interferon alpha fusion protein and demonstrate that the fusion protein had an approximate 140-fold slower clearance such that the half-life of the fusion was 18-fold longer than for the interferon alpha protein alone. Other examples of therapeutic proteins recently under development that are albumin fusions include Albulin-G™, Cardeva™ and Albugranin™ (Teva Pharmaceutical Industries, fused to Insulin, b-type natriuretic, or GCSF, respectively), Syncria® (GlaxoSmithKline, fused to Glucagon-like peptide-1) and Albuferon α-2B, fused to IFN-alpha (see *Current Opinion in Drug Discovery and Development*, (2009), vol 12, No. 2. p. 288). In these cases, Albulin-G™, Cardeva™ and Syncria® are all fusion proteins where the albumin is found on the N-terminus of the fusion, while Albugranin™ and Albuferon alpha 2G are fusions where the albumin is on the C-terminus of the fusion.

[0015] Thus, there remains a need for additional methods and compositions that can be used to express a desired transgene at a therapeutically relevant level, while avoiding any associated toxicity, and which may limit expression of the transgene to the desired tissue type, for example to treat genetic diseases such as hemophilias, diabetes, lysosomal storage diseases and A1AT deficiency. Additionally, there remains a need for additional methods and compositions to express a desired transgene at a therapeutically relevant level for the treatment of other diseases such as cancers.

SUMMARY

[0016] Disclosed herein are methods and compositions for creating a safe harbor in the genome of cells, for targeted insertion and subsequence expression of a transgene, for example expression of the transgene from a secretory tissue such as liver. In one aspect, the present invention refers to a non-naturally occurring fusion protein comprising a zinc finger protein that binds to an endogenous albumin gene and a cleavage domain, wherein the fusion protein modifies the endogenous albumin gene. Also described herein is a non-naturally occurring zinc-finger protein (ZFP) that binds to target site in a region of interest (*e.g.,* an albumin gene) in a genome, wherein the ZFP comprises one or more engineered zinc-finger binding domains. In one ~~embodiment~~example, the ZFP is a zinc-finger nuclease (ZFN) that cleaves a target genomic region of interest, wherein the ZFN comprises one or more engineered zinc-finger binding domains and a nuclease cleavage domain or cleavage half-domain. Cleavage domains and cleavage half domains can be obtained, for example, from

various restriction endonucleases and/or homing endonucleases. In one ~~embodiment~~example, the cleavage half-domains are derived from a Type IIS restriction

endonuclease *(e.g., Fok* I). In certain ~~embodiments~~examples, the zinc finger domain recognizes a target site in an albumin gene, for example a zinc finger protein with the recognition helix domains ordered as shown in a single row of Tables 1, 3, 5 or 8. .

[0017] In another ~~aspect~~example, described herein is a Transcription Activator Like Effector (TALE) protein that binds to target site in a region of interest (*e.g.,* an albumin gene) in a genome, wherein the TALE comprises one or more engineered TALE binding domains. In one ~~embodiment~~example, the TALE is a nuclease (TALEN) that cleaves a target genomic region of interest, wherein the TALEN comprises one or more engineered TALE DNA binding domains and a nuclease cleavage domain or cleavage half-domain. Cleavage domains and cleavage half domains can be obtained, for example, from various restriction endonucleases and/or homing endonucleases. In one ~~embodiment~~example, the cleavage half-domains are derived from a Type IIS restriction endonuclease *(e.g., Fok* I). In certain ~~embodiments~~examples, the TALE DNA binding domain recognizes a target site in an albumin gene, for example TALE DNA binding domain having the target sequence shown in a single row of Table 12.

[0018] The ZFN and/or TALEN as described herein may bind to and/or cleave the region of interest in a coding or non-coding region within or adjacent to the gene, such as, for example, a leader sequence, trailer sequence or intron, or within a non-transcribed region, either upstream or downstream of the coding region. In certain ~~embodiments~~examples, the ZFN binds to and/or cleaves an albumin gene. In other ~~embodiments~~examples, the ZFN and/or TALEN binds to and/or cleaves a safe-harbor gene, for example a CCR5 gene, a PPP1R12C (also known as AAVS1) gene or a *Rosa* gene. *See, e.g.,* U.S. Patent Publication Nos. 20080299580; 20080159996 and 201000218264. In another ~~aspect~~example, described herein are compositions comprising one or more of the zinc-finger and/or TALE nucleases described herein. In certain ~~embodiments~~examples, the composition comprises one or more zinc-finger and/or TALE nucleases in combination with a pharmaceutically acceptable excipient.

