US 20250257352A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2025/0257352 A1**

MCNEILL et al. (43) **Pub. Date: Aug. 14, 2025**

(54) **METHODS FOR NOMINATION OF NUCLEASE ON-/OFF-TARGET EDITING LOCATIONS, DESIGNATED "CTL-SEQ" (CRISPR TAG LINEAR-SEQ)**

(71) Applicant: **INTEGRATED DNA TECHNOLOGIES, INC.**, Coralville, IA (US)

(72) Inventors: **Matthew MCNEILL**, Iowa City, IA (US); **Rolf TURK**, Iowa City, IA (US); **Garrett RETTIG**, Coralville, IA (US); **Ellen BLACK**, Swisher, IA (US); **Yongming SUN**, San Ramon, CA (US); **Chris SAILOR**, Cedar Rapids, IA (US); **Yu WANG**, North Grafton, MA (US); **Keith GUNDERSON**, Iowa City, IA (US); **Kyle KINNEY**, Iowa City, IA (US)

(21) Appl. No.: **19/067,623**

(22) Filed: **Feb. 28, 2025**

Related U.S. Application Data

(63) Continuation of application No. 17/382,945, filed on Jul. 22, 2021, now abandoned.

(60) Provisional application No. 63/055,460, filed on Jul. 23, 2020.

Publication Classification

(51) **Int. Cl.**

C12N 15/11	(2006.01)
C12N 9/22	(2006.01)
C12Q 1/6853	(2018.01)

(52) **U.S. Cl.**

CPC ***C12N 15/111*** (2013.01); ***C12N 9/22*** (2013.01); ***C12Q 1/6853*** (2013.01); *C12N 2310/20* (2017.05)

(57) **ABSTRACT**

Described herein are methods for identifying and nominating on- and off-target CRISPR editing sites with improved accuracy and sensitivity.

Specification includes a Sequence Listing.

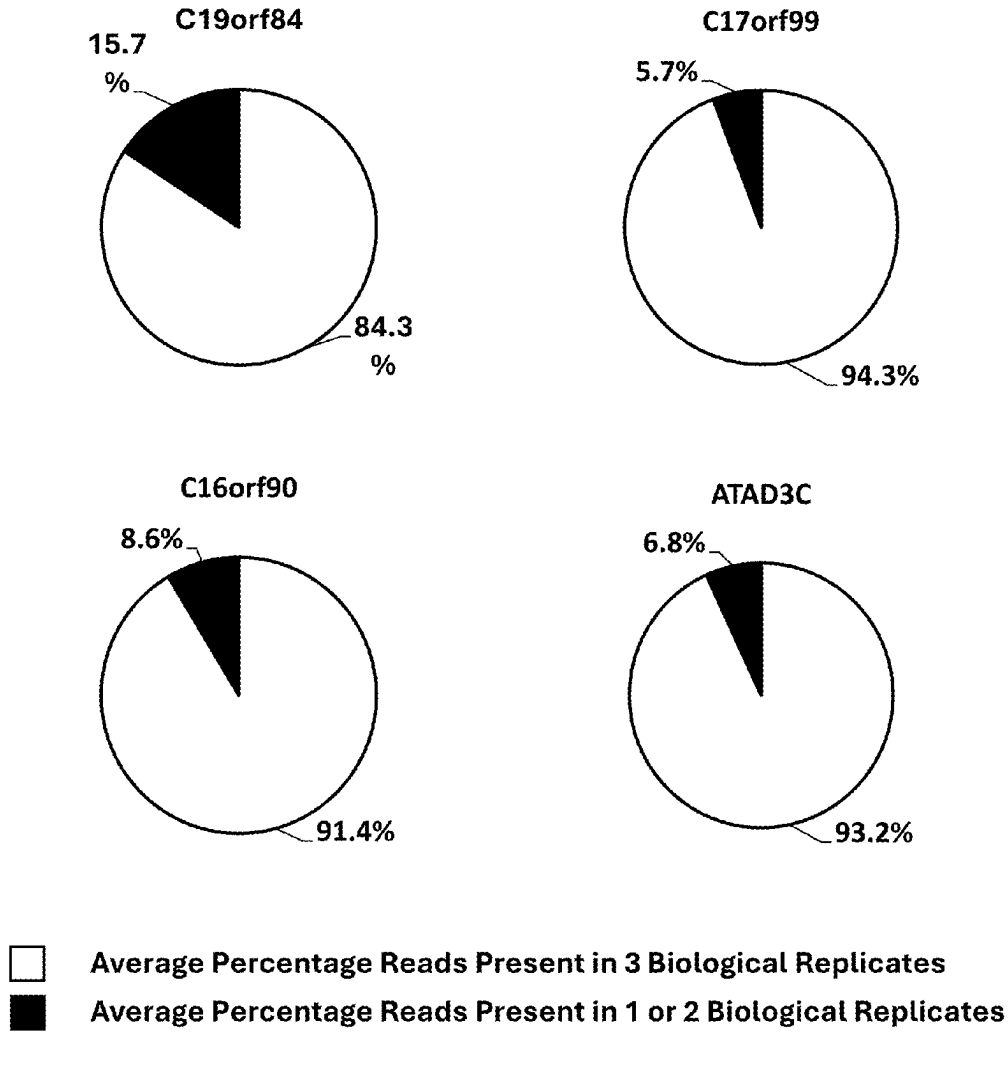


FIG. 1

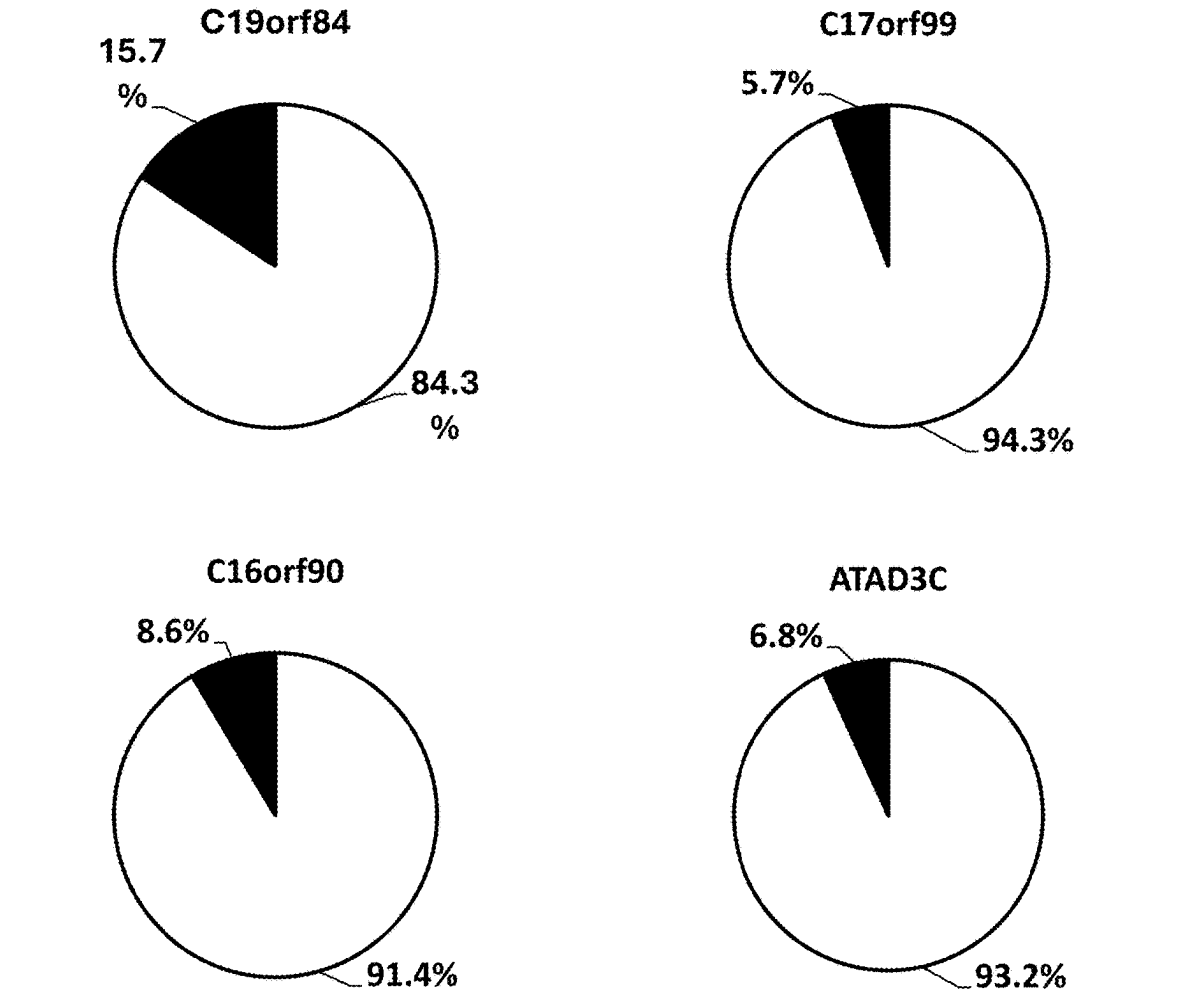

C19orf84
15.7 %
84.3 %
C17orf99
5.7%
94.3%
C16orf90
8.6%
91.4%
ATAD3C
6.8%
93.2%
Average Percentage Reads Present in 3 Biological Replicates
Average Percentage Reads Present in 1 or 2 Biological Replicates

FIG. 2

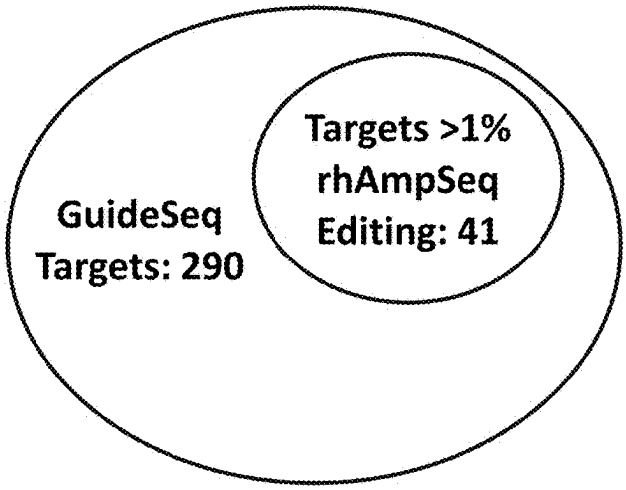

Targets >1%
rhAmpSeq
Editing: 41
GuideSeq
Targets: 290

FIG. 3

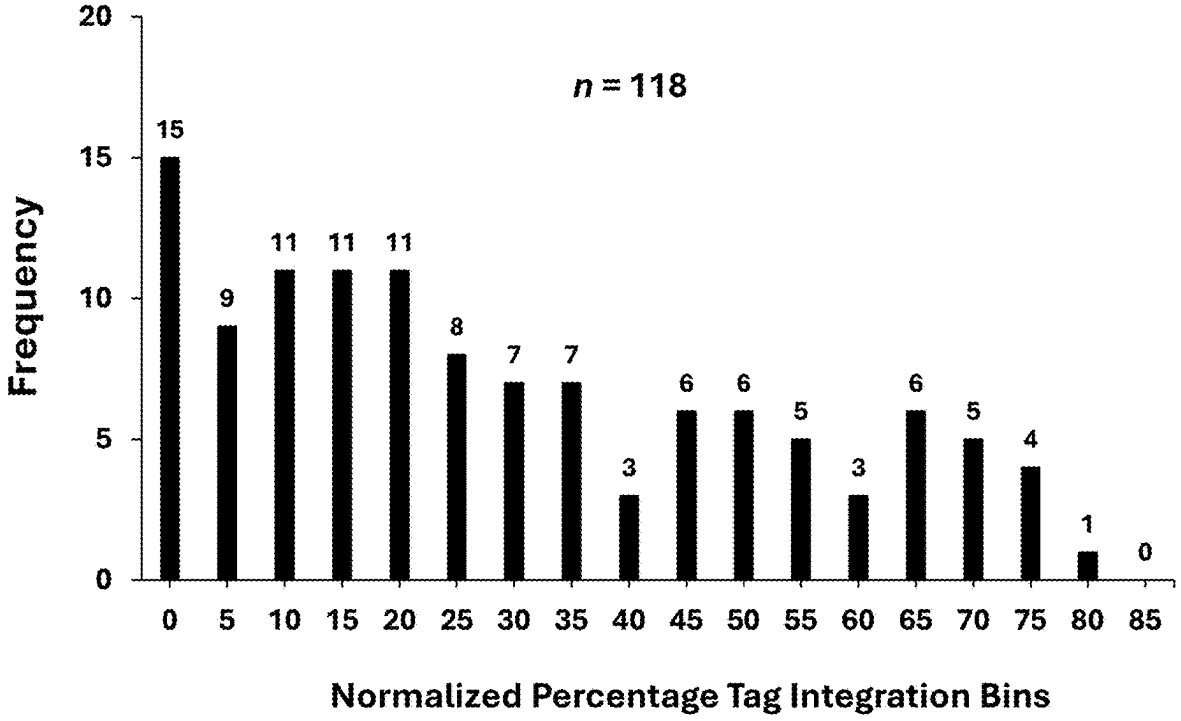

n = 118
Frequency
Normalized Percentage Tag Integration Bins
0
5
10
15
20
15
9
11
11
11
8
7
7
3
6
6
5
3
6
5
4
1
0
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85

FIG. 4

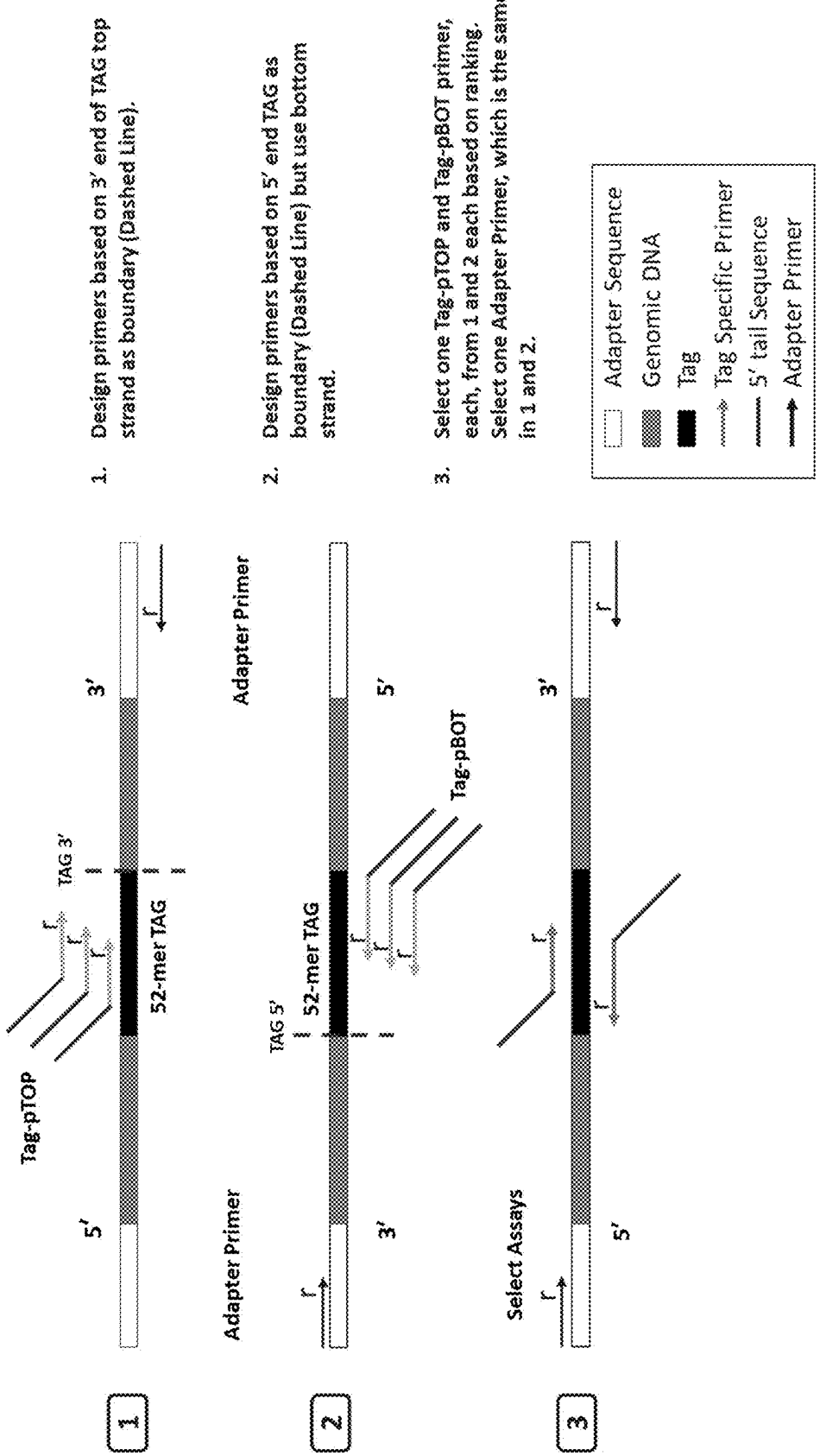

1. Design primers based on 3' end of TAG top strand as boundary (Dashed Line).
2. Design primers based on 5' end TAG as boundary (Dashed Line) but use bottom strand.
3. Select one Tag-pTOP and Tag-pBOT primer, each, from 1 and 2 each based on ranking. Select one Adapter Primer, which is the same in 1 and 2.
Adapter Sequence
Genomic DNA
Tag
Tag Specific Primer
5' tail Sequence
Adapter Primer
1
Tag-pTOP
TAG 3'
52-mer TAG
5'
3'
2
Adapter Primer
TAG 5'
52-mer TAG
Tag-pBOT
3'
5'
3
Select Assays
3'
5'

FIG. 5

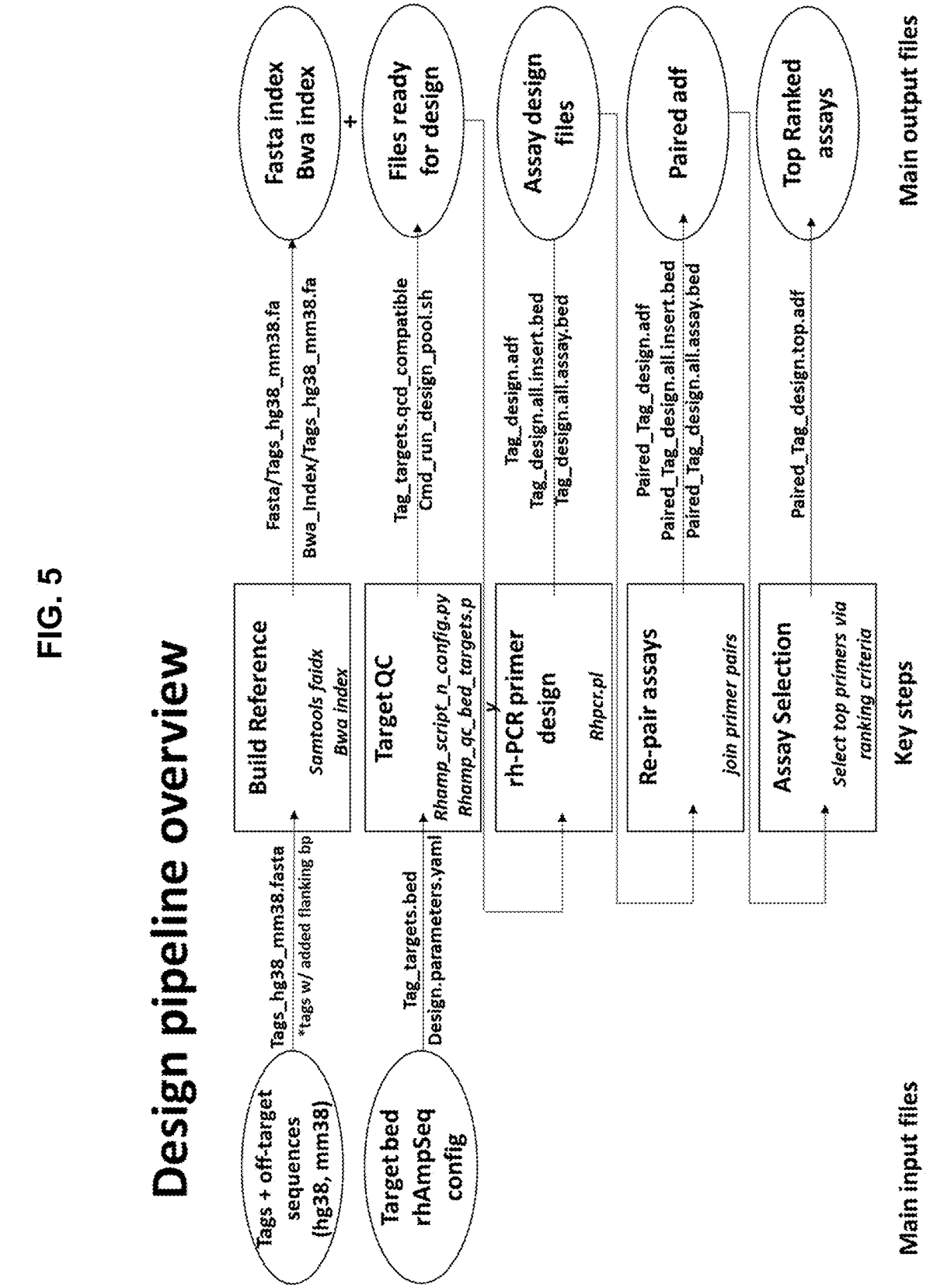

Design pipeline overview
Tags + off-target sequences (hg38, mm38)
Target bed rhAmpSeq config
Main input files
Tags_hg38_mm38.fasta
*tags w/ added flanking bp
Tag_targets.bed
Design.parameters.yaml
Build Reference
Samtools faidx
Bwa index
Target QC
Rhamp_script_n_config.py
Rhamp_qc_bed_targets.p
rh-PCR primer design
Rhpcr.pl
Re-pair assays
join primer pairs
Assay Selection
Select top primers via ranking criteria
Key steps
Fasta/Tags_hg38_mm38.fa
Bwa_Index/Tags_hg38_mm38.fa
Tag_targets.qcd_compatible
Cmd_run_design_pool.sh
Tag_design.adf
Tag_design.all.insert.bed
Tag_design.all.assay.bed
Paired_Tag_design.adf
Paired_Tag_design.all.insert.bed
Paired_Tag_design.all.assay.bed
Paired_Tag_design.top.adf
Fasta index
Bwa index
Files ready for design
Assay design files
Paired adf
Top Ranked assays
Main output files

