Article

Holographic Imaging of Insect Cell Cultures: Online Non-Invasive Monitoring of Adeno-Associated Virus Production and Cell Concentration

Daniel A. M. Pais [1,2], **Paulo R. S. Galrão** [1], **Anastasiya Kryzhanska** [1], **Jérémie Barbau** [3], **Inês A. Isidro** [1,2] **and Paula M. Alves** [1,2,*]

[1] Animal Cell Technology Unit, iBET, Instituto de Biologia Experimental e Tecnológica, Apartado 12, 2780-901 Oeiras, Portugal; dpais@ibet.pt (D.A.M.P.); paulo.galrao@ibet.pt (P.R.S.G.); a.kryzhanska@gmail.com (A.K.); iaisidro@ibet.pt (I.A.I.)

[2] ITQB-NOVA, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal

[3] OVIZIO Imaging Systems SA/NV, 1180 Brussels, Belgium; jeremie.barbau@ovizio.com

* Correspondence: marques@ibet.pt; Tel.: +351-21-4469416

Abstract: The insect cell-baculovirus vector system has become one of the favorite platforms for the expression of viral vectors for vaccination and gene therapy purposes. As it is a lytic system, it is essential to balance maximum recombinant product expression with harvest time, minimizing product exposure to detrimental proteases. With this purpose, new bioprocess monitoring solutions are needed to accurately estimate culture progression. Herein, we used online digital holographic microscopy (DHM) to monitor bioreactor cultures of Sf9 insect cells. Batches of baculovirus-infected Sf9 cells producing recombinant adeno-associated virus (AAV) and non-infected cells were used to evaluate DHM prediction capabilities for viable cell concentration, culture viability and AAV titer. Over 30 cell-related optical attributes were quantified using DHM, followed by a forward stepwise regression to select the most significant ($p < 0.05$) parameters for each variable. We then applied multiple linear regression to obtain models which were able to predict culture variables with root mean squared errors (RMSE) of 7×10^5 cells/mL, 3% for cell viability and 2×10^3 AAV/cell for 3-fold cross-validation. Overall, this work shows that DHM can be implemented for online monitoring of Sf9 concentration and viability, also permitting to monitor product titer, namely AAV, or culture progression in lytic systems, making it a valuable tool to support the time of harvest decision and for the establishment of controlled feeding strategies.

Keywords: AAV-adeno-associated virus; insect cell-baculovirus; cell culture monitoring; digital holographic microscopy; process analytical technology

1. Introduction

After the FDA launched the Process Analytical Technology (PAT) initiative in 2004 [1], an increased effort was put in place by the manufacturers of biological products to comply with PAT requirements. The PAT initiative is a guidance for the pharmaceutical industry for the development of new products and production processes, with the main focuses on: **(i)** increasing product and process knowledge through the identification of the product critical quality attributes and the process parameters affecting it; and **(ii)** monitoring in real-time the identified critical process parameters and the product quality characteristics, ensuring manufacturing robustness and an increased quality assurance to achieve the required levels of compliance [1–3].

Label-free methodologies are preferred, especially in biopharmaceutical processes, since they allow the monitoring of cell culture without adding any compounds which would influence cellular behavior. Most cell culture monitoring methods employing label-free methodologies are based on spectroscopic techniques, which have been widely used for cell culture process monitoring. Examples include the use of dielectric spectroscopy and turbidimetry/light scattering probes for the determination of cell concentration [4,5], as well as the use of Raman [6,7], infrared [8] and fluorescence [9] spectroscopy, which allow the quantification of metabolites based on direct spectra quantification, but also the indirect determination of cell concentration and product formation based on chemometric analysis.

A label-free alternative to spectroscopic techniques is imaging-based cell culture monitoring. Since cells are mostly transparent, these systems rely on several strategies to generate the needed image contrast [10,11]. One example of an imaging technique with proven demonstrations for live cell imaging is Digital Holographic Microscopy (DHM) [12]. Briefly, DHM provides quantitative phase imaging (QPI), quantifying the phase shift of the light after it has passed through the object of focus, such as cells. This light phase difference is encoded in a hologram which is used to construct high-resolution intensity and quantitative-phase images of the cell while also providing quantitative parameters related with light phase and intensity [11,13]. The way light is scattered after interacting with cells depends on factors such as cell thickness, circularity or intracellular composition [10,11,14,15]. As such, DHM can be used to extract important information from the cell state and has proven useful for several cell-based applications: identification of morphological parameters distinguishing between epithelial and mesenchymal cells [13], detecting cell division in endothelial cells [15] and developing cell proliferation [12] or cytotoxic assays [16]. In particular, infected cells will have different intracellular structure than uninfected cells [3,17,18]. Furthermore, as demonstrated by Ugele and colleagues, DHM-based detection of the intracellular composition of infected erythrocytes even allowed to distinguish between different infection phases in the malaria *P. falciparum* life cycle [17]. The ability to detect infected cells as well as cell concentration and viability makes DHM inherently attractive to monitor the progress of infection-based biopharmaceutical production systems, such as the insect cell-baculovirus system [19].

Insect cells are one of the preferred hosts for viral vector manufacturing for vaccines and gene therapy purposes, since they can be grown in suspension to high cell densities in serum free media [20,21]. However, to maximize product yields it is determinant to infect cells at low cell concentration, to prevent the so-called "cell density effect", a drop on the specific productivity of the cell when infection takes place at a high cell concentration, reviewed in Palomares et al. [22]. The optimal cell concentration for infection and the definition of "low" and "high" cell concentration are dependent on the cell type, culture medium used and recombinant product being expressed [23,24]. Moreover, baculovirus is a lytic virus, which can lead to the release of intracellular proteases into the culture medium, possibly degrading the recombinant product after it has been released into the medium. As such, both culture viability and cell concentration are critical process parameters for this system.

Our group and others have addressed ways to monitor this system using fluorescence [9] or dielectric [3,20,25,26] spectroscopies, as well as using image-based technologies, in particular for measuring the progress of baculovirus infection [27–29]. DHM can go one step further, by monitoring not only the cell diameter increase after baculovirus infection, but also the evolution profile of several cell characteristics, allowing to explore the possibility to observe baculovirus or AAV-induced changes in suspension insect cells in real-time.

In this work, we used the iLine F differential DHM system (DDHM) (Ovizio Imaging Systems SA/NV) for real-time monitoring of a Sf9 culture infected with baculovirus, expressing recombinant adeno-associated virus (AAV) type 2. AAV is widely used as a gene therapy viral vector, due to its lack of known pathogenicity, broad tissue tropism coupled with long-term transgene expression and ability to withstand harsh manufacturing conditions [30]. Estimation of AAV titer in real-time is desirable in order to harvest when its concentration is higher. Moreover, monitoring this system in real-time can support the time of harvest decision, an important process variable to consider giving the lytic nature of the baculovirus and consequential release of proteases to the medium when cells start lysing.

Since DDHM can be used to detect infected cells, we further explored this capability for monitoring the AAV titer in our cultures along with the development of predictive models for viable cell concentration and viability. Using the culture-related morphologic and optical attributes quantified with iLine F, we used forward stepwise regression to find the attributes associated with viable cell concentration, viability and intra and extracellular AAV titer. We validated this approach using leave one batch out (LOBO) and 3-fold cross-validation strategies. As such, we demonstrate that DDHM can be used not only for monitoring Sf9 cell concentration and viability but also for assessing AAV production kinetics in the insect cell system.

