



Rapid Screening for High Expressing Multicopy Recombinants and Enhanced Epidermal Growth Factor (EGF) Protein Production Using *Pichia Pastoris*

Zhouliang Guo¹ · Zhipeng Li¹ · Lu Li¹ · Qing Zhao¹ · Wei Li² · Aipo Diao¹

Accepted: 7 July 2022 / Published online: 4 August 2022

© The Author(s), under exclusive licence to Springer Nature B.V. 2022, Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Abstract

Epidermal growth factor (EGF) is a highly effective stimulator which promotes cell proliferation and this has led it to being proposed as a prospective healing agent for various treatments of wounds and many other clinical applications. With such a high demand for this protein, it is therefore important to firstly develop an expression system that efficiently generates sufficient amounts of the protein but also of high purity and biological activity. In this study, we have developed an efficient recombinant EGF production route that includes rapid multicopy recombinant screening using *Pichia pastoris*. In addition, we have optimized protein expression conditions and a simple protein purification procedure. Firstly, quantification of the ratio of *hEGF* PCR band brightness to *AOX1* was used to screen the multicopy transformants. Furthermore, Real-time PCR was used to detect the copy number of *hEGF* integrated into the *Pichia pastoris* genome, which is consistent with the normal PCR results. Optimization of recombinant hEGF expression using the selected multicopy *Pichia pastoris* strain was performed, hEGF protein was highly expressed under pH 7 conditions for a culture period of 60 h. Moreover, ammonium sulfate precipitation in combination with cationic affinity and hydrophobic interaction chromatography were used to produce 99% pure recombinant hEGF, with high biological activity of 2×10^6 IU/mg. Thus we have developed a method using *Pichia pastoris* to produce high levels of hEGF of high purity and specific activity. Our route may also be suitable for the production of other therapeutic recombinant proteins.

Keywords EGF · *pichia pastoris* · Multicopy · PCR · purification

Introduction

Epidermal growth factor (EGF) is a peptide consisting of 53 amino acid residues and three disulfide bonds, which was first discovered in the mouse submaxillary gland (Cohen

1962). The human epidermal growth factor (hEGF) was first isolated from urine (Starkey et al. 1975) and functional studies demonstrated that it promotes cell proliferation and inhibits gastric acid secretion. Accordingly, hEGF has been widely used in wound repair, corneal injuries and gastric ulcers (Cohen and Carpenter 1975). In particular, hEGF proteins conjugated to hyaluronate or loaded onto heparin-based hydrogel sheets showed a long-term effect to heal chronic skin wounds (Kim et al. 2018; Goh et al. 2016). Traditionally, EGF was isolated from fresh animal urine, blood, breast milk and gastric juice via a complicated method (Starkey et al. 1975; Parries et al. 1995). The yield and purity of the EGF extracted from this process were also low and could not meet the demand of clinical applications. Therefore, several studies have attempted to increase the production of recombinant hEGF using genetic engineering techniques. Over the last two decades, various host systems

✉ Wei Li
liivywei@126.com

✉ Aipo Diao
diaoaipo@tust.edu.cn

¹ School of Biotechnology, Key Lab of Industrial Fermentation Microbiology of the Ministry of Education, Tianjin University of Science and Technology, 300457 Tianjin, China

² School of Basic Medical Science, Inner Mongolia Medical University, 010110 Hohhot, Inner Mongolia, China

have been used to produce hEGF, including *Escherichia coli* (Lee et al. 2000, 2003; Sharma et al. 2008), *Bacillus brevis* (Ebisu et al. 1996), *Saccharomyces cerevisiae* (Topczewska and Bolewska 1993), *baculovirus* (Wei et al. 2006), *Arachis Hypogaea* L (Yao et al. 2019), rice and tobacco (Wu et al. 2014; Bai et al. 2007). However, most of these systems have limitations mainly involving a low level of protein production and or low biological activity (Allen et al. 1985; Abdull Razis et al. 2006).

The methylotrophic yeast *Pichia pastoris* has been widely used in heterologous protein expression due to its many advantages, such as achieving high cell density, post-translational protein modifications and possession of an efficient secretory system (Juturu & Wu 2017). Generally, the copy number of the gene of interest in the host cells determines the overall protein expression level. Therefore, methods describing screening for high-yielding strains of recombinant protein expression have been reported. The positive transformants of *Pichia pastoris* GS115 were screened on Yeast Extract Peptone Dextrose Medium (YPD) agar supplemented with geneticin to select for hyper-resistant clones (Eissazadeh et al. 2017). Specific radioimmunoassay polyclone antiserum against human luteinizing hormone was used to detect the ability to secrete luteinizing hormone (Gadkari et al. 2003). However, hyper-resistant clones selected using YPD agar supplemented with geneticin could not determine the gene copy number, and a specific RIA method required different antibodies according to target proteins. Moreover, in the industrial production process of recombinant proteins, the importance of expression cassettes quantitation in *Pichia pastoris* was emphasized (Inan et al. 2006), and Real-time PCR has been used to determine the gene copy number in *Pichia pastoris* (Hartner et al. 2008; Schroer et al. 2010; Marx et al. 2009; Abad et al. 2010). Although this method could accurately determine gene copy number, it was not suitable for industrial application due to its complex operation and the high cost of equipment and consumables.

In order to enhance production of functional recombinant EGF proteins for wide range of clinical applications, in this study, we have developed a rapid screening method to discover high copy colonies highly expressing recombinant hEGF in *Pichia pastoris* yeast using a traditional PCR method, as well as optimized protein expression conditions and an efficient untagged hEGF protein purification protocol using cationic affinity followed by hydrophobic interaction chromatography.

Materials and Methods

Materials

Reagents for PCR amplification and restriction enzymes were purchased from Thermo (Shanghai, China). The hEGF standard for determination of hEGF activity was purchased from NIFDC (Beijing, China). All chemicals and reagents for expression, purification and protein assays were from Sangon Biotech (Shanghai, China). The hEGF gene (GeneBank Accession No. M15672) sequencing was performed by Genewiz (Tianjin, China). Rabbit anti-EGF polyclonal antibody (Cat. No. D320578; 1:5000) was bought from Sangon Biotech (China), mouse anti-PI3K monoclonal antibody (Cat. No. sc-374,534; 1:200) and mouse anti-Akt monoclonal antibody (Cat. No. sc-81,434; 1:200) were purchased from Santa Cruz Biotechnology (USA), mouse anti-GAPDH monoclonal antibody (Cat. No. KM9002; 1:5000), goat anti-Mouse IgG (H+L)-HRP (Cat. No. LK2003; 1:5000) and goat anti-Rabbit IgG (H+L)-HRP (Cat. No. LK2001; 1:5000) were purchased from Sungene Biotech (China).

Construction of pPICZ α -hEGF Recombinant Plasmid and Transformation into *Pichia Pastoris*X33

The sequence corresponding to human EGF was amplified using the primers 5'CTCTCGAGA-AAAGAGAGGCTGAAGCTAATAG3' and 5'GTTCTAG-ATTAGCGCAGTTCCTCCCACTTC3' to introduce *Xho*I and *Xba*I sites for cloning. Additionally, a stop codon was introduced at the end of the gene to prevent expression of the optional hexa-histidine tag in the vector. The amplified fragment was cloned into the *Pichia pastoris* expression vector pPICZ α A (Invitrogen) downstream of the Alcohol Oxidase 1 (*AOX1*) promoter, the resulting recombinant DNA construct (pPICZ α -hEGF) was verified by DNA sequencing using *AOX1* specific primers. 10 μ g pPICZ α -hEGF plasmid was linearized using *Sac*I and electroporated into 80 μ l competent *Pichia pastoris* X33 cells using a standard yeast transformation method (Zhao et al. 2018). 1ml ice-cold 1 M sorbitol was added into the transformed cells and incubated at 30 °C for 60 min. The yeasts were then plated on YPD agar plates containing 600 μ g/ml, 800 μ g/ml and 1000 μ g/ml Zeocin respectively and incubated at 30 °C for 3–5 days.

