Front. Med. 2015, 9(1): 90–99
DOI 10.1007/s11684-015-0390-2

RESEARCH ARTICLE

Optimized human factor IX expression cassettes for hepatic-directed gene therapy of hemophilia B

Ru Zhang, Qiang Wang, Lin Zhang, Saijuan Chen (✉)

State Key Laboratory for Medical Genomics, Shanghai Institute of Hematology, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

Abstract **Gene therapy provides a potential cure for hemophilia B, and significant progress has been achieved in liver-directed gene transfer mediated by adeno-associated viral vectors. Recent clinical trials involving the use of a self-complementary adeno-associated virus serotype 8-human codon-optimized factor IX (AAV8-hFIXco) vector demonstrated encouraging efficacy with hFIX expression stabilized at 1% to 6% of normal level in patients, but safety concerns related to high vector doses are still present. Thus, further improvement of AAV vectors and hFIX expression cassette may positively contribute to the ultimate success of hemophilia B gene therapy. In this study, to obtain a higher expression level of hFIX that potentiates the coagulant capacity of recipients, human *FIX* expression vector was optimized by upgrading the codon adaption index and adjusting the GC content, inserting a Kozak sequence (GCCACC), and introducing a gain-of-function mutation, R338L (FIX Padua). The efficiency of the published and the presently constructed cassettes was compared through *in vivo* screening. In addition, the regulatory elements that control the *FIX* gene expression in these cassettes were screened for liver-specific effectiveness. Among all the constructed cassettes, scAAV-Pre-hFIXco-SIH-R338L, which was the construct under the control of the prothrombin enhancer and prealbumin promoter, resulted in the highest level of coagulant activity, and the expression levels of two constructed cassettes (scAAV-Chi-hFIXco-SIH-R338L and scAAV-Pre-hFIXco-SIH-R338L) were also higher than that of the published cassette (scAAV-LP1-hFIXco-SJ). In summary, our strategies led to a substantial increase in hFIX expression at the protein level or a remarkably elevated coagulant activity. Thus, these reconstructs of h*FIX* with AAV vector may potentially contribute to the creation of an efficacious gene therapy of hemophilia B.**

Keywords factor IX; hemophilia B; liver-specific regulatory elements; hydrodynamic gene transfer

Introduction

Hemophilia B is an X-linked, monogenic hereditary disorder. It is characterized by a deficiency of blood coagulation factor IX (FIX), resulting in a diminished clotting capability of blood. It occurs in 1 of every 25 000 to 30 000 live males who were born with the most common mutation on chromosome Xq27.1. The disease has been broadly classified as mild, moderate, or severe according to the level of FIX coagulant activity, which corresponds to 5% to 40%, 1% to 5% and < 1% of normal, respectively [1].

The major clinical feature of hemophilia B is continuous bleeding after or even without trauma or injury. Bleeding can occur in various tissue, particularly in the joints. Without effective treatment, soft-tissue hematomas and hemarthrosis could lead to tissue damage and chronic hemophilic arthropathy over time. In addition, bleeding in critically closed spaces can be life-threatening. Current treatment of hemophilia B relies on protein replacement therapy (PRT). PRT restores hemostasis, thereby reducing the mortality of patients [2,3]. However, because of the short half-life of the protein in blood circulation, PRT only has transient therapeutic effects. Requiring frequently periodic injections, the treatment of PRT is expensive and is also associated with complications, such as infectious diseases. Recent clinical trials of liver-directed gene therapy for HB have been performed at St Jude Children's Research Hospital and University College, London. It involved self-complementary AAV8-FIX

Received November 7, 2014; accepted November 17, 2014
Correspondence: sjchen@stn.sh.cn

vectors and demonstrated encouraging results with relative safety [4]. Nevertheless, further investigations are needed to increase FIX coagulant activity, reduce vector doses, and eliminate host immune responses.

Many reported studies have demonstrated that the expression level of a gene could be significantly improved by codon optimization [5–7]. Ward reported that codon optimization of human factor VIII cDNA led to an unprecedented 29- to 44-fold increase in factor VIII expression [8]. Identical doses of AAV8 vector containing codon-optimized hFVIII cDNA resulted in dramatic improvements in clotting time, with average aPTTs declining from 78 ± 8 s to 37 ± 2 s [9]. Additionally, hFIX expression of scAAV2/8-LP1-hFIXco was almost 4-fold higher compared with that of scAAV2/8-LP1-hFIX by tail vein administration of 2×10^9 vector particles [5]. However, the integrated application of codon optimization technology based on upgraded codon adaption index (CAI), increased GC content, and removal of repeated sequences may further improve expression of hFIX without altering the amino acid sequence. Hydrodynamic administration represents a promising approach to study gene expression in liver and has become a common method for DNA delivery in rodents [10–12]. Recently, hydrodynamic tail vein injection has been applied to evaluate the control elements that drive reporter gene expression in the liver [13–15]. In the present work, therapeutic cassettes were constructed to direct liver-specific gene expression. The codons for human *FIX* were optimized, and R338L mutation was introduced [16] to increase the expression and activity of FIX. Then, the best regulatory sequences from liver-specific expressed proteins including albumin [17–19], prealbumin [20–22], α1-antitrypsin [23,24], and prothrombin [25] were screened for their ability to drive the transgene expression. These regulatory sequences were used in hFIX expression cassettes and were evaluated in HB mice via hydrodynamic delivery method.

Materials and methods

Construction of human FIX plasmids

Diagrams of the plasmid constructs are presented in Fig. 1A and 1B. The full sequence of scAAV-LP1-hFIXco [5] was synthesized and sequenced prior to cloning into pUC57-Kan to create scAAV-LP1-hFIXco-SJ. A new hFIX expression cassette that included a codon-optimized h*FIX* (by OptimumGene™ codon optimization analysis) was sub-cloned into pUC57-Kan to construct scAAV-Alb-hFIXco-SIH. The full sequence of hFIXco-SIH was outlined in Supplementary Material. pcDNA 3.1(-)-FIXwt was a gift from Wenbin Wang (Shanghai Institute of Hematology, Shanghai, China). The h*FIX* cDNA was obtained from pcDNA 3.1(-)-FIXwt by polymerase chain reaction (PCR) amplification and sub-cloned into scAAV-polyadenylation signal (polyA) at the *Eco*RI/*Bgl*II sites. scAAV-hFIX-polyA was digested with *Eco*RI and *Spe*I to release h*FIX* and the SV40 late polyA tail. This fragment was then inserted into scAAV-LP1-hFIXco-SJ at *Eco*RI and *Spe*I to substitute hFIXco-SJ and the SV40 late polyA to construct scAAV-LP1-hFIX. The scAAV-LP1-hFIXco-SIH plasmid was constructed by digesting scAAV-LP1-hFIX at *Eco*RI and *Bgl*II. scAAV-Alb-hFIXco-SIH was digested at *Eco*RI and *Bgl*II to release the h*FIX* fragment, which was subsequently inserted into scAAV-LP1-hFIX.