2. Materials and Methods

2.1. Cell Line and Culture Medium

Spodoptera frugiperda Sf9 cells were obtained from Thermo Fisher Scientific (No. 11496015) and routinely cultivated in 500 mL glass Erlenmeyer flasks with 50 mL working volume of SF900-II medium (Gibco^{TM}), at 27 °C with an agitation rate of 100 rpm in an Innova 44R incubator (orbital motion diameter = 2.54 cm, Eppendorf). Cell concentration and viability were determined using a Cedex HiRes Analyzer (Roche).

2.2. AAV and Baculovirus Infection and Titration

We used the two baculovirus system for AAV production (reviewed in Merten [31]).

The recombinant *Autographa californica* nucleopolyhedrovirus encoding the green fluorescence protein (GFP) transgene under the control of the cytomegalovirus promoter (CMV-GFP) and flanked by AAV2 inverted terminal repeats (ITR) regions was kindly provided by Généthon and was titrated and amplified in house, as described for the *rep/cap* baculovirus (below).

The plasmid containing AAV2 *rep* and *cap* genes was a gift from Robert Kotin (Addgene plasmid #65214) [32]. Recombinant baculovirus was produced using the Bac-to-Bac® Baculovirus Expression System (Invitrogen), according to the manufacturer's instructions. Baculovirus amplification was performed as described elsewhere [9].

Recombinant adeno-associated virus (AAV) intra and extracellular titer was estimated separately using a commercially available sandwich ELISA kit (Progen Biotechnik GmbH), according to the manufacturer's instructions. This kit detects a conformational epitope present in assembled AAV capsids.

2.3. Bioreactor Cultures and Sample Processing

Benchtop 1 L bioreactor runs were performed in BIOSTAT® DCU-3 (Sartorius), equipped with two Rushton turbines. Temperature control (27 °C) was achieved using a water recirculation jacket and gas supply was provided by a ring sparger in the bottom of the vessel. Dissolved oxygen (DO) concentration was kept at 30% by cascade controlling the stirring rate (70–270 rpm) and the N_2/air ratios in a mixture of air and N_2 (0.01 vvm).

Several runs were performed to establish the standard culture progression profile. The iLine F system (Ovizio Imaging Systems SA/NV) was then used to monitor a growth batch and a production (infected) batch, which had similar culture profiles for cell concentration, viability and AAV production titer when compared to previous culture replicates [9]. The growth batch consisted of a Sf9 batch culture monitored until cell death due to nutrient starvation, which occurred after 10 days of culture. The infected batch consisted of a Sf9 culture infected with two baculovirus vectors to express recombinant adeno-associated virus type 2, harvested on day 6 after inoculation. Sf9 cells were inoculated at 0.5×10^6 cells/mL for both reactors. The infected batch was infected 31 h after inoculation, when viable cell concentration reached 1×10^6 cells/mL, with a multiplicity of infection (MOI) of 0.05 plaque forming units per cell, for each baculovirus. The two-baculovirus strategy was used, in which one baculovirus codes for the AAV2 *rep* and *cap* genes and the other provides the GFP transgene flanked by the AAV ITRs.

In the iLine F system, a single-use autoclavable closed-loop tube is inserted into a standard 19 mm bioreactor top port. This sampling tube contains in the other end a cartridge with the imaging chamber. After sterilization, the sampling tube is connected to a pump motor. Cell culture is continuously aspirated through the sampling tube to the imaging chamber and then returned to the cell culture vessel. The setup is controlled using the *OsOne* software (Ovizio Imaging Systems SA/NV), which controls the sampling rate and image analysis by a holographic microscope. Images are acquired every minute, but image processing occurs in batches of 25, thus yielding a new timepoint every 30 min (the 5 remaining minutes are used for background elimination and attribute calculations). Image processing consists in (1) image focus, (2) holographic fingerprint acquisition for every cell present in the image, (3) computation of 66 image-related attributes for every cell. Figure S1 exemplifies the cell culture and hologram evolution profiles. Acquisition of 25 images for a culture timepoint is presented in Video S1 for the bright field images and Video S2 for the phase images.

Sampling for the determination of reference variables was performed daily for the growth batch and three times per day for the infected batch. At each sampling point, cell concentration and viability were measured using Cedex HiRes Analyzer (Roche). Additionally, for the infected batch, a clarification step was performed (200 g, 10 min, 4 °C) to recover intra and extracellular AAV. Supernatant was subjected to a further clarification step (2000 g, 20 min, 4 °C) and stored at −80 °C for offline analysis. Intracellular AAV was extracted from cell pellets with TNT buffer, consisting of 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1% Triton X-100, 10 mM $MgCl_2$ [32], to which a 0.5% solution of sodium deoxycholate was added to further increase the release of intracellular AAV from pelleted cells [33]. After 10 min of incubation at 22 °C, the suspension was centrifuged (2000 g, 20 min, 4 °C) and the supernatant stored at −80 °C for offline analysis.

2.4. Modeling Strategy and Software

2.4.1. Dataset

After run completion, for each timepoint, the average for each attribute was calculated, considering all the cells present in the 25 images acquired per timepoint. This resulted in 499 timepoints for the growth batch and 275 timepoints for the infected batch (online data). These data were smoothed using a moving average of two hours, corresponding to 4 datapoints. The reference data consisted of 14 samples for the growth batch and 23 samples for the infected batch, with determination of the four reference variables (viable cell concentration, viability, extracellular volumetric AAV titer and intracellular specific AAV titer) for each sample. The data for modeling consisted of each one of the reference datapoints time-aligned with the corresponding online datapoints, yielding a matrix of [37 rows × 4 reference variables columns × 66 columns with averaged attributes].

All analyses and modeling were performed in JMP v14 (Statistical Analysis System institute).

Potential outliers in the reference data were identified by a visual inspection of the data time-course profile and confirmed by calculating the jackknife distances for each datapoint. The JMP jackknife outlier identification method relies on estimates of the mean, standard deviation, and correlation matrix that do not include the observation itself.

2.4.2. Attribute Selection and Stepwise Regression

OsOne calculates 66 attributes per each cell. However, due to the high collinearity of some attributes and to prevent model over-fitting [34], the Pearson correlation coefficient was calculated for every attribute pair. For pairs with a high correlation (Pearson correlation coefficient absolute value > 0.95), one of the attributes was excluded from further analysis. This process was iterated until no attribute had a correlation coefficient higher than 0.95 or lower than −0.95, reducing the initial 66 attributes to 30.

For model training, the JMP "Fit model" platform was used. Briefly, the 30 attributes selected were subjected to a forward stepwise regression to find the most significant for the prediction of each

of the reference variables. In a forward stepwise regression method, the most significant attribute is identified and added to the model, followed by identification and inclusion in the model of the second most significant attribute and so on. This process was stopped when the next term added was considered not significant (*p*-value > 0.05).

Since this biological system has non-linear variables, which can be observed on the viable cell concentration and AAV titer profiles, after identification of the most significant attributes for every variable, a second model was created by performing the same forward stepwise regression technique using the significant terms ("main effects") and their interactions and quadratics. The final forward stepwise regression model (main effects only or with interactions) was chosen by comparison of prediction profiles and root mean squared error (RMSE).

2.4.3. Model Training and Validation

Multiple linear regression models were built based on the forward stepwise regression strategy. Two validation strategies were used to assess model prediction capabilities and overfitting: leave one batch out (LOBO) and 3-fold cross-validation (3CV). For LOBO models, the stepwise attribute selection strategy mentioned in the previous section was applied to one batch only. After finding the most significant parameters and determining the model coefficients for each parameter by multiple linear regression, the model was applied to the remaining batch for validation. This strategy was successfully applied to viability models but resulted in significant overfitting for viable cell concentration due to the significant differences in the variable ranges between the two batches. As such, an alternative LOBO strategy was used, in which parameter selection was performed using the reference data from both batches, followed by training of each batch separately. The obtained model was then used for predicting the remaining batch for validation purposes.

