processes such as removal of virus particles by using filters or resins, as well as inactivation steps by using low pH or detergents, are implemented to prevent the serious safety issue of viral contamination of protein drug substances. Given the complexity of therapeutic proteins with respect to their large molecular size, post-translational modifications, and the variety of biological materials involved in their manufacturing process, the ability to enhance particular functional attributes while maintaining product safety and efficacy achieved through protein-engineering strategies is highly desirable.

While the integration of novel strategies and approaches to modify protein drug products is not a trivial matter, the potential therapeutic advantages have driven the increased use of such strategies during drug development. A number of protein-engineering platform technologies are currently in use to increase the circulating half-life, targeting, and functionality of novel therapeutic protein drugs as well as to increase production yield and product purity. For example, protein conjugation and derivatization approaches, including Fc-fusion, albumin-fusion, and PEGylation, are currently being used to extend a drug's circulating half-life.

The production of protein pharmaceutical (biologics) is expensive and time consuming. What is needed in the art are more efficient tools and processes for producing this important class of drugs.

SUMMARY

The present disclosure describes host cell lines containing multiple dock sites for insertion of a nucleic acid construct, and in particular nucleic acid constructs that express an exogenous gene of interest.

In one aspect, the present disclosure refers to a host cell comprising a genome, the genome comprising from 1 to 500 integrated docking sites, each docking site comprising at least one dock site insertion element. In some preferred examples, the present disclosure describes a host cell (or population of host cells) comprising a genome, the genome comprising from 1 to 1000 integrated docking sites, each docking site comprising at least one dock site insertion element. In some preferred examples, the genome comprises from 1 to 500 integrated docking sites, each docking site comprising at least one dock site insertion element. In some preferred examples, the genome comprises from 5 to 500 integrated docking sites, each docking site comprising at least one dock site insertion element. In some preferred examples, the genome comprises from 5 to 250 integrated docking sites, each docking site comprising at least one dock site insertion element. In some preferred examples, the genome comprises from 5 to 100 integrated docking sites, each docking site comprising at least one dock site insertion element. In some preferred examples, the genome comprises

from 5 to 50 integrated docking sites, each docking site comprising at least one dock site insertion element. In some preferred examples, the integrated docking sites are independently positioned throughout the genome.

In some preferred examples, the dock site insertion element is targeted by enzyme selected from the group consisting of an integrase, a recombinase, a nuclease and a nickase. In some preferred examples, the dock site insertion element is selected from the group consisting of a recombinase dock site insertion element and a HDR dock site insertion element. In some preferred examples, the dock site insertion element is a recombinase dock site insertion element. In some preferred examples, the recombinase dock site insertion element comprises an attachment site (att). In some preferred examples, the attachment site (att) is selected from the group consisting of attB and attP. In some preferred examples, the attachment site (att) is selected from the group consisting of attR and attL. In some preferred examples, the recombinase dock site insertion element comprises a LoxP sequence. In some preferred examples, the recombinase dock site insertion element is a Flp Recombination Target (FRT) site. In some preferred examples, the dock site insertion element is a HDR dock site insertion element. In some preferred examples, the HDR dock site insertion element comprises one or two dock site homology arms. In some preferred examples, the HDR dock site insertion element further comprises one or more sequences homologous to a guide RNA sequence. In some preferred examples, the dock site homology arms are from about 30 to 1000 bases in length. In some preferred examples, the integrase dock site insertion element comprises an AAVS1 safe harbor locus sequence.

In some preferred examples, each docking site is flanked by exogenous integrating vector sequences. In some preferred examples, the exogenous integrating vector sequences are selected from the group consisting of viral vector sequences and transposon vector sequences. In some preferred examples, the docking sites each further comprise a sequence encoding a selectable maker operably linked to a promoter.

In some preferred examples, the host cell line further comprises an exogenous sequence encoding an enzyme operably linked to a promoter. In some preferred examples, the exogenous sequence encoding an enzyme operably linked to a promoter is inserted into the genome of the host cell. In some preferred examples, the exogenous sequence encoding an enzyme operably linked to a promoter is inserted at the dock site. In some preferred examples, the exogenous sequence encoding an enzyme operably linked to a promoter is provided in an episomal expression vector. In some preferred examples,

the episomal expression vector is a plasmid. In some preferred examples, the enzyme is selected from the group consisting of an integrase, a recombinase, a nuclease and a nickase. In some preferred examples, the nuclease is a Cas nuclease. In some preferred examples, the nickase is a Cas nickase.