The hFIXco-SJ R338L (AGG→CTA) mutation [16] and the hFIXco-SIH R338L (CGG→CTA) mutation were introduced using the QuickChange II-XL site-directed mutagenesis kit (Stratagene) to generate scAAV-LP1-hFIXco-SJ-R338L, scAAV-LP1-hFIXco-SIH-R338L, and scAAV-Alb-hFIXco-SIH-R338L.

The scAAV-Pre-hFIXco-SIH-R338L construct was cloned by replacing the albumin enhancer and promoter in scAAV-Alb-hFIXco-SIH-R338L with the human prothrombin enhancer and prealbumin promoter at the *Xho*I/*Hin*dIII sites. scAAV-Chi-hFIXco-SIH-R338L was constructed similarly by replacing with the chimeric albumin enhancer and the chimeric promoter. Elements of all constructed plasmids are listed in Table 1.

Animal procedures

All mice were housed in a specific pathogen-free environment with a normal diet. All experimental procedures were approved by the Studies Ethics Committee of Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine (Reference number: SYXK2011-0113). All animals were handled in accordance with Institutional Animal Care and Use Committee guidelines. Animals used in this study had a C57BL/6 genetic background. FIX-deficient (FIX^{null}) mice were obtained from Dr. Guowei Zhang (Hangzhou Normal University) [26]. Separated PCR reactions were performed to identify hemophilia-affected (−/− or −/Y), carrier (+/−), and wild type (+/+ or +/Y) animals. Genomic DNA was initially prepared with blood collected from retro-orbital bleeding and was applied to standard PCR reactions. A common reverse primer (AACAGGGATAGTAAGATTGTTCC) was applied in combination with either a mutant forward primer (TCCTGTCATCTCACCTTGCTC) or a wild type forward primer (TGGAAGCAGTATGTTGGTAAGC) to identify the mutant and wild type alleles, respectively. All primers were synthesized from Genewiz (Suzhou, China). A 35-cycle PCR program (94 °C for 30 s, 55 °C for 30 s, and 72 °C for 45 s) was performed with a 2 min elongation step at 72 °C upon completion of all cycles.

Hydrodynamic injections were performed with male HB

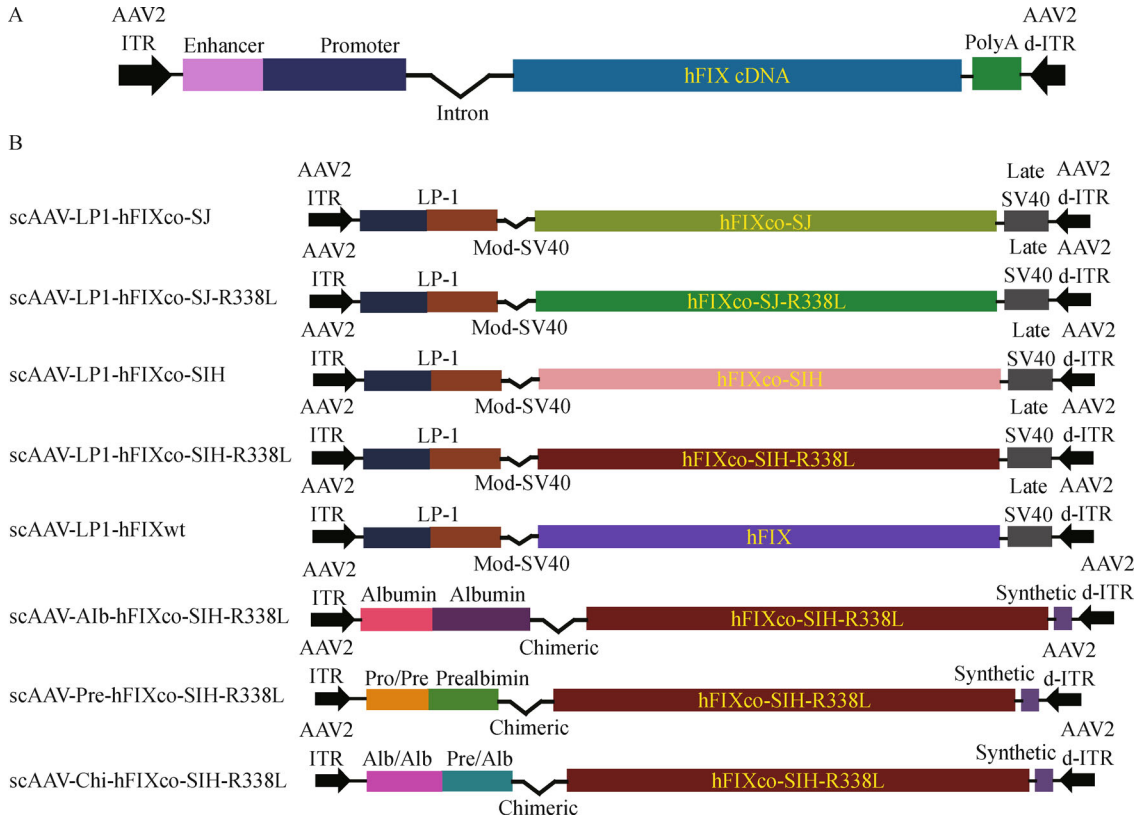

Fig. 1 Construction of hFIX expression cassettes. (A) Diagram of human FIX expression cassette: a liver-specific FIX expression cassette comprises *FIX* cDNA and liver-specific regulatory elements, including enhancer, promoter, intron, and polyA flanked by the AAV internal terminal repeats. The AAV internal terminal repeats were necessary for packaging the virion. (B) Each cassette had intact 5′ and 3′ ITRs. scAAV-LP1-hFIXco-SJ contained the LP1 promoter, which comprised core liver-specific elements from HCR and hAAT, modified SV40 intron, hFIXco-SJ, and late SV40 polyA [5]. scAAV-LP1-hFIXco-SJ-R338L, scAAV-LP1-hFIXco-SIH, scAAV-LP1-hFIXco-SIH-R338L, and scAAV-LP1-hFIXwt contained the same control elements as scAAV-LP1-hFIXco-SJ, and those cassettes had different h*FIX* as shown in the Figure above. scAAV-Alb-hFIXco-SIH-R338L, scAAV-Pre-hFIXco-SIH-R338L, and scAAV-Chi-hFIXco-SIH-R338L contained the same intron (Chimeric intron), polyA (Synthetic polyA), and hFIXco-SIH, which were driven by albumin enhancer (221 bp) and albumin promoter (292 bp), prothrombin/prealbumin enhancer (180 bp) and prealbumin promoter (216 bp), chimeric albumin enhancer (228 bp) and prealbumin/albumin promoter (215 bp), respectively.

