Pharmaceutical BIOPROCESSING

The application of systems biology to biomanufacturing

Abstract 'omics and systems biology have led to paradigm shifts in biology and medicine. This success has drawn the attention of the bioprocessing industry where their application is increasingly more prevalent. systems biology uses system-level high dimensional data generated via 'omics technologies to provide a holistic view of the production cell lines. We discuss how systems biology drives rational process improvement and cell engineering strategies, highlighting seminal studies in prokaryotes and mammalian cell lines that combined multi-'omics and modeling to provide insights into the behavior of production cell lines. Despite its recognized potential, there are challenges and limitations to overcome to fully implement and realize benefits heralded by systems biology for biomanufacturing: increasing titer, yield, quality, process efficiency and stability.

Larissa Benavente[‡,1], Adam Goldberg[‡,1], Henrietta Myburg[1], Joseph Chiera[1,2], Michael Zapata III[1,3] & Len van Zyl[*,1,4]

[1]ArrayXpress, Inc., 617 Hutton Street, Suite 107, Raleigh, NC 27606, USA
[2]Phytotron, North Carolina State University, Campus Box 7618, Raleigh, NC 27695-7618, USA
[3]Management, Innovation & Entrepreneurship, North Carolina State University, Campus Box 7229, Raleigh, NC 27695-7229, USA
[4]Biotechnology Group, North Carolina State University, Campus Box 7247, Raleigh, NC 27695-7247
*Author for correspondence: lenvanzyl@arrayxpress.com
[‡]Authors contributed equally

The biopharmaceutical industry is a powerhouse within the biotechnology sector, with tremendous technological and economic growth since its inception. Experts forecast continuous strong market growth, with biologics accounting for more than 50% of the top 100 prescription sales by 2018 [1]. Over the past several decades, improvements in manufacturing processes and product quality were mainly driven by increased sales and tighter regulatory requirements due to initiatives such as Quality by Design (QbD) and Process Analytical Technology (PAT). As a result, titers have jumped more than 100-fold, from subsingle digit yields (in gram per liter) to today's double-digit production levels [2,3]. From a regulatory perspective, this impressive achievement must be accompanied by a quality-driven biomanufacturing framework, founded on integrative systems and data-driven methods that contribute to process understanding and where critical process parameters are identified, monitored and controlled [4]. To keep up with ever-increasing market and regulatory demands and the competition from biosimilars, the biopharmaceutical industry is continuously challenged to increase its efficiency and make better products cheaper.

Upstream methods for cell line development and process optimization are time consuming, expensive and labor intensive. These limitations represent, perhaps, the most significant bottlenecks in bioprocessing. In addition, these upstream methods lack the mechanistic understanding of how and why process conditions, or any implemented changes, bring about the desired outcome – be it increased titers, higher product quality or process stability. This laborious effort must be repeated for every new production cell line and associated protein product. At best, it results in highly variable and unpredictable bioprocesses, both in terms of productivity as well as product quality. These process inconsistencies are commonplace and cannot support the expected growth in market demand nor the economic and regulatory challenges faced by manufacturers. The biopharmaceutical industry can greatly benefit from technological innovations that drive rapid and adaptive change, ultimately providing a competitive advantage, and allowing it to focus on improving efficiency, flexibility, convenience and quality [2,5–9].

Biologics are complex molecules with unique quality attributes that require complex production systems. Mammalian cell lines are

FUTURE SCIENCE part of fsg

Key terms

Quality by Design: A scientific, risk-based, holistic and proactive approach based on a deliberate design effort from product formation through product marketing.

Process analytical technology: Part of the QbD concept that provides tools to design, analyze and control pharmaceutical manufacturing processes through the measurement of critical process parameters (CPP) that affect critical quality attributes (CQA).

Biosimilar: A subsequent version of a biologic medical product whose active drug constituent was made by a living organism, or from a living organism by means of DNA recombinant or controlled gene expression methods.

ideally suited for this purpose due to their ability to generate complex human-like glycan profiles and other post-translational modifications that are critical for product efficacy and safety. At the same time, this inherent complexity of biological systems is also a primary contributor to process variability and inconsistency. This conflicts with the quality framework enforced by QbD and PAT, which is based on a better understanding of the biomanufacturing process. It is not possible to fully understand the process(es) without considering the biology of the different cell line(s) used and their relationship to the products they synthesize. In order to overcome this variability and inconsistency, bioprocess scientists and engineers need to have a better understanding of the cells and their intracellular processes relevant to biomanufacturing, including protein translation, post-translational modifications, folding, aggregation, trafficking and secretion. Without this knowledge, any optimizations that lead to production gains observed for one cell line are not likely transferrable to another and cannot be fully implemented across the entire production portfolio. Developing an in-depth understanding of the biology of these production cell lines is key for sustained biomanufacturing. systems biology is poised to tackle many of these challenges, particularly the areas of process optimization and cell line development.

What is systems biology?

Perhaps the simplest answer to this question is that systems biology is the application of systems theory to a biological system. As a unique field of study, systems biology grew in popularity in the latter part of the 1990s with the work of Hood and Kitano [10,11]. The application has been suggested as far back as the 1950s with the work of Mesarovic and Bertalanffy on general systems theory [12–14]. This answer, however, immediately leads to the next question, what is a system? A system is 'a set of elements together with relations between them' [15]. Of key importance to this definition is the relationship between the elements of the system; it is the interplay between the components of a system that make it functional. With no relationships, there simply is no system [15]. Hence, systems theory studies how the constituent parts of the whole system interact with one another and the subsequent attempt to generate a simplified model of how the system functions.

With this in mind, one definition for systems biology is the study and subsequent mathematical modeling of the components that constitute a biological system, their interactions and the system's resulting emergent properties. Emergent properties are outcomes from the interactions among the model's constituent parts which are not obvious from looking at any of its individual components on their own. This contrasts with the traditional reductionist approach, of taking apart the system and studying its parts in isolation in an attempt to simplify and understand it. While very successful for many years, and providing answers to focused biological questions involving a limited number of biological entities, reductionism has its limitations [16]. Conversely, instead of focusing on a single system component, or subset thereof, in systems biology a system-wide approach is taken. For example, by studying the entire transcriptome and proteome of a cell or organism, and more importantly, the interactions between and across them, one can learn how systems behave under different, changing conditions and environments, both intra- and extra-cellular. It also seeks to understand the mechanisms employed by the cell to maintain its homeostasis and to prevent, or reduce, malfunction and failure. Ultimately, by mathematical modeling and simulation, systems biology aims to identify the basic functional biological circuits that give rise to the diver sity and complexity of biological phenomena [11,16].

There are currently different views on the principles, methodologies and implementation of systems biology, varying from a more pragmatic approach to a more theoretical one ([17,18] and references therein), but this falls outside of the scope of this review. Nevertheless, we believe that each view has its own merit, not necessarily being mutually exclusive, and add value to biology as the study of life. Since this review focuses on the applications of systems biology for biomanufacturing, we follow the more pragmatic line of thought in the discussion below as we feel it more immediately, and perhaps more appropriately, addresses the challenges and needs of this particular industry.

Systems biology & biomanufacturing: how can it help?

Biomanufacturing performance is determined by the interaction of its two components. The 'bio' component comprises the host cells that actually synthesize the biomolecule, while the 'manufacturing' component comprises the physical production systems, media formulations and process parameters. Instead of viewing

the cells as mere catalysts in the synthesis of the product, systems biology recognizes the central role of the host cells in the overall biomanufacturing process. In a systems biology approach, the cells are viewed as complex and dynamic systems that are constantly interacting and responding to their environment, in this case the 'manufacturing' component. systems biology can thus provide a holistic view of biomanufacturing as a single, unified system. Ultimately, the aim is to attain three main goals sought by the industry: predictability, increased quality and higher titers.

In the following sections, we discuss examples of how a systems biology approach has been employed towards these goals. A common thread to these studies is the use of metabolic models. Notwithstanding the legacy of biochemical metabolic studies in biomanufacturing, metabolic models play a key role in systems biology. These models are the closest link to relevant cellular output phenotypes, such as growth and productivity. Moreover, metabolic models are also useful scaffolds to integrate a variety of 'omics data [2,19,20]. By incorporating these additional layers of regulatory information (e.g., transcriptional, translational, post-translational, signaling networks and allosteric mechanisms), these models are constantly being refined for improved accuracy and predictability [21–23]. We will highlight a few seminal studies in prokaryotes and then focus on mammalian cells, the primary factories for the production of biologics.

Systems metabolic engineering of industrial prokaryotes

Metabolic engineering approaches have been successfully used in bioprocesses to produce high value chemicals and products in microorganisms [19,24–26]. More recently this approach has expanded to systems metabolic engineering, such as the manipulation of entire metabolic pathways, facilitated by the development of next-generation sequencing (NGS) and other high-throughput technologies and advances in computation [27–30]. Systems metabolic engineering is intricately dependent on systems biology. By incorporating various genome-scale datasets and their respective regulatory component, systems biology allows for the creation of comprehensive predictive *in silico* models of the cells. In contrast to the traditional metabolic engineering approach of random mutagenesis and screening, these *in silico* models can then be used to guide rational engineering approaches at the cellular and metabolic levels.

