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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 November 2010 (18.11.2010)

PCT

(10) International Publication Number
WO 2010/132092 A2

(51) International Patent Classification:
C12N 15/62 (2006.01) *C07K 19/00* (2006.01)
C12N 15/52 (2006.01) *C12N 15/63* (2006.01)

(21) International Application Number:
PCT/US2010/001340

(22) International Filing Date:
6 May 2010 (06.05.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/216,103 12 May 2009 (12.05.2009) US

(71) Applicant *(for all designated States except US)*: **THE SCRIPPS RESEARCH INSTITUTE** [US/US]; 10550 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **BARBAS, Carlos, F.** [US/US]; 624 N. Granados Avenue, Solana Beach, CA 92075 (US). **KIM, Dae, Hee** [KR/KR]; 405-1210, Shin-donga-APT, Sadang-dong, Dongjak-gu, Seoul 156-771 (KR).

(74) Agents: **WANG, Hugh** et al.; The Scripps Research Institute, 10550 North Torrey Pines Road, Tpc-8, La Jolla, CA 92037 (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, 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, IS, JP, KE, KG, KM, 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, PE, PG, PH, PL, PT, RO, RS, RU, 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, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, 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, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

WO 2010/132092 A2

(54) **Title**: CYTIDINE DEAMINASE FUSIONS AND RELATED METHODS

(57) **Abstract**: The present invention provides fusion molecules comprising a cytidine deaminase polypeptide and a single stranded DNA binding protein. The invention also provides methods of using the fusion molecules to induce mutations in target genes or polynucleotide sequences. The invention further provides methods of evolving target proteins or polypeptides with the cytidine deaminase fusion molecules disclosed herein.

Cytidine Deaminase Fusions and Related Methods

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] The subject patent application claims the benefit of priority to U.S. Provisional Patent Application No. 61/216,103 (filed May 12, 2009). The full disclosure of the priority application is incorporated herein by reference in its entirety and for all purposes.

BACKGROUND OF THE INVENTION

[0002] In normal cells, a low mutation rate ensures genetic stability and this depends on effective DNA repair mechanisms for repairing the many accidental changes that occur continually in DNA. However, during the generation of antibodies, point mutations occur within the V-region coding sequence of the antigen receptor loci and the rate of mutation observed, called somatic hypermutation, is about a million times greater than the spontaneous mutation rate in other genes. Directed molecular evolution typically requires introducing mutations to the target molecule. For example, for therapeutic antibody development, in vivo somatic hypermutation is a powerful tool for the directed evolution of antibodies. High affinity antibody clones can be generated by somatic hypermutation and selection of variable regions from the initial repertoire.

[0003] There is a need in the art for better means of generating genetic diversity and evolving biologically or therapeutically useful molecules. The present invention addresses this and other needs.

SUMMARY OF THE INVENTION

[0004] In one aspect, the invention provides isolated or recombinant polynucleotides which comprise a polynucleotide sequence encoding a fusion polypeptide. The fusion polypeptide is comprised of (a) a single stranded DNA binding protein and (b) a polypeptide possessing cytidine deaminase activity or enzymatically active fragment thereof. In some preferred embodiments, the single stranded DNA binding protein in the fusion binds non-specifically to single stranded DNA. For example, the single stranded DNA binding protein can be encoded by gene V (gV) of a filamentous phage or a functional derivative thereof. In some embodiments, the polypeptide possessing cytidine deaminase activity in the fusion is a eukaryotic activation-induced cytidine deaminase (AID) or APOBEC enzyme. For example, a human or mouse AID polypeptide, or an APOBEC3 enzyme, can be employed in the

fusion. Polypeptides encoded by the fusion polynucleotide sequences, as well as expression vectors harboring the polynucleotide sequences, are also provided in the invention.

[0005] In another aspect, the invention provides methods of introducing mutations into a target gene or polynucleotide sequence. The methods typically entail contacting the target gene with a fusion polypeptide comprising (a) a polypeptide possessing cytidine deaminase activity and (b) a single stranded DNA binding protein, and allowing the cytidine deaminase to generate mutations in the target gene. Preferably, the target gene and the fusion polypeptide are contacted in a host cell that harbors the target gene and expresses the fusion polypeptide. The methods can further include culturing the host cell under conditions so that the fusion polypeptide is expressed. In some embodiments, the target gene is introduced into the host cell by transfecting or transforming a host cell with a target construct harboring the target gene. In some preferred embodiments, the target construct is a phage or a phagemid vector. In some embodiments, the fusion polypeptide is expressed from the same construct harboring the target gene. In some other embodiments, the fusion polypeptide is expressed from a second construct transfected or transformed into the host cell. In some preferred embodiments, the single stranded DNA binding protein used in the methods binds non-specifically to single stranded DNA. For example, the single stranded DNA binding protein can be the protein encoded by gene V (gV) of a filamentous phage or a functional derivative thereof. The target gene can encode any protein or a peptide, such as an enzyme, an antibody chain, a protein ligand, a receptor protein, a cytokine, or a zinc finger polypeptide.

[0006] In a related aspect, the invention provides methods of evolving a target gene in a phage mediated system. These methods involve (a) generating a phage or phagemid vector expressing the target gene fused to a phage coat protein or a protein that associates with phage, (b) introducing the phage or phagemid vector into a host cell expressing a fusion polypeptide comprising (i) a single stranded DNA binding protein and (ii) a polypeptide possessing cytidine deaminase activity or enzymatically active fragment thereof, (c) culturing the host cell under conditions to allow expression of the fusion polypeptide and production of phage particles, and (d) selecting or screening the produced phage particles for phages that harbor an evolved mutant of the target gene. In some of these methods, the selecting or screening is performed in a second host cell into which vectors encapsulated by the produced phage particles are introduced. The second host cell can be, e.g., a yeast cell, a bacterial cell or a mammalian cell. In some preferred embodiments, the single stranded DNA binding

protein binds non-specifically to single stranded DNA (e.g., a protein encoded by gene V (gV) of a filamentous phage). In these methods, the fusion polypeptide can be expressed from the same phage or phagemid vector encoding the target gene. Alternatively, the fusion polypeptide can be expressed from a second construct introduced into the host cell. The target gene can encode a peptide or a protein such as an immunoglobulin domain or an enzyme. Some of these methods are directed to evolving a target gene that encodes a single chain variable region fragment (scFv) or Fab fragment of an antibody. In some preferred embodiments, the phage coat protein utilized to fuse the target sequence is encoded by gVIII or gIII of a filamentous phage.

[0007] In a further aspect, the invention provides methods of evolving a polypeptide possessing cytidine deaminase activity. The methods entail (a) generating a phage or phagemid vector expressing a fusion protein comprising the polypeptide possessing cytidine deaminase activity and a single stranded DNA binding protein, (b) transfecting the phage or phagemid vector into a host cell, (c) culturing the host cell under conditions to allow expression of the fusion protein and production of phage particles, and (d) selecting and screening the produced phage particles for phages that harbor a genome encoding an evolved mutant of the polypeptide possessing cytidine deaminase activity. Preferably, the single stranded DNA binding protein used in these methods binds non-specifically to single stranded DNA. Some of these methods are directed to evolving an activation-induced cytidine deaminase (AID) or an APOBEC enzyme. Some of the methods are directed to identifying an evolved mutant that has stronger cytidine deaminase activity than the non-evolved polypeptide possessing cytidine deaminase activity.

[0008] A further understanding of the nature and advantages of the present invention may be realized by reference to the remaining portions of the specification and claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] Figure 1 displays AID activities on a blaS70P reporter gene substrate. Cells shown in the different panels all contain the same plasmids but with different induction conditions. mAID(-) denotes the absence of AHT induction, and blaS70P(-) denotes no transcription induction for blaS70P reporter gene. The results indicate that mAID has the highest mutation activities when mAID itself and substrate blaS70P sequence are highly expressed by 0.2μg/ml AHT and 0.5mM IPTG.

[0010] Figures 2A-2B display results from a cell-based assay for the mutation activities of mAID and gV-fusion proteins. Mutational activities were tested for various mouse/human AID and gV fusion proteins. Both mouse and human AID fusion to N terminus of gV showed better and comparable mutagenic activities to wild type mouse AID. A: Frequencies of $carb^R$ mutants generated in cell-based assay; and B: Representative culture plates. Cells were grown on carbenicillin plates.

[0011] Figures 3A-3B display phage-mediated assay for the mutation activities of mouse/human AID and gV-fusion proteins. Only mouse and human AID fusion to N terminus of gV showed mutagenic activities in phage-mediated assay, which was dependent on the expression of AID-gV fusion proteins by AHT induction. A: Mutational frequencies of mouse AID and gV fusion proteins. Higher mutation frequencies were shown from mouse AID fusion protein to N terminus of gV, which were depending on the induction of fusion proteins. '(+)' indicates AHT induction at phage production from BW310 cells and '(-)' indicates un-induced condition; and B: Comparisons of mutational frequencies between mouse AID-gV and human AID-gV fusion proteins. Human AID-gV fusion protein showed more mutant cells from phage-mediated assay when it was induced with 0.2μg /ml AHT. The control construct shown in the figure was an empty vector, pASK3CΔF1.

[0012] Figure 4 displays representative plates that show AHT dose dependencies of $carb^R$ mutational frequencies. When phages were produced from BW310 cells, various AHT concentrations were applied to induce AID-gV fusion proteins. The number of mutant ER2738 cells on $carb^R$ plates showed a correlation with the AHT concentration treated to BW310 cells. Each AHT concentrations are indicated and duplicate results are presented.

[0013] Figure 5 displays mutational frequency comparison between blaS130P and blaSSP mutant reporters. To compare the mutation activities of AID-gV fusion proteins, blaS70P, blaS130P and blaSSP reporters were constructed, which have 'Ser-to-Pro' mutations at Ser70 and/or Ser 130. From phage-mediated assay, blaSSP showed about $0.8/10^6$ reversion frequency while single mutation reporter blaS70P and bla130P showed about $2/10^4$ reversion frequencies. '(+)' indicates AHT induction when phages were produced from BW310 cells.

[0014] Figure 6 displays the construct designed for evolving the Herceptin Fab. Bla gene in ptacblaS70P F1 was replaced with Fab and gIII fusion fragments. To facilitate the enrichment of mutant Herceptin Fabs, 'ACG' translation initiation codon was introduced for

'mutant Her Fab' construct, which could restore Fab display on phages after AID generate 'C-to-U' mutation.

[0015] Figure 7 displays results of clonal screening by ELISA. After 4th panning, 25 clones from wt HER Fab and mut HER Fab selections were randomly picked and their Fabs were expressed. Using their overnight culture supernatants, binding of Fab from each clones to the antigen (Her2 extracellular domain) was assessed by competitive ELISA in the presence of wildtype Herceptin IgG. The amounts of competing Herceptin used in the assays are indicated in the figure.

[0016] Figure 8 illustrates the scheme for Herceptin panning. The scheme incorporated 3 times of mutation steps in BW310 cells. After 7^{th} panning, each clones were tested for binding by ELISA.

[0017] Figure 9 displays input and output results of the Herceptin Fab panning experiment. Because off-rate selection was performed from 3^{rd} panning round, there were not apparent increases in output/input ratio for each panning rounds.

[0018] Figure 10 displays results from clonal screening by ELISA. After 7^{th} panning, 25 clones from each vector sets were randomly picked and their Fabs were expressed. Using their overnight culture supernatants, binding of each clones were assessed against Her2 extracellular domain by competitive ELISA.

[0019] Figure 11 is a schematic presentation of the construct for evolving AID itself with AID-gV fusion. Tet promoter and hAID-gV fragments were introduced to ptacblaS70P reporter plasmid.

[0020] Figure 12 displays representative culture plates (in duplicate) for cell-based assay. The AID-gV evolution constructs showed mutations on blaS70P without AHT-induction, which can be originated from unrepressed tet promoter for AID-gV protein expression.

DETAILED DESCRIPTION OF THE INVENTION

I. Overview

[0021] The present invention is predicated in part on the studies undertaken by the present inventor of fusion proteins comprising a cytidine deaminase polypeptide (e.g., AID) and their activities in generating mutations in a target polynucleotide sequence. AID is a key molecule in the class switching recombination (CSR) and somatic hypermutation (SHM). It directs C to U mutations in the consensus sequence WRC on highly transcribed genes in

mammalian cells and also has detectable activity in *E. coli* (Muramatsu et al. Cell 102:553, 2000). AID introduces mutations at a rate of ~10^{-3} mutations per base per cell division, which is 10^6–fold higher than the spontaneous mutation rate in somatic cells (Odegard et al., Nat. Rev. Immunol. 2006, 6, 573).

[0022] As detailed in the Examples below, the present inventor observed that simple coexpression of AID in a cell producing filamentous phage was insufficient to target single stranded DNA for mutagenesis. The inventor further discovered that this problem can be solved by fusing AID to the phage gV protein, a protein that nonspecifically binds to single stranded DNAs. It was found that the resulting AID-gV fusion efficiently targeted ssDNAs for mutagenesis prior to their packaging in phage.

[0023] In accordance with these studies, the present invention provides cytidine deaminase fusion polypeptides that contain a cytidine deaminase (e.g., an AID polypeptide or an APOBEC enzyme) and a single stranded DNA binding protein (SSBP), and polynucleotide sequences that encode such fusion polypeptides. The invention also provides novel methods for introducing hypermutations into target polynucleotide sequences, and for evolving target molecules such as antibodies and enzymes. As detailed below, a target gene or polynucleotide sequence can be evolved with a cytidine deaminase-SSBP fusion molecule disclosed herein either in a cell based system or a phage-mediated system. The host cells can either transiently or stably express the cytidine deaminase-SSBP fusion. Target gene subject to mutator activity of the fusion molecule can then be directly selected and screened for mutant and improved property (e.g., enhanced enzymatic activity or binding activity). Alternatively, a target gene present in a phage or phagemid vector can be selected for evolved mutants in a phage mediated screening system. Other than these in vivo selection systems, the cytidine deaminase -SSBP fusion polypeptides disclosed herein can further be used for introducing mutations into a target polynucleotide in vitro.

[0024] The following sections provide more detailed guidance for practicing the present invention.

II. Definitions

[0025] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which this invention pertains. The following references provide one of skill with a general definition of

many of the terms used in this invention: *Academic Press Dictionary of Science and Technology*, Morris (Ed.), Academic Press (1st ed., 1992); *Oxford Dictionary of Biochemistry and Molecular Biology*, Smith et al. (Eds.), Oxford University Press (revised ed., 2000); *Encyclopaedic Dictionary of Chemistry*, Kumar (Ed.), Anmol Publications Pvt. Ltd. (2002); *Dictionary of Microbiology and Molecular Biology*, Singleton et al. (Eds.), John Wiley & Sons (3rd ed., 2002); *Dictionary of Chemistry*, Hunt (Ed.), Routledge (1st ed., 1999); *Dictionary of Pharmaceutical Medicine*, Nahler (Ed.), Springer-Verlag Telos (1994); *Dictionary of Organic Chemistry*, Kumar and Anandand (Eds.), Anmol Publications Pvt. Ltd. (2002); and *A Dictionary of Biology (Oxford Paperback Reference)*, Martin and Hine (Eds.), Oxford University Press (4th ed., 2000). In addition, the following definitions are provided to assist the reader in the practice of the invention.

[0026] Standard techniques are used for molecular, genetic and biochemical methods employed in the practice of the present invention. See, generally, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, N.Y., (3rd ed., 2000); Brent et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (ringbou ed., 2003); Guthrie et al., *Guide to Yeast Genetics and Molecular Biology*, Academic Press, Inc. (2002); White et al., *PCR Protocols,* Academic Press, Inc. (1990); Freshney, *Culture of Animal Cells*, Wiley-Liss (2005); and Murray et al., *Gene Transfer and Expression Protocols*, The Humana Press Inc. (1991).

[0027] The term "antibody" or "antigen-binding fragment" refers to polypeptide chain(s) which exhibit a strong monovalent, bivalent or polyvalent binding to a given antigen, epitope or epitopes. Unless otherwise noted, antibodies or antigen-binding fragments used in the invention can have sequences derived from any vertebrate, camelid, avian or pisces species. They can be generated using any suitable technology, e.g., hybridoma technology, ribosome display, phage display, gene shuffling libraries, semi-synthetic or fully synthetic libraries or combinations thereof. Unless otherwise noted, the term "antibody" as used in the present invention includes intact antibodies, antigen-binding polypeptide fragments and other designer antibodies that are described below or well known in the art (see, e.g., Serafini, J Nucl. Med. 34:533-6, 1993).

[0028] An intact "antibody" typically comprises at least two heavy (H) chains (about 50-70 kD) and two light (L) chains (about 25 kD) inter-connected by disulfide bonds. The recognized immunoglobulin genes encoding antibody chains include the kappa, lambda,

alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively.

[0029] Each heavy chain of an antibody is comprised of a heavy chain variable region (V_H) and a heavy chain constant region. The heavy chain constant region is comprised of three domains, C_{H1}, C_{H2} and C_{H3}. Each light chain is comprised of a light chain variable region (V_L) and a light chain constant region. The light chain constant region is comprised of one domain, C_L. The variable regions of the heavy and light chains contain a binding domain that interacts with an antigen. The constant regions of the antibodies may mediate the binding of the immunoglobulin to host tissues or factors, including various cells of the immune system and the first component (Clq) of the classical complement system.

[0030] The V_H and V_L regions of an antibody can be further subdivided into regions of hypervariability, also termed complementarity determining regions (CDRs), which are interspersed with the more conserved framework regions (FRs). Each V_H and V_L is composed of three CDRs and four FRs, arranged from amino-terminus to carboxyl-terminus in the following order: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4. The locations of CDR and FR regions and a numbering system have been defined by, e.g., Kabat *et al.*, *Sequences of Proteins of Immunological Interest*, U.S. Department of Health and Human Services, U.S. Government Printing Office (1987 and 1991).

[0031] Antibodies to be used in the invention also include antibody fragments or antigen-binding fragments which contain the antigen-binding portions of an intact antibody that retain capacity to bind the cognate antigen. Examples of such antibody fragments include (i) a Fab fragment, a monovalent fragment consisting of the V_L, V_H, C_L and C_{H1} domains; (ii) a $F(ab')_2$ fragment, a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region; (iii) a Fd fragment consisting of the V_H and C_{H1} domains; (iv) a Fv fragment consisting of the V_L and V_H domains of a single arm of an intact antibody; (v) disulfide stabilized Fvs (dsFvs) which have an interchain disulfide bond engineered between structurally conserved framework regions; (vi) a single domain antibody (dAb) which consists of a V_H domain (see, e.g., Ward et al., Nature 341:544-546, 1989); and (vii) an isolated complementarity determining region (CDR).

