MULTIVARIATE DATA ANALYSIS ENABLING IMPROVED CLONE SELECTION

Stephen Goldrick, MedImmune Ltd, Cambridge, UK/ Advanced Centre for Biochemical Engineering, University College London, UK
s.goldrick@ucl.ac.uk
Nicholas J. Bond, MedImmune Ltd., Cambridge, UK
Clare Lovelady, MedImmune Ltd., Cambridge, UK
Rahul Pradhan, MedImmune Ltd., Cambridge, UK
Richard Turner, MedImmune Ltd., Cambridge, UK
Suzanne S. Farid, Advanced Centre for Biochemical Engineering, University College London, UK

Key Words: Cell line development, multivariate data analysis, critical quality attributes, critical process parameters, scale-up.

Selecting a single cell from a heterogeneous transfection pool that will scale-up appropriately from a micro-scale system to a commercial facility is a challenging and hugely important task. This clonal cell line needs to demonstrate the desired product quality attributes and ensure manufacturability throughout the entire drug manufacturing lifecycle. This process typically requires 6 to 12 months and is a time, capital and labour intensive process. High throughput (HT) methodologies are increasingly being adopted to speed up this cell line selection protocol. However, often the large quantities of data generated in combination with the increase in availability of analytics results in a daunting multivariate data analysis problem. Typically, the cell line selection strategy focuses on quality attributes recorded at point of harvest such as final concentrations of process parameters including titre and viable cell density, level of aggregates or addition product quality attributes. Time-series data such as dissolved oxygen, pH or gas flow rates are often overlooked due to challenges with visualization and interpretation of the large number of process variables recorded. This work describes a novel method that implements advanced multivariate tools including principal component analysis (PCA) to better leverage the available data to help guide this challenging decision making process. The inclusion of additional process variables was demonstrated to enhance the selection of a high-yielding mammalian cell line through inclusion of scale-up dependent process parameters related to high oxygen demands and varying nutrient uptake rates. Furthermore, this technique was demonstrated to highlight problematic product heterogeneities of parent clones that were not identified through univariate analysis of the multiple cell lines. The inclusion of this MVDA methodology demonstrated a more efficient and better decision-making protocol compared to conventional cell line selection processes.