Influence of Epigenetics on Bioprocessing

Mark A. Blenner
Clemson Univ.

Harnessing and applying epigenetics in bioprocessing systems may improve cell line development, productivity, and product quality.

Epigenetic modifications, such as DNA methylation, chromatin modification through histone post-translational modification, and RNA interference, can have a great impact on industrial mammalian cell cultures.

For example, epigenetic factors influence several aspects of Chinese hamster ovary (CHO) strain development and scaleup. CHO cell cultures are used to produce over 70% of biopharmaceuticals, including monoclonal antibodies and other therapeutics. In CHO cell lines, integration site(s) of a monoclonal antibody gene and the method used to select for monoclonal antibody gene integration both affect long-term productivity loss during serial passages. Epigenetics is also a factor in controlling the activation of secondary metabolite products in fungal and yeast cells.

This article highlights the major modes of epigenetic-based regulation as they relate to mammalian cell culture instability and the influence of epigenetics on bioproduction in fungal and yeast systems. It also covers recent advances in engineering that enable control of the epigenetic state at a particular genomic location. It envisions how these nascent technologies may be used to improve host strain development, productivity, and product quality — focusing on the specific scientific questions that are still to be answered.

Mammalian cell culture instability

Industrial CHO cell lines are made by transfecting genes that code for therapeutic proteins (*e.g.,* light and heavy chains for monoclonal antibodies) under the control of strong promoters (*e.g.,* cytomegalovirus [CMV], elongation factor), which are then screened for high-producing phenotypes. CHO cells exhibit significant heterogeneity, which arises from genetic plasticity (*i.e.,* the ability of one genotype to produce more than one phenotype) that is typical of immortalized cell cultures and cancer cells. This heterogeneity is leveraged to select for highly productive clones, but eliminates some favorable production phenotypes during scaleup and production. Note that cloning does not reduce chromosomal heterogeneity or mutation rates.

Epigenetic marks, such as histone modifications and DNA methylation, influence cells in culture in the short term and long term (Figure 1). Differences in culture conditions, such as high product and metabolic waste concentration, reactor operation mode (*e.g.,* adherent vs. suspension, batch vs. fed batch), and media composition (*e.g.,* serum-containing vs. serum-free, supplemented), cause transcriptional changes that are associated with genome-wide epigenetic changes.

Most studies of CHO cell instability have focused on reduction in the rate of product formation per cell (q_p) over time, which manifests as a loss of mRNA due to one or a combination of several epigenetic mechanisms, including transgene genomic location, gene copy number loss, repeat-induced transcriptional silencing, and promoter methylation. The impact of these mechanisms on CHO, yeast, and fungal cell line development are covered in this article.

Introduction of a transgene into the CHO cell genome both perturbs the local chromatin structure and is influenced by the local chromatin structure. Epigenetic factors associated with the transgene can be leveraged to create and maintain a favorable chromatin structure. However, the

genetic instability of the CHO cell line hinders the reliable use of transgenes.

An alternative and complementary approach is to discover regions of the genome that are favorable for transgene expression, *i.e.,* hot spots. The genomic information necessary to identify hot spots only became available in 2011. Hot spots can be identified from stably selected clones and targeted using zinc-finger (ZF) nucleases, transcription activator-like effector nucleases (TALENs), or clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9.

Promoter methylation, which causes gene silencing and loss of gene copy number, has been examined as a cause of reduction in cell line productivity. Studies used the CMV promoter, which is most common in CHO cell transgene expression. The methylated promoter reduced productivity without a change in copy number *(1, 2)*. A subsequent study across ten cell lines that expressed different proteins found that while promoter methylation was a factor in productivity loss, gene copy number loss was more common *(3)*. An important difference between these studies is the selection system used (*i.e.,* dihydrofolate reductase [DHFR] vs. glutamine synthase selection), which indicates that the particular cell line, protein, and selection system may influence the mode of epigenetic silencing. Furthermore, epigenetic silencing of the CMV promoter can occur prior to methylation through histone modifications *(4)*, suggesting methylation may be a consequence rather than a cause of silencing.

