Biotechnology Advances 79 (2025) 108506

Contents lists available at ScienceDirect

Biotechnology Advances

journal homepage: www.elsevier.com/locate/biotechadv

HEK-omics: The promise of omics to optimize HEK293 for recombinant adeno-associated virus (rAAV) gene therapy manufacturing

Sai Guna Ranjan Gurazada [a], Hannah M. Kennedy [b], Richard D. Braatz [c], Steven J. Mehrman [d], Shawn W. Polson [a,*], Irene T. Rombel [b,*,1]

[a] Center for Bioinformatics and Computational Biology, Department of Computer and Information Sciences, University of Delaware, Newark, DE, United States
[b] BioCurie, Inc., Wilmington, DE, United States
[c] Massachusetts Institute of Technology, Cambridge, MA, United States
[d] Johnson & Johnson, J&J Innovative Medicine, Spring House, PA, United States

ARTICLE INFO

Keywords:
HEK293
Adeno-associated virus
Omics
Biomanufacturing
Gene therapy
rAAV

ABSTRACT

Gene therapy is poised to transition from niche to mainstream medicine, with recombinant adeno-associated virus (rAAV) as the vector of choice. However, robust, scalable, industrialized production is required to meet demand and provide affordable patient access, which has not yet materialized. Closing the chasm between demand and supply requires innovation in biomanufacturing to achieve the essential step change in rAAV product yield and quality. Omics provides a rich source of mechanistic knowledge that can be applied to HEK293, the most commonly used cell line for rAAV production. In this review, the findings from a growing number of diverse studies that apply genomics, epigenomics, transcriptomics, proteomics, and metabolomics to HEK293 bioproduction are explored. Learnings from CHO-omics, application of omics approaches to improve CHO bioproduction, provide a framework to explore the potential of "HEK-omics" as a multi-omics-informed approach providing actionable mechanistic insights for improved transient and stable production of rAAV and other recombinant products in HEK293.

1. Introduction

Gene therapies have the unique ability to prevent, treat, or cure many of the ~7000 rare hereditary diseases that afflict 400 million people worldwide (Mendell et al., 2021; Wang et al., 2019); (Rare Diseases, 2022). Recombinant adeno-associated virus (rAAV) is a potent tool for the delivery of genes into mammalian cells for therapeutic benefit (Chancellor et al., 2023; Kuzmin et al., 2021; Wang et al., 2024), and has emerged as the most common vector for in vivo gene therapy, as evidenced by a growing list of FDA-approved products (Center for Biologics Evaluation, Research, 2023) and more than 250 candidates currently in the clinic (Burdett and Nuseibeh, 2023; Regenerative Medicine: Disrupting the Status Quo, 2022).

While the clinical and commercial opportunity is significant, rAAVs are challenging to manufacture because of their high complexity, resulting in exorbitant costs, long lead times, quality control challenges, and vector shortages that contribute to clinical trial and commercial delays (Dobrowsky et al., 2021; Escandell et al., 2022; High-dose AAV gene therapy deaths, 2020; Mullard, 2021; Sha et al., 2021; Srivastava et al., 2021). There is an acute need for scalable, robust, and cost-effective processes compatible with industrial-scale production of gene therapies, reminiscent of a comparable need faced by the biopharmaceutical industry in the early days of therapeutic antibody production (Li et al., 2010).

Human embryonic kidney (HEK) 293 cell lines are versatile systems used for the production of recombinant proteins, viral and virus-like particles. With advantages that include a high transfection rate, fast growth rate, and human post-translational modification, HEK293 is the most prevalent host cell for the manufacture of rAAV vectors (Abaandou et al., 2021a; Dobrowsky et al., 2021; Naso et al., 2017; Srivastava et al., 2021). The original parental immortalized cell line was created in 1973 by the integration of a 4 kbp adenovirus serotype 5 (Ad5) genome fragment encoding E1A and E1B genes into cells derived from the kidney of an aborted human embryo (Graham et al., 1977). In addition to enabling the continuous culturing of HEK293 (Berk, 2005), E1A and E1B gene products are necessary helper proteins for rAAV production (Qiao

* Corresponding authors.
E-mail addresses: polson@udel.edu (S.W. Polson), irene@biocurie.com (I.T. Rombel).
[1] Irene Rombel will handle correspondence for the article at all stages of the refereeing and publication process, and post-publication.

https://doi.org/10.1016/j.biotechadv.2024.108506
Received 7 August 2024; Received in revised form 14 November 2024; Accepted 15 December 2024
Available online 19 December 2024

et al., 2002). Since its creation, many derivative cell lines with a range of advantages over the parental HEK293 have been developed (Zhang et al., 2024).

HEK293 cell lines are second only to Chinese hamster ovary (CHO) cells for their use in biopharmaceutical production, and are the prevailing system for small-scale protein and viral vector propagation (Lin et al., 2014; Tan et al., 2021). As with the development of CHO to meet the demands of therapeutic antibody production, improving HEK293 host cell productivity through innovative engineering strategies will be critical to meet the growing clinical and commercial need for safe, potent, and cost-effective rAAV products (Escandell et al., 2022; Sha et al., 2021). Although the upper theoretical limit for rAAV production in HEK293 is not known, comparison with monoclonal antibody production in CHO cells implies a manufacturing productivity gap in the order of 1000 fold (Wright, 2023).

Beyond process enhancements to improve product quality, efficiency, and scale – and stable producer cell lines to simplify the ubiquitous transient triple transfection step (Dobrowsky et al., 2021; Naso et al., 2017; Srivastava et al., 2021) – mechanistic models have been developed and applied to significantly enhance processes within this step, further improving product quality and yield (Canova et al., 2023; Destro et al., 2023; Nguyen et al., 2021). The host cells are a critical raw material that governs overall productivity, and much attention is needed to ensure optimal performance.

Systems biology provides an approach to identify and alleviate bottlenecks, specifically those within the cellular machinery involved in rAAV production. Individual omics technologies such as genomics, epigenomics, transcriptomics, proteomics, and metabolomics can provide valuable insight into the optimal environment for the cells and cellular mechanisms involved in rAAV production by HEK293 cells. In concert, they provide a synergistic multi-omic systems biology approach that could enable significant gains through data-driven engineering or modeling. The CHO field provides a precedent to build on: CHO-omics (Dahodwala and Sharfstein, 2017; Kildegaard et al., 2013; Lin et al., 2020; Sorourian et al., 2024; Stolfa et al., 2018) which enabled the mature development of this cell system to meet the needs of many commercial products. We review the field's current state in respect to omics applications and identify gaps and future prospects which support "HEK-omics" as a multi-omics-informed approach to HEK293 pharmacological modulation or engineering for improved production of rAAV and other recombinant products.

2. Genomics

Genomics provides unique insight into the genetic makeup of prokaryotes, eukaryotes, and viruses (Houldcroft et al., 2017; Konstantinidis and Tiedje, 2005; Sharma et al., 2018). Cellular activity is constrained by contents and organization of the genome, representing the baseline for what can be accomplished within a cell to make it a suitable host for use in biotherapeutics (Wuest et al., 2012). Historically, assessment of a cell's genome has been difficult. Over the past two decades, high-throughput DNA sequencing has evolved through multiple generations of technological advancements driving commensurate decreases in sequencing cost which have facilitated broad access to genomics approaches including de novo genome sequencing and genome resequencing, and led to development of approaches enabling exploration of the transcriptome and epigenome (Giani et al., 2020; Satam et al., 2023). Consequently, application of genomics has expanded from basic research to a multiplicity of industrial and clinical applications, from disease diagnosis to the design and implementation of precision medicines (Denny and Collins, 2021; Green et al., 2020; Lähnemann et al., 2020). In this section, we review how genomics has been applied to HEK293 cells in the context of biotherapeutic production.

2.1. De novo genome assemblies provide a foundation for HEK-omics

DNA sequencing has been used to establish reference genome assemblies that are foundational to unraveling the molecular characteristics of any organism. For HEK293 cells, the sequencing of the human genome (Venter et al., 2001) and subsequent improvements to this reference assembly (Singh et al., 2022) have paved the way for HEK-omics research. The current human reference genome version, GRCh38 (aka hg38), established by the Genome Research Consortium (GRC) in 2017, contains 2.92 Gbp of sequence composed of 24 chromosomes including chromosomes X and Y, and contains 60,090 identified genes including 19,890 protein-encoding genes (Nurk et al., 2022; Schneider et al., 2017). In 2022, the Telomere-to-Telomere (T2T) consortium published the first "completely assembled", gapless human reference genome version T2T-CHM13, for all 22 human autosomes and chromosome X, comprising 3.05 Gbp of nuclear DNA, plus a 16.5 kbp mitochondrial genome. The draft T2T-CHM13 annotation describes 63,494 genes, of which 19,969 genes are predicted to encode proteins (Nurk et al., 2022). While either version can be used as a reference genome for HEK293 sequence analysis, the GRC reference series is more widely adopted (Chung et al., 2023; Malm et al., 2022; Malm et al., 2020; Saghaleyni et al., 2022), primarily due to its well-established suite of ancillary databases and bioinformatics resources (Frankish et al., 2021; Lee et al., 2020; Yates et al., 2020).

2.2. Genome resequencing to characterize and adapt HEK293 for production

Genome resequencing helps elucidate the evolution of an organism or cell line by characterizing the genetic changes accumulated over time due to genetic drift or adverse selective pressure. The genetic changes are typically characterized in the form of single nucleotide variants (SNVs) and polymorphisms (SNPs), insertion/deletion variants (InDels), structural variations (SVs), copy number variations (CNVs), and chromosomal recombinations.

Using whole genome resequencing, five widely used rAAV-producing HEK293 cell lines (293T, 293S, 293SG, 293SGGD, 293FTM) have been characterized, by comparing genetic variations of the parental HEK293 cell line relative to the human reference genome (Lin et al., 2014). CNV analysis identified key genomic regions driving HEK293 cell adaptation. The high copy number of the adenovirus insertion site, potentially coupled with low copy number of the fumarate hydratase (FH) gene locus, appears to be a major factor in the transformed phenotype. While these CNVs help explain diverse phenotypes across HEK293 subclones, further experimentation is needed to confirm their roles in suspension growth adaptation, cellular transformation, and metabolism.

Many production cell lines, including CHO and HEK293, were initially developed as adherent cell lines with surface attachment requirements. However, adherent cell culture systems are a challenge for scale-up to meet the high demand required for commercial supply, motivating the adaptation of adherent cell lines to suspension cell lines (Chu and Robinson, 2001). As part of a larger multi-omics study, genomics has also been used to analyze a set of HEK293 adherent (293, 293E, 293T) and suspension (293H, 293F, freestyle 293F) cell culture lines used in recombinant protein production (Malm et al., 2020) (see Sections 4, 6, and 7 for further findings on this study). Hierarchical clustering of the genomic data revealed (1) a significant divergence of all the derivative cell lines from the parental HEK293 cell line and (2) that the adherent cell lines clustered together a single group that was distinct from the group of suspension cell lines. Pairwise comparison of SNVs and InDels between the HEK293 progeny cell lines and the parental line revealed that two adherent cell lines, HEK293E and 293T, had the highest number of high- and moderate-impact SNVs, the type of variants most likely to disrupt protein function. Five genes in particular (CYFIP2, C9orf43, PPP2R4, SGCD, and CTB-47B11.3) had acquired high-impact SNVs across all the progeny cell lines, and could be potential

candidates for targeted cell line engineering. CNV analysis identified similar patterns of copy number gain or loss across the progeny cell lines, with an ~15 Mb region on chromosome 13 consistently showing copy number gain in all progeny cell lines. Understanding these fundamental genetic traits characteristic of the producing cell lines, while distinguishing the adherent lines from the suspension cell lines, is a critical first step in helping explain the phenotypic and physiological differences between the commonly used rAAV-producing HEK293 cell lines.

The aforementioned studies highlighted a number of opportunities to enhance suspended HEK293 cells. The suspension cell group displayed behavior directing enhancement of adherent properties (cellular adhesion), presumably inherited from progeny cell lines. Redirecting this effort can lead to benefits of overall efficiency and promote better cellular suspension properties. For example, it was found that cholesterol content and distribution strongly relate to secretory mechanism optimization, facilitating Golgi transport (Malm et al., 2020). By experimentally establishing a clear association of genomic markers to desirable phenotypic traits, such studies hold the promise for development of improved transient and stable production in HEK293. Recent work in CHO is illustrative of this potential, such as the identification of favored DNA repair enzyme genotypes through genomics meta-analysis of production cell lines that yielded increases in productivity upon genotypic restoration (Spahn et al., 2022). Reduction in product titers over time has been related to the loss of recombinant gene copies integrated into their genomes for various CHO cell lines, including D8H3 and D7H12 (Torres et al., 2023). The larger reductions in titer over time for the D7H12 clones was largely, although not entirely, explained by the greater loss of gene copy numbers. The advent of CRISPR-Cas9 has triggered a revolution in cell line engineering, including for production cell lines. CRISPR technology has been applied to CHO cell line engineering for a variety of key pathways and target genes (e.g., secretion, apoptosis), showing many benefits including enhanced productivity and cell growth, improved glycosylation, and elimination of problematic host cell proteins and adventitious agents (reviewed in (Glinšek et al., 2023)). The application of CRISPR-based genetic screens to enhance production is further discussed in Section 2.5.

2.3. Cell line instability

Cell line instability, which refers to unintended and unpredicted genetic, epigenetic, and phenotypic changes over time, can reduce product quantity and quality, and increase production costs (Joo et al., 2023). Genome resequencing can be used across cell line derivatives, as well as within the same cell line, to determine the effects of cell procedures on the genome. There has always been a concern whether standard cell cultivation procedures such as cell banking and cultivation introduce genomic instability, which can, in turn, cause phenotypic instability (Borsi et al., 2023; Wurm and Wurm, 2017). CHO, for example, has a propensity for genomic rearrangement that, in combination with the genomic and metabolic demand of high-producing cells, manifests as a decline in productivity and product quality (Dahodwala and Lee, 2019). Spahn and colleagues deduced that compromised double-strand breaks (DSB) repair capacity is a major contributor to CHO genome instability, and showed that restoring key DSB repair genes could improve the genome stability of the CHO cell line (Spahn et al., 2022). HEK293, in contrast, does not appear to share this degree of genomic instability. Upon comparing the HEK293T cell line (genome with highest number of SVs) to the genomic samples of the same cell line after multiple rounds of cultivation, banking, freezing, and thawing, the genome of these cells was observed to be stable throughout standard cell culturing conditions, as the samples were very similar in terms of gene copy numbers, point mutations, and SVs (Lin et al., 2014). The HEK293S line was an exception, as dramatic copy number variations were detected upon thawing. Nevertheless, genomic instability has been observed in HEK293 cell lines under selective conditions as evidenced by larger chromosomal rearrangements (for example, an 800 kb deletion was detected by Lin et al. (2014) in their analysis of SVs) rather than point mutations in the ethyl methanesulfonate-mutagenized line, HEK293SG.

Another potential consequence of genomic instability is the introduction of point mutations within promoters and enhancers, such as to abrogate transcription factor binding, or rearrangements that influence transcriptional regulation of specific HEK293 host cell genes or even the gene of interest (GOI). While expression of the GOI in HEK293 cells is highly undesirable, since it diverts resources away from rAAV synthesis, the impact of host gene mutations on rAAV production depends on the affected gene(s) (see Section 7.2). Systematic genome-wide analysis to assess the impact of these phenomena in expression cell lines has not been conducted in HEK293. However, these complex regulatory mechanisms have been studied in CHO cells using a combined genomic and epigenomic approach to assess the response to evolutionary pressure over time (Feichtinger et al., 2016) (see Section 3.3), and in cancer cell lines, where multi-omics has shed light on complex regulatory mechanisms (Ghaffari et al., 2021). These integrated approaches can be adapted to examine the genomic characteristics of HEK293 cell lines and their evolution under selective pressure providing insights for rational engineering of high-producing cell lines that are genetically stable (see Section 7).

2.4. Hot spots and safe harbor regions

When establishing a stable production cell line, not all genomic loci are equal in their capacity to facilitate and stably maintain high levels of transgene expression. Production instability is often caused by random integration of transgenes into locations within the host cell genome that are vulnerable to genetic, epigenetic, and phenotypic instability (Kim et al., 2011; Osterlehner et al., 2011; Wippermann and Noll, 2017; Wurm, 2004). The phenomenon of genomic *hot spots* is ubiquitous as cell biology processes including meiotic recombination, epigenetic modifications, viral integration, chromatin structure, and transcriptional capacity show a biased, non-random distribution across the genome (Cheng et al., 2016). Transcriptional hot spots supportive of high-level, stable transgene expression – also known as *safe harbor regions* – typically reside within euchromatin. In contrast, heterochromatin regions are often silenced by histone deacetylation, histone methylation, and promoter methylation, making them undesirable 'dead spots' to avoid. (See Sections 3.2 and 3.3 for further discussion on epigenomics.)

While the presence of hot spot and safe harbor regions in HEK293 is notionally akin to other cell lines such as CHO, their application is fundamentally different when it comes to the gene of interest (GOI) encoding a therapeutic protein production. If the product is a recombinant protein, then inserting the GOI in a safe harbor region is desirable for improving stable production in the host cell. However, if the product is rAAV encoding a therapeutic gene product, the expression of the GOI in HEK293 is undesirable since it diverts cellular resources away from useful synthesis (e.g., capsid proteins; see Section 7.2). On the other hand, it could be beneficial to have critical rAAV non-therapeutic transgenes (or even HEK293 genes) within safe harbor regions.

Transcription hot spots can be identified by random transgene integration or using side-directed techniques. Using an empirical approach, three loci – AAVS1, CCR5, and ROSA26 – were identified in separate studies as safe harbor regions for stable transgene expression in HEK293 (Shin et al., 2020). Of these loci, AAVS1 showed the highest homogeneity and expression of enhanced green fluorescent protein (EGFP), and enabled faster generation of recombinant cell lines. These safe harbor regions could be suitable loci for expressing genes encoding recombinant proteins, or AAV genes that are typically used to construct a packaging cell line (*Rep*, *Cap*). However, they are not appropriate when inserting a transgene that encodes a therapeutic GOI, which should ideally be silent during rAAV production (see Sections 2.3 and 7.2). More recently, a cell-type agnostic bioinformatic screen was used to

rationally identify two safe harbor regions in HEK293 – Rogi1 and Rogi2 – capable of safe and durable GOI expression (Aznauryan et al., 2022). While the purpose of the study was ostensibly to identify safe harbor sites for in vivo and ex vivo gene therapy, the findings could apply to producer cell line development. Arguably, the most systematic analysis to identify safe harbor regions for the purpose of improving stable high-producing cell line development has been in CHO. The epigenomes and transcriptomes of an industrially relevant monoclonal antibody-producing CHO cell line and its parental CHO-K1 host have been screened (Hilliard and Lee, 2021). Using the Hi-C technique for chromatin structure analysis coupled with RNA-seq for genome-wide gene expression analysis, the study revealed that 10.9 % of the CHO genome was marked by transcriptionally permissive chromatin structures with enhanced stability relative to the rest of the genome (technologies described in Sections 3.1 and 4.1). In a subsequent study, transgene expression from thousands of stable hot spots in CHO was characterized using the thousands of reporters integrated in parallel (TRIP) high-throughput screening method, creating a compendium of stable hot spots for CHO cell line development (Hilliard and Lee, 2021). While having a compendium of integration spots to avoid (aka "dead spots") would be useful, such a compendium is not yet available for HEK293; building such a compendium would be a useful avenue for future research.

Applying these and other systematic omics-based approaches to HEK293 could dramatically reduce the search space for safe harbor regions that are optimal for commercial cell line development, both stable and transient.

2.5. Potential of CRISPR-based genetic screens for increased production

In addition to using genomics approaches for improving rAAV production, alternative avenues such as CRISPR-based genetic screens have emerged as an effective tool for implementing large-scale target knockout screens with high efficiency and reduced off-target consequences (Abaandou et al., 2021a). For example, a CRISPR-based genome-wide screening strategy has been used to identify target genes whose modulation increased the production of rAAV vectors by 3.8-fold in HEK293 cells (Barnes et al., 2021). Genes including spindle and kinetochore associated complex subunit 2 (SKA2) and inositol 1, 4, 5-trisphosphate receptor interacting protein (ITPRIP) were identified for their potential involvement in rAAV packaging and consequent impact on rAAV manufacturing. In another example concerning CHO-based biopharmaceutical production, a large-scale CRISPR-Cas9 knockout screen was used to target ~2500 metabolic enzymes and regulators involved in CHO cell metabolism (la Karottki et al., 2021). The comprehensive pooled CRISPR screen helped narrow down to a critical glutamine response network, and revealed that deletion of one gene for a poorly characterized lipase, Abhd11, significantly increased growth in glutamine-absent media by modifying regulation of the tricarboxylic acid (TCA) cycle. Hence, the application of CRISPR-based genetic screens with complementary omics approaches could potentially offer a powerful strategy for enhancing biotherapeutic production by elucidating key genetic determinants and regulatory pathways.

3. Epigenomics

Epigenomics is the study of *epi*genetic changes in the form of chemical modifications to a cell's genomic DNA and/or chromatin that have the ability to regulate gene activity. Derived from the latin term "epi", which means "on top of", epigenomics is genomics that goes beyond the primary nucleotide sequence. Epigenetic changes can be inherited from parental cells, or acquired as a result of direct environmental exposure to nutritional, chemical, and physical factors, and can influence phenotypic variability independent of genotypic variability (Cavalli and Heard, 2019; Petronis, 2010). In contrast to genomics, which is typically defined by a fixed primary DNA sequence, epigenomics is a reversible process that affects gene expression without altering the underlying base pair sequence. The best characterized epigenetic changes include DNA base modifications (e.g., methylation of cytosine residues primarily at CpG sites) and chromatin remodeling due to modification (e.g., acetylation) of histone proteins (Allis and Jenuwein, 2016; Jeltsch et al., 2018; Zhang et al., 2020). MicroRNAs (miRNA) and other non-coding RNAs (ncRNAs), while not explicitly epigenetic mechanisms, drive similar types of effects through their post-transcriptional regulation activities (Dexheimer and Cochella, 2020; Vito and Smales, 2018). In this section, we review evidence of epigenetic and miRNA activity in HEK293 cells and potential effect on the production of recombinant proteins. Harnessing and controlling this behavior is key to building scalable and high-producing cells to deliver optimal processes to support clinical trials and commercial supply.

3.1. Epigenomics analysis techniques

Similar to other omics approaches, several high-throughput molecular techniques have been developed for epigenomic analysis (Rivera and Ren, 2013; Turunen et al., 2018). The most commonly used techniques to characterize each epigenetic modification are genome-wide DNA methylation assays such as bisulfite sequencing (BS-seq); histone modification assays such as chromatin immunoprecipitation followed by sequencing (ChIP-seq); chromatin accessibility assays such as micrococcal nuclease digestion with deep sequencing (MNase-seq), DNase I hyper-sensitive sites sequencing (DNase-seq), formaldehyde-assisted identification of regulatory elements followed by sequencing (FAIRE-seq) or assay for transposase-accessible chromatin using sequencing (ATAC-seq); and finally chromatin interaction assays such as Hi-C, which is the genome-wide version of the chromatin conformation capture (3C) method (Mehrmohamadi et al., 2021). Long read sequencing platforms (e.g. Pacific Biosciences SMRT, Oxford Nanopore) have capabilities for direct detection of modified nucleotides (e.g. 5mC) without need for bisulfite or other chemical conversions, with recent platform and software advances making these methods increasingly feasible in eukaryotic hosts (Lucas and Novoa, 2023; Warburton and Sebra, 2023; Xu and Seki, 2020). More recently, epigenomic studies at the level of individual cells, rather than bulk populations, have also been possible through respective single-cell (sc) epigenomic assays such as scBS-seq, scChIP-seq, scATAC-seq, and scHi-C (Mehrmohamadi et al., 2021; Wang and Chang, 2018).

