

## PCR assembly of synthetic human erythropoietin gene

**Yazmin Bustami**

School of Biological Sciences  
Universiti Sains Malaysia  
11800 Penang, Malaysia

**Ahmad Ramli Mohd Yahya**

School of Biological Sciences  
Universiti Sains Malaysia  
11800 Penang, Malaysia

**Tengku Sifzizul Tengku Muhammad**

Department of Biological Sciences  
Faculty of Science and Technology  
Universiti Malaysia Terengganu  
21030 Kuala Terengganu, Terengganu, Malaysia

**Alexander Chong Shu-Chien<sup>#</sup>**

School of Biological Sciences  
Universiti Sains Malaysia  
11800 Penang, Malaysia

**Amirul Al-Ashraf Abdullah<sup>#</sup>**

School of Biological Sciences  
Universiti Sains Malaysia  
11800 Penang, Malaysia

**Mohd Azizan Mohd Noor**

Malaysian Institute of Chemical and Bioengineering Technology  
Universiti Kuala Lumpur  
Lot 1988 Bandar Vendor, Tabuh Naning  
78000 Alor Gajah, Melaka, Malaysia

**Yahya Mat Arip\***

School of Biological Sciences  
Universiti Sains Malaysia  
11800 Penang, Malaysia  
Tel: 60 4 653 4013  
Fax: 60 4 656 5125  
E-mail: ymarip@usm.my

**Financial support:** This work was supported by R&D initiative grant of Malaysian Institute of Pharmaceuticals & Nutraceuticals (07-05-IFN-BPH001). Yazmin Bustami is supported by fellowship from Universiti Sains Malaysia.

**Present address:** <sup>#</sup>Malaysian Institute of Pharmaceuticals and Nutraceuticals, Ministry of Science, Technology and Innovation, SAINS@USM, 10 Persiaran Bukit Jambul, 11900, Pulau Pinang, Malaysia.

**Keywords:** cloning, erythropoietin, oligonucleotide assembly.

**Abbreviations:** CHO: Chinese hamster ovary

EPO: erythropoietin

huEPO: human erythropoietin

rhuEPO: recombinant human erythropoietin

PCR: polymerase chain reaction

**Human erythropoietin (huEPO) is a glycoprotein with important physiological functions, such as erythropoiesis, angiogenesis, and wound healing. A therapeutic protein, huEPO is commonly used to treat**

**patients suffering from renal and non-renal anemia. Recombinant human erythropoietin (rhuEPO) and endogenous huEPO are similar with respect to their biological and chemical properties. In this study, we**

\*Corresponding author

**describe the construction of synthetic huEPO gene to produce rhuEPO. The synthetic huEPO gene was constructed by overlapping oligonucleotides assembly and amplified by polymerase chain reaction (PCR). Twenty oligonucleotide sets, covering the huEPO gene sequence and two newly introduced restriction enzyme sites, were pulled together and amplified using *Pfu* DNA polymerase to produce the expected DNA products with sizes of ~500bp and ~600bp. The PCR products were ligated into pGEM-T plasmid vector to facilitate DNA sequencing process of the constructed huEPO gene and downstream cloning manipulation. DNA sequence analysis showed correctly assembled oligonucleotide sets, representing the huEPO gene sequence albeit with minor base mutations. Hence, oligonucleotides assembly and PCR amplification provide a convenient and speedy method for the synthesis of huEPO gene without depending on mRNA isolation and reverse transcription or the need to have a genomic library.**

Erythropoietin (EPO) is a member of hematopoietic growth factors involved in regulating red blood cell circulation. It was the first one to be identified and was not initially recognized as a colony stimulating factor. EPO controls proliferation and differentiation of erythroid precursor cells both *in vitro* and *in vivo* (Sasaki, 2003; Bahlmann et al. 2004; Arcasoy, 2008). The importance of EPO can be seen as anaemia develops because of the impaired or a blunted response to its production that decreases the number of circulating red blood cells. For these reasons, EPO functions as an anti-anaemic drug with high demand in medical for (Jelkmann, 2000). Unfortunately, the level of EPO in body fluids is very low with urine providing the source for the natural EPO (Cointe et al. 2000).

Realizing the need to produce EPO in high quantities had prompted the pioneering work of isolating 10 mg EPO from 2550 l of human urine (Miyake et al. 1977; Jelkmann, 2000). The preparation allowed the identification of the amino acid sequence and synthesized human EPO DNA probes for the isolation and cloning of the human EPO gene from mRNA in kidney and liver which are the site for EPO production (Jacobs et al. 1985; Dame et al. 1998; Jelkmann, 2000). Recombinant human erythropoietin (rhuEPO) has been produced in various cell lines; in particular Chinese hamster ovary (CHO) and baby hamster kidney (BHK) cells (Lin et al. 1985; Jacobs et al. 1985; Inoue et al. 1995). Currently, rhuEPO produced in CHO cell line is extensively used in the therapy to cure severe anemia (Cointe et al. 2000).