Fungal and yeast cell cultures

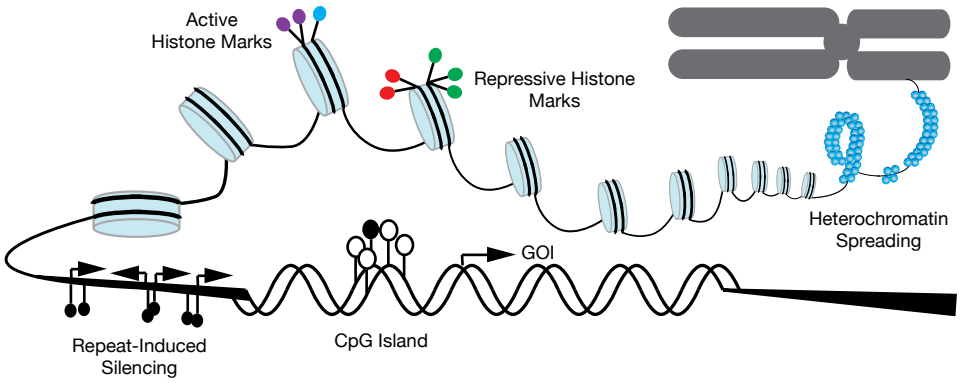
CHO cells dominate the production of monoclonal antibodies and therapeutics; however, yeast and fungi can also be used to produce certain therapeutics, as well as industrial enzymes and secondary metabolites. Eukaryotic bioprocessing systems include yeast, such as *Saccharomyces cerevisiae* and *Yarrowia lipolytica*, and filamentous fungi, such as *Aspergillus* and *Penicillium*. The effects of epigenetics on these systems have not been studied to the same degree as those on CHO cells.

However, the effect of transgene position has been studied in *S. cerevisiae (5)*. The study looked at 1,044 loci scattered across the genome, and found lack of high expression sites in the telomeric and centromeric regions. Sixteen sites tested by Wu *et al.* were among the top 88 sites of genomic interactions identified in a genome-wide associated study *(6)*. All 16 sites were extremely low expression sites, suggesting chromatin contacts are overall repressive for transcription and should be avoided.

A similar study *(7)* evaluated the effect of genome integration site on protein expression strength and the cell-to-cell variation in expression level in other identical cells — known as expression noise. Chen and Zhang found that expression levels and noise were significantly influenced by location, but more strongly controlled by the promoter strength *(7)*. The regional epigenetic state was a covariate for expression strength and noise. The effect of transgene position on mRNA level was much more pronounced than protein expression level. This suggests that while controlling transgene position may lead to better control of transcription, this will not always lead to better control of protein levels. Therefore, more focus may be needed on the influence of epigenetic state on translation rates and regulation.

Epigenetic-driven noise is responsible for cell-to-cell productivity differences in bioprocessing. It has been exploited to weed out less-productive variants and improve overall productivity. Using a microfluidic time-invariant system, researchers showed that lipid productivity of engineered *Yarrowia lipolytica* cells was sporadic with time *(8)*. However, even after eliminating translation, heterogeneity persisted, implying that metabolic heterogeneity was amplified by epigenetic heterogeneity — giving rise to bioprocessing noise.

Regulation of secondary metabolites and secreted enzymes in fungal systems has been reported *(9)*. In nearly



◀ **Figure 1.** Various epigenetic processes are known to influence gene expression, degrading production productivity. These processes include heterochromatin spreading, repeat-induced gene silencing, methylation of CpG islands (white circles are methyl-free cytosine, black circles are methylated cytosine), and histone modifications. Histone modification might include loss of active histone marks — such as H3K4 methyl groups (purple circles) and H3K27 acetyl groups (blue circle). Histone modification may also entail the addition of repressive histone modifications, such as H3K27 methyl groups (green circles) and H3K79 methyl groups (red circles).

all of these studies, environmental conditions trigger the activation of genes or gene clusters. In some studies, the histone marks associated with activation of these genes have been evaluated and, in some cases, used to enhance production. For example, histone acetyltransferase (HAT) GCN5 plays a role in controlling the expression of cellulases in *Trichoderma reesei* through H3K9 and H3K14 acetylation *(10)*. Overexpression of another HAT resulted in a two-fold increase in cellulase production *(11)*.

