A New Genetic Toolkit for Enhanced Lactoferrin Production in *Pichia pastoris*

Tamilvendan Manavalan[1], Ang Ee Lui[2], Huimin Zhao[3]

[1] Centre for Precision Fermentation and Sustainability, Illinois Advanced Research Center at Singapore.
[2]Singapore Institute of Food and Biotechnology Innovation, A*STAR.
[3]US Department of Energy Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois Urbana-Champaign.

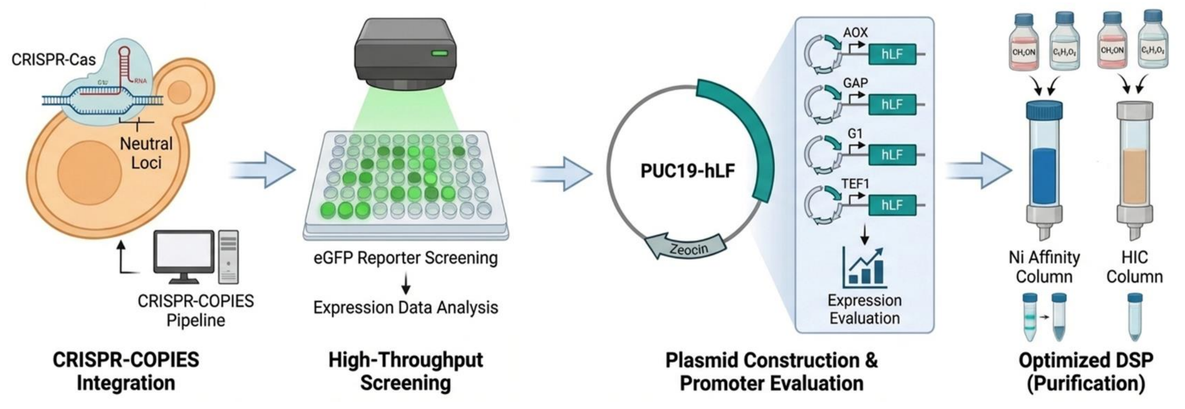
Background, Objectives, and Method

Background:
Human Lactoferrin (hLF) is a 77 kDa multifunctional glycoprotein belonging to the transferrin family [1]. Naturally occurring in breast milk and secreted by mucosal epithelial cells and neutrophils, hLF serves as a primary component of the innate immune system [1]. Its structure consists of two homologous N and C lobes, each capable of sequestering a ferric ion (Fe^{3+}) in coordination with a bicarbonate ion [2]. Beyond its established antibacterial, antifungal, and antiviral properties, hLF is a promising source of iron supplementation [3-5]. Traditional iron therapies often struggle with low absorption rates and adverse gastrointestinal side effects [3-5]. In contrast, hLF acts as a natural mineral carrier that delivers iron via specific receptors in the intestinal tract [3-5]. This mechanism significantly enhances bioavailability and minimizes oxidative stress compared to inorganic iron salts, offering a new approach to treating anemia [3-5]. To meet the demand for this therapeutic protein, *Pichia pastoris* serves as an ideal expression host. This strain offers several critical advantages over prokaryotic systems, including eukaryotic post-translational modifications for proper folding of protein, extracellular protein secretion and simplified downstream processing (DSP) [3-5].

Objectives:
The primary goal of this study is to establish a robust platform for the production and purification of recombinant hLF.

1. CRISPR-COPIES: Identify and validate high-expression neutral loci within the *P. pastoris* genome.
2. High-Throughput Validation: Utilize eGFP reporters to screen for high-expression potential of identified neutral loci.
3. Expression Profiling: Utilize high expression neutral loci to express hLF using various promoters.
4. Process Optimization: Refine DSP protocols for high-purity hLF and evaluate its yield efficiency.

Method:
This project optimizes human lactoferrin (hLF) production in *P. pastoris* through four clear steps:
Step 1: Use CRISPR-COPIES to identify list of neutral site; Step 2: Test neutral sites by using an eGFP reporter gene to rank for high expression. Step 3: Establish hLF expression using the standard promoters by replacing eGFP with hLF at the highest-ranked sites to optimize final recombinant protein titers; Step 4: Develop a robust DSP workflow.



Results and Discussion

Genomic Engineering and Neutral Loci Identification: To establish a versatile expression platform, we employed the CRISPR-COPIES computational pipeline to identify candidate neutral integration sites across the *P. pastoris* genome (Figures 1–3). To enhance targeted integration via homologous recombination (HR), the *ku70* gene was successfully deleted (Figure 4), effectively suppressing the non-homologous end joining (NHEJ) pathway.

Platform Validation via eGFP Reporter: The identified neutral sites were validated using an eGFP reporter system and plasmids were constructed as depicted in Figure 5. Fluorescence microscopy confirmed robust eGFP expression only in transformed strains (Figure 6). Quantitative analysis of the eGFP-transformants in BSMG media demonstrated significantly higher fluorescence intensity over time compared to the wild-type (WT) strain, confirming the suitability of these loci for stable recombinant protein expression (Figure 7).

