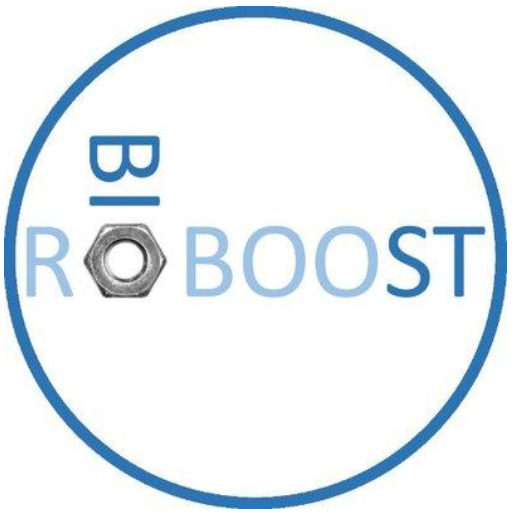

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N820699.

D 1.1

Title: Standards in synthetic biology: gaps, challenges and opportunities

Authors: **Geoff Baldwin**

DOCUMENT SHEET	
WP	1
DELIVERABLE 1.1	Standards in synthetic biology: gaps, challenges and opportunities
Due date	May 2020
Version 1	May 2020
Authors	Geoff Baldwin
Contributors	All partners involved in WP1
Dissemination level	Public
Type of Deliverable	Report

LIST OF ABREVIATIONS:

ABR	NAME
SB	Synthetic Biology
WP	Work package
RNASeq	RNA sequencing
QAcc	Accurancy metric
PoPs	Polymerases per second
RiPs	Ribosomes per second

EXECUTIVE SUMMARY

The present document describes Deliverable 1.1: *Standards in synthetic biology: gaps, challenges and opportunities*, of the H2020 project *Fostering Synthetic Biology Standardisation Through International Collaboration* (acronym BioRoboost).

The Deliverable has been generated in the context of WP1. It covers the areas of synthetic biology that the BioRoboost consortium has identified as critical areas for development, highlighting strategically important opportunities for further development that have the potential to facilitate rapid advances

TABLE OF CONTENTS

1. INTRODUCTION 4
2. DNA ASSEMBLY 4
3. METROLOGY STANDARDS 6
4. ENZYME DATA STANDARDS 9
5. DATA STANDARDS AND INTEROPERABILITY 10
6. CONCLUSION 12
7. REFERENCES 14

1. INTRODUCTION

When considering standards for synthetic biology, it is important to understand the pathway for standards development and what might be considered different 'levels' of standardization. At the lowest level, we can consider protocols and best practice guidelines while at the highest level there are certified ISO standards. It is important to recognise that for the majority of academic researchers, best practice is sufficient. Such best practice and protocol development in the form of Standard Operating Procedures (SOPs) will also enable more direct translation into industrial settings. ISO standards are uniquely important for industry where performance of critical workflows is dependent on them. The lifetime for development of ISO standards runs well beyond the lifetime of BioRoboost, so the immediate opportunity for this consortium should focus on best practices and protocols as the foundation for standards development. This review covers the areas of synthetic biology that the BioRoboost consortium has identified as critical areas for development, highlighting strategically important opportunities for further development that have the potential to facilitate rapid advances.

2. DNA ASSEMBLY

DNA assembly was identified as the area of greatest standards development and practice within synthetic biology. This applies to the construction of new DNA sequences encoding biological systems from pre-existing DNA parts, rather than the de novo chemical synthesis of DNA, which is handled by companies, each of which have their own in-house methodologies for the assembly and construction of larger DNA sequences. Methods and standards for DNA assembly have been well reviewed in the literature (1) and the key developments in the field have not progressed much since that time. We identified the greatest opportunity in this area to assess the requirements for DNA assembly in an automated workflow (Table 1).