3.2. DNA methylation reveals potential epigenetic targets

In a study to assess the impact of genome-wide epigenetic modifications in HEK293 cells, DNA methylation was introduced at several thousand CpG islands in a single experiment (Broche et al., 2021). These epigenetic modifications were shown globally to strongly correlate with a decrease in gene expression, suggesting direct impact. Transient reduction in methylation was also observed globally for H3K4me3 and H3K27ac, which are important histone tags and considered as transcription activation epigenetic biomarkers (Igolkina et al., 2019). Notably, H3K4me3 is associated with a plethora of histone modifications, and genes marked with the associated H3K4me3 domain are involved in cell identity and essential cell functions (Beacon et al., 2021). Likewise, H3K27ac is an important marker that can distinguish between active and inactive/poised enhancer elements (Creyghton et al., 2010).

As a corollary, hypomethylation of host cell DNA by treatment with 5-Aza-2′-deoxycytodine, a hypomethylation agent, was shown to have a positive influence on wild-type AAV (wtAAV) integration and rAAV transgene expression (Chanda et al., 2017). Although the study was performed in HeLa cells with the goal of assessing the potential impact of in vivo epigenetic events in the human host (i.e., the patient), this study also sheds light on the impact of cellular methylation in human-derived cell lines including HEK293. It is evident from these and other studies

that epigenetic targets present a viable and as yet untapped opportunity for HEK293 engineering to improve rAAV and other recombinant protein production (see Section 2.4 for additional examples).

3.3. Production stability assessed by epigenomics

Phenotypic stability, particularly production stability, is a prerequisite for biopharmaceutical manufacturing. However, unexpected and uncontrolled phenotypic variation remains a major challenge, and is not always explained by genetic variability alone (Moritz et al., 2016; Moritz et al., 2015; Osterlehner et al., 2011; Paredes et al., 2013; Patel et al., 2018), implying a potential role of epigenetic variability. Understanding the role of epigenetic variations in a production cell line (and processes therein) can inform strategies to improve stability through epigenetic regulation (Marx et al., 2022). Indeed, DNA and histone modifications have been demonstrated to be contributors to expression levels and evolution of instability in successive passage of CHO cells (Feichtinger et al., 2016; Huhn et al., 2022). Such epigenomic data have led to enactment of interventions to enhance production, including the use of ChIP-seq to inform CRISPR/CAS9 site-specific integrations to identify epigenetically favorable sites for stable production (Hertel et al., 2022; Hertel and Neuss, 2024). Additionally, epigenetic regulation can be responsible for the adaptation of mammalian cells to growth or media conditions that are more conducive to industrial-scale manufacturing, such as transition from adherent to suspension culture or serum-replete to serum-free media (Malm et al., 2020; Mcallister et al., 2002). To investigate these aspects, DNA methylation, histone modification, and genomic data from six related CHO cell lines, subjected to either long-term cultivation (maintained in culture for 3 to 6 months) or different evolutionary pressures (suspension growth, glutamine-free medium), were collected and the genomic and epigenomic modifications were analyzed (Feichtinger et al., 2016). Bisulfite sequencing was used to estimate genome-wide methylation profiles, and ChIP-seq was used to evaluate chromatin states and modifications. Only minor changes in DNA methylation patterns were observed when cells were cultivated under the same conditions over prolonged periods; however, methylation patterns correlated with changes in culture conditions such as a glutamine-free medium or culture methods for inducing adherent to suspension phenotype switch. This finding suggests that epigenetic regulatory mechanisms are involved in helping cells overcome evolutionary pressures, and thus could be used as potential targets for epigenetic editing and cellular engineering to engineer robust and highly productive host cell lines and processes.

3.4. Role of microRNA

MicroRNAs (miRNAs) are small, non-coding RNA molecules that play a crucial role in post-transcriptional gene expression regulation. Understanding the effect of miRNA interactions with its gene targets on HEK293 cell growth and rAAV production can illuminate potential modifications for robust rAAV manufacturing. A series of studies by Shiloach and colleagues (1) used a high-throughput miRNA screen to identify five miRNA's (hsa-miR-22-5p, hsa-miR-18a- 5p, hsa-miR-22-3p, hsa-miR-429, hsa-miR-2110) that could increase expression of the gene of interest (GOI) in HEK293; (2) showed that knocking out *HIPK1*, a target of miRNA-22-3b, increased luciferase reporter and secreted glypican-3 (GPC3) protein expression by 3.3-fold and 2.2-fold, respectively (Abaandou et al., 2021a; Inwood et al., 2020; Inwood et al., 2018; Xiao et al., 2015); and (3) created a stable cell line that over-expressed mi-22, resulting in increased recombinant protein expression by 2.4-fold, while knocking out the HIPK1 gene improved production by 4.7-fold.

In another study using miRNA microarray expression profiling, identification of differentially expressed miRNA during the exponential and stationary phases revealed miR-16 and let-7b as potentially useful targets for improving growth of HEK293 batch cultures (Koh et al., 2009). The miRNAs miR-16 and let-7b were identified to downregulate CCNE1 and CDK6, both being essential proteins for promoting the transition from the G1 to S phase. The authors postulate that, if miR-16 and let-7 were to be inhibited, the cells would divide indefinitely, resulting in higher HEK293 cell densities. A holistic view of the human miRNA-mRNA interactions has been provided by a comprehensive analysis of HEK293 and Huh7.5 datasets (Plotnikova et al., 2019). While most expressed mRNAs did not interact with any miRNA, 1 to 2 % of human transcripts interact with nine or more miRNAs each, indicating that the extent of miRNA-based regulation varying widely among the individual genes. The strategy of increasing biotherapeutics production by miRNA regulation has been demonstrated for HEK293 cells. For example, a 240-fold increase in AAV vector yields occurred by expressing an artificial miRNA that silenced toxic transgene expression (Blahetek et al., 2024; Strobel et al., 2015).

3.5. Cell density effect

The transfection efficiency for HEK293 cells have been observed to drop when transfecting at higher cell density (Lavado-García et al., 2022; Lavado-García et al., 2020b; Pérez-Rubio et al., 2024). While omics approaches have been used to gain mechanistic insights into the causes of this cell density effect, all of its specific molecular causes have not yet been elucidated. While experimental results suggest an unknown inhibitory molecule(s) may play a role (Lavado-García et al., 2022; Lavado-García et al., 2020b), epigenomic analyses have not yet been used to gain insight into the cell density effect. Herein lies a potential opportunity to increase productivity for industrial-scale manufacturing of rAAV and other biotherapeutics by investigating the role of epigenetic regulatory mechanisms, in combination with other omics approaches.

4. Transcriptomics

Transcriptomics is a field in which researchers study the complete set of all RNA transcripts produced by the genome and expressed in a given cell, tissue, or organism (Morozova et al., 2009; Wolf, 2013). Transcriptomic analysis has enabled the study of how gene expression changes in an organism under external or internal stresses, thereby giving insight into the organism's biology and gene regulatory networks at play (Lowe et al., 2017). For example, researchers that implement genetic engineering to study and evaluate phenotypic control, might utilize transcriptomics to evaluate the regulatory networks at play for gene expression at the transcriptional, post-transcriptional, translational, and post-translational layers (see Fig. 1A). Here, we review how transcriptomics has enabled a comprehensive view of gene expression dynamics, contributing to our mechanistic understanding of cellular processes and bottlenecks during recombinant proteins and viral vector production.

4.1. Transcriptomic analysis techniques

Transcriptomic studies have leveraged one of two technologies – microarrays or RNA-Seq – to provide simultaneous quantification of expression for a biological sample (Lowe et al., 2017; Wang et al., 2009). Microarrays were the first high-throughput technology to be developed, quantifying expression of a predetermined panel of genes in a sample via hybridization to an array of complementary probes. RNA-Seq provides a high resolution and global assessment of gene expression by the use of high-throughput sequencing of reverse transcribed cDNA to quantify abundance of all transcripts within a sample. With declining sequencing costs over the past decade, RNA-Seq approaches have seen wide adoption and are the method of choice for the vast majority of modern studies (McGettigan, 2013; Perez-Riverol et al., 2019). Recently the application of long read sequencing platforms (PacBio SMRT, Oxford Nanopore) for RNA-Seq, while lower resolution, are enabling characterization of full length transcripts for detection of RNA splicing variants (e.g. isoform-

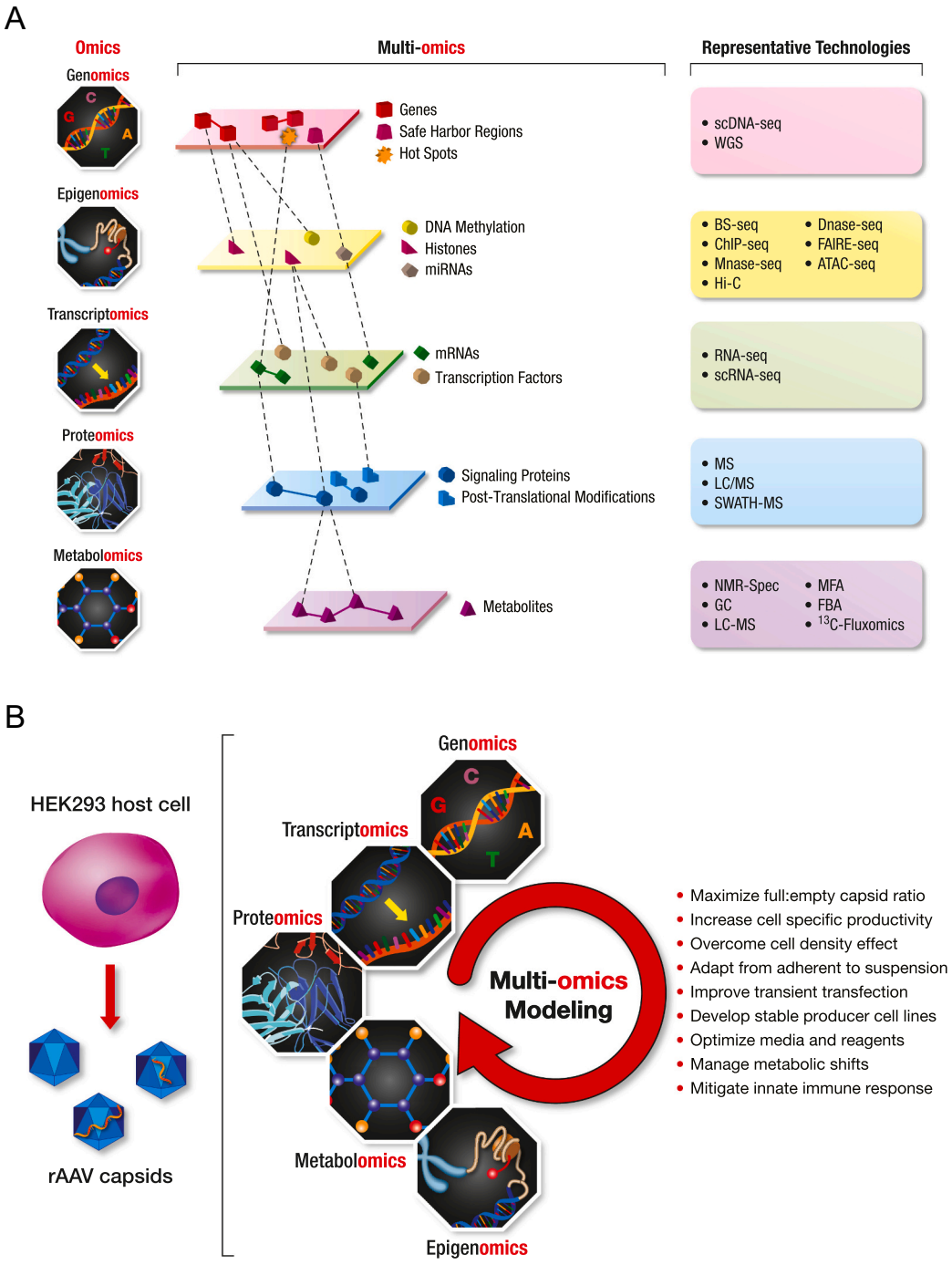

Fig. 1. A. Summary of the multi-omic cross-talk within and across omic layers, the molecular species involved, and technological approaches used to identify and/or quantify them.

scDNA-seq (single-cell DNA sequencing); WGS (whole genome sequencing); BS-seq (bisulfite sequencing); ChIP-seq (chromatin immunoprecipitation followed by sequencing); Mnase-seq (micrococcal nuclease digestion with deep sequencing); Dnase-seq (DNase I hyper-sensitive sites sequencing); FAIRE-seq (formaldehyde-assisted identification of regulatory elements followed by sequencing); ATAC-seq (assay for transposase-accessible chromatin using sequencing); Hi-C (high-throughput chromosome conformation capture); RNA-seq (RNA sequencing); scRNA-seq (single-cell RNA sequencing); MS (mass spectrometry); LC/MS (liquid chromatography coupled with mass spectrometry); SWATH-MS (sequential window acquisition of all theoretical mass spectra); NMR-Spec (nuclear magnetic resonance spectroscopy); GC (gas chromatography); MFA (metabolic flux analysis); FBA (flux balance analysis); 13C-Fluxomics (metabolic flux analysis using 13C-labeled molecules).

B. Potential opportunities to improve rAAV production in HEK293 cells through application of multi-omics technologies.

level transcriptomics), which has traditionally proven difficult with short read sequencing platforms (e.g. Illumina) (De Paoli-Iseppi et al., 2021; Warburton and Sebra, 2023). Such bulk RNA-Seq approaches have been effectively leveraged in HEK293, such as the work of Pistek and colleagues (Pistek et al., 2023), which identified characteristic gene expression patterns in three cell lines with differential rAAV production characteristics to identify potential gene targets for bioengineering. Similar to many bulk RNA-Seq analyses, significant variability was reported across replicates in their experiment. While such variability does not prevent identification of relevant differentially expressed genes, it

may well be serving to obscure additional trends within the data where heterogeneous expression patterns across cellular subpopulations may serve to confound transcriptomics resolution. Recent technological advances, including development of droplet-based microfluidics for cell separation, have enabled the advent of single-cell RNA-Seq (scRNA-Seq). By assessing gene expression patterns from thousands of isolated cells within each sample, scRNA-Seq approaches are able to overcome limitations due to heterogeneous expression within samples by enabling separate analysis of sub-populations with differential expression patterns (Aldridge and Teichmann, 2020; Tang et al., 2019; Tzani et al., 2021). While scRNA-Seq holds significant promise and has seen increasing adoption, it has associated challenges including significantly higher per sample preparation, sequencing, data storage/analysis costs, and evolving bioinformatics approaches (Ke et al., 2022; Lähnemann et al., 2020). While bulk and single-cell RNA-Seq represent the dominant approaches today and the foreseeable future, significant microarray-based studies performed in HEK provide still relevant insights into HEK-omics, and are included alongside those from RNA-Seq approaches in the coming sections.

4.2. Host cell comparisons

Comparative transcriptomic analysis has revealed opportunities, limitations, and bottlenecks within host cell production systems, providing insights on actionable steps to improve productivity. For example, gene expression dynamics within the secretory pathways of transiently expressing CHO and HEK293 host cell systems were compared during the production of 24 difficult-to-express proteins (Malm et al., 2022). Pathway enrichment analysis of differentially expressed genes within CHO and HEK293 revealed that (1) genes involved in protein secretion are more highly expressed in HEK293 than in CHO, (2) protein-folding genes are significantly suppressed in CHO cells while being significantly upregulated in HEK293 cells, and (3) both cell lines showed significant upregulation of genes involved in translation, as evidenced by a spike in ribosomal expression. In some cases, post-translational modification (PTM) machinery was observed to be key to the productivity of recombinant proteins in the two cell lines. For example, the higher titer in HEK293 for more heavily glycosylated products was associated with increased activities of N- and O-glycosyltransferases in HEK293 compared to limitations in the CHO host. Rumachik and colleagues showed that rAAV capsids have PTMs – including glycosylation, methylation, acetylation, and phosphorylation – that differ between production platforms (Rumachik et al., 2020). While the clinical implications of these PTMs are unknown, understanding these dynamics and the host's ability to control PTMs will inform the development of more robust processes and cell lines and potentially safer, more potent gene therapy products with the desired product quality attributes. This is discussed more broadly in Section 8.3 where alternative rAAV production strategies are highlighted.

4.3. Cellular targets for improving production

As discussed in Section 2.2, while HEK293 was initially developed as an adherent cell line, omics has been used to gain insights that can be used to develop suspension cell lines for more scalable production. Transcriptomics has been used for characterization of HEK293 derivatives, providing insights into molecular mechanisms that correlate with the varying phenotypic and morphological profiles of different HEK293 cell lines (Lin et al., 2014; Malm et al., 2020; Saghaleyni et al., 2022). As part of a larger multi-omics study, transcriptomics has been used to compare widely used HEK293 adherent cell lines with suspension cell lines (Malm et al., 2020) (see Sections 2, 6, and 7 for more findings on this study). The differential gene expression and functional enrichment analysis between the two groups revealed that genes involved in cellular compartment organization, cell adhesion, cell differentiation, cell morphogenesis, and cell motility, but not in extracellular matrix organization, were significantly elevated in suspension cell lines. Gene enrichment analysis further pointed to the cholesterol biosynthesis pathway as significantly enriched among the differentially expressed genes between adherent and suspension cells. Compared to adherent cell lines, five key differentially expressed genes – Lysyl Oxidase (LOX), Inhibitor of DNA Binding 1 (ID1), Zic Family Member 1 (ZIC1), Dehydrogenase/Reductase 3 (DHRS3), and Retinoic Acid Receptor Gamma (RARG) – were consistently downregulated in suspension cell lines. These results are in agreement with public transcription data collected from 47 adherent and 16 suspension human cell lines in the Human Protein Atlas database (http://www.proteinatlas.org). If linked to morphological variation between adherent and suspension cell lines, these genes could be targets for cellular engineering or pharmacological modulation.

Five widely used HEK293 cell derivatives (293-T, 293-S, 293-SG, 293-SGGD, and 293-FTM) have been compared to each other and the parental HEK293 cell lines (Lin et al., 2014). Transcriptomics revealed differential gene expression consistent with cell cycle activation and proliferation, potentially due to the faster growth rate of the derivative cell lines compared to the parental HEK293 line. Additionally, most of the 136 genes differentially expressed between derivative cell lines and the parental HEK293 cell line were found to be involved in cell adhesion and motility. This is potentially indicative of the culture characteristics of the parental adherent HEK293 cell line, which is known to be challenging to physically dissociate from culture dishes compared to its derivatives. Discovering such molecular mechanisms and key transcriptomic differences between the non-producing parental HEK293 cell line and the derivative lines could lead to new hypotheses and targets for engineering cell lines to have higher productivity.

In addition to finding new targets, transcriptomics has helped elucidate the underlying molecular mechanisms for validated target(s) that led to productivity improvement in HEK293. For example, when Caspase 8 Associated Protein 2 (CASP8AP2) gene was identified as a potential target for engineering improved expression of recombinant protein in HEK293 from a genome-wide siRNA screen (Xiao et al., 2016), the knockout of CASP8AP2 gene led to a 7-fold improvement in constitutive expression of recombinant protein luciferase and a 2.5-fold increase in transient expression of recombinant protein SEAP in HEK293 (Abaandou et al., 2021b). Transcriptomics was used to compare the CASP8AP2-knockout cell lines to the parental lines to reveal that the cell cycle regulation pathways were the most prominent among the top 25 differentially expressed pathways. This further led to the identification of the Cyclin Dependent Kinase Inhibitor 2 A (CDKN2A) gene and its significant upregulation, as a key Borsifactor for improving recombinant protein expression in HEK293 cells.

4.4. Investigating production instability and cellular heterogeneity using single-cell RNA-seq

Single-cell RNA-seq (scRNA-seq) has emerged as a powerful technique for unraveling production instability caused by cellular heterogeneity in production systems, and the impact of stochastic events. In one example, scRNA-seq provided critical insights into why the levels of mAb produced reduced dramatically in a clonally derived CHO cell line that experienced production instability (Tzani et al., 2021). Analysis of single-cell data uncovered transcriptional heterogeneity and presence of cellular subclusters that correlated with variation in the expression of mAb transgenes, including a reduction in heavy and light chain gene expression. In addition, upregulation of genes involved in oxidative stress response and apoptosis was observed in concert with the trajectory of lower productivity, highlighting the role of cellular stress as a factor linked with gradual loss of productivity. To investigate the presence of cellular heterogeneity in high cell density suspension cultures, scRNA-seq and whole mitochondrial single-cell genome sequencing were used to interrogate transgene-free CHO cells during prolonged suspension culture (Ogata et al., 2021). The study revealed that cellular

transcriptome heterogeneity was cell cycle dependent and increased with increasing culture time, while genomic mutations did not increase during suspension culture. Moreover, the culture time dependent increase in cellular heterogeneity did not show any attenuation in the increasing heterogeneity, revealing that the host CHO cell line has an intrinsic propensity towards heterogeneity, even in the absence of a transgene payload. A full-length coverage scRNA-seq approach using SMART-seq was used in a small study to assess cell-to-cell variability in populations of CHO-K1 suspension cell lines and HEK293 adherent cell lines (Borsi et al., 2023). Perhaps surprisingly, the data indicated that the gene expression patterns of cells cultured under controlled and homogeneous in-vitro conditions were relatively uniform. A small difference in cell-to-cell variability was observed between adherent and suspension cells, potentially due to differences in mixing and local cell density, but the direct comparison of adherent CHO and suspension HEK293 cells makes it difficult to draw definitive conclusions. Nonetheless, this study paves the way for a larger study of HEK293 cell lines under different conditions using more high-throughput techniques such as Drop-Seq or 10× Genomics Single Cell Gene Expression (Hwang et al., 2018; Jovic et al., 2022), as having an accurate view of the degree of stochasticity in a cell population over time can be useful to ensure consistent high productivity and quality during commercial production.

