

expression levels without excessive down-regulation, be amplifiable using metabolic selection markers such as DHFR or GS in conjunction with MTX or MSX, and will be located so that integration of transgenes does not negatively impact cell growth or protein product profiles.

[006] The S100A6 gene is part of the S100A gene cluster encoding a group of known calcium-binding proteins, e.g. S100A1, S100A13, S100A14, S100A16, S100A3, S100A2, S100A4, S100A5 and S100A6. The cluster comprises a "side cluster" including the S100A1, S100A13, S100A14 and S100A16 genes and a "main cluster", which includes the S100A3, S100A4, S100A5 and S100A6 genes.

[007] In the present invention, it is shown that the stable integration of heterologous polynucleotides within the S100A gene cluster of the CHO cell genome increases the production of a heterologous gene product. Specifically, stable integration within the upstream and downstream regions flanking the S100A3/A4/A5/A6 main gene cluster, enables a predictable, high level and stable production of a heterologous gene product, including recombinant proteins, such as antibodies and fusion proteins, or regulatory RNAs, such as shRNAs or miRNAs.

SUMMARY

[008] In an aspect, the present disclosure refers to a Chinese hamster ovary (CHO) cell, comprising at least one heterologous polynucleotide, stably integrated into the S100A gene cluster of the CHO cell genome is described, wherein the at least one heterologous polynucleotide is integrated upstream of the S100A3/A4/A5/A6 main gene cluster, into a genomic target region corresponding to the sequence of SEQ ID NO: 1; and/or the at least one heterologous polynucleotide is integrated downstream of the S100A3/A4/A5/A6 main gene cluster, into a genomic target region corresponding to the sequence of nucleotides 1 to 15,120 of SEQ ID NO: 2. Preferably the upstream genomic target region corresponds to nucleotides 30 to 19,000 of SEQ ID NO: 1, nucleotides 2,940 to 19,000 of SEQ ID NO: 1, nucleotides 4,740 to 19,000 of SEQ ID NO: 1, nucleotides 6,480 to 19,000 of SEQ ID NO: 1, nucleotides 8,280 to 19,000 of SEQ ID NO: 1, nucleotides 10,020 to 19,000 of SEQ ID NO: 1, or nucleotides 11,820 to 19,000 of SEQ ID NO: 1; and/or the downstream genomic target region corresponds to nucleotides 1 to 13,160 of SEQ ID NO: 2, nucleotides 1 to 12,000 of SEQ ID NO: 2 or nucleotides 1 to 10,260 of SEQ ID NO: 2.

[009] More preferably the upstream genomic target region corresponds to nucleotides 11,820 to 18,720 of SEQ ID NO: 1, nucleotides 13,560 to 18,720 of SEQ ID NO: 1, nucleotides 15,360 to 18,720 of SEQ ID NO: 1 or nucleotides 17,100 to 18,720 of SEQ ID NO: 1; and/or the downstream genomic target region corresponds to nucleotides 660 to 10,260 of SEQ ID NO: 2, nucleotides 1,320 to 10,260 of SEQ ID NO: 2 or nucleotides 1,480 to 10,260 of SEQ ID NO: 2. Even more preferably the upstream genomic target region corresponds to nucleotides 11,820 to 18,380 of SEQ ID NO: 1, nucleotides 13,560 to 18,380 of SEQ ID NO: 1, nucleotides 15,360 to 18,380 of SEQ ID NO: 1 or nucleotides 17,100 to 18,380 of SEQ ID NO: 1; and/or the downstream genomic target region

corresponds to nucleotides 3,180 to 10,260 of SEQ ID NO: 2, nucleotides 4,920 to 9,000 of SEQ ID NO: 2 or nucleotides 6,720 to 8,460 of SEQ ID NO: 2.

[010] In one example, the at least one heterologous polynucleotide is stably integrated into the CHO cell genome as part of an expression cassette. The at least one heterologous polynucleotide may code for a RNA and/or a protein. The RNA may be an mRNA, a miRNA or a shRNA. The protein may be a therapeutic protein, preferably a therapeutic protein selected from the group consisting of an antibody, a fusion protein, a cytokine and a growth factor.

[011] The at least one heterologous polynucleotide may also be a marker gene selected from the group consisting of a reporter gene and a selection marker gene. Preferably the marker gene is stably integrated into the CHO cell genome as part of an expression cassette and the expression cassette is flanked by recognition sites for a site specific recombinase or a sequence specific DNA editing enzyme such as a site specific nuclease.

