Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2917638 A1 2015/01/15

(21) **2 917 638**

(12) **DEMANDE DE BREVET CANADIEN**
**CANADIAN PATENT APPLICATION**

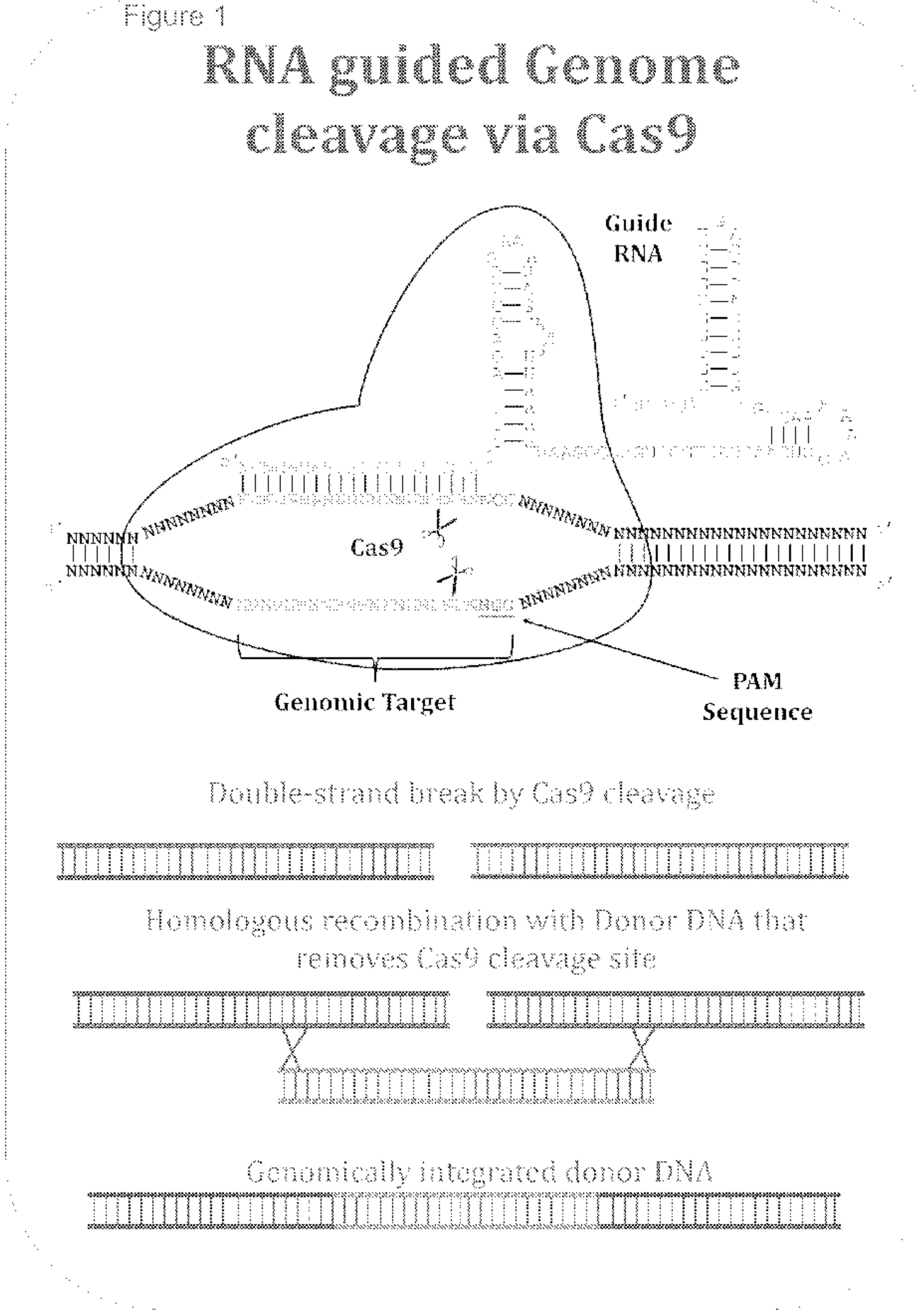
(13) **A1**

(86) **Date de dépôt PCT/PCT Filing Date:** 2014/07/08
(87) **Date publication PCT/PCT Publication Date:** 2015/01/15
(85) **Entrée phase nationale/National Entry:** 2016/01/06
(86) **N° demande PCT/PCT Application No.:** US 2014/045691
(87) **N° publication PCT/PCT Publication No.:** 2015/006290
(30) **Priorité/Priority:** 2013/07/09 (US61/844,168)

(51) **Cl.Int./Int.Cl.** ***C12N 15/63*** (2006.01), ***C12N 15/09*** (2006.01), ***C12N 15/10*** (2006.01)
(71) **Demandeur/Applicant:**
PRESIDENT AND FELLOWS OF HARVARD COLLEGE, US
(72) **Inventeurs/Inventors:**
CHURCH, GEORGE M., US;
DICARLO, JAMES, US
(74) **Agent:** KIRBY EADES GALE BAKER

(54) **Titre : INGENIERIE DES GENOMES GUIDEE PAR ARN MULTIPLEX**
(54) **Title: MULTIPLEX RNA-GUIDED GENOME ENGINEERING**

Figure 1



(57) **Abrégé/Abstract:**
Methods of multiplex genome engineering in cells using Cas9 is provided which includes a cycle of steps of introducing into the cell a first foreign nucleic acid encoding one or more RNAs complementary to the target DNA and which guide the enzyme to the target

(57) **Abrégé(suite)/Abstract(continued):**

DNA, wherein the one or more RNAs and the enzyme are members of a co-localization complex for the target DNA, and introducing into the cell a second foreign nucleic acid encoding one or more donor nucleic acid sequences, and wherein the cycle is repeated a desired number of times to multiplex DNA engineering in cells.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 January 2015 (15.01.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/006290 A1

(51) **International Patent Classification**:
*C12N 15/63* (2006.01) *C12N 15/09* (2006.01)
*C12N 15/10* (2006.01)

(21) **International Application Number**:
PCT/US2014/045691

(22) **International Filing Date**:
8 July 2014 (08.07.2014)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
61/844,168 9 July 2013 (09.07.2013) US

(71) **Applicant**: **PRESIDENT AND FELLOWS OF HARVARD COLLEGE** [US/US]; 17 Quincy Street, Cambridge, Massachusetts 02138 (US).

(72) **Inventors**: **CHURCH, George, M.**; 218 Kent Street, Brookline, Massachusetts 02446 (US). **DICARLO, James**; 15 Queensberry Street, Boston, Massachusetts 02215 (US).

(74) **Agent**: **IWANICKI, John, P.**; 28 State Street, Suite 1800, Boston, Massachusetts 02109 (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, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

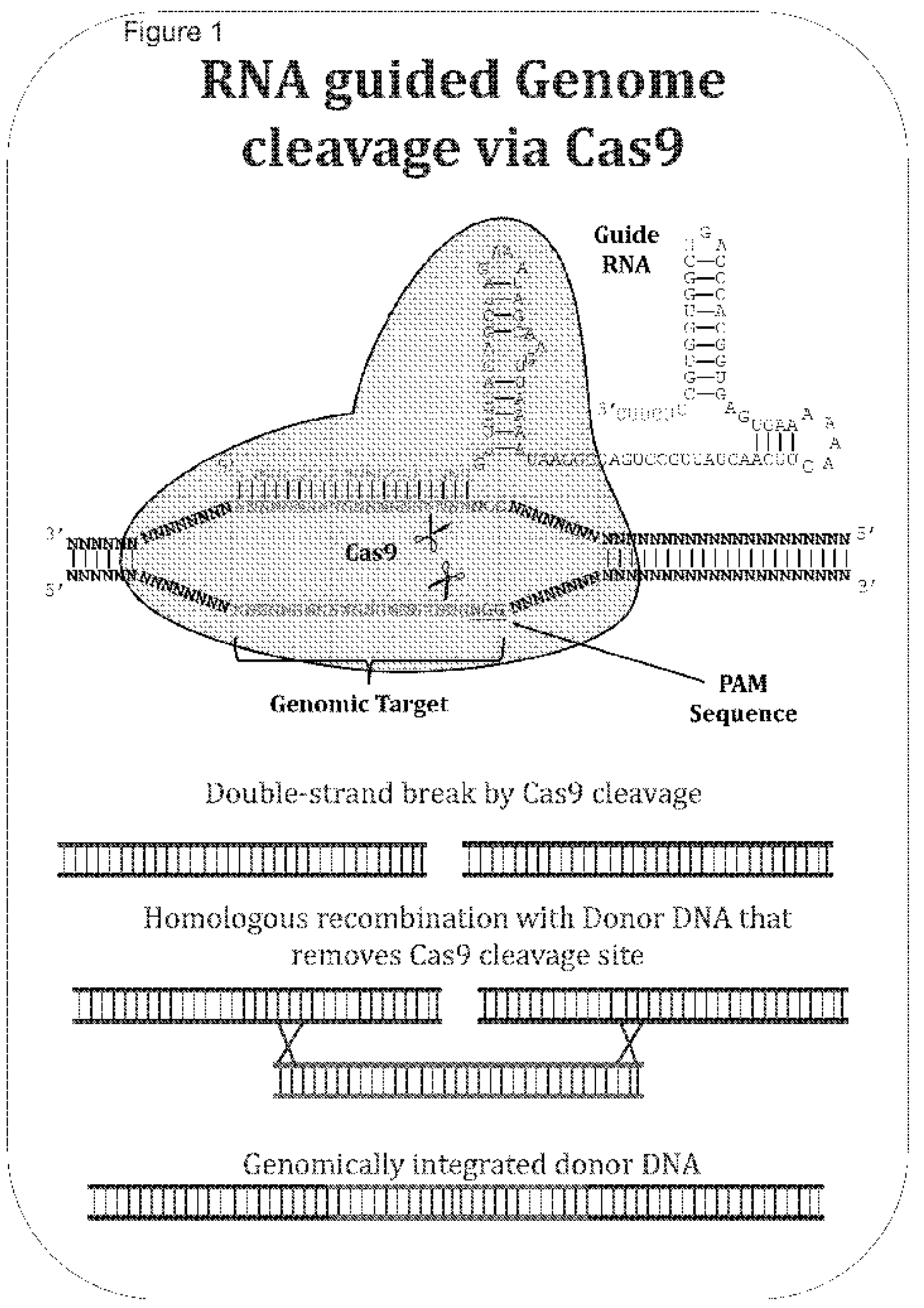
(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, 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**:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

(54) **Title**: MULTIPLEX RNA-GUIDED GENOME ENGINEERING



(57) **Abstract**: Methods of multiplex genome engineering in cells using Cas9 is provided which includes a cycle of steps of introducing into the cell a first foreign nucleic acid encoding one or more RNAs complementary to the target DNA and which guide the enzyme to the target DNA, wherein the one or more RNAs and the enzyme are members of a co-localization complex for the target DNA, and introducing into the cell a second foreign nucleic acid encoding one or more donor nucleic acid sequences, and wherein the cycle is repeated a desired number of times to multiplex DNA engineering in cells.

WO 2015/006290 A1

# MULTIPLEX RNA-GUIDED GENOME ENGINEERING

## RELATED APPLICATION DATA

This application claims priority to U.S. Provisional Patent Application No. 61/844,168 filed on July 9, 2013 and is hereby incorporated herein by reference in its entirety for all purposes.

