



- (51) **International Patent Classification:**  
C40B 30/04 (2006.01) C12Q 1/6806 (2018.01)
- (21) **International Application Number:**  
PCT/US2023/077435
- (22) **International Filing Date:**  
20 October 2023 (20.10.2023)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**  
63/380,488 21 October 2022 (21.10.2022) US  
63/516,033 27 July 2023 (27.07.2023) US
- (71) **Applicant: WATCHMAKER GENOMICS, INC.**  
[US/US]; 5744 Central Ave., Suite 100, Boulder, CO 80301 (US).
- (72) **Inventors; and**
- (71) **Applicants: ROSS, Clara;** c/o Watchmaker Genomics, Inc., 5744 Central Ave., Suite 100, Boulder, CO 80301 (US). **MARSHALL, Craig** [US/US]; c/o Watchmaker Ge-

nomics, Inc., 5744 Central Ave., Suite 100, Boulder, CO 80301 (US). **PETERKIN, Lindsay** [US/US]; c/o Watchmaker Genomics, Inc., 5744 Central Ave., Suite 100, Boulder, CO 80301 (US). **SANDERS, Travis, J.** [US/US]; c/o Watchmaker Genomics, Inc., 5744 Central Ave., Suite 100, Boulder, CO 80301 (US).

(72) **Inventors: RANIK, Martin;** c/o Watchmaker Genomics, Inc., 5744 Central Ave., Suite 100, Boulder, CO 80301 (US). **VAN DER WALT, Eric;** c/o Watchmaker Genomics, Inc., 5744 Central Ave., Suite 100, Boulder, CO 80301 (US). **KUDLOW, Brian;** c/o Watchmaker Genomics, Inc., 5744 Central Ave., Suite 100, Boulder, CO 80301 (US).

(74) **Agent: MATHIS, Andrew, D.** et al.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210-2206 (US).

(81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,

(54) **Title:** METHODS AND COMPOSITIONS FOR SEQUENCING LIBRARY NORMALIZATION

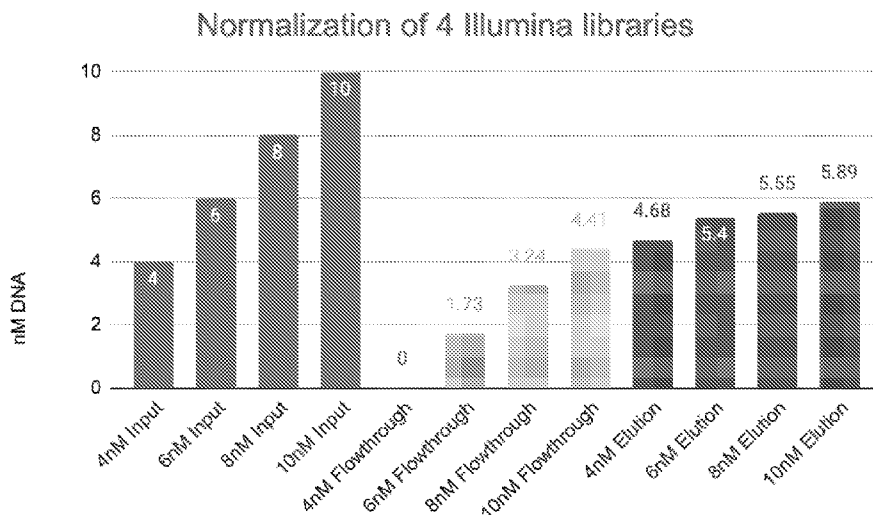


FIG. 5

(57) **Abstract:** Disclosed herein are methods and compositions for normalizing polynucleotide concentration. Normalizing may be accomplished using polynucleotide binding proteins (e.g., catalytically inactive CRISPR protein or catalytically inactive Argonaute proteins). The polynucleotide binding proteins may bind to the adapter sequences of a target polynucleotide. Thus, adding the same amount of a polynucleotide binding proteins to different samples and then extracting the polynucleotide protein is shown herein to extract similar amounts of target polynucleotides from the different samples.



HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

**(84) Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

**Published:**

- *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*
- *with sequence listing part of description (Rule 5.2(a))*

## METHODS AND COMPOSITIONS FOR SEQUENCING LIBRARY NORMALIZATION

## RELATED APPLICATIONS

This application claims the benefit under 35 U.S.C. §119(e) to U.S. Provisional  
5 Application No. 63/380488, filed October 21, 2022, entitled “Methods for Sequencing Library  
Normalization”, and to U.S. Provisional Application No. 63/516033, filed July 27, 2023, entitled  
“Methods and Compositions for Sequencing Library Normalization”, the entire contents of each  
of which are incorporated herein by reference.

## REFERENCE TO AN ELECTRONIC SEQUENCE LISTING

10 The content of the electronic sequence listing (W109470000WO00-SEQ-ARM.xml;  
Size: 53,547 bytes; and Date of Creation: October 20, 2023) is herein incorporated by reference  
in its entirety.

## BACKGROUND

15 Massively parallel, deep sequencing, or next generation sequencing (NGS) enables large-  
scale DNA and RNA sequencing. These methods involve parallel multiplexed analysis of nucleic  
acid sequences on a massive scale, allowing for millions to billions of sequences from individual  
strands to be analyzed separately but simultaneously.

20 The multiplexed analysis can be challenging because different samples to be analyzed  
often vary in the amount of nucleic acids present, which can lead to inaccuracies where low  
concentration samples are under-sequenced and high concentration samples are over-sequenced.  
Therefore, various methods for the normalization of nucleic acid concentrations between samples  
have been employed. Spectrophotometry, electrophoreses, fluorometry, and quantitative PCR  
25 (qPCR) have been used to detect the concentration of nucleic acids in a sample in order to  
normalize concentrations between samples. Each of these methods has disadvantages such as  
inaccuracy introduced by manual adjustments being made, lack of sensitivity, and multiple steps  
involving substantial time and/or expense.

30

## SUMMARY

This disclosure describes, in some aspects, methods and compositions for normalizing polynucleotide concentrations between two or more polynucleotide sequencing libraries. Polynucleotide preparation for sequencing (e.g., next-generation sequencing (NGS)) often involves adding adapter polynucleotide sequences to the polynucleotides. Certain technologies, like ILLUMINA sequencing, use the adapter sequences to perform sequencing of the polynucleotides. It is recognized herein that these adapter sequences can be leveraged to normalize polynucleotide concentrations between different polynucleotide sequencing libraries. Specifically, it is recognized that polynucleotide sequencing library normalization may be achieved by contacting a polynucleotide sequencing library with a predetermined amount of a ribonucleoprotein that binds to an adapter sequence of the polynucleotide library (e.g., a catalytically inactive CRISPR associated protein (dCAS)-guide RNA (gRNA) complex) and extracting the polynucleotides that (1) comprise the adapter and (2) are bound by the ribonucleoprotein complex. It is demonstrated that performing this method for multiple different polynucleotide sequencing libraries normalizes the concentration of polynucleotides between the polynucleotide sequencing libraries.

Accordingly, the present disclosure provides, in some aspects, method for normalizing the concentration of target polynucleotides between at least two samples each comprising target polynucleotides, the method comprising, for each sample of the at least two samples:

(i) obtaining the sample, wherein the target polynucleotides of the sample comprise an adapter sequence;

(ii) producing a solution, the producing comprising combining (a) the sample, (b) a predetermined concentration of a guide polynucleotide comprising a targeting region that is complementary to the adapter sequence of the target polynucleotides of the sample; and (c) a predetermined concentration of a dCas or a dArgonaute comprising an affinity tag, wherein the predetermined concentration of the dCas or the dArgonaute is cognate to the guide polynucleotide;

(iii) contacting the solution with a solid phase comprising an affinity tag binding molecule that is capable of binding to the affinity tag;

(iv) separating the solution from the solid phase; and

(v) extracting the target polynucleotides from the solid phase to normalize the concentration of target polynucleotides between the two or more samples. In some embodiments, the targeting region is complementary to a static region of an adapter sequence. In some embodiments, the guide polynucleotide comprises a CRISPR-associated protein (Cas) guide RNA (gRNA) polynucleotide or an Argonaute guide polynucleotide. In some 5  
embodiments, the static region is a static region of a next-generation sequencing adapter sequence. In some embodiments, the targeting region is complementary to an adapter sequence of any one of SEQ ID NOs: 27-36. In some embodiments, the targeting region is complementary to an adapter sequence of any one of SEQ ID NOs: 27-28. In some embodiments, the targeting 10  
region is complementary to an adapter sequence of any one of SEQ ID NOs: 29-30. In some embodiments, the targeting region is complementary to an adapter sequence of any one of SEQ ID NOs: 31-32. In some embodiments, the targeting region is complementary to an adapter sequence any one of SEQ ID NOs: 33-34. In some embodiments, the targeting region is complementary to an adapter sequence of SEQ ID NO: 35. In some embodiments, the targeting 15  
region is complementary to an adapter sequence of SEQ ID NO: 36.

In some embodiments, the Cas gRNA polynucleotide targeting region is a homology region. In some embodiments, the Cas gRNA polynucleotide comprises a homology region of any one of SEQ ID NOs: 1-26. In some embodiments, the Cas gRNA polynucleotide comprises a homology region of SEQ ID NO: 1.

20 In some embodiments, the Argonaute guide polynucleotide is an siRNA, miRNA, piRNA, shRNA or siDNA.

In some embodiments, the guide polynucleotide is a DNA polynucleotide. In some embodiments, the guide polynucleotide is a RNA polynucleotide. In some embodiments, the guide polynucleotide comprises a modified nucleic acid. In some embodiments, the modified 25  
nucleic acid comprises 2'F RNA, 2'OMe RNA, and/or a phosphorothioate bonds (PS).

In some embodiments, the guide polynucleotide is a Cas gRNA polynucleotide and the Cas gRNA polynucleotide does not comprise a modified nucleic acid in a position that interacts with a Cas protein cognate to the gRNA. In some embodiments, the dCas is a catalytically- 30  
inactive Cpf1, C2c1, C2c3, C2c2, CasX, or CasY protein. In some embodiments, the dCas protein comprises an amino acid sequence of SEQ ID NO: 37.

In some embodiments, the dArgonaute is a catalytically-inactive LrAgo, PfAgo, TtAgo, AaAgo, AfAgo, MjAgo, MpAgo, NgAgo, RsAgo, CpAgo, IbAgo, KmAgo, or SeAgo.

In some embodiments, the affinity tag binding molecule comprises Ni<sup>2+</sup> and the affinity tag comprises a His Tag.

5 In some embodiments, the affinity tag binding molecule comprises biotin and the affinity tag comprises avidin. In some embodiments, the affinity tag binding molecule comprises an anti-myc antibody and the affinity tag comprises a myc tag. In some embodiments, the affinity tag and corresponding affinity tag binding molecule are selected from Table 1.

10 In some embodiments, the solid phase comprises magnetic beads. In some embodiments, separating the solution from the solid phase comprises immobilizing the solid phase and washing the solid phase. In some embodiments, extracting the target polynucleotides from the solid phase comprises combining the solid phase and protease. In some embodiments, extracting the target polynucleotides from the solid phase comprises combining the solid phase and proteinase K in a solution sufficient for proteinase K activity. In some embodiments, proteinase K digests the  
15 dCas or dArgonaute that is bound to the solid phase thereby extracting the target polynucleotides.

In some embodiments, steps (i)-(v) are performed in sequential order.

In some embodiments, normalizing comprises making the concentration of the of target polynucleotides between the at least two samples within 15% of one another after normalization.

20 In some aspects, the present disclosure provides a method for normalizing the concentration of target polynucleotides between at least two samples each comprising target polynucleotides, the method comprising, for each sample of the at least two samples:

(i) obtaining the sample, wherein the target polynucleotides of the sample comprise an adapter sequence;

25 (ii) producing a solution, the producing comprising combining (a) the sample, (b) a predetermined concentration of a guide polynucleotide comprising a targeting region that is complementary to the adapter sequence of the target polynucleotides of the sample, and (c) a predetermined concentration of a dCas or a dArgonaute comprising an affinity tag; and

30 (iii) contacting the solution with a pre-determined amount of a catalytically active Cas or Argonaute protein to normalize the concentration of target polynucleotides between the two or more samples.

In some embodiments, the dCas is a catalytically-inactive Cpf1, C2c1, C2c3, C2c2, CasX, Cas9, or CasY protein. In some embodiments, the dCas protein comprises an amino acid sequence having at least 95% identity to SEQ ID NO: 37. In some embodiments, the dCas protein comprises an amino acid sequence of SEQ ID NO: 37.

5 In some embodiments, the dArgonaute is a catalytically-inactive LrAgo, PfAgo, TtAgo, AaAgo, AfAgo, MjAgo, MpAgo, NgAgo, RsAgo, CpAgo, IbAgo, KmAgo, or SeAgo.

In some embodiments, step (ii) comprises binding of dCas or dArgonaute to the adapter sequences of at least some of the target polynucleotides. In some embodiments, the catalytically active Cas9 or Argonaute digests target polynucleotides that are not bound by the dCas9 or the  
10 dArgonaute. In some embodiments, normalizing comprises making the concentration of the of target polynucleotides between the at least two samples within 15% of one another after normalization. In some embodiments, normalizing comprises making the concentration of the of target polynucleotides between the at least two samples within 10% of one another after normalization. In some embodiments, the target polynucleotides comprise a first adapter  
15 sequence and a second adapter sequence, and wherein the guide polynucleotide targeting region is complementary to the first adapter sequence.

In some embodiments, the method further comprises, after step (i) and before step (ii):

(a) contacting the sample with a primer encoding a nucleic acid sequence that is complementary to a portion of the second adapter sequence, wherein the portion of the adapter  
20 sequence is located in the proximal portion of the adapter sequence; and

(b) contacting the sample with a DNA polymerase in conditions sufficient to promote primer extension.

In some embodiments, the portion of the adapter sequence comprises a nucleic acid sequence of any one of SEQ ID NOs: 28, 30, 32, and 34.

25 In some embodiments, the second adapter sequence is a 3' adapter sequence. In some embodiments, the second adapter sequence is a 5' adapter sequence.

In some aspects, the disclosure provides a guide polynucleotide comprising a targeting region that is complementary to an adapter sequence of any one of SEQ ID NOs: 28 and 30, 32, 34-36. In some embodiments, the guide polynucleotide comprises a CRISPR-associated protein  
30 (Cas) guide RNA (gRNA) polynucleotide or an Argonaute guide polynucleotide.

In some embodiments, the static region is a static region of a next-generation sequencing adapter sequence. In some embodiments, the targeting region is complementary to an adapter sequence of SEQ ID NO: 28. In some embodiments, the targeting region is complementary to an adapter sequence of SEQ ID NO: 30. In some embodiments, the targeting region is complementary to an adapter sequence of SEQ ID NO: 32. In some embodiments, the targeting region is complementary to an adapter sequence of SEQ ID NO: 34. In some embodiments, the targeting region is complementary to an adapter sequence of SEQ ID NO: 35. In some embodiments, the targeting region is complementary to an adapter sequence of any one of SEQ ID NOs: 36.

In some embodiments, the Cas gRNA polynucleotide targeting region is a homology region. In some embodiments, the Cas gRNA polynucleotide comprises a homology region of any one of SEQ ID NOs: 1-26. In some embodiments, the Cas gRNA polynucleotide comprises a homology region of SEQ ID NO: 1.

In some embodiments, the Argonaute guide polynucleotide is an siRNA, miRNA, piRNA, shRNA or siDNA.

In some embodiments, the guide polynucleotide is a DNA polynucleotide. In some embodiments, the guide polynucleotide is a RNA polynucleotide. In some embodiments, the guide polynucleotide comprises a modified nucleic acid. In some embodiments, the modified nucleic acid comprises 2'F RNA, 2'Ome RNA, and/or a phosphorothioate bonds (PS). In some embodiments, the guide polynucleotide is a Cas gRNA polynucleotide and the Cas gRNA polynucleotide does not comprise a modified nucleic acid in a position that interacts with a Cas protein cognate to the gRNA.

In some embodiments, the present disclosure provides herein a kit comprising (a) a guide polynucleotide described herein, and (b) a Cas protein or an Argonaute protein, or a polynucleotide sequence encoding a Cas protein or an Argonaute protein. In some embodiments, the Cas protein is a catalytically-inactive Cas protein (dCas). In some embodiments, the Cas protein is a Cas9 protein that is catalytically-inactive (dCas9). In some embodiments, the Cas9 protein comprises a D10A mutation and a H840A mutation. In some embodiments, the Cas protein is a catalytically-inactive Cpf1, C2c1, C2c3, C2c2, CasX, or CasY protein. In some embodiments, the dCas protein comprises an amino acid sequence having at least 95% identity to

SEQ ID NO: 37. In some embodiments, the dCas protein comprises an amino acid sequence of SEQ ID NO: 37.

In some embodiments, the Argonaute protein is catalytically-inactive (dArgonaute). In some embodiments, the Argonaute protein is a CbAgo protein that is catalytically-inactive. In some embodiments, the Argonaute protein is a LrAgo, PfAgo, TtAgo, AaAgo, AfAgo, MjAgo, MpAgo, NgAgo, RsAgo, CpAgo, IbAgo, KmAgo, or SeAgo protein that is catalytically-inactive. In some embodiments, the dArgonaute protein comprises an amino acid sequence having at least 95% identity to any one of SEQ ID NOs: 39-45. In some embodiments, the dArgonaute protein comprises an amino acid sequence of any one of SEQ ID NOs: 39-45.

In some embodiments, the kit further comprises a catalytically active Cas protein or a catalytically active Argonaute protein. In some embodiments, the catalytically active Cas protein comprises Cpf1, C2c1, C2c3, C2c2, CasX, Cas9, or CasY. In some embodiments, the catalytically active Argonaute protein comprises LrAgo, PfAgo, TtAgo, AaAgo, AfAgo, MjAgo, MpAgo, NgAgo, RsAgo, CpAgo, IbAgo, KmAgo, or SeAgo. In some embodiments, the guide polynucleotide is capable of binding to the Cas protein or the Argonaute protein to form a ribonucleoprotein complex and the ribonucleoprotein complex is capable of binding to an adapter sequence.

In some embodiments, the kit further comprises a primer that is complementary to a portion of an adapter sequence. In some embodiments, the portion of the adapter sequence is located in the proximal portion of the adapter sequence. In some embodiments, the portion of the adapter sequence is located in the proximal portion of the adapter sequence, the adapter sequence comprising a nucleic acid sequence of any one of SEQ ID NOs: 28, 30, 32, and 34. In some embodiments, the adapter sequence is a 3' adapter sequence. In some embodiments, the adapter sequence is a 5' adapter sequence. In some embodiments, the primer is not complementary to an adapter sequence that is complementary to the guide polynucleotide targeting region.

In some aspects, the disclosure provides a reaction mixture comprising: (i) a plurality of target polynucleotides, wherein the target polynucleotides comprise an adapter sequence of any one of SEQ ID NOs: 28, 30, 32, and 34-36; (ii) a predetermined concentration of a guide polynucleotide comprising a targeting region that is complementary to the adaptor sequence; and (iii) a predetermined concentration of a dCas protein or a dArgonaute protein.

In some embodiments, the predetermined concentration of the guide polynucleotide, or the dCas protein or the dArgonaute protein is lower than the concentration of target polynucleotide in the reaction mixture.

In some embodiments, the dArgonaute protein comprises LrAgo, PfAgo, TtAgo, AaAgo, AfAgo, MjAgo, MpAgo, NgAgo, RsAgo, CpAgo, IbAgo, KmAgo, or SeAgo, or a catalytically inactive variant thereof, and the guide polynucleotides are cognate to the dArgonaute protein. In some embodiments, the dCas protein comprises catalytically inactive Cpf1, C2c1, C2c3, C2c2, CasX, Cas9, or CasY, or a variant thereof, and the guide polynucleotides are cognate to the dCas protein. In some embodiments, the adapter sequence comprises a polynucleotide sequence of any one of SEQ ID NOs: 28, 30, 32, and 34-36. In some embodiments, the reaction mixture further comprises a catalytically active Cas protein or a catalytically active Argonaute protein.

In some embodiments, the disclosure provides a ribonucleoprotein (RNP) complex comprising: (i) a guide polynucleotide described herein; (ii) a Cas protein or an Argonaute protein that is cognate to the guide polynucleotide; and (iii) an adapter sequence, wherein the targeting region of the guide polynucleotide is complementary to the adapter sequence and wherein the guide polynucleotide is cognate to the Cas protein or the Argonaute protein. In some embodiments, the Argonaute protein comprises LrAgo, PfAgo, TtAgo, AaAgo, AfAgo, MjAgo, MpAgo, NgAgo, RsAgo, CpAgo, IbAgo, KmAgo, or SeAgo, or a catalytically inactive variant thereof. In some embodiments, the Cas protein comprises Cpf1, C2c1, C2c3, C2c2, CasX, Cas9, or CasY, or a catalytically inactive variant thereof. In some embodiments, the adapter sequence comprises a polynucleotide sequence of any one of SEQ ID NOs: 28, 30, 32, and 34-36.

## BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows a representative schematic of the polynucleotide capture process by a CAS-gRNA RNP bound to a metal bead.

FIG. 2 shows normalization of two different samples each comprising i5-adapter containing polynucleotides. The two different samples of i5-containing DNA fragments were subjected to normalization with the same mass (125 femtomoles, fmols) of dCas9. Following elution after the proteinase digestion, both samples contained an equivalent number of

molecules, demonstrating the concept of normalization. The ratio of molecules between the samples changed from 2:1 to 1:1, as intended. If these samples were pooled using equivalent volumes and sequences by NGS, both samples would be expected to yield an equivalent number of clusters.

5           FIG. 3 shows an ability to tune the amount of polynucleotides extracted from a sample by titrating the amount of ribonucleotide protein that binds to the adapters of the polynucleotides (e.g., RNP) being added to the sample. RNP (dCas9-gRNA complex) titration and subsequent bead binding and elution was performed with samples that contained a constant amount of DNA (160 fmols). Increasing the amount of the RNP resulted in linearly increasing amounts of bound  
10 (and eventually recovered) DNA. There was a linear relationship in a range from 125 femtomoles to 1000 fmols of RNP at the 160 fmol DNA input level. These data demonstrate that the amount of DNA library extracted can be tuned by adjusting the amount of RNP added to the sample.

