scientific reports

Check for updates

OPEN

Development of the autonomous lab system to support biotechnology research

Keiji Fushimi[1,5], Yusuke Nakai[2,5], Akiko Nishi[1], Ryo Suzuki[1], Masahiro Ikegami[2], Risa Nimura[2], Taichi Tomono[2], Ryota Hidese[3], Hisashi Yasueda[1,3,4], Yusuke Tagawa[2]✉ & Tomohisa Hasunuma[1,3]✉

In this study, we developed the autonomous lab (ANL), which is a system based on robotics and artificial intelligence (AI) to conduct biotechnology experiments and formulate scientific hypotheses. This system was designed with modular devices and Bayesian optimization algorithms, allowing it to effectively run a closed loop from culturing to preprocessing, measurement, analysis, and hypothesis formulation. As a case study, we used the ANL to optimize medium conditions for a recombinant *Escherichia coli* strain, which overproduces glutamic acid. The results demonstrated that our autonomous system successfully replicated the experimental techniques, such as sample preparation and data measurement, and improved both the cell growth rate and the maximum cell growth. The ANL offers a versatile and scalable solution for various applications in the field of bioproduction, with the potential to improve efficiency and reliability of experimental processes in the future.

Keywords Flexible robotic system, Autonomous experiment, Bioproduction, Medium optimization, Cell growth, Bayesian optimization

Science and technology have been progressed through human creation along with accumulation of vast amounts of experimental data. Recently, automation of experimental process plays a crucial role for improving to obtain the reliable data rapidly, which widely contribute to developments of various scientific and technological fields[1]. Furthermore, with the evolution of automation technology and robots, and the advent of artificial intelligence (AI) technologies, autonomous systems can infer and discover scientific hypotheses have emerged[2–4]. The autonomous systems, which plan and execute experiments without human intervention, enable to obtain the reproducible data along with faster, more reliable, and less cost, and then allow human researchers to engage in more creative research.

Organic molecules, both natural and non-natural products, have been used as sources for functional foods, pharmaceutical, and chemical products[5–7]. Many molecules with complicated structures and/or unique chemical properties are supplied by organic synthesis. Although the chemical methods enable to product mass on industrial scales, they are involved in disposal of hazardous materials, which is often a problem of environmental pollution. In this context, as a more environmentally friendly method, supplies of these useful molecules have been tried based on metabolic functions of microorganisms, such as *Escherichia coli*[8–12] and *Saccharomyces cerevisiae*[13–17], in biotechnology field.

In the field of biotechnology, the acquisition and analysis of vast amounts of experimental data is important to design metabolic pathways for target molecules and to improve their productivity. However, the analysis is very labor-intensive because the experiment requires the measurements of the target (substrate, products, gene and protein expression, enzyme activities, and cell growth) under many conditions (concentration, temperatures, pH, and reaction and incubation times), and the analysis requires the processing of a large amount of data including various metadata such as chemical structures, properties of reagents, and experimental protocols. Autonomous systems that can perform both data acquisition and analysis have therefore attracted a significant attention. In fact, recent case studies have demonstrated that they are useful not only in the development of chemical reactions[18–20], but also in the development of bioproduction[21,22]. However, since the conditions and

[1]Graduate School of Science, Innovation and Technology, Kobe University, 1-1 Rokkodai, Nada, Kobe 657-8501, Japan. [2]Technology Research Laboratory, Shimadzu Corporation, 3-9-4 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237, Japan. [3]Engineering Biology Research Center, Kobe University, 1-1 Rokkodai, Nada, Kobe 657- 8501, Japan. [4]Research and Development Center for Precision Medicine, University of Tsukuba, 1-2 Kasuga, Tsukuba, Ibaraki 305-8550, Japan. [5]Keiji Fushimi and Yusuke Nakai contributed equally to this work. ✉email: y-tagawa@shimadzu.co.jp; hasunuma@port.kobe-u.ac.jp

targets of interest vary depending on the experiment being performed, a high degree of flexibility is required for autonomous system to be widely used in the biotechnological field. However, many autonomous systems are focused on specific experimental applications and are not scalable or versatile. Laboratory automation systems, such as the KIWI-biolab[23–25], aim to simplify plug-and-play functionality[26,27]. However, the cost and effort required for installation, expansion, and hardware reconfiguration remain significant. Additionally, the hurdle for implementing plug-ins remains high.

In this study, we developed the Autonomous Lab (ANL) as a modular autonomous experimental system with scalability and versatility. The ANL modules can be easily added, removed, or repositioned, so that the device can be configured each time to suit the experiment to be conducted. As a case study of the ANL operation, we optimized medium conditions for a glutamic acid-producing strain constructed based on *E. coli*. The ANL could find an optimized medium condition to improve the cell growth based on the relationship between the nutrient concentrations and cell densities. On the other hand, the ANL could only propose an optimized medium condition to slightly increase the produced glutamic acid concentration. It is thought that the cells intricately regulated the concentration to protect themselves from stresses of osmotic pressure and pH, and this suggests that adding intracellular osmotic pressure or pH as explanatory variables can improve glutamic acid production. The ANL, which can flexibly combine various experimental devices to obtain and analyze multiple data, would be the basis for the development of bioproduction technology.

Results

Building the autonomous lab

We developed the autonomous lab (ANL) as a modular autonomous experimental system (Fig. 1 and S1, and Sup. movies 1 and 2). The ANL is equipped with devices for culturing, preprocessing, and analysis, as well as transport robots and analysis algorithms using Bayesian optimization; thus, it can autonomously run a closed loop from culturing through preprocessing and measurement to analysis.

One important feature of ANL is high flexibility through the modularization of hardware. All devices are installed on carts with stoppers, and function as modules that can be moved, independently. In addition, these modules can be positioned anywhere within the reach of the transfer robot's arm. In other words, the ANL is not only easy to install but also simple to modify or expand its configuration as needed. For example, multiple incubators can be installed for parallel cultures, or the position of the modules can be adjusted to fit the room's layout. Furthermore, owing to the user interface (UI), which visualizes protocols and the integrated control system, scientists designing experimental systems can easily program autonomous experiments without technical assistance (Fig. S2).

In this study, we present the application of autonomous experimentation with the ANL using the optimization of medium conditions for a glutamic acid-producing strain, which harbors enhanced metabolic pathway for glutamic acid synthesis based on *Escherichia coli*, as a case study (Fig. S3). Glutamic acid is a useful molecule which has the potential for applications in food, agriculture, wastewater treatment, medicine, and cosmetics[28–30]. Understanding how to increase the glutamic acid production, which is intricately regulated by various factors related to the cell growth, could provide valuable insights into the field of bioproduction (details of the biotechnological information were shown in the Supplementary Material, "Construction of the glutamic acid-producing cells for the demonstration experiment of the ANL"). Algorithms using Bayesian optimization have been proposed for optimizing medium conditions[31,32]; however, to perform the experiment, an integrated system that performs a closed-loop of culture, pretreatment, measurement, and analysis is required. This is the first report that an actual case of optimizing medium conditions has been incorporated into an autonomous system.

As shown in Fig. 1, our system consists of a transfer robot (PF400, Brooks), plate hotels, microplate reader (SpectraMax iD3, Molecular Devices), centrifuge (HiG, BioNex), incubator (STX44-HR, LiCONiC), liquid handler (OT-2, Opentrons LabWorks), and LC-MS/MS system (Nexera XR, LCMS-8060NX, Shimadzu). The LC system comprises a sample container autoloader (ATL-40), system controller (SCL-40), high-pressure pump (LC-40BXR), degasser (DGU-403), autosampler (SIL-40CXR), and column oven (CTO-40 S).

Selection of the components to optimize the medium condition for the glutamic acid-producing cells

Based on the results of the initial investigation (Figs. S4 and S5), we used a M9 medium[33] as a base one to verify the cell growth and productivity of the glutamic acid-producing strain. Because it is a minimal medium containing only metal ions and essential nutrients, it is easy to exclusively quantify the glutamic acid produced by the cells without accounting for glutamic acid derived from medium components like yeast extract. In the preliminary experiments, we have confirmed that the glutamic acid producibility of the cells was lower in the M9 medium with yeast extract than in the M9 medium without it (Fig. S4) (details were shown in the Supplementary Material, "Initial investigation of media conditions for the glutamic acid-producing cells"). Additional components into the M9 medium to optimize the cell growth and the product amount were selected; basic components (Na_2HPO_4, KH_2PO_4, NH_4Cl, NaCl, $CaCl_2$, $MgSO_4$, glucose, and thiamine) constituted the M9 medium, and trace elements (H_3BO_3, $(NH_4)_6Mo_7O_{24}$, $MnCl_2$, $CoCl_2$, $FeSO_4$, $CuSO_4$, $ZnSO_4$, and flavin adenine dinucleotide (FAD)) constituted the MOPS (3-(*N*-morpholino)propanesulfonic acid) medium[34] or were cofactors and activators related to the glutamic acid biosynthesis (Fig. S3). In short, these components not only affect the cell growth, but also act as the cofactors for diverse enzymes related to the glutamic acid production. Therefore, optimizing the diversity of these components was necessary to regulate the multi-step enzymatic reactions in the long pathway. The ANL was ideal for performing the optimization, which is considered difficult for humans to carry out empirically. The capabilities of the cells cultivated in the adjusted media containing