## STATEMENT OF GOVERNMENT INTERESTS

This invention was made with government support under DE-FG02-02ER63445 from the Department of Energy, NSF-SynBERC from the National Science Foundation and SA5283-11210 from the National Science Foundation. The government has certain rights in the invention.

## BACKGROUND

Bacterial and archaeal CRISPR-Cas systems rely on short guide RNAs in complex with Cas proteins to direct degradation of complementary sequences present within invading foreign nucleic acid. See Deltcheva, E. et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 471, 602-607 (2011); Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proceedings of the National Academy of Sciences of the United States of America 109, E2579-2586 (2012); Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816-821 (2012); Sapranauskas, R. et al. The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. Nucleic acids research 39, 9275-9282 (2011); and Bhaya, D., Davison, M. & Barrangou, R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annual review of genetics 45, 273-297 (2011). A recent in vitro reconstitution of the S. pyogenes type II CRISPR system demonstrated that crRNA (“CRISPR RNA”) fused to a normally trans-encoded tracrRNA (“trans-activating CRISPR RNA”) is sufficient to direct Cas9 protein to sequence-specifically cleave target DNA sequences matching the crRNA. Expressing a gRNA homologous to a target site results in Cas9 recruitment and degradation of the target DNA. See H. Deveau et al., Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. Journal of Bacteriology 190, 1390 (Feb, 2008).

## SUMMARY

Aspects of the present disclosure are directed to the multiplex modification of DNA in a cell using one or more guide RNAs (ribonucleic acids) to direct an enzyme having nuclease activity expressed by the cell, such as a DNA binding protein having nuclease activity, to a target location on the DNA (deoxyribonucleic acid) wherein the enzyme cuts the DNA and an exogenous donor nucleic acid is inserted into the DNA, such as by homologous recombination. Aspects of the present disclosure include cycling or repeating steps of DNA modification on a cell to create a cell having multiple modifications of DNA within the cell. Modifications may include insertion of exogenous donor nucleic acids.

Multiple exogenous nucleic acid insertions can be accomplished by a single step of introducing into a cell, which expresses the enzyme, nucleic acids encoding a plurality of RNAs and a plurality of exogenous donor nucleic acids, such as by co-transformation, wherein the RNAs are expressed and wherein each RNA in the plurality guides the enzyme to a particular site of the DNA, the enzyme cuts the DNA and one of the plurality of exogenous nucleic acids is inserted into the DNA at the cut site. According to this aspect, many alterations or modification of the DNA in the cell are created in a single cycle.

Multiple exogenous nucleic acid insertions can be accomplished in a cell by repeated steps or cycles of introducing into a cell, which expresses the enzyme, one or more nucleic acids encoding one or more RNAs or a plurality of RNAs and one or more exogenous nucleic acids or a plurality of exogenous nucleic acids wherein the RNA is expressed and guides the enzyme to a particular site of the DNA, the enzyme cuts the DNA and the exogenous nucleic acid is inserted into the DNA at the cut site, so as to result in a cell having multiple alterations or insertions of exogenous DNA into the DNA within the cell. According to one aspect, the cell expressing the enzyme can be a cell which expresses the enzyme naturally or a cell which has been genetically altered to express the enzyme such as by introducing into the cell a nucleic acid encoding the enzyme and which can be expressed by the cell. In this manner, aspects of the present disclosure include cycling the steps of introducing RNA into a cell which expresses the enzyme, introducing exogenous donor nucleic acid into the cell, expressing the RNA, forming a co-localization complex of the RNA, the enzyme and the DNA, enzymatic cutting of the DNA by the enzyme, and insertion of the donor nucleic acid into the DNA. Cycling or repeating of the above steps results in multiplexed genetic modification of a cell at multiple loci, i.e., a cell having multiple genetic modifications.

According to certain aspects, a method of increasing rate of homologous recombination is provided by the cycling method described above. In one embodiment, genomic Cas9 directed DNA cutting stimulates exogenous DNA via dramatically increasing the rate of homologous recombination. According to a certain additional aspect, the exogenous donor nucleic acid includes homology sequences or arms flanking the cut site. According to a certain additional aspect, the exogenous donor nucleic acid includes a sequence to remove the cut sequence. According to a certain additional aspect, the exogenous donor nucleic acid includes homology sequences or arms flanking the cut site and a sequence to remove the cut site. In this manner, Cas9 can be used as a negative selection against cells that do not incorporate exogenous donor DNA. Accordingly, a negative selection method is provided for identifying cells having high recombination frequency.

According to certain aspects, DNA binding proteins or enzymes within the scope of the present disclosure include a protein that forms a complex with the guide RNA and with the guide RNA guiding the complex to a double stranded DNA sequence wherein the complex binds to the DNA sequence. According to one aspect, the enzyme can be an RNA guided DNA binding protein, such as an RNA guided DNA binding protein of a Type II CRISPR System that binds to the DNA and is guided by RNA. According to one aspect, the RNA guided DNA binding protein is a Cas9 protein.

This aspect of the present disclosure may be referred to as co-localization of the RNA and DNA binding protein to or with the double stranded DNA. In this manner, a DNA binding protein-guide RNA complex may be used to cut multiple sites of the double stranded DNA so as to create a cell with multiple genetic modifications, such as multiple insertions of exogenous donor DNA.

According to certain aspects, a method of making multiple alterations to target DNA in a cell expressing an enzyme that forms a co-localization complex with RNA complementary to the target DNA and that cleaves the target DNA in a site specific manner is provided including (a) introducing into the cell a first foreign nucleic acid encoding one or more RNAs complementary to the target DNA and which guide the enzyme to the target DNA, wherein the one or more RNAs and the enzyme are members of a co-localization complex for the target DNA, introducing into the cell a second foreign nucleic acid encoding one or more donor nucleic acid sequences, wherein the one or more RNAs and the one or more donor nucleic acid sequences are expressed, wherein the one or more RNAs and the enzyme co-localize to the target DNA, the enzyme cleaves the target DNA and the donor nucleic acid is inserted into the target DNA to produce altered DNA in the cell, and repeating step (a) multiple times to produce multiple alterations to the DNA in the cell.

According to one aspect, the cell is a eukaryotic cell. According to one aspect, the cell is a yeast cell, a plant cell or an animal cell. According to one aspect, the cell is a mammalian cell.

According to one aspect, the RNA is between about 10 to about 500 nucleotides. According to one aspect, the RNA is between about 20 to about 100 nucleotides.

According to one aspect, the one or more RNAs is a guide RNA. According to one aspect, the one or more RNAs is a tracrRNA-crRNA fusion.

According to one aspect, the DNA is genomic DNA, mitochondrial DNA, viral DNA, or exogenous DNA.

Further features and advantages of certain embodiments of the present invention will become more fully apparent in the following description of embodiments and drawings thereof, and from the claims.

## BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other features and advantages of the present embodiments will be more fully understood from the following detailed description of illustrative embodiments taken in conjunction with the accompanying drawings in which:

FIG. 1 is a schematic of RNA guided genome cleavage via Cas9.

FIG. 2 is a schematic depicting multiplexed genome engineering in yeast using Cas9.

FIG. 3 is a schematic depicting allele replacement using oligonucleotides targeting four loci crucial in thermotolerance in yeast.

FIG. 4 is a graph depicting number of modifications per cell after one cycle and after two cycles.

FIG. 5A is a table of strains having mutations. FIG. 5B shows thermotolerance to heat shock for the various strains.

FIG. 6A depicts graphical data for transformation frequency. FIG. 6B depicts graphical data for individual recombination frequency. FIG. 6C depicts graphical data for co-recombination frequency at can1 and KanMX locus.

FIG. 7 depicts graphical data for multiplex linear cassette incorporation for two loci.

FIG. 8A depicts graphical data for fold change in double time at 30°C. FIG. 8B depicts graphical data for fold change in double time at 37°C. FIG. 8C depicts graphical data for fold change in double time at 42°C with cells inoculated from the late stationary phase culture. FIG. 8D depicts graphical data for fold change in double time at 42°C with cells inoculated from the late log phase culture.

## DETAILED DESCRIPTION

Embodiments of the present disclosure are based on the repeated use of exogenous DNA, nuclease enzymes such as DNA binding proteins and guide RNAs to co-localize to DNA and digest or cut the DNA with insertion of the exogenous DNA, such as by homologous recombination. Such DNA binding proteins are readily known to those of skill in the art to bind to DNA for various purposes. Such DNA binding proteins may be naturally occurring. DNA binding proteins included within the scope of the present disclosure include those which may be guided by RNA, referred to herein as guide RNA. According to this aspect, the guide RNA and the RNA guided DNA binding protein form a co-localization complex at the DNA. Such DNA binding proteins having nuclease activity are known to those of skill in the art, and include naturally occurring DNA binding proteins having nuclease activity, such as Cas9 proteins present, for example, in Type II CRISPR systems. Such Cas9 proteins and Type II CRISPR systems are well documented in the art. See Makarova et al., Nature Reviews, Microbiology, Vol. 9, June 2011, pp. 467-477 including all supplementary information hereby incorporated by reference in its entirety.

Exemplary DNA binding proteins having nuclease activity function to nick or cut double stranded DNA. Such nuclease activity may result from the DNA binding protein having one or more polypeptide sequences exhibiting nuclease activity. Such exemplary DNA binding proteins may have two separate nuclease domains with each domain responsible for cutting or nicking a particular strand of the double stranded DNA. Exemplary polypeptide sequences having nuclease activity known to those of skill in the art include the McrA-HNH nuclease related domain and the RuvC-like nuclease domain. Accordingly, exemplary DNA binding proteins are those that in nature contain one or more of the McrA-HNH nuclease related domain and the RuvC-like nuclease domain.

An exemplary DNA binding protein is an RNA guided DNA binding protein of a Type II CRISPR System. An exemplary DNA binding protein is a Cas9 protein.