In some preferred examples, the dock site insertion element is positioned to facilitate cassette exchange. In some preferred examples, each docking site comprises two dock site insertion elements. In some preferred examples, the two dock site insertion elements are positioned to facilitate cassette exchange. In some preferred examples, the two dock site insertion elements flank sequences encoding a selectable marker, an enzyme, or a combination thereof.

In some preferred examples, the host cells further comprise nucleic acid expression constructs inserted at 1 or more of the docking sites. In some preferred examples, the nucleic acid expression construct comprises a second (i.e., internal) promoter operably linked to a sequence encoding a protein of interest. In some preferred examples, the nucleic acid expression construct further comprises a 5' promoter (or first promoter) operably linked and 5' to a selectable marker.

In some preferred examples, the nucleic acid expression construct further comprises the following elements in operable association in 5' to 3' order: a first (or 5') promoter sequence; a selectable marker sequence; a second (or internal) promoter sequence; a nucleic acid sequence encoding a first protein of interest that is operably linked to the internal promoter; and a poly A signal sequence. In some preferred examples, the nucleic acid construct further comprising at least one insertion element at a position or positions selected from the group consisting of 5' to the first promoter, 3' to the poly A signal sequence, between the first promoter and the poly A signal sequence, between the selectable marker and the second promoter sequence, and both 5' to the first promoter and 3' to the poly A signal sequence. In some preferred examples, the nucleic acid expression construct does not comprise a poly A signal sequence between the selectable marker and the second promoter. In some preferred examples, the selectable marker is adjacent to the second promoter. In some preferred examples, the second promoter is adjacent to the nucleic acid sequence encoding the first protein of interest. In some preferred examples, the nucleic acid construct comprises a non-coding region between the first promoter and the selectable marker. In some preferred examples, the non-coding region comprises multiple potential Kozak sequences and/or ATG translation start sites. In some preferred examples, the nucleic acid construct comprises an extending packaging region (EPR) between the first

promoter and the selectable marker. In some preferred examples, the EPR comprises multiple potential Kozak sequences and/or ATG translation start sites.

In some preferred examples, the first promoter sequence (i.e., the 5' promoter sequence) is selected from the group consisting of SIN-LTR, SV40, EF1α, *E. coli* lac, *E. coli* trp, phage lambda PL, phage lambda PR, T3, T7, cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, alpha-lactalbumin, and mouse metallothionein-I promoter sequences. In some preferred examples, the first promoter sequence is a weak promoter sequence. In some preferred examples, the first promoter sequence is not a retroviral LTR promoter.

In some preferred examples, the integrated docking sites further comprise an integrated exogenous promoter. In some preferred examples, the promoter is a weak promoter. In some preferred examples, the integrated exogenous promoter is selected from the group consisting of SIN-LTR, SV40, EF1α, *E. coli* lac, *E. coli* trp, phage lambda PL, phage lambda PR, T3, T7, cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, alpha-lactalbumin, and mouse metallothionein-I promoter sequences. In some preferred examples, the integrated promoter is a retroviral LTR. In some preferred examples, the retroviral LTR is a SIN LTR. In some preferred examples, each docking site comprises a SIN LTR EPR 5' of the docking site and a SIN LTR 3' of the docking site. In some preferred examples, the host cell further comprise nucleic acid expression constructs inserted at 1 or more of the docking sites with an integrated promoter.

In some preferred examples, nucleic acid expression constructs for use in cell lines with an integrated exogenous promoter in the docking site comprise an internal (or second) promoter sequence operably linked to a sequence encoding a protein of interest. In some preferred examples, the nucleic acid expression construct comprises the following elements in operable association in 5' to 3' order: a selectable marker sequence; an internal promoter sequence; a nucleic acid sequence encoding a first protein of interest that is operably linked to the internal promoter; and a poly A signal sequence. In some preferred examples, the nucleic acid construct further comprises at least one insertion element at a position or positions selected from the group consisting of 5' to the selectable marker sequence, 3' to the poly A signal sequence, between the selectable marker sequence and the poly A signal sequence, between the selectable marker and the internal promoter sequence, and both 5' to the selectable marker sequence and 3' to the poly A signal sequence. In some preferred examples, the nucleic acid expression construct does not comprise a poly A

signal sequence between the selectable marker and the second promoter. In some preferred examples, the selectable marker is adjacent to the internal promoter sequence. In some preferred examples, the internal promoter sequence is adjacent to the nucleic acid sequence encoding the first protein of interest. In some preferred examples, the internal promoter sequence is selected from the group consisting of SV40, *E. coli* lac, *E. coli* trp, phage lambda PL, phage lambda PR, T3, T7, cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, alpha-lactalbumin, and mouse metallothionein-I promoter sequences.