For 3CV, the significant parameters were identified by applying forward stepwise regression to the reference data for both batches, followed by multiple linear regression for model fitting using both batches. Model validation was performed by dividing the dataset (37 timepoints) into 3 random partitions, using two for model training with the selected parameters and predicting the third partition. The process was repeated for the two remaining partitions.

The contribution of each parameter to the final model was calculated by dividing the logworth value for each parameter by the sum of the logworth for all parameters (logworth is defined as $-\log_{10}(p\text{-value})$).

RMSEs for calibration (RMSEC) and validation (RMSEV) were calculated for all models (Equation (1)). In Equation (1), $\hat{y}$ represents a vector of model-predicted values and y represents the corresponding reference data; *ncal* and *nval* represent the number of samples in the calibration or validation set, respectively; max(y) and min(y) refer to the maximum and minimum values for the reference data, respectively. Normalized RMSE (nRMSE) was obtained by dividing the RMSE by the variable range.

The correlation coefficients of calibration and validation were calculated according to Equation (2) using calibration (R^2) or validation (Q^2) data. R^2 is a measure of how well the chosen model fits the calibration data while Q^2 measures how the obtained model fits the validation dataset, which is not used to fit the model, being indicative of the model predictive power for new data. σ^2 represents sample variance.

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{ncal}(\hat{y}-y)^2}{ncal}}. \quad RMSEV = \sqrt{\frac{\sum_{i=1}^{nval}(\hat{y}-y)^2}{nval}}. \quad nRMSE = \frac{RMSE}{\max(y)-\min(y)} \qquad (1)$$

$$R^2 = 1 - \frac{RMSEC^2}{\sigma^2}. \quad Q^2 = 1 - \frac{RMSEV^2}{\sigma^2}. \qquad (2)$$

3. Results

3.1. Digital Holographic Microscopy Can Be Used for Monitoring Viable Cell Concentration and Viability

Here, we studied the applicability of the iLine F system for monitoring critical process variables in the insect cell-baculovirus system, for the production of recombinant adeno-associated viral vectors (AAV). The critical process variables under analysis were viable cell concentration, cell viability and intra and extracellular AAV titers.

Models were trained using two batches, one infected (AAV production) and one uninfected (cell growth). These have similar viability profiles (Figure 1A), differing only in the time to onset of viability decrease, but are distinct in the viable cell concentration ranges achieved (Figure 1B), as well as the AAV production profiles (Figure 2).

The preferential validation strategy consisted in using one batch for model calibration and the other one as validation set (leave one batch out, LOBO). The high Q^2 obtained for viability (0.72 and 0.92 for validation with growth and infected batches, Figure 1A) supports the feasibility of using iLine F for monitoring viability in this process, even using only one batch for model calibration. The lower Q^2 score obtained for growth batch is mainly due to an underestimation of viability in the growth phase, but the prediction profiles for the death phase (more relevant for this system) are accurate for both runs. For viable cell concentration, the large range difference between runs causes the model to overfit the calibration batch, therefore severely underestimating viable cell concentration when predicting the growth batch, although with the correct viable cell concentration profile ($Q^2 = 0.66$) and failing to capture the correct trend for the infected batch ($Q^2 = 0.34$) (Figure 1B).
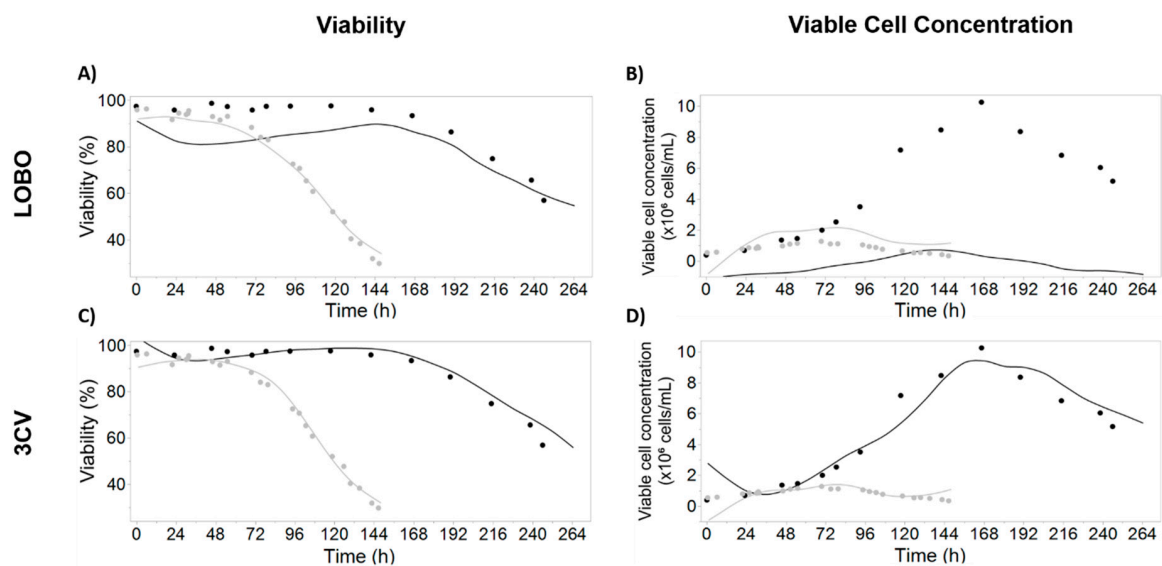

Figure 1. Viability (left) and viable cell concentration (right) predictions using leave one batch out (LOBO, top) and 3-fold cross-validation (3CV, bottom) models. Growth batch is represented in black and infected batch is colored in grey. The lines represent model-predicted values; the filled circles represent reference data; the empty circles were considered outliers and excluded from modeling. For LOBO models, the lines represent the prediction obtained with the model calibrated in the remaining batch. For 3CV models, the lines represent the model built using data from both batches. (**A**) Observed and predicted values for viability using LOBO for model validation; (**B**) Observed and predicted values for viable cell concentration using LOBO for model validation; (**C**) Observed and predicted values for viability using 3CV for model validation; (**D**) Observed and predicted values for viable cell concentration using 3CV for model validation. Model parameters and coefficients are presented in Table S1.
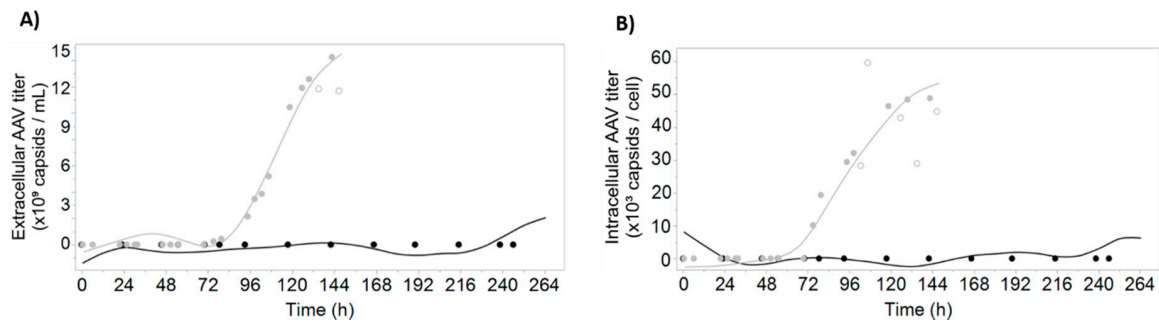
Figure 2. AAV titer predictions for both batches. The growth batch is represented in black and the infected batch is colored in grey. The lines represent model-predicted values; the filled circles represent reference data; the empty circles represent datapoints considered outliers and excluded from modeling. Models were calibrated using the reference data for both batches (filled circles). The prediction data represented by the smooth lines were obtained by applying the model to the real-time differential digital holographic microscopy data. (**A**) Observed and predicted values for extracellular AAV titer; (**B**) observed and predicted values for intracellular specific AAV titer. Model parameters and coefficients are presented in Table S1.