In S. pyogenes, Cas9 generates a blunt-ended double-stranded break 3bp upstream of the protospacer-adjacent motif (PAM) via a process mediated by two catalytic domains in the protein: an HNH domain that cleaves the complementary strand of the DNA and a RuvC-like domain that cleaves the non-complementary strand. See Jinke et al., Science 337, 816-821 (2012) hereby incorporated by reference in its entirety. Cas9 proteins are known to exist in many Type II CRISPR systems including the following as identified in the supplementary information to Makarova et al., Nature Reviews, Microbiology, Vol. 9, June 2011, pp. 467-477: Methanococcus maripaludis C7; Corynebacterium diphtheriae; Corynebacterium efficiens YS-314; Corynebacterium glutamicum ATCC 13032 Kitasato; Corynebacterium glutamicum ATCC 13032 Bielefeld; Corynebacterium glutamicum R; Corynebacterium kroppenstedtii DSM 44385; Mycobacterium abscessus ATCC 19977; Nocardia farcinica IFM10152; Rhodococcus erythropolis PR4; Rhodococcus jostii RHA1; Rhodococcus opacus B4 uid36573; Acidothermus cellulolyticus 11B; Arthrobacter chlorophenolicus A6; Kribbella flavida DSM 17836 uid43465; Thermomonospora curvata DSM 43183; Bifidobacterium dentium Bd1; Bifidobacterium longum DJO10A; Slackia heliotrinireducens DSM 20476; Persephonella marina EX H1; Bacteroides fragilis NCTC 9434; Capnocytophaga ochracea DSM 7271; Flavobacterium psychrophilum JIP02 86; Akkermansia muciniphila ATCC BAA 835; Roseiflexus castenholzii DSM 13941; Roseiflexus RS1; Synechocystis PCC6803; Elusimicrobium minutum Pei191; uncultured Termite group 1 bacterium phylotype Rs D17; Fibrobacter succinogenes S85; Bacillus cereus ATCC 10987; Listeria innocua;Lactobacillus casei; Lactobacillus rhamnosus GG; Lactobacillus salivarius UCC118; Streptococcus agalactiae A909; Streptococcus agalactiae NEM316; Streptococcus agalactiae 2603;

Streptococcus dysgalactiae equisimilis GGS 124; Streptococcus equi zooepidemicus MGCS10565; Streptococcus gallolyticus UCN34 uid46061; Streptococcus gordonii Challis subst CH1; Streptococcus mutans NN2025 uid46353; Streptococcus mutans; Streptococcus pyogenes M1 GAS; Streptococcus pyogenes MGAS5005; Streptococcus pyogenes MGAS2096; Streptococcus pyogenes MGAS9429; Streptococcus pyogenes MGAS10270; Streptococcus pyogenes MGAS6180; Streptococcus pyogenes MGAS315; Streptococcus pyogenes SSI-1; Streptococcus pyogenes MGAS10750; Streptococcus pyogenes NZ131; Streptococcus thermophiles CNRZ1066; Streptococcus thermophiles LMD-9; Streptococcus thermophiles LMG 18311; Clostridium botulinum A3 Loch Maree; Clostridium botulinum B Eklund 17B; Clostridium botulinum Ba4 657; Clostridium botulinum F Langeland; Clostridium cellulolyticum H10; Finegoldia magna ATCC 29328; Eubacterium rectale ATCC 33656; Mycoplasma gallisepticum; Mycoplasma mobile 163K; Mycoplasma penetrans; Mycoplasma synoviae 53; Streptobacillus moniliformis DSM 12112; Bradyrhizobium BTAi1; Nitrobacter hamburgensis X14; Rhodopseudomonas palustris BisB18; Rhodopseudomonas palustris BisB5; Parvibaculum lavamentivorans DS-1; Dinoroseobacter shibae DFL 12; Gluconacetobacter diazotrophicus Pal 5 FAPERJ; Gluconacetobacter diazotrophicus Pal 5 JGI; Azospirillum B510 uid46085; Rhodospirillum rubrum ATCC 11170; Diaphorobacter TPSY uid29975; Verminephrobacter eiseniae EF01-2; Neisseria meningitides 053442; Neisseria meningitides alpha14; Neisseria meningitides Z2491; Desulfovibrio salexigens DSM 2638; Campylobacter jejuni doylei 269 97; Campylobacter jejuni 81116; Campylobacter jejuni; Campylobacter lari RM2100; Helicobacter hepaticus; Wolinella succinogenes; Tolumonas auensis DSM 9187; Pseudoalteromonas atlantica T6c; Shewanella pealeana ATCC 700345; Legionella pneumophila Paris; Actinobacillus succinogenes 130Z; Pasteurella multocida; Francisella tularensis novicida U112; Francisella tularensis holarctica; Francisella tularensis FSC 198; Francisella tularensis tularensis; Francisella tularensis WY96-3418; and Treponema denticola ATCC 35405. Accordingly, aspects of the present disclosure are directed to a Cas9 protein present in a Type II CRISPR system.

The Cas9 protein may be referred by one of skill in the art in the literature as Csn1. The S. pyogenes Cas9 protein sequence that is the subject of experiments described herein is shown below. See Deltcheva et al., Nature 471, 602-607 (2011) hereby incorporated by reference in its entirety.

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE
ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFG
NIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD
VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGN
LIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA

GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELH
AILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEE
VVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL
SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWG
RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL
HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER
MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH
IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVR
K
MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA
YSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE
QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD-

According to one aspect, the RNA guided DNA binding protein includes homologs and orthologs of Cas9 which retain the ability of the protein to bind to the DNA, be guided by the RNA and cut the DNA. According to one aspect, the Cas9 protein includes the sequence as set forth for naturally occurring Cas9 from S. pyogenes and protein sequences having at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98% or 99% homology thereto and being a DNA binding protein, such as an RNA guided DNA binding protein.

According to one aspect, an engineered Cas9-gRNA system is provided which enables RNA-guided genome cutting in a site specific manner, if desired, and modification of the genome by insertion of exogenous donor nucleic acids. The guide RNAs are complementary to target sites or target loci on the DNA. The guide RNAs can be crRNA-tracrRNA chimeras. The Cas9 binds at or near target genomic DNA. The one or more guide RNAs bind at or near target genomic DNA. The Cas9 cuts the target genomic DNA and exogenous donor DNA is inserted into the DNA at the cut site.

Accordingly, methods are directed to the use of a guide RNA with a Cas9 protein and an exogenous donor nucleic acid to multiplex insertions of exogenous donor nucleic acids into DNA within a cell expressing Cas9 by cycling the insertion of nucleic acid encoding the RNA and exogenous donor nucleic acid, expressing the RNA, colocalizing the RNA, Cas9 and DNA in a manner to cut the DNA, and insertion of the exogenous donor nucleic acid. The method steps can be cycled in any desired number to result in any desired number of DNA modifications. Methods of the present disclosure are accordingly directed to editing target genes using the Cas9 proteins and guide RNAs described herein to provide multiplex genetic and epigenetic engineering of cells.

Further aspects of the present disclosure are directed to the use of DNA binding proteins or systems in general for the multiplex insertion of exogenous donor nucleic acids into the DNA, such as genomic DNA, of a cell, such as a human cell. One of skill in the art will readily identify exemplary DNA binding systems based on the present disclosure.

Cells according to the present disclosure include any cell into which foreign nucleic acids can be introduced and expressed as described herein. It is to be understood that the basic concepts of the present disclosure described herein are not limited by cell type. Cells according to the present disclosure include eukaryotic cells, prokaryotic cells, animal cells, plant cells, fungal cells, archael cells, eubacterial cells and the like. Cells include eukaryotic cells such as yeast cells, plant cells, and animal cells. Particular cells include mammalian cells, such as human cells. Further, cells include any in which it would be beneficial or desirable to modify DNA.

Target nucleic acids include any nucleic acid sequence to which a co-localization complex as described herein can be useful to nick or cut. Target nucleic acids include genes. For purposes of the present disclosure, DNA, such as double stranded DNA, can include the target nucleic acid and a co-localization complex can bind to or otherwise co-localize with the DNA at or adjacent or near the target nucleic acid and in a manner in which the co-localization complex may have a desired effect on the target nucleic acid. Such target nucleic acids can include endogenous (or naturally occurring) nucleic acids and exogenous (or foreign) nucleic acids. One of skill based on the present disclosure will readily be able to identify or design guide RNAs and Cas9 proteins which co-localize to a DNA including a target nucleic acid. One of skill will further be able to identify transcriptional regulator proteins or domains which likewise co-localize to a DNA including a target nucleic acid. DNA includes genomic DNA, mitochondrial DNA, viral DNA or exogenous DNA. According to one aspect, materials and methods useful in the practice of the present disclosure include those described in Di Carlo, et al., Nucleic Acids Research, 2013, vol. 41, No. 7 4336-4343 hereby incorporated by reference in its entirety for all purposes including exemplary strains and media, plasmid construction, transformation of plasmids, electroporation of transcient gRNA cassette and donor nucleic acids, transformation of gRNA plasmid with donor DNA into Cas9-expressing cells, galactose induction of Cas9, identification of CRISPR-Cas targets in yeast genome, etc. Additional references including information, materials and methods useful to one of skill in carrying out the invention are provided in Mali,P., Yang,L., Esvelt,K.M., Aach,J., Guell,M., DiCarlo,J.E., Norville,J.E. and Church,G.M. (2013) RNA-Guided human genome engineering via Cas9. Science, 10.1126fscience.1232033; Storici,F., Durham,C.L., Gordenin,D.A.

and Resnick,M.A. (2003) Chromosomal site-specific double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast. PNAS, 100, 14994-14999 and Jinek,M., Chylinski,K., Fonfara,I., Hauer,M., Doudna,J.A. and Charpentier,E. (2012) A programmable dual-RNA-Guided DNA endonuclease in adaptive bacterial immunity. Science, 337, 816-821 each of which are hereby incorporated by reference in their entireties for all purposes.

Foreign nucleic acids (i.e. those which are not part of a cell's natural nucleic acid composition) may be introduced into a cell using any method known to those skilled in the art for such introduction. Such methods include transfection, transduction, viral transduction, microinjection, lipofection, nucleofection, nanoparticle bombardment, transformation, conjugation and the like. One of skill in the art will readily understand and adapt such methods using readily identifiable literature sources.

The following examples are set forth as being representative of the present disclosure. These examples are not to be construed as limiting the scope of the present disclosure as these and other equivalent embodiments will be apparent in view of the present disclosure, figures and accompanying claims.

## EXAMPLE I

### General Process for Multiplexed Gene Editing Using CRISPR-Cas9 in Yeast

Cas9 from the CRISPR immune system of Streptococcous pyogenes is used to stimulate homologous recombination and to select against cells that do not recombine transformed DNA in Saccharaomyces cerevisiae. A general method of RNA-guided DNA cleavage using Cas9 is presented in FIG. 1. A co-localization complex is formed between Cas9, a guide RNA and the target DNA. A double stranded break is created in the target DNA by Cas9. Donor DNA is then inserted into the DNA by homologous recombination. The donor DNA includes flanking sequences on either side of the cut site and a sequence that removes the Cas9 cleavage site. The result is integration of the donor DNA into the DNA, which may be genomic DNA.

A general method for high frequency donor DNA recombination using multiplexed DNA engineering in yeast using Cas9 is provided as follows and with reference to FIG. 2. Cells not having a naturally present Cas9 RNA guided endonuclease may be transformed with DNA to allow the cell to express a Cas9 RNA guided endonuclease. Cells are grown that express a Cas9 RNA-guided endonuclease. A plasmid including one or more nucleic acids encoding one or more guide RNAs and a selection marker known to those of skill in the art is created for introduction into a cell and expression of the one or more guide RNAs. As shown in FIG. 2, a pool of plasmids is shown each with a nucleic acid encoding a guide RNA to be used for a different gene to be inserted into the genomic DNA of the cell, i.e. gene A, gene B, gene C, gene D and gene E. A pool of donor DNA is also provided including double stranded donor DNA for gene A, gene B, gene C, gene D and gene E.

Cells are washed and conditioned with lithium acetate. Cells may be further washed and mixed with a pool of exogenous donor nucleic acids, such as double stranded oligonucleotides, for example a DNA cassette, and the plasmids including the nucleic acids encoding the guide RNAs. As shown in FIG. 2, the cells are transformed with the exogenous donor nucleic acids and the plasmids using PEG 3350 and lithium acetate.