In some preferred examples, the selectable marker sequence in any of the foregoing expression constructs is an amplifiable selectable marker sequence selected from the group consisting of the Glutamine Synthase (GS) sequence and the Dihydrofolate Reductase (DHFR) sequence. In some preferred examples, the selectable marker sequence is an antibiotic resistance marker sequence selected from the group consisting of neomycin resistance gene (neo), hygromycin B phosphotransferase gene and puromycin N-acetyl transferase gene sequences.

In some preferred examples, the second promoter sequence in any of the foregoing expression constructs is selected from the group consisting of SV40, *E. coli* lac, *E. coli* trp, phage lambda PL, phage lambda PR, T3, T7, cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, alpha-lactalbumin, and mouse metallothionein-I promoter sequences.

In some preferred examples, the nucleic acid sequence encoding a protein of interest encodes a protein selected from the group consisting of heavy and light chain immunoglobulin sequences.

In some preferred examples, the insertion element in any of the foregoing expression constructs is selected from the group consisting of a recombinase insertion element and a HDR insertion element. In some preferred examples, the expression construct insertion element is a recombinase expression construct insertion element compatible with the recombinase dock site insertion element. In some preferred examples, the recombinase expression construct insertion element is an attachment site (att). In some preferred examples, the attachment site (att) is selected from the group consisting of attP and attB. In some preferred examples, the recombinase expression construct insertion element is a Flp Recombination Target (FRT) site. In some preferred examples, the recombinase expression construct insertion element is a LoxP sequence. In some preferred examples, the expression construct insertion element is a HDR expression

construct insertion element compatible with the HDR dock site insertion element. In some preferred examples, the HDR expression construct insertion element comprises one or two expression construct homology arms compatible with or homologous to the one or two dock site homology arms. In some preferred examples, the homology arms are from about 30 to 1000 bases in length. In some preferred examples, the recombinase expression construct comprises two homology arms positioned 5' to the first promoter and 3' to the poly A signal sequence. In some preferred examples, the HDR expression construct insertion element comprises an AAVS1 safe harbor locus sequence.

In some preferred examples, the nucleic acid expression constructs in any of the preceding embodiments further comprise a first RNA export element. In some preferred examples, the first RNA export element is located 3' or 5' to the nucleic acid sequence encoding the protein of interest. In some preferred examples, the first RNA export element is a pre-mRNA processing enhancer (PPE). In some preferred examples, the first RNA export element is a posttranscriptional regulatory element (PRE). In some preferred examples, the PRE RNA export element is a Woodchuck hepatitis virus post-transcriptional regulatory element (WPRE).

In some preferred examples, the nucleic acid expression constructs described in the preceding examplesfurther comprise a nucleic acid sequence encoding at least a second protein of interest (e.g., including third, fourth, fifth, etc. proteins of interest). In some preferred examples, the nucleic acid sequence encoding the second protein of interest is operably associated with a construct element selected from the group consisting of a third promoter, an intron, a second RNA export element and an IRES sequence and combinations thereof. In some preferred examples, the second RNA export element is a pre-mRNA processing enhancer (PPE). In some preferred examples, suitable third promoters include, but are not limited to, SV40, EF1α, *E. coli* lac, *E. coli* trp, phage lambda PL, phage lambda PR, T3, T7, cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, alpha-lactalbumin, and mouse metallothionein-I promoter sequences. In some preferred examples, the second RNA export element is a posttranscriptional regulatory element (PRE). In some preferred examples, the PRE RNA export element is a Woodchuck hepatitis virus post-transcriptional regulatory element (WPRE). In some preferred examples, the second nucleic acid sequence encoding a second protein of interest is positioned 3' to the nucleic acid sequence encoding the first protein of interest. In some preferred examples, the second nucleic acid sequence encoding a second protein of

interest is operably associated with IRES sequence is selected from the group consisting of foot and mouth disease virus (FDV), encephalomyocarditis virus and poliovirus IRES sequences.