Table 1 Structures of different FIX expression cassettes applied in this study

Plasmid name	Enhancer	Promoter	Intron	hFIX	PolyA
LP1-hFIXwt	ApoE-HCR	hAAT	Mod-SV40	hFIXwt	Late SV40
LP1-hFIXco-SJ	ApoE-HCR	hAAT	Mod-SV40	hFIXco-SJ	Late SV40
LP1-hFIXco-SIH	ApoE-HCR	hAAT	Mod-SV40	hFIXco-SIH	Late SV40
LP1-hFIXco-SJ-R338L	ApoE-HCR	hAAT	Mod-SV40	hFIXco-SJ-R338L	Late SV40
LP1-hFIXco-SIH-R338L	ApoE-HCR	hAAT	Mod-SV40	hFIXco-SIH-R338L	Late SV40
Alb-hFIXco-SIH-R338L	Alb	Alb	Chimeric	hFIXco-SIH-R338L	Synthetic
Pre-hFIXco-SIH-R338L	Pro/Pre[a]	Prealbumin	Chimeric	hFIXco-SIH-R338L	Synthetic
Chi-hFIXco-SIH-R338L	Alb/Alb[b]	Pre/Alb[c]	Chimeric	hFIXco-SIH-R338L	Synthetic

[a]Prothrombin/prealbumin enhancer.
[b]Chimeric albumin enhancer.
[c]Chimeric prealbumin/albumin promoter.

mice (−/Y), as described previously [27,28]. HB mice were injected with 120 μg of endotoxin-free plasmid DNA in 2.0 ml saline via the tail vein within 5 s to 10 s. Blood was collected at 24 h to 48 h after hydrodynamic injection by retro-orbital bleeding with sodium citrate as an anticoagulant at a final concentration of 0.38% (wt/vol).

The blood sample was then centrifuged at 4 °C for 10 min at 9000 rpm in a microcentrifuge. The plasma was collected and stored at − 80 °C for use in the FIX assays.

Determination of hFIX levels and coagulant activity

Plasma levels of hFIX in mice were detected by enzyme-linked immunosorbent assay (ELISA) according to a previously described method [29]. A monoclonal antibody against the hFIX heavy chain (AHIX-5041 monoclonal antibody, Hematologic Technologies Inc., US) was used to capture the protein. A goat anti-human FIX polyclonal antibody was used as the secondary antibody (GAFIX-APHRP, Affinity Biological Inc.). A standard curve was generated with serial dilutions of normal human pooled plasma. Human FIX coagulant activity in collected plasma samples were determined through a chromogenic assay (BIOPHEN Factor IX, Hyphen Biomed, Neuville-sur-Oise, France) and calculated with a standard curve generated with a series of diluted normal human pooled plasma.

Results

Codon optimization of h*FIX* cDNA

Codon optimization of *FIX* was performed with an OptimumGene™ algorithm in which codon adaptation index, codon usage frequency, and mRNA structure optimization were considered. To reach a higher level of expression, the codon usage bias in human was increased by upgrading the CAI from 0.72 to 0.87. The GC content and unfavorable peaks were modulated to prolong the half-life of FIX mRNA. The stem-loop structures affecting ribosomal binding efficiency and stability of mRNA were broken. In addition, the negative cis-acting sites were successfully modified by the optimization process. The information on codon optimization is shown in Fig. 2, including the CAI (Fig. 2A), GC contents (Fig. 2B), and the frequency of codon usage (Fig. 2C).

For this study, a previously described hFIX expression cassette (scAAV-LP1-hFIXco-SJ) was selected as a positive control [4,5]. Our codon-optimized h*FIX* and wild type h*FIX* were constructed under the same control elements in LP1-hFIXco-SIH and LP1-hFIX-WT plasmids, respectively. To test the expression mediated by these constructs *in vivo*, the LP1-hFIX-WT, LP1-hFIXco-SIH, and LP1-hFIXco-SJ plasmids were delivered into HB mice at a dose of 200 μg per mouse by hydrodynamic injection, and blood was collected 36 h after injection via retro-orbital eye bleeding.

Protein levels of hFIX in mouse plasma were detected by ELISA. As shown in Fig. 3, the transgene expression mediated by the hFIXco-SIH cassette was 3- to 4-fold higher than that by the wild type counterpart. Hydrodynamic administration of the LP1-hFIX-WT, LP1-hFIXco-SIH and LP1-hFIXco-SJ plasmids resulted in 47%, 162% and 157% of the normal hFIX protein individually, as compared with the pooled human plasma. In the control mice injected with saline, FIX protein was hardly detected. These data indicated that hFIX expression was increased at the protein level because of our codon optimization strategy.

Further increase of clotting activity in hFIXco-SIH-harbing HB mice by R338L mutation

To further enhance the clotting activity of hFIX to reach therapeutic effect with lower doses, a previously described hyper-functional R338L mutation was introduced into the hFIXco-SIH transgene. Human *FIX* with R338L mutation was evaluated in HB mice via hydrodynamic injection. The FIXco-SIH-R338L transgene displayed 6- to 8-fold coagulant activity compared with FIXco-SIH and the previously published FIXco-SJ (Fig. 4A and 4B), and such activity was also greater than that of FIXco-SJ-R338L. In addition, our codon-optimized transgene further promoted the protein expression to 120% of normal, whereas the positive control hFIXco-SJ and its R338L mutant transgenes resulted in about 109% of normal protein (Fig. 4C and 4D). Considering the coagulant efficiency, our codon-optimized h*FIX* transgene with an R388L mutation demonstrated the highest coagulant activity among the h*FIX* transgenes tested in HB mice.

Comparison of liver-specific hFIX expression cassettes driven by various regulatory elements

The transgene with the best performance, hFIXco-R338L, was selected for the construction of liver-specific expression cassettes. Each of the five plasmids (LP1-hFIXco-SJ, LP1-hFIXco-SIH-R338L, Alb-hFIXco-SIH-R338L, Chi-hFIXco-SIH-R338L, and Pre-hFIXco-SIH-R338L) was hydrodynamically delivered into mice using 120 μg of plasmid DNA at a final volume of 2 ml per mouse. Compared with LP1-hFIXco-SJ, which served as the positive control, the Pre-hFIXco-SIH-R338L cassette exhibited the highest level of hFIX clotting activity (Fig. 5). The coagulant activities of the hFIX protein of the 5 transgene cassettes were ranked as follows: LP1-hFIXco-SJ < Alb-hFIXco-SIH-R338L < Chi-hFIXco-SIH-R338L < LP1-hFIXco-SIH-R338L < Pre-hFIXco-SIH-R338L. In addition, higher protein expression levels were obtained from all newly constructed cassettes (except for Alb-hFIXco-SIH-R338L) compared with the control cassette.