As shown in FIG. 2, cells are selected for the one or more guide RNAs using the selection marker. The selected cells express the one or more guide RNAs. One or more co-localization complexes are formed between a guide RNA, a Cas9 RNA-guided endonuclease and DNA in the cell. The endonuclease cuts the DNA and a donor nucleic acid is inserted into the cell by recombination, such as homologous recombination. The cells are then cured for the plasmid and the cells are then optionally subjected to one or additional cycles of the above steps. A plurality of cycles may be performed. A cell subjected to a plurality of cycles exhibits high recombination frequency. Alternatively, the cells are deselected for plasmid maintenance or otherwise the cells are placed in media to select against cells with the plasmid. The process is then repeated beginning with the cell growth step. Accordingly, methods include cycling of cells already modified by a prior cycle or selecting cells from a prior cycle which have not been modified and further cycling the unmodified cells to effect modification of DNA as described herein.

## EXAMPLE II

## Detailed Cycling Protocol

Cells are grown (uracil auxotrophs, with constitutive Cas9 expression) to an optical density of 0.8 to 1.0 in 5 ml SC yeast media or of SC + FOA (100 µg/ml). The cells are spun at 2250 x g for 3 minutes, and are washed once with 10 ml water. the cells are sun and resuspended in 1 ml of 100 mM lithium acetate. The cells are pelleted and resuspended in 500 µl 100 mM lithium acetate. A transformation mixture is created by adding in the following order, 50 µl of cells; DNA mixture including 1 nmol of double stranded oligonucleotide pool, 5 µg each of guide RNA (p426 vector, with uracil marker) and fill to 70 µl with water to achieve desired final volume; 240 µl 50% PEG 3350; and 36 µl 1 M lithium acetate. The mixture is incubated at 30°C for 30 minutes. The mixture is then vortexed and the cells are heat shocked by incubating the mixture at 42°C for 20 minutes. The cells are then pelleted and the supernatant is removed. The cells are inoculated with 5 ml SC-uracil to select for uracil gene containing gRNA plasmid. The cells are allowed to recover for 2 days. After two days, 100 µl of the cell culture is inoculated into 5 ml fresh SC and allowed to grow for 12 hours to deselect for plasmid maintenance. 100 µl of the SC culture cells are then inoculated into 5 ml of SC + FOA (100µg/mL) media to select against cells with the plasmid. This completes one cycle of the process. The process is repeated for any number of desired cycles. The total process may include 1 cycle, 2 cycles, 3 cycles, 4 cycles, 5 cycles, 6 cycles, 7 cycles, 8 cycles, 9 cycles, 10 cycles, 15 cycles, 20, cycles, 25 cycles, etc.

## EXAMPLE III

## Thermotolerance to Heat Shock in Select Mutants

Using the methods described herein, thermotolerance to heat shock in select mutants has been shown. Genes that have been shown to increase thermotolerance in yeast upon knockout or point mutation were targeted by the guide RNA-Cas9 system described herein. Four genes were selected for mutation: UBC1, SCH9, TFS1, and RAS2. SCH9 is a protein kinase that regulates osmostress, nutrient and environmental stress genes. TFS1 inhibits carboxypeptidase Y and Ira2p, inhibits Ras GAP activity and responds to DNA replicative stress. RAS2 is a GTP binding protein that regulates nitrogen starvation and is involved in stress response pathways. For each of SCH9, TFS1 and RAS2, a donor DNA was created which is an allele containing a serine to alanine mutation in the coding region. UBC1-E2 is a ubiquitin-conjugating enzyme. A donor DNA including a point mutation that removes a phosphorylation site resulting in thermotolerance was created.

Using the methods described herein the genes were targeted using guide RNA designed to direct Cas9 cleavage to the loci of the genes along with double stranded oligonucleotide to impart the changes. As shown in FIG. 3, allele replacement was achieved using oligonucleotides targeting four loci responsible for thermotolerance in yeast. According to the schematic, four plasmids each incorporating a nucleic acid encoding a guide RNA for one of the genes were created: UBC1 gRNS plasmid, TFS1 gRNA plasmid, SCH9 gRNA plasmid and RAS2 gRNA plasmid. Each plasmid had a corresponding double stranded donor oligonucleotide: ubc1 (S97A) double stranded oligonucleotide, tfs1 (tag) double stranded oligonucleotide, sch9 (tag) double stranded oligonucleotide and ras (tag) double stranded oligonucleotide. The plasmids and the corresponding double stranded donor oligonucleotides were co-transformed into yeast as a pool. Two cycles were performed and the number of modifications per cell as a function of percentage of cells in the cell population is shown at FIG. 4. A significant number of cells included one and two modifications after cycle 2. One triple mutant was able to be isolated (data not shown.)

FIG. 5A is a table of the strains resulting from the methods described herein showing strains transformed with one donor oligonucleotide, strains transformed with two donor oligonucleotides and a strain transformed with three donor oligonucleotides. FIG. 5B shows the effect of incubation at 42°C for three hours compared to no incubation and s slight decrease in wild type cell number. FIG. 5B also shows the effect of incubation at 55°C for two hours compared to no incubation. The mutants most tolerant to heat shock at 55°C were sch9, sch9 tfs1 and tfs1 ubc1(s97a).

FIG. 6 in general provides graphical information on the optimization of multiplex oligonucleotide incorporation for two loci. FIG. 6A depicts the transformation frequency versus the amount of each plasmid transformed (μg). FIG. 6B depicts the individual recombination frequency versus the amount of each plasmid transformed (μg). FIG. 6C depicts the co-recombination frequency at can1 and KanMX locus versus the amount of each plasmid transformed (μg).

FIG. 7 in general provides graphical information on the multiplex linear cassette incorporation for two loci. The graph charts for the first left most bar, transformation frequency for p426 gRNA ADE2 + HygR Cassette; for the next bar, transformation frequency for p426 gRNA CAN1 + G418R cassette, for the next three bars, transformation frequency for p426 gRNA + ADE2 p426 gRNA CAN1 + HygR Cassette + G418R cassette.

FIG. 8 in general is a growth rate analysis showing double time in exponential growth in elevated temperatures for select mutants. FIG. 8A graphs the fold change in double time at 30°C for the wild type and the mutants identified. FIG. 8B graphs the fold change in double time at 37°C for the wild type and the mutants identified. FIG. 8C graphs the fold change in double time at 42°C for the wild type and the mutants identified as inoculated from the late stationary phase culture. FIG. 8D graphs the fold change in double time at 42°C for the wild type and the mutants identified as inoculated from the late log phase culture. The graphical data shows a lower doubling time at 37°C for sch9 tfs1 and tfs1 ubc1(S97A). The graphical data shows lower doubling time at 42°C for ras2 tfs1, sch9 ubc1(S97A), tfs1 ubc1(S97A) and ras2 tfs1 ubc1(S97A).

Claims:

1. A method of making multiple alterations to target DNA in a cell expressing an enzyme that forms a co-localization complex with RNA complementary to the target DNA and that cleaves the target DNA in a site specific manner comprising

(a) introducing into the cell a first foreign nucleic acid encoding one or more RNAs complementary to the target DNA and which guide the enzyme to the target DNA, wherein the one or more RNAs and the enzyme are members of a co-localization complex for the target DNA,

introducing into the cell a second foreign nucleic acid encoding one or more donor nucleic acid sequences,

wherein the one or more RNAs and the one or more donor nucleic acid sequences are expressed,

wherein the one or more RNAs and the enzyme co-localize to the target DNA, the enzyme cleaves the target DNA and the donor nucleic acid is inserted into the target DNA to produce altered DNA in the cell, and

repeating step (a) multiple times to produce multiple alterations to the DNA in the cell.

2. The method of claim 1 wherein the enzyme is an RNA-guided DNA binding protein.

3. The method of claim 1 wherein the enzyme is Cas9.

4. The method of claim 1 wherein the cell is a eukaryotic cell.

5. The method of claim 1 wherein the cell is a yeast cell, a plant cell or an animal cell.

6. The method of claim 1 wherein the RNA is between about 10 to about 500 nucleotides.

7. The method of claim 1 wherein the RNA is between about 20 to about 100 nucleotides.

8. The method of claim 1 wherein the one or more RNAs is a guide RNA.

9. The method of claim 1 wherein the one or more RNAs is a tracrRNA-crRNA fusion.

10. The method of claim 1 wherein the DNA is genomic DNA, mitochondrial DNA, viral DNA, or exogenous DNA.

11. The method of claim 1 wherein the one or more donor nucleic acid sequences are inserted by recombination.

12. The method of claim 1 wherein the one or more donor nucleic acid sequences are inserted by homologous recombination.

13. The method of claim 1 wherein the one or more RNAs and the one or more donor nucleic acid sequences are present on one or more plasmids.

Figure 1

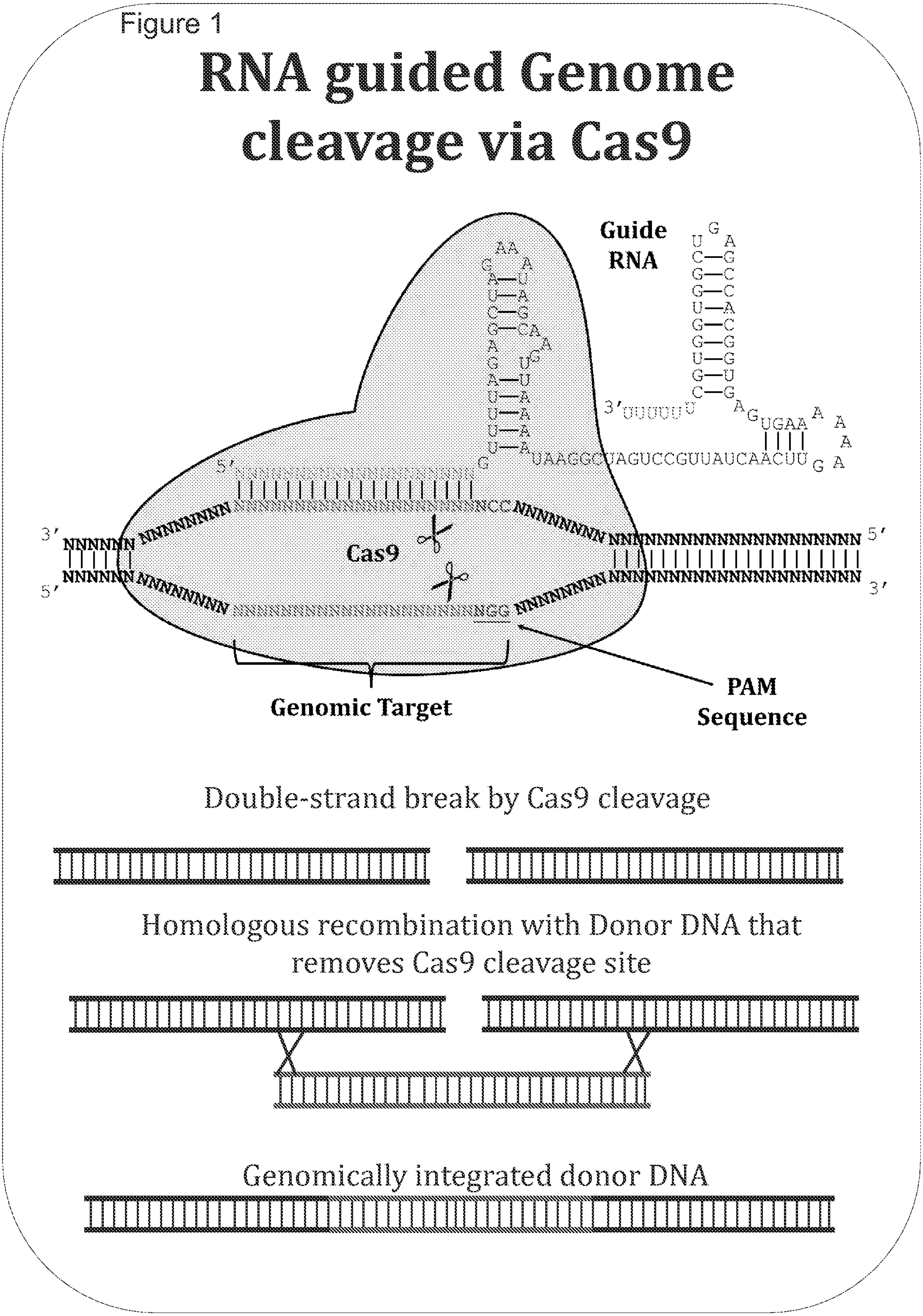

RNA guided Genome cleavage via Cas9
Guide RNA
3′UUUUUU
UAAGGCUAGUCCGUUAUCAACUU
5′
NCC
3′
5′
Cas9
NGG
5′
3′
Genomic Target
PAM Sequence
Double-strand break by Cas9 cleavage
Homologous recombination with Donor DNA that removes Cas9 cleavage site
Genomically integrated donor DNA

