CEA comité español de automática

XLIV Jornadas de Automática 2023

XLIV Jornadas de AUTOMÁTICA 2023 — ZARAGOZA

https://jautomatica.es

Towards automation of the Design-Build-Test-Learn (DBTL) bioengineering cycle: Application to the testing and characterization of standard bioparts.

Pushkareva, Anna[a], Beltran, Jaime[a], Díaz-Iza, Harold[a], Arboleda-García, Andrés[a], Boada, Yadira[a], Vignoni, Alejandro[a,*], Picó, Jesús[a]

[a] *Systems Biology and Biosystems Control Lab, Instituto de Automática e Informática Industrial, Universitat Politécnica de Valencia, Camino de Vera s/n, 46022, Valencia, España.*

To cite this article: Pushkareva, A., Beltran, J., Díaz-Iza, H., Arboleda-García, A., Boada, Y., Vignoni, A., Picó, J. 2023. Towards the automation of the Design-Build-Test-Learn (DBTL) bioengineering cycle.
XLIV Jornadas de Automática, 453-458. https://doi.org/10.17979/spudc.9788497498609.453

Resumen

El ciclo Diseño-Construcción-Prueba-Aprendizaje (DBTL) es un marco crucial en Biología Sintética para el desarrollo y optimización de sistemas biológicos. Sin embargo, la naturaleza manual del ciclo plantea limitaciones en términos de tiempo y mano de obra. Este artículo se centra en la aplicación de técnicas de automatización al ciclo DBTL, concretamente en el ensayo y caracterización de biopartes estándar, que son componentes esenciales de los circuitos genéticos. La automatización del proceso de ensayo puede mejorar significativamente el rendimiento, la fiabilidad y la reproducibilidad. En este artículo se analizan los retos asociados a los métodos de ensayo manuales y se exploran diversas estrategias y tecnologías de automatización que pueden resolverlos. Los métodos de cribado de alto rendimiento, la robótica de laboratorio y los algoritmos de análisis de datos son elementos clave en el proceso de automatización. Se examinan estudios de casos y avances recientes en la automatización del ciclo DBTL para pruebas de biopartes. La integración de la automatización en el ciclo DBTL ofrece numerosas ventajas, como una mayor eficacia, estandarización y control de calidad de las biopartes. También permite la exploración de espacios de diseño más amplios y la creación rápida de prototipos de sistemas genéticos complejos. Este artículo ofrece una revisión exhaustiva del estado actual de la técnica y las perspectivas de futuro en la automatización del ciclo DBTL para el ensayo y la caracterización de biopartes estándar.

Palabras clave: Automatización, Diseño-Construcción-Prueba-Aprendizaje, Estimación de parámetros, Opentrons, Circuitos genéticos.

Hacia la automatización del ciclo de bioingeniería Diseño-Construcción-Prueba-Aprendizaje (DBTL): Aplicación al ensayo y caracterización de biopartes estándar.

Abstract

The Design-Build-Test-Learn (DBTL) cycle is a crucial framework in Synthetic Biology for the development and optimization of biological systems. However, the manual nature of the cycle poses limitations in terms of time and labor. This paper focuses on the application of automation techniques to the DBTL cycle, specifically in the testing and characterization of standard bioparts, which are essential components of genetic circuits. By automating the testing process, throughput, reliability, and reproducibility can be significantly improved. This paper discusses the challenges associated with manual testing methods and explores various automation strategies and technologies that can address these challenges. High-throughput screening methods, laboratory robotics, and data analysis algorithms are key elements in the automation process. Case studies and recent advancements in automating the DBTL cycle for biopart testing are examined. The integration of automation in the DBTL cycle offers numerous advantages, including increased efficiency, standardization, and quality control of bioparts. It also enables the exploration of larger design spaces and rapid prototyping of complex genetic systems. This paper provides a comprehensive review of the current state-of-the-art and future prospects in automating the DBTL cycle for testing and characterizing standard bioparts.

Keywords: Automation, Design-Build-Test-Learn, Parameter estimation, Opentrons, Genetic circuits.

Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

1. Introduction

In the field of Synthetic Biology, the Design-Build-Test-Learn (DBTL) cycle (Figure 1) serves as a fundamental framework for the development and optimization of biological systems (Cummins et al., 2023). This iterative process involves designing genetic constructs, building them through molecular biology techniques, testing their functionality, and learning from the obtained results to refine future designs. The DBTL cycle has been pivotal in advancing synthetic biology and enabling the engineering of living organisms with desired traits and functions (Tellechea-Luzardo et al., 2022; Carbonell et al., 2020).

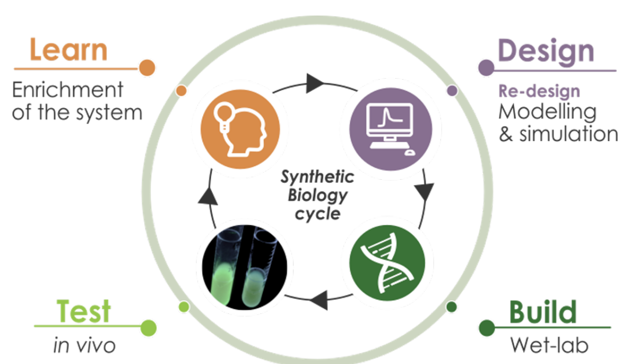

Figure 1: Design, Build, Test and Learn bioengineering cycle used in Synthetic Biology.

While the DBTL cycle has proven its effectiveness in the design and construction of biological systems (Gurdo et al., 2023), it is a labor-intensive and time-consuming process that often requires significant manual intervention. As the complexity of bioengineering projects continues to increase, there is a growing need to streamline and automate the DBTL cycle to accelerate the development of biological solutions (Cummins et al., 2023).

There is work being done towards a fully automated DBTL cycle tackling one of the steps at the time; mainly for the automation of the design step (Buecherl and Myers, 2022; Radivojević et al., 2020; Vidal et al., 2022b; Boada et al., 2021, 2022a,b), build step (Ko et al., 2022; Kang et al., 2022; Bryant Jr et al., 2023), for the test step there are advances in the calibration of the measurements (Beal et al., 2022; González-Cebrián et al., 2023), and in the automation of wetlab protocols in general(Vidal-Peña, 2023). Finally for the learn step (Vidal et al., 2022a; Yanez Feliu et al., 2020; Boada et al., 2019b). However, there are not many examples of automation of the test and learn steps combined. This paper aims to explore the application of automation techniques to the DBTL cycle steps Test and Learn, specifically focusing on the testing and characterization of standard bioparts. Standard bioparts, such as promoters, terminators, and coding sequences, form the building blocks of genetic circuits and are widely used in bioengineering projects. Efficiently testing and characterizing these bioparts is crucial for their reliable integration into larger genetic systems and the predictable behavior of engineered organisms (Boada et al., 2019a).

Automation has the potential to revolutionize the testing and characterization of standard bioparts by improving throughput, reliability, and reproducibility (Cummins et al., 2023). By leveraging advanced laboratory robotics, high-throughput screening methods, and data analysis algorithms, automation can significantly accelerate the iterative testing process and provide valuable insights for bioengineers. Furthermore, automation can enhance the standardization and quality control of bioparts (Buecherl and Myers, 2022), ensuring their compatibility and reliability across different projects and laboratories (Beal et al., 2020).

This paper will discuss the current challenges and limitations of manual testing and characterization methods for standard bioparts. It will then explore various automation strategies and technologies that can be employed to address these challenges. Additionally, we will examine with a case study the proposed advancements in the automation of the DBTL cycle for the testing and characterization of standard bioparts.

2. Automation of the Test step

This work is focused on the Test and Learn steps of the DBTL cycle, so the starting situation is where there exist several genetic constructs in the lab, and we need to test them and learn from these experiment to improve the models and characterize the used bioparts. First we deal with the Automation of the Test step. For this we implemented a automated protocol combining the Agilent Biotek plate reader (Figure 2) with the Opentrons OT-2 (Figure 3).

Figure 2: Agilent Biotek Cytation 3 plate reader. This plate reader allows us for incubation at 37ºC, agitation and measurement of both absorbance and fluorescence.

