events can result in a variety of gene expression and cell growth phenotypes. Such variation, known as "position effect variation," originates, at least in part, from the complex gene regulatory networks present in eukaryotic cell genomes and the accessibility of certain genomic loci for integration and gene expression. Second, random integration strategies generally do not offer control over the number of gene copies integrated into a host cell genome. In fact, gene amplification methods are often used to achieve high-producing cells. Such gene amplification, however, can lead to unwanted cell phenotypes such as unstable cell growth and/or product expression. Third, because of the integration loci heterogeneity inherent in the random integration process, it is time-consuming and labor-intensive to screen thousands of clones after transfection to isolate cell lines demonstrating a desirable level of expression of the polypeptides of interest. Even after isolating such cell lines, stable expression of the polypeptide of interest is not guaranteed and further screening may be required to obtain a stable commercial cell line. Finally, polypeptides produced from randomly integrated cell lines exhibit a high degree of sequence variance, which may be, in part, due to the mutagenicity of the selective agents used to select for a high level of expression of polypeptides of interest.

SUMMARY

The presently disclosed subject matter relates to targeted integration (TI) host cells suitable for the expression of recombinant proteins. The presently disclosed subject matter not only provides host cell TI sites that have high productivity, it also provides a novel method of introducing multiple sequences of interest into a single TI locus in a host cell by recombinase-mediated cassette exchange (RMCE). The advantages of the RMCE method include increased productivity and the flexibility to co-express multiple polypeptides in varying ratios from the same TI locus. For example, but not by way of limitation, the RMCE strategies disclosed herein allow for the insertion of one, two, three, four, five, six, seven, eight or more sequences (e.g., antibody heavy chain (HC) or light chain (LC) sequences) to the TI locus. With the ability to target eight or more sequences simultaneously, the two-plasmid RMCE enables modulation of HC and LC chain ratios of mAb to improve productivity, expression of complex molecules with multiple chains, and targeting of transgenes, endogenous genes or RNAi with the antibody to modify cellular pathways.

In one aspect, provided herein is a targeted integration (TI) Chinese hamster ovary (CHO) host cell comprising an exogenous nucleotide sequence integrated at an integration site within a specific locus of the genome of the host cell, wherein the locus is at least 90% homologous to the sequence of Contig NW_006874047.1.

In another aspect, provided herein is a method of preparing a TI CHO host cell expressing a polypeptide of interest comprising: (a) providing a TI CHO host cell comprising an exogenous nucleotide sequence integrated at a site within a locus of the genome of the TI host cell, wherein the locus is at least 90% homologous to the sequence of Contig NW_006874047.1, wherein the exogenous nucleotide sequence comprises two RRSs, flanking at least one first selection marker; (b) introducing into the cell provided in (a) a vector comprising two RRSs matching the two RRSs on the integrated exogenous nucleotide sequence and flanking at least one exogenous SOI and at least one second selection marker; (c) introducing a recombinase or a nucleic acid encoding a recombinase, wherein the recombinase recognizes the RRSs; and (d) selecting for TI cells expressing the second selection marker to thereby isolate a TI host cell expressing the polypeptide.

In another aspect, provided herein is a method for preparing a TI CHO host cell expressing at least a first and second polypeptide of interest comprising: (a) providing a TI CHO host cell comprising an exogenous nucleotide sequence integrated at a site within a locus of the genome of the host cell, wherein the locus is at least 90% homologous to the sequence of Contig NW_006874047.1, wherein the exogenous nucleotide sequence comprises a first and a second RRS flanking at least one first selection marker, and a third RRS located between the first and the second RRS, and all the RRSs are heterospecific; (b) introducing into the cell provided in (a) a first vector comprising two RRSs matching the first and the third RRS on the integrated exogenous nucleotide sequence and flanking at least one first exogenous SOI and at least one second selection marker; (c) introducing into the cell provided in (a) a second vector comprising two RRSs matching the second and the third RRS on the integrated exogenous nucleotide sequence and flanking at least one second exogenous SOI; optionally wherein the first vector further comprises a promoter sequence operably linked to the codon ATG positioned flanked upstream by the first SOI and downstream by an RRS, and the second vector further comprises a selection marker lacking an ATG transcription start codon flanked upstream by an RRS and downstream by the second SOI; (d) introducing one or more recombinases, or one or more nucleic acids encoding one or more recombinases, wherein the one or more recombinases recognize the RRSs; and (e) selecting for TI cells expressing the second selection marker to thereby isolate a TI host cell expressing the first and second polypeptides of interest.