Expression of hLF and Bioreactor Cultivation: hLF gene was expressed under GAP promoter and screened for initial growth kinetics using complex (BMGY/BMDY) and chemically defined (BSMG/BSMD) media (Table 1). Both media exhibited comparable growth patterns (Figure 8); however, BSMG was selected due to its reduced protease activity, minimized foaming during fermentation, and simplified DSP. Figure 9 shows the batch bioreactor cultivation reached peak biomass at 48 hours (OD_{600} ~ 235). SDS-PAGE analysis confirmed hLF expression from 48-96 hours (Figure 10).

DSP: Successful recovery of recombinant hLF was achieved from 96-hour culture supernatants using a two different purification strategy: Figure 11 shows the two different method of purification of hLF using hydrophobic Interaction Chromatography (HIC) and Ni-affinity chromatography (His-Trap). This approach yielded high-purity protein suitable for functional characterization.

Step 1:
CRISPR-COPIES (COmputational Pipeline for the Identification of CRISPR/Cas facilitated intEgration Sites)

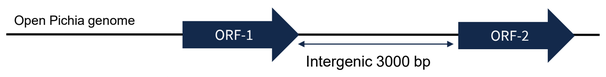


Figure 1. Scheme of CRISPR-COPIES Based Neutral Integration Site Identification
•Identified between the intergenic regions of two open reading frames (ORFs) without disrupting genes.
•Intergenic region we kept 3000 bp long to minimize impact on promoter and terminator elements.

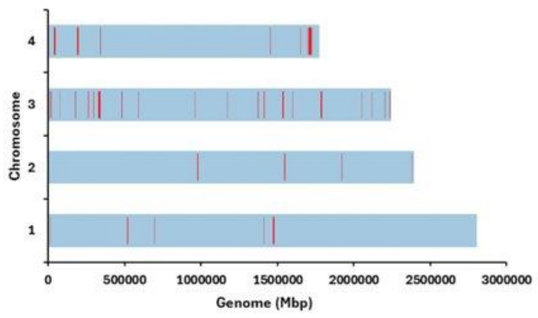


Figure 2: Genome-wide distribution of candidate neutral integration sites

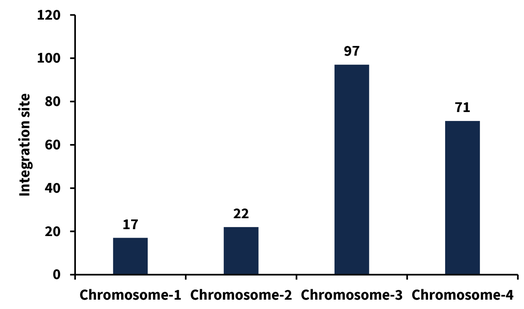


Figure 3: Frequency of identified neutral sites per chromosome

Step 2:
CRISPR-Cas9 plasmid

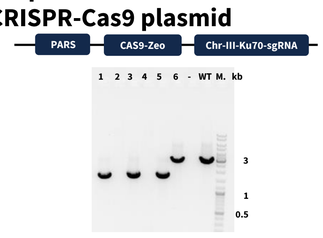


Figure 4. Deletion of *KU70* in *P. pastoris*
Lane 1, 3, 5: Δ*KU70* mutants (< 2800 bp)
Lane WT: in *P. pastoris* Wild Type (2800 bp)
Lane -: Negative control; Lane M: Marker

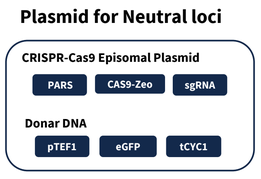


Figure 5. Scheme of CRISPR-Cas9 and donar DNA plasmids for Neutral loci validation

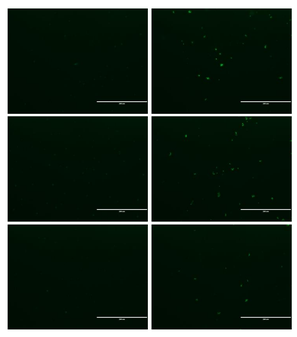


Figure 6. eGFP expression in *P. pastoris*

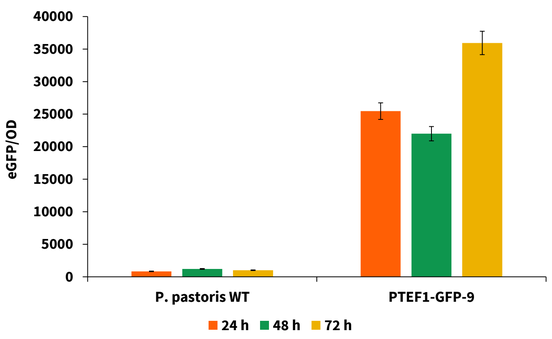


Figure 7. eGFP expression in *P. pastoris*

Step 3 and Step 4

Table 1. Expression of hLF under different promoter and media

Promoter	Inducer	Media 1	Media 2
AOX	Methanol	BMYG and BMYM	BSMG and BSMM
G1	Glucose	BMYG and BMYD	BSMG and BSMD
GAP	Constitutive	BMYG	BSMG
TEF1	Constitutive	BMYG	BSMG