Figure 2

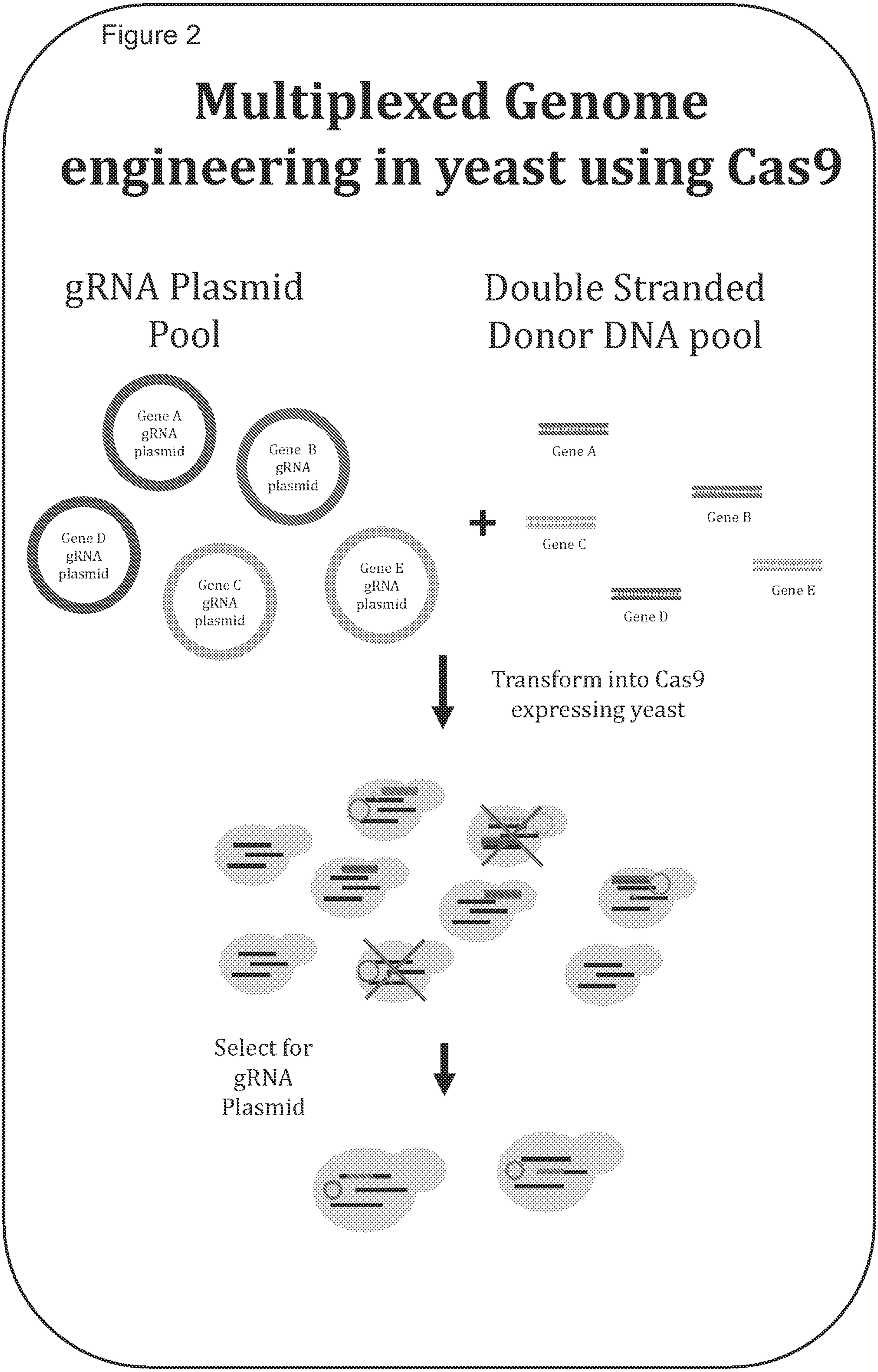

Multiplexed Genome engineering in yeast using Cas9
gRNA Plasmid Pool
Double Stranded Donor DNA pool
Gene A gRNA plasmid
Gene B gRNA plasmid
Gene D gRNA plasmid
Gene C gRNA plasmid
Gene E gRNA plasmid
+
Gene A
Gene B
Gene C
Gene D
Gene E
Transform into Cas9 expressing yeast
Select for gRNA Plasmid

FIG. 3

Allele Replacement using oligonucleotides targeting four loci crucial in thermotolerance in yeast

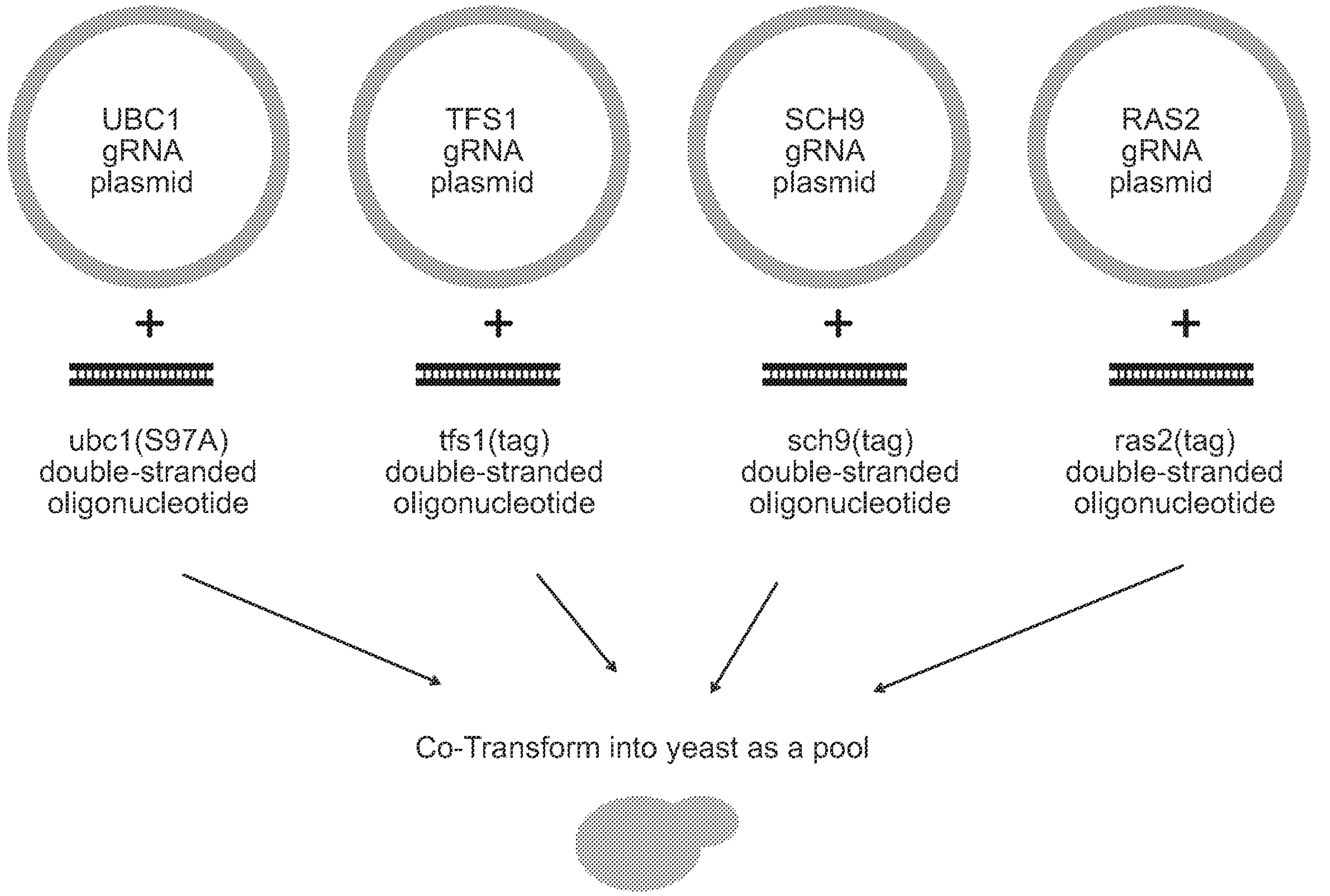
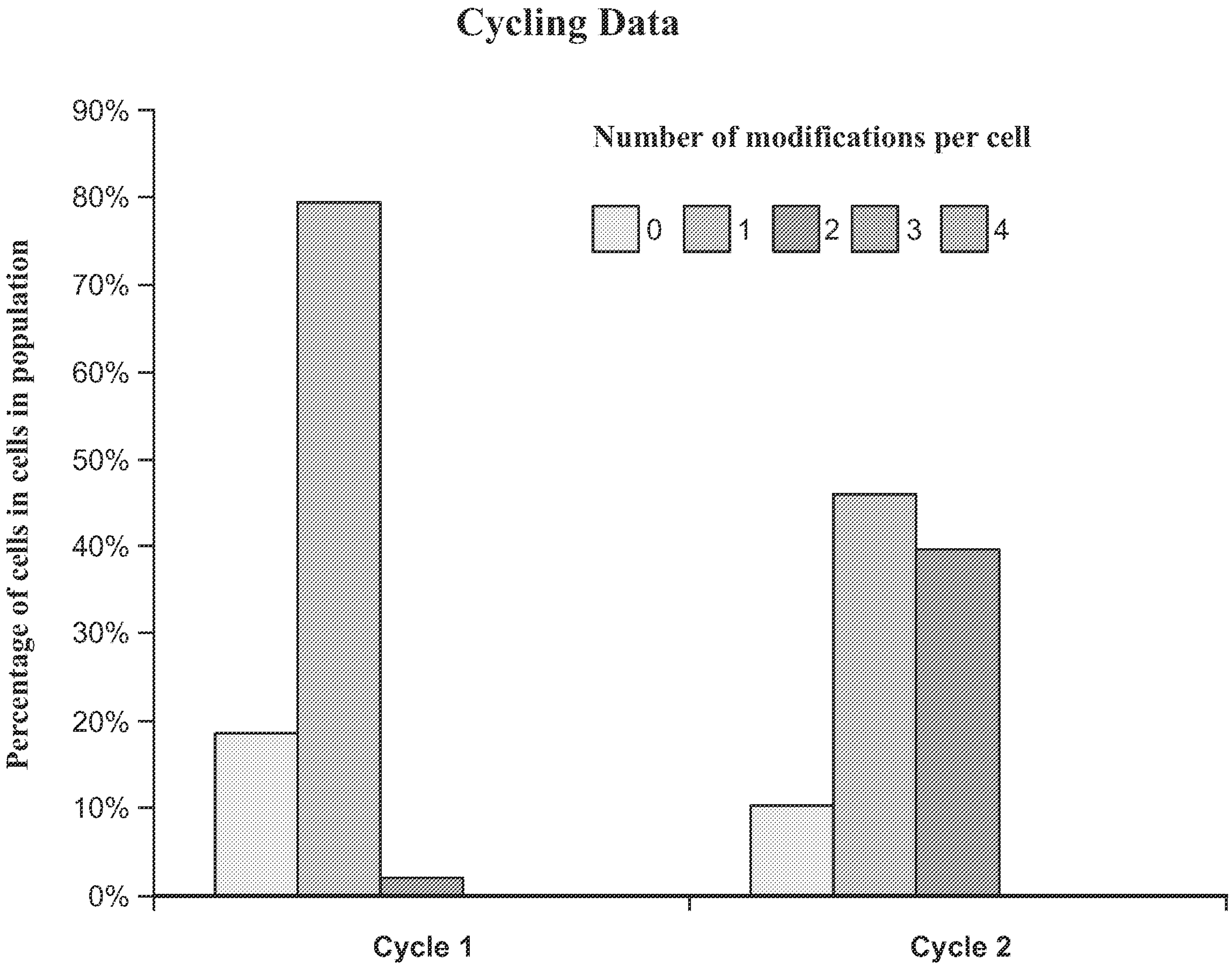