In a less complex system such as *Escherichia coli*, the application of systems-based approaches can be quickly adopted due to the wealth of genomics, transcriptomics, metabolomics and fluxomics information available for this model organism. As a result, *E. coli* has one of the most complete and detailed metabolic models comprising 1366 genes and 2251 metabolic reactions and it is continually being updated [31]. Although the models are subject to continual revision, bacterial models have already reached the point in which they are used in calculating phenotypes such as growth rates, metabolic capacity and yield. For example, using known metabolic regulatory information, new data and *in silico* simulations, industrially relevant *E. coli* strains have been created that overproduced L-valine [32]. Additionally, *in silico* models were also used to reveal that increased amino acid availability would lead to increased recombinant human IL-2 production in *E. coli* [33].

The effective application of genome-wide reconstructions is now within reach, as is the incorporation of quantitative 'omics datasets into these reconstructions. Thiele *et al.* [34] modeled the *E. coli* transcription and translational network accounting for nucleotide composition, operon association and sigma factor usage. The model was sufficiently detailed to correctly simulate ribosome production, a surrogate indicator for growth and has advanced the progress toward the genotype–phenotype correlations in microbial systems. More recently, integration of transcriptional and translational datasets with genome-scale metabolic networks produced a series of improved whole-cell models. In *E. coli*, these metabolism with gene expression (ME) models have been of particular benefit in elucidating codon usage [35] and to better predict cell growth, nutrient uptake and by-product secretion [36]. This model

Key terms

Next-generation sequencing: High-throughput methods capable of sequencing DNA in a massively parallel fashion, yielding very large amounts of data compared with traditional methods (e.g., dideoxy or Sanger, sequencing). Generally encompasses a collection of methods that involve template preparation, the actual sequencing – including signal detection and processing, and data analysis. Platforms include Illumina, Pacific Biosciences (PacBio), Ion Torrent, Oxford Nanopore and others. NGS methods have been developed for several different applications, including targeted and whole genome (re-)sequencing, transcriptome profiling (RNA-seq), DNA–protein interactions (ChIP-seq), DNA methylation profiling (Methyl-seq) and many others.

Transcriptomics: The study of the transcriptome which is represented by all RNA molecules, including mRNA, rRNA, tRNA and other noncoding RNA transcribed in one cell or a population of cells.

Metabolomics: The study of the metabolome which refers to the complete set of small chemical molecules found within a biological sample.

Fluxomics: The study of metabolic reaction rates within a biological system.

Key term

In silico modeling: Modeling or simulation performed on a computer.

was recently updated to include membrane-driven processes, including subcellular compartmentalization and protein translocation [37]. Many physiological and metabolic processes that are critical to recombinant protein production either take place at, or involve, the cellular membrane system. Yet only limited aspects of these processes have been included in previous genome-scale models. Altogether, this series of articles exemplifies the evolution of genome-scale models over time, incorporating additional biological knowledge to provide further granularity into *E. coli* cellular processes.

The ultimate goal of computational modeling is to predict phenotype from genotype. To date, a recently published model of *Mycoplasma* is the closest to accomplishing this goal [29]. At one-eighth the genetic content (genes and nucleotides) of *E. coli*, *Mycoplasma genitalium* represents the most complete genome-scale reconstruction, incorporating molecular interactions into integrated mathematical representations to create a whole-cell model of the bacterium's life cycle [29]. This combined multiscale modeling was also recently employed to construct a genome-scale model of *E. coli* based on a compendium of gene expression, signal transduction and metabolism data along with their statistical associations [30]. Although these models represent the most advanced mathematical and statistical approaches for genome-scale reconstruction, they have mostly been described in a pure academic context, rather than directly applied to biomanufacturing. Within an applied context, this strategy has been employed for in silico modeling the growth-coupled production of commodity chemicals in *E. coli* via the design of optimized pathways, and opens the door to the cell-based production of these chemicals in the near future [38].

Systems biology of industrial mammalian hosts

Systems biology approaches aim to understand four key aspects of a biological system: its structure, dynamics, control and design principles [11]. However, most of the published 'omics studies using industrially relevant eukaryote host cells however fall short of this. At best, current initial efforts in mammalian systems biology represent a compilation of one or more 'omics datasets and the correlations among them and with measured phenotypes. Some studies also included a metabolic model, but regulatory aspects that ultimately control cellular and metabolic phenotypes are generally not included. Ultimately, the key aspects of a system's dynamics, control and design are left unexplored. Without looking at these, one cannot claim to reach the core benefits of systems biology.

To our knowledge there are currently no peer-reviewed reports of a predictive genome-scale mammalian model for industrial applications in biologics manufacturing. This is not surprising; the task of constructing and validating a genome-scale whole-cell model able to predict phenotype from genotype is by no means straightforward. Even for microbial species, with much simpler intracellular structures and genomes, this has only been accomplished within the past few years [29,36,37]. The inherently greater complexity of mammalian hosts poses a significant challenge for modeling. Nonetheless, progress is being made by several groups working simultaneously with more than one 'Omic technology, while also constructing and refining metabolic models of specific mammalian host cell lines. Despite their limitations, these studies represent important first steps toward a holistic understanding of host cells as production systems, and form the foundation to the application of systems biology approaches.

Multi-'Omic approaches & metabolic models in Chinese hamster ovary (CHO) cells

The Chinese hamster ovary (CHO) cell is the workhorse of the biomanufacturing industry. Much effort is being dedicated to improve current CHO genomics resources, such as building and updating a reference genome [39] and developing a co-expression database [40]. It is important to note that CHO cell lines have genetically diverged from the wild-type Chinese hamster and from one another [41] to the extent that the various CHO cell lines are considered quasi-species [42]. Recently, it has been announced that improving the assembly and annotation of the Chinese hamster as the reference genome will be a priority within the CHO community [43]. Once accomplished, this will be a major step forward for bioprocessing and will bring us closer to the CHO systems biology era.

Several peer-reviewed 'omics studies in CHO have been published and extensively reviewed elsewhere [2,44–47]. A variety of genes and proteins have been associated or correlated with production phenotypes of interest, with little overlap across the different studies. This is not surprising given the genetic diversity of the divergent cell lines, the different molecules being synthesized by each production system, and the varying process conditions (e.g., media, scale and perfusion vs batch vs fed-batch, among others). Altogether, these studies can point us in one direction. Each study attempts at understanding only a single source of biological information, whether that is DNA, RNA, proteins or metabolites. Studied in isolation any one 'omics approach will not be able to provide complete

insight into the structure, dynamic, control and design principles of biological systems. As such, if the field of biomanufacturing wants to realize the promise of systems biology, it will be important to put an integrated multi-'omics approach in practice.

Key aspects missing from CHO's System Biology that are poorly explored are: interactions across different levels of biological information; regulatory mechanisms and a mathematical modeling framework to aggregate the inherently multiscale biological knowledge. Very few studies employ more than two 'omics to explore their production system (examples provided in Table 1). Little is known about the epigenetics, noncoding RNAs, post-translational modifications, allosteric metabolic regulation and feedback and feed-forward control within and across these different layers of biological information. Modeling efforts have been mostly focused on metabolism (reviewed in [47]; examples of more recent models are provided in Table 2). Specific to biomanufacturing, mathematical models used to describe and query the production systems must be able to account for both biological and process data as input variables [48].

Due to the incredible genetic diversity across the different cell lines, a single solution will not be applicable to all production systems. A more reasonable path forward would include a genome-scale model of the Chinese hamster that can be the basis for a model for each production system. To this aim, a standardized CHO metabolic network reconstruction is currently available that includes both a genome-scale network reconstruction (GENRE) and its derived genome-scale model (GEM), following the systems biology Markup Language (SMBL) and The Minimum Information Required in Annotation of Models (MIRIAM) standards ([71]; see section: 'Challenges, limitations, and solutions' for definition of SBML and MIRIAM). Collaborative efforts within the community to further develop genome-scale models of CHO have also recently been announced [43].

The role of post-transcriptional regulation in CHO cell growth was demonstrated using an integrated analysis of the CHO transcriptome (mRNA and miRNA) and proteome in antibody-producing sister clones [58]. The authors analyzed gene expression at both the mRNA and protein levels combined with *in silico* target prediction for the differentially expressed, growth-correlated miRNAs. Their analysis focused on a group of 158 differentially expressed proteins for which the levels of their coding mRNAs were not altered, when slow versus fast growing clones were compared. Evidence was found for potential miRNA-mediated translational repression for 41 out of those 158 differentially expressed proteins [58]. The interactions between the respective coding mRNAs and their miRNAs were, however, not confirmed *in vivo*. This study highlighted the prevalence of regulatory mechanisms controlling cellular phenotypes that are relevant to biomanufacturing. Roughly a quarter of the proteins associated with growth rate were putatively found to be regulated by classical miRNA-mediated translational repression, while the regulatory mechanisms for the remaining proteins were unaccounted for. In addition, the study was limited by the specific technologies chosen (respectively, microarrays for mRNA assessment and LC-MS for proteomics) that yielded incomplete evidence at both the mRNA and protein levels, with several molecules for which their functional counterpart was not present in the dataset. Despite its limitations, this study showcases the power of integrated multi-'omics analyses to study mammalian cell lines in industry relevant scenarios that lead to the identification of engineering targets for genetic manipulation and cell line improvement.

