

# Single-cell transcriptome analyses reveal heterogeneity in suspension cultures and clonal markers of CHO-K1 cells

Norichika Ogata<sup>1</sup> and Takeshi Omasa<sup>2</sup>

<sup>1</sup>Nihon BioData Corporation

<sup>2</sup>Osaka University

July 16, 2020

## Abstract

Cell-to-cell variability in cell populations arises from a combination of intrinsic factors and extrinsic factors related to the milieu. However, the heterogeneity of high cell density suspension cultures for therapeutic protein production remains unknown. Here, we illustrate the increasing heterogeneity in the cellular transcriptome of serum-free adapted CHO K1 cells during high cell density suspension culture over time without concomitant changes in the genomic sequence. Cell cycle-dependent subpopulations and cell clusters, which typically appear in other single-cell transcriptome analyses of adherent CHO K1 cultures, were not found in these suspension cultures. Our results indicate that cell division changes the intracellular microenvironment and leads to cell cycle-dependent heterogeneity. Whole mitochondrial single-cell genome sequencing showed cell-to-cell mitochondrial genome variation and heteroplasmy within cells. Indeed, the mitochondrial genome sequencing method developed here enables the validation of cell clonality. The culture time-dependent increase in cellular heterogeneity observed in this study did not show any attenuation in this increasing heterogeneity. Future advances in bioengineering such as culture upscaling, prolonged culturing, and complex culture systems will be confronted with the need to assess and control cellular heterogeneity, and the method described here may prove useful for this purpose.

## Keywords

Chinese hamster ovary (CHO) cells, single-cell transcriptome sequencing, heterogeneity, heteroplasmy, monoclonality

## 1. Introduction

Chinese hamster ovary (CHO) cells (Kao & Puck, 1968) play a key role as heterologous hosts for the commercial-scale production of protein pharmaceuticals (Greber & Fussenegger, 2007; Hacker, De Jesus, & Wurm, 2009; Omasa, Onitsuka, & Kim, 2010). The adaptivity of CHO cells to high cell density processes has ensured their successful application for this purpose. In the biopharmaceutical industry, a large number of products have been introduced into the market using high cell density processes such as batch and fed-batch cultures and different cell culture strategies for reducing the production cost and improving product quality (Pollock, Ho, & Farid, 2013; Tapia, Vazquez-Ramirez, Genzel, & Reichl, 2016). Cell culture times are generally long (12 days for fed-batch cultures) and advances are further prolonging culture times; perfusion continuous bioprocessing (60 days) is widely recognized as the next-generation biomanufacturing platform for achieving superior product quality and reducing manufacturing costs (Konstantinov & Cooney, 2015; Zydney, 2015).

A potential issue when using long-term cultures for therapeutic protein production is the increasing cellular heterogeneity over time (Altschuler & Wu, 2010). Cellular proliferation and gene expression are critical cell processes but genomic mutations arise during DNA replication (Furusawa, 2014) and RNA transcription (Bachl, Carlson, Gray-Schopfer, Dessing, & Olsson, 2001) inherent to cell division and gene expression,

respectively. The limited knowledge regarding cellular heterogeneity in high cell density processes poses a barrier to solving this issue.

It has been proposed that insisting on “true” clonality is scientifically unjustifiable (O’Callaghan et al., 2015; Wurm, 2013), with cell line heterogeneity being included in the FDA “Challenges to Potency Assay Development for CGT products” industry guidelines (Guidance for Industry: Potency Tests for Cellular and Gene Therapy Products) (FDA, 2011). We can only hope that the trend toward a modified gene pool can be minimized by maintaining cell populations in stable environments (Wurm, 2013). Deeper characterization such as single-cell sequencing (Tang et al., 2009) of the heterogeneity within cell line cultures would enable the study of cooperative and competitive interactions between cell populations and the mechanisms underlying related cell functions (Ben-David et al., 2018).

