targeting and generation of cell lines expressing the GOI. The RTS insertion is generally carried out by RI or into a limited number of specific genomic regions, and thus the resulting cell lines are still subject to instability and reduced production over time. Moreover, SSI generally results in a low number of integrated gene copies that could indirectly limit rP production titres.

[0005] One method to increase integrated copies of recombinant genes is referred to as cumulative or accumulative SSI (see e.g., Kameyama et al. Biotechnol. Bioeng. 105:1106-14 (2010), Kawabe et al. Cytotechnology 64:267-79 (2012) and Turan et al. J. Mol. Biol. 402:52-69 (2010)). Such a method can include repeated rounds of RMCE to load up a single site sequentially with multiple copies of rP expression cassettes.

[0006] What are needed in the art are SSI cell lines that incorporate RTS at transcriptionally active and highly stable loci within the genome of the host cell. Such cell lines would be capable of stable and long-term expression of GOI.

[0007] Publications, patents, and patent applications are cited herein, the disclosures of which are incorporated by reference herein in their entireties.

SUMMARY

[0008] The present disclosure is based upon the recognition that the transcriptional output from a transgene insertion site as well as the stability of the expression system thereof will be strongly influenced by the 3-dimensional (3D) structure of the chromatin in that region. The present disclosure describes methods based on this recognition for determination of the structure and confirmation of a genome in 3 dimensions (3D mapping of a genome). The disclosed 3D mapping methods can be carried out through utilization of techniques such as, e.g., Hi-C and other chromosome conformation capture methods (Elzo de Wit and Wouter de Laat. Genes Dev. 2012 26: 11-24) and Promoter Capture Hi-C (Schoenfelder et al. Genome Res 25:582-97 (2015), among others. Methods of utilizing information obtained by the 3D mapping protocols as well as mammalian cells that can be formed by the methods are also described. This application teaches how to generate multi-level 3D genome maps and then use that information to identify optimal genome integration sites for the expression of heterologous genes. For example, by interrogating the mapped 3D genome structure, integration sites likely to exhibit high performance can be identified.

[0009] In one example, the present disclosure describes a mammalian cell that includes an RTS at a high integrating (HI) locus. HI loci are high performance genomic sites identified by the inventors through analysis of the 3D hierarchical structure of genomic chromatin. Beneficially, HI loci are in stable, transcriptionally active environments of the

What is claimed:

1. A method for producing a recombinant cell comprising:
    a. mapping peaks in accessible chromatin of a cell genome;
    b. identifying within the mapped peaks a first set of peaks within active genomic compartments of the accessible chromatin and also within about 30,000 base pairs of a topologically associated domain (TAD) boundary;
    c. defining within the first set of peaks a first high integrating (HI) locus, the first HI locus overlapping a region of the genome that interacts with at least one enhancer element; and
    d. inserting a first recombination target site (RTS) within the first HI locus.

2. A method for producing a recombinant cell comprising:
    a. mapping peaks in accessible chromatin of a cell genome;
    b. identifying within the mapped peaks a first set of peaks within active genomic compartments of the accessible chromatin and also within about 30,000 base pairs of a topologically associated domain (TAD) boundary;
    c. identifying within the accessible chromatin regions of the genome that interact with at least one enhancer element;
    d. defining within the first set of peaks a plurality of high integrating (HI) loci, each HI locus of the plurality overlapping an identified region;
    e. integrating a recombination target site (RTS) into a plurality of cells; and
    f. selecting from the plurality of cells a cell comprising the RTS integrated at an HI locus.

3. The method of claim 1 or 2, wherein the HI loci comprise one of SEQ ID NOs: 1-125 or are within or overlapping about 5,000 base pairs of either the 5’ or 3’ end of any one of SEQ ID NOs: 1-125.

4. The method of any one of claims 1 to 3, further comprising identifying within the first set of peaks those peaks that overlap any transcription start site (TSS) for a gene, the expression product of which or lack thereof is non-vital, and defining a second set of peaks that overlap the genes and are downstream of the TSS, wherein the HI loci are defined within the second set of peaks.

5. The method of any one or claims 1 to 3, further comprising identifying within the first set of peaks a third set of peaks that that do not overlap any genes, wherein the HI loci are defined within the third set of peaks.

6. The method of any one of claims 1 to 5, further comprising transfecting the cell of claim 9 or a plurality of the cells of claim 10 with a first vector comprising an exchangeable cassette encoding a first gene of interest and integrating the first exchangeable cassette within an HI locus of the cell, the method further comprising selecting a recombinant protein producer cell comprising the first exchangeable cassette integrated into the chromosome.

7. The method of any one of claims 1 to 6, further comprising ranking the HI loci, for instance wherein the HI loci are ranked according to one or more of expression level of one or more genes associated with each locus, distance from each locus to the nearest TAD boundary, number of predicted enhancer interactions at each locus, and expression level of mRNA of one or more genes associated with each locus.