Discussion

The expression efficiency of therapeutic transgenes applied

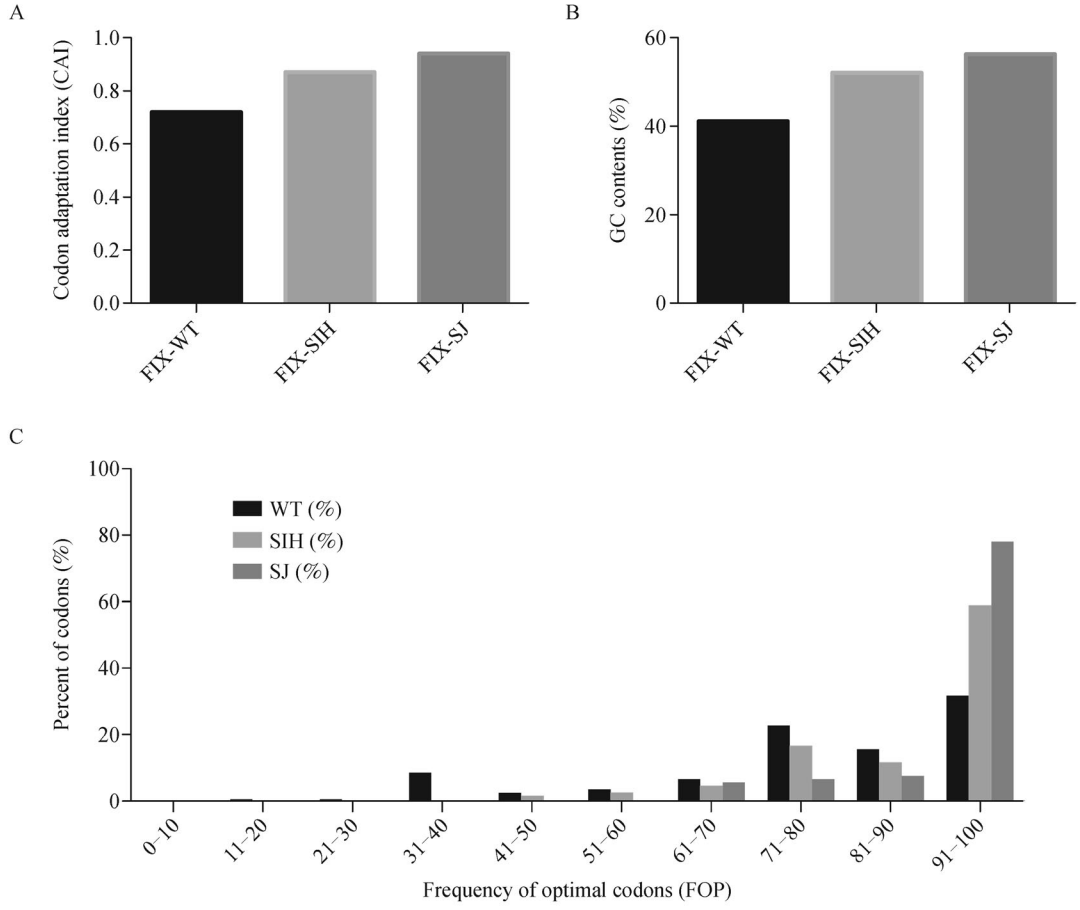

Fig. 2 Codon optimization of *FIX*: adjustment of the CAI, GC content and the frequency of codon usage for improving FIX expression. (A) A CAI of 1.0 was considered as perfect, whereas a CAI of > 0.8 was good, for high-level gene expression in the desired organism [30]. (B) The ideal range of GC content was between 30% and 70%. Peaks of GC content percentage in a 60 bp window were removed. (C) Frequency of optimal codons was the percentage distribution of codons in computed codon quality groups. The value of 100 was set for the codon with the highest usage frequency for a given amino acid in the desired expression organism.

in liver-directed gene therapy relies on not only the gene sequence, but also the regulatory elements, such as the enhancer, promoter, intron, and polyA. Previous studies demonstrated that a wide variety of factors influenced gene expression [15,31]. Ideally, a transgene cassette with high liver specificity and transcriptional activity should meet several requirements. First, it should have restricted expression only in the hepatocytes. Second, it could drive gene transcription efficiently to reach the therapeutic efficacy. Third, for the AAV vector, the size of the cassette should be sufficiently small for efficient packaging. Thus, considering the above requirements, all regulatory elements selected for driving transgene expression in this study were liver specific, small, and with high transcriptional activity.

To optimize the *FIX* transgene cassette, we made the following modifications: first, the h*FIX* coding sequence was manipulated under OptimumGene™ codon optimization analysis with high frequent codons and GC content adjustment; second, a hyper-functional FIX-R338L mutation was introduced to increase the coagulant activity of FIX; third, the strong regulatory elements was designed; and fourth, the Kozak sequence was added immediately in front of the start codon to modulate translation initiation efficiency [32].

Codon optimization is considered as an effective strategy to improve the gene expression level. It reportedly increases the steady-state mRNA level [33] and alleviates translation inefficiency caused by ribosomal pauses at rare codons [34,35]. A large number of software tools have been implemented to design synthetic genes for sequence optimization. The first generation of the tools focused

primarily on codon usage optimization and unique restriction site incorporation. In recent years, sequence design tools that aimed to construct optimal protein-coding sequences adhered to multiple objectives, such as codon bias, codon context bias, RNA secondary structure, ribosomal binding site, restriction sites, hidden stop codons, and other motif avoidance. However, the best kind of codon selection is still unknown. Thus, experimental validation is needed after the preliminary selection with software, and further refinement in the gene design software is required [36–39]. In this study, we synthesized the h*FIX*co cDNA according to the frequency of codon usage in humans by OptimumGene™ software. The experiments were then performed to compare the protein expression levels with or without codon optimization, namely, the levels of hFIXco-SIH and hFIX-WT. Results showed that the expression level of hFIXco-SIH was about 3.4 times that of hFIX-WT (162%/47%). In addition, the expression level of hFIXco-SIH was slightly higher than that of the control, hFIXco-SJ. The comparative results of hFIXco-SIH with or without the gain-of-function R338L mutation are consistent with a previous report, in which FIX-R338L had a specific activity that was 5 times to 10 times as high as the wild type *FIX* [16].