[0019] In another aspect, ~~described herein~~the present invention refers to a polynucleotide encoding one or more fusion proteins of the present invention. Also disclosed herein is a polynucleotide encoding one or more ZFNs and/or TALENs described herein. The polynucleotide may be, for example, mRNA. In some ~~aspects~~examples, the mRNA may be chemically modified (See *e.g.* Kormann *et al,* (2011) *Nature Biotechnology* 29(2):154-157).

[0020] In another ~~aspect~~example, described herein is a ZFN and/or TALEN expression vector comprising a polynucleotide, encoding one or more ZFNs and/or TALENs

described herein, operably linked to a promoter. In one ~~embodiment~~example, the expression vector is a viral vector. In one ~~aspect~~example, the viral vector exhibits tissue specific tropism.

[0021] In another aspect, the present invention refers to an isolated cell comprising one or more fusion proteins according to the present invention or one or more polynucleotides according to the present invention. Also, described herein is a host cell comprising one or more ZFN and/or TALEN expression vectors. The host cell may be stably transformed or transiently transfected or a combination thereof with one or more ZFP or TALEN expression vectors. In one ~~embodiment~~example, the host cell is an embryonic stem cell. In other ~~embodiments~~examples, the one or more ZFP and/or TALEN expression vectors express one or more ZFNs and/or TALENs in the host cell. In another ~~embodiment~~example, the host cell may further comprise an exogenous polynucleotide donor sequence. Non-limiting examples of suitable host cells include eukaryotic cells or cell lines such as secretory cells (*e.g.*, liver cells, mucosal cells, salivary gland cells, pituitary cells, etc.), blood cells (red blood cells), stem cells, etc. In any of the ~~embodiments~~ examples described herein the host cell can comprise an embryo cell, for example, of a mouse, rat, rabbit or other mammal cell embryo.

[0022] In another aspect, the present invention refers to method of cleaving an endogenous albumin gene in a cell, the method comprising introducing, into the cell, one or more expression vectors comprising at least one fusion protein according to the present invention or at least one polynucleotide according to the present invention, under conditions such that the one or more fusion proteins are expressed and the albumin gene is cleaved. Also, described herein is a method for cleaving an albumin gene in a cell, the method comprising: introducing, into the cell, one or more polynucleotides encoding one or more ZFNs and/or TALENs that bind to a target site in the one or more albumin genes under conditions such that the ZFN(s) is (are) or TALENs is (are) expressed and the one or more albumin genes are cleaved.

[0023] In other ~~embodiments~~examples, a genomic sequence in any target gene is replaced, for example using a ZFN or TALEN (or vector encoding said ZFN or TALEN) as described herein and a "donor" sequence (*e.g.*, transgene) that is inserted into the gene following targeted cleavage with the ZFN and/or TALEN. The donor sequence may be present in the ZFN or TALEN vector, present in a separate vector (*e.g.*, Ad or LV vector) or, alternatively, may be introduced into the cell using a

different nucleic acid delivery mechanism. Such insertion of a donor nucleotide sequence into the target locus (*e.g.*, albumin gene, other safe-harbor gene, etc.) results in the expression of the transgene carried by the donor under control of the target locus's (*e.g.* albumin) genetic control elements. In some ~~aspects~~examples, insertion of the transgene of interest, for example into an albumin gene results in expression of an intact exogenous protein sequence and lacks any albumin encoded amino acids. In other ~~aspects~~examples, the expressed exogenous protein is a fusion protein and comprises amino acids encoded by the transgene and by an albumin gene (*e.g.*, from the

endogenous target locus or, alternatively from albumin-encoding sequences on the transgene). In some instances, the albumin sequences will be present on the amino (N)-terminal portion of the exogenous protein, while in others, the albumin sequences will be present on the carboxy (C)- terminal portion of the exogenous protein. In other instances, albumin sequences will be present on both the N- and C-terminal portions of the exogenous protein. The albumin sequences may include full-length wild-type or mutant albumin sequences or, alternatively, may include partial albumin amino acid sequences. In certain ~~embodiments~~examples, the albumin sequences (full-length or partial) serve to increase the serum half-life of the polypeptide expressed by the transgene to which it is fused and/or as a carrier. In some ~~embodiments~~examples, the albumin-transgene fusion is located at the endogenous locus within the cell while in other embodiments, the albumin-transgene coding sequence is inserted into a safe harbor within a genome. In some ~~aspects~~examples, the safe harbor is selected from the AAVS1, Rosa, HPRT or CCR5 locus (see co-owned US patent publications Nos. 20080299580; 20080159996 and 201000218264, and US Provisional patent application No. 61/556,691).