          FIG. 4 shows specific and stoichiometric targeting and retention of target DNA  
15 molecules driven by interaction with dCas9 RNP and subsequent binding to HisPur beads via the 6His tagged dCas9. A known quantity of *i5*-sequence containing DNA fragments was bound to HisPur beads after being contacted with RNP (RNP+) or no RNP (RNP-). The majority of the library remained bound to the beads only if exposed to the RNP and eluted after proteinase K digestion (compare RNP+ Flow through to RNP+ elution). In the absence of RNP, all of the  
20 DNA was found in the flow through and none bound to the beads (compare RNP- Flow Through to RNP- Elution).

          FIG. 5 shows normalization of four different samples of PCR amplified ILLUMINA libraries using dCas9-gRNA complex. Four different samples of PCR amplified ILLUMINA libraries (input), were subjected to normalization with the same mass of dCas9. Following  
25 elution after the proteinase digestion (elution), all samples contained a significantly more uniform number of molecules. The highest concentration sample was 25% more concentrated than the lowest, post normalization (4.68 nM Elution vs 5.89 nM Elution), compared to 150% pre-normalization (4 nM Input vs 10 nM Input). In addition, the majority of the unbound library was retained in the flowthrough for all the libraries.

30           FIG. 6 shows a schematic demonstrating an exemplary embodiment of the disclosed normalization methods. In this exemplary embodiment, two samples are provided that include

end-labelled target nucleic acids of different concentrations (Library A having a concentration of 9 fmols and Library B having a concentration of 18 fmols). In this exemplary embodiment, a reaction mixture is produced by adding a predetermined concentration of 3 fmols of a guide polynucleotide and a catalytically-inactive Cas9 protein (which may be pre-assembled) to Library A and Library B and allowing the protein to bind to the adapter sequence at one end of the end-labelled target nucleic acids in the sample, protecting the end of the target nucleic acid from nuclease digestion. An exonuclease is added to the reaction mixture and digests the unprotected nucleic acids. The resulting Library A and Library B each contain 3 fmols of target nucleic acid. In other embodiments, the exonuclease in this schematic may be replaced with a nucleic acid guided nuclease (e.g., Cas9 nuclease), a nucleic acid guided nickase (e.g., Cas9 nickase), a restriction enzyme nuclease or nickase, or a transcription activator-like effector nuclease (TALEN) or another similar nuclease or nickase.

FIGs. 7A-D show a schematic of library molecule denaturation and partial sequence extension to produce a double stranded PAM site. In FIG. 7A, a library molecule is ligated with two Y-shaped adapters at the 3' and 5' ends. In FIG. 7B, the library molecule is denatured and bound to a partial primer at the 3' end. In FIG. 7C, the bound primer is extended with a polymerase, creating a complementary strand (dashed line), which does not span the entire 3' adapter, but does comprise a double-stranded, full-length 5' adapter, which contains a PAM site for dCas9 targeting. FIG. 7D illustrates the binding of a dCas9 molecule to the double-stranded PAM site.

FIG. 8 is a bar graph showing femtomoles of captured DNA (fmols out) after normalization and amplification for different amounts of starting DNA (fmols in) and two concentrations of ribonucleoprotein (RNP fmols).

FIG. 9 shows fmols of DNA captured (fmols out) by varying concentrations of ribonucleoprotein (RNP fmols). The concentration of input DNA is shown a 3,400 fmols.

FIG. 10 shows fmols of DNA captured (fmols out) by varying concentrations of ribonucleoprotein (RNP fmols) and either single guide (bottom line) or duplex guide (top line).

## DETAILED DESCRIPTION

**Guide polynucleotides**

In some aspects, this disclosure describes a guide polynucleotide comprising a targeting  
5 region that is complementary to an adapter sequence.

A “polynucleotide” refers to polymers of nucleotides (e.g., deoxyribonucleotides,  
ribonucleotides and modified nucleotides). The polymer may be in either single- or double-  
stranded form. Unless specifically limited, the term encompasses polynucleotides containing  
known nucleotide analogs or modified backbone residues or linkages, which are synthetic,  
10 naturally occurring, and non-naturally occurring, which have similar binding properties as the  
reference nucleic acid, and which are metabolized in a manner similar to the naturally occurring  
nucleotides. Examples of such analogs include, without limitation, phosphorothioates,  
phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2-O-methyl  
ribonucleotides, peptide-nucleic acids (PNAs).

A “guide polynucleotide” refers a polynucleotide that (1) comprises a targeting region  
which is complementary to a target (e.g., an adapter sequence) and (2) that is capable of  
facilitating binding of a ribonucleoprotein complex comprising the guide polynucleotide to the  
target. In some embodiments, the guide polynucleotide is a CRISPR associated protein (Cas)  
guide RNA (gRNA) polynucleotide or an Argonaute guide polynucleotide. In some  
20 embodiments, a Cas gRNA polynucleotide refers to a two polynucleotide system comprising a  
tracrRNA and a crRNA e.g., as described in Karvelis T et al., RNA Biol. 2013 May;10(5):841-  
51. PMID: 23535272. The tracrRNA comprises a sequence encoding a stem loop structure that  
associates with the Cas protein (e.g., dCas9). The crRNA comprises a homology region that is  
complementary to the target (e.g., an adaptor sequence) and a region that is complementary to  
25 the tracrRNA. The crRNA and the tracrRNA may form a complex with the Cas protein, which  
in turn binds to a target DNA or RNA polynucleotide (depending on the Cas type). In some  
embodiments, a Cas gRNA polynucleotide refers to a single guide RNA (sgRNA) polynucleotide  
e.g., as described in Jinek M et al., Science. 2012 Aug 17;337(6096):816-21. doi:  
10.1126/science.1225829. A sgRNA comprises a sequence encoding a stem loop structure that  
30 associates with the Cas protein and a homology region. Many Cas proteins bind to  
polynucleotides (e.g., double stranded DNA) in a position that is adjacent to a protospacer

adjacent motif (PAM) site. In some embodiments, the homology region is complementary to a portion of the adapter sequence that is adjacent to a PAM site (e.g., sufficiently adjacent such that Cas binding to the adapter sequence can take place). Methods of making and using gRNAs are known in the art, e.g., as described in Mohr SE, et al., FEBS J. 2016 Sep;283(17):3232-8.

5 PMCID: PMC5014588.

An Argonaute guide polynucleotide refers to a polynucleotide encoding a small interfering RNA (siRNA), microRNA (miRNA), P element-induced wimpy testis (PIWI)-interacting RNA (piRNA) and small interfering DNA (siDNA) e.g., as described in Wu J et al., J Adv Res. 2020 Apr 29;24:317-324, PMID: 32455006. In some embodiments, the Argonaute guide polynucleotide comprises a targeting region that is complementary to an adapter sequence as described herein. In some embodiments, a guide polynucleotide is single stranded (e.g., an RNA Cas gRNA polynucleotide). In some embodiments, a guide polynucleotide is double stranded (e.g., double stranded DNA encoding a Cas gRNA polynucleotide, or dsRNA used an Argonaute guide polynucleotide).

15 In some embodiments, a guide polynucleotide is an RNA molecule. In some embodiments, a guide polynucleotide is a DNA molecule. In some embodiments, the guide polynucleotide comprises modified nucleic acids (e.g., 2'F RNA, 2'OMe RNA, and/or a phosphorothioate bonds (PS)). Nucleic acids in a guide polynucleotide (e.g., an RNA guide polynucleotide like a Cas guide RNA) may be modified to increase the stability of the guide polynucleotide (e.g., reduce nuclease digestion). A Cas gRNA (e.g., a Cas9 gRNA) may comprise modified nucleic acids (e.g., 2'OMe) in regions of the gRNA that do not interact with the Cas protein (e.g., Cas9), and still maintain an ability to direct the Cas protein to the target, as described in Yin H et al., Nature Biotechnology 2017 Dec; 35(12): 1179-1187. In some embodiments, the Cas guide RNA polynucleotide comprises the modifications of the e-sgRNA as described in Yin et al. 2017. In some embodiments, a Cas9 sgRNA does not comprise a modified nucleic acid (e.g., a 2'OH modification) at one or more positions 22-27, 43-45, 47, 49, 51, 58, 59, 62, 63-65, 68-69, or 82 counted from the 5' end of the sgRNA (i.e., with reference to an sgRNA of

(GGGCGAGGAGCUGUUCACCGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGG  
30 CUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUU, SEQ ID  
NO: 46)).

In some embodiments, the sgRNA comprises 2'-O-methyl RNA bases. In 2'-O-methyl RNA bases, the 2'-hydroxyl (-OH) group of the ribose sugar in the RNA molecule is replaced by a methyl group (-CH<sub>3</sub>). This modification affects the ribose sugar, and the nitrogenous bases (adenine, guanine, cytosine, uracil) themselves are not typically modified in this context. The 2'-O-methyl RNA bases are 2'-O-methyladenosine, 2'-O-methylguanosine, 2'-O-methylcytidine, and 2'-O-methyluridine. In some embodiments, the sgRNA comprises the sequence:

5'- mA\*mG\*mA\* rUrCrG rGrArA rGrArG rCrGrU rCrGrU rGrUrG rUrUrU rUrArG  
rArGrCrUrArG rArArA rUrArG rCrArA rGrUrU rArArA rArUrA rArGrG rCrUrA rGrUrC  
rCrGrU rUrArU rCrArA rCrUrU rGrArA rArArA rGrUrG rGrCrA rCrCrG rArGrU rCrGrG  
10 rUrGrC mU\*mU\*mU\*rU-3' (SEQ ID NO: 48)

where mA and mG are 2'-O-methyladenosine, 2'-O-methylguanosine RNA bases.

A "targeting region" of a guide polynucleotide refers to a region of the guide polynucleotide that is complementary to a target polynucleotide sequence (e.g., an adapter sequence). In some embodiments, the guide polynucleotide is a Cas gRNA polynucleotide comprising a homology region. A Cas gRNA polynucleotide homology region may comprise a series of contiguous amino acids that are complementary to a target polynucleotide sequence (e.g., an adapter sequence). In some embodiments, the homology region is 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more nucleotides in length. In some embodiments, the homology region is 18-26 nucleotides in length. In some embodiments, the homology region is 17-24 nucleotides in length. In some embodiments, the homology region is 18-22 nucleotide in length. In some embodiments, the homology region is 20 nucleotides in length.

In some embodiments, the guide polynucleotide is an Argonaute guide polynucleotide. Argonaute guide polynucleotides may be RNA interference (RNAi) polynucleotides. In some embodiments, the Argonaute guide polynucleotide comprises a targeting region of a small interfering RNA (siRNA), microRNA (miRNA), P element-induced wimpy testis (PIWI)-interacting RNA (piRNA) and small interfering DNA (siDNA) e.g., as described in Wu J et al., J Adv Res. 2020 Apr 29;24:317-324, PMID: 32455006.

In some embodiments, the targeting region is complementary to an adapter sequence as described herein. In some embodiments, the targeting region is complementary to a static region (e.g., that generally does not change between adapters) of the adapter sequence. In some embodiments, the targeting region is complementary to a sequence comprising a region that was

added to the sequence for the purpose of being complementary to a guide polynucleotide targeting region. For example, an additional sequence can be added to an adaptor sequence for the purpose of being complementary to a guide polynucleotide targeting region (e.g., for use in library normalization). Such an additional sequence may be a sequence that is specifically bound  
5 by a Cas protein-gRNA complex or an Argonaute guide polynucleotide complex compared to other sequences in a polynucleotide sequencing library.

In some embodiments, the targeting region is complementary to a next-generation sequencing adapter sequence. In some embodiments, the static region is a static region of a next-generation sequencing adapter sequence. In some embodiments, the targeting region is  
10 complementary to a p5 adapter or a p7 adapter. In some embodiments, the p5 adapter and the p7 adapter are ILLUMINA sequencing adapters. In some embodiments, the targeting region is complementary to an i5 polynucleotide sequence of SEQ ID NO: 27 or SEQ ID NO: 28. In some embodiments, the targeting region is complementary to an i7 adapter sequence of SEQ ID NO: 29 or SEQ ID NO: 30. In some embodiments, the targeting region is complementary to a  
15 NEXTERA Read 1 Adapter (e.g., SEQ ID NO: 31 or SEQ ID NO: 32) or a NEXTERA Read 2 Adapter (e.g., SEQ ID NO: 33 or SEQ ID NO: 34). In some embodiments, the targeting region is complementary to an ION TORRENT A Adapter or an ION TORRENT P1 Adapter. In some embodiments, the targeting region is complementary to an ION TORRENT A Adapter of SEQ ID NO: 35. In some embodiments, the targeting region is complementary to a ION TORRENT  
20 P1 Adapter of SEQ ID NO: 36. In some embodiments, the gRNA polynucleotide comprises a homology region of any one of SEQ ID NOs: 1-26. In some embodiments, the gRNA polynucleotide comprises a homology region of SEQ ID NO: 1.

The term “complementary” as used herein refers to the degree of expected Watson-crick base pairing between a first polynucleotide (e.g., a guide polynucleotide targeting region) and a  
25 second polynucleotide (e.g., an adaptor sequence). Complementary nucleotides are, generally, A and T (or A and U), and G and C. In some embodiments, complementary refers to 100% complementarity between two polynucleotides (e.g., a guide polynucleotide targeting region and an adaptor sequence). In some embodiments, complementary refers to 70%, 80%, 90%, 95%, or 99% complementarity between two polynucleotides. For example, a targeting region may be  
30 complementary to an adaptor sequence when one, two, or three nucleic acids in the targeting region are not complementary to the adaptor sequence. In some embodiments, complementary

refers to a sufficiently degree of Watson-crick base pairing between a guide polynucleotide targeting region (of a RNP (e.g., dCas9 bound to a guide RNA) and an adaptor sequence for the RNP to bind to the adaptor sequence. In some embodiments, complementary refers to a sufficient degree of Watson Crick base pairing between an Argonaute guide polynucleotide (e.g., a miRNA, siRNA, pwrRNA, shRNA, or siDNA) and a target sequence (e.g., an adapter sequence) to direct binding of an Argonaute protein comprising the Argonaute guide polynucleotide to bind to the target.

An “adapter sequence” refers to a polynucleotide added (e.g., ligated) to an end (e.g., 3’ and/or 5’) of a target polynucleotide for use in sequencing of the polynucleotide. In some embodiments, the adapter sequence refers to a sequence of nucleic acids encoding the adapter. In some embodiments, the adapter sequence may be used for sequencing using a particular sequencing platform. For example, ILLUMINA i5/p5 and i7/p7 adapter sequences may be used for sequencing using the ILLUMINA sequencing platform. In some embodiments, the adapter sequence comprises a static region (e.g., that generally does not change between adapters) and a dynamic region (e.g., that may change between adapters). In some embodiments, the adapter sequence comprises a distal region (generally static), an index region (generally dynamic), and a proximal region (generally static). For example, an ILLUMINA adapter sequence may comprise a p5 region which is static, an index region which is dynamic, and an i5 region which is static. In some embodiments, the guide polynucleotide targeting region is complementary to a static region of the adapter sequence. In some embodiments, the guide polynucleotide targeting region is complementary to the dynamic region of the adapter sequence. In some embodiments, the guide polynucleotide comprises a targeting region that is complementary to a static region of an adapter sequence (e.g., a distal region or a proximal region of an adapter sequence). In some embodiments, the guide polynucleotide comprises a targeting region that is complementary to a dynamic region of an adapter sequence (e.g., an index). In some embodiments, at least some of the target polynucleotides comprise the same static adapter sequences (e.g., the same distal and/or proximal region). In some embodiments, the majority of the target polynucleotide comprise the same static adapter sequence. In some embodiments, all of the target polynucleotides comprise the same static adapter sequence.

In some embodiments, the adapter sequence is modified to comprise a Cas protein protospacer-adjacent motif (PAM) or a reverse complement of a PAM. A “protospacer-adjacent

motif (PAM)” is a nucleotide motif (often 3 consecutive nucleotides in the polynucleotide) that is required for Cas protein binding to the target polynucleotide. The canonical Cas9 PAM sequence is 5'-NGG, wherein N is any nucleotide A, G, C, or T. In some embodiments, the adapter sequence is modified to comprise a PAM or reverse complement of a PAM to facilitate binding of the Cas-gRNA complex to the target polynucleotide. For example, a PAM sequence may be added to the 3' end or 5' end of the adapter sequence. In some embodiments, a distal sequence can be modified to comprise a protospacer-adjacent motif (PAM) or a reverse complement thereof.

In some embodiments, the adaptor sequence is an adapter sequence used IN ILLUMINA, ION TORRENT, PACBIO, ELEMENT, ULTIMA, OMNIOME, SINGULAR, or MGI sequencing. In some embodiments, the adapter sequence is an adapter sequence from any one of the following kits: a WATCHMAKER DNA Library Prep Kit with Fragmentation or WATCHMAKER RNA Library Prep Kit with Polaris Depletion ILLUMINA TruSeq PCR-Free Library Preparation Kit, TruSeq Nano DNA Library Prep Kit, NEXTERA DNA Library Prep Kit, NEXTERA DNA XT Library Prep Kit, NEXTERA Rapid Capture Exome Kit, NEXTERA Rapid Capture Expanded Exome Kit, AmpliSeq for ILLUMINA Library Prep Kit, ILLUMINA RNA Prep with Enrichment Kit, ILLUMINA Stranded mRNA Prep Kit, TruSeq RNA Library Prep Kit, TruSeq Stranded Total RNA Kit, TruSeq Stranded mRNA Kit, TruSeq Small RNA Kit. In some embodiments, the adaptor sequence is an adapter sequence used in any one of the following kits: NEBNEXT Fast DNA for ION TORRENT, NEBNEXT Fast DNA Fragmentation & Library Prep Set for ION TORRENT, THERMOFISHER Precision ID Library Kit, THERMOFISHER Ion Xpress Plus Fragment Library Kit, THERMOFISHER Ion Xpress Barcode Adapters 1-96 Kit, and/or THERMOFISHER Ion AmpliSeq Transcriptome Human Gene Expression Kit. In some embodiments, the adaptor sequence is an adapter sequence used in any one of the following kits: a PACBIO SMRTbell Template Prep Kit 1.0, SMRTbell Barcoded Adapter Complete Prep Kit-96, or Barcoded Adapter Kit. In some embodiments, the targeting region may be complementary an adaptor sequence used in QIAGEN QIAseq Stranded RNA Library Kit or QIAseq UPX 3' Transcriptome Kit, a Perkin Elmer NEXTFLEX Rapid Directional RNA-Seq Kit or NEXTFLEX Small RNA-Seq Kit, or a Takara Bio SMART-Seq mRNA Kit or SMART-Seq mRNA LP Kit.

In some embodiments, the adapter sequence comprises a shared Y-adapter sequence. In certain embodiments, the shared Y-adapter sequence is a 13 bp sequence. In other embodiments, the shared Y-adapter sequence is a 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 bp sequence.

5 In some embodiments, the adapter sequence may comprise a sequence of any one of SEQ ID NOs: 27-36.

### Kits

10 In some aspects, this disclosure describes a kit comprising: (a) a guide polynucleotide described herein; and (b) a Cas protein or an Argonaute protein, or a polynucleotide sequence encoding a Cas protein or an Argonaute protein (e.g., as described herein).

In some embodiments, the kit comprises a Cas protein. A “Cas protein” refers to a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated protein. In some embodiments, the Cas protein has nuclease activity. In some embodiments, the Cas protein  
15 does not have nuclease activity (dCas). In some embodiments, the Cas protein has nicking activity (nickase).

In some embodiments, the Cas protein is a Cas9 protein. “Cas9” refers to a Cas9 protein or a fragment thereof present in any bacterial species that encodes a Type II CRISPR/Cas9 system. See, for example, Makarova et al., *Nature Reviews, Microbiology*, 9: 467-477 (2011),  
20 including supplemental information. Cas9 homologs are found in a wide variety of eubacteria, including, but not limited to bacteria of the following taxonomic groups: Actinobacteria, Aquificae, Bacteroidetes-Chlorobi, Chlamydiae-Verrucomicrobia, Chlorflexi, Cyanobacteria, Firmicutes, Proteobacteria, Spirochaetes, and Thermotogae. An exemplary Cas9 protein is the *Streptococcus pyogenes* Cas9 protein. Additional Cas9 proteins and homologs thereof are  
25 described in, e.g., Chylinski et al., 2013, *RNA Biol.* 10(5): 726–37; Hou et al., 2013, *Proc. Natl. Acad. Sci. USA* 110(39):15644-49; Sampson et al., 2013, 497(7448):254-57; and Jinek et al., 2012, *Science* 337(6096):816-21. Full-length Cas9 is an RNA-mediated endonuclease comprising a recognition domain and two nuclease domains (HNH and RuvC, respectively). In the amino acid sequence, HNH is linearly continuous, whereas RuvC is separated into three  
30 regions, one left of the recognition domain, and the other two right of the recognition domain flanking the HNH domain. Cas9 from *Streptococcus pyogenes* is targeted to a genomic site in a

cell by interacting with a guide RNA that hybridizes to a 20-nucleotide DNA sequence that immediately precedes an NGG motif recognized by Cas9.

In some embodiments, the Cas9 is a catalytically-inactive or nuclease-dead Cas9 (dCas9) that has been modified to inactivate Cas9 nuclease activity. Modifications include, but are not limited to, altering one or more amino acids to inactivate the nuclease activity or the nuclease domain. For example, and not to be limiting, D10A, H840A, and/or R1335K mutations can be made in Cas9 from *Streptococcus pyogenes* to inactivate Cas9 nuclease activity. In some embodiments, the Cas9 protein comprises a D10A mutation and a H840A mutation. Other modifications include removing all or a portion of the nuclease domain of Cas9, such that the sequences exhibiting nuclease activity are absent from Cas9. Accordingly, a catalytically-inactive Cas9 may include polypeptide sequences modified to inactivate nuclease activity or removal of a polypeptide sequence or sequences to inactivate nuclease activity. The catalytically-inactive Cas9 retains the ability to bind to DNA even though the nuclease activity has been inactivated. Accordingly, dCas9 includes the polypeptide sequence or sequences required for DNA binding but includes modified nuclease sequences or lacks nuclease sequences responsible for nuclease activity.