In some preferred examples, the nucleic acid expression constructs of any of the preceding examples further comprise a signal peptide sequence operably linked to the first protein of interest. In some preferred examples, the signal peptide sequence is selected from the group consisting of tissue plasminogen activator, human growth hormone, lactoferrin, alpha-casein and alpha-lactalbumin signal peptide sequences. In some preferred examples, the nucleic acid expression constructs further comprise a protein purification marker sequence. In some preferred examples, the protein purification marker sequence is a hexahistidine tag or a hemagglutinin (HA) tag.

In some preferred examples, the nucleic acid constructs encode a viral vector (i.e., instead of a protein of interest). In some preferred examples, the viral vector is selected from the group consisting of a retroviral vector, adenoviral vector and a adenovirus-associated viral vector. In some preferred examples, the retroviral vector is a lentiviral vector.

In some preferred examples, the host cell is selected from the group consisting of Chinese Hamster Ovary (CHO) cells, HEK 293 cells, CAP cells, bovine mammary epithelial cells, monkey kidney CV1 line transformed by SV40, baby hamster kidney cells, mouse sertoli cells, monkey kidney cells, African green monkey kidney cells, human cervical carcinoma cells, canine kidney cells, buffalo rat liver cells, human lung cells, human liver cells, mouse mammary tumor, TRI cells, MRC 5 cells, FS4 cells, rat fibroblasts, MDBK cells and human hepatoma line cells. In some preferred examples, the host cell is selected from the group consisting of a Chinese Hamster Ovary (CHO) cells, a HEK 293 cells and a CAP cells. In some preferred examples, the host cell line is a GS knockout cell line. In some preferred examples, the host cell line is a DHFR knockout cell line.

In some preferred examples, the host cell further comprises at least a second nucleic acid construct that encodes and allows for expression of at least a second protein of interest, and wherein said second nucleic acid construct does not include a selectable marker. In some preferred examples, the host cell further comprises at least a second nucleic acid construct that encodes and allows for expression of a second protein of interest, and wherein said second nucleic acid construct includes a selectable marker that is different from the selectable marker in the first nucleic acid construct. In some preferred examples, the first protein of interest in the first vector is one of an immunoglobulin heavy or light chain and the

second protein in the second vector is the other of an immunoglobulin heavy or light chain. In some preferred examples, the first protein of interest is an immunoglobulin heavy chain and the second protein of interest is an immunoglobulin light chain. In some preferred examples, the second nucleic acid construct is inserted at one or more dock sites.

In one aspect, the present disclosure refers to a cell culture comprising host cells as described herein. In some preferred examples, the cell culture is a master cell culture.

In one aspect, the present disclosure refers to a process for producing a protein of interest comprising culturing host cells as disclosed herein under conditions that the protein of interest is expressed and purifying the protein of interest from the host cell culture. In some preferred examples, the present disclosure describes processes for producing a protein of interest comprising culturing host cells as described above that comprise a nucleic acid expression construct encoding a protein of interest under conditions such that the protein of interest is expressed and purifying the protein of interest from the host cell culture. In some preferred examples, the host cells grown in a medium comprising an inhibitor of the selectable marker. In some preferred examples, the selectable marker is GS and the inhibitor is phosphinothricin or methionine sulphoximine (Msx). In some preferred examples, the selectable marker is DHFR and the inhibitor is methotrexate.

In some preferred examples, the present disclosure describes processes for producing a viral vector comprising culturing host cells comprising an expression construct for a viral vector as described above under conditions that the viral vector is expressed (or synthesized) and purifying the viral vectors from the host cell culture.

In one aspect, the present disclosure refers to a system comprising a host cell or population of host cells as described herein and one or more nucleic acid expression constructs, wherein the nucleic acid expression construct further comprises the following elements in operable association in 5' to 3' order: a selectable marker sequence; an internal promoter sequence; a nucleic acid sequence encoding a first protein of interest that is operably linked to the internal promoter; and a poly A signal sequence, wherein the nucleic acid construct further comprises at least one insertion element, and wherein the expression construct insertion elements are compatible with the dock site insertion elements to facilitate insertion of the nucleic acid expression construct at the dock site. In some preferred examples, the one or more nucleic acid expression constructs further comprise a 5' promotor sequence 5' to the selectable marker sequence. In some preferred examples, the 5' promoter is a weak promoter. In some preferred examples, suitable 5' promoters include, but are not limited to, SIN-LTR, SV40, EF1α, *E. coli* lac, *E. coli* trp, phage lambda PL, phage lambda PR, T3, T7, cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, alpha-lactalbumin, and mouse metallothionein-I promoter

sequences. In some preferred examples, the nucleic acid expression constructs are provided in a vector. In some preferred examples, the vector is a plasmid vector. Other suitable vectors include cosmids, YACs and small circular constructs.