[0024] In another ~~aspect~~example, the invention ~~describes~~ disclosure methods and compositions that can be used to express a transgene under the control of an albumin promoter *in vivo* (*e.g.,* endogenous or exogenous albumin promoter). In some ~~aspects~~examples, the transgene may encode a therapeutic protein of interest. The transgene may encode a protein such that the methods ~~of the invention~~disclosed herein can be used for production of protein that is deficient or lacking (*e.g.,* "protein replacement"). In some instances, the protein may be involved treatment for a lysosomal storage disease. Other therapeutic proteins may be expressed, including protein therapeutics for conditions as diverse as epidermolysis bullosa or AAT deficient emphysema. In other ~~aspects~~examples, the transgene may comprise sequences (*e.g.,* engineered sequences) such that the expressed protein has characteristics which give it novel and desirable features (increased half-life, changed plasma clearance characteristics etc.). Engineered sequences can also include amino acids derived from the albumin sequence. In some ~~aspects~~examples, the transgenes encode therapeutic proteins, therapeutic hormones, plasma proteins, antibodies and the like. In some ~~aspects~~examples, the transgenes may encode proteins involved in blood disorders such as clotting disorders. In some ~~aspects~~examples, the transgenes encode structural nucleic acids (shRNAs, miRNAs and the like).

[0025] In some ~~embodiments~~examples, the methods ~~of the invention~~disclosed herein may be used *in vivo* in transgenic animal systems. In some ~~aspects~~examples, the transgenic animal may be used in model development where the transgene encodes a human gene. In some instances, the transgenic animal may be knocked out at the corresponding endogenous locus, allowing the development of an *in vivo* system where the human protein may be studied in isolation. Such transgenic models may be used for screening purposes to identify small molecule, large biomolecules or other entities which may interact or modify the human protein of interest. In other ~~aspects~~examples, the transgenic animals may be used for production purposes, for example, to produce antibodies or other biomolecules of interest. In certain ~~embodiments~~examples, the animal is a small mammal, for example a dog, rabbit or a rodent such as rat, a mouse or a guinea pig. In other ~~embodiments~~examples, the animal is a non-human primate. In yet further embodiments, the animal is a farm animal such as a cow, goat or pig. In some ~~aspects~~examples, the transgene is integrated into the selected locus (*e.g.,* albumin or safe-harbor) into a stem cell (*e.g.,* an embryonic stem cell, an induced pluripotent stem cell, a hepatic stem cell, etc.) or animal embryo obtained by any of the methods described herein, and then the embryo is implanted such that a live animal is born. The animal is then raised to sexual maturity and allowed to produce offspring wherein at least some of the offspring comprise the integrated transgene.

[0026] In a still further example~~aspect~~, provided herein is a method for site specific integration of a nucleic acid sequence into an endogenous locus (*e.g.,* albumin gene) of a chromosome, for example into the chromosome of an embryo. In certain examples~~embodiments~~, the method comprises: (a) injecting an embryo with (i) at least one DNA vector, wherein the DNA vector comprises an upstream sequence and a downstream sequence flanking the nucleic acid sequence to be integrated, and (ii) at least one RNA molecule encoding a zinc finger and/or TALE nuclease that recognizes the site of integration in the target locus (*e.g.,* albumin locus), and (b) culturing the embryo to allow expression of the zinc finger and/or TALE nuclease, wherein a double stranded break introduced into the site of integration by the zinc finger nuclease or TALEN is repaired, via homologous recombination with the DNA vector, so as to integrate the nucleic acid sequence into the chromosome.