[0032] Antibodies suitable for practicing the present invention also encompass single chain antibodies. The term "single chain antibody" refers to a polypeptide comprising a V_H domain and a V_L domain in polypeptide linkage, generally linked via a spacer peptide, and which may comprise additional domains or amino acid sequences at the amino- and/or carboxyl-termini. For example, a single-chain antibody may comprise a tether segment for linking to the encoding polynucleotide. As an example, a single chain variable region fragment (scFv) is a single-chain antibody. Compared to the V_L and V_H domains of the Fv fragment which are coded for by separate genes, a scFv has the two domains joined (e.g., via recombinant methods) by a synthetic linker. This enables them to be made as a single protein chain in which the V_L and V_H regions pair to form monovalent molecules.

[0033] Antibodies that can be used in the practice of the present invention also encompass single domain antigen-binding units which have a camelid scaffold. Animals in the camelid family include camels, llamas, and alpacas. Camelids produce functional antibodies devoid of light chains. The heavy chain variable (V_H) domain folds autonomously and functions independently as an antigen-binding unit. Its binding surface involves only three CDRs as compared to the six CDRs in classical antigen-binding molecules (Fabs) or single chain variable fragments (scFvs). Camelid antibodies are capable of attaining binding affinities comparable to those of conventional antibodies.

[0034] The various antibodies or antigen-binding fragments described herein can be produced by enzymatic or chemical modification of the intact antibodies, or synthesized de novo using recombinant DNA methodologies, or identified using phage display libraries. Methods for generating these antibodies or antigen-binding molecules are all well known in the art. For example, single chain antibodies can be identified using phage display libraries or ribosome display libraries, gene shuffled libraries (see, e.g., McCafferty et al., Nature 348:552-554, 1990; and U.S. Pat. No. 4,946,778). In particular, scFv antibodies can be obtained using methods described in, e.g., Bird et al., Science 242:423-426, 1988; and Huston et al., Proc. Natl. Acad. Sci. USA 85:5879-5883, 1988. Fv antibody fragments can be generated as described in Skerra and Plückthun, Science 240:1038-41, 1988. Disulfide-stabilized Fv fragments (dsFvs) can be made using methods described in, e.g., Reiter et al., Int. J. Cancer 67:113-23, 1996. Similarly, single domain antibodies (dAbs) can be produced by a variety of methods described in, e.g., Ward et al., Nature 341:544-546, 1989; and Cai and Garen, Proc. Natl. Acad. Sci. USA 93:6280-85, 1996. Camelid single domain antibodies

can be produced using methods well known in the art, e.g., Dumoulin et al., Nature Struct. Biol. 11:500–515, 2002; Ghahroudi et al., FEBS Letters 414:521–526, 1997; and Bond et al., J Mol Biol. 332:643-55, 2003. Other types of antigen-binding fragments (e.g., Fab, $F(ab')_2$ or Fd fragments) can also be readily produced with routinely practiced immunology methods. See, e.g., Harlow & Lane, *Using Antibodies, A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1998.

[0035] Binding affinity is generally expressed in terms of equilibrium association or dissociation constants (K_a or K_d, respectively), which are in turn reciprocal ratios of dissociation and association rate constants (k_d and k_a, respectively). Thus, equivalent affinities may correspond to different rate constants, so long as the ratio of the rate constants remains the same.

[0036] The term "contacting" has its normal meaning and refers to combining two or more agents (e.g., a compound and a phage-displayed polypeptide) or combining agents and cells. Contacting as used herein can occur in vitro, e.g., mixing a library of phages with a binding partner or mixing a phage displayed polypeptide with a library of candidate agents in a test tube or other container. Contacting can also occur in vivo, e.g., between a protein or polypeptide and another molecule such as a polynucleotide in a host cell.

[0037] As used herein, the term "a cytidine deaminase polypeptide" or "a polypeptide possessing cytidine deaminase activity" broadly refers to any enzyme (or functional fragments thereof) that is capable of catalyzing the irreversible hydrolytic deamination of cytidine and deoxycytidine to uridine and deoxyuridine, respectively (cytidine deaminase activity). Typically, it encompasses members of enzymes in the cytidine deaminase superfamily, and in particular, enzymes of the AID/APOBEC family. Members of the AID/APOBEC enzyme family include activation-induced deaminase (AID) and APOBEC1, APOBEC2, APOBEC4, and APOBEC3 subgroups of enzymes (see, e.g., Conticello et al., Mol. Biol. Evol. 22:367-77, 2005; and Conticello, Genome Biol. 9:229, 2008). The cytidine deaminase superfamily additionally includes cytidine deaminases and CMP deaminases (Muramatsu et al., J. Biol. Chem. 274:18470-6, 1999).

[0038] Somatic hypermutation (or SHM) is a mechanism inside cells that is part of the way the immune system adapts to the new foreign elements which confront it (for example, microbes). SHM involves a programmed process of mutation affecting the variable regions of

immunoglobulin light chain and heavy genes. Unlike many other types of mutation, SHM affects only individual immune cells, and the mutations are not transmitted to offspring.

[0039] "Class switching" is the recombination process in V gene rearrangement that leads to a change in the constant region of the expressed antibody. "Gene conversion" is an additional mechanism in the recombination process which is found to occur in chicken and rabbits (but not in human or mouse) and contributes to V gene diversification.

[0040] A functional derivative of a reference polypeptide (e.g., a cytidine deaminase polypeptide or a SSBP) usually has amino acid deletions and/or insertions and/or substitutions while maintaining one or more of the bioactivities (e.g., cytidine deaminase activity of AID or ssDNA binding activity of an SSBP) and therefore can also be used in practicing the present invention.

[0041] A "fusion" protein or polypeptide refers to a polypeptide comprised of at least two polypeptides and a linking sequence or a linkage to operatively link the two polypeptides into one continuous polypeptide. The two polypeptides linked in a fusion polypeptide are typically derived from two independent sources, and therefore a fusion polypeptide comprises two linked polypeptides not normally found linked in nature.

[0042] "Linkage" refers to means of operably or functionally connecting two biomolecules (e.g., polypeptides or polynucleotides encoding two polypeptides), including, without limitation, recombinant fusion, covalent bonding, disulfide bonding, ionic bonding, hydrogen bonding, and electrostatic bonding. "Fused" refers to linkage by covalent bonding. A "linker" or "spacer" refers to a molecule or group of molecules that connects two biomolecules, and serves to place the two molecules in a preferred configuration with minimal steric hindrance.

[0043] The term "mutagenesis" or "mutagenizing" refers to a process of introducing changes (mutations) to the base pair sequence of a coding polynucleotide sequence and consequential changes to its encoded polypeptide. Unless otherwise noted, the term as used herein refers to mutations artificially introduced to target molecules via the cytidine deaminase activity of a cytidine deaminase polypeptide described herein.

[0044] The term "mutation activity" or "mutator activity" of a cytidine deaminase (e.g., AID or APOBEC) refers to the enzymatic activity of the cytidine deaminase or its homologues to increase mutation frequency of a target gene or polynucleotide above

background. Mutation frequency is the frequency at which mutants of the target molecules are found in a population of the target molecules.

[0045] The term "operably linked" when referring to a nucleic acid, means a linkage of polynucleotide elements in a functional relationship. A nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For instance, a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the coding sequence. Operably linked means that the DNA sequences being linked are typically contiguous and, where necessary to join two protein coding regions, contiguous and in reading frame.

[0046] The term "polynucleotide" or "nucleic acid" as used herein refers to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides, that comprise purine and pyrimidine bases, or other natural, chemically or biochemically modified, non-natural, or derivatized nucleotide bases. Polynucleotides of the embodiments of the invention include sequences of deoxyribopolynucleotide (DNA), ribopolynucleotide (RNA), or DNA copies of ribopolynucleotide (cDNA) which may be isolated from natural sources, recombinantly produced, or artificially synthesized. A further example of a polynucleotide is polyamide polynucleotide (PNA). The polynucleotides and nucleic acids may exist as single-stranded or double-stranded. The backbone of the polynucleotide can comprise sugars and phosphate groups, as may typically be found in RNA or DNA, or modified or substituted sugar or phosphate groups. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and nucleotide analogs. The sequence of nucleotides may be interrupted by non-nucleotide components. The polymers made of nucleotides such as nucleic acids, polynucleotides and polynucleotides may also be referred to herein as nucleotide polymers.

[0047] Polypeptides are polymer chains comprised of amino acid residue monomers which are joined together through amide bonds (peptide bonds). The amino acids may be the L-optical isomer or the D-optical isomer. In general, polypeptides refer to long polymers of amino acid residues, e.g., those consisting of at least more than 10, 20, 50, 100, 200, 500, or more amino acid residue monomers. However, unless otherwise noted, the term polypeptide as used herein also encompass short peptides which typically contain two or more amino acid monomers, but usually not more than 10, 15, or 20 amino acid monomers.

[0048] Proteins are long polymers of amino acids linked via peptide bonds and which may be composed of two or more polypeptide chains. More specifically, the term "protein" refers to a molecule composed of one or more chains of amino acids in a specific order; for example, the order as determined by the base sequence of nucleotides in the gene coding for the protein. Proteins are essential for the structure, function, and regulation of the body's cells, tissues, and organs, and each protein has unique functions. Examples are hormones, enzymes, and antibodies. In some embodiments, the terms polypeptide and protein may be used interchangeably.

[0049] The term "target molecule" or "target polypeptide" refers to a molecule of interest in which mutations are intended to be generated. Preferably, the target molecule for practicing the present invention is a polypeptide (e.g., an enzyme or an antibody chain). In the practice of the present invention, a target molecule is mutated by the cytidine deaminase activity of a cytidine deaminase polypeptide in a gene or polynucleotide sequence encoding the target molecule ("target gene").

[0050] A cell has been "transformed" by exogenous or heterologous polynucleotide when such polynucleotide has been introduced inside the cell. The transforming polynucleotide may or may not be integrated (covalently linked) into the genome of the cell. In prokaryotes, yeast, and mammalian cells for example, the transforming polynucleotide may be maintained on an episomal element such as a plasmid. With respect to eukaryotic cells, a stably transformed cell is one in which the transforming polynucleotide has become integrated into a chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the transforming polynucleotide. A "clone" is a population of cells derived from a single cell or common ancestor by mitosis. A "cell line" is a clone of a primary cell that is capable of stable growth in vitro for many generations.

[0051] A "variant" of a reference polypeptide (e.g., a cytidine deaminase polypeptide such as an AID polypeptide) refers to a molecule which has a structure that is derived from or similar to that of the reference polypeptide. Typically, the variant is obtained by mutagenesis of the reference polypeptide in a controlled or random manner. Methods for performing mutagenesis of a polypeptide are well known in the art, e.g., site-specific mutagenesis, error-prone PCR, restriction digestion and relegation, and polynucleotide shuffling.

[0052] A "vector" or "construct" is a replicon, such as plasmid, phage or cosmid, to which another polynucleotide segment may be attached so as to bring about the replication of the attached segment. Vectors capable of directing the expression of genes encoding for one or more polypeptides are referred to as "expression vectors".

III. Polypeptides possessing cytidine deaminase activity

[0053] The invention provides fusion proteins or polypeptides that comprise a polypeptide possessing cytidine deaminase activity (a cytidine deaminase polypeptide). The cytidine deaminase-SSBP fusion polypeptides are capable of introducing somatic hypermutations to a target polynucleotide sequence. Various cytidine deaminase polypeptides (or "proteins possessing cytidine deaminase activity") can be used in the present invention. A preferred example is activation-induced deaminase (AID). While the invention is illustrated in some detail herein using AID as an example of "cytidine deaminase polypeptides" (or "polypeptides possessing cytidine deaminase activity"), other polypeptides or enzymes with similar activities can also be employed in the fusion proteins of the invention. Thus, based on the subject disclosure, a skilled artisan can readily appreciate that suitable polypeptides possessing cytidine deaminase activity include other enzymes in the AID/APOBEC family and the cytidine deaminase superfamily. Enzymes in the AID/APOBEC family include AID, APOBEC1, APOBEC2, APOBEC3 and APOBEC4 subgroups of enzymes. Members of APOBEC3 enzymes include APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D, APOBEC3F, and APOBEC3G. In addition to AID and the APOBEC family of enzymes, the cytidine deaminase superfamily also broadly encompasses other RNA-editing deaminases such as cytidine deaminases (CDAs) and CMP deaminases (see, e.g., Muramatsu et al., J. Biol. Chem. 274:18470-6, 1999; and Demontis et al., Biochim. Biophys. Acta. 1443:323–33, 1999).

[0054] In some preferred embodiments, the polypeptide possessing cytidine deaminase activity employed in the present invention is an enzyme from the AID/APOBEC enzyme family. Members of the AID/APOBEC family are capable of deaminating cytidine in RNA and/or DNA and exhibiting diverse physiological functions (see, e.g., Conticello et al., Mol. Biol. Evol. 22:367–377, 2005). For example, AID and APOBEC3 enzymes can deaminate DNA to trigger pathways in adaptive and innate immunity. Structures and functions of these molecules have been characterized in the art. See, e.g., Wedekind et al., Trends Genet.

19:207–16, 2003; Mariani et al., Cell 114:21–31, 2003; Bogerd et al., Nucleic Acids Res. 34:89–95, 2006; Yu et al., J. Biol. Chem. 279:53379–86, 2005; Hakata et al., J. Biol. Chem. 281:36624–31, 2007; Zheng et al., J. Virol. 78:6073–6, 2004; Hirano et al., J. Lipid Res. 38: 847–59, 1997; Liao et al., Biochem. Biophys. Res. Commun. 260:398–404, 1999; and Rogozin et al., Cell Cycle 4:1281-5, 2005.

[0055] AID is specifically expressed in germinal center B cells and catalyzes the deamination of cytidine residues in single-stranded DNA (ssDNA), thereby producing U:G mismatches which can be repaired by alternative DNA repair pathways to produce either SHM or CSR. AID can bind ssDNA and RNA *in vitro* but it only deaminates ssDNA. It has been known that APOBEC1 is also able to deaminate ssDNA *in vitro*, although with low efficiency. Whereas ssDNA is deaminated by AID *in vitro*, double-stranded DNA (dsDNA) substrates can only serve as substrates when transcribed and the substrate strand is the nontemplate strand. Similarly in *E. coli*, it is the nontemplate strand which is preferentially mutated by AID. On the other hand, unlike results from biochemical and *E. coli* experiments, mutations are observed on both strands of Ig genes *in vivo*.

[0056] The present invention can employ a polypeptide possessing cytidine deaminase activity or a variant, homolog or a functionally active fragment thereof from various sources. AID and APOBEC enzymes have been identified in many eukaryotic species (see, e.g., Conticello et al., Mol. Biol. Evol. 22:367–377, 2005). For example, AID is highly conserved from cartilaginous fish to humans. Sequencing and functional characterization of many of these AID proteins have been reported in the art. For example, the sequence of human AID and its functional characterization were reported in Muramatsu et al., J. Biol. Chem. 274:18470-6, 1999; Muto et al., Genomics 68:85-88, 2000; Muramatsu et al., Cell 102:553-63, 2000; and Muto et al., Cell 102:565-575, 2000. Mouse AID gene is reported and characterized in Muramatsu et al., J. Biol. Chem. 274:18470-6, 1999; and Wakae et al., Int. Immunol. 18:41-7, 2006. AID genes from a number of other eukaryotic species are disclosed in, e.g., Zhao et al., Dev. Comp. Immunol. 29:61–71, 2005 (zebrafish AID); Saunders et al., Dev. Comp. Immunol. 28:657–663, 2004 (channel fish AID); and Conticello et al., Mol. Biol. Evol. 22:367–377, 2005 (dogfish AID). The genomic and cDNA sequences of AID, as well as the enclosed polypeptide sequences, are also available from GenBank. For example, human AID sequences have accession numbers AB040431, AB040430 and BAB12721. Mouse AID accession numbers are AF132979, AB091291 and AAD41793. Accession

numbers for AID from various other species include, e.g., AB122019, AY436507, AB159559 and NM_001095712.

[0057] The domain structure of AID has also been delineated in the art. See, e.g., Durand et al., Hum. Mutat. 27:1185-91, 2006; and Xie et al., Proc. Natl. Acad. Sci. U.S.A. 101:8114-9, 2004. For example, as is well known in the art, the human AID gene encodes a 198-amino acid 24-kDa protein that contains a functional cytidine deaminase domain (amino acid residues 55-94). Other than the cytidine deaminase domain, AID also contains a putative nuclear localization signal (NLS) at the N-terminus, and a nuclear export signal (NES) at the C-terminus. In addition, a leucine-rich region at the C-terminus of the protein appears to be involved in protein-protein interaction (Durand et al., Hum. Mutat. 27:1185-91, 2006). Further, it has been reported that the C-terminal region (residues 189-198) of AID is required for CSR but dispensable for SHM and therefore targeting to Ig loci, and the N-terminal region of AID is more important for SHM than for CSR.

[0058] To make and use the fusion proteins or polypeptides of the invention, any of the genes or cDNA coding sequences described herein or well known in the art that encode a full length protein possessing cytidine deaminase activity (e.g., AID or APOBEC3) can be utilized. In addition to full length sequences, a functional or enzymatically active fragment of the protein (e.g., AID) or domain sequence that harbors the cytidine deaminase activity may also be employed. Using AID as an example, a partial AID sequence encoding the cytidine deaminase domain can be fused to a ssDNA-binding protein. Other than the wildtype AID sequences, variants of the AID homologs from various species can also be used. AID variants suitable for the invention include sequences that encode the AID proteins disclosed herein but that differ conservatively because of the degeneracy of the genetic code. They also encompass sequences that are substantially identical to an AID gene or cDNA sequence disclosed herein. Thus, suitable AID variants encompass natural allelic variants, e.g., due to sequence polymorphisms that lead to changes in the amino acid sequences of the encoded proteins. Such variants may exist among individuals within a population, e.g., the human population. The natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the AID gene. To practice the present invention, the nucleotide sequence polymorphism or amino acid sequence variance of the employed AID molecule should do not alter the functional activity of AID.