While scRNA-seq has yet been used to interrogate expression during rAAV production in HEK293, it has been applied to a dual baculovirus expression vector system (BEVS), also known as the Sf9 insect cell system (Virgolini et al., 2023b). Heterogeneity within the cell population was observed prior to infection, showing that it is intrinsic to the host cell and is dependent on the cell cycle. Upon infection of the Sf9 cells with rAAV serotype 2 (rAAV2) using a low multiplicity of infection, an increase in cell-to-cell transcriptional heterogeneity was observed that was linked to baculovirus gene expression, as well as to differences in the three transgenes, namely the rAAV vector genes (*rep* and *cap*) and the gene of interest (*gfp*). Notably, at 24 h post-infection, only 29.4 % of cells carried all three essential transgenes to produce a full rAAV capsid. This disconnect in timing of expression of essential genes for rAAV production has been shown in HEK293, where the poorly coordinated timeline between capsid synthesis and Rep protein production results in inefficient production of full capsids (Srinivasan et al., 2024). The scRNA-seq study highlights the value of leveraging this technique to explore potential consequences of heterogeneity specific to rAAV production in HEK293 as well as Sf9, such as the stochasticity of the infection process and differences in infection kinetics, and sheds light on ways to improve rAAV viral vector production through rational cell line design and process engineering.

4.5. Identifying bottlenecks in protein and viral vector production

The impact of the secretion, protein folding, and stress of production in HEK293 transient and expression systems is significant, as evidenced by a growing number of studies in recent years (Zehetner et al., 2024a). To identify bottlenecks in protein secretion for improving recombinant protein production, transcriptomics was used to compare HEK293F stable cell lines expressing either the non-secretory green fluorescent protein (GFP) or the secretory protein erythropoietin (EPO) at varying translational rates (Saghaleyni et al., 2022). EPO-producing cells showed a higher ATP production capacity, indicated by upregulated genes in the oxidative phosphorylation pathway. Additionally, ribosomal gene expression patterns varied based on the type and rate of recombinant protein production. For example, the expression of mitochondrial ribosomal genes was positively correlated to EPO secretion; however, no such evidence was found with GFP production, whereas the expression of cytosolic ribosomal genes was negatively correlated to production rate in both cases. Furthermore, high EPO-producing clones significantly upregulated the ATF6B gene, potentially enhancing the endoplasmic reticulum (ER) stress response to manage high protein secretion. ATF6B is a transcription factor known to be active under ER stress conditions due to accumulation of unfolded proteins. This gene is normally integrated within the ER membrane but enters the nucleus to activate ER stress response genes under stress conditions (Iurlaro and Muñoz-Pinedo, 2016; Thuerauf et al., 2004). Understanding the effects of cargo proteins on gene expression and protein secretion can guide cargo engineering to enhance protein secretion.

Given the central role of the ER in secretion, folding, and stress, applying omics techniques to investigate ER processing pathways and the unfolded protein response (UPR) could reveal new opportunities to enhance productivity. Using transcriptomics to investigate these phenomena in the context of rAAV production, four candidate genes were identified: X-box binding protein 1 (*XPP1)*, which is involved in ER protein translocation and UPR signaling; protein phosphatase 1 regulatory subunit 15 A *(GADD34/PPP1R15A)*, which is involved in overcoming the UPR during protein synthesis; heat shock protein family member 6 (*HSPA6*), which can promote ER protein translocation, folding, secretion, and misfolded protein degradation; and B-cell lymphoma 2 *(BCL2)*, which was previously shown to increase lentiviral vector volumetric productivity by 53 % in HEK293 cells (Wang et al., 2023). All four genes were integrated into HEK293T host cells, which were then transfected using transient triple transfection with rAAV carrying *gfp* as the GOI, resulting in up to 78 % improvement in specific productivity (Fu et al., 2024). Moreover, improvement in viral titer was confirmed in different cell lines (HEK293 and HEK293T) and serotypes (AAV2 and AAV8), albeit at different levels depending on the system. The value of transcriptomics in exploiting metabolic and cellular processes and bottlenecks to improve HEK293-based rAAV production is compelling, and especially powerful when part of a multi-omic strategy (see Section 7).

4.6. Impact of antiviral and innate immune responses

Understanding the cellular response of HEK293 cells to rAAV production, and identifying genes and pathways that influence rAAV production, is pivotal to developing robust manufacturing platforms that provide safe, scalable, and affordable rAAV-based gene therapies. An antiviral response has been shown to affect protein synthesis, cellular metabolism, cell proliferation, and even programmed cell death (Fritsch and Weichhart, 2016), and could be a significant factor in curtailing rAAV production in HEK293 cells. To understand the underlying mechanism, RNA-Seq was used to interrogate the transcriptional response of suspension HEK293 cells to rAAV production following transient transfection, under manufacturing-relevant conditions (Chung et al., 2023). Systematic analysis indicated that viral vector synthesis actively triggers an innate infectious response in the host cells, as evidenced by the upregulation of pathways involving antiviral and inflammatory responses (Kahlig et al., 2024). In an orthogonal study, transcriptomic comparison between rAAV-producing and non-producing HEK293 cultures over time revealed that the antiviral innate immune response is a significant bottleneck in rAAV production. Specifically, the RIG-I-like receptor (RLR) signaling pathway, the Toll-like receptor (TLR) signaling pathway, the cytosolic DNA sensing pathway, and the downstream Janus kinase-signal transducer and activator of transcription (JAK-STAT) signaling pathway were enriched and upregulated in the host cells producing rAAV (Wang et al., 2023). For example, the transcriptome data revealed an overexpression of DDX58, IFIH1, and POLR3C (a subunit of RNA Polymerase III) suggesting the activation of the RLR signaling pathway. Wang and colleagues showed these innate immune responses can be addressed by supplementing the medium with small-molecule inhibitors, or by gene silencing or knockout. J. Fraser Wright postulates that the production efficiency gap between rAAV and monoclonal antibodies could be, in part, explained by the difference in innate antiviral activity in human-derived HEK293 cells and non-human derived CHO cells (Wright, 2023). Hence developing mechanisms to limit these responses in host cells could potentially close the gap and help improve rAAV production, and are further

discussed in Section 7.

4.7. Identification of epigenetic and cell cycle regulators and druggable targets

To gain a better mechanistic understanding of key pathways during rAAV production, a transcriptomic analysis combined with pharmacological perturbation was applied to five HEK293 cell lines with variable capacities for rAAV production (Tworig et al., 2024). An RNA-Seq time-course was used to identify genes that are both differentially expressed between base and high rAAV producers, and modulated during rAAV production. This revealed core cellular functions that are upregulated (e.g., cellular response, signal transduction, cell cycle regulation), or downregulated (e.g., mitochondrial electron transport, glutathione metabolic modulators, proteasomal proteins), reflecting a complex shift in cell state, with widespread modulation of energy production and biosynthesis. With this knowledge of potential druggable pathways and targets, small-molecule compounds were selectively tested. Up to 19 % of the tested compounds improved rAAV production by 1.5-fold above baseline, which is substantially higher than the typical hit rate of 0.1 to 1 % using an unbiased high-throughput screen. Of the drug classes that significantly enhance rAAV production, those targeting epigenetics, transcriptional activation, and cell cycle modulation were particularly compelling. Notably, two HDAC inhibitors produced titers of more than 4-fold and 9-fold above baseline, underscoring the importance of epigenetics and chromatin remodeling in regulating rAAV expression and productivity in HEK293.

4.8. Back to basics: transcription initiation, elongation, and termination

Beyond the active area of transcriptomics research and its applications to improving rAAV production in HEK293 described above, another area that warrants further exploration is leveraging omics techniques and insights to enhance the process of transcription by mitigating bottlenecks in initiation, elongation, and termination (Lenstra et al., 2016). Transcription factors, which comprise just over 7 % of the human proteome, serve as gatekeepers for these processes, and as such provide opportunities for improving production. For example, the binding of transcription factors (TFs) to their cognate sites within promoters could be modulated by engineering transcription factors to be insensitive to DNA methylation, thereby overcoming gene down-regulation or silencing. Alternatively, or in combination with, the binding of TFs to promoters and enhancers could be regulated through post-translational modification (e.g., phosphorylation, O-GlcNAcylation, acetylation, methylation, sumoylation, ubiquitination) of the TF proteins themselves (Filtz et al., 2014). Beyond characterization studies, the manipulation of gene expression levels in mammalian cell factories using a plethora of natural and synthetic tools, including promoters, transcription factors, and other regulatory elements, was recently reviewed (Eisenhut et al., 2024). While the transcription factor proteome of HEK293 has been investigated for more than 15 years using classical techniques such as 2DE (Jiang et al., 2009), repeating such investigations using the latest proteomic and epigenomic techniques could shine a new light on potential avenues to improve rAAV expression in HEK293.

5. Proteomics

The proteome describes the set of proteins encoded by the genome (Wilkins et al., 1996). Proteomics is the study of the functional proteome that enables identification of expressed proteins at a given time, and sheds light on their function and structure, including spliced protein isoforms, post-translational modifications (PTMs), interaction partners, structural description, and higher order protein complexes (Tyers and Mann, 2003). While genomics provides the blueprint of all possible gene products, proteomics builds on this foundation to provide the functional context for the translated gene products. Proteomics is complementary to genomics and transcriptomics, and together can be used to model cellular behavior at the whole system level, and are key tools in the study and optimization of systems biology. In this section, we review how proteomics can be used for enhancing host cell productivity through expression profiling, PTM discovery, protein activity, and analysis of cell composition and organelle composition.

5.1. Techniques for proteomic analysis

Tools for protein identification and characterization in mammalian cells (Farrell et al., 2014) include (1) antibody-based separation methods such as western blotting (Kurien and Scofield, 2015) and enzyme-linked immunosorbent assay (ELISA) (Lequin, 2005), and (2) gel-based separation methods such as two-dimensional gel electrophoresis (2DE) (Dunn, 1987) and two-dimensional difference in-gel electrophoresis (2D-DIGE) (Marouga et al., 2005). These techniques are considered low throughput by current standards, limited to identifying and characterizing tens to hundreds of proteins. Nevertheless, their foundational contributions were instrumental in early proteomics, revealing useful insights to guide cell development, and they continue to serve as complementary techniques to enhance data generation and analysis (Marcus and Rabilloud, 2020; Tighe et al., 2015). Accordingly, for robust proteomic analysis, a variety of high-throughput techniques based on mass spectrometry (MS) have emerged. These methods allow not only protein identification and characterization, but also relative and absolute protein quantification (Ankney et al., 2018). Among these techniques are gel-free separation methods such as reversed-phase liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) (Yates, 2011). MS-based protein quantification can be achieved using label-based and label-free methods. Label-based quantitation methods make use of stable isotope labels that are incorporated within peptides to create an expected mass difference, whereas label-free methods use signal intensity and spectral counting of peptides for relative and absolute quantitation (Anand et al., 2017). Label-based methods include isobaric mass tags for relative and absolute quantitation (iTRAQ) (Wiese et al., 2007), tandem mass tags (TMT) (Thompson et al., 2003), and stable isotope labeling by amino acids in cell culture (SILAC) (Ong and Mann, 2007). Label-free methods such as spectral counting and ion intensity, are combined with downstream MS to enable quantitation (Anand et al., 2017).

5.2. Structure-function relationships

Beyond protein characterization and quantification, MS-based techniques can be used for structural and biophysical analysis at the proteome scale, providing a holistic view of proteome-wide structural alterations, folding and stability, aggregation, and molecular interactions in cell lysates or intact cells (de Souza and Picotti, 2020). For example, sequential window acquisition of all theoretical mass spectra (SWATH-MS) is a data-independent (DIA) method that combines deep proteomic coverage capabilities with quantitative consistency and accuracy (Ludwig et al., 2018). Using a massively parallel methodology combining size exclusion chromatography (SEC) to fractionate native protein complexes, *sequential window acquisition of all theoretical mass spectra (*SWATH) mass spectrometry to precisely quantify the proteins in each SEC fraction, and a computational framework *CCProfiler*, the HEK293 proteome was interrogated to identify and quantify every protein complex in each cell (Heusel et al., 2019). Of the 4916 proteins identified and quantified, 2668 proteins (66 %) and 55 % of the protein mass were in an assembled state, organized as part of a macromolecular complex. Such complex-centric analysis affords an unprecedented view of the structure-function relationships within a HEK293 cell that could inform strategies to improve productivity. Indeed, SWATH-MS proteomic analysis of transiently transfected rAAV-producing HEK293T cells was used to generate a comprehensive temporal analysis of the cellular

states under a range of conditions involving varying parameters for initial transfection cell density, total plasmid DNA concentration, and plasmid ratios (Patra et al., 2024). The study identified 124 host proteins associated with optimal transfection conditions for rAAV production. By applying Protein-Protein Interaction Functional Analysis, many of these proteins were shown to be associated with pathways involved in viral entry, viral protein processing, modulation, replication, and other key processes. Interestingly, optimal production conditions correlated with downregulation or reduced expression of proteins associated with the innate immune response, providing further evidence that the HEK293 host immune response potentially impacts viral yields during production (see Section 4.6). New proximity-dependent labeling methods that capture transient as well as stable protein-protein interactions (PPIs) in living cells, such as proximity biotinylation, have emerged as a powerful technique that complements MS-based approaches (Samoudi et al., 2020; Schaack et al., 2023). Using proximity-dependent biotin identification (BioID) to characterize the machinery involved in recombinant protein secretion in HEK293, this method identified transient interactions supporting biotherapeutic synthesis and further illuminated mechanistic details of its secretory pathway. Taken together, these studies highlight the value of applying high-throughput techniques such as SWATH-MS and BioID for assessing PPIs to generate a comprehensive and dynamic view of rAAV and recombinant production in HEK293, providing new insights for designing improved processes and cell lines.

5.3. Enhancing host cell productivity of rAAV

Understanding the host cell proteome is necessary for devising engineering strategies that could deliver improvements in the host cell productivity. LC-MS/MS and relative label-free methods have been used to compare transfected versus non-transfected cells for AAV5-producing suspension HEK293 cells (Strasser et al., 2021). Proteins involved in cell organization, cell proliferation, and biogenesis were found to be the most upregulated in AAV5-producing cells, while proteins participating in metabolic processes were significantly downregulated. Strong upregulation was also reported in the Mannosidase Alpha Class 2 A Member 2 (MAN2A2) protein in the cell pellet samples of the transfected HEK293 cells. MAN2A2 is known to be involved in N-glycosylation of proteins, which can critically influence the function and quality of recombinant proteins produced (Kaur, 2021; Rose, 2012). In addition to glycosylation being a critical quality attribute (CQA) of biotherapeutics, expression changes in certain mannosidases could be potential candidates for targeted modulation of certain PTMs in rAAV products. For a better biological interpretation of the proteomics data, gene ontology functional enrichment and pathway analysis was applied to reveal differential expression patterns in endocytosis-lysosomal proteins that led to the hypothesis that blocking endocytosis might improve rAAV production. A follow-up experiment treating the cells with chloroquine, which is known to inhibit endocytosis, increased the viral titer by more than 35 %. This work demonstrated the utility of proteomics combined with other systems biology approaches to gain understanding of the cellular processes and thereafter devise strategies for improving rAAV production. Western blotting and proximity-dependent biotin identification were used to show that downregulation of membrane-associated accessory protein (MAAP) drastically reduced the secretion rate of AAV capsids (Elmore et al., 2021). The increased intracellular retention would increase the percentage of capsids containing the gene of interest, reducing the cost of separating full from empty capsids in downstream operations.

5.4. Enhancing host cell productivity of virus-like particles

ELISA, western blotting, and multiplexed quantitative proteomics via LC-MS/MS have been used to characterize protein profiles and cellular changes in HEK293 cells producing virus-like particles (VLPs) (Lavado-García et al., 2020b). Expression changes in cells grown under three operating conditions were investigated: (1) without transfection, (2) transient transfection of an empty plasmid, and (3) standard transient transfection of plasmid with Gag::EGFP gene. Transfection efficiency was observed to dramatically decrease at cell densities higher than $\sim 3 \times 10^6$ cells/mL, partially explained by a downregulation of intracellular protein transport to the nucleus and lipid biosynthesis. Further, the decrease in cellular viability upon transfection was caused by overall disruption of homeostasis due to multiple levels of homeostatic control being altered, such as calcium regulation, oxidant detoxification, DNA detoxification, glycosphingolipids metabolism, xenobiotic metabolism, and peptidase activity. Finally, during VLP production, while above traits were maintained, specific modifications in the membrane calcium channels and extracellular matrix were observed. Such proteomics-enabled findings could help devise future strategies to reduce bottlenecks and design less stressful processes, thereby increasing cellular productivity. Follow-up work used LC-MS/MS to characterize extracellular vesicles (EVs) co-produced with VLPs (Lavado-García et al., 2020a). EVs were detected in all the three operating conditions as described above; however, during transfection, a shift in EV biogenesis was observed leading to changes in the size distribution of the EV products from large to small. Further functional analysis revealed that production of EVs reflected an overall energy homeostasis disruption via mitochondrial protein deregulation. Understanding the nature of the extracellular environment in VLP production will facilitate design of future VLP-based therapies, as it could lead to new downstream separation strategies or using copurified EVs as delivery vehicles.

6. Metabolomics

Biological processes involving complex interactions between genes, mRNA, proteins, and metabolites, combined with external environmental factors, have a profound influence on the resulting phenotype of the cell or organism. While genes determine what may happen, metabolites define what has happened, and are typically the final output of gene expression (Tan et al., 2016). Metabolites are small molecules serving as intermediates or end-products in cellular metabolic processes (Lamichhane et al., 2018). Metabolic flux is the passage of a metabolite through a pathway over time. A measurable change in a metabolic flux can occur in seconds to minutes, whereas genes and proteins take minutes to hours for measurable changes. Metabolomics and metabolic models can provide actionable insight for engineering cells to achieve optimal cell membranes for secretion of product and minimize sensitivities to suspended cell processing (e.g., shear from mixing, gassing, and aggregation).

6.1. Metabolomic analysis techniques

Metabolomics is broadly defined as the comprehensive measurement of all metabolites in a biological system, both endogenous and exogenous, that is, produced either internally by the system or derived from various external sources such as diet, microbes, or xenobiotic sources (Lamichhane et al., 2018). This approach is complementary to genomics, transcriptomics, and proteomics, and is the most recently developed omics technique that provides a sensitive measure of the phenotype of a biological system (Tan et al., 2016). High-throughput analytical techniques such as nuclear magnetic resonance (NMR) spectroscopy, gas chromatography (GC), and liquid chromatography coupled with mass spectrometry (LC-MS) are used in metabolomics to routinely measure tens to hundreds of metabolites with high accuracy and precision in a specimen, either through a targeted or non-targeted approach (Lamichhane et al., 2018). Fluxomics, on the other hand, describes the various approaches that seek to determine the complete set of metabolic fluxes in a cell or organism modeled through the metabolic network. Fluxomics can be considered a sub-branch of metabolomics that best represents the dynamic picture of the phenotype, resulting

from the interactions of the metabolome genome, transcriptome, proteome, and epigenome (Cortassa et al., 2015; Emwas et al., 2022). Common fluxomics techniques are metabolic flux analysis (MFA), flux balance analysis (FBA), and ^{13}C-fluxomics (Emwas et al., 2022). Fluxomics can be used to understand the current phenotype state and indicative of its next state(s) to help in development of control strategies for the host cells to produce highest titer and quality.

6.2. Measuring the metabolome

In practice, metabolomics presents an analytical challenge because – unlike genomics, transcriptomics, and proteomics – metabolomics aims to measure molecules that have disparate physical and chemical properties. This requires the metabolites to be analyzed as distinct subsets, often grouped based on compound polarity, common functional groups, or structural similarity, using distinct sample procedures for each (Clish, 2015). In addition to the physical properties of the molecules and appropriate sample separation procedures, media composition has a significant influence on the outcome of metabolomic experiments. An examination of the influence of five different media combinations upon the cell secretome metabolite profiles of two phenotypically different cell lines – HEK293 and L6 rat muscle cells – found that the addition of fetal bovine serum (FBS) resulted in the detection of unique metabolites (Daskalaki et al., 2018). Media differences also had an impact on expression profiles, as glutamine and pyroglutamate were found to be more abundant in incubated relative to refrigerated medium. Among all the variables that can affect the secreted metabolite profile analyzed in the study, the choice of media was found to be the most sensitive.

6.3. Metabolic models

Beyond the measurement of metabolites, a recent computational approach to map the entire metabolic information of an organism involved construction and use of genome-scale metabolic models (GEMs) (Bernstein et al., 2021). Similar to the role of a reference genome assembly in the study of genomics, transcriptomics, and proteomics, genome-scale metabolic models provide comprehensive, systems-level representations of cellular metabolic functions. GEMs depict complex cellular networks, integrating genes, enzymes, reactions, and metabolites. These models computationally map gene-protein-reaction associations, enabling flux predictions for diverse systems-level metabolic analyses (Bernstein et al., 2021; Gu et al., 2019; Passi et al., 2021). By providing a framework linking the organism's genome and gene expression to the metabolic phenotype, GEMs can be invoked to simulate cellular changes or perturbations and obtain a readout of the equivalent metabolic responses, as a unique fingerprint for the state of bioprocesses.

The first genome-scale metabolic model for human metabolism, RECON 1, was released in 2007, helping to unlock the complexity of the relationship between metabolites and cellular behavior. At the time, RECON 1 accounted for the functions of 1496 open reading frames (ORFs), 2004 proteins, 1509 unique metabolites, and 3744 metabolic, transport, and exchange reactions (Duarte et al., 2007; Thiele et al., 2013). After successive improvements and integration with Edinburgh Human Metabolic Network (EHMN), HepatoNet1, and other databases (Ryu et al., 2015), the initial version progressed into RECON 2 (Thiele et al., 2013), which included ~2× more reactions and ~ 1.7× more unique metabolites. RECON 3D, the most comprehensive human metabolic model, incorporates 3D metabolite and protein structure data. It encompasses 3288 ORFs (17 % of annotated human genes), 13,543 metabolic reactions, 4140 unique metabolites, and 12,890 protein structures (Brunk et al., 2018).

Complementary to the RECON series, Human Metabolic Reaction (HMR) (Mardinoglu et al., 2014; Mardinoglu et al., 2013) is a well-established generic GEM series for human metabolism, focused on subcellular localization and tissue-specific gene expression (Gu et al., 2019). The HMR series, more comprehensive in fatty acid and lipid metabolism than RECON, resulted from extensive manual curation. It has spawned several cell-type specific GEMs, including iAdipocytes1809 (Mardinoglu et al., 2013), iHepatocytes2322 (Mardinoglu et al., 2014), and iMyocyte2419 (Väremo et al., 2016).

Multiple iterations of a metabolic model for HEK cells were constructed over the past two decades through meticulous manual curation efforts (Henry et al., 2005; Nadeau et al., 2002; Nadeau et al., 2000a; Nadeau et al., 2000b; Zehetner et al., 2023). Tissue- and cell-type specific omics data were integrated with a generic human metabolic model to derive a high-quality, context-specific metabolic model (Ryu et al., 2015). The HEK cell-type specific GEM was derived from a generic human model RECON 2 by devising a heuristic to emphasize flux through core metabolic reactions, and removing redundant pathways through a systematic, multi-step reduction process (Quek et al., 2014). This reduced model comprising only 357 active reactions describes the "functionalization" of RECON 2 for conducting steady-state metabolic flux analysis, and enables its application in investigating the metabolism of HEK cell cultures (Abbate et al., 2020; Martínez-Monge et al., 2019). For example, flux analysis based on a reduced HEK model (comprising 354 reactions and 335 metabolites when reduced from RECON 2.2 version using the same reduction protocol) evaluated the metabolic shift in HEK293 cell cultures from glucose consumption and lactate generation to glucose and lactate co-consumption during exponential growth (Martínez-Monge et al., 2019). This shift resulted in improved, balanced metabolism, with reduced glucose and amino acid uptake rates and minimal impact on cell growth. This insight could lead to new genetic engineering strategies that avoid the lactate accumulation in cell cultures. An alternative reduced HEK metabolic model with 82 reactions and 51 internal metabolites has been used to evaluate the comparative performance of adaptive flux variability analysis (AFVA) and classical flux variability analysis (FVA) in terms of coping with uncertainties arising from measurement noise and data smoothing (Abbate et al., 2020; Abbate et al., 2019). The reduced HEK metabolic network and the biomass synthesis reactions were largely inspired from a metabolic network designed for CHO cells and not HEK cells (Fernandes et al., 2016). The biomass fluxes estimated using AFVA were reported to be more consistent than for FVA. While there is no all-encompassing GEM for HEK293 cells, there have been methods introduced to combine and learn from a range of reduced models (Zehetner et al., 2024a). Regardless, these reduced models capture core metabolic functions and have aided process optimization.