Table1: current status of methods and standards for automated DNA assembly

Assembly method	Platform for automation	Comments	Biofoundry adoption
Golden Gate assembly and cloning standards e.g. - MoClo (2–5). - fairyTALE (6).	Microfluidics (7, 8)	Pre-digested and gel purified parts achieved 100 % accuracy (8).	Yes – Broad Foundry (9, 10), iBioFAB (6), London Biofoundry (3)
	Liquid handling robotics (2, 4, 6, 8, 11)	192 TALENs were assembled using 9 parts at a cost of $2.1 each (6).	
	Acoustic Liquid handling (2, 3, 5, 12)	2-part assemblies setup for ~$0.01 in 50 nL volumes (excluding transformation). Efficiency equivalent to manual (>10^4 CFU/mL) (12). ~3 days required for plasmid generation (2). Additional platforms used for downstream processing (5).	
Gibson assembly	Liquid handling robotics (13)	Commercially available. Accuracy drops as size increases above 1 kb.	n.d.
	Microfluidics (7, 8, 14)	Cloning efficiency high for 2 and 4-part assemblies but dropped for tiered e.g. 1 of 11, 8-part tiered assemblies (8).	
	Acoustic Liquid handling (12)	2-part assemblies setup for ~$0.06 in 500 nL volumes (excluding transformation). Efficiency equivalent to manual (~10^5 CFU/mL).	
Ligase cycling reaction	Liquid handling robotics (15)	Applied to flavonoid and alkaloid production. Several constructs not assembled successfully.	Yes – SYNBIOCHEM (15)
Iso-thermal Hierarchical DNA Construction	Microfluidics (14)	Avoids thermal cycling for easier application to microfluidics. Off-chip spin-column for downstream cloning.	n.d.
POP assembly	Microfluidics (16)	Achieved 300 nL size reactions. High substitution rate given multiple PCRs.	n.d.
BASIC DNA assembly	Liquid handling robotics (17)	Assembly method, hardware and software open-source. 88, 5-part assemblies constructed with 100 % efficiency. Initial investment of $10k with 5-part constructs at $1.5 ($ 5.5 with transformation).	Yes - London Biofoundry (18)
	Tip and acoustic-based liquid handling robotics (18)	5 – 10k constructs per run reported (19) with 88 having been demonstrated (18).	

From the development of these automated platforms it is becoming evident that the metrics required to assess the suitability of different approaches for automation at scale is different from what has been applied to date in the development of methods at a research scale. In particular the accuracy and efficiency of methods is of great concern since this greatly impacts the downstream quality control and repeats. The lack of harmonization in metrics to meaningfully report these numbers impedes the ability to compare methods and standards for adoption.

To date, QMetrics have been proposed as a basis for evaluating automated workflows (20), with QTime and QCost having been proposed as a basis for assessing improvements in automated DNA assembly processes compared to manual. However, when considering the entire workflow, it becomes apparent that these metrics are only useful if the method is completely accurate, since failure rates will incur large time and cost penalties.

We therefore propose that a new accuracy metric (QAcc) should be defined, to enable better comparability between methods. Importantly QAcc needs to encompass the accuracy of the complete experimental cycle, including successful transformation and construct validation:

$$Q_{Acc} = \text{number of correct constructs / number of constructs assembles attempted}$$

The Global Biofoundries Alliance (21) is planning an interlab study that could significantly impact this field. It is the recommendation of BioRoboost that such studies are critical for meaningful development and that there also need to be agreed metrics that can be used as a useful benchmark going forwards.

Metadata was reviewed with a view to identifying key aspects of workflows that could benefit from enhanced details. The key recommendation was that microbial was an aspect of workflows that would greatly benefit from improved metadata to assist comparison and replication of experiments. A recent study (22) provides a basis on which to define such a metadata standard.

3. METROLOGY STANDARDS

Units provide the basic level of metrology. Without units, numbers are dimensionless and hence have no anchor from which to understand the scale of change. The process of metrology is the basis by which we observe and measure phenomena, it cannot define a standard, but it enables standards: without metrology the very idea of standards is virtually meaningless, imagine defining a standard based on dimensionless numbers. Metrological advances must therefore be considered as part of the endeavour in the development of standards for synthetic biology.

Units themselves have different hierarchical levels. The highest level are absolute units. There are seven SI units. E.g. ampere or ohm that are absolute measures of a physical phenomenon. In biology, absolute units can be envisioned that describe things at the actual molecular basis of function, such as describing transcription by Polymerases per second (PoPs) or translation by Ribosomes per second (RiPs). While it would be desirable to understand and measure these, after 10 years of effort in synthetic biology it is painfully apparent that such measures remain a highly specialised research endeavour and are not practical for the everyday measurement of synthetic biology parts, devices and systems.

Relative units are often considered to be an inferior form of unit since they are not defined in absolute terms. However, they are widespread and provide an essential basis for a great deal of metrology. Length, mass temperature and time are all originally defined as relative units, yet provide a profoundly useful and practical basis for many things that concern us, as well as for the engineering of complex systems. Because of their importance, all of these units have since been redefined as physical constants (https://www.bipm.org/en/measurement-units/).

While we have been unable to define a useful metrology basis for PoPs and RiPs, expression of fluorescent proteins has provided a pseudo-measure for the processes that we wish to follow in cells. Although far from perfect, and subject to a number of significant limitations, they nonetheless present a powerful, useful and easy method for following relatively complex biological behaviours, including sophisticated gene circuits and dynamic responses. Despite their widespread adoption and use, they have for the most part remained as dimensionless terms. This creates significant problems in understanding the significance and comparability of data.

