# Accelerating cell lines selection in biopharmaceutical process development through machine learning on process and metabolomic dynamics

Gianmarco Barberia, Antonio Benedettib, Paloma Diaz-Fernandezc, Daniel C. Sévind, Johanna Vappianid, Gary Finkac, Fabrizio Bezzoa, Massimiliano Baroloa, Pierantonio Faccoa,[[1]](#footnote-1)

a *CAPE-Lab – Computer-Aided Process Engineering Laboratory, Department of Industrial Engineering, University of Padova, via Marzolo, 9 – 35131 Padova PD, Italy*

b *Process Engineering & Analytics, Product Development and Supply, GlaxoSmithKline R&D, Stevenage, UK*

c *Biopharm Process Research, Biopharm Product Development and Supply, GlaxoSmithKline R&D, Stevenage, UK*

d *Cellzome GmbH, GlaxoSmithKline R&D, Heidelberg, Germany*

# Abstract

The production of therapeutic drugs based on monoclonal antibodies has grown exponentially, in particular for the treatment of immunological and oncological diseases and, more recently, Covid-19. The monoclonal antibodies are typically produced in mammalian cell cultures. The identification of the cell lines that guarantee the desired product quality attributes, such as productivity and stability, is extremely important in the development of new antibodies. The cell selection process requires and an extensive and time-consuming experimentation, which can last several months, if not years.

In this work, we show how machine learning accelerates the selection of the best performing cell lines during bioprocess development and scale-up by exploiting the wealth of information retained in the process measurements from cell cultures and in the respective biological features, such as the metabolomic data. Metabolomics provides valuable information to increase the confidence in cell lines selection. In fact, its integration with data on cell process performance allows inferring the metabolic characteristics and the biological functions related to a desired cell behavior along the culture. Furthermore, metabolomic dynamics fused with product quality attributes can be used to identify the metabolic traits of cell lines with industrially relevant phenotypes. Specifically, the accurate prediction of the product titer time trajectory from cell metabolism information and the identification of highly productive cell lines since the early stages of the culture, provided important pieces of information to improve and accelerate the cell selection.

1. Correspondence regarding this paper should be addressed to: pierantonio.facco@unipd.it [↑](#footnote-ref-1)