Received: 27 December 2021 | Revised: 19 April 2022 | Accepted: 21 April 2022

DOI: 10.1002/bit.28128

ARTICLE

BIOTECHNOLOGY AND BIOENGINEERING WILEY

Comprehensive assessment of host cell protein expression after extended culture and bioreactor production of CHO cell lines

Nathaniel K. Hamaker | Lie Min | Kelvin H. Lee

Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, Delaware, USA

Correspondence
Kelvin H. Lee, Department of Chemical and Biomolecular Engineering, Ammon-Pinizzotto Biopharmaceutical Innovation Center, University of Delaware, Newark, Delaware, 590 Ave 1743, Newark, DE 19713, USA.
Email: KHL@udel.edu

Funding information
National Institute of Standards and Technology, Grant/Award Number: 70NANB17H002; National Science Foundation, Grant/Award Numbers: 1624698, 1736123

Abstract

The biomanufacturing industry is advancing toward continuous processes that will involve longer culture durations and older cell ages. These upstream trends may bring unforeseen challenges for downstream purification due to fluctuations in host cell protein (HCP) levels. To understand the extent of HCP expression instability exhibited by Chinese hamster ovary (CHO) cells over these time scales, an industry-wide consortium collaborated to develop a study to characterize age-dependent changes in HCP levels across 30, 60, and 90 cell doublings, representing a period of approximately 60 days. A monoclonal antibody (mAb)-producing cell line with bulk productivity up to 3 g/L in a bioreactor was aged in parallel with its parental CHO-K1 host. Subsequently, both cell types at each age were cultivated in an automated bioreactor system to generate harvested cell culture fluid (HCCF) for HCP analysis. More than 1500 HCPs were quantified using complementary proteomic techniques, two-dimensional electrophoresis (2DE) and liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS). While up to 13% of proteins showed variable expression with age, more changes were observed when comparing between the two cell lines with up to 47% of HCPs differentially expressed. A small subset (50 HCPs) with age-dependent expression were previously reported to be problematic as high-risk and/or difficult-to-remove impurities; however, the vast majority of these were downregulated with age. Our findings suggest that HCP expression changes over this time scale may not be as dramatic and pose as great of a challenge to downstream processing as originally expected but that monitoring of variably expressed problematic HCPs remains critical.

KEYWORDS
biomanufacturing, CHO cells, host cell proteins, monoclonal antibody, proteomics

1 | INTRODUCTION

Of the mammalian hosts available for therapeutic protein production, Chinese hamster ovary (CHO) cells are by far the most commonly utilized as they possess a variety of desirable qualities for the production of monoclonal antibodies (mAbs) (J. Y. Kim et al., 2012; Walsh, 2018). A key manufacturing challenge in the use of any cell-based production process is the generation of impurities originating from the native proteome of the host organism, collectively referred to as host cell proteins (HCPs). HCPs are process-related impurities that may co-purify with a therapeutic protein product, and their concentration is considered to be a critical quality attribute (CQA)

that influences product safety and efficacy (Baik et al., 2019). While exact guidance for acceptable HCP levels in the final product are not defined, 100 ppm is considered an acceptable benchmark for commercial processes (Chon & Zarbis-Papastoitsis, 2011). Achieving this level of impurity clearance remains a primary challenge, especially as increased burden has been pushed downstream due to recent advances in upstream process intensification (i.e., increased cell densities), which have enabled multigram per liter product titers, but have also contributed to increased HCP levels (Gilgunn & Bones, 2018; Kornecki et al., 2017).

A typical downstream processing platform for mAb purification consists of a product isolation step using Protein A affinity chromatography followed by product polishing using additional chromatography steps such as ion exchange chromatography (IEX) or hydrophobic interaction chromatography (HIC) (Shukla & Thömmes, 2010). Because mAbs are secreted extracellularly and harvested from cell culture supernatant, the impurity burden for mAb processes mainly consists of HCPs that are natively secreted (the secretome) but also includes some cytoplasmic proteins released due to cell lysis or shearing (Hogwood et al., 2014). Consequently, HCPs are a heterogeneous mixture of thousands of unique proteins at varying amounts and with diverse biophysical properties (Kumar et al., 2015). Enzyme-linked immunosorbent assay (ELISA) is the traditional method for quantifying and monitoring total HCP levels throughout downstream processing (Zhu-Shimoni et al., 2014) but is not capable of tracking a large number of individual HCPs. Liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) and two-dimensional electrophoresis (2DE) have enabled proteomic profiling with quantitative output and have been applied to determine the relative or absolute abundance of individual HCPs (Valente et al., 2018). These methods have been used to track HCP clearance through multiple stages of downstream processing (Chiverton et al., 2016; Falkenberg et al., 2019; Zhang et al., 2014), as well as identify and quantify residual HCPs present at ppm levels in final drug substances (Falkenberg et al., 2019; Kreimer et al., 2017; Molden et al., 2021).

Substantial effort has been dedicated to cataloging difficult-to-remove CHO HCPs that persist through chromatography steps in standard downstream purification platforms because of co-elution or product-association mechanisms. Multiple studies have identified the HCPs that are most likely to co-purify with mAbs in Protein A eluate and have found appreciable overlap across many of the mAbs tested despite differences in upstream process conditions (Chiverton et al., 2016; Farrell et al., 2015; Zhang et al., 2014, 2016). Other work has elucidated HCP clearance vulnerabilities for polishing chromatography operations, finding several HCPs that can be difficult to remove across an entire downstream process (Joucla et al., 2013; Levy et al., 2016; Pezzini et al., 2011). The establishment of co-purification mechanisms and lists of frequently encountered HCPs has helped to guide diverse strategies in both upstream and downstream processes to aid HCP elimination. Cell line engineering tools have been leveraged to inactivate native genes encoding detrimental proteins, such as lipoprotein lipase (LPL) and cathepsin D (CTSD) (Chiu et al., 2017; Dovgan et al., 2021). Alternatively, HCP tracking with MS has informed the development of Protein A wash steps that use additives to disrupt mAb-HCP interactions, thus improving clearance (Aboulaich et al., 2014; Bee et al., 2015; Farrell et al., 2015).

Some HCPs can have adverse effects on product quality or formulation, while others have the potential to elicit immunogenicity or a direct biological function in human patients (Jones et al., 2021). The presence of endogenous proteins exhibiting proteolytic activity such as CTSD and HtrA serine peptidase (HTRA1) can lead to mAb fragmentation or clipping, which reduces product recovery, therapeutic efficacy, and shelf life (Bee et al., 2015; Dorai et al., 2011). Additionally, some HCPs such as hamster phospholipase B-like 2 (PLBL2) can cause an immune response in patients (Fischer et al., 2017). By determining which HCPs pose the greatest risk, quality systems can be implemented to specifically monitor those HCPs when assessing downstream platform robustness.

Divergent CHO cell lines are generally similar in breadth and total number of HCP species detected in culture supernatant samples (Yuk et al., 2015). However, individual HCP concentrations can vary significantly between cell lines even under similar growth conditions with some HCPs exceeding a 10-fold difference in abundance (Hogwood et al., 2016; Madsen et al., 2015; Park et al., 2017; Zhang et al., 2014). Cell viability at harvest has been repeatedly identified as a crucial factor affecting the distribution and total amount of HCPs due to elevated levels of intracellular proteins released upon cell death or breakage (Farrell et al., 2015; Hogwood et al., 2016; Tait et al., 2012; Wilson et al., 2019). While the impact of process conditions such as media composition (Gronemeyer et al., 2016) and operating temperature (Goey et al., 2017) have been studied to determine subsets of differentially abundant HCPs for mAb-producing cells, the effect of cell age has received less attention.

With the current push toward continuous bioprocessing and intensified cell culture systems, downstream processes must be able to cope with potential changes in HCP expression, such as an increase in the abundance of certain difficult-to-remove HCPs (Croughan et al., 2015). CHO cell lines are widely accepted to be unstable, exhibiting chromosomal rearrangements and epigenetic variation that has often been studied in the context of decreased specific productivity of the recombinant product over time but extends to changes in the expression of endogenous proteins as well (Baumann et al., 2019). Indeed, our group has previously documented quantitative changes in extracellular HCPs for a null CHO cell line across 1 year in culture (Valente et al., 2015). Building upon these findings, we sought to apply 2DE and SWATH-MS proteomics to quantify differential expression of an expanded list of HCPs in harvested cell culture fluid (HCCF) for CHO cells over a time period that may be representative in length of a commercial seed train used to support intensified continuous processes or conventional bioreactor production campaigns. The mAb-producing cell line and its parental host were both cultivated in an automated bioreactor system to simulate a scaled-down production run. For the more than 1500 proteins quantified by proteomics, only a small subset (13% or less) showed differential expression between any two time points. Of

these, 50 proteins have been previously characterized as problematic with the majority showing decreased abundance over time. We expect these findings can be generalized to other recombinant CHO cell lines and could support platform development for improved clearance in downstream purification.

2 | METHODS

2.1 | CHO cell lines and aging

Representatives from more than a dozen companies collaborated to design the key attributes of this study to ensure maximum industry impact, including the selection of cell lines, cell ages of interest, and viability cut-offs as well as interpretation of results. An overview of the study design is shown in Supporting Information: Figure S1. Two CHO-K1 cell lines, a producer cell line expressing the VRC01 broadly neutralizing HIV antibody and its parental host, were donated by the National Institutes of Health (NIH) for this study as part of the Advanced Mammalian Biomanufacturing Innovation Center (AMBIC) consortium. Briefly, the VRC01 mAb-expressing cell line clone A11 was previously derived from the null CHO-K1 host by stable integration and methotrexate (MTX) mediated amplification of a cassette encoding the immunoglobulin G1 (IgG1) heavy chain and light chain of VRC01 (Wu et al., 2010). A starting cell bank of the host cell line was made from cells cultured for roughly the same amount of time that the producer cell line underwent selection and expansion before generation of a master cell bank so that both cell lines would initiate aging from a similar reference point at a population doubling level (PDL) of 0.

Vials of each cell line were thawed in duplicate and aged in parallel in the absence of MTX selection. Cells were cultured at 37°C with 5% CO_2 and 80% relative humidity in Actipro medium with Poloxamer-188 (Cytiva) supplemented with 6mM L-glutamine (Gibco). Cells were passaged continuously at 2–3-day intervals by seeding at a viable cell density (VCD) of 0.3–0.4 × 10^6 cells ml^{-1} into a working volume of 20–30 ml in 125-ml polycarbonate Erlenmeyer flasks (Corning) at 135 rpm (25 mm throw diameter). During aging, the cumulative PDL was tracked using Equation (1), where VCD_i is the seed density, VCD_f is the viable cell density at the time of passaging, PDL_i is the population doubling level from the previous passage, and PDL_f is the cumulative PDL:

$$PDL_f = \frac{ln\left(\frac{VCD_f}{VCD_i}\right)}{ln(2)} + PDL_i. \quad (1)$$

Cultures were aged until PDL 90 and banked every 10 PDL to make working cell banks. Banking vials were prepared using 10 × 10^6 viable cells suspended in 1 ml media supplemented with 10% dimethyl sulfoxide (Sigma-Aldrich) and cryopreserved in liquid nitrogen until later use. Cell density, viability, and size were determined using measurements from a Vi-Cell XR automated cell counter (Beckman Coulter).

2.2 | Fed-batch bioreactor runs

Working cell bank vials of independently aged duplicates of each cell line at ages PDL 20, 50, and 80 were thawed concurrently and passaged in flasks for 6–7 days to PDL 30, 60, and 90, respectively. A 3-day expansion step (N-1 stage) was performed in an ambr®250 High Throughput (Sartorius Stedim) automated bioreactor system and used for inoculation of production cultures (N stage) in ambr®250 bioreactor vessels at 0.4 × 10^6 cells ml^{-1}. Disposable standard mammalian vessels (part number: 001-5G25) were used for all runs with process setpoints of 36.5°C, pH 7, and 50% dissolved oxygen (DO) calibrated at process temperature and 100% air saturation. Impellers were initially set to 300 rpm up-stir. The DO setpoint was maintained with a control loop that first increased the sparge rate of pure oxygen to 100 ml min^{-1} before increasing the stir speed up to a maximum of 600 rpm. EX-CELL antifoam (Sigma-Aldrich) was added as needed as per manufacturer's usage recommendations. The pH was controlled by CO_2 sparging and automated additions of 1 M sodium carbonate. All gases were supplied via submerged sparger with no overlays applied. The working volume within bioreactors was maintained at 210 ml throughout the fed-batch with daily sampling balanced by bolus additions of Cell Boost 7a and 7b (Cytiva) starting on Day 3 (Supporting Information: Table S1). Starting on Day 5, bolus additions of a 45% Glucose (Sigma-Aldrich) solution were added daily as needed to avoid depletion based on cellular glucose consumption rates. Daily samples were taken to measure glucose and lactate levels with a YSI 2950 biochemistry analyzer (Xylem), as well as VCD and cell viability with a Vi-Cell XR (Beckman Coulter). HCCF samples for proteomics analysis were harvested based on a cell viability cut-off of 80% which occurred on Days 11 and 9 for the VRC01 and host cells, respectively. Cells and debris were first separated from the supernatant by centrifugation at 3400 g for 30 min before clarification with a 0.2 µm vacuum filtration unit (Thermo Fisher Scientific). Clarified HCCF supernatant samples were frozen at −80°C until further analysis. Fed-batch cultures for VRC01 and host cells were terminated on Days 14 and 11, respectively.