The second validation strategy tested for viability and viable cell concentration was the 3-fold cross-validation (3CV) (Figure 1C,D). Models were built using data from both batches and a 3CV strategy was applied to measure the model predictive power and confirm that these are not overfitting while simultaneously allowing the identification of the most important DDHM attributes for variable prediction.

Applying the 3CV model to iLine F real-time data yields greatly improved predictions when compared with LOBO models (Figure 1, $Q^2 = 0.98$ for viability and $Q^2 = 0.93$ for viable cell concentration). Although less robust, this strategy was necessary so that model coefficients could account for the differences in the variable range between the two batches. The final model parameters and coefficients are presented in Table S1. For comparison purposes, the predictions for viable cell concentration and viability using Ovizio proprietary models are shown in Figure S2. Except for the viable cell concentration LOBO model calibrated in the infected batch, no models consider parameter interactions, since the models containing only the main effects possess an equal or better predictive score than the ones considering interactions and quadratics.

3.2. Prediction of AAV Titers Using Digital Holographic Microscopy

Given that the used dataset consists of two batches, from which only one is expressing AAV, the LOBO strategy cannot be used for modeling AAV-related variables. As such, the 3-fold cross-validation (3CV) strategy described in the previous section was used to calibrate prediction models for extracellular volumetric AAV titer ($Q^2 = 0.97$) and intracellular specific AAV titer ($Q^2 = 0.99$) (Figure 2). The AAV production trend is captured with our modeling strategy, highlighting the potential of using multiple linear regression for identification of the most important optical attributes measured with DDHM and monitoring AAV production profiles in the insect cell system.

To confirm that the obtained models are not overfitting the data, the coefficients of correlation for the calibration and validation set for every partition were calculated for the four variables under study (Table S2). For each variable, the nRMSE for each partition are comparable in magnitude. Moreover, for each partition, the nRMSE values obtained for validation are on average 1.8% higher than the ones obtained for calibration, confirming that the 3CV models are not overfitting the data.

The high adjusted coefficients of correlation for calibration and validation for the models shown in Figures 1 and 2 indicate that good prediction models were obtained, with the exception of the LOBO viable cell concentration model (Figure 3, Table S3). The feasibility of using DDHM for bioprocess monitoring is demonstrated by the acceptable Q^2 (0.74) using LOBO for viability prediction, and by the

high cross-validation Q^2 for all variables (0.93 to 0.98). For the LOBO viable cell concentration models, the negative value was obtained when considering the Q^2 for both batches simultaneously, due to the high discrepancy in the variable range and the overfitting in each calibration model. Individual Q^2 are 0.66 for prediction of growth batch and 0.34 for prediction of infected batch. The Q^2 values for 3CV models are very close to the corresponding R^2, demonstrating that the chosen model is appropriate to describe both the calibration data and new datapoints. Altogether, this demonstrates that using only two batches with different AAV production profiles is enough to find the DDHM attributes likely relevant for AAV production.
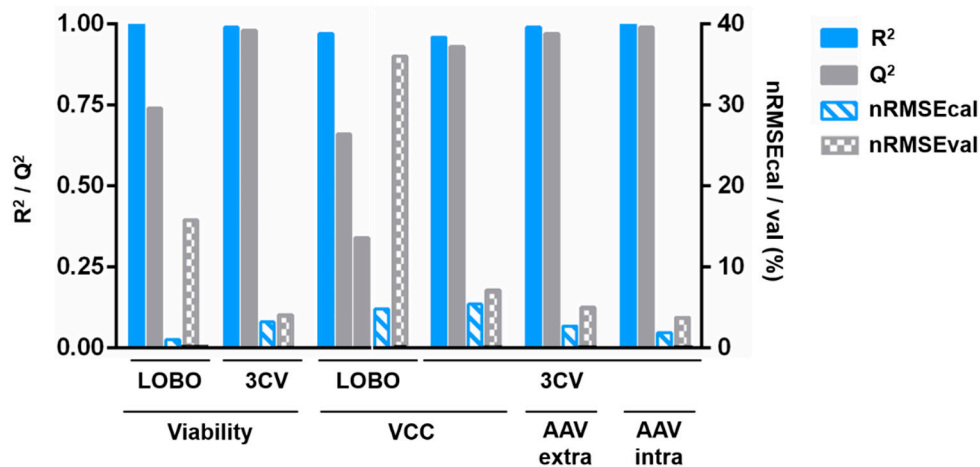

Figure 3. Quality characteristics overview for the models presented in Figure 1 and Figure 2. R^2 and Q^2 are the correlation coefficients of calibration and validation, respectively. Also depicted are the normalized root mean squared errors (nRMSE) for calibration and validation which are scaled by the variable range. For the LOBO viable cell concentration models, the difference in the cell concentration ranges and the fact that the prediction models overfit the calibration batch result in a negative Q^2 (-0.69) when data from both batches are considered. As such, we chose to depict the Q^2 for each batch separately (0.66 for prediction of growth batch and 0.34 for prediction of infected batch). CV—3-fold cross-validation; LOBO—leave one batch out; R^2—correlation coefficient of calibration; nRMSE—normalized root mean squared error; Q^2—correlation coefficient of validation; VCC—viable cell concentration. Raw data are provided in Table S3.