FIG. 4


Cycling Data
Number of modifications per cell
0
1
2
3
4
Percentage of cells in cells in population
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Cycle 1
Cycle 2

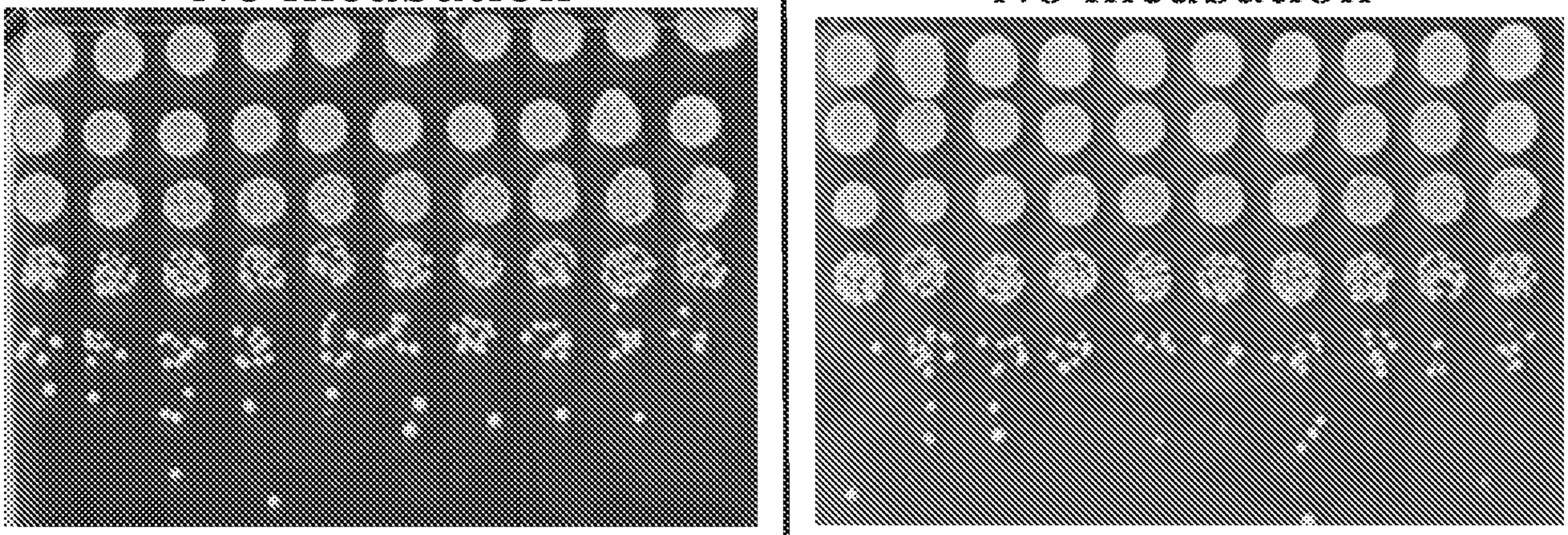
# Thermotolerance to heat shock in select mutants

Figure 5A

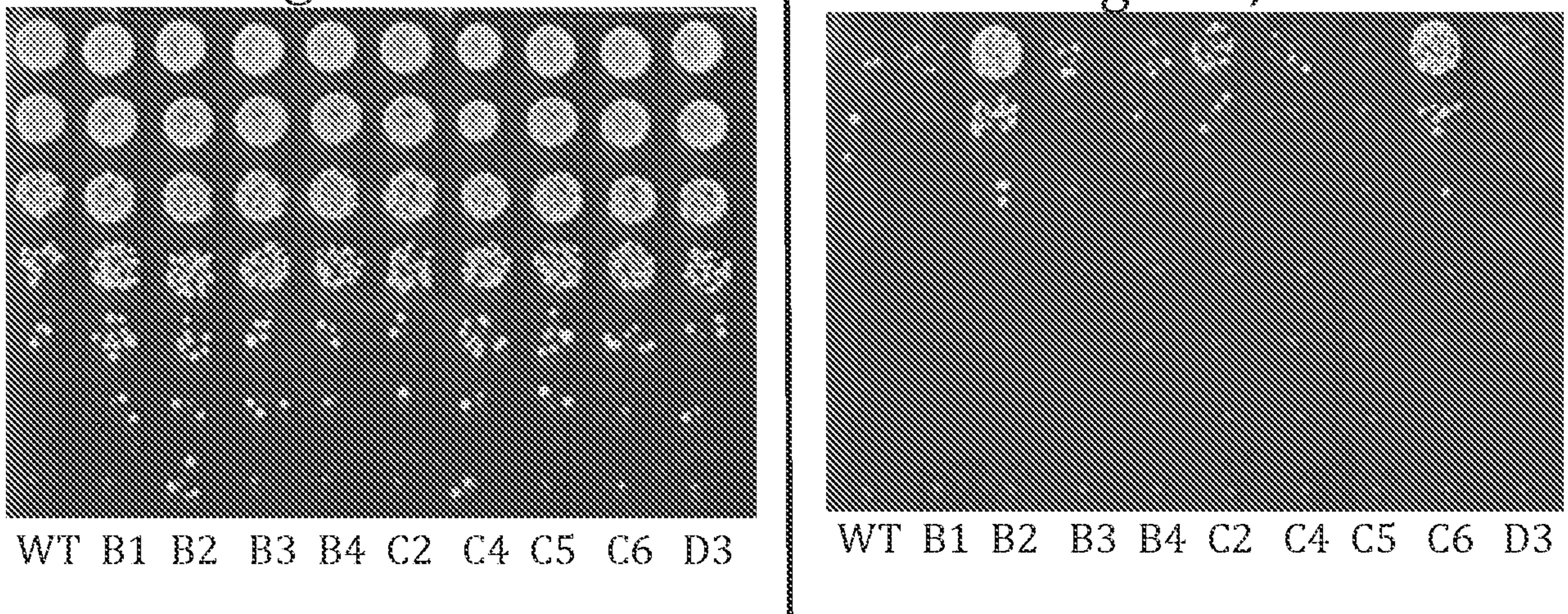
| Strain | Code |
|---|---|
| Wild-Type | WT |
| ras2 | B1 |
| sch9 | B2 |
| ubc1(S97A) | B3 |
| tfs1 | B4 |
| sch9 tfs1 | C2 |
| ras2 tfs1 | C4 |
| sch9 ubc1(S97A) | C5 |
| tfs1 ubc1(S97A) | C6 |
| ras2 tfs1 ubc1(S97A) | D3 |

Figure 5B

No Incubation No Incubation



42 degrees; 3 hours 55 degrees; 2 hours



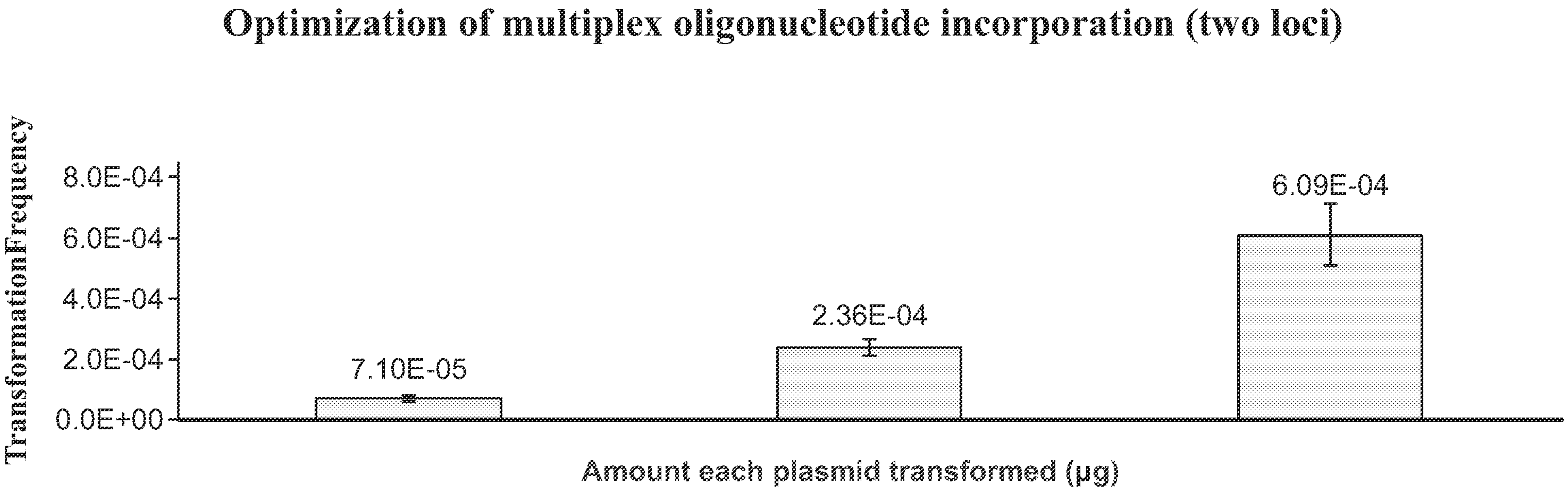

Optimization of multiplex oligonucleotide incorporation (two loci)
TransformationFrequency
8.0E-04
6.0E-04
4.0E-04
2.0E-04
0.0E+00
7.10E-05
2.36E-04
6.09E-04
Amount each plasmid transformed (µg)