[0027] Suitable embryos may be derived from several different vertebrate species, including mammalian, bird, reptile, amphibian, and fish species. Generally

speaking, a suitable embryo is an embryo that may be collected, injected, and cultured to allow the expression of a zinc finger or TALE nuclease. In some embodiments, suitable embryos may include embryos from small mammals (*e.g.*, rodents, rabbits, etc.), companion animals, livestock, and primates. Non-limiting examples of rodents may include mice, rats, hamsters, gerbils, and guinea pigs. Non-limiting examples of companion animals may include cats, dogs, rabbits, hedgehogs, and ferrets. Non-limiting examples of livestock may include horses, goats, sheep, swine, llamas, alpacas, and cattle. Non-limiting examples of primates may include capuchin monkeys, chimpanzees, lemurs, macaques, marmosets, tamarins, spider monkeys, squirrel monkeys, and vervet monkeys. In other embodiments, suitable embryos may include embryos from fish, reptiles, amphibians, or birds. Alternatively, suitable embryos may be insect embryos, for instance, a *Drosophila* embryo or a mosquito embryo.

[0028] In any of the methods or compositions described herein, the cell containing the engineered locus (*e.g.*, albumin locus) can be a stem cell. Specific stem cell types that may be used with the methods and compositions ~~of the invention~~disclosed herein include embryonic stem cells (ESC), induced pluripotent stem cells (iPSC) and hepatic or liver stem cells. The iPSCs can be derived from patient samples and from normal controls wherein the patient derived iPSC can be mutated to normal gene sequence at the gene of interest, or normal cells can be altered to the known disease allele at the gene of interest. Similarly, the hepatic stem cells can be isolated from a patient. These cells are then engineered to express the transgene of interest, expanded and then reintroduced into the patient.

[0029] In any of the methods described herein, the polynucleotide encoding the zinc finger nuclease(s) and/or TALEN(s) can comprise DNA, RNA or combinations thereof. In certain examples~~embodiments~~, the polynucleotide comprises a plasmid. In other examples~~embodiments~~, the polynucleotide encoding the nuclease comprises mRNA.

[0030] Also ~~provided~~ described is an embryo comprising at least one DNA vector, wherein the DNA vector comprises an upstream sequence and a downstream sequence flanking the nucleic acid sequence to be integrated, and at least one RNA molecule encoding a zinc finger nuclease that recognizes the chromosomal site of integration. Organisms derived from any of the embryos as described herein are also

provided (*e.g.*, embryos that are allowed to develop to sexual maturity and produce progeny).

[0031] In another examplearpect provided by the methods and compositions of the inventiondisclosed herein is the use of cells, cell lines and animals (*e.g.*, transgenic animals) in the screening of drug libraries and/or other therapeutic compositions (*i.e.*, antibodies, structural RNAs, etc.) for use in treatment of an animal afflicted with a clotting factor disorder. Such screens can begin at the cellular level with manipulated cell lines or primary cells, and can progress up to the level of treatment of a whole animal (*e.g.*, human).

[0032] A kit, comprising the ZFPs and/or TALENs of the invention, is also provideddisclosed. In one aspect, the present invention refers to a kit comprising a fusion protein according to the present invention of a polynucleotide according to the present invention. The kit may comprise nucleic acids encoding the ZFPs or TALENs, (*e.g.* RNA molecules or ZFP or TALEN encoding genes contained in a suitable expression vector), donor molecules, suitable host cell lines, instructions for performing the methods of the invention, and the like.

[0033] Thus, the disclosure herein includes, but is not limited to, the following examplesembodiments:

[0034] 1. A non-naturally occurring fusion protein comprising a DNA-binding protein that binds to an endogenous albumin gene and a cleavage domain, wherein the fusion protein modifies the endogenous albumin gene.

[0035] 2. The fusion protein of exampleembodiment 1, wherein the DNA-binding protein comprises a zinc finger protein.

[0036] 3. The fusion protein of exampleembodiment 2, wherein the zinc finger protein comprises 4, 5 or 6 zinc finger domains comprising a recognition helix region, wherein the zinc finger proteins comprise the recognition helix regions shown in a single row of Table 1, Table 3, Table 5 or Table 8.

[0037] 4. The fusion protein of exampleembodiment 1, wherein the DNA-binding protein comprises a TALE DNA-binding domain.

[0038] 5. The fusion protein of exampleembodiment 4, wherein the TALE DNA-binding domain binds to a target sequence shown in a single row of Table 12.

[0039] 6. A polynucleotide encoding one or more fusion proteins of examplesembodiments 1 to 5.