In some preferred examples, the systems further comprise a nucleic acid construct encoding an enzyme that facilitates insertion of the nucleic acid expression construct at the dock site. In some preferred examples, the enzyme is selected from the group consisting of an integrase, a recombinase, a nuclease and a nickase. In some preferred examples, the nuclease is a Cas nuclease. In some preferred examples, the nickase is a Cas nickase. In some preferred examples, the systems further comprise one or more RNA guide sequences.

In some preferred examples, the nucleic acid construct encoding an enzyme that facilitates insertion of the nucleic acid expression construct at the dock site is provided in a vector. In some preferred examples, the vector is a different vector from the vector comprising the nucleic acid expression construct. In some preferred examples, the vector is the same vector as the vector comprising the nucleic acid expression construct.

In some preferred examples, the systems further comprise at least a second nucleic acid expression construct encoding a different protein of interest. In some preferred examples, the second nucleic acid expression construct further comprises the following elements in operable association in 5' to 3' order: a selectable marker sequence; an internal promoter sequence; a nucleic acid sequence encoding a first protein of interest that is operably linked to the internal promoter; and a poly A signal sequence, wherein the nucleic acid construct further comprises at least one insertion element, and wherein the expression construct insertion elements are compatible with the dock site insertion elements to facilitate insertion of the nucleic acid expression construct at the dock site. In some preferred examples, the at least a second nucleic acid expression construct encoding a different protein of interest is provided on a separate vector. In some preferred examples, suitable internal promoters include, but are not limited to, SV40, EF1α, *E. coli* lac, *E. coli* trp, phage lambda PL, phage lambda PR, T3, T7, cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, alpha-lactalbumin, and mouse metallothionein-I promoter sequences.

In some preferred examples, the nucleic acid expression construct included in the system further comprises a nucleic acid sequence encoding at least a second protein of interest (e.g., including third, fourth, fifth, etc. proteins of interest). In some preferred examples, the nucleic acid sequence encoding the second protein of interest is operably

associated with a construct element selected from the group consisting of a third promoter, an intron, a second RNA export element and an IRES sequence and combinations thereof. In some preferred examples, suitable third promoters include, but are not limited to, SV40, EF1α, *E. coli* lac, *E. coli* trp, phage lambda PL, phage lambda PR, T3, T7, cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, alpha-lactalbumin, and mouse metallothionein-I promoter sequences. In some preferred examples, the second RNA export element is a pre-mRNA processing enhancer (PPE). In some preferred examples, the second nucleic acid sequence encoding a second protein of interest is positioned 3' to the nucleic acid sequence encoding the first protein of interest. In some preferred examples, the IRES sequence is selected from the group consisting of foot and mouth disease virus (FDV), encephalomyocarditis virus and poliovirus IRES sequences.

In one aspect, the present disclosure refers to a method comprising: providing a host cell(s) as described herein and introducing into the host cell(s) one or more nucleic acid expression constructs encoding a first protein of interest under conditions such that the nucleic acid expression constructs are inserted at the dock sites, wherein the nucleic acid expression construct further comprises at least the following elements in operable association in 5' to 3' order: a selectable marker sequence; an internal promoter sequence; a nucleic acid sequence encoding the first protein of interest that is operably linked to the internal promoter; and a poly A signal sequence, wherein the nucleic acid construct further comprises at least one insertion element, and wherein the expression construct insertion elements are compatible with the dock site insertion elements to facilitate insertion of the nucleic acid expression construct at the dock site. In some preferred examples, suitable internal promoters include, but are not limited to, SV40, EF1α, *E. coli* lac, *E. coli* trp, phage lambda PL, phage lambda PR, T3, T7, cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, alpha-lactalbumin, and mouse metallothionein-I promoter sequences.

In some preferred examples, the one or more nucleic acid expression constructs further comprise a 5' promotor sequence 5' to the selectable marker sequence. In some preferred examples, the nucleic acid expression constructs are provided in a vector. In some preferred examples, the vector is a plasmid vector. Other suitable vectors include cosmids, YACs and small circular constructs. In some preferred examples, the vector is transiently introduced into the host cell. In some preferred examples, the 5' promoter is a weak promoter. In some preferred examples, suitable 5' promoters include, but are not limited to, SIN-LTR, SV40, EF1α, *E. coli* lac, *E. coli* trp, phage lambda PL, phage lambda

PR, T3, T7, cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, alpha-lactalbumin, and mouse metallothionein-I promoter sequences.