[0059] Suitable variants of a protein possessing cytidine deaminase activity (e.g., AID or APOBEC3) further include synthetically derived nucleotide sequences that have been generated, for example, by using site-directed mutagenesis but which still harbor the enzymatic activities of the protein as disclosed herein. These include, e.g., variants which have conservative amino acid substitutions at one or more nonessential residues that do not lead to an alteration in the cytidine deaminase activity. A conservative amino acid substitution is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Generally, nucleotide sequence variants of the invention will have at least about 45%, 55%, 65%, 75%, 85%, 95%, or 98% identity to the nucleotide sequence encoding a particular polypeptide possessing cytidine deaminase activity as disclosed herein (e.g., human or mouse AID sequence). Typically, such a variant polynucleotide sequence encodes an polypeptide that has an amino acid sequence having at least about 45%, 55%, 65%, 75%, 85%, 95%, or 98% identity to the amino acid sequence of the polypeptide possessing cytidine deaminase activity disclosed herein.

[0060] Enzymatic activities of a homolog, variant or functional fragment of a polypeptide possessing cytidine deaminase activity (e.g., AID) can be readily assayed by measuring its cytidine deaminase activity in vitro or monitoring hypermutations generated in a target reporter gene. Such assays are described and exemplified in the Examples below. Similar assays have been described and routinely practiced in the art. See, e.g., Muramatsu et al., J. Biol. Chem. 274:18470–6, 1999. For example, the enzymatic activity of a cytidine deaminase variant or fragment can be determined using a thin layer chromatographic (TLC) assay as detailed in MacGinnitie et al., J. Biol. Chem. 270:14768-75, 1995.

IV. Fusions containing a cytidine deaminase polypeptide and ssDNA-binding protein

[0061] The fusion molecules of the invention typically comprise (i) a polypeptide possessing cytidine deaminase activity (e.g., an AID polypeptide) and (ii) a polypeptide or protein that binds to single stranded DNA (ssDNA). As noted above, the polypeptide possessing cytidine deaminase activity in the fusion products can be a full length cytidine deaminase enzyme (e.g., AID or an APOBEC enzyme), variant or homolog, as well as a functional domain or fragment thereof. Typically, the ssDNA binding polypeptide or protein (SSBP) employed in the invention can bind non-specifically to any ssDNA sequences. This is opposed to some other DNA binding proteins or domains which recognize only specific polynucleotide sequences. Thus, the SSBPs encompassed by the AID fusions of the invention do not include, e.g., transcription factors that bind to DNA transcription regulatory elements of a target gene or that bind to specific response element sequences of a target gene (e.g., heat shock response elements and hormone response elements). They also do not encompass DNA-binding domains or motifs such as the zinc finger, the helix-turn-helix, and the leucine zipper protein.

[0062] Proteins that non-specifically bind to ssDNAs are well known in the art. Some of these proteins function by binding to single stranded regions of DNA to prevent premature reannealing. The DNA strands have a natural tendency to revert to the duplex form, but SSBPs bind to the single strands, keeping them separate and allowing the DNA replication machinery to perform its function. SSBPs have been identified in organisms from viruses to humans. Some SSBPs (e.g., many phage and viral SSBPs) function as monomers or dimers, while eukaryotes tend to encode heterotrimeric RPA (Replication Protein A). Most bacterial SSBPs exist as a tetramer. One example is the single stranded DNA binding protein of *E. coli* , SSB (Krauss et al., Biochem. 20:5346-5352, 1981; and Weiner et al., J. Biol. Chem. 250:1972-1980, 1975). This tetrameric protein binds with high affinity in a cooperative manner to ssDNA and does not bind well to double-stranded DNA. It is involved in DNA replication and recombination *in vivo*. This protein has also been used in conjunction with RecA protein for carrying out site-directed mutagenesis (Shortle et al., Proc. Natl. Acad. Sci. USA 77:5375-5379, 1980). Any of these SSBPs well known in the art can be employed in the practice of the present invention.

[0063] Some preferred embodiments of the invention employ a SSBP that is from a filamentous phage such as M13, fd, f1, or Ike. For example, the genome of M13 phage encodes a ssDNA binding protein that can be used in the practice of the present invention

(see, e.g., Alma et al., Biochem. 22:2104-15, 1983; and Alma et al., J. Mol. Biol. 163:47-62, 1983; Bulsink et al., Eur. J. Biochem. 176:597-60; and Bulsink et al., Eur. J. Biochem. 176:589-596, 1988). The protein, encoded by gene V (gV) of filamentous phage, is a 87 amino acid homodimer that binds cooperatively to single-stranded DNA. An active gV protein is required for propagation of the phage. The primary function of gV protein is to bind to ssDNA, regulating the switch from double-stranded RF form to the production of ssDNA form. In addition to M13 phage, SSBPs encoded by gV of the other filamentous phages are all well known in the art. Any of these SSBPs can be readily employed in the present invention. These include, e.g., gV of the F1 phage (Michel et al., Nucleic Acids Res. 17:7333 – 44, 1989), gV of the Ike phage (van Duynhoven et al., Biochem. 31:1254-62, 1992), and gV of Fd phage (Sang et al., J. Biotnol. Strucl. & Dyn. 7:693-706, 1987). It should be noted that the gene V protein has been extensively characterized structurally and functionally in the art (see, e.g., Guan et al., Prot. Sci. 4:187-197, 1995). Therefore, other than an intact gV sequence, functional derivatives of a gV sequence (e.g., variant sequences, fragments or substantially identical sequences that possess ssDNA binding activity) can also be used in the present invention.

[0064] The method for generating the fusion protein of the invention is not subject to any particular limitation. The fusion protein of the invention may be a fusion protein synthesized by chemical synthesis, or a recombinant fusion protein produced by a genetic engineering technique. If the fusion protein of the invention is to be chemically synthesized, synthesis may be carried out by, for example, the Fmoc (fluorenylmethyloxycarbonyl) process or the tBoc (t-butyloxycarbonyl) process. In addition, peptide synthesizers available from, for example, Advanced ChemTech, PerkinElmer, Pharmacia, Protein Technology Instrument, Syntheceh-Vega, PerSeptive and Shimadzu Corporation may be used for chemical synthesis. If the fusion protein of the invention is to be produced by a genetic engineering technique, production may be carried out using the conventional recombination techniques routinely practiced in the art. Such techniques are described, e.g., in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, N.Y., (3rd ed., 2000); and Brent et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (ringbou ed., 2003). As detailed below, the fusion protein can be produced by inserting a polynucleotide (e.g., DNA) encoding the fusion protein into a suitable expression system.

[0065] In some preferred embodiments, the AID-SSBP fusion proteins of the invention are generated in accordance with the routinely practiced recombination technology. Some specific exemplifications are discussed in detail in the Examples below. Typically, the methods involve removing the stop codon from a polynucleotide sequence (e.g., a cDNA sequence) coding for the cytidine deaminase polypeptide (e.g., an AID polypeptide), then appending a polynucleotide sequence (e.g., a cDNA sequence) encoding the SSBP in frame through ligation or overlap extension PCR. To ensure proper folding and maintain the biological activities of the fusion partners, a linker or spacer peptides is often used for linking the two components of the fusion proteins. The fusion protein of the invention may additionally include a peptide sequence for purification. Peptide sequences for purification that may be used are also known in the art. Examples of peptide sequences for purification include histidine tag sequences having an amino acid sequence in which at least four, and preferably at least six, continuous histidine residues, and the amino acid sequence of the glutathione-binding domain in glutathione S-transferase.

[0066] Any of the above-described cytidine deaminase polypeptides, homologs, variants, or enzymatically active domains or fragments can be used to produce the cytidine deaminase-SSBP fusions of the invention. The SSBP can be fused at either the N- or C-terminus of the cytidine deaminase polypeptide. In some preferred embodiments, the SSBP is fused or operably linked to the C-terminus of the cytidine deaminase polypeptide in the fusion protein. By operably linking the SSBP to the cytidine deaminase polypeptide, the enzymatic activity (e.g., its deaminase activity) of the cytidine deaminase polypeptide is not affected by the SSBP, and the ssDNA-binding activity of the SSBP is also not substantially altered by its fusion to the cytidine deaminase polypeptide.

[0067] In addition to the cytidine deaminase-SSBP fusion proteins disclosed above, related embodiments of the invention also include polynucleotide sequences that encode such fusions, expression constructs for expressing the fusion proteins, and host cells that harbor the polynucleotides or expression constructs. The polynucleotide sequences of the invention can be any polynucleotide having a nucleotide sequence that encodes the fusion protein of the invention, although DNA is preferred. The recombinant constructs or expression vectors of the invention harbor a polynucleotide sequence of the invention that encodes a cytidine deaminase-SSBP fusion polypeptide. The recombinant constructs of the invention may be obtained by ligating (inserting) the polynucleotide (DNA) of the invention into a suitable

vector. More specifically, the recombinant vector may be obtained by cleaving purified polynucleotide (DNA) with a suitable restriction enzyme, then inserting the cleaved polynucleotide to a restriction enzyme site or multicloning site on a suitable vector, and ligating the polynucleotide to the vector. The vector for inserting the polynucleotide sequence is not subject to any particular limitation, provided it is capable of replication in an appropriate host. The expression vectors of the invention are not subject to any particular limitation, and may be, for example, bacteriophages, plasmids, cosmids or phagemids. Examples of recombinant bacteriophage or phagemid vectors include that based on a filamentous phage such as M13. Plasmid vectors include those based on plasmids from, e.g., E. coli (e.g., pBR322, pBR325, pUC118 and pUC119), plasmids from Bacillus subtilis (e.g., pUB110 and pTP5), and plasmids from yeasts (e.g., YEp13, YEp24 and YCp50). The expression vectors can also include animal viruses such as retroviruses, vaccinia viruses and insect viruses (e.g., baculoviruses).

[0068] In the expression constructs of the invention, the polynucleotide encoding the cytidine deaminase-SSBP fusion is generally ligated downstream from the promoter in a suitable vector in such a way as to be expressible. For example, if the host during transformation is an animal cell, preferred promoters include promoters from SV40, retrovirus promoters, metallothionein promoters, heat shock promoters, cytomegalovirus promoters and the SRα promoter. If the host is a genus Escherichia organism, preferred promoters include the tetracycline promoter, the Trp promoter, the T7 promoter, the lac promoter, the recA promoter, the λ promoter and the lpp promoter. If the host is a genus Bacillus organism, preferred promoters include the SPO1 promoter, the SPO2 promoter and the penP promoter. If the host is a yeast, preferred promoters include the PHO5 promoter, the PGK promoter, the GAP promoter, the ADH1 promoter and the GAL promoter. If the host is an insect cell, preferred promoters include the polyhedrin promoter and the P10 promoter.

[0069] In addition to the above, the recombinant vector used in the invention may contain, if desired, an enhancer, a splicing signal, a poly(A) addition signal, a ribosome binding sequence (SD sequence), a selective marker and the like. Examples of selective markers include the tetracycline resistance gene, the carbencillin resistance gene, the dihydrofolate reductase gene, the ampicillin resistance gene and the neomycin resistance gene. The recombinant vector of the invention may additionally include a polynucleotide

having a nucleotide sequence encoding an amino acid sequence for enhancing translation and/or a polynucleotide having a nucleotide sequence encoding a peptide sequence for purification. For example, the vectors can employ a translational enhancer element (TEE) sequence (see, e.g., Batten et al., FEBS Lett. 580:2591-7, 2006).

[0070] The invention further provides host cells that express the cytidine deaminase-SSBP fusion polypeptides described herein. The host cells are genetically engineered (transduced, transformed or transfected) with the recombinant constructs or expression vectors disclosed herein for production of the fusion protein or examination of its activity. To generate such cells, a recombinant construct which harbors and expresses a cytidine deaminase-SSBP fusion sequence can be introduced into a suitable host. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying particular genes such as the fusion gene encoding a cytidine deaminase-SSBP fusion polypeptide. The culture conditions for particular host cells selected for expression, such as temperature, pH and the like, will be readily apparent to the ordinarily skilled artisan. In some embodiments, the cytidine deaminase-SSBP fusion sequence is stably integrated into the chromosome of the host cells. With such host cells, the cytidine deaminase-SSBP sequence and its expression are substantially maintained in successive generations of cells. They are distinguished from host cells which transiently express the fusion polypeptide as detailed herein.

[0071] The host cell for production or expression of a construct of the invention, for example, can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. The selection of an appropriate host is within the scope of those skilled in the art and also exemplified in the Examples herein. Preferably, the host cell employed is suitable for expression of the cytidine deaminase-SSBP fusion as well as induction of hypermutations in the target gene or polynucleotide by the cytidine deaminase polypeptide. Representative examples of appropriate host cells suitable for practicing the present invention include, but need not be limited to, bacterial cells, such as E. coli, Streptomyces, Salmonella tvphimurium; fungal cells, such as yeast; insect cells, such as Drosophila S2 and Spodoptera Sf9; animal cells, such as CHO, COS or 293 cells; adenoviruses; plant cells, or any suitable cell already adapted to in vitro propagation or so established de novo. If the expression construct is a phage or a phagemid vector, many suitable bacterial host cells can be used, e.g.,

the *E. coli* ER2738 cell line as detailed in the Examples below. Another example of such host cells is E. coli strain BW310 as exemplified in the Examples below. This cell line is available from the Coli Genetic Stock Center at Yale University (New Haven, CT). BW310 cell doesn't have ung gene (i.e., ung- genotype). The ung gene encodes uracil-DNA glycosylase which prevents mutagenesis by eliminating from DNA molecules uracil bases produced by cytosine deamination or misincorporation of dUMP residues. This BW310 cell line has been routinely employed in the art to study expression and DNA mutator activity of AID or its orthologs (see, e.g., Ichikawa et al., J. Immunol. 177:355-361, 2006; and Haché et al., J. Biol. Chem. 280:10920-4, 2005).

[0072] Introduction of the vector or expression construct into the host cell can be effected by a variety of methods with which those skilled in the art will be familiar, including but not limited to, for example, calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (see, e.g., Brent et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (ringbou ed., 2003)). Expression and, if desired, purification, of a cytidine deaminase-SSBP fusion polypeptide in a transfected or transformed host cell can be carried out in accordance with any of the routinely practiced methods in the art, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, N.Y., (3rd ed., 2000); and Brent et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (ringbou ed., 2003). Typically, to produce the fusion protein of the invention, the host cell harboring the expression vector is cultured under appropriate conditions that allow the polynucleotide (DNA) encoding the fusion protein to be expressed, thereby inducing formation and accumulation of the fusion polypeptide, then isolating and purifying the fusion polypeptide. The fusion protein expressed in the host cell can be readily isolated and purified. Specifically, when the fusion protein of the invention accumulated within cultured bacteria or within cultured cells, following the completion of cultivation, an extract of the fusion protein of the invention may be obtained by a conventional method such as centrifugation or filtration after using a conventional technique (e.g., ultrasound, lysozymes, freezing and thawing) to disrupt the bacteria or cells. When the cytidine deaminase-SSBP fusion polypeptide accumulates in the periplasmic space, following the completion of cultivation, an extract containing the target protein may be obtained by a conventional method such as osmotic shock. When the fusion protein of the invention accumulates in the culture broth, following the completion of cultivation, a culture

supernatant containing the inventive fusion protein may be obtained by using a conventional method such as centrifugation or filtration to separate the culture supernatant from the bacteria or cells.

V. Evolving a target polynucleotides with cytidine deaminase-SSBP fusion

[0073] Using the cytidine deaminase-SSBP fusion polypeptides described herein, the invention provides methods of introducing hypermutations into a target gene or polynucleotide sequence. This can be accomplished in vitro, or more preferably, in vivo with either a cell based system or a phage based system described herein. Mutants of target molecules thus evolved can be further selected and screened in accordance with the present disclosure or methods well known in the art. In some embodiments, mutations in the target gene is generated via the presence or expression of one cytidine deaminase-SSBP fusion polypeptide of the invention. In some other embodiments, more than one cytidine deaminase-SSBP fusion polypeptides of the invention are employed. For example, in the in vivo systems detailed below, the host cell harboring a target gene can co-express an AID-SSBP fusion polypeptide (such as AID-gV fusion) and an APOBEC-SSBP fusion polypeptide (such as APOBEC-gV fusion). The presence or expression of multiple (e.g., 2, 3 or more) cytidine deaminase-SSBP fusion polypeptides in the gene evolving systems of the invention can lead to higher mutagenesis rates and a broader spectrum of mutations.

[0074] In vitro, some cytidine deaminase polypeptides (e.g., AID or an APOBEC enzyme) can introduce mutations randomly into ssDNAs, and also generate mutations to the nontemplate strand of dsDNA substrates under transcription. Thus, in some embodiments, mutations in a target polynucleotide sequence induced by the cytidine deaminase-SSBP fusion polypeptide are generated in vitro, e.g., by contacting in vitro a cytidine deaminase-SSBP fusion polypeptide of the invention and a target polynucleotide sequence. The target polynucleotide can be either single stranded DNAs or double stranded DNAs.

[0075] In some preferred embodiments, generation of hypermutations in the target gene and subsequent selection are performed in vivo, e.g., with a cell based system. In vivo, some cytidine deaminase polypeptides such as AID or APOBEC enzymes can introduce mutations to ssDNA substrates as well both strands of dsDNA substrates. The target polynucleotide sequence can be provided directly as single stranded or double stranded DNA, or preferably on a single stranded or double stranded vector as detailed below. Generation of

hypermutations in the target gene is achieved by introducing a vector expressing the cytidine deaminase-SSBP fusion polypeptide into a host cell. The host cell can stably or transiently express the cytidine deaminase-SSBP fusion polypeptide. The target gene or polynucleotide sequence can be exogenous or endogenous to the host cell. It can be present on the same vector that expresses the fusion polypeptide. Thus, in some embodiments, the target gene is heterologous to the host cell and is introduced into the cell on the same vector that expresses the cytidine deaminase-SSBP fusion polypeptide. Alternatively, the host cell can harbor the target gene or polynucleotide separately from the vector expressing the cytidine deaminase-SSBP fusion polypeptide (e.g., on the host genome or on a second vector). For example, the target polynucleotide can be an endogenous immunoglobulin gene or a heterologous gene that has been stably integrated into the host genome. In some other embodiments, the target gene is introduced into the host cell via a second vector harboring the target gene sequence (target gene construct). In some of these embodiments, expression of a target gene from a double stranded construct allows the target sequence transiently exposed as ssDNA. When the target gene construct is a phage or phagemid vector, ssDNA is also generated during phage production in a phage based system.