The promoters of genes with liver-specific expression have always been chosen for the construction of transgene cassettes [40]. However, such promoters were typically large, with low and short-lasting transcriptional activity [41,42]. These disadvantages have limited their applications in liver-targeted gene therapy. Previous studies on such promoters have shown that a cluster of cis-acting regulatory elements located immediately upstream of the transcription start site (TSS) were generally capable of directing the transcription of an already identified hepatocyte-specific gene [43]. Analysis of the regulatory components of albumin indicated that binding sites for hepatocyte nuclear factor 1 (HNF-1), NF-Y, and C/EBP were present at −107/−99, −90/−84, and −65/−49, respectively, upstream of the encoding sequence [17,18,44,45]. Four binding sites in prealbumin promoter are present for C/EBP or C/EBP-related factors, two sites for HNF-3, two sites for HNF-4, and one site for HNF-1 [20,22,46,47]. The 5′ proximal regulatory region of the human α1-antitrypsin (*α1-AT*) gene contained multiple or single binding sites for transcription factors, including C/EBP, HNF-3, and HNF-1 for liver-specific expression [23,24,48]. HNF-3 and HNF-4 were identified as crucial to prealbumin and α1-antirypsin (α1-AT) expression, but neither of them had binding sites in the albumin promoter proximal region [20]. The lack of binding sites for these two transcription factors might explain the mediocre hFIX expression and activity by the transgenes driven by the albumin promoter/enhancer elements (Fig. 5).

Based on the previous studies, transcription factor binding sites upstream of liver-specific genes were chosen to design regulatory sequences of high liver specificity and activity for expression cassettes [43,49]. We designed a new chimeric promoter composed of the distant regulatory region of prealbumin [22] and the proximal region of albumin [17]. This chimeric promoter contains four kinds of binding sites for critical transcription factors, including HNF-1, HNF-3, HNF-4, and C/EBP. Chi-hFIXco-SIH-R338L, under the control of the chimeric promoter and the core albumin enhancer elements, achieved a higher hFIX

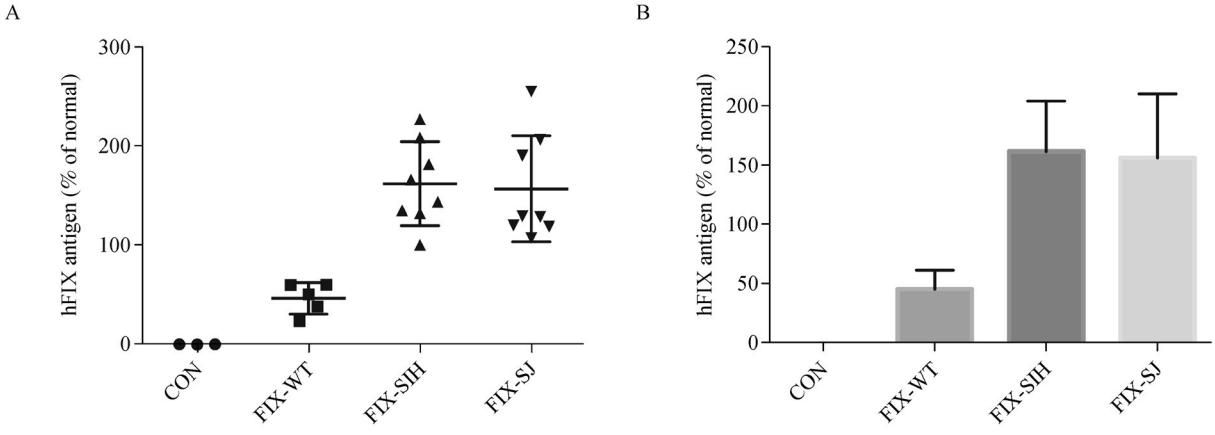

Fig. 3 Effects of codon optimization on hFIX expression under the control of the same elements. ELISA was used to detect plasma hFIX antigen at 36 h after tail-vein hydrodynamic administration of saline (CON, $n = 3$), scAAV-LP1-hFIX (FIX-WT, $n = 5$), scAAV-LP1-hFIXco-SIH (FIX-SIH, $n = 8$), and scAAV-LP1-hFIXco-SJ (FIX-SJ, $n = 8$). (A + B) hFIX antigen levels in the plasma of individual mice are indicated in (A) and presented as mean ± SD in (B). Saline-injected mice produced nearly undetectable hFIX antigen (0.2% of normal). FIX-WT, FIX-SIH, and FIX-SJ had a mean hFIX antigen levels of 47% ± 7% (2.35 ± 0.14 μg/ml), 162% ± 15% (8.1 ± 0.3 μg/ml), and 157% ± 19% (7.85 ± 0.95 μg/ml) of normal, respectively.