In one aspect, provided herein is a method for expressing a polypeptide of interest comprising: (a) providing a TI CHO host cell comprising at least one exogenous SOI and at least one selection marker where the at least one exogenous SOI and at least one selection marker are

flanked by two RRSs integrated within a locus of the genome of the TI host cell, wherein the locus is at least 90% homologous to the sequence of Contig NW_006874047.1; and (b) the cell in (a) under conditions suitable for expressing the polypeptide of interest and recovering a polypeptide of interest therefrom.

In another aspect, provided herein is a method for expressing a polypeptide of interest comprising: (a) providing a TI CHO host cell comprising at least two exogenous SOIs and at least one selection marker integrated within a locus of the genome of the host cell, wherein the locus is at least 90% homologous to the sequence of Contig NW_006874047.1, wherein at least one exogenous SOI and one selection marker is flanked by a first and a third RRS and at least one exogenous SOI is flanked by a second and the third RRS; and (b) culturing the cell in (a) under conditions suitable for expressing the polypeptide of interest and recovering a polypeptide of interest therefrom.

What is claimed is:

1. A targeted integration (TI) Chinese hamster ovary (CHO) host cell comprising an exogenous nucleotide sequence integrated at an integration site within a specific locus of the genome of the host cell, wherein the locus is at least 90% homologous to the sequence of Contig NW_006874047.1.

2. The TI host cell of claim 1, wherein the nucleotide sequence immediately 5' of the integrated exogenous nucleotide sequence is at least 90% homologous to nucleotides 41190-45269 of NW 006874047.1.

3. The TI host cell of claim 1 wherein the nucleotide sequence immediately 5' of the integrated exogenous nucleotide sequence is at least 15 base pairs, at least 20 base pairs, at least 30 base pairs, at least 40 base pairs, at least 50 base pairs, at least 75 base pairs, at least 100 base pairs, at least 150 base pairs, at least 200 base pairs, at least 300 base pairs, at least 400 base pairs, at least 500 base pairs, at least 1,000 base pairs, at least 1,500 base pairs, at least 2,000 base pairs, at least 3,000 base pairs from nucleotide 45269 of NW_006874047.1.

4. The TI host cell of claim 1, wherein the nucleotide sequence immediately 3' of the integrated exogenous nucleotide sequence is at least 90% homologous to nucleotides 45270-45490 of NW 006874047.1.

5. The TI host cell of claim 1 wherein the nucleotide sequence immediately 3' of the integrated exogenous nucleotide sequence is at least 15 base pairs, at least 20 base pairs, at least 30 base pairs, at least 40 base pairs, at least 50 base pairs, at least 75 base pairs, at least 100 base pairs, at least 150 base pairs, at least 200 base pairs, at least 300 base pairs, at least 400 base pairs, at least 500 base pairs, at least 1,000 base pairs, at least 1,500 base pairs, at least 2,000 base pairs, at least 3,000 base pairs from nucleotide 45270 of NW_006874047.1

6. The TI host cell of any one of claims 1-5, wherein the TI host cell is a CHO K1 host cell, a CHO K1SV host cell, a DG44 host cell, a DUKXB-11 host cell, a CHOK1S host cell, or a CHO K1M host cell.

