NUCLEIC ACID INSIGHTS

DNA
ENGINEERING AND MANUFACTURING

SPOTLIGHT

COMMENTARY

Incorporating digitalization in the conceptual design, research and development of plasmid biomanufacturing

Duarte Miguel de França Teixeira dos Prazeres, Ana Margarida Azevedo, Sofia Oliveira Duarte, Ana Rita Silva-Santos, Krist V Gernaey, Carina L Gargalo, Rosa Hassfurther, Annina Kemmer, and M Nicolas Cruz Bournazou

The importance of large-scale production of plasmid DNA (pDNA) has increased steadily over the years due to the development of a growing number of direct and indirect applications. To meet the growing demand for pDNA, significant efforts must be made towards improving its manufacturing. In particular, the digitalization of pDNA manufacturing could enable faster process optimization, support data-driven decision-making, and contribute to waste reduction and more sustainable operations. In this commentary article, we further contend that the benefits of digitalization should be captured early on at the research and development stage of the manufacturing process. To support this vision, we present a conceptual framework for incorporating digitalization into pDNA process development, discuss technological enablers, explain how digital methods could overcome traditional limitations, and delve into implementation considerations.

Nucleic Acid Insights 2025; 2(4), 109–121 · DOI: 10.18609/nuc.2025.018

PLASMIDS AND THE DIGITALIZATION OF BIOMANUFACTURING

Plasmids are pervasive across the gene and cell therapy industry of today [1,2]. As biologicals, they are used to deliver genetic information to patient target cells or as vehicles to deliver the molecular components of gene editor systems. Moreover, plasmids serve as essential raw materials for the manufacturing of engineered cell products (e.g., CAR-T cells) or of other biologicals (e.g., viral vectors and mRNA). The ability to manufacture plasmids cost-effectively on a large scale is thus critical for many

biopharmaceutical companies and research institutions [3]. In other circumstances, for example, in the production of lentiviral vectors for cell therapies, the actual challenge may be to develop GMP-compliant scale-down models capable of producing pDNA in a cost-effective manner [4]. One approach to increase efficiency, throughput and scalability, conserve resources, and minimize environmental impact in pDNA manufacturing is to embrace digitalization [3].

A fundamental principle of digitalization is the mapping of the physical space in a digital object via a digital twin (DT) [5]. A DT is a continuously updated *in silico* representation of a real-world system or process that acts as an identical counterpart in the digital space. An essential feature of a fully functional DT is a two-way dataflow between the physical system and its digital counterpart [6]. Ultimately, the DT generates a dynamic or static profile of the process based on historical and near-real-time measurements across an array of dimensions [6]. DTs are valuable for system simulation, integration, testing, monitoring, maintenance and even training, and are an essential building block of model-based systems engineering. Furthermore, in conjunction with mathematical modeling, DTs are likely essential for the successful implementation of continuous biomanufacturing, as they enable real-time process control, predictive decision-making, and rapid optimization [7,8].

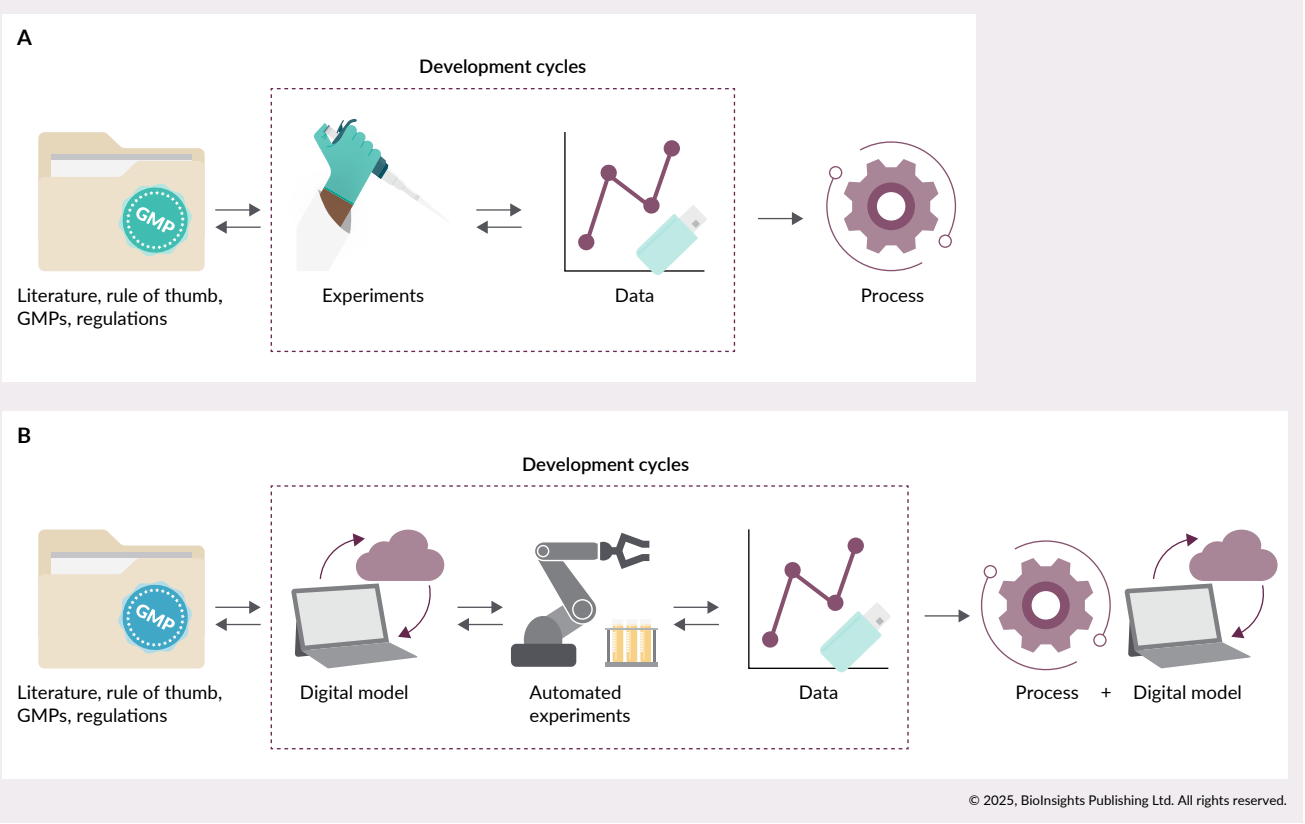
The creation of a DT of a biomanufacturing process has been advocated as one of the most compelling benefits of digitalization [6–9]. While as a first approach this will involve the digitalization of well-established manufacturing processes that are already in routine operation, several authors argue that the benefits of biomanufacturing should be captured early on at the research and development stage [10,11]. The development of a process compliant with Industry 4.0, which is characterized by the integration of digital technologies—such as the Internet of Things (IoT), artificial intelligence, data analytics, and automation—into manufacturing systems, inherently also demands the development of its DT [12]. The research question underlying this approach is therefore 'How to develop and incorporate a digitalization framework in the conceptual design, research and development of (pDNA) biomanufacturing processes?'. Such a framework is currently lacking, as most efforts are focused on digitizing established biomanufacturing processes.

