
-   Heat-resistant sealing film

#### Equipment
-   Plate sealer
-   Freezer (Optional)

#### Input
-   WPL0001, a Bubble-free destination plate prepared for Golden Gate assembly

#### Method
-   Seal the 96-well skirted plate using heat-resistant film with a plate sealer.
-   Seal the source plate and store it in the refrigerator. (optional)

#### Output
-   WPL0001, a Sealed 96-well skirted plate ready for DNA assembly (not assembled)



### [Thermocycling] Golden Gate Assembly Reaction
#### Meta
- Operator: User
- Date: DD.MM.YYYY
- Time: 00:00

#### Reagents
-   None
```

#### Consumable
- None

#### Equipment
- Thermocycler

#### Input
- WPL0001, a Sealed 96-well skirted plate ready for DNA assembly (not assembled)

#### Method
- Turn on the thermocycler
- Put the plate on the thermocycler
- Thermocycling as a following method

| **Step** | **Temperature** | **Time** | **Description** |
|:---:|:---:|:---:|:---|
| 1 | 37°C | 10 min | Initial cutting DNA with restriction enzymes |
| 2 | 37°C | 5 min | Restriction enzyme active step |
| 3 | 16°C | 5 min | Ligation, repeat step 2 for 5 cycles |
| 4 | 75°C | 5 min | Inactivation of ligase |
| 5 | 80°C | 10 min | Inactivation of restriction enzyme |
| 6 | 4°C | - | Hold at 4°C for storage |

#### Output
- WPL0001, a Sealed 96-well skirted plate contains Golden Gate-assembled DNA

### [Centrifugation] Removing Bubbles from Golden Gate assembly mixture

#### Meta
- Operator: User
- Date: DD.MM.YYYY
- Time: 00:00

#### Reagent
- None

#### Consumables
- None

#### Equipment
- Micro plate centrifuge

#### Input
- WPL0001, a Sealed 96-well skirted plate contains Golden Gate-assembled DNA

#### Method
- Slightly spin down the destination plate for removing bubbles

#### Output
- WPL0001, a Bubbles removed 96-well skirt plate contains Golden Gate-assembled DNA

### [Peeling] Plate peeling for Transformation
#### Meta
- Operator: User
- Date: DD.MM.YYYY
- Time: 00:00

```
#### Reagent
-   None

#### Consumable
-   None

#### Equipment
-   Peeler

#### Input
-   WPL0001, a Bubbles removed 96-well skirt plate contains Golden Gate-assembled DNA

#### Method
-   Peeling selected plate

#### Output
-   WPL0001, a Film-peeled 96-well skirt contains Golden Gate-assembled DNA

## Biology-mediated DNA Transfers
### [Sample Storage] Retrieve Competent Cell Plate

#### Meta
-   Operator: User
-   Date: DD/MM/YYYY
-   Time: 00:00

#### Reagents
-   None

#### Consumable
-   None

#### Equipment
-   Sample storage system
-   Ice in Rubber basket

#### input
-   Competent cell (30 µL per well) contained 96-well skirt plate

#### Method
-   Put competent cell plate from sample storage system
-   Thawing the competent cell on Ice (96-well plate)

#### Output
-   WPL0003, a Competent cell (30 µL per well) contained 96-well skirt plate


### [Liquid handling] Transformation Reagent Preparation
#### Meta
-   Operator: User
-   Date: DD/MM/YYYY
-   Time: 00:00

#### Reagents:
-   SOC media

#### Consumables:
-   Single-well plate (WPL0004)
-   Sterile pipette tips compatible with liquid handling system
```

#### Equipment:
- Liquid handler

#### Input:
- Experiment design documents

#### Method:
- Bring plates out for experiment as designed
- Dispense >10 mL SOC medium into the single-well reservoir plate (WPL0004)

#### Output:
- WPL0004, 1-well plate contained SOC media

### [Liquid handling] Heat-shock Transformation

#### Meta
- Operator: User
- Date: DD/MM/YYYY
- Time: 00:00

#### Reagents
- None

#### Consumable
- WPL0005, LB agar plate supplemented with chloramphenicol
- WPL0006, LB agar plate supplemented with chloramphenicol for 1/10 dilution spotting
- Sterile pipette tips compatible with liquid handling system

#### Equipment
- Liquid handler

#### Input
- WPL0001, a Film-peeled 96-well skirt contains Golden Gate-assembled DNA
- WPL0003, a Competent cell (30 µL per samples) contained 96-well skirt plate
- WPL0004, 1-well plate contained SOC media

#### Method
- Position the DNA input plate (WPL0001), competent cell plate (WPL0003), SOC medium reservoir (WPL0004), and sterile LB agar plates (WPL0005, WPL0006) on the automated liquid handling deck
- Execute the automated transformation protocol
    - Transfer Competent cell to DNA product
    - Cold incubation for 30 min
    - 42℃ heat-shock for 90 sec
    - Incubate immediately on ice for an additional 2 min
    - Add SOC medium to each transformed cell mixture
    - Allow cell recovery at 37℃ with agitation for 45 min
    - Spot 9 µL of the transformation mixture onto LB agar plates (WPL0005)
    - Perform a 1/10 dilution of the transformation mixture using SOC medium and spot 9 µL onto separate LB agar plates (WPL0006)

#### Output
- WPL0005, a Transformants spotted LB media plate
- WPL0006, a 1/10 diluted Transformants spotted LB media plate

## Solid Media Cell Culture
### [Incubation] Plate Incubation