FIG. 6

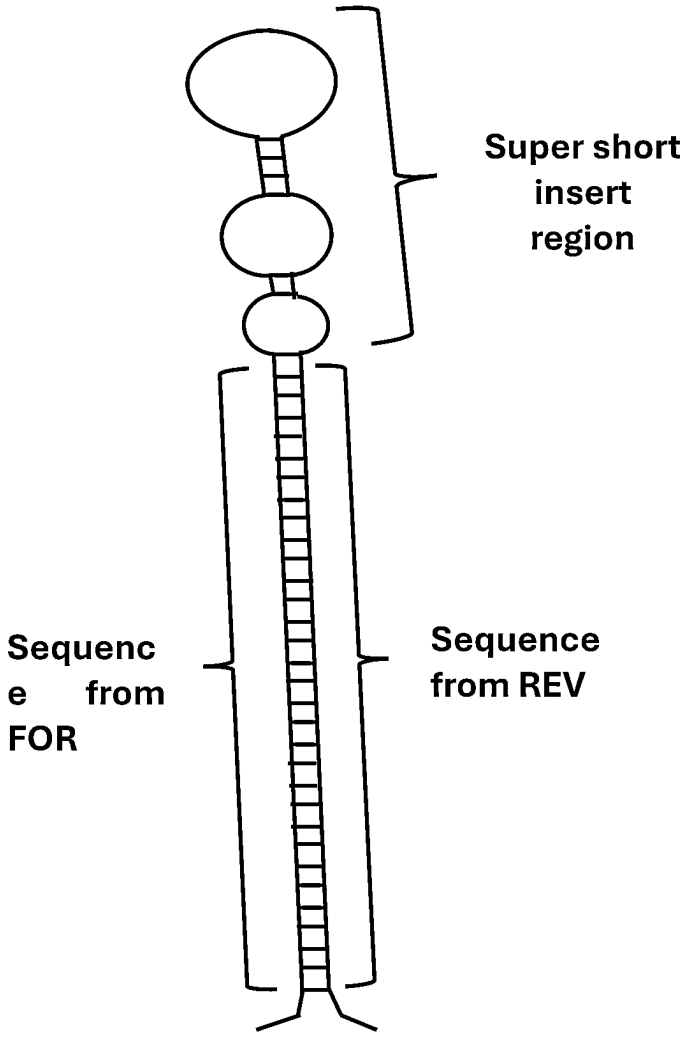

Super short insert region
Sequence from FOR
Sequence from REV

FIG. 7

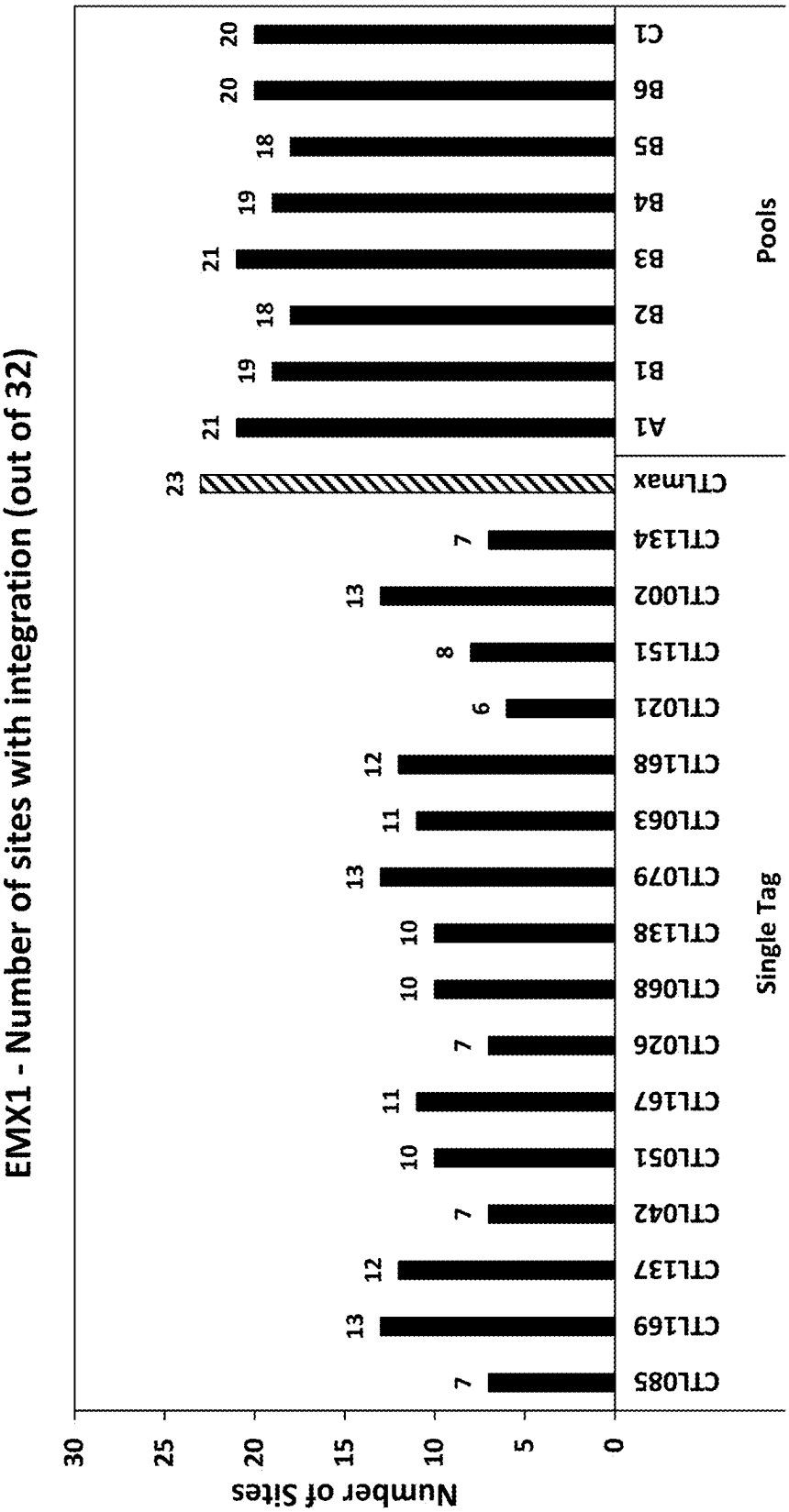
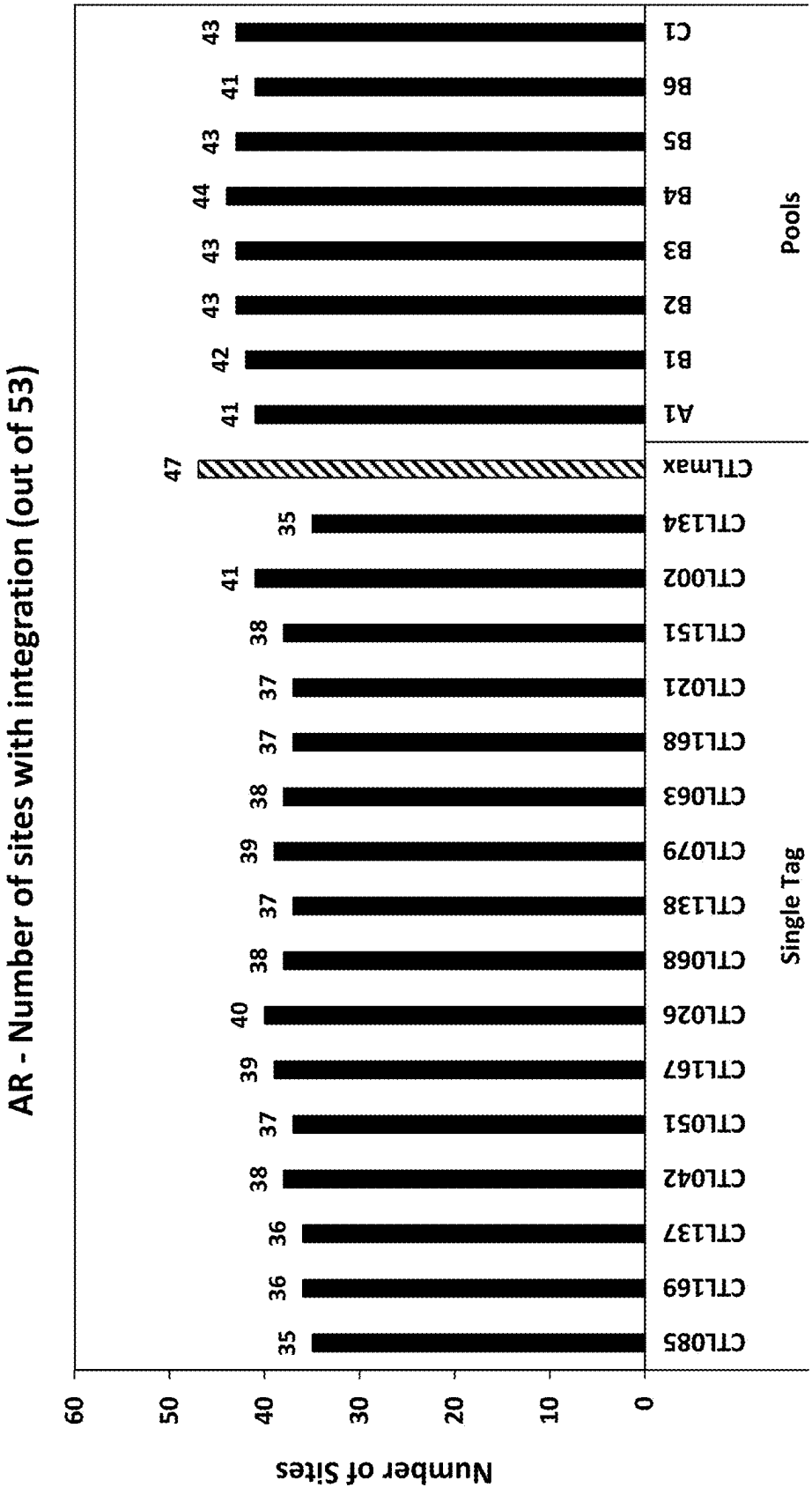

EMX1 - Number of sites with integration (out of 32)
Number of Sites
0
5
10
15
20
25
30
Single Tag
CTL085
CTL169
CTL137
CTL042
CTL051
CTL167
CTL026
CTL068
CTL138
CTL079
CTL063
CTL168
CTL021
CTL151
CTL002
CTL134
CTLmax
Pools
A1
B1
B2
B3
B4
B5
B6
C1
7
13
12
7
10
11
7
10
10
13
11
12
6
8
13
7
23
21
19
18
21
19
18
20
20

FIG. 8


AR - Number of sites with integration (out of 53)
Number of Sites
0
10
20
30
40
50
60
Single Tag
CTL085
CTL169
CTL137
CTL042
CTL051
CTL167
CTL026
CTL068
CTL138
CTL079
CTL063
CTL168
CTL021
CTL151
CTL002
CTL134
CTLmax
Pools
A1
B1
B2
B3
B4
B5
B6
C1
35
36
36
38
37
39
40
38
37
39
38
37
37
38
41
35
47
41
42
43
43
44
43
43
41
43

METHODS FOR NOMINATION OF NUCLEASE ON-/OFF-TARGET EDITING LOCATIONS, DESIGNATED "CTL-SEQ" (CRISPR TAG LINEAR-SEQ)

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application is a continuation of U.S. patent application Ser. No. 17/382,945, filed on Jul. 22, 2021, which claims priority to U.S. Provisional Patent Application No. 63/055,460, filed on Jul. 23, 2020, which is incorporated by reference herein in its entirety.

REFERENCE TO SEQUENCE LISTING

[0002] This application was filed with a Sequence Listing XML in ST.26 XML format in accordance with 37 C.F.R. § 1.831. The Sequence Listing XML file submitted in the USPTO Patent Center, "013670-9056-US03_sequence_listing_xml_1 May 2025.xml," was created on May 1, 2025, contains 273 sequences, has a file size of 248.0 kilobytes (253,952 bytes), and is incorporated by reference in its entirety into the specification.

TECHNICAL FIELD

[0003] Described herein are methods for identifying and nominating on- and off-target CRISPR editing sites with improved accuracy and sensitivity.

BACKGROUND

[0004] CRISPR (clustered regularly interspaced short palindromic repeats) has revolutionized genomics by permitting the simple introduction of changes to the genetic code. CRISPR systems, such as Cas9 and Cas12a proteins, are guided to their target by RNA oligonucleotide sequences bound by the Cas proteins (forming ribonucleoprotein protein; RNP), where the enzyme creates double stranded breaks (DSBs) in DNA sequences. Native cellular machinery repairs DSBs, generally using non-homologous end joining (NHEJ) or homology directed repair (HDR) molecular pathways. DNA repaired through NHEJ, which occurs at on- and off-target locations, often contains indels (insertions/deletions), which can lead to mutations and change the function of encoded genes. Thus, identifying these locations is critical to deconvoluting the impact of on- and off-target editing on biological phenotypes.

[0005] To date, no "gold standard" method exists to identify or nominate off-target editing locations for CRISPR or other nucleases. Many methods have been developed. These methods use a variety of strategies, including the detection of endogenous repair machinery assembled at DSBs (Discover-Seq [1]), the integration of a DNA tag sequence into the host cell genome (GUIDE-Seq; see U.S. Pat. No. 9,822,407), iGUIDE [2, 3]), or by cutting DNA in vitro (BLISS [4], CIRCLE-Seq [5], SiteSeq [6]).

[0006] Cellular or cell based (sometimes referred to as in vivo) and biochemical (sometimes referred to as in vitro) off-target assay nomination systems each have their advantages. Proteins bound to the DNA and epigenetic marks modify the function of nuclease activity, suggesting that cellular or cell based methods may better identify actual editing targets [7]. However, biochemical methods have nominated sites not identified through cellular or cell based methods, suggesting biochemical methods may be more comprehensive [5, 6]. Nevertheless, these current tools tend to have imperfect sensitivity [5, 6] (see FIG. **1**).

[0007] What is needed is a method for detecting and nominating on- and off-target CRISPR editing sites with improved accuracy and sensitivity.

SUMMARY

[0008] One embodiment described herein is a method for identifying and nominating on- and off-target CRISPR edited sites with improved accuracy and sensitivity, the process comprising the steps of: (a) co-delivering a guide sequence RNA (sgRNA) or a two-part CRISPR RNA: trans-activating crRNA (crRNA: tracrRNA) duplex, one or more tag sequences, and an RNA-guided endonuclease to cells; (b) incubating the cells for a period of time sufficient for double strand breaks to occur; (c) isolating genomic DNA from the cells, fragmenting the genomic DNA, and ligating the fragmented genomic DNA to a unique molecular index containing a universal adapter sequence; (d) amplifying the ligated DNA fragments using primers targeting the tag and universal adapter sequences to produce a first set of amplified sequences; (e) amplifying the first set of amplified sequences using universal sequencing primers targeting the tails of Tag-pTOP or Tag-pBOT primers to produce a second set of amplified sequences; (f) sequencing the pooled sequences and obtaining sequencing data; and (g) identifying on-/off-target CRISPR editing loci. In one aspect, the universal sequencing primers target SP1 or SP2 sequence (SEQ ID NO: 7, 8) tails on the Tag-pTOP or Tag-pBOT primers to produce a second set of amplified sequences. In another aspect, the universal sequencing primers target predesigned non-homologous sequence (SEQ ID NO: 269-273) tails on the Tag-pTOP or Tag-pBot primers to produce a second set of amplified sequences. In another aspect, the universal sequencing primers target predesigned 13-mer tails on the Tag-pTOP or Tag-pBot primers to produce a second set of amplified sequences. In another aspect, step (g) comprises executing on a processor: (i) aligning the sequence data to a reference genome; (ii) identifying on-/off-target CRISPR editing loci; and (iii) outputting the alignment, analysis, and results data as custom-formatted files, tables or graphics. In another aspect, the method further comprises a step following step (e) comprising: (e1) normalizing the second set of amplified sequences to produce concentration normalized libraries, pooling the normalized libraries with other samples to produce pooled libraries; and continuing with steps (f)-(i). In another aspect, step (d) uses a supression PCR method. In another aspect, the RNA-guided endonuclease comprises an endogenously-expressed Cas enzyme, a Cas expression vector, a Cas protein, or a Cas RNP complex. In another aspect, the RNA-guided endonuclease comprises an endogenously-expressed Cas9 enzyme, a Cas9 expression vector, a Cas9 protein, or a Cas9 RNP complex. In another aspect, the cells comprise human or mouse cells. In another aspect, the period of time is about 24 hours to about 96 hours. In another aspect, multiple tag sequences are co-delivered. In another aspect, the tag sequences comprise double-stranded deoxyribooligonucleotides (dsDNA) comprising 52-base pairs. In another aspect, the tag sequences comprise a 5'-terminal phosphate, and phosphorothioate linkages between the 1^{st} and 2^{nd}, 2^{nd} and 3^{rd}, 50^{th} and 51^{st}, and 51^{st} and 52^{nd} nucleotides. In another aspect, the tag sequences comprise a

double stranded DNA comprising the complementary top and bottom strand pairs of SEQ ID NO: 1-2 or 7-268.

[0009] Other embodiments described herein are on- and off-target CRISPR editing sites identified or nominated using the methods described herein.

[0010] Another embodiment described herein is a method for designing 52-base pair tag sequences, the method comprising, executing on a processor: (a) randomly generating 13-nucleotide sequences with 40-90% GC content, max homopolymer length A: 2, C: 3, G: 2, T: 2, weighted homopolymer rate <20, self-folding Tm<50° C., and self-dimer Tm<50° C.; (b) removing sequences that perfectly align to a particular genome or that are homopolymers or GG or CC dinucleotide motifs and obtaining a set of 13-mers; (c) selecting a subset of the 13-mer sequences that contain one or less CC or GG dinucleotide motifs; (d) concatenating four of the of 13-mer subset sequences to form random 52-mer sequences; (e) aligning the random 52-mer sequences to a genome; (f) removing the random 52-mer sequences that have similarity to the genome to produce a subset of 52-mer sequences; and (h) outputting the subset of 52-mer sequences and generating the complementary strands to produce double stranded 52-base pair tag sequences. In one aspect, the genome is human or mouse. In another aspect, the 52-base pair tag sequences are-non complementary to the genome. In another aspect, the method further comprises designing primers for the 52-base pair tag sequences. In another aspect, the 52-base pair tag sequences comprise a 5'-terminal phosphate, and phosphorothioate linkages between the 1^{st} and 2^{nd}, 2^{nd} and 3^{rd}, 50^{th} and 51^{st}, and 51^{st} and 52^{nd} nucleotides of the 52-base pair tag sequences. In another aspect, the method further comprises synthesizing oligonucleotides comprising the 52-base pair tag sequences, the complement of the 52-base pair tag sequences, or primers for the 52-base pair tag sequences.

[0011] Other embodiments described herein are one or more 52-base pair tag sequences designed using the methods described herein. In one aspect, the 52-base pair tag sequence comprises a double stranded DNA comprising the top and bottom strand pairs of SEQ ID NO: 1-2 or 7-268.

[0012] Another embodiment described herein is a method for designing primers partially complementary to the 52-base pair tag sequences of claim **23** and an adapter primer, the method comprising, executing on a processor: (a) designing tag primers that are partially complementary to the top and bottom strands of tag sequences; and (b) designing an adapter primer that is partially complementary to the top strand of the adapter sequence; wherein: the tag primers comprise a 5'-universal tail sequence; and the adapter primer comprises a sequence complementary to the tails of Tag-pTOP or Tag-pBOT primers. In one aspect, the 5'-universal tail sequence is complementary to an SP1 or SP2 sequence (SEQ ID NO: 7, 8), a locus specific segment, a ribonucleotide (rN) 6-nucleotides from the 3'-end, a 3'-end mismatch, a 3'-end block (3'-C_3 spacer), a predesigned non-homologous sequence (SEQ ID NO: 269-273), or a predesigned 13-mer sequence. In another aspect, the primers partially complementary to top and bottom strands of the tag sequences comprise a tail sequence complementary to the SP1 sequence (SEQ ID NO: 7) and the adapter primer comprises a sequence complementary to the SP2 sequence (SEQ ID NO: 8) tail on the Tag-pTOP or Tag-pBOT primers; or the primers partially complementary to top and bottom strands of the tag sequences comprise a tail sequence complementary to the SP2 sequence (SEQ ID NO: 8) and the adapter primer comprises a sequence complementary to the SP1 sequence (SEQ ID NO: 7) tail on the Tag-pTOP or Tag-pBOT primers. In another aspect, the amplification of a nucleic acid molecule with the primers that are complementary to the top and bottom strands of tag sequences and primers that are complementary to the top strand of the adapter sequence produces a PCR product that comprises a portion of the tag sequence, a sgDNA sequence, and the adapter sequence. In another aspect, the method further comprises synthesizing oligonucleotides comprising the sequences of the forward and reverse tag primers and the adapter primer. In another aspect, the 52-base pair tag sequences and primers partially complementary to the 52-base pair tag sequences are designed and selected using an algorithm predicting whether the primers are likely to be partially complementary and have a propensity to form primer-dimers.

[0013] Other embodiments described herein are one or more primers partially complementary to the 52-base pair tag sequences and one or more adapter primers designed using the methods described herein. In one aspect, the primers comprise the sequences of SEQ ID NO: 3, 4; and the adapter primer, wherein the adapter primer comprises the sequence of SEQ ID NO: 5.

[0014] Another embodiment described herein is the use of one or more double-stranded 52-base pair tag sequences for identifying on- and off-target CRISPR editing sites.

DESCRIPTION OF THE DRAWINGS

[0015] FIG. **1** shows fraction of reads shared by three biological replicates are shown in white sectors; whereas reads shared by two replicates, or present in a single replicate, are shown in black sectors. Table 1 shows GUIDE-seq [3] based nomination for 4 different gRNAs in triplicate in a 96-well format. gRNA complexes were generated by mixing equimolar amounts of Alt-R crRNA-XT and Alt-R tracrRNA. HEK293 cells stably expressing Cas9 were transfected with 10 UM gRNA and 0.5 UM dsODN GUIDE-seq tag using the Nucleofector™ system (Lonza). After 72 hrs, genomic DNA (gDNA) was isolated. Genomic DNA was fragmented, and adapters were ligated using the Lotus DNA library preparation kit (IDT). Libraries were generated by amplification from the inserted tag to the ligated adapters [3]. Libraries were then sequenced in paired-end fashion on an Illumina® platform.

[0016] FIG. **2** shows that GUIDE-Seq finds more off-target locations than can be validated through rhAmpSeq targeted amplification. Presented results are an aggregate of 331 GUIDE-Seq nominated sites when delivering gRNA sequences (internally named: AR, CTNNB1, EMX1, GRHPR, HPRT38087, HPRT38285, VEGFA) into HEK293 cells stably expressing WT Cas9. GUIDE-seq nominated off-targets assigned >0.1% of the total reference genome aligned reads for each guide were designed and targeted by one rhAmpSeq panel all reference genome aligned. In subsequent experiments, gRNAs were again delivered to the same cells, and editing was assayed with rhAmpSeq. Targets were called "edited" if the treated condition had observed indels ≥ the untreated control sample at ≥1%.

[0017] FIG. **3** illustrates that GUIDE-Seq tag integration rate varies. The graph shows the percentage of Tag integration (normalized to % Editing) for 118 unique Cas9 on/off-target sites that had InDel editing in rhAmpSeq panels

targeting GUIDE-Seq nominated on/off-target loci for guide sequences targeting the RAG1, RAG2, and EMX1 genes. Each guide was co-delivered with the 34-base pair GUIDE-Seq, dsODN tag into HEK293 cells stably expressing Cas9 by nucleofection. DNA was extracted 72 hrs later, amplified by rhAmpSeq multiplex PCR, sequenced on an Illumina® MiSeq, and analyzed through a custom pipeline. The normalized tag integration rate is calculated as the percentage of sequenced reads at each target containing the tag sequence divided by the total reads containing an allele divergent from the reference genome (indicating Cas9 editing).