THE LIMITATIONS OF TRADITIONAL PROCESS DEVELOPMENT

A real shift to a biomanufacturing scenario, where a physical process and its digital counterpart communicate, interact two ways, and operate in synchrony without interruption, requires digitalization to be embedded early in the biomanufacturing research and development stage [10]. This entails replacing the traditional process development pipeline, which follows a linear, step-by-step methodology known to be time consuming and laborious [10,11,13], with a digitally centered process development approach **(Figure 1)**. A process draft is usually designed based on the available literature, in-house experience, rules of thumb and GMPs. Key information to bear in mind pertains to final product specifications (e.g., pDNA topology, biological potency, impurity limits), some of which are established with guidance from regulations [14]. Examples of process-related impurities in pDNA manufacturing include host cell components (proteins, genomic DNA, RNA, endotoxins), residual reagents (solvents, salts, enzymes), and leachables from equipment, resins or filters. The final specifications will differ depending on the final application of the target pDNA [14]. For example, more stringent quality requirements regarding impurities will be in place if the pDNA is to be used in therapeutic applications, as opposed to cases

FIGURE 1

Replacing the (A) traditional approach to biomanufacturing process development by a (B) digitally centered process development.



where it serves as a raw material for the manufacture of a viral vector. Today, several pDNA manufacturing platforms have been developed, which can be readily adopted for the production of various pDNA molecules. Nevertheless, the introduction of new methodologies or process modifications targeted at generating more efficient processes will still require several process development cycles, relying heavily on human operators performing lab-based experimentation at small scale (typically 100 mL–10 L cell culture). The disadvantages of this approach are well recognized and include:

- Time and cost inefficiencies
- Limited process understanding and data utilization
- Poor scalability
- Human error and variability
- Regulatory compliance challenges

A digitally centered process development approach, paired with automated experimentation, could contribute to delivering new methodologies that generate more efficient processes, ultimately mitigating some of these limitations [10,11].

ENVISIONING A DIGITALLY CENTERED PDNA PROCESS DEVELOPMENT

A digitally centered approach to process development relies heavily on

incorporating digitalization concepts and computational tools at the early stages of process conceptualization, design, and development. The technological enablers of this approach include mathematical modelling (mechanistic, hybrid, and data-driven models), computational fluid dynamics, machine learning and AI, generative AI, automation and smart sensors, high throughput (HT) experimentation, workflow management systems, and edge and cloud computing **(Table 1)**. This approach offers several benefits, including accelerating development, reducing consumables by avoiding uninformative experiments, requiring fewer experiments, lowering error rates, and enhancing process understanding. Furthermore, by the end of the process development stage, digital models will be readily available to support technology transfer, process scale-up, and subsequently routine operation and control.

Here we present our view on how a digitally centered alternative can be utilized to

TABLE 1

Scientific and technological tools for digitally centered plasmid DNA process development.