Selection of Transformants and Analysis of Gene Copy Number

The genomic DNA of the transformants were extracted to be used as the PCR templates. Gene copy numbers were analyzed by PCR using the upstream/downstream *AOX1*

gene primers 5'-GACTGGTTCCAATTGACAAGC-3' and 5'-GCAAATGGCATTCTGACATCC-3'. The *AOX1* locus is the site for the *hEGF* gene integrated into the *Pichia pastoris* genome, therefore, theoretically the PCR products of positive transformants contain two bands at 2200 bp and 688 bp. The PCR products were detected by agarose electrophoresis. The 2200 bp and 688 bp band was further analyzed as unit gray value using ImageJ software, and the gray ratio was calculated according to the formula, $\text{gray ratio} = \frac{\text{gray value}_E}{\text{gray value}_A}$, gray value_E indicates the value of *hEGF* gene band, gray value_A indicates the value of *AOX1* gene band. Real-time PCR mixtures were prepared using SYBR Green Master Mix (ABI, U.S.A). Real-time PCR amplification was performed using an ABI StepOne instrument. The reaction was carried out at 95°C for 2 min, followed by 40 cycles of 95°C for 10 sec, 60°C for 30 sec, 72°C for 30 sec and 95°C for 1 min. The sequences (written from 5' to 3') of primers used for RT-PCR were as follows: *GAP-F*, 5'-ATATTAACGGTTTCGGACGTATTG-3', *GAP-R*, 5'-GATGTTGACAGGGTCTCTCTTTGG-3'; *AOX-F*, 5'-GAGACATGGCTCCTATGGTTTGG-3', *AOX-R*, 5'-CGTTCCTTGCAGTTGGCTTCTTC-3'; *hEGF-F*, 5'-CTCGAGAAAAGAGAGGCTGAAGCT-3', *hEGF-R*, 5'-TCTAGATTAGCGCAGTTCCCACCAC-3.

Optimization of Protein Expression Conditions

Selected colonies were inoculated and grown in flasks containing 10 ml BMGY medium (100 mM potassium phosphate buffer, 13.4 g/l YNB, 4×10^{-4} g/l biotin, 10 g/l glycerol) for 24–36 h. The culture broth was chilled after centrifugation at 8000 rpm for 5 min at RT, the supernatant was removed and the cells were then transferred to 400 ml BMMY medium (100 mM potassium phosphate buffer, 13.4 g/l YNB, 4×10^{-4} g/l biotin) for 6 h. They were then harvested at the following times: 0, 24, 36, 48, 60, 72, 84, 96 h at pH (3, 5, 7, 9) and screened for the most optimal protein expression.

Purification of Recombinant hEGF

Solid ammonium sulfate was added to the culture supernatants at different concentrations and incubated for 30 min at 4 °C, then centrifuged at 20,000 rpm for 10 min. The 60% ammonium sulfate precipitation was dissolved in 20 mM citrate buffer (pH 2.2), then dialyzed against the same buffer overnight. For experiments with RTU Fast SP FF, a cation-exchange chromatography column (Gadkari et al. 2003), the column was first equilibrated with 20 mM citrate buffer (pH 2.2). After column equilibration, the sample was injected, then 20 mM sodium citrate buffer (pH 7) used as elution buffer was mixed with a 0–100% linear gradient. The eluted

fractions were collected separately and analyzed by 12% SDS-PAGE. For RTU Fast Phenyl HP, hydrophobic interaction chromatography column (Gadkari et al. 2003), the column was first equilibrated with 20 mM phosphate buffer (pH 7) containing 1 M NaCl, then the collected fraction from SP column was injected. Elution was then carried out in 0–100% linear gradient mixing with 20 mM phosphate buffer (pH 7). Each eluted fraction was collected separately and analyzed by SDS-PAGE. Finally, the most promising fraction was dialyzed and dissolved in 20 mM phosphate buffer solution (pH 7.4). The columns were regenerated after each chromatographic run according to the manufacturer's instructions. Purification experiments were carried out using AKTA Pure chromatographic platforms (GE Healthcare). The RTU Fast SP FF column and RTU Fast Phenyl HP column were purchased from HuaChun Biological Technology (Tianjin, China).

Cell Culture and Analysis of the Biological Activity of Recombinant hEGF

NIH/3T3, a fibroblast cell line that was isolated from a mouse NIH/Swiss embryo, was obtained from the American Type Culture Collection (ATCC), and cultured in DMEM (Gibco, Cat. No. 12,100,046) supplemented with 10% FBS (EVERY GREEN), 100 U/ml penicillin/0.1 mg/ml streptomycin (Beyotime) in a humidified atmosphere with 5% CO₂ at 37 °C.

To determine the biological activity of purified recombinant hEGF, firstly the proliferation rate of NIH/3T3 cells under hEGF stimulation was determined using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) assay (Robinson et al., 1996). The activity of the hEGF standard stock was 200 IU. The NIH/3T3 cells were cultured in complete medium (DMEM medium with 10% fetal bovine albumin) under 5% CO₂ at 37 °C for 24 h. The cells were then seeded into 96-well plates (5×10^3 cells/well). After incubation at 37 °C and 5% CO₂ for 24 h, the medium was replaced with maintenance medium, and the cells were incubated under the same conditions for 24 h. Then both the purified recombinant hEGF and hEGF standards were added into the media to make a series of concentrations. After 72 h incubation, 20 µl MTT solution (0.5 mg/ml in PBS) was added and incubated for an additional 4 h at 37 °C. Finally, the supernatant was discarded and 200 µl dimethyl sulfoxide (DMSO) was added to fully dissolve the purple crystal, then measured at an absorbance of 490 nm using a microplate reader (Bio-Rad, USA). Moreover, to measure EGFR signaling, the NIH/3T3 cells were incubated with different concentrations (0, 0.25, 0.5 and 1.0 ng/ml) of hEGF for 24 h. The cells were collected and analyzed by Western blot using antibodies against PI3K and Akt,

Western Blot

Cell lysates and protein samples were extracted and separated by 12% SDS-PAGE and transferred onto PVDF membranes. After blocking with 5% non-fat milk in Tris-buffered saline containing 0.1% Tween-20 (TBST), the membranes were incubated with primary antibody at 4 °C overnight. After washing with TBST buffer, the membranes were incubated for an additional 2 h with the appropriate secondary antibodies conjugated to HRP. The immunoblot bands were visualized using a chemiluminescence analysis system (SageCreation, Beijing, China). The density of the bands was analyzed using ImageJ software (Wayne asband, NIH, Bethesda, MD, USA).

Statistical Analysis

Quantitative data were expressed as mean \pm SD. Student's unpaired test were used to compare the differences between groups using GraphPad Prism 8.3.0 (GraphPad Software, USA). P values < 0.05 were considered statistically significant.