3.3. Time-Course Profiles of Morphological and Optical Parameters Measured with DDHM

One of the advantages of using DDHM for monitoring cell culture processes in real-time is the number of cell and image attributes that are calculated and the possibility to analyze the attribute evolution profile over culture time. While some of these attributes have an obvious biological meaning (for instance "Cell Radius"), most of them do not have a direct biological meaning per se. Still, some of the attributes show an evolution over culture progression and some are clearly correlated with the critical process variables studied in this work, such as culture viability (Figure 4C,E,F), viable cell concentration (Figure 4C,G) and extracellular AAV titer profiles (Figure 4A,B,D). These attributes were included in the final multiple linear regression prediction models with varying contributions for the overall model (Figure 5 and Table S1). Our final models have between 5 and 12 parameters, excluding the intercept term (Figure 5 and Table S2).
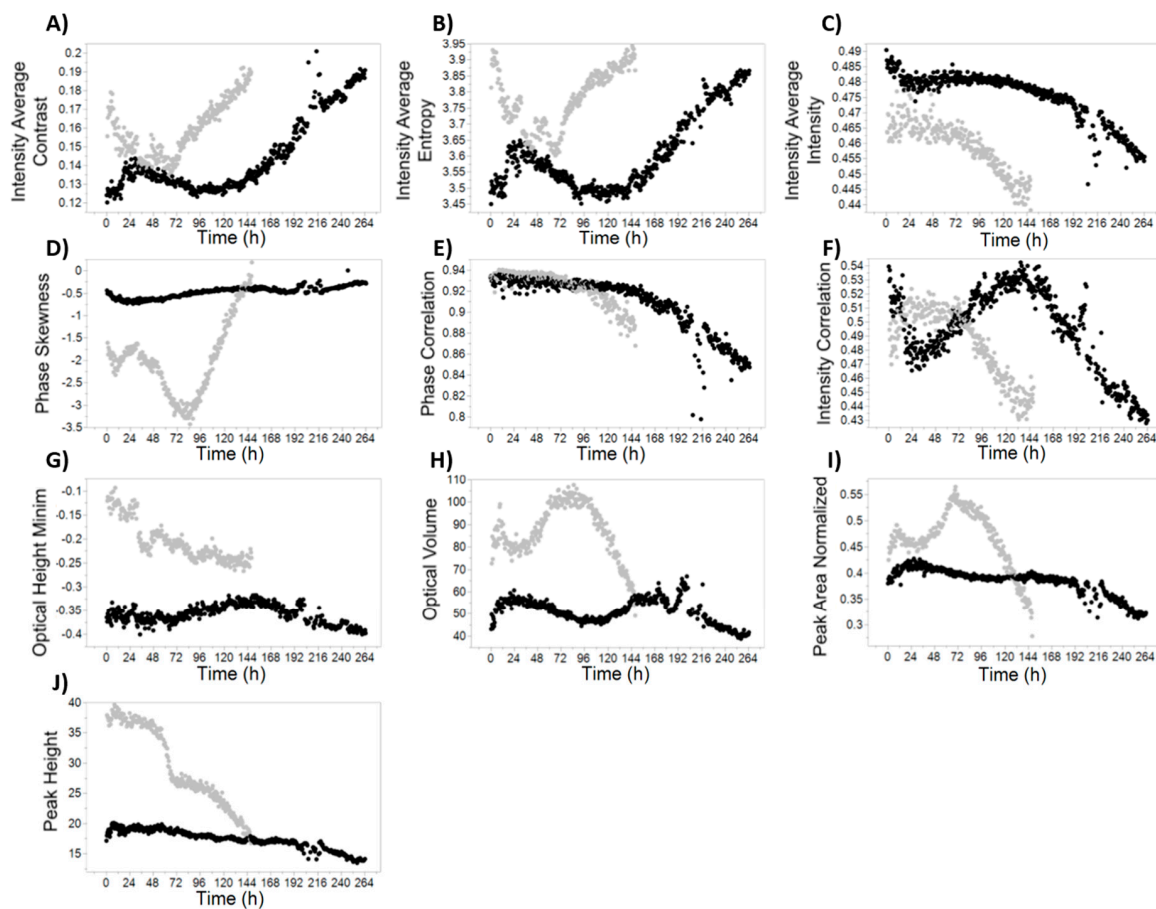
Figure 4. Time-course profiles for selected DDHM attributes. The growth batch is represented in black, while the infected batch is colored in grey. Measurements were obtained every 30 min. (**A**) Intensity Average Contrast; (**B**) Intensity Average Entropy; (**C**) Intensity Average Intensity; (**D**) Phase Skewness; (**E**) Phase Correlation; (**F**) Intensity Correlation; (**G**) Optical Height Minimum; (**H**) Optical Volume; (**I**) Peak Area Normalized; (**J**) Peak Height.

3.4. Model Parameters Have Biological Significance

With iLine F, more than 60 attributes are calculated per cell. These are related with the cell morphology (e.g., "circularity"), the light optical characteristics (e.g., "maximum intensity"), the light phase texture (e.g., "phase skewness") or the light intensity texture (e.g., "intensity correlation"). Overall, the parameters with a larger contribution for the obtained models are related with light intensity and phase characteristics (Figure 5).

Regardless of their relative contribution, some parameters are present in most of the models. Examples include "optical height maximum", "phase average uniformity", "intensity correlation", "intensity average intensity" and "phase skewness". The parameters present in the predictive models for viability, viable cell concentration and AAV extracellular titer are specially interesting because the respective variables are also correlated: viability is the quantitative measurement of the decrease in viable cell concentration, and AAV extracellular titer increase is mostly due to cell lysis [35].
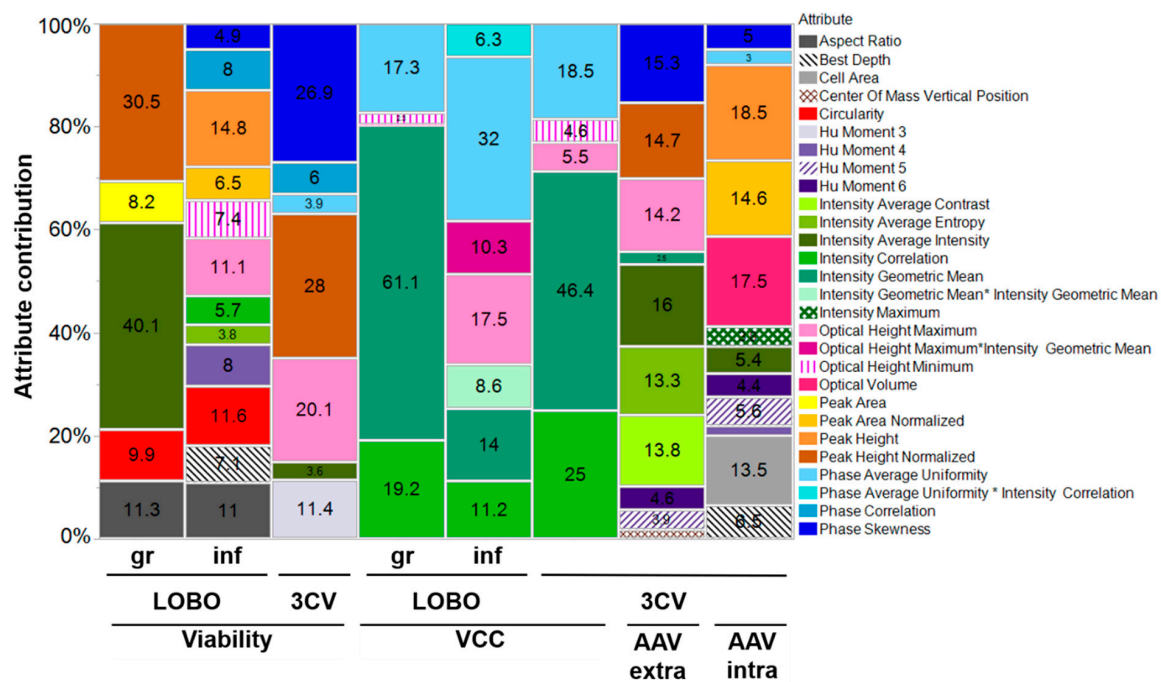
Figure 5. Relative contribution of each parameter to the final models. For the leave one batch out (LOBO) models, the batch used for model calibration is indicated (gr—growth; inf—infected). For the 3CV models, the coefficients presented are related to the model using both batches. Relative importance was calculated using the logworth for each parameter (Table S1).

"Phase skewness" was considered significant for both AAV models, with a total contribution for the overall model of 15% for extracellular AAV and 5% for specific AAV (Figure 5). Although a much higher contribution for the intracellular specific AAV model was expected, the fact that "phase skewness" is also present in some viability models may explain its high contribution for extracellular AAV. As expected, this parameter has negative coefficients for viability models and positive for the extracellular AAV prediction model (Table S1).

Another important consideration is the presence of highly correlated attributes, which may confound biological interpretation of the model contributions. For instance, "phase average uniformity", a measure of the uniformity of the light phase in each cell, is strongly correlated ($R^2 = 0.91$) with "radius variance", the variance of the cell radius, which is inversely correlated with circularity ($R^2 = -0.97$). In conclusion, a cell with an increased "phase average uniformity" has a less spherical shape ($R^2 = -0.88$). The pairwise Pearson correlations for every pair of attributes are shown in Figure S3.