[0018] FIG. **4** shows the design of rhAmpSeq primers against alien sequence tags. A cartoon diagram shows the steps of the design process using the rhAmpSeq design pipeline including design of forward primers against the top (1) and bottom (2) strands, discarding unneeded primers, and selecting tag-targeting primers that have 5'-overlapping, but not 3'-overlapping sequences, so that the top/bottom strand primer dimers would hairpin (3).

[0019] FIG. **5** shows an overview of the rhAmpSeq design pipeline used to construct the overlapping primer designs. In the pipeline, a known sequence is appended onto the 5'-end and 3'-end of each tag sequence, the inputs are quality-controlled and assays (shown in FIG. **4**A) are designed against the top and bottom strand of each tag. Primers targeting each tag strand are paired such that at least 4-nucelotides 3' of the RNA nucleotide do not overlap between primers targeting the same tag, and primer pairs are ranked and selected. Hg38 and mm38 acronyms represent versions of the human and mouse genomes, respectively.

[0020] FIG. **6** illustrates hairpin formation if overlapping primers generate PCR amplicons. The diagram shows a representative target sequence and hairpin PCR product of undesired short amplicons from overlapping primer regions with complementary 5' primer tail ends at the 3'- and 5'-end of the PCR product.

[0021] FIG. **7** shows the number of target sites (black bars) with integration of the specified single tag (SEQ ID NO: 9-40) or pools of tags described in Table 5 (SEQ ID NO: 9-40, 45-268). The striped bar (CTLmax) shows the maximum number of target sites that theoretically can be found if a combination of the single tags (SEQ ID NO: 9-40) is used (23 sites out of a maximum of 32 sites). Pool A1 contains all the single tags (SEQ ID NO: 9-40). Pools B1-6 contain 16 different tags each (SEQ ID NO: 45-268). Pool C1 contains all tags tested (SEQ ID NO: 9-40, 45-268). Integration events were determined using an in-house data analysis tool.

[0022] FIG. **8** shows the number of target sites (black bars) with integration of the specified single tag (SEQ ID NO: 9-40) or pools of tags described in Table 5 (SEQ ID NO: 9-40, 45-268). The striped bar (CTLmax) shows the maximum number of target sites that theoretically can be found if a combination of the single tags (SEQ ID NO: 9-40) is used (47 sites out of a maximum of 53 sites). Pool A1 contains all the single tags (SEQ ID NO: 9-40). Pools B1-6 contain 16 different tags each (SEQ ID NO: 45-268). Pool C1 contains all tags tested (SEQ ID NO: 9-40, 45-268). Integration events were determined using an in-house data analysis tool.

DETAILED DESCRIPTION

[0023] Described herein are methods for detecting and nominating on- and off-target CRISPR editing sites with improved accuracy and sensitivity. The intracellular context information is maintained by building upon prior in vivo nomination methods. The sensitivity is expanded by co-delivering a set of unique, predefined sequence tags. In one aspect, the co-delivered set of predefined unique tags may range from 13-80 base pairs. In another aspect, the co-delivered set of predefined tags may be comprised of 13 base pair tag sequence tags, 26 base pair tag sequence tags, 39 base pair tag sequence tags, 52 base pair tag sequence tags, 65 base pair tag sequence tags, or 78 base pair tag sequence tags. In another aspect, the unique predefined tags are a set of 52-base pair tag sequence tags (the increased length of the sequence tags improves the ability to find good primer landing sites for rhPrimers). This limitation is believed to be mitigated by using a diversity of tag sequences that are distinct from human and mouse genomes. The specificity is improved by building upon Integrated DNA Technologies (IDT)'s rhAmp technology that uses RNAaseH2 (*Pyrococcus abyssi*) to unblock primers that have correctly annealed to their target; this yields lower rates of false priming. Specificity can be further enhanced by only nominating targets using reads that contain an expected tag sequence at the 5'-end. The incorporation of suppression PCR into this method permits ease of use. The prior in vivo methods (e.g., GUIDE-seq and iGUIDE) require parallel PCR reactions (2 pool amplification) to amplify by annealing to and extending from the top and bottom strand of the tags. Here, suppression PCR is used to allow both pools to be amplified simultaneously without causing problematic dimer sequences.

[0024] A GUIDE-Seq dsDNA tag was co-delivered with one guide RNA to HEK293 cells constitutively expressing Cas9 using nucleofection. See U.S. Pat. No. 9,822,407, which is incorporated by reference herein for such teachings. A total of four different guide RNAs were tested in this fashion. Ribonucleoprotein complexes (RNPs) between the expressed Cas9 and guide RNA form within the cells, introducing double stranded breaks. Repaired breaks can contain the co-delivered tags. After delivery, cells were incubated, and the resulting DNA was extracted. Target amplification was performed according to the GUIDE-Seq protocol and assayed with a modified version of the GUIDE-Seq analytical pipeline (github.com/aryeelab/guideseq). Nominated targets were compared between three biological replicates (unique guideRNA+Tag co-deliveries). Not all nominated targets were common to all biological replicates (commonly/total nominated targets: 7/31, 6/19, 2/4, 3/5 respectively; see Table 1). However, >90% of the total reads, attributed to any target, were attributed to common targets (on average; see FIG. **1**).

TABLE 1

Identified off-target sites for four different gRNAs and relative level of editing at off-target sites compared to the on-target site

Location	C19orf84_BR1	C19orf84_BR2	C19orf84_BR3
chr19_51389306	100.00%	100.00%	100.00%
chr9_20224748	38.55%	16.43%	29.00%
chr4_28036434	16.33%	13.05%	14.36%
chr15_74256506	14.30%	18.18%	25.17%
chr2_171312919	11.40%	8.51%	7.93%
chr8_65742269	10.82%	1.17%	10.40%
chr13_96554656	8.70%	0.00%	0.00%
chr4_86807920	8.50%	9.21%	1.92%
chr3_124485356	6.57%	0.00%	0.00%

TABLE 1-continued

Identified off-target sites for four different gRNAs and relative level of editing at off-target sites compared to the on-target site

chr9_20330398	5.60%	0.00%	0.00%
chr11_71298123	5.12%	0.00%	0.00%
chr7_101729696	4.83%	0.00%	9.58%
chr19_10923882	3.67%	3.03%	0.00%
chr10_15548456	3.57%	15.38%	0.00%
chr12_117097457	2.80%	0.00%	2.60%
chr22_33493900	2.13%	0.00%	4.79%
chrX_149763439	2.13%	0.00%	3.83%
chr17_7435217	1.93%	0.00%	0.55%
chr12_26286721	1.74%	0.00%	5.06%
chr16_49704848	1.26%	5.01%	7.11%
chr12_51288216	1.06%	0.00%	0.00%
chr12_56010621	0.87%	0.00%	0.00%
chr13_29717148	0.48%	0.00%	0.00%
chr1_3088065	0.29%	0.00%	0.00%
chr15_73442915	0.19%	0.00%	0.55%
chr10_118045968	0.19%	0.00%	0.00%
chr14_102199972	0.00%	0.00%	0.68%
chr18_56334679	0.00%	0.00%	2.33%
chr21_36426137	0.00%	0.00%	2.19%
chr5_139002763	0.00%	0.00%	3.83%
chrX_58291642	0.00%	0.00%	3.83%
Location	C17orf99_BR1	C17orf99_BR2	C17orf99_BR3
chr17_78164110	100.00%	100.00%	100.00%
chr22_24471716	15.00%	13.24%	10.86%
chr10_101156881	6.22%	11.07%	9.79%
chr3_170476431	5.86%	3.97%	4.57%
chr17_17692965	4.94%	0.66%	8.62%
chr15_73400031	3.93%	4.63%	5.73%
chr19_15238775	0.00%	0.00%	2.56%
chr2_18362316	0.00%	0.00%	1.59%
chr2_171087784	0.00%	0.54%	0.84%
chr22_19959968	0.00%	1.26%	0.19%
chr22_32114104	0.00%	0.00%	4.06%
chr4_129034015	0.00%	0.00%	0.33%
chr5_61219030	0.00%	0.00%	0.33%
chr5_66209615	0.00%	0.00%	1.86%
chr7_69709389	0.00%	0.12%	2.75%
chr7_158662844	0.00%	1.44%	5.27%
chrX_9567397	0.00%	0.00%	0.23%
chr19_55657073	0.00%	0.66%	0.00%
chr22_43788032	0.00%	2.47%	0.00%
Location	C16orf90_BR1	C16orf90_BR2	C16orf90_BR3
chr16_3494817	100.00%	100.00%	100.00%
chr2_109189307	75.32%	4.27%	52.05%
chr22_24586001	45.45%	0.00%	0.00%
chr10_104736568	0.00%	0.00%	8.22%
Location	ATAD3C_BR1	ATAD3C_BR2	ATAD3C_BR3
chr1_1450685	100.00%	100.00%	100.00%
chr1_1503588	11.73%	10.07%	9.27%
chr1_1516015	2.47%	1.86%	5.14%
chr19_32167960	26.34%	0.93%	0.00%
chr2_111077960	0.00%	1.12%	0.00%

[0025] Additionally, nominated targets may not be replicable or detectable using orthogonal methods. Using the GUIDE-Seq method, the GUIDE-Seq DNA tag was co-delivered with each of 6 guides (each tag is delivered with one guide RNA) to HEK293 cells constitutively expressing Cas9 using nucleofection. rhAmpSeq multiplex amplicon panels were designed to amplify the nominated targets, and we quantified editing in biological replicates. Of the 331 targets nominated by GUIDE-Seq, only 41 (12%) could be verified with rhAmpSeq (see FIG. **2**).

[0026] dsDNA tag sequences co-delivered with the guide RNAs into a stably expressing CRISPR cell line, which are used in the NHEJ repair, are incorporated at varying rates. Here, the GUIDE-Seq dsDNA tag was co-delivered with each of 6 guides into HEK293 cells constitutively expressing Cas9. In another aspect, the dsDNA tag sequences co-delivered with CRISPR RNP, which are used in the NHEJ repair, are incorporated at varying rates. Here, the GUIDE-Seq dsDNA tag was co-delivered with each of 6 guides into HEK293 cells constitutively expressing Cas9. rhAmpSeq panels were developed to amplify nominated targets, and in biological replicates, the rates of tag integration were analyzed using a custom analytical pipeline. These results demonstrate that tags are incorporated at 0-85% of edited genomic copies, varying by target (see FIG. **3**). Without being bound by any theory, it is hypothesized that the rate varies by sequence context.

[0027] Described herein are methods to improve the signal to noise ratio by combining Integrated DNA Technology's rhAmpSeq™ technology, suppression PCR, and novel alien DNA sequence designs to nominate nuclease off-target editing locations within a host genome.

[0028] In this method, Cas9, a sgRNA or a two-part CRISPR RNA: trans-activating crRNA (crRNA: tracrRNA) duplex, and one or more double stranded DNA (dsDNA) tag sequences are delivered to cells. Co-delivering multiple tags permits improved tag integration at off-target sites (see below). The tag sequences have sequence content significantly different (i.e., alien) to the host genome. After nuclease introduced DSBs, NHEJ repair will insert the tag sequence(s) into the target site, forming known primer landing sites. After cells have time to repair the DSBs and possibly further divide (such as after 72 hr), genomic DNA is isolated, fragmented (e.g., Covaris® shearing, enzyme-based shearing, Tn5, etc.), ligated a unique molecular index (UMI)-containing universal adapter sequence to the fragmented DNA, and the un-ligated material is removed. Next, the DNA fragments are amplified by targeting primers to the tag and universal adapter sequences (Round 1 PCR). Using universal primers, a sample index (PCR2) is added, the amplified material is concentration normalized, pooled with other samples, and the pooled material is sequenced on an Illumina® (or similar) machine. The sequenced reads are aligned to a reference genome, and loci where large numbers of reads map may nominate on/off-target locations.

[0029] Alien sequences were designed by generating >1 M random 13-mer sequences with 40-90% GC content, max homopolymer length A: 2, C: 3, G: 2, T: 2, weighted homopolymer rate <20, self-folding Tm<50° C., and self-dimer Tm<50° C. From the list of sequences, sequences that aligned perfectly against human (GRCh38.p2; hg38) or mouse (GRCh38.p4; mm38) reference genomes or had troubling motif sequences (homopolymers, most G-G or C-C dinucleotide motifs) were removed, resulting in 479 sequences.

[0030] To design the 52-base pair tag sequences described herein, 49 13-mer oligo sequences were selected that contain≤1 C or G dinucleotide, and 10,000 unique combinations of four 13-mer sequences were generated. The length of each concatenated sequence (e.g., pasting four 13-mer sequences in a row using software) is 52-nucleotides. Next, each 52-nucleotide tag sequence was aligned against the human (GRCh38.p2) and mouse (GRChm38.p4) genomes using an internally modified version of bwa, called bwa-psm. Implementation of bwa-psm returns all possible secondary matches up to a defined threshold. A set of tag

sequences (SEQ ID NO:1-2) were designed that were intended to work as a group, that had no similarity to the human or mouse genomes (max seed size: 7, seed edit distance: 2, max edit distance: 21, max gap open: 2, max gap extension: 3, mismatch penalty: 1, gap open penalty: 1, gap extension penalty: 1).

[0031] Overlapping rhAmpSeq V1 primers (SEQ ID NO: 3-4) were designed complementary to the top and bottom strands of the tag and 5'-end of the adapter sequence (SEQ ID NO: 6) (FIG. **4**). The tag-specific primers (SEQ ID NO: 3-4) contain a 5'-universal tail sequence matching the SP1 and SP2 primer sequences (SEQ ID NO: 7-8), a locus specific segment, a ribonucleotide (rN) 6-nucleotides from the 3'-end, a 3'-end mismatch, and a 3'-end block (3'-C3 spacer). The adapter-specific primer (SEQ ID NO: 5) targets the 5'-end of the 5'-P5 adapter sequence (SEQ ID NO: 6), and the adapter sequence contains unique molecular index (UMI) sequence (Table 2). The primers were designed to target the plus and minus strands of the annealed tag such that, if these primers unexpectedly form a dimer, the formed product will hairpin, removing the oligo from the available reaction templates (e.g., supression PCR). (FIG. **6**A-B). Primer sequences targeting the tags were chosen based on a proprietary design algorithm designed and implemented by IDT (internal copy of the algorithm with a public-facing UI: www.idtdna.com/site/account?ReturnURL=/site/order/designtool/index/RHAMPSEQ), which selects the most optimally performing primer pairs to amplify the intended template sequence. (FIG. **5**). Primer sequences were assessed for non-specific binding to all other tag sequences and both human and mouse primary genome assemblies to verify they were unlikely to form off-target amplicons when combined with a universal adapter sequence and the presence of human or mouse genomic DNA.

[0032] The primers were desired to work in pairs where one tag-specific primer (top or bottom strand) pairs with the adapter-specific primer (SEQ ID NO:5). This results in the amplification of a molecule that contains a portion of the tag, gDNA, and the adapter sequence when amplified using supression PCR methods (FIG. **4**).

TABLE 2

Sequences Used for First Proof of Concept

Type	Name	Sequence (5'→3')	SEQ ID NO
Tag	9022179029169042579 04625907201907281	T*C*GTTCGTTC CGCTCTAACCGG CGAATCTACCGC GCATATCTACGC CGCA*A*T	SEQ ID NO: 1
Tag	9022179029169042579 04625907201907281_rev	A*T*TGCGGCGT AGATATGCGCGG TAGATTCGCCGG TTAGAGCGGAAC GAAC*G*A	SEQ ID NO: 2
Tag Primers	pFWD.ID_Target1: 9022179029169042579 04625907201907281.12 7.150.1.SP1	acactctttccc tacacgacgctc ttccgatctTCT ACCGCGCATATC TACrGCCGCT/ 3SpC3/	SEQ ID NO: 3

TABLE 2-continued

Sequences Used for First Proof of Concept

Type	Name	Sequence (5'→3')	SEQ ID NO
Tag Primers	pFWD.ID_Target2: 9022179029169042579 04625907201907281.11 6.140.-1.SP1	acactctttccc tacacgacgctc ttccgatctATA TGCGCGGTAGAT TCGCrCGGTTT/ 3SpC3/	SEQ ID NO: 4
Adapter Primer	Adapter Primer	gtgactggagtt cagacgtgtgct cttccgatctAA TGATACGGCGAC CACCGAGATCTA CArCAAGGC/ 3SpC3/	SEQ ID NO: 5
P5 Adapter	Example Sequence	AATGATACGGCG ACCACCGAGATC TACACTAGATCG CNNWNNWNNACA CTCTTTCCCTAC ACGACGCTCTTC CGATC*T	SEQ ID NO: 6
SP1	Sequencing Primer 1	acactctttccc tacacgacgctc ttccgatct	SEQ ID NO: 7
SP2	Sequencing Primer 2	gtgactggagtt cagacgtgtgct cttccgatct	SEQ ID NO: 8

"*" indicates a phosphorothioate linkage; "rN" indicates a ribonucleotide, where N is the nucleotide preceeded by the "r"; "/3SpC3/" indicates a 3'-C_3 spacer.

[0033] One embodiment described herein is a method for identifying and identifying and nominating on- and off-target CRISPR editing sites with improved accuracy and sensitivity, the process comprising the steps of: (a) co-delivering a guide sequence RNA (sgRNA) or a two-part CRISPR RNA: trans-activating crRNA (crRNA: tracrRNA) duplex and one or more tag sequences to cells; (b) incubating the cells for a period of time; (c) isolating genomic DNA from the cells, fragmenting the genomic DNA, and ligating the fragmented genomic DNA to a unique molecular index containing a universal adapter sequence; (d) amplifying the ligated DNA fragments using primers targeting the tag and universal adapter sequences to produce a first set of amplified sequences; (e) amplifying the first set of amplified sequences using universal sequencing primers targeting the tails of Tag-pTOP or Tag-pBOT primers to produce a second set of amplified sequences; (f) sequencing the pooled sequences and obtaining sequencing data; and (g) identifying on-/off-target CRISPR editing loci. In one embodiment, the universal sequencing primers target SP1 or SP2 sequence (SEQ ID NO: 7, 8) tails on the Tag-pTOP or Tag-pBOT primers to produce a second set of amplified sequences. In another embodiment, the universal sequencing primers target predesigned non-homologous sequence (Table 6; SEQ ID NO: 269-273) tails on the Tag-pTOP or Tag-pBot to produce a second set of amplified sequences. In yet another embodiment, the universal primers target predesigned 13-mer tails on the Tag-pTOP or Tag-pBOT primers to produce a second set of amplified sequences. In one

embodiment, step (g) comprises executing on a processor: (i) aligning the sequence data to a reference genome; (ii) identifying on-/off-target CRISPR editing loci; and (iii) outputting the alignment, analysis, and results data as tables or graphics. In another embodiment, the method further comprises a step following step (e) comprising: (e1) normalizing the second set of amplified sequences to produce concentration normalized libraries, pooling the normalized libraries with other samples to produce pooled libraries; and continuing with steps (f)-(i). In one aspect, step (d) uses a supression PCR method. In another aspect, the cells constitutively express a Cas enzyme, are co-delivered with a Cas expression vector, are co-delivered with a Cas protein, or are co-delivered with a Cas RNP complex. In another aspect, the cells constitutively express a Cas9 enzyme, are co-delivered with a Cas9 expression vector, are co-delivered with a Cas9 protein, or are co-delivered with a Cas9 RNP complex. In another aspect, the cells comprise human or mouse cells. In another aspect, the period of time is about 24 hours to about 96 hours. In another aspect, multiple tag sequences are co-delivered. In another aspect, the tag sequences comprise double-stranded deoxyribooligonucleotides (dsDNA) comprising 52-base pairs. In another aspect, the tag sequences comprise a 5'-terminal phosphate, and phosphorothioate linkages between the 1^{st} and 2^{nd}, 2^{nd} and 3^{rd}, 50^{th} and 51^{st}, and 51^{st} and 52^{nd} nucleotides. In another aspect, the tag sequences comprise a double stranded DNA comprising the top and bottom strand pairs of SEQ ID NO: 9-40 or 45-268.

[0034] Another embodiment described herein is on- and off-target CRISPR editing sites identified or nominated using the methods described herein.

[0035] Another embodiment described herein is a method for designing 52-base pair tag sequences, the method comprising, executing on a processor: (a) randomly generating 13-nucleotide sequences with 40-90% GC content, max homopolymer length A: 2, C: 3, G: 2, T: 2, weighted homopolymer rate <20, self-folding Tm<50° C., and self-dimer Tm<50° C.; (b) removing sequences that perfectly align to a particular genome or that are homopolymers or GG or CC dinucleotide motifs and obtaining a set of 13-mers; (c) selecting a subset of the 13-mer sequences that contain one or less CC or GG dinucleotide motifs; (d) concatenating four of the of 13-mer subset sequences to form random 52-mer sequences; (e) aligning the random 52-mer sequences to a genome; (f) removing the random 52-mer sequences that have similarity to the genome to produce a subset of 52-mer sequences; and (h) outputting the subset of 52-mer sequences and generating the complementary strands to produce double stranded 52-base pair tag sequences. In one aspect, the genome is human or mouse. In one aspect, the 52-base pair tag sequences are not complementary to the genome. In another aspect, the method further comprises designing primers for the 52-base pair tag sequences. In another aspect, the 52-base pair tag sequences comprise a 5'-terminal phosphate, and phosphorothioate linkages between the 1^{st} and 2^{nd}, 2^{nd} and 3^{rd}, 50^{th} and 51^{st}, and 51^{st} and 52^{nd} nucleotides of the 52-base pair tag sequences. In another aspect, the method further comprises synthesising oligonucleotides comprising the 52-base pair tag sequences, the complement of the 52-base pair tag sequences, or primers for the 52-base pair tag sequences.

[0036] Another embodiment described herein is one or more 52-base pair tag sequences designed using the methods described herein. In one aspect, the 52-base pair tag sequence comprises a double stranded DNA comprising the complementary top and bottom strand pairs of SEQ ID NO: 9-40 or 45-268.

[0037] Another embodiment described herein is a method for designing primers partially complementary to the 52-base pair tag sequences described herein and an adapter primer, the method comprising, executing on a processor: (a) designing tag primers that are partially complementary to the top and bottom strands of tag sequences; and (b) designing an adapter primer that is partially complementary to the top strand of the adapter sequence; wherein: the tag primers comprise a 5'-universal tail sequence complementary to an SP1 or SP2 sequence (SEQ ID NO: 7, 8), a locus specific segment, a ribonucleotide (rN) 6-nucleotides from the 3'-end, a 3'-end mismatch, and a 3'-end block (3'-C3 spacer); and the adapter primer comprises a sequence complementary to the SP1 or SP2 sequence (SEQ ID NO: 7, 8). In one aspect, the primers partially complementary to top and bottom strands of the tag sequences comprise a sequence complementary to the SP1 sequence and the adapter primer comprises a sequence complementary to the SP2 sequence; or the primers partially complementary to top and bottom strands of the tag sequences comprise a sequence complementary to the SP2 sequence and the adapter primer comprises a sequence complementary to the SP1 sequence. In another aspect, amplification of a nucleic acid molecule with the primers that are complementary to the top and bottom strands of tag sequences and primers that are complementary to the top strand of the adapter sequence produces a PCR product that comprises a portion of the tag sequence, a sgDNA sequence, and the adapter sequence. In another aspect, the method further comprises synthesising oligonucleotides comprising the sequences of the forward and reverse tag primers and the adapter primer.