Results

Screening of Multi-Copy Recombinants Highly Expressing Recombinant hEGF

To produce recombinant human EGF in *Pichia pastoris* X33, the full-length human *EGF* cDNA was cloned into expression vector pPICZ α for secreted expression using the α -factor secretion signal induced by methanol. The recombinant plasmid pPICZ α -hEGF was firstly linearized by the restriction enzyme *SacI* and then transformed into *Pichia pastoris* X33 by electroporation (Fig. 1 A). The transformants were grown and selected on YPD plates with 600 and 800 μ g/ml zeocin. 65 zeocin-resistant transformants were chosen and further analyzed by PCR using the upstream/downstream *AOXI* gene primers. 63 were positive recombinants, which produced a band of 688 bp (*hEGF* gene 162 bp plus the 526 bp sequence from the expression vector), and 2200 bp PCR band corresponding to the amplification of the *AOXI* gene in the *Pichia pastoris* X33 genome. Interestingly, we found that the ratio of *hEGF* PCR band brightness to *AOXI* was quite different among the positive recombinants (Fig. 1B).

Therefore, we speculated that the ratio might be related to the *hEGF* gene copy number integrated into the *Pichia pastoris* genome, suggesting the higher ratio of *hEGF* to *AOXI*, the more copies of *hEGF* gene integrated into the yeast genome. Thus, among the 62 positive recombinants, 5

representative recombinants were selected for further analysis of the ratio of two PCR bands using ImageJ software, and it was found that No. 4 recombinant showed the highest ratio, followed by No. 3 recombinant, and then No. 1 recombinant showing the lowest (Fig. 2 A). Furthermore, Real-time PCR was used to detect the copy numbers of *hEGF* and *AOXI* integrated into the yeast genome using the housekeeper gene *GAP* as the internal reference gene. Real-time PCR result indicated that the No. 4 recombinant had the highest *hEGF* gene copy number of 10, No. 3 recombinant had 6 *hEGF* gene copies, No. 1 recombinant had no gene copies and the other recombinants had only one copy, which was consistent with the traditional PCR results and ImageJ quantification data (Fig. 2B).

It has been usually considered that recombinants containing multi-copy integrations of the expression cassettes were associated with high level expression of the desired protein. Therefore, these 5 selected recombinants were used for further protein expression trials. The expressed and secreted recombinant protein in the culture supernatant was evaluated by SDS-PAGE followed by silver staining. A protein band at a molecular weight of about 6 kDa corresponding to the calculated molecular weight of the recombinant hEGF was consistently detected. The No. 4 recombinant containing the highest *hEGF* gene copy number produced much higher expression levels (0.23 mg/ml) than others (Fig. 2 C). The expressed protein was further verified by Western blot using an anti-EGF antibody (Fig. 2D). These results suggest that the rapid screening method using PCR can be used to efficiently select multi-copy recombinants that highly express recombinant proteins such as hEGF in this case using *Pichia pastoris*.

Optimization of Recombinant hEGF Expression

To optimize the expression conditions for recombinant hEGF protein, the selected multicopy recombinant No.4 strain was first inoculated into BMGY media and the cultures were then transferred to BMMY media and 1% of methanol was added to induce protein expression. It has been previously well characterized that the period of culture expression time is conducive to obtain high-quality target protein so the culture samples were collected every 12 h and analyzed by SDS-PAGE. The expression of hEGF was increased with the culture period and reached a maximum of 0.26 mg/ml at 84 h then decreased to 0.25 mg/ml at 96 h (Fig. 3 A). However, after 60 h expression to 0.24 mg/ml, the accumulation rate of hEGF protein was only 0.02 mg/ml within 24 h. Considering the increasing non-target proteins in the later stage of culture cycle of *Pichia pastoris*, as well as the industrial production efficiency, this suggested

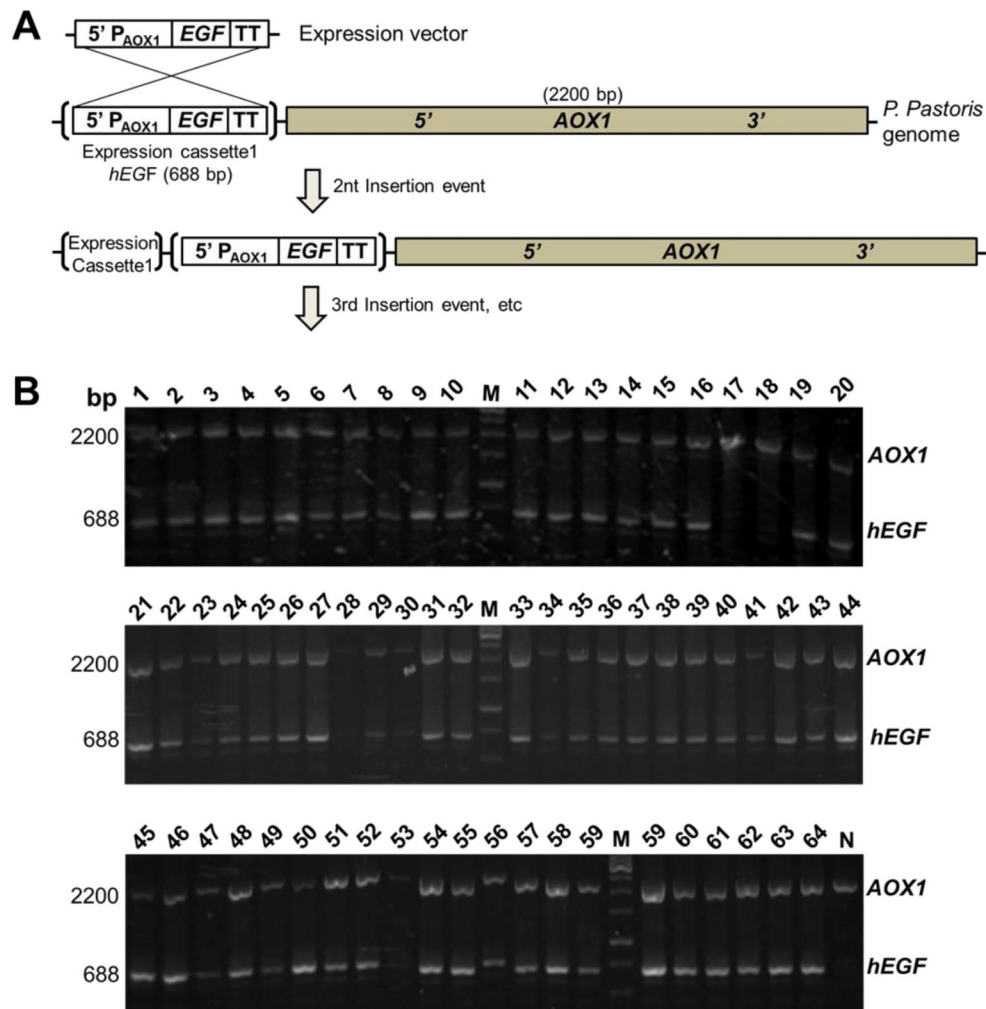


Fig. 1 **A** Schematic representation of integration events in *Pichia pastoris*. 5' P_{AOX1} : 5' *AOX1* promoter region, TT: *AOX1* transcription termination region. **B** Rapid screening of multi-copy transformants using PCR. M: DNA markers, N: negative control using X33

that 60 h culture was the optimized culture period for hEGF expression.