4. Discussion

The aim of this study was to explore the applicability of differential digital holographic microscopy (DDHM) to monitor important process parameters in the insect cell-baculovirus system, including the AAV production kinetics. Specifically, the Ovizio iLine F system was used. A forward stepwise regression technique combined with multiple linear regression was applied to the morphological and physiological attributes quantified by DDHM, successfully identifying candidates relevant for viable cell concentration, viability and intra and extracellular AAV titer.

Currently, there is a lack of methods available for monitoring of viral particles production during cell culture [3]. The existing methods explore chemometrics approaches, by measuring process variables related with the viral production kinetics [9] or changes in the morphological and physiological alterations of the cells [3,36]. In particular, for the baculovirus system, these methods are mostly based on the known increase of cell diameter upon baculovirus infection [27–29], although they were used as an assay rather than for in-culture determination.

Viability is one of the most important process variables to consider in many viral-based systems, being related with product quality and influencing harvest decision [9,36,37]. In both batches, cell viability decreases in the end of the culture. However, the onset of viability decrease occurs with different biological triggers: while in the infected batch cell viability decreases due to baculovirus-induced cell lysis, in the growth batch cells died by nutrient starvation. This validates the applicability of DDHM, but also provides a possible explanation to why the parameters present in each LOBO viability model are different (Figure 5), since the biological reason for the cell death was different. While some of the identified model parameters have a clear similarity with viability profiles (e.g., Figure 4C,E,F); in general, these are not the most important for the viability prediction models. Given the small dataset used, the parameters more important for the models may be in fact distinguishing between infected and growth batch (e.g., Figure 4J, "peak height") followed by fine-tuning using the attributes with the similar viability profile. While addition of more calibration batches would increase the confidence in the determination of the parameters associated with viability, the prediction profiles using LOBO (Figure 1A) show DDHM possess enough predictive power for prediction of viability using only one batch for calibration, and additional batches are expected to further improve the prediction accuracy.

Although the lack of an independent testing set for viable cell concentration and AAV predictions prevents assessment of model validation for new batches, our aim was to explore iLine F applicability to study this production system. Furthermore, the identification and analysis of the parameters correlated with the modeled variables provides valuable biological insights for AAV production in insect cells.

Most of the attributes calculated with DDHM have no biological meaning per se, but can be used to characterize a dynamic phenotype, indicative of the cell adaptation to different biological situations [10,34]. However, some of these parameters may have a possible biological explanation. For instance, "phase correlation", a measure of how neighboring pixels are correlated, has a time-profile very similar to the culture viability profiles (Figure 4E). A possible explanation may be related with the increase in intracellular complexity during baculovirus infection. The cellular phenotype alterations occurring throughout baculovirus infection and the release of intracellular compounds to the culture supernatant during lysis will increase the entropy inside the cell, consequently resulting in less correlation of each pixel with its neighbors and a decrease in the phase correlation profiles. For viable cell concentration, it is expected that the attributes more predictive for viable cell concentration are related to light intensity, due to light dispersion caused by suspension cells, analogous to turbidimetry-based measurements (Figure 5). In fact, one of the parameters common to all three viable cell concentration models is "intensity correlation" (Figure 4F), a measure of how correlated the intensity of one pixel is to the intensity of its neighbors over the cell surface.

Interestingly, "phase skewness" has a time-course profile very similar with extracellular AAV production (Figure 4D) for both batches. We believe this increase in "phase skewness" concomitant with AAV production is due to a combination of several factors: the cell nucleus and nucleolus possess a higher molecular density than surrounding regions, and are likely the cell organelles better detected using QPI due to their higher phase contrast [15]. Additionally, AAV capsid assembly takes place in the nucleolus [38]. We hypothesize that AAV production in the nucleolus of infected cells increases the phase contrast of that nuclear region but not in the surrounding regions, creating an asymmetry. The attribute "phase skewness" measures the lack of symmetry for the phase histogram of the cell and would therefore increase. A similar explanation can be derived for baculovirus, which also assembles in the nucleus [39]. Moreover, infection at low baculovirus multiplicity of infection (MOI) yields a first round of baculovirus release from infected cells, approximately 24 h after infection. The released baculovirus will then infect more insect cells, originating a second round of infection. In the phase skewness profiles shown in Figure 4D, all these phases can be observed, likely validating our hypothesis: first baculovirus infection cycle from 24 h (infection time) to 48 h; and second infection from 48 to 72 hours. The fact that baculovirus and AAV induce a different phase skewness profile (decrease for baculovirus and increase for AAV) may be due to their different shape (rod vs icosahedral,

respectively), and the fact that baculovirus nucleocapsid is assembled in another nucleus region, the virogenic stroma [40], among other factors.

Finally, it is important to consider the influence of biological factors such as cell passage or similar. Since we have a small dataset, we cannot be sure whether some of the model parameters are accounting with biological variability between the two runs.

Comparison of the number of parameters in the 3CV models allows to have a sense of the difficulty in measuring the AAV signals when compared to viable cell concentration and viability, which have a more "macroscopic" change. More simple models (with 5 and 7 terms) were enough to describe viable cell concentration and viability, respectively, while for AAV, models with 10 and 12 parameters were needed (for extracellular AAV and intracellular specific AAV, respectively). This is also expected due to the complexity of measuring viral-induced cell changes, in which a combination of methods (measuring nucleus, diameter, cell intracellular complexity) is needed. Another possibility relies in the very different ranges and time profiles for viable cell concentration in the two batches, while for AAV, only one range is available. Higher range variations allow to better discriminate between significant and non-significant attributes. We expect these models to be refined with more batches, excluding parameters which are less relevant and clearly highlighting the attributes relevant for each variable. After identification of the relevant attributes for each quality parameter, it would be interesting to assess how those attributes would change for other production systems, AAV serotypes or packaged transgenes.

Other authors have monitored the insect cell-baculovirus system using real-time monitoring tools, mainly using dielectric spectroscopy [3,9,20,25,26,41]. Compared with other published reports using real-time monitoring in this system, DHM provides a simpler workflow: First, iLine F assembly in the bioreactor is straightforward and no preliminary calibrations are needed; data analysis is in real-time (every 30 min) and immediate (no preprocessing needed) and, in *OsOne*, there is a beta-version algorithm to estimate the percentage of baculovirus-infected cells, which we tried for the infected batch (Figure S2). Further optimization of this algorithm could be helpful to monitor the baculovirus replication kinetics and optimize the production conditions, such as the overall multiplicity of infection to use, and contribute to understanding how this parameter correlates with infection progression. Moreover, the attribute stepwise selection coupled with the multiple linear regression methodology presented in this work has the advantage of generating more interpretable models, when compared with partial least squares (PLS) or other projection-based methods: multiple linear regression models are easier to interpret regarding the biological meaning of each parameter, enabling process understanding under the PAT initiative. This is because in multiple linear regression the coefficients of the parameters itself are analyzed, differing from PLS in which the focus is on the principal components, which are linear combinations of several parameters.