Enabler	Description
Mathematical modeling	Mathematical process models of diverse nature (mechanistic, surrogate, data-driven, hybrid) are set up to provide information about key properties, variables and performance parameters/indicators (e.g., yields) of the different sub-processes (e.g., cell culture, unit operations), the interactions between process parameters, and product quality attributes (e.g., purity); these models are a key component of a DT, providing deep insights into the current state of the process through simulation
CFD	Software tools for performing CFD dynamics simulations can play a crucial role in bioprocess scale-up by enabling the simulation and analysis of fluid flows within bioreactors; this facilitates the optimization of mixing, mass transfer, and overall reactor design, which are essential for efficient scale-up
Machine learning and AI	AI and machine learning contribute to smart automation and analytics through the identification of optimal process parameters, automation of complex tasks, prediction of potential issues; enabling the shift to predictive rather than reactive process control
Generative AI (large-language models)	If trained on large and adequate datasets of bioprocess parameters, LLMs can suggest improvements to increase efficiency and product quality, or assist in designing more effective experiments, potentially reducing the number of iterations required in bioprocess development
Automation and smart sensors	Bioprocessing workflows can be optimized, monitored and controlled in real-time by integrating advanced technologies such as sensors and IoT devices for data acquisition, and AI and machine learning for predictive modeling and decision-making; automated systems can handle tasks such as sampling, analysis, and equipment maintenance and contribute to enhance process efficiency, improve product quality, and reduce variability by minimizing human intervention
HT experimentation	HT experimentation using robotic platforms enable rapid, parallel execution of numerous experiments, significantly accelerating process optimization and development; these systems can dispense reagents, mix solutions, and transfer samples, minimizing human error and increasing experimental throughput; it is thus possible to explore a broader range of parameters and conditions simultaneously, leading to faster identification of optimal production conditions; if integrated with advanced data analytics and computational modeling, HT experimentation can enhance decision-making capabilities and reduce development timelines
Workflow management systems (WMS)	By implementing WMS, processes can become fully documented, traceable and reproducible, allowing for reuse of the generated data; WMS enhance interoperability, thus enabling better collaboration between scientists; they allow for the seamless choreographing of tasks, ensuring that complex workflows are executed efficiently and in the correct sequence; additionally, WMS facilitate structured storage for data and metadata, preserving essential context for future analyses; built-in error detection mechanisms help identify issues early, triggering automated error handling procedures to maintain workflow reliability and data integrity
Edge and cloud computing	Edge computing enables real time data processing and control of the biomanufacturing facilities, empowering quick adjustments; cloud computing provides scalable storage, big data analytics, and collaborative platforms for long term data analysis, process optimization and predictive modelling

CFD: computational fluid dynamics. HT: high throughput. WMS: Workflow Management Systems.

aid, guide and accelerate the development and establishment of a pDNA manufacturing process. The overall goal is to develop an integrated model toolset that examines the entire biomanufacturing process, providing clarity on bottlenecks, highlighting optimization opportunities, and ultimately enhancing superior product quality and efficiency in laboratory operations. Specifically, we propose an approach that involves synergies between:

- Experimentation
- Digitalization
- HT model-assisted experimentation activities **(Figure 2)**

These intertwined collaborative research activities should cover the upstream and downstream processing stages of plasmid manufacturing.

Experimental setting up of a benchmark process

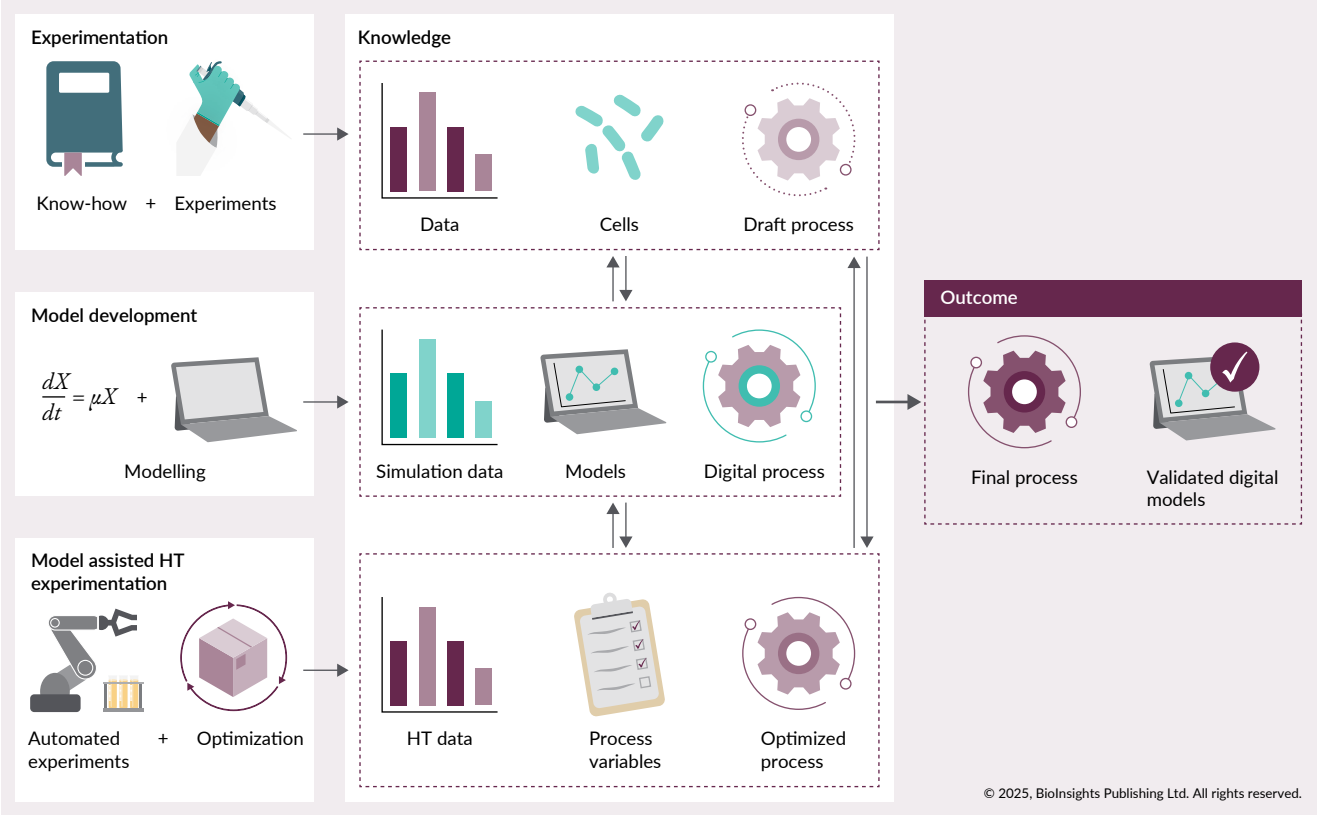
An experimental benchmark process is initially defined based on available knowledge and rules of thumb [15]. This heuristic approach involves the selection of a strain of the producer *Escherichia coli* with genotypes suitable for pDNA amplification, the preparation of banks of cells transformed with the target pDNA, and the set-up of key analytics (e.g., gel electrophoresis, HPLC, ELISA). Then, a working pDNA manufacturing process should be drafted and established at lab scale. This entails cultivating cells to amplify pDNA and then setting up a downstream processing train of operations to recover, isolate, and purify the pDNA. The goal is to quickly obtain initial datasets (e.g., time series data describing microbial cell culture and pDNA amplification, recovery yields of unit operations, etc.) that can be used to jump start and advance model development and guide HT experimentation for process optimization.