[0038] In another embodiment described herein, the 52-base pair tag sequences and primers partially complementary to the 52-base pair tag sequences are designed and selected using an algorithm predicting whether the primers are likely to be partially complementary and have a propensity to form primer-dimers.

[0039] Another embodiment described herein is one or more primers partially complementary to the 52-base pair tag sequences and one or more adapter primers designed using the methods described herein. In one aspect, the primers partially complementary to the 52-base pair tag sequence comprise the sequences of SEQ ID NO: 3, 4; and the adapter primer comprises the sequence of SEQ ID NO:5.

[0040] Another embodiment described herein is the use of one or more double-stranded 52-base pair tag sequences for identifying on- and off-target CRISPR editing sites.

[0041] It will be apparent to one of ordinary skill in the relevant art that suitable modifications and adaptations to the compositions, formulations, methods, processes, and applications described herein can be made without departing from the scope of any embodiments or aspects thereof. The compositions and methods provided are exemplary and are not intended to limit the scope of any of the specified embodiments. All the various embodiments, aspects, and options disclosed herein can be combined in any variations or iterations. The scope of the methods and processes described herein include all actual or potential combinations of embodiments, aspects, options, examples, and preferences herein described. The methods described herein may omit any component or step, substitute any component or

step disclosed herein, or include any component or step disclosed elsewhere herein. It should also be understood that embodiments may include and otherwise be implemented by a combination of various hardware, software, and electronic components. For example, various microprocessors and application specific integrated circuits ("ASICs") can be utilized, as can software of a variety of languages. Also, servers and various computing devices can be used and can include one or more processing units, one or more computer-readable mediums, one or more input/output interfaces, and various connections (e.g., a system bus) connecting the components. Should the meaning of any terms in any of the patents or publications incorporated by reference conflict with the meaning of the terms used in this disclosure, the meanings of the terms or phrases in this disclosure are controlling. Furthermore, the specification discloses and describes merely exemplary embodiments. All patents and publications cited herein are incorporated by reference herein for the specific teachings thereof.

[0042] Various embodiments and aspects of the inventions described herein are summarized by the following clauses:

[0043] Clause 1. A method for identifying and nominating on- and off-target CRISPR edited sites with improved accuracy and sensitivity, the process comprising the steps of:

[0044] (a) co-delivering a guide sequence RNA (sgRNA) or a two-part CRISPR RNA: trans-activating crRNA (crRNA: tracrRNA) duplex, one or more tag sequences, and an RNA-guided endonuclease to cells;

[0045] (b) incubating the cells for a period of time sufficient for double strand breaks to occur;

[0046] (c) isolating genomic DNA from the cells, fragmenting the genomic DNA, and ligating the fragmented genomic DNA to a unique molecular index containing a universal adapter sequence;

[0047] (d) amplifying the ligated DNA fragments using primers targeting the tag and universal adapter sequences to produce a first set of amplified sequences;

[0048] (e) amplifying the first set of amplified sequences using universal sequencing primers targeting the tails of Tag-pTOP or Tag-pBOT primers to produce a second set of amplified sequences;

[0049] (f) sequencing the pooled sequences and obtaining sequencing data; and

[0050] (g) identifying on-/off-target CRISPR editing loci.

[0051] Clause 2. The method of clause 1, wherein the universal sequencing primers target SP1 or SP2 sequence (SEQ ID NO: 7, 8) tails on the Tag-pTOP or Tag-pBOT primers to produce a second set of amplified sequences.

[0052] Clause 3. The method of clause 1 or 2, wherein the universal sequencing primers target predesigned non-homologous sequence (SEQ ID NO: 269-273) tails on the Tag-pTOP or Tag-pBot primers to produce a second set of amplified sequences.

[0053] Clause 4. The method of any one of clauses 1-3, wherein the universal sequencing primers target predesigned 13-mer tails on the Tag-pTOP or Tag-pBot primers to produce a second set of amplified sequences.

[0054] Clause 5. The method of any one of clauses 1-4, wherein step (g) comprises executing on a processor:

[0055] Clause 6. aligning the sequence data to a reference genome;

[0056] (a) (ii) identifying on-/off-target CRISPR editing loci; and

[0057] (b) (iii) outputting the alignment, analysis, and results data as custom-formatted files, tables or graphics.

[0058] Clause 7. The method of any one of clauses 1-5, further comprising a step following step (e) comprising:

[0059] (a) (e1) normalizing the second set of amplified sequences to produce concentration normalized libraries, pooling the normalized libraries with other samples to produce pooled libraries; and continuing with steps (f)-(i).

[0060] Clause 8. The method of any one of clauses 1-6, wherein step (d) uses a supression PCR method.

[0061] Clause 9. The method of any one of clauses 1-7, wherein the RNA-guided endonuclease comprises an endogenously-expressed Cas enzyme, a Cas expression vector, a Cas protein, or a Cas RNP complex.

[0062] Clause 10. The method of any one of clauses 1-8, wherein the RNA-guided endonuclease comprises an endogenously-expressed Cas9 enzyme, a Cas9 expression vector, a Cas9 protein, or a Cas9 RNP complex.

[0063] Clause 11. The method of any one of clauses 1-9, wherein the cells comprise human or mouse cells.

[0064] Clause 12. The method of any one of clauses 1-10, wherein the period of time is about 24 hours to about 96 hours.

[0065] Clause 13. The method of any one of clauses 1-11, wherein multiple tag sequences are co-delivered.

[0066] Clause 14. The method of any one of clauses 1-12, wherein the tag sequences comprise double-stranded deoxyribooligonucleotides (dsDNA) comprising 52-base pairs.

[0067] Clause 15. The method of any one of clauses 1-13, wherein the tag sequences comprise a 5'-terminal phosphate, and phosphorothioate linkages between the 1^{st} and 2^{nd}, 2^{nd} and 3^{rd}, 50^{th} and 51^{st}, and 51^{st} and 52^{nd} nucleotides.

[0068] Clause 16. The method of any one of clauses 1-14, wherein the tag sequences comprise a double stranded DNA comprising the complementary top and bottom strand pairs of SEQ ID NO: 1-2 or 7-268.

[0069] Clause 17. On- and off-target CRISPR editing sites identified or nominated using the method of any one of clauses 1-15.

[0070] Clause 18. A method for designing 52-base pair tag sequences, the method comprising, executing on a processor:

[0071] (a) randomly generating 13-nucleotide sequences with 40-90% GC content, max homopolymer length A: 2, C: 3, G: 2, T: 2, weighted homopolymer rate <20, self-folding Tm<50° C., and self-dimer Tm<50° C.;

[0072] (b) removing sequences that perfectly align to a particular genome or that are homopolymers or GG or CC dinucleotide motifs and obtaining a set of 13-mers;

[0073] (c) selecting a subset of the 13-mer sequences that contain one or less CC or GG dinucleotide motifs;

[0074] (d) concatenating four of the of 13-mer subset sequences to form random 52-mer sequences;

[0075] (e) aligning the random 52-mer sequences to a genome;

[0076] (f) removing the random 52-mer sequences that have similarity to the genome to produce a subset of 52-mer sequences; and

[0077] (g) outputting the subset of 52-mer sequences and generating the complementary strands to produce double stranded 52-base pair tag sequences.

[0078] Clause 19. The method of clause 17, wherein the genome is human or mouse.

[0079] Clause 20. The method of clause 17 or 18, wherein the 52-base pair tag sequences are-non complementary to the genome.

[0080] Clause 21. The method of any one of clauses 17-19, further comprising designing primers for the 52-base pair tag sequences.

[0081] Clause 22. The method of any one of clauses 17-20, wherein the 52-base pair tag sequences comprise a 5'-terminal phosphate, and phosphorothioate linkages between the 1^{st} and 2^{nd}, 2^{nd} and 3^{rd}, 50^{th} and 51^{st}, and 51^{st} and 52^{nd} nucleotides of the 52-base pair tag sequences.

[0082] Clause 23. The method of any one of clauses 17-21, further comprising synthesizing oligonucleotides comprising the 52-base pair tag sequences, the complement of the 52-base pair tag sequences, or primers for the 52-base pair tag sequences.

[0083] Clause 24. One or more 52-base pair tag sequences designed using the methods of clauses 17-22.

[0084] Clause 25. The 52-base pair tag sequences of clause 23, wherein the 52-base pair tag sequence comprises a double stranded DNA comprising the top and bottom strand pairs of SEQ ID NO: 1-2 or 7-268.

[0085] Clause 26. A method for designing primers partially complementary to the 52-base pair tag sequences of clause 23 and an adapter primer, the method comprising, executing on a processor:

[0086] (a) designing tag primers that are partially complementary to the top and bottom strands of tag sequences; and

[0087] (b) designing an adapter primer that is partially complementary to the top strand of the adapter sequence;

[0088] (c) wherein:

[0089] (d) the tag primers comprise a 5'-universal tail sequence; and

[0090] (e) the adapter primer comprises a sequence complementary to the tails of Tag-pTOP or Tag-pBOT primers.

[0091] Clause 27. The method of clause 25, wherein the 5'-universal tail sequence is complementary to an SP1 or SP2 sequence (SEQ ID NO: 7, 8), a locus specific segment, a ribonucleotide (rN) 6-nucleotides from the 3'-end, a 3'-end mismatch, a 3'-end block (3'-C3 spacer), a predesigned non-homologous sequence (SEQ ID NO: 269-273), or a predesigned 13-mer sequence.

[0092] Clause 28. The method of clause 25 or 26, wherein the primers partially complementary to top and bottom strands of the tag sequences comprise a tail sequence complementary to the SP1 sequence (SEQ ID NO: 7) and the adapter primer comprises a sequence complementary to the SP2 sequence (SEQ ID NO: 8) tail on the Tag-pTOP or Tag-pBOT primers; or the primers partially complementary to top and bottom strands of the tag sequences comprise a tail sequence complementary to the SP2 sequence (SEQ ID NO: 8) and the adapter primer comprises a sequence complementary to the SP1 sequence (SEQ ID NO: 7) tail on the Tag-pTOP or Tag-pBOT primers.

[0093] Clause 29. The method of any one of clauses 25-27, wherein the amplification of a nucleic acid molecule with the primers that are complementary to the top and bottom strands of tag sequences and primers that are complementary to the top strand of the adapter sequence produces a PCR product that comprises a portion of the tag sequence, a sgDNA sequence, and the adapter sequence.

[0094] Clause 30. The method of any one of clauses 25-28, further comprising synthesizing oligonucleotides comprising the sequences of the forward and reverse tag primers and the adapter primer.

[0095] Clause 31. The method of any one of clauses 17-21 and 25-29, wherein the 52-base pair tag sequences and primers partially complementary to the 52-base pair tag sequences are designed and selected using an algorithm predicting whether the primers are likely to be partially complementary and have a propensity to form primer-dimers.

[0096] Clause 32. One or more primers partially complementary to the 52-base pair tag sequences and one or more adapter primers designed using the method of clauses 22-25.

[0097] Clause 33. The primers of clause 32, wherein the primers comprise the sequences of SEQ ID NO: 3, 4; and the adapter primer, wherein the adapter primer comprises the sequence of SEQ ID NO: 5.

[0098] Clause 34. Use of one or more double-stranded 52-base pair tag sequences for identifying on- and off-target CRISPR editing sites.

REFERENCES

[0099] 1. Wienert et al., "Unbiased detection of CRISPR off-targets in vivo using DISCOVER-seq," Science 364 (6437): 286-289 (2019).

[0100] 2. Nobles et al., "IGUIDE: An improved pipeline for analyzing CRISPR cleavage specificity," Genome Biol. 20 (14): 4-9 (2019).

[0101] 3. Tsai et al., "GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases," Nature Biotechnol. 33 (2): 187-197 (2015).

[0102] 4. Yan et al., "BLISS is a versatile and quantitative method for genome-wide profiling of DNA double-strand breaks," Nature Commun. 8:15058 (2017).

[0103] 5. Tsai et al., "CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets," Nature Methods 14 (6): 607-614 (2017).

[0104] 6. Cameron et al., "Mapping the genomic landscape of CRISPR-Cas9 cleavage," Nature Methods 14 (6): 600-606 (2017).

[0105] 7. Char and Moosburner, "Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach," Nature Methods 12 (9): 823-826 (2015).

[0106] 8. Rand et al., "Headloop suppression PCR and its application to selective amplification of methylated DNA sequences," Nucleic Acids Res. 33 (14): e127 (2005).

EXAMPLES

Example 1

[0107] This experiment demonstrates the increased efficiency in tag integration when using double-stranded DNA tags with a length of 52-base pairs and varying genetic sequence. The sequences used are shown in Tables 3-5. Double-stranded tags were generated by hybridization of a top strand and a complementary bottom strand (Tables 3-4; SEQ ID NO: 9-40 or 45-268). Sixteen different tag designs were introduced separately into HEK293 cells constitutively expressing Cas9 together with a guideRNA which targets the

EMX1 locus. Alternatively, either pools of 16 tags or one pool of 112 tags were introduced into HEK293 cells constitutively expressing Cas9 together with a guideRNA which targets the EMX1 locus. GuideRNAs were electroporated at a concentration of 10 UM, whereas the single Tag or pooled Tags were delivered at a final concentration of 0.5 μM. Tag integration levels were determined by targeted amplification using rhAmpSeq primers (SEQ ID NO: 3-4), enriching for known on- and off-target sites of the EMX1 guideRNA. The rhAmpSeq pool for EMX1 consists of 32 sites, which represent empirically determined ON and OFF target loci. Amplified products were sequenced on an Illumina® MiSeq, and tag integration levels were determined using custom software. This example shows that tag integration efficiency varies among single tag constructs individually with a range between 6 (CTL021) and 13 (CTL169, CTL079, CTL002) sites out of a maximum of 32 sites, and is therefore sequence dependent (Single Tags, FIG. **7**). By taking the mathematical union of the single tag results, a hypothetical number of 23 sites was calculated (CTLmax, FIG. **7**). The hypothesis that combining a pool of tags would increase the likelihood of tag integration was tested and was demonstrated (Pooled Tags, Table, FIG. **7**). Pool A1 consists of the tags represented in the Single Tags (see Table 5) and demonstrated that 21 tag integration events were detected out of a maximum of 32 sites, which is higher than achieved with any of the single tags. Similarly, Pool B3 demonstrated integration of a tag at 21 sites out of a maximum of 32 sites. Again, variability between pools was shown (Pooled Tags, FIG. **7**), indicating optimization of tag designs can potentially maximize tag integration.

TABLE 3

Sequences Used for Second Proof of Concept

Name	Sequence (5'→3')	SEQ ID NO
CTL085_TOP_tag	/5Phos/A*C*GAGCGGTAGTCACCTAGTCGTCGTACCAATTCGACGCACACTACTCGC*G*C	SEQ ID NO: 9
CTL085_BOT_tag	/5Phos/G*C*GCGAGTAGTGTGCGTCGAATTGGTACGACGACTAGGTGACTACCGCTC*G*T	SEQ ID NO: 10
CTL169_TOP_tag	/5Phos/T*A*GCGCGAGTAGTCGGACGAGCGGTTACCAATACGCCGCACCTTAATCCG*C*G	SEQ ID NO: 11
CTL169_BOT_tag	/5Phos/C*G*CGGATTAAGGTGCGGCGTATTGGTAACCGCTCGTCCGACTACTCGCGC*T*A	SEQ ID NO: 12
CTL137_TOP_tag	/5Phos/T*C*GCGACAGTAGTCGTTCGGCTAGGTACCTATTACCGCGTAGTTAGCGGC*G*T	SEQ ID NO: 13
CTL137_BOT_tag	/5Phos/A*C*GCCGCTAACTACGCGGTAATAGGTACCTAGCCGAACGACTACTGTCGC*G*A	SEQ ID NO: 14

TABLE 3-continued

Sequences Used for Second Proof of Concept

Name	Sequence (5'→3')	SEQ ID NO
CTL042_TOP_tag	/5Phos/C*G*CGCTACTAGGTGCGTCGAATTGGTACCGATCCGCAATACACTACTCGC*G*C	SEQ ID NO: 15
CTL042_BOT_tag	/5Phos/G*C*GCGAGTAGTGTATTGCGGATCGGTACCAATTCGACGCACCTAGTAGCG*C*G	SEQ ID NO: 16
CTL051_TOP_tag	/5Phos/G*G*TAACGAGCGGTGCGTCGAATTGGTAACCGCTCGTCCGACCTTAATCGC*G*C	SEQ ID NO: 17
CTL051_BOT_tag	/5Phos/G*C*GCGATTAAGGTCGGACGAGCGGTTACCAATTCGACGCACCGCTCGTTA*C*C	SEQ ID NO: 18
CTL167_TOP_tag	/5Phos/T*T*CGGCGCTAGGTGCGGCGTATTGGTAACCGCTCGTCCGTTCGGCGCTAG*G*T	SEQ ID NO: 19
CTL167_BOT_tag	/5Phos/A*C*CTAGCGCCGAACGGACGAGCGGTTACCAATACGCCGCACCTAGCGCCG*A*A	SEQ ID NO: 20
CTL026_TOP_tag	/5Phos/T*A*CGCGACTAGGTGCGCGATTAAGGTACCTATTACCGCGCGACTATGTGC*G*C	SEQ ID NO: 21
CTL026_BOT_tag	/5Phos/G*C*GCACATAGTCGCGCGGTAATAGGTACCTTAATCGCGCACCTAGTCGCG*T*A	SEQ ID NO: 22
CTL068_TOP_tag	/5Phos/G*T*CGCGCAGTGTAGCGCGATTAAGGTACCTATTACCGCGTCGCGACAGTA*G*T	SEQ ID NO: 23
CTL068_BOT_tag	/5Phos/A*C*TACTGTCGCGACGCGGTAATAGGTACCTTAATCGCGCTACACTGCGCG*A*C	SEQ ID NO: 24
CTL138_TOP_tag	/5Phos/A*A*CCGTCGATCCGCGCGTAGTATGGTACCGATCCGCAATACTAGCGCGAC*A*A	SEQ ID NO: 25
CTL138_BOT_tag	/5Phos/T*T*GTCGCGCTAGTATTGCGGATCGGTACCATACTACGCGCGGATCGACGG*T*T	SEQ ID NO: 26
CTL079_TOP_tag	/5Phos/T*C*GCTCGATTGGTTACGCGCACTACTTATGCGCTCGACTCGTTCGGCTAG*G*T	SEQ ID NO: 27
CTL079_BOT_tag	/5Phos/A*C*CTAGCCGAACGAGTCGAGCGCATAAGTAGTGCGCGTAACCAATCGAGC*G*A	SEQ ID NO: 28

TABLE 3-continued

Sequences Used for Second Proof of Concept

Name	Sequence (5'→3')	SEQ ID NO
CTL063_TOP_tag	/5Phos/A*C*TGCGAGCGTACTTGTCGCGCTAGTACCAATTCGACGCAACCGCTCGTC*C*G	SEQ ID NO: 29
CTL063_BOT_tag	/5Phos/C*G*GACGAGCGGTTGCGTCGAATTGGTACTAGCGCGACAAGTACGCTCGCA*G*T	SEQ ID NO: 30
CTL168_TOP_tag	/5Phos/C*G*CATTAGTCGGTGCGGCGTATTGGTAACCGCTCGTCCGACGCGCTACCT*A*T	SEQ ID NO: 31
CTL168_BOT_tag	/5Phos/A*T*AGGTAGCGCGTCGGACGAGCGGTTACCAATACGCCGCACCGACTAATG*C*G	SEQ ID NO: 32
CTL021_TOP_tag	/5Phos/A*T*TGCGGATCGGTGCGTCGAATTGGTAACCGCTCGTCCGTACGCGCACTA*C*T	SEQ ID NO: 33
CTL021_BOT_tag	/5Phos/A*G*TAGTGCGCGTACGGACGAAGCGGTTACCAATTCGCGCACCGATCCGCA*A*T	SEQ ID NO: 34
CTL151_TOP_tag	/5Phos/T*C*GGCGAGTAGTTGCGCGGTTATGGTACCATAACCGCGCAGTAGTACGCG*G*T	SEQ ID NO: 35
CTL151_BOT_tag	/5Phos/A*C*CGCGTACTACTGCGCGGTTATGGTACCATAACCGCGCAACTACTCGCC*G*A	SEQ ID NO: 36
CTL002_TOP_tag	/5Phos/A*C*TAGCGATCGGTACCTAGCGCCGAAACCTATTACCGCGACCTAGCGTTG*C*G	SEQ ID NO: 37
CTL002_BOT_tag	/5Phos/C*G*CAACGCTAGGTCGCGGTAATAGGTTTCGGCGCTAGGTACCGATCGCTA*G*T	SEQ ID NO: 38
CTL134_TOP_tag	/5Phos/T*A*GCGCGTCAAGAGCGCGGTTATGGTTTCGGCGCTAGGTTAACAGCGCGT*C*G	SEQ ID NO: 39
CTL134_BOT_tag	/5Phos/C*G*ACGCGCTGTTAACCTAGCGCCGAAACCATAACCGCGCTCTTGACGCGC*T*A	SEQ ID NO: 40
GuideSeq_TOP_tag	/5Phos/G*T*TTAATTGAGTTGTCATATGTTAATAACGGT*A*T	SEQ ID NO: 41
GuideSeq_BOT_tag	/5Phos/A*T*ACCGTTATTAACATATGACAACTCAATTAA*A*C	SEQ ID NO: 42

TABLE 3-continued

Sequences Used for Second Proof of Concept

Name	Sequence (5'→3')	SEQ ID NO
EMX1 protospacer	GAGTCCGAGCAGAAGAAGAA	SEQ ID NO: 43
AR protospacer	GTTGGAGCATCTGAGTCCAG	SEQ ID NO: 44

"/5Phos/" indicates a 5'-phosphate moiety; "*" indicates a phosphorothioate linkage.

Example 2

[0108] This experiment demonstrates the increased efficiency in tag integration when using double-stranded DNA tags with a length of 52-base pairs and varying genetic sequence. The sequences used are shown in Tables 3-5. Double-stranded tags were generated by hybridization of a top strand and a complementary bottom strand (SEQ ID NO: 9-40 or 45-268). Sixteen different tag designs were introduced separately into HEK293 cells constitutively expressing Cas9 together with a guideRNA which targets the AR locus. Alternatively, either pools of 16 tags or one pool of 112 tags were introduced into HEK293 cells constitutively expressing Cas9 together with a guideRNA which targets the AR locus. GuideRNAs were electroporated at a concentration of 10 μM, whereas the single Tag or pooled Tags were delivered at a final concentration of 0.5 μM. Tag integration levels were determined by targeted amplification using rhAmpSeq primers (SEQ ID NO: 3-4), enriching for known on- and off-target sites of the AR guideRNA. The rhAmpSeq pool for AR consists of 53 sites which represent empirically determined ON and OFF target loci. Amplified products were sequenced on an Illumina® MiSeq, and tag integration levels were determined using custom software. This example shows that tag integration efficiency varies among single tag constructs individually with a range between 35 (CTL085, CTL134) and 41 sites (CTL002) out of a maximum of 53 sites, and is therefore sequence dependent (Single Tags, Table 5, FIG. **8**).