This workflow implements the protocol for the

*Corresponding Author: vignoni@isa.upv.es
Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

normalization of initial concentrations of 7 bacterial culture samples for a 16 hour incubation/measurement experiment. Specifically, the protocol is divided into two parts, the first is a 1:4 dilution of the culture in the culture medium Minimal Media M9 salts plus 20 % of glucose (blank), this dilution will be measured and the optical density (OD) of the 7 samples will be taken so that in the second part of the protocol all the culture are normalized to an OD of 0.1. Later on, this dilutions are distributed into the measurement 96 well plate. To carry out this whole process it is necessary to have a specific configuration of the OT2 and OT-2 deck (Figure 4), to make use of the Agilent Biotek Cytation measurement program and to use the Jupyter notebook linked to the OT2 to run the protocol.

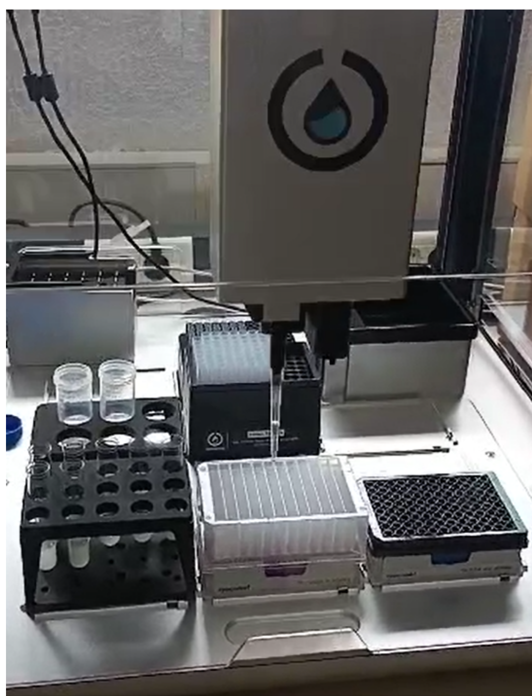

Figure 3: Opentrons OT-2 liquid-handling robot. The OT-2 already prepared to execute the protocol with all the labware in the right position. Pipettes used are the Multichannel P300 in the right mount, and the Single Channel P1000 in the left mount.

2.1. Procedure

1. Set up the OT2 as shown in the Figure 3. Pipettes used are the Multichannel P300 in the right mount, and the Single Channel P1000 in the left mount.
2. Set up the deck as follows (also shown in Figure 4):
   SLOT 1: Opentrons 15 tube rack with falcon 14 ml round
   SLOT 2: Porvair 96 deep well plate 2ml conical
   SLOT 3: Genier bio coldblock 96 well plate 400ul
   SLOT 4: Opentrons 6 tube rackwith falcon 50 ml conical
   SLOT 8: Opentrons 96 tip rack 1000ul
   SLOT 11: Opentrons 96 tip rack 300ul

3. In the third step the different liquids that the OT2 will handle will be added. The order in which the samples are placed is of great importance for the correct functioning of the protocol, in this case we have chosen the sequence A1, A2, A3, A4, A5, B1 and B2 (Figure 4).

4. The fourth step is the calibration and adjustment of the offsets. Calibration is performed by loading a python script with a small protocol in the OT2 APP with the necessary labware previously loaded. Once this protocol is selected the user has only to run the Labware Position Check to start the manual calibration. Once the offsets are obtained, do not close the window with the data, because they have to be copied to each of the protocols to be executed in Jupyter (Figure 5).

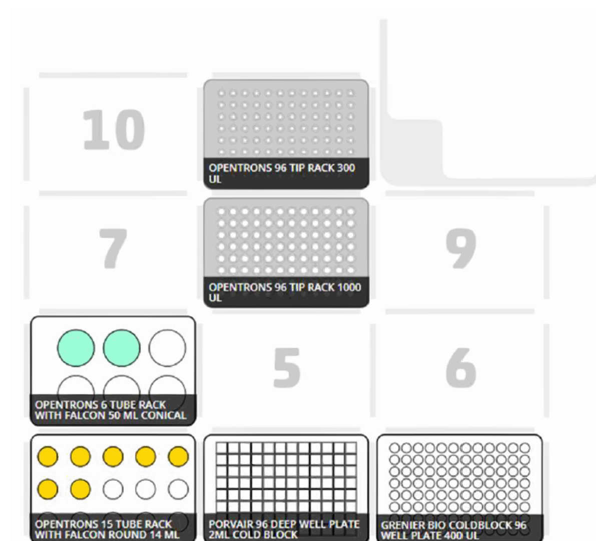

Figure 4: Deck of the Opentrons OT-2 liquid-handling robot prepared for the protocol.