Modelling of upstream and downstream processing

Mathematical models are developed to represent, analyze, and predict the complex system surrounding pDNA manufacturing. This calls for selecting an appropriate model structure that aligns with the biological and physical nuances of the different operations in the manufacturing train. The models should be able to describe and predict the dynamics of cell growth and pDNA amplification. This requires the establishment of time course relations between variables such as the concentration of key nutrients (e.g., carbon source), biomass concentration, and pDNA titers [16,17]. Stoichiometric models of *E. coli* metabolism can also be useful in this context [18,19]. Models used to describe the isolation and purification of pDNA from the *E. coli* cells should predict the performance metrics of various operations (e.g., tangential flow filtration, precipitation, chromatography), especially in terms of yield and purification efficiency. Draft models for a particular operation are first tested using the corresponding initial data sets. Simulation results are then used to guide the design of additional experiments, such as model-based design of experiments [20–22]. The new sets of experimental data are further used to refine and validate the models. These experimental/modelling development cycles should be repeated until a satisfactory model is obtained. An illustration of this approach is provided by Muller *et al.* in the context of rAAV production [23]. Starting with shaker flask data, satisfactory process models were obtained after two to three iterative cycles combining high-throughput (HT) runs in a fully automated microbioreactor system with hybrid model refinement. Benchmarking this approach against a statistical Design

FIGURE 2

Synergies between (A) experimental setting up of a benchmark process, (B) digitalization, and (C) high-throughput model-assisted experimentation activities.



of Experiments method showed that the model-based experimental design consistently produced higher rAAV titers with fewer total experiments. Rigorous validation should be made by comparing model outputs against separate experimental datasets. The outcome is a set of robust models that reflect experimental observations, provide insights into the process and support subsequent process optimization.

High-throughput model-assisted experimentation activities

Optimization of pDNA manufacturing can then be performed by resorting to automated HT screening platforms [24–26], guided by the predictive models developed. Such platforms, which are being introduced into modern process development labs, can be used to screen process conditions and operating variables that maximize the performance of the unit operation being tested [27–29]. Experimental workflows can be integrated and automated in a laboratory environment by resorting to heterogeneous devices, including liquid handling stations, parallel cultivation systems, and mobile robots [30]. For example, advanced liquid handling stations with embedded parallel mini bioreactors can be used to run up to 48 parallel cell culture/pDNA amplification dynamic experiments in a process-wide design and optimization scheme [31,32]. The integration of a workflow management system [33] ensures the flexible yet reliable handling of complex HT experiments and FAIR data storage—findable, accessible,

interoperable, and reusable (FAIR) [34]. Here, model-based tools can enhance information gain and process robustness by enabling, for example, real-time process monitoring, the selection of the most relevant sampling times, and the optimization of process control. Many downstream processing unit operations used in pDNA manufacturing such as precipitation [35,36] and chromatography [37,38], can also be optimized using HT platforms [28,29]. One important aspect to mention is that HT experimentation critically depends on HT analytics to rapidly evaluate multiple conditions [39]. For example, an evaluation of 48 parallel cell culture/pDNA amplification experiments run on a mini bioreactor platform would undoubtedly involve assessing pDNA titers and topology. Since most scale-down reactor systems only incorporate DO and pH measurements, this would require collecting samples, performing miniprep isolation, and running agarose or capillary electrophoresis analysis in parallel, which is not trivial to implement [40]. Further challenges include the small volumes of scale-down reactors, which restrict the sampling frequency and volume, as well as the large number of samples generated. Reality shows that, unfortunately, analytical capacity often lags behind experimental throughput, creating a significant bottleneck [39]. This mismatch between experimental throughput and analytical capacity can slow down decision-making and delay process optimization, particularly in complex biological systems. Notwithstanding the analytical challenge, the large amounts of heterogeneous experimental data generated by HT platforms contain valuable information that can be explored using a wide variety of machine learning (ML) approaches [41–43]. Examples of data-driven methods that may be useful for optimization purposes include artificial neural networks [44], Bayesian optimization [45–47], deep reinforcement learning [48], and others [49].