Metabolic models have been used in cell culture development for some time primarily as a natural extension of traditional biochemical data (e.g., glucose, lactate, ammonia, amino acids) collected during process development and optimization. The topic has been thoroughly reviewed [72,73], and herein we provide additional models published within the past 2 years to study and describe mammalian cell metabolism (Table 2). Some key aspects of the central metabolism of CHO cells were recently outlined in two publications. A comprehensive study by Wahrheit *et al.* characterized the different growth phases of CHO cells and their corresponding metabolic states using time-resolved dynamic metabolic flux analysis [74]. One key contribution of this study was the characterization of compartmentalized enzymatic activities, differentiating metabolic activities within the cytosol and the mitochondria. Combined with the coupled growth-metabolic state analysis performed by the authors, this study effectively provided a functional, spatial and temporal resolution of the CHO metabolism. The authors further suggested that this information can be used as additional constraints in metabolic network reconstructions, potentially leading to more accurate metabolic models [74]. The relevance of metabolism compartmentalization was subsequently confirmed using isotopic labeling and dynamic flux analysis. This follow-up study also showed the high degree of metabolic reversibility and exchange with the extracellular environment, particularly for alanine and pyruvate, of CHO-K1 cells in batch culture [69].

Key term

Proteomics: The study of the proteome which includes all proteins produced and modified by an organism or system.

Table 1. Examples of multi-'omics studies in industrial mammalian cell lines.

Cell line	Phenotype	'Omics	Ref.
CHO	Metabolic shift	Transcriptomics and proteomics	[49]
CHO	Low culture temperature	Transcriptomics and proteomics	[50]
CHO	High productivity	Transcriptomics and proteomics	[51]
NS0	High productivity	Transcriptomics and proteomics	[52]
NS0	High cell density	Transcriptomics and proteomics	[53]
CHO	Increased productivity	Transcriptomics and proteomics	[54]
CHO	Cellular growth rate	Transcriptomics and proteomics	[55]
CHO	High productivity	Transcriptomics and proteomics	[56]
CHO	Growth and productivity	Transcriptomics and epigenomics	[57]
CHO	Growth rate	Transcriptomics, epigenomics and proteomics	[58]
HEK-293	Productivity	Transcriptomics, metabolomics and fluxomics	[59]
HEK-293	Protein expression	Transcriptomics and proteomics	[60]
CHO	Productivity, growth and cell size	Transcriptomics and proteomics	[61]
CHO	Productivity	Transcriptomics and epigenomics	[62]

CHO: Chinese hamster ovary.

The Jolicoeur lab has also published a kinetic model of the CHO central carbon metabolism that was used to describe the metabolism of CHO clones producing a monoclonal antibody (mAb) using an inducible gene switch [75]. Comparing the parental line and two clones with variable mAb levels (low and high producers), the authors found that differences in metabolic flux were mostly related to clonal variation and not correlated with mAb production. A similar finding was reported by another group working with various CHO cell lines that found no correlation between mAb productivity and the overall transcriptomic and proteomic space when analyzed by principal component analysis [61]. The metabolic model from the Jolicoeur lab was further developed to study energy metabolism by including known regulatory mechanisms of the glycolysis pathway as related to oxygen availability [76]. The *in silico* model was utilized to simulate early responses to hypoxia under different metabolic regulation scenarios, including known positive and negative feedback and feed-forward loops affecting key glycolytic enzymes. In addition to these, the inclusion of a regulatory parameter to reflect the cell's energetic state (AMP-to-ATP ratio) improved the model's predictions of the responses of CHO cells to anaerobic conditions, and particularly the shift from aerobic (oxidative) to anaerobic metabolism [76]. The study highlighted the importance of accounting for metabolic regulation when developing *in silico* models in order to properly simulate, and predict, cell behavior in response to changing environmental conditions.

Multi-'omic approaches in HEK-293

Although commonly reported using other cell-based systems [45–46,49,55,58], initial multi-'Omic, partial systems biology studies with HEK-293 cell lines are now appearing in the literature [59,60]. Evidence from these studies clearly demonstrates the benefits that could be gained by addressing a biological problem applying a systems biology approach. 'omics technologies have advanced to the point where it will become routine to integrate various highly dimensional datasets in order to distill the results into comprehensible and relevant components.

In an early study with HEK-293, Lee *et al.* [77] attempted to unravel significant metabolic changes between batch and low-glutamine fed-batch cultures using DNA microarray technology. The microarray data definitively characterized nutrient-related stress and captured the transcriptional changes occurring in both cultures. Significant transcriptional differences were observed mainly in the amino acid metabolism, tRNA synthetase, TCA cycle, electron transport chain and glycolysis pathways and led to the conclusion that fed-batch cultures were more efficient in amino acid metabolism and energy production. Lower expression for genes involved in glutamine/glutamate metabolism and serine/glycine/cysteine metabolism in combination with altered amino acid production and consumption rates provided evidence that the fed-batch process increased efficiency in amino acid metabolism especially during the later phases. Increased efficiency in energy metabolism of fed-batch cultures was implied by the downregulation of TCA cycle genes and the upreg-

ulation of genes in the electron transport chain. The conclusions relied on the assumption that decreased TCA gene expression and increased gene expression in components of the electron transport machinery was a direct indication of the TCA cycle and electron transport chain activities. Even with the lack of extensive metabolic data in this study, by integrating it with the transcriptomic data provided strong evidence for the changes in the amino acid metabolism. On the other hand, the interpretation of an increased energy metabolism may very well be valid, but was based on more circumstantial evidence without the actual levels of energy-related metabolites.

The incorporation of data from multiple 'Omic is becoming a necessity. The integration of these datasets provides more informative value than the individual 'omics technology. The complex biology of production cell lines can only be appreciated in light of integrating multiple 'omics datasets as illustrated in Dietmair *et al.* [59]. The authors set out to identify target pathways for cellular engineering that could lead to higher productivity rates. Metabolic differences between a HEK-293 cell line, stably producing a fusion protein, and the parental cell line were investigated using metabolic data for 52 metabolites, transcriptomic data generated from microarrays, and metabolite consumption and production rates revealed through a flux model. Neither cell line showed any outward phenotypic difference (growth, viability or morphology) and alleviated any potential confounding factors unrelated to recombinant protein production. The complexity of cell metabolism could only be revealed using a multi-'Omic approach. For example, the metabolomic and fluxomic data showed that all the production cell lines used in this study were characterized with a decrease in glucose uptake rate, while the transcriptomic data indicated that the expression of glucose transporters was actually increased in the same cell lines. This result suggested that glucose uptake rates were mediated by other factors and were not limited by the decrease in expression of the glucose transporter genes. This finding may be unique to HEK-293 cells since studies in other cell lines [78,79] have found positive correlations between glucose transporter expression and glucose uptake rates. In conjunction with the reduced glucose uptake, there was a corresponding reduction in the glycolytic flux. Analysis of the expression levels showed that most of the glycolytic pathway genes were reduced in the production cell line. However, no correlation was found between the genes encoding the central rate-limiting glycolytic enzymes (i.e., HK, PFK and PK). Together these data indicated that at least a portion of the reduced glycolytic flux was regulated at transcript level and the reduced uptake could be a consequence of downstream regulation [59]. The complexity of the regulation becomes clear when multiple 'omics datasets are integrated and included with the biological interpretation of the data.

More recently, Dumaual *et al.* [60] applied a multi-'omics approach to gain a basic understanding of the molecular alterations associated with the expression of the *PRL-1* gene. The PRL-gene family of enzymes is a potential tumor biomarker and a potential anticancer target. Increased expression of PRL-1 has a causal role in cellular transformation and tumor advancement, and although many interactions have been shown, little is known about its biological function. By integrating transcriptomic and proteomic analyses, the current knowledge of interactions and the signaling network of PRL-1 in a stable PRL-1 overexpressing cell line and its parental HEK-293 cell line was expanded upon. Application of functional enrichment analyses to the mRNA and protein datasets supported PRL-1's putative role in cytoskeletal remodeling, cell adhesion and transcription. In particular, they were able to identify a number of differentially expressed transcripts and proteins and focus on a few new high probability targets for future functional studies. Additionally, by integrating and comparing the two 'omics datasets, Dumaual *et al.* [60] showed the coordinated regulation occurring at the

Table 2. Examples of recently published metabolic models of industrial mammalian cell lines.

Cell line	Analysis method	Ref.
HEK-293	^{13}C-MFA + isotopomer balancing	[63]
HEK-293	Flux balance analysis	[64]
CHO	Constraints-based flux analysis	[65]
CHO	MFA	[66]
CHO	Flux balance analysis	[67]
CHO	Minimal elementary flux modes	[68]
CHO	Nonstationary ^{13}C-MFA	[69]
HEK-293	MFA	[70]

CHO: Chinese hamster ovary; MFA: Metabolic flux analysis.

mRNA and protein levels. In their study, 91% of the mRNA expression levels were changing in the same direction as their corresponding protein levels, suggesting that protein abundance was directly related to mRNA levels in their system.

As with recombinant protein production, a myriad of virus-based biopharmaceutical products are manufactured in mammalian cell based systems [80]. To gain a better understanding of the physiological changes during recombinant viral biopharmaceutical production, the authors followed a functional genomics approach, including an integrated transcriptomic and metabolomic analysis, to gain a better understanding of the molecular and metabolic events taking place during the transition of the human parental cell line to its producer cell line. The study also reported on a novel approach to mine large transcriptome datasets produced from high productivity systems. This approach entails comparing high versus low productivity in two human cell lines that were from different genetic backgrounds in order to identify the transcriptional changes due to retrovirus production and not intrinsic genetic cell line properties. Several genes were identified to be limiting factors in the low producer cell line and gene manipulations of these targets showed promising increases in infectious virus-specific productivity [80].