6.4. Metabolomics for optimizing process engineering, culturing conditions, and cell lines

Metabolic flux analysis has proven useful in discovering key metabolic pathways and their potential role in driving productivity improvements in HEK293. The underlying metabolic changes in pyruvate carboxylase 2 (PYC2)-expressing HEK293 cells compared to parental HEK293 cells have been investigated by estimating intracellular fluxes using various ^{13}C-labeled substrates. Insertion of the PYC2 gene in HEK293 cells has led to marked reductions in lactate and ammonia production (which are considered toxic metabolic waste) during batch cultures, in addition to a 2-fold increase in maximum cell density and 33 % increase in the yield of recombinant glycoprotein interferon-α2b production, while maintaining maximum specific growth rate (Henry and Durocher, 2011). Flux analysis was performed to understand these improvements. The analysis indicated significant pyruvate carboxylase activity in the PYC-expressing cells, that was previously missing in the parental cells. Additionally, key flux differences in the PYC-expressing cells came to light, for example, at the pyruvate branch point, that the cells converted 56 % of the pyruvate pool into acetyl-CoA (compared to 51 % in parental cells), 26 % into oxaloacetate via the pyruvate carboxylase (not reported in parental cells), and only 9 % into lactate (compared to 48 % via lactate dehydrogenase in the parental cells). Flux

analysis also indicated an overall reduction in glucose uptake rate in PYC-expressing cells, and found that the amount of glucose channeled through the pentose phosphate pathway (PPP) was reduced, representing 7 % of the glucose uptake rate, compared to 28 % for the parental cells. Such detailed insights can help guide future cellular engineering approaches and design of improved feeding strategies and/or media formulations to maximize host cell productivity and enhanced product quality.

Metabolic flux analysis has also been used to determine HEK293 perfusion culture conditions that enhanced productivity of adenoviral vectors (Henry et al., 2005). Intracellular flux analysis showed that experiments with highest specific product yields displayed increased glycolytic and TCA cycle fluxes, along with enhanced ATP production rates at infection time. A feeding strategy was established that maintained favorable metabolic state throughput during the growth phase and allowed successful infection at a cell density of 5×10^6 cells/mL, resulting in specific productivity close to the maximum achieved while having a 2-fold increase in cell density.

Careful control of amino acids during the growth phase were identified from metabolomic analysis of fed-batch CHO process leading to higher cell densities and increased productivity (Mulukutla et al., 2017). These results were achieved by conducting a systematic study (DOE) of adding known inhibitors at specific concentrations during the cell growth phase and comparing growth kinetics. Analysis led to an optimal feed strategy that minimized lactate and osmolality build up. These studies and analysis are challenging and highlight the difficulty to turn metabolomic insights into definable actionable process improvements, subsequent work by Pereira et al. (2018) and O'Flaherty et al. (2020) have built on this work.

6.5. Lipidomics

Lipidomics, a subfield of metabolomics focused on the lipidome, is relatively understudied and underutilized. Lipids are potentially important regulators of protein production and secretion because of their involvement in energy metabolism, vesicular transport, membrane structure, dynamics, and signaling (van Meer et al., 2008). Combined lipidomic and transcriptomic analysis of HEK293, CHO, and mouse myeloma SP2/0 cell lines revealed that phosphatidylethanolamine (PE) and phosphatidylcholine (PC) are the major lipid components in all three systems (Zhang et al., 2017). HEK293 cells had 4–10 times more lysophosphatidylethanolamine (LPE) and 2–4 times more lysophosphatidylcholine (LPC) than CHO and SP2/0 cells, consistent with the transcription levels of the relevant enzymes. As previously mentioned, indirect lipidomic analysis using transcriptomic and metabolomic gene analysis to pinpoint differences between adherent and suspension HEK293 cell lines, revealed pathway changes related to cholesterol biosynthesis and metabolism (Malm et al., 2020). Transcriptomic analysis of three HEK293 cell lines (high, intermediate, and low producers) showed that cholesterol biosynthesis levels were inversely correlated with rAAV production efficiency, suggesting that high cellular cholesterol levels may be disadvantageous (Pistek et al., 2023). Given that cholesterol is an important component of the cell membrane and is integral to membrane structure, cell signaling, and protein trafficking, its impact on plasmid transfection and rAAV synthesis warrant investigation. In the future, systematic lipidomic analysis could help elucidate the physiological and secretory machinery of HEK293 involved in rAAV production to inform improvements in process design and cell line engineering.

7. Multi-omics

Applications of omics such as genomics, transcriptomics, proteomics, and metabolomics have, independently, revealed valuable insights into the cellular mechanisms of mammalian host cell systems such as CHO and HEK293 for producing recombinant proteins. Single omics approaches, although effective, are overly simplistic and fall short in explaining the complex intracellular and systems biology workings of a cell or organism. Combining different omics data in the form of multi-omics analysis can bring together different biological layers and components in a cell to synthesize a more complete mechanistic understanding. Using multi-omics, insights drawn from genomics can be connected to gene expression differences through transcriptomics, for example, and, in other cases, help understand complex molecular pathways that would be difficult or impossible to ascertain with single-omics approaches. In this section, we highlight examples of multi-omics analysis applied to HEK293 cell systems and the molecular insights that they reveal for recombinant protein production, while highlighting the challenges and opportunities that lie ahead.

7.1. Multi-omic analysis of cellular dynamics and responses

By integrating transcriptomic with targeted quantitative proteomic analysis, Lu and colleagues provided new insights into the biology and kinetics of rAAV production in HEK293 (Lu et al., 2024). Using RNA-Seq, absolute quantification through targeted proteomics (AQUA), and tandem mass tags (TMTs), analysis of 20,000 unique transcripts and 5000 proteins was used to measure the kinetics of rAAV triple transient transfection over time (0, 24, 48, and 72 h post-transfection). This analysis revealed that only 5.7 % total genomes (TG) per cell became encapsidated as full viral genomes (VG), suggesting that capsid assembly, genome encapsidation, or both are likely limiting factors for rAAV production. This is consistent with the findings of Braatz and colleagues, whose mechanistic model for rAAV production attributed the low filled capsid fraction to a poorly coordinated timeline between capsid synthesis and viral DNA replication, and the repression of later phase capsid formation by Rep proteins (Srinivasan et al., 2024).

The same study leveraged transcriptomics and proteomics to functionally assess cellular responses, shedding new insights into known responses to rAAV production in HEK293 that negatively or positively impact productivity (e.g., inflammation and the host innate immune response, the antiviral response, the unfolded protein response, MAPK cascade, cell cycle inhibition, DNA damage). Further, it revealed that certain host defense responses were a general consequence of plasmid DNA transfection that was independent of rAAV production. By providing a nuanced systematic mechanistic understanding of these complex phenomena, multi-omics enables a precise means to increase productivity through chemical intervention (e.g., by the addition of specific inhibitor or enhancer molecules) or through synthetic biology (e.g., by engineering the HEK293 host cell line or the rAAV vector).

7.2. Insights into production constraints

Using multi-omics to identify differences between high-producing (HP) and low-producing (LP) HEK293 cell lines can shed light on engineering strategies to increase productivity. Zehetner and colleagues constructed genome-scale metabolic models (GEMs *aka* GSMMs) for HP and LP HEK293 cell lines based on the integration of the biomass composition with transcriptomic, lipidomic, and exometabolomic measurements collected at different conditions and different time points (Zehetner et al., 2024b). Based on the most recent atlas of human metabolism, 20 condition-specific GEMs were reconstructed from the two strains to identify potential bottlenecks in rAAV production. Following transient triple transfection of the LP and HP strains, analysis revealed that over 99.8 % of the resources are used for growth instead of rAAV production in both strains. Insights from the multi-omics analysis were used to identify potential targets that could reallocate resources from cell growth towards rAAV synthesis. In the LP strain, Hypoxia-inducible factor 1-alpha (HIF-1α), a transcription factor that helps cells respond to low oxygen levels, was identified as a potential target. The overexpression of HIF-1α has previously been exploited in CHO cells to enhance recombinant protein production (Zeh et al., 2021). In

HEK293, inhibition of HIF-1α using a small-molecule inhibitor, PX-478, resulted in cessation of growth and a two-fold increase in rAAV production. However, the number of viral genomes was decreased, leading to a reduced full-to-empty capsid ratio, likely a consequence of nucleotide synthesis inhibition via the pentose phosphate pathway by HIF-1α. This highlights the importance of coordinating the kinetics of capsid and viral genome production to maximize the production of full capsids (Srinivasan et al., 2024), and that increasing the production of full viable rAAV products by reallocating resources away from cell growth requires understanding and pre-empting potential unintended consequences. Another aspect of resource reallocation that should be considered is reducing (ideally eliminating) the wasted resources expended by suppressing expression of the transgene – typically under the control of a constitutive promoter (e.g., CMV, CAG) – in HEK293 during manufacturing. This could leverage strategies that have been developed to produce rAAV carrying toxic transgenes, such as using miRNAs, riboswitches, or inducible promoters (Blahetek et al., 2024; Strobel et al., 2015). Arguably one of the most intriguing findings from the GEMs analysis was that both the HP and LP strains provide rAAV at rates that are two orders of magnitude lower than their theoretical production capacity.

Multi-omics analysis of HEK293 stable producer cells has revealed broad cellular network adaptations that likely support recombinant protein production (Dietmair et al., 2012). Transcriptomics, metabolomics, and fluxomics data showed downregulation of genes involved in cell growth and proliferation, explaining how producer cells maintained growth rates despite lower glucose consumption and the burden of recombinant protein production. Simultaneously, upregulation of genes related to ER stress suggested constraints in protein folding and assembly. Similarly, combined proteomics and transcriptomics analysis in CHO revealed ER stress during multispecific antibody production as a contributing factor to incorrect chain assembly and impaired biological performance, providing mechanistic insights on ways to improve clone selection during early process development (Sebastião et al., 2023).

Through a comparative transcriptomics and proteomic analysis, bottlenecks in rAAV production were investigated to enhance the genetics and processes for producing high-quality viral vectors. Using RNA-seq coupled with AQUAs, a HEK293 stable producer cell line (PCL) carrying EGFP was compared with two wtAAV production systems, one by multi-plasmid transient transfection and the other by virus co-infection (Lin et al., 2024). Genes involved in inflammation, innate immunity, antiviral response, and MAPK signaling were upregulated in all three scenarios, while those involved in mitochondrial function were downregulated, identifying potential targets for genetic intervention in both transiently stably transfected HEK293 cell lines. While all three systems produced comparable copies of AAV genomes, the kinetics of viral capsid production differed dramatically, resulting in a 10-fold lower titer for the PCL due to capsid insufficiency and inefficient genome packaging. Another notable difference between the stable PCL versus the two transient systems was that the GFP cargo in the PCL was the most abundant transcript of all cellular mRNA, despite being under the control of an inducible *Lac* Switch. Constituting ~20 % of total cellular transcripts at 72 h, this undesirable cargo expression is a significant waste of cellular resources within the stable cell line. While the scalability and low-cost potential of rAAV PCLs are appealing, they still significantly underperform transient transfection in terms of titer and yield, despite major investment in time and resources over 30+ years. Furthermore, cell line stability is unproven (Escandell et al., 2022; Shupe et al., 2022), and not a single rAAV product produced in a PCL has entered the clinic to date, reflecting caution across the industry.

7.3. Multi-omics for cell line characterization

Combining genomics, transcriptomics, and metabolic pathway analysis, multi-omics was used for cell line characterization of 6 widely used HEK293 cell lines. In addition to providing a cross-validation of results, this holistic approach revealed fundamental evolutionary differences for adapting to industrial production cell lines (Malm et al., 2020)). Notably, key common changes between HEK293 and its progeny production cell lines involved integral membrane proteins and processes related to cell adhesion, motility and the organization of various cellular components such as the cytoskeleton and extracellular matrix.

In the aforementioned integrated GEMs and multi-omic study comparing a high-producing (HP) and low-producing (LP) HEK293 cell line, both strains exhibited nearly identical production capabilities, even though the actual production rate in the HP is up to 6-fold higher, indicating that the most significant differences were between the strains themselves rather than between the production states (Zehetner et al., 2024b). Comparing the multi-omic profiles of the the two strains revealed that the LP strain exhibited preferential consumption of aspartate and glutamate, pseudohypoxia, and upregulation of reactions linked to hypoxic environments, while the HP strain had higher ceramide concentrations in the HP strain. Taken together, these and other findings in the study indicate that metabolic changes in the LP strain resemble hallmarks of cancer metabolism, which could provide further mechanistic insights to reallocate resources away from cell proliferation to increased rAAV production.

7.4. Integrating three or more omics

While combining two branches of omics is powerful for uncovering actionable information for improving production in HEK293, integrating data from three or more omic disciplines can reveal additional useful insights. This has proven to be a successful strategy for understanding complex biological processes, including constraints of recombinant protein production in HEK293. To our knowledge, there are no examples of an integrated multiomics approach for producing rAAV from HEK293 cells, therefore, we look to examples producing lentivirus and other recombinant proteins. For example, a functional genomics study used transcriptional profiling and central carbon metabolism analysis to identify pathways and bottlenecks in the production of lentivirus by investigating the transition from parental to production cell lines (Rodrigues et al., 2013). The cell lines investigated were two human-derived cell lines, HEK293 FLEX (low producer) and Te Fly (high producer), and their parental cell lines (HEK293 and Te671, respectively). Since the producing cell lines have different transcriptomes, it was difficult to find a common expression signature to compare the producing vs. parental cell lines through a single omics approach. Therefore, applying three omics approaches to evaluate genes, gene expression changes, and metabolites across the four cell lines helped identify eight distinct pathways to be recruited in the virus production state. To confirm their findings, the culture medium was supplemented to positively influence these pathways and results showed improved titer by up to 6-fold. Further, by investigating two producing cell lines, with different genetic backgrounds, several genes were identified as possible targets for future engineering strategies as they were limiting production in the low production phenotype. Applying and integrating multiple omics approaches proved to reveal actionable insights into the pathways involved in lentiviral production, such as demonstrating that manipulation of nucleotides and polyamine metabolism through media supplementation can be used to increase productivity.

In a separate study, a stable HEK293 producer cell line was compared to its non-producing parental cell line using transcriptomics, metabolomics, and fluxomics, including flux variability analysis (FVA) to measure the metabolite consumption and production rates (Dietmair et al., 2012). Integrating multiple omics approaches revealed the producer cells consumed less glucose than the non-producing cells and the majority of genes involved with oxidative phosphorylation were downregulated, with no consistent evidence of glycolytic energy compensation. Dietmair and colleagues speculated that a broad adaptation of the cellular network freed resources for other mechanisms to rise for production of recombinant proteins. Additionally, transcript

abundance often correlated positively with fluxes and metabolite concentrations. However, some discrepancies were also noted between transcript levels and fluxes, with occasional counterintuitive outcomes like increased glucose transporter expression paired with reduced glycolytic flux in producer cells. Comparing cellular behavior at multiple levels enabled by an integrated omics approach revealed techniques for rational engineering of cell lines to improve production.

7.5. Multi-omics for enhanced mechanistic understanding

Multi-omics enables a unique comprehensive view of the underlying mechanisms employed by the host cell system, building on decades of biochemistry, genetic, and cell biology that informs our current mechanistic understanding of recombinant protein production in HEK293. Highlights into the mechanisms employed by the host cell system have been described throughout this review. For example, as previously mentioned, Tworig and colleagues conducted a transcriptomics analysis resulting in mechanistic knowledge of druggable pathways in HEK293 cells for an informed experiment using small molecules to regulate core cellular functions (Tworig et al., 2024) (see Section 4.7).

Additionally, transcriptomics has been combined with high-throughput lipidomics to compare HEK293 with two other widely used mammalian expression systems, CHO and SP2/0, during both exponential and stationary growth phases (Zhang et al., 2017) (see Section 6.5). Analyzing transcriptomic expression profiles together with lipidomics data led to four mechanistic findings: (1) LPE and LPC levels were significantly higher in HEK293 cell lines, likely promoting membrane curvature, which can facilitate membrane tubule formation and membrane vesicle fusion and fission, (2) lysophospholipids (LPLs) may play an close role in exocytosis and secretion processes, (3) group IV sPLA2 enzymes are unique to the HEK293 cell line, which may enhance its intracellular membrane trafficking and fusion events, and (4) further promotion of membrane trafficking and tubule formation events may be contributed to low expression of lysophospholipid acyltransferase (LPEAT) and lysophosphatidylcholine acyltransferase (LPCAT) in HEK. This study demonstrates the power of integrating multiple omics approaches for elucidating the mechanisms that limit protein expression, exocytosis, and secretion in HEK293 cells.

While small, rationally designed multi-omics studies can lead to valuable mechanistic insights, mechanistic models provide the most robust actionable information for furthering rAAV production in HEK293 cells (Masson et al., 2023). Informed by omics data, mechanistic models can encompass a range of information from the genome scale to bioreactor scale, and from cell culturing to protein production (Canova et al., 2023; Masson et al., 2023). For those phenomena in which the existing mechanistic phenomena are not sufficiently well understood to build purely mechanistic models, data analytics/machine learning (DA/ML) tools can be applied to build those relationships, which can be integrated with the mechanistic model components to produce hybrid models (Tsopanoglou and Jiménez del Val, 2021). A general systematic modeling framework for building hybrid models from multi-omics data, however, is not yet available.

Mechanistic modeling has been an invaluable tool to support process development for biotherapeutic production in CHO cells, and more recently, HEK293 cells (Sha et al., 2018). For example, Braatz and colleagues developed a mechanistic model of rAAV via transient triple transfection of HEK293 cells growing in suspension has been developed and used to identify bottlenecks and propose improvements to the system (Nguyen et al., 2021). As previously mentioned, application of this model provided insight into a timeline disconnect between capsid synthesis and viral DNA replication, ultimately leading to a novel multi-dose transfection method to coordinate these timelines (Srinivasan et al., 2024). Additionally, utilizing the underlying mechanisms to produce rAAV incorporated in the mechanistic model, a first-in-class mechanistic model for continuous rAAV production via transfection in a perfusion bioreactor was constructed (Nguyen et al., 2024). The model was used to design a continuous process with high cell density transfection and experimentally validated the approach of using multiple dosing for re-transfection. The resulting high-density transfection and re-transfection resulted in increased batch runtime efficiency by 8-fold (i.e., 8 times the amount of a conventional batch in the same amount of time) and increased titer production per unit plasmid by ~32 %. Integrating mechanistic models with omics datasets has untapped potential in identifying bottlenecks that may not be obvious from the omics datasets alone; therefore, furthering the toolbox of mechanistic understandings for production of rAAV from HEK293 cells (Lazarou et al., 2020; Masson et al., 2023).

7.6. Challenges and opportunities

Analysis of multi-omics data can be profoundly insightful in understanding the cellular mechanisms; however, the extraction of meaningful biological knowledge from complex large datasets representing different layers of the cellular machinery can be challenging. The process of generating biological knowledge typically involves four key steps: data generation, processing, integration, and analysis (Palsson and Zengler, 2010). Although today's low costs of sequencing and automated bioinformatics pipelines make the first two steps fairly straightforward for most omics data, the third and fourth steps can be difficult in delivering data-driven insights as multi-omics data differ in type, scale, and distribution with thousands of variables and only few samples (Picard et al., 2021). For preliminary insights, cross-omics data integration can be achieved through basic statistical approaches such as correlation analysis across enriched biological features from individual omics datasets. However, to extract deeper and more profound and nuanced insights, artificial intelligence (AI)-based approaches, in particular machine learning (ML), are increasingly applied (Kang et al., 2022; Picard et al., 2021; Reel et al., 2021). By being able to better model the inter-omics interactions between complementary omics datasets, ML approaches have the potential to increase understanding of the underlying biological mechanisms. This has been demonstrated by the recent surge in their application for large-scale multi-omics analysis of cancer datasets (Chaudhary et al., 2018; Li et al., 2022; Nicora et al., 2020; Opdam et al., 2017; Poirion et al., 2021; Wang et al., 2021). In the context of biotherapeutic production, ML has been applied for genome-scale modeling of CHO metabolic networks (Schinn et al., 2021; Zampieri et al., 2019) and for predicting protein glycosylation patterns in CHO cells (Antonakoudis et al., 2021; Kotidis and Kontoravdi, 2020).

As new tools emerge for improving rAAV and recombinant protein production, multi-omics can be leveraged to elucidate the underlying mechanism and further improve productivity. The recent emergence of small-molecule (SM) chemical additives is a case in point, where a growing number of SMs has been empirically shown to increase rAAV production (Maznyi, 2023; Reese et al., 2023; Scarrott et al., 2023). One such class of SMs, known as viral sensitizers, enhance viral production by transiently attenuating antiviral pathways in the cell, such as the innate immune response (Diallo and De Jong, 2023). With a growing number of commercially available SM additives for enhancing viral vector manufacturing, multi-omics analysis could enable the systematic elucidation of the mechanism of action for each molecule, and shed light on the pathways and kinetics. Such omics-enabled insights could shed light on potential synergies, inform new mechanistic models for rAAV production, and fuel new engineering strategies to improve HEK293 cell lines, such as attenuating the innate immune response.

The ever-growing ability to generate a panoply of large data sets from multiplicity of sources and techniques combined with the unprecedented power of AI-enabled approaches can lead to the temptation to apply AI/ML to analyze the immediate data to look for patterns. However, this simplistic tendency overlooks decades of information about fundamental biological processes and leads to minor incremental improvements, at best. To gain novel, useful, and actionable insights, multi-omics should be combined with additional layers from our

Table 1
HEK-omics: Gaps and Opportunities.