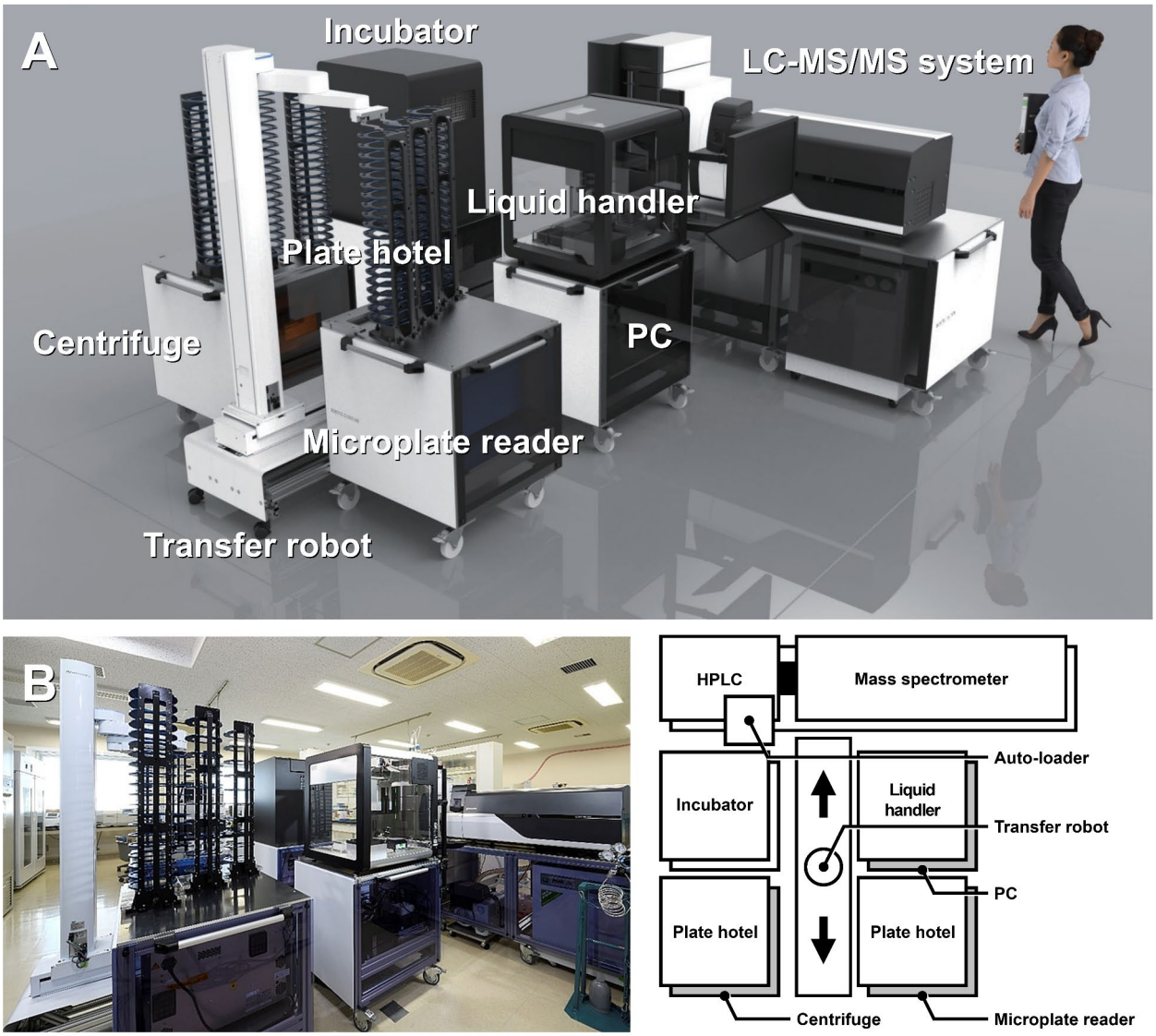

Fig. 1. Concept of the Autonomous Lab (ANL). (**A**) Computer-generated image and (**B**) photograph with the ANL layout drawing. The ANL is composed of a transfer robot, plate hotels, microplate reader, centrifuge, incubator, liquid handler, and LC-MS/MS system. The transfer robot can move the side of the instruments and transfer microplates there. Samples in the plates are treated (e.g., incubation for cell cultivation and enzymatic reactions, centrifugation, removing culture media, and adding solvents) and measured (e.g., measurements of optical density, absorption and fluorescence, and qualitative/quantitative analysis of intra- or extracellular metabolites) on the instruments. These instruments can be used in any combination and are controlled under the automation system according to a prepared protocol. The ANL learns from experimental data that are obtained automatically by the ANL, and then suggests optimized experimental conditions to researchers. The concept and demonstration of ANL are shown in Supplementary movies 1 and 2, respectively.

each nutrient at various concentrations were automatically evaluated using the ANL, according to the prepared protocol (Fig. 2).

Under the conditions for adding large amounts of $CoCl_2$ and $ZnSO_4$ of the trace elements (0.1 to 1 μM addition), the cell growth was promoted (Fig. 3, upper panel), whereas under the conditions for adding small amounts of $CaCl_2$ and $MgSO_4$ of the basic components (0.2 to 4 mM addition), the glutamic acid production was promoted (Fig. 3, middle panel). However, glutamic acid productivities, accompanied by increases in both cell growth parameters and glutamic acid production, were not improved (Fig. 3). Because Fe^{2+} cations in the culture media precipitated, especially at high concentrations, the optical densities of the cells were not measured (Fig. 3, upper panel). Under the conditions for adding large amounts of Na_2HPO_4, KH_2PO_4, NH_4Cl, and NaCl of the basic components (40 to 400 mM addition), the optical densities and the glutamic acid concentrations from the cells were low compared with controls (Fig. 3, upper and middle panels). These salts may affect the osmotic

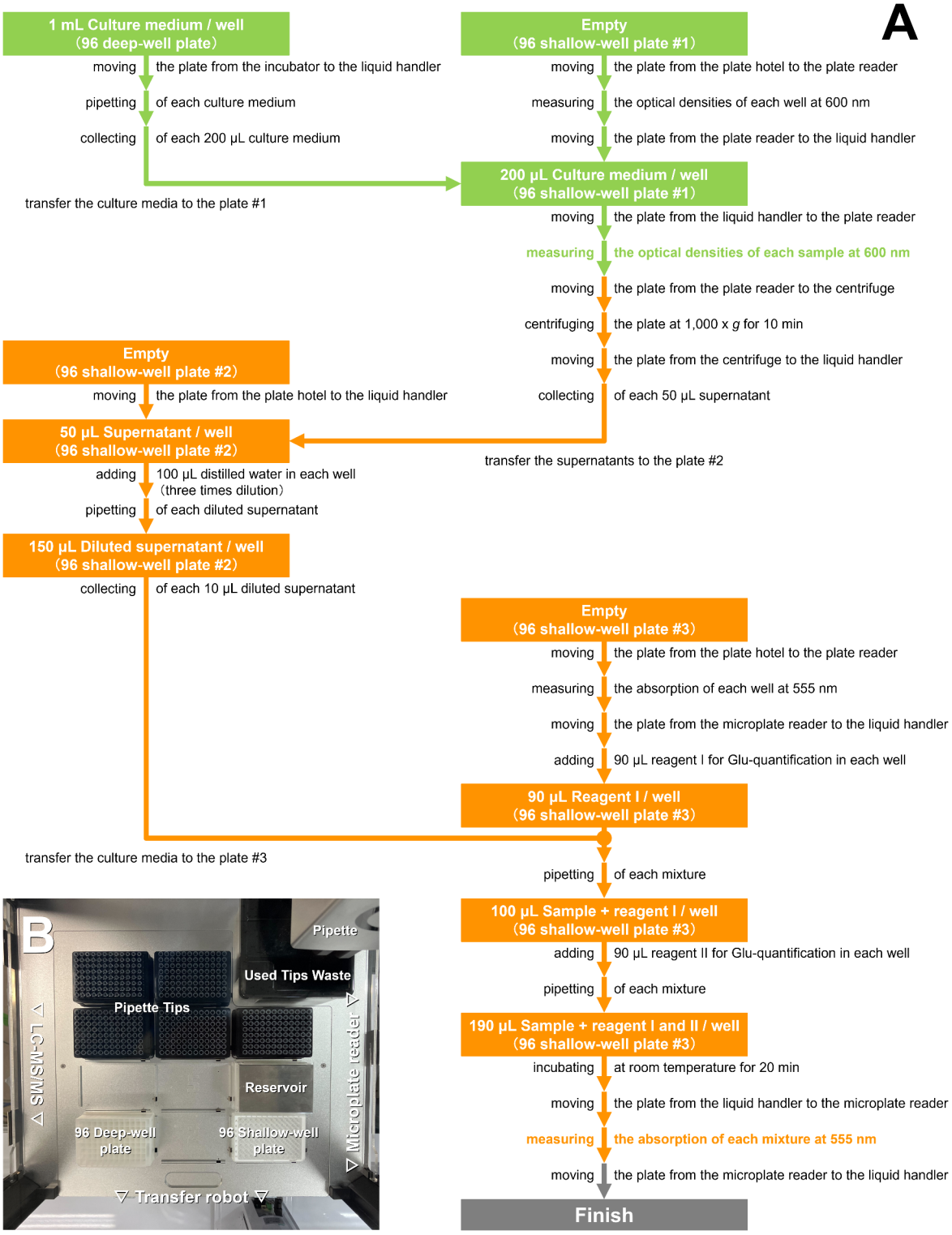

Fig. 2. Experimental process in the ANL. (**A**) Workflow to measure the cell growth (shown in yellow-green) and of the product amount (shown in orange) of glutamic acid-producing cells. The cell growth was evaluated by the optical densities from the cells at 600 nm. The product amount was evaluated by the colorimetric determination: The diluted supernatants were reacted with reagents I and II of the glutamic acid-quantification kit, and then the absorption of the mixtures was monitored at 555 nm after incubation at room temperature for 20 min. (**B**) Workspace on the liquid handler. The transfer robot transfers the plates from some devices to the liquid handler. The pipette dispenses samples in the well plate and solvents and reagents in the reservoir into each well of the other empty well plate for dilution or assay.