A key method for the calibration of measurements is to use standard reference materials. By creating standard curves of relative measurements against reference materials, output can readily be converted into calibrated units. This enables direct comparison of data generated by different instruments and different laboratories. The principle of using reference materials has been demonstrated by the iGEM interlab studies (23, 24), which demonstrated that both plate reader and flow cytometry data can be calibrated and report the same value of fluorescence per cell to within an acceptable margin of error.

The flow cytometry calibration is based on fluorescent beads that are supplied as NIST certified reference material and which cover a wide range of wavelengths (25). This approach is also very versatile in that it can be applied to both mammalian and microbial cells and since the measurement is by definition per cell, no correction for population is required.

For plate reader measurement of fluorescence, widely used for measuring microbial cells, there are no such certified reference materials available. The iGEM study demonstrated that a solution of sodium fluorescein provided an excellent reference material for calibration. A limitation of this study was that the reference was not certified nor commercially available. Repetition by an independent laboratory would therefore require the reference material to be remade, with associated errors. It would

also be beneficial to extend this approach to fluorescent proteins of other wavelengths in the blue and red range. In principle it should also be extended to cover any fluorescent protein with compensation approaches, such as those used in flow cytometry (26), to enable multicolour quantitation. A further challenge in this area is to extend calibration to microscopy.

There is thus an opportunity to extend plate reader calibration to other fluorescent proteins and a challenge in developing routes to enhance the uptake and acceptance of the standards, as well as in making suitable reference materials available to facilitate this. Reference materials should include reference dyes as well as genetic constructs and cell types for validation. Further interlab studies to enhance capabilities in this area are a priority.

Cell number is another critical measure in understanding cellular behaviour and is particularly pertinent to the widely used plate reader measurements, which only measure population level responses, rather than individual cells. Typically, Abs600 is used as a proxy for cell density, since the turbidity of microbial growth solutions scatter this wavelength of light. Absorbance is usually based on the physical absorbance of photons by a species and this gives a linear correlation to concentration based on the Beer-Lambert law. However, cell density is based on scattering, the fact that this light is not actually absorbed and this means that the Beer-Lambert law does not apply and the phenomenon is based on the physical geometry of the instrument.

Silica beads scatter light in a similar way to microbial cells (27) and it has been demonstrated that they can act as a as a proxy for cells in Absorbance 600 (Abs600) measurements. Silica microspheres thus provide a useful calibration reference material and it has been demonstrated by the iGEM interlab study that they provide a consistent and useful proxy for cell number. Fluorescence plate reader data normalized against fluorescein and silica beads enabled data in calibrated units of MEFL/Particle that correlated with the MEFL/Cell derived from flow cytometry (23). The demonstration of comparable data across different measurement devices and groups was a powerful demonstration of the applicability of normalized units.

There is currently a knowledge gap relating to the universality of this approach, its applicability to other microbial strains and sources of variance associated with changes in cell size during growth. However, there is both a need and an opportunity to develop certified reference materials for Abs600 calibration and to address wider adoption amongst the community.

A further measurement stream identified by the consortium is RNA quantitation since information at the RNA level would help deconvolute the transcription and translation processes. RNA sequencing (RNASeq), has the potential to provide calibrated measurements of cellular RNA (28) and is preferable to RT-PCR since it provides whole cell information yet can also still be highly quantitative. Since it is a whole cell approach, RNASeq can also provide a global view of cellular regulation when constructing complex synthetic biology systems or biosynthetic pathways. It is thus a critical workflow that will

become increasingly important, especially with the availability of low cost sequencing platforms like the Illumina iSEQ 100 and Oxford Nanopore MinION.
The development of DNA microarrays created the first technology for the quantitation of whole cell RNA. Early on in this effort the industry manufacturers ecogniz that that their results were not reproducible and could not be reconciled across different sites. This led to the External RNA Controls Consortium (29) that developed a spike in reference material to enable RNA quantitation which has led to their adoption and development of methods for validation (30). Their application to RNASeq has demonstrated the ability to facilitate comparable analysis across samples, protocols and platforms (31).

The consortium ecognizes that this is a key area for development and one in which we can learn from these previous advances and reference materials. Pipelines for the routine analysis of bacterial RNASeq data will be made available on a Galaxy platform. An interlab study has been initiated to assess and further develop the applicability of standard pipelines for bacterial RNASeq data analysis.