FIG. 6A

FIG. 6B

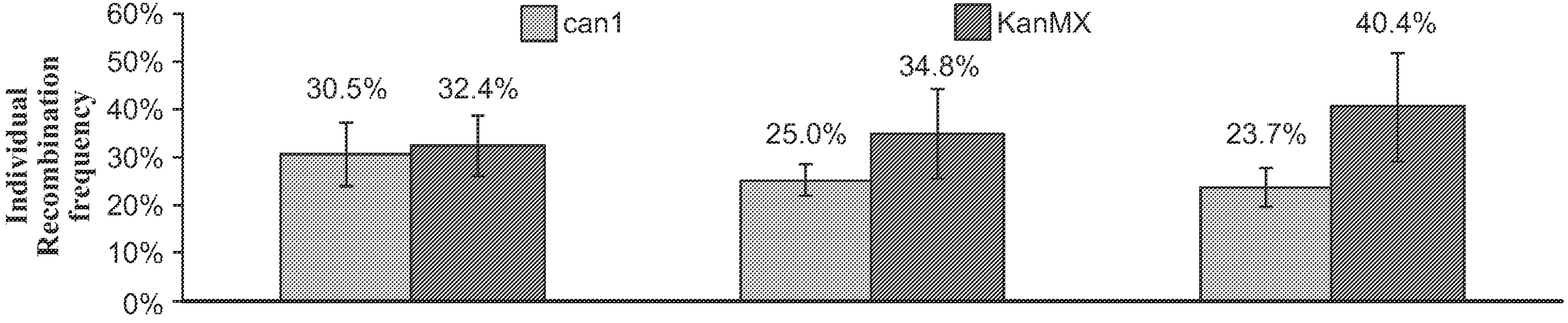
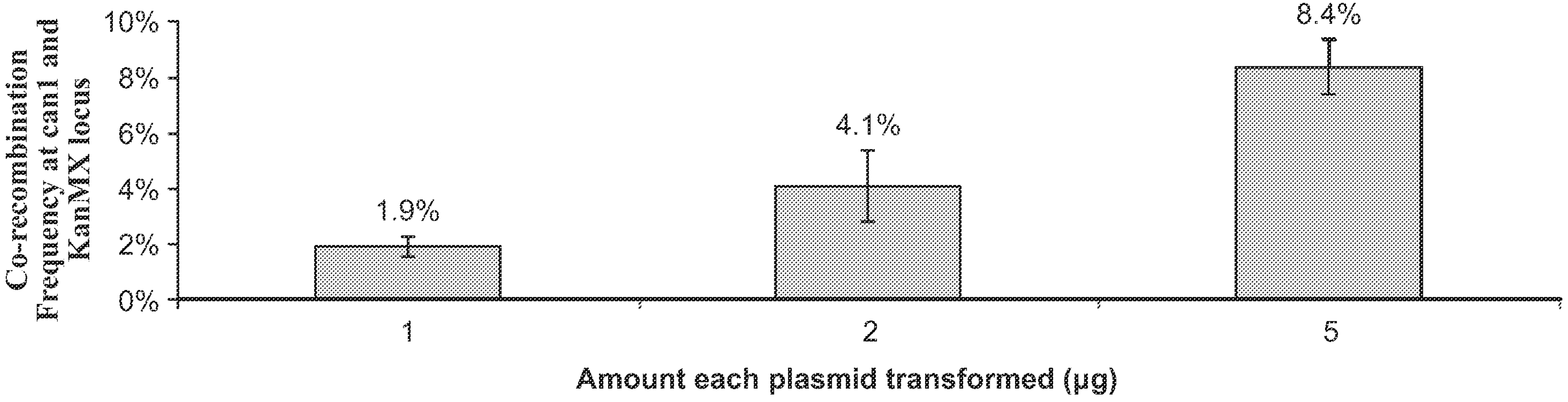

Individual Recombination frequency
60%
50%
40%
30%
20%
10%
0%
can1
KanMX
30.5%
32.4%
25.0%
34.8%
23.7%
40.4%

FIG. 6C


Co-recombination Frequency at can1 and KanMX locus
10%
8%
6%
4%
2%
0%
1.9%
4.1%
8.4%
1
2
5
Amount each plasmid transformed (μg)

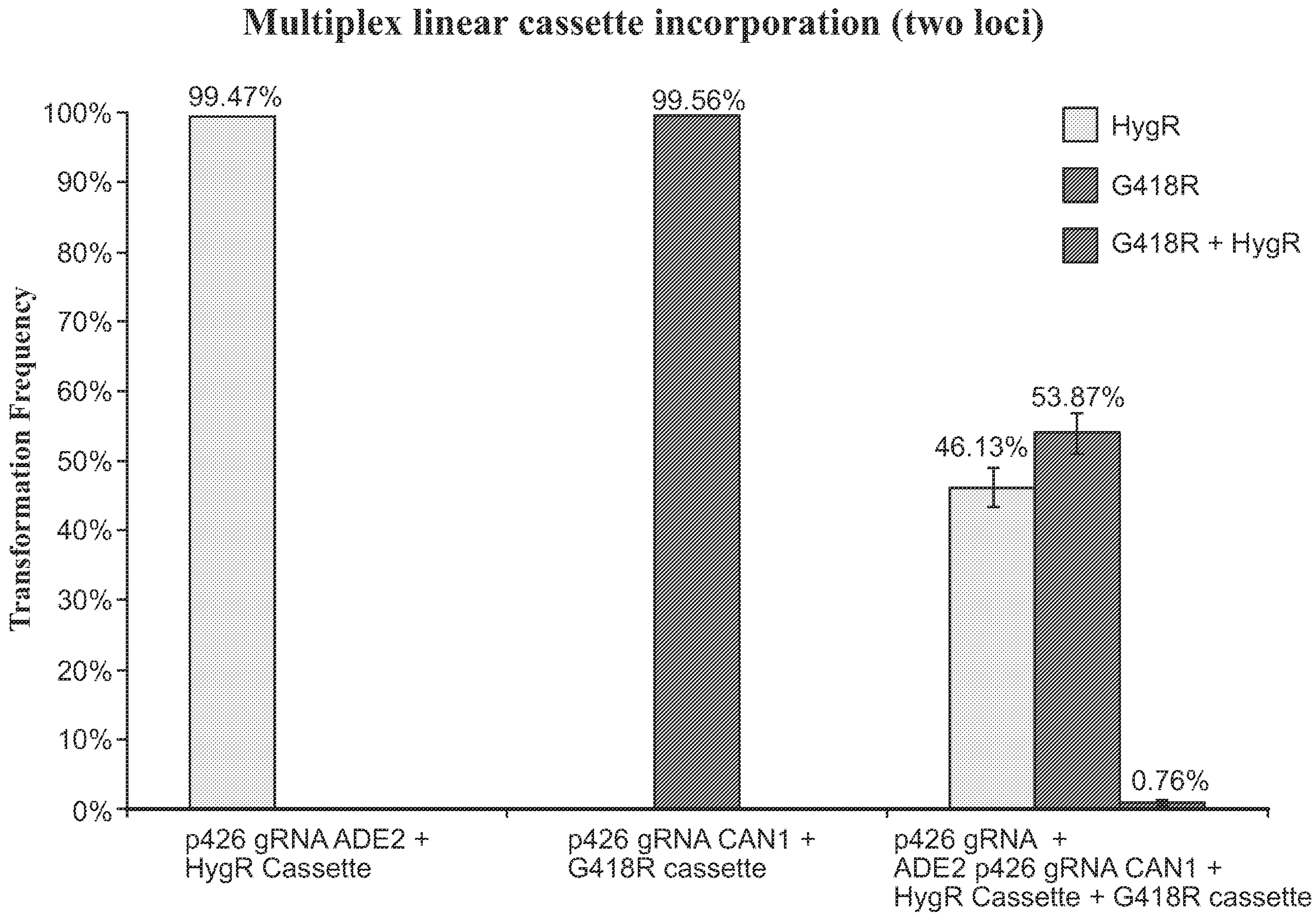

Multiplex linear cassette incorporation (two loci)
Transformation Frequency
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
99.47%
99.56%
46.13%
53.87%
0.76%
HygR
G418R
G418R + HygR
p426 gRNA ADE2 + HygR Cassette
p426 gRNA CAN1 + G418R cassette
p426 gRNA + ADE2 p426 gRNA CAN1 + HygR Cassette + G418R cassette

FIG. 7

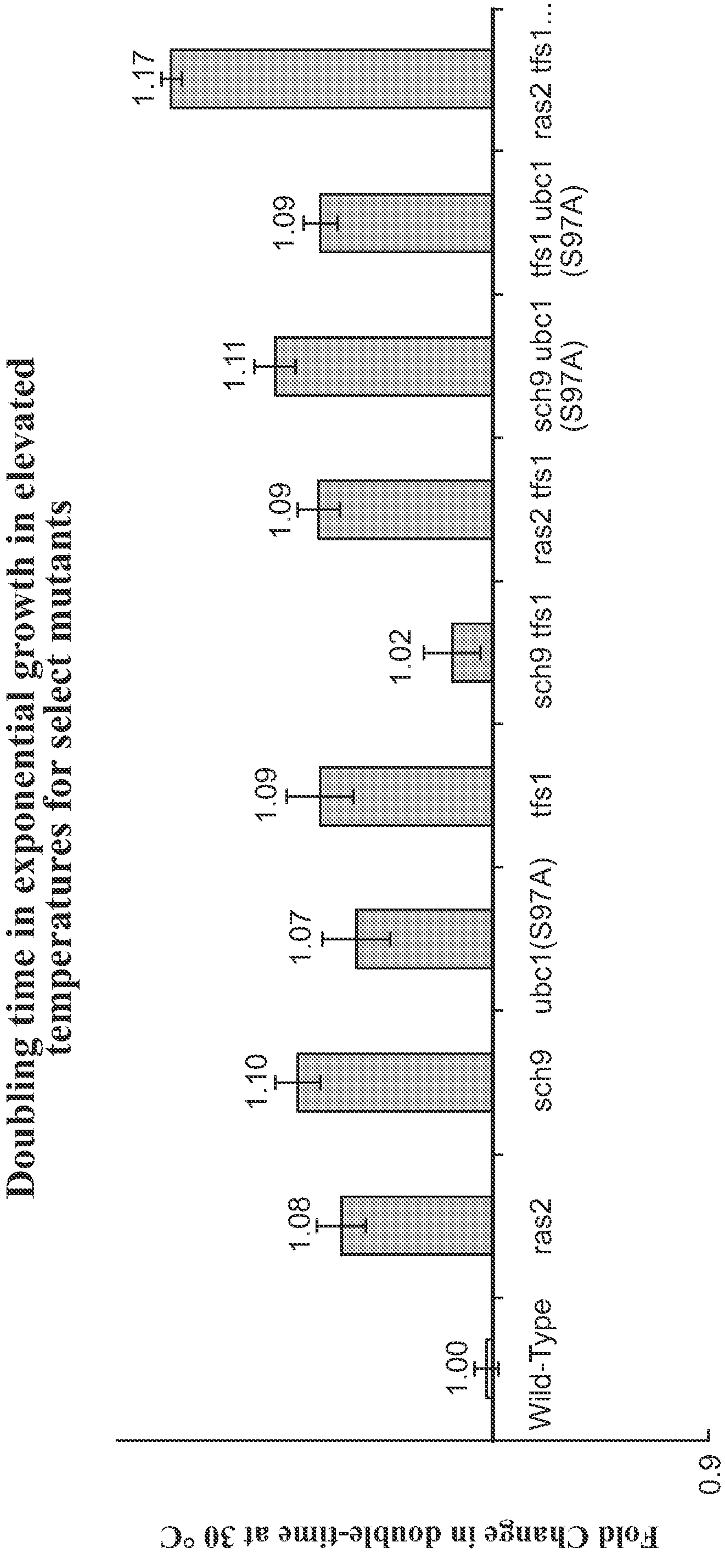

Doubling time in exponential growth in elevated temperatures for select mutants
Fold Change in double-time at 30 °C
0.9
1.00
Wild-Type
1.08
ras2
1.10
sch9
1.07
ubc1(S97A)
1.09
tfs1
1.02
sch9 tfs1
1.09
ras2 tfs1
1.11
sch9 ubc1 (S97A)
1.09
tfs1 ubc1 (S97A)
1.17
ras2 tfs1...