Multi-'omics approaches in other mammalian cell lines

Many different mammalian cell lines are currently being used in the production of biopharmaceutical products including mouse myeloma lymphoblastoid-like cells (e.g., NS0), [81,82], baby hamster kidney cells (e.g., BHK-21) [83] and human retina-derived cells (e.g., Per.C6) [84–86]. Unlike CHO and HEK-293, little has been done to apply multi-'omics approaches. Most of the current publications only evaluated a single 'omics dataset, investigating either the transcriptomic or sometimes proteomic analyses by itself.

Productivity of cell lines is an important factor in bioprocessing. Seth *et al.* [52] used a systematic approach to evaluate the production of the same antibody in 11 different NS0 cell lines. These cell lines were categorized into low-producing and high-producing groups. The systematic approach included transcriptomic data gathered from DNA microarray analysis and proteomic data derived from two-dimensional gel electrophoresis and Isobaric Tagging for Relative and Absolute Protein Quantification (iTRAQ). This integrated approach identified differentially expressed genes at both the transcriptomic and proteomic level for each of the low and high producing NS0 cell lines. Seth *et al.* [52] found that in the high producers, protein synthesis pathways were changed at both the transcriptome and proteome level while cell growth and cell death pathways were affected at the transcriptional level only. Ultimately, integrated transcriptomic and proteomic data will provide valuable information for understanding and designing high-producer cell lines.

In a similar fashion, Krampe *et al.* [53] combined gene and protein expression profiling to better understand intracellular responses of NS0 cells grown in perfusion culture. More specifically, factors such as growth rate, production rate, metabolic activity and cell viability were assessed as cell density increased. DNA microarray, real-time quantitative PCR and Western blot analysis revealed that a balance among factors involved in energy metabolism are responsible for fine tuning the choice of a cell to either go into survival mode or apoptosis. This simultaneous modulation of several physiological functions was also observed by Charaniya *et al.* [87] when the transcriptome of several NS0 cell lines with a broad range of antibody production levels were analyzed. Using Gene Set Enrichment Analysis, Gene Set Analysis and MAPPFinder between the high and low producers, they found that protein processing and transport, including protein modification, vesicle trafficking and protein turnover were significant. Also, mitochondrial ribosomal function, cell cycle regulation and cytoskeleton-related elements were altered in high-producing cell lines.

The safety and efficacy of biologics is tied to their quality, glycosylation being a key quality aspect for many classes of recombinant proteins [88–90]. A modeling framework comprising the *in silico* reconstruction of nucleotide sugar donor metabolism was developed for a hybridoma cell culture and linked to an existing glycosylation model [91]. The *in silico* metabolic model considered cell growth (as a function of extracellular glucose and glutamine), nucleotide metabolism and nucleotide sugar donor metabolism. To link the extracellular environment to antibody glycosylation profiles, the outputs of this metabolic network were fed to the glycosylation model and the simulation results obtained by this combined approach were in agreement with experimental data [91]. This study showcased how an integrated approach can be developed to address cell culture traits that are relevant to commercial biomanufacturing.

Next-generation sequencing has been a powerful tool to investigate the genome, transcriptome and epigenome of organisms. Johnson *et al.* [92] investigated the transcriptomic space of a recombinant BHK cell line using NGS with the aim to establish a well-characterized reference genome to assist future research projects with BHK cell lines. With the extremely high coverage that was achieved through RNA-seq of high abundant genes, the authors were able to identify low-frequency nucleotide variants in these genes that were not previously detectable at the genome level. Such studies

are providing the stepping stones for systems-based genome engineering of cell lines that are important to biopharmaceutical processes.

Challenges, limitations & solutions

A critical step for any systems-level analysis is to identify the components comprising the system and collect empirical data for each component. A system-wide biological model cannot be efficiently produced by only sequencing one gene or measuring the expression of one mRNA, protein or metabolite at a time. The ability to generate comprehensive quantitative datasets for the biological system in question is critical and a system-wide measurement of all components is now possible with high-throughput quantification techniques. While great strides have been made in all areas of 'omics data generation, there are still many technical limitations. RNA-seq analyses will measure many more unique genes than proteomics will measure of proteins, or metabolomics will of metabolites [93]. A complete view of the biological system will require the comprehensive study of each of the different 'omics components. There are technical challenges that must be overcome to ultimately accomplish this. For example, there are many proteins that have traditionally been problematic to identify. From a sample preparation view point, membrane-bound or low-abundance proteins are known to be difficult to extract and purify [94]. Recent developments in shotgun proteomics are overcoming these limitations [93]. Even as late as 2009, only a few thousand proteins were typically identified in any given proteomics study [94]. Conversely, with technological advancements in proteomics, two recent studies identified approximately 10,000 proteins each from human cancer cell lines, which is considered to be nearly complete coverage of the human proteome [95–97]. Though it is becoming increasingly higher-throughput, proteomics is still lagging behind nucleotide sequencing in its ability to generate a system-wide expression profile. A similar problem exists in metabolomics where any single platform that is used to study metabolites is not capable of identifying all metabolites in a sample. This requires the use of multiple platforms to analyze samples, which increases not only costs but also adds the problem of how to combine data from these different sources [98].

After data generation, the next step in systems biology is data integration, analysis and visualization. With the massive accumulation of 'omics data, advanced computational software is required to make sense of it all [99]. Data integration requires unraveling the relationships between the different levels of genetic information. A number of sophisticated statistical techniques have been developed to facilitate the analysis/integration of highly dimensional 'omics datasets including regularized canonical correlation analysis and sparse partial least squares regression [100]. Additional multivariate data analyses have been proposed for the integrated analyses of multiple 'omics datasets, often originating from other areas such as translational research [101–105]. In addition to these techniques, a more standard Pearson correlation coefficient can be computed in order to find relationships between different 'omics datasets. These coefficients can then be used to produce network maps or overlay the data on pathway maps such as KEGG or HumanCyc to visualize the relationships between the different components [106]. The numerical results of data analysis alone may not be sufficient for obtaining biologically meaningful information from the analyses, thus requiring additional visualization tools to help clarify their biological significance [100,107]. The integration, analysis and visualization of these highly dimensional datasets is a current bottleneck in systems biology. A lot of effort is being dedicated to address this gap, with several public and commercial software platforms available. In the short- and mid-term, there will be no single solution that applies to all possible scenarios. The best approach will be specific to each production platform and associated datasets.

As more and more data are being generated, and with improvements in NGS sequencing chemistry and technology, data storage capacity of both raw and analyzed data has to be taken into consideration when planning an experiment. Specifically for system-wide genomics and transcriptomics, reads from next-generation sequencers can produce files that are many gigabytes in size and can easily close in on a terabyte worth of data from a single experiment. Moreover, systems biology is inherently multidisciplinary and the transfer and sharing of these massive datasets bring their own unique set of challenges. This also requires the standardization of data reporting, storage and curation. For example, the minimum information about a microarray experiment (MIAME) standards have been implemented since 2001 [108]. The concept has been expanded to other fields of study. Since 2012, this was applied to NGS with the establishment of minimum information about a high-throughput nucleotide sequencing experiment (MINSEQE) standards [109]. The proteomics equivalent is minimum information about a proteomics experiment (MIAPE) [110] while for metabolomics, a few initiatives are under way. The minimum information about a metabolomics experiment (MIAMET) [111] standards have been proposed but are not widely adopted. The Metabolomics Standard Initiative [112] is a group under the umbrella of the Metabolomics Society that are examining standardization. Their goal is to ultimately make recommendations to researchers in the field.

Not only does the study of each molecule have its own set of standards, systems biology has begun to

develop a 'language' of its own in order to standardize the notation for biological modeling. There are a number of contenders for the standardization of systems biology models such as the SBML [113], CellML [114,115], the MIRIAM [116] and the KEGG Genome Markup Language (KGML) [117]. Without standards in place, data sharing becomes increasingly hindered such that each research facility would potentially have to recreate every experiment of interest.

Data usage

The systems biology approach addresses four key aspects of a biological system: its structure, dynamics, control and design principles [11]. Mathematical modeling is central to achieving these four goals. In general terms, a mathematical framework is employed to create *in silico* models that attempt to describe the structure and functional dynamics of biological systems. These models can then be used to simulate the system's behavior, for instance, in response to external perturbations. This iterative process of simulation and validation allows scientists to identify and study the systems' control mechanisms and design principles. For example, it seeks to answer questions such as, how does the system maintain its steady state? What are the systems' functional modules?

The system-wide, preferably quantitative, data collected are critical to create and refine the mathematical model representing a biological system [20]. For a bioprocessing system, data collected means not only measuring expressed RNA, proteins, metabolites, sequencing the genome, etc., but also recording all other biological and physical, process-related, information about the system. These can include, for instance, mutations in the genome, post-translational modifications to proteins, media composition, flask type, feeding strategies and phenotypic outcomes, among others. These diverse, highly dimensional and rich datasets are a prerequisite to generate a model [99]. After a model is created, it needs to be experimentally tested to verify and refine its accuracy. *In silico* predictions are thus compared with experimental observations or pre-existing knowledge. If the results predicted by the model being tested do not reflect what is found experimentally, the model needs to be updated with data collected from the test experiment to reflect the knowledge gained.