2.3 | Postharvest quantitation of mab and HCP titer

Starting on Day 4 of culture, VRC01 mAb quantitation was performed using an Octet RED96E system (ForteBio) using Protein A biosensors (part number: 18-5010). Samples were diluted 10-fold in phosphate-buffered saline with 0.1% bovine serum albumin (BSA; Sigma-Aldrich), and concentrations were interpolated from a 4-parameter logistic (4PL) regression model for a set of IgG calibrators. Between readings, probes were regenerated with 10 mM glycine (Biorad) at pH 1.5. Total protein concentrations were measured by Bradford assay using a Pierce Coomassie Plus assay kit (Thermo Fisher Scientific) with a BSA standard curve as per manufacturer's instructions. For VRC01 fed-batch samples, HCP concentrations were estimated from the Bradford assay results by subtracting the

mAb titer from the total protein concentration. Total immunoreactive HCP concentrations were determined using a CHO HCP ELISA 3 G kit (Cygnus) according to manufacturer's instructions. Samples were diluted 100–100,000-fold in Sample Diluent (Cygnus; part number: I028) with a consistent matrix for all samples. After stopping the colorimetric reaction, absorbance values were measured using a SpectraMax i3x (Molecular Devices) plate reader. Before all protein measurements, thawed samples were sonicated for 2–4 min at 120 W to disrupt and resuspend protein aggregates.

Cell-specific productivity (qP) was determined using the relationship in Equation (2), where the integral of VCD as a function of time can be approximated using the trapezoidal rule as the integrated viable cell density (IVCD), which is directly proportional to the concentration of protein ([Protein]):

$$[Protein]\ (t) = qP_{Protein}\int_0^t VCD(t)dt \approx qP_{Protein} \times IVCD(t). \quad (2)$$

Thus, the qP value of a protein is equal to the slope obtained from the linear regression when plotting [Protein] against IVCD. Values for qP were determined in units of pg cell^{-1} day^{-1} (pcd) from the slope of [Protein] in μg ml^{-1} as a function of IVCD in 1×10^6 viable cells · day ml^{-1}. Only data from fed-batch time points that showed an approximately linear relationship between IVCD and protein concentration were used, which generally corresponded to the exponential growth phase before any decline in culture viability.

2.4 | Quantitative SWATH-MS proteomics

CHO HCPs were precipitated from end-of-run HCCF samples with methanol as previously described (Valente et al., 2014), and pellets were resuspended in triethylammonium bicarbonate buffer (Sigma-Aldrich) via sonication. Total protein concentrations were determined by Bradford assay. Trypsin digestion was performed as described previously (Hou et al., 2013; Valente et al., 2015). Briefly, HCP samples (100 μg) were reduced in 2.5 mM tris(2-carboxyethyl) phosphine (Thermo Fisher Scientific) at 60°C for one hour, alkylated with 7.5 mM iodoacetamide (Sigma-Aldrich), digested with trypsin (Promega) at an enzyme to substrate mass ratio of 1:50, and then acidified with 20% formic acid (FA, Thermo Fisher Scientific).

Digested samples (50 μg) were prepared for LC-MS/MS with OMIX C18 pipette tips (Agilent) per manufacturer's instructions with substitution of formic acid for heptafluorobutyric acid. Tips were conditioned with 50% acetonitrile (ACN; Fisher Scientific) and equilibrated with 1% FA. Washes were performed with 0.1% FA and elution with 50% ACN, 0.1% FA. Tips were regenerated and binding/elution cycles repeated twice with the same tip for each sample. Eluates were dried with a SpeedVac™ vacuum concentrator (Thermo Fisher Scientific).

Dried samples were resuspended in 50 μl 2% ACN, 0.1% FA. Each LC-MS/MS analysis used an injection of 5 μg digested HCPs. LC-MS/MS analysis was performed on a TripleTOF 6600 (Sciex) equipped with an Eksigent nano 425 LC operating in microLC flow mode. LC separation was performed on a ChromXP C18CL column (3 μm, 120 Å, 150 mm × 0.3 mm; Sciex) with mobile phase A (0.1% FA) and mobile phase B (0.1% FA in ACN) at a flow rate of 5 μL/min. A program of 3% B to 25% B over 68 min, 25% B to 35% B over 5 min, 35% B to 80% B over 2 min, and 80% B for 3 min was used to elute peptides.

Data-dependent acquisition (DDA) was performed in positive ion mode with a full MS1 scan over a mass range of 400–1250 m/z with a scan time of 250 ms followed by MS/MS over a mass range of 100–1500 m/z with a scan time of 50 ms. The top 30 precursor ions were selected for fragmentation. Data independent acquisition (DIA) sequential windowed acquisition of all theoretical fragment ion spectra (SWATH) experiments were performed with an MS1 full scan followed by 64 MS/MS acquisitions with variable window sizes (Gillet et al., 2012; Husson et al., 2018; Ludwig et al., 2018).

Database searches were performed as previously described (Hou et al., 2013; Valente et al., 2015). ProteinPilot (v5.1; Sciex) was used to submit DDA data for searches against a local copy of the Chinese hamster RefSeq. 2019 database (Rupp et al., 2018), supplemented with common contaminants, using the Paragon Algorithm (Sciex) search engine. Search parameters were specified to include cysteine modifications by iodoacetamide, digestion with trypsin, and a detected protein threshold at 10%. Replicate DDA data from both VRC01 and host cell line HCCF were submitted for a combined search in preparation for building a spectral library.

The group file from ProteinPilot was used to generate a spectral library in Skyline (v20.2.0.343; MacCoss Lab, University of Washington) (MacLean et al., 2010). Triplicate SWATH data for each sample was extracted and processed as a combined unit using Skyline command-line. Endogenous CHO HCP peptides were used for RT alignment. Peaks were automatically picked and integrated with the mProphet algorithm based on a target-decoy approach (Reiter et al., 2011). A detection q value was assigned for each peak, and peptides with q values greater than 0.01 in more than 50% of samples were filtered out. Peptide peak areas were exported to an MSstats (v3.18.5; Olga Vitek Lab, Northeastern University) input format (Choi et al., 2014). Protein peak areas were calculated with MSstats following summarization by Tukey's median polish, and equal median normalization was applied per comparison. The MS run data from both replicates (or individual replicates where indicated) were aggregated to calculate fold change values. Fold change p values were determined by linear mixed effect models via MSstats and adjusted for multiple comparisons (Benjamini-Hochberg). Proteins were considered to be differentially expressed (with respect to cell age or cell line depending on the comparison) if their minimum adjusted p value among pairwise comparisons was less than 0.05 and also had a fold change greater than 1.5-fold (up) or less than 0.67-fold (down). Only proteins that were detected in all biological replicates (i.e., detected and quantified in at least one MS run per replicate) used for comparisons were considered.

2.5 | 2DE proteomics

Briefly, CHO HCPs (300 μg) were precipitated from end-of-run HCCF samples with methanol, resolubilized, and separated by isoelectric

focusing on pH 3–10 nonlinear gradient strips as previously described (Valente et al., 2012). Second dimension separation with SDS-PAGE was performed on 13% T, 2.6% C polyacrylamide gels, followed by SYPRO Ruby (Thermo Fisher Scientific) staining and fluorescence scanning on a Typhoon FLA 7000 (Cytiva). Scanned images were imported into ImageMaster 2D Platinum software (v5.0; Cytiva) for image analysis. Automatic spot detection was performed with manual edits to remove artifacts, and relative spot volume was calculated by normalizing the volume of each protein spot to the total spot volume detected in each image as described previously (Valente et al., 2012). Spots that could be manually excised and exhibited at least a 1.5-fold change in relative spot volume between PDL 90 and PDL 30 samples as well as some unchanging reference spots were collected for identification.

Excised spots were digested, prepared with C18 ZipTips (Sigma-Aldrich), dried with SpeedVac™ and resuspended in 24 µl 2% ACN, 0.1% FA. Each LC-MS/MS run used a sample injection of 12 µl. LC/MS-MS analysis was performed on a TripleTOF 6600 coupled to an Eksigent nano LC 425 as described above for DDA runs, but with a program of 3% B to 35% B in 5 min, 35% B to 80% B in 1 min, and 80% B for 2 min to elute peptides. As before, LC-MS/MS data were searched against the Chinese hamster RefSeq. 2019 database (Rupp et al., 2018), and proteins that matched with common contaminants were filtered out as hits. Search parameters were specified as following: modifications by iodoacetamide, trypsin digestion, and gel identification as a special factor. Proteins with a false discovery rate less than 1% and with at least two peptides were considered confident identifications.

2DE analysis was performed on two biological replicates for each cell age. The relative spot volumes were log2-transformed and searched for statistically significant changes across ages by analysis of variance (ANOVA) in R. Only protein spots that were detected on both replicates of each cell age were considered. Spots were only considered differentially expressed if their p value by ANOVA was less than 0.1 and also had a fold change greater than 1.5-fold (up) or less than 0.67-fold (down). For spots with $p < 0.1$ by ANOVA, statistical significance of pairwise comparisons was calculated by Tukey's posthoc test and considered significant if $p < 0.05$.

2.6 | Statistical analysis

All plots and figures were generated and data analyses were conducted in Microsoft Excel or R (v4.1.1). Linear regressions were tested for a statistically significant fit ($p < 0.01$) and coefficient of determination (R^2) by one-way ANOVA. Analysis of covariance (ANCOVA) homogeneity of slopes test was first applied to all slopes within a given comparison to search for significant differences ($p < 0.05$). If the group of slopes met the significance criteria, then a similar pairwise slope test was applied by fitting each pair of slopes to a linear regression model that included an interaction term for the covariate (e.g., PDL) and independent variable (e.g., IVCD or culture duration). Pairs of slopes were then considered to be significantly different if the significance of the fit with the interaction term showed $p < 0.01$.

Unsupervised clustering was performed in R on log2-transformed SWATH-MS peak area data (i.e., HCP abundance), which was normalized per protein to the average peak area at PDL 30 or the average across both cell lines, accordingly. The Euclidean distance of the scaled data was used for hierarchical clustering with Ward's minimum variance method, which was the clustering method found to produce an agglomerative coefficient closest to 1 for all data sets. For all clustering analyses, HCPs grouped optimally into three clusters according to the gap statistic.

Copy number data for the VRC01 cell line and its parental host at approximately PDL 30 were reported previously (Hilliard & Lee, 2021) and tabulated for HCPs that were differentially expressed between the two cell lines. As described previously, the same cut-offs for VRC01-to-host copy ratio of less than 0.65 or greater than 1.35 were applied to categorize copy number loss or gain, respectively. One-tailed hypergeometric tests were performed in R to determine enrichment of HCPs with copy number loss or gain in sets of differentially expressed HCPs in the SWATH-MS PDL 30 cell line comparison that were downregulated or upregulated, respectively.

3 | RESULTS AND DISCUSSION

3.1 | Growth rate of producer cells accelerates during aging

To determine the effect of cell age on growth rate, the VRC01 mAb-producing cell line was aged in parallel with its parental host over a period of time that would span the typical duration of a commercial seed train process from cell bank thaw to production. Cells were aged without MTX selection to PDL 90, which corresponded to approximately 60 days of continuous passaging, and doubling time was calculated at regular passage intervals (Figure 1a). Whereas the doubling time of the null host cells remained nearly constant, the doubling time of the VRC01 cells was negatively correlated with cell age. Both cell lines had a comparable doubling time of about 16.7 h after 70 doublings. Increased growth over time has been documented previously for CHO cells (Beckmann et al., 2012; Chusainow et al., 2009) and serves as a hallmark of instability due to mechanisms that include outgrowth of subpopulations that exhibit chromosomal aberrations (Baik & Lee, 2018). In this case, we postulate that the initial growth rate differences followed by gradually faster growth in the VRC01 cell line was due to MTX exposure during the cell line development process (but not during cell aging) as has been observed for other MTX-selected cell lines (T. K. Kim et al., 2001).

