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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 February 2008 (21.02.2008)

PCT

(10) International Publication Number
WO 2008/020331 A2

(51) International Patent Classification:
C07K 14/43 (2006.01)

(21) International Application Number:
PCT/IB2007/003345

(22) International Filing Date: 16 August 2007 (16.08.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/837,912 16 August 2006 (16.08.2006) US

(71) Applicants *(for all designated States except US)*: **INSTITUT PASTEUR** [FR/FR]; 25-28, rue du Docteur Roux, F-75015 Paris (FR). **CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS)** [FR/FR]; 3, rue Michel Ange, F-75016 Paris (FR). **UNIVERSITA' DEGLI STUDI DI MILANO** [IT/IT]; Via Festa del Perdono, 7, I-20122 Milano (IT).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SALA, Monica** [FR/FR]; 3 rue François Mouthon, F-75015 Paris (FR). **GRECO, Raffaella** [IT/IT]; Via Togliatti 9, I-20067 Paullo (MI) (IT). **MICHEL, Marie** [FR/FR]; 2 rue Chalres Bossut, F-75012 Paris (FR). **GUETARD, Denise** [FR/FR]; 4B rue Anselme Payen, F-75015 Paris (FR). **WAIN-HOBSON, Simon** [GB/FR]; 3 rue Jean de la Fontaine, F-78180 Montigny-le-Bretonneux (FR). **SALA, Francesco** [IT/IT]; Via Genova 29, I-27100 Pavia (PV) (IT).

(74) Agent: **SANTARELLI**; B.P. 237, 14, avenue de la Grande-Armée, F-75822 Paris Cedex 17 (FR).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, 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, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv))*

Published:
— *without international search report and to be republished upon receipt of that report*
— *the filing date of the international application is within two months from the date of expiration of the priority period*

(54) **Title:** POLYNUCLEOTIDES ALLOWING THE EXPRESSION AND SECRETION OF RECOMBINANT HSBAG VIRUS-LIKE PARTICLES CONTAINING A FOREIGN PEPTIDE, THEIR PRODUCTION AND USE

(57) **Abstract:** The hepatitis B surface antigen (HBsAg) can assemble into sub-virion virus like particles (VLPs). Vectors comprised of polynucleotides, identified as GA1xFlag-M and GA3xFlag-M, are provided for the expression of foreign proteins in such VLPs, and for release of the VLPs from host cells containing the vectors. In one example, an HIV-1 polyepitope-HBsAg recombinant fusion protein assembled into VLPs and was efficiently secreted. In another example, the successful expression of novel recombinant HIV-1/HBV virus-like particles (VLPs) in Nicotiana tabacum and Arabidopsis thaliana is described. The production levels and quality of the recombinant VLPs were comparable in the two plants, showing that parameters intrinsic to the recombinant proteins determined their assembly into VLPs. Theses recombinant transgenes represent an innovative tool to set up a bivalent anti-HIV-1/-HBV vaccine based on oral administration of crude extracts from transgenic plants. In a final example, it is demonstrated that by oral administration of transgenic plant crude extracts to humanized HSB mice it is possible to induce the activation of anti- HIV-1 specific CD8+ T cells in peripheral lymph nodes and spleen.

WO 2008/020331 A2

Polynucleotides allowing the expression and secretion of recombinant hbsag virus-like particles containing a foreign peptide, their production and use

[001] This invention relates to polynucleotides for the expression of recombinant hepatitis B surface antigen (HBsAg) virus-like particles (VLPs) and to the secretion of the particles from host cells. The recombinant HBsAg virus-like particles can contain a foreign peptide or polypeptide, such as foreign amino acid residues of a pathogen. The HBsAg virus-like particles are particularly useful in immunogenic compositions and as vaccines.

[002] Many viral structural proteins have the intrinsic ability to assemble into virus-like particles (VLPs) independently of nucleic acids. VLPs can elicit potent anti-viral humoral and cellular immune responses directed against viruses they derive from (10, 24, 36, 37). They are efficiently taken up, rapidly internalised, and processed by antigen presenting cells (APCs) of myeloid origin, leading to MHC class I-associated antigen cross-presentation (1, 17, 33-35, 38). Indeed, MHC class I cross-presentation of VLP epitopes by APCs can be exploited to induce anti-viral $CD8^+$ cytotoxic T lymphocyte (CTL) responses. VLPs are powerful antigen delivery systems, the most developed examples being hepatitis B surface antigen (HBsAg), the yeast Ty retrotransposon structural protein "a" (Tya), the VP2 capsid protein of porcine parvovirus (PPV), and the papillomavirus capsid L1 protein. The generation of recombinant VLPs bearing relevant antigens opens up the way to the development of bivalent vaccine candidates (19, 21, 30).

[003] Even though the three envelope proteins of hepatitis B virus (HBV), (large, middle, and small: L, M and S protein, respectively) are encoded by a sole open reading frame (orf), they are encoded by distinct regions of the orf as a result of two different mRNA transcripts (L by preS1 + pre S2 + S regions of a first mRNA; M and S by pre S2 + S regions and S region of the second mRNA, respectively). The preS2 translation initiation codon is less efficient than the S region one (HBsAg) meaning that the mRNA is bicistronic (14).

[004] HBsAg carries all the information necessary for membrane translocation, particle assembly, and secretion from mammalian cells (5). S ubstitutions within HBsAg that impair VLPs assembly are generally characterized by HBsAg accumulation in the endoplasmic reticulum (ER) and Golgi apparatus (8).

[005] By fusing foreign DNA to the S protein gene, HBsAg has been used as carrier for a wide panel of antigens (12, 19, 21, 27, 30). In a notable example, a series of 13 HIV-1 epitopes restricted by the HLA-A*0201 class I allele, which is present at ~15-30% of Black, Caucasian, and Oriental populations, was incorporated into the preS2 region as a polyepitope (polHIV-1) fused to HBsAg. Although the study reported the induction of HIV-1 specific CTL responses by DNA vaccination (12) of humanised HLA-A*0201 transgenic mice (11), it was not shown whether the recombinant HBsAg actually formed VLPs.

[006] Thus, there exists a need in the art for vectors for the expression of recombinant peptides or polypeptides, such as HIV polyepitopes, compatible with VLPs formation and secretion of recombinant VLPs from host cells. Preferably, recombinant VLPs secretion should result in the induction of robust neutralising anti-HBsAg humoral immune responses and the enhancement of the activation state of foreign, sequence specific $CD8^+$ T lymphocytes so that the VLPs can be employed in therapeutic applications.

A previous HLA.A2.1-restricted HIV-1 polyepitope was constructed with the aim of triggering an antiviral cellular immune response. It has been discovered by inventors of the present patent application that fused to the HBsAg protein, this polyepitope impairs the secretion of virus-like particles (VLPs). This invention involves the design of polynucleotides and expression vectors for cloning and expressing foreign peptides or polypeptides, such as HIV-1 polyepitopes, as tagged HBsAg fusion proteins in HBsAg VLP. Polynucleotides and expression vectors comprising these polynucleotides have been designed, all preserving recombinant HBsAg VLPs formation and secretion.

[007] Thus, in one aspect, this invention concerns: i) the GA1xFlag-M and GA3xFlag-M polynucleotides, which once inserted in HBsAg expression vectors induce optimal expression of recombinant HBsAg VLPs; ii) new expression vectors comprising the GA1xFlag-M or GA3xFlag-M polynucleotide for optimal expression of recombinant HBsAg VLPs; iii) the use of these new expression vectors for the production of recombinant HBsAg VLPs that are secreted from host cells; and iv) the recombinant expression vectors obtained assembling the nucleic acids encoding foreign peptides or polypeptides to the new expression vectors.

[008] More particularly, this invention provides a polynucleotide comprising GA1xFlag-M or GA3xFlag-M polynucleotide. Specifically, this invention provides a polynucleotide comprising:

CAGGCCATGCAGTGGAACTCCACAcccgggGCTGGAGCAGGAG
CTGATTACAAGGACGACGACGACAAGgaattcCTGCAGGCTAGC
AGATCTctcgagCTGAACATG [SEQ ID NO: 1];

and a polynucleotide comprising:

CAGGCCATGCAGTGGAACTCCACAcccgggGCTGGAGCAGGAG
CTGACTACAAAGACCACGACGGTGATTATAAAGATCACGACAT
TGATTACAAGGACGACGACGACAAGgaattcCTGCAGGCTAGCA
GATCTctcgagCTGAACATG [SEQ ID NO: 2].

[009] The polynucleotide can further comprise a eukaryotic promoter sequence, a nucleotide sequence encoding hepatitis B surface antigen protein (HBsAg), a polyadenylation sequence, or combinations of these elements.

[010] In one embodiment of the invention, the polynucleotide comprises a nucleotide sequence encoding hepatitis B surface antigen protein (HBsAg) devoid of translation initiation ATG and positioned downstream and in frame with the GA1xFlag-M or the GA3xFlag-M polynucleotide sequence. The polynucleotide can comprise a polyadenylation sequence operably linked to the other sequences.

[011] This invention provides a polynucleotide of the invention comprising the polynucleotide sequence cloned between HindIII and AvrII restriction sites in pGA1xFlag-M plasmid deposited at the CNCM on December 16, 2005, under the Accession Number I-3543.

[012] This invention also provides a polynucleotide of the invention comprising the polynucleotide cloned between HindIII and AvrII restriction sites in pGA3xFlag-M plasmid deposited at the CNCM on December 16, 2005, under the Accession Number I-3545.

[013] In addition, this invention provides a polynucleotide hybridizing under stringent conditions to a polynucleotide of the invention, or its complement, such as variant polynucleotides resulting from degeneracy of the genetic code. Another example of a variant polynucleotide is a polynucleotide having a different restriction site or sites in the polylinker, provided that any removed or modified restriction site does not disrupt the translation frame.

[014] Further, this invention provides a polynucleotide of the invention, which further comprises a foreign coding polynucleotide inserted in any of restriction sites of the GA1xFlag-M or GA3xFlag-M polynucleotide and in frame with the ATG at position 7 in the GA1xFlag-M or GA3xFlag-M polynucleotide sequence. The polynucleotide is useful for preparing recombinant DNA constructs prior to insertion into a vector.

[015] A cloning and/or expression vector comprising a polynucleotide of the invention is also provided.

[016] Further, this invention provides a eukaryotic host cell comprising a vector of the invention. In one embodiment, the vector in the eukaryotic host cell can comprise an eukaryotic promoter sequence operably linked to a nucleotide sequence encoding HBsAg protein for expression of HBsAg virus-like particles. Optionally, the vector can comprise a nucleotide sequence encoding a HBsAg fusion protein comprising a foreign polypeptide and HBsAg protein, wherein the eukaryotic host cell produces HBsAg virus-like particles constituted by the HBsAg fusion protein and HBsAg protein.

[017] This invention also provides a method of producing HBsAg virus-like particles. The method comprises providing a host cell of the invention, and expressing the fusion proteins and HBsAg proteins under conditions in which the proteins assemble into virus-like particles, which are released from the host cell into extracellular space. Optionally, the method comprises recovering the virus-like particles.

[018] The invention also provides Virus Like Particules (VLP) susceptible to be produced by the method of the invention and the use of such VLP in the preparation of a composition intended to immunize a human or an animal.

[019] Further, the invention provides a method of preparing a HBsAg fusion protein, wherein the method comprises providing a host cell of the invention, and expressing a tagged HBsAg fusion protein and HbsAg protein under conditions in which the proteins assemble into virus-like particles. The particles are released from the host cell into extracellular space. The VLP bearing tagged HBsAg fusion proteins from the bacteria culture can be separated by capture with Flag-M antibodies, HBsAg antibodies, or by use of both of these types of antibodies.

[020] This invention provides an expression vector selected from the deposited, recombinant vectors pGA1xFlag-M (CNCM No. I-3543), pGA3xFlag-M (CNCM No. I-

3545), pGA1xFlag-M pol.opt (CNCM No. I-3544), pGA3xFlag-M pol.opt (CNCM No. I-3546), pGA1xFlag-M.pol1A2 (CNCM No. I-3579), pGA1xFlag-M.pol2A2 (CNCM No. I-3580), pGA1xFlag-M.pol1B7 (CNCM No. I-3581), and pGA1xFlag-M.pol2B7 (CNCM No. I-3582). A preferred embodiment of this invention provides a polynucleotide comprising the sequence between HindIII and AvrII restriction sites of one of these expression vectors.

[021] This invention provides a polypeptide encoded by a polynucleotide or by a vector according to the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[022] This invention will be described with reference to the drawings in which:

Figure 1(A) is a schematic representation of plasmid pCMV-B10.

Figure 1(B) is the complete nucleic acid sequence of pCMV-basic.

Figure 2 is the nucleic acid sequence of pGA1xFlag-M.

Figure 3 is the nucleic acid sequence of pGA3xFlag-M.

Figure 4 is a more detailed nucleic acid sequence for the GA1xFlag-Mpol.opt insert in pGA1xFlag-M to produce pGA1xFlag-Mpol.opt.

Figure 5 is a more detailed nucleic acid sequence for the GA3xFlag-Mpol.opt insert in pGA3xFlag-M to produce pGA3xFlag-Mpol.opt.

Figure 6 is the nucleic acid sequence of pGA1xFlag-Mpol.opt.

Figure 7 is the nucleic acid sequence of pGA3xFlag-Mpol.opt.

Figure 8 depicts the secretion kinetics corresponding to pGA1xFlag-M, pGA1xFlag-Mpol.opt, pGA3xFlag-M and pGA3xFlag-Mpol.opt.

Figure 9 depicts the results of a semi-quantitative anti-Flag-M ELISA on transfected SW480 supernatants.

Figure 10 is pGA1xFlag-M.pol1A2 nucleic acid sequence (in bold: pol1A2 polyepitope).

Figure 11 is pGA1xFlag-M.pol2A2 nucleic acid sequence (in bold: pol2A2 polyepitope).

Figure 12 is pGA3xFlag-M.pol1A2 nucleic acid sequence (in bold: pol1A2 polyepitope).

Figure 13 is pGA3xFlag-M.pol2A2 nucleic acid sequence (in bold: pol2A2 polyepitope).

Figure 14 is pGA1xFlag-M.pol1B7 nucleic acid sequence (in bold: pol1B7 polyepitope).

Figure 15 is pGA1xFlag-M.pol2B7 nucleic acid sequence (in bold: pol2B7 polyepitope).

Figure 16 is pGA3xFlag-M.pol1B7 nucleic acid sequence (in bold: pol1B7 polyepitope).

Figure 17 is pGA3xFlag-M.pol2B7 nucleic acid sequence (in bold: pol2B7 polyepitope).

Figure 18 depicts the secretion kinetics corresponding to pGA1xFlag-Mpol1.A2 and pGA1xFlag-Mpol2.A2.

Figure 19 depicts the secretion kinetics corresponding to pGA3xFlag-Mpol1.A2 and pGA3xFlag-Mpol2.A2.

Figure 20 depicts the secretion kinetics corresponding to pGA1xFlag-Mpol1.B7 and pGA1xFlag-Mpol2.B7.

Figure 21 depicts the secretion kinetics corresponding to pGA3xFlag-Mpol1.B7 and pGA3xFlag-Mpol2.B7.

Figure 22: oligonucleotide used for engineering the pGA3xFlagbasic, pGA1xFlag-Mbasic and pGA3xFlag-Mbasic plasmids

Figure 23: ppolHIV-1.opt plasmid (a); ppolHIV-1.opt plasmid redesigned (b) ; "1xFlag-M" tag (c) and "3xFlag-M" tag (d)

Figure 24: Hydropathy profiles of the amino acid sequences between the preS2 ATG codon and the HBsAg stop codon from (a) the GA1xFlag-Mbasic transgene, (b) the GA3xFlag-Mbasic transgene, (c) the GA1xFlag-Mpol.opt transgene, (d) the GA3xFlag-Mpol.opt transgene, (e) the preS2-HBsAg polyproteine from accession number U95551 HBV ayw isolate, and (f) the HIV-1/HBV transgene from Shchelkunov et al. [19]. On the x axis: position in the protein amino acid sequence. On the y axis: hydrophobicity scores. A score of 4.5 is the most hydrophobic and a score of -4.5 is the most hydrophilic.

Figure 25: Anti-HBsAg and anti-Flag-M ELISA analyses on samples from mammalian cells transient transfections.

Figure 26: Schematic representation of the Flag-M constructs used for plant transformation.

Figure 27: Southern blot and anti-HBsAg ELISA analyses on transgenic *Nicotiana tabacum* T0 plants. Protein data refer to protein extraction E1-A. The 14 plants selected for further analyses are highlighted in violet. [a]: TSP : total soluble protein; [b]:reported values are the mean among three independent measurements; [c]: n.t. : not tested; [d]:n.d.: not detectable.

Figure 28: Southern blot and anti-HBsAg ELISA analyses on transgenic *Arabidopsis thaliana* plants. The 16 plants selected for further analyses are highlighted in green. [a]: TSP : total soluble protein; [b]: reported values are the mean among three independent measurements; [c]: n.t. : not tested; [d]: n.d.: not detectable.

Figure 29: VLPs production in transgenic Tobacco and Arabidopsis plants.

Figure 30: Characterization of the 14 selected transgenic tobacco plants.

Figure 31: Anti-Flag-M ELISA on *Nicotinia tabacum* protein extracts

Figure 32: Southern blot and anti-HBsAg ELISA analyses on T1 transgenic *Nicotiana tabacum* plants. [a]: TSP: total soluble protein; [b]: reported values are the mean among three independent measurements; [c]: n.d.: not detectable; [d]: z-correlation test.

Figure 33: Anti-Flag-M ELISA tests on T1 *Nicotiana tabacum* protein extracts from (a) GA1xFlag-Mbasic, (b) GA3x Flag-Mbasic, (c) GA1xFlag-Mpol.opt and (d) GA3xFlag-Mpol.opt plants from Figure 32.

Figure 34: Characterization of 16 selected transgenic Arabidopsis plants

Figure 35: Anti-Flag-M ELISA on *Arabidopsis thaliana* proteins extracts.

Figure 36: Schematic representation of vaccination protocol 1 (**A**) and 2 (**B**). Cardiotoxin, plasmid DNA and lyophilised transgenic plants administration timing (detailed per day: d) and quantity are indicated. The plasmid DNA was the previously described pGA1xFlag-Mpol.opt [Michel, 2007 #142].

Figure 37: INF-γ *ex vivo* secretion assay on lymphocytes from mesenteric lymph nodes pooled from 3 mice having received tobacco stock #5 following protocol 1. Presented data correspond to over night stimulation with. **A**) the pool of S9L, L9V, L10V and Y/I9V HIV-1 peptides; **B**) the pool of V11V, Y/P9L, Y/V9L and Y/T9V HIV-1 peptides; or **C**) the irrelevant G9L peptide [Michel, 2007 #80]. Percentages of CD8+ T cell subsets secreting INF-γ in total lymphocytes are highlighted in bold.

Figure 38: Foxp3 intra-cellular labeling on cells from spleen (**A**) and pool of peripheral lymph nodes from 3 mice (**B**). The percentages of Foxp3+ cells among CD3+CD4+ T lymphocytes are indicated on the y axis. On the x axis the different groups of mice are indicated: naïve: not receiving any treatment; wt: mice primed with plasmid DNA and boosted by wild type tobacco; #4: mice primed with plasmid DNA and boosted by tobacco stock 4; #5: mice primed with plasmid DNA and boosted by tobacco stock 5. Mean values for each group are represented by horizontal lines. In boxes, mean values are calculated without taking into account external points. Significant ($p < 0.05$) non-parametric Mann-Whitney tests are indicated by a star (*) on horizontal bars indicating compared groups.

Figure 39: INF-γ *ex vivo* secretion assay on CD8+ T lymphocytes following cell sorting by magnetic beads. The percentages of INF-γ secreting cells among CD8+CD3+ T lymphocytes are indicated on the y axis. The CD8+CD3+ T lymphocytes were put in the presence of feeder cells charged either with relevant (HIV) or irrelevant (G9L) peptides as indicated on the x axis. Relevant peptides correspond to the pool of S9L, L9V, L10V, Y/I9V, V11V, Y/P9L, Y/V9L and Y/T9V peptides. Mean values for each group are represented by horizontal lines. Significant ($p < 0.05$) non-parametric Wilcoxon signed-rank tests are indicated by a star (*) on horizontal bars indicating compared groups.

[023] The hepatitis B surface antigen (HBsAg) can assemble into sub-virion virus like particles (VLPs). By fusing immunogenic peptides to the amino-terminus of HBsAg, several bivalent vaccines have been developed. Notably, a polyepitope bearing HIV-1 epitopes restricted to the HLA-A*0201 class I allele elicited a significant HIV-1 specific $CD8^+$ cytotoxic T lymphocyte (CTL) response *in vivo* (12). This recombinant HBsAg failed to form VLPs due to retention in the Golgi apparatus.

[024] In contrast, this invention provides polynucleotides and expression vectors for the production of recombinant proteins as tagged HBsAg fusion proteins, which assemble into VLPs and which are efficiently secreted by host cells. It is thus possible to make self-assembling recombinant HBsAg VLPs with residues of another protein. This is demonstrated for HIV-1 polyepitopes, and thus provides efficient bivalent HBV/HIV vaccines, which are particularly apposite given that these two viruses are frequently associated.

[025] More particularly, this invention provides two polynucleotide motifs (GA1xFlag-M and GA3xFlag-M) for cloning and expressing foreign sequences fused in frame to HBsAg. In a particular embodiment of the present patent application, these polynucleotides have been inserted into two distinct expression vectors designated herein as pGA1xFlag-M and pGA3xFlag-M. These plasmids were constructed as follows.

[026] The plasmid pCMV-B10 was previously known. *See Eur. J. Immunol.* 2001, 31:3064-3074. Fig. 1(A) is a schematic representation of the pCMV-B10 plasmid.

[027] Specifically, the pCMV-B10 plasmid vector is a pcDNA3 derivative (Invitrogen, Costa Mesa, CA), in which the nucleotide sequences of the hepatitis B middle [initiation at Position 900 (ATG), termination at Pos. 1744(TAA)] and small [initiation at Pos. 1066 (ATG), termination at Pos. 1744 (TAA)] envelope proteins have been inserted downstream of a human CMV immediate early promoter. A polyadenylation signal is provided by the HBV untranslated sequence (nucleotides 1744-2899). The central part of the coding preS2 segment sequence was replaced by a polylinker in which a polyepitope DNA can be inserted. An HIV-1 derived (MN isolate) V3 loop tag was inserted, downstream of the polyepitope.

[028] Using the pCMV-B10 plasmid, a plasmid designated ppolHIV-1.opt was constructed by cloning an HIV-1 polyepitopic sequence between the EcoRI and XhoI restriction sites in pCMV-B10. The HIV-1 polyepitopic sequence is identified as polHIV-1.opt and has the following amino acid sequence:

YLKEPVHGVRAKT**YLNAWVKVV**RDTA**VLDVGDAYFSV**RAKT**YLVKLWYQL**RADT**RLY NTVATL**RTKA**LLDTGADDTV**RAKT**LLWKGEGAV**RTDA**YIYQYMDDL**R.

[029] Using the ppolHIV-1.opt plasmid, a control plasmid, pCMV-basic (Fig. 1B), was constructed for use in expression studies involving the vectors of the invention. Specifically, the polHIV-1.opt polyepitope in ppolHIV-1.opt was removed by digestion with EcoRI and XhoI and substituted by a polylinker comprising EcoRI, NheI, EcoRV, SmaI, and XhoI restriction sites. In between the EcoRI and XhoI restriction sites, the NheI, EcoRV, SmaI restriction sites follow one another as shown below.

Polylinker of the pCMV-basic plasmid (EcoRI-XhoI)

Nucleic acid sequence

GAATTC-A-GCTAGC-GATATC-CCCGGG-CTCGAG

Restriction enzyme sequence :

EcoRI-NheI-EcoRV-SmaI-XhoI.

[030] The pGA1xFlag-M (Figure 2) and pGA3xFlag-M (Figure 3) plasmids of the invention were then constructed from the pCMV-basic plasmid. Specifically, the newly created pGA1xFlag-M and pGA3xFlag-M constructions only maintained the pCMV-basic plasmid backbone, with its CMV promoter and HBV polyadenylation signal. The nucleic acid sequence between the HindIII and the AvrII restriction sites in the pCMV-basic plasmid was eliminated. Nucleic acid sequences were then cloned between the HindIII and AvrII restriction sites. Nucleic acid sequences corresponding to the cloned sequences in between the HindIII and AvrII of the pCMV-basic plasmid are given below for the pGA1xFlag-M and pGA3xFlag-M plasmids:

pGA1xFlag-M and pGA3xFlag-M nucleic acids sequences between the HindIII (aagctt) and AvrII (cctagg) restrictions sites

Nucleic acid sequence of pGA1xFlag-M

aagcttCAGGCCATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTG
ATTACAAGGACGACGACGACAAGgaattcCTGCAGGCTAGCAGATCTctcgagCT
GAACATGGAGAACATCACATCAGGATTcctagg

Nucleic acid sequence of pGA3xFlag-M

aagcttCAGGCCATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTG
ACTACAAAGACCACGACGGTGATTATAAAGATCACGACATTGATTACAAGGA
CGACGACGACAAGgaattcCTGCAGGCTAGCAGATCTctcgagCTGAACATGGA
GAACATCACATCAGGATTcctagg

EcoRI (gaattc) and XhoI (ctcgag) restriction sites are provided in these sequences for cloning foreign sequences in the vectors.

[031] More particularly, starting from the HindIII restriction site in 5', nucleotides encoding the QA peptide corresponding to C-terminal sequence of the HBV preS1 region were introduced. These nucleotides were inserted to preserve the native context of ATG start codon of preS2, hence the strength of this translation initiation codon. Then nucleotides encoding the MQWNSTP peptide corresponding to the N-terminal portion of the HBV preS2 region were

introduced. This modification was made to reintroduce in the pGA1xFlag-M and pGA3xFlag-M plasmids the glycosylation site (N4*), which was absent in the pCMV-basic construction.

[032] The preS2 peptide is followed by the GA motif (Gly-Ala amino acids repeated three times), here located to prevent steric impairment for the binding of anti-Flag antibodies to the tag (1xFlag-M or 3xFlag-M) by sugar molecules covalently linked to preS2 N4*.

[033] The GA motif is followed by the 1xFlag-M or 3xFlag-M tag (modified from SIGMA-ALDRICH), giving the pGA1xFlag-M and pGA3xFlag-M plasmids, respectively. Specifically, the 1xFlag and 3xFlag nucleic acid sequences from SIGMA-ALDRICH were modified to eliminate the ATG codons in the second and third possible reading frames of the sequences. Nucleic acid sequences of the original SIGMA-ALDRICH1xFlag and 3xFlag were modified preserving amino acids sequences, but eliminating methionine residues in secondary and tertiary phases.

[034] In the pGA1xFlag-M and pGA3xFlag-M plasmids, between the EcoRI and XhoI restriction sites, a new polylinker was inserted, where the PstI, NheI, BglII restriction sites follow one the others.

[035] In the pGA1xFlag-M and pGA3xFlag-M plasmids, the preS2 C-terminal peptide preceding the HBsAg ATG start codon (M_1) was reduced to two amino acids (Leu-Asn), preserving the nucleic acid context, hence the strength, of the HBsAg translation initiation ATG codon. Finally, nucleic acid sequence of HBsAg protein from the ATG start codon to the seventh codon was inserted in the pGA1xFlag-M and pGA3xFlag-M plasmids in order to preserve the N-terminal sequence of HBsAg deleted by HindIII-AvrII digestion.

[036] The sequence cloned between HindIII and AvrII sites in pGA1xFlag-M and pGA3xFlag-M plasmids were obtained by "atypical" PCR, as described below, using long primers (60-80 nucleic acids) all corresponding to the 5'-3' strand of the final product, and cloning in between the HindIII and AvrII of the pCMV-basic plasmid.

[037] The nucleic acid sequences of pGA1xFlag-M and pGA3xFlag-M are given in Figs. 2 and 3, respectively. In Fig. 2, the GA1xFlag-M sequence is underlined. The GA1xFlag-M nucleic acid sequence in bold corresponds to the

following amino acid sequence, which represents in the order: the preS1 C-terminal sequence (QA)- the preS2 ATG - the glycosylation site (N* in the N-terminal protion of preS2 polypeptide QWNSTP)– the GA motif – the 1xFlag-M tag – the polylinker – the preS2 C-terminal portion (LN) – the HBsAg ATG :

QAMQWN*STPGAGAGADYKDDDDKEFLQASRSLELNM.

The 1xFlag-M sequence is underlined above. The remainder of the sequence depicted is the vector.

[038] In Fig. 3, the GA3xFlag-M sequence is underlined. The GA3xFlag-M nucleic acid sequence in bold corresponds to the following amino acid sequence which represents in the order: the preS1 C-terminal sequence (QA) - the preS2 ATG - the glycosylation site (N* in the N-terminal protion of preS2 polypeptide QWNSTP) – the GA motif – the 3xFlag-M tag – the polylinker – the preS2 C-terminal portion (LN) – the HBsAg ATG :

QAMQWN*STPGAGAGADYKDHDGDYKDHDIDYKDDDDKEFLQASRSLELNM.

The 3xFlag-M sequence is underlined above.

[039] The polynucleotides and plasmids of the invention are particularly useful for cloning and expressing foreign sequences. Polynucleotides and plasmids provide the nucleic sequence context to produce a fusion polypeptide where the HBsAg protein is the carrier, referred herein as HBsAg fusion protein, which is efficiently secreted by host cells and assembles with HBsAg protein as HBsAg virus-like particles. HBsAg fusion proteins produced by using the polynucleotides of the invention and the expression vectors of the invention have the relevant characteristic to carry the 1xFlag or 3xFlag tag. For this reason, HBsAg fusion proteins of the invention are also named tagged HBsAg fusion proteins.

[040] The HBsAg fusion proteins are characterized herein as "tagged" because of the presence of 1xFlag-M and 3xFlag-M tags in the GA1xFlag-M and GA3xFlag-M motifs, respectively. This tag is very relevant because 1) it makes it possible to follow the incorporation of HBsAg fusion proteins in VLP; 2) it can be used to purify recombinant VLPs from the cell culture medium; and 3) it can be used to further purify or isolate the HBsAg fusion protein from the VLP and from the HBsAg proteins of the recombinant VLP.

[041] It is interesting to note that even without foreign nucleotide sequence insertion, two types of proteins are synthesized by the expression vectors of the invention as pGA1xFlag-M and pGA3xFlag-M : the HBsAg protein and a tagged HBsAg protein translated from the preS2 ATG and carrying the Flag tag. The tagged HBsAg VLP obtained by assembling these two proteins makes it possible to follow the formation of the VLP in host cells and to study the kinetics of secretion of HBsAg VLP.

[042] The term "peptide" is generally understood in the art to refer to a small amino acid molecule, whereas the term "polypeptide" is generally understood to refer to a larger amino acid molecule. Both peptides and polypeptides are within the scope of this invention. Thus, for example, the foreign sequence can be either a peptide or a polypeptide. The terms are used interchangeably herein.

[043] The HBsAg proteins and fusion polypeptides can assemble with host cell derived lipids into multimeric particles that are highly immunogenic in comparatively low concentrations. T he HBsAg fusion proteins containing the foreign sequence are exposed on the surface of the virus-like particles. The resulting virus-like particles provide excellent configurational mimics for protective epitopes as they exist in pathogens, such as an infectious virus. For these reasons, the virus-like particles are suitable for exploitation as carriers for foreign peptides or polypeptides, such as protective determinants of etiologic agents. The foreign peptides and polypeptides are comprised of sequences other than HBsAg sequences. These highly immunogenic virus-like particles display epitopes of the foreign peptides or polypeptides while retaining the protective response to HBsAg determinants.

[044] Fusion proteins containing a foreign peptide or polypeptide and a very small part of preS2 region and HBsAg protein are alternatively referred to herein as HBsAg fusion protein or the recombinant HBsAg fusion protein.