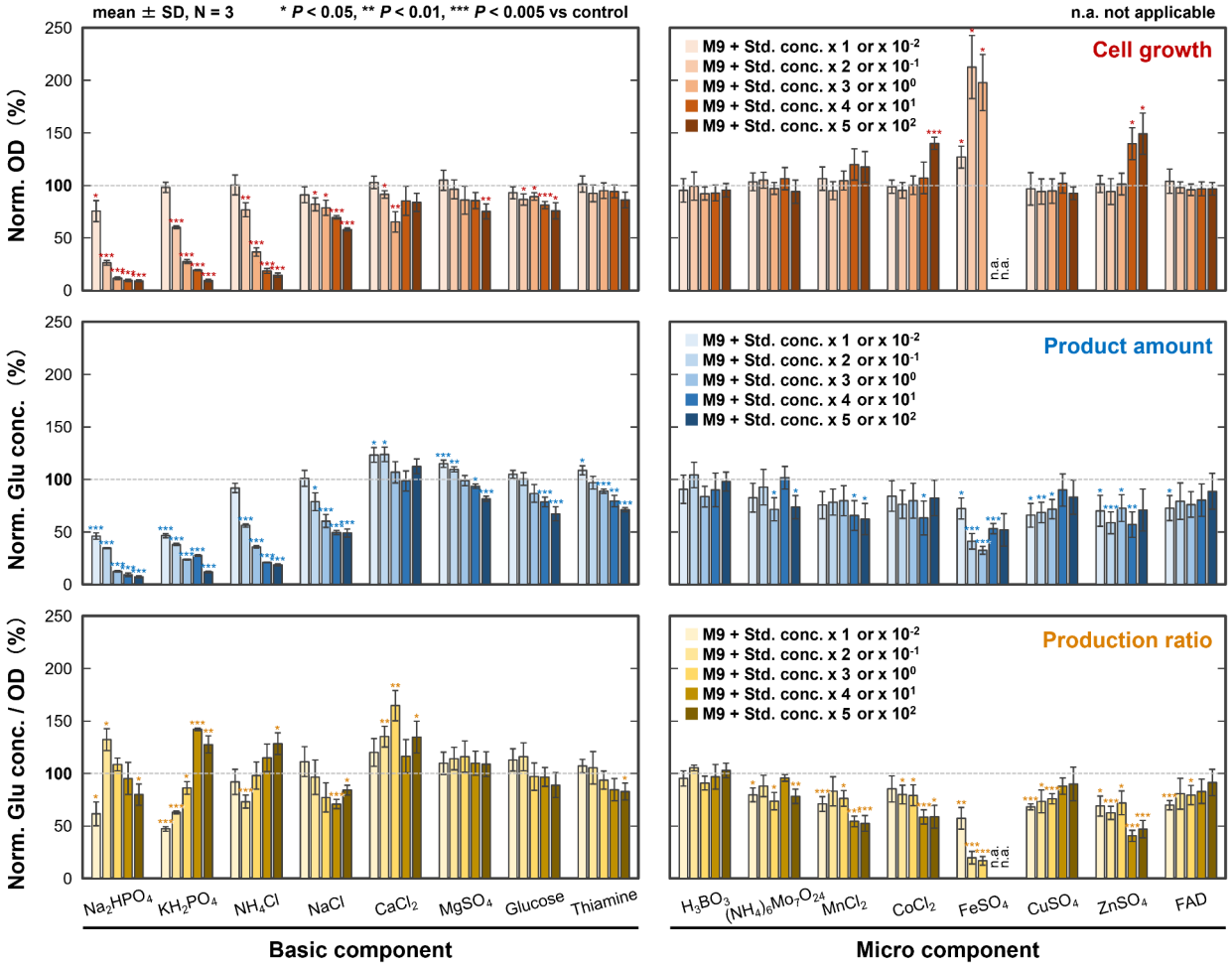

Fig. 3. Effect of the basic and trace elements added into the M9 medium on the glutamic acid-producing cells. The standard concentrations (Std. conc.) of the components were determined from the compositions of the minimal media[33,34]. Additional amounts of the basic components and flavin adenine dinucleotide (FAD) were adjusted from equal to 5 times these standard amounts. Additional amounts of the trace elements, except for FAD, were adjusted from 0.01 to 100 times these standard amounts. Details of their concentrations were shown in (Table S1). The optical densities (shown in red) and the extracellular glutamic acid concentrations (shown in blue) from the cells cultivated in the 80 adjusted media for 24 h were measured. The glutamic acid productivities (shown in yellow) were calculated from the measurements of the cell growth and the product amounts. These data were normalized by those of the cells cultivated in the control medium (shown as Norm. OD, Norm. Glu conc., and Norm. Glu conc./OD; mean ± SD, $n = 3$). Significant differences compared to the control were shown statistically by Student's t-test (* $P < 0.05$, ** $P < 0.01$, and *** $P < 0.005$). "n.a." means not applicable. These data were obtained automatically under the ANL.

pressure of the cells and inhibit their function. Other components (glucose, thiamine, H_3BO_3, $(NH_4)_6Mo_7O_{24}$, $MnCl_2$, $CuSO_4$, and FAD) were ineffective at improving glutamic acid production (Fig. 3). Therefore, we focused on the following four components: $CaCl_2$, $MgSO_4$, $CoCl_2$, and $ZnSO_4$.

Optimization of medium conditions using the ANL

We used a Bayesian optimization algorithm to adjust the concentrations of $CaCl_2$, $MgSO_4$, $CoCl_2$, and $ZnSO_4$ in the culture medium to maximize objective variables, such as cell growth and glutamic acid production (Figs. 4 and 5). First, the initial dataset consisting of their concentrations and objective variables (Fig. 3) was input to the optimization algorithm, and then the next combinations of the concentrations to be tested were obtained (Fig. 4A-i, and 5A-i). The combinations of concentrations were then tested using the ANL to obtain the objective variables (Fig. 4B-i, and 5B-i). The obtained experimental data were added to the dataset, and the dataset was again subjected to the optimization algorithm to obtain the next combination of concentrations (Fig. 4A-ii to iv, 4B-ii and iii, 5 A-ii to iv, and 5B-ii and iii). This cycle was repeated to allow for autonomous and effective improvement of the objective variable. In addition, we used batch Bayesian optimization, which can provide multiple promising combinations of four-component concentrations and allow for efficient experiments with simultaneous processing of multiple conditions using a 96-well plate. We obtained $N = 3$ experimental data