Cell size was also monitored over time as a potential marker of instability (Figure 1b). Cell size tends to correlate positively with qP as more volume and biomass supports greater capacity for recombinant protein expression and secretion (T. K. Kim et al., 2001; Lloyd et al., 2000). Although we did not observe a decrease in cell size over time, the average diameter of the VRC01 cell line was significantly larger

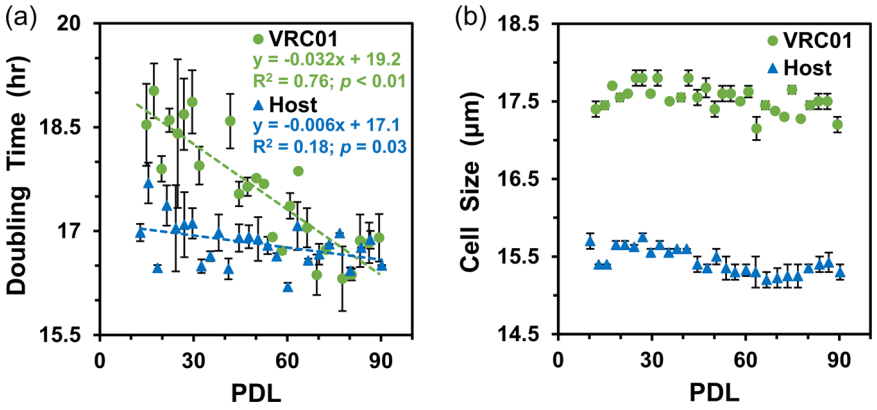

FIGURE 1 Comparison of cellular characteristics for the VRC01 producer and its parental CHO-K1 host during extended culture in shake flasks. (a) Doubling time during aging with cell age measured in terms of population doubling level (PDL) starting once cells reached PDL 10 and had fully recovered from thaw. Linear regressions of doubling time versus PDL for each cell line are provided with *p* value determined by one-way ANOVA. (b) Average cell diameter during aging. Error bars represent the standard error of the mean for two independently aged biological replicates. ANOVA, analysis of variance; CHO, Chinese hamster ovary; PDL, population doubling level.

than that of the null host ($p < 0.01$ by one-tailed Student's *t* test) by 2.1 µm.

3.2 | Mab productivity decline is accompanied by increased HCP productivity with age

The aged cells were used in bioreactor runs to investigate age-dependent productivity of both mAb and HCPs. An ambr®250 bioreactor system was leveraged for both the N-1 stages and production runs to ensure consistent implementation of an internally developed process. This bioreactor system has been applied previously as a scale-down model for high-throughput CHO process development (Bareither et al., 2013; Clark et al., 2017). Cell ages at the time of bioreactor inoculation were targeted to be PDL 30, 60, and 90. Culture viability (Figure 2a,d) and VCD (Figure 2b,e) were monitored during the fed-batch production runs. Overall, highly similar viability and growth profiles were observed across cell ages despite differences in the growth rate of producer cells during the N-1 stage (Supporting Information: Figure S2). HCCF samples for proteomic analysis of HCP composition were collected according to a pre-determined 80% viability cut-off, which occurred on the same culture day for each cell line at all three ages (Days 11 and 9 for VRC01 and host cell line, respectively). The viability cut-off was used to ensure an even comparison among samples as culture viability at harvest is known to be a critical factor that impacts the composition of HCPs present in the feed-stream for downstream purification (Hogwood et al., 2016; Tait et al., 2012). The earlier expiration of the null host was likely caused by runaway lactate production that accumulated to 9–12 g/L (Supporting Information: Figure S3) whereas VRC01 cells demonstrated a shift from lactate production to lactate consumption after 3 days in culture. For both cell lines, VCD peaked on Day 7 at 39×10^6 and 54×10^6 cells ml^{-1} for VRC01 and host cells, respectively, followed by a sharp decline in cell density. A reduction in the effective number of cells in suspension due to clumping and bioreactor fouling was responsible for the VCD drop rather than cell death. All conditions reached an IVCD of more than 260×10^6 viable cells · day ml^{-1} by completion of the run.

Despite showing relatively little phenotypic difference with respect to growth and metabolism, the VRC01 cells demonstrated a decline in mAb titer with age (Figure 2c). There was 70% titer retention at PDL 60% and 57% titer retention at PDL 90 relative to the initial titer (3.1 g L^{-1} at PDL 30). Correspondingly, the calculated qP values (Figure 2f, Supporting Information: Figure S4) significantly declined with each age ($p < 0.01$ for all pairwise slope test comparisons). Productivity instability is commonly observed in CHO producer cell lines, which generally show high clone-to-clone variability in stability levels due to mechanisms such as epigenetic silencing and copy number loss (Bailey et al., 2012; Chusainow et al., 2009; Paredes et al., 2013). Extended culture in the absence of drug selection represents a worst-case scenario in which the major selection pressure becomes cell growth, enabling nonproducing subpopulations with a metabolic advantage to overgrow and outcompete producers. Depending on the sequence structure at the site of transgene integration, spontaneous chromosomal re-arrangements can give rise to cells with reduced transgene copies that persist as long as MTX selection is withheld (Baik & Lee, 2017). In agreement with this mechanistic basis for instability, the productivity decline exhibited by the VRC01 cell line correlated with a loss of both the light chain and heavy chain gene copies over time (Supporting Information: Figure S4).

HCP concentration during the bioreactor runs was quantified by both Bradford assay and CHO-specific HCP ELISA (Figure 3a,b). Although both methods indicated similar time course trends across cell ages, Bradford assay measurements were approximately 3–4-fold higher than those measured with the ELISA. Because this multiplier was consistent for both cell lines, the discrepancy between methods was unrelated to the presence of mAb but instead caused by

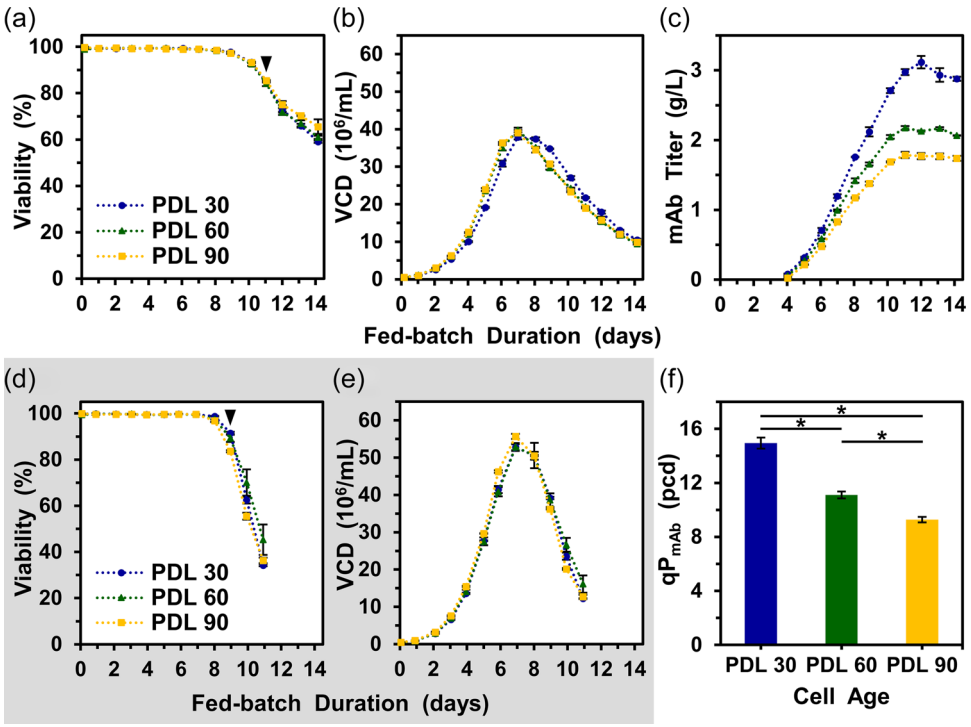

FIGURE 2 Fed-batch culture data from bioreactor runs for cells at three population doubling levels (PDLs). Viability profiles for (a) VRC01 producer cells and (d) host cells. Black arrows indicate days that samples were collected for proteomic analysis, where cells reached the viability cut-off before declining below 80%. Viable cell density (VCD) profiles for (b) VRC01 cells and (e) host cells. (c) Titer data for VRC01 monoclonal antibody (mAb) over days in fed-batch culture. Error bars in scatter plots represent the standard error of the mean for two independently aged biological replicates. (f) Cell-specific productivity (qP) measured in pg $cell^{-1}$ day^{-1} (pcd) for mAb production from aged VRC01 cells. Error bars in bar charts represent the 95% confidence interval for the qP slope determined by linear regression with statistical significance calculated for pairwise slope comparisons denoted as * $p < 0.01$.

substances registered by the Bradford assay that were not immunoreactive HCPs. The ELISA is widely accepted to be the gold standard for routine tracking of HCP levels during bioprocess steps despite certain limitations (Zhu-Shimoni et al., 2014). From ELISA, the HCP-to-mAb ratio on Day 11 for VRC01 runs at PDL 30 was 0.27 g HCP per g mAb, which was consistent with prior literature for a similar process (Goey et al., 2019), before rising to 0.49 and 0.76 at PDL 60 and PDL 90, respectively. The HCP productivity (Figure 3c,d, Supporting Information: Figure S5) of the VRC01 producer cell line increased from PDL 30 to PDL 60 and then remained constant ($p < 0.01$ for only pairwise slope test comparisons between PDL 30 and PDL 60/90, Figure 3c,d). The same pattern was true for the HCP productivity of host cells but only for Bradford assay measurements ($p < 0.01$ and $p = 0.28$ by ANCOVA for Bradford and ELISA measurements, respectively). Notably, the HCP qP values reported correspond to the exponential phase of the fed-batches; more HCPs were generated during the latter part of the runs for both cell lines even before dropping in viability. The nonlinear relationship between HCP concentration and IVCD upon the onset of viability decline indicated a loss in HCP productivity during later days in culture. This observation contradicts conventional assumptions that a considerable amount of HCP would be released upon loss of culture viability and thereby impact downstream processing.

3.3 | Age-dependent differential expression of HCPs measured by SWATH-MS

Label-free SWATH-MS proteomics was used to determine the effect of cell age on HCP composition. This DIA method for LC-MS/MS enables high-throughput and reproducible quantitation of diverse protein species within complex samples like HCCF (Sim et al., 2020). SWATH-MS has been previously leveraged to characterize CHO cell lines and bioprocesses by tracking the levels of individual HCPs (Orellana et al., 2015; Walker et al., 2017). In the present study, HCCF samples from both cell lines were first used to build a spectral library of approximately 33,000 peptides representing over 2000 HCP groups for extraction of quantitative measurements from the SWATH-MS data. 1518 and 1660 unique HCPs could be identified and quantified across all end-of-run samples from the VRC01 and host cell line bioreactors, respectively, with substantial overlap (>94%) in proteins identified between biological replicates (Supporting Information: Figure S6). The similarity of age-dependent patterns of protein expression across independently aged biological replicates was assessed by hierarchical clustering (Figure 4a,b). Biological duplicates of the VRC01 producer cell line clustered together at each age in pairs with the PDL 60 and PDL 90 samples also being somewhat similar to one another. In contrast, clustering of host cell

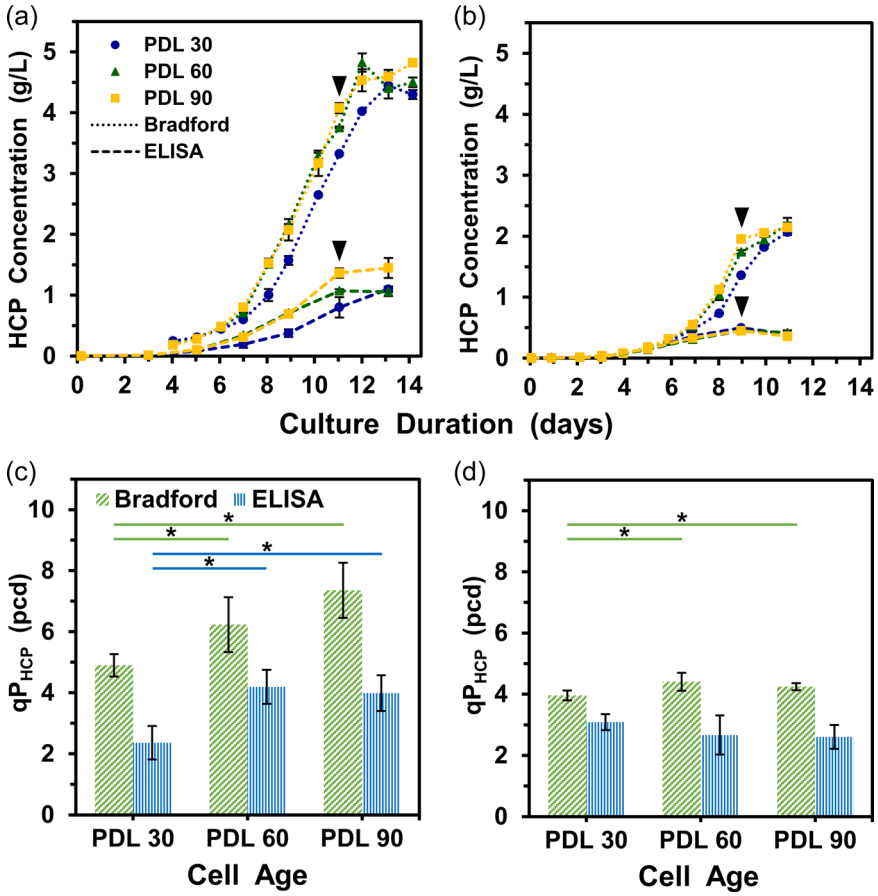

FIGURE 3 Total host cell protein (HCP) levels during fed-batch cultures with aged cells at three population doubling levels (PDLs). Total HCP concentrations measured by Bradford assay and CHO-specific HCP ELISA over culture days for (a) VRC01 producer cells and (b) host cells. Black arrows indicate days that samples were collected for proteomics analysis. Error bars in scatter plots represent the standard error of the mean for two independently aged biological replicates. Cell-specific productivity (qP) levels measured in pg cell^{-1} day^{-1} (pcd) for HCP production during exponential phase for (c) VRC01 cells and (d) host cells. Error bars in bar charts represent the 95% confidence interval for the qP slope determined by linear regression with statistical significance calculated for pairwise slope comparisons denoted as * $p < 0.01$. CHO, Chinese hamster ovary.