Key benefits of HT experimentation include accelerated development, and the ability to perform a higher number of experiments while keeping the number of needed consumables low due to the smaller volumes. Applying model-based methods to design experiments with optimal information gain ensures that only the minimum number of experiments is performed [50,51]. The new data generated can be used to refine and validate the models that have been developed. The goal of these activities is to determine the optimal conditions for pDNA manufacturing and to develop a reliable digital model of the process. On the other hand, one should be aware that miniaturized systems may not replicate large-scale pDNA manufacturing (e.g., bioreactor dynamics and substrate heterogeneities/gradients [52]), and that analytical and data handling limitations can hinder the translation of results. Furthermore, the complexity of integrating automated platforms and the resources required to ensure regulatory compliance cannot be overstated.

Model integration and process validation

Ideally, models describing both upstream and downstream processing sections should be merged into a singular, unified model. This integration is still perceived as a bottleneck, often because upstream and downstream process models have focused on describing different sets of variables. Once integration is achieved, the consolidated model should be rigorously validated against lab-scale datasets (e.g., at the 1–2 L lab scale), ensuring it reflects real system dynamics, and that it is robust and reliable [53].

TRANSLATION INSIGHT

Embracing digitalization concepts and tools at the early stages of process conceptualization, design, and development

can accelerate development, reduce consumables and error rates, increase the number of informative experiments, and ultimately improve process understanding. In the context of plasmid manufacturing, this digitally centered approach to process development requires synergies and interconnection of:

- Experimentation
- Digitalization
- HT model-assisted experimentation activities

However, the field is in its infancy, with several areas requiring further study or pilot testing.

For once, many of the digitalization tools at our disposal **(Table 1)** are still underexplored in the context of process development. For example, there is clearly room for the development of LLMs tailored to the conceptual development of processes for the biomanufacturing of a particular class of bioproducts (e.g. nucleic acids, pDNA, mRNA), leveraging existing literature data and pre-existing knowledge (e.g., company data, expertise). Such dedicated LLMs could be invaluable, for example, in the initial drafting of a manufacturing process. The use of CFD in the context of process scale-up can also be considered sub-optimal due to its high computational cost, reliance on simplifications that may not fully capture complex interactions, and challenges in accurately predicting scale-dependent phenomena (e.g. turbulence, mixing, and heat transfer). Another important area that requires investment is the development of more advanced and refined mathematical models capable of accurately representing complex biological systems, for example, microbial cell culture and pDNA amplification. The importance of mathematical models in conjunction with the adoption of digitalization will be especially relevant in the context of continuous manufacturing, which is an industry trend likely to change the way plasmids are manufactured in the future [54–56].

Additionally, the full technical integration of the digitalization tools available **(Table 1)** in the context of process development is still a bottleneck. Clearly, we need to improve our ability to manage the loop of hypothesis formulation, model-based experimental design, high-throughput experimentation, data evaluation, model adaptation, conclusion, and new hypothesis generation, which still requires considerable human intervention. Although we are far from creating a 'Robot Process Development Scientist' designed to autonomously automate process development, akin to the Robot Scientist discussed by King *et al.* [57], the potential for digitalization to contribute to the generation of process knowledge is huge. The necessity to upgrade technological infrastructure for real-time data integration in process development laboratories is also imperative. Examples include the integration of HT experimentation and advanced analytics capabilities, the implementation of integrated Laboratory Information Management Systems (LIMS) or Electronic Lab Notebooks (ELN) [58], the replacement of legacy laboratory instruments with digitally enabled, IoT-compatible sensors and Process Analytical Technology tools (e.g., Raman, NIR, FTIR, and *in situ* microscopy) [59], and the installation of systems to ensure data integrity, traceability, and regulatory compliance in digital environments [60].

The implementation of digitalization in biomanufacturing—both in process development and operation—further requires a fundamental shift in how data are acquired and managed, aligning with the FAIR principles to ensure seamless integration, traceability, and utility across digital systems [61,62]. For example, this requires transforming heterogeneous data

formats (e.g., PDFs, Excel sheets) into structured, machine-readable formats (e.g., XML, JSON) to enable real-time synchronization between physical systems and their digital counterparts. Furthermore, the thorough tracking and recording of all tasks performed throughout experimentation at both experimental and computational levels is critical to ensure data reproducibility [33]. Another important aspect of digitalization is data safety, also known as cybersecurity, which involves managing data in a responsible manner to minimize the risk of a data breach. However, users are often not sufficiently aware of such safety aspects [63].

One significant challenge in embracing digitally centered process development is resistance to change among stakeholders. This can be addressed by demonstrating clear return on investment, ensuring data security, and fostering cross-disciplinary collaboration to build trust in digital innovations. This resistance may be exacerbated further by the lack of user knowledge—many potential users simply do not know how to use digital tools effectively or where to begin—as well as by the lack of tools specifically tailored for bioengineering. Clearly, a skilled workforce with competencies that differ from those of the past must be trained to understand the importance and value of digitalization tools, to utilize the new methodologies and associated devices in the laboratory, and to handle complex data outputs. This requires universities and research institutes to develop world-class educational programs in digital biomanufacturing, which are currently not widely available.

Although quantitative data on the digitalization of pDNA manufacturing is still scarce, it is reasonable to anticipate benefits comparable to those reported in other biomanufacturing domains where AI and advanced analytics have been integrated—such as improvements of throughput upstream (15–30%) and downstream (up to 60%) and significant improvements in resource efficiency and process robustness [64]. The digital shift in pDNA production is thus expected to enhance efficiency, sustainability, and decision-making in a similar manner.

