Database Power

The development of new cell lines can be inhibited by ineffective data management systems that are slowing down progress. Upgrading to a more sophisticated and advanced approach would optimise productivity while reducing development timelines

Eliot Randle at IDBS

Despite significant investment in new automated technologies to streamline cell line development activities, the continued over-reliance on paper and spreadsheets results in unnecessary inefficiencies within the development of novel biologics, thus inhibiting progress. The full benefit of increasing throughput can only be realised with effective management of the data generated.

Most biopharmaceutical organisations have invested in sophisticated equipment and robotics to automate many aspects of their cell line development activities, and to increase the throughput of their selection processes. However, the majority have not invested in data management platforms to support this, and often still rely on the seemingly low-tech approach of using paper records, spreadsheets and other data silos. Electronic data management systems improve productivity, compliance and insight, while also providing a platform for collaboration, ultimately reducing development timelines.

Identifying the Best Candidates

The development of a new cell line to produce a biologic drug is a multi-stage process. Often starting with molecular biological techniques and transfection, it then progresses through several screening rounds of selection and enrichment before generating a number of likely candidates for further assessment.

In the early stages of selection, it is common to only take a limited number of measurements, such as a viable cell count, productivity and activity, and then to assess the performance of a candidate cell line based solely on its specific productivity and activity. Even if robotics and automation have increased the number of candidates screened, storing the data from related experiments in several places makes it hard to identify patterns and the best candidates.

It is only once the initial rounds of screening have yielded a more manageable number of candidates that a wider range of assays and measurements are employed. This is typically due to the need to balance the amount of information collected with the desire for increased throughput and decreased running costs.

However, the ultimate performance of a cell line is defined by the complex interaction of an extended range of factors, many of which are not considered until later in the process, if at all. For example, a cell line that displays high productivity in the early selection stages may ultimately result in a product that does not fold correctly after larger scale cell culture and purification. Hence, reliance on a small initial set of measures is sub-optimal and leads to the possibility of discarding otherwise good cell lines. Efforts may be focused on a subset that have high productivity but could easily fail other criteria on scale-up. So what really constitutes good performance? Surely it needs to consider developability and manufacturability too?

Organisations rarely try to identify correlations between data from the screening and optimisation stages and the performance of a cell line in pilot-scale culture and beyond. This is sometimes a result of incorrectly perceiving the initial screening stage as a self-contained, discrete activity in the development process. More often it is due to necessity, as the barriers of gathering all the information required to perform such an assessment are difficult to overcome. The process and analytical data required for the analysis are usually spread across the organisation in various folders, files and formats. Since the process can span research, development and pilot plant groups, and even contract research companies or other third-party collaborators, the data is sometimes held at geographically separated sites, in different disconnected systems. Not only is it a matter of finding out where the relevant data resides, but the problem can be exacerbated by not knowing what specific data has been produced by scientists in other departments.

Collating the necessary information is significantly simplified by employing a single data management platform across the organisation, and ensuring all data is captured with its full context. Such an approach makes it more straightforward to find and retrieve any relevant information, both current and historical, and reveal correlations. In addition to providing the basis for a better definition of both desirable and undesirable characteristics, it also enables scientists to select which assays or tests to perform and which criteria they need to apply, resulting in a more effective selection process.

Traceability and Compliance

The creation of a regulatory document to support the use of a new cell line and detailing its development process