Omics Type	Gaps and Opportunities	Key References
Genomics	• Establish comprehensive HEK293 cell line-specific reference map and gene annotation set. • Investigate genetic changes in the form of SNPs, CNVs, and structural variants between HEK293 clonal populations with different phenotypes, e.g., linking varying titer productivity levels to genetic characteristics.	(Lin et al., 2014; Malm et al., 2020)
Epigenomics	• Characterize epigenetic differences among HEK293 cell lines under various culture and media conditions, e.g., cell cultures adapted to suspension versus adherent medium. • Investigate unpredictable outcomes occurring in HEK293 cells producing rAAV, e.g., cell density effect.	(Broche et al., 2021; Tóth et al., 2019; Chanda et al., 2017; Feichtinger et al., 2016; Marx et al., 2022)
Transcriptomics	• Investigate transcriptional changes in transiently transfected HEK293 cell cultures and identify key regulatory pathways invoked during the production of recombinant AAV. • Investigate druggable pathways within HEK293 cells to enhance transfection and optimize small-molecule usage.	(Chung et al., 2023; Malm et al., 2020, 2022; Saghaleyni et al., 2022; Wang et al., 2023; Tworig et al., 2024)
Proteomics	• Identify and utilize structure-function relationships for enhancing engineered HEK293 producing cell lines. • Analyze proteomic characteristics and differences among widely used rAAV-producing HEK293 cell lines, e.g., HEK293E, HEK293T (adherent lines), HEK293H, HEK293F, and Freestyle HEK293F (suspension lines). • Proteomic profiling of rAAV-producing HEK293 cell lines in manufacturing environments to study the effects of process conditions on productivity.	(Strasser et al., 2021; Lavado-García et al., 2020a, 2020b; Malm et al., 2020; Heusel et al., 2019)
Metabolomics	• Develop comprehensive metabolic model for HEK293 cells, derived from human genome-scale metabolic models such as RECON or HMR series.	(Henry et al., 2005; Henry and Durocher, 2011; Martínez-Monge et al., 2019; Quek et al., 2014)
Multi-omics	• Multi-omics study of HEK293 cells during rAAV production process to identify host cellular mechanisms favorable for higher productivity and better quality (to parallel similar studies for CHO, e.g., Lee et al., 2020). • Integrated multi-omics study of HEK293 cells producing rAAV using 3 or more omics approaches. • Develop mechanistic or hybrid models for HEK293 producing rAAV.	(Dietmair et al., 2012; Malm et al., 2020; Nguyen et al., 2021; Lu et al., 2024)

"institutional knowledge." Today, the analysis and integration of multi-omics data is still relatively untapped for HEK293, and therein lies an opportunity. Doing so could enable exponential improvements in productivity that would have a major impact on rAAV gene therapy production in terms of yield, speed, and cost (Fig. 1).

8. Conclusion and future directions

For rAAV gene therapy to cross the chasm from niche to mainstream medicine, there must be a step change in scalable, cost-effective production. We propose exploiting HEK-omics – ranging from single- to multi-omics – to address knowledge gaps and identify opportunities to optimize rAAV production by revealing actionable insights (Fig. 1B and Table 1). This section discusses some opportunities beyond omics and HEK293 for scalable production of rAAV-based gene therapies.

8.1. Trade-offs between process development strategies

The optimization of the production of a biotherapeutic can focus on cell line improvements, process improvements, or both. The advantages of improving a cell line are that the genome has a large number of degrees of freedom for optimization, and any changes baked into the genome are easier to transfer to larger scale bioreactor production. The advantage of process improvement is that its available degrees of freedom can be optimized much more quickly. The number of degrees of freedom in process optimization can include media composition, flow rates (for fed-batch or perfusion), target cell density, plasmid designs, and choice of transfection reagent – many of which can be changed dynamically over time. Going the route of cell line development, such as creating a cell line in which the genome of the biotherapeutic is baked into the genome, opens up handles the huge number of degrees of freedom in the genome but also removes other degrees of freedom, such as the ability to transiently add multiple plasmids such as occurs in the transient transfection-based approach for biotherapeutics production. Of course some process optimization variables, such as media composition, have the same number of degrees of freedom irrespective of whether cell line improvement is being explored as part of the biotherapeutics process development. In either case, the degrees of freedom are chosen to maximize overall productivity, including maintenance of production stability, while ensuring that all specifications on critical product quality attributes (CQAs) are met for ensuring efficacy and patient safety.

8.2. Underexplored omics: glycomics and epigenomics

Relative to genomics, transcriptomics, proteomics, epigenomics, and metabolomics, the area of glycomics in HEK293 is the least explored. Given the impact of N-linked glycosylation on the potency, safety, immunogenicity, and pharmacokinetic clearance of therapeutic proteins including monoclonal antibodies (Heffner et al., 2021; Narimatsu et al., 2019), this is a critical knowledge gap. As a case in point, integrating glycome data with transcriptomic and metabolomic data measured during a time course of mAb production in a CHO fed-batch culture revealed that the steps involving galactose and sialic acid addition were temporal bottlenecks (Sumit et al., 2019).

The relevance of glycomics to rAAV production in HEK293 was underscored by the glycomic profiling of capsid proteins from AAV serotypes (Xie and Butler, 2023). Using a high-sensitivity N-glycan profiling platform, analysis of 9 serotypes (AAV1–9) showed that all had comparable glycosylation profiles, characterized by high mannosylated N-glycans and lower fucosylated and sialylated N-glycan structures. However, the precise glycan compositions differed across the different serotypes. Given that many virus-host interactions are driven or influenced by glycans (Miller et al., 2021; Zhou et al., 2024), these data suggest that the differences could play a role in tissue tropism, cell surface receptor binding, intracellular uptake and processing.

Leveraging transcriptomics, Huang and colleagues developed GlycoMaple, a glycosylation mapping tool for visualizing glycosylation pathways and estimating glycan structures in human-derived cells (Huang et al., 2021). This tool was verified using HEK293 cells, determining 38 and 14 composition structures in N-glycan and O-glycan analysis, respectively. Characterization of these glycan structures further informed engineering strategies to knock out certain genes involved in LacdiNAc formation to eliminate LacdiNAc structures. This

tool has proven to be useful in predicting changes in glycans to genetic mutations, but further experimentation can be completed to investigate limiting or critical steps in rAAV glycosylation pathways.

Beyond glycomics, epigenomics and the role of epigenetics in HEK293-producing cell lines are relatively unexplored to date. Methylation and chromatin analysis of CHO cell lines (Hilliard and Lee, 2021; Moritz et al., 2016; Osterlehner et al., 2011; Patel et al., 2018), combined with preliminary analysis of rAAV production in HEK293, points to epigenetic engineering as a promising means to improve productivity. Further, epigenetics can potentially be driving the adaptation of HEK293 cells to serum-free suspension medium from adherent medium, as in the case of CHO cells (Feichtinger et al., 2016), and yet a comprehensive epigenomics study investigating these mechanisms in rAAV-producing HEK293 cells is missing.

8.3. Alternative rAAV production systems

When exploring approaches to solve the fundamental gap in scalable rAAV production, it is prudent to consider alternatives. Here, we discuss two alternative production systems – BEVS and HeLA S2 – as viable systems for clinical and commercial production, although other host production systems are also emerging, including plant cells (Gibbs and Connors, 2021).

The Sf9 insect cell baculovirus expression vector system (BEVS) has emerged as a viable alternative rAAV production system (Joshi et al., 2021), with the first BEVS-based gene therapy (HEMGENIX®) approved in 2022 (Herzog et al., 2023). While BEVS has several advantages over HEK293 (Destro et al., 2023), many challenges remain that would benefit from insights through omics analysis. Transcriptome analysis of dual-baculovirus infection of Sf9 cells revealed genes that were differentially expressed during rAAV production, revealing potential engineering targets for process engineering (Virgolini et al., 2023a). Also, mechanistic models have been developed to describe the production dynamics of rAAV in the One-Bac, Two-Bac, and Three-Bac BEVS and applied to identify bottlenecks that limit full capsid formation (Destro et al., 2023). Integrating the mechanistic models for rAAV production in Sf9 cells with the corresponding omics data could improve their predictive accuracy and broaden their utility.

Since little is known about the variability of insect cell populations and the potential effects of heterogeneity on viral titer and/or quality, single-cell RNA sequencing (scRNA-seq) was performed to shed light on transcriptional variation within a cell population following infection (Virgolini et al., 2023b). At 24 h post-infection, only 29.4 % of cells expressed all of the three rAAV transgenes needed to produce full rAAV capsids. This study highlights the power of single-cell omics, revealing production bottlenecks and how to address them, such as using synchronization strategies to reduce cell heterogeneity. The study also underscores an intrinsic limitation of current omic-based analysis: most approaches measure averages across cell populations and a limited number of time points, obscuring the stochasticity and temporal variation for every step of a process, particularly one as complex as rAAV production. As single-cell multi-omics technologies improve in terms of throughput, data analytics, and affordable pricing (Baysoy et al., 2023; Heumos et al., 2023), they will become increasingly used to characterize cell-to-cell variability and use that information to increase overall system productivity.

Another alternative production system to HEK293 is the HeLa S3 stable cell line platform. First described in 1995, the HeLa S3 production system proved to be scalable up to 250 L and was applied in clinical manufacturing of AAV gene therapies (Clark et al., 1995; Tatalick et al., 2005; Thorne et al., 2009). To develop an rAAV-expressing HeLa S3 producer cell line, *rep*, *cap*, and *ITR* genes were stably integrated into the HeLa S3 genome, with wt-Ad5 providing the helper elements (Escandell et al., 2023; Liu et al., 2000). However, the cell density effect in the HeLa S3 system was observed at lower cell concentrations than for HEK293, with significantly lower volumetric productivity of $>3 \times 10^{11}$ vg/mL at a 2 L bioreactor scale, and purified yield of $>1 \times 10^{11}$ vg/mL. More recently, a novel HeLa S3-based system, the Pinnacle PCL platform, in combination with an abbreviated perfusion production process was able to achieve higher maximum cell densities, resulting in titers approaching 1×10^{12} GC/mL that could scale to a 250 L bioreactor (Xue et al., 2024). The 6- to 8-fold increase in volumetric titer was largely attributable to the new process rather than the new cell line. Given the relative speed, flexibility, and many degrees of freedom for process engineering compared with the high cost, time, and risky nature of cell line engineering, this example highlights the power of process improvement.

Finally, in light of the HEK293 productivity gap for rAAV gene therapy production, it bears re-examining why CHO cells were not adopted as the system of choice. There is speculation that cellular restriction factors in CHO cells could affect rAAV production and interfere with virus packaging, as reported for other viruses (Nagy et al., 2023). Nevertheless, two new CHO-based systems for rAAV production recently came to light. Nagy and colleagues engineered a stable CHO cell line to express entry receptors so that they can be transfected using the herpes simplex virus (HSV) (Nagy et al., 2023). Although the titers resulting from transfection of the CHO cells were low (1.6×10^9–4.1×10^9 vg/mL obtained 24 h post-transfection), the resulting rAAVs had comparable in vitro transduction, infectivity, and biodistribution titers to rAAVs produced by triple transient transfection of HEK293 cells. In contrast to the stable cell line approach, Cao and colleagues developed a helper-free, transient 5-plasmid transfection CHO system for rAAV production (Cao et al., 2023). Despite low productivity, the study identified the catabolite activator protein (cap) expression as a limiting factor for rAAV production in CHO cells, laying the groundwork for future research. While one of the themes of this article is that HEK-omics can learn from studies on CHO-omics, the converse is true for increasing titer for rAAV production in CHO. To borrow from Shakespeare, the "wheel is come full circle".

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by an EPoC Grant from the Delaware Bioscience Center for Advanced Technologies to SWP in collaboration with ITR. GG received additional support from a Doctoral Fellowship for Excellence from the University of Delaware Graduate College.

References

Abaandou, L., Quan, D., Shiloach, J., 2021a. Affecting HEK293 cell growth and production performance by modifying the expression of specific genes. Cells 10, 1667. https://doi.org/10.3390/cells10071667.

Abaandou, L., Sharma, A.K., Shiloach, J., 2021b. Knockout of the caspase 8-associated protein 2 gene improves recombinant protein expression in HEK293 cells through up-regulation of the cyclin-dependent kinase inhibitor 2A gene. Biotechnol. Bioeng. 118, 186–198. https://doi.org/10.1002/bit.27561.

Abbate, T., Dewasme, L., Bogaerts, P., Wouwer, A.V., 2019. Adaptive flux variability analysis: a tool to deal with uncertainties. IFAC-PapersOnLine 52, 70–75. https://doi.org/10.1016/j.ifacol.2019.06.039.

Abbate, T., Dewasme, L., Vande Wouwer, A., Bogaerts, P., 2020. Adaptive flux variability analysis of HEK cell cultures. Comput. Chem. Eng. 133, 106633. https://doi.org/10.1016/j.compchemeng.2019.106633.

Aldridge, S., Teichmann, S.A., 2020. Single cell transcriptomics comes of age. Nat. Commun. 11, 4307. https://doi.org/10.1038/s41467-020-18158-5.

Allis, C.D., Jenuwein, T., 2016. The molecular hallmarks of epigenetic control. Nat. Rev. Genet. 17, 487–500. https://doi.org/10.1038/nrg.2016.59.

Anand, S., Samuel, M., Ang, C.-S., Keerthikumar, S., Mathivanan, S., 2017. Label-based and label-free strategies for protein quantitation. In: Keerthikumar, S., Mathivanan, S. (Eds.), Proteome Bioinformatics, Methods in Molecular Biology. Humana Press, New York, NY, pp. 31–43. https://doi.org/10.1007/978-1-4939-6740-7_4.

Ankney, J.A., Muneer, A., Chen, X., 2018. Relative and absolute quantitation in mass spectrometry-based proteomics. Annu. Rev. Anal. Chem. 11, 49–77. https://doi.org/10.1146/annurev-anchem-061516-045357.

Antonakoudis, A., Strain, B., Barbosa, R., Jimenez del Val, I., Kontoravdi, C., 2021. Synergising stoichiometric modelling with artificial neural networks to predict antibody glycosylation patterns in Chinese hamster ovary cells. Comput. Chem. Eng. 154, 107471. https://doi.org/10.1016/j.compchemeng.2021.107471.

Aznauryan, E., Yermanos, A., Kinzina, E., Devaux, A., Kapetanovic, E., Milanova, D., Church, G.M., Reddy, S.T., 2022. Discovery and validation of human genomic safe harbor sites for gene and cell therapies. Cell Rep. Meth. 2, 100154. https://doi.org/10.1016/j.crmeth.2021.100154.

Barnes, C.R., Lee, H., Ojala, D.S., Lewis, K.K., Limsirichai, P., Schaffer, D.V., 2021. Genome-wide activation screens to increase adeno-associated virus production. Mol. Ther. Nucleic Acids 26, 94–103. https://doi.org/10.1016/j.omtn.2021.06.026.

Baysoy, A., Bai, Z., Satija, R., Fan, R., 2023. The technological landscape and applications of single-cell multi-omics. Nat. Rev. Mol. Cell Biol. 24, 695–713. https://doi.org/10.1038/s41580-023-00615-w.

Beacon, T.H., Delcuve, G.P., López, C., Nardocci, G., Kovalchuk, I., van Wijnen, A.J., Davie, J.R., 2021. The dynamic broad epigenetic (H3K4me3, H3K27ac) domain as a mark of essential genes. Clin. Epigenetics 13, 138. https://doi.org/10.1186/s13148-021-01126-1.

Berk, A.J., 2005. Recent lessons in gene expression, cell cycle control, and cell biology from adenovirus. Oncogene 24, 7673–7685. https://doi.org/10.1038/sj.onc.1209040.

Bernstein, D.B., Sulheim, S., Almaas, E., Segrè, D., 2021. Addressing uncertainty in genome-scale metabolic model reconstruction and analysis. Genome Biol. 22, 64. https://doi.org/10.1186/s13059-021-02289-z.

Blahetek, G., Mayer, C., Zuber, J., Lenter, M., Strobel, B., 2024. Suppression of toxic transgene expression by optimized artificial miRNAs increases AAV vector yields in HEK-293 cells. Mol. Ther. Methods Clin. Dev. 32, 101280. https://doi.org/10.1016/j.omtm.2024.101280.

Borsi, G., Motheramgari, K., Dhiman, H., Baumann, M., Sinkala, E., Sauerland, M., Riba, J., Borth, N., 2023. Single-cell RNA sequencing reveals homogeneous transcriptome patterns and low variance in a suspension CHO-K1 and an adherent HEK293FT cell line in culture conditions. J. Biotechnol. 364, 13–22. https://doi.org/10.1016/j.jbiotec.2023.01.006.

Broche, J., Kungulovski, G., Bashtrykov, P., Rathert, P., Jeltsch, A., 2021. Genome-wide investigation of the dynamic changes of epigenome modifications after global DNA methylation editing. Nucleic Acids Res. 49, 158–176. https://doi.org/10.1093/nar/gkaa1169.

Brunk, E., Sahoo, S., Zielinski, D.C., Altunkaya, A., Dräger, A., Mih, N., Gatto, F., Nilsson, A., Preciat Gonzalez, G.A., Aurich, M.K., Prlić, A., Sastry, A., Danielsdottir, A.D., Heinken, A., Noronha, A., Rose, P.W., Burley, S.K., Fleming, R.M. T., Nielsen, J., Thiele, I., Palsson, B.O., 2018. Recon3D enables a three-dimensional view of gene variation in human metabolism. Nat. Biotechnol. 36, 272–281. https://doi.org/10.1038/nbt.4072.

Burdett, T., Nuseibeh, S., 2023. Changing trends in the development of AAV-based gene therapies: a meta-analysis of past and present therapies. Gene Ther. 30, 323–335. https://doi.org/10.1038/s41434-022-00363-0.

Canova, C.T., Inguva, P.K., Braatz, R.D., 2023. Mechanistic modeling of viral particle production. Biotechnol. Bioeng. 120, 629–641. https://doi.org/10.1002/bit.28296.

Cao, T.M., Chen, D., Barnard, G.C., Shen, A., 2023. Recombinant adeno-associated virus production evaluation in Chinese hamster ovary cells. Biotechnol. Bioeng. 121, 395–402. https://doi.org/10.1002/bit.28578.

Cavalli, G., Heard, E., 2019. Advances in epigenetics link genetics to the environment and disease. Nature 571, 489–499. https://doi.org/10.1038/s41586-019-1411-0.

Center for Biologics Evaluation, Research, 2023. Approved Cellular and Gene Therapy Products [WWW Document]. U.S. Food and Drug Administration. URL. https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/approved-cellular-and-gene-therapy-products (accessed 7.31.23).

Chancellor, D., Barrett, D., Nguyen-Jatkoe, L., Millington, S., Eckhardt, F., 2023. The state of cell and gene therapy in 2023. Mol. Ther. 31, 3376–3388. https://doi.org/10.1016/j.ymthe.2023.11.001.

Chanda, D., Hensel, J.A., Higgs, J.T., Grover, R., Kaza, N., Ponnazhagan, S., 2017. Effects of cellular methylation on transgene expression and site-specific integration of adeno-associated virus. Genes (Basel) 8, 232. https://doi.org/10.3390/genes8090232.

Chaudhary, K., Poirion, O.B., Lu, L., Garmire, L.X., 2018. Deep learning-based multi-omics integration robustly predicts survival in liver cancer. Clin. Cancer Res. 24, 1248–1259. https://doi.org/10.1158/1078-0432.CCR-17-0853.

Cheng, J.K., Lewis, A.M., Kim, D.S., Dyess, T., Alper, H.S., 2016. Identifying and retargeting transcriptional hot spots in the human genome. Biotechnol. J. 11, 1100–1109. https://doi.org/10.1002/biot.201600015.

Chu, L., Robinson, D.K., 2001. Industrial choices for protein production by large-scale cell culture. Curr. Opin. Biotechnol. 12, 180–187. https://doi.org/10.1016/s0958-1669(00)00197-x.

Chung, C.-H., Murphy, C.M., Wingate, V.P., Pavlicek, J.W., Nakashima, R., Wei, W., McCarty, D., Rabinowitz, J., Barton, E., 2023. Production of rAAV by plasmid transfection induces antiviral and inflammatory responses in suspension HEK293 cells. Mol. Ther. Meth. Clin. Dev. 28, 272–283. https://doi.org/10.1016/j.omtm.2023.01.002.

Clark, K.R., Voulgaropoulou, F., Fraley, D.M., Johnson, P.R., 1995. Cell lines for the production of recombinant adeno-associated virus. Hum. Gene Ther. 6, 1329–1341. https://doi.org/10.1089/hum.1995.6.10-1329.

Clish, C.B., 2015. Metabolomics: an emerging but powerful tool for precision medicine. Cold Spring Harb. Mol. Case Stud. 1, a000588. https://doi.org/10.1101/mcs.a000588.

Cortassa, S., Caceres, V., Bell, L.N., O'Rourke, B., Paolocci, N., Aon, M.A., 2015. From metabolomics to fluxomics: a computational procedure to translate metabolite profiles into metabolic fluxes. Biophys. J. 108, 163–172. https://doi.org/10.1016/j.bpj.2014.11.1857.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., Boyer, L.A., Young, R.A., Jaenisch, R., 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. USA 107, 21931–21936. https://doi.org/10.1073/pnas.1016071107.

Dahodwala, H., Lee, K.H., 2019. The fickle CHO: a review of the causes, implications, and potential alleviation of the CHO cell line instability problem. Curr. Opin. Biotechnol. 60, 128–137. https://doi.org/10.1016/j.copbio.2019.01.011.

Dahodwala, H., Sharfstein, S.T., 2017. The 'omics revolution in CHO biology: Roadmap to improved CHO productivity. In: Meleady, P. (Ed.), Heterologous Protein Production in CHO Cells, Methods in Molecular Biology. Humana Press, New York, NY, pp. 153–168. https://doi.org/10.1007/978-1-4939-6972-2_10.

Daskalaki, E., Pillon, N.J., Krook, A., Wheelock, C.E., Checa, A., 2018. The influence of culture media upon observed cell secretome metabolite profiles: the balance between cell viability and data interpretability. Anal. Chim. Acta 1037, 338–350. https://doi.org/10.1016/j.aca.2018.04.034.

De Paoli-Iseppi, R., Gleeson, J., Clark, M.B., 2021. Corrigendum: isoform age - splice isoform profiling using long-read technologies. Front. Mol. Biosci. 8, 767743. https://doi.org/10.3389/fmolb.2021.767743.

de Souza, N., Picotti, P., 2020. Mass spectrometry analysis of the structural proteome. Curr. Opin. Struct. Biol. 60, 57–65. https://doi.org/10.1016/j.sbi.2019.10.006.

Denny, J.C., Collins, F.S., 2021. Precision medicine in 2030—seven ways to transform healthcare. Cell 184, 1415–1419. https://doi.org/10.1016/j.cell.2021.01.015.

Destro, F., Joseph, J., Srinivasan, P., Kanter, J.M., Neufeld, C., Wolfrum, J.M., Barone, P. W., Springs, S.L., Sinskey, A.J., Cecchini, S., Kotin, R.M., Braatz, R.D., 2023. Mechanistic modeling explains the production dynamics of recombinant adeno-associated virus with the baculovirus expression vector system. Mol. Ther. Meth. Clin. Dev. 30, 122–146. https://doi.org/10.1016/j.omtm.2023.05.019.

Dexheimer, P.J., Cochella, L., 2020. MicroRNAs: from mechanism to organism. Front. Cell Dev. Biol. 8, 409. https://doi.org/10.3389/fcell.2020.00409.

Diallo, J.-S., De Jong, J., 2023. Overcoming the challenge of antiviral defense in viral vector manufacturing. Cell Gene Ther. Insights 9, 1097–1104. https://doi.org/10.18609/cgti.2023.144.