In this study, we performed single-cell transcriptome sequencing of serum-free suspension-adapted CHO K1 cells in shake flasks. The degree of cellular heterogeneity and the existence of subpopulations were determined based on gene expression analysis and whole mitochondrial genome sequencing. To facilitate comparisons of our data to those from other single-cell transcriptome studies, we also performed single-cell transcriptome sequencing of adherent CHO K1 cells in culture dishes. Additionally, we developed clonal markers to validate the clonality of the cells based on their mitochondrial genome sequences.

## 2. Materials and Methods

### 2.1. Cell Lines and Culture Conditions

CHO K1 cells (Kao & Puck, 1968) (ATCC, Manassas, VA, USA) and CHO K1-derived cells adapted to serum-free suspension culture were used in this study. The CHO K1 cells were maintained in Dulbecco’s Modified Eagle’s Medium: Nutrient Mixture F-12 (DMEM/F-12; Gibco, Paisley, UK) supplemented with 10% dialyzed fetal bovine serum (SAFC Biosciences, Lenexa, KS, USA). The CHO K1-derived, serum-free suspension-adapted cells were cultured in suspension using T-25 flasks (Greiner, Nürtingen, Germany) on a rotary shaker at 120 rpm (Taitec, Saitama, Japan). A total of  $0.1 \times 10^6$ /mL CHO cells were cultivated in 7 mL BalanCD CHO Growth A media (Irvine Scientific, Santa Ana, CA, USA). The suspension culture conditions were assessed using computational fluid and particle dynamics analyses with OpenFOAM version 2.3.0. (H. G. Weller, 1998). We modified the stirringInterPTFoam application to compute the culture conditions in box-shaped T-25 flasks on a rotary shaker (<https://github.com/akionux/OpenFOAM-2.3.x/tree/master/applications/solvers/multiphase/stirringInterFoam/stirringInterPTFoam>). All cells were cultivated at 37°C in a humidified atmosphere containing 5% CO<sub>2</sub>.

### 2.2. Library Preparation and Sequencing (Serum-free Suspension Cultures)

After days 4 (D4), 8 (D8), and 11 (D11), the CHO cells in three replicate serum-free suspension culture flasks were harvested for single-cell sequencing, and the remainder was discarded. Samples of the CHO cells in suspension culture were directly applied to the SH800 Cell Sorter (Sony, Tokyo, Japan), and individual cells were collected directly into PCR tubes as previously reported (Hayashi et al., 2010). A total of  $0.5 \times 10^6$  cells were stained with 1 mL phosphate-buffered saline (PBS) containing 10 µg/mL Hoechst 33342 at 37°C for 15 min. The stained cells were sorted based on the Hoechst 33342-stained cell area of the fluorescence-activated cell sorting (FACS) distribution (Hayashi et al., 2010). Library preparation for single-cell Quartz-Seq was performed as previously reported (Sasagawa et al., 2013). We prepared a library for conventional RNA-seq using a commercial kit (TruSeq RNA Library Prep Kit; Illumina, San Diego, CA, USA) in accordance with a previous study (Yamano-Adachi et al., 2019). Conventional RNA-seq and single-cell Quartz-Seq were performed with a commercial sequencer (NextSeq 500; Illumina) in accordance with the manufacturer’s protocols. We performed single-cell Quartz-Seq as previously reported (Sasagawa et al., 2013) on 85 D4, 70 D8, and 78 D11 CHO cells. The resulting short-read data have been deposited in the Short Read Archive of the DNA Data Bank of Japan (DDBJ) under project ID DRA004159.

### 2.3. Library Preparation and Sequencing (Adherent Cultures)

Adherent CHO cell cultures in the log phase of growth were harvested for single-cell sequencing. A total