5. Execution of the first part of the protocol from Jupyter. For this it is necessary to change the variable action to 2 of the protocol, to start executing the code and then launch the code by pressing Ctrl+Enter.

6. Once the whole process has been completed, the SLOT 3 plate will carry the samples in the 12 column, and it must be transfered to the plate reader to perform the Absorbance measurement. The OD measurements obtained should be saved into the provided Template Spreadsheet as a comma separated values files. This template calculates the desired dilutions and volumes to be used later by the OT-2.

7. Upload the Template csv files with the measured values to the OT-2 Jupyter environment.

8. Execution of the second part of the protocol as done in the fourth step by putting back the 96 well plate to SLOT 3 without the lid, and adding the offsets and changing the action variable to 2 of the protocol to start executing the code and then launch the code by pressing Ctrl+Etner.

Figure 5: Using a simple protocol with the OT-2 APP allows us to get the calibration values for each SLOT.

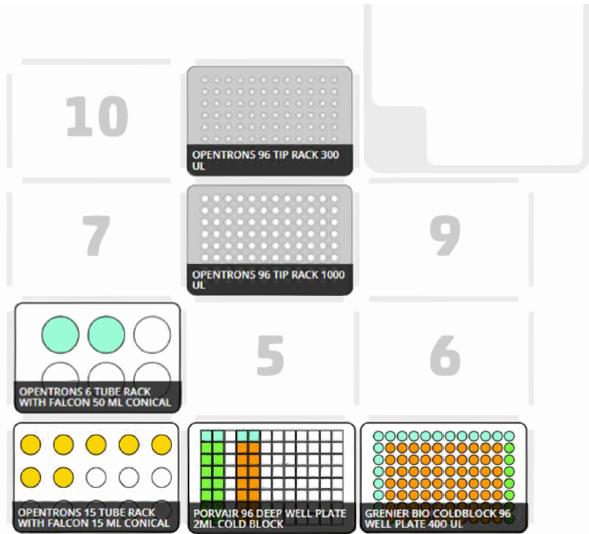

Figure 6: Deck of the Opentrons OT-2 liquid-handling robot after the completion of the two protocols. The sample 96 well-plate is prepared for growth/measurement experiment.

The use of this protocol establishes the systematic preparation of the 96 well-plate for the growth/measurement experiment, by ensuring a constant initial concentration of cultures across the plate.

Once the experiment is done, we have a big dataset of absorbance and fluorescent measurements from the selected sample.

3. Using data in the Learn step

Using the previously explained workflow for the Test step of the DBTL cycle, we obtain a dataset of measurements for the two genetic circuits shown in Figure 7.

We have 10 technical replicates in 4 experiments performed on different days for each one of the devices. While incubating at 37°C and 230 rpm in a high-speed double orbital shaking, the absorbance was recorded at 600nm, and the fluorescence was measured at 530nm with an excitation of 488nm for 16 hours. The calibration of both absorbance and fluorescence was done using our calibration protocol (Beal et al., 2022; González-Cebrián et al., 2023). In Figure 8, the replicate-averaged data show: the number of bacteria (Particles), the number of molecules (MEFL), the number of molecules per cell (MEFL/Part), and the growth rate calculated with an anti-causal filter in Matlab (filtfilt command).

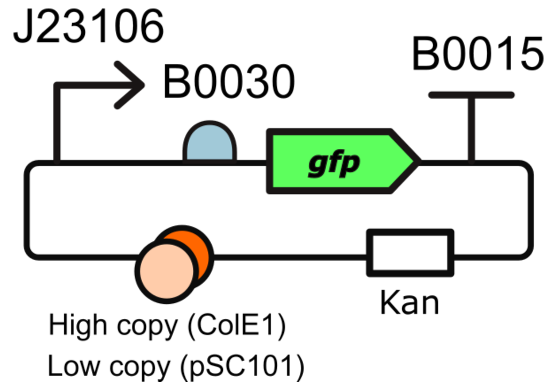

Figure 7: A constitutively expressed GFPmut3b in a high copy plasmid (ColE1 ori) and a low copy plasmid (pSC101 ori) both with the same RBS (BBa_B0030) and Promoter (BBa_J23106).