line duplicates indicated early divergence, with samples from later cell ages showing similar degrees of variability such that they did not cluster neatly in pairs. For both cell lines, the HCPs grouped into three clusters according to similar expression patterns. These three patterns could be generalized as expression that correlated negatively with cell age, correlated positively, or showed little to no change (ordered from top to bottom in Figure 4). A key question we sought to address in this study was whether cell populations derived from clonal origins show consistent longitudinal changes when aged independently. In the absence of perturbations, the underlying mechanisms influencing HCP expression changes would ostensibly involve stochastic processes, such as genetic mutation (e.g., copy number loss/gain and single nucleotide changes), epigenetic instability, and phenotypic drift. However, the clustering analysis provided preliminary support that some HCPs do indeed have reproducible age-dependent expression patterns, even if the majority do not.

To further characterize HCP expression, data from duplicates were aggregated. Significantly changing HCPs (adjusted p value < 0.05 with fold change >1.5 or <0.67) were categorized according to the direction of their largest change in expression (Figure 5a). For VRC01 cells, only 62 (4%) HCPs were differentially expressed with 40 (65%) of those categorized as downregulated and 22 (35%) upregulated. The parental host cells had more HCPs with variable expression: 213 (13%) total with 175 (82%) downregulated and 38 (18%) upregulated. When analyzed individually, the biological replicates show that ~30% of differentially expressed proteins are identified in both VRC01 replicates whereas ~55% of differentially expressed proteins are identified in both host cell replicates (Supporting Information: Figure S6). Lactate dehydrogenase (LDHA) levels across ages, which are a well-established indicator of cell health and membrane integrity (Legrand et al., 1992), were constant with age for both cell lines (Supporting Information: Figure S7), indicating that the differential abundance of HCPs was not simply due to varying degrees of intracellular release across ages but rather changes in expression or secretion. Between PDL 30 and PDL 90, VRC01 cell line HCPs showed fold-change increases up to 3.1-fold (annexin A6, ANXA6) and decreases up to 8-fold (annexin A4, ANXA4), while the range for host was narrower (2.3-fold increase and 3.1-fold decrease) (Figure 5b, Supporting Information: Figure S8).

The fraction of changing HCPs determined for the parental host in this study was comparable to the fraction we found previously for null CHO-K1 cells (13% of 630 proteins using iTRAQ™ shotgun proteomics) (Valente et al., 2015). However, previously observed fold-change magnitudes were much greater than those observed here, probably due to the large difference in the aging time scales between the two studies in addition to the subcloning of the parental host after suspension adaptation. Another contributing factor may be the change in upstream process (i.e., fed-batch culture in bioreactor vs. batch culture in flask), which could have helped to stabilize HCP

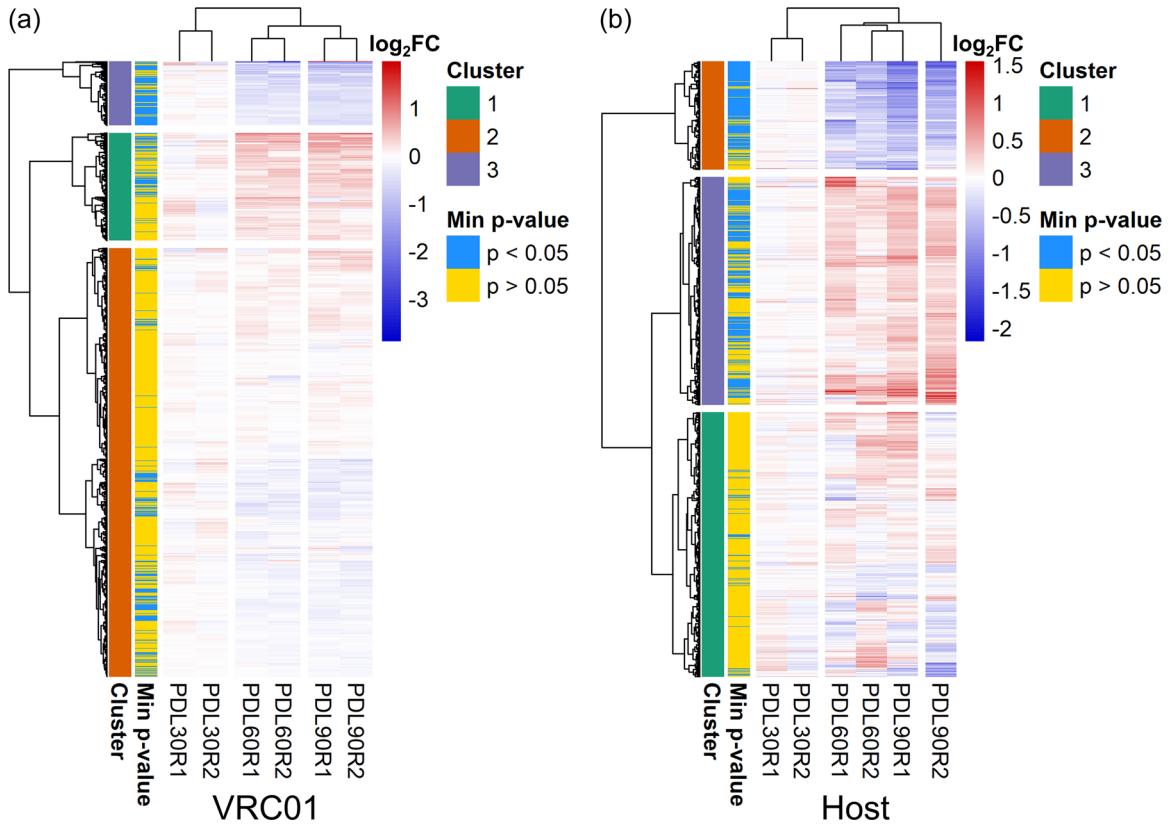

FIGURE 4 Unsupervised clustering analysis and heat map of host cell protein (HCP) abundance data obtained from SWATH-MS proteomics. Hierarchical clustering of log2-fold change (FC) values for each HCP relative to its average abundance of replicates at PDL 30 for (a) VRC01 and (b) host bioreactor samples. Cluster number labels are arbitrarily assigned to each cell line clustering set. Replicates (R1 and R2) at each PDL correspond to samples from independently aged biological duplicates processed in technical triplicate by SWATH-MS. Each HCP (row) is annotated with the minimum adjusted *p* value obtained from pairwise comparisons among cell ages using data averaged for replicates.

levels in this case, and has been previously determined to influence HCP levels (Goey et al., 2019; Gronemeyer et al., 2016; Park et al., 2017). We were also able to reliably measure more HCPs due to improved MS instrumentation and the use of SWATH-MS, yielding a population that more closely agrees with the putative size of the CHO supernatant proteome (Kumar et al., 2015). It is uncertain how the differential expression of individual HCPs measured by relative quantification with SWATH-MS would translate to the changes noted in the bulk HCP concentration. Still, the majority of the top 10% most abundant HCPs were not changing; within this subset, 7 (5%) and 58 (35%) HCPs were differentially expressed for VRC01 and host cells, respectively. Coupled with relatively low fold-change magnitudes, this is consistent with the total HCP concentration at harvest remaining roughly constant across ages despite changes in HCP qP with age. For both cell lines, the two most abundant HCPs that showed constant expression levels were histone H2B type 1 and histone H4, suggesting that sustained high-level expression of these histones may be crucial for cellular growth.

Eleven (4%) proteins were changing in both cell lines between the two sets of HCPs exhibiting age-dependent differential expression (Figure 5c). All but one of these HCPs (C-C motif chemokine 7) were changing expression in opposite directions for the two cell lines (Figure 5d), suggesting that changes observed consistently when aging one cell line are not generalizable even between cell lines with a common lineage. The proteins changing in both cell lines may represent a subset of HCPs with highly variable and unpredictable expression, and therefore should be of particular concern during downstream processing. One member of this list, decorin (DCN), was also found to be differentially expressed in previous work (Valente et al., 2015) and has been linked to cell stress and apoptosis, as well as contact-mediated inhibition of cell proliferation (Albrecht et al., 2018; Yamaguchi & Ruoslahti, 1988).

3.4 | HCPs are more differentially expressed between cell lines than across ages

As a reference for age-dependent expression changes, we wanted to directly compare HCP expression levels between the two cell lines. 82% of all proteins measured were common to both cell lines (Figure 6a). This degree of overlap matches what has been observed previously among CHO cell lines (Yuk et al., 2015). As expected, the VRC01 mAb light chain and heavy chain were two proteins that were unique to the producer cell line. Their relative expression closely

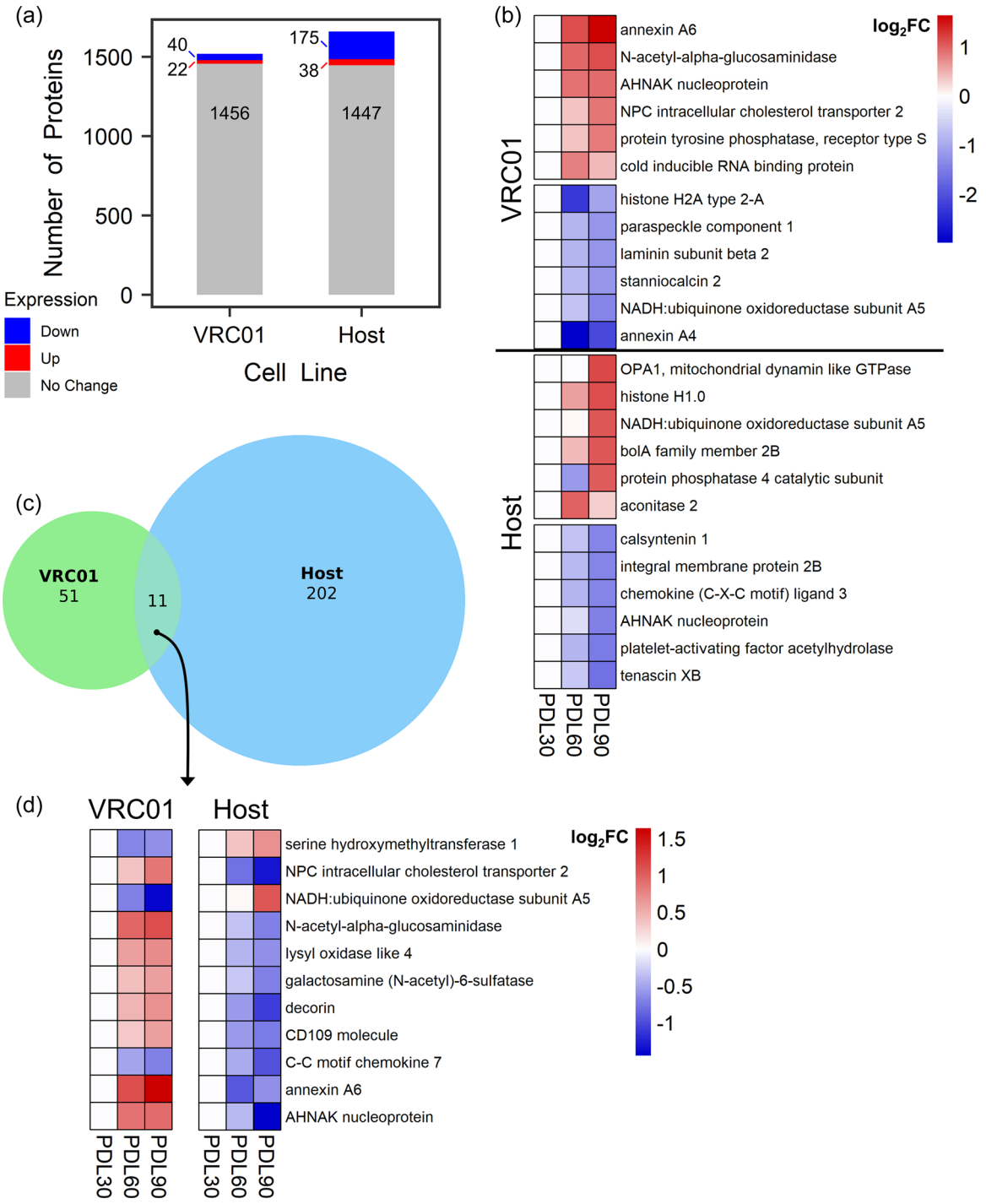

FIGURE 5 Overview of CHO host cell proteins (HCPs) that demonstrated age-dependent expression changes measured by SWATH-MS. (a) HCPs categorized according to the direction of their largest fold change (FC) among cell age comparisons for each cell line. FC comparisons were relative to initial time points (either PDL 30 or PDL 60). Proteins were only considered differentially expressed if their minimum adjusted *p* value among pairwise comparisons was less than 0.05 and also had a fold change greater than 1.5-fold (up) or less than 0.67-fold (down). (b) Heat map of log2-FC values for the top six most up- and downregulated HCPs over time in each cell line relative to PDL 30. (c) Overlap between cell lines for HCPs that exhibited age-dependent expression. (d) Heat map of log2-FC values for the 11 HCPs that were differentially expressed as a function of cell age in both cell lines. CHO, Chinese hamster ovary.