4. ENZYME DATA STANDARDS

For the design and realization of synthetic metabolic networks and pathways (32), it is very important to know the kinetic parameters of the respective candidate enzymes. While a great body of biochemical data on more than 40,000 enzymes has been collected and made accessible through different databases, such as BRENDA (33), the kinetic parameters of even homologous enzymes are unfortunately not directly comparable with each other, as the individual test conditions differ very often between labs (e.g. buffer composition, pH, enzyme and substrate concentrations, etc.). Therefore, the current data does not allow researchers to select ideal candidates for a given application, nor does it allow the accurate prediction of metabolic fluxes and/or behavior of a new design. Moreover, for the development of synthetic metabolic routes it is important to understand allosteric control mechanisms of individual enzymes, as well as the promiscuity of active sites, which provides the basis to engineer new-to-nature reactions in the scaffold of existing enzymes (34, 35). However, neither allosteric control, nor enzymatic promiscuity is systematically accessed in many studies, which provides a great bottleneck for the design and realization of new metabolic routes.

The lack of high-quality, comparable enzyme kinetic data prompts the consortium to support more systematic studies on the catalytic capacities of enzyme families and/or to standardize reporting procedures of enzyme kinetics, such as the STRENDA standard (36, 37) to facilitate the development of novel metabolic pathways in a standardized and rational fashion.

5. DATA STANDARDS AND INTEROPERABILITY

The Design-Build-Test-Learn cycle requires interconnectivity between each phase of the cycle. Data standards are critical to these interfaces and the interoperability of both digital and physical platforms. The utilization of data standards enables the capturing of in silico design information in a formalized manner making standards the key prerequisite for data sharing, reproducibility and reusability. This also facilitates the exploration of larger design spaces with enhanced predictability and control, thus moving the field closer to automated processes.

Cello perhaps provides the best example of predictable genetic circuit design (38), using a Verilog based code as the input, the described function is first parsed into logic and then translated into DNA sequence, based on functional DNA parts. The success of this process is based firstly on the logic parsing, but critically it is dependent on the accurate characterization of the bioparts used in the construction. While impressive, it is limited to the library of parts available and to the application of genetic circuit designs and is thus not a generalized solution for synthetic biology.

We have compiled a list of open source synthetic biology design tools developed by the community and are available within an integrated platform for the purposes of metabolic engineering and biosynthetic pathway design and implementation. These tools have been encapsulated into easy-to-share docker images and wrappers have been published to bind the tools within the SynBioCAD Galaxy web-based system. As of today, the tools enable one to produce a list of DNA parts to be cloned or synthetized to produce a molecule in a chassis organism provided an assembly protocol. Any molecule of the chemical space can be entered using a standard InChI code. Any chassis organism from the BiGG database can be used as long as a metabolic model is provided in a SBML standard format. Assembly protocols that can be used are GoldenGate, Gibson and LCR. The SynBioCAD Galaxy platform is open to the public and accessible at https://galaxy-synbiocad.org after registration, the platform allows users to write their own workflows. The development of the platform necessitated to standardize the input and output of each tool. The chosen standards are SBML for strains and pathways and SBOL for genetic constructs.

Tools in SynBioCAD with standard Inputs -> Outputs (IOs) and references:
Metabolic Reaction Rules (RetroRules) – IOs: None -> SMARTS – (39)
Retrosynthesis (RetroPath) – IOs: SMARTS, InChI, SBML -> SBML – (40)
Flux Balance Analysis (COBRApy) – IOs: SBML -> SBML – (41)
Reaction thermodynamics (eQuilibrator) – IOs: SBML -> SBML – (42)
Enzyme sequence search (Selenzyme) – IOs: SBML -> SBML – (43)
RBS calculator (PartGenie) – IOs: SBML -> SBOL – (44)
Design of Experiment (OptDoE) – IOs: SBML, SBOL -> SBOL – (45)
DNA parts calculation (DNA Weaver and LCR Genie) – IOs: SBOL -> csv

There are also a wide range of other computational tools that address synthetic biology digital workflows. These have typically been developed as stand alone tools and one issue for users is that they typically have to implemented individually and as yet there is no seamless workflow. *It is worth noting that, importantly, there are key underlying data standards that are evident and which are becoming central to synthetic biology. SBOL is the primary data standard for DNA parts and sequences and SBML is the primary data standard for modelling.*

Other tools:

Software tool to facilitate automated construction of DNA using the BASIC methodology on an Opentrons robot (DNABOT) – IOs: csv -> csv (17)
SynBioHub (46) – parts registry: SBOL
JBEI-ICE – parts registry (47)
SBOLDesigner – IOs: SBOL -> SBOL, csv (48)
iBioSim (49)
Cello (38)
Eugene (50)
Double Dutch (51)

Integration of digital and physical standards

Standardization coupled with abstraction permits the separation of in silico biological design from physical assembly, which in principle enables the uncoupling of complexity in both hierarchies (Endy 2005). Additionally, uncoupling fosters the separation and hence accelerates independent development of digital and physical domains. In practice, however, this can create problems when 're-coupling'. If biological design takes place in a purely abstract space then it can lead to a myriad of problems when it comes to creating the physical manifestation of the design, and this will likely lead to a loss of performance from the designed intention.