In some embodiments, the catalytically-inactive Cas9 protein is a full-length Cas9 sequence from *S. pyogenes* lacking the polypeptide sequence of the RuvC nuclease domain and/or the HNH nuclease domain and retaining the DNA binding function. In other embodiments, the catalytically-inactive Cas9 protein sequences have at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98% or 99% identity to Cas9 polypeptide sequences lacking the RuvC nuclease domain and/or the HNH nuclease domain and retains DNA binding function. In some embodiments, the Cas protein is Alt-R™ *S.p.* dCas9 Protein V3 (IDT).

In some embodiments, the Cas protein is a Cpf1 (Cas12a), C2c1, C2c3, C2c2, CasX, or CasY protein. In some embodiments, the Cas protein has been modified to inactivate Cas9 nuclease activity. Modifications include, but are not limited to, altering one or more amino acids to inactivate the nuclease activity or the nuclease domain of the Cas protein. For example, D908, D832, E993, R1226, and/or D1235 mutations can be made in Cpf1 from *Acidaminococcus sp.* BV3L6, *Lachnospiraceae* ND2006, or *Francisella tularensis* subsp. *novicida* U112, to inactivate Cpf1 nuclease activity. Other modifications include removing all or a portion of the nuclease domain, such that the sequences exhibiting nuclease activity are absent. Accordingly, a

catalytically-inactive Cas protein may include polypeptide sequences modified to inactivate nuclease activity or removal of a polypeptide sequence or sequences to inactivate nuclease activity. The catalytically-inactive Cas protein retains the ability to bind to DNA even though the nuclease activity has been inactivated. Accordingly, catalytically-inactive Cas protein includes the polypeptide sequence or sequences required for DNA binding but includes modified nuclease sequences or lacks nuclease sequences responsible for nuclease activity. In some embodiments, the catalytically-inactive Cas protein is a full-length Cas sequence lacking the polypeptide sequence of the nuclease domain and retaining the DNA binding function. In other embodiments, the catalytically-inactive Cas protein sequences have at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98% or 99% identity to Cas polypeptide sequences lacking the nuclease domain and retains DNA binding function.

In some embodiments, the Cas protein comprises an amino acid sequence having at least 95% identity to SEQ ID NO: 37 or SEQ ID NO: 38. In some embodiments, the Cas protein comprises an amino acid sequence of SEQ ID NO: 37 or SEQ ID NO: 38.

In some embodiments, the dCas protein comprises an amino acid sequence having at least 95% identity to SEQ ID NO: 37. In some embodiments, the dCas protein comprises an amino acid sequence of SEQ ID NO: 37.

In some embodiments, the Cas protein comprise a nuclease localization sequence.

In some embodiments, the kit comprising an Argonaute protein. An "Argonaute protein" refers to a protein that binds small noncoding nucleic acids (e.g., an Argonaute guide polynucleotide) and utilizes them for the guided cleavage of complementary nucleic acid targets or indirect gene silencing by recruiting additional factors. Catalytically-active Argonaute proteins are capable of nucleic acid guided binding to a complementary nucleic acid target (e.g., DNA) and cleavage of the nucleic acid target. See, e.g., Kaya et al., 2016, PNAS 113(5):4057-62. The prokaryotic Argonaute (AGO) gene family encodes several domains: N-terminal (N), PAZ, MID, and a C-terminal PIWI domain. The MID and PAZ domains are responsible for binding of the 5'-end and 3'-end, respectively, of a guide polynucleotide. In contrast to eukaryotic Agos that use exclusively small RNA guides, the majority of characterized prokaryotic Agos bind single-stranded DNA (ssDNA) guides. The Ago PIWI domain comprises nuclease activity. In some embodiments, the Argonaute protein is catalytically-inactive. In some embodiments, the Argonaute protein is a CbAgo protein that is catalytically-inactive. In some

embodiments, the Argonaute protein is a LrAgo, PfAgo, TtAgo, AaAgo, AfAgo, MjAgo, MpAgo, NgAgo, RsAgo, CpAgo, IbAgo, KmAgo, or SeAgo protein that is catalytically-inactive. In certain embodiments, the Argonaute protein has been modified to inactivate nuclease activity or inherently has reduced nuclease activity (see, e.g., Kaya et al., 2016, PNAS 113(5):4057-62).  
5 Modifications may include, but are not limited to, altering one or more amino acids to inactivate the nuclease activity or the nuclease domain of the Argonaute protein. For example, in some embodiments, the catalytically-inactive Argonaute protein is a CbAgo protein that comprises a D541A mutation and a D611A mutation. See, e.g., Hegge et al., 2019, Nuc. Acids Res. 47(11):5809-21. Alternatively, all or a portion of the nuclease domain may be removed, such  
10 that the sequences exhibiting nuclease activity are absent. Accordingly, a catalytically-inactive Argonaute protein may include polypeptide sequences modified to inactivate nuclease activity or removal of a polypeptide sequence or sequences to inactivate nuclease activity. The catalytically-inactive Argonaute protein retains the ability to bind to DNA even though the nuclease activity has been inactivated. Accordingly, catalytically-inactive Argonaute protein  
15 includes the polypeptide sequence or sequences required for DNA binding but includes modified nuclease sequences or lacks nuclease sequences responsible for nuclease activity. In some embodiments, the catalytically-inactive Argonaute protein is a full-length Argonaute sequence lacking the polypeptide sequence of the nuclease domain and retaining the DNA binding function. In other embodiments, the catalytically-inactive Argonaute protein sequences have at  
20 least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98% or 99% identity to the Argonaute polypeptide sequences lacking the nuclease domain and retains DNA binding function.

In some embodiments, the dArgonaute protein comprises an amino acid sequence having at least 95% identity to any one of SEQ ID NOs: 40-45. In some embodiments, the dArgonaute protein comprises an amino acid sequence of any one of SEQ ID NOs: 40-45.

25 In some embodiments, the kit comprises a catalytically inactive Cas protein (dCas protein) or a catalytically inactive Argonaute protein (e.g., dArgonaute) as described herein. In some embodiments, the kit comprises a catalytically inactive Cas protein (dCas protein) and a catalytically active Cas protein as described herein. In some embodiments, the kit comprises a catalytically a catalytically inactive Argonaute protein (e.g., dArgonaute) a catalytically active  
30 Argonaute protein as described herein. In some embodiments, the kit comprises a dCas and a cognate dCas guide RNA polynucleotide. In some embodiments, the kit comprises a dCas9 and

a cognate dCas9 guide RNA. In some embodiments, the kit comprises a dArgonaute and a cognate dArgonaute guide polynucleotide.

In some embodiments, the kit comprises a Cas protein comprising an affinity tag or a Argonaute protein comprising an affinity tag. In some embodiments, the affinity tag is Albumin-binding protein (ABP), Alkaline Phosphatase (AP), AU1 epitope, AU5 epitope, Bacteriophage T7 epitope (T7-tag), Bacteriophage V5 epitope (V5-tag), Biotin-carboxy carrier protein (BCCP), Bluetongue virus tag (B-tag), Calmodulin binding peptide (CBP), Chloramphenicol Acetyl Transferase (CAT), Cellulose binding domain (CBP), Chitin binding domain (CBD), Choline-binding domain (CBD), Dihydrofolate reductase (DHFR), E2 epitope, FLAG epitope, Galactose-binding protein (GBP), Green fluorescent protein (GFP), Glu-Glu (EE-tag), Glu-Glu (EE-tag), Human influenza hemagglutinin (HA), HaloTag®, Histidine affinity tag (HAT), Horseradish Peroxidase (HRP), HSV epitope, Ketosteroid isomerase (KSI), KT3 epitope, LacZ, Luciferase, Maltose-binding protein (MBP), Myc epitope, Nus , PDZ domain, PDZ ligand, Polyarginine (Arg-tag), Polyaspartate (Asp-tag), Polyhistidine, (His-tag), Polyphenylalanine (Phe-tag), Profinity eXact, Protein C, S1-tag, S-tag, Streptavidin-binding peptide (SBP), Staphylococcal protein A (Protein A), Staphylococcal protein G (Protein G), Strep-tag, Streptavidin, Small Ubiquitin-like Modifier (SUMO), Tandem Affinity Purification (TAP), T7 epitope, Thioredoxin (Trx), TrpE, Ubiquitin, Universal, and VSV-G as described in Kimple ME et al., Curr Protoc Protein Sci. 2013 Sep 24;73:9.9.1-9.9.23. doi: 10.1002/0471140864.ps0909s73.

In some embodiments, the affinity tag and corresponding affinity tag binding molecule, in a kit or for use in the method, are selected from Table 1.

Table 1: Affinity tag and affinity tag binding molecule.

<u>Affinity tag</u>	<u>Fusion MW</u>	<u>Affinity tag binding molecule</u>
GST	27 kDa	Glutathione beads
Protein A	49 kDa	IgG-conjugated beads
Protein G	65 kDa	IgG-conjugated beads
Streptavidin	16.5 kDa	Biotin beads
Biotin	post-translational	Streptavidin beads

Cell Surface Vimentin (CSV)	57 kDa	CSV monoclonal Ab beads
Human PSMA	82.5 kDa	PMSA monoclonal Ab beads
IgG Fc heavy chain	~25-50 kDa	Protein A/G beads
Chitin binding domain	27 kDa	Chitin beads
Maltose binding protein (MBP)	42.5 kDa	Amylose or anti-MBP beads
Protein L 'B' repeats	36 kDa	IgG-conjugated beads

In some embodiments, the kit comprises a guide polynucleotide that is capable of binding to the Cas protein (e.g., cognate the guide polynucleotide is cognate to the Cas protein) or the Argonaute protein of the kit to form a ribonucleoprotein complex and the ribonucleoprotein complex is capable of binding to an adapter sequence.

The term “cognate” as used in the context of the guide polynucleotide and the Cas protein or Argonaute protein refers to a guide polynucleotide that is compatible with the Cas protein or the Argonaute protein (e.g., the guide polynucleotide is capable of directing the Cas protein or the Argonaute protein to a target polynucleotide). For example, a Cas9 sgRNA is cognate to Cas9 and dCas9.

The terms “binds” or “specifically binds,” and like terms, refer to a molecule (e.g., a Cas-gRNA complex or a Argonaute-guide polynucleotide complex) that binds to a target nucleic acid with at least 2-fold greater affinity than non-target nucleic acids, e.g., at least any of 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 20-fold, 25-fold, 50-fold, 100-fold, 1,000-fold, 10,000-fold, or greater affinity for the target nucleic acid compared to an unrelated nucleic acid when assayed under the same binding affinity assay conditions. In some embodiments, the term “binds” or “specifically binds,” as used herein, can be exhibited, for example, by a molecule (e.g., a Cas-gRNA complex or a Argonaute-guide polynucleotide complex) having an equilibrium dissociation constant  $K_D$  for the target nucleic acid of, e.g.,  $10^{-2}$  M or smaller, e.g.,  $10^{-3}$  M,  $10^{-4}$  M,  $10^{-5}$  M,  $10^{-6}$  M,  $10^{-7}$  M,  $10^{-8}$  M,  $10^{-9}$  M,  $10^{-10}$  M,  $10^{-11}$  M, or  $10^{-12}$  M. In some embodiments, an antibody has a  $K_D$  of less than 10 nM or less than 100 nM.

In some embodiments, the kit further comprises a primer that is complementary to a portion of an adapter sequence (e.g., as described herein). In some embodiments, the primer is

complementary to a proximal portion of an adapter sequence. This primer may be used for creating double stranded target nucleotides comprising a double stranded proto-spacer adjacent motif (PAM) site that a dCas protein (e.g., as described herein) can use for binding. In some embodiments, the primer binds to a proximal portion of the adapter sequences such that the new polynucleotide (which is complementary to the target polynucleotide) does not comprise the distal portion of the adapter sequence and thus cannot be sequenced using next-generation sequencing (e.g., ILLUMINA Sequencing). Without being bound to theory, this method may be advantageous compared to PCR based methods because errors (e.g., mutations) introduced during elongation of the primer are not sequenced, but double stranded target polynucleotides may still be produced and bound by dCas proteins.

In some embodiments, the primer is a DNA primer. In some embodiments, the DNA primer comprises nucleic acid modifications (e.g., as described herein). In some embodiments, the portion of the adapter sequence is located in the proximal portion of the adapter sequence. In some embodiments, the proximal portion of the adapter sequence comprises a nucleic acid sequence of any one of SEQ ID NOs: 28, 30, 32, and 34. In some embodiments, the primer is complementary to a 3' adapter sequence. In some embodiments, the primer is complementary to a 5' adapter sequence. In some embodiments, the primer is not complementary to an adapter sequence that is complementary to the guide polynucleotide targeting region. In some embodiments, the primer when bound to the adapter sequence and elongated does not elongate the entire adapter sequence.

### Reaction Mixtures

In some aspects, this disclosure describes a reaction mixture comprising: (i) a plurality of target polynucleotides, wherein the target polynucleotides comprise an adapter sequence; (ii) a predetermined concentration of a guide polynucleotide comprising a targeting region that is complementary to the adapter sequence; and (iii) a predetermined concentration of a dCas protein or a dArgonaute protein. In some embodiments, the reaction mixture comprises an aqueous solution (e.g., a buffer that is suitable for performing a method described herein). In some embodiments, the reaction mixture comprises a predetermined concentration of dCas9 protein). In some embodiments, the reaction mixture comprises a predetermined concentration of dArgonaute protein. In some embodiments, the reaction mixture comprises target

polynucleotides comprising an adapter sequence of any one of SEQ ID NOs: 27-36. In some embodiments, the reaction mixture comprises a predetermined concentration of a guide polynucleotide of any one of SEQ ID NOs: 27-36.

In some embodiments, the reaction mixture comprises a plurality of target polynucleotides comprising an adapter sequence at a concentration of 50 femtomoles (fmols) to 3,400 fmols. In some embodiments, the reaction mixture comprises a plurality of target polynucleotides comprising an adapter sequence at a concentration of 200 fmols to 3,400 fmols. In some embodiments, the reaction mixture comprises a plurality of target polynucleotides comprising an adapter sequence at a concentration of 300 fmols to 3,400 fmols. In some 10 embodiments, the reaction mixture comprises a plurality of target polynucleotides comprising an adapter sequence at a concentration of 500 fmols to 3,400 fmols. In some embodiments, the reaction mixture comprises a plurality of target polynucleotides comprising an adapter sequence at a concentration of 750 fmols to 3,400 fmols. In some embodiments, the reaction mixture comprises a plurality of target polynucleotides comprising an adapter sequence at a concentration 15 of 1,000 fmols to 3,400 fmols. In some embodiments, the reaction mixture comprises a plurality of target polynucleotides comprising an adapter sequence at a concentration of 1,000 fmols to 3,400 fmols. In some embodiments, the reaction mixture comprises a plurality of target polynucleotides comprising an adapter sequence at a concentration of 2,000 fmols to 3,400 fmols. In some embodiments, the reaction mixture comprises a plurality of target 20 polynucleotides comprising an adapter sequence at a concentration of 50 fmols. In some embodiments, the reaction mixture comprises a plurality of target polynucleotides comprising an adapter sequence at a concentration of 3,400 fmols. In some embodiments, the reaction mixture comprises a plurality of target polynucleotides comprising an adapter sequence at a concentration of 500 fmols to 3,400 fmols.

25 In some embodiments, the reaction mixture comprises a predetermined concentration of dCas9 of SEQ ID NO: 38, target polynucleotides comprising an adapter sequence of SEQ ID NO: 28, and a predetermined concentration of a guide polynucleotide (e.g., an RNA guide polynucleotide) that comprises a homology region that is complementary to the adapter sequence (e.g., the homology region comprises SEQ ID NO: 37).

30 In some embodiments, the reaction mixture comprises a predetermined concentration of dArgonaute of any one of SEQ ID NOs: 40-45 (e.g., a catalytically inactive variant of any one of

SEQ ID NOs: 40-45), target polynucleotides comprising an adapter sequence of SEQ ID NO: 28, and a predetermined concentration of a dArgonaute guide polynucleotide that comprises a targeting region which is complementary to the adapter sequence.

In some embodiments, the reaction mixture further comprises a catalytically active  
5 nuclease as described herein.

### **Ribonucleoprotein (RNP) complexes**

In some aspects this disclosure provides a ribonucleoprotein (RNP) complex comprising:  
10 (i) a guide polynucleotide; (ii) a Cas protein (e.g., dCas) or an Argonaute protein (e.g., dArgonaute); and (iii) an adapter sequence; wherein the targeting region of the guide polynucleotide is complementary to the adapter sequence.

In some embodiments, the RNP complex comprises (i) a RNA Cas gRNA polynucleotide  
15 (ii) a Cas protein as described herein, and (iii) an adapter sequence; wherein the homology region of the RNA Cas gRNA polynucleotide is complementary to the adapter sequence.

In some embodiments, the RNP complex comprises (i) a RNA Cas gRNA polynucleotide  
comprising a homology region of any one of SEQ ID NOs: 1-26 (ii) a dCas protein of SEQ ID  
NO: 37, and (iii) an adapter sequence of any one of SEQ ID NOs: 27-36; wherein the homology  
region of the RNA Cas gRNA polynucleotide is complementary to the adapter sequence.

In some embodiments, the RNP complex comprises (i) a RNA Cas gRNA polynucleotide  
20 comprising a homology region of any one of SEQ ID NO: 1 (ii) a dCas protein of SEQ ID NO:  
37, and (iii) an adapter sequence of any one of SEQ ID NOs: 28; wherein the homology region  
of the RNA Cas gRNA polynucleotide is complementary to the adapter sequence.

In some embodiments, the RNP complex comprises (i) an Argonaute guide  
25 polynucleotide(ii) an Argonaute protein, and (iii) an adapter sequence; wherein the targeting  
region of the Argonaute guide polynucleotide is complementary to the adapter sequence.

In some embodiments, the RNP complex comprises (i) an Argonaute guide  
polynucleotide (e.g., an siDNA) (ii) an Argonaute protein of any one of SEQ ID NOs: 39-45 and  
(iii) an adapter sequence of SEQ ID NO: 28; wherein the targeting region of the Argonaute guide  
30 polynucleotide is complementary to the adapter sequence.

In some embodiments, the RNP complex comprises (i) an Argonaute guide polynucleotide (e.g., an siDNA) (ii) an Argonaute protein of any one of SEQ ID NOs: 40-45 and (iii) an adapter sequence of SEQ ID NO: 28; wherein the targeting region of the Argonaute guide polynucleotide is complementary to the adapter sequence.

5 In some embodiments of RNP complexes provided herein, the adapter sequence is unmodified. In some embodiments, the adapter sequence is unmodified at the 5' end.

In some embodiments, the RNP complexes are present at a concentration of 250 femtomoles (fmols) to 14,000 fmols. In some embodiments, the RNP complexes are present at a concentration of 300 fmols to 14,000 fmols. In some embodiments, the RNP complexes are present at a concentration of 2,000 fmols to 14,000 fmols. In some embodiments, the RNP complexes are present at a concentration of 6,000 fmols to 14,000 fmols. In some embodiments, the RNP complexes are present at a concentration of 10,000 fmols to 14,000 fmols. In some embodiments, the RNP complexes are present at a concentration of at least 250 fmols. In some embodiments, the RNP complexes are present at a concentration of at least 300 fmols. In some embodiments, the RNP complexes are present at a concentration of at least 2,000 fmols. In some embodiments, the RNP complexes are present at a concentration of at least 6,000 fmols. In some embodiments, the RNP complexes are present at a concentration of at least 10,000 fmols. In some embodiments, the RNP complexes are present at a concentration of 250 fmols. In some embodiments, the RNP complexes are present at a concentration of 300 fmols. In some embodiments, the RNP complexes are present at a concentration of 2,000 fmols. In some embodiments, the RNP complexes are present at a concentration of 6,000 fmols. In some embodiments, the RNP complexes are present at a concentration of 10,000 fmols. In some embodiments, the RNP complexes are present at a concentration of 14,000 fmols.