Based on this holistic understanding of their host-cell lines, the goal is that bioprocessing scientists and engineers can manipulate current and/or design new production systems (i.e., cell hosts) that outperform those currently in use in terms of titers, quality, process stability and predictability. Specifically to bioprocessing, the *in silico* experiments can be conducted to simulate the cell's response to various process conditions. In traditional process optimization, these conditions might represent different media formulations, feed strategies, scale-up/-down processes or culture conditions. In cell line development, these same models can guide the design of optimized cell engineering strategies. Here, the goal is to guide targeted genetic manipulations to address genetic, metabolic or cellular processes to improve phenotypic outputs of interest (e.g., cell viability, titer, glycosylation profile and by-product accumulation, among others). One advantage of modeling is that it allows scientists to evaluate a greater number of variables at one time. Based on the model's outputs, they are able to prioritize strategies that will yield greater benefits, streamlining efforts and ultimately, saving time, effort and valuable resources. Overall, modeling is a very active area of research and we refer the interested reader to more in-depth reviews on this topic [118,119].

Conclusion & future perspective

Ultimately, individual 'omics technologies will need to achieve parity with one another in terms of comprehensiveness, throughput and quantitativeness. Additionally, improvements in analytical software and hardware tools will become essential in order to handle increasingly larger datasets and we need to improve the ability to store and transfer these massive datasets. More sophisticated statistical and bioinformatics methodologies will also be needed to assist with improved data integration and system modeling, with better predictive capabilities.

However advanced our data generation, analyses, and storage abilities become, the biological interpretation and utilization of the data will undoubtedly be the major bottleneck. This is due to our limited understanding of the cell as a system, and its structure, dynamics, control and design principles. Genome-scale models are important tools to formally aggregate the biological knowledge around these principles within a mathematical framework designed to represent the cell as a complex system. These models are designed taking into account the variables that control cellular behavior and function, providing further insights into the biology of the organism being studied. As these models are developed and refined over time, they will continue to improve in accuracy and predictive power.

In simpler prokaryotic organisms, such as *Mycoplasma* and *E. coli* that have an extensive knowledge base, such established models have evolved incrementally. Information on additional cellular processes are being constantly incorporated, resulting in more sophisticated models. Mammalian cells are inherently more complex, both genetically and structurally, making modeling much more challenging. Similar to the path followed by prokaryote sys-

tems, the 'omics efforts currently underway within the mammalian cell culture community will serve the important role of establishing and growing the knowledge base required to accomplish a full systems biology based understanding of mammalian production platforms.

The wide adoption and utilization of 'omics technologies as part of the overall systems biology approach to cell culture and biomanufacturing optimization will be highly reliant on technological advancements capable of driving down the overall costs of experimentation and data analysis. Implementing the required infrastructure and know-how internally is not a simple undertaking, often requiring new partnerships to be established. There are very few commercial platforms developed specifically toward this application. One such example is ArrayXpress' (Raleigh, NC) iCOP – integrated Cellular 'omics platform, which is our suite of tools for building and mining systems biology knowledge bases.

Next-Generation Sequencing and 'omics are becoming pervasive in almost every field of biological and medical research, including cancer research, diagnostics and drug development. Regulatory agencies are already familiar with the application of these technologies in these areas, and indeed are keen on their utilization as a tool to improve human health [120]. It is also a particularly relevant tool for biosimilars, to help ensure their safety and efficacy. We fully anticipate the same trend to biomanufacturing, with regulatory endorsement of the application of NGS and 'omics to support new submissions. There is increasing recognition in the field of the benefits of NGS and 'omics as important tools within the US FDA's PAT and QbD framework [4,48].

Financial & competing interests disclosure

Work to develop a systems biology methodology and the iCOP product was funded through paid projects for biomanufacturing and biopharmaceutical companies through ArrayXpress, Inc (Raleigh, NC). The authors continue to work with paying commercial entities to develop and implement these methodologies. Authors with university affiliations also contribute the knowledge to their respective departments. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript beyond those disclosed.

No writing assistance beyond the listed authors was used in the production of this manuscript.

Executive summary

Background

- Systems biology is the study and subsequent mathematical modeling of the components that constitute a biological system, their interactions and the system's resulting emergent properties.
- Advances in 'omics technologies have propelled systems biology to mainstream biomedical/biological research.
- Biomanufacturing is harvesting some benefits of applied 'omics approaches for increasing titer, yield, quality, process efficiency and stability.
- The biopharmaceutical industry is starting to realize the limitations of single 'omics approaches and is slowly moving toward a true systems biology approach for data-driven, rational bioprocesses optimization and cell line development.
- Most applications of systems biology to date have been performed on prokaryotes with predictive genome-scale whole-cell models already having been developed for many species.

Mammalian cell lines

- A number of successful projects integrating two or more 'omics datasets on Chinese hamster ovary cells and HEK-293 have been published.
- There are currently no peer-reviewed reports of predictive genome-scale mammalian models having been implemented in an industrially relevant context.
- Systems biology resources geared toward mammalian cell lines are under development, priming production hosts for systems-wide analyses in the near future.

Working with the data

- Currently, metabolomic and proteomic data production lag behind transcriptomics, as a result, information obtained from data integration can be limited.
- Once data have been generated, integration, analysis, data mining and visualization are necessary to gain an understanding of the biological system.
- The massive amount of data being generated is creating new requirements in infrastructure, software and computing power for accessing, analyzing and storage.
- Wider utilization of systems biology by the industry is hampered by high cost in technology and data analyses.
- While the barriers for production of data have lowered, working with it and developing meaningful, actionable outcomes represent the major bottleneck for most companies.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 Evaluate Pharma. World Preview 2013, Outlook to 2018: Returning to Growth 1–39 (2013). http://info.evaluategroup.com/worldpreview2018_fp_lp.html

2 Carinhas N, Oliveira R, Alves PM, Carrondo MJT, Teixeira AP. Systems biotechnology of animal cells: the road to prediction. *Trends Biotechnol.* 30(7), 377–385 (2012).

3 Kim JY, Kim Y-G, Lee GM. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Appl. Microbiol. Biotechnol.* 93(3), 917–930 (2012).

4 Glassey J, Gernaey K V, Clemens C *et al.* Process analytical technology (PAT) for biopharmaceuticals. *Biotechnol. J.* 6(4), 369–377 (2011).

5 Gottschalk U, Brorson K, Shukla AA. The need for innovation in biomanufacturing. *Nat. Biotechnol.* 30(6), 489–492 (2012).

6 Langer E. A decade of biomanufacturing. *Bioprocess Int.* 10(6), 65–68 (2012).

7 Gottschalk U. Biomanufacturing: time for change? *Pharm. Bioprocess.* 1(1), 7–9 (2013).

8 Rader RA. FDA biopharmaceutical product approvals and trends in 2012. *Bioprocess Int.* 11(3), 18–27 (2013).

9 Davidson A, Farid SS. Innovation in biopharmaceutical manufacture. *Bioprocess Int.* 12(1), 12–19 (2014).

10 Ideker T, Thorsson V, Ranish JA *et al.* Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* 292(5518), 929–934 (2001).

11 Kitano H. Systems biology: a brief overview. *Science* 295(5560), 1662–1664 (2002).

12 Von Bertalanffy L. General system theory. *Gen. Syst.* 1(1), 11–17 (1956).

13 MesarovićMD. *Systems Theory And Biology – View Of A Theoretician.* Springer, Berlin, Heidelberg, Germany (1968).

14 De Schutter E. Why are computational neuroscience and systems biology so separate? *PLoS Comput. Biol.* 4(5), e1000078 (2008).

15 Wierzbicki AP. Systems theory, theory of chaos, emergence. In: *Technen: Elements of Recent History of Information Technologies with Epistemological Conclusions*. Springer International Publishing, Switzerland (2015).

16 Bose B. Systems biology: a biologist's viewpoint. *Prog. Biophys. Mol. Biol.* 113(3), 358–368 (2013).

17 Bizzarri M, Palombo A, Cucina A. Theoretical aspects of systems biology. *Prog. Biophys. Mol. Biol.* 112(1–2), 33–43 (2013).

18 Conesa A, Mortazavi A. The common ground of genomics and systems biology. *BMC Syst. Biol.* 8(Suppl. 2), S1 (2014).

19 Blazeck J, Alper H. Systems metabolic engineering: genome-scale models and beyond. *Biotechnol. J.* 5(7), 647–659 (2010).

20 Hyduke DR, Lewis NE, Palsson BØ. Analysis of omics data with genome-scale models of metabolism. *Mol. Biosyst.* 9(2), 167–74 (2013).

21 Gajadhar AS, White FM. System level dynamics of post-translational modifications. *Curr. Opin. Biotechnol.* 28, 83–87 (2014).

22 Kim J, Reed JL. Refining metabolic models and accounting for regulatory effects. *Curr. Opin. Biotechnol.* 29, 34–38 (2014).

23 Saha R, Chowdhury A, Maranas CD. Recent advances in the reconstruction of metabolic models and integration of omics data. *Curr. Opin. Biotechnol.* 29, 39–45 (2014).

24 Lee SK, Chou H, Ham TS, Lee TS, Keasling JD. Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. *Curr. Opin. Biotechnol.* 19(6), 556–563 (2008).

25 Park JH, Lee SY, Kim TY, Kim HU. Application of systems biology for bioprocess development. *Trends Biotechnol.* 26, 404–412 (2008).

26 Sagt CMJ. Systems metabolic engineering in an industrial setting. *Appl. Microbiol. Biotechnol.* 97(6), 2319–2326 (2013).

27 Caspeta L, Shoaie S, Agren R, Nookaew I, Nielsen J. Genome-scale metabolic reconstructions of *Pichia stipitis* and *Pichia pastoris* and in-silico evaluation of their potentials. *BMC Syst. Biol.* 6(1), 24 (2012).