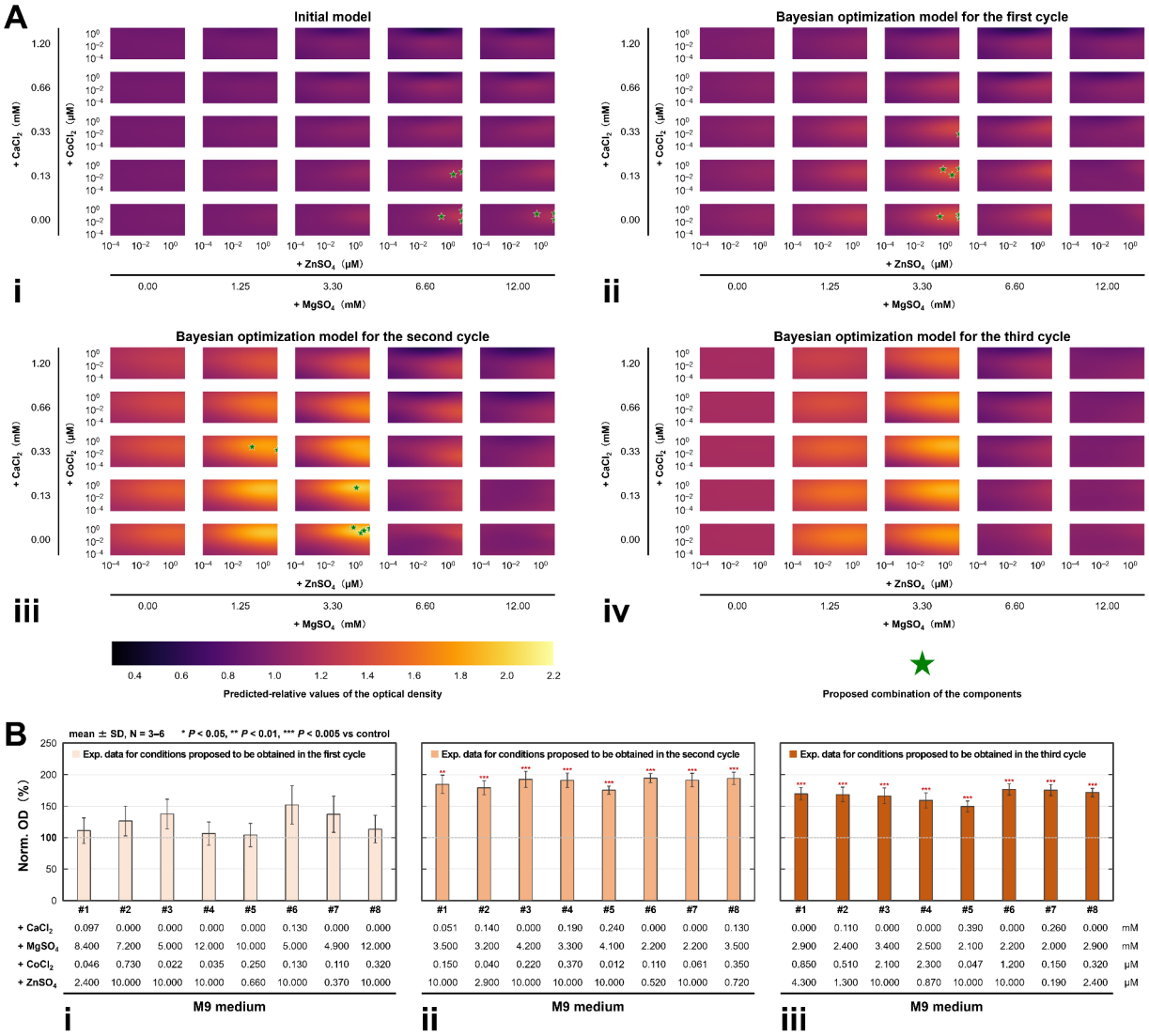

Fig. 4. Development of an algorithm to derive the optimized medium conditions for the cell growth of the glutamic acid-producing cells. (**A**) Bayesian-optimized models for the cell growth: (i) Model constructed on the basis of the initial dataset. (ii)–(iv) Bayesian-optimized models successively updated on the cell growth data shown in (Fig. 4B). Additional amounts of the four components ($CaCl_2$, $MgSO_4$, $CoCl_2$, and $ZnSO_4$) into the M9 medium are shown on the X- and Y-axes of the heatmaps. Color scales are shown as predicted-relative values of the optical densities (vs. the control) when the cells were cultivated under the adjusted medium conditions. Star symbols indicate the proposed combinations of the four components. (**B**) Cell growth data obtained from the cells cultivated under the adjusted medium conditions proposed by the Bayesian-optimized models shown in (Fig. 4A). These data were normalized by those of the cells cultivated in the control medium (shown as Norm. OD, mean ± SD, $n = 3–6$). Significant differences compared to the control were shown statistically by Student's *t*-test (* $P < 0.05$, ** $P < 0.01$, and *** $P < 0.005$). These data were obtained automatically under the ANL.

points under 256 medium conditions consisting of four-component combinations in a grid search to compare the performance of the optimization algorithm (Figs. S6 and S7).

We compared the results of searching for medium condition that increases cell growth using two algorithms. Figure 4B shows the results of three experimental cycles with $N = 3$ measurements for the eight conditions obtained by the Bayesian optimization of cell growth using eight data pairs as the initial dataset. The best condition (M9 + 0.000 mM $CaCl_2$, 2.200 mM $MgSO_4$, 0.110 μM $CoCl_2$, 0.520 μM $ZnSO_4$) out of the 24 conditions obtained by Bayesian optimization algorithm was 94.7% better than the M9 medium for cell growth (Fig. 4B-ii, #6), while the best condition (M9 + 0.000 mM $CaCl_2$, 0.000 mM $MgSO_4$, 0.100 μM $CoCl_2$, 10.000 μM $ZnSO_4$) out of the 256 conditions obtained by grid search was 77.3% better than the M9 medium (Fig. S6). These data suggest that combination and balance of the $MgSO_4$, $ZnSO_4$ and $CoCl_2$ concentrations might be important for

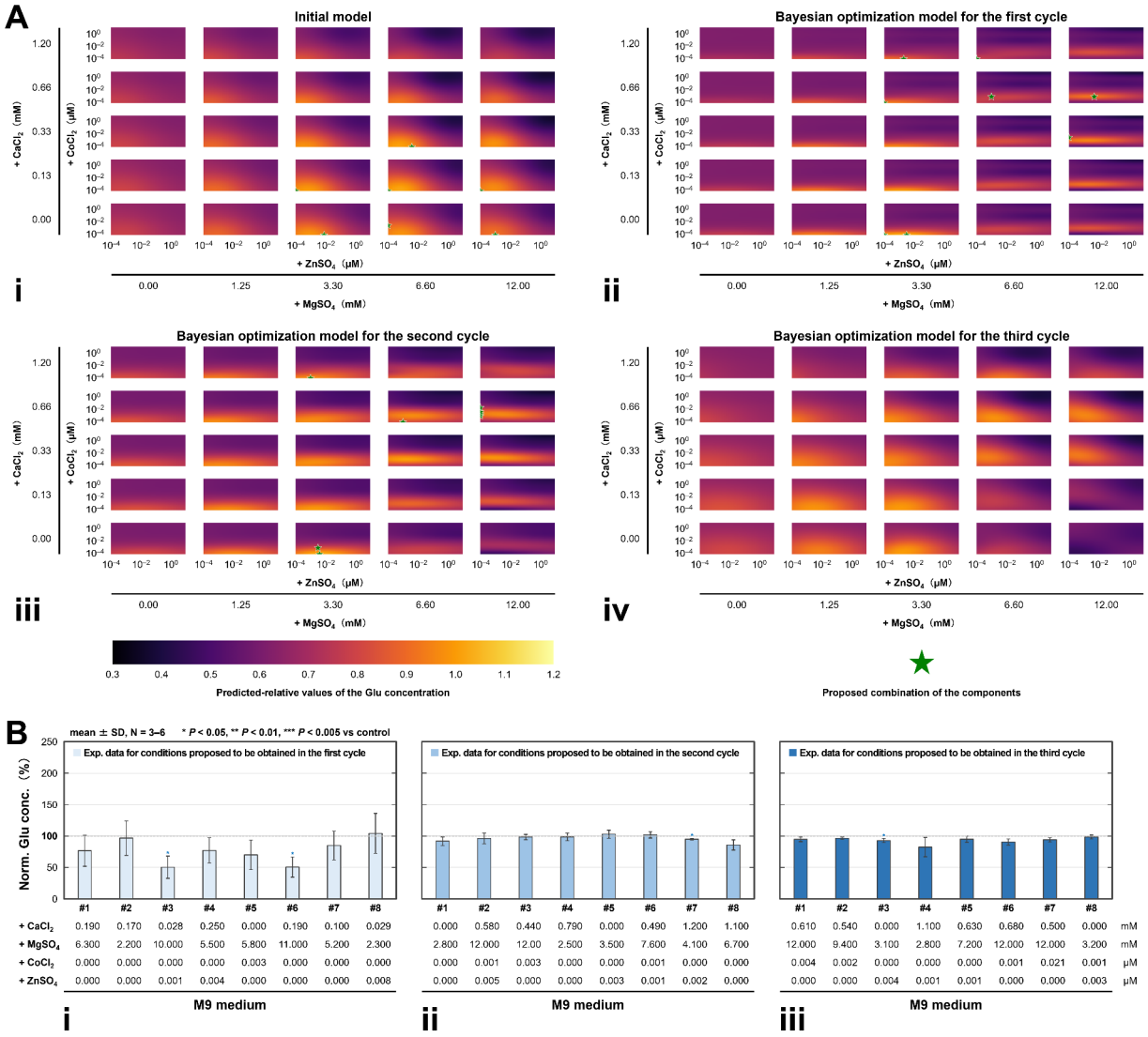

Fig. 5. Development of an algorithm to derive the optimized medium conditions for the product amount of the glutamic acid-producing cells. (**A**) Bayesian-optimized models for the glutamic acid production: (i) Model constructed based on the initial dataset. (ii)–(iv) Bayesian-optimized models successively updated on the glutamic acid production data shown in (Fig. 5B). Additional amounts of the four components ($CaCl_2$, $MgSO_4$, $CoCl_2$, and $ZnSO_4$) into the M9 medium are shown on the X- and Y-axes of the heatmaps. Additional amounts of the four components ($CaCl_2$, $MgSO_4$, $CoCl_2$, and $ZnSO_4$) into the M9 medium are shown on the X- and Y-axes of the heatmaps. Color scales are shown as predicted-relative values of the glutamic acid concentrations (vs. the control) when the cells were cultivated under the adjusted medium conditions. Star symbols indicate the proposed combinations of the four components. (**B**) Glutamic acid production data obtained from the cells cultivated under the adjusted medium conditions proposed by the Bayesian-optimized models shown in (Fig. 5A). These data were normalized by those of the cells cultivated in the control medium (shown as Norm. Glu conc., mean ± SD, $n = 3–6$). Significant differences compared to the control were shown statistically by Student's *t*-test (* $P < 0.05$, ** $P < 0.01$, and *** $P < 0.005$). These data were obtained automatically under the ANL.

cell growth. It has been reported that: (i) magnesium plays a role in population doubling time, morphological phenotypes, and osmotic pressure of *E. coli*[35]; (ii) zinc causes dramatic changes in cell transcripts, antibiotic response, heat stress, growth regulation, cell shape, and biofilm formation of *E. coli*[36]; and (iii) although cobalt exhibits toxicity to *E. coli* through its competition with iron in various metabolic processes[37], it may subtly influence the regulation of growth. In this context, growth of the glutamic acid-producing cells derived from the *E. coli* strain may be regulated in response to the metal ions under these regulatory mechanisms. In summary, the Bayesian optimization algorithm was able to find the medium conditions that resulted in higher cell growth with fewer search points than grid search.