[0076] Evolving a target gene with the cytidine deaminase-SSBP fusion based system of the invention is advantageous over methods that would have been conceived by a skilled artisan in the art. For example, there is no need to induce expression of the target gene in order to allow the fusion cytidine deaminase polypeptide to generate mutations in the target gene. Previously, in order for a skilled artisan to efficiently generate mutations on a target gene with a cytidine deaminase such as AID, an induced high level expression of the target gene is needed. As demonstrated in Example 2 below, in the absence of an induced high level target gene expression, efficiency for AID alone to generate mutations in the target gene is much lower relative to that obtained when target gene expression is induced (even if expression of AID itself is induced) (see Figure 1). In contrast, as exemplified with an AID-gV fusion in Example 4 below, cytidine deaminase-SSBP fusion polypeptides of the invention are able to generate mutations with a high frequency on a target gene present in a phagemid vector without an induced expression of the target gene (see, Figure 12).

[0077] To allow production of candidate evolved mutants of a target gene or polynucleotide sequence, the host cell harboring the vector(s) comprising the cytidine deaminase-SSBP fusion sequence and the target gene is cultured under suitable conditions.

Target genes with desired mutations can be selected and screened with appropriate functional or phenotypic assays. The selection and screening can be performed in various formats, and are typically adapted in a high-throughput fashion. For example, in some embodiments, the evolved mutants can be selected and screened directly by examining colonies of the host cell (e.g., bacterial colonies). Alternatively, vectors harboring candidate mutants of the target gene identified from the bacterial host cell can be reintroduced into a mammalian cell for expression and subsequent selection and screening. In some other embodiments, screening for evolved mutants with desired properties can be performed with phage display as detailed below. Other platforms well known in the art for selecting and screening biological molecules can also be employed in the practice of the present invention. These include, e.g., yeast display. See, e.g., Boder and Wittrup, Nat Biotechnol. 15:553-7, 1997; Feldhaus et al., Nat Biotechnol. 21:163-70, 2003; and Boder et al., Proc. Nat. Acad. Sci. 97:10701-10705, 2000.

[0078] Any of the above described cytidine deaminase-SSBP fusion constructs can be used for evolving a target molecule encoded by a target gene or polynucleotide sequence. With one vector system, the target gene is additionally inserted into the construct. This can be performed with standard recombinant techniques that are well known in the art and also exemplified herein. In some of these embodiments, the construct harboring the target gene and cytidine deaminase-SSBP fusion sequence is a double stranded construct that allows production of single stranded DNAs, e.g., phagemid vectors. When two vectors are used, the target gene construct can be a double stranded construct such as phagemid vectors that allow production of single stranded DNAs. Alternatively, the target gene construct can be either a double stranded vector that can be converted into single-stranded sequences (e.g., through enzymatic action such as helicase or transciptase activity) or a single stranded vector. For example, the target gene can be provided in single stranded construct such as filamentous phage based vector M13mp18.

[0079] In some preferred embodiments, a phage or phagemid construct is used to express the cytidine deaminase-SSBP fusion and/or the target gene or polynucleotide. In some of these embodiments, a phage is used. Preferably, phage constructs used in the practice of the invention are based on a filamentous phage (e.g., M13). Construction and manipulation of a filamentous phage expressing a cytidine deaminase-SSBP fusion and/or a target polynucleotide can be performed with techniques routinely practiced in the art. See, e.g.,

Barbas et al., *Phage Display: A Laboratory Manual*, Cold Spring Harbor Laboratory Press (2001); and Sidhu, *Phage Display In Biotechnology and Drug Discovery*, CRC (2005). In some other embodiments, phagemid constructs are employed to express the cytidine deaminase-SSBP fusion and/or the target polynucleotide. A phagemid can grow as a plasmid, and also be packaged as single stranded DNA in viral particles. Phagemids contain an origin of replication (ori) for double stranded replication, as well as the origin of replication of the filamentous phage f1 (f1 ori) to enable single stranded replication and packaging into phage particles. Many commonly used plasmids contain an F1 ori and are thus phagemids. Similar to plasmids, a phagemid can be used to clone DNA fragments and be introduced into a bacterial host by a range of techniques (transformation, electroporation).

[0080] In addition to the target polynucleotide sequence (as well as the cytidine deaminase-SSBP fusion sequence if one vector system is employed), the target gene construct typically also harbors a promoter operably linked to the target gene sequence. It can further contain other elements necessary for selection and propagation, e.g., selection marker, origin of replication and enhancer. For example, the construct can optionally contain enhancer elements that are known to play a role in somatic hypermutation, e.g., the Ei/MAR and E3' elements in the Ig kappa gene (Betz et al., Cell 77: 239–248, 1994; and Yelamos et al., Nature 376: 225–229, 1995). The construct can further contain other sequence motifs that are known in the art to recruit somatic hypermutations to a heterologous sequence, e.g., the sequence motifs reported in Klix et al., Eur. J. Immunol. 28:317-326, 1998; and Sharpe et al., EMBO J. 10:2139-2145, 1991. Depending on the specific target gene to be mutated, the target gene construct can be readily generated with standard recombinant technologies. Any of the eukaryotic or prokaryotic promoters described herein or routinely used in the art can be utilized to generate the target gene construct.

[0081] Host cells that are proper for introduction of the expression construct(s) are also described herein. With two vector system, the two vectors can preferably contain different selection markers (e.g., different antibiotic resistance genes) to allow for easy selection of transfectants or transformants that harbor the two constructs. The two vectors can also optionally employ promoters that can be induced under different conditions. For example, the cytidine deaminase-SSBP fusion can be placed under the control of tetracycline (tet) promoter that can be induced with, e.g., anhydrotetracycline (AHT). Expression of the target gene can be placed under the control of a different promoter, e.g., the *tac* promoter which can

be induced with IPTG. Further, in some embodiments, to allow ssDNA production only from the target gene construct, the cytidine deaminase-SSBP construct should preferably not contain an f1 origin of replication. This can facilitate introduction of hypermutations only in the target gene construct. Increased production of ssDNA from the target gene construct can also be achieved by convergent transcription or by a combination of transcription and DNA bending caused by promoter activation. Thus, in some embodiments, the target gene construct can contain convergent promoters upstream and downstream of the target gene that induce transcription in both directions (see, e.g., Bierei et al., Mol. Breeding 10:107-117, 2002). Convergent promoters are promoters that are situated on either side of the isolated nucleic acid fragment of interest such that the direction of transcription from each promoter is opposing each other. Examples of suitable promoter combinations include the trp and lac UV5 promoters (Horiwitz et al., Nucleic Acids Res. 10:5447-65, 1982), and the H1 and U6 promoters (Jian et al., Eur. J. Cell Biol. 85:433-40, 2006).

[0082] Some embodiments of the invention are directed to evolving a target gene or polynucleotide sequence that encodes an enzyme (e.g., a kinase, a protease, or a cytidine deaminase). A target gene encoding an enzyme can be selected and screened for altered or improved enzymatic activity as a result of hypermutations induced by the cytidine deaminase-SSBP fusion. Selection of the target gene bearing mutations can be performed with appropriate assays that are depending on the specific enzymes encoded by the target gene and the desired property to be screened for. By way of example, the present inventors demonstrated that hypermutations can be induced into the AID enzyme itself by an AID-SSBP fusion disclosed herein, and screened for improved enzymatic activities (e.g., more potent or enhanced cytidine deaminase activity). The Examples below described studies directed to a screening assay to select AID-SSBP mutants that can generate higher mutation frequency in a reporter gene or polynucleotide sequence. As exemplified herein, the target construct harbors a mutant antibiotic resistance gene, e.g., a mutant carbenicillin resistance gene *blaS70P* which encodes a Ser-Pro mutation at codon 70 (AGC changed to CCT). Due to the mutation at the important Ser70 residue, the protein encoded by the *blaS70P* gene has lost the antibiotic resistance activity. However, the mutant gene can regain the activity upon a C-T mutation at the CCT codon induced by a functional AID enzyme (generating a Ser encoding TCT codon). Relative to wild type AID, evolved AID mutants isolated by the present inventors were found to be able to generate reversion in the blaS70P reporter gene at

higher frequency or at lower levels of AID expression. Other enzymes can be similarly evolved with the methods described herein, and screened using standard enzymology techniques. See, e.g., *Methods in Enzymology, Preparation and Assay of Enzymes and Substrates*, Sidney et al. (eds.), Academic Press, Inc. (1963); Bisswanger, *Practical Enzymology*, Wiley-VCH (2004); and Price & Stevens, *Fundamentals of Enzymology: The Cell and Molecular Biology of Catalytic Proteins*, Oxford University Press (3rd ed., 1999).

[0083] In some other embodiments of the invention, the target polynucleotide to be evolved encodes an immunoglobulin chain or antibody domain, or a polypeptide that is expressed and displayed on a cell surface (e.g., T cell receptor). The methods involve generating mutations in a target nucleic acid which encodes an immunoglobulin chain. A vector expressing the cytidine deaminase-SSBP fusion is introduced into a host cell that harboring the target gene. The target gene can be naturally expressed from the cell or separately introduced into the host cell on an expression vector. The target gene also can be present on the same vector that expresses the cytidine deaminase-SSBP fusion (i.e., one vector system). The host cells can secrete immunoglobulins into the culture medium or display the expressed immunobglobulin chain on the cell surface. Mutations induced by the cytidine deaminase-SSBP fusion that result in a disruption of the immunoglobulin chain expression can be readily detected with any suitable techniques well known in the art. For example, if the immunoglobulin is naturally present on the cell surface, its absence in a host cell can be detected by FACS analysis, immunofluorescence microscopy or ligand immobilization to a support (e.g., antigen-coated magnetic beads). Mutations that lead to an altered immunoglobulin chain (e.g., altered binding affinity) can be identified by a suitable activity assay (e.g., a binding assay with a cognate antigen). Similarly, immunoglobulin chains bearing hypermutations that are induced by the cytidine deaminase-SSBP fusion can be screened for altered binding specificity (e.g., by measuring binding to a panel of different antigens). Mutations induced by the cytidine deaminase-SSBP fusion may also be detected by directly sequencing the target gene, e.g., the variable region of the immunoglobulin gene. All these screening methods can be performed in high-throughput and automated format.

[0084] Other than enzymes and antibodies, various other types of target genes or polynucleotides can also be similarly evolved with the methods disclosed herein. These include polynucleotides encode receptor proteins (e.g., Fc-receptor or its fragments), cytokines, hormones, toxins, enzyme inhibitors, DNA-binding domains (e.g., zinc fingers),

protein ligands or any other peptides or proteins (e.g., protein A or protein L). Some specific examples of such target molecules include, e.g., tumor necrosis factor α (TNFα), vascular endothelial growth factor (VEGF), transforming growth factor (TGF), fibroblast growth factor (FF), platelet derived growth factor (PDGF), angiotensin II, IL-2, IL-10, insulin-like growth factor, insulin receptor, MHC proteins (e.g. class I MHC and class II MHC protein), CD3 receptor, cytokine receptors (e.g., interleukin receptors), G-protein coupled receptors, chemokine receptors, Maurotoxin, Agitoxin, and transcription factor Zif268. Selection and screening of evolved mutants of any of these target molecules can be readily performed in accordance with methods exemplified herein or well known in the art.

VI. Evolving a target polynucleotide via phage mediated system

[0085] In addition to introducing hypermutations to a target gene in a cell based system, the invention also provides methods with which the various target genes or target polynucleotide sequences described above can be evolved with a phage mediated system. As demonstrated in the Examples below, the phage based system of the present invention allows phage display of mutants of a target molecule generated by a cytidine deaminase protein such as AID. Previously, in order for a skilled artisan to generate mutations on a target gene with a cytidine deaminase such as AID, an induced high level expression of the target gene is needed to allow AID access the ssDNA substrate transiently exposed during expression. On the other hand, it is well known in the art that, if the target gene is fused to a phage coat protein (e.g., gIII or gVIII protein) for display, an induced high level expression of the phage coat protein is usually toxic to the cell. As a result of the toxicity to the host cell, no phage particles can be produced. Thus, the dilemma faced by the skilled artisan is that, on the one hand, a high level expression of the target gene is necessary in order for AID or another cytidine deaminase to generate mutations, and on the other hand, if expression of the target gene-phage coat protein fusion is induced at a high level, phage production is impaired due to toxicity to the host cell. This dilemma could be a reason why there have been no reports of successful incorporation of AID-mediated mutation into phage display procedure. As detailed below, the phage mediated system of the present invention overcomes the problem previously encountered in the art.

[0086] There are a number of advantages in the cytidine deaminase-SBP fusion based phage system of the invention. First, it was observed by the present inventor that, when a

cytidine deaminase such as AID alone is expressed (even at high level) in a phage system to evolve a target gene, the enzyme cannot efficiently generate mutations in the phagemid vector bearing the target gene (Example 2). Not intended to be bound in theory, this could be because ssDNA generated during phagemid replication is highly protected by gV. The employment of a cytidine deaminase-SSBP fusion (e.g., an AID-gV fusion) solves this problem by allowing AID access to gV-protected ssDNA phagemid. In addition, with the phage based system of the invention, there is no need to induce expression of the target gene in order to generate mutations in the target gene. This is because, unlike a cell based system wherein mutation on the target gene is generated when ssDNA is transiently exposed during expression of the gene, mutations on the target sequence are also generated during the replication of the phagemid vector and phage production with a phage based system of the invention. A cytidine deaminase-SSBP fusion polypeptide of the invention (e.g., an AID-gV fusion) is able to generate mutations on a target gene present in a phagemid vector without an induced high level expression of the target gene. As exemplified herein (see Example 4), phage particles with a genome bearing mutant target gene sequences were successfully produced from host cells transfected with a vector expressing the AID-gV fusion and the target gene.

[0087] The principle underlying all phage display platforms is the physical linkage of a polypeptide's phenotype to its corresponding genotype. In practice, the proteins or peptides to be displayed are usually expressed as fusions with the phage coat protein pIII or pVIII (or other coat proteins as described in Sidhu, Biomol. Eng. 18:57-63, 2001). Such fusion proteins are directed to the bacterial periplasm or inner cell membrane by an appropriate signal sequence that is added to their N terminus. During the phage assembly process the fusion proteins are incorporated into the nascent phage particle. The genetic information encoding the displayed fusion protein is packaged inside the same phage particle in the form of a single-stranded DNA (ssDNA) molecule. Hence, the genotype-phenotype coupling occurs before the phages are released into the extracellular environment, ensuring that phages produced from the same bacteria cell clone are identical.

[0088] With phage display, huge display libraries containing up to 10^{10} individual members can be created from batch-cloned gene libraries. Most applications of phage display libraries aim at identifying polypeptides that bind to a given target molecule. The enrichment of phages that present a binding protein (or peptide) is achieved by affinity

selection of a phage library on the immobilized target. In this "panning" process, binding phages are captured whereas nonbinding ones are washed off. In the next steep, the bonded phages are eluted and amplified by reinfection of suitable host (e.g., E. coli cells). The amplified phage population can, in turn, be subjected to the next round of panning. See, e.g., Barbas et al., *Phage Display: A Laboratory Manual*, Cold Spring Harbor Laboratory Press (2001); WO 91/19818; WO 91/18989; WO 92/01047; WO 92/06204; WO 92/18619; Han et al., Proc. Natl. Acad. Sci. USA 92: 9747-51, 1995; and Donovan et al., J. Mol. Biol. 196: 1-10, 1987.

[0089] The phage mediated system of the present invention preferably employs E. coli filamentous phage such as M13, fd, f1, Ike, and engineered variants thereof. An example of engineered variants of these phages is fd-tet, which has a 2775-bp *Bgl*II fragment of transposon Tn10 inserted into the *Bam*HI site of wild-type phage fd. Because of its Tn10 insert, fd-tet confers tetracycline resistance on the host and can be propagated like a plasmid independently of phage function as the displaying replicable genetic package. Using M13 as an exemplary filamentous phage, the phage virion consists of a stretched-out loop of single-stranded DNA (ssDNA) sheathed in a tube composed of several thousand copies of the major coat protein pVIII (product of gene VIII or "gVIII"). Four minor coat proteins are found at the tips of the virion, each present in about 4-5 copies/virion: pIII (product of gene III or "gIII"), pIV (product of gene IV or "gIV"), pVII (product of gene VII or "gVII"), and pIX (product of gene IX or "gIX"). Of these, pIII and pVIII (either full length or partial length) represent the most typical fusion protein partners for polypeptides of interest. A wide range of polypeptides, including random combinatorial amino acid libraries, randomly fragmented chromosomal DNA, cDNA pools, antibody binding domains, receptor ligands, etc., may be expressed as fusion proteins, e.g., with pIII or pVIII, for selection in phage display methods. In addition, methods for the display of multichain proteins (where one of the chains is expressed as a fusion protein) are also well known in the art. In addition to gIII and gVIII, other phage proteins such as gVI, gVII and gIX can also be used to display a target molecule in the present invention. Further, other than phage coat proteins, some non-phage proteins that associate with phage can also be used for displaying a target molecule (see, e.g., Russel et al., *Introduction to phage biology and phage display*, in "Phage Display – A Practical Approach", Clackson, T. and Lowman, H. B., eds., Oxford University Press, 2004, pp. 1 -26).

[0090] Phage system has been employed successfully for the display of functional proteins such as antibody fragments (scFv or Fab'), hormones, enzymes, and enzyme inhibitors, as well as the selection of specific phage on the basis of functional interactions (antibody - antigen; hormone - hormone receptor; enzyme - enzyme inhibitor). See, e.g., Paschke, Appl. Micbiol. Biotechnol. 70:2-11, 2006; and Kehoe and Kay, Chem. Rev. 105:4056-72, 2005. In general, phage display platforms can be grouped into two classes on the basis of the vector system used for the production of phages. True phage vectors are directly derived from the genome of filamentous phage (M13, f1, or fd) and encode all the proteins needed for the replication and assembly of the filamentous phage (Cwirla et al., Proc. Natl. Acad. Sci. USA 87:6378–6382, 1990; Scott and Smith, Science 249:386-390, 1990; Petrenko et al., Protein. Eng. 9:797-801, 1996; and McLafferty et al., Gene 128:29-36, 1993). In these vectors, the target gene or polynucleotide is either cloned as a fusion with the coat protein originally present in the phage genome or inserted as fusion gene cassette with an additional copy of the coat protein. The former vector system produces phages exclusively presenting the fusion coat protein, whereas the latter system yields phages that present the wild type and the fusion coat protein on the same phage particle.