The pH of culture media is another important factor affecting protein expression. In order to determine the optimal pH value of culture, BMMY media were adjusted to different pHs (3, 5, 7 and 9) and cultured for 60 h to allow for protein expression. The cultures were collected and analyzed by SDS-PAGE. Under pH 7 condition, the expression of hEGF was significantly higher than that under pH 3 condition and was not significantly different from that under pH 5 and pH 9 conditions. Infact the proportion of hEGF in total protein terms reached 76% under pH 7 conditions, which was significantly higher than that of 35% and 24% under pH 5 and pH 9 conditions respectively. Considering the expression level and the proportion of hEGF, pH 7 was chosen as the optimized culture pH for hEGF expression (Fig. 3B).

Purification of Recombinant hEGF Protein

Recombinant hEGF protein was expressed under the optimized conditions. In order to concentrate hEGF in the culture supernatant for the next stage of purification, the expressed proteins were firstly purified by ammonium sulfate fractionation. To determine the range of ammonium sulfate concentration for fractional precipitation of the recombinant protein, firstly, the fermentation supernatant was precipitated in 45–100% saturated ammonium sulfate and analyzed by SDS-PAGE followed by silver staining. As shown in Fig. 4 A, about 70% target protein was precipitated in 55–60% saturated ammonium sulfate. Therefore, the culture media was firstly precipitated in 50% ammonium sulfate, then the supernatant was further precipitated when the concentration of ammonium sulfate reached 60%. The precipitated proteins were dissolved in citric acid sodium citrate buffer and desalted by dialysis. Ion exchange

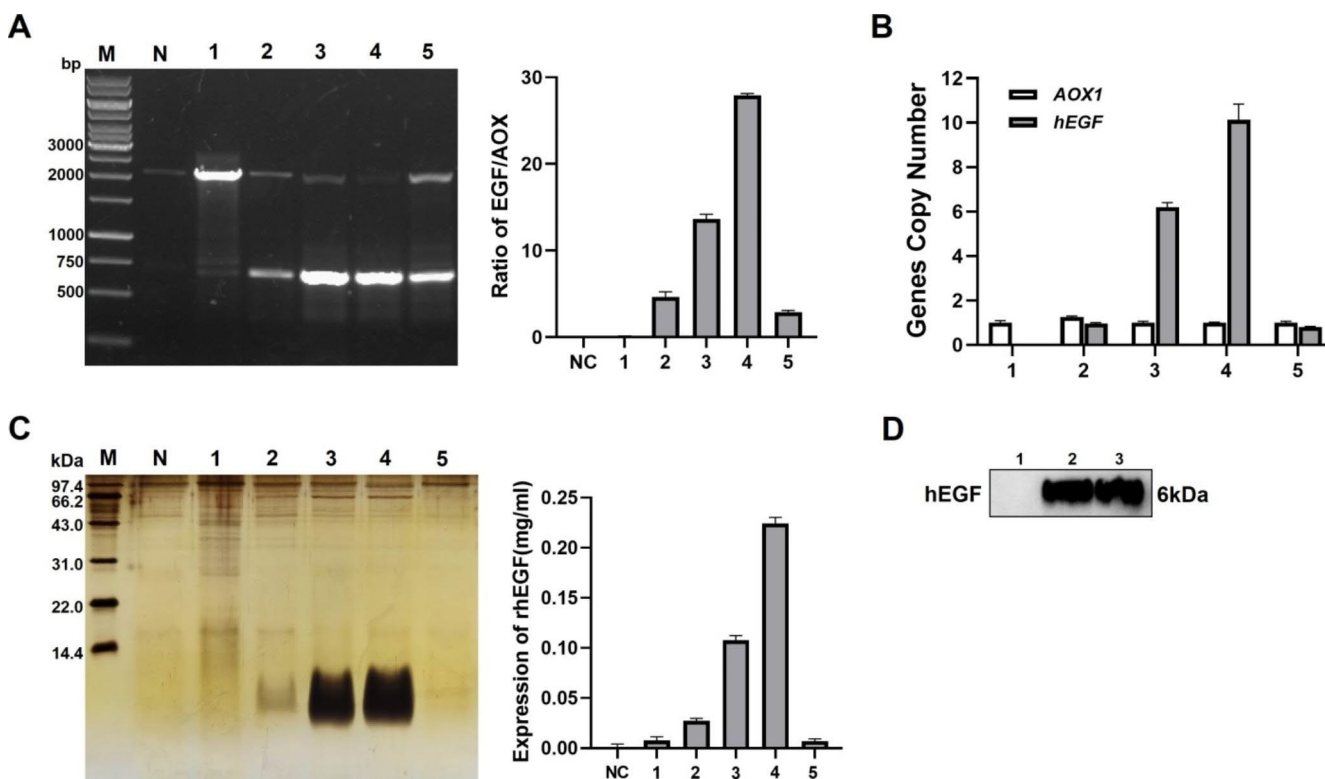


Fig. 2 Selection of multicopy recombinants highly expressing recombinant hEGF. **A** PCR analysis of 5 representative recombinants, the ratio of *EGF* PCR band brightness to *AOX* was quantified using ImageJ software. M: DNA markers; N: negative control using X33. **B** Real-time PCR analysis of the copy numbers of *hEGF* and *AOX1* genes. **C** Expression of recombinant hEGF was analyzed by SDS-PAGE. Following silver staining, the hEGF protein expression level was quantified using ImageJ software. M: Protein markers; N: negative control using X33, 1–5: 5 representative recombinants in the same order as **A**. **D** Expressed recombinant hEGF was analyzed by Western blot using an anti-EGF antibody. 1: negative control using X33; 2: hEGF standard; 3: sample collected from No. 4 recombinant

chromatography and hydrophobic interaction chromatography were used for further protein purification.

The RTU Fast SP FF column was employed to capture the hEGF protein in the citric acid sodium citrate buffer (20 mM pH 2), the target protein being eluted using 0–100% gradient of phosphate buffer (20 mM, pH 7). The collected fractions were analyzed by SDS-PAGE followed by silver staining which showed that hEGF was mainly in the second peak. The purity of the collected protein was increased from 95 to 97% compared to the loading protein after ammonium sulfate precipitation (Fig. 4B). The SP FF column purified protein was adjusted to phosphate buffer (20 mM, pH 7) containing 1 M NaCl, and injected into a RTU Fast Phenyl HP column, the target protein was eluted using 0–100% gradient of phosphate buffer (20 mM, pH 7). The collected protein fractions were analyzed by SDS-PAGE followed by silver staining which showed that hEGF was mainly in the second peak with a purity of 99% (Fig. 4C). Increased loading of 10 μ g purified hEGF protein was further analyzed by SDS-PAGE and following silver staining, no obvious non-target protein bands were detected (Fig. 4D). The recovery of the purified hEGF was 33% (Table 1).

Measuring the Biological Activity of Purified hEGF

EGF is a growth factor that stimulates the proliferation of epithelial and epidermal cells. The activity of EGF is determined by the dose-dependent proliferation of NIH/3T3 cells. MTT assay was used to detect the proliferation of NIH/3T3 cells treated with hEGF standard or our purified hEGF respectively. The activity curves were drawn by Graphpad software, and showed that the EC₅₀ value of the hEGF standard sample was 0.1 IU (Fig. 5A), while the EC₅₀ value of our purified hEGF sample was 0.0017 ng/ml (Fig. 5B), equivalently, the biological activity of the purified hEGF protein was calculated to be 2×10^6 IU/mg.