Dietmair, S., Hodson, M.P., Quek, L.-E., Timmins, N.E., Gray, P., Nielsen, L.K., 2012. A multi-omics analysis of recombinant protein production in Hek293 cells. PLoS One 7, e43394. https://doi.org/10.1371/journal.pone.0043394.

Dobrowsky, T., Gianni, D., Pieracci, J., Suh, J., 2021. AAV manufacturing for clinical use: insights on current challenges from the upstream process perspective. Current Opinion in Biomedical Engineering 20, 100353. https://doi.org/10.1016/j.cobme.2021.100353.

Duarte, N.C., Becker, S.A., Jamshidi, N., Thiele, I., Mo, M.L., Vo, T.D., Srivas, R., Palsson, B.Ø., 2007. Global reconstruction of the human metabolic network based on genomic and bibliomic data. Proc. Natl. Acad. Sci. USA 104, 1777–1782. https://doi.org/10.1073/pnas.0610772104.

Dunn, M.J., 1987. Two-dimensional gel electrophoresis of proteins. J. Chromatogr. 418, 145–185. https://doi.org/10.1016/0378-4347(87)80008-7.

Eisenhut, P., Marx, N., Borsi, G., Papež, M., Ruggeri, C., Baumann, M., Borth, N., 2024. Manipulating gene expression levels in mammalian cell factories: An outline of synthetic molecular toolboxes to achieve multiplexed control. New Biotechnol. 79, 1–19. https://doi.org/10.1016/j.nbt.2023.11.003.

Elmore, Z.C., Patrick Havlik, L., Oh, D.K., Anderson, L., Daaboul, G., Devlin, G.W., Vincent, H.A., Asokan, A., 2021. The membrane associated accessory protein is an adeno-associated viral egress factor. Nat. Commun. 12, 6239. https://doi.org/10.1038/s41467-021-26485-4.

Emwas, A.-H., Szczepski, K., Al-Younis, I., Lachowicz, J.I., Jaremko, M., 2022. Fluxomics - new metabolomics approaches to monitor metabolic pathways. Front. Pharmacol. 13, 805782. https://doi.org/10.3389/fphar.2022.805782.

Escandell, J.M., Pais, D.A., Carvalho, S.B., Vincent, K., Gomes-Alves, P., Alves, P.M., 2022. Leveraging rAAV bioprocess understanding and next generation bioanalytics development. Curr. Opin. Biotechnol. 74, 271–277. https://doi.org/10.1016/j.copbio.2021.12.009.

Escandell, J., Moura, F., Carvalho, S.B., Silva, R.J.S., Correia, R., Roldão, A., Gomes-Alves, P., Alves, P.M., 2023. Towards a scalable bioprocess for rAAV production using a HeLa stable cell line. Biotechnol. Bioeng. 120, 2578–2587. https://doi.org/10.1002/bit.28394.

Farrell, A., McLoughlin, N., Milne, J.J., Marison, I.W., Bones, J., 2014. Application of multi-omics techniques for bioprocess design and optimization in chinese hamster ovary cells. J. Proteome Res. 13, 3144–3159. https://doi.org/10.1021/pr500219b.

Feichtinger, J., Hernández, I., Fischer, C., Hanscho, M., Auer, N., Hackl, M., Jadhav, V., Baumann, M., Krempl, P.M., Schmidl, C., Farlik, M., Schuster, M., Merkel, A., Sommer, A., Heath, S., Rico, D., Bock, C., Thallinger, G.G., Borth, N., 2016. Comprehensive genome and epigenome characterization of CHO cells in response to evolutionary pressures and over time. Biotechnol. Bioeng. 113, 2241–2253. https://doi.org/10.1002/bit.25990.

Fernandes, S., Robitaille, J., Bastin, G., Jolicoeur, M., Wouwer, A.V., 2016. Application of dynamic metabolic flux convex analysis to CHO-DXB11 cell fed-batch cultures. IFAC-PapersOnLine 49, 466–471. https://doi.org/10.1016/j.ifacol.2016.07.386.

Filtz, T.M., Vogel, W.K., Leid, M., 2014. Regulation of transcription factor activity by interconnected post-translational modifications. Trends Pharmacol. Sci. 35, 76–85. https://doi.org/10.1016/j.tips.2013.11.005.

Frankish, A., Diekhans, M., Jungreis, I., Lagarde, J., Loveland, J.E., Mudge, J.M., Sisu, C., Wright, J.C., Armstrong, J., Barnes, I., Berry, A., Bignell, A., Boix, C., Carbonell Sala, S., Cunningham, F., Di Domenico, T., Donaldson, S., Fiddes, I.T., García Girón, C., Gonzalez, J.M., Grego, T., Hardy, M., Hourlier, T., Howe, K.L., Hunt, T., Izuogu, O.G., Johnson, R., Martin, F.J., Martínez, L., Mohanan, S., Muir, P., Navarro, F.C.P., Parker, A., Pei, B., Pozo, F., Riera, F.C., Ruffier, M., Schmitt, B.M., Stapleton, E., Suner, M.-M., Sycheva, I., Uszczynska-Ratajczak, B., Wolf, M.Y., Xu, J., Yang, Y.T., Yates, A., Zerbino, D., Zhang, Y., Choudhary, J.S., Gerstein, M., Guigó, R., Hubbard, T.J.P., Kellis, M., Paten, B., Tress, M.L., Flicek, P., 2021. GENCODE 2021. Nucleic Acids Res. 49, D916–D923. https://doi.org/10.1093/nar/gkaa1087.

Fritsch, S.D., Weichhart, T., 2016. Effects of interferons and viruses on metabolism. Front. Immunol. 7, 630. https://doi.org/10.3389/fimmu.2016.00630.

Fu, Q., Wang, Y., Qin, J., Xie, D., McNally, D., Yoon, S., 2024. Enhanced ER protein processing gene expression increases rAAV yield and full capsid ratio in HEK293 cells. Appl. Microbiol. Biotechnol. 108, 459. https://doi.org/10.1007/s00253-024-13281-5.

Ghaffari, S., Hanson, C., Schmidt, R.E., Bouchonville, K.J., Offer, S.M., Sinha, S., 2021. An integrated multi-omics approach to identify regulatory mechanisms in cancer metastatic processes. Genome Biol. 22, 19. https://doi.org/10.1186/s13059-020-02213-x.

Giani, A.M., Gallo, G.R., Gianfranceschi, L., Formenti, G., 2020. Long walk to genomics: history and current approaches to genome sequencing and assembly. Comput. Struct. Biotechnol. J. 18, 9–19. https://doi.org/10.1016/j.csbj.2019.11.002.

Gibbs, D., Connors, J.O., 2021. Recombinant adeno-associated viral vectors in plants. WO2021158648. August 12, 2021.

Glinšek, K., Bozovičar, K., Bratkovič, T., 2023. CRISPR technologies in Chinese hamster ovary cell line engineering. Int. J. Mol. Sci. 24. https://doi.org/10.3390/ijms24098144.

Graham, F.L., Smiley, J., Russell, W.C., Nairn, R., 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J. Gen. Virol. 36, 59–74. https://doi.org/10.1099/0022-1317-36-1-59.

Green, E.D., Gunter, C., Biesecker, L.G., Di Francesco, V., Easter, C.L., Feingold, E.A., Felsenfeld, A.L., Kaufman, D.J., Ostrander, E.A., Pavan, W.J., Phillippy, A.M., Wise, A.L., Dayal, J.G., Kish, B.J., Mandich, A., Wellington, C.R., Wetterstrand, K.A., Bates, S.A., Leja, D., Vasquez, S., Gahl, W.A., Graham, B.J., Kastner, D.L., Liu, P., Rodriguez, L.L., Solomon, B.D., Bonham, V.L., Brody, L.C., Hutter, C.M., Manolio, T. A., 2020. Strategic vision for improving human health at the forefront of genomics. Nature 586, 683–692. https://doi.org/10.1038/s41586-020-2817-4.

Gu, C., Kim, G.B., Kim, W.J., Kim, H.U., Lee, S.Y., 2019. Current status and applications of genome-scale metabolic models. Genome Biol. 20, 121. https://doi.org/10.1186/s13059-019-1730-3.

Heffner, K.M., Wang, Q., Hizal, D.B., Can, Ö., Betenbaugh, M.J., 2021. Glycoengineering of mammalian expression systems on a cellular level. Adv. Biochem. Eng. Biotechnol. 175, 37–69. https://doi.org/10.1007/10_2017_57.

Henry, O., Durocher, Y., 2011. Enhanced glycoprotein production in HEK-293 cells expressing pyruvate carboxylase. Metab. Eng. 13, 499–507. https://doi.org/10.1016/j.ymben.2011.05.004.

Henry, O., Perrier, M., Kamen, A., 2005. Metabolic flux analysis of HEK-293 cells in perfusion cultures for the production of adenoviral vectors. Metab. Eng. 7, 467–476. https://doi.org/10.1016/j.ymben.2005.08.002.

Hertel, O., Neuss, A., 2024. Enhancing cell line stability by CRISPR/Cas9-mediated site-specific integration based on histone modifications. Methods Mol. Biol. 2810, 211–233. https://doi.org/10.1007/978-1-0716-3878-1_14.

Hertel, O., Neuss, A., Busche, T., Brandt, D., Kalinowski, J., Bahnemann, J., Noll, T., 2022. Enhancing stability of recombinant CHO cells by CRISPR/Cas9-mediated site-specific integration into regions with distinct histone modifications. Front. Bioeng. Biotechnol. 10, 1010719. https://doi.org/10.3389/fbioe.2022.1010719.

Herzog, R.W., VandenDriessche, T., Ozelo, M.C., 2023. First hemophilia B gene therapy approved: more than two decades in the making. Mol. Ther. 31, 1–2. https://doi.org/10.1016/j.ymthe.2022.12.001.

Heumos, L., Schaar, A.C., Lance, C., Litinetskaya, A., Drost, F., Zappia, L., Lücken, M.D., Strobl, D.C., Henao, J., Curion, F., Consortium, Single-cell Best Practices, Schiller, H. B., Theis, F.J., 2023. Best practices for single-cell analysis across modalities. Nat. Rev. Genet. 24, 550–572. https://doi.org/10.1038/s41576-023-00586-w.

Heusel, M., Bludau, I., Rosenberger, G., Hafen, R., Frank, M., Banaei-Esfahani, A., van Drogen, A., Collins, B.C., Gstaiger, M., Aebersold, R., 2019. Complex-centric proteome profiling by SEC-SWATH-MS. Mol. Syst. Biol. 15, e8438. https://doi.org/10.15252/msb.20188438.

High-dose AAV gene therapy deaths, 2020. Nat. Biotechnol. 38, 910. https://doi.org/10.1038/s41587-020-0642-9.

Hilliard, W., Lee, K.H., 2021. Systematic identification of safe harbor regions in the CHO genome through a comprehensive epigenome analysis. Biotechnol. Bioeng. 118, 659–675. https://doi.org/10.1002/bit.27599.

Houldcroft, C.J., Beale, M.A., Breuer, J., 2017. Clinical and biological insights from viral genome sequencing. Nat. Rev. Microbiol. 15, 183–192. https://doi.org/10.1038/nrmicro.2016.182.

Huang, Y.-F., Aoki, K., Akase, S., Ishihara, M., Liu, Y.-S., Yang, G., Kizuka, Y., Mizumoto, S., Tiemeyer, M., Gao, X.-D., Aoki-Kinoshita, K.F., Fujita, M., 2021. Global mapping of glycosylation pathways in human-derived cells. Dev. Cell 56, 1195–1209.e7. https://doi.org/10.1016/j.devcel.2021.02.023.

Huhn, S., Chang, M., Kumar, A., Liu, R., Jiang, B., Betenbaugh, M., Lin, H., Nyberg, G., Du, Z., 2022. Chromosomal instability drives convergent and divergent evolution toward advantageous inherited traits in mammalian CHO bioproduction lineages. iScience 25, 104074. https://doi.org/10.1016/j.isci.2022.104074.

Hwang, B., Lee, J.H., Bang, D., 2018. Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp. Mol. Med. 50, 1–14. https://doi.org/10.1038/s12276-018-0071-8.

Igolkina, A.A., Zinkevich, A., Karandasheva, K.O., Popov, A.A., Selifanova, M.V., Nikolaeva, D., Tkachev, V., Penzar, D., Nikitin, D.M., Buzdin, A., 2019. H3K4me3, H3K9ac, H3K27ac, H3K27me3 and H3K9me3 histone tags suggest distinct regulatory evolution of open and condensed chromatin landmarks. Cells 8, 1034. https://doi.org/10.3390/cells8091034.

Inwood, S., Buehler, E., Betenbaugh, M., Lal, M., Shiloach, J., 2018. Identifying HIPK1 as target of miR-22-3p enhancing recombinant protein production from HEK 293 cell by using microarray and HTP siRNA screen. Biotechnol. J. 13. https://doi.org/10.1002/biot.201700342.

Inwood, S., Abaandou, L., Betenbaugh, M., Shiloach, J., 2020. Improved protein expression in HEK293 cells by over-expressing miR-22 and knocking-out its target gene, HIPK1. New Biotechnol. 54, 28–33. https://doi.org/10.1016/j.nbt.2019.08.004.

Iurlaro, R., Muñoz-Pinedo, C., 2016. Cell death induced by endoplasmic reticulum stress. FEBS J. 283, 2640–2652. https://doi.org/10.1111/febs.13598.

Jeltsch, A., Broche, J., Bashtrykov, P., 2018. Molecular processes connecting DNA methylation patterns with DNA methyltransferases and histone modifications in mammalian genomes. Genes (Basel) 9, 566. https://doi.org/10.3390/genes9110566.

Jiang, D., Jarrett, H.W., Haskins, W.E., 2009. Methods for proteomic analysis of transcription factors. J. Chromatogr. A 1216, 6881–6889. https://doi.org/10.1016/j.chroma.2009.08.044.

Joo, Y., Kim, H., Baik, J.Y., 2023. Chinese hamster ovary cell line instability: causes, mitigation, and prediction. Biotechnol. Bioprocess Eng. 28, 750–760. https://doi.org/10.1007/s12257-023-0120-6.

Joshi, P.R.H., Venereo-Sanchez, A., Chahal, P.S., Kamen, A.A., 2021. Advancements in molecular design and bioprocessing of recombinant adeno-associated virus gene delivery vectors using the insect-cell baculovirus expression platform. Biotechnol. J. 16, e2000021. https://doi.org/10.1002/biot.202000021.

Jovic, D., Liang, X., Zeng, H., Lin, L., Xu, F., Luo, Y., 2022. Single-cell RNA sequencing technologies and applications: a brief overview. Clin. Transl. Med. 12, e694. https://doi.org/10.1002/ctm2.694.

Kahlig, C.-I., Moser, S., Micutkova, L., Grillari, J., Kraus, B., Hernandez Bort, J.A., 2024. Enhancement of rAAV titers via inhibition of the interferon signaling cascade in transfected HEK293 suspension cultures. Biotechnol. J. 19, e2300672. https://doi.org/10.1002/biot.202300672.

Kang, M., Ko, E., Mersha, T.B., 2022. A roadmap for multi-omics data integration using deep learning. Brief. Bioinform. 23. https://doi.org/10.1093/bib/bbab454.

Kaur, H., 2021. Characterization of glycosylation in monoclonal antibodies and its importance in therapeutic antibody development. Crit. Rev. Biotechnol. 41, 300–315. https://doi.org/10.1080/07388551.2020.1869684.

Ke, M., Elshenawy, B., Sheldon, H., Arora, A., Buffa, F.M., 2022. Single cell RNA-sequencing: a powerful yet still challenging technology to study cellular heterogeneity. Bioessays 44, e2200084. https://doi.org/10.1002/bies.202200084.

Kildegaard, H.F., Baycin-Hizal, D., Lewis, N.E., Betenbaugh, M.J., 2013. The emerging CHO systems biology era: harnessing the 'omics revolution for biotechnology. Curr. Opin. Biotechnol. 24, 1102–1107. https://doi.org/10.1016/j.copbio.2013.02.007.

Kim, M., O'Callaghan, P.M., Droms, K.A., James, D.C., 2011. A mechanistic understanding of production instability in CHO cell lines expressing recombinant monoclonal antibodies. Biotechnol. Bioeng. 108, 2434–2446. https://doi.org/10.1002/bit.23189.

Koh, T.-C., Lee, Y.-Y., Chang, S.-Q., Nissom, P.M., 2009. Identification and expression analysis of miRNAs during batch culture of HEK-293 cells. J. Biotechnol. 140, 149–155. https://doi.org/10.1016/j.jbiotec.2009.01.021.

Konstantinidis, K.T., Tiedje, J.M., 2005. Genomic insights that advance the species definition for prokaryotes. Proc. Natl. Acad. Sci. USA 102, 2567–2572. https://doi.org/10.1073/pnas.0409727102.

Kotidis, P., Kontoravdi, C., 2020. Harnessing the potential of artificial neural networks for predicting protein glycosylation. Metab. Eng. Commun. 10, e00131. https://doi.org/10.1016/j.mec.2020.e00131.

Kurien, B.T., Scofield, R.H., 2015. Western blotting: an introduction. Methods Mol. Biol. 1312, 17–30. https://doi.org/10.1007/978-1-4939-2694-7_5.

Kuzmin, D.A., Shutova, M.V., Johnston, N.R., Smith, O.P., Fedorin, V.V., Kukushkin, Y.S., van der Loo, J.C.M., Johnstone, E.C., 2021. The clinical landscape for AAV gene therapies. Nat. Rev. Drug Discov. 20, 173–174. https://doi.org/10.1038/d41573-021-00017-7.

la Karottki, K.J.C., Hefzi, H., Li, S., Pedersen, L.E., Spahn, P.N., Joshi, C., Ruckerbauer, D., Bort, J.A.H., Thomas, A., Lee, J.S., Borth, N., Lee, G.M., Kildegaard, H.F., Lewis, N.E., 2021. A metabolic CRISPR-Cas9 screen in Chinese hamster ovary cells identifies glutamine-sensitive genes. Metab. Eng. 66, 114–122. https://doi.org/10.1016/j.ymben.2021.03.017.

Lähnemann, D., Köster, J., Szczurek, E., McCarthy, D.J., Hicks, S.C., Robinson, M.D., Vallejos, C.A., Campbell, K.R., Beerenwinkel, N., Mahfouz, A., Pinello, L., Skums, P., Stamatakis, A., Attolini, C.S.-O., Aparicio, S., Baaijens, J., Balvert, M., de Barbanson, B., Cappuccio, A., Corleone, G., Dutilh, B.E., Florescu, M., Guryev, V., Holmer, R., Jahn, K., Lobo, T.J., Keizer, E.M., Khatri, I., Kielbasa, S.M., Korbel, J.O., Kozlov, A.M., Kuo, T.-H., Lelieveldt, B.P.F., Mandoiu, I.I., Marioni, J.C., Marschall, T., Mölder, F., Niknejad, A., Raczkowski, L., Reinders, M., de Ridder, J., Saliba, A.-E., Somarakis, A., Stegle, O., Theis, F.J., Yang, H., Zelikovsky, A., McHardy, A.C., Raphael, B.J., Shah, S.P., Schönhuth, A., 2020. Eleven grand