[0109] By taking the mathematical union of the single tag results, a hypothetical number of 47 sites was calculated (CTLmax, FIG. **8**). The hypothesis that combining a pool of tags would increase the likelihood of tag integration was tested and was demonstrated (Pooled Tags, Table 5, FIG. **8**). Pool B4 (see Table 5) demonstrated that 44 tag integration events were detected out of a maximum of 53 sites, which is higher than achieved with any of the single tags. Again, variability between pools was shown (Pooled Tags, Table 5, FIG. **8**), indicating optimization of tag designs can potentially maximize tag integration.

TABLE 4

Tag Sequences

Name	Sequence (5'→3')	SEQ ID NO
CTL085_TOP_tag	/5Phos/A*C*GAGCGGTAGTCACCTAGTCGTCGTACCAATTCGA CGCACACTACTCGC*G*C	SEQ ID NO: 45
CTL169_TOP_tag	/5Phos/T*A*GCGCGAGTAGTCGGACGAGCGGTTACCAATACGC CGCACCTTAATCCG*C*G	SEQ ID NO: 46
CTL137_TOP_tag	/5Phos/T*C*GCGACAGTAGTCGTTCGGCTAGGTACCTATTACC GCGTAGTTAGCGGC*G*T	SEQ ID NO: 47
CTL042_TOP_tag	/5Phos/C*G*CGCTACTAGGTGCGTCGAATTGGTACCGATCCGC AATACACTACTCGC*G*C	SEQ ID NO: 48
CTL051_TOP_tag	/5Phos/G*G*TAACGAGCGGTGCGTCGAATTGGTAACCGCTCGT CCGACCTTAATCGC*G*C	SEQ ID NO: 49
CTL167_TOP_tag	/5Phos/T*T*CGGCGCTAGGTGCGGCGTATTGGTAACCGCTCGT CCGTTCGGCGCTAG*G*T	SEQ ID NO: 50
CTL026_TOP_tag	/5Phos/T*A*CGCGACTAGGTGCGCGATTAAGGTACCTATTACC GCGCGACTATGTGC*G*C	SEQ ID NO: 51
CTL068_TOP_tag	/5Phos/G*T*CGCGCAGTGTAGCGCGATTAAGGTACCTATTACC GCGTCGCGACAGTA*G*T	SEQ ID NO: 52
CTL138_TOP_tag	/5Phos/A*A*CCGTCGATCCGCGCGTAGTATGGTACCGATCCGC AATACTAGCGCGAC*A*A	SEQ ID NO: 53
CTL079_TOP_tag	/5Phos/T*C*GCTCGATTGGTTACGCGCACTACTTATGCGCTCGS ACTCGTTCGGCTAG*G*T	SEQ ID NO: 54
CTL063_TOP_tag	/5Phos/A*C*TGCGAGCGTACTTGTCGCGCTAGTACCAATTCGA CGCAACCGCTCGTC*C*G	SEQ ID NO: 55
CTL168_TOP_tag	/5Phos/C*G*CATTAGTCGGTGCGGCGTATTGGTAACCGCTCGT CCGACGCGCTACCT*A*T	SEQ ID NO: 56
CTL021_TOP_tag	/5Phos/A*T*TGCGGATCGGTGCGTCGAATTGGTAACCGCTCGT CCGTACGCGCACTA*C*T	SEQ ID NO: 57
CTL151_TOP_tag	/5Phos/T*C*GGCGAGTAGTTGCGCGGTTATGGTACCATAACCG CGCAGTAGTACGCG*G*T	SEQ ID NO: 58
CTL002_TOP_tag	/5Phos/A*C*TAGCGATCGGTACCTAGCGCCGAAACCTATTACC GCGACCTAGCGTTG*C*G	SEQ ID NO: 59
CTL134_TOP_tag	/5Phos/T*A*GCGCGTCAAGAGCGCGGTTATGGTTTCGGCGCTA GGTTAACAGCGCGT*C*G	SEQ ID NO: 60
CTL085_BOT_tag	/5Phos/G*C*GCGAGTAGTGTGCGTCGAATTGGTACGACGACTA GGTGACTACCGCTC*G*T	SEQ ID NO: 61
CTL169_BOT_tag	/5Phos/C*G*CGGATTAAGGTGCGGCGTATTGGTAACCGCTCGT CCGACTACTCGCGC*T*A	SEQ ID NO: 62
CTL137_BOT_tag	/5Phos/A*C*GCCGCTAACTACGCGGTAATAGGTACCTAGCCGA ACGACTACTGTCGC*G*A	SEQ ID NO: 63
CTL042_BOT_tag	/5Phos/G*C*GCGAGTAGTGTATTGCGGATCGGTACCAATTCGA CGCACCTAGTAGCG*C*G	SEQ ID NO: 64
CTL051_BOT_tag	/5Phos/G*C*GCGATTAAGGTCGGACGAGCGGTTACCAATTCGA CGCACCGCTCGTTA*C*C	SEQ ID NO: 65
CTL167_BOT_tag	/5Phos/A*C*CTAGCGCCGAACGGACGAGCGGTTACCAATACGC CGCACCTAGCGCCG*A*A	SEQ ID NO: 66
CTL026_BOT_tag	/5Phos/G*C*GCACATAGTCGCGCGGTAATAGGTACCTTAATCGS CGCACCTAGTCGCG*T*A	SEQ ID NO: 67
CTL068_BOT_tag	/5Phos/A*C*TACTGTCGCGACGCGGTAATAGGTACCTTAATCGS CGCTACACTGCGCG*A*C	SEQ ID NO: 68

TABLE 4-continued

Tag Sequences

Name	Sequence (5'→3')	SEQ ID NO
CTL138_BOT_tag	/5Phos/T*T*GTCGCGCTAGTATTGCGGATCGGTACCATACTACGCGATCGACGG*T*T	SEQ ID NO: 69
CTL079_BOT_tag	/5Phos/A*C*CTAGCCGAACGAGTCGAGCGCATAAGTAGTGCGCGTAACCAATCGAGC*G*A	SEQ ID NO: 70
CTL063_BOT_tag	/5Phos/C*G*GACGAGCGGTTGCGTCGAATTGGTACTAGCGCGACAAGTACGCTCGCA*G*T	SEQ ID NO: 71
CTL168_BOT_tag	/5Phos/A*T*AGGTAGCGCGTCGGACGAGCGGTTACCAATACGCCGCACCGACTAATG*C*G	SEQ ID NO: 72
CTL021_BOT_tag	/5Phos/A*G*TAGTGCGCGTACGGACGAGCGGTTACCAATTCGACGCACCGATCCGCA*A*T	SEQ ID NO: 73
CTL151_BOT_tag	/5Phos/A*C*CGCGTACTACTGCGCGGTTATGGTACCATAACCGCGCAACTACTCGCC*G*A	SEQ ID NO: 74
CTL002_BOT_tag	/5Phos/C*G*CAACGCTAGGTCGCGGTAATAGGTTTCGGCGCTAGGTACCGATCGCTA*G*T	SEQ ID NO: 75
CTL134_BOT_tag	/5Phos/C*G*ACGCGCTGTTAACCTAGCGCCGAAACCATAACCGCGCTCTTGACGCGC*T*A	SEQ ID NO: 76
CTL161_TOP_tag	/5Phos/T*A*CACTGCGCGACACTGCGAGCGTACACCTTAATCGCGCTAGTTAGCGGC*G*T	SEQ ID NO: 77
CTL164_TOP_tag	/5Phos/A*A*CCGTCGAGTGCACCGCGTACTACTAATGTCGAACCGCTACGCGCACTA*C*T	SEQ ID NO: 78
CTL030_TOP_tag	/5Phos/C*G*CGGACTAAGGTGCGCGAGTAGTGTTACGCGCACTCACTAATCTAGCCGC*G*A	SEQ ID NO: 79
CTL088_TOP_tag	/5Phos/A*C*TAGTGCGACGAACTACTCGCGCTAACCAATTCGACGCACCGATCGCTA*G*T	SEQ ID NO: 80
CTL148_TOP_tag	/5Phos/A*A*TGTCGAACCGCGCGCGAGTAGTGTACCATAACCGCGCACCTTAGTCCG*C*G	SEQ ID NO: 81
CTL152_TOP_tag	/5Phos/G*C*GTCGAATTGGTACCGCCGACTTATACCAATACGCCGCATAGGTAGCGC*G*T	SEQ ID NO: 82
CTL007_TOP_tag	/5Phos/A*C*CTAGTAGCGCGGCGTCGAATTGGTACTAGCGCGACAACGCGTAGTATG*G*T	SEQ ID NO: 83
CTL141_TOP_tag	/5Phos/A*C*CGCTCGTTACCGCGCGATTAAGGTACGCCGCTAACTACGGTACGGTCG*G*T	SEQ ID NO: 84
CTL064_TOP_tag	/5Phos/A*C*CGCCGACTTATCGTTCGGCTAGGTACCAATTCGACGCACTGCGAGCGT*A*C	SEQ ID NO: 85
CTL158_TOP_tag	/5Phos/A*C*CTTAATCCGCGACTGCGAGCGTACACCTATTACCGCGCGACGCGCTGT*T*A	SEQ ID NO: 86
CTL066_TOP_tag	/5Phos/A*C*GACGACTAGGTACCGCTCGTTACCTCTTGACGCGCTAACCAATTCGAC*G*C	SEQ ID NO: 87
CTL144_TOP_tag	/5Phos/A*C*CATACTACGCGGCGGTTCGACATTACCATAACCGCGCTAGTGCGAGCG*T*A	SEQ ID NO: 88
CTL107_TOP_tag	/5Phos/C*T*TGTACGGCGGTGCGGCGTATTGGTACCAATACGCCGCTCGTCGCACTA*G*T	SEQ ID NO: 89
CTL149_TOP_tag	/5Phos/G*T*ACGCTCGCAGTACCGCCGACTTATACCTTAATCGCGCACTAGCGCGAC*A*A	SEQ ID NO: 90
CTL008_TOP_tag	/5Phos/A*C*GACGACTAGGTTATGGTACGGCGTTAGCGCGAGTCGAGTACCTTAGTCCG*C*G	SEQ ID NO: 91
CTL099_TOP_tag	/5Phos/A*C*GAGCGGTAGTCATAGGTAGCGCGTTCTTGACGCGCGCTAACCGATCGCTA*G*T	SEQ ID NO: 92

TABLE 4-continued

Tag Sequences

Name	Sequence (5'→3')	SEQ ID NO
CTL089_TOP_tag	/5Phos/A*C*CGATCCGCAATGCGTCGAATTGGTACCATAACCGCGCACCGCCGTACA*A*G	SEQ ID NO: 93
CTL081_TOP_tag	/5Phos/A*C*TAGTGCGACGAACTACTGTCGCGAACCTATTACCGCGACCAATCGAGC*G*A	SEQ ID NO: 94
CTL075_TOP_tag	/5Phos/A*C*CGCCGTACAAGTCGCGACAGTAGTAACCGCTCGTCCGTTCGGCGCTAG*G*T	SEQ ID NO: 95
CTL160_TOP_tag	/5Phos/T*C*GTCGCACTAGTCGCATTAGTCGGTAGTAGTACGCGGGTATAGGTAGCGC*G*T	SEQ ID NO: 96
CTL133_TOP_tag	/5Phos/A*C*CAATTCGACGCTAGTTAGCGGCGTACACTACTCGCGCGCACTCGACGG*T*T	SEQ ID NO: 97
CTL076_TOP_tag	/5Phos/C*G*CGGTAATAGGTCGCGGTAATAGGTACGAGCGGTAGTCACACTACTCGC*G*C	SEQ ID NO: 98
CTL024_TOP_tag	/5Phos/T*C*GGCGAGTAGTTTAGTGCGAGCGTAAGTAGTGCGCGTAACCAATCGAGC*G*A	SEQ ID NO: 99
CTL045_TOP_tag	/5Phos/G*T*CGCGCAGTGTAGCGCGGTTATGGTACCATAACCGCGCACTAGTGCGAC*G*A	SEQ ID NO: 100
CTL009_TOP_tag	/5Phos/T*A*TGCGCTCGACTGCGCGATTAAGGTAATGTCGAACGCGCAGTAGTACGCG*G*T	SEQ ID NO: 101
CTL055_TOP_tag	/5Phos/A*C*TAGCGCGACAACGACTATGTGCGCACCAATTCGAGCGCTACGCGCACTA*C*T	SEQ ID NO: 102
CTL101_TOP_tag	/5Phos/A*A*CTACTCGCCGACTTGTACGGCGGTACCAATTCGAGCGCAACTAATCCGC*G*C	SEQ ID NO: 103
CTL135_TOP_tag	/5Phos/C*G*CGGATTAAGGTCTTGTACGGCGGTACCTAGCCGAGCACGTACGCGCACTA*C*T	SEQ ID NO: 104
CTL155_TOP_tag	/5Phos/T*A*GCGCGTCAAGACTTGTACGGCGGTACCGATCCGCGCAATGCACTCGACGG*T*T	SEQ ID NO: 105
CTL122_TOP_tag	/5Phos/C*G*CATTAGTCGGTGCGGCGTATTGGTACGACGACTAGCGGTACCAATACGCC*G*C	SEQ ID NO: 106
CTL080_TOP_tag	/5Phos/A*C*CTAGTAGCGCGGCGCGGTTATGGTACCGACTAATGCGACTAGCGATCG*G*T	SEQ ID NO: 107
CTL126_TOP_tag	/5Phos/A*C*TACTCGCGCTAACCTAGTCGTCGTAATCTAGCCGGCGATACGCTCGCAC*T*A	SEQ ID NO: 108
CTL098_TOP_tag	/5Phos/A*C*CGCCGCTATACGCGCGATTAAGGTGTACGCTCGCGCAGTCGCGGACTAAG*G*T	SEQ ID NO: 109
CTL038_TOP_tag	/5Phos/T*A*CGCGCACTACTAACCGTCGAGTGCGTACGCTCGCGCAGTACCGATCGCTA*G*T	SEQ ID NO: 110
CTL139_TOP_tag	/5Phos/G*T*CGCGCAGTGTATAACAGCGCGTCGTTAGTGCGCGGCAGAACGACGACTAG*G*T	SEQ ID NO: 111
CTL010_TOP_tag	/5Phos/G*C*GTCGAATTGGTCGCGTAGTATGGTACCGCCGCTAGCTACACCAATACGCC*G*C	SEQ ID NO: 112
CTL034_TOP_tag	/5Phos/T*A*CGCGCACTACTTACGCGACTAGGTACCGATCGCTCAGTCGACGCGCTGT*T*A	SEQ ID NO: 113
CTL117_TOP_tag	/5Phos/A*C*GCCGCTAACTATAGTTAGCGGCGTACCAATTCGAGCCGCAACTAATCCGC*G*C	SEQ ID NO: 114
CTL035_TOP_tag	/5Phos/C*G*CGGACTAAGGTTAGTTAGCGGCGTTACGCGCACTCGACTACCGATCCGCA*A*T	SEQ ID NO: 115
CTL121_TOP_tag	/5Phos/A*C*GACGACTAGGTACCGCCGACTTATACGCCGCTAAGCCTAATAGGTAGCGC*G*T	SEQ ID NO: 116

TABLE 4-continued

Tag Sequences

Name	Sequence (5'→3')	SEQ ID NO
CTL106_TOP_tag	/5Phos/C*G*GATCGACGGTTGCGCGAGTAGTGTAGTAGTACGC GGTTACACTGCGCG*A*C	SEQ ID NO: 117
CTL059_TOP_tag	/5Phos/A*T*TGCGGATCGGTACCGCCGACTTATACCGATCCGC AATTCGCTCGATTG*G*T	SEQ ID NO: 118
CTL157_TOP_tag	/5Phos/A*C*TGCGAGCGTACACTGCGAGCGTACACCTTAATCGS CGCACCGCTCGTTA*C*C	SEQ ID NO: 119
CTL015_TOP_tag	/5Phos/A*C*TACTGTCGCGATCGTCGCACTAGTTACGCTCGCAS CTAATTGCGGATCG*G*T	SEQ ID NO: 120
CTL110_TOP_tag	/5Phos/G*G*TAACGAGCGGTTCTCGCGCACTAATTAGTGCGCGS AGAACCATACTACG*C*G	SEQ ID NO: 121
CTL123_TOP_tag	/5Phos/A*C*TACTCGCGCTAGCGCGATTAAGGTACCTTAATCGS CGCAACTACTCGCC*G*A	SEQ ID NO: 122
CTL014_TOP_tag	/5Phos/T*A*CGCGCACTACTCTTGTACGGCGGTACCAATTCGAS CGCAACCGTCGAGT*G*C	SEQ ID NO: 123
CTL131_TOP_tag	/5Phos/A*A*CCGTCGATCCGATTGCGGATCGGTACCTTAATCGS CGCACTAGTGCGAC*G*A	SEQ ID NO: 124
CTL062_TOP_tag	/5Phos/A*G*TAGTGCGCGTATACACTGCGCGACACACTACTCGS CGCACCTTAATCCG*C*G	SEQ ID NO: 125
CTL044_TOP_tag	/5Phos/A*C*GCCGTACCATACGCGGTAATAGGTAGTAGTGCGCS GTATTCGGCGCTAG*G*T	SEQ ID NO: 126
CTL043_TOP_tag	/5Phos/T*A*GCGCGTCAAGAACCTAGCGTTGCGATAAGTCGGCS GGTAGTAGTACGCG*G*T	SEQ ID NO: 127
CTL118_TOP_tag	/5Phos/C*G*CATTAGTCGGTAATCTAGCCGCGAACCATAACCGS CGCACCGATCGCTA*G*T	SEQ ID NO: 128
CTL128_TOP_tag	/5Phos/T*A*TGGTACGGCGTGCGGCGTATTGGTACGCCGCTAAS CTAATAAGTCGGCG*G*T	SEQ ID NO: 129
CTL067_TOP_tag	/5Phos/G*C*GCGGTTATGGTGCGGCGTATTGGTACGAGCGGTAS GTCAACCGCTCGTC*C*G	SEQ ID NO: 130
CTL020_TOP_tag	/5Phos/C*G*ACTATGTGCGCAACTACTCGCCGAACCATAACCGS CGCTATGCGCTCGA*C*T	SEQ ID NO: 131
CTL006_TOP_tag	/5Phos/T*A*GTTAGCGGCGTACCGCTCGTTACCACCTTAATCGS CGCACCATACTACG*C*G	SEQ ID NO: 132
CTL017_TOP_tag	/5Phos/C*G*CATTAGTCGGTAGTAGTGCGCGTAAACCGCTCGTS CCGTTAGTGCGCGA*G*A	SEQ ID NO: 133
CTL057_TOP_tag	/5Phos/T*A*GCGCGAGTAGTACCGACTAATGCGTCTCGCGCACS TAAGACTACCGCTC*G*T	SEQ ID NO: 134
CTL078_TOP_tag	/5Phos/T*A*CGCTCGCACTATCGCTCGATTGGTACCGCCGCTAS TACACCATAACCGC*G*C	SEQ ID NO: 135
CTL031_TOP_tag	/5Phos/A*C*CAATCGAGCGAAGTCGAGCGCATAACGCGCTACCS TATACGCCGCTAAC*T*A	SEQ ID NO: 136
CTL136_TOP_tag	/5Phos/A*C*CTTAATCCGCGACTGCGAGCGTACACCGACTAATS GCGACTACTGTCGC*G*A	SEQ ID NO: 137
CTL165_TOP_tag	/5Phos/A*G*TAGTGCGCGTATCGCTCGATTGGTTCTTGACGCGS CTAGTATAGCGGCG*G*T	SEQ ID NO: 138
CTL039_TOP_tag	/5Phos/T*C*GTCGCACTAGTCGGTACGGTCGGTGCGCACATAGS TCGTATGGTACGGC*G*T	SEQ ID NO: 139
CTL036_TOP_tag	/5Phos/C*G*CGGATTAAGGTAGTCGAGCGCATAACCGCGTACTS ACTACGACGACTAG*G*T	SEQ ID NO: 140

TABLE 4-continued

Tag Sequences

Name	Sequence (5'→3')	SEQ ID NO
CTL048_TOP_tag	/5Phos/C*G*ACTATGTGCGCTACGCTCGCACTAACACTACTCGCGCACCTAGCGCCG*A*A	SEQ ID NO: 141
CTL053_TOP_tag	/5Phos/A*C*CGCCGACTTATTCTCGCGCACTAATCGTCGCACTAGTAACCGTCGATC*C*G	SEQ ID NO: 142
CTL072_TOP_tag	/5Phos/A*C*CTAGCGTTGCGACCGACTAATGCGGGTAACGAGCGGTTATGGTACGGC*G*T	SEQ ID NO: 143
CTL096_TOP_tag	/5Phos/C*G*CGCTACTAGGTCGCGGTAATAGGTACCTAGCGTTGCGACCTAGTCGCG*T*A	SEQ ID NO: 144
CTL150_TOP_tag	/5Phos/C*G*TTCGGCTAGGTACTACTCGCGCTACGCATTAGTCGGGTTCGCGACAGTA*G*T	SEQ ID NO: 145
CTL084_TOP_tag	/5Phos/C*G*GACGAGCGGTTCGCGGTAATAGGTACGACGACTAGGTTAGTTAGCGGC*G*T	SEQ ID NO: 146
CTL142_TOP_tag	/5Phos/T*A*CGCTCGCACTAATTGCGGATCGGTACCGACTAATGCGACCGCGTACTA*C*T	SEQ ID NO: 147
CTL102_TOP_tag	/5Phos/A*C*CGACCGTACCGTATGGTACGGCGTTCTTGACGCGCTAACCTAGCGCCG*A*A	SEQ ID NO: 148
CTL154_TOP_tag	/5Phos/G*C*GCGGATTAGTTAACCGTCGAGTGCACACTACTCGCGCACTGCGAGCGT*A*C	SEQ ID NO: 149
CTL112_TOP_tag	/5Phos/A*C*CTTAATCCGCGACCGACTAATGCGTACGCGCACTACTATAAGTCGGCG*G*T	SEQ ID NO: 150
CTL145_TOP_tag	/5Phos/A*C*CTTAATCCGCGGCGCGGTTATGGTACCGACTAATGCGAACCGCTCGTC*C*G	SEQ ID NO: 151
CTL060_TOP_tag	/5Phos/A*C*TGCGAGCGTACCTTGTACGGCGGTACCTAGTAGCGCGATAAGTCGGCG*G*T	SEQ ID NO: 152
CTL016_TOP_tag	/5Phos/T*T*CGGCGCTAGGTACCTTAGTCCGCGTTCGGCGCTAGGTACCTAGCGTTG*C*G	SEQ ID NO: 153
CTL159_TOP_tag	/5Phos/A*C*CTAGTCGCGTACTTGTACGGCGGTACCTAGCCGAACGAACCGTCGAGT*G*C	SEQ ID NO: 154
CTL056_TOP_tag	/5Phos/A*C*CATAACCGCGCTACACTGCGCGACACCAATACGCCGCTATGGTACGGC*G*T	SEQ ID NO: 155
CTL162_TOP_tag	/5Phos/A*C*ACTACTCGCGCTACGCGACTAGGTAATGTCGAACCGCACGCCGCTAAC*T*A	SEQ ID NO: 156
CTL018_TOP_tag	/5Phos/A*C*CGACTAATGCGTAACAGCGCGTCGTTAGTGCGCGSEQ ID NO: 157 AGAACCTTAATCGC*G*C	SEQ ID NO: 157
CTL115_TOP_tag	/5Phos/A*C*GCCGTACCATAACCGACTAATGCGATAAGTCGGCSEQ ID NO: 158 GGTACCAATACGCC*G*C	SEQ ID NO: 158
CTL033_TOP_tag	/5Phos/G*T*ACGCTCGCAGTCGCGGTAATAGGTTCGGCGAGTASEQ ID NO: 159 GTTACCATAACCGC*G*C	SEQ ID NO: 159
CTL047_TOP_tag	/5Phos/C*G*GACGAGCGGTTGCGCGGTTATGGTACTAGTGCGASEQ ID NO: 160 CGAGCGCACATAGT*C*G	SEQ ID NO: 160
CTL108_TOP_tag	/5Phos/A*C*TACTCGCGCTAGCGCGATTAAGGTACGCCGCTAASEQ ID NO: 161 CTATCGCGGCTAGA*T*T	SEQ ID NO: 161
CTL041_TOP_tag	/5Phos/A*C*CAATTCGACGCAACTAATCCGCGCACCAATTCGASEQ ID NO: 162 CGCAGTAGTGCGCG*T*A	SEQ ID NO: 162
CTL061_TOP_tag	/5Phos/A*C*CGCCGCTATACACCTAGCGCCGAAGTACGCTCGCSEQ ID NO: 163 AGTGTATAGCGGCG*G*T	SEQ ID NO: 163
CTL166_TOP_tag	/5Phos/A*C*ACTACTCGCGCCGGACGAGCGGTTACCAATACGCSEQ ID NO: 164 CGCTAGCGCGAGTA*G*T	SEQ ID NO: 164