Oligonucleotide and polymerase chain reaction (PCR) provide the alternative in cloning, characterization and expression of the gene of interest. The availability of gene sequences from GenBank provides the convenience to synthetically construct any particular gene. This eliminates the need to have the original gene source: the tissue sample, the cell sample or the organism itself, as well as, the necessary processing of the gene sources. Methods of DNA

sequences assembly from oligonucleotides, such as, DNA ligase assembly, *FokI* gene synthesis method, self-priming PCR method and DNA shuffling are quite laborious, troublesome and time consuming (Stemmer et al. 1995).

We describe here the construction of a synthetic human erythropoietin (huEPO) gene using PCR amplified single-step assembly of forty overlapping oligonucleotides. The synthetic huEPO gene was successfully constructed in a fast and convenient process. In addition, this PCR gene assembly method would not require mRNA isolation and reverse transcription or the availability of genomic library.

## MATERIALS AND METHODS

In our work, we adapted and optimized the previously described PCR assembly techniques (Stemmer et al. 1995; Moore, 2001; Xiong et al. 2006; Mehrnejad et al. 2008). The constructed synthetic huEPO gene was cloned into a plasmid vector as a preparation for sub-cloning the synthetic gene into *Pichia pastoris* expression vector.

## Materials

Desalting oligonucleotides ranging from 39 mer to 57 mer was purchased from Research Biolabs. Upon arrival, the lyophilized oligonucleotides were dissolved in water to a final concentration of 100 µM. Restriction enzymes used in this work were purchased from New England Biolabs and Fermentas. Thermostable *Taq* DNA polymerase and *Pfu* proofreading DNA polymerase were obtained from Promega and Fermentas, respectively. Competent *E. coli* JM109, DNA ligation kit and pGEMT easy vector system were obtained from Promega.

## Oligonucleotide design

The oligonucleotide sets were designed based on the nucleotide sequence of huEPO mRNA (Accession no.: NM\_000799), obtained from the National Center for Biotechnology Information, USA. The oligonucleotide sets were designed according to *Pichia Pastoris* codons preference. Overlapping complimentary oligonucleotide sets were used to produce two constructed huEPO genes: i) 166 amino acids huEPO A that lacks the hydrophobic leader sequence and ii) 193 amino acids huEPO B with the presence of an additional 27 hydrophobic amino acids, acting as the signal peptide. A single-step gene assembly allowed the introduction of two unique restriction enzyme sites, *EcoRI* and *AvrII* at the 5'and 3' ends of huEPO gene, for directional cloning. There are many computer programs available for oligonucleotide design, such as DNA works (Hoover and Lubkowski, 2002), Gene2Oligo (Rouillard et al. 2004), DNABuilder, Assembly PCR Oligo Maker(Rydzanicz et al. 2005) and GeneDesign. These computer programs were used to design the optimal oligonucleotides sets that would enhance the hybridization efficiencies between the complimentary oligonucleotides sets; making sure the G + C contents remain in the range of

**1      EcoRI\***  
**5'-AAAAAAGAATTCA**GGCCCCACCGCCTCATCTGTGACAGCCGACTC  
**AGGTACCTTGGAGGCCAAGGAG**  
**3'-OGGGGTGGTGCGGAGTAGACACTGTGCGCTAGGACCTCTCCATG**  
**AACCTCGGGTTCCTC**  
**72                          126**  
**GCCGAGAAATATCACGACGGGC**  
**GAACACTCGAGCTTGAATGAGAATATCACTGTCCCA**  
**CGGCTCTTATACTGTGCGCGACAAGACTT**  
**TCGAACTTACTTTATAAGTGACAGGGT**  
**135                          189**  
**GACACCAA**  
**TTCATGCCTGGAAGAGGATGGAGGTGCGGCAGCAGGCCGTAGAACGTC**  
**CTGTGGTTCAATTAAAGATA**  
**TTCTCCACCTCCAGCCCGTGTGCGGCCAGGCCCTGTTAG**  
**198                          243**  
**GGCCTGGCCCTGCTGTGCGAACGTCGCTGCCGGGGCAGGCCCTGTTG**  
**ACOGTCCCGGAC**  
**GACGACGCGCTTCGACAGGACGCCCGGTCGGGACAACCAAGTTGAGA**  
**261                          297                          315**  
**TCCCAGCCGTGGGAGCCCTGAGCTGCATGTGGATAAAGCCGTAGTGGC**  
**AGCCTC**  
**ACOCTGGGGACGTGACGTACACCTATTCGGCAGTCACCGGAAGCGTGC**  
**324                          351                          378**  
**ACCAACTGCTTCGGGCTCTGGGAGCCCAGAAGGAAGCC**  
**CCTCCAGATGCGGCCCA**  
**TGGTGAAGACGAAGCCCGAGACCCCTCGGGTCTTCTCGGTAGAGGGGA**  
**CGCCGGAGT**  
**387                          405                          432**  
**GCTGCTCCACTCCGAACAATCACTGCT**  
**CGCAAACCTTCCGAGTCACTCCAAT**  
**CGACGAGGTTGAGGCTTGTAGTGACGACTGTGAAAG**  
**GAGAAGGCTCAGATGAGGTTA**  
**450                          \* 504**  
**TTCCCTCCGGGAAAG**  
**AAGCTGTACACAGGGGAGGGCTGAGGACAGGGGACAGATAATT-3'**  
**AAGGAGGCCCCTTTOGACTTGTGAC**  
**CCCCCTGGAGCTGTCTGTCCTGCTATTAAAGGATOCAAAAAA-5'**  
**AvrII**

**Figure 1. Gene assembly for huEPO gene A.** The gaps would be filled-in during the PCR cycles. (\*) represents the start and the stop codons. The same gene assembly scheme would apply for huEPO gene B.