Next, we compared the results of searching for medium condition that increases glutamic acid production using two algorithms. Figure 5B shows the results of three experimental cycles with $N=3$ measurements for the eight conditions obtained by the Bayesian optimization of glutamic acid production using eight data pairs as the initial dataset. The best condition (M9 + 0.029 mM $CaCl_2$, 2.300 mM $MgSO_4$, 0.000 µM $CoCl_2$, 0.008 µM $ZnSO_4$) out of the 24 conditions obtained by the Bayesian optimization algorithm was 3.5% better than the M9 medium for glutamic acid production (Fig. 5B-i, #8), while the best condition (M9 + 0.400 mM $CaCl_2$, 0.000 mM $MgSO_4$, 0.100 µM $CoCl_2$, 0.000 µM $ZnSO_4$) out of 256 conditions obtained by grid search was 18.1% better than the M9 medium (Fig. S7). By continuing more experimental cycles to improve the Bayesian optimization algorithm, a best condition for high glutamic acid production may propose better than the grid search. However, even with a grid search that is searching widely with 256 points, at most only an 18.1% improvement was seen, so unless there is a large increase or decrease in production within the local range of the explanatory variables, it may not be possible to expect a large increase in production (Fig. S7). In general, glutamic acid plays a vital role in nutrition, amino acid metabolism, and protein synthesis. The balance between the production and consumption of glutamic acid is regulated during the cell growth process in *E. coli*. Furthermore, a balance of the intra- and extracellular glutamic acid concentrations is very important for microorganisms to protect from osmotic pressure[38,39] and pH[40,41] stress. Therefore, synthesis, degradation, and release of glutamic acid should be strictly and intricately regulated in the glutamic acid-producing cells. However, in this study, we were unable to artificially manipulate these regulatory mechanisms through the optimization of the nutrients related to the glutamic acid synthesis. This suggests that clearly demonstrating the correlation between the nutrient, which is only one explanatory variable of medium conditions, and glutamic acid productivity is challenging. In other words, the amounts of glutamic acid products should be explored with considering several explanatory variables, not only nutrients but also osmotic pressure and pH. In conclusion, the ANL could automatically obtain the large amounts of experimental data and efficiently suggest the medium condition for high cell growth but could not suggest the medium condition for high glutamic acid, which is thought to involve many explanatory variables, based only on nutrient concentration.

Factor analysis of cell growth

To compare the specific growth rate and maximum growth of the glutamic acid-producing cells in the M9 medium (basic condition) and optimized medium (for high cell growth condition; M9 + 0.000 mM $CaCl_2$, 2.200 mM $MgSO_4$, 0.110 µM $CoCl_2$, 0.520 µM $ZnSO_4$), the cells were incubated at 20, 22, 23, 24, 25, 26, 28, and 48 h in these media in 1-to-100-mL scales: The specific growth rates were obtained by the regression with 20–26 h as the log phase (Fig. 6A–C). The maximum growths were obtained for 48-hour incubation (Fig. 6D). We first compared these factors in the case of the 1-mL scale cultivation, which is the same condition as the case of the previous verification tests. The specific growth rate under the optimized condition ($\mu = 0.0529 \pm 0.0015\ h^{-1}$) was 8.4% higher than that under the basic condition ($\mu = 0.0488 \pm 0.0029\ h^{-1}$) (Fig. 6A). The maximum growth under the optimized condition (OD = 0.2984 ± 0.0092) was significantly 31.7% higher than that under the basic condition (OD = 0.2266 ± 0.0081) (Fig. 6D). The results indicate that the medium found using the optimization algorithm has increased both the growth rate and the maximum growth of the glutamic acid production, and it is suggested that higher cell growth was obtained due to these factors.

Similarly, in the cases of the 10- and 100-mL scale cultivations, the specific growth rates and maximum growths under the optimized condition (specific growth rates, $\mu = 0.0353 \pm 0.0036\ h^{-1}$ (10-mL scale) and $0.0346 \pm 0.0017\ h^{-1}$ (100-mL scale); maximum growths, OD = 0.2816 ± 0.0087 (10-mL scale) and 0.3062 ± 0.0027 (100-mL scale)) were higher than those under the basic condition (specific growth rates, $\mu = 0.0268 \pm 0.0026\ h^{-1}$ (10-mL scale) and $0.0284 \pm 0.0047\ h^{-1}$ (100-mL scale); maximum growths, OD = 0.1887 ± 0.0102 (10-mL scale) and 0.1855 ± 0.0101 (100-mL scale)) (Fig. 6B–D). Since the optimized medium conditions at the 1-mL scale showed an increase in both growth rate and maximum growth rate regardless of the scale of culture, it is expected that ANL can propose useful medium conditions even for actual production scales.

Discussion

We developed the ANL, which is an autonomous experimental system combining robotics and AI based on Bayesian optimization algorithms (Fig. 1 and S1), and built a tool adaptable to any biotechnology experimental system, such as microbial culture and high-throughput evaluation experiments.

We conducted an experiment to optimize the medium conditions for the glutamic acid producing strain as a demonstration to show the usefulness of the ANL (Figs. 2 and 3). The ANL could propose the optimized medium condition which improved the cell growth in 1-mL scale (Fig. 4), and it was also confirmed that this medium condition showed high growth rate and maximum cell growth in 10- and 100-mL scale (Fig. 6). On the other hand, an optimized condition to improve the glutamic acid productivity could not be found (Fig. 5). This is thought to be due to glutamic acid concentration was intricately regulated in the cells[38–41].

From these results, at least, the ANL should be able to optimize a condition based on the relationship between some explanatory and objective variables. The Bayesian optimization method can assist in exploring an optimized condition from a combination of multiple explanatory variables[42–44], and its improvement will expect to suggest much better one. The ANL based on the Bayesian optimization algorithm introduced with multiple explanatory variables should support to obtain better medium conditions for the cell growth and bioproduction.

Regarding cell growth, Bayesian optimization was able to find conditions that were 94.7% better than M9 basic medium (Fig. 4), but in the factor analysis experiment conducted under the same conditions, cell growth was only 2.9% better than M9 basic medium (Fig. 6A). One possible reason for the differences in measurements is that the culture time was set at 24 h, which is the log phase, therefore a small difference in time may have caused a large difference in cell growth. If the culture time can be adjusted during the optimization cycle to obtain cell growth during the stationary phase, more stable medium optimization can be expected. For a more

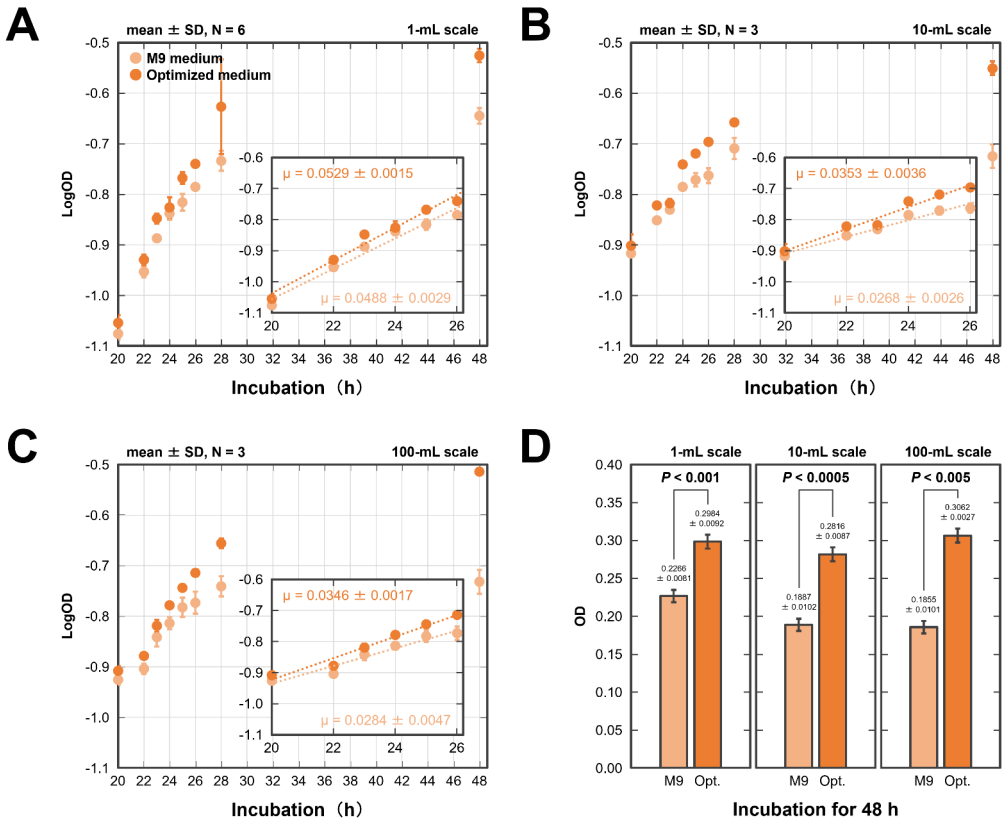

Fig. 6. Comparison of cell growth between the glutamic acid-producing cells cultivated under the M9 medium condition (light orange) and those under the optimized medium condition (dark orange). (**A**–**C**) Cell growth curves of the cells cultivated for 20–48 h in 1 mL (**A**), 10 mL (**B**), and 100 mL (**C**) scale containers. Y-axes were shown as ordinary logarithms of the optical densities of the cells, which were measured at 20, 22, 23, 24, 25, 26, 28 and 48 h. The specific growth rates (μ) were calculated from these data on the log growth phases for 20–26 h incubations of the cells (shown in the insets). (**D**) Optical densities of the cells cultivated for 48 h in the scale containers. These data were shown as mean ± SD, $n = 3$–6. Significant differences between these conditions were statistically calculated by Student's t-test. These data were obtained automatically under the ANL.

stable measurement and optimization cycle, it would be desirable to include incubation time as an explanatory variable.