TABLE 4-continued

Tag Sequences

Name	Sequence (5'→3')	SEQ ID NO
CTL012_TOP_tag	/5Phos/T*C*GTCGCACTAGTACCTTAATCCGCGCGCAACGCTA GGTACACTACTCGC*G*C	SEQ ID NO: 165
CTL052_TOP_tag	/5Phos/C*G*CGCTACTAGGTACCGACTAATGCGCGCAACGCTA GGTAATGTCGAACC*G*C	SEQ ID NO: 166
CTL153_TOP_tag	/5Phos/A*C*GAGCGGTAGTCACTACTGTCGCGACGCAACGCTA GGTTACACTGCGCG*A*C	SEQ ID NO: 167
CTL094_TOP_tag	/5Phos/A*C*CTAGTCGCGTACGCGTAGTATGGTACCGATCGCT AGTGGTAACGAGCG*G*T	SEQ ID NO: 168
CTL095_TOP_tag	/5Phos/G*C*GGTTCGACATTACCGACTAATGCGTATGCGCTCG ACTACCTAGCGTTG*C*G	SEQ ID NO: 169
CTL105_TOP_tag	/5Phos/A*C*TGCGAGCGTACTCTCGCGCACTAAACGCCGCTAA CTACGCGCTACTAG*G*T	SEQ ID NO: 170
CTL109_TOP_tag	/5Phos/C*G*GTACGGTCGGTAATCTAGCCGCGAACCTTAGTCC GCGACCGCCGTACA*A*G	SEQ ID NO: 171
CTL032_TOP_tag	/5Phos/T*C*GGCGAGTAGTTACGCGCTACCTATTCGCGGCTAG ATTACGCCGCTAAC*T*A	SEQ ID NO: 172
CTL161_BOT_tag	/5Phos/A*C*GCCGCTAACTAGCGCGATTAAGGTGTACGCTCGC AGTGTCGCGCAGTG*T*A	SEQ ID NO: 173
CTL164_BOT_tag	/5Phos/A*G*TAGTGCGCGTAGCGGTTCGACATTAGTAGTACGC GGTGCACTCGACGG*T*T	SEQ ID NO: 174
CTL030_BOT_tag	/5Phos/T*C*GCGGCTAGATTAGTAGTGCGCGTAACACTACTCG CGCACCTTAGTCCG*C*G	SEQ ID NO: 175
CTL088_BOT_tag	/5Phos/A*C*TAGCGATCGGTGCGTCGAATTGGTTAGCGCGAGT AGTTCGTCGCACTA*G*T	SEQ ID NO: 176
CTL148_BOT_tag	/5Phos/C*G*CGGACTAAGGTGCGCGGTTATGGTACACTACTCG CGCGCGGTTCGACA*T*T	SEQ ID NO: 177
CTL152_BOT_tag	/5Phos/A*C*GCGCTACCTATGCGGCGTATTGGTATAAGTCGGC GGTACCAATTCGAC*G*C	SEQ ID NO: 178
CTL007_BOT_tag	/5Phos/A*C*CATACTACGCGTTGTCGCGCTAGTACCAATTCGA CGCCGCGCTACTAG*G*T	SEQ ID NO: 179
CTL141_BOT_tag	/5Phos/A*C*CGACCGTACCGTAGTTAGCGGCGTACCTTAATCG CGCGGTAACGAGCG*G*T	SEQ ID NO: 180
CTL064_BOT_tag	/5Phos/G*T*ACGCTCGCAGTGCGTCGAATTGGTACCTAGCCGA ACGATAAGTCGGCG*G*T	SEQ ID NO: 181
CTL158_BOT_tag	/5Phos/T*A*ACAGCGCGTCGCGCGGTAATAGGTGTACGCTCGC AGTCGCGGATTAAG*G*T	SEQ ID NO: 182
CTL066_BOT_tag	/5Phos/G*C*GTCGAATTGGTTAGCGCGTCAAGAGGTAACGAGC GGTACCTAGTCGTC*G*T	SEQ ID NO: 183
CTL144_BOT_tag	/5Phos/T*A*CGCTCGCACTAGCGCGGTTATGGTAATGTCGAAC CGCCGCGTAGTATG*G*T	SEQ ID NO: 184
CTL107_BOT_tag	/5Phos/A*C*TAGTGCGACGAGCGGCGTATTGGTACCAATACGC CGCACCGCCGTACA*A*G	SEQ ID NO: 185
CTL149_BOT_tag	/5Phos/T*T*GTCGCGCTAGTGCGCGATTAAGGTATAAGTCGGC GGTACTGCGAGCGT*A*C	SEQ ID NO: 186
CTL008_BOT_tag	/5Phos/C*G*CGGACTAAGGTACTACTCGCGCTAACGCCGTACC ATAACCTAGTCGTC*G*T	SEQ ID NO: 187
CTL099_BOT_tag	/5Phos/A*C*TAGCGATCGGTTAGCGCGTCAAGAACGCGCTACC TATGACTACCGCTC*G*T	SEQ ID NO: 188

TABLE 4-continued

Tag Sequences

Name	Sequence (5'→3')	SEQ ID NO
CTL089_BOT_tag	/5Phos/C*T*TGTACGGCGGTGCGCGGTTATGGTACCAATTCGA CGCATTGCGGATCG*G*T	SEQ ID NO: 189
CTL081_BOT_tag	/5Phos/T*C*GCTCGATTGGTCGCGGTAATAGGTTCGCGACAGTS AGTTCGTCGCACTA*G*T	SEQ ID NO: 190
CTL075_BOT_tag	/5Phos/A*C*CTAGCGCCGAACGGACGAGCGGTTACTACTGTCGS CGACTTGTACGGCG*G*T	SEQ ID NO: 191
CTL160_BOT_tag	/5Phos/A*C*GCGCTACCTATACCGCGTACTACTACCGACTAATS GCGACTAGTGCGAC*G*A	SEQ ID NO: 192
CTL133_BOT_tag	/5Phos/A*A*CCGTCGAGTGCGCGCGAGTAGTGTACGCCGCTAAS CTAGCGTCGAATTG*G*T	SEQ ID NO: 193
CTL076_BOT_tag	/5Phos/G*C*GCGAGTAGTGTGACTACCGCTCGTACCTATTACCS GCGACCTATTACCG*C*G	SEQ ID NO: 194
CTL024_BOT_tag	/5Phos/T*C*GCTCGATTGGTTACGCGCACTACTTACGCTCGCAS CTAAACTACTCGCC*G*A	SEQ ID NO: 195
CTL045_BOT_tag	/5Phos/T*C*GTCGCACTAGTGCGCGGTTATGGTACCATAACCGS CGCTACACTGCGCG*A*C	SEQ ID NO: 196
CTL009_BOT_tag	/5Phos/A*C*CGCGTACTACTGCGGTTCGACATTACCTTAATCGS CGCAGTCGAGCGCA*T*A	SEQ ID NO: 197
CTL055_BOT_tag	/5Phos/A*G*TAGTGCGCGTAGCGTCGAATTGGTGCGCACATAGS TCGTTGTCGCGCTA*G*T	SEQ ID NO: 198
CTL101_BOT_tag	/5Phos/G*C*GCGGATTAGTTGCGTCGAATTGGTACCGCCGTACS AAGTCGGCGAGTAG*T*T	SEQ ID NO: 199
CTL135_BOT_tag	/5Phos/A*G*TAGTGCGCGTACGTTCGGCTAGGTACCGCCGTACS AAGACCTTAATCCG*C*G	SEQ ID NO: 200
CTL155_BOT_tag	/5Phos/A*A*CCGTCGAGTGCATTGCGGATCGGTACCGCCGTACS AAGTCTTGACGCGC*T*A	SEQ ID NO: 201
CTL122_BOT_tag	/5Phos/G*C*GGCGTATTGGTACCTAGTCGTCGTACCAATACGCS CGCACCGACTAATG*C*G	SEQ ID NO: 202
CTL080_BOT_tag	/5Phos/A*C*CGATCGCTAGTCGCATTAGTCGGTACCATAACCGS CGCCGCGCTACTAG*G*T	SEQ ID NO: 203
CTL126_BOT_tag	/5Phos/T*A*GTGCGAGCGTATCGCGGCTAGATTACGACGACTAS GGTTAGCGCGAGTA*G*T	SEQ ID NO: 204
CTL098_BOT_tag	/5Phos/A*C*CTTAGTCCGCGACTGCGAGCGTACACCTTAATCGS CGCGTATAGCGGCG*G*T	SEQ ID NO: 205
CTL038_BOT_tag	/5Phos/A*C*TAGCGATCGGTACTGCGAGCGTACGCACTCGACGS GTTAGTAGTGCGCG*T*A	SEQ ID NO: 206
CTL139_BOT_tag	/5Phos/A*C*CTAGTCGTCGTTCTCGCGCACTAACGACGCGCTGS TTATACACTGCGCG*A*C	SEQ ID NO: 207
CTL010_BOT_tag	/5Phos/G*C*GGCGTATTGGTGTATAGCGGCGGTACCATACTACS GCGACCAATTCGAC*G*C	SEQ ID NO: 208
CTL034_BOT_tag	/5Phos/T*A*ACAGCGCGTCGACTAGCGATCGGTACCTAGTCGCS GTAAGTAGTGCGCG*T*A	SEQ ID NO: 209
CTL117_BOT_tag	/5Phos/G*C*GCGGATTAGTTGCGTCGAATTGGTACGCCGCTAAS CTATAGTTAGCGGC*G*T	SEQ ID NO: 210
CTL035_BOT_tag	/5Phos/A*T*TGCGGATCGGTAGTAGTGCGCGTAACGCCGCTAAS CTAACCTTAGTCCG*C*G	SEQ ID NO: 211
CTL121_BOT_tag	/5Phos/A*C*GCGCTACCTATTAGTTAGCGGCGTATAAGTCGGCS GGTACCTAGTCGTC*G*T	SEQ ID NO: 212

TABLE 4-continued

Tag Sequences

Name	Sequence (5'→3')	SEQ ID NO
CTL106_BOT_tag	/5Phos/G*T*CGCGCAGTGTAACCGCGTACTACTACACTACTCGCGAACCGTCGATC*C*G	SEQ ID NO: 213
CTL059_BOT_tag	/5Phos/A*C*CAATCGAGCGAATTGCGGATCGGTATAAGTCGGCGGTACCGATCCGCA*A*T	SEQ ID NO: 214
CTL157_BOT_tag	/5Phos/G*G*TAACGAGCGGTGCGCGATTAAGGTGTACGCTCGCAGTGTACGCTCGCA*G*T	SEQ ID NO: 215
CTL015_BOT_tag	/5Phos/A*C*CGATCCGCAATTAGTGCGAGCGTAACTAGTGCGACGATCGCGACAGTA*G*T	SEQ ID NO: 216
CTL110_BOT_tag	/5Phos/C*G*CGTAGTATGGTTCTCGCGCACTAATTAGTGCGCGAGAACCGCTCGTTA*C*C	SEQ ID NO: 217
CTL123_BOT_tag	/5Phos/T*C*GGCGAGTAGTTGCGCGATTAAGGTACCTTAATCGCGCTAGCGCGAGTA*G*T	SEQ ID NO: 218
CTL014_BOT_tag	/5Phos/G*C*ACTCGACGGTTGCGTCGAATTGGTACCGCCGTACAAGAGTAGTGCGCG*T*A	SEQ ID NO: 219
CTL131_BOT_tag	/5Phos/T*C*GTCGCACTAGTGCGCGATTAAGGTACCGATCCGCAATCGGATCGACGG*T*T	SEQ ID NO: 220
CTL062_BOT_tag	/5Phos/C*G*CGGATTAAGGTGCGCGAGTAGTGTGTCGCGCAGTGTATACGCGCACTA*C*T	SEQ ID NO: 221
CTL044_BOT_tag	/5Phos/A*C*CTAGCGCCGAATACGCGCACTACTACCTATTACCGCGTATGGTACGGC*G*T	SEQ ID NO: 222
CTL043_BOT_tag	/5Phos/A*C*CGCGTACTACTACCGCCGACTTATCGCAACGCTAGGGTTCTTGACGCGC*T*A	SEQ ID NO: 223
CTL118_BOT_tag	/5Phos/A*C*TAGCGATCGGTGCGCGGTTATGGTTCGCGGCTAGATTACCGACTAATG*C*G	SEQ ID NO: 224
CTL128_BOT_tag	/5Phos/A*C*CGCCGACTTATTAGTTAGCGGCGTACCAATACGCCGCACGCCGTACCA*T*A	SEQ ID NO: 225
CTL067_BOT_tag	/5Phos/C*G*GACGAGCGGTTGACTACCGCTCGTACCAATACGCCGCACCATAACCGC*G*C	SEQ ID NO: 226
CTL020_BOT_tag	/5Phos/A*G*TCGAGCGCATAGCGCGGTTATGGTTCGGCGAGTAGTTGCGCACATAGT*C*G	SEQ ID NO: 227
CTL006_BOT_tag	/5Phos/C*G*CGTAGTATGGTGCGCGATTAAGGTGGTAACGAGCGGTACGCCGCTAAC*T*A	SEQ ID NO: 228
CTL017_BOT_tag	/5Phos/T*C*TCGCGCACTAACGGACGAGCGGTTTACGCGCACTACTACCGACTAATG*C*G	SEQ ID NO: 229
CTL057_BOT_tag	/5Phos/A*C*GAGCGGTAGTCTTAGTGCGCGAGACGCATTAGTCGGTACTACTCGCGC*T*A	SEQ ID NO: 230
CTL078_BOT_tag	/5Phos/G*C*GCGGTTATGGTGTATAGCGGCGGTACCAATCGAGCGATAGTGCGAGCG*T*A	SEQ ID NO: 231
CTL031_BOT_tag	/5Phos/T*A*GTTAGCGGCGTATAGGTAGCGCGTTATGCGCTCGGACTTCGCTCGATTG*G*T	SEQ ID NO: 232
CTL136_BOT_tag	/5Phos/T*C*GCGACAGTAGTCGCATTAGTCGGTGTACGCTCGCAGTCGCGGATTAAG*G*T	SEQ ID NO: 233
CTL165_BOT_tag	/5Phos/A*C*CGCCGCTATACTAGCGCGTCAAGAACCAATCGAGCGATACGCGCACTA*C*T	SEQ ID NO: 234
CTL039_BOT_tag	/5Phos/A*C*GCCGTACCATACGACTATGTGCGCACCGACCGTAGCCGACTAGTGCGAC*G*A	SEQ ID NO: 235
CTL036_BOT_tag	/5Phos/A*C*CTAGTCGTCGTAGTAGTACGCGGTTATGCGCTCGGACTACCTTAATCCG*C*G	SEQ ID NO: 236

TABLE 4-continued

Tag Sequences

Name	Sequence (5'→3')	SEQ ID NO
CTL048_BOT_tag	/5Phos/T*T*CGGCGCTAGGTGCGCGAGTAGTGTTAGTGCGAGCGTAGCGCACATAGT*C*G	SEQ ID NO: 237
CTL053_BOT_tag	/5Phos/C*G*GATCGACGGTTACTAGTGCGACGATTAGTGCGCGAGAATAAGTCGGCG*G*T	SEQ ID NO: 238
CTL072_BOT_tag	/5Phos/A*C*GCCGTACCATAACCGCTCGTTACCCGCATTAGTCGGTCGCAACGCTAG*G*T	SEQ ID NO: 239
CTL096_BOT_tag	/5Phos/T*A*CGCGACTAGGTCGCAACGCTAGGTACCTATTACCGCGACCTAGTAGCG*C*G	SEQ ID NO: 240
CTL150_BOT_tag	/5Phos/A*C*TACTGTCGCGAACCGACTAATGCGTAGCGCGAGTAGTACCTAGCCGAA*C*G	SEQ ID NO: 241
CTL084_BOT_tag	/5Phos/A*C*GCCGCTAACTAACCTAGTCGTCGTACCTATTACCGCGAACCGCTCGTC*C*G	SEQ ID NO: 242
CTL142_BOT_tag	/5Phos/A*G*TAGTACGCGGTCGCATTAGTCGGTACCGATCCGCAATTAGTGCGAGCG*T*A	SEQ ID NO: 243
CTL102_BOT_tag	/5Phos/T*T*CGGCGCTAGGTTAGCGCGTCAAGAACGCCGTACCATACGGTACGGTCG*G*T	SEQ ID NO: 244
CTL154_BOT_tag	/5Phos/G*T*ACGCTCGCAGTGCGCGAGTAGTGTGCACTCGACGGTTAACTAATCCGC*G*C	SEQ ID NO: 245
CTL112_BOT_tag	/5Phos/A*C*CGCCGACTTATAGTAGTGCGCGTACGCATTAGTCGGGTCGCGGATTAAG*G*T	SEQ ID NO: 246
CTL145_BOT_tag	/5Phos/C*G*GACGAGCGGTTCGCATTAGTCGGTACCATAACCGCGCCGCGGATTAAG*G*T	SEQ ID NO: 247
CTL060_BOT_tag	/5Phos/A*C*CGCCGACTTATCGCGCTACTAGGTACCGCCGTACAAGGTACGCTCGCA*G*T	SEQ ID NO: 248
CTL016_BOT_tag	/5Phos/C*G*CAACGCTAGGTACCTAGCGCCGAACGCGGACTAAGGTACCTAGCGCCG*A*A	SEQ ID NO: 249
CTL159_BOT_tag	/5Phos/G*C*ACTCGACGGTTCGTTCGGCTAGGTACCGCCGTACAAGTACGCGACTAG*G*T	SEQ ID NO: 250
CTL056_BOT_tag	/5Phos/A*C*GCCGTACCATAGCGGCGTATTGGTGTCGCGCAGGTAGCGCGGTTATG*G*T	SEQ ID NO: 251
CTL162_BOT_tag	/5Phos/T*A*GTTAGCGGCGTGCGGTTCGACATTACCTAGTCGCGGTAGCGCGAGTAGT*G*T	SEQ ID NO: 252
CTL018_BOT_tag	/5Phos/G*C*GCGATTAAGGTTCTCGCGCACTAACGACGCGCTGTTACGCATTAGTCG*G*T	SEQ ID NO: 253
CTL115_BOT_tag	/5Phos/G*C*GGCGTATTGGTACCGCCGACTTATCGCATTAGTCGGGTTATGGTACGGC*G*T	SEQ ID NO: 254
CTL033_BOT_tag	/5Phos/G*C*GCGGTTATGGTAACTACTCGCCGAACCTATTACCGCGACTGCGAGCGT*A*C	SEQ ID NO: 255
CTL047_BOT_tag	/5Phos/C*G*ACTATGTGCGCTCGTCGCACTAGTACCATAACCGGCGAACCGCTCGTC*C*G	SEQ ID NO: 256
CTL108_BOT_tag	/5Phos/A*A*TCTAGCCGCGATAGTTAGCGGCGTACCTTAATCGCGCTAGCGCGAGTA*G*T	SEQ ID NO: 257
CTL041_BOT_tag	/5Phos/T*A*CGCGCACTACTGCGTCGAATTGGTGCGCGGATTAGTTGCGTCGAATTG*G*T	SEQ ID NO: 258
CTL061_BOT_tag	/5Phos/A*C*CGCCGCTATACACTGCGAGCGTACTTCGGCGCTAGGTGTATAGCGGCG*G*T	SEQ ID NO: 259
CTL166_BOT_tag	/5Phos/A*C*TACTCGCGCTAGCGGCGTATTGGTAACCGCTCGTCCGGCGCGAGTAGT*G*T	SEQ ID NO: 260

TABLE 4-continued

Tag Sequences

Name	Sequence (5'→3')	SEQ ID NO
CTL012_BOT_tag	/5Phos/G*C*GCGAGTAGTGTACCTAGCGTTGCGCGCGGATTAAGGTACTAGTGCGAC*G*A	SEQ ID NO: 261
CTL052_BOT_tag	/5Phos/G*C*GGTTCGACATTACCTAGCGTTGCGCGCATTAGTCGGTACCTAGTAGCG*C*G	SEQ ID NO: 262
CTL153_BOT_tag	/5Phos/G*T*CGCGCAGTGTAACCTAGCGTTGCGTCGCGACAGTAGTGACTACCGCTC*G*T	SEQ ID NO: 263
CTL094_BOT_tag	/5Phos/A*C*CGCTCGTTACCACTAGCGATCGGTACCATACTACGCGTACGCGACTAG*G*T	SEQ ID NO: 264
CTL095_BOT_tag	/5Phos/C*G*CAACGCTAGGTAGTCGAGCGCATACGCATTAGTCGGGTAATGTCGAACC*G*C	SEQ ID NO: 265
CTL105_BOT_tag	/5Phos/A*C*CTAGTAGCGCGTAGTTAGCGGCGTTTAGTGCGCGSAGAGTACGCT CGCA*G*T	SEQ ID NO: 266
CTL109_BOT_tag	/5Phos/C*T*TGTACGGCGGTCGCGGACTAAGGTTCGCGGCTAGSATTACCGACCGTAC*C*G	SEQ ID NO: 267
CTL032_BOT_tag	/5Phos/T*A*GTTAGCGGCGTAATCTAGCCGCGAATAGGTAGCGSCGTAACTACTCGCC*G*A	SEQ ID NO: 268

"/5Phos/" indicates a 5'-phosphate moiety; "*" indicates a phosphorothioate linkage.