25 **Sequences**

Table 2: Sequences

SEQ ID NO	Description	Sequence
1	ILLUMINA i5 gRNA-1 polynucleotide homology region	GAUCGGAAGAGCGUCGUGUA

2	ILLUMINA i5 gRNA-2 polynucleotide homology region	AGAUCGGAAGAGCGUCGUGU
3	ILLUMINA i5 gRNA-3 polynucleotide homology region	AGCGUCGUGUAGGGAAAGAG
4	ILLUMINA i5 gRNA-4 polynucleotide homology region	AGAGCGUCGUGUAGGGAAAG
5	ILLUMINA i5 gRNA-5 polynucleotide homology region	GAAGAGCGUCGUGUAGGGAA
6	ILLUMINA i5 gRNA-6 polynucleotide homology region	AUCGGAAGAGCGUCGUGUAG
1	ILLUMINA i5 gRNA-7 polynucleotide homology region	GAUCGGAAGAGCGUCGUGUA
2	ILLUMINA i5 gRNA-8 polynucleotide homology region	AGAUCGGAAGAGCGUCGUGU
9	ILLUMINA i7 gRNA-1 polynucleotide homology region	AGUUCAGACGUGUGCUCUUC
10	ILLUMINA i7 gRNA-2 polynucleotide homology region	GUGACUGGAGUUCAGACGUG
11	NEXTERA Read 1 gRNA-1 polynucleotide homology region	CACAUCUGAGACGCUGCCGA
12	NEXTERA Read 1 gRNA-2 polynucleotide homology region	AUACACAUCUGAGACGCUGC
13	NEXTERA Read 2 gRNA-1	GGCUCGGAGAUGUGUAUAAG

	polynucleotide homology region	
14	NEXTERA Read 2 gRNA-2 polynucleotide homology region	UGGGCUCGGAGAUGUGUAUA
15	NEXTERA Read 2 gRNA-3 polynucleotide homology region	GGGCUCGGAGAUGUGUAUAA
16	NEXTERA Read 2 gRNA-4 polynucleotide homology region	GUGGGCUCGGAGAUGUGUAU
17	Ion Torrect A adapter gRNA-1 homology region	GGAGACACGCAGGGAUGAGA
18	Ion Torrect A adapter gRNA-2 homology region	AGUCGGAGACACGCAGGGAU
19	Ion Torrect A adapter gRNA-3 homology region	GAGACACGCAGGGAUGAGAU
17	Ion Torrect A adapter gRNA-4 homology region	GGAGACACGCAGGGAUGAGA
21	Ion Torrect A adapter gRNA-5 homology region	GUCGGAGACACGCAGGGAUG
22	Ion Torrect A adapter gRNA-6 homology region	GAGUCGGAGACACGCAGGGA
23	Ion Torrect P adapter gRNA-1 homology region	ACUGCCCAUAGAGAGGAAAG
24	Ion Torrect P adapter gRNA-2 homology region	AUCACCGACUGCCCAUAGAG
25	Ion Torrect P adapter gRNA-3 homology region	GCCCAUAGAGAGGAAAGCGG
26	Ion Torrect P adapter gRNA-4 homology region	GCCUCCGCUUCCUCUCUAU

27	ILLUMINA i5 adapter sequence - Distal	AATGATACGGCGACCACCGAGATCTACAC
28	ILLUMINA i5 adapter sequence - Proximal	ACACTCTTTCCCTACACGACGCTCTTCCGATCT
29	ILLUMINA i7 adapter sequence - Distal	ATCTCGTATGCCGTC'TTCTGCTTG
30	ILLUMINA i7 adapter sequence - Proximal	AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC
31	ILLUMINA NEXTERA Read 1 adapter - Distal	AATGATACGGCGACCACCGAGATCTACAC
32	ILLUMINA NEXTERA Read 1 adapter - Proximal	TCGTCCGGCAGCGTCTCAGATGTGTATAAGAGACAG
33	ILLUMINA NEXTERA Read 2 adapter sequence - Distal	ATCTCGTATGCCGTC'TTCTGCTTG
34	ILLUMINA NEXTERA Read 2 adapter sequence - Proximal	CTGTCTCTTATACACATCTCCGAGCCCACGAGAC
35	ION TORRENT A Adapter - distal	CCATCTCATCCCTGCGTGTCTCCGACTCAG
36	ION TORRENT PI Adapter	ATCACCGACTGCCCATAGAGAGGAAAGCGGAGGCGTAGTGG
37	Sp_dCas9	MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTDRH SIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE IFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFL IEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAIL SARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSN FDLAEDAKLQLSKDQYDDDLNLLAQIGDQYADLFLAAKNL SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALV RQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT RKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL

		<p>GTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIER  LKTIAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS  GKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQG  DSLHEHIANLAGSPAIIKKGILQTVKVVDLKVVMGRHKPENI  VIEMARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHPV  ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVAI  VPQSFLKDDSIDNKVLTRSDKNRKGSDNVPSEEVVKKMKNY  WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE  TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSD  FRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKYPKLES  EFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFK  TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMP  QVNIKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY  GGFDSPTVAYSVLVVAKVEKKGKSKLKS VKELLGITIMERS  FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS  AGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQKQL  FVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHRD  KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKR YTSTKEV  LDATLIHQ SITGLYETRIDLSQLGGD</p>
38	spCas9	<p>MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRH  SIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE  IFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVA  YHEKYPTIYHLRKKLADSTDKADLRIYLALAHMIKFRGHFL  IEGDLNPDNSVDKLFQILVQTYNQLFEENPINASRVDAKAIL  SARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKN  FDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADLFLAAKNL  SDATLLSDILRVNSEITKAPLSASMIKRYDEHHQDLTLLKALV  RQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK  MDGTEELLAKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR  RQEDFPFLKDNREKIEKILTRIPYYVGPLARGNSRFAWMT  RKSEETITPWNFEVVDK GASAQSFIERMTNFDKNLPNEKVL  PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV  DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL  GTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDKEMIEE  RLKTYAHLFDDKVMKQLKRRHYTGWGRLSRKLINGIRDKQ  SGKTILDFLKSDGFANRNFQIQLIHDDSLTFKEAIQKAQVSGQG  HSLHEQIANLAGSPAIIKKGILQSVKVVDLKVVMGHKPENIV  IEMARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHPVE  NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP  QSFIDDDSIDNKVLTRSDKNRKGSDNVPSEEVVKKMKNYWR  QLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQ  ITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK  DFQFYKVRINNYHHAHDAYLNAVVGTAIIKYPKLESEFEV  YGDYKVYDVRKMIKSEQEIGKATAKRFFYSNIMNFFKTEIT  LANGEIRKRPLIETNEETGEIVWDKGRDFATVRKVL SMPQVN  IVKKTEVQTGALTNESIYARGSFDKLISRKH RFESSKYGGFGS</p>

		PTVTYSVLVVAKSKVQDGKVKKIKTGKELIGITLLDKLVFEK NPLKFIEDKGYGNVQIDKCIKLPKYSLEFENGTRRMLASVM ANNNSRGDLQKANEMFLPAKLVTLTYHHAHKIESSKELEHEA YILDHYNDLYQLLSYIERFASLYVDVEKNISKVKELFSNIESY SISEICSSVINLLTLTASGAPADFKFLGTTIPRKRYGSPQSILSST LIHQSI TGLYETRIDLSQLGSD
39	Argo-1 Piwi domain-containing protein (WP_058142162.1)	MNNLTFEAFEGIGQLNELNFYKYRLIGKGGQIDNVHQAIWSVK YKLQANNFFKPVFVKGEILYSLDELKVIPEFENVEVILDGNIIL SISENTDIYKDVIVFYINNALKNIKDITNYRKYITKNTDEIICKS ILTTNLKYQYMKSEKGFKLQRKFKISPVVFRNGKVLILYLNCS SDFSTDKSIYEMLNGLDVVGLQVKNRWTNSNGNIFIEVLD KSISEPGTSGKLGQSLIDYYINGNQKYRVEKFTDEDKAKVI KAKIKNKTYNYIPQALTPVITREYLSHTDKKFSKQIENVIKMD MNYRYQTLKSFVEDIGVIKELNNLHFKNQYYTNFDFMGFES GILEEPVLMGANGKIKDKKQIFINGFFKNPKENVKFGVLYPE GCMENAQSIARSILDFATAGKYNKQENKYISKNL MNIGFKPS ECIFESYKLG DITEYKATARKLKEHEK VGFVIAVIPDMNESEV ENPYNPFKKVWAKLNIPSQMITLKTTEKFKNIVDKSGLYLYH NIALNILGKIGGIPWIIKDMPGNIDCFIGLDVGTREKGIHFPAC SVLFDKYGKLINYYKPTIPQS GEKIAETILQEIFDNVLISYKBE NGEYPKNIVIH RDGFSRENIDWYKEYFDKKGKIFNIEVKKNI PVKIAKV VGSNICNPIKGSYVLKNDKAFIVTTDIKDGVASPNP LKIEKTYGDVEMKSILEQIYSLSQIHVGSTKSLRLPITTYAD KICKAIEYIPQGVVDNRLFFL
40	Argo-2 (EHP50500.1)	MSNLTVEAFEGIGSVNPMLFYQYKVTGKGKYDNVYKIIKSA RYKMHSKNRFRKPVFIKDDKLYTLEKLPDIEDLDFANINFVKS EVL SIEDNMSIYGEVVEYYINLKLKKVKVLGKYPKYRINYSK EILSNTLLTRELKDEFKSNKGFNLKRKFRISPVVNKMGKVIL YLSCSADFSTNKNIYEMLKEGLEVEGLAVKSEWSNISGNLVI ESVLETKISEPTSLGQSLIDYYKNNNQGYRVKDFTDEDLNAN IVNVRGNKKIYMYIPHALKPIITREYLA KNDPEFSKEIEQLIKM NMNYRYETLKS FVNDIGVIEELNNLSFKNKYYEDVKLLGYSS GKIDEPVLMGAKGIIKNKMQIFSNGFYKLPEGKVRFGVLYPK EFDGVSRAIRAIYDFSKEGKYHGESNKYIAEHLINVEFNPKE CIFEGYELGDITEYKKAALKLN NYNNVDFVIAIVPNMSDEEIE NSYNPFKKIWAELNLP SQMISVKTAEIFANSRDNTALYYLHN IVLGILGKIGGIPWVVKDMKGDVDCFVGLDVG TREKGIHYP ACSVVFDKYGKLINYYKPNIPQNGEKINTEILQEIFDKVLISYE EENGA YPKNIVIH RDGFSREDLDWYENYFGKKNIKFNIEVK KSTPLKIASINEGNITNPEKGSYILRGNKAYMVTTDIKENLGS PKPLKIEKSYGDIDMLTALSQIYALTQIHVGATKSLRLPITTY YADKICKAIEFIPQGRVDNRLFFL
41	Argo-3 (IZ25_A Chain A)	SMKAIVVINLVKINKKIIPDKIYVYRLFNDPEEELQKEGYSIYR LAYENVGIVIDPENLIIATTKELEYEGEPIEPEISFSEL RNDYQ SKLVLRLKENGIGEYELSKLLRKFRKPKTFGDYKVIPS VEM SVIKHDEDFYLVIHIIHQIQSMKTLWELVNKDPKELEEFMTH

		<p>KENLMLKDIASPLKTVYKPCFEEYTKKPKLDHNDQEIWKYWY          NYHIERYWNTPEAKLEFYRKFGQVDLQKQPAILAKFASKIKKN          KNYKIYLLPQLVVPTYNAEQLESDVAKEILEYTKLMPEERKE          LLENILAEVDSDIIDKSLSEIEVEKIAQELENKIRVRDDKNSV          PISQLNVQKSQLLLWTNYSRKYPVILPYEVPEKFRKIREIPMFI          ILDSGLLADIQNFATNEFREL VKSMYYSLAKKYNLSLAKKARS          TNEIGLPFLDFRGKEKVITEDLNSDKGHEVVEQVSSFMKGKE          LGLAFIAARNKLSSEKFEEIKRRLFNLNVISQVVNEDTLKNKR          DKYDRNRLDLFVRHNLFFQVLSKLGVKYVLDYRFNYDYII          GIDVAPMKRSEGYIGGSVVMFDSQGYIRKIVPIKIGEQRGESV          DMNEFFKEMVDKFKFNKLDNKKILLLRDGRITNNEEEGLK          YISEMFDIEVVTMDVIKHPVRAFANMKMYFNLGGAIYLIPH          KLKQAKGTPIPIKLAKKRIKNGKVEKQSITRQDVLDFILTRL          NYGSISADMRLPAPVHYAHKFANAIKNEWKIKEEFLAEGFLY          FV</p>
42	Argo-4 (WP_015791216.1)	<p>MKNLRYKINAYRIKKDYIPKEVYRYRIRSFIEININIYRFVGFY          GGVALNQSEFILPYPVENLVLEYDGGKDVKLEHIDTLNLEDIE          NKDKEKAEKLVRYLTSIYKLPILYKILRDVRESKIINDIRV          DPIPDFTVKRHNNEYLLVIDFNHTATVLKNLWDFVGRDKLK          LEDYIGKKIIFKPNPKRYTIKSIEKQNKDIDDIVEHIEIYK          WTEEEIKSTFGEIDYTQPIIHCEGIPYPFAPQFCNIVFTMEDLD          ENTLKDLQSYWRLPNEIKGNIINQIAKKLRFVENEPIELEFIKF          NNTPLIVKDENGKPTKIYTTNRLFRWNYDSKSKLYLPYDIPDI          IKNKTLTTFVLIDENLKNVSGKIKRKVYQMFKNYNKIASKTE          LPKFDFAKWKYFSNNNIRDVIRKIKDEFNEELGFALIIGNRY          YENDYYETLKMQLFNLNII SQNILWENWSKDDNNFMTNLL          IQIMGKLGKIFYFALDAKVNYDYIMGLDSGLGAFKSNRVSGC          TVIYDSEGKIRRIQPIDVPSGERIPIHLVVEFLETKTINMENK          NILFLRDGFVQNSEREELKLSKELNSNIEVISIRKNNKYKVF          TSDYGIGSIFGNDGIFLPHKTTFGSNPVKLSWLRFNNGNEEK          LKINESIMQLLYDLTKMNY SALYGEGRNLRIPAPIHYADKFV          KALGKNWKIDEELLKHGFLYFI</p>
43	Argo-5 (NPB03901.1)	<p>MLVNGFKLKIPLIAKDHENRVVEIYSTSTVLKEKFVPEIPQFI          RGREIETWILIDKEIKDNFQEIKKTVIEFFRYNSIRNNLLPYF          KFSYNYVIFSRQNVFKVLTKMNGVDSDAVGFALIIGKQKYR          NSDYEEIKRILFNKNIISQNLWDQGTLLKNNFARNNIIIIQILS          KLGKIFYVFLKYNAEYDYIFGVDIGKEKFSGSSLGGCTIYDYK          GELKKIVPIEILAKKETLNLERIFETLQLDMNLEGKKILLLRDG          SIKNFEKEQLKNISKRYHVEITLNIKKYSKFRIGNDEGGIGVL          IEDIALLLPHHYPGSKPIKIDNKILFKDGNYTELEINANDLEL          IYGLSKLNYSSLSSEERILRLPAPVHYAHKFVKALGKGWKIR          KDLLEEGCLYFI</p>
44	Argo-6 (WP_012966655.1)	<p>MMLNIFEVDKNRVEVPQDVYLYKVHLKTLEQRKRDMCIAV          LRNSFGYLDVNSFVIYSYKEIRDLPRIKKYCDLEPTGKVKM          NEVDENVRNALVKTFLRSKIRKDIKLLKFKKFKQSVGRW          TVALDLERIELVEHNGEILVSFNVKLNISMINLWDIHERDVN</p>

		<p>RLKGLCWSPENLNDNRIWFRYIPHLIIEEVEDFEESAKSFILSD  VHTGEEFGGWTTEDLKKYPENEYGLSEDAIKKIGNYTQFDST  QPIIKGVTWSGKEYPFLPQHCIPAYNPMLATABEKKRIIEIKIN  LKNKKDEIIRKIIELPYLKQPDNIEIKKVQETARLRAKFVKV  EVTGKITKTLSPYEKPVSSSTLDLFGWISRIIDGGDGIIEICIP  DYIPENLSRIKEIEAFLLIENDLNENERKVGDKLLNDAIYVYN  FVRSVCLRCGINIPYLNKGNRFYFENSKEGIRDYKRITTSLS  GEIGFALIFGKRDNYEDEEGEDSFDYYNPLKSALFRNNILSQN  FDVTNYVRGDGKINKNTIKYA VSNIYNIFGKLGVKFFVLEED  VPYDYILGIDVGYGEAYTGKVAGCTTVHDSEGRLRNLPIEK  QNYPKETARIKALLEEIEQKKKIYNIDFENKSILILRDGRINK  EEINQLMEFSEERNCRITYIEIRKNIVHQFLVNSSQACYVKIGD  YYILKAHNPRIGFPRAIKIARKIVIEGDAWRESSLTEDDILLIY  KLTALNYSTIGRDSNLRIPAPIYYADKLVKALKKGWKFDERF  LRYGILYFL</p>
45	Argo-7 (HIP43457.1)	<p>MKGRNNSLKLNIFRVNLEQLSIPEKIYVYNISANILDSSKLTS  FYKRLNIYGFMDTQNRKLYSYKEIDFISKNMKENLNLKLEK  TIKLSELESSFRNQILKTYIRNSVKMDINRIKEIQHQFLGKEK  KRTGRWNLSIIPERINVEHINDAFYVAFNVKLRILANKNLWD  FIGRDLEKLTLCWFPDKSKDIKWFRYVPDLKEENERSYLL  TYIQSKDEAKNSGFSFEDLKYYPQEKRNTTYGELKEIAKFED  FDENQPIIVGVSSDMDRNPLYFLPQYCIPAYNPVLASENESKK  IQKVYESVLFNRKYEHYKIYDKIPYLELNYEDISFKELDNHRK  GKLKVNFKAKLYLGKDKKDKQGKEIKCKVVGKPKQKRTIE  NTADLFSWIHKLEDLRGKKKKELIVDIPIPEYVPEYLQKLDEI  GTFLLVESGNPSSDIDMIKNFLLLIADVYRVIREASEGFNKIPR  LKFIKNPETNRFEBFLFKSSEGIDKTVRELQQLKKGKKElse  KELGFAFIFGSQDDFVEEHEDFDYYIPLKRNLFNNILSQNFLI  DTYTNKNKIKFALSNIYVNLFGKLGKIFFALEEKVYYDYILGI  DTGLAEAYTGRVAGCTTVHDSNGRLKNIPIEKLNPARRTV  RIKALLEEIHIDADYNMDFSNKKILILRDGKIQPEELKQLVEFT  KSKKCKITMIDVRKHTVYQWLEKGNDKHLSEIKVGDFFCLLKP  HSPRRGYPRMLKISQKVEIDENGFTYKDLTDYDILLIYKLTLL  NYSTIGRPSNLKLP GPIYYADKLVKALKRGWKLEPKFLKEGF  LYFL</p>
46	Exemplarily sgRNA of Yin et al. 20017	<p>GGGCGAGGAGCUGUUCACCGGUUUUAGAGCUAGAAAUA  GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA  AAGUGGCACCGAGUCGGUGCUUUUU</p>
47	His Tag	HHHHHH

Methods of Polynucleotide Library Normalization

In some aspects, this disclosure provides methods of normalizing the concentration of target polynucleotides between two or more samples (e.g., normalizing between two or more polynucleotide libraries).

“Normalizing”, “normalization”, and similar terms used herein refer to the process of producing a subsequent sample with a desired concentration of target polynucleotides from an initial sample having a different starting concentration of target polynucleotides than the subsequent sample. In some embodiments, “normalization” of two or more samples results in the production of two or more subsequent samples each having a concentration of target polynucleotides that are more similar to each other after performing the normalization methods than before performing the normalization methods. For example, a first sample (e.g., a first polynucleotide library) may comprise 5-fold more target polynucleotide than a second sample (e.g., a second polynucleotide library). In this example, after normalization of the first sample and the second sample, the difference in concentration between the first sample and the second sample would be less than 5-fold (e.g., less than 4-fold, less than 3-fold, or less than 2-fold). In some embodiments, normalization of two or more samples results in the production of two or more subsequent samples each having equimolar (i.e., 1:1) concentrations. In some embodiments, the method of normalization results in differences in target polynucleotide concentration between two or more samples being less than 50% (e.g., less than 40%, less than 30%, less than 25%, less than 20%, less than 15%, less than 10%, less than 5%, or less than 2.5%). In some embodiments, the method of normalization results in differences in target polynucleotide concentration between two or more samples being 5%-40%, 5%-30%, 5%-20%, 5%-10%, 10%-40%, 10%-30%, or 10%-20% in concentration. In some embodiments, the difference in starting target polynucleotide concentration between two or more samples may be relatively small (e.g., less than 5-10%) prior to performing a method of normalization described herein. In such embodiments, the method of normalization may be performed but the two or more samples may not have a detectably more similar concentration after normalization than before normalization.

“Target polynucleotide” or “target polynucleotides” refers to a polynucleotide that comprises nucleic acids encoding an adapter sequence or a static region of an adapter sequence. In some embodiments, the target polynucleotide comprises a polynucleotide that comprises (1) nucleic acids encoding an adapter sequence or a sequence that has been added to the

polynucleotide for the purpose of normalizing using a method described herein (e.g., a Zinc-finger binding domain or a Talen binding domain) and (2) nucleic acids encoding a sequence of interest. In some embodiments, the target polynucleotide comprises nucleic acids encoding an adapter sequence described herein and nucleic acids encoding a sequence of interest. The  
5 sequence of interest may be any sequence for which normalization is desired. For example, a sequence of interest may comprise, but is not limited to, DNA, genomic DNA, circulating tumor DNA, RNA, rRNA, mRNA, miRNA, or cDNA. The adapter sequence may be ligated to the 5' end and/or 3' end of the sequence of interest. In some embodiments, the target polynucleotide comprises a first adapter sequence (e.g., a p5 adapter) on the 5' end of the sequence of interest  
10 and a second adapter sequence (e.g., a p7 adapter) on the 3' end of target sequence.

A "sample" refers to a composition or solution comprising one or more target polynucleotides. In some embodiments, the sample comprises a plurality of target polynucleotides. A "plurality" refers to two or more. For example, a plurality of target polynucleotide comprises two or more target polynucleotides. In some embodiments, the  
15 plurality of target polynucleotides comprises multiple copies of a single target polynucleotide sequence (e.g., at least 10, at least 100, at least 1000, at least 10,000, at least 100,000, or at least 1,000,000, or at least 10,000,000, at least 100,000,000 target polynucleotides). In some embodiments, the plurality of target polynucleotides comprises different target polynucleotide sequences (e.g., at least 10, at least 100, at least 1000, at least 10,000, at least 100,000, or at least  
20 1,000,000, or at least 10,000,000, at least 100,000,000 different target polynucleotides), but at least some of the target polynucleotides (e.g., all of the target polynucleotides) comprise the same adapter sequence. In some embodiments, the sample comprises a target polynucleotide and other molecules (e.g., polynucleotides that are not target polynucleotides and do not comprise an adapter sequence or comprise a different adapter sequence from the one complementary to the  
25 guide polynucleotide). In some embodiments, the sample comprises target polynucleotides comprising nucleic acids encoding genomic DNA. In some embodiments, the sample comprises target polynucleotides comprising nucleic acids encoding an RNA. In some embodiments, the sample comprises target polynucleotides comprising nucleic acids encoding a cDNA (e.g., of mRNA or lncRNA). In some embodiments, the sample comprises target polynucleotides  
30 comprising nucleic acids encoding an exon. In some embodiments, the sample comprises target polynucleotides comprising nucleic acids encoding an intron.