[012] The CHO cell according to the invention may be a CHO-DG44 cell, a CHO-K1 cell, a CHO-DXB11 cell, a CHO-S cell, a CHO glutamine synthetase (GS)-deficient cell or a derivative of any of these cells.

[013] In one example, the genomic target region consists of any one of the sequences defined in SEQ ID NO: 1 and/or SEQ ID NO: 2 above or a sequence having at least 80% sequence identity thereto.

[014] The at least one heterologous polynucleotide may be stably integrated into one or both alleles of the S100A gene cluster of the CHO cell genome.

[015] In another aspect, the present disclosure refers to a method for the production of a CHO cell, comprising the steps of (a) providing a CHO cell; (b) introducing a heterologous polynucleotide into said CHO cell, wherein the heterologous polynucleotide is stably integrated into the S100A gene cluster of the CHO cell genome, wherein said heterologous polynucleotide is integrated upstream of the S100A3/A4/A5/A6 main gene cluster, into a genomic target region corresponding to the sequence of SEQ ID NO: 1; and/or said heterologous polynucleotide is integrating downstream of the S100A3/A4/A5/A6 main gene cluster, into a genomic target region corresponding to the sequence of nucleotides 1 to 15,120 of SEQ ID NO: 2.

[016] Preferably the upstream genomic target region corresponds to nucleotides 30 to 19,000 of SEQ ID NO: 1, nucleotides 2,940 to 19,000 of SEQ ID NO: 1, nucleotides 4,740 to 19,000 of SEQ ID NO: 1, nucleotides 6,480 to 19,000 of SEQ ID NO: 1, nucleotides 8,280 to 19,000 of SEQ ID NO: 1, nucleotides 10,020 to 19,000 of SEQ ID NO: 1, or nucleotides 11,820 to 19,000 of SEQ ID NO: 1; and/or the downstream genomic target region corresponds to nucleotides 1 to 13,160 of SEQ ID NO: 2, nucleotides 1 to 12,000 of SEQ ID NO: 2 or nucleotides 1 to 10,260 of SEQ ID NO: 2. More preferably the upstream genomic target region corresponds to nucleotides 11,820 to 18,720 of SEQ ID NO: 1, nucleotides 13,560 to 18,720 of SEQ ID NO: 1, nucleotides 15,360 to 18,720 of SEQ ID NO: 1 or nucleotides 17,100 to 18,720 of SEQ ID NO: 1; and/or the downstream genomic target region corresponds to nucleotides 660 to 10,260 of SEQ ID NO: 2, nucleotides 1,320 to 10,260 of

SEQ ID NO: 2 or nucleotides 1,480 to 10,260 of SEQ ID NO: 2. Even more preferably the upstream genomic target region corresponds to nucleotides 11,820 to 18,380 of SEQ ID NO: 1, nucleotides 13,560 to 18,380 of SEQ ID NO: 1, nucleotides 15,360 to 18,380 of SEQ ID NO: 1, nucleotides 17,100 to 18,380 of SEQ ID NO: 1; and/or the downstream genomic target region corresponds to
5 nucleotides 3,180 to 10,260 of SEQ ID NO: 2, nucleotides 4,920 to 9,000 of SEQ ID NO: 2 or nucleotides 6,720 to 8,460 of SEQ ID NO: 2.

[017] In one example, the genomic target region consists of any one of the sequences defined in SEQ ID NO: 1 and/or SEQ ID NO: 2 above or a sequence having at least 80% sequence identity thereto.

10 [018] In another example, the at least one heterologous polynucleotide is stably integrated into the CHO cell genome as part of an expression cassette and the expression cassette may be flanked by recognition sites for a site specific recombinase or a sequence specific DNA editing enzyme (e.g., a site specific nuclease)

[019] In one example, the at least one heterologous polynucleotide is stably integrated into the CHO
15 cell genome as part of an expression cassette. The at least one heterologous polynucleotide may code for a RNA and/or a protein. The RNA may be an mRNA, a miRNA or a shRNA. The protein may be a therapeutic protein, preferably a therapeutic protein selected from the group consisting of an antibody, a fusion protein, a cytokine and a growth factor.

[020] The at least one heterologous polynucleotide may also be a marker gene selected from the
20 group consisting of a reporter gene and a selection marker gene. Preferably the marker gene is stably integrated into the CHO cell genome as part of an expression cassette and the expression cassette is flanked by recognition sites for a site specific recombinase or a sequence specific DNA editing enzyme (e.g., a site specific nuclease).