TABLE 5

Pools of Tag Sequences

Tags Present in Pools	Pools							
	Pool A1	Pool B1	Pool B2	Pool B3	Pool B4	Pool B5	Pool B6	Pool C1
	CTL085	CTL161	CTL089	CTL098	CTL062	CTL048	CTL018	Pool A1
	CTL169	CTL164	CTL081	CTL038	CTL044	CTL053	CTL115	Pool B1
	CTL137	CTL030	CTL075	CTL139	CTL043	CTL072	CTL033	Pool B2
	CTL042	CTL088	CTL160	CTL010	CTL118	CTL096	CTL047	Pool B3
	CTL051	CTL148	CTL133	CTL034	CTL128	CTL150	CTL108	Pool B4
	CTL167	CTL152	CTL076	CTL117	CTL067	CTL084	CTL041	Pool B5
	CTL026	CTL007	CTL024	CTL035	CTL020	CTL142	CTL061	Pool B6
	CTL068	CTL141	CTL045	CTL121	CTL006	CTL102	CTL166	
	CTL138	CTL064	CTL009	CTL106	CTL017	CTL154	CTL012	
	CTL079	CTL158	CTL055	CTL059	CTL057	0TL112	CTL052	
	CTL063	CTL066	CTL101	CTL157	CTL078	0TL145	CTL153	
	CTL168	CTL144	CTL135	CTL015	CTL031	CTL060	CTL094	
	CTL021	CTL107	CTL155	CTL110	CTL136	CTL016	CTL095	
	CTL151	CTL149	CTL122	CTL123	CTL165	CTL159	CTL105	
	CTL002	CTL008	CTL080	CTL014	CTL039	CTL056	CTL109	
	CTL134	CTL099	CTL126	CTL131	CTL036	CTL162	CTL032	

TABLE 6

Non-homologous tails

Name	Sequence (5'→3')	SEQ ID NO:
H1	ACGCGACTATACGCGCAATATGGT	SEQ ID NO: 269
H2	CTAGCGATACTACGCGATACGAGAT	SEQ ID NO: 270
H3	CATAGCGGTATTACGCGAGATTACGA	SEQ ID NO: 271

TABLE 6-continued

Non-homologous tails

Name	Sequence (5'→3')	SEQ ID NO:
H4	CGCGAGTACGTACGATTACCG	SEQ ID NO: 272
H5	ACGCGCGACTATACGCGCCTC	SEQ ID NO: 273

SEQUENCE LISTING

Sequence total quantity: 273

SEQ ID NO: 1 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 1
tcgttcgttc cgctctaacc ggcgaatcta ccgcgcatat ctacgccgca at 52

SEQ ID NO: 2 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 2
attgcggcgt agatatgcgc ggtagattcg ccggttagag cggaacgaac ga 52

SEQ ID NO: 3 moltype = DNA length = 57
FEATURE Location/Qualifiers
source 1..57
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 3
acactctttc cctacacgac gctcttccga tcttctaccg cgcatatcta cgccgct 57

SEQ ID NO: 4 moltype = DNA length = 58
FEATURE Location/Qualifiers
source 1..58
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 4
acactctttc cctacacgac gctcttccga tctatatgcg cggtagattc gccggttt 58

SEQ ID NO: 5 moltype = DNA length = 68
FEATURE Location/Qualifiers
source 1..68
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 5
gtgactggag ttcagacgtg tgctcttccg atctaatgat acggcgacca ccgagatcta 60
cacaaggc 68

SEQ ID NO: 6 moltype = DNA length = 78
FEATURE Location/Qualifiers
source 1..78
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 6
aatgatacgg cgaccaccga gatctacact agatcgcnnw nnwnnacact ctttccctac 60
acgacgctct tccgatct 78

SEQ ID NO: 7 moltype = DNA length = 33
FEATURE Location/Qualifiers
source 1..33
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 7
acactctttc cctacacgac gctcttccga tct 33

SEQ ID NO: 8 moltype = DNA length = 34
FEATURE Location/Qualifiers

-continued

```
source                  1..34
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 8
gtgactggag ttcagacgtg tgctcttccg atct                                34

SEQ ID NO: 9            moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 9
acgagcggta gtcacctagt cgtcgtacca attcgacgca cactactcgc gc             52

SEQ ID NO: 10           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 10
gcgcgagtag tgtgcgtcga attggtacga cgactaggtg actaccgctc gt             52

SEQ ID NO: 11           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 11
tagcgcgagt agtcggacga gcggttacca atacgccgca ccttaatccg cg             52

SEQ ID NO: 12           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 12
cgcggattaa ggtgcggcgt attggtaacc gctcgtccga ctactcgcgc ta             52

SEQ ID NO: 13           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 13
tcgcgacagt agtcgttcgg ctaggtacct attaccgcgt agttagcggc gt             52

SEQ ID NO: 14           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 14
acgccgctaa ctacgcggta ataggtacct agccgaacga ctactgtcgc ga             52

SEQ ID NO: 15           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 15
cgcgctacta ggtgcgtcga attggtaccg atccgcaata cactactcgc gc             52

SEQ ID NO: 16           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 16
gcgcgagtag tgtattgcgg atcggtacca attcgacgca cctagtagcg cg             52

SEQ ID NO: 17           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 17
```

-continued

```
ggtaacgagc ggtgcgtcga attggtaacc gctcgtccga ccttaatcgc gc          52

SEQ ID NO: 18          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 18
gcgcgattaa ggtcggacga gcggttacca attcgacgca ccgctcgtta cc          52

SEQ ID NO: 19          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 19
ttcggcgcta ggtgcggcgt attggtaacc gctcgtccgt tcggcgctag gt          52

SEQ ID NO: 20          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 20
acctagcgcc gaacggacga gcggttacca atacgccgca cctagcgccg aa          52

SEQ ID NO: 21          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 21
tacgcgacta ggtgcgcgat taaggtacct attaccgcgc gactatgtgc gc          52

SEQ ID NO: 22          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 22
gcgcacatag tcgcgcggta ataggtacct taatcgcgca cctagtcgcg ta          52

SEQ ID NO: 23          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 23
gtcgcgcagt gtagcgcgat taaggtacct attaccgcgt cgcgacagta gt          52

SEQ ID NO: 24          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 24
actactgtcg cgacgcggta ataggtacct taatcgcgct acactgcgcg ac          52

SEQ ID NO: 25          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 25
aaccgtcgat ccgcgcgtag tatggtaccg atccgcaata ctagcgcgac aa          52

SEQ ID NO: 26          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 26
ttgtcgcgct agtattgcgg atcggtacca tactacgcgc ggatcgacgg tt          52

SEQ ID NO: 27          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
```

-continued

```
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 27
tcgctcgatt ggttacgcgc actacttatg cgctcgactc gttcggctag gt          52

SEQ ID NO: 28           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 28
acctagccga acgagtcgag cgcataagta gtgcgcgtaa ccaatcgagc ga          52

SEQ ID NO: 29           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 29
actgcgagcg tacttgtcgc gctagtacca attcgacgca accgctcgtc cg          52

SEQ ID NO: 30           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 30
cggacgagcg gttgcgtcga attggtacta gcgcgacaag tacgctcgca gt          52

SEQ ID NO: 31           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 31
cgcattagtc ggtgcggcgt attggtaacc gctcgtccga cgcgctacct at          52

SEQ ID NO: 32           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 32
ataggtagcg cgtcggacga gcggttacca atacgccgca ccgactaatg cg          52

SEQ ID NO: 33           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 33
attgcggatc ggtgcgtcga attggtaacc gctcgtccgt acgcgcacta ct          52

SEQ ID NO: 34           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 34
agtagtgcgc gtacggacga gcggttacca attcgacgca ccgatccgca at          52

SEQ ID NO: 35           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 35
tcggcgagta gttgcgcggt tatggtacca taaccgcgca gtagtacgcg gt          52

SEQ ID NO: 36           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 36
```

-continued

```
accgcgtact actgcgcggt tatggtacca taaccgcgca actactcgcc ga          52

SEQ ID NO: 37          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 37
actagcgatc ggtacctagc gccgaaacct attaccgcga cctagcgttg cg          52

SEQ ID NO: 38          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 38
cgcaacgcta ggtcgcggta ataggtttcg gcgctaggta ccgatcgcta gt          52

SEQ ID NO: 39          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 39
tagcgcgtca agagcgcggt tatggtttcg gcgctaggtt aacagcgcgt cg          52

SEQ ID NO: 40          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 40
cgacgcgctg ttaacctagc gccgaaacca taaccgcgct cttgacgcgc ta          52

SEQ ID NO: 41          moltype = DNA  length = 34
FEATURE                Location/Qualifiers
source                 1..34
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 41
gtttaattga gttgtcatat gttaataacg gtat                              34

SEQ ID NO: 42          moltype = DNA  length = 34
FEATURE                Location/Qualifiers
source                 1..34
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 42
ataccgttat taacatatga caactcaatt aaac                              34

SEQ ID NO: 43          moltype = DNA  length = 20
FEATURE                Location/Qualifiers
source                 1..20
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 43
gagtccgagc agaagaagaa                                              20

SEQ ID NO: 44          moltype = DNA  length = 20
FEATURE                Location/Qualifiers
source                 1..20
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 44
gttggagcat ctgagtccag                                              20

SEQ ID NO: 45          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 45
acgagcggta gtcacctagt cgtcgtacca attcgacgca cactactcgc gc          52

SEQ ID NO: 46          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
```

-continued

```
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 46
tagcgcgagt agtcggacga gcggttacca atacgccgca ccttaatccg cg          52

SEQ ID NO: 47           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 47
tcgcgacagt agtcgttcgg ctaggtacct attaccgcgt agttagcggc gt          52

SEQ ID NO: 48           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 48
cgcgctacta ggtgcgtcga attggtaccg atccgcaata cactactcgc gc          52

SEQ ID NO: 49           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 49
ggtaacgagc ggtgcgtcga attggtaacc gctcgtccga ccttaatcgc gc          52

SEQ ID NO: 50           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 50
ttcggcgcta ggtgcggcgt attggtaacc gctcgtccgt tcggcgctag gt          52

SEQ ID NO: 51           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 51
tacgcgacta ggtgcgcgat taaggtacct attaccgcgc gactatgtgc gc          52

SEQ ID NO: 52           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 52
gtcgcgcagt gtagcgcgat taaggtacct attaccgcgt cgcgacagta gt          52

SEQ ID NO: 53           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 53
aaccgtcgat ccgcgcgtag tatggtaccg atccgcaata ctagcgcgac aa          52

SEQ ID NO: 54           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 54
tcgctcgatt ggttacgcgc actacttatg cgctcgactc gttcggctag gt          52

SEQ ID NO: 55           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 55
```

-continued

actgcgagcg tacttgtcgc gctagtacca attcgacgca accgctcgtc cg 52

SEQ ID NO: 56 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 56
cgcattagtc ggtgcggcgt attggtaacc gctcgtccga cgcgctacct at 52

SEQ ID NO: 57 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 57
attgcggatc ggtgcgtcga attggtaacc gctcgtccgt acgcgcacta ct 52

SEQ ID NO: 58 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 58
tcggcgagta gttgcgcggt tatggtacca taaccgcgca gtagtacgcg gt 52

SEQ ID NO: 59 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 59
actagcgatc ggtacctagc gccgaaacct attaccgcga cctagcgttg cg 52

SEQ ID NO: 60 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 60
tagcgcgtca agagcgcggt tatggtttcg gcgctaggtt aacagcgcgt cg 52

SEQ ID NO: 61 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 61
gcgcgagtag tgtgcgtcga attggtacga cgactaggtg actaccgctc gt 52

SEQ ID NO: 62 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 62
cgcggattaa ggtgcggcgt attggtaacc gctcgtccga ctactcgcgc ta 52

SEQ ID NO: 63 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 63
acgccgctaa ctacgcggta ataggtacct agccgaacga ctactgtcgc ga 52

SEQ ID NO: 64 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 64
gcgcgagtag tgtattgcgg atcggtacca attcgacgca cctagtagcg cg 52

SEQ ID NO: 65 moltype = DNA length = 52
FEATURE Location/Qualifiers

-continued

```
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 65
gcgcgattaa ggtcggacga gcggttacca attcgacgca ccgctcgtta cc          52

SEQ ID NO: 66           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 66
acctagcgcc gaacggacga gcggttacca atacgccgca cctagcgccg aa          52

SEQ ID NO: 67           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 67
gcgcacatag tcgcgcggta ataggtacct taatcgcgca cctagtcgcg ta          52

SEQ ID NO: 68           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 68
actactgtcg cgacgcggta ataggtacct taatcgcgct acactgcgcg ac          52

SEQ ID NO: 69           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 69
ttgtcgcgct agtattgcgg atcggtacca tactacgcgc ggatcgacgg tt          52

SEQ ID NO: 70           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 70
acctagccga acgagtcgag cgcataagta gtgcgcgtaa ccaatcgagc ga          52

SEQ ID NO: 71           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 71
cggacgagcg gttgcgtcga attggtacta gcgcgacaag tacgctcgca gt          52

SEQ ID NO: 72           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 72
ataggtagcg cgtcggacga gcggttacca atacgccgca ccgactaatg cg          52

SEQ ID NO: 73           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 73
agtagtgcgc gtacggacga gcggttacca attcgacgca ccgatccgca at          52

SEQ ID NO: 74           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 74
```

-continued

```
accgcgtact actgcgcggt tatggtacca taaccgcgca actactcgcc ga          52

SEQ ID NO: 75          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 75
cgcaacgcta ggtcgcggta ataggtttcg gcgctaggta ccgatcgcta gt          52

SEQ ID NO: 76          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 76
cgacgcgctg ttaacctagc gccgaaacca taaccgcgct cttgacgcgc ta          52

SEQ ID NO: 77          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 77
tacactgcgc gacactgcga gcgtacacct taatcgcgct agttagcggc gt          52

SEQ ID NO: 78          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 78
aaccgtcgag tgcaccgcgt actactaatg tcgaaccgct acgcgcacta ct          52

SEQ ID NO: 79          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 79
cgcggactaa ggtgcgcgag tagtgttacg cgcactacta atctagccgc ga          52

SEQ ID NO: 80          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 80
actagtgcga cgaactactc gcgctaacca attcgacgca ccgatcgcta gt          52

SEQ ID NO: 81          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 81
aatgtcgaac cgcgcgcgag tagtgtacca taaccgcgca ccttagtccg cg          52

SEQ ID NO: 82          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 82
gcgtcgaatt ggtaccgccg acttatacca atacgccgca taggtagcgc gt          52

SEQ ID NO: 83          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 83
acctagtagc gcggcgtcga attggtacta gcgcgacaac gcgtagtatg gt          52

SEQ ID NO: 84          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
```

-continued

```
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 84
accgctcgtt accgcgcgat taaggtacgc cgctaactac ggtacggtcg gt          52

SEQ ID NO: 85           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 85
accgccgact tatcgttcgg ctaggtacca attcgacgca ctgcgagcgt ac          52

SEQ ID NO: 86           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 86
accttaatcc gcgactgcga gcgtacacct attaccgcgc gacgcgctgt ta          52

SEQ ID NO: 87           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 87
acgacgacta ggtaccgctc gttacctctt gacgcgctaa ccaattcgac gc          52

SEQ ID NO: 88           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 88
accatactac gcggcggttc gacattacca taaccgcgct agtgcgagcg ta          52

SEQ ID NO: 89           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 89
cttgtacggc ggtgcggcgt attggtacca atacgccgct cgtcgcacta gt          52

SEQ ID NO: 90           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 90
gtacgctcgc agtaccgccg acttatacct taatcgcgca ctagcgcgac aa          52

SEQ ID NO: 91           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 91
acgacgacta ggttatggta cggcgttagc gcgagtagta ccttagtccg cg          52

SEQ ID NO: 92           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 92
acgagcggta gtcataggta gcgcgttctt gacgcgctaa ccgatcgcta gt          52

SEQ ID NO: 93           moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 93
```

-continued

```
accgatccgc aatgcgtcga attggtacca taaccgcgca ccgccgtaca ag          52

SEQ ID NO: 94          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 94
actagtgcga cgaactactg tcgcgaacct attaccgcga ccaatcgagc ga          52

SEQ ID NO: 95          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 95
accgccgtac aagtcgcgac agtagtaacc gctcgtccgt tcggcgctag gt          52

SEQ ID NO: 96          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 96
tcgtcgcact agtcgcatta gtcggtagta gtacgcggta taggtagcgc gt          52

SEQ ID NO: 97          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 97
accaattcga cgctagttag cggcgtacac tactcgcgcg cactcgacgg tt          52

SEQ ID NO: 98          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 98
cgcggtaata ggtcgcggta ataggtacga gcggtagtca cactactcgc gc          52

SEQ ID NO: 99          moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 99
tcggcgagta gtttagtgcg agcgtaagta gtgcgcgtaa ccaatcgagc ga          52

SEQ ID NO: 100         moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 100
gtcgcgcagt gtagcgcggt tatggtacca taaccgcgca ctagtgcgac ga          52

SEQ ID NO: 101         moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 101
tatgcgctcg actgcgcgat taaggtaatg tcgaaccgca gtagtacgcg gt          52

SEQ ID NO: 102         moltype = DNA  length = 52
FEATURE                Location/Qualifiers
source                 1..52
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 102
actagcgcga caacgactat gtgcgcacca attcgacgct acgcgcacta ct          52

SEQ ID NO: 103         moltype = DNA  length = 52
FEATURE                Location/Qualifiers
```

-continued

```
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 103
aactactcgc cgacttgtac ggcggtacca attcgacgca actaatccgc gc          52

SEQ ID NO: 104          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 104
cgcggattaa ggtcttgtac ggcggtacct agccgaacgt acgcgcacta ct          52

SEQ ID NO: 105          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 105
tagcgcgtca agacttgtac ggcggtaccg atccgcaatg cactcgacgg tt          52

SEQ ID NO: 106          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 106
cgcattagtc ggtgcggcgt attggtacga cgactaggta ccaatacgcc gc          52

SEQ ID NO: 107          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 107
acctagtagc gcggcgcggt tatggtaccg actaatgcga ctagcgatcg gt          52

SEQ ID NO: 108          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 108
actactcgcg ctaacctagt cgtcgtaatc tagccgcgat acgctcgcac ta          52

SEQ ID NO: 109          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 109
accgccgcta tacgcgcgat taaggtgtac gctcgcagtc gcggactaag gt          52

SEQ ID NO: 110          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 110
tacgcgcact actaaccgtc gagtgcgtac gctcgcagta ccgatcgcta gt          52

SEQ ID NO: 111          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 111
gtcgcgcagt gtataacagc gcgtcgttag tgcgcgagaa cgacgactag gt          52

SEQ ID NO: 112          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 112
```

-continued

```
gcgtcgaatt ggtcgcgtag tatggtaccg ccgctataca ccaatacgcc gc          52

SEQ ID NO: 113          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 113
tacgcgcact acttacgcga ctaggtaccg atcgctagtc gacgcgctgt ta          52

SEQ ID NO: 114          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 114
acgccgctaa ctatagttag cggcgtacca attcgacgca actaatccgc gc          52

SEQ ID NO: 115          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 115
cgcggactaa ggttagttag cggcgttacg cgcactacta ccgatccgca at          52

SEQ ID NO: 116          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 116
acgacgacta ggtaccgccg acttatacgc cgctaactaa taggtagcgc gt          52

SEQ ID NO: 117          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 117
cggatcgacg gttgcgcgag tagtgtagta gtacgcggtt acactgcgcg ac          52

SEQ ID NO: 118          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 118
attgcggatc ggtaccgccg acttataccg atccgcaatt cgctcgattg gt          52

SEQ ID NO: 119          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 119
actgcgagcg tacactgcga gcgtacacct taatcgcgca ccgctcgtta cc          52

SEQ ID NO: 120          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 120
actactgtcg cgatcgtcgc actagttacg ctcgcactaa ttgcggatcg gt          52

SEQ ID NO: 121          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 121
ggtaacgagc ggttctcgcg cactaattag tgcgcgagaa ccatactacg cg          52

SEQ ID NO: 122          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
```

-continued

```
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 122
actactcgcg ctagcgcgat taaggtacct taatcgcgca actactcgcc ga          52

SEQ ID NO: 123          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 123
tacgcgcact actcttgtac ggcggtacca attcgacgca accgtcgagt gc          52

SEQ ID NO: 124          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 124
aaccgtcgat ccgattgcgg atcggtacct taatcgcgca ctagtgcgac ga          52

SEQ ID NO: 125          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 125
agtagtgcgc gtatacactg cgcgacacac tactcgcgca ccttaatccg cg          52

SEQ ID NO: 126          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 126
acgccgtacc atacgcggta ataggtagta gtgcgcgtat tcggcgctag gt          52

SEQ ID NO: 127          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 127
tagcgcgtca agaacctagc gttgcgataa gtcggcggta gtagtacgcg gt          52

SEQ ID NO: 128          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 128
cgcattagtc ggtaatctag ccgcgaacca taaccgcgca ccgatcgcta gt          52

SEQ ID NO: 129          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 129
tatggtacgg cgtgcggcgt attggtacgc cgctaactaa taagtcggcg gt          52

SEQ ID NO: 130          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 130
gcgcggttat ggtgcggcgt attggtacga gcggtagtca accgctcgtc cg          52

SEQ ID NO: 131          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 131
```

-continued

```
cgactatgtg cgcaactact cgccgaacca taaccgcgct atgcgctcga ct          52

SEQ ID NO: 132          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 132
tagttagcgg cgtaccgctc gttaccacct taatcgcgca ccatactacg cg          52

SEQ ID NO: 133          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 133
cgcattagtc ggtagtagtg cgcgtaaacc gctcgtccgt tagtgcgcga ga          52

SEQ ID NO: 134          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 134
tagcgcgagt agtaccgact aatgcgtctc gcgcactaag actaccgctc gt          52

SEQ ID NO: 135          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 135
tacgctcgca ctatcgctcg attggtaccg ccgctataca ccataaccgc gc          52

SEQ ID NO: 136          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 136
accaatcgag cgaagtcgag cgcataacgc gctacctata cgccgctaac ta          52

SEQ ID NO: 137          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 137
accttaatcc gcgactgcga gcgtacaccg actaatgcga ctactgtcgc ga          52

SEQ ID NO: 138          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 138
agtagtgcgc gtatcgctcg attggttctt gacgcgctag tatagcggcg gt          52

SEQ ID NO: 139          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 139
tcgtcgcact agtcggtacg gtcggtgcgc acatagtcgt atggtacggc gt          52

SEQ ID NO: 140          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 140
cgcggattaa ggtagtcgag cgcataaccg cgtactacta cgacgactag gt          52

SEQ ID NO: 141          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
```