[045] The HBsAg virus-like particles thus comprise a mixture of the HBsAg proteins and fusion polypeptides comprising the foreign peptide inserted in the preS2 part of M protein. In one embodiment of the invention, the foreign peptide or polypeptide is a peptide or polypeptide other than a peptide or polypeptide from HBsAg.

[046] The foreign peptide or polypeptide can contain from 8-11 to 138-140 amino acid residues, preferably from about 20-26 to about 138-140 amino acid residues, especially from about 63-64 to about 138-140 amino acid residues. Preferably, the foreign peptide or polypeptide inserted in the preS2 region is free of cysteine residues and contains 0 to 1 methionine residues apart from the methionine required for initiation of preS2 translation. Flanking residues on either the N-terminal, C-terminal, or both N- and C-terminal ends may be added to the foreign peptide or polypeptide to generate the virus-like particles. In one aspect, the invention provides virus-like particles comprising epitope-bearing portions of foreign peptides or polypeptides. These epitopes are immunogenic or antigenic epitopes of the foreign peptides or polypeptides. An "immunogenic epitope" is defined as a part of a protein that elicits a humoral or cellular response *in vivo* when the whole polypeptide or fragment thereof, is the immunogen. A region of a polypeptide to which an antibody can bind is defined as an "antigenic determinant" or "antigenic epitope." Included in the present invention are VLPs containing both immunogenic epitopes and antigenic epitopes. Foreign peptides or polypeptides comprising immunogenic or antigenic epitopes are at least 7 amino acids residues in length.

[047] The foreign peptide or polypeptide can also be derived from any number of foreign proteins, i.e. proteins other than the envelope proteins of HBV. The foreign peptide or polypeptide can be derived from any protein of any plant, animal, bacterial, viral or parasitic organism. In one embodiment the foreign peptide or polypeptide can be derived from a polypeptide of a pathogen. The term “pathogen” as used herein, means a specific causative agent of disease, and may include, for example, any bacteria, virus, or parasite. The term “disease” as used herein, means an interruption, cessation, or disorder of body function, system, or organ. Typical diseases include infectious diseases. For example, the foreign peptide or polypeptide can be from the immunogenic proteins of an RNA virus, such as HIV-1, HIV-2, SIV, HCV, Ebola virus, Marbourg virus, HTLV-I, and HTLV-II. Specific examples are the structural or NS1 proteins of Dengue virus; the G1, G2, or N proteins of Hantaan virus; the HA proteins of Influenza A virus; the Env proteins of Friend murine leukemia virus; the Env proteins of HTLV-1 virus; the preM, E, NS1, or NS2A proteins of Japanese encephalitis virus; the N

or G proteins of Lassa virus; the G or NP proteins of lymphocytic choriomeningitis virus; the HA or F proteins of measles virus; the F or HN proteins of parainfluenza 3 virus; the F or HN proteins of parainfluenza SV5 virus; the G proteins of Rabies virus; the F or G proteins of respiratory syncytial virus; the HA or F proteins of Rinderpest; or the G proteins of vesicular stomatitis virus. These are just some of the possibilities and do not represent an exhaustive or restricted list.

[048] The foreign peptide or polypeptide can also be derived from the immunogenic proteins of a DNA virus, such as gp89 of cytomegalovirus; gp340 of Epstein-Barr; gp13 or 14 of equine herpesvirus; gB of herpes simplex 1; gD of Herpes simplex 1; gD of herpes simplex 2; or gp50 of pseudorabies. These are just some of the possibilities and do not represent an exhaustive or restricted list.

[049] Further, the foreign peptide or polypeptide can be derived from the immunogenic proteins of bacteria, such as Streptococci A M6 antigens, or tumor antigens, such as human melanoma p97, rat *Neu* oncogene p185, human epithelial tumor ETA, or human papillomavirus antigens. These are just some of the possibilities and do not represent an exhaustive or restricted list.

[050] In a preferred embodiment of this invention, the foreign peptide or polypeptide is derived from a human immunodeficiency virus. Following are HIV-1 epitopes that can be employed in designing the foreign peptide or polypeptide.

GAG	P17	(77-85)	SLYNTVATL (S9L)
	P24	(19-27)	TLNAWVKW (T9V)
POL		(79-88)	LLDTGADDTV (L10V)
		(263-273)	VLDVGDAYFSV (V11V)
		(334-342)	VIYQYMDDL (V9L)
		(464-472)	ILKEPVHGV (I9V)
		(576-584)	PLVKLWYQL (P9L)
		(669-679)	ESELVNQIIEQ (E11Q)
		(671-680)	ELVNQIIEQL (E10
		(956-964)	LLWKGEGAV (L9V)
ENV	Gp41	(260-268)	RLRDLLLIV (R9V)
NEF		(188-196)	AFHHVAREL (A9L)

Numbering is based on the amino acid sequence of the HIV-1 WEAU clone 1.60 (Genbank accession no. U21135). The WEAU sequence may not be always identical to that of the reactive peptide and simply indicates its location in the viral proteins.

[051] The foreign peptide or polypeptide can comprise a multiplicity of epitopes linked to each other. It will be understood that the virus-like particles of the invention can contain multiple epitopes of one or more origins, such as epitopes from different immunogenic proteins of the same pathogen. It will also be understood that the virus-like particles can contain one or more epitopes from different pathogens. In addition, mixtures of virus-like particles having different epitopes in different particles are contemplated by this invention.

[052] Recombinant expression vectors containing a nucleic acid encoding the foreign peptide or polypeptide in VLPs can be prepared using well known methods. The expression vectors include the sequence encoding the foreign peptide or polypeptide operably linked to suitable transcriptional or translational regulatory nucleotide sequences, such as those derived from a mammalian, viral, or insect gene. A transcriptional or translational regulatory nucleotide sequence is operably linked if the nucleotide sequence controls the transcription or translation of another coding DNA sequence. Examples of regulatory sequences include transcriptional promoters, operators, or enhancers, an mRNA ribosomal binding site, and appropriate sequences which control transcription and translation initiation or termination. The ability to replicate in the desired host cells, usually conferred by an origin of replication, and a selection gene by which transformants are identified can additionally be incorporated into the expression vector.

[053] Among eukaryotic vectors for use in the preparation of vectors of the invention are pWLNEO, pSV2CAT, pOG44, pXT1, and pSG available from Stratagene; and pSVK3, pBPV, pMSG, and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

[054] Among vectors for use in the preparation of vector of the invention, non-integrative eucaryotic vectors for a transient expression are not only useful, but integrative/transformant vectors (i.e. vectors that integrate a part of their nucleic acid material in the genome of the eukaryotic host cell) for a stable expression can also be employed. Typical of these vectors are lentiviral vector

Trips, adenovirus for transformation of mammalian cells, yeast integrative vectors, and *Agrobacterium tumefaciens* Ti-based vectors for transformation of plant cells.

[055] The nucleic acid construct encoding the recombinant virus-like particles of the invention can be inserted in a variety of different types of expression vectors for a host cell. The resulting vectors are herein referred to as the recombinant expression vectors of the invention. These vectors include vectors for use in eukaryotic expression systems and preferably for mammalian expression systems, such as recombinant poxvirus expression vectors, for example, vaccinia virus, fowlpox virus, or canarypox virus; animal DNA viruses, for example, herpes simplex 1 and 2, varicella zoster, pseudorabies, human cytomegalovirus, murine cytomegalovirus, Esptein-Barr virus, Karposi's sarcoma virus, or murine herpes virus. Animal RNA viruses can also be employed as vectors for expression of the nucleic acid construct of the invention. Suitable animal RNA viruses include positive-strand RNA viruses, such as the picornaviruses, for example, poliovirus, the flaviviruses, for example, hepatitis C virus, or coronaviruses. Examples of other suitable vectors are lentiviral vectors, adenoviral vectors, and adeno-associated viral vectors. Other suitable eukaryotic vectors are expression vectors for yeast cells, expression vectors for insect cells, such as baculoviruses, or even expression vectors for plant cells chosen for example from *Agrobacterium tumefaciens* Ti-based vectors, such pBIN-Plus (43), vectors of the pCAMBIA family (http://www.patentlens.net/daisy/bios/585.html), or pBI121 (Clontech). Plasmid and phage vectors can also be employed.

[056] The expression vectors of the invention can include at least one selectable marker. Such markers include, for example, dihydrofolate reductase, G418, ampicilin or neomycin resistance for eukaryotic cell culture, kanamycin , hygromycin, or phosphinothricin for plant cells selection and culture.

[057] Any strong promoter known to those skilled in the art can be used for driving expression. Suitable promoters include adenoviral promoters, such as the adenoviral major late promoter; heterologous promoters, such as the cytomegalovirus (CMV) promoter; the respiratory syncytial virus (RSV) promoter; inducible promoters, such as the MMT promoter, the metallothionin promoter; heat shock promoters; the albumin promoter; the ApoAI promoter; human globin

promoters; viral thymidine kinase promoters, such as the Herpes Simplex thymidine kinase promoter; retroviral LTRs; the β-actin promoter; and human growth hormone promoters. The promoter also can be a native promoter from HBV. For expression in plant cells, a promoter which enables the highly-efficient expression of the chimeric HBSAg gene encoding HBsAg and HBsAg fusion proteins is assembled at the 5' end of the gene, and the promoter is preferably the doubled cauliflower mosaic virus 35S (CaMV35S) promoter (Nature 313 (6005): 810-812 (1985)); a terminator which enhances the expression of the said chimeric HBsAg gene can be assembled at the 3'end of the gene, and the terminator is preferably the CaMV 35S terminator.

[058] Suitable host cells for expression of VLP include yeast or higher eukaryotic cells. Appropriate cloning and expression vectors for use with plant, fungal, yeast, insect, and mammalian cellular hosts are described, for example, in Pouwels et al. *Cloning Vectors: A Laboratory Manual,* Elsevier, New York, (1985). Representative examples of appropriate hosts include, but are not limited to, fungal cells, such as yeast cells; insect cells, such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells, such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

[059] Introduction of the vector of the invention or the nucleic acid construct encoding the recombinant virus-like particles of the invention into the host cell may result in transient or stable expression of the nucleic acid under control of the promoter. The introduction can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986).

[060] Therefore, the invention is also concerned with recombinant cells, such as recombinant eucaryotic cells, infected, transformed, or transfected by a polynucleotide or vector of the invention for expressing the HBsAg virus-like particles.

[061] Transformation of plant cells and regeneration of plants with polynucleotides or expression vectors of the invention

Whenever a plant cell is employed, it is preferred that the polynucleotide is integrated into the nuclear genome of the plant cell to ensure its stability and passage into the germline, although transient expression can serve an important purpose, particularly when the plant under investigation is slow-growing. A polynucleotide of the invention can also in some cases be maintained outside the chromosome, such as in the mitochondrion, chloroplast or cytoplasm.

Plant tissue suitable for transformation include leaf tissue, root tissue, meristems, zygotic and somatic embryos, callus, protoplasts, tassels, pollen, embryos, anthers, and the like. The means of transformation chosen is that most suited to the tissue to be transformed.

Transient expression in plant tissue is often achieved by particle bombardment (Klein et al.,"High-Velocity Microprojectiles for Delivering Nucleic Acids Into Living Cells,"Nature 327: 70-73 (1987), which is hereby incorporated by reference in its entirety). In this method, tungsten or gold microparticles (1 to 2 um in diameter) are coated with the DNA of interest and then bombarded at the tissue using high pressure gas. In this way, it is possible to deliver foreign DNA into the nucleus and obtain a temporal expression of the gene under the current conditions of the tissue. Biologically active particles (e. g. dried bacterial cells containing the vector and heterologous DNA) can also be propelled into plant cells. Other variations of particle bombardment, now known or hereafter developed, can also be used.

An appropriate method of stably introducing the nucleic acid construct into plants is the *Agrobacterium tumefaciens* (*A. tumefaciens*) transformation technique. This method is based upon the etiologic agent of crown gall, which afflicts a wide range of dicotyledons and gymnosperms. Where the target plant host is susceptible to infection, the *A. tumefaciens* system provides high rates of transformation and predictable chromosome integration patterns.

A. tumefaciens, which normally infects a plant at wound sites, carries a large extrachromosomal element called Ti (tumor inducing) plasmid. Ti plasmids contain two regions required for tumor induction. One region is the T-DNA (transferred DNA), which is the DNA sequence that is ultimately stably transferred to plant genomic DNA. The other region is the *vir* (virulence) region, which has been implicated in the transfer mechanism. Although the *vir* region is

required for stable transformation, the *vir* region DNA is not transferred to the infected plant.

Transformation of plant cells mediated by infection with *A. tumefaciens* and subsequent transfer of the T-DNA have been well documented. Bevan et al., Int. Rev. Genet. 16: 357 (1982). The *A. tumefaciens* system is well developed and permits routine transformation of DNA into the plant genome of a variety of plant. For example, Arabidopsis thaliana, tobacco, tomato, potato, sunflower, cotton, rapeseed, potato, poplar, and soybean can be transformed with the *A. tumefaciens* system.

Preferably, where *A. tumefaciens*-mediated transformation of plants with a polynucleotide of the invention is used, flanking T-DNA border regions of *A. tumefaciens* are provided. T-DNA border regions are 23-25 base pair direct repeats involved in the transfer of T-DNA to the plant genome. The flanking T-DNA border regions bracket the T-DNA and signal the polynucleotide that is to be transferred and integrated into the plant genome. Preferably, a polynucleotide or expression vector of the invention comprises at least one T-DNA border, particularly the right T-DNA border. Optionally, a polynucleotide to be delivered to a plant genome is sandwiched between the left and right T-DNA borders. The borders may be obtained from any Ti plasmid and may be joined to an expression vector or polynucleotide by any conventional means.

Typically, a vector containing the polynucleotide to be transferred is first constructed and replicated in E. coli. This vector contains at least one right T-DNA border region, and preferably a left and right border region flanking the desired polynucleotide. A selectable marker (such as a gene encoding resistance to an antibiotic such as kanamycin hygromycin, or phosphinothricin) can also be present to permit ready selection of transformed cells. The E. coli vector is next transferred to *A. tumefaciens*, which can be accomplished via a conjugation mating system or by direct uptake. Once inside the *A. tumefaciens*, the vector containing the polynucleotide can undergo homologous recombination with a Ti plasmid of the *A. tumefaciens* to incorporate the T-DNA into a Ti plasmid. A Ti plasmid contains a set of inducible *vir* genes that effect transfer of the T-DNA to plant cells.

Alternatively, the vector comprising the polynucleotide can be subjected in trans to the *vir* genes of the Ti plasmids. In a preferred aspect, a Ti plasmid of a given strain is "disarmed", whereby the *onc* genes of the T-DNA is eliminated or suppressed to avoid formation of tumors in the transformed plant, but the *vir* genes provided in trans still effect transfer of T-DNA to the plant host. See, e. g., Hood, Transgenic Res. 2: 208-218 (1993); Simpson, Plant Mol. Biol. 6: 403-415 (1986). For example, in a binary vector system, an E. coli plasmid vector is constructed comprising a polynucleotide of interest flanked by T-DNA border regions and a selectable marker. The plasmid vector is transformed into E. coli and the transformed E. coli is then mated to *A. tumefaciens* by conjugation. The recipient *A. tumefaciens* contains a second Ti plasmid (helper Ti plasmid) that contains *vir* genes, but has been modified by removal of its T-DNA fragment. The helper Ti plasmid will supply proteins necessary for plant cell infection, but only the E. coli modified T-DNA plasmid will be transferred to the plant cell.

The *A. tumefaciens* system permits routine transformation of a variety of plant tissues. See, e. g., Chilton, Scientific American 248: 50 (1983); Gelvin, Plant Physiol. 92: 281-285 (1990); Hooykaas, Plant Mol Biol. 13: 327-336 (1992); Rogers et al., Science 227: 1229-1231 (1985).

Methods of inoculation of the plant tissue vary depending upon the plant species and the *A. tumefaciens* delivery system. A convenient approach is the leaf disc procedure which can be performed with any tissue explant that provides a good source for initiation of whole plant differentiation [45]. The addition of nurse tissue may be desirable under certain conditions. Other procedures such as in vitro transformation of regenerating protoplasts with *A. tumefaciens* may be followed to obtain transformed plant cells as well. Specifically for *Arabidopsis thaliana* species, *in planta* transformation methods were developed which avoid plant tissue culture and regeneration (Bechtold et al. C. R. Acad. Sci. paris, Life Sciences 326:1194-1199 (1993), Chang et al., Plant J. 55:551-558 (1994), Feldmann and Marks, Mol. Gen. Genet. 208:1-9 (1987), Feldmann, Methods in Arabidopsis Research (Koncz, C., Chua, N.-H. and Scell, J. eds.) Singapore World Scientific :274-289, Katavic Mol. Gen. Genet. 245:363-370 (1994), [46]). Two of these reliable methods, the "*Agrobacterium* vaccum infiltration" (Bechtold et al. C. R. Acad. Sci. paris, Life Sciences 326:1194-1199 (1993), and "floral

tissues dipping" ([46] involved the growth of Arabidopsis to flowering stage, application of *A. tumefaciens* to the whole plant or floral tissues, collection of seed a few weeks later and identification of transformed progeny by selection on media containing antibiotic or herbicide. With the *in planta* transformation procedures, most transformed progeny are genetically uniform (non-chimeric) and the somaclonal variation associated with tissue culture and regeneration is minimized. Transformed progeny are typically hemizigous for the transgene at a given locus.

Another approach to transform plant cells and plants without the use of *A. tumefaciens* plasmids is the direct gene transfer procedures (Potrykus, Bio/Technology. 8: 535-542 (1990); Smith et al. CropSci., 35: 01-309 (1995)). Direct transformation involves the uptake of exogenous genetic material into plant cells or protoplasts. Such uptake can be enhanced by use of chemical agents or electric fields. For example, a polynucleotide of the invention can be transformed into protoplasts of a plant by treatment of the protoplasts with an electric pulse in the presence of the protoplast using electroporation. For electroporation, the protoplasts are isolated and suspended in a mannitol solution. Supercoiled or circular plasmid DNA comprising a polynucleotide of the invention is added. The solution is mixed and subjected to a pulse of about 400 V/cm at room temperature for about 10 to 100 microseconds. A reversible physical breakdown of the membrane occurs such that the foreign genetic material is transferred into the protoplasts. The foreign genetic material can then be integrated into the nuclear genome. Several monocotyledon protoplasts have also been transformed by this procedure including rice and maize. Liposome fusion is also an effective method for transformation of plant cells. In this method, protoplasts are brought together with liposomes carrying a polynucleotide of the invention. As the membranes merge, the foreign gene is transferred to the protoplasts (Dehayes et al., EMBO J. 4: 2731 (1985)). Similarly, direct gene transfer using polyethylene glycol (PEG) mediated transformation has been carried out in *N. tabacum* (a dicotyledon) and *Lolium multiflorum* (a monocotyledon). Direct gene transfer is effected by the synergistic interaction between Mg^{+2}, PEG, and possibly Ca^{+2} (Negrutiu et al., Plant Mol. Biol. 8: 363 (1987)). Alternatively, exogenous DNA can be introduced into cells or protoplasts

by microinjection of a solution of plasmid DNA comprising a polynucleotide of the invention directly into the cell with a finely pulled glass needle. Direct gene transfer can also be accomplished by particle bombardment (or microparticle acceleration), which involves bombardment of plant cells by microprojectiles carrying a polynucleotide of the invention (Klein et al., Nature 327:70 (1987); Sanford, Physiol. Plant. 79:206-209 (1990)). In this procedure, chemically inert metal particles, such as tungsten or gold, are coated with a polynucleotide of the invention and accelerated toward the target plant cells. The particles penetrate the cells, carrying with them the coated polynucleotide. Microparticle acceleration has been shown to lead to both transient expression and stable expression in cells suspended in cultures, protoplasts, and immature embryos of plants, including onion, maize, soybean, and tobacco (McCabe et al., Bio/Technology. 6:923 (1988)). Additionally, DNA viruses can be used as gene vectors in plants. For example, a cauliflower mosaic virus carrying a modified bacterial methotrexate-resistance gene has been used to infect a plant. The foreign gene systematically spreads throughout the plant (Brisson et al., Nature 301:511 (1984)). The advantages of this system are the ease of infection, systemic spread within the plant, and multiple copies of the gene per cell.

Once plant cells have been transformed, there are a variety of methods for regenerating plants. The particular method of regeneration will depend on the starting plant tissue and the particular plant species to be regenerated. Many plants can be regenerated from callus tissue derived from plant explants, including, but not limited to corn, rice, barley, wheat, rye, sunflower, soybean, cotton, rapeseed, and tobacco. Regeneration of plants from tissue transformed with *A. tumefaciens* has been demonstrated in plants including, but not limited to sunflower, tomato, white clover, rapeseed, cotton, tobacco, potato, maize, rice, and numerous vegetable crops. Plant regeneration from protoplasts is a particularly useful technique and has been demonstrated in plants including, but not limited to tobacco, potato, poplar, corn, and soybean (Evans et al., Handbook of Plant Cell Culture 1,124 (1983)).

Preferably, transformed cells are first identified using a selection marker simultaneously introduced into the host cells along with the nucleic acid construct of the present invention. Suitable selection markers include, without limitation,

markers encoding for antibiotic resistance, such as the nptII gene which confers kanamycin resistance (Fraley et al., Proc Natl Acad Sci USA 80: 4803- 4807 (1983)), and the genes which confer resistance to gentamycin, G418, hygromycin, streptomycin, spectinomycin, tetracycline, chloramphenicol, and the like. Cells or tissues are grown on a selection medium containing the appropriate antibiotic, whereby generally only those transformants expressing the antibiotic resistance marker continue to grow. Other types of markers are also suitable for inclusion in the expression cassette of the present invention. For example, a gene encoding for herbicide tolerance, such as tolerance to sulfonylurea is useful, or the dhfr gene, which confers resistance to methotrexate (Bourouis et al.,EMBO J 2 : 1099-1104 (1983)). Similarly,"reporter genes" which encode for enzymes providing for production of an identifiable compound are suitable. The most widely used reporter gene for gene fusion experiments has been uidA, a gene from *Escherichia coli* that encodes the ß-glucuronidase protein, also known as GUS. Jefferson et al.,"GUS Fusions: ß Glucuronidase as a Sensitive and Versatile Gene Fusion Marker in Higher Plants," EMBO J 6: 3901-3907 (1987)). Similarly, enzymes providing for production of a compound identifiable by luminescence, such as luciferase, are useful. The selection marker employed will depend on the target species; for certain target species, different antibiotics, herbicide, or biosynthesis selection markers are preferred. Plant cells and tissues selected by means of an inhibitory agent or other selection marker are then tested for the acquisition of the recombinant VLP-coding gene of the present invention by Southern blot hybridization analysis, using a probe specific to the genes contained in the given cassette used for transformation (Sambrook et al., "Molecular Cloning : A Laboratory Manual," Cold Spring Harbor, New York: Cold Spring Harbor Press (1989)). After the recombinant VLP-encoding gene of the present invention is stably incorporated in transgenic plants, the transgene can be transferred to other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed. Once transgenic plants of this type are produced, the plants themselves can be cultivated in accordance with conventional procedure so that the nucleic acid construct is present in the resulting plants. Alternatively, transgenic seeds are

recovered from the transgenic plants. These seeds can then be planted in the soil and cultivated using conventional procedures to produce transgenic plants. The invention includes whole plants, plant cells, plant organs, plant tissues, plant seeds, protoplasts, callus, cell culture and any group of plant cells organized into structural and/or functional units capable of expressing recombinant VLP of the invention.

[062] Methods for using the recombinant host cells in the production of proteins or peptides are well known in the art. The virus-like particles can be recovered and purified from recombinant host cell cultures by well-known methods, including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography, and lectin chromatography.

[063] In another aspect, the invention is directed to a method for producing, *in vitro* HBsAg virus-like particles, comprising : culturing *in vitro*, in a suitable culture medium, a cell incorporating an expression vector of the invention or a polynucleotide of the invention, and collecting in the culture medium HBsAg virus-like particles produced by these cells. The virus-like particles are released from the host cell into the extracellular space.

[064] According to the invention, an *in vivo* method for producing HBsAg virus-like particles involves providing a transgenic plant or plant seed transformed with a recombinant VLP-encoding polynucleotide, and growing the transgenic plant or a transgenic plant from a plant seed of the transgenic plant under conditions effective to produce the recombinant HBsAg VLP. Extracts of a transgenic plant tissue can be assayed for expression of recombinant HBsAg VLP by ELISA-type immunoassay. Recombinant HBsAg VLP can also be purified from any tissue (for example leaf) of a transgenic plant transformed with the recombinant VLP-encoding polynucleotide of the invention by any extraction protocol well-known by the skilled in the art (Huang et al. Vaccine 23:1851-1858 (2005)).

[065] Another aspect of the present invention is a method of immunizing a subject against disease resulting from infection by a pathogen. This method involves administering the plant of the present invention producing recombinant

HBsAg VLP harbouring an antigen or an epitope from the pathogen or a component part of the plant to the subject under conditions effective to immunize the subject against disease. Administering is desirably carried out by feeding the transgenic plant, or a component part. Pre ferably, the transgenic plant or a component thereof is administered as a crude extract, a freeze-dried extract, or an intact part, as a fruit, of the plant. An immunization schedule may consist in one feeding per week during 4 weeks, each feeding comprising 0.05 to 0.1 gr of transgenic plant crude extract prepared as described in Example C below.

[066] The ability of an immunogen to establish a memory response is a key element in the design of an efficacious vaccine. Responses to oral boosting observed in mice (describing in the Examples below) indicate the generation of antigen-specific memory cells. Thus, following an initial feeding administration, a vaccine "booster" may be administrated by further feeding of recombinant HBsAg VLP-producing plant of the present invention or a component thereof.

[067] Another aspect of the invention is an oral vaccine for immunization of a subject against infection by HBV or/and another pathogen. This vaccine is a component of the transgenic plant according to the present, where the plant produces recombinant HBsAg VLP harbouring an antigen or an epitope from the pathogen for immunization of a subject, and a pharmaceutical adjuvant. This transgenic plant is prepared as described above, such that the nucleic acid molecule coding for recombinant HBSAg is operably linked to a promoter and a terminator suitable for expression in plants is inserted in the plant. When a constitutive promoter is used, the recombinant HBsAg VLP will be expressed through the plants tissues. Therefore, vaccine of the present invention can be made from any plant component, including but not limited to, leaf tissue of the plant. Alternatively, the promoter may be an inducible or tissue specific promoter if desired. To obtain the immunogenic recombinant HBsAg from the plant component , the desired plant tissue of the plant can be subjected to a lysing process, or a drying process, for example freeze-drying or air-drying following procedures known in the art or as described in the Examples below. The vaccine may additionally include any suitable adjuvant to enhance the immunogenicity of the vaccine when administered to a patient.

[068] While this invention relates to HBsAg virus-like particles carrying one or more polyepitopes of foreign peptides or polypeptides on their surfaces, this invention contemplates the use of polyepitopes that have been optimized for incorporation in virus-like particles. The polyepitope nucleic and amino acid sequences can be optimized in view of increasing the overall hydrophilicity of polyepitope and ensuring an optimal processing of epitopes. Epitopes in the polyepitope can be permutated in order to obtain the best hydrophilic profile. Hydrophilic spacers can be added to counterbalance the generally hydrophobic class I epitopes. Epitopes bearing cysteine residues can be eliminated, the number of internal methionine residues can be limited to a minimum, and optionally *Homo sapiens* codon usage can be adopted. Epitopes can be positioned in head-tot-tail and an arginine residue can be inserted at the epitope C1 terminal position. As an example, the surface antigen (HBsAg) of the Hepatitis B virus (HBV) carries all the information required for membrane translocation, particle assembly, and secretion from mammalian cells. HBsAg assembles into VLPs polymeric structure that enhances antigenic stability. It is only if assembled in VLPs that HBsAg can be secreted out of cells. In this system, secretion provides high-density HBsAg presentation to antigen presenting cells (APCs). It is preferable to utilize criteria for optimizing the polyepitope sequence, which ensure the conservation of recombinant virus-like particle structures and secretion, once the particle is used as carrier of a polyepitope. Procedures for optimizing epitopes for incorporation in VLPs are described in detail in the U.S. Provisional application filed concurrently herewith by the same inventors and entitled RECOMBINANT HBsAg VIRUS-LIKE PARTICLES CONTAINING POLYEPITOPES OF INTEREST, THEIR PRODUCTION AND USE, attorney docket No. 03495.6115, the entire disclosure of which is relied upon and incorporated by reference herein.

[069] Several polyepitopic sequences of HIV-1 were prepared for incorporation in the vectors of the invention and HBsAg virus-like particles were produced and assayed for activity. One such polyepitopic sequence was derived from epitopes of HIV-1 and has been designated polHIV-1.opt. Following is the polHIV-1.opt polyepitope amino acid sequence:

polHIV-1.opt

Nucleic acid sequence

CTACTTGAAAGAGCCAGTTCATGGGGTGAGAGCCAAGACCTACCTGAATGC
ATGGGTGAAAGTTGTCAGAGACACCGCAGTGCTGGATGTGGGGGATGCCT
ACTTCTCAGTGAGAGCTAAGACTTATCTGGTCAAACTCTGGTACCAGTTGAG
GGCTGACACTCGTCTTTACAACACTGTGGCCACCCTTAGGACCAAGGCTCT
TCTGGACACTGGAGCAGATGACACTGTGAGGGCTAAGACCCTGCTGTGGAA
GGGAGAGGGAGCAGTTAGGACTGATGCTTACATCTACCAGTATATGGATGA
CCTTAGA

Amino acid sequence

YLKEPVHGVRAKTYLNAWVKVVRDTAVLDVGDAYFSVRAKTYLVKLWYQLRADTRLYN
TVATLRTKALLDTGADDTVRAKTLLWKGEGAVRTDAYIYQYMDDLR.

[070] The polHIV-1.opt polyepitope was cloned through the EcoRI and XhoI restriction sites into the pGA1xFlag-M and pGA3xFlag-M plasmids to obtain plasmid constructions designated pGA1xFlag-Mpol.opt and pGA3xFlag-Mpol.opt, respectively. Figure 4 is a detailed nucleic acid sequence for the GA1xFlag-Mpol.opt insert in pGA1xFlag-M to produce pGA1xFlag-Mpol.opt. Figure 5 is a detailed nucleic acid sequence for the GA3xFlag-Mpol.opt insert in pGA3xFlag-M to produce pGA3xFlag-Mpol.opt. Figs. 6 and 7 show nucleic acid sequences of resulting pGA1xFlag-Mpol.opt and pGA3xFlag-Mpol.opt, respectively. In FigureFigures 4 and 5, preS2 and HBsAg ATG codons are highlighted in bold and in FigureFigures 6 and 7, the polHIV-1.opt polyepitope is highlighted in bold. In Figs. 3, 4, 5 and 6 the EcoRI and XhoI restriction sites are in lower case.

[071] The secretion kinetics corresponding to pGA1xFlag-M, pGA1xFlag-Mpol.opt, pGA3xFlag-M and pGA3xFlag-Mpol.opt are shown in Figure 8.

[072] Similarly, four additional polyepitopic sequences were designed. These polyepitopic sequences have been designated pol1A2, pol2A2, pol1B7, and pol2B7. The nucleic acid and amino acid sequences, as well as epitope name and epitope sequences, are as follows.