The ANL can autonomously pretreatment samples and monitor various data such as absorption, fluorescence, and molecular mass. In this study, we presented the ANL combined with a microplate reader and a liquid chromatography-mass spectrometry (LC-MS) (Fig. 1). In this experiment, the ANL successfully replicated the experimental techniques such as cell incubation, sample preparation (i.e., centrifugation, dilution, and mixing), optical density measurement, and qualification and quantification of the produced glutamic acid (Fig. 2). To verify the ANL's performance, although we limited the microorganism (*E. coli*), the medium (M9), and the target (glutamic acid) to a single type each, inherently, the ANL is a system to provide a versatile and scalable solution for various applications in the field of bioproduction. The ANL has a potential to improve efficiency and reliability of experimental processes in the future. To further demonstrate the utility, we are now continuing verifications, focusing on other target molecules. Some of the broader applications of the ANL could include sample preparation and formulation development in the pharmaceutical field, as well as process development in the materials field. Although the ANL currently lacks modules for dispensing powders, bioreactors, or highly viscous liquids, the development of these modules would expand its range of applications.

In this context, we further have improved the system to be able to combine not only with the microplate reader and LC-MS but also with a gas chromatography-mass spectrometry (GC-MS) to detect various biomolecules (Fig. S8). The ANL combined with various experimental devices and multiple explanatory variables will be able to measure a wide range of biomolecules and to analyze objective variables which are complicatedly regulated. The improved system will then suggest the best conditions based on these vast amounts of information. Further, the system should be applicable to design various biomolecules, such as enzymes and metabolisms, in the field of bioproduction. A wide range of organic compounds have been produced based on bioproduction (e.g., lycopene[45], β-carotene[46], 6-methylsalicylic acid[47], erythromycin[48], (*S*)-reticuline[49], L-DOPA[50], yersiniabactin[51], and echinomycin[52]), but in many cases, their producibilities have not yet reached on a practical level. These problems could be improved by optimization based on the vast amounts of data analysis. In this context, it

is expected that the ANL will enable the production of the useful compounds which have been difficult to synthesize and will improve their yields. Further, the ANL would be the basis for spurring the development of bioproduction technology to produce various compounds.

Materials and methods

Creation of autonomous lab system

Autonomous Lab system was composed of a transfer robot (PF400, Brooks, Chelmsford, MA, USA) with a plate hotels, a microplate reader (SpectraMax iD3, Molecular Devices, San Jose, CA, USA), a centrifuge (HiG, BioNex, San Jose, CA, USA), an incubator (STX44, LiCONiC, Mauren, Liechtenstein), a liquid handler (OT-2, Opentrons, NY, USA) and LC-MS/MS system (Nexera XR, LCMS-8060NX, Shimadzu, Kyoto, Japan).

These instruments were integrated into the automation system constructed by the control programs (Fig. S1). The instruments that configure this system can be freely rearranged and added around the robot. The system used in this study was designed for the metabolic analysis of bacterial cells. The control software consists of a web application and a local control application, both of which are in-house developed. In the web application, it is possible to set a protocol on the GUI (graphical user interface) that describes the order and conditions in which the instruments constituting the system perform preprocessing and measurement of the samples (Fig. S2). The local control application transports the samples and controls the instruments according to the protocol. The web application associates and manages the information of strains and media with the measurement results of the amount of substances produced by the bacterial cells. By associating and collecting the strain and medium conditions and the amount of the substance produced, it is possible to search for the optimal strain and medium conditions by the artificial intelligence.

Batch bayesian optimization with datasets

Batch Bayesian optimization was used with the concentrations of $CaCl_2$, $MgSO_4$, $CoCl_2$, and $ZnSO_4$ as explanatory variables, and either cell growth or glutamic acid production as objective variables. The concentrations of $CoCl_2$ and $ZnSO_4$ were used as logarithmic values to reduce the influence of a wider range of parameters. The lower and upper bounds were set at 0.2–1.4 mM for $CaCl_2$, at 2–14 mM for $MgSO_4$, and at 0.0001-10 μM for $CoCl_2$ and $ZnSO_4$ (the trace elements were not added in the case of the calculated concentrations were below 0.0001). The explanatory and objective variables were normalized and standardized, respectively, and then these parameters were restored to their original scales at the time of output. The Gaussian process regression and the acquisition function were performed using SingleTaskGP from BoTorch[53] and qNoisyExpectedImprovement, respectively.

Bacterial strains, media and plasmid construction

The *Escherichia coli* (*E. coli*) strain DH5α (TaKaRa, Shiga, Japan) was used for cloning plasmid DNA. The *E. coli* strain BW25113 (National Institute of Genetics, Shizuoka, Japan) was used for based-cells to construct the glutamic acid-producing cells. The Lysogeny Broth (LB) medium (0.5% (w/v) yeast extract, 1.0% (w/v) tryptone and 0.5% (w/v) NaCl) was used for the bacterial cell cultivation. M9 medium[34] (a synthetic minimal one for *E. coli*; 84.5 mM Na_2HPO_4, 44.1 mM KH_2PO_4, 37.4 mM NH_4Cl, 17.1 mM NaCl, 0.2 mM $CaCl_2$, 2.0 mM $MgSO_4$, 22.4 mM glucose and 59.3 μM thiamine) was used for the base one to optimize the bacterial cell growth and the glutamic acid production. The adjusted media, based on the M9 medium added with basic components and/or trace elements, were used for the evaluation for the cell's capabilities: The basic components (Na_2HPO_4, KH_2PO_4, NH_4Cl, NaCl, $CaCl_2$, $MgSO_4$, glucose and thiamine) are constituents of the M9 medium. The trace elements (H_3BO_3, $(NH_4)_6Mo_7O_{24}$, $MnCl_2$, $CoCl_2$, $FeSO_4$, $CuSO_4$, $ZnSO_4$ and flavin adenine dinucleotide (FAD)) are constituents of the MOPS (3-(*N*-morpholino)propanesulfonic acid) medium[33] (a synthetic minimal one for the high growth of *E. coli*) or cofactors and activators related to the glutamic acid biosynthesis. Their standard concentrations were decided based on the compositions of minimal media[33,34]. Additional amounts of these components into the M9 medium were shown in Table S1.

pET44a vector (Novagen, Madison, WI, USA) and pGETS118 vector which is a shuttle one between *Bacillus subtilis* (*B. subtilis*) and *E. coli*[54] were used for overexpression of glycolysis- and tricarboxylic acid (TCA) cycle-related genes. The pET44a-based plasmid was constructed using PrimeSTAR Max DNA polymerase (TaKaRa, Shiga, Japan) and In-Fusion HD Cloning kit (TaKaRa, Shiga, Japan). The pGETS118-based plasmid was constructed by Ordered Gene Assembly in *B. subtilis* (OGAB) method which enables to insert multiple genes with or without promotor/terminator sets into the shuttle vector[54–56]. In this study, we inserted 10 genes from *E. coli* into the pET44a and the pGETS118 vectors (details were shown below). These genes were amplified by polymerase chain reaction (PCR) with the *E. coli* genome or synthetic gene templates.

Construction of the glutamic acid-producing cells

To enhance the amount of glutamic acid production from glucose, glycolysis- and TCA cycle-related gene functions in *E. coli* strain BW25113 were regulated: the *pts*HI (encoding sugar-transport molecules on phosphoenolpyruvate-dependent phosphotransferase system (PTS)), *pta* (encoding a phosphotransacetylase), *ack*A (encoding an acetate kinase) and *zwf* (encoding a glucose-6-phosphate dehydrogenase) genes in the cell genome were destructed by lambda-red recombination system[57,58]. The *gal*P-*glk* (encoding a glucose symporter, and a glucokinase, respectively) genes were overexpressed on the pET44a-based plasmid in the cells. These gene expression was regulated by PR promotor from lambda phage and rrnB terminator (these nucleotide sequences were shown in Fig. S9). The *pgi* (encoding a glucose-6-phosphate isomerase), *pfk*A (encoding a 6-phosphofructokinase), *gap*A (encoding a glyceraldehyde-3-phosphate dehydrogenase), *gpm*A (encoding a 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase), *eno* (encoding an enolase), *ppc* (encoding a phosphoenolpyruvate carboxylase), mutated *lpd*A[59,60] (encoding a dihydrolipoamide dehydrogenase) and *ace*F (encoding a dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex)

genes were overexpressed on the pGETS118-based plasmid in the ceotolls. These gene expression was regulated by each artificial promotor/terminator (these nucleotide sequences were shown in Figs. S10–S17). The metabolic pathway in the glutamic acid-producing cells with these genes were shown in (Fig. S3).