Currently there are only very limited examples of the interaction of digital workflows with physical workflows. Where this has been successful, as in Cello (38), it is based on the specification of specific parts and with an inherent knowledge regarding the build processes required for those parts. In principle, with this standardization in place the well-characterized bioparts and systems that are the products of one workflow, can be re-used in other workflows while retaining control and predictability, thus establishing modularity. However, this inevitably becomes predicated on the use of the same physical build standards.

Software tools for DNA assembly will likely be inherently linked to specific standards and we can now see such tools emerging with the like of DNABOT (17). Although many in the software community would like to have a computational design space that is abstracted from specific methodologies, this seems unlikely in the first wave of such tools since our ability to construct specific designs is inherently linked to their physical

standards and creating integrated and automated workflows requires software tools to drive these processes.

To better realise the objectives of the synthetic biology community in the application of robust Design-Build-Test-Learn engineering cycles there is an opportunity to interface physical standards with data standards, when combined with metrology standards, this will greatly enhance the ability for closed-loop learning. BioRoboost has an important role to play in bringing together the software and wetware communities in an effort to bridge this gap.

6. CONCLUSION

In the last 2 decades, the field of synthetic biology has made significant progress towards its goals of the rational engineering of biological systems. The increasing complexity of biologic designs now clearly demonstrates the need for standardized digital and physical workflows combined with automation. However, the development and application of standards, whilst being one of the key tenets of synthetic biology, has lagged behind. There are many reasons for this, but its lack of prioritization by funders and the over-riding priority of high-profile science and publications by researchers are key contributing factors. There are three key areas that need to be integrated: physical standards, metrology standards and data standards. Rather than developing new approaches, the community should focus on consolidating the development of key foundational standard workflows, their robustness, modularity and extensibility.

Physical standards for accurate automated DNA assembly will provide a paradigm shift in capability. This needs to be combined with well designed, curated and characterised part libraries.

Metrology standards are required to enable data comparability between experiments, sites and over time. The development of robust SOPs will enable adoption by a wider range of research labs and industry. New protocols will enable measurement of different aspects of workflows, facilitating enhanced learning of biological responses.

Data standards are key to all interfaces of the Design-Build-Test-Learn cycle. Improved integration of digital and physical workflows will provide a rapid advance in capability. Digital information needs to be interoperable with instruments for design, physical construction and data acquisition.

Addressing the standards issues above would greatly facilitate the automation and exploration of large design spaces. Current biological data streams, whilst very rich and in many cases very large, are primarily focused on the behaviour of natural biological systems and so inherently have limited information regarding the discreet engineering of biological sub-systems within those organisms. In synthetic biology, high dimensionality in genetic designs as well as biological are crucial to enhance understanding, accuracy and predictability of biological systems and hence Learn the rules of biological system design.

However, there is a risk that too much standardization too soon could remove flexibility through relying on standard processes that lack information and characteristics required for particular applications. For example, unresolved context-dependency could result in assuming that models work perfectly in different conditions, resulting in non-functional systems (Marguet et al. 2007), also suggesting that synthetic biology is not ready for end-to-end automation. Nonetheless, acknowledging limitations in standardization allows to proceed with caution towards improved standards and automation. Community-wide eorts are required to achieve this.

By applying standards to deep, multi-dimensional data sets, there will be a significant improvement in data comparability. This will gradually enable full convergence of design, build, test and active learning, and eventually allow computational methods to drive biofoundries, making standardization a key pillar for synthetic biology to deliver its promise of rational, predictable engineering of biological systems.