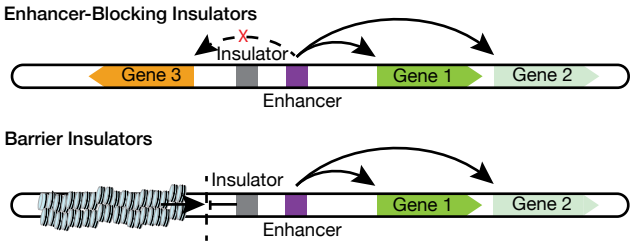
Improving strain development and productivity

Epigenetics is rarely considered as part of the design strategy of cell line engineering campaigns. This is not surprising given that we have only recently begun to understand the epigenetic basis of strain engineering and genomic stability problems. However, we now have the knowledge and tools to avoid epigenetic problems in strain engineering and, perhaps, take advantage of epigenetics to improve strain stability and productivity. This section outlines three areas that are poised to impact bioprocess strain development in the near future.

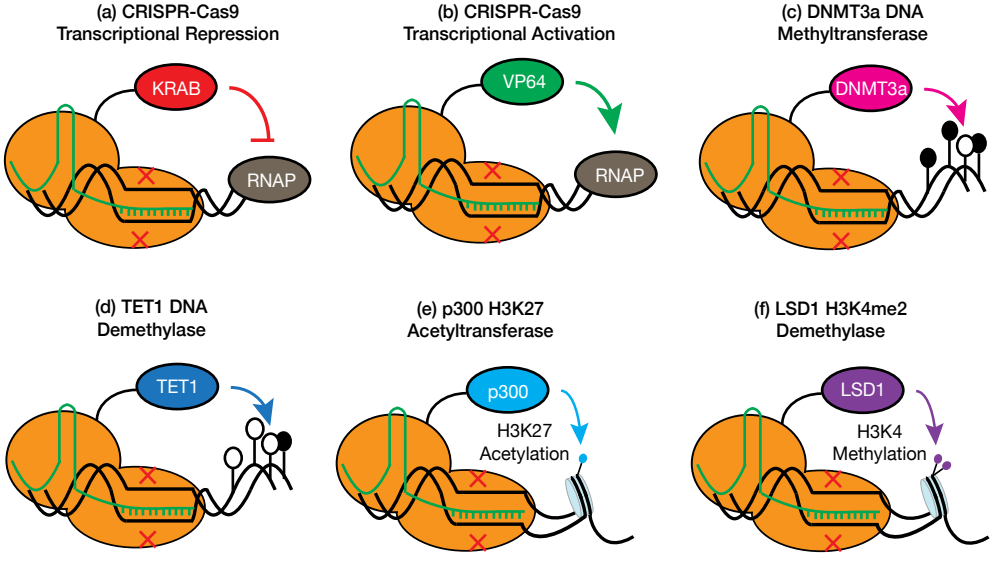
Incorporating genetic elements with favorable epigenetic properties. A genetic insulator is a DNA element that prevents promoters from being influenced by the surrounding DNA structure and epigenetic state. Two types of insulators have been characterized: enhancer-blocking insulators and barrier insulators (Figure 2). The main purpose of enhancer-blocking insulators is to keep enhancers contained in so-called "insulated neighborhoods" from affecting gene expression of surrounding genes *(12)*. These types of insulators could be useful for preventing strongly expressed transgenes from affecting neighboring gene expression. Barrier insulators prevent heterochromatin from spreading into actively transcribed regions.

In animals, enhancer-blocking insulators are believed to create compartments of chromatin loops, forming higher-ordered structures in either active or inactive states. Such looping is not known to be prominent in yeast. However, barrier insulators in yeast stop heterochromatin from spreading from centromeres and telomeres.

Researchers often use the CMV promoter for transgene expression in mammalian cells due to its high activity; however, the CMV promoter contains CpG islands (*i.e.,* short regions of DNA that contains a repeat sequence of cytosine and guanine), which can be transcriptionally silenced by DNA methylation. Hypermethylation and hypomethylation of CpG islands in gene promoters is often associated with silencing or activation of transcription in cancer cells,



▲ **Figure 2.** DNA insulators prevent epigenetic regulation of surrounding genes. Enhancer-blocking insulators prevent the action of transcriptional enhancers on spatially local genes that would otherwise be influenced by active enhancer activity. Barrier insulators prevent heterochromatin spreading into regions of active transcription, blocking associated transcriptional silencing.