agreed with mAb titer measured via biolayer interferometry (Figure 6b), supporting the accuracy of our MS quantification results. Hierarchical clustering of HCPs in both cell lines at PDL 30 and PDL 90 separated proteins into three groups according to expression patterns that corresponded to downregulation in the producer cell line, upregulation, or little to no change in expression between cell lines (Supporting Information: Figure S9). Notably, the number and proportion of HCPs with differential expression between cell lines

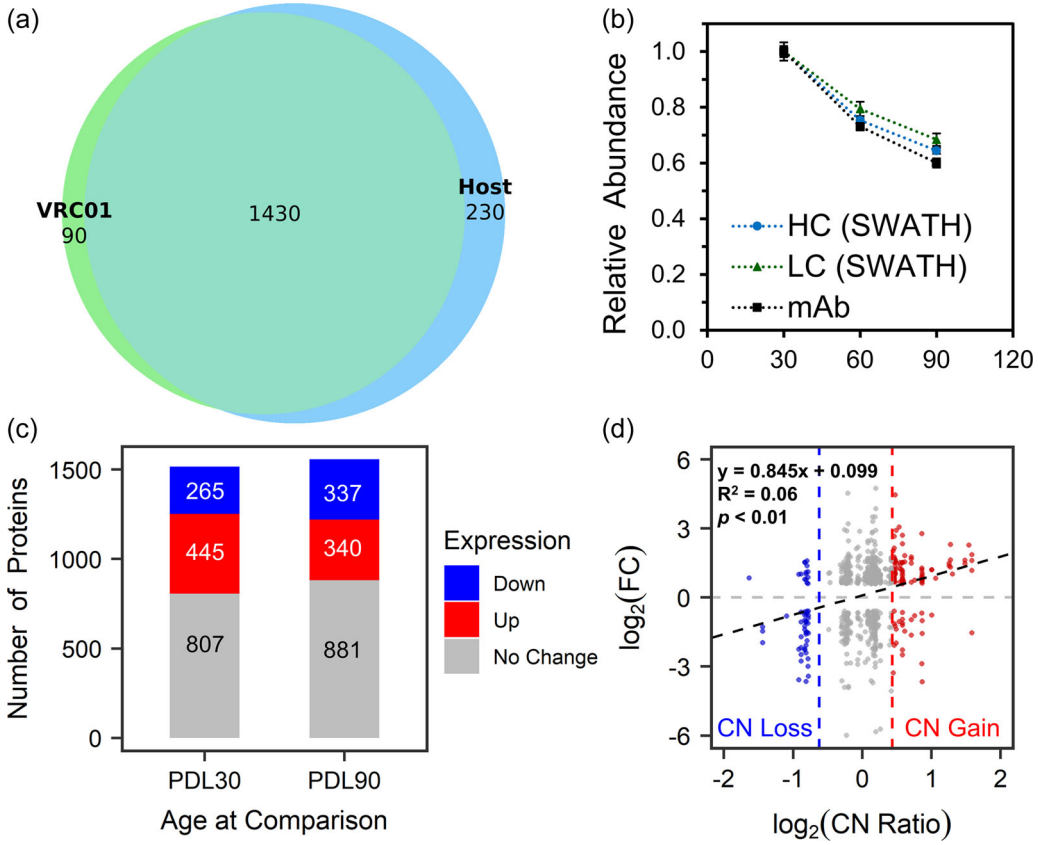

FIGURE 6 Comparison of host cell proteins (HCPs) that were differentially expressed between CHO cell lines measured by SWATH-MS. (a) Overlap between cell lines for all proteins that were identified and quantified in bioreactor harvest samples across all three cell ages. (b) SWATH-MS abundance of the VRC01 mAb light chain and heavy chain across cell ages relative to PDL 30 compared to relative mAb abundance as determined by biolayer interferometry with Octet. Error bars represent the standard error of the mean for two independently aged biological replicates. VRC01 mAb peptides were not detected in host cell samples. (c) HCPs categorized according to the direction of their fold change (FC) in VRC01 relative to host samples at PDL 30 and also at PDL 90. Proteins were only considered differentially expressed between cell lines if their adjusted *p* value for the comparison was less than 0.05 and also had a fold change greater than 1.5-fold (up) or less than 0.67-fold (down). (d) Cell line comparison FC values at PDL 30 versus copy number (CN) ratio for differentially expressed HCPs. CN ratio was calculated as HCP gene CN in VRC01 cells over CN in host cells. Linear regression of log2-FC versus log2-CN is provided with *p* value determined by one-way ANOVA. ANOVA, analysis of variance; CHO, Chinese hamster ovary; PDL, population doubling level.

(677–710 proteins, 43–47% of all HCPs) was much greater than had been observed within any one cell line across ages (Figure 6c). The HCPs that demonstrated the greatest difference in abundance between cell lines spanned a range of up to 27-fold increase (transcobalamin 2, TCN2) and up to 63-fold decrease (DCN) for the PDL 30 comparison (Supporting Information: Figure S10); however, more than 90% of differentially expressed proteins exhibited less than a 5-fold change. We found a 49% overlap between the sets of HCPs changing between the two cell lines at PDL 30 and PDL 90, but the differentially expressed proportion relative to the total number of proteins measured remained constant over time. Surprisingly, this indicates that changes in expression were balanced such that the two cell lines did not become more divergent with time. We leveraged available genome sequencing data for these cell lines to investigate the mechanism of differential expression (Hilliard & Lee, 2021). Although the sequenced cell samples had been cultivated to approximately PDL 30 in shake flasks, they did not undergo bioreactor production, but we assumed the dominant genetic distribution of the cell populations was unlikely to change over such short time scales. By aligning copy number data to each HCP gene, we found that genes with increased and decreased copy number were significantly enriched within sets of differentially expressed genes that were upregulated and downregulated, respectively ($p < 0.01$ for both associations by one-tailed hypergeometric test). Despite these associations, only a small portion (6%) of the variance in expression could be explained by corresponding changes in gene copy number (Figure 6d). Differential expression between CHO cell lines could be caused by a combination of posttranscriptional regulation mechanisms, such as protein translation, secretion, and degradation. These findings support the mounting body of evidence that CHO HCP populations are qualitatively similar regardless of cell lineage, but that a subset of individual HCPs may display profoundly different levels of expression (Hogwood et al., 2016; Madsen et al., 2015; Park et al., 2017; Zhang et al., 2014).

3.5 | Age-dependent differential expression of HCPs measured by 2DE proteomics

2DE gel analysis was performed as a complement to SWATH-MS to characterize additional changes in HCP expression. Differences are expected between the two methods because LC-MS/MS provides peptide-level resolution, while 2DE provides protein-level resolution based on changes in spot volume across gels. Visual inspection of the gels provided a qualitative indication that spot patterns did not differ noticeably across ages (Figure 7a,b). Image analysis data from independently aged biological duplicates were combined to calculate fold-change values associated with 408 and 371 anonymous protein spots for the VRC01 and host cells, respectively. Of these spots, 41 (10%) for VRC01 and 12 (3%) for host were significantly differentially expressed ($p < 0.1$ by ANOVA and with fold change greater than 1.5 or less than 0.67). More of the changing spots were increasing (78%) than decreasing (22%) for the producer cell line (Figure 8a), which differed from SWATH-MS results. Still, the majority of HCPs were not changing in either cell line across ages when measured using either method, while those that did change exhibited a similar range of fold changes, measuring between 2.8-fold down and 4.7-fold up by 2DE.

A subset of 52 spots (28 and 24 from VRC01 and host gels, respectively), some of which showed variable expression with age, were excised and identified by MS (Supporting Information: Figures S11 and S12). Fifty of these spots could be confidently identified as one or multiple protein groups. The association of fold change values with individual HCPs was complicated by the detection of multiple HCPs (3.5 proteins per spot on average, up to 13) within 90% of excised spots. 61 HCPs were identified within 18 spots from producer cell line samples that met the criteria for differential expression, but none of these HCPs overlapped with proteins that had been identified as differentially expressed by SWATH-MS. Similarly, 28 HCPs were identified within 5 spots from host samples, and 7 of these were also identified as differentially expressed by SWATH-MS. Only 3 spots, all from the producer cell line, were both differentially expressed and matched to unique HCPs: alpha-enolase (ENO1), peroxiredoxin-1 (PRDX1), and vimentin (VIM). The increased abundance of these particular protein species over time (Figure 8b) may be related to oxidative stress, which could indicate a heightened response and ability to cope with cellular stressors (UniProt Consortium, 2021).

Co-migration of protein species that originated from separate genes but happen to have similar biophysical characteristics can occur with 2DE, especially as the limit of detection for MS instrumentation continues to improve (Campostrini et al., 2005; Thiede et al., 2013). Accurate assignment of differential expression becomes challenging in such cases and may contribute to an apparent disagreement in the observed changes as measured between proteomic methods. Solutions exist to address co-migration, including three-dimensional separation (Colignon et al., 2013) and isotopic

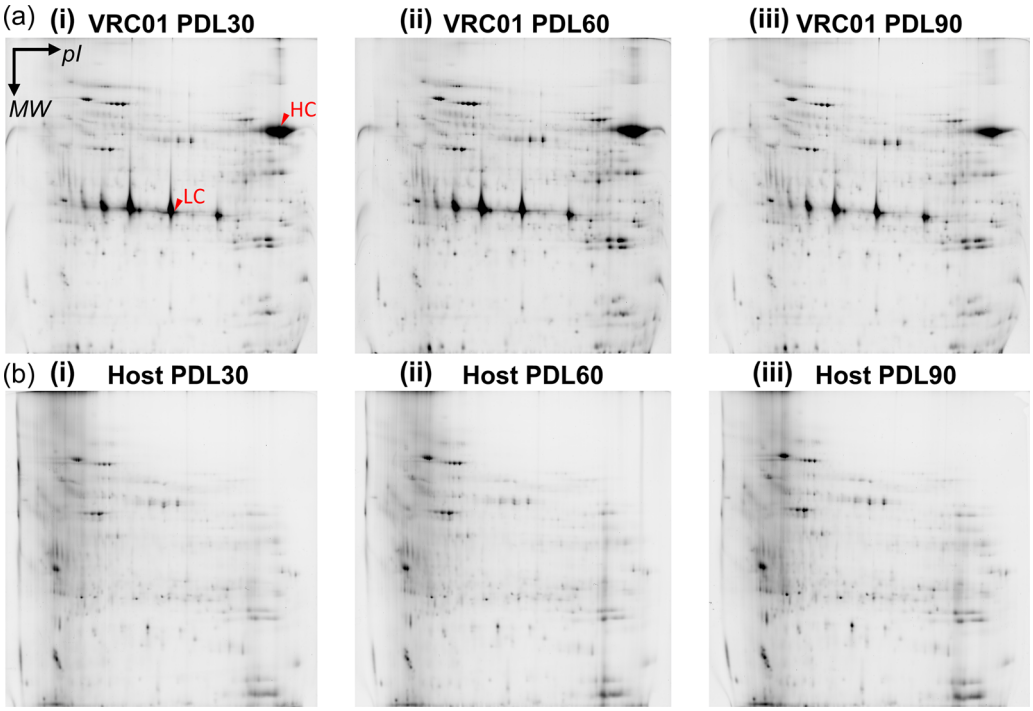

FIGURE 7 Comparison of representative two-dimensional electrophoresis (2DE) images for harvested cell culture fluid (HCCF) samples from (a) VRC01 and (b) host cell line bioreactor runs. Only one biological replicate gel image stained with SYPRO Ruby is shown for each cell age: (i) PDL 30, (ii) PDL 60, and (iii) PDL 90. The approximate locations of the VRC01 mAb heavy chain (HC) and light chain (LC) are annotated. See Supporting Information: Figures S11 and S12 for all gel images of sample replicates and locations of excised spots. PDL, population doubling level.