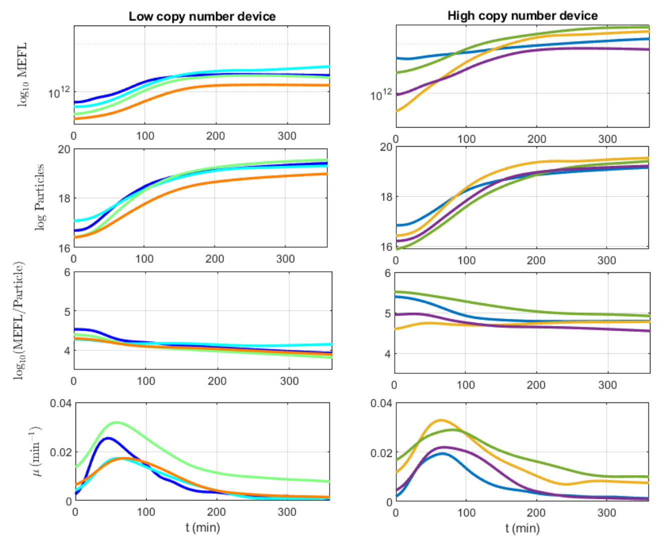

Figure 8: Experimental data for the two devices obtained using the automated workflow for the Test step. From up to bottom: Number of bacteria (Particles), Number of molecules (MEFL), number of molecules per cell (MEFL/Part), and growth rate. Left, low copy number device, right: high copy number device.

With this dataset, we perform a parameter identification by taking a simple growth-independent protein production model:

$$\frac{dN}{dt} = \mu(t)N \tag{1}$$

$$\frac{dG}{dt} = \frac{\mathrm{C_N\,k_G\,p_G}}{\mathrm{dm_G} + \mu_{\max}} - \mathrm{d_G}G - \mu_{\max}G \tag{2}$$

where $\mu(t)$ is the growth rate, N is the number of particles, $\mathrm{C_N}$ is the copy number, $\mathrm{k_G}$ is the transcription rate of the protein, $\mathrm{p_G}$ is the translation rate of the protein, $\mathrm{dm_G}$ is the dilution rate of mRNA, $\mathrm{d_G}$ is the dilution rate of the protein, μ_{max} is the maximal growth rate, and G is the protein concentration. In the first place, we assume as five the value for the low copy number construct ($C_N = 5$), then we perform the identification obtaining the parametric values shown in the first row of Table 1 using the Genetic Algorithm of MathWorks (2023). After that, we just evaluate the same cost function assuming that the only change from the low copy number device to the high copy number one is precisely the copy number, from 5 in the case of the low copy to 35 in the case of the high copy (Morgan, 2023).

In particular from the dataset, we take a $\mu_{\max} = 0.02321$ $[min^{-1}]$ for both low and high copy. By doing this, we can use the same identified parameter values from the low copy number device to predict the protein production of the high copy number device. As shown in Table 2, we obtain a similar error in prediction and optimization.

Table 1: Estimated parameters for Low copy number device.

	k_G	p_G	d_{mG}	d_G	C_N
Low copy	9.91	2.18	0.36	0.0021	5
High copy	9.91	2.18	0.36	0.0021	35

In Figure 9, the experimental data for each device is shown together with the optimized model for the low copy number device, which was used to obtain values for the parameter. In addition, the figure shows the predicted protein production for the high copy number device (using the low copy number parameters with a change in the copy number).

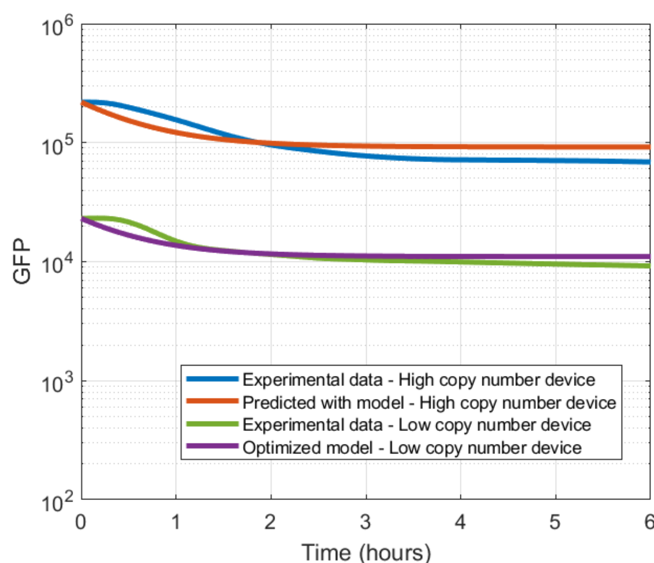

Figure 9: Optimized and predicted protein production for the low copy number device and high copy number device respectively.