To further verify whether purified hEGF stimulates the EGFR (epidermal growth factor receptor) signaling pathway, different concentrations of hEGF were added into the culture medium respectively and NIH/3T3 cells treated for 24 h. The treated cells were then collected for Western blot analysis. EGFR signaling pathway is involved in PI3K/AKT signaling cascades, Western blot analysis showed that the levels of PI3K and Akt were significantly increased in a dose-dependent manner in the hEGF treated NIH/3T3 cells

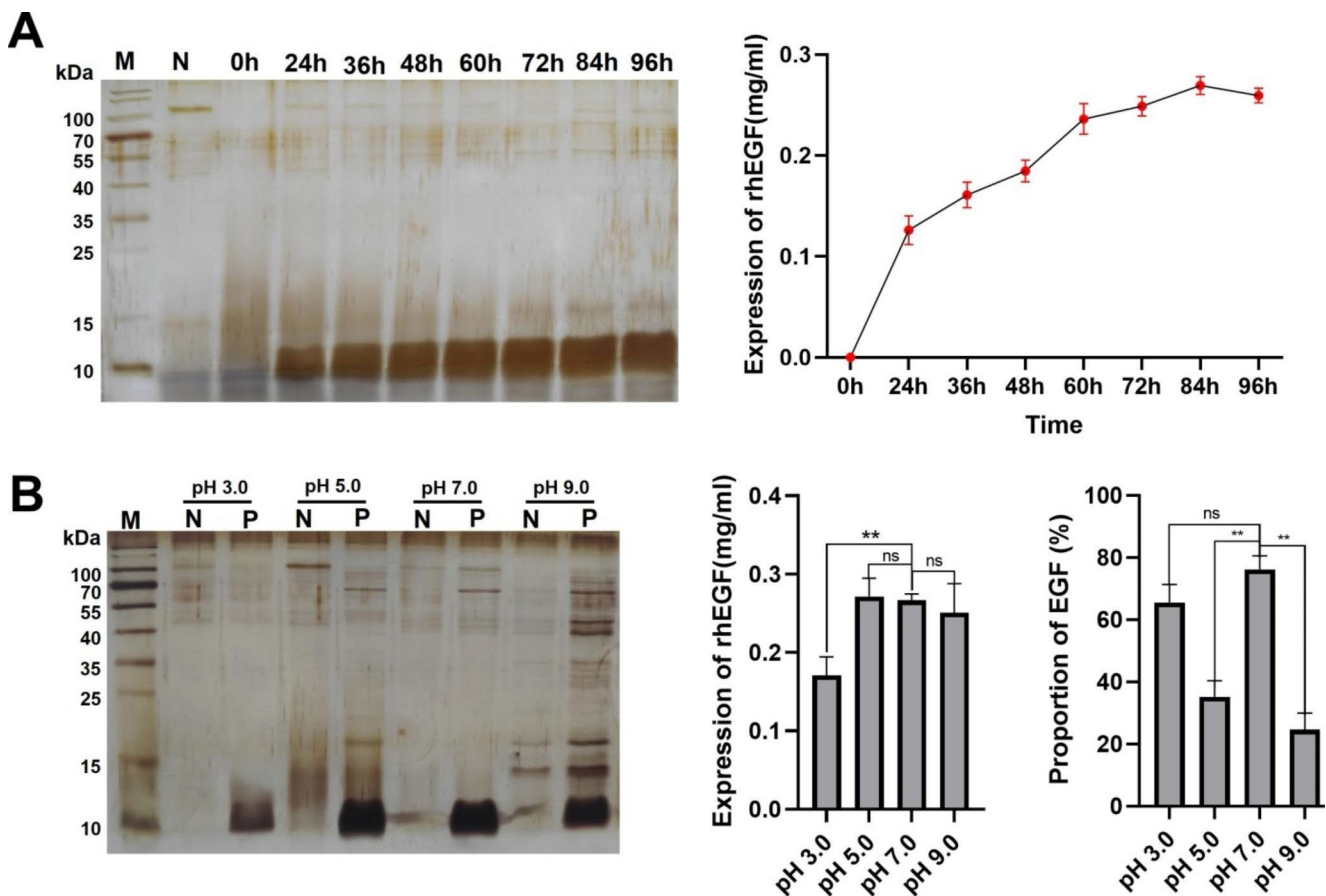


Fig. 3 **A** Optimization of culture times for hEGF expression. M: Protein markers; N: negative control using X33. **B** Optimization of culture media pH for hEGF expression (60 h culture). N indicates negative control using X33, P indicates No. 4 recombinant strain. Proteins were analyzed by SDS-PAGE following silver staining, expression of hEGF or proportion of hEGF was quantified using ImageJ software. The quantification was presented as mean \pm SD of three independent experiments (** $P < 0.01$, ns > 0.05)

(Fig. 5 C). This result indicates that our recombinant hEGF proteins are functional by activating the EGFR signaling.

Discussion

EGF is an effective stimulator to promote proliferation of a wide range type of cells, leading to a prospective healing agent for various corneal and skin wounds. Recombinant hEGF chemically conjugated to P64 protein derived from the Meningitis B bacteria has been developed as a therapeutic vaccine to treat advanced non-small cell lung cancer (Saavedra et al. 2018; Rodriguez et al. 2016). Therefore, the clinical demand of recombinant hEGF has boomed in recent years. hEGF is a small and yet complex protein containing of three disulfide bridges with glycosylation and so proper folding is important for its biological activity. The *Pichia pastoris* expression system is a promising host for the high production of active hEGF.

Generally, high target gene copy number strains lead to optimal expression rates using the *Pichia pastoris* protein

expression system (Inan et al. 2006; Hartner et al. 2008; Schroer et al. 2010; Marx et al. 2009; Abad et al. 2010). Multiple gene insertion events at a single locus in a *Pichia pastoris* cell do occur spontaneously with a low frequency and normally hundreds to thousands of Zeocin-resistant transformants need to be screened. Therefore, it is important to apply a high-throughput and more efficient way to screen and identify multicopy recombinants after Zeocin selection. Quantitative dot blots or Southern hybridization have been reported as methods to analyze gene copy number (Clare et al. 1991; Scorer et al. 1993), however these methods are time-costly. Real-time PCR (RT-PCR) was successfully used to evaluate the relative expression level of *miR-377-3p* in healthy and diseased groups (Amini Khorasgani et al. 2019). Labh used RT-PCR to detect the expression level of growth genes from liver and immune genes from head-kidney in Silver Carp (S. N. Labh. 2020). In this study, we have developed a rapid screening method to efficiently select multicopy recombinants using normal PCR, measuring the ratios of *hEGF* gene copies to *AOX1* as a means of monitoring for highly expressed recombinants. Further