of  $2.5 \times 10^6$  cells were stained with 500  $\mu$ L DMEM/F-12 containing L-glutamine (Gibco; Thermo Fisher Scientific, Waltham, MA, USA) and 10  $\mu$ g/mL Hoechst 33342 (Funakoshi, Tokyo, Japan) by stirring at 600 rpm for 40 min at 37°C. Cells were centrifuged for 5 min at  $300 \times g$  and the supernatant was removed. Next, 500  $\mu$ L of  $1 \times$  PBS (no calcium, no magnesium, pH 7.2, Gibco; Thermo Fisher Scientific) was added to the cell pellet, which was pipetted up and down five times for thorough resuspension. Cells were centrifuged for 5 min at  $300 \times g$ , the supernatant was removed, and the cells were thoroughly resuspended in 500  $\mu$ L of 30% (w/v) Albumin D-PBS(-) Solution (fatty acid-free from bovine serum albumin; FujiFilm Wako Pure Chemical Corporation, Osaka, Japan) by pipetting the cell pellet up and down five times. Cells in the albumin-containing PBS were filtered with a 35- $\mu$ m nylon mesh filter (Falcon, Corning, NY, USA). Cells were applied to the SH800S Cell Sorter and individual cells were sorted based on the Hoechst 33342-stained cell area of the FACS distribution. RNA isolation and complementary DNA (cDNA) synthesis were performed with the SMART-Seq v4 Ultra Low Input RNA Kit (Clontech, Palo Alto, CA, USA) for sequencing according to the manufacturer's protocol (Ziegenhain et al., 2017). Library preparation was performed with the Nextera XT DNA Library Prep Kit (Illumina) according to the manufacturer's protocol. The finished cDNA libraries were quantified with the Agilent 2200 TapeStation system (Agilent Technologies, Santa Clara, CA, USA) and the QuantiFluor dsDNA System (Promega, Fitchburg, WI, USA) and then sequenced on an Illumina NextSeq 500 platform using 75-bp single-end reads. We performed single-cell RNA-seq analysis with 24 G1 cells and 8 G2 cells.

#### 2.4. Gene Expression Heterogeneity Analyses

All short reads were mapped to the CHO K1 RefSeq assembly (22,516 sequences, RefSeq Assembly ID: GCF\_000223135.1) and CHO K1 mitochondrial DNA (1 sequence, RefSeq Assembly ID: GCF\_000055695.1) using Bowtie2 v. 2.3.0 (Langmead, Wilks, Antonescu, & Charles, 2018). Estimating gene and isoform expression levels from the RNA-Seq data was performed with RSEM v. 1.2.31 (B. Li & Dewey, 2011). Principal component analyses (PCA) of the single-cell RNA-seq data were performed with R v. 3.0.1 and the R Stats Package v. 3.2.2 (RCoreTeam, 2013). The Bartlett test for the homogeneity of variances in principal components 1 between different culture times (Days) was performed with the R Stats Package v. 3.2.2.

#### 2.5. Genome Mutation Heterogeneity Analyses

All short reads were mapped to the *Cricetulus griseus* mitochondrion complete genome (gi|186694636|gb|EU660217.1|) using HISAT2 v. 2.0.5 (Kim, Langmead, & Salzberg, 2015). PCR duplicate removal, and SNP (single nucleotide polymorphism) and short indel (insertion or deletion) calling were performed with the SAMtools v. 1.8.0 utilities, rmdup and mpileup, respectively (H. Li et al., 2009). The called variant loci were annotated as "heteroplasmy," "homo alternative," "homo reference," or "low coverage." A correlation test between the culture time (Days) and the hetero/homo ratio was performed with the R Stats Package v. 3.2.2.

### 3. Results and Discussion

#### 3.1. Time Course Transcriptome Heterogeneity in Suspension Cultures

We performed single-cell Quartz-Seq as previously reported (Sasagawa et al., 2013) of 85 D4, 70 D8, and 78 D11 CHO cells. To compare the transcriptomes of the cells cultured for different lengths of time (days), we performed PCA of all the expression data (Figure 1). The proportion of variance in principal component 1 (PC1) was 0.3102, while that of PC2 was 0.077. Using this approach, we could quantitatively evaluate the cellular heterogeneity, which previously was not possible. The standard deviations in PC1 were 1634 for D4, 3024 for D8, and 5029 for D11. We rejected the null hypothesis that heterogeneity did not increase during culture, based on the Bartlett test for the homogeneity of variances in the PC1 between different culture times ( $p < 2.2 \times 10^{-16}$ ). Thus, the cellular heterogeneity increased during cultivation and this increase showed no sign of attenuation. Hence, the maximum attainable heterogeneity of the CHO cells in a long-term high cell density culture was not predictable from these data. To assess cellular heterogeneity in future bioengineering processes, very long-term cultures such as perfusion continuous cultures need to be studied. We can control

it if it can be measured.