In some embodiments, the plurality of target polynucleotides in the sample is present at a concentration of 50 femtomoles (fmols) to 3,400 fmols. In some embodiments, the plurality of target polynucleotides is present at a concentration of 200 fmols to 3,400 fmols. In some embodiments, the plurality of target polynucleotides is present at a concentration of 300 fmols to 3,400 fmols. In some embodiments, the plurality of target polynucleotides is present at a concentration of 500 fmols to 3,400 fmols. In some embodiments, the plurality of target polynucleotides is present at a concentration of 750 fmols to 3,400 fmols. In some embodiments, the plurality of target polynucleotides is present at a concentration of 1,000 fmols to 3,400 fmols. In some embodiments, the plurality of target polynucleotides is present at a concentration of 1,000 fmols to 3,400 fmols. In some embodiments, the plurality of target polynucleotides is present at a concentration of 2,000 fmols to 3,400 fmols. In some embodiments, the plurality of target polynucleotides is present at a concentration of 50 fmols. In some embodiments, the plurality of target polynucleotides is present at a concentration of 3,400 fmols. In some embodiments, the plurality of target polynucleotides is present at a concentration of 500 fmols to 3,400 fmols.

In some embodiments, the method of normalizing is performed for at least 2 samples (e.g., at least 3 samples, at least 4 samples, at least 5 samples, at least 6 samples, at least 7 samples, at least 8 samples, at least 9 samples, at least 10 samples, at least 25 samples, at least 50 samples, at least 75 samples, at least 100 samples, at least 150 samples, or at least 200 samples, at least 300 samples, at least 384 samples, at least 400 samples, at least 500 samples, at least 600 samples, at least 700 samples, at least 800 samples, at least 900 samples, at least 1000 samples, at least 1500 samples, at least 2000 samples, at least 3000 samples, or at least 5000 samples or more). In some embodiments, the method of normalizing is performed for 2-384 samples. In some embodiments, the method of normalizing is performed for 24-384 samples. In some embodiments, the method of normalizing is performed for 24-500 samples. In some embodiments, the method of normalizing is performed for 24-750 samples. In some embodiments, the method of normalizing is performed for 24-1000 samples. In some embodiments, the method of normalizing is performed for 24-2000 samples. In some embodiments, the method of normalizing is performed for 24-3000 samples.

In some embodiments, the samples comprise target polynucleotides at a pre-normalization concentration of within a 70-fold concentration of each other. In some

embodiments, the samples comprise target polynucleotides at a pre-normalization concentration of within a 60-fold concentration of each other. In some embodiments, the samples comprise target polynucleotides at a pre-normalization concentration of within a 50-fold concentration of each other. In some embodiments, the samples comprise target polynucleotides at a pre-normalization concentration of within a 40-fold concentration of each other. In some  
5 embodiments, the samples comprise target polynucleotides at a pre-normalization concentration of within a 30-fold concentration of each other. In some embodiments, the samples comprise target polynucleotides at a pre-normalization concentration of within a 20-fold concentration of each other. In some embodiments, the samples comprise target polynucleotides at a pre-normalization concentration of within a 10-fold concentration of each other. In some  
10 embodiments, the samples comprise target polynucleotides at a pre-normalization concentration of within a 5-fold concentration of each other. In some embodiments, the samples comprise target polynucleotides at a pre-normalization concentration of within a 2-fold concentration of each other. In some embodiments, the samples comprise target polynucleotides at a pre-normalization concentration of within a 1.5-fold concentration of each other.  
15

In some embodiments, this disclosure provides a method for normalizing the concentration of target polynucleotides between at least two samples (e.g., two polynucleotide libraries) each comprising a target polynucleotide, the method comprising, for each sample of the at least two samples: (i) obtaining the sample, wherein the target polynucleotides of the sample  
20 comprise an adapter sequence; (ii) producing a solution, the producing comprising combining (a) the sample, (b) a predetermined concentration of a guide polynucleotide comprising a targeting region that is complementary to a static region of the adapter sequence of the target polynucleotides of the sample, and (c) a predetermined concentration of a dCas or a dArgonaute; (iii) contacting the solution with a solid phase comprising a dCas or a dArgonaute binding  
25 molecule; (iv) separating the solution from the solid phase; and (v) extracting the target polynucleotides from the solid phase to normalize the concentration of target polynucleotides between the two or more samples.

In some embodiments, this disclosure provides a method for normalizing the concentration of target polynucleotides between at least two samples (e.g., two polynucleotide libraries) each comprising a target polynucleotide, the method comprising, for each sample of the  
30 at least two samples: (i) obtaining the sample, wherein the target polynucleotides of the sample

comprise an adapter sequence; (ii) producing a solution, the producing comprising combining (a) the sample, (b) a predetermined concentration of a guide polynucleotide comprising a targeting region that is complementary to the adapter sequence of the target polynucleotides of the sample, and (c) a predetermined concentration of a dCas or a dArgonaute; (iii) contacting the solution  
5 with a solid phase comprising a dCas or a dArgonaute binding molecule; (iv) separating the solution from the solid phase; and (v) extracting the target polynucleotides from the solid phase to normalize the concentration of target polynucleotides between the two or more samples.

In some embodiments, a method for normalizing the concentration of target polynucleotides between at least two samples comprises a step of obtaining a sample comprising  
10 target polynucleotides. The sample may be obtained from any suitable source, including, but not limited to cells (e.g., eukaryotic or prokaryotic cells), tissues (e.g., brain, heart, lung, liver, kidney, adipose, skin, gall bladder, or breast), diseased tissues (e.g., tumor), bodily fluid (e.g., saliva, sweat, blood, urine, mucus, or cerebral spinal fluid), or from *in vitro* production. In some embodiments, obtaining a sample comprises obtaining polynucleotides of interest (e.g.,  
15 extracting the polynucleotide of interest from a biological sample) and modifying the polynucleotides of interest to comprise adaptor sequences (e.g., modifying the polynucleotides of interest to become target polynucleotides). Modifying the polynucleotides of interest to comprise adaptor sequences may be performed using any suitable methods, including, but not limited to using a kit for attaching adapter sequences to polynucleotides described herein. In  
20 some embodiments, the sample comprises 1-100 nM of target polynucleotides. In some embodiments, the sample comprises 1-1000 nM of target polynucleotides.

In some embodiments, producing a solution, in the context of the method of normalization, refers to combining, the sample and the predetermined concentration of a guide polynucleotide and the predetermined concentration of a dCas or a dArgonaute in an aqueous  
25 solution (e.g., a buffer suitable for the method). In some embodiments, producing the solution comprises combining a pre-determined amount of dCas protein or dArgonaute protein and a predetermined amount of guide polynucleotide prior to combining with the sample. In some embodiments, producing a solution comprises combining a predetermined amount of a guide RNA polynucleotide that is RNA with a predetermined amount of dCas9 protein or dArgonaute  
30 protein. In some embodiments, producing a solution comprises combining a predetermined amount of a guide polynucleotide that is DNA with a predetermined amount of dArgonaute

protein. In some embodiments, producing a solution comprises combining a predetermined amount of a guide RNA comprising a homology region of any one of SEQ ID NOs: 1-26 with a predetermined amount of dCas9 of SEQ ID NO: 37.

A “predetermined concentration” refers to a concentration (e.g., nanomolar) or amount (e.g., nanomoles) of a reagent (e.g., a guide polynucleotide, a dCas protein or a dArgonaute protein) that has been selected for extracting a particular or consistent amount of target polynucleotides from a sample. In some embodiments, the same or similar amounts of a predetermined concentration of a dCas protein or a dArgonaute protein is added to each sample that is to be normalized. In some embodiments, the same or similar amounts of a pre-determined concentration of a guide RNA polynucleotide is added to each sample that is to be normalized. In some embodiments, a predetermined concentration of a dCas protein or a dArgonaute protein is added to the sample, and a predetermined amount of guide RNA polynucleotide added to the sample is in excess of the predetermined amount of dCas protein or a dArgonaute protein added to the sample. In some embodiments, the predetermined amount of dCas protein or dArgonaute protein is 1-2000 femtomoles. In some embodiments, the predetermined amount of dCas protein or dArgonaute protein is 60-2000 femtomoles. In some embodiments, the predetermined amount of dCas protein or dArgonaute protein is 60-1000 femtomoles. In some embodiments, the predetermined amount of dCas protein or dArgonaute protein is 125-1000 femtomoles. In some embodiments, the predetermined amount of dCas protein or dArgonaute protein is at least 60 femtomoles (e.g., at least 60 femtomoles, at least 125 femtomoles, at least 250 femtomoles, at least 500 femtomoles, at least 750 femtomoles, or at least 1000 femtomoles). In some embodiments, the predetermined amount of dCas protein or dArgonaute protein is 50, 60, 125, 250, 750, 1000, 1500, or 2000 femtomoles. In some embodiments, the predetermined concentration of guide RNA polynucleotides is in excess to the predetermined concentration of dCas9 protein or dArgonaute protein.

In some embodiments, ribonucleoprotein (RNP) complex (e.g., the guide RNA, dCas protein or dArgonaute protein, and adapter sequence) is present at in the sample at a concentration of 250 femtomoles (fmols) to 14,000 fmols. In some embodiments, the RNP complex is present in the sample at a concentration of 300 fmols to 14,000 fmols. In some embodiments, the RNP complex is present in the sample at a concentration of 2,000 fmols to 14,000 fmols. In some embodiments, RNP complex is present in the sample at a concentration

of 6,000 fmols to 14,000 fmols. In some embodiments, RNP complex is present in the sample at a concentration of 10,000 fmols to 14,000 fmols. In some embodiments, the RNP complex is present in the sample at a concentration of at least 250 fmols. In some embodiments, RNP complex is present in the sample at a concentration of at least 300 fmols. In some embodiments, the RNP complex is present in the sample at a concentration of at least 2,000 fmols. In some embodiments, RNP complex is present in the sample at a concentration of at least 6,000 fmols. In some embodiments, the RNP complex is present in the sample at a concentration of at least 10,000 fmols. In some embodiments, the RNP complex is present in the sample at a concentration of 250 fmols. In some embodiments, the RNP complex is present in the sample at a concentration of 300 fmols. In some embodiments, the RNP complex is present in the sample at a concentration of 2,000 fmols. In some embodiments, the RNP complex is present in the sample at a concentration of 6,000 fmols. In some embodiments, the RNP complex is present in the sample at a concentration of 10,000 fmols. In some embodiments, the RNP complex is present in the sample at a concentration of 14,000 fmols.

15 In some embodiments, the method comprising producing a solution comprising a predetermined concentration of a dCas protein comprising an affinity tag or dArgonaute protein comprising an affinity tag as described herein.

In some embodiments, the affinity tag binding molecule comprises a metal ion (e.g., Ni<sup>2+</sup>) and the affinity tag comprises a His Tag. In some embodiments, the affinity tag binding molecule comprises biotin and the affinity tag comprises avidin. In some embodiments, the affinity tag binding molecule comprises an anti-myc antibody and the affinity tag comprises a myc tag. In some embodiments, the affinity tag binding molecule and a corresponding affinity tag is selected from Table 1.

25 In some embodiments, the dArgonaute or dCas of the method does not comprise an affinity tag. In some embodiments, the method comprises contacting the solution with a solid phase comprising a molecule that binds to the dCas or the dArgonaute (e.g., an antibody that binds to the dCas or dArgonaute).

30 In some aspects, the method comprises contacting the solution and a solid phase. In some embodiments, the solid phase comprises a molecule that binds to dArgonaute or dCas (e.g., an antibody or an affinity tag binding molecule). In some embodiments, the solid phase comprises a beads (e.g., a microparticle or a nanoparticle). In some embodiments, the bead is a

metal bead, a polymer bead, a protein bead, or a lipid bead. In some embodiments, the bead (e.g., metal bead) comprises metal ions (e.g., Ni<sup>2+</sup>, Co<sup>2+</sup>, Cu<sup>2+</sup>, and/or Zn<sup>2+</sup> ions) on the surface of the bead. In some embodiments, the metal bead is magnetic (e.g., paramagnetic, diamagnetic, or ferromagnetic).

5 In some embodiments, the method comprises incubating the solid phase and the solution. In some embodiments, the incubation is at room temperature. In some embodiments, the incubation is at 30-40 °C. In some embodiments, the incubation is at 35-38 °C. In some  
embodiments, the incubation is at or about 37 °C. In some embodiments, the incubation is at  
least 15 minutes (e.g., at least 30 minutes, at least 45 minutes, at least 60 minutes, or at least 2  
10 hours). In some embodiments, the incubation is at least 60 minutes. In some embodiments, the  
incubation is 15 minutes to 2 hours.

In some embodiments, the method further comprises separating the solution from the  
solid phase. "Separating" as used in the context of this method, refers to removing the solution  
from the solid phase or removing the solid phase from the solution. In some embodiments,  
15 separating is not a perfect separation, e.g., there may small amount of solution (e.g., less than  
about 1 microliter of solution) and solid phase that are still in contact with one another after  
separation. In some embodiments, the purpose of the separating step is to separate the unbound  
target polynucleotide of the solution from the target polynucleotides bound to the solid phase,  
which is part of normalizing concentration. This may be done by (1) capturing the magnetic  
20 beads bound to target polynucleotides (e.g., using a magnet) (2) removing the solution from the  
solid phase (e.g., using a pipette) and (3) repeatedly rinsing the solid phase with a buffer (e.g., a  
buffer that does not denature the Cas protein or the Argonaute protein). In some embodiments,  
the method comprises washing the solid phase at least 1 (e.g., at least 2, at least 3, at least 4, or at  
least 5, or more times).

25 In some embodiments, the method further comprises extracting the target polynucleotide  
from the solid phase. In some embodiments, the extracting comprising contacting the solid  
phase with a protease, which proteolyzes the Cas protein or Argonaute protein releasing the  
target polynucleotides. The protease may be any suitable protease, including, but not limited to  
trypsin, chymotrypsin, endoproteinase Lys-C, endoprotease AspN, endoprotease GluC, elastase,  
30 proteinase K, or papain. Corresponding protease reaction conditions and incubation times are  
known in the art. In some embodiments, the extracting comprising contacting the solid phase

with a protein denaturing agent (e.g., a detergent, and organic solvent, or a chaotropic agent). In some embodiments, extracting comprising increasing the temperature of the solid phase (e.g., by warming a liquid in which the solid phase is located) to denature the Cas protein or Argonaute protein). In some embodiments, extracting comprises increasing or decreasing the pH of liquid  
5 in which the solid phase is located.

In some embodiments, the method of normalizing comprises, performing, in sequential order: (i) obtaining the sample, wherein the target polynucleotides of the sample comprise an adapter sequence; (ii) producing a solution, the producing comprising combining (a) the sample,  
10 (b) a predetermined concentration of a guide polynucleotide comprising a targeting region that is complementary to the adapter sequence of the target polynucleotides of the sample, and (c) a predetermined concentration of a dCas or a dArgonaute comprising an affinity tag, wherein the predetermined concentration of the dCas or the dArgonaute corresponds to the guide  
polynucleotide; (iii) contacting the solution with a solid phase comprising an affinity tag binding molecule that is capable of binding to the affinity tag; (iv) separating the solution from the solid  
15 phase; and (v) extracting the target polynucleotides from the solid phase to normalize the concentration of target polynucleotides between the two or more samples.

In some embodiments, a method for normalizing the concentration of target polynucleotides between at least two samples comprises, for each sample of the at least two samples: (i) obtaining the sample, wherein target polynucleotides of the sample comprise an  
20 adapter sequence of any one of SEQ ID NOs 27-36; (ii) producing a solution, the producing comprising combining (a) the sample, (b) a predetermined concentration of a Cas gRNA polynucleotide comprising a homology region that is complementary to the adapter sequence, and (c) a predetermined concentration of a dCas9 comprises a His-Tag of SEQ ID NO: 47 (e.g., a his-tag on the n-terminus of the dCas9); (iii) contacting the solution with a magnetic solid phase  
25 comprising an Ni<sup>2+</sup> ion; (iv) separating the solution from the solid phase; and (v) extracting the target polynucleotides from the solid phase to normalize the concentration of target polynucleotides between the two or more samples.

In some embodiments, a method for normalizing the concentration of target polynucleotides between at least two samples comprises, for each sample of the at least two  
30 samples: (i) obtaining the sample, wherein target polynucleotides of the sample comprise an adapter sequence of any one of SEQ ID NOs: 27-36; (ii) producing a solution, the producing

comprising combining (a) the sample, (b) a predetermined concentration of a Argonaute guide polynucleotide (e.g., an siDNA) comprising a targeting region that is complementary to the adapter sequence, and (c) a predetermined concentration of a an Argonaute protein comprising a His-Tag of SEQ ID NO: 47; (iii) contacting the solution with a magnetic solid phase comprising an Ni<sup>2+</sup> ion; (iv) separating the solution from the solid phase; and (v) extracting the target polynucleotides from the solid phase to normalize the concentration of target polynucleotides between the two or more samples.

In some aspects, this disclosure provides a method for normalizing the concentration of target polynucleotides between at least two samples each comprising target polynucleotides, wherein the target polynucleotides that are not captured during normalization are digested. In some embodiments, the method comprises, for each sample of the at least two samples: (i) obtaining the sample (e.g., as described herein), wherein the target polynucleotides of the sample comprise an adapter sequence; (ii) producing a solution, the producing comprising combining (a) the sample, (b) a predetermined concentration of a guide polynucleotide comprising a targeting region that is complementary to the adapter sequence of the target polynucleotides of the sample, and (c) a predetermined concentration of a dCas or a dArgonaute comprising an affinity tag; and (iii) contacting the solution with a nuclease (e.g., a catalytically active Cas or Argonaute protein 9) to normalize the concentration of target polynucleotides between the two or more samples (e.g., digest the target polynucleotides that are not bound to the solid phase).

In some embodiments, the method comprises contacting the solution with a nuclease. In some embodiments, the nuclease is an exonuclease. In certain embodiments, the exonuclease is an exonuclease I, exonuclease II, exonuclease III, exonuclease IV, exonuclease V, exonuclease VI, exonuclease VII, or exonuclease VIII. In some embodiments, the nuclease is a catalytically-active Cas RNP or Argonaute RNP. In some embodiments, the catalytically-active Cas RNP or Argonaute RNP comprises a guide polynucleotide that is complementary to an adapter sequence. In some embodiments, the Cas endonuclease is a Cas9 endonuclease. In some embodiments, the Cas endonuclease is a Cpf1, C2c1, C2c3, C2c2, CasX, or CasY endonuclease. In some embodiments, the catalytically-active is an Argonaute protein (e.g., an Argonaute endonuclease). In some embodiments, the Argonaute protein is a CbAgo endonuclease. In certain embodiments, the Argonaute protein is a LrAgo, PfAgo, TtAgo, AaAgo, AfAgo, MjAgo, MpAgo, NgAgo, RsAgo, CpAgo, IbAgo, KmAgo, or SeAgo endonuclease.

In some aspects, the two more samples being normalized (e.g., using the methods described herein) comprise target polynucleotides comprising a first adapter sequence and a second adapter sequence. In some embodiments, the guide polynucleotide targeting region is complementary to the first adapter sequence. In some embodiments, the method further  
5 comprises contacting the sample with a primer encoding a nucleic acid sequence that is complementary to a portion of the second adapter sequence and a polymerase (e.g., a DNA polymerase) in conditions sufficient for primer elongation. As described in the “Kits” section, this primer may be used to elongate a target polynucleotide such that first adapter sequence comprises a double stranded PAM for Cas (e.g., dCas) binding. In some embodiments, the  
10 primer is complementary to the proximal portion of an adapter sequence (e.g., the second adapter sequence). In such embodiments, elongating the primer to produce a double stranded polynucleotide may not produce an elongated strand that is capable of being sequenced by next-generation sequence because the elongated strand does not comprise the distal portion of the second adapter sequence. This may be advantageous because elongation (e.g., DNA  
15 amplification), can introduce artifacts (e.g., mutations) into polynucleotides that may bias or distort sequencing results.

In some embodiments, the primer comprises a nucleic acid sequence that is complementary to any one of SEQ ID NOs: 28, 30, 32, and 34. In some embodiments, the second adapter sequence is a 3' adapter sequence. In some embodiments, the second adapter  
20 sequence is a 5' adapter sequence.

As used in herein, the singular forms “a,” “an,” and “the” include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to “an antibody” optionally includes a combination of two or more such molecules, and the like.

The use of any and all examples or exemplary language (e.g., “such as”) provided herein,  
25 is intended merely to better illustrate the invention and does not pose a limitation on the scope of the invention unless otherwise claimed.

The terms “may,” “may be,” “can,” and “can be,” and related terms are intended to convey that the subject matter involved is optional (that is, the subject matter is present in some examples and is not present in other examples), not a reference to a capability of the subject  
30 matter or to a probability, unless the context clearly indicates otherwise.

The terms “optional” and “optionally” mean that the subsequently described event, circumstance, or material may or may not occur or be present, and that the description includes instances where the event, circumstance, or material occurs or is present as well as instances where it does not occur or is not present.

5           The use herein of the terms “including,” “comprising,” or “having,” and variations thereof, is meant to encompass the elements listed thereafter and equivalents thereof as well as additional elements. Embodiments recited as “including,” “comprising,” or “having” certain elements are also contemplated as “consisting essentially of” and “consisting of” those certain elements. As used herein, “and/or” refers to and encompasses any and all possible combinations  
10 of one or more of the associated listed items, as well as the lack of combinations where interpreted in the alternative (“or”).

## EXAMPLES

### 15 **Example 1 - Library normalization by Cas9-guide mediated capture of library molecules.**

This Example demonstrates that target polynucleotide concentration can be normalized between samples using a predetermined amount of dCas protein and a gRNA that binds to an adapter sequence of the target polynucleotides.

20           Polynucleotide library (also called a NGS library) normalization before equimolar pooling of multiple different polynucleotide libraries is a challenge in next generation sequencing workflows. While various strategies for normalization exist, these strategies often require additional steps such as qPCR, or special primers or oligo nucleotides, and are often destructive to unused material.