Cultivation of the glutamic acid-producing cells

The glutamic acid-producing cells were pre-cultured in 2 mL LB medium containing 50 μg/mL carbenicillin with 12.5 μg/mL chloramphenicol overnight at 37°C at 180 rpm using a BR-23FP Bioshaker (Taitec, Saitama, Japan). The culture medium was removed by centrifugation at 10,000 x rpm for 2 min. The cells were washed with 1 mL M9 medium containing the antibiotics, and then collected by centrifugation as the same condition. The collected cells were suspended in 2 mL M9 medium containing the antibiotics. After dispensing 1 mL adjusted media containing the antibiotics in each well of the sterilized 96 deep-well plates, 10 μL suspension was added to these media. After sealing on these plates, the cells were cultured for 0–48 h at 37°C at 1000 rpm using a M·BR-420FL maximizer (Taitec, Saitama, Japan). The samples were set in the incubator of the ANL-equipment after removing the seals on the plates.

In the case of tube- and flask-scale cultivation, 2 to 100 mL adjusted media were packed in these sterilized containers, and then 20 to 1,000 μL suspension of the pre-cultured cells was added to these media, respectively. The cells were cultured for 0–48 h at 37°C at 180 rpm using the shaker.

Evaluation of the glutamic acid-producing cell's capabilities by the ANL

200 μL culture broth on 96 deep-well plates were dispensed to shallow-well plates after pipetting them. The optical densities from the cells at 600 nm were monitored by the microplate reader. Furthermore, the shallow-well plates were centrifuged at 1,000 x *g* for 10 min to remove the cultivated cells. 50 μL supernatants of them were collected in other shallow-well plates, and then diluted to 3 times. Finally, the 10 μL diluted samples were dispensed to other shallow-well plates, and then mixed with each 90 μL of reagent I and II of Yamasa NEO which is a glutamic acid-quantification kit[61,62] (Yamasa, Chiba, Japan). The absorption of these mixtures was monitored at 555 nm after incubation at room temperature for 20 min. The amount of glutamic acid was quantified using a standard curve.

These measured values were subtracted from those of each empty plate, as a blank. These data were obtained from 3 to 6 independent experiments which were performed automatically under the ANL according to the created protocol (the workflow was shown in Fig. 2).

Identification and quantification of intra- and extracellular glutamic acids of the glutamic acid-producing cells cultivated under the nutrient-rich conditions

The quantitative evaluation of the intra- and extracellular glutamic acid amounts of the glutamic acid-producing cells cultivated under the nutrient-rich conditions was performed by using the LC-MS/MS system of the ANL-equipment. The cells were cultured in the 2 mL M9 media containing the antibiotics with or without 0.005 to 0.5% (w/v) yeast extract packed in sterilized tubes (cultivation condition was shown above). After the optical densities from the cells at 600 nm were monitored by the microplate reader, the culture broth was centrifuged at 10,000 x rpm for 2 min to separate the cells and the culture medium.

To detect the intracellular glutamic acid, the cells were washed with 500 μL distilled water containing 10 μM phenethylamine (internal standard (IS)), and then centrifuged at 15,000 x rpm for 2 min. After removing the supernatants, the cells were suspended in the 100 μL IS solution, and then transferred to microtubes with glass beads. The cells were homogenized at 50 Hz for 3 min using a TissueLyser LT (QIAGEN, Venlo, Netherlands). The mixtures were further added with the 400 μL IS solution, and then suspended. The mixtures were centrifuged at 15,000 x rpm for 3 min to remove the pellets and to collect the supernatants, which were diluted to 50 times with the IS solution. On the other hand, to detect extracellular glutamic acid, the culture media were diluted to 100 or 200 times with the IS solution.

After these samples were filtered using a 0.22 μm PTFE filter, 1 μL of the samples were eluted by the HPLC system under the following conditions: column (SunShell Biphenyl column, 2.1 i.d. x 50 mm, 2.6 μm (ChromaNik Technologies, Osaka, Japan)), temperature (40°C), solvent (5% (v/v) MeOH / 0.1% (v/v) HCOOH in water) and flow rate (0.5 mL/min in isocratic mode). Glutamic acid in these samples was detected by the MS/MS system under the following conditions: ionization (electrospray ionization (ESI)), ionization parameters (3.0 L/min nebulizing gas flow, 10.0 L/min heating gas flow, 10.0 L/min drying gas flow, interface temperature at 300°C, DL temperature at 250°C, heat block temperature at 400°C) and detection mode (multiple reaction monitoring (MRM)). MRM transitions of m/z 148 (+) > m/z 56, 84 and 130 (+) from glutamic acid were used for identification and quantification[63]. The amount of glutamic acid was quantified using a standard curve, which was generated from the peak area ratio (glutamic acid/IS).

Data availability

Data is provided within the manuscript or supplementary information files.

Received: 6 December 2024; Accepted: 3 February 2025
Published online: 24 February 2025

References

1. Stephenson, A. et al. Physical laboratory automation in synthetic biology. *Science* **12**, 3156–3169 (2023).
2. King, R. D. et al. The automation of science. *Science* **324**, 85–89 (2009).
3. Shimizu, R., Kobayashi, S., Watanabe, Y., Ando, Y. & Hitosugi, T. Autonomous materials synthesis by machine learning and robotics. *APL Mater.* **8**, (2020).