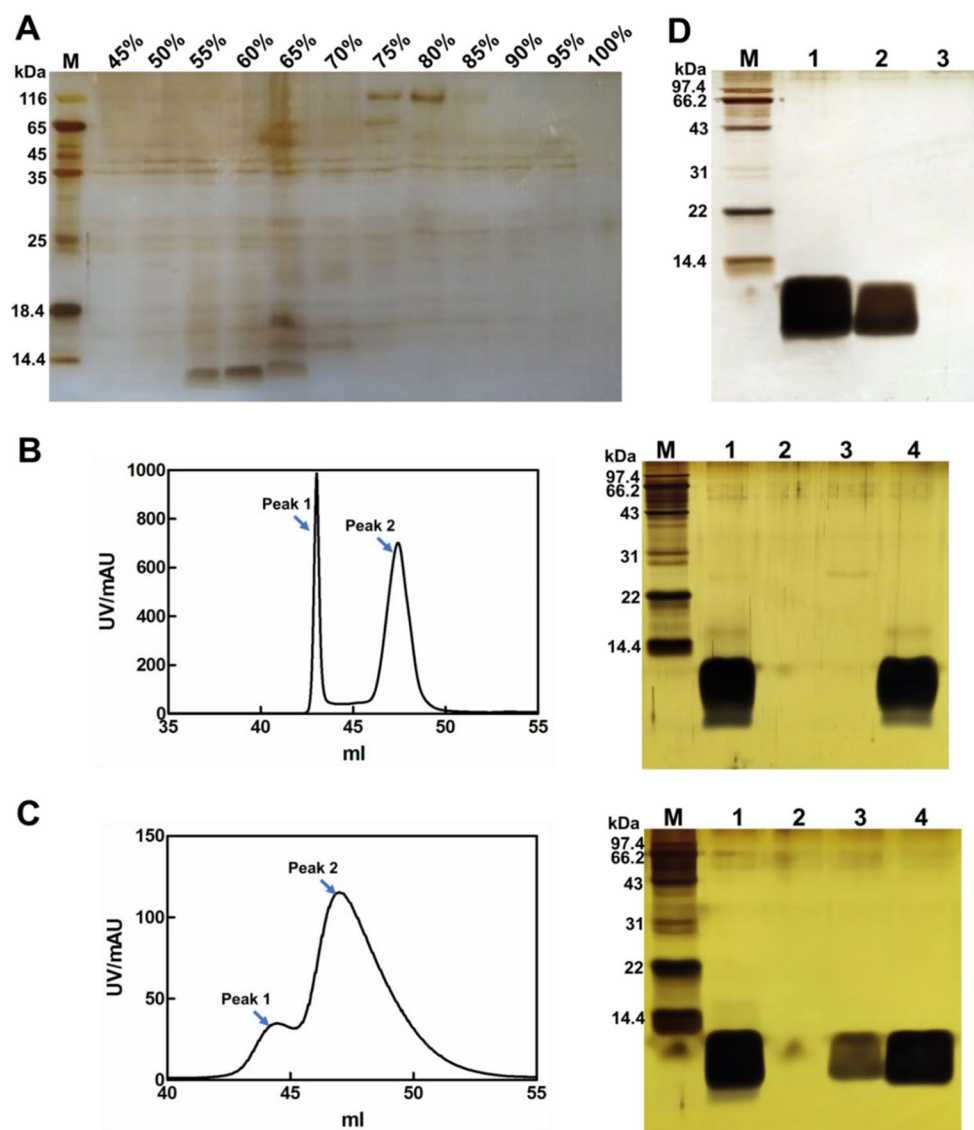


Fig. 4 **A** ammonium sulfate fractionation of expressed protein in the culture, proteins were analyzed by SDS-PAGE followed by silver staining. M: protein markers. **B** Expressed protein was purified using SP FF ion exchange chromatography, proteins were analyzed by SDS-PAGE followed by silver staining. Lane 1: Loading protein samples; 2: Flow through sample; 3: Peak 1 fraction; 4: Peak 2 fraction. **C** Purified protein using SP FF ion exchange chromatography was further purified using Phenyl HP hydrophobic interaction chromatography, proteins were analyzed by SDS-PAGE followed by silver staining. Lane 1: Loading protein samples; 2: Flow through sample; 3: Peak 1 fraction; 4: Peak 2 fraction. **D** Purified hEGF was analyzed by SDS-PAGE followed by silver staining. Lane 1: Purified hEGF; 2: hEGF standard; 3: Background without sample loaded

Table 1 Purification of recombinant hEGF produced in *Pichia pastoris*

Purification step	Yield (mg)	Yield (%)	Purity (%)
Fermentation supernatant	200	100	76
Ammonium sulfate precipitation	140	70	95
Ion exchange chromatography	97.5	48	97.5
Hydrophobic interaction chromatography	66	33	99.3

real-time PCR determining the actual number of gene copies reflected the normal PCR results, which indicates that the rapid screening method developed here using traditional PCR is feasible. For future work, we would like to monitor

the stability of selected multicopy recombinants after long time passage of the *Pichia pastoris* strains.

In addition, the purity of industrially produced recombinant hEGF is crucial for clinical application. It is well known that recombinant proteins fused with an affinity tag can be easily purified using affinity chromatography. In most cases, the recombinant hEGF fused with affinity tag was removed by treatment with a site-specific protease during purification (Ferrer Soler et al. 2003). However, during enzymatic processing, residual amino acids from the fused tag might be immunogenic and have negative impacts on clinical utilization (Kaukoranta-Tolvanen et al. 1996). Moreover,

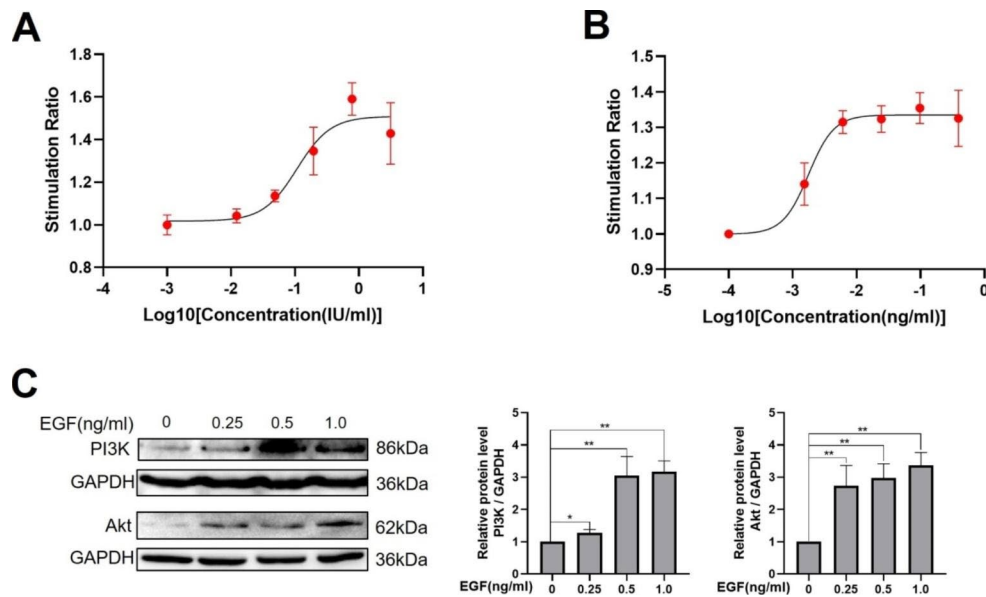


Fig. 5 Measurement of the biological activity of rhEGF. **A** MTT assay was used to detect the proliferation of NIH/3T3 cells treated with hEGF standard, the curve showed the hEGF standard concentration dependent stimulation of cell proliferation. **B** MTT assay was used to detect the proliferation of NIH/3T3 cells treated with purified hEGF, the curve showed the purified hEGF concentration dependent stimulation of cell proliferation. **C** The PI3K and Akt protein levels in the different concentrations (0, 0.25, 0.5 and 1.0 ng/ml) of hEGF treated NIH/3T3 cells were analyzed by Western blot using antibodies against PI3K and Akt, and quantified using ImageJ software. The quantification was presented as mean \pm SD of three independent experiments (* $P < 0.05$, ** $P < 0.01$)

highly specific proteases are typically expensive and could be cost-prohibitive (Fong et al. 2010). These problems limit the large-scale industrial production of high-quality hEGF production in *Pichia pastoris*.