In some preferred examples, the host cell line used in the method comprises a nucleic acid construct encoding an enzyme that facilitates insertion of the nucleic acid expression construct at the dock site. In some preferred examples, the methods further comprise transiently introducing into the host cell a nucleic acid construct encoding an enzyme that facilitates insertion of the nucleic acid expression construct at the dock site. In some preferred examples, the nucleic acid construct encoding an enzyme that facilitates insertion of the nucleic acid expression construct at the dock site is provided in a vector. In some preferred examples, the vector is a plasmid vector. In some preferred examples, the ratio of the ratio of the nucleic acid constructs encoding an enzyme that facilitates insertion of the nucleic acid expression construct at the dock site to the nucleic acid expression constructs encoding a first protein of interest that are transiently introduced into the host cell line is from 1:1000 to 1:10. In some preferred examples, the ratio of the ratio of the nucleic acid constructs encoding an enzyme that facilitates insertion of the nucleic acid expression construct at the dock site to the nucleic acid expression constructs encoding a first protein of interest that are transiently introduced into the host cell line is from 1:100 to 1:750. In some preferred examples, the ratio of the ratio of the nucleic acid constructs encoding an enzyme that facilitates insertion of the nucleic acid expression construct at the dock site to the nucleic acid expression constructs encoding a first protein of interest that are transiently introduced into the host cell line is from 1:400 to 1:600.

In some preferred examples, the enzyme is selected from the group consisting of an integrase, a recombinase, a nuclease and a nickase. In some preferred examples, the nuclease is a Cas nuclease. In some preferred examples, the nickase is a Cas nickase. In some preferred examples, the methods further comprise introducing into the host cell one or more RNA guide sequences.

In some preferred examples, the nucleic acid construct encoding an enzyme that facilitates insertion of the nucleic acid expression construct at the dock site is provided in a vector. In some preferred examples, the vector is a different vector from the vector comprising the nucleic acid expression construct. In some preferred examples, the vector is the same vector as the vector comprising the nucleic acid expression construct.

In some preferred examples, the methods further comprise introducing into the host cell at least a second nucleic acid expression construct encoding a different protein of interest. In some preferred examples, the at least a second nucleic acid expression

construct encoding a different protein of interest is provided on a separate vector. In some preferred examples, the second nucleic acid expression construct comprises the following elements in operable association in 5' to 3' order: a selectable marker sequence; an internal promoter sequence; a nucleic acid sequence encoding a first protein of interest that is operably linked to the internal promoter; and a poly A signal sequence, wherein the nucleic acid construct further comprises at least one insertion element, and wherein the expression construct insertion elements are compatible with the dock site insertion elements to facilitate insertion of the nucleic acid expression construct at the dock site. In some preferred examples, the nucleic acid expression construct further comprises a nucleic acid sequence encoding at least a second protein of interest. In some preferred examples, suitable internal promoters include, but are not limited to, SV40, EF1α, *E. coli* lac, *E. coli* trp, phage lambda PL, phage lambda PR, T3, T7, cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, alpha-lactalbumin, and mouse metallothionein-I promoter sequences.

In some preferred examples, the nucleic acid sequence encoding the second protein of interest is operably associated with a construct element selected from the group consisting of a third promoter, an intron, a second RNA export element and an IRES sequence and combinations thereof. In some preferred examples, suitable third promoters include, but are not limited to, SV40, EF1α, *E. coli* lac, *E. coli* trp, phage lambda PL, phage lambda PR, T3, T7, cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, alpha-lactalbumin, and mouse metallothionein-I promoter sequences. In some preferred examples, the second RNA export element is a pre-mRNA processing enhancer (PPE). In some preferred examples, the second nucleic acid sequence encoding a second protein of interest is positioned 3' to the nucleic acid sequence encoding the first protein of interest. In some preferred examples, the IRES sequence is selected from the group consisting of foot and mouth disease virus (FDV), encephalomyocarditis virus and poliovirus IRES sequences.