FIG. 8A

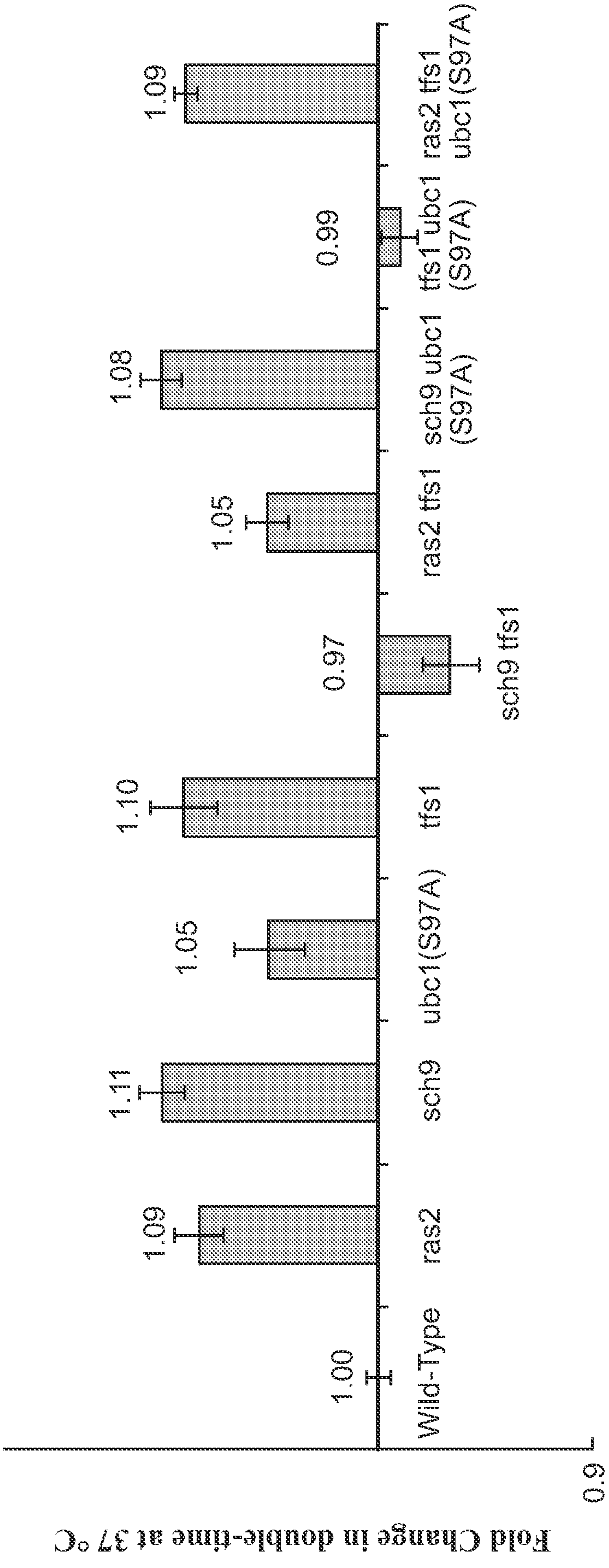

Fold Change in double-time at 37 °C
0.9
1.00
Wild-Type
1.09
ras2
1.11
sch9
1.05
ubc1(S97A)
1.10
tfs1
0.97
sch9 tfs1
1.05
ras2 tfs1
1.08
sch9 ubc1 (S97A)
0.99
tfs1 ubc1 (S97A)
1.09
ras2 tfs1 ubc1(S97A)

FIG. 8B

FIG. 8C

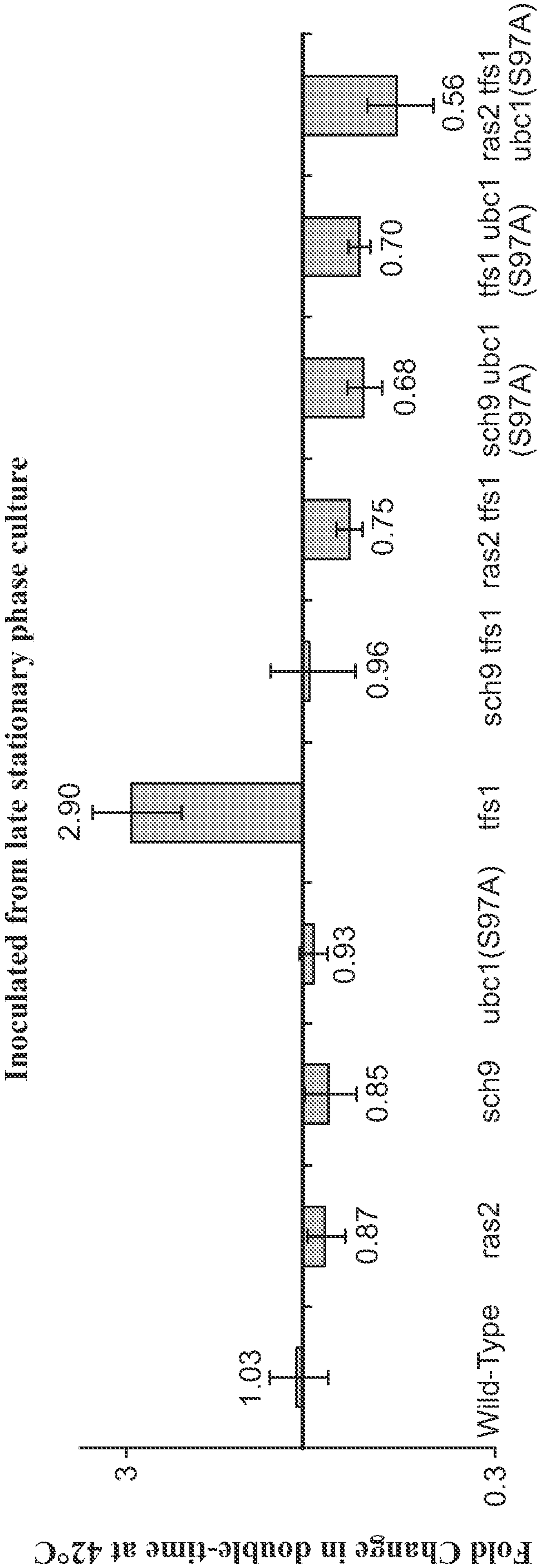

Inoculated from late stationary phase culture
Fold Change in double-time at 42°C
3
0.3
1.03
0.87
0.85
0.93
2.90
0.96
0.75
0.68
0.70
0.56
Wild-Type
ras2
sch9
ubc1(S97A)
tfs1
sch9 tfs1
ras2 tfs1
sch9 ubc1 (S97A)
tfs1 ubc1 (S97A)
ras2 tfs1 ubc1(S97A)

FIG. 8D

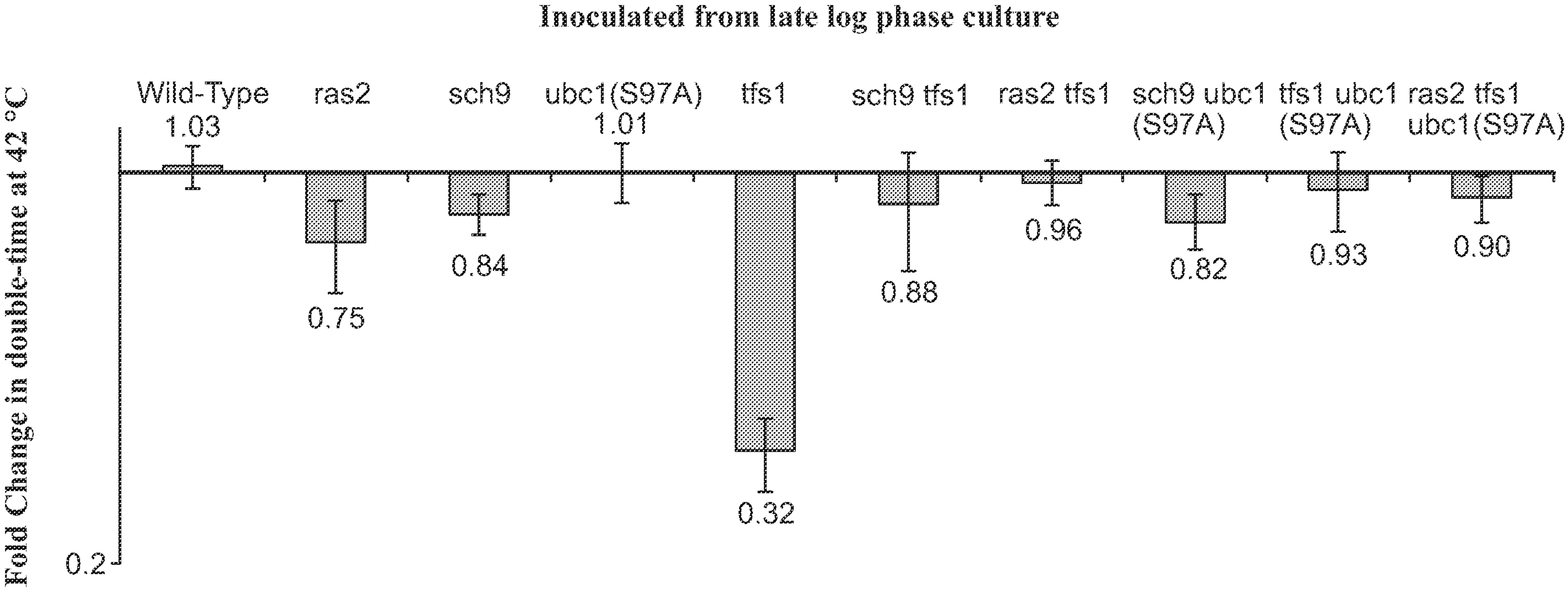

Inoculated from late log phase culture
Fold Change in double-time at 42 °C
0.2
Wild-Type
1.03
ras2
0.75
sch9
0.84
ubc1(S97A)
1.01
tfs1
0.32
sch9 tfs1
0.88
ras2 tfs1
0.96
sch9 ubc1 (S97A)
0.82
tfs1 ubc1 (S97A)
0.93
ras2 tfs1 ubc1(S97A)
0.90

Figure 1

# RNA guided Genome cleavage via Cas9