Due to the different interactions between proteins and chromatographic column materials, retention time of proteins in the chromatographic columns are different, thus separation and purification can be performed according to the retention time of different proteins (Liu et al. 2020). Ammonium sulfate precipitation in combination with chromatography purification process is of low-cost, which can obtain about half of the target proteins from the fermentation broth, and the purity of target protein can reach 95%. According to the isoelectric point (pI) of hEGF i.e. 4.6, the ammonium sulfate precipitated expressed hEGF without tags was adjusted to 20 mM citrate buffer (pH 2.2) and captured by cation exchange chromatography. Using the RTU Fast SP FF column, about 2% of the non-target proteins at 26 kDa were removed and the recovery rate of the target protein with 97% purity was 48%. The remaining non-target proteins at 16 kDa, which had similar charge properties to hEGF, were then further divided by employing hydrophobic interaction chromatography. The RTU Fast Phenyl HP column could help to obtain 99% purity target protein with a recovery of 33%. Using these two chromatographic approaches we have produced a method to generate biologically active recombinant hEGF protein on a large scale.

EGF promotes cell growth and differentiation by binding to its receptor EGFR. The PI3K/Akt signaling pathway is closely related to cell growth and proliferation, and activated recombinant hEGF promotes cell proliferation via the EGFR/PI3K/Akt pathway (Wang et al. 2021; Ilowski et al. 2011). Therefore, the biological activity of EGF can be measured by detecting the proliferation rate of EGF treated cells or the activation of EGFR signaling. NIH/3T3 is a fibroblast cell line that was isolated from a mouse NIH/Swiss embryo and was usually used to detect biological activity of hEGF. Our results showed that the endogenous PI3K and Akt protein levels in the hEGF treated cells were significantly up-regulated. In addition, the activation of standard hEGF and purified hEGF on cell proliferation reached the maximum under concentration of 1.5 IU and 1.3 ng/ml respectively, and increasing amount of hEGF could not further promote cell proliferation. These results indicate the purified hEGF using our developed method is functional.

In summary, we have developed a rapid screening method to discover high expressing EGF multicopy recombinants in *Pichia pastoris*. The highly produced recombinant EGF proteins were functional and can be used for clinical applications. We believe that our *Pichia pastoris* method may also be used for the high expression of other recombinant proteins.

Abbreviations

EGF Epidermal growth factor

EGFR	Epidermal growth factor receptor
AOX1	Alcohol Oxidase 1
YPD	Yeast extract peptone dextrose
YNB	Yeast nitrogen base
BMGY	Buffered Glycerol-complex Medium
BMMY	Buffered Methanol-complex Medium
PI3K	Phosphoinositide 3-kinases
AKT	RAC(Rho family)-alpha serine/threonine-protein kinase

Acknowledgements We are grateful to Dr. Edward McKenzie (University of Manchester, UK) for critical reading of the manuscript. This research was supported by the Scientific and Technological Research Program of Tianjin Municipal Education Commission.

Data Availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare no conflict of interest.

Statement of Human and Animal Rights This article does not contain studies with human or animal subjects performed by any of the authors that should be approved by Ethics Committee

Statement of Informed Consent The article does not contain any studies in patients by any of the authors

References

- Abad S, Kitz K, Hormann A, Schreiner U, Hartner FS, Glieder A (2010) Real-time PCR-based determination of gene copy numbers in *Pichia pastoris*. *Biotechnol J* 5:413–420. <https://doi.org/10.1002/biot.200900233>
- Abdull Razis AF, Ismail EN, Hambali Z, Abdullah MNH, Ali AM, Lila MAM (2006) The Periplasmic Expression of Recombinant Human Epidermal Growth Factor (hEGF) in *Escherichia coli*, Asia-Pac. *J Mol Biol* 14:41–45. <https://www.researchgate.net/publication/230727612>
- Allen G, Paynter C, Winther M (1985) Production of Epidermal Growth Factor in *Escherichia coli* from a Synthetic Gene. *J Cell Sci Suppl* 3:29–38. <https://doi.org/10.1242/jcs>
- Bai JY, Zeng L, Hu YL, Li YF, Lin ZP, Shang SC, Shi YS (2007) Expression and characteristic of synthetic human epidermal growth factor (hEGF) in transgenic tobacco plants. *Biotechnol Lett* 29:2007–2012. <https://doi.org/10.1007/s10529-007-9438-y>
- Clare JJ, Rayment FB, Ballantine SP, Sreekrishna K, Romanos MA (1991) High-Level Expression of Tetanus Toxin Fragment C in *Pichia Pastoris* Strains Containing Multiple Tandem Integrations of the Gene. *Nat Biotechnol* 9:455–460. <https://doi.org/10.1038/nbt0591-455>
- Cohen SN (1962) Isolation of a Mouse Submaxillary Gland Protein Accelerating Incisor Eruption and Eyelid Opening in the New-born Animal. *J Biol Chem*. 237:1555–62. <http://dx.doi.org>
- Cohen S, Carpenter G (1975) Human epidermal growth factor: isolation and chemical and biological properties. *Proc. Nat. Acad. Sci.* 72(4):1317–1321. <https://doi.org/10.1073/pnas.72.4.1317>
- Ebisu S, Takagi H, Kadowaki K, Yamagata H, Udaka S (1996) The Efficient Production of Human Epidermal Growth Factor by *Bacillus brevis*. *Ann N Y Acad Sci* 782:115–122. <https://doi.org/10.1111/j.1749-6632.1996.tb40553.x>
- Eissazadeh S, Moeini H, Dezfouli MG, Heidary S, Nelofer R, Abdull MP (2017) Production of recombinant human epidermal growth factor in *Pichia pastoris*. *Braz J Microbiol* 48:286–293. <https://doi.org/10.1016/j.bjm.2016.10.017>
- Ferrer Soler L, Cedano J, Querol E (2003) Cloning, expression and purification of human epidermal growth factor using different expression systems. *J Chromatog B* 788:113–123. [https://doi.org/10.1016/s1570-0232\(02\)01035-8](https://doi.org/10.1016/s1570-0232(02)01035-8). de Llorens R.
- Fong BA, Wu WY, Wood DW (2010) The potential role of self-cleaving purification tags in commercial-scale processes. *Trends Biotechnol* 28:272–279. <https://doi.org/10.1016/j.tibtech.2010.02.003>
- Gadkari R, Deshpande R, Rajan R, Dighe RR (2003) Hyperexpression and purification of biologically active human luteinizing hormone and human chorionic gonadotropin using the methylotropic yeast, *Pichia pastoris*. *Protein Expr Purif* 32:175–184
- Goh M, Hwang Y, Tae G (2016) Epidermal growth factor loaded heparin-based hydrogel sheet for skin wound healing. *Carbohydr Polym* 147:251–260. <https://doi.org/10.1016/j.carbpol.2016.03.072>
- Hartner FS, Ruth C, Langenegger D, Johnson SN, Hyka P, Lin-Cereghino GP, Glieder A (2008) Promoter library designed for fine-tuned gene expression in *Pichia pastoris*. *Nucleic Acids Res* 36:76–90. <https://doi.org/10.1093/nar/gkn369>
- Ilowski M, Kleespies A (2011) Augmenter of liver regeneration (ALR) protects human hepatocytes against apoptosis. *Biochem Biophys Res Commun* 404(1):148–152 De Toni ENDONabauer, B., Jauch, K.W., Hengstler, J.G., Thasler, W.E. ()
- Inan M, Aryasomayajula D, Sinha J, Meagher MM (2006) Enhancement of protein secretion in *Pichia pastoris* by overexpression of protein disulfide isomerase. *Biotechnol Bioeng* 93:771–778. <https://doi.org/10.1002/bit.20762>
- Juturu V, Wu JC (2017) Heterologous Protein Expression in *Pichia pastoris*: Latest Research Progress and Applications. *Chembiochem* 19:7–21. <https://doi.org/10.1002/cbic.201700460>
- Kaukoranta-Tolvanen S-SE, Teppo A-M, Laitinen K, Saikku P, Linnavuori K, Leinonen M (1996) Growth of *Chlamydia pneumoniae* in cultured human peripheral blood mononuclear cells and induction of a cytokine response. *Microbiol Rev* 60:512–538. <https://doi.org/10.1006/mpat.1996.0056>
- Khorasgani MA, Nejad MP, Bashi MMM, Hedayati M (2019) Evaluation of mir-377-3p Expression in Patients with Multiple Sclerosis. *SciMed J* 1(2):48–54
- Kim YS, Sung DK, Kong WH, Kim H, Hahn SK (2018) Synergistic effects of hyaluronate-epidermal growth factor conjugate patch on chronic wound healing. *Biomater Sci* 6(5):1020–1030. <https://doi.org/10.1039/c8bm00079d>
- Labh SN (2020) Effects of Lapsi Choerospondias axillaris on Growth and Immune-Related Genes in Silver Carp (*Hypophthalmichthys molitrix*). *SciMed J* 2(2):86–99
- Lee JY, Yoon CS, Chung IY, Lee YS, Lee EK (2000) Scale-up process for expression and renaturation of recombinant human epidermal growth factor from *Escherichia coli* inclusion bodies. *Biotechnol Appl Chem* 31(3):245. <https://doi.org/10.1042/ba19990101>
- Lee YS, Suh CW, Park SK, Lee EK (2003) Purification of soluble human epidermal growth factor (hEGF) from recombinant *Escherichia coli* culture broth by using expanded-bed adsorption chromatography. *Biotechnol Appl Biochem* 31:245–248. <https://doi.org/10.1042/ba20020113>
- Liu S, Li Z, Yu B, Wang S, Shen Y, Cong H (2020) Recent advances on protein separation and purification methods. *Adv Colloid Interface* 284:102254–102276. <https://doi.org/10.1016/j.cis.2020.102254>