challenges in single-cell data science. Genome Biol. 21, 31. https://doi.org/10.1186/s13059-020-1926-6.
Lamichhane, S., Sen, P., Dickens, A.M., Hyötyläinen, T., Orešič, M., 2018. An overview of metabolomics data analysis: Current tools and future perspectives. In: Jaumot, J., Bedia, C., Tauler, R. (Eds.), Comprehensive Analytical Chemistry, Comprehensive Analytical Chemistry, vol. 82. Elsevier, pp. 387–413. https://doi.org/10.1016/bs.coac.2018.07.001.
Lavado-García, J., González-Domínguez, I., Cervera, L., Jorge, I., Vázquez, J., Gòdia, F., 2020a. Molecular characterization of the coproduced extracellular vesicles in HEK293 during virus-like particle production. J. Proteome Res. 19, 4516–4532. https://doi.org/10.1021/acs.jproteome.0c00581.
Lavado-García, J., Jorge, I., Cervera, L., Vázquez, J., Gòdia, F., 2020b. Multiplexed quantitative proteomic analysis of HEK293 provides insights into molecular changes associated with the cell density effect, transient transfection, and virus-like particle production. J. Proteome Res. 19, 1085–1099. https://doi.org/10.1021/acs.jproteome.9b00601.
Lavado-García, J., Pérez-Rubio, P., Cervera, L., Gòdia, F., 2022. The cell density effect in animal cell-based bioprocessing: questions, insights and perspectives. Biotechnol. Adv. 60, 108017. https://doi.org/10.1016/j.biotechadv.2022.108017.
Lazarou, G., Chelliah, V., Small, B.G., Walker, M., van der Graaf, P.H., Kierzek, A.M., 2020. Integration of omics data sources to inform mechanistic modeling of immune-oncology therapies: a tutorial for clinical pharmacologists. Clin. Pharmacol. Ther. 107, 858–870. https://doi.org/10.1002/cpt.1786.
Lee, C.M., Barber, G.P., Casper, J., Clawson, H., Diekhans, M., Gonzalez, J.N., Hinrichs, A.S., Lee, B.T., Nassar, L.R., Powell, C.C., Raney, B.J., Rosenbloom, K.R., Schmelter, D., Speir, M.L., Zweig, A.S., Haussler, D., Haeussler, M., Kuhn, R.M., Kent, W.J., 2020. UCSC genome browser enters 20th year. Nucleic Acids Res. 48, D756–D761. https://doi.org/10.1093/nar/gkz1012.
Lenstra, T.L., Rodriguez, J., Chen, H., Larson, D.R., 2016. Transcription dynamics in living cells. Annu. Rev. Biophys. 45, 25–47. https://doi.org/10.1146/annurev-biophys-062215-010838.
Lequin, R.M., 2005. Enzyme immunoassay (EIA)/enzyme-linked immunosorbent assay (ELISA). Clin. Chem. 51, 2415–2418. https://doi.org/10.1373/clinchem.2005.051532.
Li, F., Vijayasankaran, N., Shen, A.Y., Kiss, R., Amanullah, A., 2010. Cell culture processes for monoclonal antibody production. MAbs 2, 466–479. https://doi.org/10.4161/mabs.2.5.12720.
Li, X., Ma, J., Leng, L., Han, M., Li, M., He, F., Zhu, Y., 2022. MoGCN: a multi-omics integration method based on graph convolutional network for cancer subtype analysis. Front. Genet. 13, 806842. https://doi.org/10.3389/fgene.2022.806842.
Lin, Y.-C., Boone, M., Meuris, L., Lemmens, I., Van Roy, N., Soete, A., Reumers, J., Moisse, M., Plaisance, S., Drmanac, R., Chen, J., Speleman, F., Lambrechts, D., Van de Peer, Y., Tavernier, J., Callewaert, N., 2014. Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. Nat. Commun. 5, 4767. https://doi.org/10.1038/ncomms5767.
Lin, D., Yalamanchili, H.B., Zhang, X., Lewis, N.E., Alves, C.S., Groot, J., Arnsdorf, J., Bjørn, S.P., Wulff, T., Voldborg, B.G., Zhou, Y., Zhang, B., 2020. CHOmics: a web-based tool for multi-omics data analysis and interactive visualization in CHO cell lines. PLoS Comput. Biol. 16, e1008498. https://doi.org/10.1371/journal.pcbi.1008498.
Lin, Y.-C., Lu, M., Cai, W., Hu, W.-S., 2024. Comparative transcriptomic and proteomic kinetic analysis of adeno-associated virus production systems. Appl. Microbiol. Biotechnol. 108, 385. https://doi.org/10.1007/s00253-024-13203-5.
Liu, X., Voulgaropoulou, F., Chen, R., Johnson, P.R., Clark, K.R., 2000. Selective rep-cap gene amplification as a mechanism for high-titer recombinant AAV production from stable cell lines. Mol. Ther. 2, 394–403. https://doi.org/10.1006/mthe.2000.0132.
Lowe, R., Shirley, N., Bleackley, M., Dolan, S., Shafee, T., 2017. Transcriptomics technologies. PLoS Comput. Biol. 13, e1005457. https://doi.org/10.1371/journal.pcbi.1005457.
Lu, M., Lee, Z., Hu, W.-S., 2024. Multi-omics kinetic analysis of recombinant adeno-associated virus production by plasmid transfection of HEK293 cells. Biotechnol. Prog. 40, e3428. https://doi.org/10.1002/btpr.3428.
Lucas, M.C., Novoa, E.M., 2023. Long-read sequencing in the era of epigenomics and epitranscriptomics. Nat. Methods 20, 25–29. https://doi.org/10.1038/s41592-022-01724-8.
Ludwig, C., Gillet, L., Rosenberger, G., Amon, S., Collins, B.C., Aebersold, R., 2018. Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. Mol. Syst. Biol. 14, e8126. https://doi.org/10.15252/msb.20178126.
Malm, M., Saghaleyni, R., Lundqvist, M., Giudici, M., Chotteau, V., Field, R., Varley, P.G., Hatton, D., Grassi, L., Svensson, T., Nielsen, J., Rockberg, J., 2020. Evolution from adherent to suspension: systems biology of HEK293 cell line development. Sci. Rep. 10, 18996. https://doi.org/10.1038/s41598-020-76137-8.
Malm, M., Kuo, C.-C., Barzadd, M.M., Mebrahtu, A., Wistbacka, N., Razavi, R., Volk, A.-L., Lundqvist, M., Kotol, D., Tegel, H., Hober, S., Edfors, F., Gräslund, T., Chotteau, V., Field, R., Varley, P.G., Roth, R.G., Lewis, N.E., Hatton, D., Rockberg, J., 2022. Harnessing secretory pathway differences between HEK293 and CHO to rescue production of difficult to express proteins. Metab. Eng. 72, 171–187. https://doi.org/10.1016/j.ymben.2022.03.009.
Marcus, K., Rabilloud, T., 2020. How do the different proteomic strategies cope with the complexity of biological regulations in a multi-omic world? Critical appraisal and suggestions for improvements. Proteomes 8, 23. https://doi.org/10.3390/proteomes8030023.
Mardinoglu, A., Agren, R., Kampf, C., Asplund, A., Nookaew, I., Jacobson, P., Walley, A. J., Froguel, P., Carlsson, L.M., Uhlen, M., Nielsen, J., 2013. Integration of clinical data with a genome-scale metabolic model of the human adipocyte. Mol. Syst. Biol. 9, 649. https://doi.org/10.1038/msb.2013.5.
Mardinoglu, A., Agren, R., Kampf, C., Asplund, A., Uhlen, M., Nielsen, J., 2014. Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. Nat. Commun. 5, 3083. https://doi.org/10.1038/ncomms4083.
Marouga, R., David, S., Hawkins, E., 2005. The development of the DIGE system: 2D fluorescence difference gel analysis technology. Anal. Bioanal. Chem. 382, 669–678. https://doi.org/10.1007/s00216-005-3126-3.
Martínez-Monge, I., Albiol, J., Lecina, M., Liste-Calleja, L., Miret, J., Solà, C., Cairó, J.J., 2019. Metabolic flux balance analysis during lactate and glucose concomitant consumption in HEK293 cell cultures. Biotechnol. Bioeng. 116, 388–404. https://doi.org/10.1002/bit.26858.
Marx, N., Eisenhut, P., Weinguny, M., Klanert, G., Borth, N., 2022. How to train your cell - towards controlling phenotypes by harnessing the epigenome of Chinese hamster ovary production cell lines. Biotechnol. Adv. 56, 107924. https://doi.org/10.1016/j.biotechadv.2022.107924.
Masson, H.O., la Cour Karottki, K.J., Tat, J., Hefzi, H., Lewis, N.E., 2023. From observational to actionable: rethinking omics in biologics production. Trends Biotechnol. 41, 1127–1138. https://doi.org/10.1016/j.tibtech.2023.03.009.
Maznyi, G., 2023. Development of a high-throughput screening approach to identify production enhancers of Adeno-associated virus. https://doi.org/10.20381/RUOR-29685.
Mcallister, R., Schofield, C., Pettman, G., Mannix, C., 2002. Adaptation of recombinant HEK-293 cells to growth in serum free suspension. In: Bernard, A., Griffiths, B., Noé, W., Wurm, F. (Eds.), Animal Cell Technology: Products from Cells, Cells as Products: Proceedings of the 16th ESACT Meeting April 25–29, 1999, Lugano, Switzerland. Springer Netherlands, Dordrecht, pp. 367–369. https://doi.org/10.1007/0-306-46875-1_79.
McGettigan, P.A., 2013. Transcriptomics in the RNA-seq era. Curr. Opin. Chem. Biol. 17, 4–11. https://doi.org/10.1016/j.cbpa.2012.12.008.
Mehrmohamadi, M., Sepehri, M.H., Nazer, N., Norouzi, M.R., 2021. A comparative overview of epigenomic profiling methods. Front. Cell Dev. Biol. 9, 714687. https://doi.org/10.3389/fcell.2021.714687.
Mendell, J.R., Al-Zaidy, S.A., Rodino-Klapac, L.R., Goodspeed, K., Gray, S.J., Kay, C.N., Boye, S.L., Boye, S.E., George, L.A., Salabarria, S., Corti, M., Byrne, B.J., Tremblay, J. P., 2021. Current clinical applications of in vivo gene therapy with AAVs. Mol. Ther. 29, 464–488. https://doi.org/10.1016/j.ymthe.2020.12.007.
Miller, N.L., Clark, T., Raman, R., Sasisekharan, R., 2021. Glycans in virus-host interactions: a structural perspective. Front. Mol. Biosci. 8, 666756. https://doi.org/10.3389/fmolb.2021.666756.
Moritz, B., Becker, P.B., Göpfert, U., 2015. CMV promoter mutants with a reduced propensity to productivity loss in CHO cells. Sci. Rep. 5, 16952. https://doi.org/10.1038/srep16952.
Moritz, B., Woltering, L., Becker, P.B., Göpfert, U., 2016. High levels of histone H3 acetylation at the CMV promoter are predictive of stable expression in Chinese hamster ovary cells. Biotechnol. Prog. 32, 776–786. https://doi.org/10.1002/btpr.2271.
Morozova, O., Hirst, M., Marra, M.A., 2009. Applications of new sequencing technologies for transcriptome analysis. Annu. Rev. Genomics Hum. Genet. 10, 135–151. https://doi.org/10.1146/annurev-genom-082908-145957.
Mullard, A., 2021. Gene therapy community grapples with toxicity issues, as pipeline matures. Nat. Rev. Drug Discov. 20, 804–805. https://doi.org/10.1038/d41573-021-00164-x.
Mulukutla, B.C., Kale, J., Kalomeris, T., Jacobs, M., Hiller, G.W., 2017. Identification and control of novel growth inhibitors in fed-batch cultures of Chinese hamster ovary cells. Biotechnol. Bioeng. 114, 1779–1790. https://doi.org/10.1002/bit.26313.
Nadeau, I., Jacob, D., Perrier, M., Kamen, A., 2000a. 293SF metabolic flux analysis during cell growth and infection with an adenoviral vector. Biotechnol. Prog. 16, 872–884. https://doi.org/10.1021/bp000098l.
Nadeau, I., Sabatié, J., Koehl, M., Perrier, M., Kamen, A., 2000b. Human 293 cell metabolism in low glutamine-supplied culture: interpretation of metabolic changes through metabolic flux analysis. Metab. Eng. 2, 277–292. https://doi.org/10.1006/mben.2000.0152.
Nadeau, I., Gilbert, P.A., Jacob, D., Perrier, M., Kamen, A., 2002. Low-protein medium affects the 293SF central metabolism during growth and infection with adenovirus. Biotechnol. Bioeng. 77, 91–104. https://doi.org/10.1002/bit.10128.
Nagy, A., Chakrabarti, L., Kurasawa, J., Mulagapati, S.H.R., Devine, P., Therres, J., Chen, Z., Schmelzer, A.E., 2023. Engineered CHO cells as a novel AAV production platform for gene therapy delivery. Sci. Rep. 13, 19210. https://doi.org/10.1038/s41598-023-46298-3.
Narimatsu, Y., Joshi, H.J., Nason, R., Van Coillie, J., Karlsson, R., Sun, L., Ye, Z., Chen, Y.-H., Schjoldager, K.T., Steentoft, C., Furukawa, S., Bensing, B.A., Sullam, P. M., Thompson, A.J., Paulson, J.C., Büll, C., Adema, G.J., Mandel, U., Hansen, L., Bennett, E.P., Varki, A., Vakhrushev, S.Y., Yang, Z., Clausen, H., 2019. An atlas of human glycosylation pathways enables display of the human glycome by gene engineered cells. Mol. Cell 75, 394–407.e5. https://doi.org/10.1016/j.molcel.2019.05.017.
Naso, M.F., Tomkowicz, B., Perry III, W.L., Strohl, W.R., 2017. Adeno-associated virus (AAV) as a vector for gene therapy. BioDrugs 31, 317–334. https://doi.org/10.1007/s40259-017-0234-5.
Nguyen, T.N.T., Sha, S., Hong, M.S., Maloney, A.J., Barone, P.W., Neufeld, C., Wolfrum, J., Springs, S.L., Sinskey, A.J., Braatz, R.D., 2021. Mechanistic model for production of recombinant adeno-associated virus via triple transfection of HEK293 cells. Mol. Ther. Meth. Clin. Dev. 21, 642–655. https://doi.org/10.1016/j.omtm.2021.04.006.
Nguyen, T.N.T., Park, D., Canova, C.T., Sangerman, J., Srinivasan, P., Ou, R.W., Barone, P., Neufeld, C., Wolfrum, J., Springs, S., Sinskey, A., Braatz, R., 2024.

Continuous production of rAAV via an intensified transient transfection process. Authorea Inc. https://doi.org/10.22541/au.172704244.48626142/v1
Nicora, G., Vitali, F., Dagliati, A., Geifman, N., Bellazzi, R., 2020. Integrated multi-omics analyses in oncology: a review of machine learning methods and tools. Front. Oncol. 10, 1030. https://doi.org/10.3389/fonc.2020.01030.
Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A.V., Mikheenko, A., Vollger, M. R., Altemose, N., Uralsky, L., Gershman, A., Aganezov, S., Hoyt, S.J., Diekhans, M., Logsdon, G.A., Alonge, M., Antonarakis, S.E., Borchers, M., Bouffard, G.G., Brooks, S. Y., Caldas, G.V., Chen, N.-C., Cheng, H., Chin, C.-S., Chow, W., de Lima, L.G., Dishuck, P.C., Durbin, R., Dvorkina, T., Fiddes, I.T., Formenti, G., Fulton, R.S., Fungtammasan, A., Garrison, E., Grady, P.G.S., Graves-Lindsay, T.A., Hall, I.M., Hansen, N.F., Hartley, G.A., Haukness, M., Howe, K., Hunkapiller, M.W., Jain, C., Jain, M., Jarvis, E.D., Kerpedjiev, P., Kirsche, M., Kolmogorov, M., Korlach, J., Kremitzki, M., Li, H., Maduro, V.V., Marschall, T., McCartney, A.M., McDaniel, J., Miller, D.E., Mullikin, J.C., Myers, E.W., Olson, N.D., Paten, B., Peluso, P., Pevzner, P.A., Porubsky, D., Potapova, T., Rogaev, E.I., Rosenfeld, J.A., Salzberg, S. L., Schneider, V.A., Sedlazeck, F.J., Shafin, K., Shew, C.J., Shumate, A., Sims, Y., Smit, A.F.A., Soto, D.C., Sović, I., Storer, J.M., Streets, A., Sullivan, B.A., Thibaud-Nissen, F., Torrance, J., Wagner, J., Walenz, B.P., Wenger, A., Wood, J.M.D., Xiao, C., Yan, S.M., Young, A.C., Zarate, S., Surti, U., McCoy, R.C., Dennis, M.Y., Alexandrov, I.A., Gerton, J.L., O'Neill, R.J., Timp, W., Zook, J.M., Schatz, M.C., Eichler, E.E., Miga, K.H., Phillippy, A.M., 2022. The complete sequence of a human genome. Science 376, 44–53. https://doi.org/10.1126/science.abj6987.
O'Flaherty, R., Bergin, A., Flampouri, E., Mota, L.M., Obaidi, I., Quigley, A., Xie, Y., Butler, M., 2020. Mammalian cell culture for production of recombinant proteins: a review of the critical steps in their biomanufacturing. Biotechnol. Adv. 43, 107552.
Ogata, N., Nishimura, A., Matsuda, T., Kubota, M., Omasa, T., 2021. Single-cell transcriptome analyses reveal heterogeneity in suspension cultures and clonal markers of CHO-K1 cells. Biotechnol. Bioeng. 118, 944–951. https://doi.org/10.1002/bit.27624.
Ong, S.-E., Mann, M., 2007. Stable isotope labeling by amino acids in cell culture for quantitative proteomics. Methods Mol. Biol. 359, 37–52. https://doi.org/10.1007/978-1-59745-255-7_3.
Opdam, S., Richelle, A., Kellman, B., Li, S., Zielinski, D.C., Lewis, N.E., 2017. A systematic evaluation of methods for tailoring genome-scale metabolic models. Cell Syst. 4, 318–329.e6. https://doi.org/10.1016/j.cels.2017.01.010.
Osterlehner, A., Simmeth, S., Göpfert, U., 2011. Promoter methylation and transgene copy numbers predict unstable protein production in recombinant Chinese hamster ovary cell lines. Biotechnol. Bioeng. 108, 2670–2681. https://doi.org/10.1002/bit.23216.
Palsson, B., Zengler, K., 2010. The challenges of integrating multi-omic data sets. Nat. Chem. Biol. 6, 787–789. https://doi.org/10.1038/nchembio.462.
Paredes, V., Park, J.S., Jeong, Y., Yoon, J., Baek, K., 2013. Unstable expression of recombinant antibody during long-term culture of CHO cells is accompanied by histone H3 hypoacetylation. Biotechnol. Lett. 35, 987–993. https://doi.org/10.1007/s10529-013-1168-8.
Passi, A., Tibocha-Bonilla, J.D., Kumar, M., Tec-Campos, D., Zengler, K., Zuniga, C., 2021. Genome-scale metabolic modeling enables in-depth understanding of big data. Metabolites 12, 14. https://doi.org/10.3390/metabo12010014.
Patel, N.A., Anderson, C.R., Terkildsen, S.E., Davis, R.C., Pack, L.D., Bhargava, S., Clarke, H.R.G., 2018. Antibody expression stability in CHO clonally derived cell lines and their subclones: role of methylation in phenotypic and epigenetic heterogeneity. Biotechnol. Prog. 34, 635–649. https://doi.org/10.1002/btpr.2655.
Patra, A.T., Tan, E., Kok, Y.J., Ng, S.K., Bi, X., 2024. Temporal insights into molecular and cellular responses during rAAV production in HEK293T cells. Mol. Ther. Methods Clin. Dev. 32, 101278. https://doi.org/10.1016/j.omtm.2024.101278.
Pereira, S., Kildegaard, H.F., Andersen, M.R., 2018. Impact of CHO metabolism on cell growth and protein production: An overview of toxic and inhibiting metabolites and nutrients. Biotechnol. J. 13, 1700499.
Perez-Riverol, Y., Zorin, A., Dass, G., Vu, M.-T., Xu, P., Glont, M., Vizcaíno, J.A., Jarnuczak, A.F., Petryszak, R., Ping, P., Hermjakob, H., 2019. Quantifying the impact of public omics data. Nat. Commun. 10, 3512. https://doi.org/10.1038/s41467-019-11461-w.
Pérez-Rubio, P., Lavado-García, J., Bosch-Molist, L., Romero, E.L., Cervera, L., Gòdia, F., 2024. Extracellular vesicle depletion and UGCG overexpression mitigate the cell density effect in HEK293 cell culture transfection. Mol. Ther. Methods Clin. Dev. 32, 101190. https://doi.org/10.1016/j.omtm.2024.101190.
Petronis, A., 2010. Epigenetics as a unifying principle in the aetiology of complex traits and diseases. Nature 465, 721–727. https://doi.org/10.1038/nature09230.
Picard, M., Scott-Boyer, M.-P., Bodein, A., Périn, O., Droit, A., 2021. Integration strategies of multi-omics data for machine learning analysis. Comput. Struct. Biotechnol. J. 19, 3735–3746. https://doi.org/10.1016/j.csbj.2021.06.030.
Pistek, M., Kahlig, C.-I., Hackl, M., Unterthurner, S., Kraus, B., Grabherr, R., Grillari, J., Hernandez Bort, J.A., 2023. Comprehensive mRNA-sequencing-based characterization of three HEK-293 cell lines during an rAAV production process for gene therapy applications. Biotechnol. J. 18, e2200513. https://doi.org/10.1002/biot.202200513.
Plotnikova, O., Baranova, A., Skoblov, M., 2019. Comprehensive analysis of human microRNA-mRNA interactome. Front. Genet. 10, 933. https://doi.org/10.3389/fgene.2019.00933.
Poirion, O.B., Jing, Z., Chaudhary, K., Huang, S., Garmire, L.X., 2021. DeepProg: an ensemble of deep-learning and machine-learning models for prognosis prediction using multi-omics data. Genome Med. 13, 112. https://doi.org/10.1186/s13073-021-00930-x.
Qiao, C., Li, J., Skold, A., Zhang, X., Xiao, X., 2002. Feasibility of generating adeno-associated virus packaging cell lines containing inducible adenovirus helper genes. J. Virol. 76, 1904–1913. https://doi.org/10.1128/jvi.76.4.1904-1913.2002.
Quek, L.-E., Dietmair, S., Hanscho, M., Martínez, V.S., Borth, N., Nielsen, L.K., 2014. Reducing Recon 2 for steady-state flux analysis of HEK cell culture. J. Biotechnol. 184, 172–178. https://doi.org/10.1016/j.jbiotec.2014.05.021.
Rare Diseases, 2022. IFPMA. URL [WWW Document]. https://www.ifpma.org/areas-of-work/advancing-sustainable-health-systems/rare-diseases/ (accessed 5.2.23).
Reel, P.S., Reel, S., Pearson, E., Trucco, E., Jefferson, E., 2021. Using machine learning approaches for multi-omics data analysis: a review. Biotechnol. Adv. 49, 107739. https://doi.org/10.1016/j.biotechadv.2021.107739.
Reese, B., Swanson, J., Storck, A., Llc, L.J.M., Usa, W.I., 2023. RevIT™ AAV enhancer: Rev up AAV genome production in upstream manufacturing [WWW document]. URL. https://scientifix.com.au/media/promotions/pdf/m/i/mirus-bio-revit_aav_enhancer-white-paper-sfx.pdf.
Regenerative Medicine: Disrupting the Status Quo, 2022. Alliance for Regenerative Medicine. URL. https://alliancerm.org/sector-report/2021-annual-report/ [WWW Document] (accessed 5.2.23).
Rivera, C.M., Ren, B., 2013. Mapping human epigenomes. Cell 155, 39–55. https://doi.org/10.1016/j.cell.2013.09.011.
Rodrigues, A.F., Formas-Oliveira, A.S., Bandeira, V.S., Alves, P.M., Hu, W.S., Coroadinha, A.S., 2013. Metabolic pathways recruited in the production of a recombinant enveloped virus: mining targets for process and cell engineering. Metab. Eng. 20, 131–145. https://doi.org/10.1016/j.ymben.2013.10.001.
Rose, D.R., 2012. Structure, mechanism and inhibition of Golgi α-mannosidase II. Curr. Opin. Struct. Biol. 22, 558–562. https://doi.org/10.1016/j.sbi.2012.06.005.
Rumachik, N.G., Malaker, S.A., Poweleit, N., Maynard, L.H., Adams, C.M., Leib, R.D., Cirolia, G., Thomas, D., Stamnes, S., Holt, K., Sinn, P., May, A.P., Paulk, N.K., 2020. Methods matter: standard production platforms for recombinant AAV produce chemically and functionally distinct vectors. Mol. Ther. Methods Clin. Dev. 18, 98–118. https://doi.org/10.1016/j.omtm.2020.05.018.
Ryu, J.Y., Kim, H.U., Lee, S.Y., 2015. Reconstruction of genome-scale human metabolic models using omics data. Integr. Biol. 7, 859–868. https://doi.org/10.1039/c5ib00002e.
Saghaleyni, R., Malm, M., Moruzzi, N., Zrimec, J., Razavi, R., Wistbacka, N., Thorell, H., Pintar, A., Hober, A., Edfors, F., Chotteau, V., Berggren, P.-O., Grassi, L., Zelezniak, A., Svensson, T., Hatton, D., Nielsen, J., Robinson, J.L., Rockberg, J., 2022. Enhanced metabolism and negative regulation of ER stress support higher erythropoietin production in HEK293 cells. Cell Rep. 39, 110936. https://doi.org/10.1016/j.celrep.2022.110936.
Samoudi, M., Kuo, C.C., Robinson, C.M., Shams-Ud-Doha, K., Schinn, S.M., Kol, S., Weiss, L., Petersen Bjorn, S., Voldborg, B.G., Rosa Campos, A., Lewis, N.E., 2020. In situ detection of protein interactions for recombinant therapeutic enzymes. Biotechnol. Bioeng. 118, 890–904. https://doi.org/10.1002/bit.27621.
Satam, H., Joshi, K., Mangrolia, U., Waghoo, S., Zaidi, G., Rawool, S., Thakare, R.P., Banday, S., Mishra, A.K., Das, G., Malonia, S.K., 2023. Next-generation sequencing technology: current trends and advancements. Biology (Basel) 12, 997. https://doi.org/10.3390/biology12070997.
Scarrott, J.M., Johari, Y.B., Pohle, T.H., Liu, P., Mayer, A., James, D.C., 2023. Increased recombinant adeno-associated virus production by HEK293 cells using small molecule chemical additives. Biotechnol. J. 18, e2200450. https://doi.org/10.1002/biot.202200450.
Schaack, G.A., Sullivan, O.M., Mehle, A., 2023. Identifying protein-protein interactions by proximity Biotinylation with AirID and splitAirID. Curr Protoc. 3, e702. https://doi.org/10.1002/cpz1.702.
Schinn, S.-M., Morrison, C., Wei, W., Zhang, L., Lewis, N.E., 2021. A genome-scale metabolic network model and machine learning predict amino acid concentrations in Chinese Hamster ovary cell cultures. Biotechnol. Bioeng. 118, 2118–2123. https://doi.org/10.1002/bit.27714.
Schneider, V.A., Graves-Lindsay, T., Howe, K., Bouk, N., Chen, H.-C., Kitts, P.A., Murphy, T.D., Pruitt, K.D., Thibaud-Nissen, F., Albracht, D., Fulton, R.S., Kremitzki, M., Magrini, V., Markovic, C., McGrath, S., Steinberg, K.M., Auger, K., Chow, W., Collins, J., Harden, G., Hubbard, T., Pelan, S., Simpson, J.T., Threadgold, G., Torrance, J., Wood, J.M., Clarke, L., Koren, S., Boitano, M., Peluso, P., Li, H., Chin, C.-S., Phillippy, A.M., Durbin, R., Wilson, R.K., Flicek, P., Eichler, E.E., Church, D.M., 2017. Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. Genome Res. 27, 849–864. https://doi.org/10.1101/gr.213611.116.
Sebastião, M.J., Hoffman, M., Escandell, J., Tousi, F., Zhang, J., Figueroa, B., DeMaria, C., Gomes-Alves, P., 2023. Identification of mispairing omic signatures in Chinese hamster ovary (CHO) cells producing a tri-specific antibody. Biomedicines 11, 2890. https://doi.org/10.3390/biomedicines11112890.
Sha, S., Huang, Z., Wang, Z., Yoon, S., 2018. Mechanistic modeling and applications for CHO cell culture development and production. Curr. Opin. Chem. Eng. 22, 54–61. https://doi.org/10.1016/j.coche.2018.08.010.
Sha, S., Maloney, A.J., Katsikis, G., Nguyen, T.N.T., Neufeld, C., Wolfrum, J., Barone, P. W., Springs, S.L., Manalis, S.R., Sinskey, A.J., Braatz, R.D., 2021. Cellular pathways of recombinant adeno-associated virus production for gene therapy. Biotechnol. Adv. 49, 107764. https://doi.org/10.1016/j.biotechadv.2021.107764.
Sharma, V., Hecker, N., Roscito, J.G., Foerster, L., Langer, B.E., Hiller, M., 2018. A genomics approach reveals insights into the importance of gene losses for mammalian adaptations. Nat. Commun. 9, 1215. https://doi.org/10.1038/s41467-018-03667-1.
Shin, S., Kim, S.H., Shin, S.W., Grav, L.M., Pedersen, L.E., Lee, J.S., Lee, G.M., 2020. Comprehensive analysis of genomic safe harbors as target sites for stable expression