Table 2: Prediction errors

	MSE	RSE
Low copy (optimization)	3.5005e+08	4.0695
High copy (prediction)	3.4719e+08	3.1733

4. Conclusions

This paper proposes an automation of the Test step of the DBTL and using the experimental data obtained for characterization of standard bioparts. With a simple model of protein production and the experimental data we can obtain parameter values that allows us to accurately predict the protein production level of another device only using the maximal growth rate of that construct. This paves the road for modeling complex devices and using these models to obtain sensible prediction of their outputs. The integration of automation into the DBTL cycle holds great promise for advancing the field of bioengineering and synthetic biology. It not only increases the efficiency and reliability of biopart testing but also enables the exploration of larger design spaces and the rapid prototyping of complex genetic systems. Ultimately, automation can facilitate the development of novel bioengineered solutions with enhanced functionality and applicability in areas such as healthcare, biomanufacturing, and environmental sustainability. By embracing automation, we can overcome the limitations of manual approaches and unlock the full potential of bioengineering to address pressing societal challenges and pave the way for innovative biological solutions.

Acknowledgments

This research was funded by the Spanish Ministry of Science and Innovation MCIN/AEI/10.13039/501100011033 grants number PID2020-117271RB-C21, TED2021-131049B-I00 and GVA grant CIAICO/ 2021/159. J.B. is a recipient of a ”Práctica de Empresa - Plan de ayudas ai2 a la I+D+i 2022”. A.P. is a recipient of a ”Práctica de Empresa - Plan de ayudas ai2 a la I+D+i 2021”. H.D.I. holds an ”Contrato predoctoral para la formación de doctores, convocatoria 2021” (PRE2021-098767) from the Agencia Estatal de Investigación. A.A.G. thanks Grant PAID-01-21 Programa de Ayudas de Investigación y Desarrollo - Universitat Politècnica de València. Y.B. thanks Grant PAID-10-21 Acceso al Sistema Español de Ciencia e Innovación-Universitat Politècnica de València. Y.B. also thanks to Secretaría de Educación Superior, Ciencia, Tecnología e Innovación of Ecuador (Scholarship Convocatoria Abierta 2011).

References

Beal, J., Farny, N. G., Haddock-Angelli, T., Selvarajah, V., Baldwin, G. S., Buckley-Taylor, R., Gershater, M., Kiga, D., Marken, J., Sanchania, V., et al., 2020. Robust estimation of bacterial cell count from optical density. Communications biology 3 (1), 512.

Beal, J., Telmer, C. A., Vignoni, A., Boada, Y., Baldwin, G. S., Hallett, L., Lee, T., Selvarajah, V., Billerbeck, S., Brown, B., et al., 2022. Multicolor plate reader fluorescence calibration. Synthetic Biology 7 (1), ysac010.

Boada, Y., Picó, J., Vignoni, A., 2021. Multi-objective optimization tuning framework for kinetic parameter selection and estimation. Methods in Molecular Biology, vol 2385. Springer US, New York, NY. DOI: 10.1007/978-1-0716-1767-0

Boada, Y., Santos-Navarro, F., Vignoni, A., Picó, J., 2022a. Optimization of the dynamic regulation in a branch-in metabolic pathway. IFAC-PapersOnLine 55 (7), 119–124.

Boada, Y., Santos-Navarro, F. N., Picó, J., Vignoni, A., 2022b. Modeling and optimization of a molecular biocontroller for the regulation of complex metabolic pathways. Frontiers in Molecular Biosciences 9.