[021] The at least one heterologous polynucleotide may be stably integrated into one or both alleles
25 of the S100A gene cluster of the CHO cell genome.

[022] In one example, the heterologous polynucleotide is introduced into the CHO cell genome using (a) a sequence specific DNA editing enzyme, preferably a site specific nuclease, more preferably selected from the group consisting of zinc finger nucleases (ZFNs), meganucleases, transcription activator-like effector nucleases (TALENs) and CRISPR associated nucleases; or (b) a
30 site-specific recombinase, preferably selected from the group consisting of lambda integrase, PhiC31 integrase, Cre, Dre and Flp.

[023] In another example, the method may further comprise the steps of (a) providing a CHO cell; (aa) introducing a first heterologous polynucleotide into said CHO cell, wherein the first heterologous polynucleotide is a marker gene and is stably integrated into the S100A gene cluster of the CHO cell
35 genome as part of an expression cassette flanked by recognition sites for a site specific recombinase or a sequence specific DNA editing enzyme (e.g., a site specific nuclease), wherein (i) said heterologous polynucleotide is integrated upstream of the S100A3/A4/A5/A6 main gene cluster, into a genomic target region corresponding to the sequence of SEQ ID NO: 1; and/or (ii) said

heterologous polynucleotide is integrated downstream of the S100A3/A4/A5/A6 main gene cluster, into a genomic target region corresponding to the sequence of nucleotides 1 to 15,120 of SEQ ID NO: 2; and introducing an expression cassette comprising a second heterologous polynucleotide into said CHO cell by replacing the expression cassette comprising the first heterologous polynucleotide of step aa).

[024] In yet another aspect, the present disclosure refers to a method for the production of a protein of interest in a CHO cell comprising (a) providing the CHO cell of the invention; (b) culturing the CHO cell of step a) in a cell culture medium at conditions allowing production of the protein of interest; (c) harvesting the protein of interest, and (d) optionally purifying the protein of interest.

[025] The CHO cell used in the methods according to the invention may be a CHO-DG44 cell, a CHO-K1 cell, a CHO-DXB11 cell, a CHO-S cell, a CHO glutamine synthetase (GS)-deficient cell or a derivative of any of these cells.

[026] In yet another aspect, the present disclosure refers to a use of the CHO cell as disclosed herein for producing a protein of interest at high yield is described.

DESCRIPTION OF THE FIGURES

FIGURE 1: Random versus targeted integration via ZFN (Pool data) in CHO cells. (A) Shown are IgG1 antibody concentrations from randomly integrated (black bars) versus targeted integrated (white bars) CHO-DG44 cell pools after 3-7 days of fed-batch culture. (B) Shown are IgG1 antibody concentrations from randomly integrated (black bars) versus targeted integrated (white bars) CHOZN GS cell pools after 8-10 days of fed-batch culture. TI pools were enriched using FACS cell sorting, metabolic selections and a second round of FACS. Targeted integration was zinc finger nuclease (ZFN) mediated using zinc finger nuclease pair (ZFN) 13 designed to integrate downstream of the S100A3/A4/A5/A6 main gene cluster.

FIGURE 2: Productivity assessment of independent single CHOZN GS clones for homogeneity of antibody production following (A) targeted integration via ZFN or (B) random integration. Shown are 20 - 24 independent clones, which were obtained via limiting dilution following the respective transfection protocol (TI or RI, respectively). Cells were passaged over 60 days in TTP tubes. The bars represent pooled data from IgG titers in $\mu\text{g/ml}$ of individual clones in fed-batch cultures after 8 days following 0 ($n = 2$) and 60 days ($n = 2$) of passaging. Error bars indicate stability of clones passaged for 0 to 60 days. Targeted integration downstream of the S100A3/A4/A5/A6 main gene cluster using ZNF 13 resulted in more homogenous clonal IgG expression levels and more stable expression over 60 days in culture of the single clones.

FIGURE 3: Effect of integration site on antibody pool titers after TI. (A) Illustration of the location of individual ZFNs and hot spot loci in the S100A gene cluster. Numbers indicate boundaries based on the *Cricetulus griseus* scaffold of CHOZN GS cells having the NCBI Reference Sequence: NW_003613854.1. The arrows indicate the integration site of ZNFs 7 to 14 and are classified into