▶ **Figure 3.** In addition to the use of CRISPR-Cas9 to introduce a gene into a favorable genomic location, nuclease-deficient variants (dCas9) could be used to engineer the epigenetics of a transgene to prevent loss of productivity. CRISPR-based repression (CRISPRi) and activation (CRISPRa) can be achieved by fusing the tool to chromatin modifiers such as the (a) Kruppel-associated box (KRAB) and (b) herpes simplex viral protein 64 (VP64) that recruit several epigenome-modifying enzymes. More specific epigenome modification is possible using fusions of dCas9 to epigenetic writers and erasers of DNA methylation, such as the (c) DNMT3a methyltransferase and the (d) TET1 demethylase, respectively. dCas9 can also be fused to writers and erasers of histone modifications, such as the (e) p300 acetyltransferase and the (f) LSD1 demethylase, respectively.

respectively. Some researchers have switched to promoters with fewer CpG islands, such as the elongation factor alpha promoter. Another improvement is to mutate the commonly methylated cytosine nucleotide that is located 179 base pairs upstream of the transcriptional start site. According to a study by Mariati *et al.*, inserting a DNA-methylation-preventing sequence known as a "CpG island element" into the CMV promoter improved expression stability *(13)*. Nevertheless, a CpG-free promoter did not provide long-term expression stability due to histone methylation and deacetylation.

The ubiquitous chromatin-opening element (UCOE) is perhaps the most promising genetic tool that could offer stable long-term high gene expression. UCOE sequences are found in the divergent promoter regions between two housekeeping genes, such as the HNRPA2B1-CRX3 locus (*i.e.,* A2UCOE). These promoter regions have notably methylation-free CpG islands that prevent gene silencing; however, they do not typically have high levels of transcription. When combined with the CMV promoter, the construct confers site-independent open-chromatin conformation that resists transgene silencing. The use of A2UCOE and the RPS3 UCOE both provided higher expression of a variety of transgenes, more stable integration, and faster clonal selection than the CMV promoter *(14, 15)*. The relative performance of the two UCOEs, however, was dependent on the type of protein and cell being evaluated *(16)*.

Targeted integration into hot spots. CRISPR-Cas9 is a powerful tool for editing the genome (Figure 3). Transfecting cells with both a targeted CRISPR-Cas9 and transgenes with homologous flanking DNA sequences has made site-specific genome integration much easier and faster.

It is well understood, but often ignored, that certain regions of the genome are better for high expression levels. Hot spots have been studied in *S. cerevisiae* most recently through characterization of over 1,000 genomic loci. Wu *et al.* identified regions that resulted in 13-fold higher expression of red fluorescent protein (RFP) transgene than the mean *(5)*.

Smaller-scale studies have been performed with human embryonic kidney (HEK-293) cells, identifying eight "safe harbors" where random integration resulted in stable expression and minimal perturbation of the surrounding genes. A single integration site was identified as a hot spot with four-fold higher expression of green fluorescent protein (GFP) transgene *(17)*.

Transcriptional hot spots are usually identified through random integration; however, random integration is not truly random and is biased toward sites apt to integrate DNA. Therefore, while the hot spots that have been identified thus far do contribute to higher expression, these sites only explore a subset of the possible integration loci.

CRISPR-Cas9 shows promise for hot spot identification through genome-wide transgene integration but has not yet been used to create large libraries of gene integration sites to identify additional hot spots. The selection of highly productive cells and identification of genome-integration sites, as well as epigenetic marks or context relevant to the expression phenotype, is a complementary approach.

Writing and editing epigenetic marks. DNA-binding domains, including ZFs, transcription activator-like effectors (TALE), and nuclease deficient mutants of Cas9 (dCas9), have been fused to chromatin modifiers, allowing epigenetic marks to be specifically targeted to certain loci. For example, the H3K9 methyltransferase was fused with a ZF, which provided strong evidence that H3K9 methylation causes transcriptional repression *(18)*. Similarly, dCas9 fused to the catalytic core of acetyltransferase p300 targets the acetylation of H3K27, which causes transcriptional activation *(19)*.