25           The normalization method described herein includes a sample comprising a target polynucleotide library, the concentration of which is variable and unknown, and results in the reduction of the concentration of this library to a predetermined value. This enables the target polynucleotide concentration of several unknown target polynucleotide libraries to be normalized to more similar molar concentrations and then pooled by volume prior to loading on the  
30 sequencer. The general method relies on capturing a known number of library molecules using an adjustable amount of non-cleaving nucleic acid guided nuclease (e.g., a dCas9 or dArgonaute)

attached to magnetic beads (FIG. 1). Excess library molecules are removed by pipette removal of unbound material after magnetic bead separation from the initial binding incubation and subsequent buffer washes. After washes, the magnetic beads from each sample are resuspended in the same amount of buffer. Captured DNA can then be obtained by denaturing the non-cleaving nucleic acid guided nuclease. The resulting solutions are then pooled in equal volume with other libraries prepared in the same manner prior to sequencing.

In this Example, a magnetic bead-compatible form of a nucleic acid guided nuclease (NAGN), capable of specifically binding the target determined by the guide nucleic acid (e.g. guide RNA or DNA) but not cleaving the target (e.g. dCas9 D10A/H840A double mutant), was pre-loaded with Cas guide RNAs that are specific to the adapters which are common to the NGS library polynucleotide (e.g. ILLUMINA i5/i7 sequences). Guide RNA Sequences AA (corresponding with SEQ ID NO: 1) and AB (corresponding with SEQ ID NO: 2) were tested and AA showed higher efficiency. In this case the guide RNA sequences target the i5 region of the ILLUMINA Tru-Seq adapter.

Guide RNA Sequence AA: /AltR1/rArG rArUrC rGrGrA rArGrA rGrCrG rUrCrG rUrGrU rGrUrU rUrUrA rGrArG rCrUrA rUrGrC rU/AltR2/

Guide RNA Sequence AB: /AltR1/rGrA rUrCrG rGrArA rGrArG rCrGrU rCrGrU rGrUrA rGrUrU rUrUrA rGrArG rCrUrA rUrGrC rU/AltR2/

DNA containing the i5 adapter sequence (target) was normalized using a dCas-Nickel bead pull down system with guide RNA designed to target the i5 adapter sequence. Two samples, a first containing 88 femtomoles of target DNA and second containing 44 femtomoles of target DNA were normalized using with 125 femtomoles of dCas enzyme (FIG. 2). After normalization, the first and second samples were normalized to 4.3 femtomoles and 4.2 femtomoles of target DNA, respectively.

Normalization was further demonstrated by normalizing a dilution series of ILLUMINA DNA sequencing libraries comprising 4 nM, 6 nM, 8 nM and 10 nM of input DNA could be normalized to 4.68 nM, 5.4 nM, 5.55 nM, and 5.89 nM, respectively. (FIG. 5). These results further demonstrate that this method can be used to normalize the concentration of target polynucleotides in samples where the samples have more than a 2-fold difference in starting concentration. Results also showed that the capture of target-containing DNA molecules was driven specifically by the presence of dCas9 RNP (FIG. 4). Additionally, the amount of target

DNA extracted per sample using the normalization method could be tuned in a linear manner by increasing the dCas9 RNP (dCas9 + Cas gRNA) concentration (FIG. 3). Thus, one may be able to predict the amount of target polynucleotides that will be extracted from sample based on the amount of the dCas9 RNP used.

5

## METHODS

### 1. Form the ribonucleoprotein (RNP) binding complex

Combine 1uM guide RNA, 1uM dCas9 enzyme (6His-tagged), and 1X Cas9 dilution buffer and allow the reaction to incubate at room temperature for 5-10 minutes. This creates the RNP specific to the i5 sequence of the library adapter and allows specific targeting of library molecules with precise amounts of dCas9.

### 2. Perform DNA binding reaction.

Combine 10X Cas9 Reaction Buffer, a desired nM of DNA Library, 1uM RNP complex, and dilute the 10X Cas9 Reaction Buffer to 1x using milliQ water. Incubate the reaction for 1hr at 37°C. This results in dCas9 bound to library molecules in a tight and specific manner, allowing the subsequent capture using Nickel magnetic beads e.g., HisPur (Thermo Scientific).

### 3. Bead pulldown

To each DNA binding reaction add HisPur magnetic beads and 1X Cas9 reaction buffer and incubate for 15 minutes. Remove supernatant and wash samples twice with 1X Cas Reaction buffer to remove unbound DNA.

### 4. Proteinase K digestion (elution)

Incubate samples with Proteinase K for 10 minutes at 56°C to release (elute) the bound DNA from the dCas enzyme.

### 5. Quantification

Flowthrough and final eluted library were quantified using library-specific primers via qPCR.

30

Step 1.1 PCR-free library normalization by Cas9-guide mediated capture of library molecules.

5           Optionally, this method may include an additional step (step 1.1) for producing a double stranded PAM site on a single stranded target polynucleotide without requiring PCR amplification (FIG. 7A-D). PCR amplification is a popular strategy in the construction of NGS libraries, in particular because it allows for NGS sequencing on small amounts of sample.

          However, PCR amplification can be problematic for a number of NGS applications. For  
10       example, PCR amplification can introduce GC bias, which interferes with data analysis, such as the identification of novel single-nucleotide polymorphisms (SNPs). To allow library normalization as described in Example 1 while avoiding PCR amplification, the workflow described in this Example introduces a step of library molecule denaturation and partial extension. Step 1.1 is performed before the RNP binding complex is combined with the target  
15       nucleotides.

          This step begins with a sample having target polynucleotide comprising a 3' and a 5' adapter sequence (FIG. 7A). The reverse complement of the PAM site is encoded on the 5' adapter. A partial primer—which is complementary to a portion of the 3' adapter sequence such that when the partial primer is extended it does produce a reverse complement of the entire 3' adapter sequence—is contacted with the target polynucleotide (FIG. 7B) in conditions such that  
20       primer extension can occur (FIG. 7C), e.g., the conditions include the presence of a DNA polymerase. This produces a double stranded PAM site that can be bound an dCas protein used in the normalization method, as described herein (FIG. 7D). However, the elongated primer does not comprise a complete 3' adapter and therefore is not sequenced during next generation  
25       sequences. Thus, sequencing results are not biased by amplified DNA (e.g., the elongated primer).

## METHODS

30       In order to produce a partial second strand which will enable, for example, an Illumina PCR-free library to be normalized the following approach can be used:

Combine a ligated, unamplified library with an excess of a primer specific to the i7 adapter sequence, for example: 5'GTGACTGGAGTTCAGACGTGT'3 (SEQ ID NO: 49). This primer binds to the proximal part of the i7 adapter and crucially does not include the flow cell binding sequences from the distal adapter part. After primer binding, extend the primer using a polymerase in the presence of dNTPs. The resulting extended strand will not be capable of binding to the flow cell and cluster and is thus inert. Following extension, subject the library to the normalization workflows as described in this specification.

Advantages of this method:

- Only the original PCR-free molecules will be capable of clustering.
- This method ensures that only fully ligated library molecules (molecules which have a 5' and 3' adapter ligated to the insert) will be targeted by the dCas9 (or other nucleic acid guided nuclease) in the subsequent normalization step.

Detailed protocol.

1. Combine a 100 femtomoles of ligated library with 500 femtomoles of i7 complement primer ( 5'GTGACTGGAGTTCAGACGTGT'3 (SEQ ID NO: 49)) in the presence of 1X PCR buffer (Tris-HCl pH 8: 20 mM, Magnesium chloride (MgCl<sub>2</sub>): 2 mM, Potassium chloride (KCl): 50 mM, dNTPs (each): 200 μM together with 1 unit of Taq polymerase. Any other polymerase can be used here, such as Bst, KOD etc.
2. Subject the sample to a single denaturation-extension step: 95C for 30s followed by 60C for 1 minute.
3. Optionally, purify the extended duplex using SPRI beads (Ampure XP or equivalent) and subject the extended library to normalization using the methods described in the examples.
4. Only the original ligated library molecule will be capable of clustering on NGS sequencers as it contains both 5' and 3' adapters.

**Example 2 – Normalizing Samples for NGS with Catalytically-Inactive Cas9 Protein Binding and Exonuclease Digestion**

A NGS library is provided for which the exact concentration of nucleic acids is unknown  
5 but that is at least 20 nM (20 fmol/uL). The nucleic acids in the library have an adapter sequence  
(e.g., ILLUMINA P5/P7 sequence and a shared Y-adapter sequence (e.g., 13 bp)) on the first and  
second end of the nucleic acid. The library is combined with a predetermined amount (80 fmol) a  
catalytically-inactive Cas9 protein having D10A and H840A substitutions. The catalytically-  
inactive Cas9 protein is preloaded with two different guide RNAs or guide DNAs that are each  
10 specific to one of the adapter sequences present on the nucleic acids in the NGS library (e.g., the  
P5 sequence, the P7 sequence, or the shared Y-adapter sequence). Alternatively, the  
catalytically-inactive Cas9 protein is preloaded with a single guide RNA or guide DNA that is  
specific to the Y-adapter sequence present on the nucleic acids in the NGS library or a single  
guide RNA or guide DNA specific to another sequence that is the same on both ends of the  
15 nucleic acids in the NGS library). The reaction mixture is incubated under conditions that  
promote the specific and strong (guide-directed) association of the catalytically-inactive Cas9  
protein with the adapter sequence. This incubation results in a specific number of library  
molecules (80 fmol) being bound by the catalytically-inactive Cas9 protein at their termini. A  
non-targeted nuclease, exonuclease III, is added to the reaction mixture under conditions which  
20 maintain a specific binding of the catalytically-inactive Cas9 protein with the adapter sequence  
but that also allow for the non-targeted exonuclease activity. Only the specific number of library  
molecules (80 fmol) bound by the catalytically-inactive Cas9 protein at their termini remain in  
the resulting NGS library, with the remaining unprotected molecules partially or fully digested  
by the nuclease.

25 These steps are performed simultaneously or sequentially on one or more additional NGS  
libraries for which the exact concentration of nucleic acids is unknown but that is at least 20 nM  
(20 fmol/uL). The resulting NGS libraries (which now have concentrations that are more similar  
between the NGS libraries than the starting concentrations) are pooled, and a purification is  
performed to remove the catalytically-inactive Cas9 protein from the nucleic acids. The nucleic  
30 acids in the libraries are sequenced if at the desired concentration. If the nucleic acids in the  
library are not yet at the desired concentration, they are amplified and then sequenced.

**Example 3 – Normalizing Samples for NGS with Catalytically-Inactive CbAgo Protein Binding and Exonuclease Digestion**

5           A NGS library is provided for which the exact concentration of nucleic acids is unknown but that is at least 20 nM (20 fmol/uL). The nucleic acids in the library have an adapter sequence (e.g., ILLUMINA P5/P7 sequence and a shared Y-adapter sequence (e.g., 13 bp)) on the first and second end of the nucleic acid. The library is combined with a predetermined amount (80 fmol) of a catalytically-inactive CbAgo protein. The catalytically-inactive CbAgo protein is preloaded  
10           with one or more guide DNAs that are specific to the adapter sequence present on the nucleic acids in the NGS library (e.g., the P5 sequence, the P7 sequence, or the shared Y-adapter sequence). The reaction mixture is incubated under conditions that promote the specific and strong (guide-directed) association of the catalytically-inactive CbAgo protein with the adapter sequence. This incubation results in a specific number of library molecules (80 fmol) being  
15           bound by the catalytically-inactive CbAgo protein at their termini. A non-targeted nuclease, exonuclease III is added to the reaction mixture under conditions which maintain a tight binding of the catalytically-inactive CbAgo protein with the adapter sequence but that also allow for the non-targeted exonuclease activity. Only the specific number of library molecules (80 fmol) bound by the catalytically-inactive CbAgo protein at their termini remain in the resulting NGS  
20           library.

          These steps are performed simultaneously or sequentially on one or more additional NGS libraries for which the exact concentration of nucleic acids is unknown but that is at least 20 nM (20 fmol/uL). The resulting NGS libraries (which now have concentrations that are more similar between the NGS libraries than the starting concentrations) are pooled, and a purification is  
25           performed to remove the catalytically-inactive CbAgo protein from the nucleic acids. The nucleic acids in the libraries are sequenced if at the desired concentration. If the nucleic acids in the library are not yet at the desired concentration, they are amplified and then sequenced.

**Example 4 – Alternate Methods for Normalizing Samples for NGS Using Other Nucleases**

30

The normalization methods are performed as described in Example 2 or 3, with the exception that the non-targeted nuclease, exonuclease III, is replaced with exonuclease I, a restriction enzyme, or a sequence-specific nuclease. For example, the nuclease is a 5' to 3' exonuclease or a 3' to 5' exonuclease, a single strand-specific nuclease, a double strand-specific  
5 nuclease, or a mixture thereof. Alternatively, the non-targeted nuclease is replaced with a catalytically-active nucleic acid guided nuclease (e.g., Cas9, CbAgo). In some embodiments of the methods, the non-targeted nuclease is replaced with a nucleic acid guided nickase (e.g., Cas9 nickase). In certain embodiments, the nucleic acid guided nickase is a Cas9 nickase that has a D10A mutation or an H840A mutation. In some embodiments, the non-targeted nuclease is  
10 replaced with a transcription activator-like effector nuclease (TALEN) or TALE nickase. In certain embodiments, the TALEN or TALE nickase targets the same adapter sequence as the nucleic acid binding protein (e.g., dCas protein, dAgo, etc.) used in the method.

**Example 5 – Normalizing Samples for NGS with Catalytically-Inactive Cas9 Protein Binding and Cas9 Nickase Digestion**  
15

A NGS library is provided for which the exact concentration of nucleic acids is unknown but that is at least 20 nM (20 fmol/uL). The nucleic acids in the library have an adapter sequence (e.g., ILLUMINA P5/P7 sequence and a shared Y-adapter sequence (e.g., 13 bp)) on the first and  
20 second end of the nucleic acid. The library is combined with a predetermined amount (80 fmol) of a nucleic acid guided binding protein to create a reaction mixture. The nucleic acid guided binding protein is a catalytically-inactive Cas9 protein having D10A and H840A substitutions. The catalytically-inactive Cas9 protein is preloaded with a single guide RNA or guide DNA that is specific to the shared Y-adapter sequence present on the nucleic acids in the NGS library. The  
25 reaction mixture is incubated under conditions that promote the specific and strong (guide-directed) association of the catalytically-inactive Cas9 protein with the adapter sequence. This incubation results in a specific number of library molecules (80 fmol) being bound by the catalytically-inactive Cas9 protein at their termini. A Cas9 nickase is added to the reaction mixture under conditions which maintain a specific binding of the catalytically-inactive Cas9  
30 protein with the adapter sequence but that also allow for the targeted Cas9 nickase activity. Only the specific number of library molecules (80 fmol) bound by the catalytically-inactive Cas9

protein at their termini remain in the resulting NGS library, with the remaining unprotected molecules partially or fully digested by the Cas9 nickase.

These steps are performed simultaneously or sequentially on one or more additional NGS libraries for which the exact concentration of nucleic acids is unknown but that is at least 20 nM (20 fmol/uL). The resulting NGS libraries (which now have concentrations that are more similar between the NGS libraries than the starting concentrations) are pooled, and a purification is performed to remove the proteins from the nucleic acids. The nucleic acids in the libraries are sequenced if at the desired concentration. If the nucleic acids in the library are not yet at the desired concentration, they may be amplified and then sequenced.

10

#### **Example 6 – Alternate Methods for Normalizing Samples for NGS Including Amplification Before Digestion**

A NGS library is provided for which the exact concentration of nucleic acids is unknown. The nucleic acids in the library have an adapter sequence (e.g., ILLUMINA P5/P7 sequence and a shared Y-adapter sequence (e.g., 13 bp)) on the first and second end of the nucleic acid. An amplification is performed with primers that bind to the adapter sequence, producing an amplified library.

The amplified library is combined with a predetermined amount (80 fmol) of a nucleic acid guided binding protein to create a reaction mixture. The nucleic acid guided binding protein is a dCas9, a catalytically-inactive Cas9 protein having D10A and H840A substitutions. The catalytically-inactive Cas9 protein is preloaded with a guide RNA or guide DNA that is specific to the adapter sequence present on the nucleic acids in the NGS library (e.g., the P5 sequence, the P7 sequence, or the shared Y-adapter sequence). The reaction mixture is incubated under conditions that promote the specific and strong (guide-directed) association of the catalytically-inactive Cas9 protein with the adapter sequence. This incubation results in a specific number of library molecules being bound by the catalytically-inactive Cas9 protein at their termini. A non-targeted nuclease, exonuclease III, is added to the reaction mixture under conditions which maintain a tight binding of the catalytically-inactive Cas9 protein with the adapter sequence but that also allow for the non-targeted exonuclease activity. Only the specific number of library

30

molecules bound by the catalytically-inactive Cas9 protein at their termini remain in the resulting NGS library.

These steps are performed simultaneously or sequentially on one or more additional NGS libraries. That is, the dCas9 is bound to individual libraries, the protected libraries are pooled, the pool of protected libraries is exposed to an excess of nuclease (e.g., exonuclease III, exonuclease I, restriction enzyme, sequence-specific nuclease, 5' to 3' exonuclease, 3' to 5' exonuclease, single strand-specific nuclease, double strand-specific nuclease, catalytically-active nucleic acid guided nuclease (e.g., Cas9, CbAgo), nucleic acid guided nickase (e.g., Cas9 nickase), transcription activator-like effector nuclease (TALEN), or TALE nickase, or a mixture thereof). The resulting NGS libraries (which now have concentrations that are more similar between the NGS libraries than the starting concentrations) are pooled, and a standard SPRI purification is performed on the pooled NGS libraries. The nucleic acids in the libraries are then sequenced.

#### **Example 7 – Normalization of sequencing libraries post amplification**

Typical Illumina sequencer library preparation chemistries target a final amplified library molar amount between 200 and 1,000 femtomoles (fmols). To demonstrate the capability of library normalization across a range of input, an Illumina TruSeq adapter-ligated double stranded DNA (dsDNA) library was amplified to produce a high concentration DNA input. DNA input amounts of 50, 500, 750, or 1,000 fmols of this DNA was used as input into two different normalization reactions containing either 250 fmols or 300 fmols of ribonucleoprotein (RNP, e.g., dCas9-guide RNA complex). The resulting normalized library was quantified using an HSD1000 DNA tape on an Agilent TapeStation with a region of 100-1,000 base pairs. Both 250 fmols and 300 fmols of RNP normalized the range of DNA library input, with the resulting amount of normalized output being dependent on the RNP amount. Between the low DNA input of 50 fmols and the high DNA input of 1,000 fmols, there was a 20-fold difference in DNA amount. After normalization with 250 fmols of RNP, there was only a 1.5 fold difference, and after normalization with 300 fmols of RNP, there was only a 1.3 fold difference. Between all samples normalized with 250 fmols of RNP, there was an average output of 37 fmols, while a normalization with 300 fmols of RNP resulted in an average output of 56 fmols, demonstrating that the amount of library retained during normalization can be modulated by the amount of RNP used (FIG. 8).

**Example 8 – Normalization of sequencing libraries pre targeted capture**

Targeted capture/enrichment panels typically require high masses of amplified library, often ranging from 100ng-1,000ng of library molecules. Though dependent on library size, such a mass range broadly corresponds to 300-2,000 fmols of DNA. To demonstrate the capabilities of normalization to achieve this final output range, an Illumina TruSeq adapter-ligated dsDNA library was amplified to produce a high concentration of DNA input and 3,400 fmols of this DNA was used as input. Normalization was performed with 2,000, 6,000, 10,000, or 14,000 fmols of RNP. The output DNA was analyzed using a D1000 tape on an Agilent Tapestation with a region setting of 180-1,000 base pairs. Increasing RNP resulted in the capture of more DNA library. By increasing RNP to 6,000 fmols, about 2,000 fmols of library was captured, supporting normalization for use upstream of target enrichment and similar applications (FIG. 9).

**Example 9. Normalization using a single guide RNA (sgRNA) in the RNP**

Using a single guide RNA (sgRNA) as opposed to a separate CRISPR RNA (crRNA) and trans-activating CRISPR RNA (tracrRNA) offers a simpler format for performing normalization. To demonstrate similar performance to crRNA/tracrRNA RNPs, an Illumina TruSeq adapter-ligated dsDNA library was amplified to produce a high concentration of DNA input, and 500 fmols of this DNA was used. Normalization was performed with 250, 600, or 1,000 fmols of RNP containing either a crRNA/tracrRNA (duplex guide) or a sgRNA. The sgRNA sequence was:

5'- mA\*mG\*mA\* rUrCrG rGrArA rGrArG rCrGrU rCrGrU rGrUrG rUrUrU rUrArG rArGrCrUrArG rArArA rUrArG rCrArA rGrUrU rArArA rArUrA rArGrG rCrUrA rGrUrC rCrGrU rUrArU rCrArA rCrUrU rGrArA rArArA rGrUrG rGrCrA rCrCrG rArGrU rCrGrG rUrGrC mU\*mU\*mU\*rU-3' (SEQ ID NO: 48) (mA/mG/mU are 2'-O-methyl RNA bases; \* indicates phosphorothioate linkages).

The crRNA sequence was:

5'-/AltR1/rGrCrA rCrGrG rArGrA rCrGrG rArUrG rUrUrA rUrUrGrUrUrU rUrArG rArGrC rUrArUrGrCrU/AltR2/-3' (SEQ ID NO: 50) (/AltR1/ and /AltR2/ are proprietary IDT modifications).

The output DNA was analyzed using a D1000 tape on an Agilent TapeStation with a region setting of 180-1,000 base pairs. RNP containing a sgRNA was found to perform similarly to RNP containing a crRNA/tracrRNA. The sgRNA containing RNPs retained about 20% less DNA than RNPs containing crRNA/tracrRNA. However, increasing fmols of sgRNA RNP  
5 resulted in more captured DNA, allowing for ranges demonstrated with crRNA/tracrRNA to be achieved with an sgRNA (FIG. 10).

The examples are provided to illustrate the disclosure but not to limit its scope. Other variations of the disclosure will be readily apparent to one of ordinary skill in the art and are  
10 encompassed by the appended claims.