Nucleic sequence of pol1A2

GTGCTGGATGTGGGAGATGCCTACTTCTCAGTGAGAGCTGACACCTACCTGAATGCCTGG
GTGAAGGTGGTCAGAGCCAAGACCTACCTGGTGAAGCTGTGGTACCAGCTGAGGACAGAT
GCCTCCCTGGTGAAGCATCACATGTATGTGAGAGACACAGCCTACATCTACCAGTACATG
GATGACCTGAGA

Amino acid sequence of pol1A2

VLDVGDAYFSVRADTYLNAWVKVVRAKTYLVKLWYQLRTDASLVKHHMYVRDTAYIYQYM
DDLR

Name	aa seqnuc seq
V11V	VLDVGDAYFSVGTGCTGGATGTGGGAGATGCCTACTTCTCAGTG
Y/T9V	YLNAWVKVVTACCTGAATGCCTGGGTGAAGGTGGTC
Y/P9L	YLVKLWYQLTACCTGGTGAAGCTGTGGTACCAGCTG
Vif23	SLVKHHMYVTCCCTGGTGAAGCATCACATGTATGTG
Y/V9L	YIYQYMDDLTACATCTACCAGTACATGGATGACCTG

Nucleic sequence of pol2A2

CTGCTTGACACAGGAGCTGATGACACAGTGAGGACAGATGCCAGCCTGTATAACACAGTG
GCCACCCTGAGAGCTGACACCTACCTGAAGGAGCCTGTGCATGGAGTGAGAGCTAAGACC
CTCCTGTGGAAGGGAGAGGGAGCAGTGAGAACCAAGGCAGTGCTGGCTGAGGCCATGTCC
CAGGTGAGA

Amino acid sequence of pol2A2

LLDTGADDTVRTDASLYNTVATLRADTYLKEPVHGVRAKTLLWKGEGAVRTKAVLAEAMS
QVR

Name	aa seqnuc seq
L10V	LLDTGADDTVCTGCTTGACACAGGAGCTGATGACACAGTG
S9L	SLYNTVATLAGCCTGTATAACACAGTGGCCACCCTG
Y/I9V	YLKEPVHGVTACCTGAAGGAGCCTGTGCATGGAGTG
L9V	LLWKGEGAVCTCCTGTGGAAGGGAGAGGGAGCAGTG
Gag 362	VLAEAMSQVGTGCTGGCTGAGGCCATGTCCCAGGTG

Nucleic sequence of pol1B7

TCCCCTAGGACCCTGAATGCCTGGGTGAGAGCTAAGACCAGACCTAACAATAACACAAGG
AAGTCCATCAGAGACACAGCCTTCCCTGTGAGACCACAGGTGCCTCTGAGGAGAACCAAG
GCCCACCCTGTGCATGCTGGCCCTATTGCCAGAGCTGATACAGCACCCACTAAGGCCAAA
AGGAGAGTGGTCAGG

Amino acid sequence of pol1B7

SPRTLNAWVRAKTRPNNNTRKSIRDTAFPVRPQVPLRRTKAHPVHAGPIARADTAPTKAK
RRVVR

Name	aa seqnuc seq
S9WV	SPRTLNAWVTCCCCTAGGACCCTGAATGCCTGGGTG

R10SI RPNNNTRKSIAGACCTAACAATAACACAAGGAAGTCCATC

F10LR FPVRPQVPLRTTCCCTGTGAGACCACAGGTGCCTCTGAGG

Gag237 HPVHAGPIACACCCTGTGCATGCTGGCCCTATTGCC

A10VV APTKAKRRVVGCACCCACTAAGGCCAAAAGGAGAGTGGTC

Nucleic sequence of pol2B7

AAGCCTGTGGTCTCCACACAGCTGCTTCTCAGGGCCAAGACCTTCCCTGTGAGACCCCAA GTGCCACTGAGAAGGGCTGATACACAGCCCAGGAGTGACACCCATGTGTTCAGAACCAAG GCCATTCCTAGGAGAATTAGGCAGGGCCTGAGAGATACAGCTACACCTCAGGACCTGAAC ACCATGCTGAGA

Amino acid sequences of pol2B7

KPVVSTQLLLRAKTFPVRPQVPLRRADTQPRSDTHVFRTKAIPRRIRQGLRDTATPQDLN TMLR

Name aa seqnuc seq

K10LL KPVVSTQLLLAAGCCTGTGGTCTCCACACAGCTGCTTCTC

F10LR FPVRPQVPLRTTCCCTGTGAGACCCCAAGTGCCACTGAGA

Q9VF QPRSDTHVFCAGCCCAGGAGTGACACCCATGTGTTC

I9GL IPRRIRQGLATTCCTAGGAGAATTAGGCAGGGCCTG

T9ML TPQDLNTMLACACCTCAGGACCTGAACACCATGCTG

[073] Each of the polyepitopes pol1A2, pol2A2, pol1B7, and pol2B7 was similarly inserted into pGA1xFlag-M and pGA3xFlag-M plasmids. A detailed nucleic acid sequence for each of the resulting constructs is shown in Figs. 10 to 17. The polyepitopic sequence inserted in the plasmid is shown in bold in each Figure.

[074] The recombinant HBsAg VLPs secretion kinetics corresponding to pGA1xFlag-M.pol1A2, pGA1xFlag-M.pol2A2, pGA3xFlag-M.pol1A2, pGA3xFlag-M.pol2A2, pGA1xFlag-M.pol1B7, pGA1xFlag-M.pol2B7, pGA3xFlag-M.pol1B7, and pGA3xFlag-M.pol2B7 transfections are shown in Figs. 18 to 21. All constructions give rise to VLPs secretion from transfected cells. Th e lowest values are obtained by pol1B7 and pol2B7 bearing constructions. This is due to the fact that HLA-B7 restricted epitopes are more hydrophobic peptides, when compared to HLA-A2 restricted ones.

[075] All *in vitro* analyses employed the pCMV-S2.S positive control plasmid, which expresses the wild type preS2-HBsAg fusion protein (23).

[076] This invention will now be further described in the following Examples.

[077] EXAMPLE A : Production of recombinant HIV-1/HBV virus-like particles in human cell line .

EXAMPLE A1 : Expression vectors and constructions

[078] The plasmid referred to as pCMV-basic (Figure 1A and B) is a derivative of the pCMV-B10 construction. It was made with a small polylinker (NheI, EcoRV, SmaI) replacing the pCMV-B10 polylinker between the EcoRI and XhoI restriction sites.

[079] The pGA1xFlag-M and pGA3xFlag-M plasmids are based on the expression vector pCMV-basic plasmid backbone (Figure 1) and obtained by substitution of the nucleic acids sequence in between the HindIII and the AvrII restriction sites by the insertion of the GA1xFlag-M and GA3xFlag-M motifs. In the GA1xFlag-M and GA3xFlag-M motifs, starting from the HindIII restriction site in 5', were introduced the QA peptide corresponding to the C-terminal sequence of the HBV preS1 region, the MQWNSTP peptide corresponding to the N-terminal portion of the HBV preS2 protein. This latter introduced into the construction the glycosylation site (N_4^*), which is highly conserved among primate HBV isolates. The preS2 peptide is followed by the GA motif (Gly-Ala amino acids repeated three time), here located to prevent steric impairment for the binding of anti-Flag antibodies to the tag (1xFlag-M or 3xFlag-M) by sugar molecules covalently linked to preS2 N_4*. The GA motif is followed by the 1xFlag-M or 3xFlag-M tags (modified from SIGMA-ALDRICH), giving the pGA1xFlag-M and pGA3xFlag-M, respectively (FigureFigures 2 and 3). Nucleic sequences of the original SIGMA-ALDRICH 1xFlag and 3xFlag have been modified preserving amino acids sequences but eliminating methionine residues in secondary and tertiary phases. In between the EcoRI and XhoI restriction sites, a new polylinker has been inserted, where the PstI, NheI, BglII restriction sites follow one the others. In the new constructions, the preS2 C-terminal peptide preceding the HBsAg methionine start codon (M_1) has been reduced to two amino acids (Leu-Asn), preserving the nucleic acid context, hence the strength, of the HBsAg ATG codon. Finally several N-terminal amino acid residues of HBsAg (ENITSG) were introduced.

[080] The two constructions have been obtained by "atypical PCR" using long primers (60-80 nucleic acids) detailed in the Table I (nucleic acids in HindIII and AvrII restriction sites are highlighted in small case).

Table 1 Oligonucleotides used for GA1XFlag-M and GA3XFlag-M motifs construction

Oligonucleotide	Sequence
1Xflag1	5'-CAGGCCATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTGATTACAAGGACG-3'
1Xflag2	5'-TGGAGCAGGAGCTGATTACAAGGACGACGACGACAAGgaattcCTGCAGGCTAGCAGATCTctcg-3'
1Xflag3	5'-tcCTGCAGGCTAGCAGATCTctcgagCTGAACATGGAGAACATCACATCAGGATTcctag-3'
3Xflag-M.1	5'-CAGGCCATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTGACTACAAAG-3'
3Xflag-M.2	5'-TGGAGCAGGAGCTGACTACAAAGACCACGACGGTGATTATAAAGATCACGACATTGATTACAAG-3'
3Xflag-M.3	5'-TGATTATAAAGATCACGACATTGATTACAAGGACGACGACAAGgaattcCTGCAGGCTAGCAGATCTctcg-3'
3Xflag-M.4	5'-TGCAGGCTAGCAGATCTctcgagCTGAACATGGAGAACATCACATCAGGATTcctag-3'
5' flag	5'-AGACCCaagcttCAGGCCATGCAGTGGAACTCCACA-3'
3' flag	5'-AGGGGTcctaggAATCGTGATGTGATGTTCTCCATG-3'

[081] Two separate reactions (A and B), for GA1xFlag-M and GA3xFlag-M motifs respectively, were performed using 50 pmols of 1xFlag1, -2 and -3, in reaction A, and 3xFlag-M .1, .2, .3, and .4 in B. In both reactions A and B, 10 pmols 3' flag were added. Six cycles of PCR were then performed.

[082] Then, 25 cycles of classical PCR were performed, adding 100 pmols of the 5' flag and the 3' flag primers.

[083] To obtain the pGA1xFlag-M and pGA3xFlag-M plasmids, the reaction A and B PCR fragments were then digested by HindIII and AvrII to be cloned into the corresponding cloning sites of the pCMV-basic plasmid.

[084] To obtain the pGA1xFlag-Mpol.opt and pGA3xFlag-Mpol.opt plasmids (FigureFigures 4, 5, 6 and 7), the polHIV-1.opt polyepitope was cloned between the EcoRI and XhoI restriction sites of the pGA1xFlag-M and pGA3xFlag-M plasmids, respectively. Codon usage was optimized according to the *Homo sapiens* table (http://www.kazusa.or.ip/codon). Hydrophathy profiles were obtained by DNA StriderTM 1.2 (Kyte-Doolittle option).

[085] The polHIV-1.opt polyepitope was assembled by multiple rounds of "atypical PCR". Briefly, a series of six 70-80-mer oligonucleotides were synthesized corresponding to the plus strand and overlapped one another by ~20 bases at both 5' and 3' ends (the oligonucleotides used in this invention are shown in Table 2).

Table 2 Oligonucleotides used for polHIV-1.opt polyepitope construction

Oligo-nucleotide	Sequence
1-HIVPOLY	5'-GAATTCCTACTTGAAAGAGCCAGTTCATGGGGTGAGAGCCAAGACCTACCTGAATGCATGGGTGAAAGTTG-3'
2-HIVPOLY	5'-CTGAATGCATGGGTGAAAGTTGTCAGAGACACCGCAGTGCTGGATGTGGGGGATGCCTACTTCTCAGTGAGAG-3'
3-HIVPOLY	5'-ATGCCTACTTGTCAGTGAGAGCTAAGACTTATCTGGTCAAACTCTGGTACCAGTTGAGGGCTGACACTCG-3'
4-HIVPOLY	5'-CAGTTGAGGGCTGACACTCGTCTTTACAACACTGTGGCCACCCTTAGGACCAAGGCTCTTCTGGACACTGGAGCAGATG-3'
5-HIVPOLY	5'-CTTCTGGACACTGGAGCAGATGACACTGTGAGGGCTAAGACCCTGCTGTGGAAGGGAGAGGGAGCAGTTAGGACTG-3'
6-HIVPOLY	5'-AAGGGAGAGGGAGCAGTTAGGACTGATGCTTACATCTACCAGTATATGGATGACCTTAGACTCGAG-3'
5'conPmodifpc	5'-CATGAACTGGCTCTTTCAAGTAGGAATTCCACTG-3'
5'modifPoly	5'-GCAGTGGAATTCCTACTTGAAAGAGCCAGTTCATG-3'
3'modifPoly	5'-CTATATGCTCGAGTCTAAGGTCATCCATATACTG-3'

[086] Two separate reactions (A and B) were performed using 50 pmols of HIVPOLY-1, -2 and -3, in reaction A, and HIVPOLY-4, -5 and -6 in B, respectively. Then, 25 pmols of 5'conpmodifpc and 3'modifpoly were added in reactions A and B, respectively. Fifteen cycles of PCR were then performed.

[087] PCR products from reactions A and B were assembled as follows: 0.5 µl of each reaction were put in 20 µl of H_2O at 95°C for 30 seconds and then to room temperature (r.t.). Five units of Klenow fragment and 1 µl of dNTPs (40 mM) were added and reaction performed for 15 minutes at 37°C.

[088] Then, 25 cycles of classical PCR were performed, adding 100 pmols of the 5'modifpoly and the 3'modifpoly primers.

EXAMPLE A2: In vitro evaluation of VLPs secretion by the plasmids bearing the polHIV-1.opt polyepitope

[089] The pGA1xFlag-M, pGA3xFlag-M, pGA1xFlag-Mpol.opt and pGA3xFlag-Mpol.opt plasmids were transiently transfected into SW480 cells, along with pCMV-S2.S as positive control for HBsAg VLPs formation and secretion (Figure 8). The pCMV-S2.S plasmid expresses the wild type pres2-HBsAg fusion protein (23).

[090] More particularly, the SW480 human cell line was maintained in Dulbecco medium supplemented with 5% foetal calf serum (FCS) and 1% streptomycin and penicillin, according to recommendations of the manufacturer. The pCMV-S2.S plasmid was kindly provided by Dr. Marie-Louise Michel (23).

[091] Cells were transiently transfected by FuGENE6™ transfection reagent (Roche). Out of 2 ml, 500 µl of supernatant were collected and renewed at each time point. HBsAg concentration in supernatants was estimated by the Monolisa® Ag HBsAg Plus Kit (BIORAD). To detect fusion protein on HBsAg

VLPs, an anti-Flag-M ELISA was performed using the M2 monoclonal antibody (SIGMA-ALDRICH), which recognizes a linear epitope. Briefly, 96 well plates were coated with M2 mAb, and 2.5 ng/ml of HBsAg positive samples tested per well.

[092] The anti-HBsAg ELISA test used allows detection and quantification of HBsAg antigenic units only if the protein is assembled into VLPs. The pGA1xFlag-M plasmid resulted in VLPs secretion, which was only ~2-3 fold down from the pCMV-S2.S (Figure 8). The pGA1xFlag-Mpol.opt plasmid resulted in VLPs secretion ~20-35 fold down from the pGA1xFlag-M. These data clearly show a gradual impact of HBsAg fusion protein complexity on the inhibition of recombinant VLPs assembly. This impact is more drastic in the case of the pGA3xFlag-M, which resulted in VLPs secretion ~60-70 fold down from the pCMV-S2.S. Nevertheless, over a 14 days period, recombinant HBsAg VLPs could be detected at comparable levels in pGA1xFlag-Mpol.opt and pGA3xFlag-Mpol.opt samples. This suggests that once the polHIV-1.opt polyepitope is inserted into the pGA1xFlag-M or the pGA3xFlag-M plasmid, the inhibition by the GA1xFlag-M and GA3xFlag-M motifs on VLPs assembly is eliminated and replaced by the impact of polyepitope next to the N-terminal ATG of the HBsAg, here of the polHIV-1.opt polyepitope.

[093] To verify that the pGA1xFlag-M, pGA3xFlag-M, pGA1xFlag-Mpol.opt and pGA3xFlag-Mpol.opt plasmids could give rise to recombinant VLPs bearing fusion proteins, an anti-Flag-M ELISA assay was performed. (Figure 9.) The Flag-M is a linear epitope present in N-terminal to fusion proteins, hence N-terminal to the polHIV-1.opt polyepitope and the HBsAg genes, in the order. The Flag-M ELISA was performed on the equivalent of 2.5 and 5 ng HBsAg/ml of supernatants determined by the Monolisa® Ag HBsAg Plus Kit (BIORAD). The Flag-M ELISA was made by coating over night 96 well plates with 200µl of 4 µg/ml of M2 monoclonal anti-Flag antibody (SIGMA-ALDRICH). Supernatants are incubated for 2 hours at 37°C and then revealed by the R6-R7-R8-R9 reagents from the Monolisa® Ag HBsAg Plus Kit (BIORAD). By this procedure, recombinant VLPs bearing fusion proteins are trapped in ELISA plate wells by the anti-Flag antibody and then identified by polyclonal anti-HBsAg antibodies recognizing conformational epitopes in the VLP structure. As negative samples,

N-terminal Flag-BAP™ and 3xFlag-BAP™ control proteins (SIGMA-ALDRICH) and supernatants from pCMV.S2.S transfected cultures are used. Wells were revealed by alkaline phosphatase reaction and read at 450 nm and 620 nm. Limit of detection corresponds to 0.05 OD at 450 nm.

[094] Results showed that the all recombinant VLPs quantified by anti-HBsAg ELISA bear on their surfaces fusion proteins, which are highlighted by the presence of the Flag-M tag (Figure 9 A and B). This semi-quantitative test allows comparison of GA1xFlag-M (Figure 9A) or GA3xFlag-M (Figure 9B) samples as far as their fusion protein content with respect to a given HBsAg input. This test does not allow direct comparison of GA1xFlag-M samples to GA3xFlag-M ones, as M2 antibody affinity for 3xFlag-M is assumed to be higher than for 1xFlag-M. In other words, at comparable HBsAg input and OD_{450nm}. output, more fusion protein is present in the GA1xFlag-M sample when compared to an equivalent GA3xFlag-M sample.

EXAMPLE A3: Expression vectors and constructions

[095] To obtain the pGA1xFlag-Mpol1.A2, pGA1xFlag-Mpol2.A2, pGA3xFlag-Mpol1.A2 and pGA3xFlag-Mpol2.A2 plasmids (FigureFigures 10, 11, 12, and 13, respectively), the pol1.A2, pol2.A2 polyepitopes were cloned between the EcoRI and XhoI restriction sites of the pGA1xFlag-M and pGA3xFlag-M plasmids, respectively. Codon usage was optimized according to the *Homo sapiens* table (http://www.kazusa.or.ip/codon). Hydrophathy profiles were obtained by DNA StriderTM 1.2 (Kyte-Doolittle option).

[096] The pol1.A2, pol2.A2 polyepitopes were assembled by multiple rounds of "atypical PCR." Briefly, a series of four 60-80-mer oligonucleotides were synthesized corresponding to the plus strand of each polyepitope and overlapped one another by ~20 bases at both 5' and 3' ends (the oligonucleotides used in this invention are shown in Table 3.

Table 3 Oligonucleotides used for pol1.A2 and pol2.A2 polyepitopes construction

Oligonucleotide	Sequence
pol1.A2-1	5'-attcGTGCTGGATGTGGGAGATGCCTACTTCTCAGTGAGAGCTGACACCTACCTGAATGCCTGGGTGAAGGTG-3'
pol1.A2-2	5'-ACCTGAATGCCTGGGTGAAGGTGGTCAGAGCCAAGACCTACCTGGTGAAGCTGTGGTACCAGCTGAGGACAG-3'
pol1.A2-3	5'-AGCTGTGGTACCAGCTGAGGACAGATGCCTCCCTGGTGAAGCATCACATGTATGTGAGAGACACAG-3'
pol1.A2-4	5'-AGCATCACATGTATGTGAGAGACACAGCCTACATCTACCAGTACATGGATGACCTGAG-3'
5' pol1.A2	5'-GAGAATgaattcGTGCTGGATGTGGGAGATG-3'
3' pol1.A2	5'-CTATATctcgagTCTCAGGTCATCCATGTACTGGTAG-3'
pol2.A2-1	5'-attcCTGCTTGACACAGGAGCTGATGACACAGTGAGGACAGATGCCAGCCTGTATAACACAGTGGCCACCCTG-3'
pol2.A2-2	5'-AGCCTGTATAACACAGTGGCCACCCTGAGAGCTGACACCTACCTGAAGGAGCCTGTGCATGGAGTGAGAG-3'
pol2.A2-3	5'-AGCCTGTGCATGGAGTGAGAGCTAAGACCCTCCTGTGGAAGGGAGAGGGAGCAGTGAGAACCAAGGCAGTG-3'
pol2.A2-4	5'-AGCAGTGAGAACCAAGGCAGTGCTGGCTGAGGCCATGTCCCAGGTGAGActcgag-3'
5' pol2.A2	5'-GAGAATgaattcCTGCTTGACACAGGAGCTG-3'
3' pol2.A2	5'-CTAGATctcgagTCTCACCTGGGACATG-3'

[097] Two separate reactions (A and B) for pol1.A2 and pol2.A2 polyepitopes, respectively, were performed using 50 pmols of pol1.A2-1, -2, -3, -4 and 10 pmols of 3' pol1.A2 in reaction A, and 50 pmols of pol2.A2-1, -2, -3, -4 and 10 pmols of 3' pol2.A2 in B. Six cycles of PCR were then performed.

[098] Then, 25 cycles of classical PCR were performed, adding 100 pmols of the 5' pol1.A2 and 3' pol1.A2 for reactions A, and 5' pol2.A2 and 3' pol2.A2 for reaction B.

[099] To obtain the pGA1xFlag-Mpol1.B7, pGA1xFlag-Mpol2.B7, pGA3xFlag-Mpol1B7 and pGA3xFlag-Mpol2B7 plasmids (FigureFigures 14, 15, 16, and 17, respectively), the pol1.B7, pol2.B7 polyepitopes were cloned between the EcoRI and XhoI restriction sites of the pGA1xFlag-M and pGA3xFlag-M plasmids, respectively. Codon usage was optimized according to the *Homo sapiens* table (http://www kazusa.or.ip/capon). Hydrophathy profiles were obtained by DNA StriderTM 1.2 (Kyte-Doolittle option).

[0100] The pol1.B7, pol2.B7 polyepitopes were assembled by multiple rounds of "atypical PCR." Briefly, a series of four 60-80-mer oligonucleotides were synthesized corresponding to the plus strand of each polyepitope and overlapped one another by ~20 bases at both 5' and 3' ends (The oligonucleotides used in this invention are shown in Table 4).

[0101] More particularly, the pol1.B7, pol2.B7 polyepitopes were synthesized by multiple rounds of "atypical" PCR using the long primers detailed in the Table 4:

Table 4 Oligonucleotides used for pol1.B7 and pol2.B7 polyepitopes construction

Oligonucleotide	Sequence
pol1.HIVB7-1	5'-AATTCCCCTAGGACCCTGAATGCCTGGGTGAGAGCTAAGACCAGACCTAACAATAACACAAGGAAG-3'
pol1.HIVB7-2	5'-ACCAGACCTAACAATAACACAAGGAAGTCCATCAGAGACACAGCCTTCCCTGTGAGACCACAGGTGCCTCTGAG-3'
pol1.HIVB7-3	5'-AGACCACAGGTGCCTCTGAGGAGAACCAAGGCCCACCCTGTGCATGCTGGCCCTATTGCCAGAGCTG-3'
pol1.HIVB7-4	5'-ATGCTGGCCCTATTGCCAGAGCTGATACAGCACCCACTAAGGCCAAAAGGAGAGTGGTCAGGCTCGAG-3'
5' pol1.HIVB7	5'-GTATAAGaattcTCCCCTAGGACCCTGAATGCCTG-3'
3' pol1.HIVB7	5'-CTATATGCTCGAGCCTGACCACTCTCCTTTTG-3'
pol2.HIVB7-1	5'-AATTCCAAGCCTGTGGTCTCCACACAGCTGCTTCTCAGGGCCAAGACCTTCCCTGTGAGACCCCAAGTG-3'
pol2.HIVB7-2	5'-ACCTTCCCTGTGAGACCCCAAGTGCCACTGAGAAGGGCTGATACACAGCCCAGGAGTGACACCCATGTGTTCAG-3'
pol2.HIVB7-3	5'-AGGAGTGACACCCATGTGTTCAGAACCAAGGCCATTCCTAGGAGAATTAGGCAGGGCCTGAGAGATACAG -3'
pol2.HIVB7-4	5'-ATTAGGCAGGGCCTGAGAGATACAGCTACACCTCAGGACCTGAACACCATGCTGAGACTCGAG-3'
5' pol2.HIVB7	5'-GTCCTGaattcAAGCCTGTGGTCTCCACACAG-3'
3' pol2.HIVB7	5'-CTATATGCTCGAGTCTCAGCATGGTGTTCAG-3'

[0102] Two separate reactions (A and B), for pol1.B7 and pol2.B7 polyepitopes, respectively, were performed using 50 pmols of pol1.HIVB7-1, -2, -3, -4, and 10 pmols of 3'pol1.HIVB7 in reaction A, and 50 pmols of pol2.HIVB7-1, -2, -3, -4, and 10 pmols of 3' pol2.HIVB7 in B. Six cycles of PCR were then performed.

[0103] Then, 25 cycles of classical PCR were performed, adding 100 pmols of the 5' pol1.HIVB7 and 3' pol1.HIVB7 for reactions A, and 5' pol2.HIVB7 and 3' pol2.HIVB7 for reaction B.

EXAMPLE A4

In vitro evaluation of VLPs secretion by plasmids bearing the pol1.A2.pol2A2, pol1B7 and pol2.B7 polyepitopes

[0104] pGA1xFlag-Mpol1.A2, pGA1xFlag-Mpol2.A2, pGA3xFlag-Mpol1.A2, pGA3xFlag-Mpol2.A2, pGA1xFlag-Mpol1.B7, pGA1xFlag-Mpol2.B7, pGA3xFlag-Mpol1.B7 and pGA3xFlag-Mpol2.B7 plasmids were transiently transfected into SW480 cells, along with pCMV-S2.S as positive controls for HBsAg VLPs formation and secretion (FigureFigures 18, 19, 20, and 21). The pCMV-S2.S plasmid expresses the wild type preS2 HBsAg fusion proteins (23)

[0105] More particularly, the SW480 human cell line was maintained in Dulbecco medium supplemented with 5% foetal calf serum (FCS) and 1% streptomycin and penicillin, according to recommendations of the manufacturer. The pCMV-S2.S plasmid was kindly provided by Dr. Marie-Louise Michel (23).

[0106] Cells were transiently transfected by FuGENE6™ transfection reagent (Roche). Out of 2 ml, 500 µl of supernatant were collected and renewed at each

time point. HBsAg concentration in supernatants was estimated by the Monolisa® Ag HBsAg Plus Kit (BIORAD).

[0107] The ELISA test used allows detection and quantification of HBsAg antigenic units only if the protein is assembled into VLPs. pGA1xFlag-Mpol1.A2, pGA1xFlag-Mpol2.A2, pGA3xFlag-Mpol1.A2, pGA3xFlag-Mpol2.A2 plasmids resulted in VLPs secretion (FigureFigures 18 and 19) when compared to pGA1xFlag-Mpol1.B7, pGA1xFlag-Mpol2.A2B7 pGA3xFlag-Mpol1.B7, pGA3xFlag-Mpol2.B7 plasmids (FigureFigures 20 and 21). This is in keeping with the fact that HLA.A2.1 epitopes are generally more hydrophilic than the HLA.B7 ones. Moreover, the pGA1xFlag-Mpol1.A2, pGA1xFlag-Mpol2.A2, pGA3xFlag-Mpol1.A2, pGA3xFlag-Mpol2.A2 plasmids resulted in VLPs production comparable to that obtained by pGA1xFlag-Mpol.opt, pGA3xFlag-Mpol.opt (Figure 8), all these constructions sharing HLA.A2.1 restricted epitopes.

[0108] Thus, this invention provides the GA1xFlag-M nucleotide sequence (5'->3'):

CAGGCCATGCAGTGGAACTCCACAcccgggGCTGGAGCAGGAGCTGATTACA AGGACGACGACGACAAGgaattcCTGCAGGCTAGCAGATCTctcgagCTGAACAT G.

[0109] The GA1xFlag-M nucleotide sequence can be comprised of the following elements:

1) CAGGCC corresponding to preS1 C-terminal sequence of HBV strain U95551 (nucleotides 3168-3173),

2) ATGCAGTGGAACTCCACA corresponding to preS2 N-terminal sequence of HBV strain U95551 (nucleotides 3174-3182; 1-9). Note: the T nucleotide in HBV strain U95551 position 3 has been substituted by C in the present invention,

3) cccggg is the SmaI restriction site (overlapping the following motif),

4) gggGCTGGAGCAGGAGCT encoding the GAGAGA amino acid sequence of the spacer motif,

5) GATTACAAGGACGACGACGACAAG corresponding to the 1xFlag-M nucleotide sequence,

6) gaattcCTGCAGGCTAGCAGATCTctcgag corresponding to EcoRI, PstI, NheI, BglI, and XhoI polylinker,

7) CTGAAC corresponding to preS2 C-terminal sequence of HBV strain U95551 (nucleotides 151-156), and

8) ATG is the first ATG codon of any S HBV protein.

[0110] This invention also provides the GA3xFlag-M nucleotide sequence (5'->3'):

CAGGCCATGCAGTGGAACTCCACAcccgggGCTGGAGCAGGAGCTGACTACA
AAGACCACGACGGTGATTATAAAGATCACGACATTGATTACAAGGACGACG
ACGACAAGgaattcCTGCAGGCTAGCAGATCTctcgagCTGAACATG

[0111] The GA3xFlag-M nucleotide sequence can be comprised of the following elements:

1) CAGGCC corresponding to preS1 C-terminal sequence of HBV strain U95551 (nucleotides 3168-3173),

2) ATGCAGTGGAACTCCACA corresponding to preS2 N-terminal sequence of HBV strain U95561 (nucleotides 3174-3182; 1-9). Note: the T nucleotide in HBV strain U95551 position 3 has been substituted by C in the present invention,

3) cccggg corresponding to the SmaI restriction site (overlapping the following motif),

4) gggGCTGGAGCAGGAGCT encoding GAGAGA amino acid sequence of the spacer motif,

5) GACTACAAAGACCACGACGGTGATTATAAAGATCACGAC ATTGATTACAA GGACGACGACGACAAG corresponding to 3xFlag-M nucleotide sequence,

6) gaattcCTGCAGGCTAGCAGATCTctcgag corresponding to EcoRI, PstI, NheI, BglI, and XhoI polylinker,

7) CTGAAC corresponding to preS2 C-terminal sequence of HBV strain U95551 (nucleotides 151-156), and

8) ATG is the first ATG codon of any S HBV protein.

[0112] *E. coli* strains carrying the following plasmids were deposited at the Collection Nationale de Cultures de Microorganismes (C.N.C.M.), of Institut Pasteur, 25, rue du Docteur Roux, F-75724 Paris, Cedex 15, France, and assigned the following Accession Nos.:

<u>PlasmidAccession No.</u>

pGA1xFlag-MCNCM I-3543 filed on December 16, 2005

pGA1xFlag-Mpol.optCNCM I-3544 filed on December 16, 2005

pGA3xFlag-MCNCM I-3545 filed on December 16, 2005
pGA3xFlag-Mpol.optCNCM I-3546 filed on December 16, 2005
pGA1xFlag-M.pol.1A2CNCM I-3579 filed on February 28, 2006
pGA1xFlag-M.pol.2A2CNCM I-3580 filed on February 28, 2006
pGA1xFlag-M.pol.1B7CNCM I-3581 filed on February 28, 2006
pGA1xFlag-M.pol.2B7CNCM I-3582 filed on February 28, 2006.