28 De Jong B, Siewers V, Nielsen J. Systems biology of yeast: enabling technology for development of cell factories for production of advanced biofuels. *Curr. Opin. Biotechnol.* 23(4), 624–30 (2012).

29 Karr JR, Sanghvi JC, Macklin DN *et al.* A whole-cell computational model predicts phenotype from genotype. *Cell* 150(2), 389–401 (2012).

30 Carrera J, Estrela R, Luo J, Rai N, Tsoukalas A, Tagkopoulos I. An integrative, multi-scale, genome-wide model reveals the phenotypic landscape of *Escherichia coli*. *Mol. Syst. Biol.* 10, 735 (2014).

31 Orth JD, Conrad TM, Na J *et al.* A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism–2011. *Mol. Syst. Biol.* 7(535), 1–9 (2011).

32 Park JH, Lee KH, Kim TY, Lee SY. Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. *Proc. Natl Acad. Sci. USA* 104(19), 7797–7802 (2007).

33 Yegane-Sarkandy S, Farnoud AM, Shojaosadati SA *et al.* Overproduction of human interleukin-2 in recombinant *Escherichia coli* BL21 high-cell-density culture by the determination and optimization of essential amino acids using a simple stoichiometric model. *Biotechnol. Appl. Biochem.* 54(1), 31–39 (2009).

34 Thiele I, Jamshidi N, Fleming RMT, Palsson BO. Genome-scale reconstruction of escherichia coli's transcriptional and translational machinery: a knowledge base, its mathematical formulation, and its functional characterization. *PLoS Comput. Biol.* 5(3), e1000312 (2009).

35 Thiele I, Fleming RMT, Que R, Bordbar A, Diep D, Palsson BØ. Multiscale modeling of metabolism and macromolecular synthesis in *E. coli* and its application to the evolution of codon usage. *PLoS ONE* 7(9), 7(9), e45635 (2012).

36 O'Brien EJ, Lerman J, Chang RL, Hyduke DR, Palsson BØ. Genome-scale models of metabolism and gene expression

extend and refine growth phenotype prediction. *Mol. Syst. Biol.* 9(693), 693 (2013).

37 Liu JK, Brien EJO, Lerman JA, Zengler K, Palsson BØ, Feist AM. Reconstruction and modeling protein translocation and compartmentalization in Escherichia coli at the genome-scale. *BMC Syst. Biol.* 8, 110 (2014).

38 Campodonico MA, Andrews BA, Asenjo JA, Palsson BØ, Feist AM. Generation of an atlas for commodity chemical production in *Escherichia coli* and a novel pathway prediction algorithm, GEM-Path. *Metab. Eng.* 25, 140–158 (2014).

• **A systematic analysis of all possible metabolic routes to produce commodity chemicals in *Escherichia coli*. The authors further develop GEM-Path, a pathway prediction algorithm that integrates genome-scale models for evaluating the growth-coupled metabolic synthesis of both native and non-native compounds. A workflow for the creation of production strains is also provided.**

39 CHO genome. www.CHOgenome.org

40 CGCDB: CHO gene coexpression database. www.cgcdb.org

41 Lewis NE, Liu X, Li Y *et al.* Genomic landscapes of Chinese hamster ovary cell lines as revealed by the *Cricetulus griseus* draft genome. *Nat. Biotechnol.* 31(8), 759–765 (2013).

•• **Landmark publication on the characterization of the Chinese hamster genome and six derived production lines. The study highlights their genetic diversity, including single-nucleotide polymorphisms, indels and copy number variants. A key resource for Chinese hamster ovary (CHO) genomics.**

42 Wurm F. CHO quasispecies – implications for manufacturing processes. *Processes* 1(3), 296–311 (2013).

43 Borth N. Opening the black box: Chinese hamster ovary research goes genome scale. *Pharm. Bioprocess.* 2(5), 367–369 (2014).

44 Datta P, Linhardt RJ, Sharfstein ST. An 'omics approach towards CHO cell engineering. *Biotechnol. Bioeng.* 110(5), 1255–1271 (2013).

45 Kildegaard HF, Baycin-Hizal D, Lewis NE, Betenbaugh MJ. The emerging CHO systems biology era: harnessing the 'omics revolution for biotechnology. *Curr. Opin. Biotechnol.* 24(6), 1102–1107 (2013).

46 Farrell A, McLoughlin N, Milne JJ, Marison IW, Bones J. Application of multi-omics techniques for bioprocess design and optimization in chinese hamster ovary cells. *J. Proteome Res.* 13(7), 3144–3159 (2014).

47 Kyriakopoulos S, Kontoravdi C. Insights on biomarkers from Chinese hamster ovary "omics" studies. *Pharm. Bioprocess* 2(5), 389–401 (2014).

48 Carrondo MJT, Alves PM, Carinhas N *et al.* How can measurement, monitoring, modeling and control advance cell culture in industrial biotechnology? *Biotechnol. J.* 7(12), 1522–1529 (2012).

49 Korke R, Gatti MDL, Lau ALY *et al.* Large scale gene expression profiling of metabolic shift of mammalian cells in culture. *J. Biotechnol.* 107(1), 1–17 (2004).

50 Baik JY, Lee MS, An SR *et al.* Initial transcriptome and proteome analyses of low culture temperature-induced expression in CHO cells producing erythropoietin. *Biotechnol. Bioeng.* 93(2), 361–371 (2006).

51 Nissom PM, Sanny A, Kok YJ *et al.* Transcriptome and proteome profiling to understanding the biology of high productivity CHO cells. *Mol. Biotechnol.* 34(2), 125–140 (2006).

52 Seth G, Philp R, Lau A. Molecular portrait of high productivity in recombinant NS0 cells. *Biotechnol. Bioeng.* 97(4), 933–951 (2007).

53 Krampe B, Swiderek H, Al-Rubeai, M. Transcriptome and proteome analysis of antibody-producing mouse myeloma NS0 cells cultivated at different cell densities in perfusion culture. *Biotechnol. Appl. Biochem.* 50(Pt 3), 133–141 (2008).

54 Yee JC, de Leon Gatti M, Philp RJ, Yap M, Hu W-S. Genomic and proteomic exploration of CHO and hybridoma cells under sodium butyrate treatment. *Biotechnol. Bioeng.* 99(5), 1186–1204 (2008).

55 Doolan P, Meleady P, Barron N *et al.* Microarray and proteomics expression profiling identifies several candidates, including the valosin-containing protein (VCP), involved in regulating high cellular growth rate in production CHO cell lines. *Biotechnol. Bioeng.* 106(1), 42–56 (2010).

56 Kantardjieff A, Jacob NM, Yee JC *et al.* Transcriptome and proteome analysis of Chinese hamster ovary cells under low temperature and butyrate treatment. *J. Biotechnol.* 145(2), 143–159 (2010).

57 Hernández Bort JA, Hackl M, Höflmayer H *et al.* Dynamic mRNA and miRNA profiling of CHO-K1 suspension cell cultures. *Biotechnol. J.* 7(4), 500–515 (2012).

58 Clarke C, Henry M, Doolan P *et al.* Integrated miRNA, mRNA and protein expression analysis reveals the role of post-transcriptional regulation in controlling CHO cell growth rate. *BMC Genomics* 13(1), 656 (2012).

•• **This multi-'omics study identified widespread miRNA-mediated post-transcriptional regulation of cell growth, ultimately providing mechanistic insights into and potential targets for manipulation of this trait.**

59 Dietmair S, Hodson MP, Quek L-E, Timmins NE, Gray P, Nielsen LK. A multi-omics analysis of recombinant protein production in Hek293 cells. *PLoS ONE* 7(8), e43394 (2012).

• **Integrated transcriptomics, metabolomics and fluxomics analyses in Hek-293 revealed complex layers of regulatory mechanisms that ultimately influence metabolic output and recombinant protein production. Instead of specific metabolic changes, a system-wide adaptation of the cellular network was proposed to free metabolic resources for recombinant protein production.**

60 Dumaual CM, Steere BA, Walls CD, Wang M, Zhang Z-Y, Randall SK. Integrated analysis of global mRNA and protein expression data in HEK293 cells overexpressing PRL-1. *PLoS ONE* 8(9), e72977 (2013).

61 Kang S, Ren D, Xiao G *et al.* Cell line profiling to improve monoclonal antibody production. *Biotechnol. Bioeng.* 111(4), 748–760 (2014).

62 Maccani A, Hackl M, Leitner C *et al.* Identification of microRNAs specific for high producer CHO cell lines using steady-state cultivation. *Appl. Microbiol. Biotechnol.* 98(17), 7535–7548 (2014).

63 Henry O, Jolicoeur M, Kamen A. Unraveling the metabolism of HEK-293 cells using lactate isotopomer analysis. *Bioprocess Biosyst. Eng.* 34(3), 263–273 (2011).

64 Dietmair S, Hodson MP, Quek L-E *et al.* Metabolite profiling of CHO cells with different growth characteristics. *Biotechnol. Bioeng.* 109(6), 1404–1414 (2012).

65 Selvarasu S, Ho YS, Chong WPK *et al.* Combined in silico modeling and metabolomics analysis to characterize fed-batch CHO cell culture. *Biotechnol. Bioeng.* 109(6), 1415–1429 (2012).

66 Ghorbaniaghdam A, Henry O, Jolicoeur M. A kinetic-metabolic model based on cell energetic state: study of CHO cell behavior under Na-butyrate stimulation. *Bioprocess Biosyst. Eng.* 36(4), 469–487 (2013).