Moving forward, academia can play a crucial role in exploring innovative digitalization approaches for early-stage biomanufacturing research, while industry should focus on pilot-testing digital tools in process development to assess their practical applications. Policymakers, on the other hand, must work to develop clear guidelines and regulatory frameworks that support the adoption of digitalization in biomanufacturing, ensuring both compliance and technological advancement. In conclusion, incorporating digitalization into manufacturing development is a strategic move towards efficiency and sustainability; however, its full potential depends on further research, industry validation, and supportive regulatory frameworks to ensure seamless integration and long-term impact.

REFERENCES

1. Ohlson J. Plasmid manufacture is the bottleneck of the genetic medicine revolution. *Drug Discov. Today* 2020; 25(11), 1891–1893.
2. Schmeer M, Buchholz T, Schleef M. Plasmid DNA manufacturing for indirect and direct clinical applications. *Hum. Gene Ther*. 2017; 28(10), 856–861.
3. Prazeres DMF. The supporting role of plasmids in gene and cell therapy. *Cell Gene Therapy & Insights* 2023; 9(5), 755–762.
4. Bakker NAM, De Boer R, Marie C, *et al*. Small-scale GMP production of plasmid DNA using a simplified and fully disposable production method. *J. Biotechnol*. 2019; 306, 100007.

5. Narayanan H, Luna MF, Von Stosch M, *et al*. Bioprocessing in the digital age: the role of process models. *Biotechnol. J*. 2020; 15(1), 1900172.
6. Gargalo CL, De Las Heras SC, Jones MN, *et al*. Towards the Development of Digital Twins for the Bio-manufacturing Industry. In: *Digital Twins* (Editors: Herwig C, Pörtner R, Möller J). 2020; Springer International Publishing, 1–34.
7. Udugama IA, Lopez PC, Gargalo CL, Li X, Bayer C, Gernaey KV. Digital twin in biomanufacturing: challenges and opportunities towards its implementation. *Syst. Microbiol. Biomanuf*. 2021; 1(3), 257–274.
8. Tiwari A, Masampally VS, Agarwal A, Rathore AS. Digital twin of a continuous chromatography process for mAb purification: design and model-based control. *Biotechnol. Bioeng*. 2023; 120(3), 748–766.
9. Zobel-Roos S, Schmidt A, Uhlenbrock L, Ditz R, Köster D, Strube J. Digital Twins in Biomanufacturing. In: *Digital Twins* (Editors: Herwig C, Pörtner R, Möller J). 2020; Springer International Publishing, 181–262.
10. von Stosch M, Portela RM, Varsakelis C. A roadmap to AI-driven in silico process development: bioprocessing 4.0 in practice. *Curr. Op. Chem. Eng*. 2021; 33, 100692.
11. Isoko K, Cordiner JL, Kis Z, Moghadam PZ. Bioprocessing 4.0: a pragmatic review and future perspectives. *Digital Discov.* 2024; 3(9), 1662–1681.
12. Hermann M, Pentek T, Otto B. Design Principles for Industrie 4.0 Scenarios. In: *2016 49th Hawaii International Conference on System Sciences (HICSS)*. 2016; IEEE, 3928–3937.
13. Lye G, Hubbuch J, Schroeder T, Willimann E. Shrinking the costs of bioprocess development. *BioProcess Int*. 2009; 38(2), e3230.
14. Prazeres DMF. *Plasmid Biopharmaceuticals: Basics, Applications, and Manufacturing*. 2011; Wiley, 590.
15. Baumann P, Hubbuch J. Downstream process development strategies for effective bioprocesses: trends, progress, and combinatorial approaches. *Eng. Life Sci.* 2017; 17(11), 1142–1158.
16. Bohle K, Ross A. Plasmid DNA production for pharmaceutical use: role of specific growth rate and impact on process design. *Biotechnol. Bioeng*. 2011; 108(9), 2099–2106.
17. Lopes MB, Martins G, Calado CRC. Kinetic modeling of plasmid bioproduction in *Escherichia coli* DH5α cultures over different carbon-source compositions. *J. Biotechnol*. 2014; 186, 38–48.
18. Gotsmy M, Strobl F, Weiß F, *et al.* Sulfate limitation increases specific plasmid DNA yield and productivity in *E. coli* fed-batch processes. *Microb. Cell Fact*. 2023; 22(1), 242.
19. Xu Z, Zhu X, Mohsin A, *et al.* A machine learning-based approach for improving plasmid DNA production in *Escherichia coli* fed-batch fermentations. *Biotechnol. J.* 2024; 19(6), 2400140.
20. Franceschini G, Macchietto S. Model-based design of experiments for parameter precision: state of the art. *Chem. Eng. Sci*. 2008; 63(19), 4846–4872.
21. Chakrabarty A, Buzzard GT, Rundell AE. Model-based design of experiments for cellular processes. *WIREs Mech. Dis*. 2013; 5(2), 181–203.
22. Krämer D, Wilms T, King R. Model-based process optimization for the production of macrolactin D by *Paenibacillus polymyxa*. *Processes* 2020; 8(7), 752.
23. Müller C, Siegwart G, Heider S, *et al.* Iterative hybrid model-based optimization of rAAV production. *Biotechnol. Prog.* 2025; e70006.
24. Kemmer A, Fischer N, Wilms T, *et al.* Nonlinear state estimation as tool for online monitoring and adaptive feed in high throughput cultivations. *Biotechnol. Bioeng*. 2023; 120(11), 3261–3275.
25. Janzen NH, Striedner G, Jarmer J, Voigtmann M, Abad S, Reinisch D. Implementation of a fully automated microbial cultivation platform for strain and process screening. *Biotechnol. J.* 2019; 14(10), 1800625.
26. Velez-Suberbie ML, Betts JPJ, Walker KL, Robinson C, Zoro B, Keshavarz-Moore E. High throughput automated microbial bioreactor system used for clone selection and rapid scale-down process optimization. *Biotechnol. Prog.* 2018; 34(1), 58–68.