[0091] The second group of phage display platforms utilizes phagemid vectors (see, e.g., Marks et al., J. Mol. Biol. 222:581–597, 1991; and Barbas et al., Proc. Natl. Acad. Sci. USA 88:7978-7982, 1991) which produce the fusion coat protein. As noted above, a phagemid is a plasmid that bears a phage-derived origin of replication in addition to its plasmid origin of replication. The phage-derived origin of replication is also known as intergenic region. Besides its function in DNA replication, the intergenic region contains a 78-nucleotide hairpin section (packaging signal), which promotes the packaging of the ssDNA in the phage coat. However, the production of phages containing the phagemid genome can only be achieved when additional phage derived proteins are present. For the purpose of phage display, these proteins are simply provided by superinfecting phagemid-carrying cells with a helper phage. In this procedure, often called "phage rescue," the helper phage provides all the proteins and enzymes required for phagemid replication, ssDNA production and packaging, and also the structural proteins forming the phage coat. The replication and packaging machinery supplied by the helper phage acts on the phagemid DNA and on the helper phage genome itself. Therefore, two distinct types of phage particles with different genotypes are produced from cells bearing phagemid and helper phage DNA: (1) those

carrying the phagemid genome and (2) those carrying the helper phage genome. Phage particles containing the helper phage genome are useless in phage display processes even if they present the desired phenotype because they do not contain the required genetic information. The fraction of phages containing helper phage genome can be reduced to ~1/1,000 by using a helper phage with a defective origin of replication or packaging signal, which leads to preferential packaging of the phagemid DNA over the helper phage genome. Independent of the genotype, phagemid-based display platforms usually yield phages with a hybrid phenotype displaying wild type and fusion coat protein on the same particle.

[0092] Detailed procedures for using phage display platforms are provided in the art. See, e.g., Barbas et al., *Phage Display: A Laboratory Manual*, Cold Spring Harbor Laboratory Press (2001). Only routinely practiced standard recombinant DNA techniques are required to evolve a target gene with a cytidine deaminase-SSBP fusion in a phage mediated system, as demonstrated in the Examples below. Such techniques are described, e.g., in Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, N.Y., (3rd ed., 2000); and Brent et al., Current Protocols in Molecular Biology, John Wiley & Sons, Inc. (ringbou ed., 2003). Fusion of the target polynucleotide and the phage polynucleotide can be accomplished by inserting the phage polynucleotide into a particular site on a plasmid that also contains the target polynucleotide gene, or by inserting the target polynucleotide into a particular site on a plasmid that also contains the phage polynucleotide. In some embodiments, the fusion polypeptides typically comprise a signal sequence, usually from a secreted protein other than the phage coat protein, the target molecule to be displayed, and either the gene III or gene VIII protein or a fragment thereof effective to display the polypeptide. The gene III or gene VIII protein used for display is preferably from (i.e., homologous to) the phage type selected as the display vehicle. Exogenous coding sequences are often inserted at or near the N-terminus of gene III or gene VIII although other insertion sites are possible.

[0093] In some preferred embodiments, a filamentous phage based phagemid vector such as M13 phage vectors is employed to express the target gene. With one vector system, the same vector also expresses the cytidine deaminase-SSBP fusion. Examples of vectors suitable for the invention include, but are not limited to, pASK-IBA3C, pOM5, fUSE5, fAFF1, fd-CAT1, m663, pHEN1, pComb3, pComb8, and p8V5. Some filamentous phage vectors have been engineered to produce a second copy of either gene III or gene VIII. In

such vectors, exogenous sequences are inserted into only one of the two copies. Expression of the other copy effectively dilutes the proportion of fusion protein incorporated into phage particles and can be advantageous in reducing selection against polypeptides deleterious to phage growth. In another variation, target polypeptide sequences are cloned into phagemid vectors which encode a phage coat protein and phage packaging sequences but which are not capable of replication. Phagemids are transfected into cells and packaged by infection with helper phage. Use of phagemid system also has the effect of diluting fusion proteins formed from coat protein and displayed polypeptide with wildtype copies of coat protein expressed from the helper phage. See, e.g., Garrard, WO 92/09690.

[0094] The choice of expression vector depends on the intended host cells in which the vector is to be expressed. Typically, the vector includes a promoter and other regulatory sequences in operable linkage to the inserted coding sequences that ensure the expression of the latter. Use of an inducible promoter is advantageous to prevent expression of inserted sequences except under inducing conditions. Examples of inducible promoters include *lac* promoter, *tac* promoter, arabinose promoter, metallothionein promoter or heat shock promoters. Cultures of transformed host cells can be expanded under noninducing conditions without biasing the population for coding sequences whose expression products are better tolerated by the host cells. The vector may also provide a secretion signal sequence positioned to form a fusion protein with polypeptides encoded by the inserted sequences, although often inserted polypeptides are linked to a signal sequences before inclusion in the vector. If the target polynucleotide encodes antibody light and heavy chain variable domains, the vectors to be used can sometimes encode constant regions or parts thereof that can be expressed as fusion proteins with inserted chains, thereby leading to production of intact antibodies or fragments thereof.

[0095] In some embodiments, the target sequences to be displayed on the surface of phage particles can comprise amino acids encoding one or more tag sequences. Such tag sequences can facilitate identification and/or purification of fusion proteins. Such tag sequences include, but are not limited to, glutathione S transferase (GST), maltose binding protein (MBP), thioredoxin (Tax), calmodulin binding peptide (CBP) , poly-His, FLAG, c-myc, and hemagglutinin (HA). GST, MBP, Trx, CBP, and poly-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, c-myc, and hemagglutinin (HA)

enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. Other suitable tag sequences will be apparent to those of skill in the art.

[0096] The vector with inserted target gene can be transformed into a suitable host cell. Prokaryotes are the preferred host cells for phage vectors. Suitable prokaryotic host cells include the E. coli BW301 cell or ER2738 cell as exemplified in the Examples below. These cell lines as well as suitable helper phage (e.g., a M13K07 helper phage or VCSM13 helper phage) can be obtained commercially, e.g., from New England Biolabs (Beverly, Mass.). Other suitable prokaryotic host cells for practicing the present invention include E coli strain JM109, E coli strain JM101, E. coli K12 strain 294 (ATCC number 31,466), E. coli strain W3110 (ATCC number 27,325), E. coli strain X1776 (ATCC number 31,537), and E. coli XLl-Blue cells (Stratagene, La Jolla, California). In addition, other strains of E. coli, such as HB101, NM522, NM538, NM539, and cells from many other species and genera of prokaryotes can also be used. For example, bacilli such as Bacillus subtilis, other enterobacteriaceae such as Salmonella or Serratia marcesans, and various Pseudomonas species may all be used as hosts.

[0097] Transformation of prokaryotic cells can be readily accomplished using methods well known in the art, e.g., as discussed in Sambrook et al., supra; and Brent et al., supra. For example, the calcium chloride method is a suitable method for transforming a prokaryotic host cell with a phage display vector. Alternatively, electroporation (Neumann et al., EMBO J., 1:84, 1982) may be used to transform these cells. The transformed cells are selected by growth on an antibiotic, e.g., tetracycline (tet) or ampicillin (amp), to which they are rendered resistant due to the presence of tet and/or amp resistance genes on the vector.

[0098] To evolve a target gene or polynucleotide sequence with the phage mediated system, a phage or phagemid vector expressing the target gene as fusion with a phage display anchor protein (e.g., gIII or gVIII protein) and a cytidine deaminase-SSBP fusion of the invention is introduced into a host cell for induction of mutations and production of phage particles (if needed, in the presence of a helper phage). Alternatively, with two vector system, the target gene construct and a separate cytidine deaminase-SSBP fusion construct of the invention are co-introduced into a host cell for inducing hypermutations in the target gene by cytidine deaminase activity, and for producing phage particles. In some other embodiments, the vector(s) and the helper phage can be transformed into a first host cell

(e.g., BW310 cell) for cytidine deaminase-induced mutagenesis of the target gene. Phages produced in the first host cell are then harvested and re-introduced into a second host cell that is better suitable for phage production and panning (e.g., the ER2738 cells).

[0099] Phage particles displaying a target molecule or mutant thereof (e.g., an Fab molecule) can be produced by culturing host cells that have been transformed with the expression vector(s) in accordance with the procedures described herein or that is well known in the art. For example, host cells (e.g., XL1-Blue E. coli cells) harboring vectors encoding the fusion polypeptides can be grown under suitable conditions (e.g., at 37°C in superbroth-medium containing 1% glucose and appropriate antibiotics) to allow propagation of phage particles. The phage particles released into the growth medium (cell supernatant) can be then harvested in the form of phage medium at that time. If desired, several rounds of enrichment procedures can be carried out, e.g., under conditions with increasingly higher stringency. As described herein, this is typically achieved by affinity selection or panning, using a target compound (e.g., an antigen) to which the displayed molecules (e.g., antibodies) are intended to bind. The harvested phage particles can then be used in subsequent selection for evolved mutants. Detailed techniques and procedures for enrichment and panning, as well as selection of evolved mutants of the target molecule for improved properties (e.g., enhanced binding affinity), are well known in the art. See, e.g., Barbas et al., *Phage Display: A Laboratory Manual*, Cold Spring Harbor Laboratory Press (2001). Exemplified conditions and procedures for enrichment, propagation and harvest of phage display particles are also provided in the Examples below.

[00100] As with the cell based system explained above, various types of target genes or polynucleotides can be evolved in the phage mediated system. In some exemplified embodiments, the methods are directed to evolving a target polypeptide that encodes an antibody or antigen-binding fragments of antibodies. Antigen binding fragments that can be used include, e.g., single chain variable region fragments (scFvs), single domain antibodies (dAbs), Fab fragments, F(ab')2 fragments, Fv fragments and Fd fragments. Antibodies have been displayed on phage in form of scFv or Fab' fragments using either the phage or the phagemid system. For example, in the latter case, either the V_H-C_{H1} or V_L-C_L chain is fused to gene III or gene VIII while the other chain is expressed without fusion.

[00101] By way of example, the Examples below describe evolving an Fab fragment with the cytidine deaminase-SSBP fusion of the invention in a phage mediated system. With this

system, the target Fab encoding sequence (heavy chain and light chain V-C1 regions) were first fused to a phage anchor protein (e.g., gIII or gVIII) in a phage display vector that is typically a phagemid vector. To increase the number of clones that likely bear cytidine deaminase-induced mutations in the phage population, the translation initiation codon of the target polynucleotide can be altered to function only after having a C to T mutation (e.g., an ACG codon). The cytidine deaminase-SSBP fusion can be expressed from the same vector (one vector system) or from a separate construct (two vector system). Similar to the above described cell based system, this target construct (and the cytidine deaminase-SSBP construct if two vectors are used) can then be introduced into a host cell (e.g., BW310 cells as exemplified in the Examples) suitable for expression of the cytidine deaminase-SSBP fusion, introduction of hypermutations into a ssDNA produced from the target construct, and production of phage particles harboring ssDNA of the target construct. Phages thus generated can then be directly subject to panning and selection for evolved proteins encoded by mutants of the target polynucleotide sequence. Alternatively, the phages can be harvested from the first host cell and re-infected a second host cell (e.g., ER2738 cell) that is more suitable for phage display and subsequent panning, selection and screening. If needed, the evolution cycle containing the step of introducing mutations into the target construct and the step of phage display and selection can be repeated several times until mutants with desired properties are identified from the selection and screening.

EXAMPLES

[00102] The following examples are provided to further illustrate the invention but not to limit its scope.

Example 1 Materials and methods

[00103] Plasmids and constructs. ptacblaS70PF1 and ptacblaSSPF1 reporters were generated by PCR-mediated site-directed mutagenesis (Barettino et al. NAR 22;541,1994). First, 4 mutagenic primers synthesized from IDT DNA Inc. (Coralville, IA) were used to amplify fragments of bla gene from pComb3XTT vector; '5-CCAATGATGCCTACTTTTAAAGTTCTGCTATG-3' (SEQ ID NO: 1), '5-CTTTAAAAGTAGGCATCATTGGAAAACGTTCTT-3' (SEQ ID NO: 2), '5-ATAACCATGCCTGATAACACTGCGGCCAACTTA-3' (SEQ ID NO: 3) and '5-

AGTGTTATCAGGCATGGTTATGGCAGCACTGCA-3' (SEQ ID NO: 4). Underlined sequences represent Pro residues which impairs lactamase functionality. Overlap-PCR products were amplified and subcloned into pGBM2 vector (Manen et al., Gene 186:197-200, 1997) using BamHI and NotI, together with lacIq gene, tac promoter and F1 origin. The direction of F1 origin was designed to generate a sense strand of lactamase gene as ssDNA using 'gcg aagctt GGGAAATTGTAAGCGTTAATAT-3' (SEQ ID NO: 5) and '5-cgcgcggccgcACGCGCCCTGTAGCGGCGCA-3' (SEQ ID NO: 6) primers, which has HindIII and NotI site, respectively, for cloning. Mouse AID cDNA was synthesized based on the E coli. codon optimized sequence from Ramiro et al., Nat. Immunol. 4:452-6, 2003. Human AID gene was obtained from ATCC MGC-12911 cDNA clone. Original pASK-IBA3C (IBA, Germany) vector was modified for removal of F1 origin by connecting 'lpp terminator' and 'chloramphenicol gene' on it, which resulted in 'pASK3CΔ'. Mouse/human AID genes were amplified by PCR and subcloned into pASK3CΔ using XbaI and HindIII sites. Mouse AID and gV fusion constructs were generated by overlap PCR, where short GGGS liker was introduced. gV fragment was amplified from VCSM13 helper phage. For Herceptin mutagenesis, blaS70P gene on ptacblaS70PF1 was replaced with Herceptin-Fab and gIII sequences from pComb3X-based plasmid by PCR-cloning.

[00104] Cell-based assay and phage-mediated assay. Cell-based assay procedure has been adapted from Ramiro et al., Nat. Immunol. 4:452-6, 2003. Briefly, BW310 UDG-deficient strain (λ−ung-1 relA1 spoT1 thi-1, 4. AB1255, ilvA201 metB1 hisG1 argH1 Hfr), a kind gift from Dr. Bernard Weiss, MD, was transformed with reporter plasmids (ptacblaS70P F1 or ptacblaSSP F1) and other plasmids that express various forms of mouse/human AID proteins (pASK3CΔmAID, pASK3CΔmAIDgV, pASK3CΔgVmAID, pASK3CΔhAID and pASK3CΔhAIDgV). Cells were grown in selective medium containing 100μg /ml spectinomycin(Spc) and 34μg /ml chloramphenicol (Chl) overnight. Overnight-grown cells were inoculated with 1:100 dilution into 3ml fresh selective medium and grown to an optical density of 0.6~0.8 at 600nm and induced with 0.2μg /ml AHT (Sigma) plus 1 mM IPTG for 3 h at 37 °C. Cell cultures were centrifuged, washed, and spread in LB plates containing 100μg /ml Spc, and 34μg /ml Chl (cell viability plates) or 100μg /ml Amp, 100μg /ml Spc, and 34μg /ml Chl (Amp^R plates). The mutation frequencies were calculated as the ratio of Amp^R colonies to the total number of viable cells, identified for each individual colony.

[00105] Phage-mediated assay was performed using phage particle as a mutation transporter to phenotypic secondary cells (ER2738). Briefly, transformed cells for cell-based assay were grown overnight in selective medium containing 100 μg/ml spectinomycin(Spc) and 34 μg/ml chloramphenicol(Chl). Overnight-grown cells were inoculated with 1:100 dilution into 10ml SB medium containing antibiotics and grown to an optical density of 0.6~0.8 at 600nm, and helper phage (at a titer of 10^{11} phage particles) were inoculated to infect growing cells. After 15 min of incubation at RT, 90ml SB was added and cultured more for 1.5~2hr at 37°C shaker. 70μg/ml Kanamycin and 0.2μg/ml AHT were added and cells were grown overnight at 37°C to produce phage particles. On next day, produced phages were harvested by mixing with 1/4 volume of PEG/NaCl solution (20% polyethylene glycol -8000, 2.5 M NaCl) and 30 min on ice-bath. After 20min of centrifugation at 14,000xg, supernatants were discarded and phage pellet were resuspended with 1ml PBS and filtered through 0.45μm sterile syringe filter. Fresh ER2738 (NEB, MA) cells were grown overnight for phage infection. Overnight cultured ER2738 cells were inoculated with 1:100 dilution into 5ml SB medium and grown to an optical density of 0.8~1.0 at 600nm. These cells were infected with 50~100μl of phage samples and left for 15 min at RT. After 30min of further incubation at 37°C shaker, infected cells were spread on LB plates containing 100μg /ml Spc (cell viability plates) or 100μg /ml Amp and 100μg /ml Spc (Amp^R plates). On next day, the mutation frequencies were calculated as the ratio of Amp^R colonies to the total number of viable cells, identified for each individual colony.

[00106] Herceptin Fab display and mutant selection. The procedure for Herceptin Fab display was mostly based on phage-mediated assay. Two plasmids, pASK3CΔhAIDgV and ptacwtHER RF1, were used to produce phage particles from BW310 cells. After phages from BW310 cells were resuspended with PBS, 5ml fresh ER2738 cells were diluted and kept growing to an optical density of 0.8~1.0 at 600nm for about 2hr. 50~100ul of harvested phages were inoculated to ER2738 cells and allow 15min at RT for the cells to get infected. 10ml SB and 60μl of Spec were added, and sample cells were taken to measure phage titers and then, allow 1hr of shaking at 37°C . After 1 hr, 90μl of spec was added once more and the 1 hr of shaking followed. Then, 300 μl of VCSM13 helper phage were inoculated with following incubation at RT for 15min. Helper phage infected cells were transfered to 500ml centrifuge bottle and add 135ml SB which contained 100μg /ml Spc, then cultured for 1.5~2hr at 37°C. After that, 70μg/ml kanamycin was added to cells. Next day, phages from

ER2738 cells were harvested by PEG/NaCl solution and centrifugation. Phages resuspended with PBS were used for bio-panning.