- Marx H, Mecklenbrauker A, Gasser B, Sauer M, Mattanovich D (2009) Directed gene copy number amplification in *Pichia pastoris* by vector integration into the ribosomal DNA locus. *FEMS. Yeast Res* 9:1260–1270. <https://doi.org/10.1111/j.1567-1364.2009.00561.x>
- Parries G, Chen K, Misono KS, Cohen S (1995) The Human Urinary Epidermal Growth Factor (EGF) Precursor. *J Biol Chem* 270(46):27954–27960. <https://doi.org/10.1074/jbc.270.46.27954>
- Robinson CJ, Gaines-Das R (1996) The international standard for epidermal growth factor (EGF): comparison of candidate preparations by in vitro bioassays and immunoassays. National Institute for Biological Standards and Control. *Growth Factors* 13:163–170. <https://doi.org/10.3109/08977199609003218>
- Rodriguez PC, Popa X, Martinez O, Mendoza S, Santiesteban E, Crespo T, Neninger E (2016) A Phase III Clinical Trial of the Epidermal Growth Factor Vaccine CIMAvax-EGF as Switch Maintenance Therapy in Advanced Non-Small Cell Lung Cancer Patients. *Clin Cancer Res* 22:3782–3790. <https://doi.org/10.1158/1078-0432.ccr-15-0855>
- Saavedra D, Neninger E, Rodriguez C, Viada C, Mazorra Z, Lage A, Crombet T (2018) CIMAvax-EGF: Toward long-term survival of advanced NSCLC. *Semin Oncol* 45:34–40. <https://doi.org/10.1053/j.seminoncol.2018.04.009>
- Schroer K, Peter Luef K, Stefan Hartner F, Glieder A, Pscheidt B (2010) Engineering the *Pichia pastoris* methanol oxidation pathway for improved NADH regeneration during whole-cell biotransformation. *Metab Eng* 12:8–17. <https://doi.org/10.1016/j.ymben.2009.08.006>
- Scorer CA, Buckholz RG, Clare JJ, Romanes MA (1993) The intracellular production and secretion of HIV-1 envelope protein in the methylotrophic yeast *Pichia pastoris*. *Gene* 136:111–119. [https://doi.org/10.1016/0378-1119\(93\)90454-b](https://doi.org/10.1016/0378-1119(93)90454-b)
- Sharma K, Babu PVC, Sasidhar P, Srinivas VK, Mohan VK, Krishna E (2008) Recombinant human epidermal growth factor inclusion body solubilization and refolding at large scale using expanded-bed adsorption chromatography from *Escherichia coli*. *Protein Exprs Purif* 60:7–14. <https://doi.org/10.1016/j.pep.2008.02.020>
- Starkey R, Cohen S, Orth D (1975) Epidermal growth factor: identification of a new hormone in human urine. *Science* 189(4205):800–802. <https://doi.org/10.1126/science.1172293>
- Topczewska J, Bolewska K (1993) Cloning and expression of the hEGF gene in *Saccharomyces cerevisiae*. *Acta Biochim Pol* 40:4–7. https://doi.org/10.18388/abp.1993_4831
- Wang Q, Tao C, Hannan A, Yoon S, Min X, Peregrin J, Qu X, Li H, Yu H, Zhao J, Zhang X (2021) Lacrimal gland budding requires PI3K-dependent suppression of EGF signaling. *Sci Adv* 30(27):eabf10687
- Wei Y, Jian C, Zhao X, Lv Z, Zhang Y (2006) Expression of polyhedrin-hEGF fusion protein in cultured cells and larvae of *Bombyx mori*. *Afr J Biotechnol* 5:1034–1040. <https://doi.org/10.4314/ajb.v5i11.42955>
- Wu CS, Kuo WT, Chang CY, Kuo JY, Tsai YT, Yu SM, Wu HT, Chen PW (2014) The modified rice alphaAmy8 promoter confers high-level foreign gene expression in a novel hypoxia-inducible expression system in transgenic rice seedlings. *Plant Mol Biol* 85:147–161. <https://doi.org/10.1007/s11103-014-0174-0>
- Yao Q, Yu Z, Liu P, Zheng H, Xu Y, Sai S, Zheng C (2019) High Efficient Expression and Purification of Human Epidermal Growth Factor in *Arachis Hypogaea* L. *Int J Mol Sci* 20:2045–2056. <https://doi.org/10.3390/ijms20082045>
- Zhao T, Li ZP, Guo ZL, Wang AL, Liu ZX, Zhao Q, Li YY, McKenzie EA, Diao AP (2018) Functional recombinant human Legumain protein expression in *Pichia pastoris* to enable screening for Legumain small molecule inhibitors. *Protein Exprs Purif* 150:12–16

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.