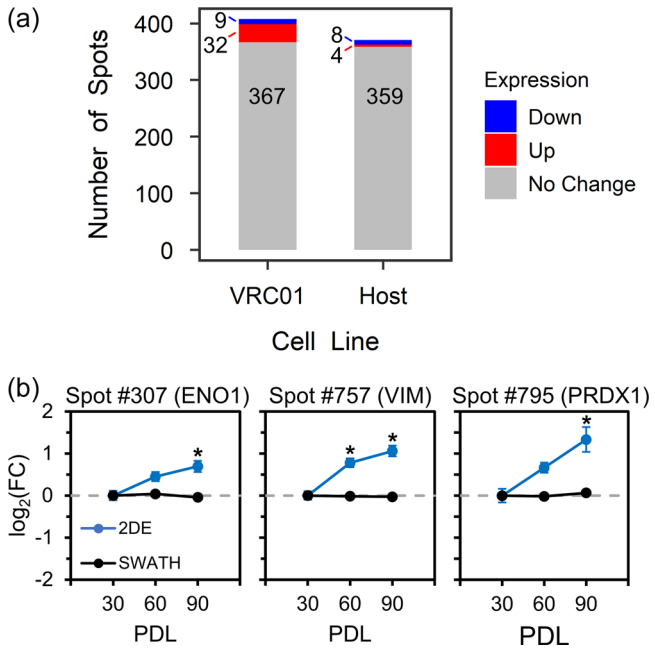

FIGURE 8 Overview of CHO host cell proteins (HCPs) that demonstrated age-dependent expression changes measured by 2DE. (A) Gel spots (HCP species) categorized according to the direction of their largest fold change (FC) among cell age comparisons for each cell line. FC comparisons were relative to initial time points (either PDL 30 or PDL 60). Spots were only considered differentially expressed if their p value by ANOVA was less than 0.1 and also had a fold change greater than 1.5-fold (up) or less than 0.67-fold (down). (B) Relative protein expression by 2DE and SWATH-MS proteomics for 3 spots from VRC01 samples that were differentially expressed (by 2DE only) and matched to unique HCPs: alpha-enolase (ENO1), vimentin (VIM), and peroxiredoxin-1 (PRDX1). Error bars represent standard error of the mean of relative spot volumes from two independently aged biological replicates. Statistical significance calculated by Tukey's posthoc test with respect to expression at PDL 30 and denoted as * $p < 0.05$. 2DE, two-dimensional electrophoresis; CHO, Chinese hamster ovary; PDL, population doubling level.

labeling (Thiede et al., 2013). 2DE is also limited relative to SWATH-MS with respect to its ability to resolve low-abundance HCPs or HCPs with less typical biophysical properties (e.g., very high/low molecular weight or isoelectric point). A feature of 2DE that makes it complementary to shotgun proteomics is that protein species (also called proteoforms) can be resolved such that modified protein forms originating from a single gene manifest as multiple spots across the gel, whereas these species would be collated within a single protein group in the SWATH-MS analysis. Some of the observed modifications that can potentially alter protein species migration include glycosylation (and glycation), phosphorylation, deamidation, and proteolysis (Marcus et al., 2020; Smith & Kelleher, 2013). Consequently, quantitative differences for a single protein species (spot) may not necessarily correspond to changes in the total level of its protein group as was likely the case for the three uniquely matched HCPs mentioned above. However, the potential for the concentrations of certain HCP species to vary significantly in HCCF has unknown implications for downstream processing. As a hypothetical example, a particular modified form of an HCP might exhibit product-association, while other forms of that same HCP might not display the same problematic interactions. To the best of our knowledge, no such instances of this scenario have been reported thus far, but we speculate that more granular characterization of residual HCP species detected after purification steps may prove beneficial.

3.6 | Behavior of especially problematic HCPs

Certain difficult-to-remove CHO HCPs are known to be especially problematic during downstream processing, while others pose considerable risk if present in the final therapeutic product. 62 and 213 HCPs exhibited age-dependent expression measured by SWATH-MS in the VRC01 and host cell lines, respectively. Three additional protein species with variable expression were uniquely identified in the 2DE analysis of the producer cell line. These sets of HCPs were classified according to whether they had been previously determined to be difficult to remove due to product-association with mAbs (Levy et al., 2014), co-purification in Protein A eluate (Zhang et al., 2016), co-elution during polishing chromatography operations (Levy et al., 2016), and/or have been detected as impurities in drug substance (Falkenberg et al., 2019; Kreimer et al., 2017). Differentially expressed HCPs were also classified according to a recent industry-wide perspective piece (Jones et al., 2021) that designated certain CHO HCPs as high-risk based on their potential adverse impacts: (1) immunogenicity, (2) reduced drug quality via drug aggregation, modification, or degradation; (3) effect on formulation via polysorbate degradation, or (4) direct biological function in humans. The producer cell line had substantially fewer problematic HCPs that varied with age (8 HCPs, Figure 9a) compared to its parental host (44 HCPs, Figure 9b) with 50 unique HCPs between the two sets. Notably, each cell line had a subset of changing HCPs that are known to be both difficult-to-remove and high-risk (Supporting Information: Table S2). ENO1 and PRDX1 were differentially expressed in the producer cell line, although both HCPs exhibited changes by 2DE only. Cathepsin B (CTSB), clusterin (CLU), HTRA1, and PLBL2 were differentially expressed in the host cell line. AHNAK nucleoprotein (AHNAK), previously identified as a drug substance impurity in five out of six different mAbs (Falkenberg et al., 2019), and ANXA6, a member of the potentially immunogenic annexin family of HCPs (Bailey-Kellogg et al., 2014; Farrell et al., 2015; Fukuda et al., 2019), were problematic HCPs that demonstrated age-dependent expression in both cell lines. There were 25 problematic HCPs that were differentially expressed in at least one of the two cell lines and that also exhibited age-dependent expression changes in our previous study (Valente et al., 2015). Overall, 50% of the 92 HCPs that had previously shown variable expression in null CHO-K1 cells over 1 year in culture were also found to be changing in this study for the parental host even after just 60 days of aging.

Remarkably, both high-risk and difficult-to-remove HCPs seemed to be globally downregulated with age in the parental host, whereas

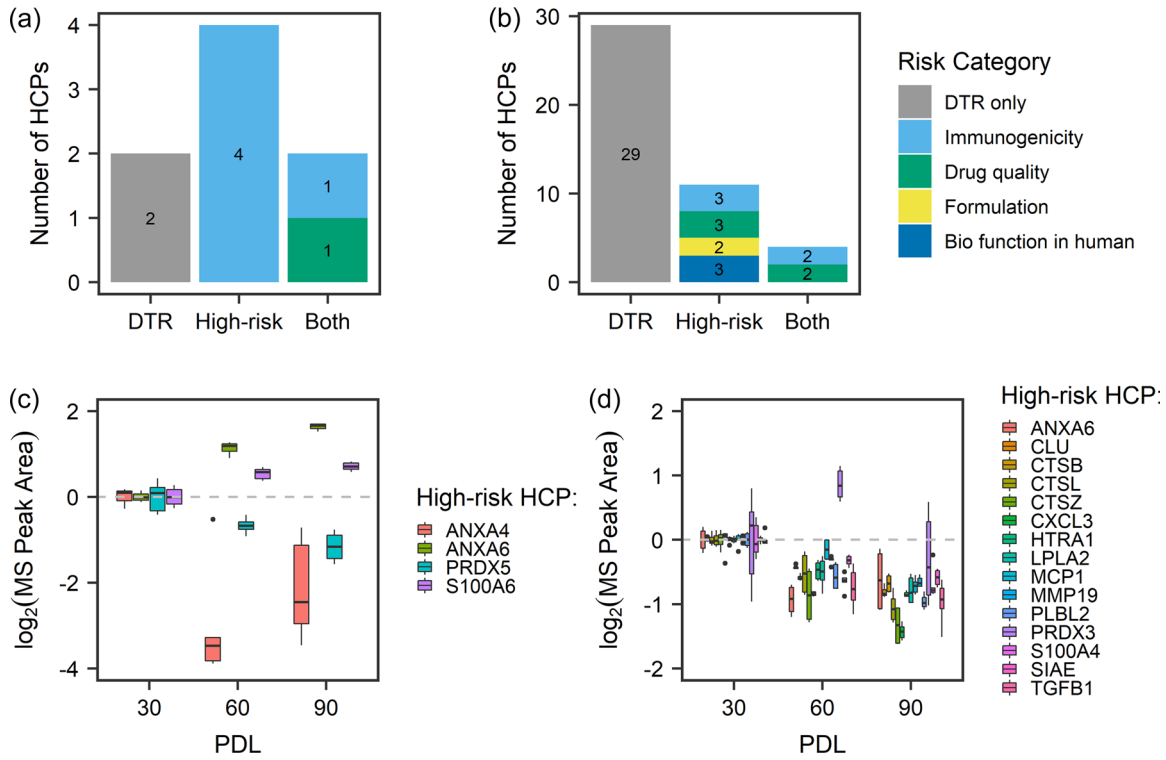

FIGURE 9 Age-dependent differential expression of problematic CHO host cell proteins (HCPs). Counts of HCPs that exhibited variable expression with age by SWATH-MS or 2DE in (a) VRC01 and (b) host HCCF samples that have been previously reported as difficult-to-remove (DTR) and/or high-risk. HCPs were considered DTR if they had previously been determined to associate with mAbs (Levy et al., 2014), co-purify in Protein A eluate (Zhang et al., 2016), co-elute during polishing chromatography steps (Levy et al., 2016), or have been detected as impurities in drug substance (Falkenberg et al., 2019; Kreimer et al., 2017). High-risk HCPs were subcategorized according to the risk classifications used by Jones et al. (2021). SWATH-MS abundance (MS peak area) plots for differentially expressed high-risk HCPs in (c) VRC01 and (d) host HCCF samples relative to average abundance of replicates at PDL 30. Boxplots represent data from SWATH-MS technical replicates for both independently aged biological duplicates at each cell age. 2DE, two-dimensional electrophoresis; CHO, Chinese hamster ovary; PDL, population doubling level.

VRC01 cells showed an even mix of both up- and downregulation (Figure 9c,d and Supporting Information: Figure S13). However, the relative abundance changes for these HCPs were less than 4-fold in all cases except for ANXA4 so therefore might have little practical significance for downstream processing. We acknowledge that the age-dependent expression patterns of problematic HCPs documented in this study may not readily translate to cultivation systems that use perfusion bioreactors due to the differences in culture conditions compared to conventional bioreactors. The proteomic analysis framework detailed here could be applied to study the effects of other culture conditions and perturbations on the HCP composition in downstream processing feed streams. The surprising potential for problematic HCPs to generally decrease over time in null CHO suspension cultures might have beneficial applications for cell line development strategies (e.g., initiating producer cell line generation from aged CHO hosts) and could complement existing efforts to genetically engineer "clean" CHO hosts via multi-HCP knock-out (Kol et al., 2020). This finding challenges the conventional practice of starting the cell line generation process from an early PDL of the host. Implementation of this strategy would require careful consideration for other quality attributes beyond the scope of this study, such as product glycosylation and aggregation, which may also be influenced by cell age due to changes in the expression levels of intracellular glycosyltransferases and chaperone proteins. Additionally, host cell population heterogeneity would increase over time, yielding more diverse panels of production clones with potentially unfavorable productivity characteristics.