67 Martínez VS, Dietmair S, Quek L-E, Hodson MP, Gray P, Nielsen LK. Flux balance analysis of CHO cells before and after a metabolic switch from lactate production to consumption. *Biotechnol. Bioeng.* 110(2), 660–666 (2013).

68 Zamorano F, Vande Wouwer A, Jungers RM, Bastin G. Dynamic metabolic models of CHO cell cultures through minimal sets of elementary flux modes. *J. Biotechnol.* 164(3), 409–422 (2013).

69 Nicolae A, Wahrheit J, Bahnemann J, Zeng A-P, Heinzle E. Non-stationary 13C metabolic flux analysis of Chinese hamster ovary cells in batch culture using extracellular labeling highlights metabolic reversibility and compartmentation. *BMC Syst. Biol.* 8, 50 (2014).

• **Comprehensive characterization of the CHO metabolism in batch culture using dynamic flux analysis in combination to *in situ* enzymatic measurements focusing on the intersection of the cytosolic glycolysis and the mitochondrial tricarboxylic acid cycle as a key energetic hub. The ratio between these two pathways is utilized as an indicator of metabolic efficiency in CHO.**

70 Quek LE, Dietmair S, Hanscho M, Martínez VS, Borth N, Nielsen LK. Reducing Recon 2 for steady-state flux analysis of HEK cell culture. *J. Biotechnol.* 184, 172–178 (2014).

71 Smallbone K. Standardized network reconstruction of CHO cell metabolism (2013). http://arxiv.org/pdf/1304.3146v1.pdf

•• **A CHO genome-scale network reconstruction is provided that includes both a genome-scale network reconstruction (GENRE) and its derived genome-scale model (GEM) using systems biology markup language and minimum information required in annotation of models standards. The model contains 1065 genes, 1545 metabolic reactions and 1218 unique metabolites.**

72 Ahn WS, Antoniewicz MR. Towards dynamic metabolic flux analysis in CHO cell cultures. *Biotechnol. J.* 7(1), 61–74 (2012).

73 Young JD. Metabolic flux rewiring in mammalian cell cultures. *Curr. Opin. Biotechnol.* 24(6), 1108–1115 (2013).

74 Wahrheit J, Niklas J, Heinzle E. Metabolic control at the cytosol-mitochondria interface in different growth phases of CHO cells. *Metab. Eng.* 23, 9–21 (2014).

75 Ghorbaniaghdam A., Chen J, Henry O, Jolicoeur M. Correction: analyzing clonal variation of monoclonal antibody-producing CHO cell lines using an *in silico* metabolomic platform. *PLoS ONE* 9(3) (2014).

76 Ghorbaniaghdam A, Henry O, Jolicoeur M. An in-silico study of the regulation of CHO cells glycolysis. *J. Theor. Biol.* 357, 112–122 (2014).

77 Lee Y, Wong K, Nissom P. Transcriptional profiling of batch and fed-batch protein-free 293-HEK cultures. *Metab. Eng.* 9(1), 52–67 (2007).

78 Flier JS, Mueckler MM, Usher P, Lodish HF. Elevated levels of glucose transport and transporter messenger RNA are induced by ras or src oncogenes. *Sci.* 235(4795), 1492–1495 (1987).

79 Paredes C, Prats E, Cairó JJ, Azorín F, Cornudella L, Gòdia F. Modification of glucose and glutamine metabolism in hybridoma cells through metabolic engineering. *Cytotechnology* 30(1–3), 85–93 (1999).

80 Rodrigues AF, Formas-Oliveira AS, Bandeira VS, Alves PM, Hu WS, Coroadinha AS. Metabolic pathways recruited in the production of a recombinant enveloped virus: mining targets for process and cell engineering. *Metab. Eng.* 20, 131–145 (2013).

81 Bebbington CR, Renner G, Thomson S, King D, Abrams D, Yarranton GT. High-level expression of a recombinant antibody from myeloma cells using a glutamine synthetase gene as an amplifiable selectable marker. *Biotechnology* 10, 169–175 (1992).

82 Barnes LM, Bentley CM, Dickson AJ. Advances in animal cell recombinant protein production: GS-NS0 expression system. *Cytotechnology* 32, 109–123 (2000).

83 Wurm FM. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat. Biotechnol.* 22, 1393–1398 (2004).

84 Pau MG, Ophorst C, Koldijk MH, Schouten G, Mehtali M, Uytdehaag F. The human cell line PER.C6 provides a new manufacturing system for the production of influenza vaccines. *Vaccine* 19, 2716–2721 (2001).

85 Jones D, Kroos N, Anema R *et al.* High-level expression of recombinant IgG in the human cell line PER.C6. *Biotechnol. Prog.* 19, 163–168 (2003).

86 Swiech K, Picanço-Castro V, Covas DT. Human cells: new platform for recombinant therapeutic protein production. *Protein Expr. Purif.* 84(1), 147–153 (2012).

87 Charaniya S, Karypis G, Hu W-S. Mining transcriptome data for function-trait relationship of hyper productivity of recombinant antibody. *Biotechnol. Bioeng.* 102(6), 1654–1669 (2009).

88 Eon-Duval A, Broly H, Gleixner R. Quality attributes of recombinant therapeutic proteins: an assessment of impact on safety and efficacy as part of a quality by design development approach. *Biotechnol. Prog.* 28(3), 608–622 (2012).

89 Lingg N, Zhang P, Song Z, Bardor M. The sweet tooth of biopharmaceuticals: importance of recombinant protein glycosylation analysis. *Biotechnol. J.* 7(12), 1462–1472 (2012).

90 Costa AR, Rodrigues ME, Henriques M, Oliveira R, Azeredo J. Glycosylation: impact, control and improvement during therapeutic protein production. *Crit. Rev. Biotechnol.* 8551, 1–19 (2013).

91 Jedrzejewski PM, del Val IJ, Constantinou A *et al.* Towards controlling the glycoform: a model framework linking extracellular metabolites to antibody glycosylation. *Int. J. Mol. Sci.* 15, 4492–4522 (2014).

•• **The authors develop an *in silico* model particularly focused on nucleotide sugar donor metabolism that connects extracellular environment to the cell's metabolism and recombinant protein glycosylation.**

92 Johnson KC, Yongky A, Vishwanathan N *et al.* Exploring the transcriptome space of a recombinant BHK cell line through next generation sequencing. *Biotechnol. Bioeng.* 111(4), 770–781 (2014).

93 Mann M, Kulak NA, Nagaraj N, Cox J. The coming age of complete, accurate, and ubiquitous proteomes. *Mol. Cell.* 49, 583–590 (2013).

94 Wiśniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. *Nat. Methods* 6(5), 359–362 (2009).

95 Beck M, Schmidt A, Malmstroem J *et al.* The quantitative proteome of a human cell line. *Mol. Syst. Biol.* 7(549), 1–8 (2011).

96 Nagaraj N, Wisniewski JR, Geiger T *et al.* Deep proteome and transcriptome mapping of a human cancer cell line. *Mol. Syst. Biol.* 7(548), 1–8 (2011).

97 Breker M, Schuldiner M. The emergence of proteome-wide technologies: systematic analysis of proteins comes of age. *Nat. Rev. Mol. Cell Biol.* 15(7), 453–464 (2014).

98 Monteiro MS, Carvalho M, Bastos ML, Guedes de Pinho P. Metabolomics analysis for biomarker discovery: advances and challenges. *Curr. Med. Chem.* 20, 257–271 (2013).

99 Prokop A, Csukas B. *Systems Biology: Integrative Biology and Simulation Tools (Volume 1).* Springer, Dordrecht, The Netherlands (2013).

100 González I, Cao K-AL, Davis MJ, Déjean S. Visualising associations between paired "omics" data sets. *BioData Min.* 5(1), 19 (2012).

101 Boccard J, Rutledge DN. A consensus orthogonal partial least squares discriminant analysis (OPLS-DA) strategy for multiblock omics data fusion. *Anal. Chim. Acta* 769, 30–39 (2013).

102 Boccard J, Rutledge DN. Iterative weighting of multiblock data in the orthogonal partial least squares framework. *Anal. Chim. Acta* 813, 25–34 (2014).

103 Conesa A, Hernandez R. Omics data integration in systems biology: methods and applications. In: *Applications of Advanced Omics Technologies: From Genes to Metabolites (Volume 64).* García-Cañas V, Cifuentes A, Simó C (Eds). Elsevier, Oxford, UK, 441–459 (2014).

104 Kohl M, Megger DA, Trippler M *et al.* A practical data processing workflow for multi-OMICS projects. *Biochim. Biophys. Acta* 1844(1 Pt A), 52–62 (2014).

• **Presents a sample workflow for analyzing, incorporating and visualizing transcriptomic and proteomic data from hepatocellular carcinoma patients that can be used as a template for other high-throughput techniques.**

105 Madhavan S, Gauba R, Clarke R, Gusev Y. Integrative analysis workflow for untargeted metabolomics in translational research. *Metabolomics* 4(130), 2153–0769 (2014).

106 Kuo T-C, Tian T-F, Tseng YJ. 3Omics: a web-based systems biology tool for analysis, integration and visualization of human transcriptomic, proteomic and metabolomic data. *BMC Syst. Biol.* 7(1), 64 (2013).