[00107] For enrichment of high affinity Herceptin mutants clones, competitive and off-rate panning was performed. 200ng of Erb2 ECD-Fc fusion protein was coated to 96 well plate at 4°C overnight and coated wells were blocked with 5% BSA. 70μl of phage samples together with Herceptin, 0 to10μg increasing amount at each rounds, were incubated for 1hr at RT. From 3rd round, overnight incubation with Herceptin was applied for off-rate selection. After 5~10 times washing with 0.05% Tween20 TBS, bound phages were eluted with 100μl of 0.1M HCl-glycine (pH 2.2), and neutralized with 20 μl of 1M Tris-Cl (pH 9.1). The eluted phages from each panning round were infected to exponentially growing ER2738 cells and those infected cells were used for phage production for next round of panning or cultured for plasmid DNA preparation.

[00108] ELISA assay. From output tittering plates for the last round of panning, 25~50 colonies were picked. Then expression of the Fab antibody fragments was induced with 0.5mM IPTG in culture medium. To analyze relative binding affinities of each Fab, ELISA plates (Nunc) were coated with 15ng of Erb2 protein and incubated overnight at 4°C. Coated wells were blocked for 1 h at 37°C with 5% BSA in PBS. Then, culture supernatants containing Fab were incubated for 1 h at 30 °C. After washing, HRP-conjugated anti-HA monoclonal antibody (Roche) was used for detection. The results were measured by optical density at 405 nm.

Example 2 Development of assay systems

[00109] This Example describes the vectors and assays developed for evolving a target polynucleotide sequence in a cell based system or a phage based system.

[00110] Reporter vector ptacblaS70PF1 was designed to assay mutator activity of AID or AID-gV fusion. For efficient selection of gain-of-function mutants by 'C-to-T' AID mutation, Ser70 in bla gene were mutated to Pro (AGC → CCT), so that blaS70P mutants could retain Carbenicillin resistance after C-to-T reversion by AID activity. Ser70 is known as key residue in lactamase core site. To generate this vector, lacIq, tac promoter, blaS70P gene and F1 origin were introduced into pGBM2 (Manen D et al. Gene 186, 197-200 ,1997).

[00111] With a two plasmids system for assessing hypermutations generated by AID, AID protein is produced from a second construct. In this construct, AID expression is under control of tetracycline-inducible promoter. This construct was constructed on the pASK-IBA3C cloning vector available from Cayman Chemical Company (Ann Arbor, Michigan) with either human AID or mouse AID cDNA sequences. The human AID sequence was obtained from ATCC (cDNA clone MGC:12911). The Mouse AID sequence used was described in the supplementary data of Ramiro et al., Nat. Immunol., 4:452-6, 2003. To allow ssDNA production only from ptacblaS70PF1 reporter vector, F1 origin in pASK-IBA3C was removed and named as pASK3CΔmAID for expressing mouse AID or pASK3CΔhAID for expressing human AID.

[00112] AID activity was tested with the blaS70P reporter construct and the pASK3CΔmAID construct. After co-transformation of two plasmids, AID-driven reversion on blaS70P was tested by cell-based assay system. As a result, relatively a bit more background colonies showed up in controls in comparison with the kanL94P assay system reported in Ramiro et al., Nat. Immunol., 4:452-6, 2003. In the kanL94P system reported in Ramiro et al., Pro-to-Ser or Pro-to-Leu mutations generated by mouse AID can provide kanamycin resistance to the cells. However, as shown in Figure 1, the results indicate that induction of mAID expression and increased transcription of the reporter gene increased the reversion rate of the reporter gene, similar to results as reported in other reports (Ramiro et al., Nat. Immunol. 4:452-6, 2003).

[00113] Phage-mediated assay using ptacblaS70PF1 construct was examined in BW310 host cells (ung-) by co-transforming the cells with ptacblaS70PF1 and pASK3CΔh/mAID constructs. The cells were propagated for phage production overnight in AID induced condition. Phage particles were harvested the next day and introduced into fresh target cells, ER2738(F+, ung+) or BW310(Hfr, ung-). The infected cells were spread on $carb^+$ plates. The results indicate that intact AID expression (both mouse AID and human AID) didn't show increased number of colonies on the $carb^+$ plates. This suggests that AID doesn't lead to increased mutation rate in the blaS70P gene on the single stranded phage DNA.

[00114] Because AID itself did not appear to efficiently function on the single stranded phage DNA, a fusion construct expressing a gV/AID fusion protein was constructed. The fusion proteins were designed with gV protein fused to either the N-terminus or the C-terminus of AID (mouse and human version) in the pASK3CΔh/mAID background. For

separation of the two domains, a short linker 'GGGS' was used. Human AID and gV fusion constructs were similarly constructed.

mAID-gV : ~~ DAFRMLGF GGGS MIKVEI ~~ (SEQ ID NO: 7)
mAID linker gV

gV-mAID : ~~ RLRLVPAK GGGS MDSLLM ~~ (SEQ ID NO: 8)
gV linker mAID

[00115] The mutator activity of the AID-gV fusion proteins were then tested in a cell-based assay. To ensure that the fusion forms of AID maintain their mutation activity, the fusion proteins were expressed together with ptacblaS70PF1 reporter plasmids in BW310 host cells. As shown in Figure 2, results from the reversion rate of BlaS70P from this study revealed that the AID-gV fusion didn't affect the AID mutator activity, and that the gV-AID fusion showed decreased AID activity. The latter result could be due to toxic effect on cell growth that might affect the number of mutated clones. Based on this study, only vectors with the AID-gV fusion (pASK3CΔhAIDgV and pASK3CΔmAIDgV) were used in subsequent assays.

[00116] The mutator activity of the AID-gV fusion proteins were also examined in a phage-mediated assay. Upon transforming BW310 cells (ung-) the ptacblaS70PF1 phagemid construct along with either the AID construct or the AID-gV fusion construct, phage production (having ssDNA of ptacblaS70PF1) was carried out in the presence of a helper phage (VCSM13 helper phage from Stratagene) with or without the induction of AID expression. Phage particles were harvested after o/n production. These phage particles were then introduced to fresh target cell, BW310 (ung-) or ER2738 (ung+). The infected cells were then spread on $carb^+$ plates. As shown in Figure 3, AID-gV fusion construct showed about 10-fold increased reversion on the blaS70P reporter gene, even though the background revertant frequency is about 1 per 10^5. When mouse AID and human AID were compared, human AID-gV fusion showed a higher reversion frequency. However, the sequence analysis for the single colonies from $carb^+$ plates showed that they have mixed peak on S70P site, which might have come from multiple phage infection.

[00117] As shown in Figure 4, revertant frequencies of single- and double-mutation reporters showed a dose-dependency with the increase of hAID-gV induction. Also, sequence analysis for $carb^r$ clones from double-mutation reporter showed several mutations at hot-spot target sequences of AID, which is known as 'WRC' (W=A or T, R=A or G). These results support that these mutations on the reporter gene were introduced by C-to-T mutation activity of AID-gV when phage particles were produced from BW310 cells.

[00118] In order to compare the frequency of single-site and dual-sites mutations by AID, double mutation reporter, which has two 'Ser-to-Pro' mutations at Ser70 and Ser130, were generated by overlap PCR and cloned as 'ptacblaSSPF1' reporter vector. Double mutation reporter with hAID-gV expression in phage-mediated assay showed about 0.8 per 10^6 reversion frequency while single site reversion showed about 2 per 10^4 frequency (Figure 5). Compared to the well-known *mutD5* mutator cell which has shown 12~220 per 10^6 mutation frequency on Nal^r, Str^r and Arg^+ phenotype, the mutation activity of hAID-gV is similar to *mutD5* activity. The mutD (or dnaQ) gene encodes the ε-subunit of the Pol III holoenzyme, which carries the 3' – 5' proofreading exonuclease of this enzyme. Pol III from mutD5 or dnaQ49 strains was shown to be defective in 3'→ 5' exonuclease activity (Degnen et al., J Bacteriol. 117:477,1974 & Schaaper RM, PNAS. 85:8126, 1988).

Example 3 Evolving an antibody using AID-mediated mutagenesis

[00119] This Example describes the vectors and assays developed for evolving a Fab fragment of the antibody Herceptin with AID-gV fusion in a phage mediated system.

[00120] A vector was constructed to display the Herceptin Fab on the phage particle, which is derived from 'ptacblaS70PF1' vector for bla-reversion assay (Figure 6). The blaS70P gene was replaced with the Fab fragments of Herceptin and gIII fusion protein. To harvest more clones mutated by AID after selection, mutant Fab was also constructed, which contains 'ACG' translation initiation codon on the heavy chain. This mutant construct was expected to display Herceptin Fab on phage coat only after having 'ACG to ATG' mutation. New vectors were named as 'ptacwtHER RF1' and 'ptacmutHER RF1'. From both vectors, phage particles were efficiently produced, just 3~5 times lower titer than pComb3X vector, and Fab display on phage particle were confirmed through western blot analysis.

[00121] Phage panning for Herceptin evolution: General scheme for phage display and selection cycle was modified a little for the mutation step by AID. The wt/mut Herceptin Fab

vectors were transformed into BW310 cell together with pASK3CΔhAIDgV. During phage production from BW310 cell, expression of the hAIDgV fusion was induced for producing mutations in the Herceptin vectors. Next day, phages were harvested and re-infected to ER2738 cell, common cell for phage display. General phage panning experiments were followed at the next following day; 1) Phage incubation with HER2 extracellular domain, 2) washing to remove non-binders, 3) elution of antigen-bound phages, and 4) reinfection to ER2738 cell to produce phages. 4 rounds of panning were conducted.

[00122] Clone analysis: After 4 rounds of panning, 25 selected clones from wt HER or mut HER selection were picked from the plates, and Fab molecules from each clone were expressed for ELISA. For affinity comparison, whole IgG form of Herceptin was co-incubated with the Fab clones. Results from the ELISA assays are shown in Figure 7. The assays are competitive ELISA designed to check whether if there is any Fab clone that can bind to HER2 in a competing condition with Herceptin® (max. amount of Herceptin® was 5 μg/well, which is about 50 μg/ml). However, the results indicate that there were no better binders showing stronger signal than Her Fab. So from each pools, clones with relatively high signal were picked and analyzed for their sequences. From the sequence analysis, 6 out of 16 clones showed 'C to A' or 'G to A' mutations and 4 of them were 'TGA' stop codon generation.

[00123] New panning strategy for Herceptin evolution: To check if ptacwtHER RF1 vector has generated a bias on the previous selection result, pComb3X vector based constructs, pC3XHER RF1 and pC3XHER original, lac promoter containing vector (placwtHER), and modified tac promoter containing vector (pΔtacHER), were tested for a new panning cycles. To increase selection efficiency, off-rate selection was applied and AID mutation step was increased 2 more times. For off-rate selection, phage bound to antigen in 96 well was incubated with 75ug of whole IgG Herceptin, 37'C, overnight and then eluted with trypsin. The scheme for selection cycle is shown in Figure 8. Panning results from the selection are shown in Figure 9. The data indicate that input phage titers were quite as normal as common phage display experiments. But output didn't show any significant increases in eluted phage titers. This is because the assay performed was an 'off-rate selection' using Herceptin as a competing antibody to enrich strong binders.

[00124] Clone analysis: After 7 rounds of panning, 25 clones were picked from each vector sets and tested for binding on Her2 by ELISA. Total 40 clones showing positive

signals were picked and their sequence information were analyzed (V_L-C_L1-V_H region). Overall, 65 mutations were found from 46,560 bp (40×1,164bp). Mainly, they were 'G to A' and 'C to T' mutations as reported (Yoshikawa et al., Science, 296:2033-6, 2002; and Peterson-Mahrt et al., Nature, 418:99-103, 2002), and 34 clones showed at least one mutation. 4 clones showed 'TGA' stop codon mutation and 3 clones showed 'TAG' amber codon mutation, which was interesting because these amber mutations appeared only at originally glutamine sites (in ER2738 cell, amber codon is suppressed by $tRNA^{glu}$). Other mutations were mostly silent, but 7 clones showed mutations on amino acid level, 6 of them expressed intact Fab and D to N mutation at VH-CDR3 appeared twice. Interestingly, D98(V_H) and H91(V_L) were reported as key residues for Herceptin binding (Gerstner et al., J. Mol. Biol. 321:851-62, 2002).

[00125] Results from ELISA study of some of the clones identified with the new panning strategy are shown in Figure 10. As indicated in the figure, each vector system showed different aspect on enrichment of binders. Obviously, phage based system using the pC3XHER RF1 and pΔtacHER vectors for evolving Herceptin Fab produced more binders based on the ELISA analysis.

Example 4 Evolving AID with AID-gV fusion

[00126] This Example describes studies of evolving AID itself with AID-gV fusion in a cell based system or phage-mediated system.

[00127] One vector system design for phage based system: In order to isolate more active AID mutants through phage-mediated selection, one vector system was developed. The phagemid vector has the hAIDgV fusion gene expression under the control of tet promoter and the blaS70P gene under either one of tac promoter and lac promoter (Figure 11). At first, after allowing AHT induction of AID-gV fusion expression and generation of mutations in the target gene, phage particles were produced from BW310 cell and then re-infected into ER2738 cell. Infected ER2738 cells were incubated for at least 30 minutes, and then spread on either carb(+) or spectinomycin(+) plates. For spectinomycin(+) plates, cells were diluted about 10^4 times for spreading around 100~2000 cells on each plate. The cells which have 'C to T' mutation at Pro70(CCT) showed up next day. Cells growing on spectinomycin(+) plates represent viable cells harboring the vector that expresses spectinomycin resistance gene. Cells growing on Carb(+) plates represent those which harbor a mutation that restores

carbenicillin resistance to the blaS70P gene. The results indicate that the frequency of phage particles bearing the mutation in the hAIDgV-ptacblaS70PF1 construct is about $72/10^6$ (the ratio of colonies on carb plate (mutated cells) over colonies on spectinomycin plate (viable cells). Unlike cell based system discussed below, there was no IPTG induction of the target gene blaS70P with phage-mediated assay. This is because, with a phage based system, expression of the target gene is not necessary in order to produce ssDNA substrate for AID activity. Rather, production of ssDNA occurs during replication of the phagemid vector. Therefore, generation of mutations in the target gene by AID with the phage based system likely also takes place during the process of vector replication and phage production, alternative or additional to the process of target gene expression. In addition, a high level expression of a target gene fused to a phage coat protein (e.g., gIII protein) can have a toxic effect to the cells. As such, it also would not be desirable to induce expression of a target gene fused to a phage coat protein in a phage based system.

[00128] Cell-based bla mutation assay: When cell-mediated bla mutation assay was tested, it was found that the 'hAIDgV-ptacblaS70PF1' construct showed bla reversion even without induction of hAIDgV by adding AHT (Figure 12). It was later found that the *tet* promoter in the evolution construct doesn't have repressor binding on its operator. So, it seems that hAIDgV was constantly expressed and restored Pro70 to Ser on bla gene without AHT induction. In the absence of AHT induction, the 'hAIDgV-ptacblaS70PF1' construct produced a mutation frequency of $\sim 14/10^6$, while control vector ptacblaS70PF1 (w/o hAIDgV) showed a mutation frequency of $\sim 0.033/10^6$. When examined with the two vector cell based system (ptacblaS70PF1 & pASK3CΔhAIDgV), the mutant frequency is $\sim 0.72/10^6$ (Figure 12). When hAIDgV expression in the cells was induced with AHT, mutation frequency obtained with the 'hAIDgV-ptacblaS70PF1' construct was slightly increased from $\sim 14/10^6$ to $\sim 19/10^6$ (data not shown).

[00129] As additional controls, one vector constructs which express under tet promoter either (1) hAID without gV or (2) an inactive AID mutant, hAIDE58QgV were also examined. These two constructs produced a blaS70P mutation frequency of $3.3/10^6$ and $0.039/10^6$, respectively (Figure 12). It is noted that in these 'cell-based' assays, both AID and AID-gV target double strand DNA as a substrate. However, AID seems to access target sequence when the target sequence is transiently exposed as ssDNA during expression.

[00130] Thus, the results indicate that the phage-mediated evolution assay (one vector system) appears to generate about 4 fold more mutations than the cell-based assay with AHT induction (mutation frequency of $72/10^6$ versus $19/10^6$). However, if both IPTG and AHT are used in the cell-based assay, a mutation frequency higher than that obtained with the phage mediated system is produced ($103/10^6$). IPTG induces a high level blaS70P expression from the tac promoter in the vector. The increased mutation frequency in the blaS70P sequence by enhanced blaS70P expression is likely because induction of blaS70P expression leads to more ssDNA substrate for AID activity.

***

[00131] Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to one of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

[00132] All publications, databases, GenBank sequences, patents, and patent applications cited in this specification are herein incorporated by reference as if each was specifically and individually indicated to be incorporated by reference.

WHAT IS CLAIMED IS:

1. An isolated or recombinant polynucleotide which comprises a polynucleotide sequence encoding a fusion polypeptide comprising (a) a single stranded DNA binding protein and (b) a polypeptide possessing cytidine deaminase activity or enzymatically active fragment thereof.

2. The polynucleotide of claim 1, wherein the single stranded DNA binding protein binds non-specifically to single stranded DNA.

3. The polynucleotide of claim 1, wherein the single stranded DNA binding protein is encoded by gene V (gV) of a filamentous phage or a functional derivative thereof.

4. The polynucleotide of claim 1, wherein the polypeptide possessing cytidine deaminase activity is a eukaryotic activation-induced cytidine deaminase (AID) or APOBEC enzyme.

5. The polynucleotide of claim 1, wherein the polypeptide possessing cytidine deaminase activity is a human or mouse AID polypeptide.

6. The polynucleotide d of claim 1, wherein the polypeptide possessing cytidine deaminase activity is an APOBEC3 enzyme.

7. The polynucleotide of claim 1, wherein the activation-induced cytidine deaminase is fused at the N-terminus of the single stranded DNA binding protein.