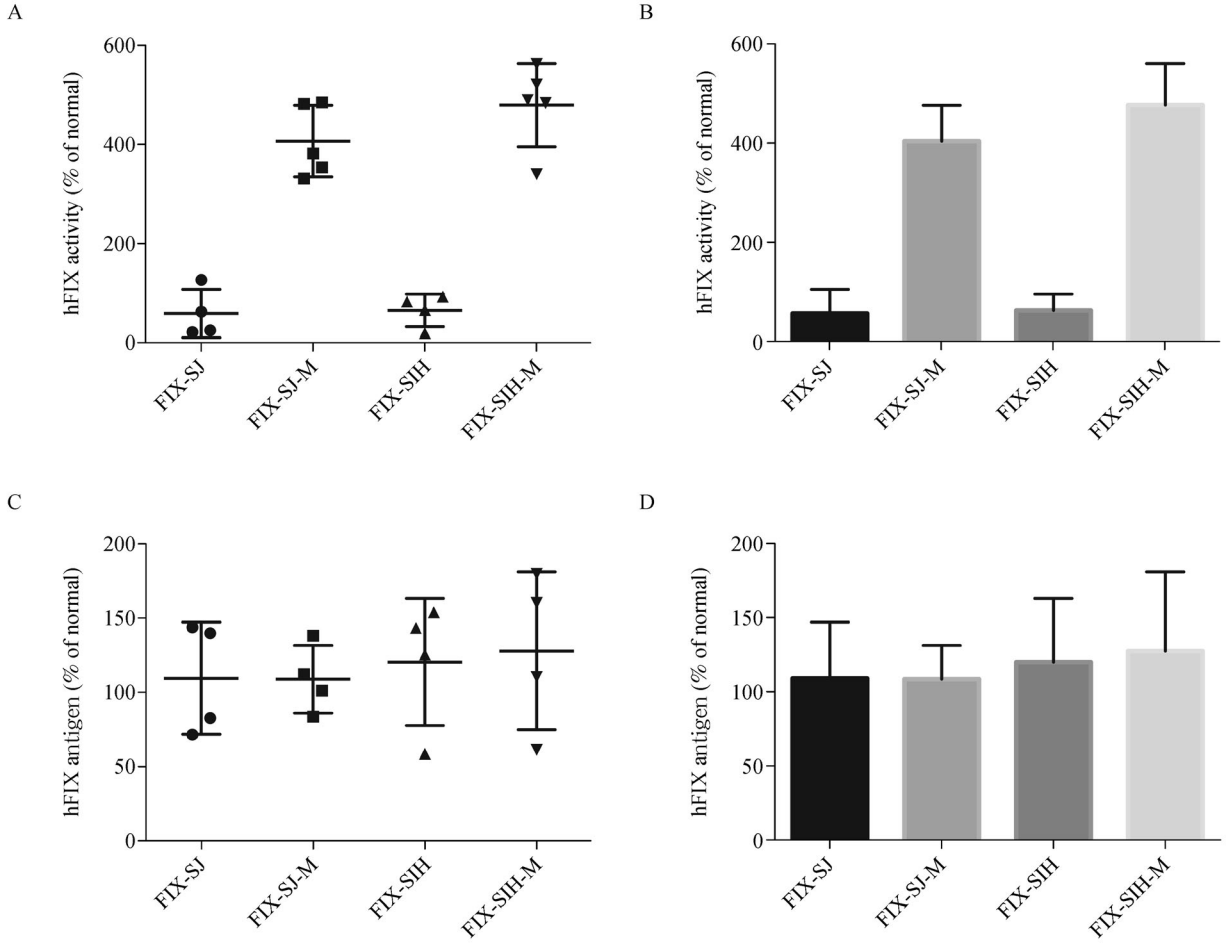

Fig. 4 Comparison of hFIX coagulant activity and expression level mediated by transgene cassettes in HB mice after hydrodynamic injection. FIX-SJ: scAAV-LP1-hFIXco-SJ; FIX-SJ-M: scAAV-LP1-hFIXco-SJ-R338L; FIX-SIH: scAAV-LP1-hFIXco-SIH; and FIX-SIH-M: scAAV-LP1-hFIXco-SIH-R338L. (A + B) Chromogenic assay to determine hFIX activity in individual mice in (A) is presented as mean ± SD in (B). The activity is calculated as a percentage of normal hFIX activity in the pooled plasma. The average hFIX activities mediated by FIX-SJ, FIX-SJ-M, FIX-SIH, and FIX-SIH-M were about 60% ± 24%, 406% ± 32%, 66% ± 16%, and 479% ± 37%, respectively. (C + D) ELISA results of hFIX expression in individual mice in (C) are presented as mean ± SD in (D). The expression of codon-optimized h*FIX* with or without R338L mutation showed no significant difference.

expression at protein level and coagulant activity. Notably, the hFIXco-SIH-R338L construct driven by the prothrombin/prealbumin enhancer and prealbumin promoter displayed the greatest coagulant activity of hFIX in HB mice with high hFIX protein level. Compared with scAAV-LP1-hFIXco-SJ, both scAAV-Chi-hFIXco-SIH-R338L and scAAV-Pre-hFIXco-SIH-R338L displayed higher levels of expression and hFIX coagulant activity. Thus, better therapeutic efficacy may be achieved with both cassettes above via AAV-mediated gene therapy in hemophilia B. However, excess level of FIX activity and AAV vector may lead to thrombosis and/or immune disease. Thus, we need further investigations to achieve ideal therapeutic efficacy and avoid side effects.

In summary, we have constructed a number of liver-specific hFIX expression cassettes. These cassettes have demonstrated relatively high hFIX expression and activity compared with an hFIX cassette that is currently in trial. This design is promising for HB gene therapy mediated by AAV vectors. In addition, regarding the recent progress in clinical application of non-viral constructs targeting the human liver, our configuration could be an optional therapeutic strategy that delivers naked DNA to the liver. Our data showed that the FIX expression level correlated well with the types and amounts of liver-specific transcription factor binding sites in the cassettes. In conclusion, this study provides information on the optimization of transgene cassettes for gene therapy.

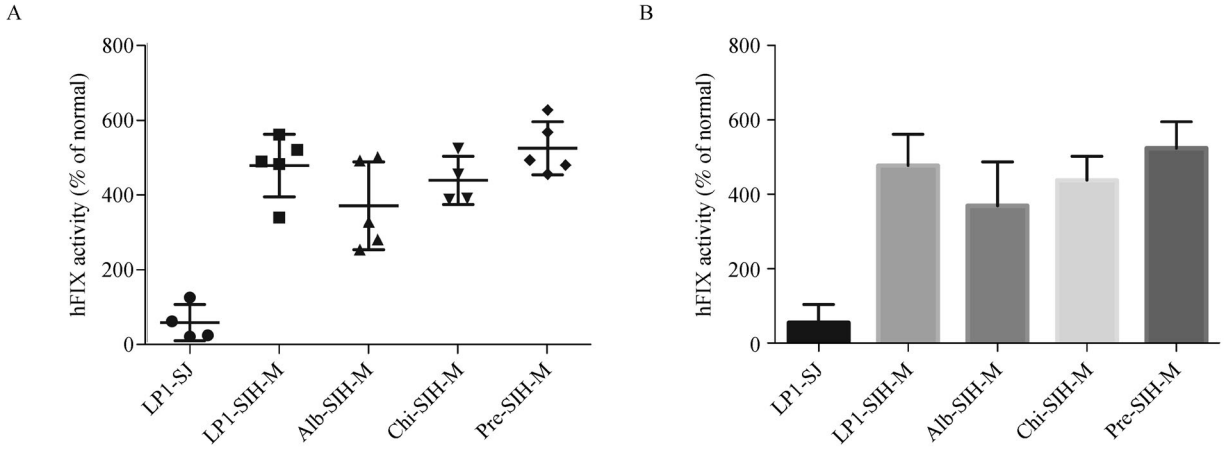

Fig. 5 Comparison of liver-specific regulatory elements driving hFIX expression *in vivo* after hydrodynamic delivery. (A + B) Chromogenic assay to determine hFIX coagulant activity, which was calculated as percentage of normal hFIX activity in pooled human plasma. The activities in individual mice are indicated in (A) and presented as mean ± SD in (B). LP1-SJ: LP1-hFIXco-SJ ($n = 4$); LP1-SIH-M: LP1-hFIXco-SIH-M ($n = 5$); Alb-SIH-M: Alb-hFIXco-SIH-M ($n = 5$); Chi-SIH-M: Chi-hFIXco-SIH-M ($n = 4$); and Pre-SIH-M: Pre-hFIXco-SIH-M ($n = 5$).