[0040] 7. An isolated cell comprising one or more fusion proteins according to examples~~embodiments~~ 1 to 5 or one or more polynucleotides according to example~~embodiment~~ 6.

5

[0041] 8. The cell of example~~embodiment~~ 7, wherein the cell is a stem cell or an embryo cell.

[0042] 9. The cell of example~~embodiment~~ 8, wherein the stem cell is selected from the group consisting of an embryonic stem cell (ESC), an induced pluripotent stem cell (iPSC), a hepatic stem cell and a liver stem cell.

[0043] 10. A kit comprising a fusion protein according to examples~~embodiments~~ 1 to 5 or a polynucleotide according to example~~embodiment~~ 6 or a cell according to example~~embodiment~~ 7-9.

[0044] 11. A method of cleaving an endogenous albumin gene in a cell, the method comprising:

introducing, into the cell, one or more expression vectors comprising at least one polynucleotide according to example~~embodiment~~ 6, under conditions such that the one or more fusion proteins are expressed and the albumin gene is cleaved.

[0045] 12. The method of example~~embodiment~~ 11, wherein the polynucleotide comprises an AAV vector.

[0046] 13. The method of example~~embodiment~~ 11, wherein the cell is a liver cell.

[0047] 14. A method of introducing a transgene into an endogenous albumin gene, the method comprising:

cleaving the endogenous albumin gene according to the method of any of examples~~embodiments~~ 15-17 in the presence of an exogenous polynucleotide comprising the transgene such that the transgene is integrated into the endogenous albumin gene.

[0048] 15. The method of example~~embodiment~~ 14, wherein the transgene expresses a therapeutic protein.

[0049] 16. The method of example~~embodiment~~ 15, wherein the therapeutic protein is involved in treating a lysosomal storage disease, epidermolysis bullosa, AAT deficient emphysema or blood disorders such as clotting disorders.

[0050] 17. The method of examples~~embodiments~~ 15 or 16, wherein expression of the transgene is driven by the endogenous albumin control sequences.

[0051] 18. The method of any of examples~~embodiments~~ 15-17, wherein the transgene further comprises albumin sequences.

[0052] 19. The method of example~~embodiment~~ 18, wherein the albumin sequences are present on the amino (N)-terminal and/or carboxy (C)-terminal portion of the protein.

[0053] 20. A method of increasing the serum half-life of a polypeptide expressed from a transgene integrated into an endogenous albumin gene, the method comprising introducing the transgene into the endogenous albumin gene according to the method of example~~embodiment~~ 18 or example~~embodiment~~ 19, wherein the transgene expresses the polypeptide and albumin sequences such that the serum half-life of the polypeptide in increased.

[0054] 21. A method of treating a subject having a disease caused by a deficiency of a polypeptide, the method comprising,

introducing a transgene encoding the polypeptide into an isolated cell according to the method of examples~~embodiments~~ 14-19 such that the transgene is expressed in the isolated cell; and introducing the isolated cell into the subject, thereby treating the disease.

[0055] 22. The method of example~~embodiment~~ 21, wherein the cell is a liver cell or a stem cell.

[0056] 23. The cell of example~~embodiment~~ 22, wherein the stem cell is selected from the group consisting of an embryonic stem cell (ESC), an induced pluripotent stem cell (iPSC), a hepatic stem cell and a liver stem cell.

[0057] These and other aspects will be readily apparent to the skilled artisan in light of disclosure as a whole.

BRIEF DESCRIPTION OF THE DRAWINGS

[0058] **Figure 1, panels A and B,** are gels depicting the results of a Cel-I mismatch assay (Surveyor™, Transgenomic) that quantifies the degree to which ZFN cleavage of an endogenous chromosomal target, followed by imperfect repair via NHEJ, has yielded small insertions or deletions ("indels") of the targeted locus. For a description of the assay see Horton *et al. Methods Mol Biol.* (2010) 649:247-56. Figure 1A shows results using expression constructs for ZFNs targeted to the mouse albumin gene which were transfected into Neuro2A cells, where the cells were treated for 3 days at 37°C following transfection, and then analyzed for the fraction of modified target sites via Cel-I analysis. Figure 1B shows results for the same ZFNs and cells as Figure 1A except cells were subjected to hypothermic shock (30°C) during their 3 days of growth following transfection. The percent mismatch, or % indels shown at the bottom on the lanes, is a measure of the ZFN activity and