8. A fusion polypeptide encoded by the isolated or recombinant polynucleotide of claim 1.

9. An expression vector harboring the isolated or recombinant polynucleotide of claim 1.

10. A method of introducing mutations to a target gene, comprising contacting the target gene with at least one fusion polypeptide comprising (a) a polypeptide possessing cytidine deaminase activity and (b) a single stranded DNA binding protein, thereby introducing mutations to the target gene.

11. The method of claim 10, wherein the target gene and the fusion polypeptide are contacted in a host cell that harbors the target gene and expresses the fusion polypeptide.

12. The method of claim 11, further comprising culturing the host cell under conditions so that the fusion polypeptide is expressed.

13. The method of claim 11, wherein the target gene is introduced into the host cell by transfecting or transforming a host cell with a target construct harboring the target gene.

14. The method of claim 13, wherein the target construct is a phage or a phagemid vector.

15. The method of claim 13, wherein the fusion polypeptide is expressed from the construct harboring the target gene.

16. The method of claim 13, wherein the fusion polypeptide is expressed from a second construct transfected or transformed into the host cell.

17. The method of claim 10, wherein the single stranded DNA binding protein binds non-specifically to single stranded DNA.

18. The method of claim 10, wherein the single stranded DNA binding protein is encoded by gene V (gV) of a filamentous phage or a functional derivative thereof.

19. The method of claim 10, wherein the polypeptide possessing cytidine deaminase activity is fused at the N-terminus of the single stranded DNA binding protein.

20. The method of claim 10, wherein the polypeptide possessing cytidine deaminase activity is an activation-induced cytidine deaminase (AID) or an APOBEC enzyme.

21. The method of claim 10, wherein the polypeptide possessing cytidine deaminase activity is human AID or mouse AID.

22. The method of claim 10, wherein the polypeptide possessing cytidine deaminase activity is an APOBEC3 enzyme.

23. The method of claim 10, wherein the target gene is contacted with (a) a first fusion polypeptide comprising an activation-induced cytidine deaminase (AID) and a single stranded DNA binding protein; and (b) a second fusion polypeptide comprising an APOBEC enzyme and a single stranded DNA binding protein.

24. The method of claim 23, wherein the first and second fusion polypeptides are co-expressed in a host cell that harbors the target gene.

25. The method of claim 10, wherein the target gene encodes a protein or a peptide.

26. The method of claim 25, wherein the protein or peptide is an enzyme, an antibody chain, a protein ligand, a receptor protein, a cytokine, or a zinc finger polypeptide.

27. A method of evolving a target gene in a phage mediated system, comprising (a) generating a phage or phagemid vector expressing the target gene fused to a phage coat protein or a protein that associates with phage, (b) introducing the phage or phagemid vector into a host cell expressing at least one fusion polypeptide comprising (i) a single stranded DNA binding protein and (ii) a polypeptide possessing cytidine deaminase activity or enzymatically active fragment thereof, (c) culturing the host cell under conditions to allow expression of the fusion polypeptide and production of phage particles, and (d) selecting or screening the produced phage particles for phages that harbor an evolved mutant of the target gene.

28. The method of claim 27, wherein the selecting or screening is performed in a second host cell into which vectors encapsulated by the produced phage particles are introduced.

29. The method of claim 28, wherein the second host cell is a yeast cell, a bacterial cell or a mammalian cell.

30. The method of claim 27, wherein the single stranded DNA binding protein binds non-specifically to single stranded DNA.

31. The method of claim 27, wherein the single stranded DNA binding protein is encoded by gene V (gV) of a filamentous phage or a functional derivative thereof

32. The method of claim 27, wherein the fusion polypeptide is also expressed from the phage or phagemid vector.

33. The method of claim 27, wherein the fusion polypeptide is expressed from a second construct introduced into the host cell.

34. The method of claim 27, wherein the target gene encodes a peptide or a protein.

35. The method of claim 27, wherein the target gene encodes at least one immunoglobulin domain or an enzyme.

36. The method of claim 35, wherein the target gene encodes a single chain variable region fragment (scFv).

37. The method of claim 35, wherein the target gene encodes the V_H-C_{H1} chain of a Fab fragment, and wherein the phagemid construct further expresses the V_L-C_{L1} chain of the Fab fragment.

38. The method of claim 27, wherein the phage coat protein is encoded by gVIII or gIII of a filamentous phage.

39. A method of evolving a polypeptide possessing cytidine deaminase activity, comprising (a) generating a phage or phagemid vector expressing a fusion protein comprising the polypeptide possessing cytidine deaminase activity and a single stranded DNA binding protein, (b) transfecting the phage or phagemid vector into a host cell, (c) culturing the host cell under conditions to allow expression of the fusion protein and production of phage particles, and (d) selecting and screening the produced phage particles for phages that harbor a genome encoding an evolved mutant of the polypeptide possessing cytidine deaminase activity.

40. The method of claim 39, wherein the single stranded DNA binding protein binds non-specifically to single stranded DNA.

41. The method of claim 39, wherein the single stranded DNA binding protein is encoded by gene V (gV) of a filamentous phage or a functional derivative thereof

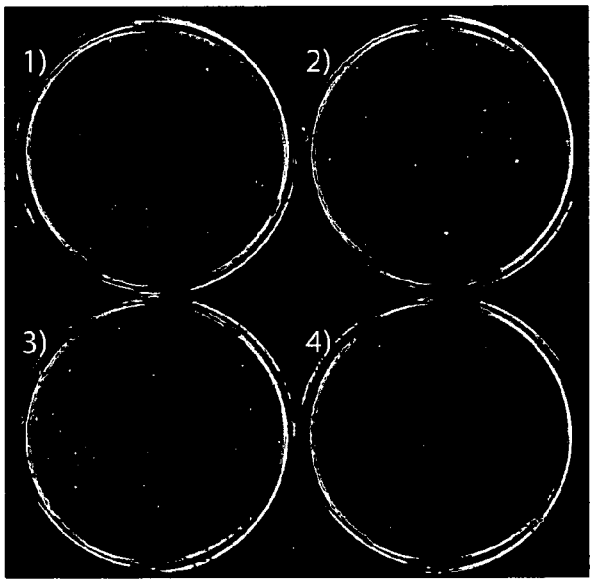
42. The method of claim 39, wherein the polypeptide possessing cytidine deaminase activity is an activation-induced cytidine deaminase or an APOBEC enzyme.

43. The method of claim 39, wherein the polypeptide possessing cytidine deaminase activity is human AID or mouse AID.

44. The method of claim 39, wherein the polypeptide possessing cytidine deaminase activity is an APOBEC3 enzyme.

45. The method of claim 39, wherein the polypeptide possessing cytidine deaminase activity is fused at the N-terminus of the single stranded DNA binding protein.

46. The method of claim 39, wherein the produced phage particles are selected for an evolved mutant of the polypeptide possessing cytidine deaminase activity that has stronger cytidine deaminase activity than the non-evolved polypeptide possessing cytidine deaminase activity.



1) mAID (-), blaS70P(-)
2) mAID (-), blaS70P(+)
3) mAID (+), blaS70P(-)
4) mAID (+), blaS70P(+)

FIG. 1

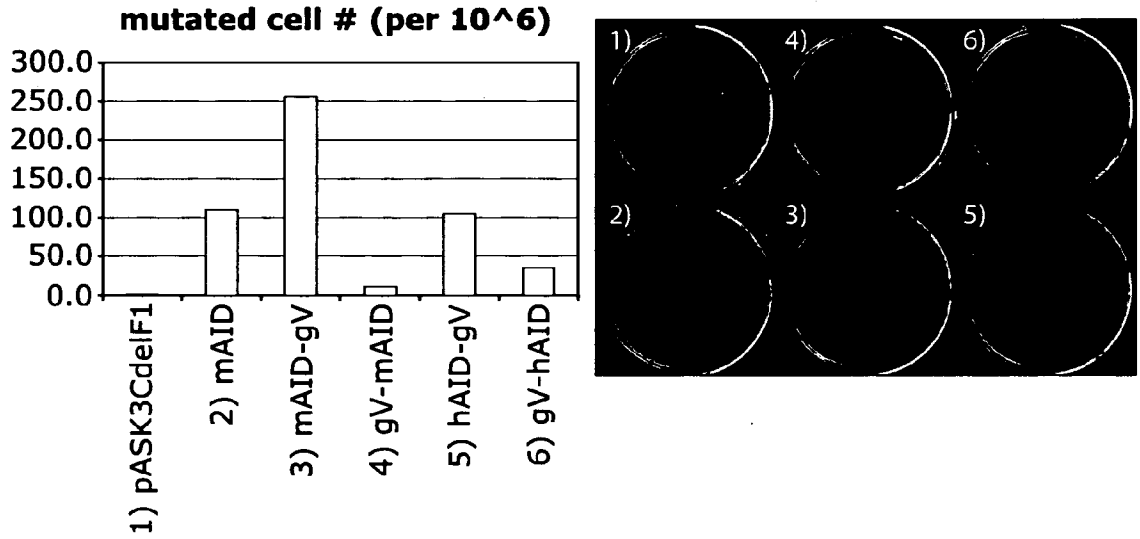

mutated cell # (per 10^6)
300.0
250.0
200.0
150.0
100.0
50.0
0.0
1) pASK3CdelF1
2) mAID
3) mAID-gV
4) gV-mAID
5) hAID-gV
6) gV-hAID
1)
4)
6)
2)
3)
5)

FIG. 2

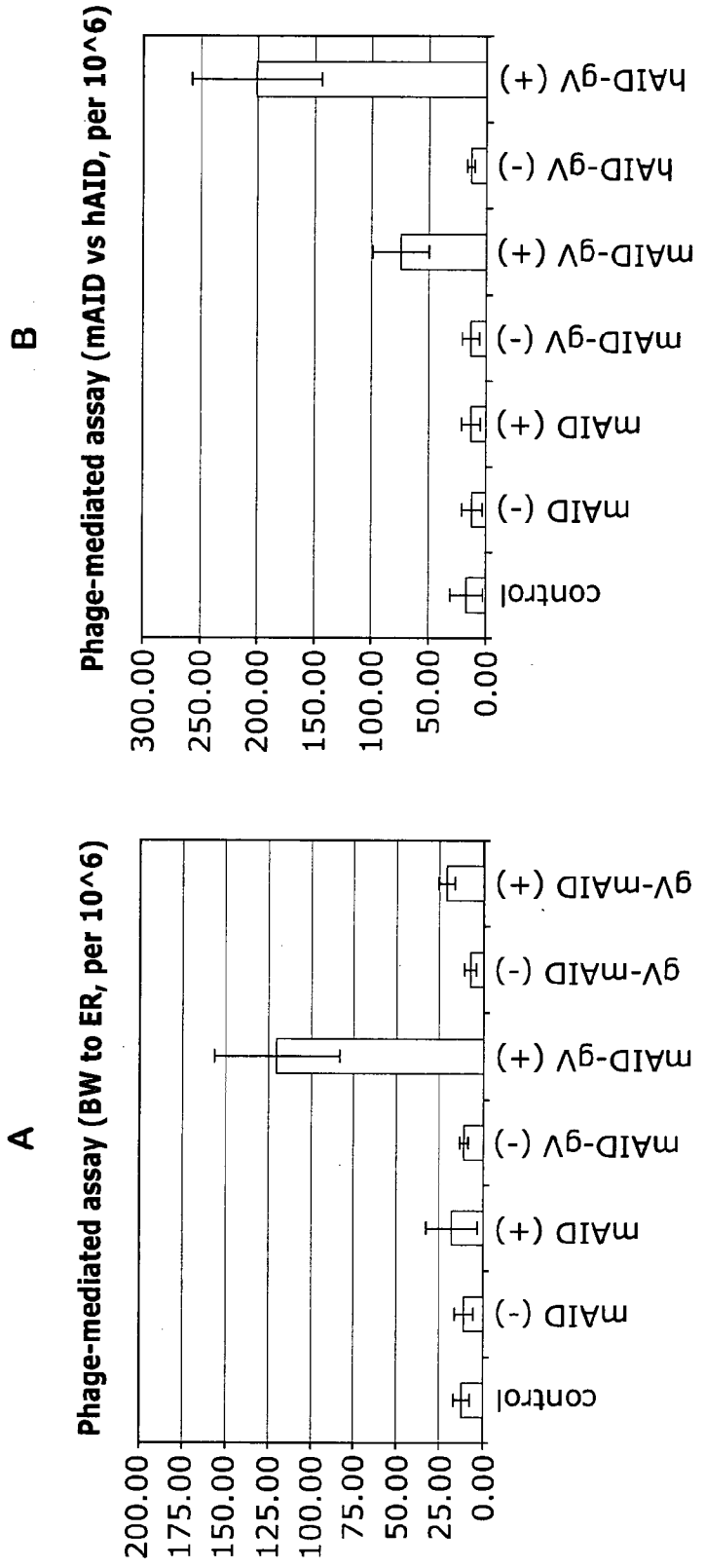

A
Phage-mediated assay (BW to ER, per 10^6)
200.00
175.00
150.00
125.00
100.00
75.00
50.00
25.00
0.00
control
mAID (-)
mAID (+)
mAID-gV (-)
mAID-gV (+)
gV-mAID (-)
gV-mAID (+)
B
Phage-mediated assay (mAID vs hAID, per 10^6)
300.00
250.00
200.00
150.00
100.00
50.00
0.00
control
mAID (-)
mAID (+)
mAID-gV (-)
mAID-gV (+)
hAID-gV (-)
hAID-gV (+)

FIG. 3

Dose-dependent bla reversion (phage-mediated assay)

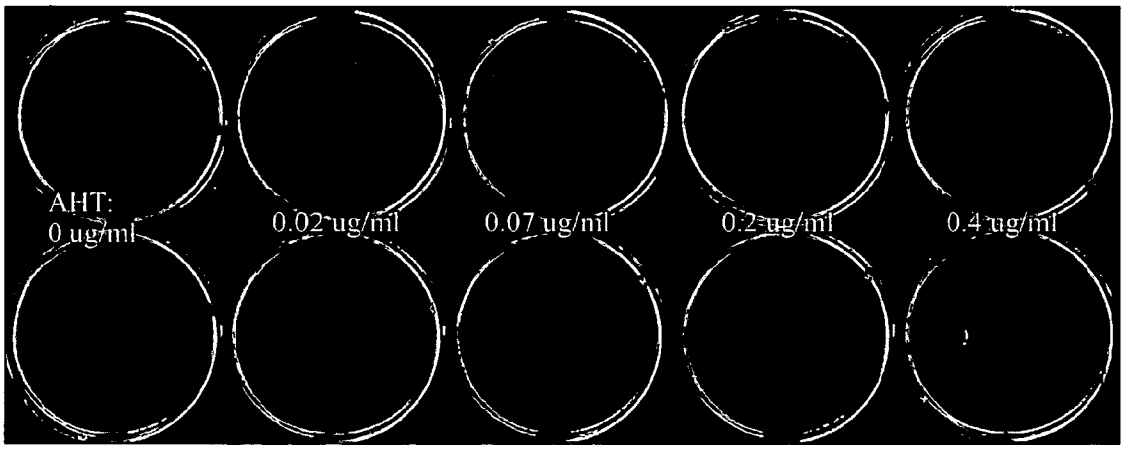


FIG. 4

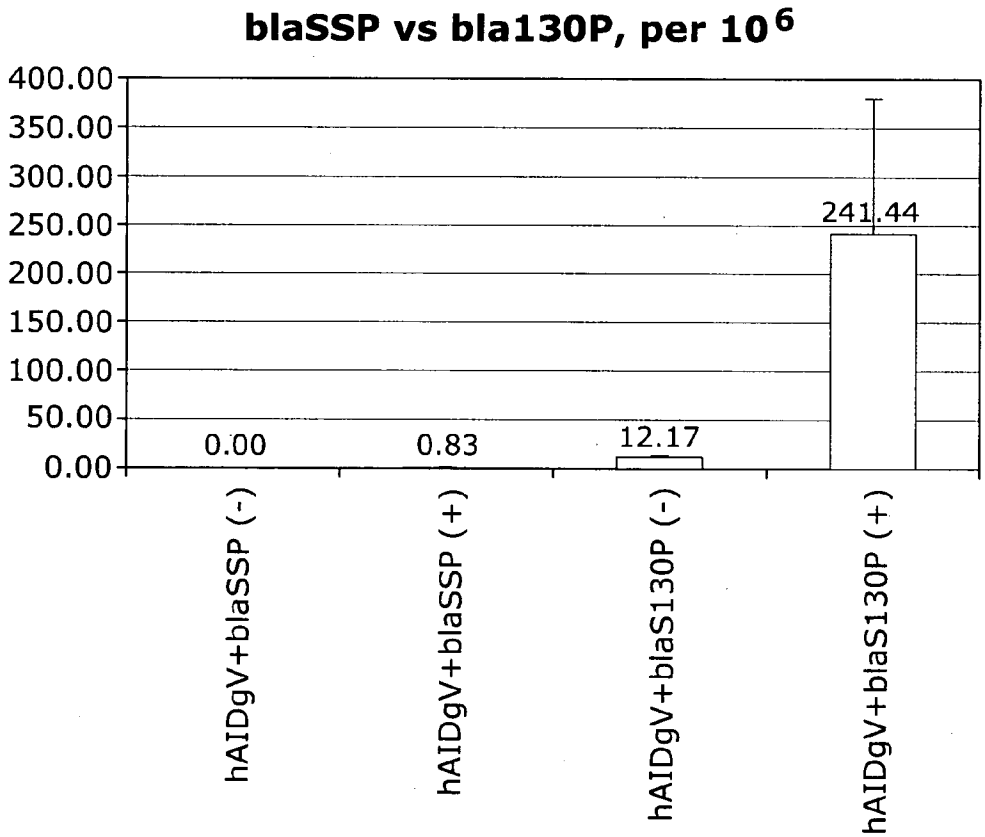

blaSSP vs bla130P, per 10 6
400.00
350.00
300.00
250.00
200.00
150.00
100.00
50.00
0.00
0.00
0.83
12.17
241.44
hAIDgV+blaSSP (-)
hAIDgV+blaSSP (+)
hAIDgV+blaS130P (-)
hAIDgV+blaS130P (+)