```
#### Meta
-   Operator: User
-   Date: DD/MM/YYYY
-   Time: 00:00

#### Reagents
-   None

#### Consumable
-   None

#### Equipment
-   Incubator

#### input
-   WPL0005, a Transformants spotted LB media plate
-   WPL0006, a 1/10 diluted Transformants spotted LB media plate

#### Method
-   Place the plate in an incubator set to 37°C.
-   Incubate for overnight (~16 h)

#### Output
-   WPL0005, a Transformants spotted LB media plate with colonies
-   WPL0006, a 1/10 diluted Transformants spotted LB media plate with colonies

#### Result
-   image of colonies

![WPL0005](image-3.png)
![WPL0006](image-4.png)
```

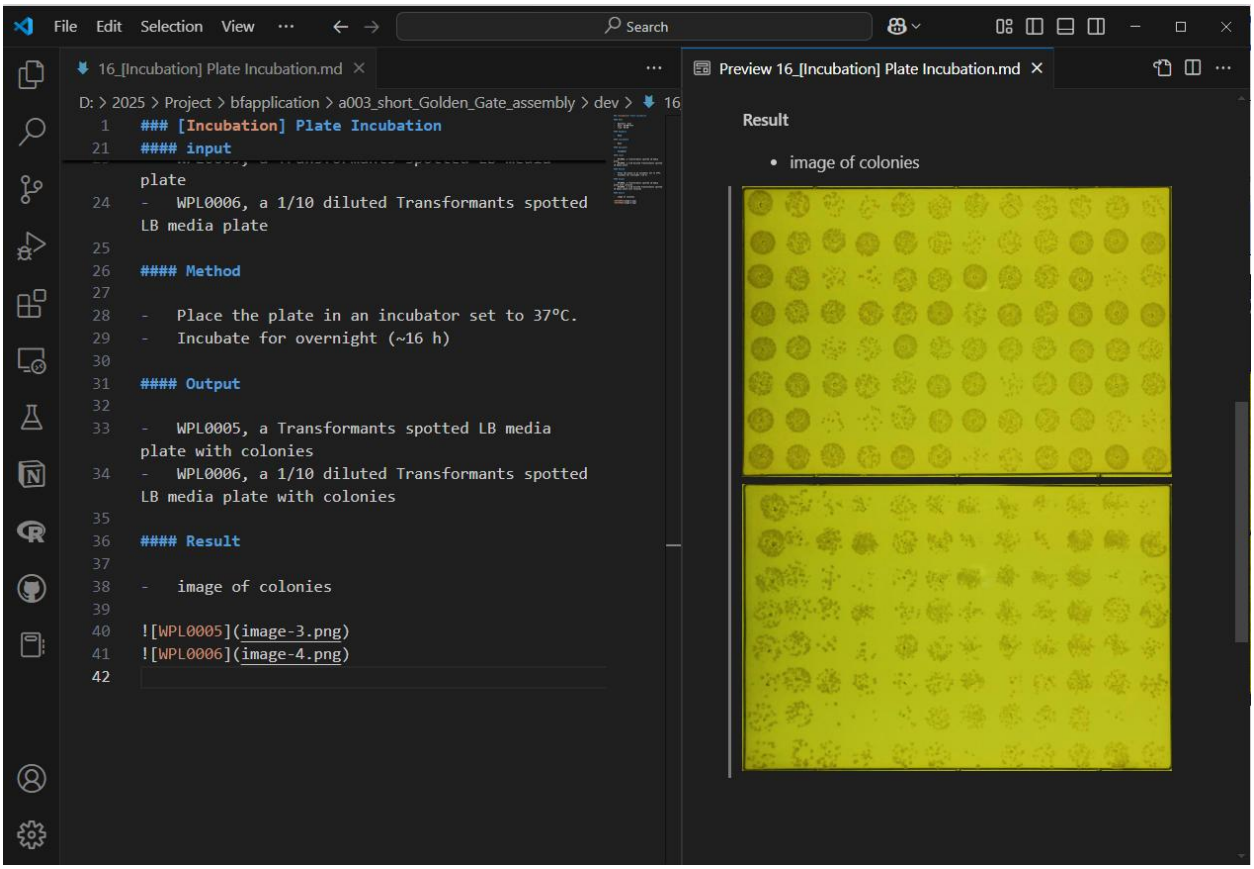


Fig S4. Capture image of "[Incubation] Plate incubation" markdown file and its preview by Visual Studio code.

## Supplementary References


1. Dauparas, J. *et al.* Robust deep learning–based protein sequence design using ProteinMPNN. *Science* **378**, 49–56 (2022).
2. Hutchison, C. A. *et al.* Design and synthesis of a minimal bacterial genome. *Science* **351**, aad6253 (2016).
3. Kanigowska, P., Shen, Y., Zheng, Y., Rosser, S. & Cai, Y. Smart DNA Fabrication Using Sound Waves: Applying Acoustic Dispensing Technologies to Synthetic Biology. *SLAS Technology* **21**, 49–56 (2016).

4. Wang, Y., Zhao, Y., Bollas, A., Wang, Y. & Au, K. F. Nanopore sequencing technology, bioinformatics and applications. *Nature biotechnology* **39**, 1348–1365 (2021).
5. Pryor, J. M., Potapov, V., Bilotti, K., Pokhrel, N. & Lohman, G. J. S. Rapid 40 kb Genome Construction from 52 Parts through Data-optimized Assembly Design. *ACS Synth. Biol.* **11**, 2036–2042 (2022).