7. REFERENCES

1. Casini,A., Storch,M., Baldwin,G.S. and Ellis,T. (2015) Bricks and blueprints: Methods and standards for DNA assembly. *Nat. Rev. Mol. Cell Biol.*, **16**.
2. Ortiz,L., Pavan,M., McCarthy,L., Timmons,J. and Densmore,D.M. (2017) Automated Robotic Liquid Handling Assembly of Modular DNA Devices. *J. Vis. Exp.*, 10.3791/54703.
3. Rajakumar,P.D., Gowers,G.-O.O.F.F., Suckling,L., Foster,A., Ellis,T., Kitney,R.I., McClymont,D.W. and Freemont,P.S. (2019) Rapid Prototyping Platform for Saccharomyces cerevisiae Using Computer-Aided Genetic Design Enabled by Parallel Software and Workcell Platform Development. *SLAS Technol.*, **24**, 291–297.
4. Walsh,D.I., Pavan,M., Ortiz,L., Wick,S., Bobrow,J., Guido,N.J., Leinicke,S., Fu,D., Pandit,S., Qin,L., *et al.* (2019) Standardizing Automated DNA Assembly: Best Practices, Metrics, and Protocols Using Robots. *SLAS Technol. Transl. Life Sci. Innov.*, **24**, 247263031882533.
5. Mann,D.G.J., Bevan,S.A., Harvey,A.J. and Leffert-Sorenson,R.A. (2019) The use of an automated platform to assemble multigenic constructs for plant transformation. In *Methods in Molecular Biology*. Humana Press Inc., Vol. 1864, pp. 19–35.
6. Chao,R., Liang,J., Tasan,I., Si,T., Ju,L. and Zhao,H. (2017) Fully Automated One-Step Synthesis of Single-Transcript TALEN Pairs Using a Biological Foundry. *ACS Synth. Biol.*, **6**, 678–685.
7. Shih,S.C.C., Goyal,G., Kim,P.W., Koutsoubelis,N., Keasling,J.D., Adams,P.D., Hillson,N.J., Singh,A.K., C.C.,S.S., Garima,G., *et al.* (2015) A Versatile Microfluidic Device for Automating Synthetic Biology. *ACS Synth. Biol.*, **4**, 1151–1164.
8. Linshiz,G., Stawski,N., Goyal,G., Bi,C., Poust,S., Sharma,M., Mutalik,V., Keasling,J.D. and Hillson,N.J. (2014) PR-PR: Cross-platform laboratory automation system. *ACS Synth. Biol.*, **3**, 515–524.
9. Woodruff,L.B.A., Gorochowski,T.E., Roehner,N., Mikkelsen,T.S., Densmore,D., Gordon,D.B., Nicol,R. and Voigt,C.A. (2017) Registry in a tube: multiplexed pools of retrievable parts for genetic design space exploration. *Nucleic Acids Res.*, **45**, 1553–1565.
10. Casini,A., Chang,F.Y., Eluere,R., King,A.M., Young,E.M., Dudley,Q.M., Karim,A., Pratt,K., Bristol,C., Forget,A., *et al.* (2018) A Pressure Test to Make 10 Molecules in 90 Days: External Evaluation of Methods to Engineer Biology. *J. Am. Chem. Soc.*, **140**, 4302–4316.
11. Beal,J., Weiss,R., Densmore,D., Adler,A., Appleton,E., Babb,J., Bhatia,S., Davidsohn,N., Haddock,T., Loyall,J., *et al.* (2012) An end-to-End workflow for engineering of biological networks from high-level specifications. *ACS Synth. Biol.*, **1**, 317–331.
12. Kanigowska,P., Shen,Y., Zheng,Y., Rosser,S. and Cai,Y. (2016) Smart DNA Fabrication Using Sound Waves: Applying Acoustic Dispensing Technologies to Synthetic Biology. *J. Lab. Autom.*, **21**, 49–56.
13. Sgi-dna (2018) Expected Results of BioXp ™ Tiles-Gel Electrophoresis.
14. Linshiz,G., Jensen,E., Stawski,N., Bi,C., Elsbree,N., Jiao,H., Kim,J., Mathies,R., Keasling,J.D. and Hillson,N.J. (2016) End-to-end automated microfluidic platform