4 | CONCLUSION

The push toward commercial implementation of extended bioprocesses warrants a better understanding of CHO platforms with respect to HCP expression levels across relevant time scales. We have performed an in-depth characterization of HCP profiles in samples derived from high-cell-density (>10^7 cells/ml) fed-batch bioreactor runs for a mAb-producing CHO cell line and its parental host. Across ages, cells showed a high degree of phenotypic similarity for growth and metabolism during production, yet exhibited significant instability with respect to mAb and HCP cell-specific productivity. Profiling of over 1500 HCPs by quantitative proteomics indicated that the composition of HCP impurities can change with cell

age, in agreement with previous work. However, relatively few changes (4%–13% of all HCPs were differentially expressed) were observed over a 60-day time scale, representative of the maximum length of a typical seed train or perfusion process. Moreover, differential expression between cell lines was proportionally higher and spanned greater fold change magnitudes despite both cell types having qualitatively similar HCP populations. Still, a small number of problematic HCPs that are considered high-risk impurities and/or challenging for downstream purification showed variable expression with age. Given their relatively low magnitude (less than 4-fold for all but one HCP), the age-dependent expression changes for these problematic HCPs are unlikely to pose a major challenge to downstream purification platforms, but should be considered nonetheless. We anticipate the general trends observed here can be generalized to other recombinant CHO cell lines subjected to extended cell culture without additional perturbations (e.g., change in culture medium, genotoxic stress, nutrient deprivation, etc.). The use of reference cell lines (the VRC01 producer and its parental host) in this study will facilitate future work to study CHO cell biology and elucidate mechanisms that influence age-dependent expression differences. Ultimately, characterization of CHO instability and HCP expression dynamics will prove beneficial in guiding rational cell line engineering efforts and downstream process development to alleviate the problems caused by HCP impurities in biomanufacturing.

ACKNOWLEDGMENTS

We wish to thank Bramie Lenhoff from University of Delaware for important contributions; Inn Yuk from Genentech for important discussions as well as leading the industry-wide team that supported this study; Scott Davis and Leila Choe for assistance with 2DE data generation; and Lauren Cordova for performing qPCR. We would also like to thank the NIH for sharing of cell lines. NH was supported in part by Award Number T32GM008550 from the National Institute of General Medical Sciences. This study was funded in part by NSF grants 1736123 and 1624698 as well as by the Department of Commerce, National Institute of Standards and Technology agreement 70NANB17H002.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data are available in article supplementary material. Other data that may be of interest to the community will be made available on request of the authors.

ORCID

Kelvin H. Lee http://orcid.org/0000-0003-0908-2981

REFERENCES

Aboulaich, N., Chung, W. K., Thompson, J. H., Larkin, C., Robbins, D., & Zhu, M. (2014). A novel approach to monitor clearance of host cell proteins associated with monoclonal antibodies. *Biotechnology Progress*, *30*, 1114–1124. https://doi.org/10.1002/btpr.1948

Albrecht, S., Kaisermayer, C., Gallagher, C., Farrell, A., Lindeberg, A., & Bones, J. (2018). Proteomics in biomanufacturing control: Protein dynamics of CHO-K1 cells and conditioned media during apoptosis and necrosis. *Biotechnology and Bioengineering*, *115*, 1509–1520. https://doi.org/10.1002/bit.26563

Baik, J. Y., Guo, J., & Lee, K. H. (2019). Host cell proteins during biomanufacturing. *Cell Cult. Eng*, 295–311.

Baik, J. Y., & Lee, K. H. (2017). A framework to quantify karyotype variation associated with CHO cell line instability at a single-cell level. *Biotechnology and Bioengineering*, *114*, 1045–1053.

Baik, J. Y., & Lee, K. H. (2018). Growth rate changes in CHO host cells are associated with karyotypic heterogeneity. *Biotechnology Journal*, *13*, 1700230.

Bailey, L. A., Hatton, D., Field, R., & Dickson, A. J. (2012). Determination of Chinese hamster ovary cell line stability and recombinant antibody expression during long-term culture. *Biotechnology and Bioengineering*, *109*, 2093–2103.

Bailey-Kellogg, C., Gutiérrez, A. H., Moise, L., Terry, F., Martin, W. D., & De Groot, A. S. (2014). CHOPPI: A web tool for the analysis of immunogenicity risk from host cell proteins in CHO-based protein production. *Biotechnology and Bioengineering*, *111*, 2170–2182.

Bareither, R., Bargh, N., Oakeshott, R., Watts, K., & Pollard, D. (2013). Automated disposable small scale reactor for high throughput bioprocess development: A proof of concept study. *Biotechnology and Bioengineering*, *110*, 3126–3138.

Baumann, M., Klanert, G., Vcelar, S., Weinguny, M., Marx, N., & Borth, N. (2019). Genome variation, the epigenome and Acellular phenotypes. In G. M. Lee, & H. F. Kildegaard (Eds.), *Cell culture engineering* (pp. 97–126). Wiley. https://doi.org/10.1002/9783527811410.ch5

Beckmann, T. F., Krämer, O., Klausing, S., Heinrich, C., Thüte, T., Büntemeyer, H., Hoffrogge, R., & Noll, T. (2012). Effects of high passage cultivation on CHO cells: A global analysis. *Applied Microbiology and Biotechnology*, *94*, 659–671. http://www.ncbi.nlm.nih.gov/pubmed/22331235

Bee, J. S., Tie, L., Johnson, D., Dimitrova, M. N., Jusino, K. C., & Afdahl, C. D. (2015). Trace levels of the CHO host cell protease cathepsin D caused particle formation in a monoclonal antibody product. *Biotechnology Progress*, *31*, 1360–1369. https://doi.org/10.1002/btpr.2150

Campostrini, N., Areces, L. B., Rappsilber, J., Pietrogrande, M. C., Dondi, F., Pastorino, F., Ponzoni, M., & Righetti, P. G. (2005). Spot overlapping in two-dimensional maps: A serious problem ignored for much too long. *Proteomics*, *5*, 2385–2395.

Chiu, J., Valente, K. N., Levy, N. E., Min, L., Lenhoff, A. M., & Lee, K. H. (2017). Knockout of a difficult-to-remove CHO host cell protein, lipoprotein lipase, for improved polysorbate stability in monoclonal antibody formulations. *Biotechnology and Bioengineering*, *114*, 1006–1015.

Chiverton, L. M., Evans, C., Pandhal, J., Landels, A. R., Rees, B. J., Levison, P. R., Wright, P. C., & Smales, C. M. (2016). Quantitative definition and monitoring of the host cell protein proteome using iTRAQ—A study of an industrial mAb producing CHO-S cell line. *Biotechnology Journal*, *11*, 1014–1024. https://doi.org/10.1002/biot.201500550

Choi, M., Chang, C. Y., Clough, T., Broudy, D., Killeen, T., MacLean, B., & Vitek, O. (2014). MSstats: An R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics*, *30*, 2524–2526.

Chon, J. H., & Zarbis-Papastoitsis, G. (2011). Advances in the production and downstream processing of antibodies. *New Biotechnology*, *28*, 458–463. https://doi.org/10.1016/j.nbt.2011.03.015

Chusainow, J., Yang, Y. S., Yeo, J. H. M., Ton, P. C., Asvadi, P., Wong, N. S. C., & Yap, M. G. S. (2009). A study of monoclonal antibody-producing CHO cell lines: What makes a stable high producer? *Biotechnology and Bioengineering*, *102*, 1182–1196.

Clark, C., Ryder, T., Sparks, C., Wang, M., Russell, R., Scott, C., Xu, P., Clark, C., Ryder, T., Sparks, C., Zhou, J., Wang, M., Russell, R., & Scott, C. (2017). Characterization of TAP Ambr 250 disposable bioreactors, as a reliable scale-down model for biologics process development. *Biotechnology Progress*, *33*, 478–489.

Colignon, B., Raes, M., Dieu, M., Delaive, E., & Mauro, S. (2013). Evaluation of three-dimensional gel electrophoresis to improve quantitative profiling of complex proteomes. *Proteomics*, *13*, 2077–2082.

Croughan, M. S., Konstantinov, K. B., & Cooney, C. (2015). The future of industrial bioprocessing: Batch or continuous? *Biotechnology and Bioengineering*, *112*, 648–651.

Dorai, H., Santiago, A., Campbell, M., Tang, Q. M., Lewis, M. J., Wang, Y., Lu, Q., Wu, S., & Hancock, W. (2011). Characterization of the proteases involved in the N-terminal clipping of glucagon-like-peptide-1-antibody fusion proteins. *Biotechnology Progress*, *27*, 220–231. https://doi.org/10.1002/btpr.537

Dovgan, T., Golghalyani, V., Zurlo, F., Hatton, D., Lindo, V., Turner, R., Harris, C., & Cui, T. (2021). Targeted CHO cell engineering approaches can reduce HCP-related enzymatic degradation and improve mAb product quality. *Biotechnology and Bioengineering*, *118*, 3821–3831. https://doi.org/10.1002/bit.27857

Falkenberg, H., Waldera-Lupa, D. M., Vanderlaan, M., Schwab, T., Krapfenbauer, K., Studts, J. M., Flad, T., & Waerner, T. (2019). Mass spectrometric evaluation of upstream and downstream process influences on host cell protein patterns in biopharmaceutical products. *Biotechnology Progress*, *35*, e2788. https://doi.org/10.1002/btpr.2788

Farrell, A., Mittermayr, S., Morrissey, B., Mc Loughlin, N., Navas Iglesias, N., Marison, I. W., & Bones, J. (2015). Quantitative host cell protein analysis using two dimensional data independent LC-MS(E). *Analytical Chemistry*, *87*, 9186–9193. https://doi.org/10.1021/acs.analchem.5b01377

Fischer, S. K., Cheu, M., Peng, K., Lowe, J., Araujo, J., Murray, E., McClintock, D., Matthews, J., Siguenza, P., & Song, A. (2017). Specific immune response to phospholipase B-like 2 protein, a host cell impurity in lebrikizumab clinical material. *AAPS Journal*, *19*, 254–263. https://doi.org/10.1208/s12248-016-9998-7

Fukuda, N., Senga, Y., & Honda, S. (2019). Anxa2- and Ctsd-knockout CHO cell lines to diminish the risk of contamination with host cell proteins. *Biotechnology Progress*, *35*, 1–8.

Gilgunn, S., & Bones, J. (2018). Challenges to industrial mAb bioprocessing—Removal of host cell proteins in CHO cell bioprocesses. *Current Opinion in Chemical Engineering*, *22*, 98–106. https://doi.org/10.1016/j.coche.2018.08.001

Gillet, L. C., Navarro, P., Tate, S., Röst, H., Selevsek, N., Reiter, L., Bonner, R., & Aebersold, R. (2012). Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: A new concept for consistent and accurate proteome analysis. *Molecular & Cellular Proteomics*, *11*, 1–17.

Goey, C. H., Bell, D., & Kontoravdi, C. (2019). CHO cell cultures in shake flasks and bioreactors present different host cell protein profiles in the supernatant. *Biochemical Engineering Journal*, *144*, 185–192.

Goey, C. H., Tsang, J. M. H., Bell, D., & Kontoravdi, C. (2017). Cascading effect in bioprocessing—The impact of mild hypothermia on CHO cell behavior and host cell protein composition. *Biotechnology and Bioengineering*, *114*, 2771–2781.

Gronemeyer, P., Ditz, R., & Strube, J. (2016). DoE based integration approach of upstream and downstream processing regarding HCP and ATPE as harvest operation. *Biochemical Engineering Journal*, *113*, 158–166. https://doi.org/10.1016/j.bej.2016.06.016

Hilliard, W., & Lee, K. H. (2021). Systematic identification of safe harbor regions in the CHO genome through a comprehensive epigenome analysis. *Biotechnology and Bioengineering*, *118*, 659–675. https://doi.org/10.1002/bit.27599

Hogwood, C. E. M., Ahmad, S. S., Tarrant, R. D., Bracewell, D. G., & Smales, C. M. (2016). An ultra scale-down approach identifies host cell protein differences across a panel of mAb producing CHO cell line variants. *Biotechnology Journal*, *11*, 415–424. https://doi.org/10.1002/biot.201500010

Hogwood, C. E. M., Bracewell, D. G., & Smales, C. M. (2014). Measurement and control of host cell proteins (HCPs) in CHO cell bioprocesses. *Current Opinion in Biotechnology*, *30*, 153–160. https://doi.org/10.1016/j.copbio.2014.06.017

Hou, S., Jones, S. W., Choe, L. H., Papoutsakis, E. T., & Lee, K. H. (2013). Workflow for quantitative proteomic analysis of *Clostridium acetobutylicum* ATCC 824 using iTRAQ tags. *Methods*, *61*, 269–276. https://doi.org/10.1016/j.ymeth.2013.03.013

Husson, G., Delangle, A., O'Hara, J., Cianferani, S., Gervais, A., Van Dorsselaer, A., Bracewell, D., & Carapito, C. (2018). Dual data-independent acquisition approach combining global HCP profiling and absolute quantification of key impurities during bioprocess development. *Analytical Chemistry*, *90*, 1241–1247.

Jones, M., Palackal, N., Wang, F., Gaza-Bulseco, G., Hurkmans, K., Zhao, Y., Chitikila, C., Clavier, S., Liu, S., Menesale, E., Schonenbach, N. S., Sharma, S., Valax, P., Waerner, T., Zhang, L., & Connolly, T. (2021). "High-risk" host cell proteins (HCPs): A multi-company collaborative view. *Biotechnology and Bioengineering*, *118*, 2870–2885.