EXAMPLE B : Production of recombinant HIV-1/HBV virus-like particles in *Nicotiana tabacum* and *Arabidopsis thaliana* plants for a bivalent plant-based vaccine

EXAMPLE B1 Materials and Methods

B.1.1. Engineering the Flag and Flag-M plasmids

[0113] The ppolHIV-1.opt plasmid, which is a derivative of the pCMV-B10 plasmid was constructed as described in example 1, where the polHIV-1.opt polyepitope has been inserted between the *Eco*RI and *Xho*I restriction sites. In the pCMV-basic control plasmid, the polHIV-1.opt polyepitope was substituted in the ppolHIV-1.opt plasmid by a small polylinker (*Nhe*I, *Eco*RV, *Sma*I) between the *Eco*RI and *Xho*I restriction sites. The pCMV-S2.S control plasmid was kindly provided by Dr Marie Louise Michel [23] and expresses the wild type preS2-HBsAg fusion protein. The pGA3xFlagbasic, pGA1xFlag-Mbasic and pGA3xFlag-Mbasic plasmids have been engineered from the pCMV-basic plasmid by replacing the nucleic acid sequence between the *Hind*III and the *Avr*II unique restriction sites localised at 7-12 nucleotides upstream the preS2 ATG codon and 21-26 nucleotides downstream the HBsAg ATG codon, respectively. The new nucleic acid inserts have been obtained by "atypical PCR".

[0114] Briefly, a series of 50-70-mer oligonucleotides was synthesised corresponding to the gene plus strand and overlapped one another by ~20 bases at both 5' and 3' ends (oligonucleotide sequences are detailed in Figure 22). Six cycles were performed using 50 pmols of oligonucleotides specific for each construction and 10 pmols of 3'flag primer for each separate reaction. Then, 100 pmols of 5'flag and 3'flag oligos were added and 25 cycles of classical PCR were performed. The pGA1xFlag-Mpol.opt and pGA3xFlag-Mpol.opt plasmids have been engineered by inserting the polHIV-1.opt polyepitope between the *Eco*RI and *Xho*I restriction sites of the pGA1xFlag-Mbasic and pGA3xFlag-Mbasic

plasmids. Hydrophathy profiles were obtained by DNA StriderTM 1.2 (Kyte-Doolittle option), [44].

B.1.2. Mammalian transient cell transfection

[0115] The SW480 human adherent cell line was maintained in D-MEM medium supplemented with 5% fetal calf serum (FCS) and 1% streptomycin and penicillin. Cells were transiently transfected by FuGENE6™ transfection reagent (Roche), according to the manufacturer's recommendations. Out of 2 ml, 500 µl of supernatant were collected and renewed at each time point. At day 14, transfected cells were trypsinated, counted, aliquotted at 3.5 10^6 cell per sample and lysed by three rounds of freezing (-70°C) and thawing (50°C).

B.1.3. Anti-HBsAg and anti-Flag-M ELISA tests

[0116] Anti-HBsAg and anti-Flag-M ELISA tests were performed either on SW480 cell culture supernatants and lysates or plant protein extracts. HBsAg concentration was estimated by the Monolisa® HBsAg Ultra Kit (BIORAD). The anti-Flag-M ELISA was performed using the M2 monoclonal antibody (M2 mAb; SIGMA Aldrich), which recognizes both the 1xFlag-M and 3xFlag-M amino acid sequences. Briefly, 96 well plates were coated over night in carbonate buffer (pH 9.6) with 200 l of M2 mAb (4µg/ml) and then washed with 250µl 1xPBS/0.1% Tween. Saturation was obtained by putting 200µl of carbonate buffer/10% fetal calf serum for 1 hour at 37°C. Samples were diluted in a total of 100µl 1xPBS, incubated for 2 hours at 37°C and then extensively washed with 1xPBS/0.1% Tween. Wells were filled with 150µl of R6-R7 1/3 diluted in 1xPBS/0.1% Tween, incubated for 2 hours at 37°C and then washed. 100µl of R8-R9 reagents were incubated for 30 minutes in the dark and then 100µl of R10 were added to stop reactions and wells were read at OD620nm. R6 to R10 reagents were from the Monolisa® HBsAg Ultra Kit (BIORAD). In the anti-Flag-M ELISA tests, the Flag-BAP and 3xFlag-BAP proteins (SIGMA Aldrich) and the pCMV-S2.S supernatant were used as negative controls. The pGA1xFlag-Mbasic supernatant was the positive control. For the pGA1xFlag-Mbasic plants, E1-A extracts could not be tested, as they were entirely used for setting the different protocols involved in plant analyses.

B.1.4. Subcloning of the Flag-M transgenes into a plant expression vector and plant transformation

[0117] The two "preS2-GAmotif-(1x or 3x) Flag-M-polHIV-1.opt-HBsAg" sequences, as well as the two control "preS2-GAmotif-(1x or 3x)Flag-M-polylinker-HBsAg" sequences, were amplified from the pGA1xFlag-Mpol.opt, pGA3xFlag-Mpol.opt, pGA1xFlag-Mbasic and pGA3xFlag-Mpol.opt, respectively, by PCR using *Pfu* Turbo DNA polymerase (Stratagene) with the primers HIV-F1 (5'-CCaagcttCAGGCCATGCAGT) and HIV-R2 (5'-ATgatatcCCCATCTCTTTGTTTTGTTAGG) and subcloned into the pGEM-T Easy Vector (Promega). All the four amplified regions span from the pre-S2 ATG start codon to the HBsAg stop codon. Each cloned sequence was then excised from the pGEM-T Easy vector as a *Hin*dIII-*Eco*RV fragment and inserted into the pAMPAT-MCS binary vector (GeneBank accession number AY436765), between the doubled cauliflower mosaic virus 35S (CaMV35S) promoter and terminator. The resulting plasmids were introduced into the *Agrobacterium tumefaciens* strain GV3101 (pMP90RK) via electroporation and utilized for plant transformation. Tobacco (*Nicotiana tabacum* cv. Samsun) was transformed by the leaf disc transformation procedure [45]. Phosphinotricin (DUCHEFA) was added at the concentration of 5mg/L during plant transformation and regeneration of plantlets.

[0118] Regenerated transgenic T0 plants were transferred to soil and grown in a fully climatized greenhouse. Arabidopsis (*Arabidopsis thaliana* cv. Columbia) transformation was performed by floral dip [46]. Seeds from T0 primary transformants were selected on MS medium [47] containing 1% (w/v) sucrose and 10mg/L phosphinotricin. T1 seedlings that survived selection were then transferred to soil in a growth chamber under standard conditions.

B.1.5. PCR and Southern analyses

[0119] Genomic DNA was isolated from leaves of tobacco and Arabidopsis plants using an urea-phenol extraction procedure, as previously described [53]. PCR analysis to confirm the presence of the transgenes in the transformed plants was performed on about 50 ng of genomic DNA, using primers annealing to the CaMV35S promoter (p35S-F1: 5'-CCACTATCCTTCGCAAGACCC) and terminator (t35S-R2: 5'-TCAACACATGAGCGAAACCC) and standard PCR conditions. To determine the number of integrated copies of the transgenes, about 8 µg of tobacco or 0.5µg of Arabidopsis genomic DNA were restricted with

*Eco*RI and subjected to Southern blot analysis as described previously [48]. A 0.59-kb fragment of the HBsAg gene was used as a probe.

B.1.6. Northern analysis

[0120] Total RNA extraction was performed on tobacco and Arabidopsis leaves using Trizol (Invitrogen) according to the manufacturer's instructions. As far as transformed T0 tobacco plants is concerned, extractions were performed on three weeks old plants following transfer from tissue culture to greenhouse (extraction time point E1) and on five months old greenhouse plants (extraction time point E2). Extraction from T1 tobacco progeny was made on three weeks old greenhouse plants. In the case of Arabidopsis, one extraction was performed on five weeks old transformed plants. For Northern blot analysis, total RNA was fractionated on a 1.5% formaldehyde agarose gel and blotted in 10X SSC onto a Hybond-N+membrane (GE Healthcare). About 10µg of total RNA from tobacco or 5µg from Arabidopsis samples were loaded on the gel. Pre-hybridisation and hybridisation were made as previously described [49]. The membranes were hybridized with the same HBsAg specific probe used for Southern blot, and then re-hybridized with a tobacco 18S rRNA specific probe for loading control. Hybridisation was quantified using a Typhoon Phosphor-Imager and ImageQuant software (GE Healthcare). Transgene mRNA expression levels were normalized by calculating for each sample the ratio between HBsAg and 18S rRNA signals.

B.1.7. Protein analysis

[0121] Plant crude extracts were obtained from tobacco or Arabidopsis leaves collected at the times specified for RNA analyses. Leaves were grinded in liquid nitrogen with the following extraction buffer (1 ml buffer / 0.35g of fresh leaves): 1xPBS pH 7.4, 10mM EDTA, 0.1% Triton X-100 and 1mM phenylmethylsulphonyl fluoride (PMSF). The homogenates were centrifuged at 10,000 rpm for 10 min at 4°C and then supernatants stored at -80°C for protein analyses. Total soluble protein (TSP) was quantified by Bradford analysis (Bio-Rad) performed in 96 wells plates, according to the manufacturer's instructions. Recombinant HIV-1/HBV VLPs production was assayed by anti-HBsAg and anti-Flag-M ELISA, as previously described.

EXAMPLE B.2. Results

B.2.1. HIV-1/HBV transgenes engineering

[0122] In a previous study, a peptide containing 8 epitopes from the HIV-1 Gag and Pol open reading frames has been engineered in N-terminal to the self-assembling HBsAg protein to obtain efficient recombinant HIV-1/HBV VLPs production [55]. The polyepitope (polHIV-1.opt) was cloned between the preS2 and HBsAg ATG codons, to mimic the wild type HBV preS2-HBsAg fusion protein (ppolHIV-1.opt plasmid; Fig. 23a). In the design of this HIV-1 polyepitope, a major effort had been made to preserve protein hydrophilicity in order to ensure recombinant HIV-1/HBV VLPs secretion from any production system. Indeed, in wild type HBV VLPs, HBsAg constitutes the backbone of the particles and the N-terminal hydrophilic preS2 peptide is exposed on VLPs surfaces. The hydrophilicity of the preS2 region has been demonstrated to be essential for VLPs production/secretion [50].

[0123] In the present work, the preS2 region surrounding the HIV polyepitope in the ppolHIV-1.opt plasmid [50] was extensively redesigned, with the aim of increasing recombinant HIV-1/HBV VLPs production and improve their detection by simpler methods (Fig. 23b). Firstly, the highly conserved preS2 N-glycosylation site (N*ST) [50] was reintroduced by adding a Threonine after the preS2 MQWNS motif. Indeed, the attachment of an oligosaccharide unit to a polypeptide at the site of N-glycosylation can enhance solubility, improves folding, facilitates secretion, modulates antigenicity and increases half-life of a glycoprotein *in vivo* [51]. Secondly, to improve detection of recombinant proteins, the HIV-1 MN V3loop tag has been replaced by the 1xFlag or 3xFlag tags (SIGMA Aldrich). Indeed, the commercial anti-Flag M2 monoclonal antibody (mAb) allows high sensitive detection in both ELISA and Western blots. Moreover, the 1xFlag and 3xFlag amino acid sequences are highly hydrophilic (Fig.23a and 23b) and do not contain cysteine residues, which are elective parameters for HBsAg N-terminal peptides to obtain efficient recombinant VLPs production [50]. In the new constructs, the Flag tags have been inserted N-terminal to the HIV-1 polyepitope to ensure detection of the polyepitopic sequence in recombinant HIV/HBV proteins. Thirdly, a motif of six amino acids (GAGAGA) has been introduced between the preS2 MQWNST motif and the tags (preceded by the P amino acid to give the *Sma*I restriction site) to preserve antibody recognition of the tags in N-glycosylated fusion proteins. Finally, the

preS2 C-terminal portion has been reduced to the two amino acids (LN) positioned just upstream the HBsAg ATG start codon, to conserve the "strong efficiency" of this ATG codon in promoting protein translation by the ribosomal machinery [52].

[0124] In the wild type bicistronic HBV mRNA coding for both the pre-S2-HBsAg and the HBsAg proteins, no additional ATG codons are present in any of the three open reading frames (ORFs) between the pre-S2 and HBsAg start codons. While in the 1xFlag and the 3xFlag tags, three and six "medium efficiency" ATG codons [52] are present in their respective second ORF. To analyse the impact of these codons on recombinant VLPs production, we engineered the pGA3xFlagbasic plasmid, bearing the commercial 3xFlag tag nucleic acid sequence (SIGMA Aldrich). Supernatants from three independent transient transfections of the human SW480 established cell line with this plasmid were analysed by an anti-HBsAg ELISA, which detects the HBsAg protein only when assembled into VLPs [50].

[0125] Data showed that the ATG codons even in the second ORF have a major impact on VLPs assembling, as the test gave completely negative results for the pGA3xFlagbasic plasmid aside to positive results for the pCMV-S2.S controls.

[0126] Therefore, the 1xFlag and 3xFlag nucleic acid sequences were modified in order to remove the ATG codon in the second ORFs, thus obtaining the "1xFlag-M" and "3xFlag-M"tags with amino acid sequences identical to the originals (Fig. 23c and 23d). Once inserted in the pCMV-basic vector, they gave rise to the pGA1xFlag-Mbasic and pGA3xFlag-Mbasic plasmids, from which the pGA1xFlag-Mpol.opt and the pGA3xFlag-Mpol.opt were derived by cloning the polHIV-1.opt polyepitope through the *Eco*RI and *Xho*I restriction sites (Fig. 23b). Hydropathy profiles of the recombinant proteins clearly showed their hydrophilicity (Fig. 24 a, b, c and d), which parallels the general feature of the wild type HBV preS2 polypeptide N-terminal to HBsAg (Fig. 24e).

B.2.2. VLPs produced in an in vitro mammalian expression system bear HIV-1/HBV fusion proteins

[0127] To verify the production of recombinant VLPs, all the four Flag-M constructs were transfected in the SW480 cell line, and supernatants and cell

lysates were analysed by anti-HBsAg conformational ELISA (Fig. 25a and 25b). Data showed that nucleic acid modifications from the original Flag sequences determined the recovery of VLPs secretion from undetectable to significant levels. Notably, the Flag-M plasmids bearing the HIV-1 polyepitope gave 3 to 4 times higher VLPs secretion than the original ppolHIV-1.opt plasmid, showing that vector redesign was successful. At all time point, the pGA3xFlag-Mbasic samples showed a decrease of one log in VLPs secretion with respect to the pGA1xFlag-Mbasic construct, while both HIV-1 polyepitope bearing constructs were comparable with the pGA3xFlag-Mbasic samples (Fig. 25a). In the supernatants, VLPs secretion paralleled VLPs detection in SW480 cell lysates (Fig. 25b). This indicates that VLPs secretion is directly proportional to intracellular VLPs assembling into structured particles.

[0128] To verify that the secreted VLPs bear recombinant proteins, an anti-Flag-M ELISA was set up that combines Flag-M to conformational HBsAg detection. In this test, the anti-Flag M2 mAb traps any Flag-M bearing protein, which is revealed only if assembled into HBsAg VLPs. As the M2 mAb has a two log higher affinity for the 3xFlag than the 1xFlag tag (http://www.sigmaaldrich.com), it ensures more sensitive detection of recombinant VLPs bearing the 3xFlag-M constructs. For this reason and in the absence of a standard for recombinant VLPs detection, the anti-Flag-M ELISA could only be considered as a semiquantitative analysis, allowing robust comparison only between samples sharing the same tag (i.e. 1xFlag-M plasmids among themselves). Results from the anti-Flag-M ELISA on the SW480 supernatants showed that the VLPs detected by the anti-HBsAg ELISA (Fig. 25a and 25b) did contain recombinant fusion proteins (Fig. 25c and 25d). Basic 1xFlag-M and 3xFlag-M constructs could be tested at lower HBsAg concentrations than the respective HIV-1 polyepitope bearing constructs (Fig. 25c and 25d). This shows that longer polypeptides placed in the N-terminal position of HBsAg destabilise VLPs and that, as a consequence, there is a less efficient inclusion of recombinant proteins into VLPs. Given the M2 mAb different affinity for the two tags, values obtained in the anti-Flag-M ELISA for the 1xFlag-M construct have to be increased by a factor of 10^2 to be compared to 3xFlag-M values. Hence, for the same quantity of HBsAg protein, the 1xFlag-M constructs

gave rise to VLPs bearing a higher amount of recombinant fusion proteins than the 3xFlag-M plasmids.

B.2.3. Stable expression of recombinant HIV-1/HBV transgenes in plants

[0129] The GA1xFlag-Mbasic-HBsAg, GA1xFlag-Mpol.opt-HBsAg, GA3xFlag-Mbasic-HBsAg, GA3xFlag-Mpol.opt-HBsAg transgenes (Fig. 23b, 23c and 23d) were subcloned into the pAMPAT-MCS binary vector for expression in plant cells to obtain four different Flag-M constructs. Expression of the inserted sequence is driven by a doubly enhanced cauliflower mosaic virus 35S promoter (p35Sde) known to give strong and constitutive expression in plant tissues (Fig. 26). The four constructs were then used for stable nuclear transformation of *Nicotiana tabacum* and *Arabidopsis thaliana* through infection with *Agrobacterium tumefaciens*. The plants obtained were analysed by PCR to verify the nuclear integration of the transgenes. This analysis showed that 85 out of 95 (89%) regenerated T0 tobacco plants (Fig.27) and 137 plants out of 147 (93%) Arabidopsis T1 plants (Fig. 18) were transgenic. To determine the number of integrated transgène copies and the independence of the integration events, plants were analysed by Southern blot (Fig. 27 and 28). Approximately, 88% of tobacco transgenic plants contained between 1 to 3 copies of integrated transgene, with about 51% showing a single insertion. Arabidopsis revealed a higher percentage of transgenic plants having multiple transgene copies, as 19% bear a single transgene copy, 54% up to 3 copies and 46% more than 4 copies.

B.2.4. VLPs produced in tobacco plants bear the HIV-1/HBV fusion protein

[0130] The production of the recombinant HIV-1/HBV peptides and their assembly into VLPs in the 85 T0 transgenic tobacco plants was firstly verified by the anti-HBsAg ELISA on crude total protein extracts obtained from leaves at time point E1 (protein extraction E1-A).

[0131] Total soluble protein (TSP) was quantified in each sample to normalize the data obtained from the anti-HBsAg ELISA (Fig. 27). As expected, variability in transgene copy numbers and insertion sites (modulating gene transcription) resulted in nonhomogeneous VLPs production levels in different tobacco plants from the same construct.

[0132] The best expressing transgenic plant containing the pGA1xFlag-Mbasic construct revealed a VLPs production up to ~8-fold higher than the best

pGA3xFlag-Mbasic plant. Both the 1x and 3x HIV-1 polyepitope bearing constructs gave VLPs production in a comparable range, with a decrease from the best expressing basic constructs of 33- and 4-fold, respectively (Fig. 29).

[0133] Out of this broad initial screening, 14 plants were selected as a representative sample for further analysis on the basis of their higher level of HBsAg expression. Four plants per construct were retained except for the pGA3xFlag-Mpol.opt, for which only two plants gave significant VLPs production (Fig. 30 and Fig. 27). In the 14 selected plants, the transgene mRNA expression level was analysed by Northern blot, using a probe which could detect all HBsAg recombinant transcripts (with and without the HIV-1 polyepitope), (Fig. 30 and Fig. 27). The analysis was performed on RNA isolated from young (E1) and mature (E2) plants. In both experiments, plants bearing the transgenes lacking the HIV-1 polyepitope revealed mRNA expression levels at least two-fold higher than the corresponding HIV-1 counterparts (Fig. 30). Transcription remained constant throughout plant growth, as no statistically relevant differences could be observed between E1 and E2 hybridization experiments (Wilcoxon signed-rank test: $p>0.05$).

[0134] For the 14 plants, an anti-HBsAg ELISA was then performed on a second protein extraction made from young plants (E1-B; Fig. 30). The HBsAg values were statistically comparable with those obtained in E1-A (Wilcoxon signed-rank test: $p>0.05$) demonstrating that, at a given time point, estimations of HBsAg VLPs production on crude plants extracts give robust results. To investigate the levels of VLPs production in mature plants, an anti-HBsAg ELISA test was performed on protein extracts from the 14 selected tobacco plants at time point E2. Data showed increased VLPs production in mature plants, with HBsAg values that were on average 3-fold higher than mean E1 values and up to 11-fold higher for pGA3xFlag-Mbasic plants (Fig. 30). The differences between HBsAg measurements in E1 (the mean between E1A and E1B values was taken into account) and E2 were statistically significant (Wilcoxon signed rank test: $p<0.05$). Transgene mRNA expression levels did not statistically correlate with data from the anti-HBsAg ELISA at both E1 and E2 time points (zcorrelation test: $p>0.05$) (Fig. 30).

[0135] To verify that the VLPs produced in tobacco contain the HIV-1/HBV fusion proteins, aliquots of the E1 protein extracts previously used in anti-HBsAg ELISA were tested by the semi-quantitative anti-Flag-M ELISA (Fig. 31). For the Flag-Mbasic constructs it was possible to analyse different dilutions of the extracts, corresponding to the HBsAg content indicated on the X-axes (Figs. 31a and 31b). By contrast, for the HIV-1 polyepitope plus constructs undiluted extracts had to be used in order to get results significantly above the cut-off value (Figs. 31c and 31d). The results showed almost linear correlations among HBsAg content and Flag-M OD620nm values in the Flag-Mbasic samples, with higher amounts of HBsAg corresponding to increased detection of Flag-M recombinant protein (Figs. 31a and 31b). The analysis by the Wilcoxon signed-rank test of the differences between the E1-A and E1-B data from plants bearing the HIV-1 polyepitope (Figs. 31c and 31d) showed that the apparent trait towards less fusion protein detection in the E1-B samples than in the E1-A was not statistically significant ($p>0.05$). The apparent discrepancy between E1-A and E1-B data can be explained by variability in ELISA detection at low positive values and by the fact that the ratio between fusion protein and HBsAg into VLPs is intrinsically not constant, as it is the case in the wild type HBV context (HBsAg versus preS2-HBsAg). On the basis of the anti-Flag-M ELISA, the presence of recombinant HIV-1/HBV proteins in VLPs produced in tobacco could be demonstrated for all the 14 analysed plants. As the affinity of the M2 mAb for the 3xFlag-M is two logs higher than for the 1xFlag-M, OD620nm values for 1xFlag-Mbasic plants have to be increased by a factor of 10^2 to be comparable with 3xFlag-M values. Hence, as in the mammalian expression system, the 1xFlag-M constructs gave rise to VLPs bearing a higher amount of recombinant proteins than the 3xFlag-M plasmids.

[0136] To investigate the stability of transgene expression and recombinant VLPs production in subsequent plant generations, eight out of the 14 T0 plants listed in Table 2 (two plants for each construct) were seed propagated and a total of 40 T1 progeny plants was analysed by Northern blot, anti-HBsAg and anti-Flag ELISA tests (Fig. 32). Transgene mRNA expression levels remained stable in the sexual progeny, as values for T1 plants did not differ significantly from the T0

parent plants at both E1 and E2 time points (Wilcoxon signed-rank test performed comparing mean E1 and E2 values to T1 mean values: $p>0.05$).

[0137] However, VLPs production in T1 progeny plants bearing the Flag-Mbasic constructs increased up to 6-fold when compared to HBsAg values of the relative T0 parents at time point E2, and up to 47-fold when compared to T0 E1 values. By contrast, T1 plants bearing the HIV-1 polyepitope showed a reduced VLPs production (up to 5-fold and 17-fold less versus T0 E1 and E2 values, respectively).

[0138] As for T0 plants, statistical analysis showed no correlation between the level of transgene mRNA expression and relative HBsAg VLPs production (zcorrelation test: $p> 0.05$). Notably, in six out of 20 T1 plants bearing the HIV-1 polyepitope, VLPs could not be detected by anti-HBsAg ELISA, while all the plants were positive in the more sensitive anti-Flag-M ELISA (Fig. 33), demonstrating that VLPs bear recombinant proteins in all the T1 plants. Taken together, these data showed that transgène expression was retained in the T1 generation, with a general trend towards higher recombinant VLPs production in plants bearing the 1x and 3x basic constructs than in the polyepitope bearing ones, as observed in T0.

B.2.5. Stable production of VLPs bearing the HIV-1/HBV fusion protein in Arabidopsis plants

[0139] Five weeks following transfer to soil, the 137 transgenic Arabidopsis plants were screened for recombinant VLP production by anti-HBsAg ELISA (Fig. 28 and Fig. 29). The HBsAg concentrations obtained were normalized with respect to TSP content determined by Bradford. In Arabidopsis, as in tobacco and in the mammalian expression systems, VLP production among the best 1xFlag-Mbasic and 3xFlag-Mbasic transgenic plants differed by one log, while it was comparable among plants bearing the HIV-1 polyepitope. Moreover, a reduced VLPs production was found in Arabidopsis plants expressing the HIV-1 polyepitope as compared with the 1x or 3x Flag-Mbasic counterparts, as it was the case for tobacco at any time point (T1 and T0 E1-A, E1-B and E2).

[0140] On the basis of this analysis, the best HBsAg VLP producing plants (four plants for each construct) were selected and further characterized by Northern blot (Fig.33 and Fig. 27) and anti-Flag-M ELISA (Fig. 34). As for

tobacco, reduced mRNA levels (about 2-fold) were found in plants expressing the HIV-1 polyepitope and no correlation could be established between the transgene mRNA expression levels and the HBsAg concentrations by the z-correlation test ($p > 0.05$; Fig. 33).

[0141] The presence of recombinant proteins in VLPs from the 16 Arabidopsis selected plants was verified by anti-Flag-M ELISA (Fig. 35). Different HBsAg VLPs dilutions could be tested for the Flag-Mbasic protein extracts (Fig. 35a), while for the HIV-1 polyepitope bearing plants undiluted lysates had to be analysed to get reliable output data (Fig. 35b). Taking into account the 10^2 log higher affinity of the M2 mAb for the 3xFlag-M than for the 1xFlag-M, the 1xFlag-M constructs gave rise to VLPs bearing a higher amount of recombinant proteins than the 3xFlag-M plasmids, as it was the case for tobacco. By comparing OD620nm values from anti-Flag-M ELISA on tobacco and Arabidopsis, data showed that values directly corresponded for basic constructs (Figs. 25a and b; Fig. 26a). This was also the case for the HIV-1 polyepitope bearing plants (Figs. 25c and 25d; Fig. 26b), once OD620nm values were compared taking into account the different HBsAg concentration in the anti-Flag-M ELISA tests. Hence, Arabidopsis and tobacco resulted comparable plant expression systems for recombinant HIV-1/HBV VLPs, since fusion protein contents in VLPs were found to be at the same levels in the two plants.

Example C: Anti-HIV-1 cellular immune responses elicited *in vivo* by a transgenic plant-based oral vaccine

Example C 1.: Materials and Methods

C.1.1. Transgenic tobacco

[0142] Selected T_0 plants, together with a wild type negative control, were maintained in the greenhouse through several rounds of cuttings to maintain maximum vegetative production, in order to constitute leaf stocks for multiple oral administration to mice. Collected leaves were lyophilised and then mechanically ground to powder. From a fresh weight (FW) of leaves to lyophilised powder, a 10-fold reduction was obtained. Lyophilised material from plants bearing the same construct was then mixed and the resulting stocks stored at 4°C. From each stock, 0.1g of powder was resuspended in 2ml of the following extraction buffer: 1xPBS pH 7.4, 10mM EDTA, 0.1% Triton X-100 and 1mM

phenylmethylsulphonyl fluoride (PMSF), centrifuged at 10,000 rpm for 10 min at 4°C and then supernatants were submitted to anti-HBsAg, anti-Flag-M ELISA and Bradford analyses as previously described (Example B). For the anti-Flag-M ELISA, plant stocks 4 and 18 were tested at different HBsAg concentration following anti-HBsAg analysis (from 0.3 to 2ng/ml HBsAg). For stocks 5 and R, the maximum volume of 100□l for the test was analysed. Anti-Flag-M ELISA limit of detection corresponded to 0.25 OD_{620nm}.

C.1.2. Mice immunisation and collection of biological tissues

[0143] Immunisation was performed on 6 to 8 week old HLA-A*0201 and HLA-DR1 double transgenic female HSB mice ($HHD^{+/+}$ $b2m^{-/-}$ $HLA\text{-}DR1^{+/+}$ $IAb^{-/-}$; [26]) according with institutional guidelines. The pGA1xFlag-Mpol.opt plasmid DNA (Example B) for immunisation was prepared by endotoxin-free giga-preparation kit (QIAGEN) and re-suspended in endotoxin-free PBS (Sigma). Five days before DNA injection (d-5), an inflammatory reaction was induced by inoculating 1nmol of cardiotoxin (Latoxan) per hind leg [Loirat, 1999 #164]. At day 0 (d0), intramuscular prime immunisation was performed by injecting 50□g of the pGA1xFlag-Mpol.opt plasmid DNA per hind leg into regenerating tibialis anterior muscles [Loirat, 1999 #164]. Then, immunization boosts were performed according to experimental protocols by administrating 0.1g of transgenic plants mixed to 5g (Monday, Tuesday and Thursday) or 2g (Friday) of hydrated paste normally used to feed mice during their transfer between animal houses. Wild type lyophilized tobacco plants were used as negative boost control. According to protocols, mice were sacrificed at day 24 (d24) or 31 (d31) and different biological tissues were taken by microchirurgical intervention. Blood was collected by intra-heart puncture, heparinized and centrifuged 5 minutes at 3,000rpm. Serum was stored at 4°C. The spleen and the small gut were recovered in RPMI medium supplemented with 5% foetal calf serum (FCS), 2% streptomycin and penicillin and 1% glutamine (complete medium). Peripheral lymph nodes (mesenteric, maxillary, axillary and inguinal) were recovered in this complete medium supplemented with 2mM EDTA. Lymph nodes and spleens were mechanically crushed. Splenocyte suspensions were submitted to FicollYL purification procedures as previously described [50] and re-suspended at 10 x 10^6 cells in 1ml of RPMI supplemented with 3% FCS. The small gut was

extensively washed with 1xPBS and either the intestinal epithelial lymphocytes associated with the gut mucosa (IELs) were isolated or it was directly lysed. IELs were collected from the epidermis of the intestinal mucosa as described previously (Buzoni-Gatel et al. J Immunol 1999;162(10):5846-52.; Mennechet et al. Eur J Immunol 2004;34(4):1059-67) and then resuspended in complete RPMI medium and counted.

[0144] At sacrifice, feces were collected from the colon in 2ml of 1xPBS/antiprotease (1 pastille of antiprotease cocktail from Roche in 50ml of 1xPBS) and let 10 minutes at room temperature. The suspensions were hardly vortexed, centrifuged for 10 minutes at 1,500 rpm, supernatants were recovered into clean tubes, recentrifuged for 15 minutes at 5,000g and stored at -20 °C.

[0145] On serum and feces supernatants, detection by ELISA of anti-HBsAg antibodies was performed as previously described [50], using for IgG detection the mouse polyclonal anti-IgG (Amersham: NXA931) 10^{-3} diluted, and for IgA detection the mouse polyclonal anti-IgA diluted at 1/500 (STAR 85P; AbDSeroTec), both antibodies being labeled with peroxidase.

C.1.3. Cell sorting by magnetic beads

[0146] Spleens and peripheral lymph nodes from 16 mice boosted with stock 5 of tobacco transgenic plants were pooled to obtain 5 samples, 4 deriving from three mice and 1 from four mice. Organs were treated as previously described [Michel, 2007 #80] and then $CD8^+$ T lymphocytes were negatively selected by the CD8a+ T Cell Isolation Kit (Myltenyi Biotec) applied on LS columns (Myltenyi Biotec) according to manufacture's instructions. By this technique, mouse $CD8a^+$ T cells were isolated by depletion of non-$CD8a^+$ T cells by the following antibodies cocktails (for CD4, clone L3T4; for CD45R, B220; for CD49b, DX5; for CD11b, Mac-1; and Ter-119). Once recovered the $CD8a^+$ population (for spleen, 60% of estimated purity once assembled with feeders, and for peripheral lymph nodes, 30% in the same experimental conditions), the non-$CD8a^+$ cells were separately eluted from columns. Both cell populations were resuspended in 1xPBS and counted.