33% to 42% and the melting temperature within 60°C to 77°C. Additionally, primer-dimer formations were minimized and hairpin loops formations were avoided. All these would create a favourable environment for the overlapping complimentary oligonucleotide sets to correctly assemble the huEPO gene.

### Gene assembly

Equal volumes of the overlapping complimentary oligonucleotides were mixed together in a tube. The mixture was further diluted to give a final concentration of 5 μM for each of the oligonucleotide. The mixture was added into a 50 μl reaction solution (1 X *Pfu* buffer, 1.5 mM MgSO<sub>4</sub>, 2 mM dNTPs mixed, 5% DMSO, and 1.25 unit of *Pfu* DNA polymerase) in a PCR tube. This would further dilute the oligonucleotides 10-fold. The oligonucleotides were assembled on a thermocycler: i) One cycle at 94°C for 2 min; ii) Forty cycles at 94°C, 75°C, 70°C, 65°C, 60°C, 57°C, 55°C and 72°C for 35 sec each, consecutively; iii) One cycle at 72°C for 10 min.

### Gene amplification

1.25 μl of the gene assembly mixture was added into a 50 μl PCR reaction (1 X *Pfu* buffer, 1.5 mM MgSO<sub>4</sub>, 2 mM dNTPs mixed, 5% DMSO, 5 μM each of the outermost 5'-

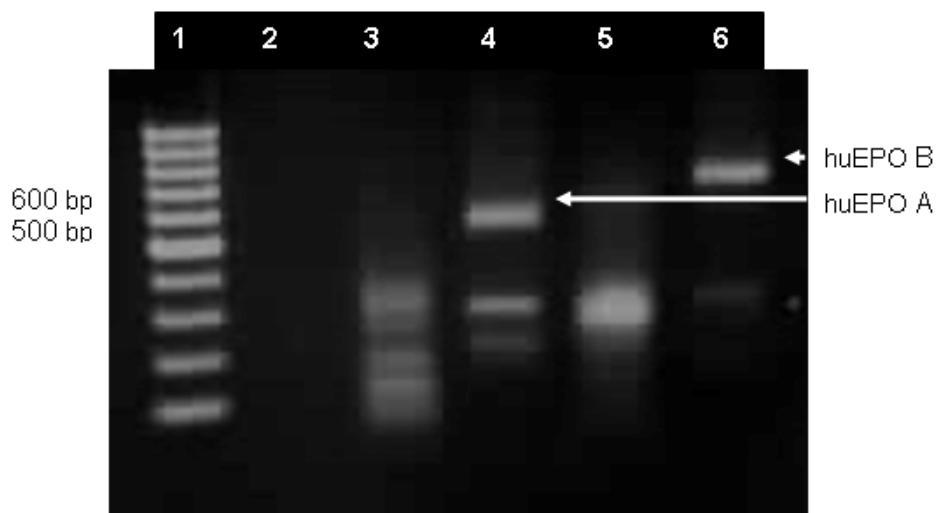
end and 3'-end primers and 1.25 unit of *Pfu* DNA polymerase). This would further dilute the gene assembly mixture 40-fold. Then, the mixture was subjected to twenty cycles of amplification and the reaction was set up as in the assembly PCR program. The PCR product was desalting using a PCR purification kit (Promega).

### A-tailing reaction

The thermostable DNA polymerase with proofreading activity, *Pfu*, generates blunt-ended fragments during PCR amplification. PCR fragments generated using proofreading polymerase can be modified using the A-tailing procedure to allow ligation into the pGEM<sup>®</sup>-T plasmid vector. Briefly, 5 μl of the gene amplified product was mixed with 1 X *Taq* buffer, 20 mM MgCl<sub>2</sub>, 0.2 mM dATP and 1.25 unit of *Taq* DNA polymerase. The mixture was incubated at 70°C for 30 min.

### Ligation and transformation

The A-tailing reaction mixture was ligated into pGEM<sup>®</sup>-T plasmid vector according to manufacturer's protocol. Specifically, 3 μl of the A-tailing reaction was mixed with 5 μl of 2 X Rapid ligation Buffer T4 DNA ligase, 1 μl pGEM<sup>®</sup>-T plasmid vector and 1 μl T4 DNA ligase. The ligation reaction was incubated overnight at 4°C. The



**Figure 2. 1.5% agarose gel electrophoresis showing the results of gene assembly and gene amplification processes.**

Lane 1: 100bp DNA ladder.

Lane 2: skip.

Lane 3: gene assembly of huEPO A gene.

Lane 4: gene amplification of huEPO A gene (~500bp).

Lane 4: gene assembly of huEPO B gene.