107 Gehlenborg N, O'Donoghue SI, Baliga NS *et al.* Visualization of omics data for systems biology. *Nat. Methods* 7(3), S56–S68 (2010).

108 Brazma A, Hingamp P, Quackenbush J *et al.* Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat. Genet.* 29(4), 365–371 (2001).

109 Functional Genomics Data Society. www.fged.org/projects/minseqe/

110 Taylor CF, Paton NW, Lilley KS *et al.* The minimum information about a proteomics experiment (MIAPE). *Nat. Biotechnol.* 25(8), 887–893 (2007).

111 Bino RJ, Hall RD, Fiehn O *et al.* Potential of metabolomics as a functional genomics tool. *Trends Plant Sci.* 9(9), 418–425 (2004).

112 Fiehn O, Robertson D, Griffin J *et al.* The metabolomics standards initiative (MSI). *Metabolomics* 3, 175–178 (2007).

113 Hucka M, Finney A, Sauro HM *et al.* The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19(4), 524–531 (2003).

114 Lloyd CM, Halstead MDB, Nielsen PF. CellML: its future, present and past. *Prog. Biophys. Mol. Biol.* 85, 433–450 (2–3).

115 The CellML project. www.cellml.org

116 Le Novère N, Finney A, Hucka M *et al.* Minimum information requested in the annotation of biochemical models (MIRIAM). *Nat. Biotechnol.* 23(12), 1509–1515 (2005).

117 Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 32, D277–D280 (2004).

118 Shirsat NP, English NJ, Glennon B. Modelling of mammalian cell cultures. In: *Animal Cell Culture.* Al-Rubeai M (Ed.). Springer International Publishing, Switzerland, 259–326 (2015).

119 Gerdtzen ZP. Modeling metabolic networks for mammalian cell systems: general considerations, modeling strategies, and available tools. In: *Genomics and Systems Biology of Mammalian Cell Culture (Volume 127).* Hu WS, Zeng, A-P (Eds.). Springer, Berlin, Germany, 71–108 (2012).

120 Food and Drug Administration. Guidance for industry: clinical pharmacogenomics: premarket evaluation in early-phase clinical studies and recommendations for labeling. www.fda.gov/downloads/drugs

Author Guidelines

Our complete Author Guidelines are available at www.future-science.com

Audience
The audience for Future Science titles consists of research scientists, decision-makers and a range of professionals in the scientific community.

Submission
We accept unsolicited manuscripts. If you are interested in submitting an article, or have any queries regarding article submission, please contact the Commissioning Editor directly. For new article proposals, the Editor will require a brief article outline and working title in the first instance. We also have an active commissioning program whereby the Editor, under the advice of the Editorial Advisory Panel, solicits articles directly for publication.

Peer review & revision
Once the manuscript has been received in-house, it will be peer-reviewed (usually 2–3 weeks). Following peer review, 2 weeks is allowed for any revisions (suggested by the referees/Editor) to be made.

In-house production
Following acceptance of the revised manuscript, it will undergo production in-house. Authors will receive proofs of the article to approve before going to print, and will be asked to sign a copyright transfer form (except in cases where this is not possible, i.e., government employees in some countries).

Article types
For a more detailed description of each article type, please view our author guidelines at: www.future-science.com

Reviews
Reviews aim to highlight recent significant advances in research, ongoing challenges and unmet needs.
Word limit: 4000–8000 words (excluding Abstract, Executive summary, References and Figure/Table legends). Mini-reviews are also accepted.
Required sections: (for a more detailed description of these sections go to www.future-science.com):
- Summary
- Keywords
- Future perspective
- Executive summary
- References: target of 80 maximum
- Reference annotations
- Financial disclosure

Patent reviews
Word limit: 4000–8000 words (excluding Abstract, Executive summary, References and Figure/Table legends).
Patent reviews aim to highlight and critically discuss the most important, promising and recent patents granted/filed in the chosen area. Discussions should be placed within the context of the relevant wider IP landscape, exploring the impact, significance and essential content of the inventions under discussion.

Research articles
Authors of original research must provide a supporting cover letter on submission.
Word limit: 5000–8000 words (excluding Executive summary, References and Figure/Table legends
Required sections: (for a more detailed description of these sections go to www.future-science.com):
- Structured abstract
- Introduction
- Experimental
- Results and discussion
- Conclusions
- Executive summary
- Future perspective
- References
- Reference annotations
- Financial disclosure/acknowledgments
- Defined key terms

Methodology articles
Word limit: 3000–5000 words (excluding Abstract, Executive summary, References and Figure/Table legends).
Methodology articles should provide an overview of a new experimental or computational method, test or procedure. The method described may be either completely novel, or may offer a demonstrable improvement on an existing method. The significance and potential implications of the developments must be explicit.

Preliminary Communications
Word limit: 3000–5000 words (excluding Executive summary, References and Figure/Table legends)
Preliminary Communication articles are intended for short reports of studies that present promising improvements or developments on existing areas of research. The significance and potential implications of the developments must be explicit.

Perspectives
Word limit: 4000–8000 words
Perspectives should be speculative and very forward looking, even visionary. They offer the author the opportunity to present criticism or address controversy. Authors of perspectives are encouraged to be highly opinionated. The intention is very much that these articles should represent a personal perspective. Referees will be briefed to review these articles for quality and relevance of argument only. They will not necessarily be expected to agree with the authors' sentiments.

Special reports
Word limit: 3000–5000 words
Special reports are short review-style articles that summarize a particular niche area, be it a specific technique or therapeutic method.

Editorials/Opinions
Word limit: 1000–1500 words
Editorials are short articles on issues of topical importance. We encourage our editorial writers to express their opinions, giving the author the opportunity to present criticism or address controversy. The intention is very much that the article should offer a personal perspective on a topic of recent interest. Editorials should not contain figures or tables. Maximum 20 references. **Commentaries** (Word limit: 1500–3000 words) are also accepted.

Conference reports
Word limit: 1000–3000 words (excluding References)
Conference reports aim to summarize the most important research presented at a recent conference relevant to the journal's readership. A conference report should not contain figures or tables. Maximum of 20 references.

Technology reviews
Word limit: 4000–8000 words (excluding Executive summary, References and Figure/Table legends)
Technology reviews are short review-style articles that summarize selected instrumentation, techniques or software. The article should clearly highlight the relevance of the products or technology to pharmaceutical bioprocessing and present an objective perspective on the product(s) under discussion.

Letters to the Editor
Word limit: 1500 words (excluding References)
Inclusion of Letters to the Editor in the journal is at the discretion of the Editor. All Letters to the Editor will be sent to the author of the original article, who will have 28 days to provide a response to be published alongside the letter.

Manuscript preparation
Spacing & headings: Please use double line spacing throughout the manuscript. No more than four levels of subheading should be used to divide the text and should be clearly designated.
Abbreviations: Abbreviations should be defined on their first appearance, and in any table and figure footnotes. It is helpful if a separate list is provided of any abbreviations.
Spelling: US-preferred spelling will be used in the final publication.
Figures, tables & boxes: Future Science has a charge for the printing of color figures (i.e., each color page) in the print issue of the journal. We have no page charges and aim to keep our color charge to a minimum. The charge does not apply to the online version of articles, where all figures appear in color at no charge.
Copyright: If a figure, table or box has been published previously (even if you were the author), acknowledge the original source and submit written permission from the copyright holder to reproduce the material where necessary.

As the author of your manuscript, you are responsible for obtaining permissions to use material owned by others. Since the permission-seeking process can be remarkably time-consuming, it is wise to begin writing for permission as soon as possible.

Please send us photocopies of letters or forms granting you permission for the use of copyrighted material so that we can see that any special requirements with regard to wording and placement of credits are fulfilled. Keep the originals for your files. If payment is required for use of the figure, this should be covered by the author.

Key formatting points
Please ensure your paper concurs with the following article format:
Title: concise, not more than 120 characters.
Author(s) names & affiliations: including full name, address, phone & fax numbers and e-mail.
Abstract/Summary: approximately 120 words. No references should be cited in the abstract.
Keywords: approximately 5–10 keywords for the review, including brief definitions.
Body of the article: article content under relevant headings and subheadings.
Conclusion: analysis of the data presented in the review.
Future perspective: a speculative viewpoint on how the field will evolve in 5–10 years time.
Executive summary: bulleted summary points that illustrate the main topics or conclusions made under each of the main headings of the article.
References:
- Primary literature references, and any patents or websites, should be numerically listed in the reference section in the order that they occur in the text.
- Should appear as a number i.e., [1,2] in the text.
- Any references that are cited in figures/tables/boxes that do not appear in the text should also be numerically listed in the reference section in the order that they occur in the text.
- Quote first six authors' names. If there are more than six, then quote first three *et al.*
- The Future Science Endnote style can be downloaded from our website at: www.future-science.com/page/authors.jsp

Reference annotations: please highlight 6–8 references that are of particular significance to the subject and provide a brief (1–2 line) synopsis. Papers should be highlighted as one of the following:
n of interest
nn of considerable interest
Figures/Tables/Boxes: Summary figures/tables/boxes are very useful, and we encourage their use in reviews/perspectives/special reports. The author should include illustrations and tables to condense and illustrate the information they wish to convey. Commentary that augments an article and could be viewed as 'stand-alone' should be included in a separate box. An example would be a summary of a particular trial or trial series, a case study summary or a series of terms explained. Please include scale bars where appropriate.
If any of the figures or tables used in the manuscript requires permission from the original publisher, it is the author's responsibility to obtain this. Figures must be in an editable format.