C.1.4. Foxp3 intra-cellular labeling

[0147] In 96 well plates, $5x10^6$ cells per sample were centrifuged for 5 minutes at 1,700rpm to eliminate the supernatant. Twenty microliter of mouse Fc

block anti-CD11/CD32 (BD Pharmingen) 10^{-2} diluted into PBS FACS (1xPBS 1%BSA and 10^{-4} azide) were added and incubated 10 minutes at 4°C. Cells were washed with 100µl of PBS FACS and centrifugation for 5 minutes at 1,700rpm. Cells were incubated with 20µl of mouse anti-CD4-PerCP, anti-CD3-APC and anti-CD25-FITC (BD Pharmingen: clones RM4-5, 145-2C11 and 7D4, respectively) 10^{-2} diluted in PBS FACS for 15 minutes at 4°C in the dark. Then, cells were washed as previously described and resuspended in 100µl of Fixation/Permeabilization solution, readily prepared as described by the manufacturer (FoxP3 staining buffer set; ebioscience). Cells were incubated for 30 minutes at 4°C in the dark, washed with 200µl of 1xPermeabilization Buffer and incubated with anti-mouse Foxp3-PE (clone FJK-16s; ebioscience) 10^{-2} diluted in 1xPermeabilization Buffer for 30 minutes at 4°C in the dark. Cells were washed twice in 200µl of 1xPermeabilization Buffer and resuspended in 100µl of fixation buffer (1xPBS, 1%BSA, 10^{-3} azide and 2% formaldehyde) and analysed by FACScalibur (BD Biosciences).

C.1.5. INF-γ and IL-10 secretion assays

[0148] INF- γ and IL-10 secretion assays were performed on cells from spleens or on peripheral lymph nodes (mesenteric, maxillary, axillary and inguinal). The assays were made following the manufacturer's instructions (Miltenyi Biotec) and as described previously [50]. In the INF- γ assay on CD8+ T cell sub-population obtained from cell sorting, feeder cells were put to a ratio of 1 to 1 with respect to CD8+ analysed cells. While in all the other INF- γ and IL-10 assays, different peptides were directly added to culture cell supernatants to a final global concentration of 10µg/ml. Both feeder cells and peptides were assembled to analysed cells and incubated for 16h at 37°C. To obtain feeder cells, relevant or irrelevant peptides were separately incubated for 2h at room temperature with splenocytes from naïve female HSB mice at 10µg/ml final concentration, before irradiation at 10,000rad for 45 minutes. Feeder cells charged with each of the eight relevant HIV-1 peptides were pooled following incubation and previously to be aliquoted per sample. Previously described HIV-1 relevant class I peptides [50] were provided either into two pools of four epitopes each (1^{st}: S9L, L10V, L9V and Y/I9V; and 2^{nd}: V11V, Y/P9L, Y/T9V and Y/V9L) or the eight altogether, as specified in the text. Relevant peptides for HBsAg class II

epitopes were T15Q and Q16S [Pajot, 2006 #143]. The irrelevant class I G9L peptide [Michel, 2007 #80] in INF-□ secretion assay and class II G15W peptide [Pajot, 2006 #143] in IL-10 secretion assay were used as negative controls. For positive control samples, 12.5ng/ml phorbol 12-myristate 13-acetate (PMA; SIGMA) and 1mg/ml ionomycin (SIGMA) were added to cells.

[0149] In the INF-γ secretion assay, samples were labelled with INF- γ catch reagent and with the INF-$\tilde{\gamma}$-PE (Miltenyi Biotec), the CD8a-APC (clone 53-6.7; BD Pharmingen) and the CD3-FITC (clone 145-2C11; BD Pharmingen) antibodies. In the IL-10 secretion assay, the IL-10 catch reagent and the IL-10-APC (Miltenyi Biotec), the Foxp3-PE (clone FJK-16s; ebioscience), the CD4-PerCP (clone RM4-5; BD Pharmingen) and the CD25-FITC (clone 7D4; BD Pharmingen) antibodies were used. Samples were analysed by the flow cytometry analysis using a FACScalibur (BD Biosciences). *p* values were obtained by the StatView F-4.5 using the non-parametric Mann-Whitney or Wilcoxon signed-rank tests.

C.1.6. Proliferation test by CFSE labeling

[0150] Carboxyfluorescein diacetate (CFSE; 10nM; Invitrogen) proliferation test was performed on $1x10^{-6}$ cells in 24 well plates. At J0, cells were centrifuged at 1,200rpm for 5 minutes, resuspended in 5ml of RPMI containing CFSE 1/5,000 diluted, incubated 10 minutes at 37°C, washed with 2ml of RPMI and plated in 1ml of RPMI complete medium (where FCS was at 10%). The T15Q and Q16S (Pajot et al. Microbes Infect 2006;8(12-13):2783-90) class II HBsAg peptides were put to a final global concentration of 10 μg/ml in the cultures supernatants of $CD8^+$ depleted samples from spleens and peripheral lymph nodes. Positive control was obtained by adding 25µl of PMA (1µg/ml; SIGMA) and 10µl of ionomycin (100µg/ml; SIGMA) and negative controls were represented by cells stimulated with the class II irrelevant peptide G15W (Pajot et al. Microbes Infect 2006;8(12-13):2783-90) at 10µg/ml final concentration or samples never put in the presence of peptides. At J4, cells were centrifuged for 5 minutes at 1,700rpm, resuspended in 100□l PBS FACS and transferred to 96 well plates for Foxp3 labeling as previously described. As CFSE is visible on the FL1 channel of FACScalibur, the anti-CD3-APC was not used, and the anti-CD25-APC (BD Pharmingen: clone 7D4) replaced the anti-CD25-FITC.

C.1.7. Real time quantitative PCR analysis

[0151] The whole small gut (limited by the stomach and the caecum) was lysed by dipping the organ into 6ml of the lyse buffer from the EPICENTRE kit (Biotechnologies) and 100µl of Proteinase K (20mg/ml; Eurobio). Samples were incubated at 37°C over night on a rolling wheel. Then, total RNA extraction was performed on 150µl from the small gut lysate diluted to one third by adding 150µl of lyse buffer and 150µl of protein precipitation reagent from the kit. Then, samples were treated according to manufacture's instructions. Total RNA concentration in these extracts was determined by spectrophotometer (Nanodrop, Biocompare). Then, 3µg of RNA per sample was taken and resuspended into 29µl of H_20 and 1µl of RNasin (20-40 u/; PROMEGA). Two microliters of polydT (16mer; EUROGENTEC) were added and samples were incubated 10 minutes at 70°C. Then, cDNA was synthesized by adding 1µl of reverse transcriptase (Super Script TM II 200u/µl; Invitrogen), 1µl of RNasin (20-40 u/µl; PROMEGA), 1µl of dNTP (40mM), 5µl of DTT (0.1M) and 10µl of 5xFirst Stand Buffer (Invitrogen) and incubating for 1h at 42°C and 10 minutes at 95°C. Samples were stored at -20°C.

[0152] Real time quantitative PCR (RQ-PCR) on different tolerance markers in the small gut was performed using 0.75µl of cDNA, 12.5µl of 2xBuffer (Taqman Universe PCR master mix; Applied Biosystem) and 1.25µl of H_20 to obtain the "cDNA mix". For Foxp3 and CD3 analyses, to the 14.5µl of "cDNA mix", 9.25µl of H_20 and 1.25µl of primer/probe specific solutions were added (Taqman gene expression assay Foxp3, 4331182/Mm 00599683-m1; Taqman gene expression assay CD3ε, 4331182/Mm 00475156-m1; Applied Biosystem). For INF-γ, Smad2, Smad3 and IL-10, to the total 14.5µl of the cDNA mix, 10µl of cytokine-specific primer (2µM) and 0.5µl of probe (20µM) were added. For these last cDNA, primers and probes were:

forward (FW) 5'-AAAGGATGCATTCATGAGTATTGC,
reverse (RV) 5'-CGCTTCCTGAGGCTGGATT and
probe (P) 5'-AGGTCAACAACCCACAGGTCCAGCG for INF-γ
FW 5'-CGGCTGAACTGTCTCCTACTCCTCT,
RV 5'-CGAGTTTGATGGGTCTGTGA and
P 5'-CATTCTGGTGTTCAATCGCATACTAT for Smad2;

FW 5'-CAAATTCCTGGTTGT TGAAGATCTT,
RV 5'-GCAACCAGCGCTATGGCT and
P 5'-CACCCGGCCACTGTCTGCAATAT for Smad3;
FW 5'-GGCGCTGTCATCGATTTCTC,
RV 5'-GACACCTTGGTCTTGGAGCTTATT and
P 5'-AAAATAAGAGCAAGGCAGTGGAGCAGGTG for IL-10.

[0153] Reported results are the mean of triplicates of copy numbers of the target gene divided by the mean of triplicates of copy numbers of the CD3ε taken as gene expression reference.

Example C 2.:. Results

C.2.1 Preparation of plant stocks expressing recombinant HIV-1/HBV VLPs

[0154] In a previous Example B, we have described the expression of HIV-1/HBV recombinant virus-like particles (VLPs) in *Nicotiana tabacum* and *Arabidopsis thaliana*. This work represented the first demonstration that it is possible to produce in plants recombinant VLPs based on the assembly of the HBsAg of HBV and of HIV-1/HBV fusion proteins where the class I restricted HIV-1 polyepitope (polHIV-1.opt) is N-terminal to HBsAg. The polHIV-1.opt was optimized in order not to impair recombinant VLPs assembly and to be exposed on HBV VLPs surface [50]. By DNA immunization, it was possible to demonstrate that HIV-1/HBV recombinant VLPs could elicit *in vivo* a HIV-1 specific activation of peripheral CD8+ T cells [50]. Subsequent comparison of recombinant VLPs produced in a mammalian established cell line and plants showed that VLPs quality, defined by the relative quantity of HIV-1/HBV fusion proteins assembled into VLPs, was similar in the two expression systems (Example B). Moreover, the production levels and quality of recombinant VLPs were comparable in the two experimental plant species: *Nicotiana tabacum* and *Arabidopsis thaliana*. All these data taken together showed that parameters intrinsic to the recombinant proteins determined their assembly into HBV VLPs whichever the expression system. Hence, once tested in a given plant species, the fusion protein of HBsAg can most likely be transposed to any plant expression system, preserving the quality of produced recombinant VLPs. The described fusion proteins (Example B and [50]) represent an innovative tool to set up an anti-HIV-1 vaccine based on oral administration of crude extracts from transgenic plants.

[0155] The HIV-1/HBV transgenes (GA1xFlag-Mpol.opt and GA3xFlag-Mpol.opt) used to transformed plants were constituted by a bicistronic open reading frame essentially expressing, from N-terminal to C-terminal, a polyprotein made by a tag (1xFlag-M or 3xFlag-M), the HIV-1 polyepitope (polHIV-1.opt) and the HBsAg (Example B). The 1x and 3xFlag-Mbasic constructs corresponded to the respective tag transgenes devoid of the HIV-1 polyepitope (GA1xFlag-Mbasic and GA3xFlag-Mbasic). Among the transgenic tobacco plants previously described, two or three plants per construct were chosen on the basis of the amount and quality of VLPs produced (GA1xFlag-Mbasic: plants 4-4 and 4-11; GA1xFlag-Mpol.opt: plants 5-15, 5-17 and 5-38; GA3xFlag-Mbasic: plants 18-2 and 18-7; GA3xFlag-Mpol.opt: plants R-12 and R-D). Plants were lyophilised and material from the same construct were pooled to yield stocks #4, #5, #18 and #R which were characterized as far as recombinant VLPs concentration and presence of HIV-1 polyepitope on their surfaces are concerned. HBsAg concentration determined by anti-HBsAg ELISA was reported to total soluble protein (TSP) concentration evaluated by the Bradford test. By these analyses, the two stocks corresponding to tobacco plants bearing the 1x and 3xFlag-Mbasic constructs (stock #4 and #18, respectively) gave 436ng HBsAg/mg TSP and 278ng HBsAg/mg TSP. In the stocks corresponding to plants bearing the HIV-1 polyepitope (stock #5 and #R), HBsAg concentration was under the limit of detection (0.2ng) of the ELISA assay. On all the stocks, an anti-Flag-M ELISA was performed to detect the Flag-M tag N-terminal to the HIV-1 polyepitope in the HIV-1/HBV fusion protein, indirectly demonstrating by its detection the presence of the polyepitope on recombinant VLP surface (Example B). For plants expressing the 1x or 3xFlag-Mbasic constructs (stock #4 and #18), fusion protein detection was positive and content in recombinant VLPs was ~5 fold less than in younger T_0 plants (E1; Example B). For HIV-1 polyepitope bearing plants (stock #5 and #R), Flag-M detection was positive for the 1xFlag-M construct and under the limit of detection for the 3xFlag-M.

C.2.2 Orally administrated transgenic plants can activate HIV-1 specific CD8+ T lymphocytes.

[0156] Following characterization of tobacco stocks, lyophilised plants bearing the 1xFlag-M constructs (stocks #4 and #5) were selected for further analyses of

their immunogenicity *in vivo*. It was decided to concentrate on the analysis of the ability of these plants to boost a classical DNA-primed vaccination (Figure 36). Following cardiotoxin injection at d-5, nine HSB mice were primed at day 0 (d0) by the pGA1xFlag-Mpol.opt plasmid which bears the 1xFlag-M tag and the polHIV-1.opt polyepitope (Example B). At d10 of a classical immunization protocol, a second DNA-injection is provided to boost a cellular response (Example B). Instead of this, in protocol number 1 (Figure 36A), mice were feed by 0.1g/day of wild type crude lyophilised tobacco or stocks #4 or #5, twice two successive days in a week for two successive weeks. At d12 and at d19 time-breaks, mice received "normal" food. When mice were provided with lyopilised plants, they were in individual cages to ensure complete uptake of the 0.1g/day crude plant extract. Mice were sacrificed at d24 and the blood, spleen and mesenteric lymph nodes were taken by microchirurgical intervention. ELISA detection of anti-HBsAg IgG antibodies in serum yielded negative results. On cells recovered from spleen and mesenteric lymph nodes, the *ex vivo* secretion IFN-γ assay was performed in order to detect HIV-1 specific activated CD8+ T lymphocytes. Spleens could be analysed separately, while mesenteric lymph nodes had to be pooled in groups of three mice belonging to the same plant administration (wt, #4 or #5). In the *ex vivo* secretion IFN- γ assay, cells were tested following stimulation with two separate pools of four HIV-1 peptides each, all corresponding to the eight epitopes in the HIV-1 polyepitope expressed in plants ([50] and Example B). Among all the performed analyses, one pool of mesenteric lymph nodes from HSB mice having received stock #5 in protocol 1 resulted clearly positive for IFN-γ secretion from CD8+ T lymphocytes, once stimulated with a pool of four epitopes out of the eight composing the HIV-1 polyepitope (Figure 37A) [50]. Stimulation with the complementary pool of four peptides gave negative results (Figure 37B) on an experimental 0% background signal obtained with the irrelevant G9L peptide (Figure 37C). The 0.14% data obtained following stimulation with the first pool of four peptides is highly relevant as it has to be considered out of the 5% of total CD8+ T lymphocytes present in naïve HSB mice (data not shown) to be compared to 20% in C57/Black/6 mice, the genetic background of HSB mice. This result represents the first

demonstration that it is possible to boost a systemic anti-HIV-1 specific cellular immune response by oral administration of transgenic plants.

[0157] *C.2.3 Treg activation can counterbalanced immunogenicity of transgenic plants*

[0158] As detection of systemic HIV-1 specific activation of CD8+ T lymphocytes was limited to one pool of mesenteric lymph nodes out of all the analysed samples, the hypothesis was made that protocol 1 could rather induce tolerance to antigens than favors antigens immunogenicity. In the aim to reduce eventual tolerance induced by tobacco and to characterize this tolerance, a second administration protocol was designed in which the two weeks in which crude plant extracts are administrated were separated by one week with normal food administration (Figure 36B). Following DNA-prime, three groups of nine mice each were submitted to different regimes: either with wild type (wt) tobacco or with stock 4 (#4) or 5 (#5). Hence, at d31 blood, spleen, peripheral lymph nodes (mesenteric, maxillary, axillary and inguinal), the small gut and the colon were collected from mice feed with tobacco and naive mice. From blood and colon, serum and feces were recovered, respectively and an ELISA to detect anti-HBsAg antibodies was performed on serum (anti-IgG and anti-IgA ELISA) and feces (anti-IgA ELISA) and did not gave values above background obtained with biological materials from naïve mice. From small guts, intestinal epithelial lymphocytes associated with the gut mucosa (IELs) were purified and had to be pooled in one sample for each of the four mouse groups (naïve, wt, #4 and #5) as in HSB mice IELs are very few (mean value of $1x10^6$/HSB mouse to be compared to mean values of $4x10^6$/C57/Black/6 mouse) mirroring low peripheral CD8+ T lymphocytes ratio. On spleen, peripheral lymph nodes and IELs, *ex vivo* INF-γ secretion assay by CD8+ T lymphocytes was performed by stimulating each sample with the pool of the eight HIV-1 peptides. Unlike protocol 1, peripheral lymph nodes could be analysed per individual mouse. None of the samples could give positive results.

[0159] In the aim to evaluate tolerance induction by administration of tobacco following protocol 2, Foxp3 intra-cellular labeling on CD3+CD4+T lymphocytes cellular subset was performed on the same cellular samples analysed in the *ex vivo* INF-γ secretion assay (Figure 38). Data showed that indeed tobacco induces

a significant activation of Tregs either in spleen (Figure 38A) or in peripheral lymph nodes (Figure 38B) as demonstrated by Mann-Whitney analysis between the naïve and the wt mouse groups. The presence of expressed viral antigens in plants could not inverse the balance between tolerance and immunogenicity in favor of the this last, as shown by not significant differences between wt, #4 and wt, #5, respectively, in the spleen. The statistically relevant differences between wt and #4 or #4 and #5 groups in peripheral lymph nodes may just be due to difference in time points of Treg activation kinetics, the sacrifice being programmed too close to the last immunization boost by plants.

C.2.4. When Tregs are depleted, HIV-1 specific INF-γ secretion by CD8+ T lymphocytes can be restored

[0160] Based on these results, the hypothesis was made that Tregs in lymphoid tissues could inhibit IFN-γ secretion by HIV-1-specific CD8+ T lymphocytes. Hence, protocol 2 was performed on 18 HSB mice by feeding them with stock 5, in order to identify HIV-1-specific cell population. From these mice, blood, peripheral lymph nodes and spleens were collected and analysed. Once again, anti-HBsAg IgG antibodies in the serum could not be detected. Peripheral lymph nodes and spleens were submitted to cell sorting by magnetic beads to negatively separate CD8a+ T lymphocytes from all other cellular populations (among the more relevant: CD4+ T lymphocytes, antigen presenting cells, red cells, natural killer cells). Then, non-CD8a+ T cells were eluted and submitted to CFSE combined to Foxp3 intracellular labeling and *ex vivo* IL-10 secretion assay. The intensity of CFSE labeling and IL-10 secretion in Foxp3+ or Foxp3- populations was compared to data obtained from the same sample submitted to stimulation with the two class II relevant (T15Q and Q16S (Pajot et al. Microbes Infect 2006;8(12-13):2783-90) or the irrelevant (G15W (Pajot et al. Microbes Infect 2006;8(12-13):2783-90)) HBsAg peptides. By these analyses, any antigen-specific proliferation or activation could be put in evidence in the non-CD8+ T cell fractions. The CD3+CD4+Foxp3 labeling showed that Tregs were represented at 30-35% in the non-CD8+ T cellular subsets of spleen and peripheral lymph nodes. All these data taken together show that oral immunization didn't have any impact on antigen-specific activation or proliferation of Tregs, but that the Tregs population was elicited by transgenic plant material as shown in Figure 38.

[0161] The CD8a+ T lymphocytes from spleen and peripheral lymph node samples were analysed *ex vivo* by IFN-γ secretion assay (Figure 39). For each samples, the pool of eight HIV-1 peptides was used as relevant antigens and compared to cells stimulated with the irrelevant G9L peptide. The non-parametric Wicoxon signed-rank test, which allows to compare each HIV-1 test to its respective G9L test, shows that following cell sorting, in both peripheral lymph nodes and spleen, CD8+ T lymphocytes could be activated by stimulation with HIV-1 specific peptides. Notably, in the G9L samples, secretion of IFN-γ was significantly above medium negative control. This may indicate that CD8+ T lymphocytes issued from immunized mice are activated *in vivo* and retain this activation state for two days in the *ex vivo* assay, where IFN-γ secretion can occur in the absence of inhibiting Tregs and be enhanced by HIV-1 specific peptides. Mean activation levels of the 8 HIV1 peptides correspond to those obtained previously by "classical" DNA-immunisation protocol by a construct (pHIV-1pol.opt [50]) bearing the HIV-1 polyepitope inserted in transgenic plants (Example B). This demonstrates that HIV-1-specific CD8+ T cells could be elicited in the periphery by mucosal immunization by using transgenic plants expressing viral antigens. Nevertheless, these HIV-1-specific CD8+ T cells can be identified and activated only in the absence of Tregs.

REFERENCES

[0162] The following references are cited herein. The entire disclosure of each reference is relied upon and incorporated by reference herein.

1. Albert, M. L., S. F. Pearce, L. M. Francisco, B. Sauter, P. Roy, R. L. Silverstein, and N. Bhardwaj. 1998. Immature dendritic cells phagocytose apoptotic cells via alphavbeta5 and CD36, and cross-present antigens to cytotoxic T lymphocytes. J. Exp. Med 188:1359-1368.

2. Bisht, H., D. A. Chugh, M. Raje, S. S. Swaminathan, and N. Khanna. 2002. Recombinant dengue virus type 2 envelope/hepatitis B surface antigen hybrid protein expressed in Pichia pastoris can function as a bivalent immunogen. J. Biotechnol. 99:97-110.

3. Boisgerault, F., G. Moron, and C. Leclerc. 2002. Virus-like particles: a new family of delivery systems. Expert Rev. Vaccines 1:101-109.

4. Boyle, J. S., C. Koniaras, and A. M. Lew. 1997. Influence of cellular location of expressed antigen on the efficacy of DNA vaccination: cytotoxic T lymphocyte and antibody responses are suboptimal when antigen is cytoplasmic after intramuscular DNA immunization. Int. Immunol. 9:1897-1906.

5. Bruss, V. 2004. Envelopment of the hepatitis B virus nucleocapsid. Virus Res. 106:199-209.

6. Bullock, T. N., T. A. Colella, and V. H. Engelhard. 2000. The density of peptides displayed by dendritic cells affects immune responses to human tyrosinase and gp100 in HLA-A2 transgenic mice. J. Immunol. 164:2354-2361.

7. Buseyne, F., M. Fevrier, S. Garcia, M. L. Gougeon, and Y. Riviere. 1996. Dual function of a human immunodeficiency virus (HIV)-specific cytotoxic T-lymphocyte clone: inhibition of HIV replication by noncytolytic mechanisms and lysis of HIV-infected CD4+ cells. Virology 225:248-253. 30

8. Chua, P. K., R. Y. Wang, M. H. Lin, T. Masuda, F. M. Suk, and C. Shih. 2005. Reduced secretion of virions and hepatitis B virus (HBV) surface antigen of a naturally occurring HBV variant correlates with the accumulation of the small s envelope protein in the endoplasmic reticulum and Golgi apparatus. J. Virol. 79:13483-13496.

9. Davis, H. L., C. L. Millan, and S. C. Watkins. 1997. Immune-mediated destruction of transfected muscle fibers after direct gene transfer with antigen-expressing plasmid DNA. Gene Ther. 4:181-188.

10. Doan, L. X., M. Li, C. Chen, and Q. Yao. 2005. Virus-like particles as HIV-1 vaccines. Rev. Med. Virol. 15:75-88.

11. Firat, H., F. Garcia-Pons, S. Tourdot, S. Pascolo, A. Scardino, Z. Garcia, M. L. Michel, R. W. Jack, G. Jung, K. Kosmatopoulos, L. Mateo, A. Suhrbier, F. A. Lemonnier, and P. Langlade-Demoyen. 1999. H-2 class I knockout, HLA-A2.1-transgenic mice: a versatile animal model for preclinical evaluation of antitumor immunotherapeutic strategies. Eur. J. Immunol. 29:3112-3121.

12. Firat, H., S. Tourdot, A. Ureta-Vidal, A. Scardino, A. Suhrbier, F. Buseyne, Y. Riviere, O. Danos, M. L. Michel, K. Kosmatopoulos, and F. A. Lemonnier. 2001. Design of a polyepitope construct for the induction of HLA-A0201-restricted HIV 1-specific CTL responses using HLA-A*0201 transgenic, H-2 class I KO mice. Eur. J. Immunol. 31:3064-3074.

13. Inaba, K., S. Turley, F. Yamaide, T. Iyoda, K. Mahnke, M. Inaba, M. Pack, M. Subklewe, B. Sauter, D. Sheff, M. Albert, N. Bhardwaj, I. Mellman, and R. M. Steinman. 1998. Efficient presentation of phagocytosed cellular fragments on the major histocompatibility complex class II products of dendritic cells. J. Exp. Med. 188:2163-2173.

14. Kozak, M. 2002. Pushing the limits of the scanning mechanism for initiation of translation. Gene 299:1-34.

15. Kypr, J., and J. Mrazek. 1987. Unusual codon usage of HIV. Nature 327:20.

16. Le Borgne, S., M. Mancini, R. Le Grand, M. Schleef, D. Dormont, P. Tiollais, Y. Riviere, and M. L. Michel. 1998. In vivo induction of specific cytotoxic T lymphocytes in mice and rhesus macaques immunized with DNA vector encoding an HIV epitope fused with hepatitis B surface antigen. Virology 240:304-315.

17. Lenz, P., C. D. Thompson, P. M. Day, S. M. Bacot, D. R. Lowy, and J. T. Schiller. 2003. Interaction of papillomavirus virus-like particles with human myeloid antigenpresenting cells. Clin. Immunol. 106:231-237.

18. Lewis, P. J., H. van Drunen Littel-van den, and L. A. Babiuk. 1999. Altering the cellular location of an antigen expressed by a DNA-based vaccine modulates the immune response. J. Virol. 73:10214-10223.

19. Li, H. Z., H. Y. Gang, Q. M. Sun, X. Liu, Y. B. Ma, M. S. Sun, and C. B. Dai. 2004. Production in Pichia pastoris and characterization of genetic engineered chimeric HBV/HEV virus-like particles. Chin. Med. Sci. J. 19:78-83.

20. Livingston, B. D., M. Newman, C. Crimi, D. McKinney, R. Chesnut, and A. Sette. 2001. Optimization of epitope processing enhances immunogenicity of multiepitope DNA vaccines. Vaccine 19:4652-4660.

21. Marsac, D., A.-L. Puaux, Y. Riviere, and M. L. Michel. 2005. In vivo induction of cellular and humoral immune response by hybrid DNA vectors encoding simian/human immunodeficiency virus/hepatitis B surface antigen virus particules in BALB/c and HLA-A2-transgenic mice. Immunobiology 210:305-319.

22. Mathet, V. L., M. Feld, L. Espinola, D. O. Sanchez, V. Ruiz, O. Mando, G. Carballal, J. F. Quarleri, F. D'Mello, C. R. Howard, and J. R. Oubina. 2003. Hepatitis B virus S gene mutants in a patient with chronic active hepatitis with circulating Anti-HBs antibodies. J Med Virol 69:18-26.

23. Michel, M. L., H. L. Davis, M. Schleef, M. Mancini, P. Tiollais, and R. G. Whalen. 1995. DNA-mediated immunization to the hepatitis B surface antigen in mice: aspects of the humoral response mimic hepatitis B viral infection in humans. Proc. Natl. Acad. Sci. USA 92:5307-5311.

24. Michel, M. L., and D. Loirat. 2001. DNA vaccines for prophylactic or therapeutic immunization against hepatitis B. Intervirology 44:78-87.

25. Morel, P. A., D. Falkner, J. Plowey, A. T. Larregina, and L. D. Falo. 2004. DNA immunization: altering the cellular localisation of expressed protein and the immunization route allows manipulation of the immune response. Vaccine 22:447-456.

26. Pajot, A., M. L. Michel, N. Fazilleau, V. Pancre, C. Auriault, D. M. Ojcius, F. A. Lemonnier, and Y. C. Lone. 2004. A mouse model of human adaptive immune functions: HLA-A2.1-/HLA-DR1-transgenic H-2 class I-/class II-knockout mice. Eur. J. Immunol. 34:3060-3069.

27. Pajot, A., V. Pancre, N. Fazilleau, M. L. Michel, G. Angyalosi, D. M. Ojcius, C. Auriault, F. A. Lemonnier, and Y. C. Lone. 2004. Comparison of HLA-

DR1-restricted T cell response induced in HLA-DR1 transgenic mice deficient for murine MHC class II and HLA-DR1 transgenic mice expressing endogenous murine MHC class II molecules. Int. Immunol. 16:1275-1282.

28. Pascolo, S., N. Bervas, J. M. Ure, A. G. Smith, F. A. Lemonnier, and B. Perarnau. 1997. HLA-A2.1-restricted education and cytolytic activity of CD8(+) T lymphocytes from beta2 microglobulin (beta2m) HLA-A2.1 monochain transgenic H-2Db beta2m double knockout mice. J. Exp. Med. 185:2043-2051.

29. Peixoto, L., A. Zavala, H. Romero, and H. Musto. 2003. The strength of translational selection for codon usage varies in the three replicons of Sinorhizobium meliloti. Gene 320:109-116.

30. Pumpens, P., R. Razanskas, P. Pushko, R. Renhof, I. Gusars, D. Skrastina, V. Ose, G. Borisova, I. Sominskaya, I. Petrovskis, J. Jansons, and K. Sasnauskas. 2002. Evaluation of HBs, HBc, and frCP virus-like particles for expression of human papillomavirus 16 E7 oncoprotein epitopes. Intervirology 45:24-32.

31. Rajcani, J., T. Mosko, and I. Rezuchova. 2005. Current developments in viral DNA vaccines: shall they solve the unsolved? Rev. Med. Virol. 15:303-325.

32. Romero, H., A. Zavala, H. Musto, and G. Bernardi. 2003. The influence of translational selection on codon usage in fishes from the family Cyprinidae. Gene 317:141-147.

33. Rudolf, M. P., S. C. Fausch, D. M. Da Silva, and W. M. Kast. 2001. Human dendritic cells are activated by chimeric human papillomavirus type-16 virus-like particles and induce epitope-specific human T cell responses in vitro. J. Immunol. 166:5917-5924.

34. Rudolf, M. P., J. D. Nieland, D. M. DaSilva, M. P. Velders, M. Muller, H. L. Greenstone, J. T. Schiller, and W. M. Kast. 1999. Induction of HPV16 capsid protein-specific human T cell responses by virus-like particles. Biol. Chem. 380:335-340.

35. Ruedl, C., T. Storni, F. Lechner, T. Bachi, and M. F. Bachmann. 2002. Cross-presentation of virus-like particles by skin-derived CD8(-) dendritic cells: a dispensable role for TAP. Eur. J. Immunol. 32:818-825.

36. Schreckenberger, C., and A. M. Kaufmann. 2004. Vaccination strategies for the treatment and prevention of cervical cancer. Curr. Opin. Oncol. 16:485-491.

37. Stern, P. L. 2005. Immune control of human papillomavirus (HPV) associated anogenital disease and potential for vaccination. J. Clin. Virol. 32 Suppl 1:S72-81.