-continued

```
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 141
cgactatgtg cgctacgctc gcactaacac tactcgcgca cctagcgccg aa          52

SEQ ID NO: 142          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 142
accgccgact tattctcgcg cactaatcgt cgcactagta accgtcgatc cg          52

SEQ ID NO: 143          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 143
acctagcgtt gcgaccgact aatgcgggta acgagcggtt atggtacggc gt          52

SEQ ID NO: 144          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 144
cgcgctacta ggtcgcggta ataggtacct agcgttgcga cctagtcgcg ta          52

SEQ ID NO: 145          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 145
cgttcggcta ggtactactc gcgctacgca ttagtcggtt cgcgacagta gt          52

SEQ ID NO: 146          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 146
cggacgagcg gttcgcggta ataggtacga cgactaggtt agttagcggc gt          52

SEQ ID NO: 147          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 147
tacgctcgca ctaattgcgg atcggtaccg actaatgcga ccgcgtacta ct          52

SEQ ID NO: 148          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 148
accgaccgta ccgtatggta cggcgttctt gacgcgctaa cctagcgccg aa          52

SEQ ID NO: 149          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 149
gcgcggatta gttaaccgtc gagtgcacac tactcgcgca ctgcgagcgt ac          52

SEQ ID NO: 150          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 150
```

-continued

```
accttaatcc gcgaccgact aatgcgtacg cgcactacta taagtcggcg gt          52

SEQ ID NO: 151          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 151
accttaatcc gcggcgcggt tatggtaccg actaatgcga accgctcgtc cg          52

SEQ ID NO: 152          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 152
actgcgagcg taccttgtac ggcggtacct agtagcgcga taagtcggcg gt          52

SEQ ID NO: 153          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 153
ttcggcgcta ggtaccttag tccgcgttcg gcgctaggta cctagcgttg cg          52

SEQ ID NO: 154          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 154
acctagtcgc gtacttgtac ggcggtacct agccgaacga accgtcgagt gc          52

SEQ ID NO: 155          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 155
accataaccg cgctacactg cgcgacacca atacgccgct atggtacggc gt          52

SEQ ID NO: 156          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 156
acactactcg cgctacgcga ctaggtaatg tcgaaccgca cgccgctaac ta          52

SEQ ID NO: 157          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 157
accgactaat gcgtaacagc gcgtcgttag tgcgcgagaa ccttaatcgc gc          52

SEQ ID NO: 158          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 158
acgccgtacc ataaccgact aatgcgataa gtcggcggta ccaatacgcc gc          52

SEQ ID NO: 159          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 159
gtacgctcgc agtcgcggta ataggttcgg cgagtagtta ccataaccgc gc          52

SEQ ID NO: 160          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
```

-continued

```
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 160
cggacgagcg gttgcgcggt tatggtacta gtgcgacgag cgcacatagt cg          52

SEQ ID NO: 161          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 161
actactcgcg ctagcgcgat taaggtacgc cgctaactat cgcggctaga tt          52

SEQ ID NO: 162          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 162
accaattcga cgcaactaat ccgcgcacca attcgacgca gtagtgcgcg ta          52

SEQ ID NO: 163          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 163
accgccgcta tacacctagc gccgaagtac gctcgcagtg tatagcggcg gt          52

SEQ ID NO: 164          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 164
acactactcg cgccggacga gcggttacca atacgccgct agcgcgagta gt          52

SEQ ID NO: 165          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 165
tcgtcgcact agtaccttaa tccgcgcgca acgctaggta cactactcgc gc          52

SEQ ID NO: 166          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 166
cgcgctacta ggtaccgact aatgcgcgca acgctaggta atgtcgaacc gc          52

SEQ ID NO: 167          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 167
acgagcggta gtcactactg tcgcgacgca acgctaggtt acactgcgcg ac          52

SEQ ID NO: 168          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 168
acctagtcgc gtacgcgtag tatggtaccg atcgctagtg gtaacgagcg gt          52

SEQ ID NO: 169          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 169
```

-continued

```
gcggttcgac attaccgact aatgcgtatg cgctcgacta cctagcgttg cg          52

SEQ ID NO: 170          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 170
actgcgagcg tactctcgcg cactaaacgc cgctaactac gcgctactag gt          52

SEQ ID NO: 171          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 171
cggtacggtc ggtaatctag ccgcgaacct tagtccgcga ccgccgtaca ag          52

SEQ ID NO: 172          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 172
tcggcgagta gttacgcgct acctattcgc ggctagatta cgccgctaac ta          52

SEQ ID NO: 173          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 173
acgccgctaa ctagcgcgat taaggtgtac gctcgcagtg tcgcgcagtg ta          52

SEQ ID NO: 174          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 174
agtagtgcgc gtagcggttc gacattagta gtacgcggtg cactcgacgg tt          52

SEQ ID NO: 175          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 175
tcgcggctag attagtagtg cgcgtaacac tactcgcgca ccttagtccg cg          52

SEQ ID NO: 176          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 176
actagcgatc ggtgcgtcga attggttagc gcgagtagtt cgtcgcacta gt          52

SEQ ID NO: 177          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 177
cgcggactaa ggtgcgcggt tatggtacac tactcgcgcg cggttcgaca tt          52

SEQ ID NO: 178          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 178
acgcgctacc tatgcggcgt attggtataa gtcggcggta ccaattcgac gc          52

SEQ ID NO: 179          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
```

-continued

```
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 179
accatactac gcgttgtcgc gctagtacca attcgacgcc gcgctactag gt          52

SEQ ID NO: 180          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 180
accgaccgta ccgtagttag cggcgtacct taatcgcgcg gtaacgagcg gt          52

SEQ ID NO: 181          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 181
gtacgctcgc agtgcgtcga attggtacct agccgaacga taagtcggcg gt          52

SEQ ID NO: 182          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 182
taacagcgcg tcgcgcggta ataggtgtac gctcgcagtc gcggattaag gt          52

SEQ ID NO: 183          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 183
gcgtcgaatt ggttagcgcg tcaagaggta acgagcggta cctagtcgtc gt          52

SEQ ID NO: 184          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 184
tacgctcgca ctagcgcggt tatggtaatg tcgaaccgcc gcgtagtatg gt          52

SEQ ID NO: 185          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 185
actagtgcga cgagcggcgt attggtacca atacgccgca ccgccgtaca ag          52

SEQ ID NO: 186          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 186
ttgtcgcgct agtgcgcgat taaggtataa gtcggcggta ctgcgagcgt ac          52

SEQ ID NO: 187          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 187
cgcggactaa ggtactactc gcgctaacgc cgtaccataa cctagtcgtc gt          52

SEQ ID NO: 188          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 188
```

-continued

actagcgatc ggttagcgcg tcaagaacgc gctacctatg actaccgctc gt 52

SEQ ID NO: 189 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 189
cttgtacggc ggtgcgcggt tatggtacca attcgacgca ttgcggatcg gt 52

SEQ ID NO: 190 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 190
tcgctcgatt ggtcgcggta ataggttcgc gacagtagtt cgtcgcacta gt 52

SEQ ID NO: 191 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 191
acctagcgcc gaacggacga gcggttacta ctgtcgcgac ttgtacggcg gt 52

SEQ ID NO: 192 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 192
acgcgctacc tataccgcgt actactaccg actaatgcga ctagtgcgac ga 52

SEQ ID NO: 193 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 193
aaccgtcgag tgcgcgcgag tagtgtacgc cgctaactag cgtcgaattg gt 52

SEQ ID NO: 194 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 194
gcgcgagtag tgtgactacc gctcgtacct attaccgcga cctattaccg cg 52

SEQ ID NO: 195 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 195
tcgctcgatt ggttacgcgc actacttacg ctcgcactaa actactcgcc ga 52

SEQ ID NO: 196 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 196
tcgtcgcact agtgcgcggt tatggtacca taaccgcgct acactgcgcg ac 52

SEQ ID NO: 197 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 197
accgcgtact actgcggttc gacattacct taatcgcgca gtcgagcgca ta 52

SEQ ID NO: 198 moltype = DNA length = 52
FEATURE Location/Qualifiers

-continued

```
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 198
agtagtgcgc gtagcgtcga attggtgcgc acatagtcgt tgtcgcgcta gt          52

SEQ ID NO: 199          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 199
gcgcggatta gttgcgtcga attggtaccg ccgtacaagt cggcgagtag tt          52

SEQ ID NO: 200          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 200
agtagtgcgc gtacgttcgg ctaggtaccg ccgtacaaga ccttaatccg cg          52

SEQ ID NO: 201          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 201
aaccgtcgag tgcattgcgg atcggtaccg ccgtacaagt cttgacgcgc ta          52

SEQ ID NO: 202          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 202
gcggcgtatt ggtacctagt cgtcgtacca atacgccgca ccgactaatg cg          52

SEQ ID NO: 203          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 203
accgatcgct agtcgcatta gtcggtacca taaccgcgcc gcgctactag gt          52

SEQ ID NO: 204          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 204
tagtgcgagc gtatcgcggc tagattacga cgactaggtt agcgcgagta gt          52

SEQ ID NO: 205          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 205
accttagtcc gcgactgcga gcgtacacct taatcgcgcg tatagcggcg gt          52

SEQ ID NO: 206          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 206
actagcgatc ggtactgcga gcgtacgcac tcgacggtta gtagtgcgcg ta          52

SEQ ID NO: 207          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 207
```

-continued

```
acctagtcgt cgttctcgcg cactaacgac gcgctgttat acactgcgcg ac          52

SEQ ID NO: 208          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 208
gcggcgtatt ggtgtatagc ggcggtacca tactacgcga ccaattcgac gc          52

SEQ ID NO: 209          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 209
taacagcgcg tcgactagcg atcggtacct agtcgcgtaa gtagtgcgcg ta          52

SEQ ID NO: 210          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 210
gcgcggatta gttgcgtcga attggtacgc cgctaactat agttagcggc gt          52

SEQ ID NO: 211          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 211
attgcggatc ggtagtagtg cgcgtaacgc cgctaactaa ccttagtccg cg          52

SEQ ID NO: 212          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 212
acgcgctacc tattagttag cggcgtataa gtcggcggta cctagtcgtc gt          52

SEQ ID NO: 213          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 213
gtcgcgcagt gtaaccgcgt actactacac tactcgcgca accgtcgatc cg          52

SEQ ID NO: 214          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 214
accaatcgag cgaattgcgg atcggtataa gtcggcggta ccgatccgca at          52

SEQ ID NO: 215          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 215
ggtaacgagc ggtgcgcgat taaggtgtac gctcgcagtg tacgctcgca gt          52

SEQ ID NO: 216          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 216
accgatccgc aattagtgcg agcgtaacta gtgcgacgat cgcgacagta gt          52

SEQ ID NO: 217          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
```

-continued

```
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 217
cgcgtagtat ggttctcgcg cactaattag tgcgcgagaa ccgctcgtta cc          52

SEQ ID NO: 218          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 218
tcggcgagta gttgcgcgat taaggtacct taatcgcgct agcgcgagta gt          52

SEQ ID NO: 219          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 219
gcactcgacg gttgcgtcga attggtaccg ccgtacaaga gtagtgcgcg ta          52

SEQ ID NO: 220          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 220
tcgtcgcact agtgcgcgat taaggtaccg atccgcaatc ggatcgacgg tt          52

SEQ ID NO: 221          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 221
cgcggattaa ggtgcgcgag tagtgtgtcg cgcagtgtat acgcgcacta ct          52

SEQ ID NO: 222          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 222
acctagcgcc gaatacgcgc actactacct attaccgcgt atggtacggc gt          52

SEQ ID NO: 223          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 223
accgcgtact actaccgccg acttatcgca acgctaggtt cttgacgcgc ta          52

SEQ ID NO: 224          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 224
actagcgatc ggtgcgcggt tatggttcgc ggctagatta ccgactaatg cg          52

SEQ ID NO: 225          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 225
accgccgact tattagttag cggcgtacca atacgccgca cgccgtacca ta          52

SEQ ID NO: 226          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 226
```

-continued

cggacgagcg gttgactacc gctcgtacca atacgccgca ccataaccgc gc 52

SEQ ID NO: 227 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 227
agtcgagcgc atagcgcggt tatggttcgg cgagtagttg cgcacatagt cg 52

SEQ ID NO: 228 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 228
cgcgtagtat ggtgcgcgat taaggtggta acgagcggta cgccgctaac ta 52

SEQ ID NO: 229 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 229
tctcgcgcac taacggacga gcggtttacg cgcactacta ccgactaatg cg 52

SEQ ID NO: 230 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 230
acgagcggta gtcttagtgc gcgagacgca ttagtcggta ctactcgcgc ta 52

SEQ ID NO: 231 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 231
gcgcggttat ggtgtatagc ggcggtacca atcgagcgat agtgcgagcg ta 52

SEQ ID NO: 232 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 232
tagttagcgg cgtataggta gcgcgttatg cgctcgactt cgctcgattg gt 52

SEQ ID NO: 233 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 233
tcgcgacagt agtcgcatta gtcggtgtac gctcgcagtc gcggattaag gt 52

SEQ ID NO: 234 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 234
accgccgcta tactagcgcg tcaagaacca atcgagcgat acgcgcacta ct 52

SEQ ID NO: 235 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 235
acgccgtacc atacgactat gtgcgcaccg accgtaccga ctagtgcgac ga 52

SEQ ID NO: 236 moltype = DNA length = 52
FEATURE Location/Qualifiers

-continued

```
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 236
acctagtcgt cgtagtagta cgcggttatg cgctcgacta ccttaatccg cg          52

SEQ ID NO: 237          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 237
ttcggcgcta ggtgcgcgag tagtgttagt gcgagcgtag cgcacatagt cg          52

SEQ ID NO: 238          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 238
cggatcgacg gttactagtg cgacgattag tgcgcgagaa taagtcggcg gt          52

SEQ ID NO: 239          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 239
acgccgtacc ataaccgctc gttacccgca ttagtcggtc gcaacgctag gt          52

SEQ ID NO: 240          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 240
tacgcgacta ggtcgcaacg ctaggtacct attaccgcga cctagtagcg cg          52

SEQ ID NO: 241          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 241
actactgtcg cgaaccgact aatgcgtagc gcgagtagta cctagccgaa cg          52

SEQ ID NO: 242          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 242
acgccgctaa ctaacctagt cgtcgtacct attaccgcga accgctcgtc cg          52

SEQ ID NO: 243          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 243
agtagtacgc ggtcgcatta gtcggtaccg atccgcaatt agtgcgagcg ta          52

SEQ ID NO: 244          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 244
ttcggcgcta ggttagcgcg tcaagaacgc cgtaccatac ggtacggtcg gt          52

SEQ ID NO: 245          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 245
```

-continued

gtacgctcgc agtgcgcgag tagtgtgcac tcgacggtta actaatccgc gc 52

SEQ ID NO: 246 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 246
accgccgact tatagtagtg cgcgtacgca ttagtcggtc gcggattaag gt 52

SEQ ID NO: 247 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 247
cggacgagcg gttcgcatta gtcggtacca taaccgcgcc gcggattaag gt 52

SEQ ID NO: 248 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 248
accgccgact tatcgcgcta ctaggtaccg ccgtacaagg tacgctcgca gt 52

SEQ ID NO: 249 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 249
cgcaacgcta ggtacctagc gccgaacgcg gactaaggta cctagcgccg aa 52

SEQ ID NO: 250 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 250
gcactcgacg gttcgttcgg ctaggtaccg ccgtacaagt acgcgactag gt 52

SEQ ID NO: 251 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 251
acgccgtacc atagcggcgt attggtgtcg cgcagtgtag cgcggttatg gt 52

SEQ ID NO: 252 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 252
tagttagcgg cgtgcggttc gacattacct agtcgcgtag cgcgagtagt gt 52

SEQ ID NO: 253 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 253
gcgcgattaa ggttctcgcg cactaacgac gcgctgttac gcattagtcg gt 52

SEQ ID NO: 254 moltype = DNA length = 52
FEATURE Location/Qualifiers
source 1..52
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 254
gcggcgtatt ggtaccgccg acttatcgca ttagtcggtt atggtacggc gt 52

SEQ ID NO: 255 moltype = DNA length = 52
FEATURE Location/Qualifiers

-continued

```
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 255
gcgcggttat ggtaactact cgccgaacct attaccgcga ctgcgagcgt ac          52

SEQ ID NO: 256          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 256
cgactatgtg cgctcgtcgc actagtacca taaccgcgca accgctcgtc cg          52

SEQ ID NO: 257          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 257
aatctagccg cgatagttag cggcgtacct taatcgcgct agcgcgagta gt          52

SEQ ID NO: 258          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 258
tacgcgcact actgcgtcga attggtgcgc ggattagttg cgtcgaattg gt          52

SEQ ID NO: 259          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 259
accgccgcta tacactgcga gcgtacttcg gcgctaggtg tatagcggcg gt          52

SEQ ID NO: 260          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 260
actactcgcg ctagcggcgt attggtaacc gctcgtccgg cgcgagtagt gt          52

SEQ ID NO: 261          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 261
gcgcgagtag tgtacctagc gttgcgcgcg gattaaggta ctagtgcgac ga          52

SEQ ID NO: 262          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 262
gcggttcgac attacctagc gttgcgcgca ttagtcggta cctagtagcg cg          52

SEQ ID NO: 263          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 263
gtcgcgcagt gtaacctagc gttgcgtcgc gacagtagtg actaccgctc gt          52

SEQ ID NO: 264          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 264
```

-continued

```
accgctcgtt accactagcg atcggtacca tactacgcgt acgcgactag gt          52

SEQ ID NO: 265          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 265
cgcaacgcta ggtagtcgag cgcatacgca ttagtcggta atgtcgaacc gc          52

SEQ ID NO: 266          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 266
acctagtagc gcgtagttag cggcgtttag tgcgcgagag tacgctcgca gt          52

SEQ ID NO: 267          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 267
cttgtacggc ggtcgcggac taaggttcgc ggctagatta ccgaccgtac cg          52

SEQ ID NO: 268          moltype = DNA  length = 52
FEATURE                 Location/Qualifiers
source                  1..52
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 268
tagttagcgg cgtaatctag ccgcgaatag gtagcgcgta actactcgcc ga          52

SEQ ID NO: 269          moltype = DNA  length = 24
FEATURE                 Location/Qualifiers
source                  1..24
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 269
acgcgactat acgcgcaata tggt                                        24

SEQ ID NO: 270          moltype = DNA  length = 25
FEATURE                 Location/Qualifiers
source                  1..25
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 270
ctagcgatac tacgcgatac gagat                                       25

SEQ ID NO: 271          moltype = DNA  length = 26
FEATURE                 Location/Qualifiers
source                  1..26
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 271
catagcggta ttacgcgaga ttacga                                      26

SEQ ID NO: 272          moltype = DNA  length = 21
FEATURE                 Location/Qualifiers
source                  1..21
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 272
cgcgagtacg tacgattacc g                                           21

SEQ ID NO: 273          moltype = DNA  length = 21
FEATURE                 Location/Qualifiers
source                  1..21
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 273
acgcgcgact atacgcgcct c                                           21
```

1. A method for identifying and nominating on- and off-target CRISPR edited sites with improved accuracy and sensitivity, the process comprising the steps of:

(a) isolating genomic DNA from a cell having one or more tag sequences incorporated into a target site within a genome of the cell;

(b) integrating a universal adapter sequence comprising a unique molecular index (UMI) into the isolated genomic DNA;

(c) providing a multiplex PCR reaction mixture comprising:

(i) one or more on-target oligonucleotide primers, each having a cleavage region comprising a ribonucleotide (rN) positioned 5' of a blocking group and a complementary region flanking the on-target genome edited locus, wherein the blocking group prevents primer extension and/or inhibits the oligonucleotide primer from serving as a template for DNA synthesis;

(ii) one or more adapter-specific oligonucleotide primers, each having a cleavage region comprising a ribonucleotide (rN) positioned 5' of a blocking group and a complementary region flanking the 5' of the universal adapter sequence; and

(iii) a cleaving enzyme, wherein the cleaving enzyme is an RNase H2 enzyme;

(d) hybridizing the on-target oligonucleotide primer to the on-target genome edited locus to form an on-target double stranded substrate and hybridizing the one or more adapter-specific oligonucleotide primers to the 5' of the universal adapter sequence;

(e) cleaving at a point within or adjacent to the cleavage region to remove the blocking group from the one or more on-target oligonucleotide primers and the one or more adapter-specific oligonucleotide primers; and

(f) simultaneously amplifying a portion of the isolated genomic DNA comprising the one or more tag sequences and the universal adapter sequence; and

(g) sequencing the amplified portion of the isolated genomic DNA, thereby identifying on- and off-target CRISPR edited sites.

2. The method of claim **1**, wherein the one or more adapter-specific oligonucleotide primers target SP1 or SP2 sequence (SEQ ID NO: 7, 8) tails on a top strand of the one or more on-target oligonucleotide primers or a bottom strand of the one or more on-target oligonucleotide primers.

3. The method of claim **1**, wherein the one or more adapter-specific oligonucleotide primers target predesigned non-homologous sequence (SEQ ID NO: 269-273) tails on a top strand of the one or more on-target oligonucleotide primers or a bottom strand of the one or more on-target oligonucleotide primers.

4. The method of claim **1**, wherein the one or more adapter-specific oligonucleotide primers target predesigned 13-mer tails on a top strand of the one or more on-target oligonucleotide primers or a bottom strand of the one or more on-target oligonucleotide primers.

5. The method of claim **1**, wherein the sequencing of step (g) further comprises executing on a processor:

(i) aligning the sequence data to a reference genome; and

(ii) outputting the alignment, analysis, and results data as custom-formatted files, tables or graphics.

6. (canceled)

7. The method of claim **1**, wherein step (d) uses a suppression PCR method.

8. The method of claim **1**, wherein the one or more on-target oligonucleotide primers comprise a first on-target oligonucleotide primer targeting a top strand of the isolated genomic DNA and a second on-target oligonucleotide primer targeting a bottom strand of the isolated genomic DNA.

9. The method of claim **1**, wherein the one or more adapter-specific oligonucleotide primers comprise a first adapter-specific oligonucleotide primer targeting a top strand of the isolated genomic DNA and a second adapter-specific oligonucleotide primer targeting a bottom strand of the isolated genomic DNA.

10. The method of claim **1**, wherein the cells comprise human or mouse cells.

11-12. (canceled)

13. The method of claim **1**, wherein the one or more tag sequences comprise double-stranded deoxyribooligonucleotides (dsDNA) comprising 52-base pairs.

14. The method of claim **1**, wherein the one or more tag sequences comprise a 5'-terminal phosphate, and phosphorothioate linkages between the 1^{st} and 2^{nd}, 2^{nd} and 3^{rd}, 50^{th} and 51^{st}, and 51^{st} and 52^{nd} nucleotides.

15. The method of claim **1**, wherein the one or more tag sequences comprise a double stranded DNA comprising the complementary top and bottom strand pairs of SEQ ID NO: 1-2 or 7-268.

16. On- and off-target CRISPR editing sites identified or nominated using the method of claim **1**.

17-33. (canceled)

34. The method of claim **1**, wherein the one or more tag sequences alien sequence content containing no sequence identity to a mouse or human genome.

35. The method of claim **1**, wherein the cleavage region comprises a ribonucleotide (rN) that is positioned 6-nucleotides from the 3'-end.

* * * * *