## CLAIMS

1. A method for normalizing the concentration of target polynucleotides between at least  
5 two samples each comprising target polynucleotides, the method comprising, for each sample of  
the at least two samples:
- (i) obtaining the sample, wherein the target polynucleotides of the sample comprise  
an adapter sequence;
  - (ii) producing a solution, the producing comprising combining (a) the sample, (b) a  
10 predetermined concentration of a guide polynucleotide comprising a targeting region that  
is complementary to the adapter sequence of the target polynucleotides of the sample;  
and (c) a predetermined concentration of a dCas or a dArgonaute comprising an affinity  
tag, wherein the predetermined concentration of the dCas or the dArgonaute is cognate to  
the guide polynucleotide;
  - 15 (iii) contacting the solution with a solid phase comprising an affinity tag binding  
molecule that is capable of binding to the affinity tag;
  - (iv) separating the solution from the solid phase; and
  - (v) extracting the target polynucleotides from the solid phase to normalize the  
concentration of target polynucleotides between the two or more samples.
- 20 2. The method of claim 1, wherein the targeting region is complementary to a static region  
of an adapter sequence.
3. The method of claim 1 or claim 2, wherein the guide polynucleotide comprises a  
CRISPR-associated protein (Cas) guide RNA (gRNA) polynucleotide or an Argonaute guide  
polynucleotide.
- 25 4. The method of claim 2, wherein the static region is a static region of a next-generation  
sequencing adapter sequence.

5. The method of any one of claims 1-4, wherein the targeting region is complementary to an adapter sequence of any one of SEQ ID NOs: 27-36.
6. The method of any one of claims 1-4, wherein the targeting region is complementary to an adapter sequence of any one of SEQ ID NOs: 27-28.
- 5 7. The method of any one of claims 1-4, wherein the targeting region is complementary to an adapter sequence of any one of SEQ ID NOs: 29-30.
8. The method of any one of claims 1-4, wherein the targeting region is complementary to an adapter sequence of any one of SEQ ID NOs: 31-32.
9. The method of any one of claims 1-4, wherein the targeting region is complementary to  
10 an adapter sequence any one of SEQ ID NOs: 33-34.
10. The method of any one of claims 1-5, wherein the targeting region is complementary to an adapter sequence of SEQ ID NO: 35.
11. The method of any one of claims 1-4, wherein the targeting region is complementary to an adapter sequence of SEQ ID NO: 36.
- 15 12. The method of any one of claims 1-11, wherein the Cas gRNA polynucleotide targeting region is a homology region.
13. The method of claim 12, wherein the Cas gRNA polynucleotide comprises a homology region of any one of SEQ ID NOs: 1-26.
14. The method of claim 13, wherein the Cas gRNA polynucleotide comprises a homology  
20 region of SEQ ID NO: 1.
15. The method of claim 1, wherein the Argonaute guide polynucleotide is an siRNA, miRNA, piRNA, shRNA or siDNA.
16. The method of any one of claims 1-15, wherein the guide polynucleotide is a DNA polynucleotide.

17. The method of any one of claims 1-16, wherein the guide polynucleotide is a RNA polynucleotide.
18. The method of claim 16 or claim 17, wherein the guide polynucleotide comprises a modified nucleic acid.
- 5 19. The method of claim 18, wherein the modified nucleic acid comprises 2'F RNA, 2'OMe RNA, and/or a phosphorothioate bonds (PS).
20. The method of claim 19, wherein the guide polynucleotide is a Cas gRNA polynucleotide and the Cas gRNA polynucleotide does not comprise a modified nucleic acid in a position that interacts with a Cas protein cognate to the gRNA.
- 10 21. The method of any one of claims 1-20, wherein the dCas is a catalytically-inactive Cpf1, C2c1, C2c3, C2c2, CasX, or CasY protein.
22. The method of claim 21, wherein the dCas protein comprises an amino acid sequence of SEQ ID NO: 37.
23. The method of any one of claims 1-11, wherein the dArgonaute is a catalytically-inactive  
15 LrAgo, PfAgo, TtAgo, AaAgo, AfAgo, MjAgo, MpAgo, NgAgo, RsAgo, CpAgo, IbAgo, KmAgo, or SeAgo.
24. The method of any one of claims 1-23, wherein the affinity tag binding molecule comprises Ni<sup>2+</sup> and the affinity tag comprises a His Tag.
25. The method of any one of claims 1-23, wherein the affinity tag binding molecule  
20 comprises biotin and the affinity tag comprises avidin.
26. The method of any one of claims 1-23, wherein the affinity tag binding molecule comprises an anti-myc antibody and the affinity tag comprises a myc tag.
27. The method of any one of claims 1-26, wherein the affinity tag and corresponding affinity tag binding molecule are selected from Table 1.

28. The method of any one of claims 1-27, wherein the solid phase comprises magnetic beads.
29. The method of any one of claims 1-28, wherein separating the solution from the solid phase comprises immobilizing the solid phase and washing the solid phase.
- 5 30. The method of any one of claims 1-29, wherein extracting the target polynucleotides from the solid phase comprises combining the solid phase and protease.
31. The method of any one of claims 1-29, wherein extracting the target polynucleotides from the solid phase comprises combining the solid phase and proteinase K in a solution sufficient for proteinase K activity.
- 10 32. The method of claim 31, wherein proteinase K digests the dCas or dArgonaute that is bound to the solid phase thereby extracting the target polynucleotides.
33. The method of any one of claims 1-32, wherein steps (i)-(v) are performed in sequential order.
34. The method of any one of claims 1-33, wherein normalizing comprises making the  
15 concentration of the of target polynucleotides between the at least two samples within 15% of one another after normalization.
35. A method for normalizing the concentration of target polynucleotides between at least two samples each comprising target polynucleotides, the method comprising, for each sample of the at least two samples:
- 20 (i) obtaining the sample, wherein the target polynucleotides of the sample comprise an adapter sequence;
- (ii) producing a solution, the producing comprising combining (a) the sample, (b) a predetermined concentration of a guide polynucleotide comprising a targeting region that is complementary to the adapter sequence of the target polynucleotides of the sample, and  
25 (c) a predetermined concentration of a dCas or a dArgonaute comprising an affinity tag; and

- (iii) contacting the solution with a pre-determined amount of a catalytically active Cas or Argonaute protein to normalize the concentration of target polynucleotides between the two or more samples.
36. The method of claim 35, wherein the dCas is a catalytically-inactive Cpf1, C2c1, C2c3,  
5 C2c2, CasX, Cas9, or CasY protein.
37. The method of claim 36, wherein the dCas protein comprises an amino acid sequence having at least 95% identity to SEQ ID NO: 37.
38. The method of claim 36, wherein the dCas protein comprises an amino acid sequence of SEQ ID NO: 37.
- 10 39. The method of claim 35, wherein the dArgonaute is a catalytically-inactive LrAgo, PfAgo, TtAgo, AaAgo, AfAgo, MjAgo, MpAgo, NgAgo, RsAgo, CpAgo, IbAgo, KmAgo, or SeAgo.
40. The method of any one of claims 35-39, wherein (ii) comprises binding of dCas or dArgonaute to the adapter sequences of at least some of the target polynucleotides.
- 15 41. The method of claim 40, wherein the catalytically active Cas9 or Argonaute digests target polynucleotides that are not bound by the dCas9 or the dArgonaute.
42. The method of any one of claims 35-40, wherein normalizing comprises making the concentration of the of target polynucleotides between the at least two samples within 15% of one another after normalization.
- 20 43. The method of any one of claims 35-40, wherein normalizing comprises making the concentration of the of target polynucleotides between the at least two samples within 10% of one another after normalization.
44. The method of any one of claims 1-43, wherein the target polynucleotides comprise a first adapter sequence and a second adapter sequence, and wherein the guide polynucleotide  
25 targeting region is complementary to the first adapter sequence.

45. The method of claim 44, further comprising, after (i) and before (ii):
- (a) contacting the sample with a primer encoding a nucleic acid sequence that is complementary to a portion of the second adapter sequence, wherein the portion of the adapter sequence is located in the proximal portion of the adapter sequence; and
  - 5 (b) contacting the sample with a DNA polymerase in conditions sufficient to promote primer extension.
46. The method of claim 45, wherein the portion of the adapter sequence comprises a nucleic acid sequence of any one of SEQ ID NOs: 28, 30, 32, and 34.
- 10 47. The method of claim 45 or claim 46, wherein the second adapter sequence is a 3' adapter sequence.
48. The method of claim 45 or claim 46, wherein the second adapter sequence is a 5' adapter sequence.
- 15 49. A guide polynucleotide comprising a targeting region that is complementary to a static region of an adapter sequence of any one of SEQ ID NOs: 28 and 30, 32, 34-36.
50. The guide polynucleotide of claim 49, wherein the guide polynucleotide comprises a CRISPR-associated protein (Cas) guide RNA (gRNA) polynucleotide or an Argonaute guide
- 20 polynucleotide.
51. The guide polynucleotide of claim 49 or claim 50, wherein the static region is a static region of a next-generation sequencing adapter sequence.
52. The guide polynucleotide of claim 51, wherein the targeting region is complementary to an adapter sequence of SEQ ID NO: 28.
- 25 53. The guide polynucleotide of claim 51, wherein the targeting region is complementary to an adapter sequence of SEQ ID NO: 30.
54. The guide polynucleotide of claim 51, wherein the targeting region is complementary to an adapter sequence of SEQ ID NO: 32.

55. The guide polynucleotide of claim 51, wherein the targeting region is complementary to an adapter sequence of SEQ ID NO: 34.
56. The guide polynucleotide of claim 51, wherein the targeting region is complementary to an adapter sequence of SEQ ID NO: 35.
- 5 57. The guide polynucleotide of claim 51, wherein the targeting region is complementary to an adapter sequence of any one of SEQ ID NOs: 36.
58. The guide polynucleotide of any one of claims 51-57, wherein the Cas gRNA polynucleotide targeting region is a homology region.
59. The guide polynucleotide of claim 58, wherein the Cas gRNA polynucleotide comprises  
10 a homology region of any one of SEQ ID NOs: 1-26.
60. The guide polynucleotide of claim 59, wherein the Cas gRNA polynucleotide comprises a homology region of SEQ ID NO: 1.
61. The guide polynucleotide of any one of claims 50-57, wherein the Argonaute guide polynucleotide is an siRNA, miRNA, piRNA, shRNA or siDNA.
- 15 62. The guide polynucleotide of any one of claims 49-61, wherein the guide polynucleotide is a DNA polynucleotide.
63. The guide polynucleotide of any one of claims 49-61, wherein the guide polynucleotide is a RNA polynucleotide.
64. The guide polynucleotide of claim 62 or claim 63, wherein the guide polynucleotide  
20 comprises a modified nucleic acid.
65. The guide polynucleotide of claim 64, wherein the modified nucleic acid comprises 2'F RNA, 2'Ome RNA, and/or a phosphorothioate bonds (PS).
66. The guide polynucleotide of claim 65, wherein the guide polynucleotide is a Cas gRNA polynucleotide and the Cas gRNA polynucleotide does not comprise a modified nucleic acid in a  
25 position that interacts with a Cas protein cognate to the gRNA.

67. A kit comprising:
- (a) the guide polynucleotide of any one of claims 49-66; and
  - (b) a Cas protein or an Argonaute protein, or a polynucleotide sequence encoding a Cas protein or an Argonaute protein.
- 5 68. The kit of claim 67, wherein the Cas protein is a catalytically-inactive Cas protein (dCas).
69. The kit of claim 67 or claim 68, wherein the Cas protein is a Cas9 protein that is catalytically-inactive (dCas9).
70. The kit of any one of claims 67-69, wherein the Cas9 protein comprises a D10A mutation and a H840A mutation.
- 10 71. The kit of any one of claims 67-68, wherein the Cas protein is a catalytically-inactive Cpf1, C2c1, C2c3, C2c2, CasX, or CasY protein.
72. The kit of any one of claims 68-71, wherein the dCas protein comprises an amino acid sequence having at least 95% identity to SEQ ID NO: 37.
- 15 73. The kit of any one of claims 68-71, wherein the dCas protein comprises an amino acid sequence of SEQ ID NO: 37.
74. The kit of claim 67, wherein the Argonaute protein is catalytically-inactive (dArgonaute).
75. The kit of claim 67, wherein the Argonaute protein is a CbAgo protein that is catalytically-inactive.
- 20 76. The kit of claim 67, wherein the Argonaute protein is a LrAgo, PfAgo, TtAgo, AaAgo, AfAgo, MjAgo, MpAgo, NgAgo, RsAgo, CpAgo, IbAgo, KmAgo, or SeAgo protein that is catalytically-inactive.
77. The kit of any one of claims 74-76, wherein the dArgonaute protein comprises an amino acid sequence having at least 95% identity to any one of SEQ ID NOs: 39-45.

78. The kit of any one of claims 74-76, wherein the dArgonaute protein comprises an amino acid sequence of any one of SEQ ID NOs: 39-45.
79. The kit of any one of claims 67-78, further comprising a catalytically active Cas protein or a catalytically active Argonaute protein.
- 5 80. The kit of claim 79, wherein the catalytically active Cas protein comprises Cpf1, C2c1, C2c3, C2c2, CasX, Cas9, or CasY.
81. The kit of claim 80, wherein the catalytically active Argonaute protein comprises LrAgo, PfAgo, TtAgo, AaAgo, AfAgo, MjAgo, MpAgo, NgAgo, RsAgo, CpAgo, IbAgo, KmAgo, or SeAgo.
- 10 82. The kit of any one of claims 67-81, wherein the guide polynucleotide is capable of binding to the Cas protein or the Argonaute protein to form a ribonucleoprotein complex and the ribonucleoprotein complex is capable of binding to an adapter sequence.
83. The kit of any one of claims 67-82, further comprising a primer that is complementary to a portion of an adapter sequence.
- 15 84. The kit of claim 83, wherein the portion of the adapter sequence is located in the proximal portion of the adapter sequence.
85. The kit of claim 84, wherein the portion of the adapter sequence is located in the proximal portion of the adapter sequence, the adapter sequence comprising a nucleic acid sequence of any one of SEQ ID NOs: 28, 30, 32, and 34.
- 20 86. The kit of any one of claims 83-85, wherein the adapter sequence is a 3' adapter sequence.
87. The kit of any one of claims 83-85, wherein the adapter sequence is a 5' adapter sequence.
88. The kit of any one of claims 83-87, wherein the primer is not complementary to an  
25 adapter sequence that is complementary to the guide polynucleotide targeting region.

89. A reaction mixture comprising:
- (i) a plurality of target polynucleotides, wherein the target polynucleotides comprise an adapter sequence of any one of SEQ ID NOs: 28, 30, 32, and 34-36;
  - 5 (ii) a predetermined concentration of a guide polynucleotide comprising a targeting region that is complementary to the adaptor sequence; and
  - (iii) a predetermined concentration of a dCas protein or a dArgonaute protein.
90. The reaction mixture of claim 89, wherein the predetermined concentration of the guide  
10 polynucleotide, or the dCas protein or the dArgonaute protein is lower than the concentration of target polynucleotide in the reaction mixture.
91. The reaction mixture of claim 89 or claim 90, wherein the dArgonaute protein comprises LrAgo, PfAgo, TtAgo, AaAgo, AfAgo, MjAgo, MpAgo, NgAgo, RsAgo, CpAgo, IbAgo, KmAgo, or SeAgo, or a catalytically inactive variant thereof, and the guide polynucleotides are  
15 cognate to the dArgonaute protein.
92. The reaction mixture of claim 89 or claim 90, wherein the dCas protein comprises catalytically inactive Cpf1, C2c1, C2c3, C2c2, CasX, Cas9, or CasY, or a variant thereof, and the guide polynucleotides are cognate to the dCas protein.
93. The reaction mixture of any one of claims 89-92, wherein the adapter sequence comprises  
20 a polynucleotide sequence of any one of SEQ ID NOs: 28, 30, 32, and 34-36.
94. The reaction mixture of any one of claims 89-93, further comprising a catalytically active Cas protein or a catalytically active Argonaute protein.
95. A ribonucleoprotein (RNP) complex comprising:
- (i) the guide polynucleotide of any one of claims 49-66;
  - 25 (ii) a Cas protein or an Argonaute protein that is cognate to the guide polynucleotide;
- and
- (iii) an adapter sequence,

wherein the targeting region of the guide polynucleotide is complementary to the adapter sequence and wherein the guide polynucleotide is cognate to the Cas protein or the Argonaute protein.

96. The RNP complex of claim 95, wherein the Argonaute protein comprises LrAgo, PfAgo, 5 TtAgo, AaAgo, AfAgo, MjAgo, MpAgo, NgAgo, RsAgo, CpAgo, IbAgo, KmAgo, or SeAgo, or a catalytically inactive variant thereof.

97. The RNP complex of claim 95, wherein the Cas protein comprises Cpf1, C2c1, C2c3, C2c2, CasX, Cas9, or CasY, or a catalytically inactive variant thereof.

98. The RNP complex of any one of claims 95-97, wherein the adapter sequence comprises a 10 polynucleotide sequence of any one of SEQ ID NOs: 28, 30, 32, and 34-36.