In another study, DNA methyltransferases (DMNT3A) and demethylases (TET1) have been targeted to specific sites through dCas9 fusion. TET1 can demethylate a promoter and activate transcription, while DNMT3A will methylate a promoter and repress transcription *(20)*. dCas9 fused to epigenetic modifiers that prevent gene repression through suppression of CpG methylation or heterochromatin spreading has not yet been applied to a bioprocess system, but seems likely in the near future.

Looking to the future

Synthetic biologists and biomolecular engineers are poised to address some of the biggest challenges in strain engineering by controlling epigenetics. The biggest questions for epigenetic engineering in bioprocessing are:

- What are the drivers of genome plasticity?
- How homogeneous can we make strains? Or, do we even want to make strains?
- Can engineered genes be shielded from unwanted epigenetic regulation?

For CHO cell line development, it is important that

MARK A. BLENNER, PhD, is the McQueen-Quattlebaum Associate Professor of Chemical and Biomolecular Engineering at Clemson Univ. (Email: blenner@clemson.edu; Address: 206 S. Palmetto Blvd., Clemson, SC 29634; Twitter: @MarkBlenner; Website: http://sites.clemson.edu/proteinengineering). His research focuses on cellular, metabolic, and protein engineering across microbial and mammalian cell culture systems. He has received several awards, including the NASA Early Career Faculty Award, the Air Force Office of Scientific Research Young Investigator Award, and the Dean's Professorship Award. He serves on the editorial board of *Microbial Cell Factories* and *Biochemical Engineering Journal*. Blenner received a BS in chemical engineering from Manhattan College, a PhD in chemical engineering from Columbia Univ., and trained as a postdoctoral fellow at Harvard Medical School and Children's Hospital Boston.

we control the drivers of genome plasticity. Eliminating the drivers of genetic variation is tempting because it may prevent specific productivity loss during scaleup and make strain development more predictable. On the other hand, genetic instability is what drives gene duplication and structural variability, which are important for identifying highly productive strains. Therefore, it might be more useful to be able to switch plasticity on and off instead of eliminating it altogether.

The heterogeneity of clonal populations is well appreciated, but it remains unclear what costs or benefits are associated with heterogeneity. Single-cell studies enabled by microfluidics have helped researchers identify over-productive and under-productive cells due to gene expression and metabolic biological noise that is influenced by microenvironmental perturbations, such as local concentration or pH gradients *(8)*. Control of epigenetic chromatin modification may prevent this noise. Conversely, if the epigenetic changes associated with over-productive strains can be identified, these epigenetic marks could be used to enforce a more homogeneous over-production phenotype.

We do not know much about how epigenetics influences production cell phenotypes. A promising approach to strain engineering could be to identify, engineer, and use orthogonal expression systems that shield genes of interest from the surrounding genomic and epigenetic context. For example, episomes carrying the scaffold/matrix attachment region (S/MAR) are easy to manipulate and could provide a platform that is not influenced by chromatin structure. However, the relative small size of such episomes might limit the copy number of transgenes and limit their utility for productive phenotypes *(21)*. Artificial chromosomes offer similar orthogonality from the native epigenetic regulation and can contain several million bases but are comparatively hard to manipulate *(22)*. Such orthogonal systems may be useful not only for strain engineering, but also for understanding the role of epigenetics in gene regulation and genomic stability. CEP