Lane 5: gene amplification of huEPO B gene (~600bp).

ligation products were transformed into *E. coli* JM109 competent cells using the standard heat-shock method. Selections of the transformed colonies were performed on Luria Bertani (LB) agar plates supplemented with ampicillin (100 µg/ml), 0.5 mM IPTG and X-gal (50 mg/ml).

### Screening and DNA sequencing

Colony PCR was performed, with modifications of Elbir et al. (2008), on white colonies to detect the presence huEPO gene. Briefly, a white colony was picked and mixed into 25 µl PCR mixture (2.5 µl 10 X Taq buffers, 2.0 mM MgCl<sub>2</sub>, 400 µM dNTPs mixed, 0.4 µM each of the outermost 5'-end and 3'-end primers, and 0.5 unit of *Taq* DNA polymerase. The mixture was subjected to a standard PCR protocol. Plasmids from the white colonies were extracted, purified and sent for DNA sequencing, performed at NHK Bioscience Solution.

## RESULTS

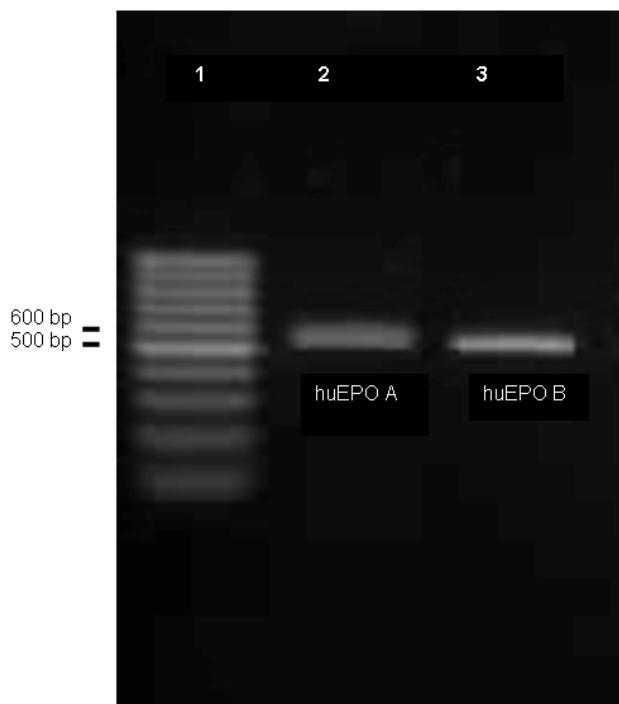
### Design and assembly of the synthetic gene

In principles, there are four steps in the construction of synthetic huEPO gene by PCR amplified single-step gene assembly: oligonucleotides design/synthesis, gene assembly, gene amplification and cloning. Since single-stranded ends of complementary DNA fragments are filled-in during the gene assembly process, cycling with DNA polymerase results in the formation of increasingly larger

DNA fragments until the full-length gene is obtained, without the need of DNA ligase.

The use of overlapping oligonucleotides necessitated the design of the oligonucleotides used for the synthesis of the huEPO gene great attention to detail, owing to the requirement for a large number to be mixed in one PCR. For these reasons the designed oligonucleotides would be screened and matched in order to meet several criteria as followed (Stemmer et al. 1995): (i) elimination of palindromic sequences, (ii) minimization of tandem or inverted repeats (< 10bp in length), (iii) optimization of the short region overlap between each primer and (iv) allowing subsequent use of the primers for DNA sequencing. Based on the deposited huEPO gene in the GenBank, two sets of twenty overlapping oligonucleotides (Table 1) were designed to give two constructs: (i) huEPO A gene that encode the 166 amino acids and (ii) huEPO B gene that encode the 193 amino acid residues. Two restriction enzyme recognition sites, namely *Eco*RI and *Avr*II, were introduced at both ends of the constructed huEPO gene to allow the directional cloning of the gene into *Pichia pastoris* expression vector (pPIC9K). In this study, the annealing temperatures for the oligonucleotides were within the range of 55°C to 75°C to facilitate optimal PCR.

The gene assembly reaction involved the construction of the full length huEPO gene from a stoichiometric mixture of overlapping oligonucleotides (Figure 1). The assembly process took the advantages of complementary overlapping regions between the sense and anti-sense strands. The



**Figure 3. Colony PCR of the white transformed colonies resolved on 1.5% agarose gel electrophoresis.** The presence of the synthetic huEPO genes was detected with ethidium bromide staining. Lane 1: 100bp DNA ladder, lane 2: huEPO A gene and lane 3: huEPO B gene.

presence of DNA polymerase and dNTPs in the PCR cycles sealed the gaps between the two strands. An aliquot of this gene assembly reaction was then used as a template for the gene amplification process, together with the outermost 5'-end and 3'-end primers. Analysis of the amplification products on 1.5% agarose gels revealed the presence of the expected DNA bands at ~500bp and ~600bp (Figure 2). There were no significant bands present in the gene assembly reaction that appeared as a 'tail', probably due to mismatching and mispairing of oligonucleotides. In addition, primer-dimer formations were detected on the agarose gel in the gene amplification reaction owing to the high concentration of oligonucleotides mixed in the assembly reaction. The DNA bands were excised from agarose gel and purified using a standard purification kit.