for synthetic biology: From design to functional analysis. *J. Biol. Eng.*, **10**, 1–15.
15. Carbonell,P., Jervis,A.J., Robinson,C.J., Yan,C., Dunstan,M., Swainston,N., Vinaixa,M., Hollywood,K.A., Currin,A., Rattray,N.J.W., *et al.* (2018) An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. *Commun. Biol.*, **1**, 1–10.
16. Yehezkel,T. Ben, Rival,A., Raz,O., Cohen,R., Marx,Z., Camara,M., Dubern,J.-F.F., Koch,B., Heeb,S., Krasnogor,N., *et al.* (2016) Synthesis and cell-free cloning of DNA libraries using programmable microfluidics. *Nucleic Acids Res.*, **44**, e35–e35.
17. Storch,M., Haines,M.C. and Baldwin,G.S. (2019) DNA-BOT: A low-cost, automated DNA assembly platform for synthetic biology. *bioRxiv*, 10.1101/832139.
18. Exley,K., Reynolds,C.R., Suckling,L., Chee,S.M., Tsipa,A., Freemont,P.S., McClymont,D. and Kitney,R.I. (2019) Utilising datasheets for the informed automated design and build of a synthetic metabolic pathway. *J. Biol. Eng.*, **13**, 8.
19. McClymont,D.W. and Freemont,P.S. (2017) With all due respect to Maholo, lab automation isn't anthropomorphic. *Nat. Biotechnol.*, **35**, 312–314.
20. Walsh,D.I., Pavan,M., Ortiz,L., Wick,S., Bobrow,J., Guido,N.J., Leinicke,S., Fu,D., Pandit,S., Qin,L., *et al.* (2019) Standardizing Automated DNA Assembly: Best Practices, Metrics, and Protocols Using Robots. *SLAS Technol.*, 10.1177/2472630318825335.
21. Hillson,N., Caddick,M., Cai,Y., Carrasco,J.A., Chang,M.W., Curach,N.C., Bell,D.J., Le Feuvre,R., Friedman,D.C., Fu,X., *et al.* (2019) Building a global alliance of biofoundries. *Nat. Commun.*, **10**.
22. Hecht,A., Filliben,J., Munro,S.A. and Salit,M. A minimum information standard for reproducing bench-scale bacterial cell growth and productivity. *Commun. Biol.*, 10.1038/s42003-018-0220-6.
23. Beal,J., Farny,N.G., Haddock-angelli,T., Selvarajah,V., Geoff,S., Buckley-taylor,R., Gershater,M., Kiga,D., Marken,J., Sanchania,V., *et al.* (2019) Robust Estimation of Bacterial Cell Count from Optical Density.
24. Beal,J., Haddock-Angelli,T., Baldwin,G., Gershater,M., Dwijayanti,A., Storch,M., de Mora,K., Lizarazo,M. and Rettberg,R. (2018) Quantification of bacterial fluorescence using independent calibrants. *PLoS One*, **13**.
25. Beal,J., Overney,C., Adler,A., Yaman,F., Tiberio,L. and Samineni,M. (2019) TASBE Flow Analytics: A Package for Calibrated Flow Cytometry Analysis. *ACS Synth. Biol.*, **8**, 1524–1529.
26. Roederer,M. (2002) Compensation in Flow Cytometry. *Curr. Protoc. Cytom.*, **22**.
27. Stevenson,K., McVey,A.F., Clark,I.B.N., Swain,P.S. and Pilizota,T. (2016) General calibration of microbial growth in microplate readers. *Sci. Rep.*, **6**.
28. Gorochowski,T.E., Chelysheva,I., Eriksen,M., Nair,P., Pedersen,S. and Ignatova,Z. Absolute quantification of translational regulation and burden using combined sequencing approaches. 10.1101/338939.
29. Baker,S.C., Bauer,S.R., Beyer,R.P., Brenton,J.D., Bromley,B., Burrill,J., Causton,H., Conley,M.P., Elespuru,R., Fero,M., *et al.* (2005) The external RNA controls consortium: A progress report. *Nat. Methods*, **2**, 731–734.
30. Pine,P.S., Munro,S.A., Parsons,J.R., McDaniel,J., Lucas,A.B., Lozach,J., Myers,T.G., Su,Q., Jacobs-Helber,S.M. and Salit,M. (2016) Evaluation of the External RNA Controls Consortium (ERCC) reference material using a modified Latin square