27. Cruz Bournazou MN, Barz T, Nickel DB, *et al.* Online optimal experimental re-design in robotic parallel fed-batch cultivation facilities. *Biotechnol. Bioeng*. 2017; 114(3), 610–619.

28. Baumann P, Hubbuch J. Downstream process development strategies for effective bioprocesses: trends, progress, and combinatorial approaches. *Eng. Life Sci.* 2017; 17(11), 1142–1158.

29. São Pedro MN, Silva TC, Patil R, Ottens M. White paper on high-throughput process development for integrated continuous biomanufacturing. *Biotechnol. Bioeng*. 2021; 118(9), 3275–3286.

30. Kaspersetz L, Englert B, Krah F, Martinez EC, Neubauer P, Cruz Bournazou MN. Management of experimental workflows in robotic cultivation platforms. *SLAS Technol.* 2024; 29(6), 100214.

31. Kemmer A, Cai L, Cruz Bournazou MN, Neubauer P. High-Throughput Expression of Inclusion Bodies on an Automated Platform. In: *Inclusion Bodies* (Editors: Kopp J, Spadiut O). 2023; Springer US, 31–47.

32. Kim JW, Krausch N, Aizpuru J, *et al.* Model predictive control and moving horizon estimation for adaptive optimal bolus feeding in high-throughput cultivation of *E. coli*. *Comput. Chem. Eng*. 2023; 172, 108158.

33. Mione FM, Kaspersetz L, Luna MF, *et al.* A workflow management system for reproducible and interoperable high-throughput self-driving experiments. *Comput. Chem. Eng*. 2024; 187, 108720.

34. Wilkinson MD, Dumontier M, Aalbersberg IJ, *et al.* The FAIR guiding principles for scientific data management and stewardship. *Sci Data* 2016; 3(1), 160018.

35. Nfor BK, Hylkema NN, Wiedhaup KR, Verhaert PDEM, Van Der Wielen LAM, Ottens M. High-throughput protein precipitation and hydrophobic interaction chromatography: salt effects and thermodynamic interrelation. *J. Chromatogr. A* 2011; 1218(49), 8958–8973.

36. Gu Q, Li Z, Coffman JL, Przybycien TM, Zydney AL. High throughput solubility and redissolution screening for antibody purification via combined PEG and zinc chloride precipitation. *Biotechnol. Prog*. 2020; 36(6), e3041. 5.

37. Coffman JL, Kramarczyk JF, Kelley BD. High-throughput screening of chromatographic separations: I. Method development and column modeling. *Biotechnol. Bioeng*. 2008; 100(4), 605–618.

38. Silva TC, Eppink M, Ottens M. Digital twin in high throughput chromatographic process development for monoclonal antibodies. *J. Chromatogr. A* 2024; 1717, 464672.

39. Konstantinidis S, Kong S, Titchener-Hooker N. Identifying analytics for high throughput bioprocess development studies. *Biotechnol. Bioeng.* 2013; 110(7), 1924–1935.

40. Fang S, Sinanan DJ, Perez MH, Cruz-Quintero RG, Jadhav SR. Development of a high-throughput scale-down model in Ambr® 250 HT for plasmid DNA fermentation processes. *Biotechnol. Prog.* 2024;40(4), e3458.

41. Duong-Trung N, Born S, Kim JW, *et al.* When bioprocess engineering meets machine learning: a survey from the perspective of automated bioprocess development. *Biochem. Eng. J*. 2023; 190, 108764.

42. Helleckes LM, Hemmerich J, Wiechert W, Von Lieres E, Grünberger A. Machine learning in bioprocess development: from promise to practice. *Trends Biotechnol*. 2023; 41(6), 817–835.

43. Albino M, Gargalo CL, Nadal-Rey G, Albæk MO, Krühne U, Gernaey KV. Hybrid modeling for on-line fermentation optimization and scale-up: a review. *Processes* 2024; 12(8), 1635.

44. Tavasoli T, Arjmand S, Ranaei Siadat SO, Shojaosadati SA, Sahebghadam Lotfi A. A robust feeding control strategy adjusted and optimized by a neural network for enhancing of alpha 1-antitrypsin production in *Pichia pastoris*. *Biochem. Eng. J.* 2019; 144, 18–27.

45. Rosa SS, Nunes D, Antunes L, Prazeres DMF, Marques MPC, Azevedo AM. Maximizing mRNA vaccine production with Bayesian optimization. *Biotechnol. Bioeng*. 2022; 119(11), 3127–3139.

46. Gisperg F, Klausser R, Elshazly M, Kopp J, Brichtová EP, Spadiut O. Bayesian optimization in bioprocess engineering—where do we stand today? *Biotechnol. Bioeng*. 2025; 122(6), 1313–1325.