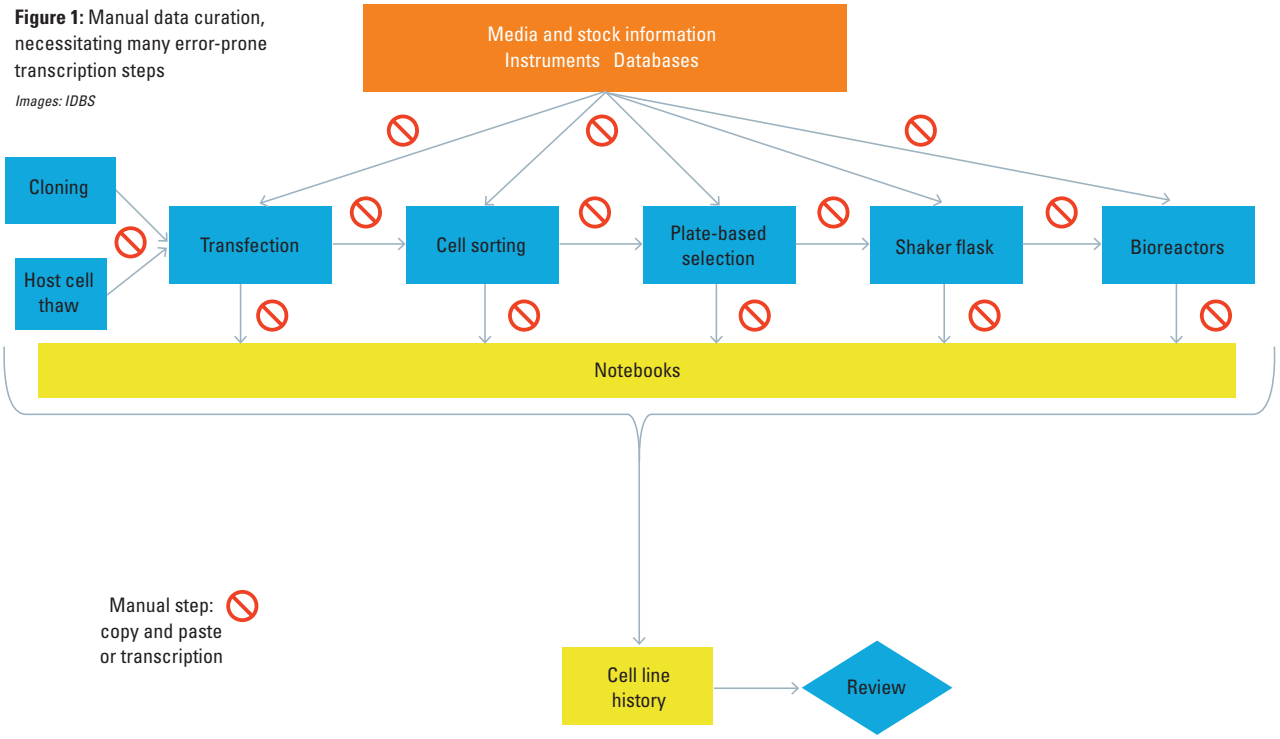

Figure 1: Manual data curation, necessitating many error-prone transcription steps

Images: IDBS

is a very time-consuming activity. Tracing the lineage of a cell line is onerous, particularly when multiple people are involved. It incorporates the error-prone manual curation of data from different people and across disparate silos of information, such as binders, paper notebooks and spreadsheets, cross-referencing information, and subsequently ensuring both the transcribed data and references are correct (see Figure 1). Identifiers (IDs) often change throughout the process, making it difficult to associate one step with another. For example, the scientist performing the amplification work may have employed a different numbering scheme to the scientist who generated the transfection pools.

As such, the generation of development histories can take days or even weeks, and is frequently a limiting factor in continuing to the next stage, particularly when it forms part of a technology transfer package. In a surprisingly high number of cases, the required information cannot be found and work has to be repeated, leading to further delays.

By reducing reliance on paper and standalone spreadsheets to store the valuable data generated during the development process, electronic data management systems ensure that everything is centralised (see Figure 2, page 26). Effective systems contextualise information and maintain relationships between all stages of the development process, making it easier to rapidly identify all experiments that led to the generation of a specific cell line of interest. When combined with comprehensive reporting capabilities, such systems generate error-free development history reports in minutes, leading to significant time savings and facilitating a better transfer of information to the next stage.

Fostering Collaboration

Cell line development involves input from different people and departments, and can even involve exchanges across organisations. When relying on paper records and standalone spreadsheets, data is disconnected, which makes it difficult not only to find key information, but also to share it with colleagues and collaborators.

Firstly, when starting a new project, reviewing prior work and understanding what worked well and what did not can be helpful. However, inaccessible, incomplete or ineffectively shared information can make it very difficult to find out how and why decisions were made during the development process. Secondly, cell line development groups often reinvent the wheel because manually trawling through previous project documentation is prohibitively time consuming. There is little opportunity to learn from past experiences, which wastes both time and money.

Electronic data management systems provide a platform to foster collaboration. They not only facilitate the sharing of data, but also support the exchange of ideas, enabling scientists to provide input and suggestions to colleagues, solicit advice and invite ideas. This provides a platform for biopharmaceutical organisations to move towards the paradigm of social science within a secure internal environment.

Selecting an Effective System

Biopharmaceutical organisations can face an overwhelming number of options to reduce reliance on paper. The following considerations suggest what is best for your organisation and