In some preferred examples, the present disclosure describes methods for making a host cell line comprising inserting into the genome of the host cell multiple docking sites, each docking site comprising at least one dock site insertion element. In some preferred examples, the inserting comprises insertion of the docking sites via a vector selected from the group consisting of a viral vector and a transposon vector. In some preferred examples, the viral vector is a retroviral vector. In some preferred examples, from 1 to 500 integrated docking sites are inserted, each docking site comprising at least one dock

site insertion element. In some preferred examples, from 5 to 500 integrated docking sites are inserted, each docking site comprising at least one dock site insertion element. In some preferred examples, the integrated docking sites are independently inserted throughout the genome.

In some preferred examples, the dock site insertion element is targeted by enzyme selected from the group consisting of an integrase, a recombinase, a nuclease and a nickase. In some preferred examples, the dock site insertion element is selected from the group consisting of a recombinase dock site insertion element and a HDR dock site insertion element. In some preferred examples, the dock site insertion element is a recombinase dock site insertion element. In some preferred examples, the recombinase dock site insertion element comprises an attachment site (att). In some preferred examples, the attachment site (att) is selected from the group consisting of attB and attP. In some preferred examples, the attachment site (att) is selected from the group consisting of attR and attL. In some preferred examples, the recombinase dock site insertion element comprises a LoxP sequence. In some preferred examples, the recombinase dock site insertion element is a Flp Recombination Target (FRT) site. In some preferred examples, the dock site insertion element is a HDR dock site insertion element. In some preferred examples, the HDR dock site insertion element comprises one or two dock site homology arms. In some preferred examples, the HDR dock site insertion element further comprises one or more sequences homologous to a guide RNA sequence. In some preferred examples, the dock site homology arms are from about 30 to 1000 bases in length. In some preferred examples, the integrase dock site insertion element comprises an AAVS1 safe harbor locus sequence.

In some preferred examples, each docking site is flanked by exogenous integrating vector sequences following insertion. In some preferred examples, the exogenous integrating vector sequences are selected from the group consisting of viral vector sequences and transposon vector sequences.

In some preferred examples, the docking sites each further comprise a sequence encoding a selectable maker operably linked to a promoter.

In some preferred examples, the docking site comprises a promoter. In some preferred examples, the promoter is a weak promoter. In some preferred examples, suitable promoters included in the docking site include, but are not limited to, SIN-LTR, SV40, EF1α, *E. coli* lac, *E. coli* trp, phage lambda PL, phage lambda PR, T3, T7,

cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, alpha-lactalbumin, and mouse metallothionein-I promoter sequences.

In some preferred examples, each docking site comprises two dock site insertion elements. In some preferred examples, the two dock site insertion elements are positioned to facilitate cassette exchange. In some preferred examples, the two dock site insertion elements flank sequences encoding a selectable marker, an enzyme, or a combination thereof.

DESCRIPTION OF THE FIGURES

Abbreviations used in figures:

AmpR= bacterial ampicillin resistance gene

attB= Bacterial Attachment Site

attP= Phage Attachment Site

attR= Recombined Upstream Attachmend Site

Backbone=Plasmid Backbone

CDS= Coding Sequence

EPR=MMLV Extended Packaging Region

GCI = Gene Copy Index

GS= Glutamine Synthetase

H or HC= Heavy Chain

hCMV= Human Cytomegalovirus immediate-early Promoter

I= intron

L or LC= Light Chain

MoMuSV 5'LTR= Moloney Murine Sarcoma Virus 5' Long Terminal Repeat

Neo= Neomycin resistance genePA or PolyA= Polyadenylation signal

ProV SIN-LTR= Proviral Self- Inactivating Long Terminal Repeat

sCMV= Simian Cytomegalovirus immediate-early Promoter

SDS-PAGE= Sodium Dodecyl Sulphate- Polyacrylamide Gel Electrophoresis

SIN-3'LTR= Self-Inactivation 3' Long Terminal Repeat

SV40= Simian Virus 40

TK= Thymidine Kinase

UTR= Untranslated Region

W or WPRE= Woodchuck Post-transcriptional Regulatory Element

FIG. 1. Nucleic acid construct design for certain embodiments of the invention

CLAIMS

What is claimed is:

1. A host cell comprising a genome, the genome comprising from 1 to 500 integrated docking sites, each docking site comprising at least one dock site insertion element.

2. The host cell of claim 1, wherein the genome comprises from 5 to 500 integrated docking sites, each docking site comprising at least one dock site insertion element.