FIG. 5

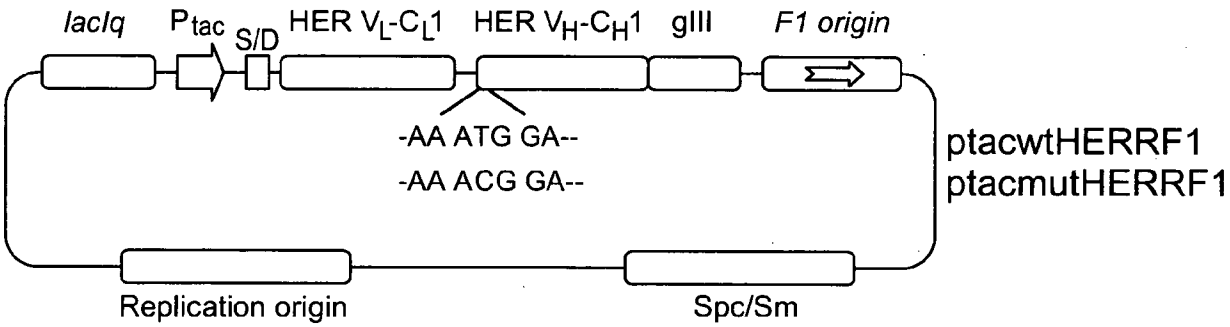

lacIq
Ptac
S/D
HER VL-CL1
HER VH-CH1
gIII
F1 origin
-AA ATG GA--
-AA ACG GA--
ptacwtHERRF1
ptacmutHERRF1
Replication origin
Spc/Sm

FIG. 6

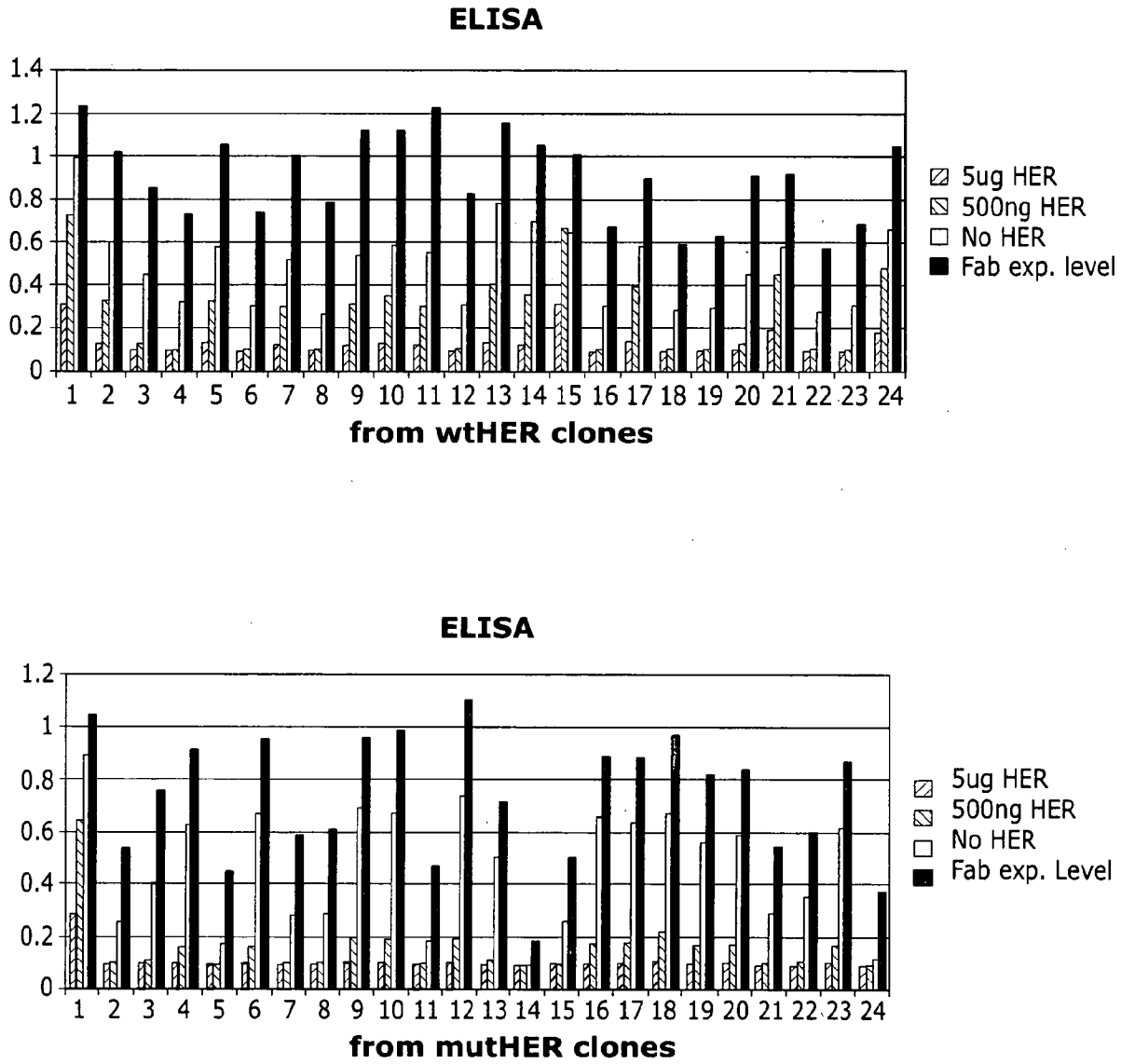

ELISA
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
from wtHER clones
5ug HER
500ng HER
No HER
Fab exp. level
ELISA
1.2
1
0.8
0.6
0.4
0.2
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
from mutHER clones
5ug HER
500ng HER
No HER
Fab exp. Level

FIG. 7

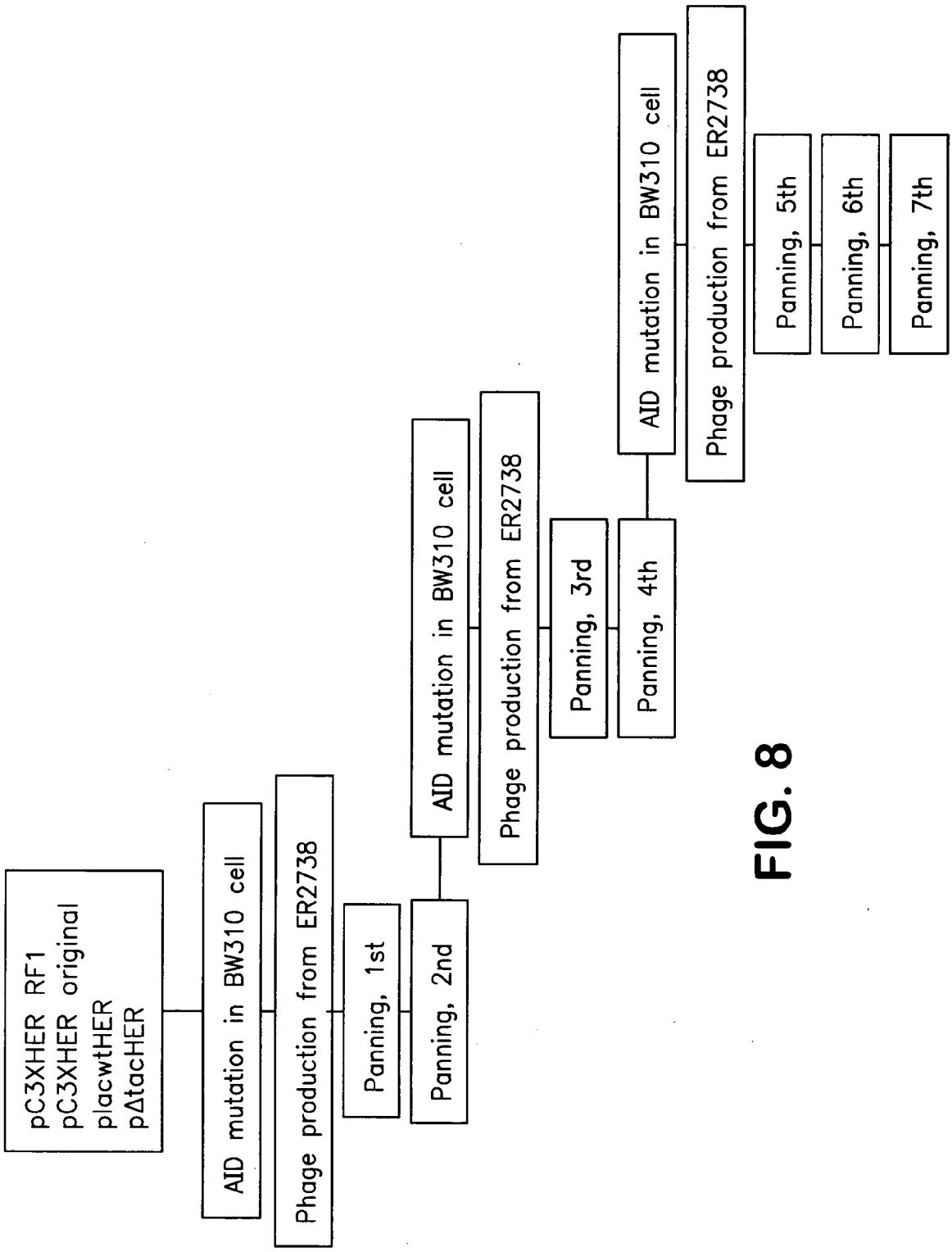

pC3XHER RF1
pC3XHER original
placwtHER
pΔtacHER
AID mutation in BW310 cell
Phage production from ER2738
Panning, 1st
Panning, 2nd
AID mutation in BW310 cell
Phage production from ER2738
Panning, 3rd
Panning, 4th
AID mutation in BW310 cell
Phage production from ER2738
Panning, 5th
Panning, 6th
Panning, 7th

FIG. 8

Panning result

	1st panning		2nd panning		3rd panning		4th panning	
	input phage	output phage	input phage	output phage	input phage	output phage	input phage	output phage
pC3XHER RF1	2.6E+11	1.1E+07	1.4E+11	1.2E+06	6.3E+11	1.3E+06	6.8E+11	2.4E+04
pC3XHER orig	3.3E+11	4.4E+06	1.7E+11	4.3E+05	5.4E+11	4.3E+05	7.1E+11	4.2E+04
placHER #1	6.0E+10	2.3E+06	4.2E+10	1.7E+05	5.8E+10	3.4E+05	1.6E+11	N/D
pdeltacHER #1	5.3E+10	2.1E+06	4.3E+10	1.7E+05	3.3E+10	1.7E+05	1.5E+11	4.2E+04

	5th panning		6th panning		7th panning	
	input phage	output phage	input phage	output phage	input phage	output phage
pC3XHER RF1	7.5E+11	2.6E+05	4.6E+11	1.9E+05	5.9E+11	1.5E+05
pC3XHER orig	5.6E+11	1.4E+05	5.4E+11	1.4E+05	5.4E+11	1.2E+05
placHER #1	2.8E+11	4.2E+04	4.7E+10	4.5E+04	4.7E+10	3.8E+04
pdeltacHER #1	1.1E+11	2.3E+05	4.2E+10	6.0E+04	4.2E+10	5.2E+04

N/D : Not Determined

FIG. 9

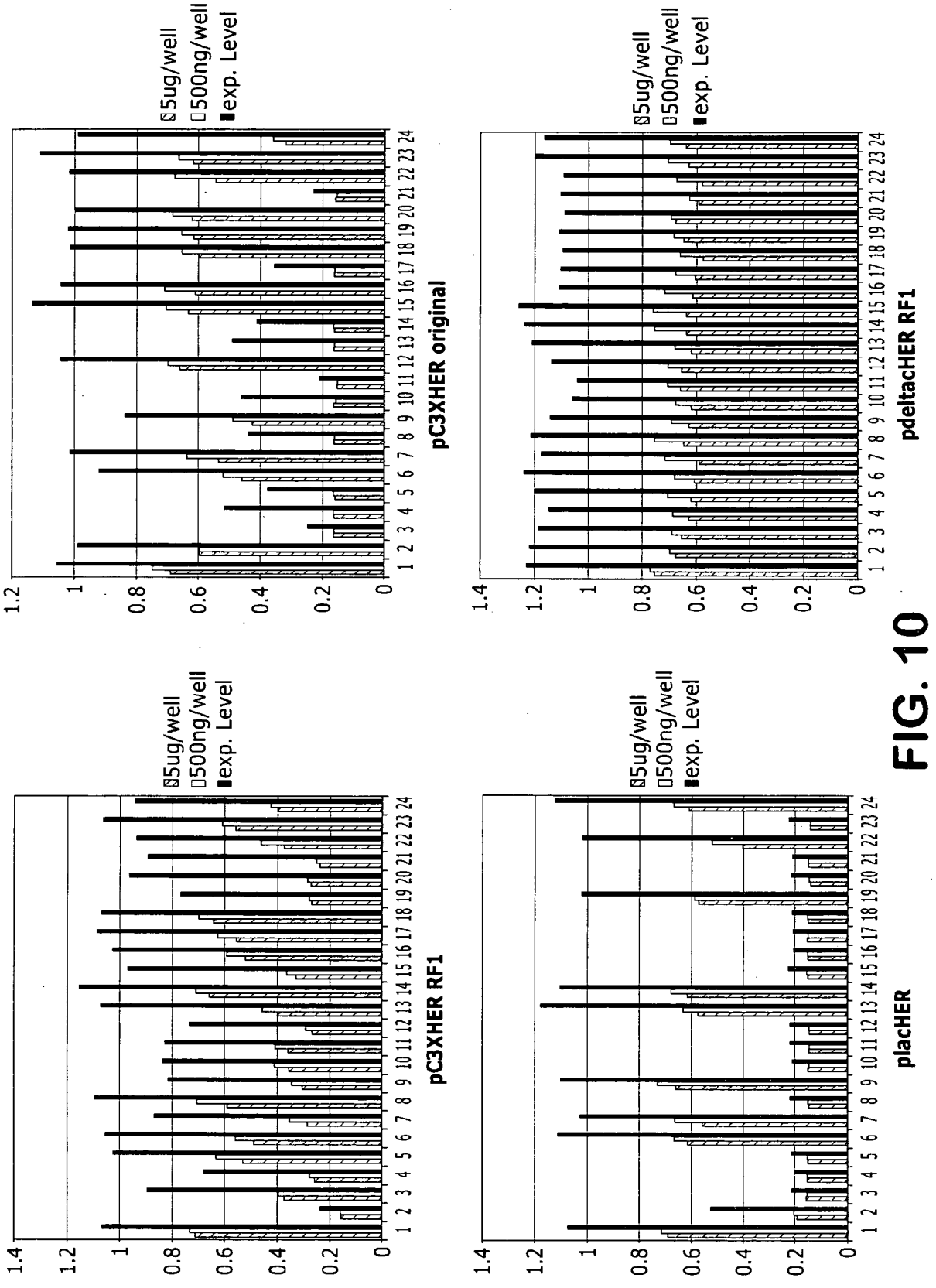

5ug/well
500ng/well
exp. Level
pC3XHER original
pdeltacHER RF1
pC3XHER RF1
placHER

FIG. 10

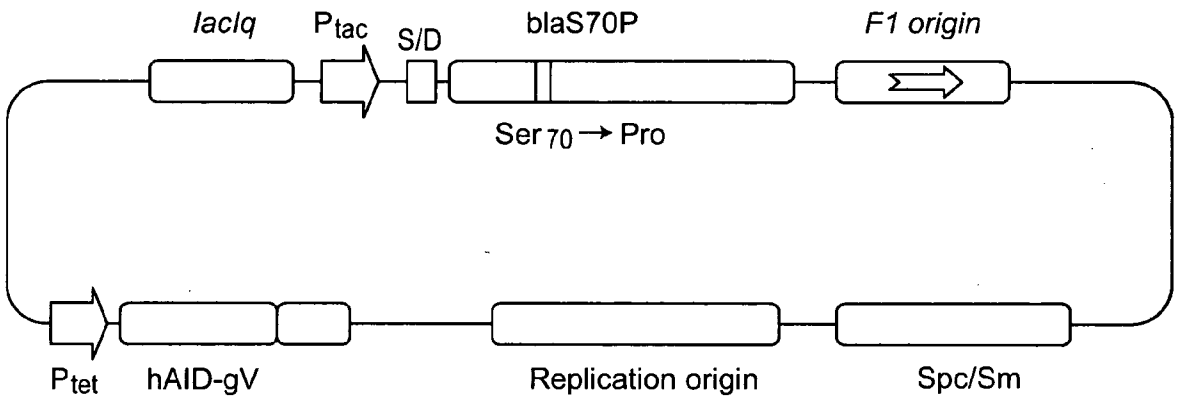

laclq
Ptac
S/D
blaS70P
F1 origin
Ser 70 → Pro
Ptet
hAID-gV
Replication origin
Spc/Sm

FIG. 11

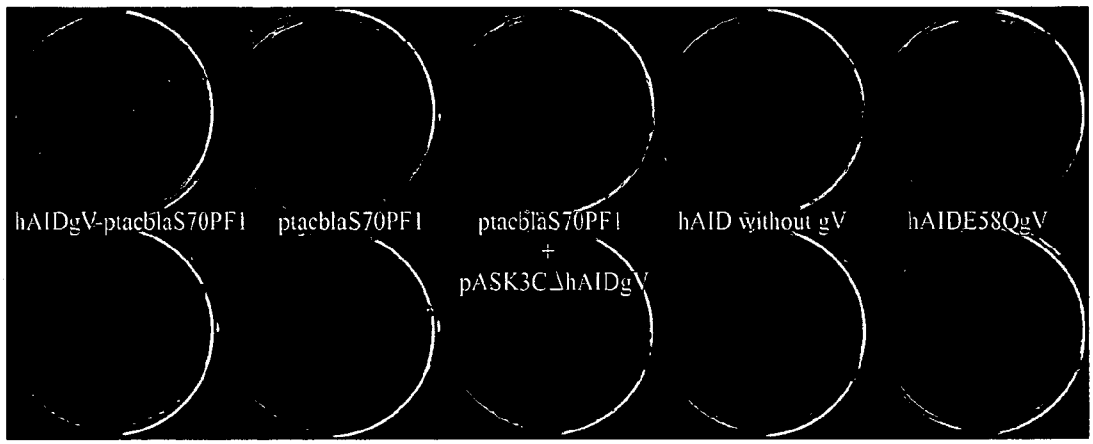

hAIDgV-ptacblaS70PF1
ptacblaS70PF1
ptacblaS70PF1
+
pASK3CΔhAIDgV
hAID without gV
hAIDE58QgV

FIG. 12