Joucla, G., Le Sénéchal, C., Bégorre, M., Garbay, B., Santarelli, X., & Cabanne, C. (2013). Cation exchange versus multimodal cation exchange resins for antibody capture from CHO supernatants: Identification of contaminating host cell proteins by mass spectrometry. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, *942–943*, 126–133. https://doi.org/10.1016/j.jchromb.2013.10.033

Kim, J. Y., Kim, Y.-G., & Lee, G. M. (2012). CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Applied Microbiology and Biotechnology*, *93*, 917–930.

Kim, T. K., Chung, J. Y., Sung, Y. H., & Lee, G. M. (2001). Relationship between cell size and specific thrombopoietin productivity in Chinese hamster ovary cells during dihydrofolate reductase-mediated gene amplification. *Biotechnology and Bioprocess Engineering*, *6*, 332–336.

Kol, S., Ley, D., Wulff, T., Decker, M., Arnsdorf, J., Schoffelen, S., Hansen, A. H., Jensen, T. L., Gutierrez, J. M., Chiang, A. W. T., Masson, H. O., Palsson, B. O., Voldborg, B. G., Pedersen, L. E., Kildegaard, H. F., Lee, G. M., & Lewis, N. E. (2020). Multiplex secretome engineering enhances recombinant protein production and purity. *Nature Communications*, *11*, 1908. https://doi.org/10.1038/s41467-020-15866-w

Kornecki, M., Mestmäcker, F., Zobel-Roos, S., Heikaus de Figueiredo, L., Schlüter, H., & Strube, J. (2017). Host cell proteins in biologics manufacturing: The good, the bad, and the ugly. *Antibodies*, *6*, 13. http://www.mdpi.com/2073-4468/6/3/13

Kreimer, S., Gao, Y., Ray, S., Jin, M., Tan, Z., Mussa, N. A., Tao, L., Li, Z., Ivanov, A. R., & Karger, B. L. (2017). Host cell protein profiling by targeted and untargeted analysis of data independent acquisition mass spectrometry data with parallel reaction monitoring verification. *Analytical Chemistry*, *89*, 5294–5302. https://doi.org/10.1021/acs.analchem.6b04892

Kumar, A., Baycin-Hizal, D., Wolozny, D., Pedersen, L. E., Lewis, N. E., Heffner, K., Chaerkady, R., Cole, R. N., Shiloach, J., Zhang, H., Bowen, M. A., & Betenbaugh, M. J. (2015). Elucidation of the CHO super-ome (CHO-SO) by proteoinformatics. *Journal of Proteome Research*, *14*, 4687–4703.

Legrand, C., Bour, J. M., Jacob, C., Capiaumont, J., Martial, A., Marc, A., Wudtke, M., Kretzmer, G., Demangel, C., Duval, D., & Hache, J. (1992). Lactate dehydrogenase (LDH) activity of the number of dead cells in the medium of cultured eukaryotic cells as marker. *Journal of Biotechnology*, *25*, 231–243.

Levy, N. E., Valente, K. N., Choe, L. H., Lee, K. H., & Lenhoff, A. M. (2014). Identification and characterization of host cell protein product-associated impurities in monoclonal antibody bioprocessing. *Biotechnology and Bioengineering*, *111*, 904–912.

Levy, N. E., Valente, K. N., Lee, K. H., & Lenhoff, A. M. (2016). Host cell protein impurities in chromatographic polishing steps for monoclonal antibody purification. *Biotechnology and Bioengineering*, *113*, 1260–1272. http://www.ncbi.nlm.nih.gov/pubmed/26550778

Lloyd, D. R., Holmes, P., Jackson, L. P., Emery, A. N., & Al-Rubeai, M. (2000). Relationship between cell size, cell cycle and specific recombinant protein productivity. *Cytotechnology*, *34*, 59–70.

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. *Molecular Systems Biology*, *14*, 1–23.

MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C., & MacCoss, M. J. (2010). Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*, *26*, 966–968.

Madsen, J. A., Farutin, V., Carbeau, T., Wudyka, S., Yin, Y., Smith, S., Anderson, J., & Capila, I. (2015). Toward the complete characterization of host cell proteins in biotherapeutics via affinity depletions, LC-MS/MS, and multivariate analysis. *mAbs*, *7*, 1128–1137. https://doi.org/10.1080/19420862.2015.1082017

Marcus, K., Lelong, C., & Rabilloud, T. (2020). What room for two-dimensional gel-based proteomics in a shotgun proteomics world? *Proteomes*, *8*, 17.

Molden, R., Hu, M., Yen, E. S., Saggese, D., Reilly, J., Mattila, J., Qiu, H., Chen, G., Bak, H., & Li, N. (2021). Host cell protein profiling of commercial therapeutic protein drugs as a benchmark for monoclonal antibody-based therapeutic protein development. *mAbs*, *13*, 1955811. https://doi.org/10.1080/19420862.2021.1955811

Orellana, C. A., Marcellin, E., Schulz, B. L., Nouwens, A. S., Gray, P. P., & Nielsen, L. K. (2015). High-antibody-producing Chinese hamster ovary cells up-regulate intracellular protein transport and glutathione synthesis. *Journal of Proteome Research*, *14*, 609–618.

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. *Biotechnology Letters*, *35*, 987–993.

Park, J. H., Jin, J. H., Lim, M. S., An, H. J., Kim, J. W., & Lee, G. M. (2017). Proteomic analysis of host cell protein dynamics in the culture supernatants of antibody-producing CHO cells. *Scientific Reports*, *7*, 1–13. http://www.nature.com/articles/srep44246

Pezzini, J., Joucla, G., Gantier, R., Toueille, M., Lomenech, A.-M., Le Sénéchal, C., Garbay, B., Santarelli, X., & Cabanne, C. (2011). Antibody capture by mixed-mode chromatography: A comprehensive study from determination of optimal purification conditions to identification of contaminating host cell proteins. *Journal of Chromatography, A*, *1218*, 8197–8208. http://www.ncbi.nlm.nih.gov/pubmed/21982448

Reiter, L., Rinner, O., Picotti, P., Hüttenhain, R., Beck, M., Brusniak, M. Y., Hengartner, M. O., & Aebersold, R. (2011). MProphet: Automated data processing and statistical validation for large-scale SRM experiments. *Nature Methods*, *8*, 430–435.

Rupp, O., MacDonald, M. L., Li, S., Dhiman, H., Polson, S., Griep, S., Heffner, K., Hernandez, I., Brinkrolf, K., Jadhav, V., Samoudi, M., Hao, H., Kingham, B., Goesmann, A., Betenbaugh, M. J., Lewis, N. E., Borth, N., & Lee, K. H. (2018). A reference genome of the Chinese hamster based on a hybrid assembly strategy. *Biotechnology and Bioengineering*, *115*, 2087–2100. https://doi.org/10.1002/bit.26722

Shukla, A. A., & Thömmes, J. (2010). Recent advances in large-scale production of monoclonal antibodies and related proteins. *Trends in Biotechnology*, *28*, 253–261. https://linkinghub.elsevier.com/retrieve/pii/S0167779910000314

Sim, K. H., Liu, L. C. Y., Tan, H. T., Tan, K., Ng, D., Zhang, W., Yang, Y., Tate, S., & Bi, X. (2020). A comprehensive CHO SWATH-MS spectral library for robust quantitative profiling of 10,000 proteins. *Scientific Data*, *7*, 1–13.

Smith, L. M., & Kelleher, N. L. (2013). Proteoform: A single term describing protein complexity. *Nature Methods*, *10*, 186–187.

Tait, A. S., Hogwood, C. E. M., Smales, C. M., & Bracewell, D. G. (2012). Host cell protein dynamics in the supernatant of a mAb producing CHO cell line. *Biotechnology and Bioengineering*, *109*, 971–982.

Thiede, B., Koehler, C. J., Strozynski, M., Treumann, A., Stein, R., Zimny-Arndt, U., Schmid, M., & Jungblut, P. R. (2013). High resolution quantitative proteomics of hela cells protein species using stable isotope labeling with amino acids in cell culture (SILAC), Two-dimensional gel electrophoresis(2DE) and nano-liquid chromatograpohy coupled to an LTQ-OrbitrapMass spectrometer. *Molecular & Cellular Proteomics*, *12*, 529–538. https://doi.org/10.1074/mcp.M112.019372

UniProt Consortium. (2021). UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Research*, *49*, D480–D489. http://www.ncbi.nlm.nih.gov/pubmed/33237286

Valente, K. N., Choe, L. H., Lenhoff, A. M., & Lee, K. H. (2012). Optimization of protein sample preparation for two-dimensional electrophoresis. *Electrophoresis*, *33*, 1947–1957.

Valente, K. N., Lenhoff, A. M., & Lee, K. H. (2015). Expression of difficult-to-remove host cell protein impurities during extended Chinese hamster ovary cell culture and their impact on continuous bioprocessing. *Biotechnology and Bioengineering*, *112*, 1232–1242.

Valente, K. N., Levy, N. E., Lee, K. H., & Lenhoff, A. M. (2018). Applications of proteomic methods for CHO host cell protein characterization in biopharmaceutical manufacturing. *Current Opinion in Biotechnology*, *53*, 144–150. https://doi.org/10.1016/j.copbio.2018.01.004

Valente, K. N., Schaefer, A. K., Kempton, H. R., Lenhoff, A. M., & Lee, K. H. (2014). Recovery of Chinese hamster ovary host cell proteins for proteomic analysis. *Biotechnology Journal*, *9*, 87–99.

Walker, D. E., Yang, F., Carver, J., Joe, K., Michels, D. A., & Yu, X. C. (2017). A modular and adaptive mass spectrometry-based platform for support of bioprocess development toward optimal host cell protein clearance. *mAbs*, *9*, 654–663. https://doi.org/10.1080/19420862.2017.1303023

Walsh, G. (2018). Biopharmaceutical benchmarks 2018. *Nature Biotechnology*, *36*, 1136–1145. https://doi.org/10.1038/nbt.4305

Wilson, L. J., Lewis, W., Kucia-Tran, R., & Bracewell, D. G. (2019). Identification of upstream culture conditions and harvest time parameters that affect host cell protein clearance. *Biotechnology Progress*, *35*, 1–10.

Wu, X., Yang, Z.-Y., Li, Y., Hogerkorp, C.-M., Schief, W. R., Seaman, M. S., Zhou, T., Schmidt, S. D., Wu, L., Xu, L., Longo, N. S., McKee, K., O'Dell, S., Louder, M. K., Wycuff, D. L., Feng, Y., Nason, M., Doria-Rose, N., Connors, M., … Mascola, J. R. (2010). Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science*, *329*, 856–861. http://www.ncbi.nlm.nih.gov/pubmed/20616233

Yamaguchi, Y., & Ruoslahti, E. (1988). Expression of human proteoglycan in Chinese hamster ovary cells inhibits cell proliferation. *Nature*, *336*, 244–246. http://www.ncbi.nlm.nih.gov/pubmed/21404259

Yuk, I. H., Nishihara, J., Walker, D., Huang, E., Gunawan, F., Subramanian, J., Pynn, A. F. J., Yu, X. C., Zhu-Shimoni, J., Vanderlaan, M., & Krawitz, D. C. (2015). More similar than different: Host cell protein production using three null CHO cell lines. *Biotechnology and Bioengineering*, *112*, 2068–2083.

Zhang, Q., Goetze, A. M., Cui, H., Wylie, J., Tillotson, B., Hewig, A., Hall, M. P., & Flynn, G. C. (2016). Characterization of the co-elution of host cell proteins with monoclonal antibodies during protein A purification. *Biotechnology Progress*, *32*, 708–717. https://doi.org/10.1002/btpr.2272

Zhang, Q., Goetze, A. M., Cui, H., Wylie, J., Trimble, S., Hewig, A., & Flynn, G. C. (2014). Comprehensive tracking of host cell proteins during monoclonal antibody purifications using mass spectrometry. *mAbs*, *6*, 659–670. https://doi.org/10.4161/mabs.28120

Zhu-Shimoni, J., Yu, C., Nishihara, J., Wong, R. M., Gunawan, F., Lin, M., Krawitz, D., Liu, P., Sandoval, W., & Vanderlaan, M. (2014). Host cell protein testing by ELISAs and the use of orthogonal methods. *Biotechnology and Bioengineering*, *111*, 2367–2379.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Hamaker, N. K., Min, L., & Lee, K. H. (2022). Comprehensive assessment of host cell protein expression after extended culture and bioreactor production of CHO cell lines. *Biotechnology and Bioengineering*, 119, 2221–2238. https://doi.org/10.1002/bit.28128