Acknowledgements

This study was financially supported by the National Basic Research Program of China (973 Program: 2013CB966800), the Special Grant of MOH (201202003), the Mega-projects of Scientific Research for the 12th Five-Year Plan (2013ZX09303302), the State Key Laboratories Project of Excellence (81123005).

Compliance with ethics guidelines

Ru Zhang, Qiang Wang, Lin Zhang, and Saijuan Chen declare that there is no conflict of interest. All institutional and national guidelines for the care and use of laboratory animals were followed.

Electronic Supplementary Material Supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s11684-015-0390-2 and is accessible for authorized users.

References

1. Thrombosis and Hemostasis Group, Chinese Society of Hematology, Chinese Medical Association/Hemophilia Treatment Center Collaborative Network of China. Consensus of Chinese experts on diagnosis and treatment of hemophilia(2013). Chin J Hematol (Zhonghua Xue Ye Xue Za Zhi) 2013; 34(5): 461–463 (in Chinese)
2. Nilsson IM, Berntorp E, Löfqvist T, Pettersson H. Twenty-five years' experience of prophylactic treatment in severe haemophilia A and B. J Intern Med 1992; 232(1): 25–32
3. Manco-Johnson MJ, Abshire TC, Shapiro AD, Riske B, Hacker MR, Kilcoyne R, Ingram JD, Manco-Johnson ML, Funk S, Jacobson L, Valentino LA, Hoots WK, Buchanan GR, DiMichele D, Recht M, Brown D, Leissinger C, Bleak S, Cohen A, Mathew P, Matsunaga A, Medeiros D, Nugent D, Thomas GA, Thompson AA, McRedmond K, Soucie JM, Austin H, Evatt BL. Prophylaxis versus episodic treatment to prevent joint disease in boys with severe hemophilia. N Engl J Med 2007; 357(6): 535–544
4. Nathwani AC, Tuddenham EG, Rangarajan S, Rosales C, McIntosh J, Linch DC, Chowdary P, Riddell A, Pie AJ, Harrington C, O'Beirne J, Smith K, Pasi J, Glader B, Rustagi P, Ng CY, Kay MA, Zhou J, Spence Y, Morton CL, Allay J, Coleman J, Sleep S, Cunningham JM, Srivastava D, Basner-Tschakarjan E, Mingozzi F, High KA, Gray JT, Reiss UM, Nienhuis AW, Davidoff AM. Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. N Engl J Med 2011; 365(25): 2357–2365
5. Nathwani AC, Gray JT, Ng CY, Zhou J, Spence Y, Waddington SN, Tuddenham EG, Kemball-Cook G, McIntosh J, Boon-Spijker M, Mertens K, Davidoff AM. Self-complementary adeno-associated virus vectors containing a novel liver-specific human factor IX expression cassette enable highly efficient transduction of murine and nonhuman primate liver. Blood 2006; 107(7): 2653–2661
6. Li YM, Li DJ, Xu XJ, Cui M, Zhen HH, Wang Q. Effect of codon optimization on expression levels of human cystatin C in *Pichia pastoris*. Genet Mol Res 2014; 13(3): 4990–5000
7. Burgess-Brown NA, Sharma S, Sobott F, Loenarz C, Oppermann U, Gileadi O. Codon optimization can improve expression of human genes in *Escherichia coli*: a multi-gene study. Protein Expr Purif 2008; 59(1): 94–102
8. Ward NJ, Buckley SM, Waddington SN, Vandendriessche T, Chuah MK, Nathwani AC, McIntosh J, Tuddenham EG, Kinnon C, Thrasher AJ, McVey JH. Codon optimization of human factor VIII cDNAs leads to high-level expression. Blood 2011; 117(3): 798–807
9. Sack BK, Merchant S, Markusic DM, Nathwani AC, Davidoff AM, Byrne BJ, Herzog RW. Transient B cell depletion or improved transgene expression by codon optimization promote tolerance to factor VIII in gene therapy. PLoS ONE 2012; 7(5): e37671