### Screening of huEPO clones

Transformed white colonies were screened for the presence of huEPO gene by colony PCR and resolved on 1.5% agarose gel (Figure 3). Two DNA bands, ~500bp and ~600bp, were detected by ethidium bromide staining that represent huEPO A gene and huEPO B gene, respectively.

### DNA sequencing

Plasmids from the transformed white colonies were extracted, purified and sent for DNA sequencing. The

sequencing process was carried out using the outermost 5'-end and 3'-end primers.

DNA sequencing results show colonies carrying the correctly assembled huEPO gene albeit with minor base mutations (Table 2). For huEPO A gene, mutations involved base substitution and base insertion. Substitution of  $A_6G_7$  to  $G_6A_7$  eliminated the *Eco*RI restriction site at the 5'-end of huEPO A gene. This was further verified by restriction enzyme analysis where *Eco*RI digestion no longer digested the plasmid at that particular site (data not shown). Other substitutions changed the amino acid codes:  $CA_{210}G$  (Gln) to  $CT_{210}G$  (Leu),  $CTG_{217}$  (Leu) to  $CTC_{217}$  (Leu) and  $AA_{264}C$  (Asn) to  $AC_{264}C$  (Thr). There was a base insertion (**G**) after base  $G_{175}$ . There were 1.1% mutations which was very low.

Almost the same DNA sequencing results were observed for huEPO B gene (Table 3). The percent of base mutation was low at 1.5% contributed by base substitution and base insertion. Similar to huEPO A gene, base substitution for huEPO B gene saw the changes in the amino acid codes. Base substitutions involved:  $TTG_{144}$  (Leu) to  $TTC_{144}$  (Phe),  $G_{157}CC$  (Ala) to  $A_{157}CC$  (Thr),  $G_{241}C_{242}C_{243}$  (Ala) to  $C_{241}G_{242}G_{243}$  (Arg),  $CA_{287}G$  (Gln) to  $CT_{287}G$  (Leu),  $CTG_{294}$  (Leu) to  $CTC_{294}$  (Leu) and  $AA_{341}C$  (Asn) to  $AC_{341}C$  (Thr). There was a base insertion (**T**) after base  $T_{53}$ , base insertion (**C**) after base  $C_{121}$  and base insertion (**G**) after base  $G_{300}$ .

### DISCUSSION

Molecular applications, such as DNA synthesis, gene expression and *in vitro* mutagenesis are the beneficiaries for oligonucleotides assembly technology. It gives advantages to researchers to construct synthetic genes and modify the genes from several factors that might affect gene expression, *i.e.*, codon usage and formation of unwanted secondary structure. The procedure is PCR-based single-step gene assembly comprises two main steps that are gene assembly and gene amplification. In the gene assembly step, proofreading DNA polymerase is used to build long DNA fragment from a pool of overlapping complementary oligonucleotides. This is followed by gene amplification step where the long DNA fragment would now serve as the template for the amplification process, using the same outer most 5'-end and 3'-end primers.

Gene synthesis is an advanced gene tool which gives researchers advantages to construct any synthetic gene using a set of large numbers of oligonucleotides. This tool allows the synthetic gene that encodes the same product as the natural gene might have variations in nucleotide sequence (Richardson et al. 2006). This would be applicable to the theory of gene design when high expression levels are desired is relatively uncomplicated. This strategy would allow codon usage to be optimized for the host organism or changed completely to accommodate a variety of constraints (Richardson et al. 2006).

In this work, the synthetic huEPO gene was designed according to the codon preference of *Pichia pastoris*, a methylotrophic yeast that has the potential to produce high level of active eukaryotic recombinant protein. The assembly step of huEPO gene was impressive such that very low percent of base mutations were detected after gene amplification step, despite random oligonucleotides being pooled together in a single reaction. The use of proofreading *Pfu* DNA polymerase certainly contributed to this phenomenon by ensuring both processivity and fidelity of the enzyme for gene assembly and amplification. The oligonucleotides used in this work were in the range of 40-mer compared to some other reported PCR-based gene synthesis methods, where oligonucleotides used were in the range of 80-mer (Stemmer et al. 1995; Zhang and Henderson, 1998). Using longer oligonucleotides has certain advantages, including the simplified gene design is and a smaller number of oligonucleotides required for gene assembly reaction (Withers-Martinez et al. 1999). Nevertheless, utilization of longer oligonucleotides contributes to low stability and specificity, tendency to form secondary structure, and the higher cost to synthesize longer oligonucleotides. Thus, using shorter oligonucleotides helps to avoid secondary structure formation and lowers the number of errors introduced during the assembly process. Since the frequency of PCR-derived errors increases with the increasing number of amplification cycles (Withers-Martinez et al. 1999), the gene amplification process was limited to twenty cycles.

The use of proofreading *Pfu* DNA polymerase did not give mutation-free synthetic huEPO gene. In fact, the presence of mutations in oligonucleotides assembly is unavoidable (Stemmer et al. 1995). The mutations were distributed randomly; suggesting that oligonucleotides were not the source of these errors and neither was the proofreading *Pfu* DNA polymerase was at fault for such mutations. Rather, these errors were most likely introduced during the gene assembly step. The lack of distinctive DNA band (Figure 2) after the gene assembly step might be indicative to the mutations due to mismatched and mispaired of oligonucleotides. Hence, it would be almost impossible to eliminate the errors in such a random assembly reaction. However, the mutations in the synthetic gene could be reduced significantly by optimizations as described above.