7. The TI host cell of any one of claims 1-6, wherein the exogenous nucleotide sequence comprises one or more recombination recognition sequences (RRSs), particularly at least two RRSs, wherein the RRS can be recognized by a recombinase; optionally wherein the recombinase is a Cre recombinase or an FLP recombinase, or a Bxb1 integrase or a φC31 integrase; and/or the RRS is

selected from the group consisting of a LoxP sequence, a LoxP L3 sequence, a LoxP 2L sequence, a LoxFas sequence, a Lox511 sequence, a Lox2272 sequence, a Lox2372 sequence, a Lox5171 sequence, a Loxm2 sequence, a Lox71 sequence, a Lox66 sequence, a FRT sequence, a Bxb1 attP sequence, a Bxb1 attB sequence, a ϕC31 attP sequence, and a ϕC31 attB sequence.

8. The TI host cell of claim 7, wherein the exogenous nucleotide sequence comprises a first and a second RRS, and at least one selection marker located between the first and the second RRS, particularly wherein first selection marker is selected from the group consisting of aminoglycoside phosphotransferase (APH), hygromycin phosphotransferase (HYG), neomycin, G418 APH), dihydrofolate reductase (DHFR), thymidine kinase (TK), glutamine synthetase (GS), asparagine synthetase, tryptophan synthetase (indole), histidinol dehydrogenase (histidinol D), blasticidin, bleomycin, phleomycin, chloramphenicol, Zeocin, and mycophenolic acid;

and optionally further comprises

(a) a second selection marker, wherein the first and the second selection markers are different, and particularly wherein the second selection marker is selected from the group consisting of aminoglycoside phosphotransferase (APH), hygromycin phosphotransferase (HYG), neomycin, G418 APH), dihydrofolate reductase (DHFR), thymidine kinase (TK), glutamine synthetase (GS), asparagine synthetase, tryptophan synthetase (indole), histidinol dehydrogenase (histidinol D), blasticidin, bleomycin, phleomycin, chloramphenicol, Zeocin, and mycophenolic acid;

(b) a second selection marker as in (a) and a third selection marker and an internal ribosome entry site (IRES), wherein the IRES is operably linked to the third selection marker, and particularly wherein the third selection marker is different from the first or the second selection marker and/or is selected from the group consisting of a green fluorescent protein (GFP) marker, an enhanced GFP (eGFP) marker, a synthetic GFP marker, a yellow fluorescent protein (YFP) marker, an enhanced YFP (eYFP) marker, a cyan fluorescent protein (CFP) marker, a mPlum marker, a mCherry marker, a tdTomato marker, a mStrawberry marker, a J-red marker, a DsRed-monomer marker, a mOrange marker, a mKO marker, a mCitrine marker, a Venus marker, a YPet marker, an Emerald6 marker, a CyPet marker, a mCFPm marker, a Cerulean marker, and a T-Sapphire marker; and/or

(c) a third RRS, wherein the third RRS is located between the first and the second RRS, and the third RRS is heterospecific relative the first or the second RRS.

9. The TI host cell of claim 7 or 8, wherein the exogenous nucleotide sequence further comprises

(a) at least one exogenous sequence of interest (SOI), particularly wherein the exogenous SOI is located between the first and the second RRS; or

(b) at least one exogenous SOI located between the first and the third RRS, and at least one exogenous SOI located between the third and the second RRS;

wherein optionally in (a) and (b) the SOI encodes a single chain antibody, an antibody light chain, an antibody heavy chain, a single-chain Fv fragment (scFv), or an Fc fusion protein.