10. Miao CH, Thompson AR, Loeb K, Ye X. Long-term and therapeutic-level hepatic gene expression of human factor IX after naked plasmid transfer *in vivo*. Mol Ther 2001; 3(6): 947–957
11. Zhang G, Song YK, Liu D. Long-term expression of human α1-antitrypsin gene in mouse liver achieved by intravenous administration of plasmid DNA using a hydrodynamics-based procedure. Gene Ther 2000; 7(15): 1344–1349
12. Aliño SF, Crespo A, Dasí F. Long-term therapeutic levels of human α1-antitrypsin in plasma after hydrodynamic injection of nonviral DNA. Gene Ther 2003; 10(19): 1672–1679
13. Nguyen AT, Dow AC, Kupiec-Weglinski J, Busuttil RW, Lipshutz GS. Evaluation of gene promoters for liver expression by hydrodynamic gene transfer. J Surg Res 2008; 148(1): 60–66
14. Al-Dosari M, Zhang G, Knapp JE, Liu D. Evaluation of viral and mammalian promoters for driving transgene expression in mouse liver. Biochem Biophys Res Commun 2006; 339(2): 673–678
15. Xu ZL, Mizuguchi H, Ishii-Watabe A, Uchida E, Mayumi T, Hayakawa T. Optimization of transcriptional regulatory elements for constructing plasmid vectors. Gene 2001; 272(1–2): 149–156
16. Simioni P, Tormene D, Tognin G, Gavasso S, Bulato C, Iacobelli NP, Finn JD, Spiezia L, Radu C, Arruda VR. X-linked thrombophilia with a mutant factor IX (factor IX Padua). N Engl J Med 2009; 361(17): 1671–1675
17. Minghetti PP, Ruffner DE, Kuang WJ, Dennison OE, Hawkins JW, Beattie WG, Dugaiczyk A. Molecular structure of the human albumin gene is revealed by nucleotide sequence within q11-22 of chromosome 4. J Biol Chem 1986; 261(15): 6747–6757
18. Frain M, Hardon E, Ciliberto G, Sala-Trepat JM. Binding of a liver-specific factor to the human albumin gene promoter and enhancer. Mol Cell Biol 1990; 10(3): 991–999
19. Hayashi Y, Chan J, Nakabayashi H, Hashimoto T, Tamaoki T. Identification and characterization of two enhancers of the human albumin gene. J Biol Chem 1992; 267(21): 14580–14585
20. Costa RH, Grayson DR, Darnell JE Jr. Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyretin and α 1-antitrypsin genes. Mol Cell Biol 1989; 9(4): 1415–1425
21. Costa RH, Lai E, Darnell JE Jr. Transcriptional control of the mouse prealbumin (transthyretin) gene: both promoter sequences and a distinct enhancer are cell specific. Mol Cell Biol 1986; 6(12): 4697–4708
22. Tsuzuki T, Mita S, Maeda S, Araki S, Shimada K. Structure of the human prealbumin gene. J Biol Chem 1985; 260(22): 12224–12227
23. Monaci P, Nicosia A, Cortese R. Two different liver-specific factors stimulate *in vitro* transcription from the human α1-antitrypsin promoter. EMBO J 1988; 7(7): 2075–2087
24. De Simone V, Ciliberto G, Hardon E, Paonessa G, Palla F, Lundberg L, Cortese R. Cis- and trans-acting elements responsible for the cell-specific expression of the human α1-antitrypsin gene. EMBO J 1987; 6(9): 2759–2766
25. Bancroft JD, McDowell SA, Degen SJ. The human prothrombin gene: transcriptional regulation in HepG2 cells. Biochemistry 1992; 31(49): 12469–12476
26. Chen Y, Schroeder JA, Kuether EL, Zhang G, Shi Q. Platelet gene therapy by lentiviral gene delivery to hematopoietic stem cells restores hemostasis and induces humoral immune tolerance in FIX (null) mice. Mol Ther 2014; 22(1): 169–177
27. Song YK, Liu F, Zhang G, Liu D. Hydrodynamics-based transfection: simple and efficient method for introducing and expressing transgenes in animals by intravenous injection of DNA. Methods Enzymol 2002; 346: 92–105
28. Liu F, Song Y, Liu D. Hydrodynamics-based transfection in animals by systemic administration of plasmid DNA. Gene Ther 1999; 6(7): 1258–1266
29. Nathwani AC, Davidoff AM, Hanawa H, Hu Y, Hoffer FA, Nikanorov A, Slaughter C, Ng CY, Zhou J, Lozier JN, Mandrell TD, Vanin EF, Nienhuis AW. Sustained high-level expression of human factor IX (hFIX) after liver-targeted delivery of recombinant adeno-associated virus encoding the hFIX gene in rhesus macaques. Blood 2002; 100(5): 1662–1669
30. Gutman GA, Hatfield GW. Nonrandom utilization of codon pairs in *Escherichia coli*. Proc Natl Acad Sci U S A 1989; 86(10): 3699–3703
31. Papadakis ED, Nicklin SA, Baker AH, White SJ. Promoters and control elements: designing expression cassettes for gene therapy. Curr Gene Ther 2004; 4(1): 89–113
32. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res 2004; 14(6): 1188–1190
33. Tokuoka M, Tanaka M, Ono K, Takagi S, Shintani T, Gomi K. Codon optimization increases steady-state mRNA levels in *Aspergillus oryzae* heterologous gene expression. Appl Environ Microbiol 2008; 74(21): 6538–6546
34. Hu S, Li L, Qiao J, Guo Y, Cheng L, Liu J. Codon optimization, expression, and characterization of an internalizing anti-ErbB2 single-chain antibody in *Pichia pastoris*. Protein Expr Purif 2006; 47(1): 249–257
35. Xia X. How optimized is the translational machinery in *Escherichia coli*, *Salmonella typhimurium* and *Saccharomyces cerevisiae*? Genetics 1998; 149(1): 37–44
36. Gvritishvili AG, Leung KW, Tombran-Tink J. Codon preference optimization increases heterologous PEDF expression. PLoS One 2010; 5(11): e15056
37. Wang L, Morizono H, Lin J, Bell P, Jones D, McMenamin D, Yu H, Batshaw ML, Wilson JM. Preclinical evaluation of a clinical candidate AAV8 vector for ornithine transcarbamylase (OTC) deficiency reveals functional enzyme from each persisting vector genome. Mol Genet Metab 2012; 105(2): 203–211
38. Kumar S, Tamura K, Nei M. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 2004; 5(2): 150–163
39. Xia X, Xie Z. DAMBE: software package for data analysis in molecular biology and evolution. J Hered 2001; 92(4): 371–373
40. Xiao W, Berta SC, Lu MM, Moscioni AD, Tazelaar J, Wilson JM. Adeno-associated virus as a vector for liver-directed gene therapy. J Virol 1998; 72(12): 10222–10226
41. Kay MA, Baley P, Rothenberg S, Leland F, Fleming L, Ponder KP, Liu T, Finegold M, Darlington G, Pokorny W. Expression of human α1-antitrypsin in dogs after autologous transplantation of retroviral transduced hepatocytes. Proc Natl Acad Sci U S A 1992; 89(1): 89–93
42. Kay MA, Li Q, Liu TJ, Leland F, Toman C, Finegold M, Woo SL. Hepatic gene therapy: persistent expression of human α1-antitrypsin in mice after direct gene delivery *in vivo*. Hum Gene Ther 1992; 3 (6): 641–647
43. Gehrke S, Jérôme V, Müller R. Chimeric transcriptional control

units for improved liver-specific transgene expression. Gene 2003; 322: 137–143

44. Wu KJ, Wilson DR, Shih C, Darlington GJ. The transcription factor HNF1 acts with C/EBPα to synergistically activate the human albumin promoter through a novel domain. J Biol Chem 1994; 269 (2): 1177–1182
45. Godbout R, Ingram R, Tilghman SM. Multiple regulatory elements in the intergenic region between the α-fetoprotein and albumin genes. Mol Cell Biol 1986; 6(2): 477–487
46. Costa RH, Lai E, Grayson DR, Darnell JE Jr. The cell-specific enhancer of the mouse transthyretin (prealbumin) gene binds a common factor at one site and a liver-specific factor(s) at two other sites. Mol Cell Biol 1988; 8(1): 81–90
47. Costa RH, Grayson DR, Xanthopoulos KG, Darnell JE Jr. A liver-specific DNA-binding protein recognizes multiple nucleotide sites in regulatory regions of transthyretin, α1-antitrypsin, albumin, and simian virus 40 genes. Proc Natl Acad Sci U S A 1988; 85(11): 3840–3844
48. Grayson DR, Costa RH, Xanthopoulos KG, Darnell JE Jr. A cell-specific enhancer of the mouse α1-antitrypsin gene has multiple functional regions and corresponding protein-binding sites. Mol Cell Biol 1988; 8(3): 1055–1066
49. Lemken ML, Wybranietz WA, Schmidt U, Graepler F, Armeanu S, Bitzer M, Lauer UM. Expression liver-directed genes by employing synthetic transcriptional control units. World J Gastroenterol 2005; 11(34): 5295–5302