To conclude, we have successfully constructed and synthesized huEPO gene from the sequence information derived from GenBank using the PCR-based single-step gene assembly and amplification. Theoretically, there would be no practical limit on the number of oligonucleotides which may be mixed together, suggesting that bigger genes could also be successfully synthesized. Therefore, this would open vast possibilities to construct any gene as long as the codes for gene product could be identified. Our work on correcting the mutated bases is currently underway before sub-cloning into a *Pichia pastoris* expression vector for recombinant huEPO production.

## REFERENCES

- ARCASOY, Murat O. The non-haematopoietic biological effects of erythropoietin. *British Journal of Haematology*, April 2008, vol. 141, no. 1, p. 14-31.
- BAHLMANN, Ferdinand H.; DE GROOT, Kirsten; SPANDAU, Jens-Michael; LANDRY, Aimee L.; HERTEL, Barbara; DUCKERT, Thorsten; BOEHM, Sascha M.; MENNE, Jan; HALLER, Hermann and FLISER, Danilo. Erythropoietin regulates endothelial progenitor cells. *Blood*, February 2004, vol. 103, no. 3, p. 921-926.
- COINTE, Didier; BÉLIARD, Roland; JORIEUX, Sylvie; LEROY, Yves; GLACET, Arnaud; VERBERT, André; BOUREL, Dominique and CHIRAT, Frédéric. Unusual N-glycosylation of recombinant human erythropoietin expressed in human lymphoblastoid cell line does not alter its biological properties. *Glycobiology*, 2000, vol. 10, no. 5, p. 511-519.
- DAME, Christof; FAHNENSTICH, Hubert; FREITAG, Patricia; HOFMANN, Dietmar; ABDUL-NOUR, Thair; BARTMANN, Peter and FANDREY, Joachim. Erythropoietin mRNA expression in human fetal and neonatal tissue. *Blood*, November 1998, vol. 92, no. 9, p. 3218-3225.
- ELBIR, H.; ABDEL-MUHSIN, A.M. and BABIKER A. A one-step DNA PCR-Based method for the detection of *mycobacterium tuberculosis* complex grown on Lowenstein-Jensen media. *The American Journal of Tropical Medicine and Hygiene*, February 2008, vol. 78, no. 2, p. 316-317.
- HOOVER, David M. and LUBKNOWSKI, Jacek. DNworks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Research*, May 2002, vol. 30, no. 10, e43.
- INOUE, Noboru; TAKEUCHI, Makoto; OHASHI, Hideya and SUZUKI, Takamoto. The production of recombinant human erythropoietin. *Biotechnology Annual Review*, 1995, vol. 1, no. 1, p. 297-313.
- JACOBS, Kenneth; SHOEMAKER, Charles; RUDERSDORF, Richard; NEILL, Suzanne D.; KAUFMAN, Randal J.; MUFSON, Allan; SEEHRA, Jasbir; JONES, Simon S.; HEWICK, Rodney; FRITSCH Edward F.; KAWAKITA, Makoto; SHIMIZU, Tomoe and MIYAKE, Takaji. Isolation and characterization of genomic and cDNA clones of human erythropoietin. *Nature*, February 1985, vol. 313, no. 6005, p. 806-810.
- JELKMANN, W. Use of recombinant human erythropoietin as an antianemic and performance enhancing drug. *Current Pharmaceutical Biotechnology*, 2000, vol. 1, p. 11-31.

LIN, Fu-Kuen; SUGGS, Sidney; LIN, Chi-Hwei; BROWNE, Jeffrey K.; SMALLING, Ralph; EGRIE, Joan C.; CHEN, Kenneth K.; FOX, Gary M.; MARTIN, Frank; STABINSKY, Zippora; BADRAWI, Sayed M.; LAI, Por-Hsiung and GOLDWASSER, Eugene. Cloning and expression of the human erythropoietin gene. *Proceedings of the National Academy of Sciences of the United States of America*, November 1985, vol. 82, no. 22, p. 7580-7584.

MEHRNEJAD, Faramarz; NADERI-MANESH, Hossein; RANJBAR, Bijan; MAROUFI, Bahman; ASOODEH, Ahmad and DOUSTDAR, Farahnoosh. PCR-based gene synthesis, molecular cloning, high level expression, purification, and characterization of novel antimicrobial peptide, brevinin-2R, in *escherichia coli*. *Applied Biochemistry and Biotechnology*, May 2008, vol. 149, no. 2, p. 109-118.

MIYAKE, Takaji; KUNG, Charles K. and GOLDWASSER, Eugene. Purification of human erythropoietin. *Journal of Biological Chemistry*, August 1977, vol. 252, no. 15, p. 5558-5564.

MOORE, David D. Gene Synthesis: Assembly of target sequences using mutually priming long oligonucleotides. In: *Current Protocols in Molecular Biology*. New Jersey; John Wiley & Sons, May 2001, Chapter 8, Unit 8.2B.