Literature Cited

1. **Chusainow, J., *et al.,*** "A Study of Monoclonal Antibody-Producing CHO Cell Lines: What Makes a Stable High Producer?," *Biotechnology and Bioengineering,* **102** (4), pp. 1182–1196 (2009).
2. **Yang, Y. S., *et al.,*** "DNA Methylation Contributes to Loss in Productivity of Monoclonal Antibody-Producing CHO Cell Lines," *Journal of Biotechnology,* **147** (3–4), pp. 180–185 (2010).
3. **Kim, M., *et al.,*** "A Mechanistic Understanding of Production Instability in CHO Cell Lines Expressing Recombinant Monoclonal Antibodies," *Biotechnology and Bioengineering,* **108** (10), pp. 2434–2446 (2011).
4. **Meilinger, D., *et al.,*** "Np95 Interacts with *De Novo* DNA Methyltransferases, Dnmt3a and Dnmt3b, and Mediates Epigenetic Silencing of the Viral CMV Promoter in Embryonic Stem Cells," *EMBO Reports,* **10** (11), pp. 1259–1264 (2009).
5. **Wu, X. L., *et al.,*** "Genome-Wide Landscape of Position Effects on Heterogeneous Gene Expression in *Saccharomyces Cerevisiae*," *Biotechnology for Biofuels,* **10,** p. 189 (2017).
6. **Duan, Z., *et al.,*** "A Three-Dimensional Model of the Yeast Genome," *Nature,* **465** (7296), pp. 363–367 (2010).
7. **Chen, X. S., and J. Z. Zhang,** "The Genomic Landscape of Position Effects on Protein Expression Level and Noise in Yeast," *Cell Systems,* **2** (5), pp. 347–354 (2016).
8. **Vasdekis, A. E., *et al.,*** "Origins of Cell-to-Cell Bioprocessing Diversity and Implications of the Extracellular Environment Revealed at the Single-Cell Level," *Scientific Reports,* **5** (2015).
9. **Aghcheh, R. K., and C. P. Kubicek,** "Epigenetics as an Emerging Tool for Improvement of Fungal Strains Used in Biotechnology," *Applied Microbiology and Biotechnology,* **99** (15), pp. 6167–6181 (2015).
10. **Xin, Q., *et al.,*** "*Trichoderma Reesei* Histone Acetyltransferase Gcn5 Regulates Fungal Growth, Conidiation, and Cellulase Gene Expression," *Current Microbiology,* **67** (5), pp. 580–589 (2013).
11. **Hakkinen, M., *et al.,*** "Screening of Candidate Regulators for Cellulase and Hemicellulase Production in *Trichoderma Reesei* and Identification of a Factor Essential for Cellulase Production," *Biotechnology for Biofuels,* **7** (2014).
12. **Gaszner, M., and G. Felsenfeld,** "Insulators: Exploiting Transcriptional and Epigenetic Mechanisms," *Nature Reviews Genetics,* **7** (9), pp. 703–713 (2006).
13. **Mariati, *et al.,*** "Insertion of Core CpG Island Element into Human CMV Promoter for Enhancing Recombinant Protein Expression Stability in CHO Cells," *Biotechnology Progress,* **30** (3), pp. 523–534 (2014).
14. **Betts, Z., *et al.,*** "Evaluating the Interaction Between UCOE and DHFR-Linked Amplification and Stability of Recombinant Protein Expression," *Biotechnology Progress,* **31** (4), pp. 1014–1025 (2015).
15. **Nair, A. R., *et al.,*** "Effect of Different UCOE-Promoter Combinations in Creation of Engineered Cell Lines for the Production of Factor VIII," *BMC Research Notes,* **4,** p. 178 (2011).
16. **Simpson, D. J., *et al.,*** "Expression Elements," U.S. Patent and Trademark Office (2009).
17. **Cheng, J. K., *et al.,*** "Identifying and Retargeting Transcriptional Hot Spots in the Human Genome," *Biotechnology Journal,* **11** (8), pp. 1100–1109 (2016).
18. **Snowden, A. W., *et al.,*** "Gene-Specific Targeting of H3K9 Methylation Is Sufficient for Initiating Repression *In Vivo,*" *Current Biology,* **12** (24), pp. 2159–2166 (2002).
19. **Hilton, I. B., *et al.,* "**Epigenome Editing by a CRISPR-Cas9-Based Acetyltransferase Activates Genes from Promoters and Enhancers," *Nature Biotechnology,* **33** (5), pp. 510-U225 (2015).
20. **Liu, X. S., *et al.,*** "Editing DNA Methylation in the Mammalian Genome," *Cell,* **167** (1), pp. 233–247 (2016).
21. **Verghese, S. C., *et al.,*** "S/MAR Sequence Confers Long-Term Mitotic Stability on Non-Integrating Lentiviral Vector Episomes without Selection," *Nucleic Acids Research,* **42** (7) (2014).
22. **Kouprina, N., *et al.,*** "A New Generation of Human Artificial Chromosomes for Functional Genomics and Gene Therapy," *Cellular and Molecular Life Sciences,* **70** (7), pp. 1135–1148 (2013).