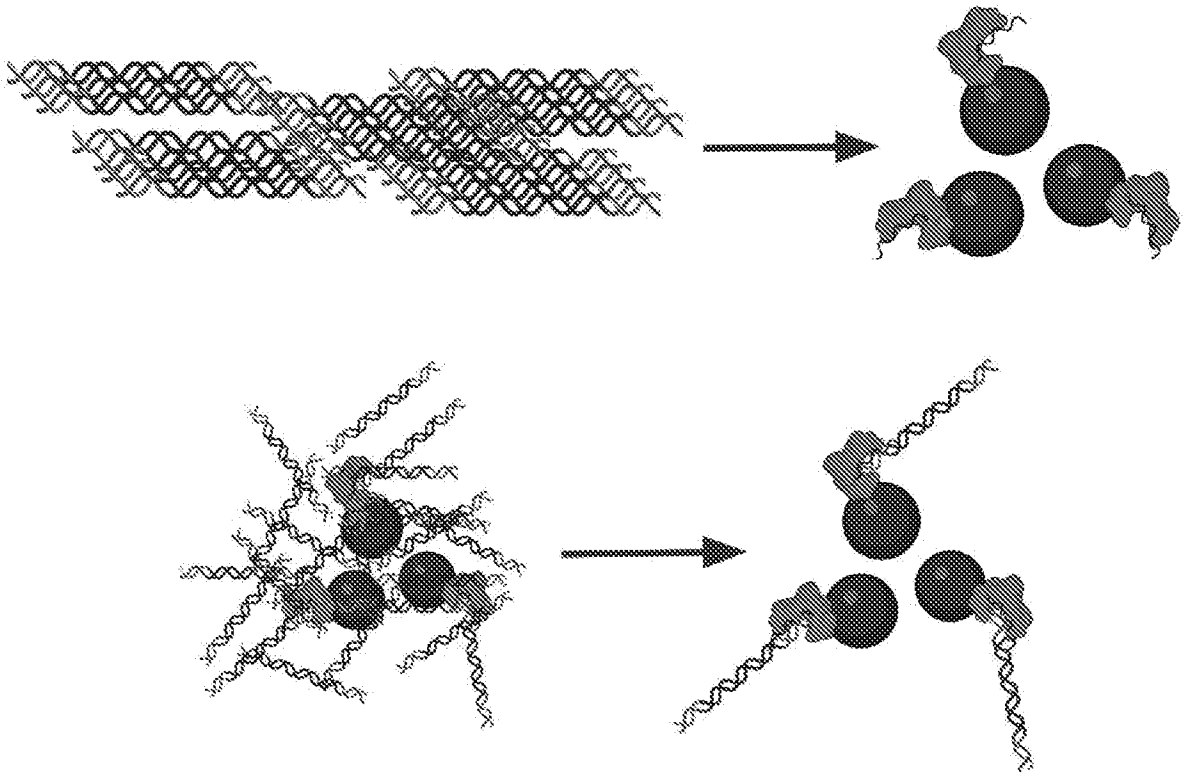


FIG. 1

Enzymatic Normalization with 125 fmoles dCAS

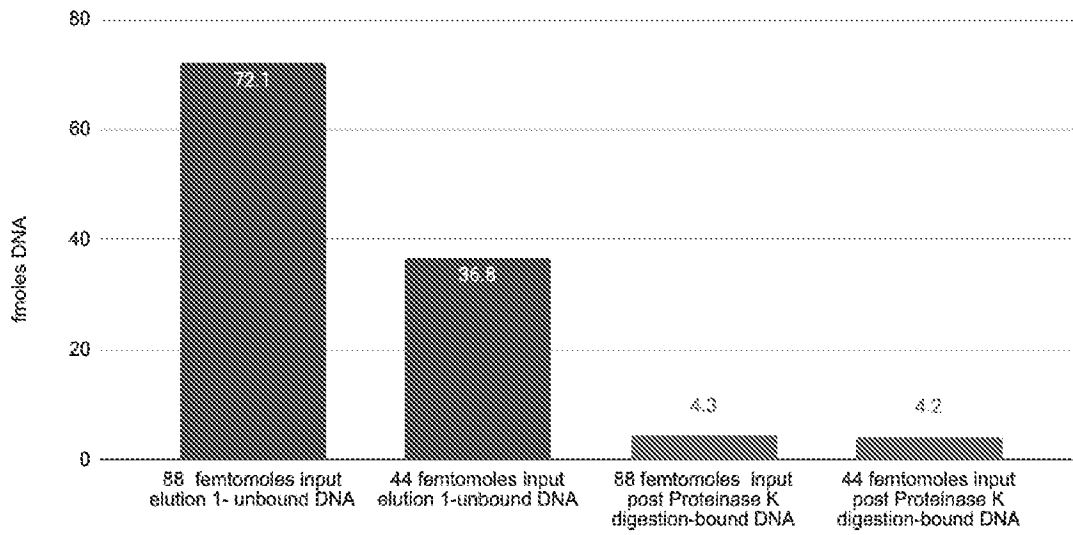


FIG. 2

160 femtomoles DNA vs. dCas enzyme titration, post proteinase K digestion

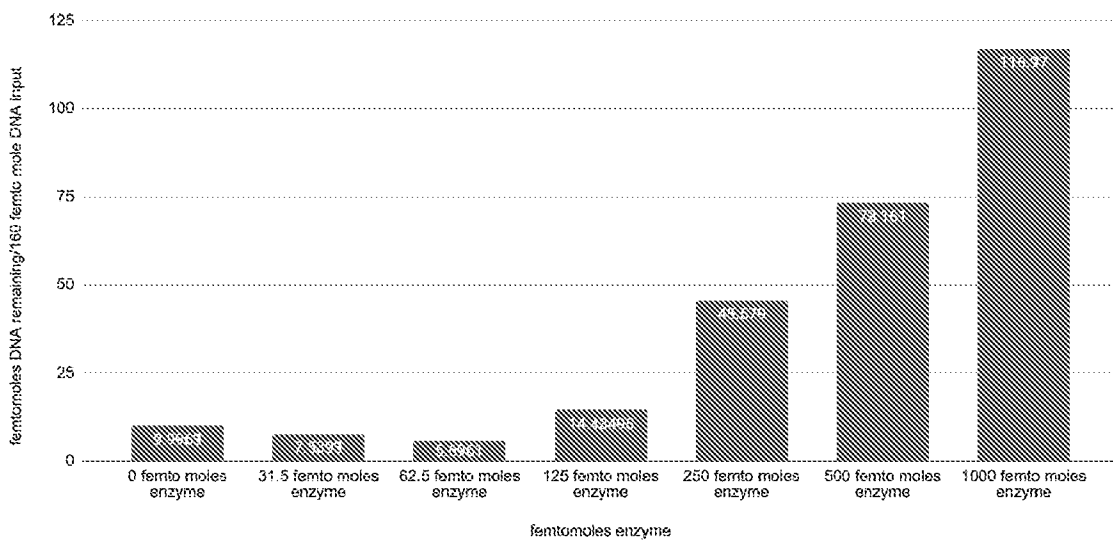


FIG. 3

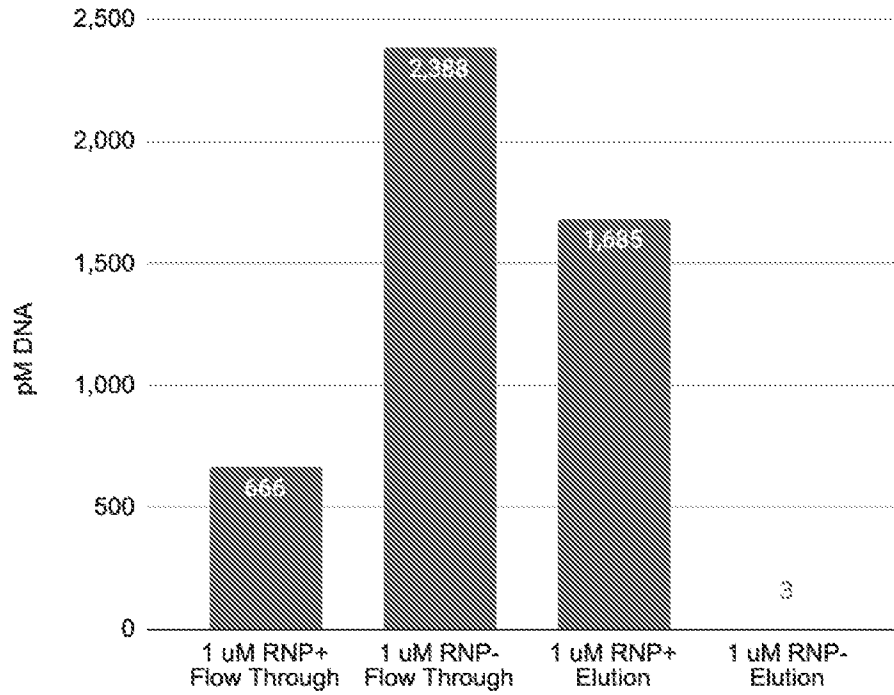


FIG. 4

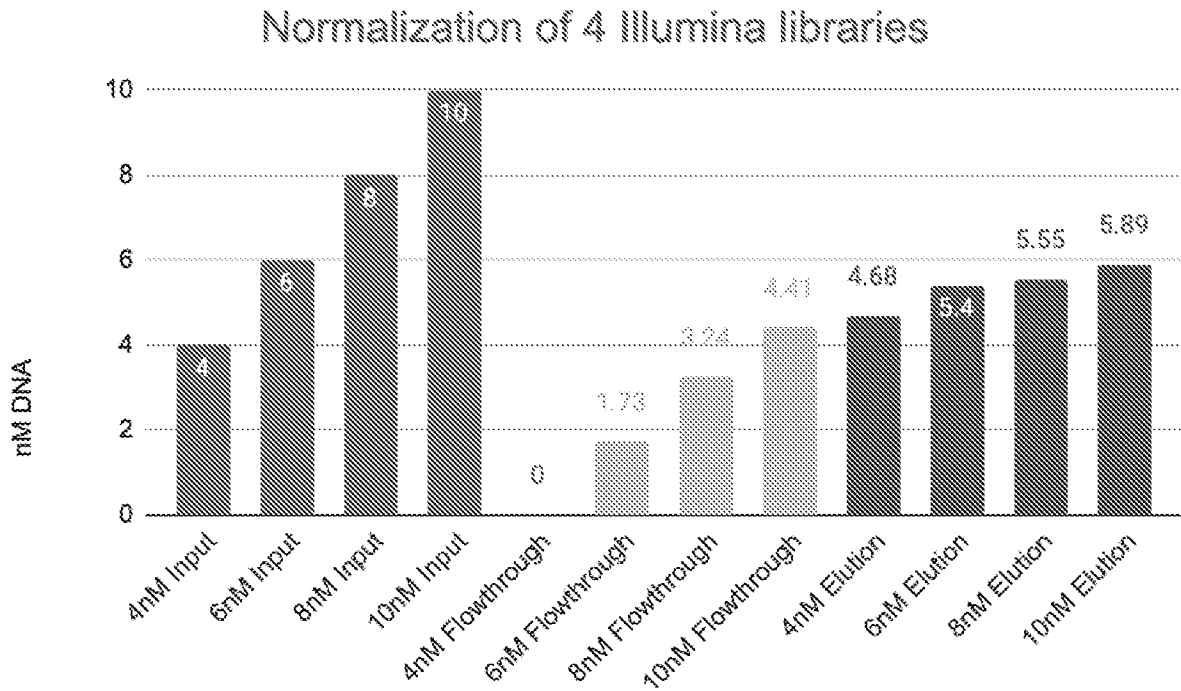


FIG. 5

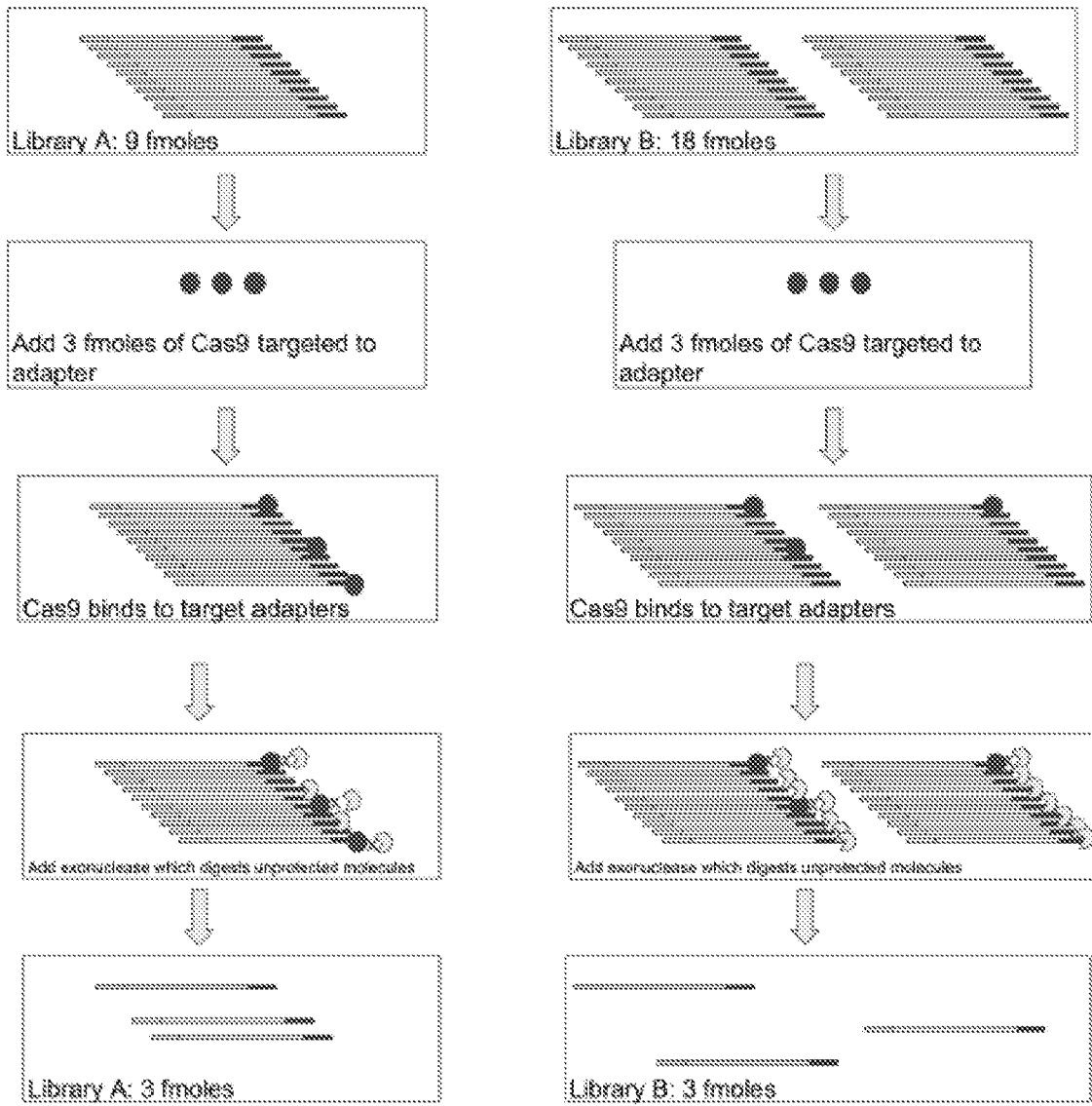


FIG. 6



FIG. 7A

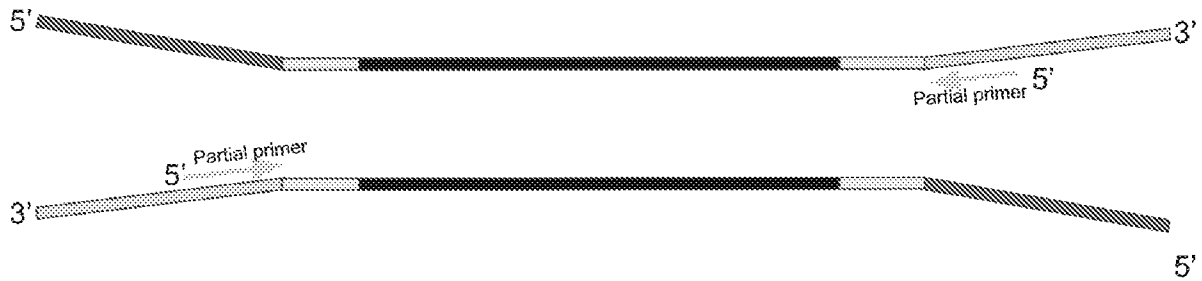


FIG. 7B

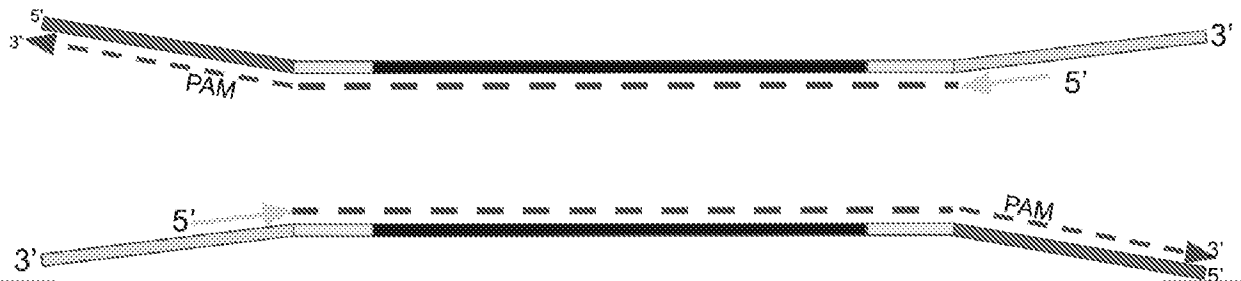


FIG. 7C

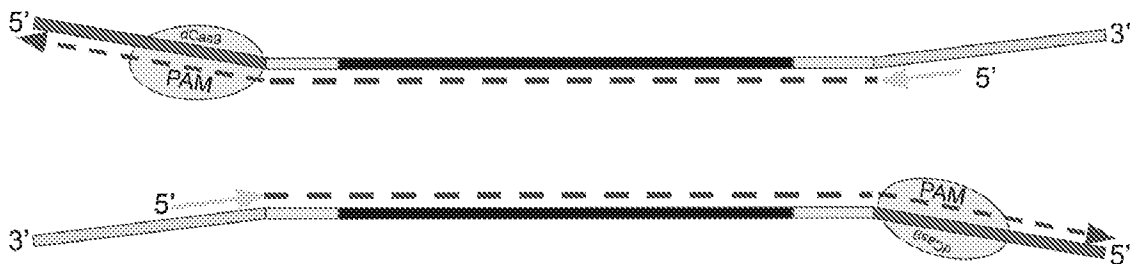


FIG. 7D

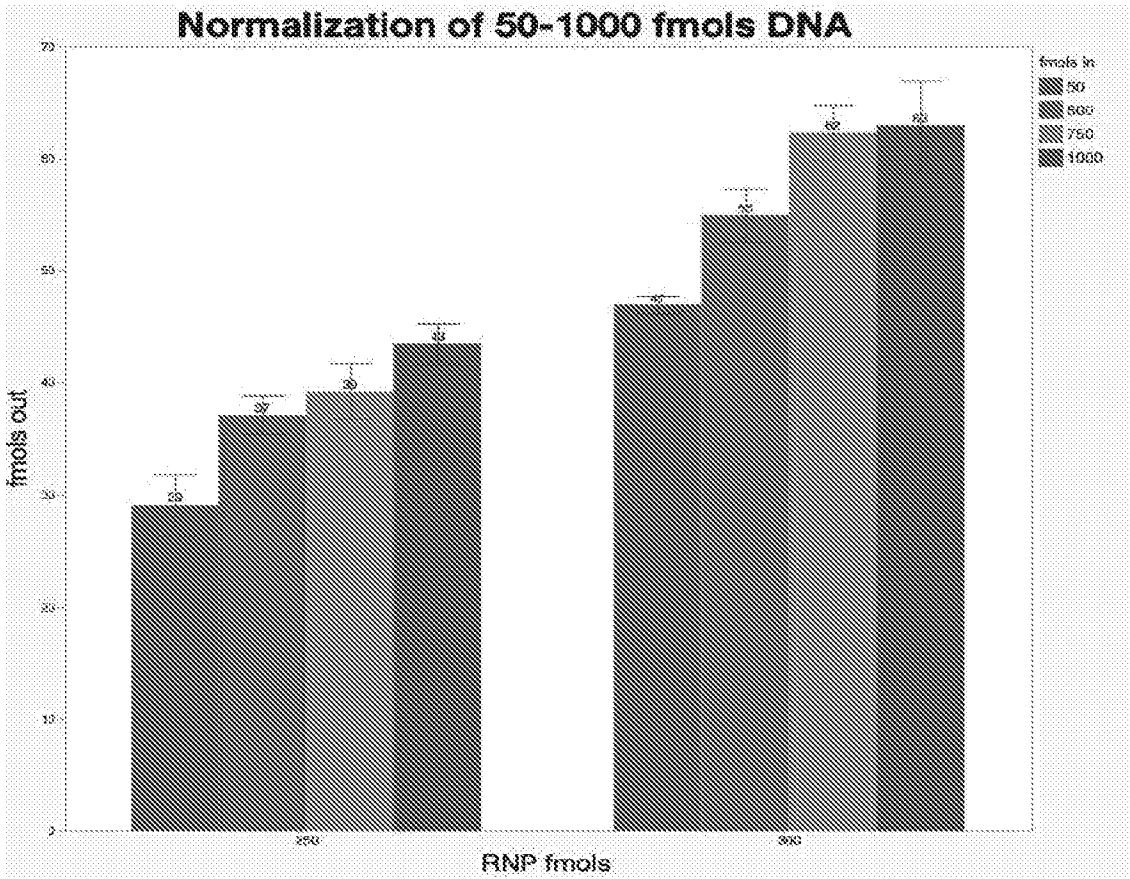


FIG. 8

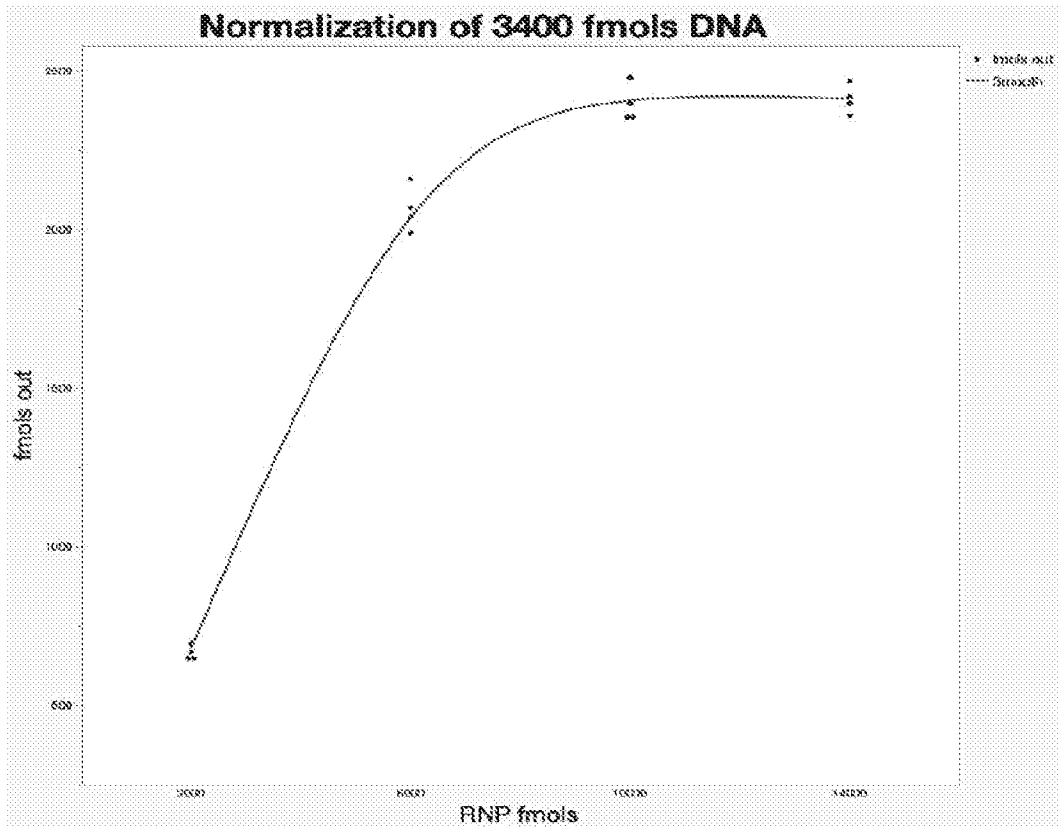


FIG. 9

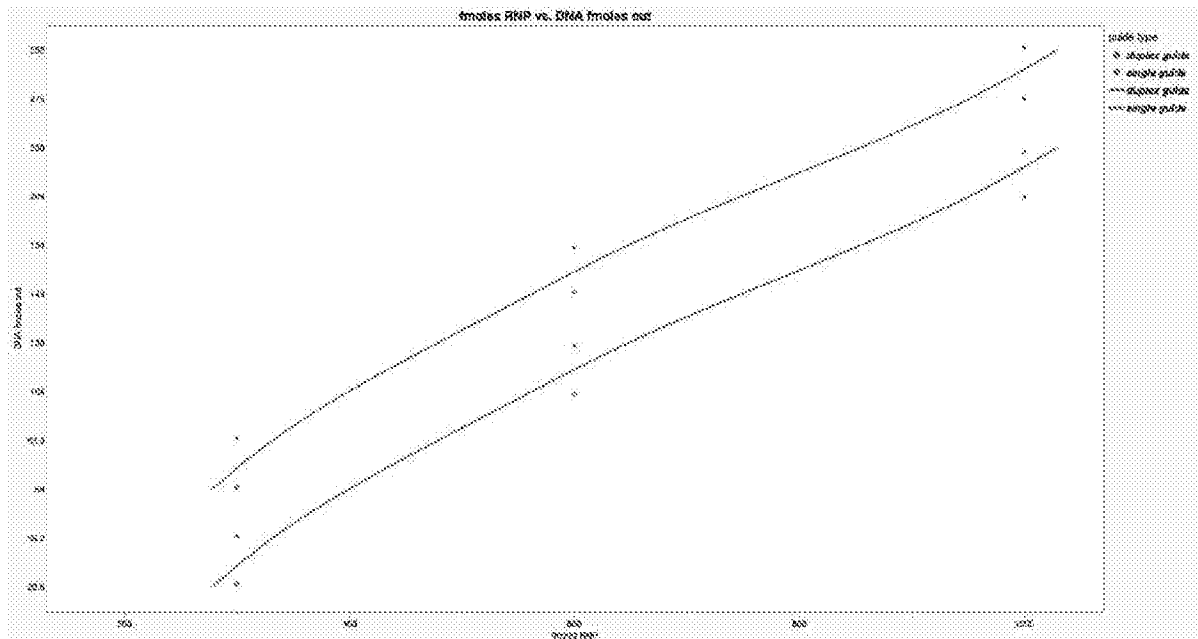


FIG. 10