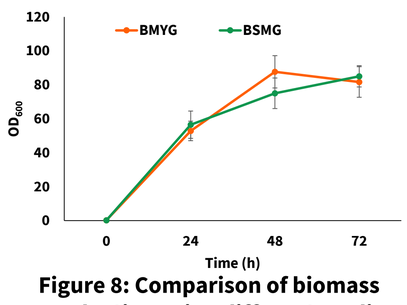


Figure 8: Comparison of biomass production using different media

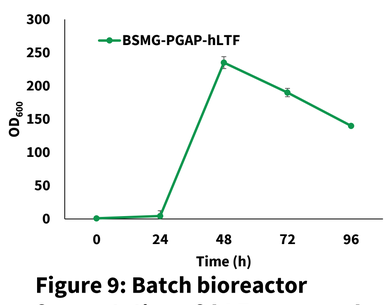


Figure 9: Batch bioreactor fermentation of hLF expressed under GAP promoter

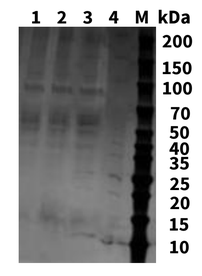


Figure 10. Expression of hLF at different time points
Lane 1: 48 h; Lane 2: 72 h
Lane 3: 96 h; Lane M: Marker

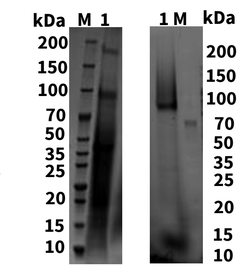


Figure 11. Purification of hLF using Hic-Trap (Left) and His-Trap (Right)
Lane 1: Purified hLF
Lane M: Marker

Conclusion

We identified 207 neutral integration sites using the CRISPR-COPIES method and successfully generated a Δ*ku70* background strain to enable highly efficient genomic engineering in *P. pastoris*. Validation using eGFP reporters confirmed these neutral loci as high-expression targets. Furthermore, we demonstrated the utility of this site by integrating and expressing hLF. By optimizing cultivation in BSMG media, we achieved high-density biomass accumulation (OD_{600} ~235) and established a robust DSP framework. Two distinct purification strategies HIC and His-Trap column chromatography were successfully used to purify hLF. This integrated genomic and bioprocessing framework provides a high-efficiency platform for the large-scale manufacturing of hLF and other high-value proteins.

Future Work

- Multi-Omics Discovery: Use RNA-seq and Proteomics to identify novel promoters and signal peptides for enhanced expression.
- Loci Validation: Systematically validate all remaining candidate neutral integration sites using eGFP reporters.
- Secretory Engineering: Identify and modulate genetic targets to optimize the protein secretion pathway.
- Yield Protection: Perform targeted protease knockouts to minimize extracellular hLF degradation.
- Industrial Scale-up: Optimize large-scale fermentation and purification protocols for engineered strains.
- Safety & Bioequivalence: Compare the allergenicity of recombinant hLF with the native human version.

References

1. Nuijens, J.H., van Berkel, P.H.C. & Schanbacher, F.L. Structure and biological actions of lactoferrin. J. Mammary Gland Biol. Neoplasia 1, 285-295 (1996); **2.** Anderson, B.F., Baker, H.M., Norris, G.E., Rice, D.W. & Baker, E.N. Structure of human lactoferrin: crystallographic structure analysis and refinement at 2.8 Å resolution. J. Mol. Biol. 209, 711-734 (1989); **3.** Conesa, C., Calvo, M., Sánchez, L. Recombinant human lactoferrin: a valuable protein for pharmaceutical products and functional foods Biotechnol. Adv., 28 (6), 831-838 (2010); **4.** Vishwanath-Deutsch, R., Dallas, D.C., Besada-Lombana, P., Katz, L., Conze, D., Kruger, C.; Clark, A.J.; Peterson, R., Malinczak, C.-A. A Review of the Safety Evidence on Recombinant Human Lactoferrin for Use as a Food Ingredient. Food Chem. Toxicol. 189, 114727 (2024). **5.** Wang, M., Sun, Z., Yu, T., Ding, F., Li, L., Wang, X. Large-scale production of recombinant human lactoferrin from high-expression, marker-free transgenic cloned cows. Sci Rep. 7(1):10733 (2017).

This research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.