10. A method of preparing a TI host cell expressing a polypeptide of interest comprising:

(a) providing a TI host cell comprising an exogenous nucleotide sequence integrated at a site within a locus of the genome of the TI host cell, wherein the locus is at least 90% homologous to the sequence of Contig NW_006874047.1, wherein the exogenous nucleotide sequence comprises two RRSs, flanking at least one first selection marker;

(b) introducing into the cell provided in (a) a vector comprising two RRSs matching the two RRSs on the integrated exogenous nucleotide sequence and flanking at least one exogenous SOI and at least one second selection marker;

(c) introducing a recombinase or a nucleic acid encoding a recombinase, wherein the recombinase recognizes the RRSs; and

(d) selecting for TI cells expressing the second selection marker to thereby isolate a TI host cell expressing the polypeptide.

11. A method for preparing a TI CHO host cell expressing at least a first and second polypeptide of interest comprising:

(a) providing a TI CHO host cell comprising an exogenous nucleotide sequence integrated at a site within a locus of the genome of the host cell, wherein the locus is at least 90% homologous to the sequence of Contig NW_006874047.1, wherein the exogenous nucleotide sequence comprises a first and a second RRS flanking at least one first selection marker, and a third RRS located between the first and the second RRS, and all the RRSs are heterospecific;

(b) introducing into the cell provided in (a) a first vector comprising two RRSs matching the first and the third RRS on the integrated exogenous nucleotide sequence and flanking at least one first exogenous SOI and at least one second selection marker;

(c) introducing into the cell provided in (a) a second vector comprising two RRSs matching the second and the third RRS on the integrated exogenous nucleotide sequence and flanking at least one second exogenous SOI;

optionally wherein the first vector further comprises a promoter sequence operably linked to the codon ATG positioned flanked upstream by the first SOI and downstream by an RRS, and the second vector further comprises a selection marker lacking an ATG transcription start codon flanked upstream by an RRS and downstream by the second SOI;

(d) introducing one or more recombinases, or one or more nucleic acids encoding one or more recombinases, wherein the one or more recombinases recognize the RRSs; and

(e) selecting for TI cells expressing the second selection marker to thereby isolate a TI host cell expressing the first and second polypeptides of interest.

12. The method of claim 10 or 11, wherein

(i) the recombinase is a Cre recombinase or an FLP recombinase, or is a Bxb1 integrase or a φC31 integrase;

(ii) the (first and/or the second) SOI encodes a single chain antibody, an antibody light chain, an antibody heavy chain, a single-chain Fv fragment (scFv), or an Fc fusion protein; and/or

(iii) the TI host cell is a CHO K1 host cell, a CHO K1SV host cell, a DG44 host cell, a DUKXB-11 host cell, a CHOK1S host cell, or a CHO K1M host cell.

13. A method for expressing a polypeptide of interest comprising:

(a) providing a TI CHO host cell comprising at least one exogenous SOI and at least one selection marker where the at least one exogenous SOI and at least one selection marker are flanked by

two RRSs integrated within a locus of the genome of the TI host cell, wherein the locus is at least 90% homologous to the sequence of Contig NW_006874047.1; and

(b) culturing the cell in (a) under conditions suitable for expressing the polypeptide of interest and recovering a polypeptide of interest therefrom.

14. A method for expressing a polypeptide of interest comprising:

(a) providing a TI CHO host cell comprising at least two exogenous SOIs and at least one selection marker integrated within a locus of the genome of the host cell, wherein the locus is at least 90% homologous to the sequence of Contig NW_006874047.1, wherein at least one exogenous SOI and one selection marker is flanked by a first and a third RRS and at least one exogenous SOI is flanked by a second and the third RRS; and

(b) culturing the cell in (a) under conditions suitable for expressing the polypeptide of interest and recovering a polypeptide of interest therefrom.

15. The method of claim 13 or 14, wherein

(i) the SOI encodes a single chain antibody, an antibody light chain, an antibody heavy chain, a single-chain Fv fragment (scFv), or an Fc fusion protein; and/or

(ii) the TI host cell is a CHO K1 host cell, a CHO K1SV host cell, a DG44 host cell, a DUKXB-11 host cell, a CHOK1S host cell, or a CHO K1M host cell.