47. Rosa SS, Nunes D, Grinsted J, *et al.* Exploring Bayesian methods in chromatographic development: increasing the capacity of the mRNA affinity ligand. *Sep. Pur. Technol.* 2025; 367, 132881.

48. Treloar NJ, Fedorec AJH, Ingalls B, Barnes CP. Deep reinforcement learning for the control of microbial co-cultures in bioreactors. *PLOS Comput. Biol*. 2020; 16(4), e1007783.

49. Khuat TT, Bassett R, Otte E, Grevis-James A, Gabrys B. Applications of machine learning in antibody discovery, process development, manufacturing and formulation: current trends, challenges, and opportunities. *Comput. Chem. Eng*. 2024; 182, 108585.

50. Abt V, Barz T, Cruz-Bournazou MN, *et al.* Model-based tools for optimal experiments in bioprocess engineering. *Curr. Op. Chem. Eng*. 2018; 22, 244–252.

51. Kim JW, Krausch N, Aizpuru J, *et al.* Model predictive control guided with optimal experimental design for pulse-based parallel cultivation. *IFAC-Pap. OnLine* 2022; 55(7), 934–939.

52. Anane E, García ÁC, Haby B, *et al.* A model-based framework for parallel scale-down fed-batch cultivations in mini-bioreactors for accelerated phenotyping. *Biotechnol. Bioeng.* 2019; 116(11), 2906–2918.

53. Neubauer P, Cruz N, Glauche F, Junne S, Knepper A, Raven M. Consistent development of bioprocesses from microliter cultures to the industrial scale. *Eng. Life Sci.* 2013; 13(3), 224–238.

54. Hernández-Beltrán JCR, San Millán A, Fuentes-Hernández A, Peña-Miller R. Mathematical models of plasmid population dynamics. *Front. Microbiol*. 2021; 12, 606396.

55. Narayanan H, Sponchioni M, Morbidelli M. Integration and digitalization in the manufacturing of therapeutic proteins. *Chem. Eng. Sci*. 2022; 248, 117159.

56. Rathore AS, Thakur G, Kateja N. Continuous integrated manufacturing for biopharmaceuticals: a new paradigm or an empty promise? *Biotechnol. Bioeng.* 2023; 120, 333–351.

57. King RD, Rowland J, Oliver SG, *et al.* The automation of science. *Science* 2009; 324(5923), 85–89.

58. Argento N. Institutional ELN/LIMS deployment: highly customizable ELN/LIMS platform as a cornerstone of digital transformation for life sciences research institutes. *EMBO Reports* 2020; 21(3), e49862.

59. Gerzon G, Sheng Y, Kirkitadze M. Process analytical technologies—advances in bioprocess integration and future perspectives. *J. Pharm. Biomed. Anal.* 2022; 207, 114379.

60. Ullagaddi, P. Digital transformation in the pharmaceutical industry: ensuring data integrity and regulatory compliance. *IJMB* 2024; 12, 110–120.

61. Wise J, De Barron AG, Splendiani A, *et al.* Implementation and relevance of FAIR data principles in biopharmaceutical R&D. *Drug Discov. Today* 2019; 24(4), 933–938.

62. Habich T, Beutel S. Digitalization concepts in academic bioprocess development. *Eng. Life Sci.* 2024; 24(4), 2300238.

63. Buyel JF. How digital transformation can influence workflows, teaching practices and curricula in (bio)process science and engineering—an interview series with stakeholders. *Fut. Ed. Res.* 2024; 2(3), 205–224.

64. Pelletier JP, Spiegl M, Esteve PO, Seliga J, Semishchenko K, Shah S. *Human–Machine Harmonization to Upgrade Biopharma Production*. 2025; McKinsey & Company.

AFFILIATIONS

Duarte Miguel de França Teixeira dos Prazeres, Ana Margarida Azevedo, Sofia Oliveira Duarte, and **Ana Rita Silva-Santos**, Institute for Bioengineering and Biosciences, Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

Krist V Gernaey, and Carina L Gargalo, Department of Chemical and Biochemical Engineering, Process and Systems Engineering Center (PROSYS), Technical University of Denmark, Kongens Lyngby, Denmark

Rosa Hassfurther, Annina Kemmer, and M Nicolas Cruz Bournazou, Technische Universität Berlin, Faculty III Process Sciences, Institute of Biotechnology, Berlin, Germany

AUTHORSHIP & CONFLICT OF INTEREST

Contributions: The named authors take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Acknowledgements: None.

Disclosure and potential conflicts of interest: The authors have no conflicts of interest.

Funding declaration: Dr Prazeres has received a research contract from Akron Biotech. Dr Kemmer and Dr Silva-Santos received financial support from Berliner Chancen-gleichheits Programm (BCP) as part of the graduate program DiGiTal. This commentary stems from research conducted within the framework of the project 'Dig4Bio—Digitalization of Biomanufacturing of Plasmids for the Development of Advanced Therapy Modalities', which was recently funded by the European Union under Horizon Europe research and innovation program to explore and advance knowledge in the field of digitalization (HORIZON-WIDERA-2023-ACCESS-02-0-101159993). The insights presented derive from the project's research activities and aim to foster further discussion and debate within the field.

ARTICLE & COPYRIGHT INFORMATION

Article source: Invited; externally peer reviewed.

Submitted for peer review: Mar 31, 2025.

Revised manuscript received: May 19, 2025.

Publication date: May 29, 2025.