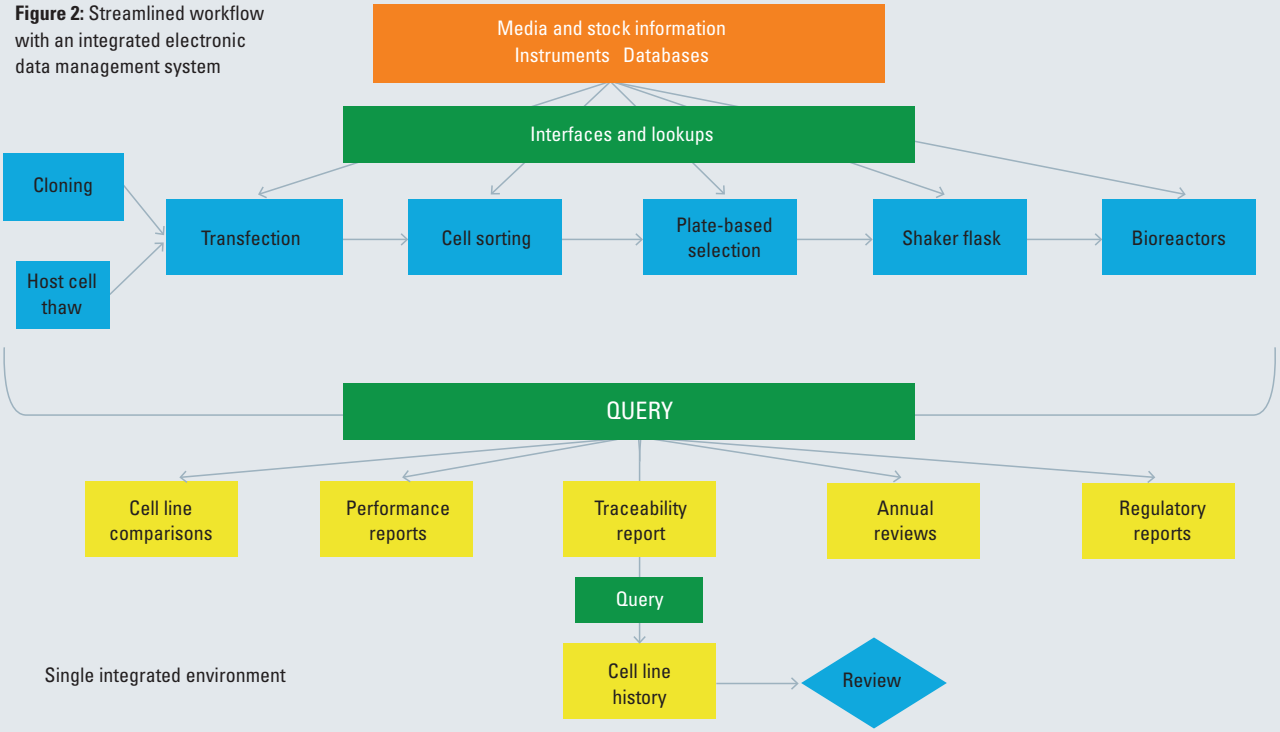

Figure 2: Streamlined workflow with an integrated electronic data management system

help you to avoid the common mistake of implementing a database system that becomes a data dump where scientists store and rarely reuse the information.

Major improvements in data quality can be achieved by systems that go beyond simply storing data and support scientific workflow. The ability to integrate with equipment and instruments enables direct capture of experimental data at the point of process execution. Additionally, to reduce transcription errors, some data management systems also provide in-built mechanisms to flag deviations and avoid mistakes – essentially providing real-time error prevention capabilities.

A data management system needs to be flexible enough to accommodate the dynamic nature of cell line development. With technology evolving rapidly, deploying new equipment and instrumentation requires modifications to the existing processes and the ability to support changes in workflow. Even organisations equipped with a platform process for cell line development need sufficient adaptability to enable variations in the number of selection rounds, changes in the order of process steps and the introduction of new technologies.

Finally, a system that supports the entire process from host cell transfection through selection in plates and lab-scale assessment outweighs the benefits of improving data management in any one aspect of cell line development. Systems-supporting cell culture and purification activities within R&D and beyond deliver valuable insight across the organisation, and form the basis for continuous improvement initiatives.

By taking this holistic view, organisations can significantly reduce the effort required to generate regulatory documentation and internal reports. Aside from realising tangible time savings, the benefits of sharing information across the organisation may be harder to measure, but providing a platform for collaboration undoubtedly helps foster a culture of learning and sharing knowledge.

The reliance on paper and disparate data files introduces unnecessary inefficiencies into the process of developing novel or optimised cell lines. The deployment of an effective electronic data management system brings wide-reaching and significant benefits, including reduced reporting effort and improved quality, traceability and insight. By extending deployment to other activities throughout the organisation, these benefits can streamline the entire process of developing biological therapeutics.

About the author

Eliot Randle is Principal Analyst at IDBS. He focuses on providing solutions for development, manufacturing and quality organisations. Eliot has spent over 10 years consulting in the bioinformatics and data management space. Prior to this, he spent five years within the biotechnology industry at British Biotech and Prolifix. Eliot completed both his BSc and PhD at the University of Manchester, and holds an MBA from Warwick Business School. **Email: info@idbs.com**