### 3.3. Genomic Mutation Analyses of Single Cells

To compare the mutation rate in cells cultured for different lengths of time, we performed mitochondrial genomic sequence analyses. Recently, many single-cell genomics technologies have been developed. However, it is difficult to obtain genomic/epigenomic information from limited numbers of cells because of sample loss during chromatin preparation and inefficient immunoprecipitation (Harada et al., 2018). Each cell has only a single nuclear genome but several mitochondrial genomes. Using our single-cell transcriptome sequence data, we obtained sufficient coverage of the mitochondrial genome to call SNPs and short indels, with 498 genomic loci identified with variation at single-cell resolution. We annotated variations as “heteroplasmy,” “homo alternative,” “homo reference,” or “low coverage” (Figure 2). Homo alternative and homo reference variations were shared between most cells. However, heteroplasmic SNPs were unique to individual cells. Stochastically, recently acquired mutations do not show a high frequency in the gene pool. The mutations that arose during our experiment should have appeared as new hetero SNPs; indeed, we could not reject the null hypothesis that the number of novel mutations did not increase during culture, based on a correlation test between the culture time and the hetero SNP/homo SNP ratio in each cell ( $p = 0.14$ ; Figure 3). This indicated that there was a low genomic mutation rate during cultivation and that the increasing heterogeneity observed in genome-wide gene expression was independent of genomic mutations.

We would like to emphasize the usefulness of the identified heteroplasmic mitochondrial mutations in the context of single-cell cloning. Regulatory authorities require that cell lines used in the commercial production of recombinant proteins must be derived from a single cell progenitor or clone (Evans et al., 2015). To provide evidence of monoclonality, multiple rounds of the limiting dilution method or cell sorting coupled with high-resolution cell imaging are necessary. However, our study indicates that the unique heteroplasmic mitochondrial mutations of individual cultured cells could potentially provide evidence of monoclonality. In other words, evidence of monoclonality could be obtained by single-cell sequencing of the cell strain in question.

### 3.4. Transcriptome Heterogeneity Analyses of Adherent Cultures

In our single-cell transcriptome analysis of suspension cultures, no subpopulations were apparent. This finding differs from that of previous studies, which indicated a strong correlation between genome-wide gene expression and the cell cycle stage of adherent cells (Buettner et al., 2015; Santos, Wernersson, & Jensen, 2015; Sasagawa et al., 2013; Scialdone et al., 2015). Cellular heterogeneity is known to be influenced by the cellular microenvironment in adherent, two-dimensional cultures (Snijder et al., 2009). It is possible that cell cycle-dependent genome-wide gene expression is an adherent culture-specific phenomenon. In this context, the origin of cell cycle-dependent cellular heterogeneity in adherent cultures would not be the product of intrinsic factors but rather the product of the extrinsic milieu, with cellular division altering genome-wide gene expression in cells by changing the microenvironment in adherent cultures. A previous study supports this interpretation; suspended T cells did not show any cell cycle-dependent cellular transcriptome heterogeneity (Mahata et al., 2014).

To further investigate cell cycle-dependent cellular transcriptome heterogeneity, we performed single-cell SMART-Seq as previously reported (Ziegenhain et al., 2017) on 24 G1 cells and 8 G2 cells from adherent cultures. To compare the transcriptomes between cells in different phases of the cell cycle, we performed PCA of all the expression data (Figure 4). The proportion of variance of PC1 was 0.86, while that of PC2 was 0.033. The G2 cells formed a distinct cluster. We, therefore, rejected the null hypothesis that cell cycle-dependent subpopulations do not appear during adherent culture, based on the Bartlett test for the homogeneity of variances in PC1 between G1 and G2 cells ( $p = 0.000245$ ). Thus, cell cycle-dependent heterogeneity in genome-wide gene expression was found to be a phenomenon specific to adherent cultures of CHO cells.