38. Subklewe, M., C. Paludan, M. L. Tsang, K. Mahnke, R. M. Steinman, and C. Munz. 2001. Dendritic cells cross-present latency gene products from Epstein-Barr virus-transformed B cells and expand tumor-reactive CD8(+) killer T cells. J. Exp. Med. 193:405-411.

39. Wang, S., D. J. Farfan-Arribas, S. Shen, T. H. Chou, A. Hirsch, F. He, and S. Lu. 2005. Relative contributions of codon usage, promoter efficiency and leader sequence to the antigen expression and immunogenicity of HIV-1 Env DNA vaccine. Vaccine. In Press.

40. Wang, Y., J. A. Smith, T. Kamradt, M. L. Gefter, and D. L. Perkins. 1992. Silencing of immunodominant epitopes by contiguous sequences in complex synthetic peptides. Cell. Immunol. 143:284-297.

41. Wherry, E. J., M. J. McElhaugh, and L. C. Eisenlohr. 2002. Generation of CD8(+) T cell memory in response to low, high, and excessive levels of epitope. J. Immunol. 168:4455-4461.

42. Yan, M., J. Peng, I. A. Jabbar, X. Liu, L. Filgueira, I. H. Frazer, and R. Thomas. 2004. Despite differences between dendritic cells and Langerhans cells in the mechanism of papillomavirus-like particle antigen uptake, both cells cross-prime T cells. Virology 324:297-310.

43. Van Engelen FA, LMolthoff JW, Conner AJ, Nap JP, Pereira A, Steikema WJ, pBIN-plus : an improved plant transformation vactor based on pBIN19!. Transgenic Res., 1995, 4(4): 288-290.

44. Kyte J, Doolittle R. A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology 1982;157:105-32.

45. Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT. A simple and general method for transferring genes into plants. Science 1985;227:1129-231.

46. Clough SJ, Bent AF. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J 1998;16(6):735-43.

47. Murashige T, Skooog F. A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant 1962;15:493-97.

48. Greco R, Ouwerkerk PB, Taal AJ, Favalli C, Beguiristain T, Puigdomenech P, et al. Early and multiple Ac transpositions in rice suitable for efficient insertional mutagenesis. Plant Mol Biol 2001;46(2):215-27.

49. Greco R, Ouwerkerk PBF, Taal AJC, Sallaud C, Guiderdoni E, Meijer AH, et al. Transcription and somatic transposition of the maize En/Spm transposon system in rice. Molecular Genetics and Genomics 2004;270((6)):514-23.

50. Michel M, Lone Y-C, Centlivre M, Roux P, Wain-Hobson S, Sala M. Optimisation of secretion of recombinant HBsAg virus-like particles: impact on the development of HIV-1/HBV bivalent vaccines. Vaccine 2007;25:1901-11.

51. Jones J, Krag SS, Betenbaugh MJ. Controlling N-linked glycan site occupancy. Biochim Biophys Acta 2005;1726(2):121-37.

52. Kozak M. Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6. Embo J 1997;16(9):2482-92.

CLAIMS

1. A polynucleotide comprising GA1xFlag-M polynucleotide of sequence : CAGGCCATGCAGTGGAACTCCACAcccgggGCTGGAGCAGGAGCTGA TTACAAGGACGACGACGACAAGgaattcCTGCAGGCTAGCAGATCTctc gagCTGAACATG [SEQ ID NO: 1];
or GA3XFlag-M polynucleotide of sequence : CAGGCCATGCAGTGGAACTCCACAcccgggGCTGGAGCAGGAGCTGA CTACAAAGACCACGACGGTGATTATAAAGATCACGACATTGATTACA AGGACGACGACGACAAGgaattcCTGCAGGCTAGCAGATCTctcgagCT GAACATG [SEQ ID NO: 2].

2. A polynucleotide as claimed in claim 1, which comprises an eukaryotic promoter operably linked to the GA1xFlag-M polynucleotide or the GA3xFlag-M polynucleotide.

3. A polynucleotide as claimed in claim 2, which comprises a nucleotide sequence encoding hepatitis B surface antigen protein (HBsAg) devoid of translation initiation ATG and positioned downstream and in frame with the GA1xFlag-M or the GA3xFlag-M polynucleotide sequence.

4. A polynucleotide as claimed in claim 2 or 3, which comprises a polyadenylation sequence operably linked to the other sequences.

5. A polynucleotide as claimed in claim 1 comprising the polynucleotide cloned between HindIII and AvrII restriction sites in pGA1xFlag-M plasmid deposited at the CNCM on December 16, 2005, under the Accession Number I-3543.

6. A polynucleotide as claimed in claim 1 comprising the polynucleotide cloned between HindIII and AvrII restriction sites in pGA3xFlag-M plasmid deposited at the CNCM on December 16, 2005, under the Accession Number I-3545.

7. A polynucleotide as claimed in claim 1 comprising the GA1xFlag-M polynucleotide or the GA3xFlag-M polynucleotide, a eukaryotic promoter sequence, a nucleotide sequence encoding hepatitis B surface antigen protein (HBsAg), and a polyadenylation sequence.

8. A polynucleotide hybridizing under stringent conditions to the polynucleotide as claimed in claim 1 or its complement.

9. A polynucleotide as claimed in any of claims 1 to 10, wherein it further comprises a foreign coding polynucleotide inserted in any of restriction sites of the GA1xFlag-M or GA3xFlag-M polynucleotide and in frame with the ATG at position 7 in the GA1xFlag-M or GA3xFlag-M polynucleotide sequence.

10. A cloning and/or expression vector comprising a polynucleotide as claimed in any one of claims 1-9.

11. A vector as claimed in claim 10, wherein the vector is chosen from

a. vectors for use in eukaryotic expression systems and preferably for mammalian expression systems, such as recombinant poxvirus expression vectors such as vaccinia virus, fowlpox virus, or canarypox virus ; animal DNA viruses such as herpes simplex 1 and 2, varicella zoster, pseudorabies, human cytomegalovirus, murine cytomegalovirus, Esptein-Barr virus, Karposi's sarcoma virus, or murine herpes virus ; animal RNA viruses such as positive-strand RNA viruses such as the picornaviruses as poliovirus, the flaviviruses as hepatitis C virus, or coronaviruses ; lentiviral vectors, adenoviral vectors, and adeno-associated viral vectors ;

b. vectors for expression in yeast cells,

c. vectors for expression in insect cells, such as baculoviruses,

d. vectors for expression in plant cells chosen from *Agrobacterium tumefaciens* Ti-based vectors;

e. plasmid and phage vectors.

12. A host cell comprising a vector as claimed in anyone claims or 11 or a polynucleotide as claimed in anyone of claims 1 to 9.

13. A host cell as claimed in claim 12, wherein the vector comprises an eukaryotic promoter sequence operably linked to a nucleotide sequence encoding HBsAg protein for expression of HBsAg virus-like particles.

14. A host cell as claimed in anyone of claims 12 and 13, wherein the vector or the polynucleotide comprises a nucleotide sequence encoding a HBsAg fusion protein comprising a foreign polypeptide and a HBsAg protein, and wherein the host cell produces HBsAg virus-like particles constituted by said HBsAg fusion protein and HBsAg protein.

15. An *Agrobacterium tumefaciens* cell transformed with a vector of anyone of claims 10 to 12.

16. A plant cell transformed with a vector of anyone of claims 10 to 12.

17. A plant cell of claim 16, wherein HBsAg virus-like particles are assembled in the cell.

18. A plant cell of claim 16, wherein the polynucleotide of any one of claims 1-9 is integrated into the nuclear genome of the plant cell.

19. A plant seed comprising the polynucleotide of anyone of claims 1 to 9.

20. An *in vitro* method for producing HBsAg virus-like particles, wherein the method comprises:

- providing a host cell as claimed in any one of claims 12 to 14; and
- expressing the fusion proteins and HBsAg proteins under conditions in which the proteins assemble into virus-like particles, which are released from the host cell into extracellular space.

21. A method as claimed in claim 20, which comprises recovering the virus-like particles.

22. An *in vivo* method for producing HBsAg virus-like particle, wherein the method comprises :

- providing a transgenic plant seed according to claim19,
- growing the transgenic plant from the plant seed under conditions effective to produce the HBsAg VLP.

23. HBsAg virus like particules susceptible to be produced by the method as claimed in anyone claims 20 to 22.

24. A virus like particle encoding by a polynucleotide as claimed in any one of claims 1-9.

25. A composition comprising a virus-like particle as claimed in anyone of claims 23 or 24 and a pharmaceutically acceptable carrier.

26. A vaccine comprising a composition of claim 25.

27. Use of the VLP as claimed in claim 23 or 24 in the preparation of a composition intended to immunize a human or an animal.

28. A method for preparing a HBsAg virus-like particle, wherein the method comprises :

- providing a host cell as claimed in any one of claims 12 to 14 or an *Agrobacterium tumefaciens* cell of claim 15 or a plant cell of anyone of claims 16 to 18;
- expressing a tagged HBsAg fusion protein and an HBsAg protein under conditions in which the proteins assemble into virus-like particles; and
- separating the virus-like particle bearing tagged HBsAg fusion proteins from the host cells by capture with Flag-M antibodies and/or HBsAg antibodies.

29. An expression vector of claim 10, wherein it is selected from pGA1xFlag-M (CNCM No. I-3543), pGA3xFlag-M (CNCM No. I-3545), pGA1xFlag-M pol.opt (CNCM No. I-3544), pGA3xFlag-M pol.opt (CNCM No. I-3546), pGA1xFlag-M.pol1A2 (CNCM No. I-3579), pGA1xFlag-M.pol2A2 (CNCM No. I-3580), pGA1xFlag-M.pol1B7 (CNCM No. I-3581), and pGA1xFlag-M.pol2B7 (CNCM No. I-3582).

30. A polynucleotide comprising the sequence between HindIII and AvrII restriction sites of an expression vector as claimed in claim 29.

31. A polypeptide encoded by a polynucleotide as claimed in any of claims 1 to 9 or by a vector according to anyone claim 10 to 11.

32. An immunogenic composition comprising the plant cell of anyone of claims 16 to 18.

33. The immunogenic composition of claim 32 wherein the plant cell is present in plant tissue selected from the group consisting of a fruit, leaf, plant organ, seed.

34. The immunogenic composition of claim 33 comprising a crude extract of the plant tissue.

35. A method of eliciting an immune response in mammal comprising the step of administering the composition of claim of claim 32 to a mammal, wherein an immune response is elicited.

36. The method of claim 35 wherein the composition is administered orally.

37. A transgenic plant wherein the plant comprises plant cells of anyone of claims 16 to 18.

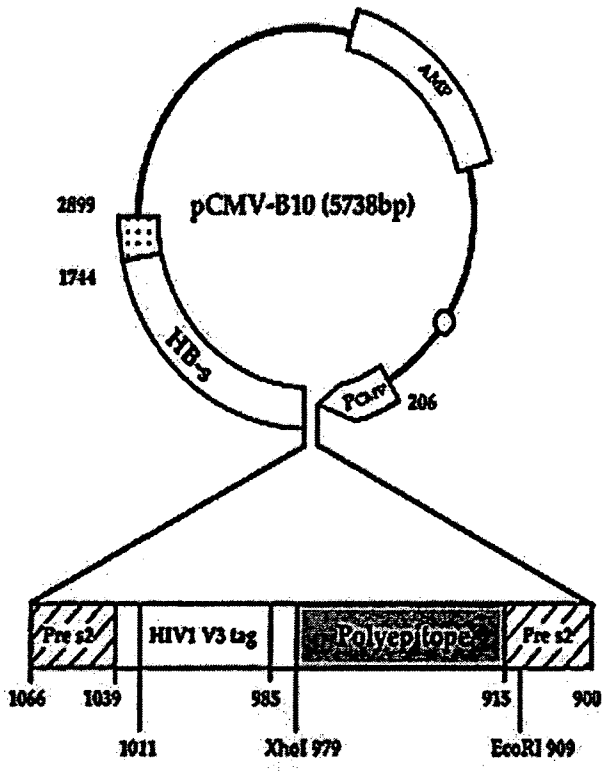

AMP
pCMV-B10 (5738bp)
2899
1744
HB-3
206
Pre s2
HIV1 V3 tag
Polycptope
Pre s2
1066
1039
985
915
900
1011
XhoI 979
EcoRI 909

Figure 1A

pCMV-basic complete nucleotide sequence (HindIII and AvrII restriction sites are highlighted in bold)

AACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCGACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAG
TACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGC
TACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATAC
GCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACT
TACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAC
TTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGA
CGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT
CGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGA
CGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCG
TGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAG
GGAGACCC**AAGCTT**CAGGCC**ATG**CAGTGGAATTCAGCTAGCGATATCCCCGGGCTCGAGCATATAGGACCAGGGAGAGCATTTGTGGTACCTC
HindIII preS2 polylinker

TAGAAGAAGCTTGGGATCCGTCGAGGATTGGGGACCCTGCGCTGAAC**ATG**GAGAACATCACATCAGGATT**CCTAGG**ACCCCTTCTCGTGTTAC
HBsAg AvrII

AGGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCG
TGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGC
GGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTC
TAATTCCAGGATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTA
CCAAACCTTCGGACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCT
CCTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGG
GGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCCTAACAAAACAAA
GAGATGGGGTTACTCTCTAAATTTTATGGGTTATGTCATTGGATGTTATGGGTCCTTGCCACAAGAACACATCATACAAAAAATCAAAGAATG
TTTTAGAAAACTTCCTATTAACAGGCCTATTGATTGGAAAGTATGTCAACGAATTGTGGGTCTTTTGGGTTTTGCTGCCCCTTTTACACAATG
TGGTTATCCTGCGTTGATGCCTTTGTATGCATGTATTCAATCTAAGCAGGCTTTCACTTTCTCGCCAACTTACAAGGCCTTTCTGTGTAAACA
ATACCTGAACCTTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGTCATGGG
CCATCAGCGCATGCGTGGAACCTTTTCGGCTCCTCTGCCGATCCATACTGCGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGC
AAACATTATCGGGACTGATAACTCTGTTGTCCTATCCCGCAAATATACATCGTTTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCG
CGGGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCTGCGGACGACCCTTCTCGGGGTCGCTTGGGACTCTCTCGTCCCCTTCTCCGTCT
GCCGTTCCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACC
TCTGCACGTCGCATGGAGACCACCGTGAACGCCCACCAAATATTGCCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACG
ACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAAGACTGGGAGGAGTTGGGGGAGGAGATTAGGTTAAAGGTCTTTGTACTAGGAGGC
TGTAGGCATAAATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGTCCTACTGTTCAAGCCTCCA
AGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACT
TCTTTCCTTCAGTACGAGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGC
GCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCG
ACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACC
GCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGG
GTTCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGT
TTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGG
TTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATC
TTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAAT
TCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCC
CGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC
TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATC
AGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCT
CCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGG
AAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATG
CTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTT
ATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTA
TGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT
TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCG
CAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCAT
GAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGA
CAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAAC
TACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCA
GCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAG
TTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGG
TTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTT
GGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTA
CTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAAC
TTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG
TGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAG
GGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT
TGAATGTATTTAGAAAAATA

Fig.1B

◆ pGA1xFlag-M nucleic acid sequence:

GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTG
CTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC
TTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGT
AAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT
ATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCaagctt**CAGGCC**
ATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTGATTACAAGGACGACGACGACAAGgaattcCTGCAGGCTAGCAGATCTctc
gagCTGAACATGGAGAACATCACATCAGGATTcctaggACCCCTTCTCGTGTTACAGGCGGGGTTTTTCTTGTTGACAAGAATCCTCACA
ATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACC
TCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTTTTATCATCTTCCTCTTCATCCTG
CTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGGATCCTCAACAACCAGC
ACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGGACGGAAATTGC
ACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG
CCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGC
ATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCCTAACAAAACAAAGAGATGGGGTTACT
CTCTAAATTTTATGGGTTATGTCATTGGATGTTATGGGTCCTTGCCACAAGAACACATCATACAAAAAATCAAAGAATGTTTTAGAAAAC
TTCCTATTAACAGGCCTATTGATTGGAAAGTATGTCAACGAATTGTGGGTCTTTTGGGTTTTGCTGCCCCTTTTACACAATGTGGTTATC
CTGCGTTGATGCCTTTGTATGCATGTATTCAATCTAAGCAGGCTTTCACTTTCTCGCCAACTTACAAGGCCTTTCTGTGTAAACAATACC
TGAACCTTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGTCATGGGCC
ATCAGCGCATGCGTGGAACCTTTTCGGCTCCTCTGCCGATCCATACTGCGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAG
CAAACATTATCGGGACTGATAACTCTGTTGTCCTATCCCGCAAATATACATCGTTTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCC
TGCGCGGGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCTGCGGACGACCCTTCTCGGGGTCGCTTGGGACTCTCTCGTCCCCTTC
TCCGTCTGCCGTTCCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACT
TCGCTTCACCTCTGCACGTCGCATGGAGACCACCGTGAACGCCCACCAAATATTGCCCAAGGTCTTACATAAGAGGACTCTTGGACTCTC
AGCAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAAGACTGGGAGGAGTTGGGGGAGGAGATTAGGTTAAAGGT
CTTTGTACTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGT
CCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTA
CTCTCGTTTTTGCCTTCTGACTTCTTTCCTTCAGTACGAGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGT
TCCCTTTAGTGAGGGTTAATTGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTG
GAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATT
GCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGC
CTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCA
CCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGT
TCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTT
CACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAAT
CATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG
GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAAT
GAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC
TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAG
GCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC
GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGA
CCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGG
TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGT
CCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT
TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAG
TTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT
CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAA
AAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC
AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGA
TACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC
CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTA
GTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCT
CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAA
GTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGA
CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGC
CACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT
CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATG
CCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT
GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACG
TC

Figure 2

◆ pGA3xFlag-M nucleic acid sequence

GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTG
CTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC
TTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGT
AAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT
ATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCaagctt**CAGGCC**
ATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTGACTACAAAGACCACGACGGTGATTATAAAGATCACGACATTGATTACAAG
GACGACGACGACAAGgaattc**CTGCAGGCTAGCAGATCT**ctcgag**CTGAACATGG**GAGAACATCACATCAGGATTcctaggACCCCTTCTC
GTGTTACAGGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGG
GGAACTACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGC
TGGATGTGTCTGCGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATG
TTGCCCGTTTGTCCTCTAATTCCAGGATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATG
TATCCCTCCTGTTGCTGTACCAAACCTTCGGACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGG
GAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCA
GTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGG
GTATACATTTAAACCCTAACAAAACAAAGAGATGGGGTTACTCTCTAAATTTTATGGGTTATGTCATTGGATGTTATGGGTCCTTGCCAC
AAGAACACATCATACAAAAAATCAAAGAATGTTTTAGAAAACTTCCTATTAACAGGCCTATTGATTGGAAAGTATGTCAACGAATTGTGG
GTCTTTTGGGTTTTGCTGCCCCTTTTACACAATGTGGTTATCCTGCGTTGATGCCTTTGTATGCATGTATTCAATCTAAGCAGGCTTTCA
CTTTCTCGCCAACTTACAAG
GCCTTTCTGTGTAAACAATACCTGAACCTTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACT
GGCTGGGGCTTGGTCATGGGCCATCAGCGCATGCGTGGAACCTTTTCGGCTCCTCTGCCGATCCATACTGCGGAACTCCTAGCCGCTTGT
TTTGCTCGCAGCAGGTCTGGAGCAAACATTATCGGGACTGATAACTCTGTTGTCCTATCCCGCAAATATACATCGTTTCCATGGCTGCTA
GGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCTGCGGACGACCCTTCTCGGGGTCGC
TTGGGACTCTCTCGTCCCCTTCTCCGTCTGCCGTTCCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCCGTCTGTGCCTTCT
CATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATGGAGACCACCGTGAACGCCCACCAAATATTGCCCAAGGTCTTAC
ATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAAGACTGGGAGGAGTTGG
GGGAGGAGATTAGGTTAAAGGTCTTTGTACTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACCTCTG
CCTAATCATCTCTTGTTCATGTCCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGACCCTTATAAAGA
ATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTTCCTTCAGTACGAGATCCACTAGTTCTAGAGCGGCCGCCACC
GCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGC
CTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACAT
AGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTC
GCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCTG
CCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAG
CGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAAT
TTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACC
TCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGA
AGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAAC
CTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGAC
TCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCA
GGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCT
GACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCC
CTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCA
CGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTA
TCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGG
TATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAG
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAG
ATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGG
ATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAG
TAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTC
CCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCA
GATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGT
TGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCG
TTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTC
GGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCA
TCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCG
TCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATC
TTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGA
GCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTAT
TGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTT
CCCCGAAAAGTGCCACCTGACGTC

Figure 3

Nucleic acid sequence of pGA1xFlag-Mpol.opt (a)

ATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTGATTACAAGGACGACGACGAC
AAGgaattcgaCTACTTGAAAGAGCCAGTTCATGGGGTGAGAGCCAAGACCTACCTGAAT
GCATGGGTGAAAGTTGTCAGAGACACCGCAGTGCTGGATGTGGGGGATGCCTACTTCTCA
GTGAGAGCTAAGACTTATCTGGTCAAACTCTGGTACCAGTTGAGGGCTGACACTCGTCTT
TACAACACTGTGGCCACCCTTAGGACCAAGGCTCTTCTGGACACTGGAGCAGATGACACT
GTGAGGGCTAAGACCCTGCTGTGGAAGGGAGAGGGAGCAGTTAGGACTGATGCTTACATC
TACCAGTATATGGATGACCTTAGActcgagCTGAAC**ATG**

Figure 4

Nucleic acid sequence of GA3xFlag-Mpol.opt
CAGGCCATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTGACTACAAAGACCAC
GACGGTGATTATAAAGATCACGACATTGATTACAAGGACGACGACGACAAGgaattcGAC
TACTTGAAAGAGCCAGTTCATGGGGTGAGAGCCAAGACCTACCTGAATGCATGGGTGAAA
GTTGTCAGAGACACCGCAGTGCTGGATGTGGGGGATGCCTACTTCTCAGTGAGAGCTAAG
ACTTATCTGGTCAAACTCTGGTACCAGTTGAGGGCTGACACTCGTCTTTACAACACTGTG
GCCACCCTTAGGACCAAGGCTCTTCTGGACACTGGAGCAGATGACACTGTGAGGGCTAAG
ACCCTGCTGTGGAAGGGAGAGGGAGCAGTTAGGACTGATGCTTACATCTACCAGTATATG
GATGACCTTAGActcgagCTGAAC**ATG**

Figure 5

◆ pGA1xFlag-Mpol.opt nucleic acid sequence:

GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTG
CTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC
TTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGT
AAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT
ATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCaagcttCAGGCC
ATGCAGTGGAACTCCACAcccgggGCTGGAGCAGGAGCTGATTACAAGGACGACGACGACAAGgaattcG**ACTACTTGAAAGAGCCAGTT**
CATGGGGTGAGAGCCAAGACCTACCTGAATGCATGGGTGAAAGTTGTCAGAGACACCGCAGTGCTGGATGTGGGGGATGCCTACTTCTCA
GTGAGAGCTAAGACTTATCTGGTCAAACTCTGGTACCAGTTGAGGGCTGACACTCGTCTTTACAACACTGTGGCCACCCTTAGGACCAAG
GCTCTTCTGGACACTGGAGCAGATGACACTGTGAGGGCTAAGACCCTGCTGTGGAAGGGAGAGGGAGCAGTTAGGACTGATGCTTACATC
TACCAGTATATGGATGACCTTAGActcgagCTGAACATGGAGAACATCACATCAGGATTcctaggACCCCTTCTCGTGTTACAGGCGGGG
TTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGTGT
CTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGG
CGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCT
CTAATTCCAGGATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGC
TGTACCAAACCTTCGGACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCC
CGTTTCTCCTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATG
TGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACC
CTAACAAAACAAAGAGATGGGGTTACTCTCTAAATTTTATGGGTTATGTCATTGGATGTTATGGGTCCTTGCCACAAGAACACATCATAC
AAAAAATCAAAGAATGTTTTAGAAAACTTCCTATTAACAGGCCTATTGATTGGAAAGTATGTCAACGAATTGTGGGTCTTTTGGGTTTTG
CTGCCCCTTTTACACAATGTGGTTATCCTGCGTTGATGCCTTTGTATGCATGTATTCAATCTAAGCAGGCTTTCACTTTCTCGCCAACTT
ACAAGGCCTTTCTGTGTAAACAATACCTGAACCTTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCC
CCACTGGCTGGGGCTTGGTCATGGGCCATCAGCGCATGCGTGGAACCTTTTCGGCTCCTCTGCCGATCCATACTGCGGAACTCCTAGCCG
CTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACATTATCGGGACTGATAACTCTGTTGTCCTATCCCGCAAATATACATCGTTTCCATGGC
TGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCTGCGGACGACCCTTCTCGGG
GTCGCTTGGGACTCTCTCGTCCCCTTCTCCGTCTGCCGTTCCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCCGTCTGTGC
CTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATGGAGACCACCGTGAACGCCCACCAAATATTGCCCAAGGT
CTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAAGACTGGGAGGA
GTTGGGGGAGGAGATTAGGTTAAAGGTCTTTGTACTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCAGCACCATGCAACTTTTTCAC
CTCTGCCTAATCATCTCTTGTTCATGTCCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGACCCTTAT
AAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTTCCTTCAGTACGAGATCCACTAGTTCTAGAGCGGCCG
CCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGC
GATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAG
GACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCC
GATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCA
ACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCC
TCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCA
CAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGT
CGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAG
CCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGG
GAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCA
CTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATA
ACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCC
CCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAA
GCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAAT
GCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCG
CCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG
CGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGC
TGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGC
AGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTT
AAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATAT
ATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCT
GACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGG
CTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTA
ATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCT
CGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCT
CCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCA
TGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCC
CGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAA
GGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTG
GGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAAT
ATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCA
CATTTCCCCGAAAAGTGCCACCTGACGTC

Figure 6

◆ pGA3xFlag-Mpol.opt nucleic acid sequence:

GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTG
CTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC
TTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGT
AAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT
ATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCaagcttCAGGCC
ATGCAGTGGAACTCCACAcccgggGCTGGAGCAGGAGCTGACTACAAAGACCACGACGGTGATTATAAAGATCACGACATTGATTACAAG
GACGACGACGACAAGgaattcGA**CTACTTGAAAGAGCCAGTTCATGGGGTGAGAGCCAAGACCTACCTGAATGCATGGGTGAAAGTTGTC**
AGAGACACCGCAGTGCTGGATGTGGGGGATGCCTACTTCTCAGTGAGAGCTAAGACTTATCTGGTCAAACTCTGGTACCAGTTGAGGGCT
GACACTCGTCTTTACAACACTGTGGCCACCCTTAGGACCAAGGCTCTTCTGGACACTGGAGCAGATGACACTGTGAGGGCTAAGACCCTG
CTGTGGAAGGGAGAGGGAGCAGTTAGGACTGATGCTTACATCTACCAGTATATGGATGACCTTAGActcgagCTGAACATGGAGAACATC
ACATCAGGATTcctaggACCCCTTCTCGTGTTACAGGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCG
TGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCT
TGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTG
TTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGGATCCTCAACAACCAGCACGGGACCATGCCGGACCTGC
ATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGGACGGAAATTGCACCTGTATTCCCATCCCATCA
TCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTA
GGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCG
CTGTTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCCTAACAAAACAAAGAGATGGGGTTACTCTCTAAATTTTATGGGTTATG
TCATTGGATGTTATGGGTCCTTGCCACAAGAACACATCATACAAAAAATCAAAGAATGTTTTAGAAAACTTCCTATTAACAGGCCTATTG
ATTGGAAAGTATGTCAACGAATTGTGGGTCTTTTGGGTTTTGCTGCCCCTTTTACACAATGTGGTTATCCTGCGTTGATGCCTTTGTATG
CATGTATTCAATCTAAGCAGGCTTTCACTTTCTCGCCAACTTACAAGGCCTTTCTGTGTAAACAATACCTGAACCTTTACCCCGTTGCCC
GGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGTCATGGGCCATCAGCGCATGCGTGGAACCT
TTTCGGCTCCTCTGCCGATCCATACTGCGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACATTATCGGGACTGATA
ACTCTGTTGTCCTATCCCGCAAATATACATCGTTTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTT
ACGTCCCGTCGGCGCTGAATCCTGCGGACGACCCTTCTCGGGGTCGCTTGGGACTCTCTCGTCCCCTTCTCCGTCTGCCGTTCCGACCGA
CCACGGGGCGCACCTCTCTTTACGCGGACTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCG
CATGGAGACCACCGTGAACGCCCACCAAATATTGCCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGACCGACCT
TGAGGCATACTTCAAAGACTGTTTGTTTAAAGACTGGGAGGAGTTGGGGGAGGAGATTAGGTTAAAGGTCTTTGTACTAGGAGGCTGTAG
GCATAAATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGTCCTACTGTTCAAGCCTCCAAG
CTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGAC
TTCTTTCCTTCAGTACGAGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATT
GCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGA
TTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAA
TGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA
GCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGT
TGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGT
TTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTT
TGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTG
TGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAAC
TCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGA
GAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC
ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACC
GTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA
ACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACC
TGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC
TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACT
TATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT
ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCA
AACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT
TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCC
TTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC
CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTG
GCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA
GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGC
GCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC
GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTAT
CACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT
CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAG
TGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC
CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGG
CGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT
TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

Figure 7

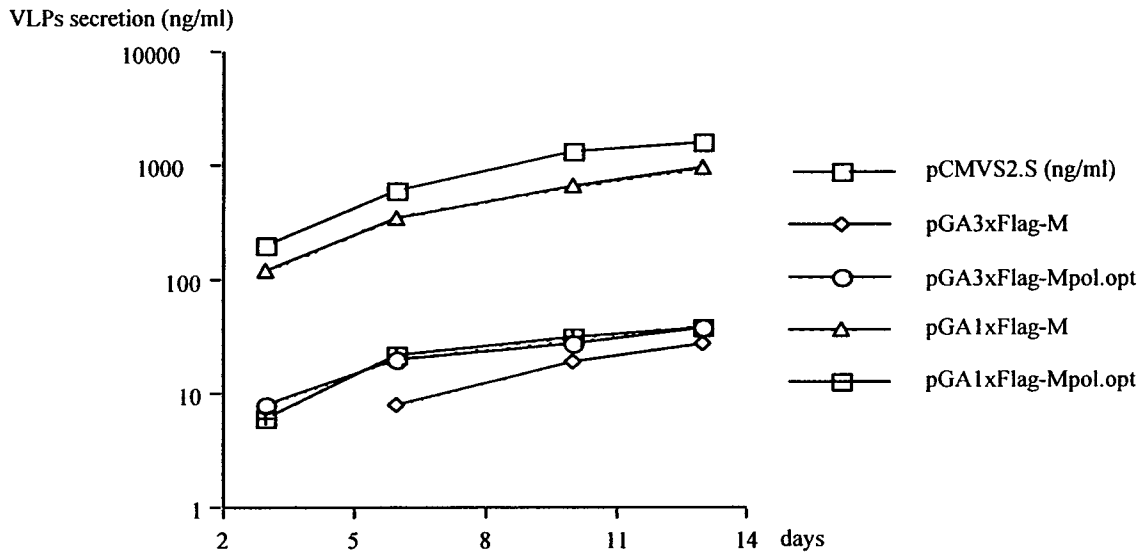

VLPs secretion (ng/ml)
10000
1000
100
10
1
2
5
8
11
14
days
pCMVS2.S (ng/ml)
pGA3xFlag-M
pGA3xFlag-Mpol.opt
pGA1xFlag-M
pGA1xFlag-Mpol.opt