3. The host cell of claim 1, wherein the integrated docking sites are independently positioned throughout the genome.

4. The host cell of claim 1, wherein the dock site insertion element is targeted by enzyme selected from the group consisting of an integrase, a recombinase, a nuclease and a nickase.

5. The host cell of claim 1, wherein the dock site insertion element is selected from the group consisting of a recombinase dock site insertion element and a HDR dock site insertion element.

6. The host cell of claim 5, wherein the dock site insertion element is a recombinase dock site insertion element.

7. The host cell of claim 6, wherein the recombinase dock site insertion element comprises an attachment site (att).

8. The host cell of claim 7, wherein the attachment site (att) is selected from the group consisting of attB and attP and attR and attL.

9. The host cell of claim 6, wherein the recombinase dock site insertion element comprises a LoxP sequence.

10. The host cell line of claim 6, wherein the recombinase dock site insertion element is a Flp Recombination Target (FRT) site.

11. The host cell line of claim 1, wherein each docking site is flanked by exogenous integrating vector sequences.

12. The host cell line of claim 11, wherein the exogenous integrating vector sequences are selected from the group consisting of viral vector sequences and transposon vector sequences.

13. The host cell line of claim 1, wherein the docking sites each further comprise a sequence encoding a selectable maker operably linked to a promoter.

14. The host cell line of claim 1, wherein the host cell line further comprises an exogenous sequence encoding an enzyme operably linked to a promoter.

15. The host cell line of claim 1, wherein the exogenous sequence encoding an enzyme operably linked to a promoter is inserted into the genome of the host cell.

16. The host cell line of claim 15, wherein the exogenous sequence encoding an enzyme operably linked to a promoter is inserted at the dock site.

17. The host cell line of claim 1, wherein the exogenous sequence encoding an enzyme operably linked to a promoter is provided in an episomal expression vector.

18. The host cell line of claim 17, wherein the exogenous enzyme is selected from the group consisting of an integrase, a recombinase, a nuclease and a nickase.

19. The host cell line of claim 1, wherein the dock site insertion element is positioned to facilitate cassette exchange.

20. The host cell line of claim 1, wherein each docking site comprises two dock site insertion elements.

21. The host cell line of claim 20, wherein the two dock site insertion elements are positioned to facilitate cassette exchange.

22. The host cell line of claim 1, further comprising nucleic acid expression constructs inserted at 1 or more of the docking sites.

23. The host cell line of claim 22, wherein the nucleic acid expression construct further comprises the following elements in operable association in 5' to 3' order:

a first promoter sequence;

a selectable marker sequence;

a second promoter sequence;

a nucleic acid sequence encoding a first protein of interest that is operably linked to the internal promoter; and

a poly A signal sequence.

24. The host cell line of claim 23, wherein the nucleic acid construct further comprising at least one insertion element at a position or positions selected from the group consisting of 5' to the first promoter, 3' to the poly A signal sequence, between the first promoter and the poly A signal sequence, between the selectable marker and the second promoter sequence, and both 5' to the first promoter and 3' to the poly A signal sequence.

25. A cell culture comprising host cells of claim 1.

26. A process for producing a protein of interest comprising culturing host cells according to claim 22 under conditions that the protein of interest is expressed and purifying the protein of interest from the host cell culture.

27. A system comprising:

a host cell of claim 1;

one or more nucleic acid expression constructs, wherein the nucleic acid expression construct further comprises the following elements in operable association in 5' to 3' order:

a selectable marker sequence;

an internal promoter sequence;

a nucleic acid sequence encoding a first protein of interest that is operably linked to the internal promoter; and

a poly A signal sequence,

wherein the nucleic acid construct further comprises at least one insertion element, and

wherein the expression construct insertion elements are compatible with the dock site insertion elements to facilitate insertion of the nucleic acid expression construct at the dock site.

28. A method comprising:

providing a host cell of claim 1;

introducing into the host cell one or more nucleic acid expression constructs encoding a first protein of interest under conditions such that the nucleic acid expression constructs are inserted at the dock sites, wherein the nucleic acid expression construct further comprises at least the following elements in operable association in 5' to 3' order:

a selectable marker sequence;

an internal promoter sequence;

a nucleic acid sequence encoding the first protein of interest that is operably linked to the internal promoter; and

a poly A signal sequence,

wherein the nucleic acid construct further comprises at least one insertion element, and

wherein the expression construct insertion elements are compatible with the dock site insertion elements to facilitate insertion of the nucleic acid expression construct at the dock site.