### 3.5. Subpopulation Analyses of Suspension Cultures

In suspension cultures, there were no clear subpopulations of cells based on PCA. However, there remained the possibility of hidden subpopulations of cells. We generated histograms of gene expression and found a gene, enolase, that had a multipeak distribution. We divided the individual D4 cells into groups with high enolase expression (giI625221705IrefIXM\_003512016.2I; 2000–4200 RPM) and low enolase expression (0–2000 RPM; Figure 5a) and used PCA to analyze these individual D4 cells. The PC1 of the individual D4 cells were plotted against those for PC2. Individual D4 cells with high and low enolase expression formed distinct clusters (Figure 5b), which represents the hidden subpopulations of CHO cells in suspension cultures. This phenomenon is contrary to expectations. Hidden cellular subpopulations such as those found in this study hint at the future discovery of intrinsic factors giving rise to cell-to-cell variability in cell populations.

#### 4. Conclusions

In summary, we sequenced the single-cell transcriptomes of high cell density suspension mammalian cell cultures. Our findings indicated that cellular transcriptome heterogeneity increased with increasing culture time, while genomic mutations did not increase during suspension culture. We also established a method for the quantitative evaluation of cellular heterogeneity, which was previously not possible. This method entails examining the evidence of monoclonality by single-cell sequencing of the mitochondria of cell strains. Further, we uncovered several differences between adherent cells and suspension cells based on single-cell genomics. Fundamental biological observation and assessment are of incremental importance in successful bioengineering. Instead, further single-cell resolution observations (Norichika Ogata, 2018) and improved measurement technologies are needed. The intercellular variations in the mitochondrial genome sequences discovered in this study are also useful for generating clonal markers for future applications.

#### Acknowledgments

This work was funded partly by the Ministry of Economy, Trade, and Industry (METI) of Japan, and the Japan Agency for Medical Research and Development (AMED) for “Developing key technology for discovering and manufacturing pharmaceuticals used for next-generation treatments and diagnoses” (JP17ae0101003 and JP18ae0101054). We thank Dr. Itoshi Nikaido and Dr. Hiroki Danno for performing Quartz-Seq library preparation. We thank Natasha Beeton-Kempen, Ph.D., from Edanz Group (<https://en-author-services.edanzgroup.com/>) for editing a draft of this manuscript.

#### Conflict of Interest

The authors declare no conflict of interest.

#### Data accessibility

The resulting short-read data have been deposited in the Short Read Archive of the DNA Data Bank of Japan (DDBJ) under project ID DRA004159.

#### References

- Altschuler, S. J., & Wu, L. F. (2010). Cellular heterogeneity: do differences make a difference? *Cell*, 141 (4), 559-563. doi:10.1016/j.cell.2010.04.033
- Bachl, J., Carlson, C., Gray-Schopfer, V., Dessing, M., & Olsson, C. (2001). Increased transcription levels induce higher mutation rates in a hypermutating cell line. *J Immunol*, 166 (8), 5051-5057.
- Ben-David, U., Siranosian, B., Ha, G., Tang, H., Oren, Y., Hinohara, K., . . . Golub, T. R. (2018). Genetic and transcriptional evolution alters cancer cell line drug response. *Nature*, 560 (7718), 325-330. doi:10.1038/s41586-018-0409-3
- Buettner, F., Natarajan, K. N., Casale, F. P., Proserpio, V., Scialdone, A., Theis, F. J., . . . Stegle, O. (2015). Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nature biotechnology*, 33 (2), 155-160. doi:10.1038/nbt.3102