RICHARDSON, Sarah M.; WHEELAN, Sarah J.; YARRINGTON Robert M. and BOEKER, Jef D. Gene design: Rapid, automated designof multi kilobase synthetic genes.*Genome Research*, February 2006, vol. 16, no. 4, p. 550-556.

ROUILLARD, Jean-Marie; LEE, Woonghee, TRUAN, Gilles; GAO, Xiaolian; ZHOU, Xiaochuan and GULARI, Erdogan. Gene2Oligo: Oligonucleotide design for in vitro gene synthesis. *Nucleic Acids Research*, March 2004, vol. 32, Web Server issue, p. W176-W180.

RYDZANICZ, Roman; ZHAO, Sharon X. and JOHNSON Philip E. Assembly PCR oligo maker: A tool for designing oligodeoxynucleotides for constructing long DNA molecules for RNA production. *Nucleic Acids Research*, July 2005, vol. 33, Web Server issue, p. W521-W525.

SASAKI, Ryuzo. Pleiotropic functions of erythropoietin. *Internal Medicine*, February 2003, vol. 42, no. 2, p. 142-149.

STEMMER, Willem P.C.; CRAMERI, Andreas; HA, Kim D.; BRENNAN, Thomas M. and HEYNEKER, Herbert L. Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene*, October 1995, vol. 164, no. 1, p. 49-53.

WITHERS-MARTINEZ, Chrislaine; CARPENTER, Elisabeth P.; HACKETT, Fiona; ELY, Barry; SAJID, Mohammed; GRAINGER, Muni and BLACKMAN,

Michael, J. PCR-based gene synthesis as an efficient approach for expression of the A+T-Rich malaria genome. *Protein Engineering*, December 1999, vol. 12, no. 12, p. 1113-1120.

XIONG, Ai-Sheng; YAO, Quan-Hong; PENG, Ri-He; DUAN, Hui; LI, Xian; FAN, Hui-Qin; CHENG, Zong-Ming and LI, Yi. PCR-based accurate synthesis of long DNA sequences. *Nature Protocols*, July 2006, vol. 1, no. 2, p. 791-797.

ZHANG, Lin-Hua and HENDERSON, Howard E. Long-distance PCR-based strategy for preparing knock-in vectors directly from ES cell genomic DNA. *BioTechniques*, November 1998, vol. 25, no. 5, p. 784-788.

## APPENDIX

### TABLES

**Table 1.** Twenty overlapping oligonucleotides utilized to construct the synthetic huEPO gene. The overlapping complementary sequences are underlined.

- (a) Oligonucleotide sets for synthetic huEPO gene A - 531 bases  
 (b) Oligonucleotide sets for synthetic huEPO gene B - 609 bases.

a) huEPO gene

Oligo	Sequence
F-1A	5'-AAAAAAAGAATTCAT <u>GGCCCCACCACGCCTCATCTGTGACAGCCGAGTC</u> -3'
R-1A	3'- <u>CGGGGTGGTGC</u> GGAGTAGACACTGT <u>CGGCTCAGGACCTCTCCATG</u> -5'
F-2A	5'-AGGTACCT <u>CTTGAGGCCAAGGAGGCCGAGAATATCACGACGGC</u> -3'
R-2A	3'- <u>AACCTCCGGTTCCCTCCGGCTCTTATAGTGCTGCCGACACGACTT</u> -5'
F-3A	5'-GAACACTGC <u>AGCTTGAATGAGAATATCACTGTCCCAGACACCAAA</u> -3'
R-3A	3'- <u>TCGAACTTACTCTTATAGTGACAGGGTCTGTGGTTCAATTAAAGATA</u> -5'
F-4A	5'-TTCTATGC <u>CTGGAAGAGGATGGAGGTGGGCAGCAGGCCGTAGAAGTC</u> -3'
R-4A	3'- <u>TTCTCCTACCTCCAGCCGTCGTCCGGCATCTCAGACCGTCCCGAC</u> -5'
F-5A	5'- <u>GGCCTGGCCCTGCTGCGAACGCTGTCCCTGCGGGGCCAGGCCCTGTTG</u> -3'
R-5A	3'- <u>GACGACAGCCTCGACAGGACGCCCGGTCCGGACAACCAGTTGAGAAGG</u> -5'
F-6A	5'- <u>TCCCAGCCGTGGAGCCCCTGCAGCTGCATGTGGATAAAGCCGTAGTGGC</u> -3'
R-6A	3'- <u>ACCTCGGGGACGTCGACGTACACCTATTCGGCAGTCACCGGAAGCGTCG</u> -5'
F-7A	5'- <u>AGCCTCACCACTCTGCTTGGCTCTGGAGGCCAGAAGGAAGCC</u> -3'
R-7A	3'- <u>TGGTGAGACGAAGCCCAGACCCCTGGGTCTTCCTCGGTAGAGGGGA</u> -5'
F-8A	5'- <u>CCTCCAGATGCGGCCTCAGCTGCTCCACTCGAACAAATCACTGCT</u> -3'
R-8A	3'- <u>CGCCGGAGTCGACGAGGTGAGGTTAGTGACGACTGTGAAAG</u> -5'
F-9A	5'- <u>TTCCGCAAACCTTCCGAGTCTACTCCAATTCCCGGGAAAG</u> -3'
R-9A	3'- <u>GAGAAGGCTCAGATGAGGTTAAAGGAGGCCCTTCGACTTCGAC</u> -5'
F-10A	5'- <u>AAGCTGTACACAGGGAGGCCTGCAGGACAGGGACAGATAATT</u> -3'
R-10A	3'- <u>CCCTCCGGACGTCTGTCCCCTGTCTATTAAAGGATCCAAAAAA</u> -5'