design. *BMC Biotechnol.*, **16**, 54.
31. Jiang,L., Schlesinger,F., Davis,C.A., Zhang,Y., Li,R., Salit,M., Gingeras,T.R. and Oliver,B. (2011) Synthetic spike-in standards for RNA-seq experiments. *Genome Res.*, **21**, 1543–1551.
32. Schwander,T., Von Borzyskowski,L.S., Burgener,S., Cortina,N.S. and Erb,T.J. (2016) A synthetic pathway for the fixation of carbon dioxide in vitro. *Science (80-.).*, **354**, 900–904.
33. Jeske,L., Placzek,S., Schomburg,I., Chang,A. and Schomburg,D. (2019) BRENDA in 2019: A European ELIXIR core data resource. *Nucleic Acids Res.*, **47**, D542–D549.
34. Erb,T.J. (2019) Back to the future: Why we need enzymology to build a synthetic metabolism of the future. *Beilstein J. Org. Chem.*, **15**, 551–557.
35. Erb,T.J., Jones,P.R. and Bar-Even,A. (2017) Synthetic metabolism: metabolic engineering meets enzyme design. *Curr. Opin. Chem. Biol.*, **37**, 56–62.
36. Swainston,N., Baici,A., Bakker,B.M., Cornish-Bowden,A., Fitzpatrick,P.F., Halling,P., Leyh,T.S., O'Donovan,C., Raushel,F.M., Reschel,U., *et al.* (2018) STRENDA DB: enabling the validation and sharing of enzyme kinetics data. *FEBS J.*, **285**, 2193–2204.
37. Gardossi,L., Poulsen,P.B., Ballesteros,A., Hult,K., Švedas,V.K., Vasić-Rački,D., Carrea,G., Magnusson,A., Schmid,A., Wohlgemuth,R., *et al.* (2010) Guidelines for reporting of biocatalytic reactions. *Trends Biotechnol.*, **28**, 171–180.
38. Nielsen,A.A.K., Der,B.S., Shin,J., Vaidyanathan,P., Paralanov,V., Strychalski,E.A., Ross,D., Densmore,D. and Voigt,C.A. (2016) Genetic circuit design automation. *Science (80-.).*, **352**.
39. Duigou,T., Du Lac,M., Carbonell,P. and Faulon,J.L. (2019) Retrorules: A database of reaction rules for engineering biology. *Nucleic Acids Res.*, 10.1093/nar/gky940.
40. Delépine,B., Duigou,T., Carbonell,P. and Faulon,J.L. (2018) RetroPath2.0: A retrosynthesis workflow for metabolic engineers. *Metab. Eng.*, **45**, 158–170.
41. Ebrahim,A., Lerman,J.A., Palsson,B.O. and Hyduke,D.R. (2013) COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Syst. Biol.*, 10.1186/1752-0509-7-74.
42. Flamholz,A., Noor,E., Bar-Even,A. and Milo,R. (2012) EQuilibrator - The biochemical thermodynamics calculator. *Nucleic Acids Res.*, 10.1093/nar/gkr874.
43. Carbonell,P., Wong,J., Swainston,N., Takano,E., Turner,N.J., Scrutton,N.S., Kell,D.B., Breitling,R. and Faulon,J.L. (2018) Selenzyme: Enzyme selection tool for pathway design. *Bioinformatics*, 10.1093/bioinformatics/bty065.
44. Swainston,N., Dunstan,M., Jervis,A.J., Robinson,C.J., Carbonell,P., Williams,A.R., Faulon,J.-L., Scrutton,N.S. and Kell,D.B. PartsGenie: an integrated tool for optimizing and sharing synthetic biology parts. 10.1093/bioinformatics/bty105.
45. Carbonell,P., Faulon,J.L. and Breitling,R. (2019) Efficient learning in metabolic pathway designs through optimal assembling. *IFAC-PapersOnLine*, **52**, 7–12.
46. McLaughlin,J.A., Myers,C.J., Zundel,Z., Mlslrll,G., Zhang,M., Ofiteru,I.D., Goñi-Moreno,A. and Wipat,A. (2018) SynBioHub: A Standards-Enabled Design Repository for Synthetic Biology. *ACS Synth. Biol.*, **7**, 682–688.
47. Ham,T.S., Dmytriv,Z., Plahar,H., Chen,J., Hillson,N.J. and Keasling,J.D. (2012) Design, implementation and practice of JBEI-ICE: An open source biological part registry platform and tools. *Nucleic Acids Res.*, 10.1093/nar/gks531.

48. Zhang,M., McLaughlin,J.A., Wipat,A. and Myers,C.J. (2017) SBOLDesigner 2: An Intuitive Tool for Structural Genetic Design. *ACS Synth. Biol.*, **6**, 1150–1160.
49. Watanabe,L., Nguyen,T., Zhang,M., Zundel,Z., Zhang,Z., Madsen,C., Roehner,N. and Myers,C.J. (2018) IBioSim 3: A Tool for Model-Based Genetic Circuit Design. *ACS Synth. Biol.*, 10.1021/acssynbio.8b00078.
50. Bilitchenko,L., Liu,A., Cheung,S., Weeding,E., Xia,B., Leguia,M., Anderson,J.C. and Densmore,D. (2011) Eugene – A Domain Specific Language for Specifying and Constraining Synthetic Biological Parts, Devices, and Systems. *PLoS One*, **6**, e18882.
51. Roehner,N., Young,E.M., Voigt,C.A., Benjamin Gordon,D. and Densmore,D. (2016) Double dutch: A tool for designing combinatorial libraries of biological systems. *ACS Synth. Biol.*, **5**, 507–517.