- Evans, K., Albanetti, T., Venkat, R., Schoner, R., Savery, J., Miro-Quesada, G., . . . Groves, C. (2015). Assurance of monoclonality in one round of cloning through cell sorting for single cell deposition coupled with high resolution cell imaging. *Biotechnology progress*, *31* (5), 1172-1178. doi:10.1002/btpr.2145
- Furusawa, M. (2014). The disparity mutagenesis model predicts rescue of living things from catastrophic errors. *Frontiers in genetics*, *5* , 421. doi:10.3389/fgene.2014.00421
- Greber, D., & Fussenegger, M. (2007). Mammalian synthetic biology: engineering of sophisticated gene networks. *Journal of biotechnology*, *130* (4), 329-345. doi:10.1016/j.jbiotec.2007.05.014
- H. G. Weller, G. T., H. Jasak, C. Fureby. (1998). A tensorial approach to computational continuum mechanics using object-oriented techniques. *COMPUTERS IN PHYSICS*, *12* , 620-631. Retrieved from [www.OpenFOAM.org](http://www.OpenFOAM.org)
- Hacker, D. L., De Jesus, M., & Wurm, F. M. (2009). 25 years of recombinant proteins from reactor-grown cells - where do we go from here? *Biotechnol Adv*, *27* (6), 1023-1027. doi:10.1016/j.biotechadv.2009.05.008
- Harada, A., Maehara, K., Handa, T., Arimura, Y., Nogami, J., Hayashi-Takanaka, Y., . . . Ohkawa, Y. (2018). A chromatin integration labelling method enables epigenomic profiling with lower input. *Nat Cell Biol* . doi:10.1038/s41556-018-0248-3
- Hayashi, T., Shibata, N., Okumura, R., Kudome, T., Nishimura, O., Tarui, H., & Agata, K. (2010). Single-cell gene profiling of planarian stem cells using fluorescent activated cell sorting and its "index sorting" function for stem cell research. *Development, growth & differentiation*, *52* (1), 131-144. doi:10.1111/j.1440-169X.2009.01157.x
- Kao, F. T., & Puck, T. T. (1968). Genetics of somatic mammalian cells, VII. Induction and isolation of nutritional mutants in Chinese hamster cells. *Proceedings of the National Academy of Sciences of the United States of America*, *60* (4), 1275-1281.
- Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nature methods*, *12* (4), 357-360. doi:10.1038/nmeth.3317
- Konstantinov, K. B., & Cooney, C. L. (2015). White Paper on Continuous Bioprocessing May 20&#x2013;21 2014 Continuous Manufacturing Symposium. *Journal of Pharmaceutical Sciences*, *104* (3), 813-820. doi:10.1002/jps.24268
- Langmead, B., Wilks, C., Antonescu, V., & Charles, R. (2018). Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* . doi:10.1093/bioinformatics/bty648
- Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics*, *12* , 323. doi:10.1186/1471-2105-12-323
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25* (16), 2078-2079. doi:10.1093/bioinformatics/btp352
- Mahata, B., Zhang, X., Kolodziejczyk, A. A., Proserpio, V., Haim-Vilmovsky, L., Taylor, A. E., . . . Teichmann, S. A. (2014). Single-cell RNA sequencing reveals T helper cells synthesizing steroids de novo to contribute to immune homeostasis. *Cell reports*, *7* (4), 1130-1142. doi:10.1016/j.celrep.2014.04.011
- Norichika Ogata, A. S., Takayuki Komiya, Yoji Iizuka, Ken Matsuse, Fuminobu Imaizumi, Tomoyuki Suwa, Akinobu Teramoto. (2018). *An Electrical Impedance Biosensor Array for Tracking Moving Cells*. Paper presented at the IEEE Sensors, New Dehli.
- O'Callaghan, P. M., Berthelot, M. E., Young, R. J., Graham, J. W., Racher, A. J., & Aldana, D. (2015). Diversity in host clone performance within a Chinese hamster ovary cell line. *Biotechnology progress* .

doi:10.1002/btpr.2097

Omasa, T., Onitsuka, M., & Kim, W. D. (2010). Cell engineering and cultivation of chinese hamster ovary (CHO) cells. *Curr Pharm Biotechnol*, 11 (3), 233-240.