b) huEPO B

Oligo	Sequence
F-1B	5'-AAAAAAGAATT <u>CATGGGGTGCACGAATGTCCTGCCTGG</u> -3'
R-1B	3'- <u>TACCCCCACGTGCTTACAGGACGGACCGAACCGAAGAGGACAGGGACGACAGC</u> -5'
F-2B R-2B	5'-CTTCTCCTGTCCCTGCTGTCGCTCCCT <u>CTGGCCTCCCAGTCCTGGCGCC</u> -3' 3'- <u>CCGGAGGGTGACGACCCGCGGGTGGTGC GGAGTAGACACTGTCGGCTCAGGAC</u> -5'
F-3B	5'-CTCATCTGTGACAGCCGAGTCCTGGAGAGGTACCT <u>CTTGAGGCCAAGGAGGCC</u> -3'
R-3B	3'- <u>GAGAACGCTCCGGTCCCG</u> CTCTTATAGTGC <u>TGCCCCACACGACTTGTGACGTTA</u> -5'
F-4B	5'-ACGACGGGCTGTGCTGAACACTGCAGCTGAAT <u>GAGAACATACGTCCC</u> A-3'
R-4B	3'- <u>CTCTTATAGTGACAGGGTCTGTGGTTCAATTAAAGATAGCCACCTCTCCCCC</u> -5'
F-5B	5'-GTTAATTCTATGCCTGGAAGAGGGATGGAGGT <u>CGGGCAGCACGCCGTAGAAC</u> TC-3'
R-5B	3'- <u>GTCGTCCGGCATCTTCAGACCGACCCGGAGCAGCACAGCCTTCGA</u> -5'
F-6B	5'-GGCCTGGCCCTGCTGCGGAAGCTGTC <u>CTGCGGGCCAGGCCCTGTTG</u> -3'
R-6B	3'- <u>GCCCCGGTCCGGACAACCAGTGGAGAAGGGTGGCACCCCTGGGACGTCGAC</u> -5'
F-7B	5'-CAGCCGTGGAGCCCCTGCAGCTGCATGTGG <u>AATAAGCCGTCAGTGGC</u> TT-3'
R-7B	3'- <u>CTATTCGGCAGTCACCGGAAGCGTCGGAGTGGAGACGAAGCCGAGAC</u> -5'
F-8B	5'-CTCACCACTCTGCTTGGCTCTGGAGCC <u>AAGGAAGCCATCTCCCTCCAGA</u> T-3'
R-8B	3'- <u>TTCCTTCGGTAGAGGGAGGTCTACGCCGGAGTCGACGAGGTGAGGCTTGTAG</u> -5'
F-9B	5'-GCTGCTCCACTCCGAACAATCACTGCTGACACT <u>TTCCGCAAAC</u> TCTCCGAGTC-3'
R-9B	3'- <u>AAGGC</u> TTGAGAAGGCTCAGATGAGGTTAAAGGAGGCCCTTCGACTTCGAC-5'
F-10B	5'-CGGGGAAAGCTGAAGCTGTACAC <u>AGGGAGGC</u> TCAGGACAGGGACAGATAA-3'
R-10B	3'- <u>CCCCTCCGGACGTCC</u> GTCCCCTGTCTATTAAAGGATCCAAA AAA-5'

**Table 2. DNA sequence alignment result showing base substitution and base insertion in synthetic huEPO A gene.** The underlined bases are the bases that were substituted with the bases underneath them. The dash (-) indicates the position of the inserted base with the base underneath it.

Location (bp)	Alignment analysis
6, 7	AAAAAA <u>G</u> AATTTC
	GA
175	AGAGG-ATGGA
	G
210, 217	CTGGC <u>A</u> GGGC <u>T</u> GGCCCT
	T C
264	GGTCA <u>A</u> CTCTT
	C

**Table 3. DNA sequence alignment result showing base substitution and base insertion in synthetic huEPO B gene.** The underlined bases are the bases that were substituted with the bases underneath them. The dash (-) indicates the position of the inserted base with the base underneath it.

Location (bp)	Alignment analysis
53	TCTCCT-GTCCCT
	T
121	CC-GAGTCC
	C
144, 157	TACCTCTT <u>GG</u> AGGCCAAGGAG <u>GG</u> CGAGAAT
	C A
241-243	AT <u>GC</u> CTGGAAG
	CGG
287, 294	GTCTGGC <u>A</u> GGGC <u>T</u> GGCCC
	T C
300	TG-CTGTCG
	G
341	TTGGTCA <u>A</u> CTCTTCC
	C