of the heterologous gene in HEK293 cells. ACS Synth. Biol. 9, 1263–1269. https://doi.org/10.1021/acssynbio.0c00097.
Shupe, J., Zhang, A., Odenwelder, D.C., Dobrowsky, T., 2022. Gene therapy: challenges in cell culture scale-up. Curr. Opin. Biotechnol. 75, 102721. https://doi.org/10.1016/j.copbio.2022.102721.
Singh, V., Pandey, S., Bhardwaj, A., 2022. From the reference human genome to human pangenome: premise, promise and challenge. Front. Genet. 13, 1042550. https://doi.org/10.3389/fgene.2022.1042550.
Sorourian, S., Behzad Behbahani, A., Rafiei Dehbidi, G., Zare, F., Farajnia, S., Najafi, H., Safari, F., 2024. Omics and CRISPR in CHO bioproduction: a state-of-the-art review. Biotechnol. Bioprocess Eng. 29, 465–482. https://doi.org/10.1007/s12257-024-00026-4.
Spahn, P.N., Zhang, X., Hu, Q., Lu, H., Hamaker, N.K., Hefzi, H., Li, S., Kuo, C.-C., Huang, Y., Lee, J.C., Davis, A.J., Ly, P., Lee, K.H., Lewis, N.E., 2022. Restoration of DNA repair mitigates genome instability and increases productivity of Chinese hamster ovary cells. Biotechnol. Bioeng. 119, 963–982. https://doi.org/10.1002/bit.28016.
Srinivasan, P., Canova, C.T., Sha, S., Nguyen, T.N.T., Joseph, J., Sangerman, J., Maloney, A.J., Katsikis, G., Ou, R.W., Hong, M.S., Ng, J., Yuan, A., Antov, D., Song, S., Chen, W., Neufeld, C., Wolfrum, J.M., Barone, P.W., Sinskey, A.J., Springs, S.L., Braatz, R.D., 2024. Multidose transient transfection of human embryonic kidney 293 cells modulates recombinant adeno-associated virus2/5 rep protein expression and influences the enrichment fraction of filled capsids. Biotechnol. Bioeng. 121, 3694–3714. https://doi.org/10.1002/bit.28828.
Srivastava, A., Mallela, K.M.G., Deorkar, N., Brophy, G., 2021. Manufacturing challenges and rational formulation development for AAV viral vectors. J. Pharm. Sci. 110, 2609–2624. https://doi.org/10.1016/j.xphs.2021.03.024.
Stolfa, G., Smonskey, M.T., Boniface, R., Hachmann, A.-B., Gulde, P., Joshi, A.D., Pierce, A.P., Jacobia, S.J., Campbell, A., 2018. CHO-omics review: the impact of current and emerging technologies on Chinese hamster ovary based bioproduction. Biotechnol. J. 13, e1700227. https://doi.org/10.1002/biot.201700227.
Strasser, L., Boi, S., Guapo, F., Donohue, N., Barron, N., Rainbow-Fletcher, A., Bones, J., 2021. Proteomic landscape of adeno-associated virus (AAV)-producing HEK293 cells. Int. J. Mol. Sci. 22, 11499. https://doi.org/10.3390/ijms222111499.
Strobel, B., Klauser, B., Hartig, J.S., Lamla, T., Gantner, F., Kreuz, S., 2015. Riboswitch-mediated attenuation of transgene cytotoxicity increases adeno-associated virus vector yields in HEK-293 cells. Mol. Ther. 23, 1582–1591. https://doi.org/10.1038/mt.2015.123.
Sumit, M., Dolatshahi, S., Chu, A.-H.A., Cote, K., Scarcelli, J.J., Marshall, J.K., Cornell, R. J., Weiss, R., Lauffenburger, D.A., Mulukutla, B.C., Figueroa Jr., B., 2019. Dissecting N-glycosylation dynamics in Chinese hamster ovary cells fed-batch cultures using time course omics analyses. iScience 12, 102–120. https://doi.org/10.1016/j.isci.2019.01.006.
Tan, S.Z., Begley, P., Mullard, G., Hollywood, K.A., Bishop, P.N., 2016. Introduction to metabolomics and its applications in ophthalmology. Eye 30, 773–783. https://doi.org/10.1038/eye.2016.37.
Tan, E., Chin, C.S.H., Lim, Z.F.S., Ng, S.K., 2021. HEK293 cell line as a platform to produce recombinant proteins and viral vectors. Front. Bioeng. Biotechnol. 9, 796991. https://doi.org/10.3389/fbioe.2021.796991.
Tang, X., Huang, Y., Lei, J., Luo, H., Zhu, X., 2019. The single-cell sequencing: new developments and medical applications. Cell Biosci. 9, 53. https://doi.org/10.1186/s13578-019-0314-y.
Tatalick, L.M., Gerard, C.J., Takeya, R., Price, D.N., Thorne, B.A., Wyatt, L.M., Anklesaria, P., 2005. Safety characterization of HeLa-based cell substrates used in the manufacture of a recombinant adeno-associated virus-HIV vaccine. Vaccine, Gene Therapy: Therapeutic Mechanisms and Strategies 23, 2628–2638. https://doi.org/10.1016/j.vaccine.2004.11.027.
Thiele, I., Swainston, N., Fleming, R.M.T., Hoppe, A., Sahoo, S., Aurich, M.K., Haraldsdottir, H., Mo, M.L., Rolfsson, O., Stobbe, M.D., Thorleifsson, S.G., Agren, R., Bölling, C., Bordel, S., Chavali, A.K., Dobson, P., Dunn, W.B., Endler, L., Hala, D., Hucka, M., Hull, D., Jameson, D., Jamshidi, N., Jonsson, J.J., Juty, N., Keating, S., Nookaew, I., Le Novère, N., Malys, N., Mazein, A., Papin, J.A., Price, N.D., Selkov Sr., E., Sigurdsson, M.I., Simeonidis, E., Sonnenschein, N., Smallbone, K., Sorokin, A., van Beek, J.H.G.M., Weichart, D., Goryanin, I., Nielsen, J., Westerhoff, H.V., Kell, D.B., Mendes, P., Palsson, B.Ø., 2013. A community-driven global reconstruction of human metabolism. Nat. Biotechnol. 31, 419–425. https://doi.org/10.1038/nbt.2488.
Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A.K.A., Hamon, C., 2003. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal. Chem. 75, 1895–1904. https://doi.org/10.1021/ac0262560.
Thorne, B.A., Takeya, R.K., Peluso, R.W., 2009. Manufacturing recombinant adeno-associated viral vectors from producer cell clones. Hum. Gene Ther. 20, 707–714. https://doi.org/10.1089/hum.2009.070.
Thuerauf, D.J., Morrison, L., Glembotski, C.C., 2004. Opposing roles for ATF6alpha and ATF6beta in endoplasmic reticulum stress response gene induction. J. Biol. Chem. 279, 21078–21084. https://doi.org/10.1074/jbc.M400713200.
Tighe, P.J., Ryder, R.R., Todd, I., Fairclough, L.C., 2015. ELISA in the multiplex era: potentials and pitfalls. Proteomi. Clin. Apps 9, 406–422. https://doi.org/10.1002/prca.201400130.
Torres, M., Betts, Z., Scholey, R., Elvin, M., Place, S., Hayes, A., Dickson, A.J., 2023. Long term culture promotes changes to growth, gene expression, and metabolism in CHO cells that are independent of production stability. Biotechnol. Bioeng. 120, 2389–2402. https://doi.org/10.1002/bit.28399.
Tóth, R., Mészáros, I., Hüser, D., Forró, B., Marton, S., Olasz, F., Bányai, K., Heilbronn, R., Zádori, Z., 2019. Methylation status of the adeno-associated virus type 2 (AAV2). Viruses 11, 38. https://doi.org/10.3390/v11010038.
Tsopanoglou, A., Jiménez del Val, I., 2021. Moving towards an era of hybrid modelling: advantages and challenges of coupling mechanistic and data-driven models for upstream pharmaceutical bioprocesses. Curr. Opin. Chem. Eng. 32, 100691. https://doi.org/10.1016/j.coche.2021.100691.
Turunen, T.A., Väänänen, M.-A., Ylä-Herttuala, S., 2018. Epigenomics. In: Vasan, R.S., Sawyer, D.B. (Eds.), Encyclopedia of Cardiovascular Research and Medicine. Elsevier, Oxford, pp. 258–265. https://doi.org/10.1016/B978-0-12-809657-4.99575-9.
Tworig, J., Grafton, F., Hörer, M., Reid, C.A., Mandegar, M.A., 2024. Transcriptomics-informed pharmacology identifies epigenetic and cell cycle regulators as enhancers of AAV production. bioRxiv. https://doi.org/10.1101/2024.06.14.599118.
Tyers, M., Mann, M., 2003. From genomics to proteomics. Nature 422, 193–197. https://doi.org/10.1038/nature01510.
Tzani, I., Herrmann, N., Carillo, S., Spargo, C.A., Hagan, R., Barron, N., Bones, J., Shannon Dillmore, W., Clarke, C., 2021. Tracing production instability in a clonally derived CHO cell line using single-cell transcriptomics. Biotechnol. Bioeng. 118, 2016–2030. https://doi.org/10.1002/bit.27715.
van Meer, G., Voelker, D.R., Feigenson, G.W., 2008. Membrane lipids: where they are and how they behave. Nat. Rev. Mol. Cell Biol. 9, 112–124. https://doi.org/10.1038/nrm2330.
Väremo, L., Scheele, C., Broholm, C., Mardinoglu, A., Kampf, C., Asplund, A., Nookaew, I., Uhlén, M., Pedersen, B.K., Nielsen, J., 2016. Proteome- and transcriptome-driven reconstruction of the human myocyte metabolic network and its use for identification of markers for diabetes. Cell Rep. 14, 1567. https://doi.org/10.1016/j.celrep.2016.01.054.
Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P., Ballew, R.M., Huson, D.H., Wortman, J.R., Zhang, Q., Kodira, C.D., Zheng, X.H., Chen, L., Skupski, M., Subramanian, G., Thomas, P.D., Zhang, J., Gabor Miklos, G.L., Nelson, C., Broder, S., Clark, A.G., Nadeau, J., McKusick, V.A., Zinder, N., Levine, A. J., Roberts, R.J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A.E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T.J., Higgins, M.E., Ji, R.R., Ke, Z., Ketchum, K.A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G.V., Milshina, N., Moore, H.M., Naik, A.K., Narayan, V.A., Neelam, B., Nusskern, D., Rusch, D.B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M.L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y.H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N.N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J.F., Guigó, R., Campbell, M.J., Sjolander, K.V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y.H., Coyne, M., Dahlke, C., Deslattes Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D., Wu, M., Xia, A., Zandieh, A., Zhu, X., 2001. The sequence of the human genome. Science 291, 1304–1351. https://doi.org/10.1126/science.1058040.
Virgolini, N., Hagan, R., Correia, R., Silvano, M., Fernandes, S., Alves, P.M., Clarke, C., Roldão, A., Isidro, I.A., 2023a. Transcriptome analysis of Sf9 insect cells during production of recombinant adeno-associated virus. Biotechnol. J. 18, e2200466. https://doi.org/10.1002/biot.202200466.
Virgolini, N., Silvano, M., Hagan, R., Correia, R., Alves, P.M., Clarke, C., Roldão, A., Isidro, I.A., 2023b. Impact of dual-baculovirus infection on the Sf9 insect cell transcriptome during rAAV production using single-cell RNA-seq. Biotechnol. Bioeng. 120, 2588–2600. https://doi.org/10.1002/bit.28377.
Vito, D., Smales, C.M., 2018. The long non-coding RNA transcriptome landscape in CHO cells under batch and fed-batch conditions. Biotechnol. J. e1800122. https://doi.org/10.1002/biot.201800122.
Wang, K.C., Chang, H.Y., 2018. Epigenomics: technologies and applications. Circ. Res. 122, 1191–1199. https://doi.org/10.1161/CIRCRESAHA.118.310998.
Wang, Z., Gerstein, M., Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. 10, 57–63. https://doi.org/10.1038/nrg2484.

Wang, D., Tai, P.W.L., Gao, G., 2019. Adeno-associated virus vector as a platform for gene therapy delivery. Nat. Rev. Drug Discov. 18, 358–378. https://doi.org/10.1038/s41573-019-0012-9.

Wang, T., Shao, W., Huang, Z., Tang, H., Zhang, J., Ding, Z., Huang, K., 2021. MOGONET integrates multi-omics data using graph convolutional networks allowing patient classification and biomarker identification. Nat. Commun. 12, 3445. https://doi.org/10.1038/s41467-021-23774-w.

Wang, Y., Fu, Q., Lee, Y.S., Sha, S., Yoon, S., 2023. Transcriptomic features reveal molecular signatures associated with recombinant adeno-associated virus production in HEK293 cells. Biotechnol. Prog. 39, e3346. https://doi.org/10.1002/btpr.3346.

Wang, J.-H., Gessler, D.J., Zhan, W., Gallagher, T.L., Gao, G., 2024. Adeno-associated virus as a delivery vector for gene therapy of human diseases. Signal Transduct. Target. Ther. 9, 78. https://doi.org/10.1038/s41392-024-01780-w.

Warburton, P.E., Sebra, R.P., 2023. Long-read DNA sequencing: recent advances and remaining challenges. Annu. Rev. Genomics Hum. Genet. 24, 109–132. https://doi.org/10.1146/annurev-genom-101722-103045.

Wiese, S., Reidegeld, K.A., Meyer, H.E., Warscheid, B., 2007. Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research. Proteomics 7, 340–350. https://doi.org/10.1002/pmic.200600422.

Wilkins, M.R., Pasquali, C., Appel, R.D., Ou, K., Golaz, O., Sanchez, J.-C., Yan, J.X., Gooley, A.A., Hughes, G., Humphery-Smith, I., Williams, K.L., Hochstrasser, D.F., 1996. From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. Biotechnology 14, 61–65. https://doi.org/10.1038/nbt0196-61.

Wippermann, A., Noll, T., 2017. DNA methylation in CHO cells. J. Biotechnol. 258, 206–210. https://doi.org/10.1016/j.jbiotec.2017.08.005.

Wolf, J.B.W., 2013. Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial. Mol. Ecol. Resour. 13, 559–572. https://doi.org/10.1111/1755-0998.12109.

Wright, J.F., 2023. AAV vector production: troublesome host innate responses in another setting. Mol. Ther. Methods Clin. Dev. 28, 412–413. https://doi.org/10.1016/j.omtm.2023.02.008.

Wuest, D.M., Harcum, S.W., Lee, K.H., 2012. Genomics in mammalian cell culture bioprocessing. Biotechnol. Adv. 30, 629–638. https://doi.org/10.1016/j.biotechadv.2011.10.010.

Wurm, F.M., 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. Nat. Biotechnol. 22, 1393–1398. https://doi.org/10.1038/nbt1026.

Wurm, F.M., Wurm, M.J., 2017. Cloning of CHO cells, productivity and genetic stability—a discussion. Processes 5, 20. https://doi.org/10.3390/pr5020020.

Xiao, S., Chen, Y.-C., Betenbaugh, M.J., Martin, S.E., Shiloach, J., 2015. MiRNA mimic screen for improved expression of functional neurotensin receptor from HEK293 cells. Biotechnol. Bioeng. 112, 1632–1643. https://doi.org/10.1002/bit.25567.

Xiao, S., Chen, Y.C., Buehler, E., Mandal, S., Mandal, A., Betenbaugh, M., Park, M.H., Martin, S., Shiloach, J., 2016. Genome-scale RNA interference screen identifies antizyme 1 (OAZ1) as a target for improvement of recombinant protein production in mammalian cells. Biotechnol. Bioeng. 113, 2403–2415. https://doi.org/10.1002/bit.26017.

Xie, Y., Butler, M., 2023. N-glycomic profiling of capsid proteins from adeno-associated virus serotypes. Glycobiology 34, cwad074. https://doi.org/10.1093/glycob/cwad074.

Xu, L., Seki, M., 2020. Recent advances in the detection of base modifications using the Nanopore sequencer. J. Hum. Genet. 65, 25–33. https://doi.org/10.1038/s10038-019-0679-0.

Xue, W., Fulco, C., Sha, S., Alden, N., Panteli, J., Hossler, P., Warren, J., 2024. Adeno-associated virus perfusion enhanced expression: a commercially scalable, high titer, high quality producer cell line process. Mol. Ther. Methods Clin. Dev. 32, 101266. https://doi.org/10.1016/j.omtm.2024.101266.

Yates, J.R., 2011. A century of mass spectrometry: from atoms to proteomes. Nat. Methods 8, 633–637. https://doi.org/10.1038/nmeth.1659.

Yates, A.D., Achuthan, P., Akanni, W., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I.M., Azov, A.G., Bennett, R., Bhai, J., Billis, K., Boddu, S., Marugán, J. C., Cummins, C., Davidson, C., Dodiya, K., Fatima, R., Gall, A., Giron, C.G., Gil, L., Grego, T., Haggerty, L., Haskell, E., Hourlier, T., Izuogu, O.G., Janacek, S.H., Juettemann, T., Kay, M., Lavidas, I., Le, T., Lemos, D., Martinez, J.G., Maurel, T., McDowall, M., McMahon, A., Mohanan, S., Moore, B., Nuhn, M., Oheh, D.N., Parker, A., Parton, A., Patricio, M., Sakthivel, M.P., Abdul Salam, A.I., Schmitt, B.M., Schuilenburg, H., Sheppard, D., Sycheva, M., Szuba, M., Taylor, K., Thormann, A., Threadgold, G., Vullo, A., Walts, B., Winterbottom, A., Zadissa, A., Chakiachvili, M., Flint, B., Frankish, A., Hunt, S.E., IIsley, G., Kostadima, M., Langridge, N., Loveland, J.E., Martin, F.J., Morales, J., Mudge, J.M., Muffato, M., Perry, E., Ruffier, M., Trevanion, S.J., Cunningham, F., Howe, K.L., Zerbino, D.R., Flicek, P., 2020. Ensembl 2020. Nucleic Acids Res. 48, D682–D688. https://doi.org/10.1093/nar/gkz966.

Zampieri, G., Vijayakumar, S., Yaneske, E., Angione, C., 2019. Machine and deep learning meet genome-scale metabolic modeling. PLoS Comput. Biol. 15, e1007084. https://doi.org/10.1371/journal.pcbi.1007084.

Zeh, N., Schlossbauer, P., Raab, N., Klingler, F., Handrick, R., Otte, K., 2021. Cell line development for continuous high cell density biomanufacturing: exploiting hypoxia for improved productivity. Metab. Eng. Commun. 13, e00181. https://doi.org/10.1016/j.mec.2021.e00181.

Zehetner, L., Széliová, D., Kraus, B., Graninger, M., Zanghellini, J., Hernandez Bort, J.A., 2023. Optimizing VLP production in gene therapy: opportunities and challenges for in silico modeling. Biotechnol. J. 18, e2200636. https://doi.org/10.1002/biot.202200636.

Zehetner, L., Széliová, D., Kraus, B., Hernandez Bort, J.A., Zanghellini, J., 2024a. Logistic PCA explains differences between genome-scale metabolic models in terms of metabolic pathways. PLoS Comput. Biol. 20, e1012236. https://doi.org/10.1371/journal.pcbi.1012236.

Zehetner, L., Széliová, D., Kraus, B., Hernandez Bort, J.A., Zanghellini, J., 2024b. Improving HEK293-based AAV-production using GSMMs, and a multi-omics approach. bioRxiv. https://doi.org/10.1101/2024.10.10.617556.

Zhang, Y., Baycin-Hizal, D., Kumar, A., Priola, J., Bahri, M., Heffner, K.M., Wang, M., Han, X., Bowen, M.A., Betenbaugh, M.J., 2017. High-throughput lipidomic and transcriptomic analysis to compare SP2/0, CHO, and HEK-293 mammalian cell lines. Anal. Chem. 89, 1477–1485. https://doi.org/10.1021/acs.analchem.6b02984.

Zhang, L., Lu, Q., Chang, C., 2020. Epigenetics in health and disease. In: Chang, C., Lu, Q. (Eds.), Epigenetics in Allergy and Autoimmunity, 1253. Springer Singapore, Singapore, pp. 3–55. https://doi.org/10.1007/978-981-15-3449-2_1.

Zhang, L., Gao, J., Zhang, X., Wang, X., Wang, T., Zhang, J., 2024. Current strategies for the development of high-yield HEK293 cell lines. Biochem. Eng. J. 205, 109279. https://doi.org/10.1016/j.bej.2024.109279.

Zhou, Y., Priya, S., Ong, J.Y., 2024. Characterizing glycosylation of adeno-associated virus serotype 9 capsid proteins generated from HEK293 cells through glycopeptide mapping and released glycan analysis. Microorganisms 12, 946. https://doi.org/10.3390/microorganisms12050946.