Pollock, J., Ho, S. V., & Farid, S. S. (2013). Fed-batch and perfusion culture processes: economic, environmental, and operational feasibility under uncertainty. *Biotechnology and bioengineering*, 110 (1), 206-219. doi:10.1002/bit.24608

RCoreTeam. (2013). R: A Language and Environment for Statistical Computing. Retrieved from <http://www.R-project.org/>

Santos, A., Wernersson, R., & Jensen, L. J. (2015). Cyclebase 3.0: a multi-organism database on cell-cycle regulation and phenotypes. *Nucleic acids research*, 43 (Database issue), D1140-1144. doi:10.1093/nar/gku1092

Sasagawa, Y., Nikaido, I., Hayashi, T., Danno, H., Uno, K. D., Imai, T., & Ueda, H. R. (2013). Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. *Genome biology*, 14 (4), R31. doi:10.1186/gb-2013-14-4-r31

Scialdone, A., Natarajan, K. N., Saraiva, L. R., Proserpio, V., Teichmann, S. A., Stegle, O., . . . Buettner, F. (2015). Computational assignment of cell-cycle stage from single-cell transcriptome data. *Methods*, 85 , 54-61. doi:10.1016/j.jymeth.2015.06.021

Snijder, B., Sacher, R., Ramo, P., Damm, E. M., Liberali, P., & Pelkmans, L. (2009). Population context determines cell-to-cell variability in endocytosis and virus infection. *Nature*, 461 (7263), 520-523. doi:10.1038/nature08282

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., . . . Surani, M. A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nature methods*, 6 (5), 377-382. doi:10.1038/nmeth.1315

Tapia, F., Vazquez-Ramirez, D., Genzel, Y., & Reichl, U. (2016). Bioreactors for high cell density and continuous multi-stage cultivations: options for process intensification in cell culture-based viral vaccine production. *Appl Microbiol Biotechnol*, 100 (5), 2121-2132. doi:10.1007/s00253-015-7267-9

Wurm, F. M. (2013). CHO Quasispecies—Implications for Manufacturing Processes. *Processes*, 1 (3), 296-311. doi:10.3390/pr1030296

Ziegenhain, C., Vieth, B., Parekh, S., Reinis, B., Guillaumet-Adkins, A., Smets, M., . . . Enard, W. (2017). Comparative Analysis of Single-Cell RNA Sequencing Methods. *Mol Cell*, 65 (4), 631-643 e634. doi:10.1016/j.molcel.2017.01.023

Zydney, A. L. (2015). Perspectives on integrated continuous bioprocessing—opportunities and challenges. *Current Opinion in Chemical Engineering*, 10 , 8-13. doi:<https://doi.org/10.1016/j.coche.2015.07.005>

## Figure legends

**Figure 1.** Increased single-cell transcriptome heterogeneity of suspension-cultured CHO cells over time. Single-cell transcriptome sequencing was performed on 85, 70, and 78 cells from cultures maintained for 4, 8, and 11 days, respectively. Principal components 1 (proportion of variance, 0.31) and 2 (proportion of variance, 0.077) are plotted. Blue, black, and red plots represent individual day 4, day 8, and day 11 cells, respectively. Variances in principal component 1 between the different culture times had a p-value < 2.2e-16.

**Figure 2.** Single-cell resolution mitochondrial genomic heterogeneity in suspension-cultured CHO cells. Genotypes of 498 variable loci in each cell were visualized.

**Figure 3.** Correlation between mutation rate and culture time. Single-cell resolution genomic mutation analysis was performed on the mitochondrial genomes of CHO cells cultured for different periods. No

statistical correlation was observed between culture time (Days) and the heteroplasmic variant rate (rate of new mutations at called variable loci;  $p = 0.14$ ).

**Figure 4.** Cell cycle-dependent subpopulations in traditional adherent CHO cultures. Single-cell transcriptome sequencing was performed on 24 G1 cells (circles) and 8 G2 cells (crosses).

**Figure 5.** Hidden subpopulations in suspension cultures of CHO K1 cells. Individual D4 cells clustered into groups with high enolase expression (2000–4200 RPM, red) and low enolase expression (0–2000 RPM, black). (a) Histogram of enolase expression degree (RPM) in suspension CHO K1 cells after 4 days. (b) A colored plot of principal components 1 and 2 in Day 4 suspension cultures of CHO K1 cells.