In future experiments using this modeling approach, more "perturbation" batches will be useful to determine an AAV-related "label-free dynamic phenotype" [10], identifying the attributes related with AAV production and gaining insights into their biological meaning. Batches that would strengthen the viable cell concentration model calculations include more "growth only" runs, at different cell seeding densities. For AAV models, examples include runs allowing to decouple AAV production signals from other signals which may be correlated with viable cell concentration or baculovirus production. For instance, infection with empty baculovirus (a baculovirus vector which is devoid of any transgene, but still can infect and replicate in insect cells, and thus generate the normal cytopathic effects expected in this system) or only with the rep-encoding baculovirus. Infection with only the cap-encoding baculovirus would possibly be useful for finding attributes associated with empty or full AAV capsid formation, which, together with the infectivity profile, is one of the most important quality attributes for AAV vectors [31]. Regarding the full to empty ratio, runs using other AAV production systems can also be performed, particularly using systems known by their high full particle ratio, as is the case of the herpes simplex production system [42]. Exploring the application of DHM to other AAV-producing systems, such as the HEK293 transfection system, could elucidate the differences

for AAV production in transfection and infection processes and between different producer cells. Moreover, DHM could provide further insight into the reason why suspension-based transfection is less efficient than adherent-transfection. An alternative DHM device with equivalent image processing capabilities, the QMod (also by Ovizio Imaging Systems), could be used to enable a similar approach in adherent cell culture. Finally, combining the DDHM attributes with process data (e.g., DO profiles, total oxygen flow) may further increase prediction capabilities due to the increase of complementary information available [43].

Overall, we demonstrate the suitability of this methodology and DDHM technology for monitoring two of the most important variables for AAV production using insect cells: cell concentration and viability, and with potential for the development of feeding strategies schemes for AAV production. The approach described in this work enables model interpretability, increasing process understanding and allowing to draw conclusions regarding the biological state of the cell at each infection stage. Moreover, models for determination of AAV production were developed, and correlations between DDHM attributes and AAV measurements were determined, identifying for the first-time attributes related with AAV production detectable using phase microscopy. For future work, it would be relevant to employ the same strategy for identification of the DDHM attributes relevant for prediction of AAV infectivity and full to empty ratio, in order to fully explore the potential of this method to optimize AAV titer and quality, in line with the PAT initiative.

Supplementary Materials: The following are available online at http://www.mdpi.com/2227-9717/8/4/487/s1: Table S1: "Estimates, corresponding standard error, t-ratio and logworth/*p*-value for all the models shown in Figures 1 and 2", Table S2: "RMSE for calibration and validation models, scaled for the variable range, using the 3-fold cross-validation strategy", Table S3: "Quality characteristics overview for the models presented in Figure 1 and Figure 2", Figure S1: "Overview of the cell culture evolution profile over time, as captured by OsOne software", Figure S2: "Evolution of the predicted process variables using Ovizio proprietary models", Figure S3: "Pearson correlation coefficients for all attributes", Video S1: "Representative video of the culture at 99h of culture. Shown are the 25 image frames for the bright field images", Video S2: "Representative video of the culture at 99h of culture. Shown are the 25 image frames for the phase images. The images shown in the video correspond to the same ones as for Video S1".

Abbreviations

3CV	3-fold cross-validation
AAV	Adeno-associated virus
DDHM	Differential digital holographic imaging
FDA	Food and Drug administration
GFP	Green Fluorescence Protein
ITR	Inverted Terminal Repeat
LOBO	Leave one batch out
PAT	Process analytical technology
Sf9	*Spodoptera frugiperda* cell line
VCC	Viable cell concentration

References

1. US Department of Health and Human Services, Food and Drug Administration. Guidance for Industry PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing and Quality Assurance. 2004. Available online: https://www.fda.gov/media/71012/download (accessed on 15 January 2020).
2. Pais, D.A.M.; Carrondo, M.J.T.; Alves, P.M.; Teixeira, A.P. Towards real-time monitoring of therapeutic protein quality in mammalian cell processes. *Curr. Opin. Biotechnol.* **2014**, *30*, 161–167. [CrossRef] [PubMed]
3. Petiot, E.; Ansorge, S.; Rosa-Calatrava, M.; Kamen, A. Critical phases of viral production processes monitored by capacitance. *J. Biotechnol.* **2016**, *242*, 19–29. [CrossRef] [PubMed]
4. Moore, B.; Sanford, R.; Zhang, A. Case study: The characterization and implementation of dielectric spectroscopy (biocapacitance) for process control in a commercial GMP CHO manufacturing process. *Biotechnol. Prog.* **2019**, *35*. [CrossRef] [PubMed]
5. Loutfi, H.; Pellen, F.; Le Jeune, B.; Lteif, R.; Kallassy, M.; Le Brun, G.; Abboud, M. Real-time monitoring of bacterial growth kinetics in suspensions using laser speckle imaging. *Sci. Rep.* **2020**, *10*, 1–11. [CrossRef] [PubMed]
6. Tulsyan, A.; Wang, T.; Schorner, G.; Khodabandehlou, H.; Coufal, M.; Undey, C. Automatic real-time calibration, assessment, and maintenance of generic Raman models for online monitoring of cell culture processes. *Biotechnol. Bioeng.* **2020**, *117*, 406–416. [CrossRef]
7. Santos, R.M.; Kessler, J.M.; Salou, P.; Menezes, J.C.; Peinado, A. Monitoring mAb cultivations with in-situ raman spectroscopy: The influence of spectral selectivity on calibration models and industrial use as reliable PAT tool. *Biotechnol. Prog.* **2018**, *34*, 659–670. [CrossRef]
8. Zavala-Ortiz, D.A.; Ebel, B.; Li, M.Y.; Barradas-Dermitz, D.M.; Hayward-Jones, P.M.; Aguilar-Uscanga, M.G.; Marc, A.; Guedon, E. Interest of locally weighted regression to overcome nonlinear effects during in situ NIR monitoring of CHO cell culture parameters and antibody glycosylation. *Biotechnol. Prog.* **2019**, 1–10. [CrossRef]
9. Pais, D.A.M.; Portela, R.M.C.; Carrondo, M.J.T.; Isidro, I.A.; Alves, P.M. Enabling PAT in insect cell bioprocesses: In situ monitoring of recombinant adeno-associated virus production by fluorescence spectroscopy. *Biotechnol. Bioeng.* **2019**, *116*, 2803–2814. [CrossRef]
10. Kasprowicz, R.; Suman, R.; O'Toole, P. Characterising live cell behaviour: Traditional label-free and quantitative phase imaging approaches. *Int. J. Biochem. Cell Biol.* **2017**, *84*, 89–95. [CrossRef]
11. Mann, C.J.; Yu, L.; Lo, C.-M.; Kim, M.K. High-resolution quantitative phase-contrast microscopy by digital holography. *Opt. Express* **2005**, *13*, 8693–8698. [CrossRef]
12. Janicke, B.; Kårsnäs, A.; Egelberg, P.; Alm, K. Label-free high temporal resolution assessment of cell proliferation using digital holographic microscopy. *Cytom. Part A* **2017**, *91*, 460–469. [CrossRef] [PubMed]
13. Kamlund, S. Not all Those Who Wander Are Lost: A Study of Cancer Cells by Digital Holographic Imaging, Fluorescence and a Combination Thereof. Ph.D. Thesis, Lund University, Lund, Sweden, 2018.
14. Rapoport, D.H.; Becker, T.; Mamlouk, A.M.; Schicktanz, S.; Kruse, C. A novel validation algorithm allows for automated cell tracking and the extraction of biologically meaningful parameters. *PLoS ONE* **2011**, *6*. [CrossRef]
15. Kemper, B.; Bauwens, A.; Vollmer, A.; Ketelhut, S.; Langehanenberg, P.; Müthing, J.; Karch, H.; von Bally, G. Label-free quantitative cell division monitoring of endothelial cells by digital holographic microscopy. *J. Biomed. Opt.* **2010**, *15*, 036009. [CrossRef] [PubMed]
16. Kühn, J.; Shaffer, E.; Mena, J.; Breton, B.; Parent, J.; Rappaz, B.; Chambon, M.; Emery, Y.; Magistretti, P.; Depeursinge, C.; et al. Label-free cytotoxicity screening assay by digital holographic microscopy. *Assay Drug Dev. Technol.* **2013**, *11*, 101–107. [CrossRef] [PubMed]
17. Ugele, M.; Weniger, M.; Leidenberger, M.; Huang, Y.; Bassler, M.; Friedrich, O.; Kappes, B.; Hayden, O.; Richter, L. Label-free, high-throughput detection of P. falciparum infection in sphered erythrocytes with digital holographic microscopy. *Lab Chip* **2018**, *18*, 1704–1712. [CrossRef] [PubMed]
18. Altschuler, S.J.; Wu, L.F. Cellular Heterogeneity: Do Differences Make a Difference? *Cell* **2010**, *141*, 559–563. [CrossRef]
19. Hidalgo, D.; Paz, E.; Palomares, L.A.; Ramírez, O.T. Real-time imaging reveals unique heterogeneous population features in insect cell cultures. *J. Biotechnol.* **2017**, *259*, 56–62. [CrossRef]