4. Nikolaev, P. et al. Autonomy in materials research: a case study in carbon nanotube growth. *NPJ Comput. Mater.* **2**, 16031 (2016).
5. Hooe, S. L. et al. Multienzymatic cascades and nanomaterial scaffolding—A potential way forward for the efficient biosynthesis of novel chemical products. *Adv. Mater.* **36**, (2023).
6. Essa, M. M. et al. Functional foods and their impact on health. *J. Food Sci. Technol.* **60**, 820–834 (2023).
7. Chopra, B. & Dhingra, A. K. Natural products: a lead for drug discovery and development. *Phytother. Res.* **35**, 4660–4702 (2021).
8. Yang, D. et al. *Escherichia coli* as a platform microbial host for systems metabolic engineering. *Essays Biochem.* **65**, 225–246 (2021).
9. Yang, D., Park, S. Y., Park, Y. S., Eun, H. & Lee, S. Y. Metabolic engineering of *Escherichia coli* for natural product biosynthesis. *Trends Biotechnol.* **38**, 745–765 (2020).
10. Wang, Z., Sun, J., Yang, Q. & Yang, J. Metabolic engineering escherichia coli for the production of lycopene. *Molecules* **25**, 3136 (2020).
11. Ferreira, S., Pereira, R., Wahl, S. A. & Rocha, I. Metabolic engineering strategies for butanol production in *Escherichia coli*. *Biotechnol. Bioeng.* **117**, 2571–2587 (2020).
12. Ward, V. C. A., Chatzivasileiou, A. O. & Stephanopoulos, G. Metabolic engineering of *Escherichia coli* for the production of isoprenoids. *FEMS Microbiol. Lett.* **365**, (2018).
13. Yin, M. Q. et al. Metabolic engineering for compartmentalized biosynthesis of the valuable compounds in Saccharomyces cerevisiae. *Microbiol. Res.* **286**, 127815 (2024).
14. Wang, S., Zhao, F., Yang, M., Lin, Y. & Han, S. Metabolic engineering of *Saccharomyces cerevisiae* for the synthesis of valuable chemicals. *Crit. Rev. Biotechnol.* **44**, 163–190 (2024).
15. Baptista, S. L., Costa, C. E., Cunha, J. T., Soares, P. O. & Domingues, L. Metabolic engineering of Saccharomyces cerevisiae for the production of top value chemicals from biorefinery carbohydrates. *Biotechnol. Adv.* **47**, 107697 (2021).
16. Guan, R. et al. Metabolic engineering for glycyrrhetinic acid production in *Saccharomyces cerevisiae*. *Front. Bioeng. Biotechnol.* **8**, (2020).
17. Kim, S. J., Kim, J. W., Lee, Y. G., Park, Y. C. & Seo, J. H. Metabolic engineering of *Saccharomyces cerevisiae* for 2,3-butanediol production. *Appl. Microbiol. Biotechnol.* **101**, 2241–2250 (2017).
18. Seifrid, M. et al. Autonomous chemical experiments: challenges and perspectives on establishing a self-driving lab. *Acc. Chem. Res.* **55**, 2454–2466 (2022).
19. Butakova, M. A., Chernov, A. V., Kartashov, O. O. & Soldatov, A. V. Data-centric architecture for self-driving laboratories with autonomous discovery of new nanomaterials. *Nanomaterials* **12**, 12 (2021).
20. Soldatov, M. A. et al. Self-driving laboratories for development of new functional materials and optimizing known reactions. *Nanomaterials* **11**, 619 (2021).
21. Rapp, J. T., Bremer, B. J. & Romero, P. A. Self-driving laboratories to autonomously navigate the protein fitness landscape. *Nat. Chem. Eng.* **1**, 97–107 (2024).
22. Freschlin, C. R., Fahlberg, S. A., Heinzelman, P. & Romero, P. A. Neural network extrapolation to distant regions of the protein fitness landscape. *Nat. Commun.* **15**, 6405 (2024).
23. Mione, F. M. et al. Closed-loop optimization of high-throughput robotic platforms for reproducible bioprocess development. *Comput. Aided Chem. Eng.* **52**, 2613–2618 (2023).
24. Kaspersetz, L. et al. Automation of experimental workflows for high throughput robotic cultivations. *Comput. Aided Chem. Eng.* **53**, 2971–2976 (2024).
25. Mione, F. M. et al. A workflow management system for reproducible and interoperable high-throughput self-driving experiments. *Comput. Chem. Eng.* **187**, 108720 (2024).
26. Helleckes, L. M. et al. From frozen cell bank to product assay: high-throughput strain characterisation for autonomous design-build-test-learn cycles. *Microb. Cell. Fact.* **22**, 130 (2023).
27. Helleckes, L. M. et al. High-throughput screening of catalytically active inclusion bodies using laboratory automation and bayesian optimization. *Microb. Cell. Fact.* **23**, 67 (2024).
28. Brosnan, J. T. & Brosnan, M. E. Glutamate: a truly functional amino acid. *Amino Acids* **45**, 413–418 (2013).
29. Zhang, Z., He, P., Cai, D. & Chen, S. Genetic and metabolic engineering for poly-γ-glutamic acid production: current progress, challenges, and prospects. *World J. Microbiol. Biotechnol.* **38**, 208 (2022).
30. Wang, D., Fu, X., Zhou, D., Gao, J. & Bai, W. Engineering of a newly isolated *Bacillus tequilensis* BL01 for poly-γ-glutamic acid production from citric acid. *Microb. Cell. Fact.* **21**, 276 (2022).
31. Cosenza, Z., Astudillo, R., Frazier, P. I., Baar, K. & Block, D. E. Multi-information source bayesian optimization of culture media for cellular agriculture. *Biotechnol. Bioeng.* **119**, 2447–2458 (2022).
32. Yoshida, K., Watanabe, K., Chiou, T. Y. & Konishi, M. High throughput optimization of medium composition for *Escherichia coli* protein expression using deep learning and bayesian optimization. *J. Biosci. Bioeng.* **135**, 127–133 (2023).
33. Neidhardt, F. C., Bloch, P. L. & Smith, D. F. Culture medium for enterobacteria. *J. Bacteriol.* **119**, 736–747 (1974).
34. Wang, L., Liu, Q., Du, Y., Tang, D. & Wise, M. Optimized M9 minimal salts medium for enhanced growth rate and glycogen accumulation of *Escherichia coli* DH5α. *Microbiol. Biotechnol. Lett.* **46**, 194–200 (2018).
35. Nepal, S. & Kumar, P. Growth, cell division, and gene expression of *Escherichia coli* at elevated concentrations of magnesium sulfate: implications for habitability of Europa and mars. *Microorganisms* **8**, 637 (2020).
36. Rihacek, M. et al. Zinc effects on bacteria: insights from *Escherichia coli* by multi-omics approach. *mSystems* 8, (2023).
37. Majtan, T., Frerman, F. E. & Kraus, J. P. Effect of cobalt on Escherichia coli metabolism and metalloporphyrin formation. *BioMetals* 24, 335–347 (2011).
38. Ogahara, T., Ohno, M., Takayama, M., Igarashi, K. & Kobayashi, H. Accumulation of glutamate by osmotically stressed *Escherichia coli* is dependent on pH. *J. Bacteriol.* **177**, 5987–5990 (1995).
39. Malcolm, H. R. & Maurer, J. A. The mechanosensitive channel of small conductance (MscS) superfamily: not just mechanosensitive channels anymore. *ChemBioChem* **13**, 2037–2043 (2012).
40. Xu, Y. et al. An acid-tolerance response system protecting exponentially growing *Escherichia coli*. *Nat. Commun.* **11**, 1496 (2020).
41. Lu, P. et al. L-glutamine provides acid resistance for Escherichia coli through enzymatic release of ammonia. *Cell. Res.* **23**, 635–644 (2013).
42. de Freitas, N. & Wang, Z. Bayesian optimization in high dimensions via random embeddings. (2013).
43. Eriksson, D. & Jankowiak, M. High-dimensional bayesian optimization with sparse axis-aligned subspaces. *Uncertain. Artif. Intell.* 493–503 (2021).
44. Binois, M. & Wycoff, N. A Survey on high-dimensional gaussian process modeling with application to bayesian optimization. *ACM Trans. Evolutionary Learn. Optim.* **2**, 1–26 (2022).
45. Sun, T. et al. Production of lycopene by metabolically-engineered *Escherichia coli*. *Biotechnol. Lett.* **36**, 1515–1522 (2014).
46. Yang, J. & Guo, L. Biosynthesis of β-carotene in engineered E. Coli using the MEP and MVA pathways. *Microb. Cell. Fact.* **13**, 160 (2014).
47. Yang, D. et al. Repurposing type III polyketide synthase as a malonyl-CoA biosensor for metabolic engineering in bacteria. *Proceedings of the National Academy of Sciences* 115, 9835–9844 (2018).
48. Zhang, H., Wang, Y., Wu, J., Skalina, K. & Pfeifer, B. A. Complete biosynthesis of erythromycin A and designed analogs using *E. Coli* as a heterologous host. *Chem. Biol.* **17**, 1232–1240 (2010).
49. Nakagawa, A. et al. A bacterial platform for fermentative production of plant alkaloids. *Nat. Commun.* **2**, 326 (2011).

50. Fordjour, E., Adipah, F. K., Zhou, S., Du, G. & Zhou, J. Metabolic engineering of *Escherichia coli* BL21 (DE3) for de novo production of l-DOPA from d-glucose. *Microb. Cell. Fact.* **18**, 74 (2019).
51. Ahmadi, M. K., Fawaz, S., Jones, C. H., Zhang, G. & Pfeifer, B. A. Total biosynthesis and diverse applications of the nonribosomal peptide-polyketide siderophore yersiniabactin. *Appl. Environ. Microbiol.* **81**, 5290–5298 (2015).
52. Watanabe, K. et al. Total biosynthesis of antitumor nonribosomal peptides in *Escherichia coli*. *Nat. Chem. Biol.* **2**, 423–428 (2006).
53. Balandat, M. et al. BoTorch: a Framework for efficient monte-carlo bayesian optimization. *Adv. Neural Inf. Process. Syst.* **33**, 21524–21538 (2019).
54. Kaneko, S., Tsuge, K., Takeuchi, T. & Itaya, M. Conversion of sub-megasized DNA to desired structures using a novel Bacillus subtilis genome vector. *Nucleic Acids Res.* **31**, e112 (2003).
55. Tsuge, K., Matsui, K. & Itaya, M. One step assembly of multiple DNA fragments with a designed order and orientation in Bacillus subtilis plasmid. *Nucleic Acids Res.* **31**, e133 (2003).
56. Tsuge, K. et al. Method of preparing an equimolar DNA mixture for one-step DNA assembly of over 50 fragments. *Sci. Rep.* **5**, 10655 (2015).
57. Murphy, K. C. Use of bacteriophage λ recombination functions to promote gene replacement in *Escherichia coli*. *J. Bacteriol.* **180**, 2063–2071 (1998).
58. Sharan, S. K., Thomason, L. C., Kuznetsov, S. G. & Court, D. L. Recombineering: a homologous recombination-based method of genetic engineering. *Nat. Protoc.* **4**, 206–223 (2009).
59. Kim, Y., Ingram, L. O. & Shanmugam, K. T. Dihydrolipoamide dehydrogenase mutation alters the NADH sensitivity of pyruvate dehydrogenase complex of *Escherichia coli* K-12. *J. Bacteriol.* **190**, 3851–3858 (2008).
60. Li, M., Ho, P. Y., Yao, S. & Shimizu, K. Effect of lpdA gene knockout on the metabolism in Escherichia coli based on enzyme activities, intracellular metabolite concentrations and metabolic flux analysis by 13 C-labeling experiments. *J. Biotechnol.* **122**, 254–266 (2006).
61. Kuksakabe, H., Midorikawa, Y., Fujishima, T., Kuninaka, A. & Yoshino, H. Purification and properties of a new enzyme, L-glutamate oxidase, from Streptomyces sp. X-119-6 grown on wheat bran. *Agric. Biol. Chem.* **47**, 1323–1328 (1983).
62. Kusakabe, H., Midorikawa, Y. & Fujishima, T. Methods for determining L-glutamate in soy sauce with L-glutamate oxidase. *Agric. Biol. Chem.* **48**, 181–184 (1984).
63. Zhang, P. et al. Revisiting fragmentation reactions of protonated α-amino acids by high-resolution electrospray ionization tandem mass spectrometry with collision-induced dissociation. *Sci. Rep.* **9**, 6453 (2019).

Acknowledgements

This study was supported by the Program for Forming Japan's Peak Research Universities (J-PEAKS) from the Japan Society for the Promotion of Science (JSPS).

Author contributions

Conceptualization, K.F., T.T., Y.T. and T.H.; methodology, K.F., Y.N. and T.H.; software, Y.N., M.I. and R.N.; validation, K.F., Y.N. and T.H.; formal analysis, K.F. and Y.N.; investigation, K.F., A.N. and R.S.; resources, K.F., Y.N., A.N., R.S., M.I. and R.N.; data curation, K.F., A.N. and R.S.; writing- original draft writing- K.F. and Y.N.; writing-review and edit, K.F., Y.N., A.N., R.S., M.I., R.N., T.T., R.H., H.Y., Y.T. and T.H.; visualization, K.F. and Y.N.; supervision, R.H., H.Y., Y.T. and T.H.; project administration, Y.T. and T.H.

Declarations

Competing interests

This study was conducted with a research fund from Shimadzu Corporation.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-89069-y.

Correspondence and requests for materials should be addressed to Y.T. or T.H.