Boada, Y., Vignoni, A., Alarcon-Ruiz, I., Andreu-Vilarroig, C., Monfort-Llorens, R., Requena, A., Picó, J., 2019a. Characterization of Gene Circuit Parts Based on Multiobjective Optimization by Using Standard Calibrated Measurements. ChemBioChem 20 (20). DOI: 10.1002/cbic.201900272

Boada, Y., Vignoni, A., Picó, J., 2019b. Multiobjective identification of a feedback synthetic gene circuit. IEEE Transactions on Control Systems Technology 28 (1), 208–223.

Bryant Jr, J. A., Kellinger, M., Longmire, C., Miller, R., Wright, R. C., 2023. Assemblytron: Flexible automation of dna assembly with opentrons ot-2 lab robots. Synthetic Biology 8 (1), ysac032.

Buecherl, L., Myers, C. J., 2022. Engineering genetic circuits: advancements in genetic design automation tools and standards for synthetic biology. Current opinion in microbiology 68, 102155.

Carbonell, P., Le Feuvre, R., Takano, E., Scrutton, N. S., 2020. In silico design and automated learning to boost next-generation smart biomanufacturing. Synthetic Biology 5 (1), ysaa020.

Cummins, B., Vrana, J., Moseley, R. C., Eramian, H., Deckard, A., Fontanarrosa, P., Bryce, D., Weston, M., Zheng, G., Nowak, J., et al., 2023. Robustness and reproducibility of simple and complex synthetic logic circuit designs using a dbtl loop. Synthetic Biology 8 (1), ysad005.

González-Cebrián, A., Borràs-Ferrís, J., Boada, Y., Vignoni, A., Ferrer, A., Picó, J., 2023. Platero: A calibration protocol for plate reader green fluorescence measurements. Frontiers in Bioengineering and Biotechnology 11.

Gurdo, N., Volke, D. C., McCloskey, D., Nikel, P. I., 2023. Automating the design-build-test-learn cycle towards next-generation bacterial cell factories. New Biotechnology 74, 1–15.
URL: https://www.sciencedirect.com/science/article/pii/S187167842300002X
DOI: https://doi.org/10.1016/j.nbt.2023.01.002

Kang, D. H., Ko, S. C., Heo, Y. B., Lee, H. J., Woo, H. M., 2022. Robomoclo: A robotics-assisted modular cloning framework for multiple gene assembly in biofoundry. ACS Synthetic Biology 11 (3), 1336–1348.

Ko, S. C., Cho, M., Lee, H. J., Woo, H. M., 2022. Biofoundry palette: Planning-assistant software for liquid handler-based experimentation and operation in the biofoundry workflow. ACS Synthetic Biology 11 (10), 3538–3543.

MathWorks, 2023. How the genetic algorithm works.
URL: https://es.mathworks.com/help/gads/how-the-genetic-algorithm-works.html

Morgan, K., 2023. Plasmids 101: Origin of replication.
URL: https://blog.addgene.org/plasmid-101-origin-of-replication

Radivojević, T., Costello, Z., Workman, K., Garcia Martin, H., 2020. A machine learning automated recommendation tool for synthetic biology. Nature communications 11 (1), 4879.

Tellechea-Luzardo, J., Otero-Muras, I., Goñi-Moreno, A., Carbonell, P., 2022. Fast biofoundries: coping with the challenges of biomanufacturing. Trends in biotechnology.

Vidal, G., Vidal-Céspedes, C., Muñoz Silva, M., Castillo-Passi, C., Yáñez Feliú, G., Federici, F., Rudge, T. J., 2022a. Accurate characterization of dynamic microbial gene expression and growth rate profiles. Synthetic Biology 7 (1), ysac020.

Vidal, G., Vidal-Céspedes, C., Rudge, T. J., 2022b. Loica: Integrating models with data for genetic network design automation. ACS Synthetic Biology 11 (5), 1984–1990.

Vidal-Peña, G., 2023. Protocol unified design unit.
URL: https://github.com/RudgeLab/PUDU

Yanez Feliu, G., Earle Gomez, B., Codoceo Berrocal, V., Munoz Silva, M., Nunez, I. N., Matute, T. F., Arce Medina, A., Vidal, G., Vidal Cespedes, C., Dahlin, J., et al., 2020. Flapjack: Data management and analysis for genetic circuit characterization. ACS Synthetic Biology 10 (1), 183–191.