20. Negrete, A.; Esteban, G.; Kotin, R.M. Process optimization of large-scale production of recombinant adeno-associated vectors using dielectric spectroscopy. *Appl. Microbiol. Biotechnol.* **2007**, *76*, 761–772. [CrossRef]

21. Cox, M.M.J. Recombinant protein vaccines produced in insect cells. *Vaccine* **2012**, *30*, 1759–1766. [CrossRef]

22. Palomares, L.; Estrada-Mondaca, S.; Ramirez, O. Principles and Applications of the Insect Cell-Baculovirus Expression Vector System. In *Cell Culture Technology for Pharmaceutical and Cell-Based Therapies*; Ozturk, S.S., Hu, W.-S., Eds.; Taylor & Francis: Milton Park, UK, 2005; pp. 627–692. [CrossRef]

23. Bernal, V.; Carinhas, N.; Yokomizo, A.Y.; Carrondo, M.J.T.; Alves, P.M. Cell density effect in the baculovirus-insect cells system: A quantitative analysis of energetic metabolism. *Biotechnol. Bioeng.* **2009**, *104*, 162–180. [CrossRef]

24. Sequeira, D.P.; Correia, R.; Carrondo, M.J.T.; Roldão, A.; Teixeira, A.P.; Alves, P.M. Combining stable insect cell lines with baculovirus-mediated expression for multi-HA influenza VLP production. *Vaccine* **2018**, *36*, 3112–3123. [CrossRef] [PubMed]

25. Zeiser, A.; Bédard, C.; Voyer, R.; Jardin, B.; Tom, R.; Kamen, A.A. On-line monitoring of the progress of infection in Sf-9 insect cell cultures using relative permittivity measurements. *Biotechnol. Bioeng.* **1999**, *63*, 122–126. [CrossRef]

26. Zeiser, A.; Elias, C.B.; Voyer, R.; Jardin, B.; Kamen, A.A. On-line monitoring of physiological parameters of insect cell cultures during the growth and infection process. *Biotechnol. Prog.* **2000**, *16*, 803–808. [CrossRef] [PubMed]

27. Palomares, L.A.; Pedroza, J.C.; Ramírez, O.T. Cell size as a tool to predict the production of recombinant protein by the insect-cell baculovirus expression system. *Biotechnol. Lett.* **2001**, *23*, 359–364. [CrossRef]

28. Janakiraman, V.; Forrest, W.F.; Chow, B.; Seshagiri, S. A rapid method for estimation of baculovirus titer based on viable cell size. *J. Virol. Methods* **2006**, *132*, 48–58. [CrossRef]

29. Laasfeld, T.; Kopanchuk, S.; Rinken, A. Image-based cell-size estimation for baculovirus quantification. *Biotechniques* **2017**, *63*, 161–168. [CrossRef]

30. Naso, M.F.; Tomkowicz, B.; Perry, W.L.; Strohl, W.R. Adeno-Associated Virus (AAV) as a Vector for Gene Therapy. *BioDrugs* **2017**. [CrossRef]

31. Merten, O. AAV vector production: State of the art developments and remaining challenges. *Cell Gene Ther.* **2016**, 521–551. [CrossRef]

32. Smith, R.H.; Levy, J.R.; Kotin, R.M. A simplified baculovirus-AAV expression vector system coupled with one-step affinity purification yields high-titer rAAV stocks from insect cells. *Mol. Ther.* **2009**, *17*, 1888–1896. [CrossRef]

33. Gray, S.J.; Choi, V.W.; Asokan, A.; Haberman, R.A.; Thomas, J.; Samulski, R.J. Production of Recombinant Adeno-Associated Viral Vectors and Use in In Vitro and In Vivo Administration. *Curr. Protoc. Neurosci.* **2011**, *57*, 4–17. [CrossRef]

34. Feng, Y.; Mitchison, T.J.; Bender, A.; Young, D.W.; Tallarico, J.A. Multi-parameter phenotypic profiling: Using cellular effects to characterize small-molecule compounds. *Nat. Rev. Drug Discov.* **2009**, *8*, 567–578. [CrossRef] [PubMed]

35. Meghrous, J.; Aucoin, M.G.; Jacob, D.; Chahal, P.S.; Arcand, N.; Kamen, A.A. Production of Recombinant Adeno-Associated Viral Vectors Using a Baculovirus / Insect Cell Suspension Culture System: From Shake Flasks to a 20-L Bioreactor. *Biotechnol. Prog.* **2005**, 154–160. [CrossRef] [PubMed]

36. Grein, T.A.; Loewe, D.; Dieken, H.; Salzig, D.; Weidner, T.; Czermak, P. High titer oncolytic measles virus production process by integration of dielectric spectroscopy as online monitoring system. *Biotechnol. Bioeng.* **2018**, *115*, 1186–1194. [CrossRef] [PubMed]

37. Nikolay, A.; Léon, A.; Schwamborn, K.; Genzel, Y.; Reichl, U. Process intensification of EB66® cell cultivations leads to high-yield yellow fever and Zika virus production. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 8725–8737. [CrossRef] [PubMed]

38. Bennett, A.; Mietzsch, M.; Agbandje-mckenna, M. Understanding capsid assembly and genome packaging for adeno-associated viruses. *Future Virol.* **2017**. [CrossRef]

39. Ohkawa, T.; Volkman, L.E.; Welch, M.D. Actin-based motility drives baculovirus transit to the nucleus and cell surface. *J. Cell Biol.* **2010**, *190*, 187–195. [CrossRef] [PubMed]

40. Zhao, S.; He, G.; Yang, Y.; Liang, C. Nucleocapsid assembly of baculoviruses. *Viruses* **2019**, *11*, 595. [CrossRef]

41. Mena, J.; Aucoin, M.; Montes, J.; Chahal, P.; Kamen, A. Improving adeno-associated vector yield in high density insect cell cultures. *J. Gene Med.* **2010**, 157–167. [CrossRef]

42. Merten, O.; Gaillet, B. Viral vectors for gene therapy and gene modification approaches. *Biochem. Eng. J.* **2016**, *108*, 98–115. [CrossRef]

43. Bayer, B.; von Stosch, M.; Melcher, M.; Duerkop, M.; Striedner, G. Soft sensor based on 2D-fluorescence and process data enabling real-time estimation of biomass in Escherichia coli cultivations. *Eng. Life Sci.* **2019**. [CrossRef]