Figure 8

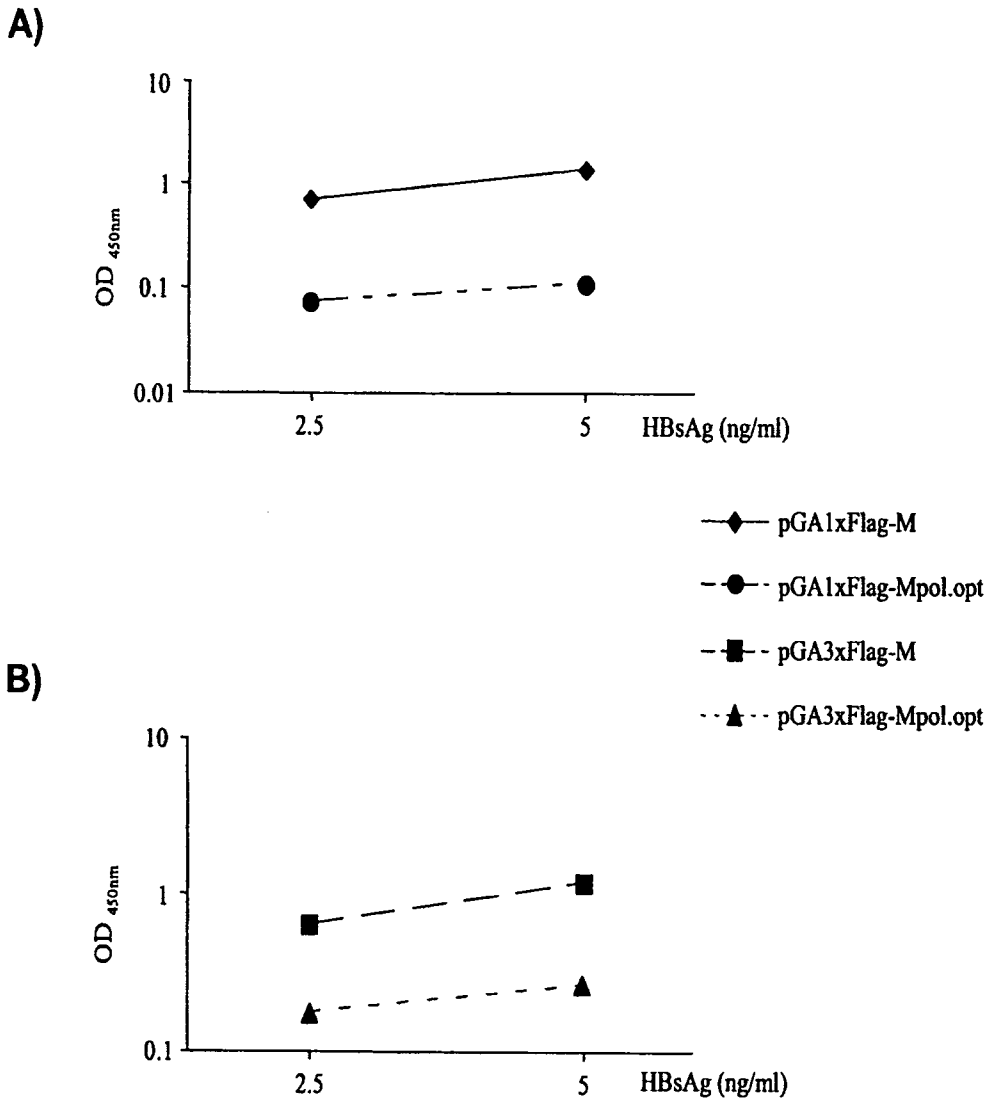

A)
10
1
0.1
0.01
OD 450nm
2.5
5
HBsAg (ng/ml)
pGA1xFlag-M
pGA1xFlag-Mpol.opt
pGA3xFlag-M
pGA3xFlag-Mpol.opt
B)
10
1
0.1
OD 450nm
2.5
5
HBsAg (ng/ml)

Figure 9

pGA1xFlag-M.pol1A2 nucleic acid sequence (in bold: pol1A2 polyepitope)

GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTG
CTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC
TTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGT
AAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT
ATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCaagcttCAGGCC
ATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTGATTACAAGGACGACGACGACAAGgaattc**GTGCTGGATGTGGGAGATGCC**
TACTTCTCAGTGAGAGCTGACACCTACCTGAATGCCTGGGTGAAGGTGGTCAGAGCCAAGACCTACCTGGTGAAGCTGTGGTACCAGCTG
AGGACAGATGCCTCCCTGGTGAAGCATCACATGTATGTGAGAGACACAGCCTACATCTACCAGTACATGGATGACCTGAGActcgagCTG
AACATGGAGAACATCACATCAGGATTcctaggACCCCTTCTCGTGTTACAGGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCG
CAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAAT
CACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTA
TGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGGATCCTCAACAACCAGCACGGGA
CCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGGACGGAAATTGCACCTGT
ATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTGCCATTT
GTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTG
AGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCCTAACAAAACAAAGAGATGGGGTTACTCTCTAA
ATTTTATGGGTTATGTCATTGGATGTTATGGGTCCTTGCCACAAGAACACATCATACAAAAAATCAAAGAATGTTTTAGAAAACTTCCTA
TTAACAGGCCTATTGATTGGAAAGTATGTCAACGAATTGTGGGTCTTTTGGGTTTTGCTGCCCCTTTTACACAATGTGGTTATCCTGCGT
TGATGCCTTTGTATGCATGTATTCAATCTAAGCAGGCTTTCACTTTCTCGCCAACTTACAAGGCCTTTCTGTGTAAACAATACCTGAACC
TTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGTCATGGGCCATCAGC
GCATGCGTGGAACCTTTTCGGCTCCTCTGCCGATCCATACTGCGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACA
TTATCGGGACTGATAACTCTGTTGTCCTATCCCGCAAATATACATCGTTTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCG
GGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCTGCGGACGACCCTTCTCGGGGTCGCTTGGGACTCTCTCGTCCCCTTCTCCGTC
TGCCGTTCCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTT
CACCTCTGCACGTCGCATGGAGACCACCGTGAACGCCCACCAAATATTGCCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAAT
GTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAAGACTGGGAGGAGTTGGGGGAGGAGATTAGGTTAAAGGTCTTTGT
ACTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGTCCTACT
GTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCG
TTTTTGCCTTCTGACTTCTTTCCTTCAGTACGAGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTT
TAGTGAGGGTTAATTGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAAT
GGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAA
GAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTT
GACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCG
CCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCG
CCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGC
ATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGT
CATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCT
AATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCG
GCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGC
GAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGC
AAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAA
GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGC
CGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGG
TCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACC
CGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGA
AGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTA
GCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAG
AAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGA
TCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCT
TAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGG
AGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCG
GAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGC
CAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTT
CCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGT
TGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTG
AGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATA
GCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGT
AACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAA
AAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCA
TGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

Figure 10

pGA1xFlag-M.pol2A2 nucleic acid sequence (in bold: pol2A2 polyepitope)

GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTG
CTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC
TTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGT
AAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT
ATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCaagcttCAGGCC
ATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTGATTACAAGGACGACGACGACAAGgaattc**CTGCTTGACACAGGAGCTGAT**
GACACAGTGAGGACAGATGCCAGCCTGTATAACACAGTGGCCACCCTGAGAGCTGACACCTACCTGAAGGAGCCTGTGCATGGAGTGAGA
GCTAAGACCCTCCTGTGGAAGGGAGAGGGAGCAGTGAGAACCAAGGCAGTGCTGGCTGAGGCCATGTCCCAGGTGAGActcgagCTGAAC
ATGGAGAACATCACATCAGGATTcctaggACCCCTTCTCGTGTTACAGGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAG
AGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCAC
TCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGC
CTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGGATCCTCAACAACCAGCACGGGACCA
TGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGGACGGAAATTGCACCTGTATT
CCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTGCCATTTGTT
CAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGT
CCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCCTAACAAAACAAAGAGATGGGGTTACTCTCTAAATT
TTATGGGTTATGTCATTGGATGTTATGGGTCCTTGCCACAAGAACACATCATACAAAAAATCAAAGAATGTTTTTAGAAAACTTCCTATTA
ACAGGCCTATTGATTGGAAAGTATGTCAACGAATTGTGGGTCTTTTGGGTTTTGCTGCCCCTTTTACACAATGTGGTTATCCTGCGTTGA
TGCCTTTGTATGCATGTATTCAATCTAAGCAGGCTTTCACTTTCTCGCCAACTTACAAGGCCTTTCTGTGTAAACAATACCTGAACCTTT
ACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGTCATGGGCCATCAGCGCA
TGCGTGGAACCTTTTCGGCTCCTCTGCCGATCCATACTGCGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACATTA
TCGGGACTGATAACTCTGTTGTCCTATCCCGCAAATATACATCGTTTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGA
CGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCTGCGGACGACCCTTCTCGGGGTCGCTTGGGACTCTCTCGTCCCCTTCTCCGTCTGC
CGTTCCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCAC
CTCTGCACGTCGCATGGAGACCACCGTGAACGCCCACCAAATATTGCCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTC
AACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAAGACTGGGAGGAGTTGGGGGAGGAGATTAGGTTAAAGGTCTTTGTACT
AGGAGGCTGTAGGCATAAATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGTCCTACTGTT
CAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTT
TTGCCTTCTGACTTCTTTCCTTCAGTACGAGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAG
TGAGGGTTAATTGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGC
CGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAG
CTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGAC
GAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCT
TCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCC
ACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATT
CTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCAT
AGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAT
GAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCC
AACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAG
CGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAA
AGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTC
AGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGC
TTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCG
TTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGG
TAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT
GGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCT
CTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAG
ATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCT
TCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAA
TCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGG
GCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAA
GGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAG
TTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCC
AACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGG
CCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGT
ACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCA
GAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAC
CCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAA
AGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGA
GCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

Figure 11

pGA3xFlag-M.pol1A2 nucleic acid sequence (in bold: pol1A2 polyepitope)

GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTG
CTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC
TTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGT
AAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT
ATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCaagcttCAGGCC
ATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTGACTACAAAGACCACGACGGTGATTATAAAGATCACGACATTGATTACAAG
GACGACGACGACAAGgaattc**GTGCTGGATGTGGGAGATGCCTACTTCTCAGTGAGAGCTGACACCTACCTGAATGCCTGGGTGAAGGTG**
GTCAGAGCCAAGACCTACCTGGTGAAGCTGTGGTACCAGCTGAGGACAGATGCCTCCCTGGTGAAGCATCACATGTATGTGAGAGACACA
GCCTACATCTACCAGTACATGGATGACCTGAGActcgagCTGAACATGGAGAACATCACATCAGGATTcctaggACCCCTTCTCGTGTTA
CAGGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACT
ACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATG
TGTCTGCGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCC
GTTTGTCCTCTAATTCCAGGATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCC
TCCTGTTGCTGTACCAAACCTTCGGACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGG
GCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATA
TGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATAC
ATTTAAACCCTAACAAAACAAAGAGATGGGGTTACTCTCTAAATTTTATGGGTTATGTCATTGGATGTTATGGGTCCTTGCCACAAGAAC
ACATCATACAAAAAATCAAAGAATGTTTTAGAAAACTTCCTATTAACAGGCCTATTGATTGGAAAGTATGTCAACGAATTGTGGGTCTTT
TGGGTTTTGCTGCCCCTTTTACACAATGTGGTTATCCTGCGTTGATGCCTTTGTATGCATGTATTCAATCTAAGCAGGCTTTCACTTTCT
CGCCAACTTACAAGGCCTTTCTGTGTAAACAATACCTGAACCTTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTG
ACGCAACCCCCACTGGCTGGGGCTTGGTCATGGGCCATCAGCGCATGCGTGGAACCTTTTCGGCTCCTCTGCCGATCCATACTGCGGAAC
TCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACATTATCGGGACTGATAACTCTGTTGTCCTATCCCGCAAATATACATCGT
TTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCTGCGGACGACC
CTTCTCGGGGTCGCTTGGGACTCTCTCGTCCCCTTCTCCGTCTGCCGTTCCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCC
CGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATGGAGACCACCGTGAACGCCCACCAAATATT
GCCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAAGA
CTGGGAGGAGTTGGGGGAGGAGATTAGGTTAAAGGTCTTTGTACTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCAGCACCATGCAA
CTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGTCCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATC
GACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTTCCTTCAGTACGAGATCCACTAGTTCTA
GAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCATGCCCGACGGCGAGGATCTCGTCGTG
ACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGAC
CGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATC
GCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAG
CGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCT
GGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCA
ATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCT
GTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACA
ACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTT
TCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTT
CCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAAT
CAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATA
GGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTC
CCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGC
TTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCG
ACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA
TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCT
GCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTG
TTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAA
ACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCT
AAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCA
TAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCAC
GCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCC
AGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGG
TGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAG
CGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTC
TTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTT
GCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAA
AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCA
GCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCC
TTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGG
TTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

Figure 12

pGA3xFlag-M.pol2A2 nucleic acid sequence (in bold: pol2A2 polyepitope)

GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTG
CTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC
TTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGT
AAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT
ATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCaagcttCAGGCC
ATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTGACTACAAAGACCACGACGGTGATTATAAAGATCACGACATTGATTACAAG
GACGACGACGACAAGgaattc**CTGCTTGACACAGGAGCTGATGACACAGTGAGGACAGATGCCAGCCTGTATAACACAGTGGCCACCCTG**
AGAGCTGACACCTACCTGAAGGAGCCTGTGCATGGAGTGAGAGCTAAGACCCTCCTGTGGAAGGGAGAGGGAGCAGTGAGAACCAAGGCA
GTGCTGGCTGAGGCCATGTCCCAGGTGAGActcgagCTGAACATGGAGAACATCACATCAGGATTcctaggACCCCTTCTCGTGTTACAG
GCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACC
GTGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGT
CTGCGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTT
TGTCCTCTAATTCCAGGATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCC
TGTTGCTGTACCAAACCTTCGGACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCC
TCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGG
ATGATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATACATT
TAAACCCTAACAAAACAAAGAGATGGGGTTACTCTCTAAATTTTATGGGTTATGTCATTGGATGTTATGGGTCCTTGCCACAAGAACACA
TCATACAAAAAATCAAAGAATGTTTTAGAAAACTTCCTATTAACAGGCCTATTGATTGGAAAGTATGTCAACGAATTGTGGGTCTTTTGG
GTTTTGCTGCCCCTTTTACACAATGTGGTTATCCTGCGTTGATGCCTTTGTATGCATGTATTCAATCTAAGCAGGCTTTCACTTTCTCGC
CAACTTACAAGGCCTTTCTGTGTAAACAATACCTGAACCTTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACG
CAACCCCCACTGGCTGGGGCTTGGTCATGGGCCATCAGCGCATGCGTGGAACCTTTTCGGCTCCTCTGCCGATCCATACTGCGGAACTCC
TAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACATTATCGGGACTGATAACTCTGTTGTCCTATCCCGCAAATATACATCGTTTC
CATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCTGCGGACGACCCTT
CTCGGGGTCGCTTGGGACTCTCTCGTCCCCTTCTCCGTCTGCCGTTCCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCCGT
CTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATGGAGACCACCGTGAACGCCCACCAAATATTGCC
CAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAAGACTG
GGAGGAGTTGGGGGAGGAGATTAGGTTAAAGGTCTTTGTACTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCAGCACCATGCAACTT
TTTCACCTCTGCCTAATCATCTCTTGTTCATGTCCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGAC
CCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTTCCTTCAGTACGAGATCCACTAGTTCTAGAG
CGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACC
CATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGC
TATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCC
GCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGA
CGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGA
TGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATA
GCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTA
TACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACA
TACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCC
AGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCT
CGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAG
GGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGC
TCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCC
CTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTT
CTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACC
GCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTA
GCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCG
CTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTT
GCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACT
CACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAA
GTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAG
TTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCT
CACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGT
CTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGT
CACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGG
TTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTA
CTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCT
CTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAAC
TCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCG
TTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTT
TTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTC
CGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

Figure 13

pGA1xFlag-M.pol1B7 nucleic acid sequence (in bold: pol1B7 polyepitope)

GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGC
TTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTT
AGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTA
CGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCG
CCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACT
GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCC
AGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACAT
CAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAA
CGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTC
TCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCaagcttCAGGCCATGCAGTGGA
ACTCCACACCCGGGGCTGGAGCAGGAGCTGATTACAAGGACGACGACGACAAGgaattc**TCCCCTAGGACCCTGAATGCCTGGGTGAGAGC**
TAAGACCAGACCTAACAATAACACAAGGAAGTCCATCAGAGACACAGCCTTCCCTGTGAGACCACAGGTGCCTCTGAGGAGAACCAAGGCC
CACCCTGTGCATGCTGGCCCTATTGCCAGAGCTGATACAGCACCCACTAAGGCCAAAAGGAGAGTGGTCAGGctcgagCTGAACATGGAGA
ACATCACATCAGGATTcctaggACCCCTTCTCGTGTTACAGGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGA
CTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACC
TCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCT
TGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGGATCCTCAACAACCAGCACGGGACCATGCCGGACCTG
CATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGGACGGAAATTGCACCTGTATTCCCATCCCATCA
TCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAG
GGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCT
GTTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCCTAACAAAACAAAGAGATGGGGTTACTCTCTAAATTTTATGGGTTATGTCA
TTGGATGTTATGGGTCCTTGCCACAAGAACACATCATACAAAAAATCAAAGAATGTTTTAGAAAACTTCCTATTAACAGGCCTATTGATTG
GAAAGTATGTCAACGAATTGTGGGTCTTTTGGGTTTTGCTGCCCCTTTTACACAATGTGGTTATCCTGCGTTGATGCCTTTGTATGCATGT
ATTCAATCTAAGCAGGCTTTCACTTTCTCGCCAACTTACAAGGCCTTTCTGTGTAAACAATACCTGAACCTTTACCCCGTTGCCCGGCAAC
GGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGTCATGGGCCATCAGCGCATGCGTGGAACCTTTTCGGC
TCCTCTGCCGATCCATACTGCGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACATTATCGGGACTGATAACTCTGTT
GTCCTATCCCGCAAATATACATCGTTTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCCGT
CGGCGCTGAATCCTGCGGACGACCCTTCTCGGGGTCGCTTGGGACTCTCTCGTCCCCTTCTCCGTCTGCCGTTCCGACCGACCACGGGGCG
CACCTCTCTTTACGCGGACTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATGGAGACCA
CCGTGAACGCCCACCAAATATTGCCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGACCGACCTTGAGGCATACTT
CAAAGACTGTTTGTTTAAAGACTGGGAGGAGTTGGGGGAGGAGATTAGGTTAAAGGTCTTTGTACTAGGAGGCTGTAGGCATAAATTGGTC
TGCGCACCAGCACCATGCAACTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGTCCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTG
GCTTTGGGGCATGGACATCGACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTTCCTTCAGTA
CGAGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCATGCCCGACG
GCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCG
GCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTC
GTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGA
AATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTC
CGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTT
ACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATC
TTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACA
ATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCAC
TGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTC
TTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCC
ACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT
TCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGC
GTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG
GCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGC
CCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG
GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATC
TGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTG
TTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAA
CTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAA
AGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAG
TTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTC
ACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCT
ATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCAC
GCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAG
CTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC
ATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCC
CGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAG
GATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGG
TGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATT
ATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATT
TCCCCGAAAAGTGCCACCTGACGTC

Figure 14

pGA1xFlag-M.pol2B7 nucleic acid sequence (in bold: pol2B7 polyepitope)

GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTG
CTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC
TTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGT
AAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT
ATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCaagcttCAGGCC
ATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTGATTACAAGGACGACGACGACAAGgaattc**AAGCCTGTGGTCTCCACACAG**
CTGCTTCTCAGGGCCAAGACCTTCCCTGTGAGACCCCAAGTGCCACTGAGAAGGGCTGATACACAGCCCAGGAGTGACACCCATGTGTTC
AGAACCAAGGCCATTCCTAGGAGAATTAGGCAGGGCCTGAGAGATACAGCTACACCTCAGGACCTGAACACCATGCTGAGActcgagCTG
AACATGGAGAACATCACATCAGGATTcctaggACCCCTTCTCGTGTTACAGGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCG
CAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAAT
CACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTA
TGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGGATCCTCAACAACCAGCACGGGA
CCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGGACGGAAATTGCACCTGT
ATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTGCCATTT
GTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTG
AGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCCTAACAAAACAAAGAGATGGGGTTACTCTCTAA
ATTTTATGGGTTATGTCATTGGATGTTATGGGTCCTTGCCACAAGAACACATCATACAAAAAATCAAAGAATGTTTTAGAAAACTTCCTA
TTAACAGGCCTATTGATTGGAAAGTATGTCAACGAATTGTGGGTCTTTTGGGTTTTGCTGCCCCTTTTACACAATGTGGTTATCCTGCGT
TGATGCCTTTGTATGCATGTATTCAATCTAAGCAGGCTTTCACTTTCTCGCCAACTTACAAGGCCTTTCTGTGTAAACAATACCTGAACC
TTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGTCATGGGCCATCAGC
GCATGCGTGGAACCTTTTCGGCTCCTCTGCCGATCCATACTGCGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACA
TTATCGGGACTGATAACTCTGTTGTCCTATCCCGCAAATATACATCGTTTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCG
GGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCTGCGGACGACCCTTCTCGGGGTCGCTTGGGACTCTCTCGTCCCCTTCTCCGTC
TGCCGTTCCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTT
CACCTCTGCACGTCGCATGGAGACCACCGTGAACGCCCACCAAATATTGCCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAAT
GTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAAGACTGGGAGGAGTTGGGGGAGGAGATTAGGTTAAAGGTCTTTGT
ACTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGTCCTACT
GTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCG
TTTTTGCCTTCTGACTTCTTTCCTTCAGTACGAGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTT
TAGTGAGGGTTAATTGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAAT
GGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAA
GAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTT
GACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCG
CCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCG
CCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGC
ATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGT
CATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCT
AATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCG
GCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGC
GAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGC
AAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAA
GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGC
CGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGG
TCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACC
CGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGA
AGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTA
GCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAG
AAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGA
TCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCT
TAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGG
AGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCG
GAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGC
CAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTT
CCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGT
TGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTG
AGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATA
GCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGT
AACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAA
AAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCA
TGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

Figure 15

pGA3xFlag-M.pol1B7 nucleic acid sequence (in bold: pol1B7 polyepitope)

GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTG
CTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC
TTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGT
AAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT
ATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCaagcttCAGGCC
ATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTGACTACAAAGACCACGACGGTGATTATAAAGATCACGACATTGATTACAAG
GACGACGACGACAAGgaattc**TCCCCTAGGACCCTGAATGCCTGGGTGAGAGCTAAGACCAGACCTAACAATAACACAAGGAAGTCCATC**
AGAGACACAGCCTTCCCTGTGAGACCACAGGTGCCTCTGAGGAGAACCAAGGCCCACCCTGTGCATGCTGGCCCTATTGCCAGAGCTGAT
ACAGCACCCACTAAGGCCAAAAGGAGAGTGGTCAGGctcgagCTGAACATGGAGAACATCACATCAGGATTcctaggACCCCTTCTCGTG
TTACAGGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGA
ACTACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGG
ATGTGTCTGCGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTG
CCCGTTTGTCCTCTAATTCCAGGATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTAT
CCCTCCTGTTGCTGTACCAAACCTTCGGACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAG
TGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTT
ATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTA
TACATTTAAACCCTAACAAAACAAAGAGATGGGGTTACTCTCTAAATTTTATGGGTTATGTCATTGGATGTTATGGGTCCTTGCCACAAG
AACACATCATACAAAAAATCAAAGAATGTTTTAGAAAACTTCCTATTAACAGGCCTATTGATTGGAAAGTATGTCAACGAATTGTGGGTC
TTTTGGGTTTTGCTGCCCCTTTTACACAATGTGGTTATCCTGCGTTGATGCCTTTGTATGCATGTATTCAATCTAAGCAGGCTTTCACTT
TCTCGCCAACTTACAAGGCCTTTCTGTGTAAACAATACCTGAACCTTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTG
CTGACGCAACCCCCACTGGCTGGGGCTTGGTCATGGGCCATCAGCGCATGCGTGGAACCTTTTCGGCTCCTCTGCCGATCCATACTGCGG
AACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACATTATCGGGACTGATAACTCTGTTGTCCTATCCCGCAAATATACAT
CGTTTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCTGCGGACG
ACCCTTCTCGGGGTCGCTTGGGACTCTCTCGTCCCCTTCTCCGTCTGCCGTTCCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACT
CCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATGGAGACCACCGTGAACGCCCACCAAAT
ATTGCCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAA
AGACTGGGAGGAGTTGGGGGAGGAGATTAGGTTAAAGGTCTTTGTACTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCAGCACCATG
CAACTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGTCCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGAC
ATCGACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTTCCTTCAGTACGAGATCCACTAGTT
CTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCATGCCCGACGGCGAGGATCTCGTC
GTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCG
GACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGT
ATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACC
AAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCG
GCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAA
GCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATG
TCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCAC
ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCG
CTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCG
CTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAG
AATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC
ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGT
TTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG
CGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGC
CCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA
GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTA
TCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT
TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACG
AAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAA
TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCAT
CCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC
CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCA
TCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG
TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAA
AAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT
CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGA
GTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC
GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCA
CCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT
TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAG
GGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

Figure 16

pGA3xFlag-M.pol2B7 nucleic acid sequence (in bold: pol2B7 polyepitope)

GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTG
CTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC
TTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGT
AAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT
ATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCaagcttCAGGCC
ATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTGACTACAAAGACCACGACGGTGATTATAAAGATCACGACATTGATTACAAG
GACGACGACGACAAGgaattc**AAGCCTGTGGTCTCCACACAGCTGCTTCTCAGGGCCAAGACCTTCCCTGTGAGACCCCAAGTGCCACTG**
AGAAGGGCTGATACACAGCCCAGGAGTGACACCCATGTGTTCAGAACCAAGGCCATTCCTAGGAGAATTAGGCAGGGCCTGAGAGATACA
GCTACACCTCAGGACCTGAACACCATGCTGAGActcgagCTGAACATGGAGAACATCACATCAGGATTcctaggACCCCTTCTCGTGTTA
CAGGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACT
ACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATG
TGTCTGCGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCC
GTTTGTCCTCTAATTCCAGGATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCC
TCCTGTTGCTGTACCAAACCTTCGGACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGG
GCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATA
TGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATAC
ATTTAAACCCTAACAAAACAAAGAGATGGGGTTACTCTCTAAATTTTATGGGTTATGTCATTGGATGTTATGGGTCCTTGCCACAAGAAC
ACATCATACAAAAAATCAAAGAATGTTTTAGAAAACTTCCTATTAACAGGCCTATTGATTGGAAAGTATGTCAACGAATTGTGGGTCTTT
TGGGTTTTGCTGCCCCTTTTACACAATGTGGTTATCCTGCGTTGATGCCTTTGTATGCATGTATTCAATCTAAGCAGGCTTTCACTTTCT
CGCCAACTTACAAGGCCTTTCTGTGTAAACAATACCTGAACCTTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTG
ACGCAACCCCCACTGGCTGGGGCTTGGTCATGGGCCATCAGCGCATGCGTGGAACCTTTTCGGCTCCTCTGCCGATCCATACTGCGGAAC
TCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACATTATCGGGACTGATAACTCTGTTGTCCTATCCCGCAAATATACATCGT
TTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCTGCGGACGACC
CTTCTCGGGGTCGCTTGGGACTCTCTCGTCCCCTTCTCCGTCTGCCGTTCCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCC
CGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATGGAGACCACCGTGAACGCCCACCAAATATT
GCCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAAGA
CTGGGAGGAGTTGGGGGAGGAGATTAGGTTAAAGGTCTTTGTACTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCAGCACCATGCAA
CTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGTCCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATC
GACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTTCCTTCAGTACGAGATCCACTAGTTCTA
GAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCATGCCCGACGGCGAGGATCTCGTCGTG
ACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGAC
CGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATC
GCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAG
CGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCT
GGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCA
ATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCT
GTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACA
ACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTT
TCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTT
CCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAAT
CAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATA
GGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTC
CCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGC
TTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCG
ACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA
TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCT
GCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTG
TTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAA
ACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCT
AAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCA
TAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCAC
GCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCC
AGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGG
TGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAG
CGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTC
TTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTT
GCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAA
AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCA
GCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCC
TTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGA

Figure 17

pGA1xFlag-Mpol1.A2 and pGA1xFlag-Mpol2.A2

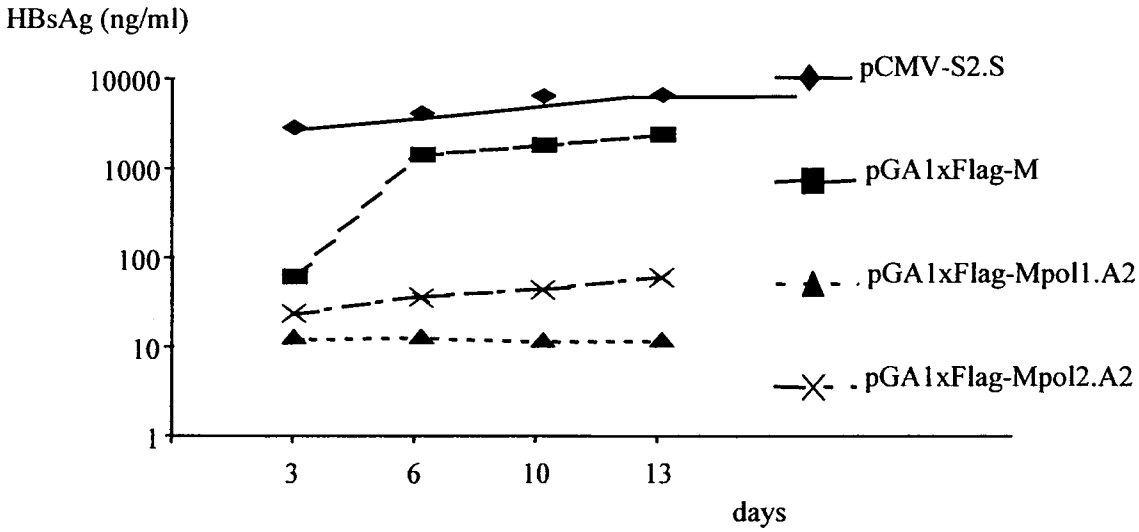


Figure 18

pGA3xFlag-Mpol1.A2 and pGA3xFlag-Mpol2.A2

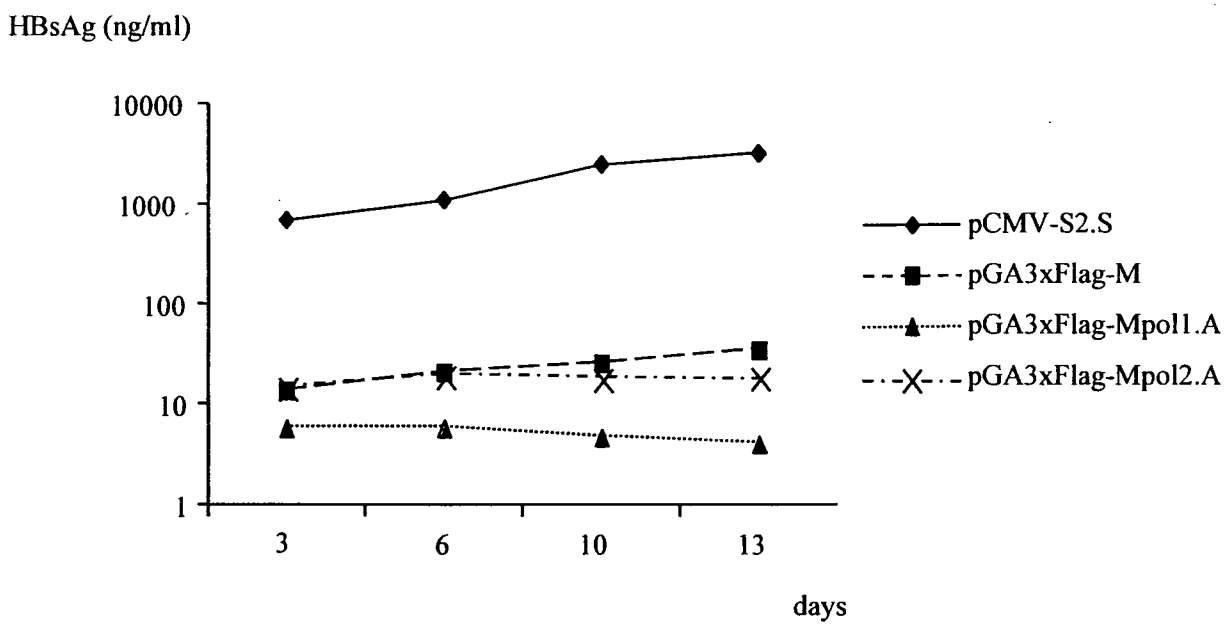


Figure 19

pGA1xFlag-Mpol1.B7 and pGA1xFlag-Mpol2.B7

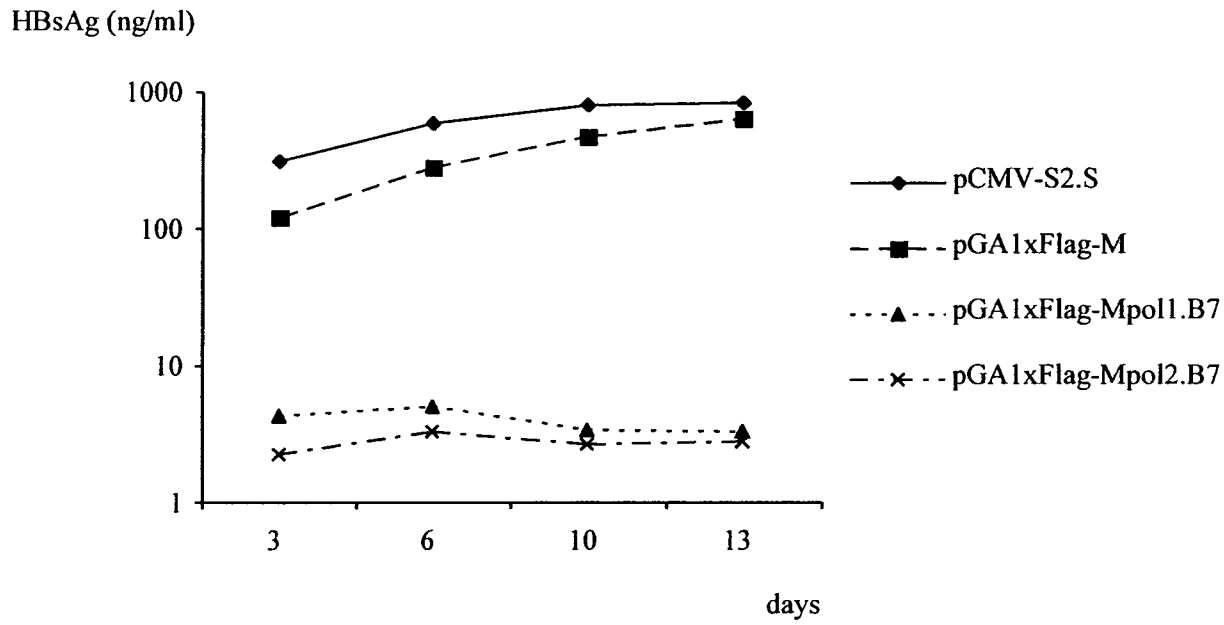


Figure 20

pGA3xFlag-Mpol1.B7 and pGA3xFlag-Mpol2.B7

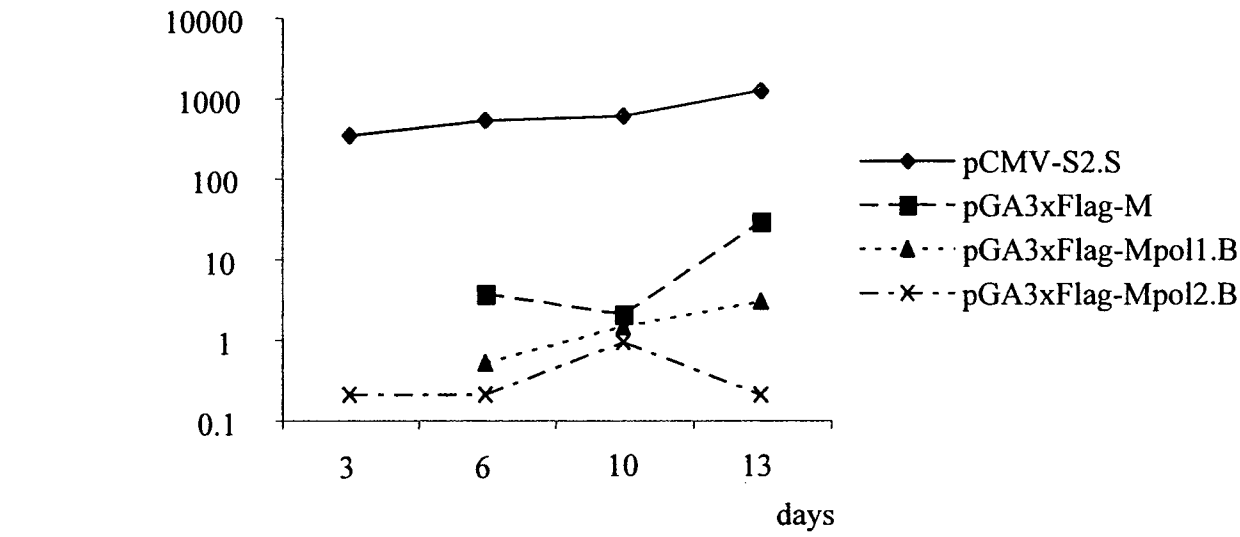


Figure 21

Oligonucleotides used for engineering the pGA3xFlagbasic, pGA1xFlag-Mbasic and pGA3xFlag-Mbasic plasmids.

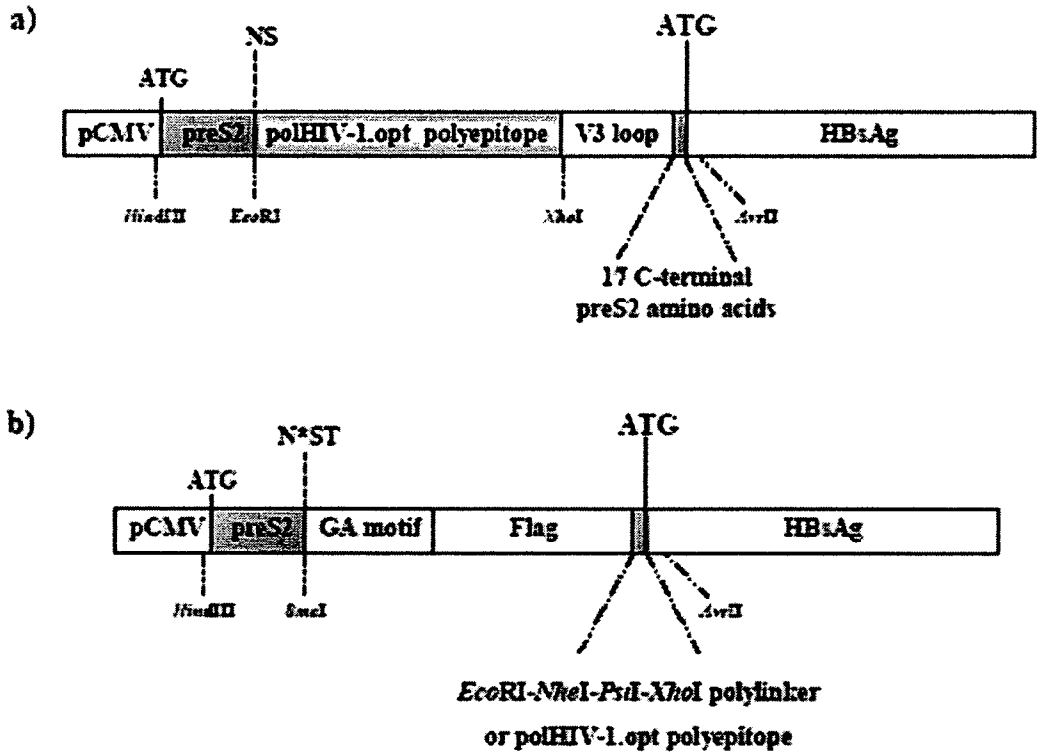
Oligonucleotide Sequence
pGA3xFlagbasic
Flag1
5'-CAGGCCATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGGGCTGACTACAAAGACCA TGACGGTGATTATAAAG-3'
Flag2
5'-TACAAAGACCATGACGGTGATTATAAAGATCATGACATTGATTACAAGGATGACGATGACA AGgaattcCTGCAG-3'
Flag3
5'-TGACGATGACAAGGAATTCCTGCAGGCTAGCAGATCTctcgagCTGAACATGGAGAACATCA CATCAGGATTcctag-3'

pGA1xFlag-Mbasic
1Xflag1
5'-CAGGCCATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTGATTACAAGGACG-3'
1Xflag2
5'-TGGAGCAGGAGCTGATTACAAGGACGACGACGACAAGgaattcCTGCAGGCTAGCAGATCTc tcg-3'
1Xflag3
5'-tcCTGCAGGCTAGCAGATCTctcgagCTGAACATGGAGAACATCACATCAGGATTcctag-3'

pGA3xFlag-Mbasic
3Xflag-M.1
5'-CAGGCCATGCAGTGGAACTCCACACCCGGGGCTGGAGCAGGAGCTGACTACAAAG-3'
3Xflag-M.2
5'-TGGAGCAGGAGCTGACTACAAAGACCACGACGGTGATTATAAAGATCACGACATTGATTA CAAG-3'
3Xflag-M.3
5'-TGATTATAAAGATCACGACATTGATTACAAGGACGACGACGACAAGgaattcCTGCAGGCTA GCAGATCTctcg-3'
3Xflag-M.4
5'-GCAGGCTAGCAGATCTctcgagCTGAACATGGAGAACATCACATCAGGATTcctag-3'

pGA3xFlagbasic, pGA1xFlag-Mbasic, pGA3xFlag-Mbasic
5'flag
5'-AGACCCaagcttCAGGCCATGCAGTGGAACTCCACA-3'
3'flag
5'-AGGGGTcctaggAATCCTGATGTGATGTTCTCCATG-3'

Figure 22



c) 5'- aagcttCAGGCCATGCAGTGGAACTCCACAcccgggGCTGGAGCAGGAGCT*GATTACAA*
*GGACGACGACGACAAG*gaattcctgcaggctagcagatctctcgagCTGAACATGGAGAACATCACATCA
GGATT*cctagg* - 3'

d) 5'- aagcttCAGGCCATGCAGTGGAACTCCACAcccgggGCTGGAGCAGGAGCT*GACTA*
CAAAGACCACGACGGTGATTATAAAGATCACGACATTGATTACAAGGACGACGACG
*ACAAG*gaattcctgcaggctagcagatctctcgagCTGAACATGGAGAACATCACATCAGGATT*cctagg* - 3'

Figure 23

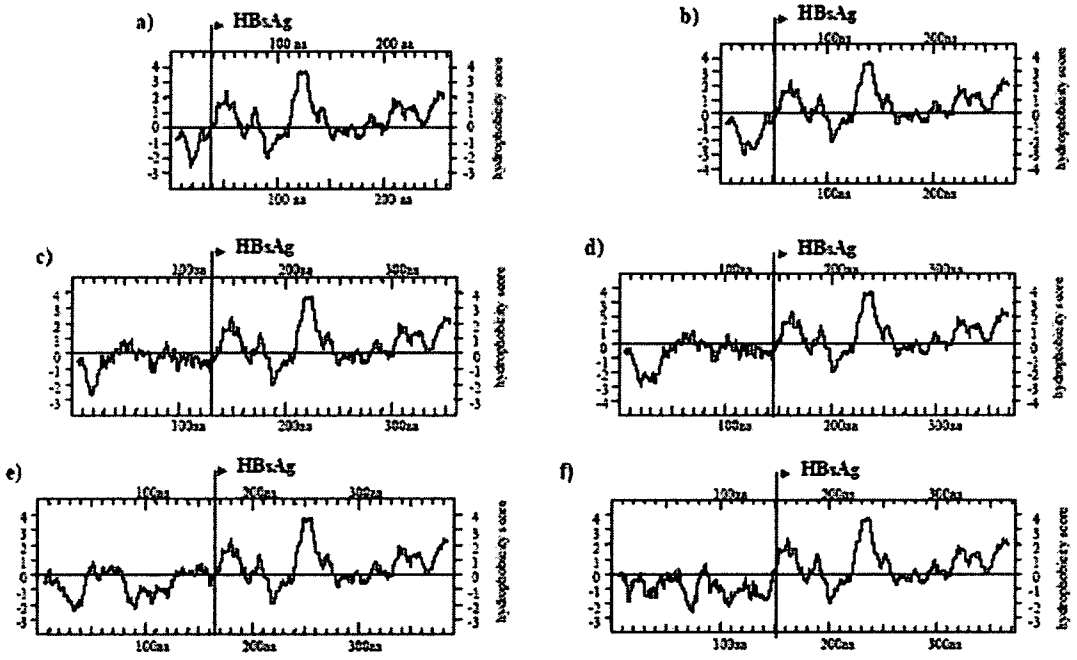


Figure 24

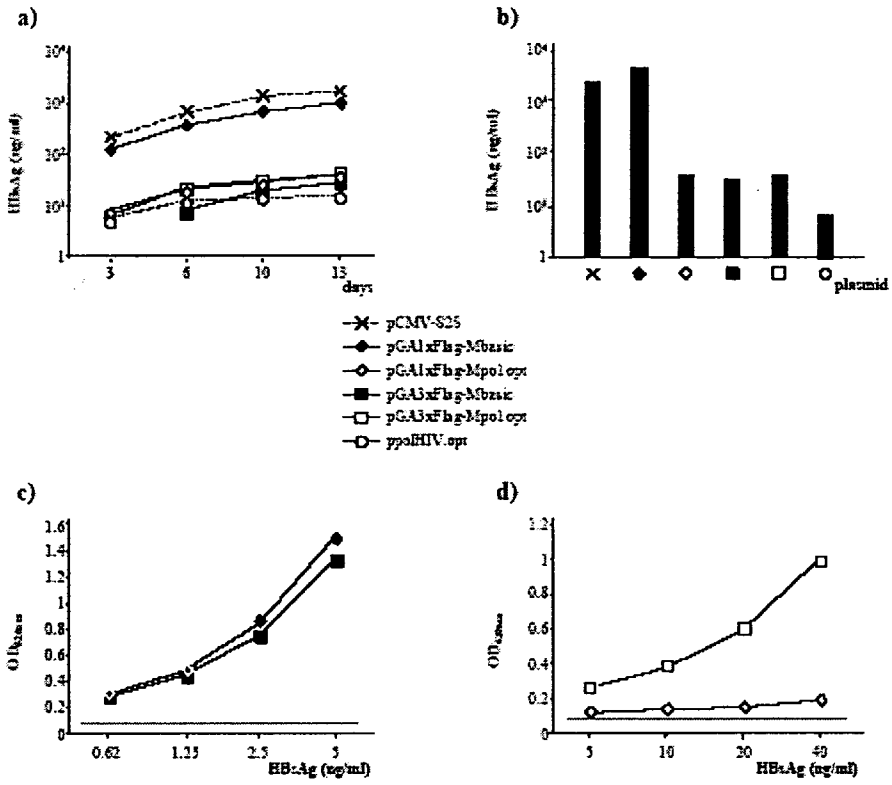


Figure 25

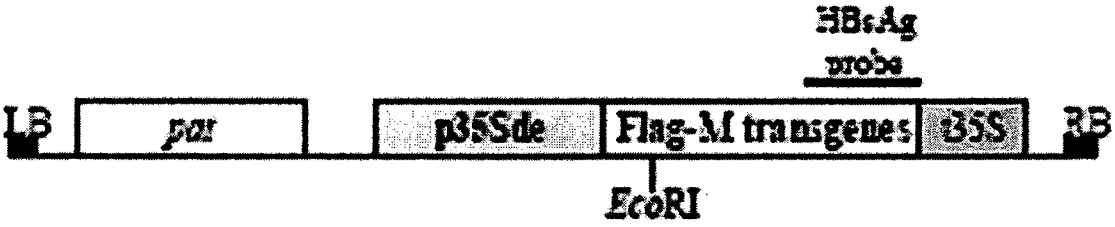


Figure 26

Construct	Plant code	Transgene Copy number	TSP[a,b] (mg/ ml)	HBsAg[b] (ng/ ml)	ngHBsAg / mgTSP[b]
GA1xFlag-Mbasic	Nt4-1	4	0.86	1.6	1.9
	Nt4-2	4	0.9	30.2	33.6
	Nt4-3	2	1.1	27.4	24.9
	Nt4-4	4	0.8	60.6	75.8
	Nt4-5	1	1.1	0.9	0.8
	Nt4-6	1	0.8	8.2	10.25
	Nt4-9	N.t.c	0.94	1.5	1.6
	Nt4-10	N.t.	0.62	4.8	7.7
	Nt4-11	3	1	21.3	21.3
	Nt4-12	1-2	0.65	3	4.6
	Nt4-13	1	0.74	9	12.1
GA1xFlag-Mpol.opt	Nt5-1	3	0.85	0.9	1.1
	Nt5-2	1	1	N.d.d	-
	Nt5-3	2	0.85	0.7	0.8
	Nt5-4	1	0.65	0.5	0.8
	Nt5-5	1	1.08	0.8	0.75
	Nt5-6	1	1.04	1.1	1.1
	Nt5-7	1	0.92	1.1	1.2
	Nt5-8	2	1.06	1.4	1.3
	Nt5-9	1	1.16	0.8	0.7
	Nt5-10	1	0.9	1.2	1.3
	Nt5-15	2	1.5	1.9	1.3
	Nt5-17	1	1.5	3.4	2.3
	Nt5-22	1	0.8	1.3	1.625
	Nt5-25	3	0.78	N.d.	-
	Nt5-29	2	1.6	1.4	0.9
	Nt5-30	1	1.12	0.8	0.7
	Nt5-31	1	0.75	N.d.	-
	Nt5-32	2	0.95	0.5	0.5
	Nt5-33	5	1.16	0.96	0.8
	Nt5-34	5	0.92	1.1	1.2
	Nt5-35	1	0.78	N.d.	-
	Nt5-36	3	1.4	0.8	0.6
	Nt5-37	1	1	0.5	0.5
	Nt5-38	1	1.4	1.7	1.2
	Nt5-39	3	1.24	1.6	1.3
	Nt5-40	2	0.72	N.d.	-
	Nt5-41	1	1	N.d.	-
	Nt5-42	5	1.1	1.5	1.4
	Nt5-43	7	1.12	1	0.9
	Nt5-44	1	1.3	1.2	0.9

Figure 27

Construct	Plant code	Transgene Copy number	TSP[a,b] (mg/ ml)	HBsAg[b] (ng/ ml)	ngHBsAg / mgTSP[b]
GA3xFlag-Mbasic	Nt18-1	1	0.82	1.7	2
	Nt18-2	3	1	9.3	9.3
	Nt18-4	1	1.02	3.1	3
	Nt18-5	1	0.66	1.3	1.9
	Nt18-7	1	1	6.6	6.6
	Nt18-8	4	0.96	3.2	3.3
	Nt18-10	3	0.56	0.6	1
	Nt18-11	2	0.82	1	1.2
	Nt18-13	1	1.2	4.2	3.5
	Nt18-14	1	0.74	2.4	3.3
	Nt18-19	2	1.6	6.1	3.8
	Nt18-21	2	0.84	2.2	2.6
	Nt18-22	1	1.18	3.5	3
GA3xFlag-Mpol.opt	NtR-1	4	0.8	0.5	0.6
	NtR-2	1	0.62	1	1.6
	NtR-3	1	1.22	1.3	1.1
	NtR-4	1	1.06	0.5	0.5
	NtR-5	2	0.96	1	1
	NtR-6	1	0.85	N.d.	-
	NtR-11	1	1.15	0.5	0.4
	NtR-12	1	1.3	3	2.3
	NtR-13	1	1.18	0.6	0.5
	NtR-14	2	0.7	N.d.	-
	NtR-16	1	0.95	N.d.	-
	NtR-17	1	0.8	1.3	1.625
	NtR-18	3	0.74	0.8	1.1
	NtR-19	2	0.95	0.85	0.9
	NtR-21	3	1.1	2	1.8
	NtR-22	3	0.94	1.1	1.2
	NtR-24	1	1.24	0.7	0.6
	NtR-25	2	0.76	0.6	0.8
	NtR-26	2	0.88	N.d.	0.4
	NtR-27	2	1.16	1	0.9
	NtR-28	1	0.78	N.d.	-
	NtR-29	1	1.05	1.1	1
	NtR-30	1	0.86	0.7	0.8
	NtR-31	3	1.42	0.9	0.65
	NtR-32	5	1.3	1.2	0.9
	NtR-33	3	0.88	0.4	0.5
	NtR-34	2	0.9	0.8	0.9
	NtR-35	2	0.9	0.7	0.8
	NtR-36	1	1.14	0.7	0.6
	NtR-37	1	1.1	1.4	1.3
	NtR-B	1	0.8	0.7	0.9
Wild type	WT	-	1.15	N.d.	-

Figure 27

Construct	Plant code	Transgene Copy number	TSP[a,b] (mg/ ml)	HBsAg[b] (ng/ ml)	ng HBsAg / mg TSP[b]
GA1xFlag-Mbasic	At4-2	n.t.[c]	0.22	1.89	8.75
	At4-4	n.t.	0.21	7.47	35.48
	At4-5	5	0.22	n.d.d	-
	At4-6	1	0.32	n.d.	-
	At4-8	2	0.18	2.66	14.61
	At4-11	2	0.21	25.60	119.94
	At4-12	2	0.15	5.78	39.09
	At4-13	2	0.14	35.00	257.84
	At4-14	1	0.18	7.69	42.62
	At4-15	7	0.35	1.19	3.43
	At4-17	8	0.16	0.29	1.87
	At4-19	n.t.	0.22	2.08	9.55
	At4-21	2	0.18	7.23	41.06
	At4-22	1	0.15	5.40	35.81
	At4-23	n.t.	0.20	n.d.	-
	At4-24	8	0.32	0.92	2.89
	At4-25	1	0.26	0.12	0.46
	At4-27	1	0.33	1.12	3.41
GA1xFlag-Mpol.opt	At5-1	6	0.38	n.d.	-
	At5-3	1	0.39	1.28	3.29
	At5-4	5	0.31	0.71	2.33
	At5-5	1	0.28	0.65	2.35
	At5-6	5	0.26	0.09	0.33
	At5-7	3	0.23	1.03	4.46
	At5-8	4	0.33	1.40	4.23
	At5-9	4	0.25	0.94	3.84
	At5-10	3	0.26	n.d.	-
	At5-11	2	0.33	4.52	13.61
	At5-12	3	0.29	n.d.	-
	At5-14	2	0.31	1.26	4.10
	At5-15	2	0.33	0.09	0.27
	At5-16	4	0.38	n.d.	-
	At5-17	6	0.33	0.08	0.25
	At5-18	3	0.39	0.29	0.73
	At5-19	1	0.39	3.23	8.31
	At5-20	n.t.	0.32	1.34	4.15
	At5-21	2	0.15	1.75	11.48
	At5-22	6	0.39	0.64	1.63
	At5-23	1	0.37	1.87	5.09
	At5-24	3	0.35	0.48	1.38
	At5-25	4	0.42	2.49	5.98
	At5-26	n.t.	0.29	0.26	0.90
	At5-30	1	0.33	1.93	5.93
	At5-31	4	0.29	n.d.	-
	At5-33	4	0.35	0.98	2.77
	At5-34	n.t.	0.35	0.37	1.08
	At5-35	6	0.32	0.56	1.78
	At5-36	4	0.41	0.97	2.34
	At5-37	n.t.	0.29	0.18	0.63
	At5-38	n.t.	0.36	1.41	3.91
	At5-40	2	0.23	1.51	6.67
	At5-41	1	0.33	0.78	2.33
	At5-42	1	0.19	n.d.	-
	At5-43	n.t.	0.38	1.65	4.37
	At5-44	n.t.	0.31	1.49	4.87
	At5-48	2	0.31	0.94	3.06
	At5-50	n.t.	0.19	n.d.	-
	At5-51	2	0.35	0.05	0.16
	At5-52	3	0.22	0.24	1.06
	At5-53	3	0.26	1.12	4.30
	At5-54	1	0.18	n.d.	-

Figure 28

Construct	Plant code	Transgene Copy number	TSP[a,b] (mg/ ml)	HBsAg[b] (ng/ ml)	ng HBsAg / mg TSP[b]
	At5-55	9	0.23	1.03	4.49
	At5-57	6	0.20	1.29	6.51
	At5-58	n.t.	0.28	0.94	3.36
	At5-59	5	0.37	1.56	4.24
	At5-60	6	0.22	0.13	0.61
	At5-61	2	0.23	0.87	3.76
	At5-62	7	0.26	0.32	1.24
	At5-63	7	0.26	0.78	3.03
	At5-65	n.t.	0.30	1.02	3.43
	At5-68	2	0.21	0.58	2.75
	At5-69	2	0.18	n.d.	-
GA3xFlag-Mbasic	At18-1	n.t.	0.18	0.27	1.48
	At18-5	2	0.19	5.57	29.34
	At18-7	n.t.	0.20	n.d.	-
	At18-8	1	0.19	7.05	37.31
	At18-9	9	0.20	1.45	7.39
	At18-10	4	0.28	1.50	5.30
	At18-11	2	0.30	3.47	11.75
	At18-12	4	0.29	9.65	33.15
	At18-13	3	0.28	1.89	6.76
	At18-14	3	0.33	12.3	36.87
	At18-15	4	0.19	0.27	1.39
	At18-16	2	0.22	n.d.	-
	At18-18	4	0.26	0.99	3.79
	At18-19	3	0.27	1.34	4.88
	At18-20	1	0.31	n.d.	-
	At18-22	2	0.28	1.45	5.26
	At18-23	9	0.18	n.d.	-
	At18-25	5	0.22	0.52	2.30
	At18-26	3	0.19	0.06	0.31
	At18-27	6	0.22	0.33	1.50
	At18-29	8	0.27	1.85	6.80
	At18-30	1	0.34	8.37	24.72
	At18-31	1-2	0.27	3.31	12.38
	At18-32	n.t.	0.17	3.36	19.91
	At18-33	n.t.	0.24	1.79	7.34
GA3xFlag-Mpol.opt	AtR-1	3	0.43	n.d.	-
	AtR-2	6	0.20	n.d.	-
	AtR-3	7-8	0.20	0.01	0.03
	AtR-9	5	0.24	n.d.	-
	AtR-10	n.t.	0.21	0.97	4.59
	AtR-11	4	0.20	0.66	3.37
	AtR-12	4	0.22	0.00	0.02
	AtR-13	1	0.22	0.01	0.04
	AtR-17	n.t.	0.21	0.37	1.74
	AtR-19	2	0.21	1.28	6.16
	AtR-20	2	0.24	1.43	5.86
	AtR-21	5	0.23	1.33	5.71
	AtR-22	3	0.40	1.58	3.93
	AtR-23	6	0.19	1.44	7.73
	AtR-25	4	0.31	0.28	0.91
	AtR-26	6	0.33	1.37	4.19
	AtR-27	4	0.34	2.23	6.64
	AtR-28	3	0.37	1.24	3.34
	AtR-29	3	0.22	0.64	2.91
	AtR-30	6	0.19	0.50	2.67
	AtR-31	4	0.33	n.d.	-
	AtR-32	1	0.40	0.93	2.30
	AtR-33	n.t.	0.30	1.97	6.55
	AtR-34	1	0.23	0.33	1.43
	AtR-35	5	0.44	0.13	0.29

Figure 28

Construct	Plant code	Transgene Copy number	TSP[a,b] (mg/ ml)	HBsAg[b] (ng/ ml)	ng HBsAg / mg TSP[b]
	AtR-36	4	0.42	0.59	1.41
	AtR-37	5	0.26	0.13	0.53
	AtR-38	5	0.32	1.87	5.87
	AtR-40	n.t.	0.39	0.02	0.04
	AtR-41	6	0.26	1.14	4.46
	AtR-43	4	0.30	2.99	9.98
	AtR-44	1	0.25	1.30	5.27
	AtR-45	1	0.20	2.41	12.31
	AtR-46	2	0.24	n.d.	-
	AtR-48	5	0.37	1.19	3.22
	AtR-49	n.t.	0.35	0.49	1.39
	AtR-51	5	0.17	n.d.	-
	AtR-52	3	0.17	0.30	1.76
	AtR-53	n.t.	0.16	0.21	1.31
	AtR-56	n.t.	0.32	n.d.	-
Wild type	WT1		0.24		
	WT2		0.32		

Figure 28

Table 1
VLPs production in transgenic Tobacco and Arabidopsis plants.

Construct	*Nicotiana tabacum*[a]			*Arabidopsis thaliana*[b]		
		HBsAg (ng / mg TSP[c])			HBsAg (ng / mg TSP[c])	
	n[d]	Min	Max	n[d]	Min	Max
GA1xFlag-Mbasic	11	0.8	75.8	18	0.5	257.8
GA1xFlag-Mpolopt	30	0.5	2.3	54	0.2	13.6
GA3xFlag-Mbasic	13	1.0	9.3	25	0.3	37.3
GA3xFlag-Mpolopt	31	0.4	2.3	40	0.0	12.3

[a]Reported values summarise data from Supplementary Table 2.
[b]Reported values summarise data from Supplementary Table 3.
[c]TSP: total soluble protein.
[d]Number of analysed transgenic plants.

Figure 29

Table 2

Characterization of the 14 selected transgenic tobacco plants.

Construct	Plant	Transgene copy number	Transgene mRNA expression (E1)[a]	Transgene mRNA expression (E2)[a]	ng HBsAg / mg TSP[b] (E1-A)[c]	ng HBsAg / mg TSP (E1-B)[c]	ng HBsAg / mg TSP (E2)[c]
GA1xFlag-Mbasic	Nt4-2	4	1.1	1.3	33.5	45.8	86.6
	Nt4-3	2	1.3	1.3	24.9	10.6	183.3
	Nt4-4	4	1.2	1.5	75.8	64.3	29.2
	Nt4-11	3	0.8	0.6	21.3	12.9	166.3
GA1xFlag-Mpolopt	Nt5-15	2	0.4	0.5	3.3	1.1	1.7
	Nt5-17	1	0.4	0.4	2.3	2.1	6.0
	Nt5-29	2	0.5	0.7	0.9	0.8	6.6
	Nt5-38	1	0.1	0.5	1.2	1.0	6.2
GA3xFlag-Mbasic	Nt18-2	3	3.6	1.9	9.3	9.4	86.2
	Nt18-7	1	1.5	0.6	6.6	5.6	34.7
	Nt18-13	1	2.0	1.9	3.5	5.1	70.5
	Nt18-19	2	2.0	2.4	3.8	4.8	70.5
GA3xFlag-Mpolopt	NtR-12	1	0.4	0.7	2.3	1.1	4.4
	NtR-21	3	0.5	1.7	3.8	2.5	6.8

$p > 0.05$[e]

$p > 0.05$[d]

$p < 0.05$[d]

$p > 0.05$[e]

$p > 0.05$[e]

[a] Normalised values.
[b] TSP: total soluble protein.
[c] Reported values are the mean among three independent measurements.
[d] Wilcoxon signed-rank test.
[e] Z-correlation test.

Figure 30

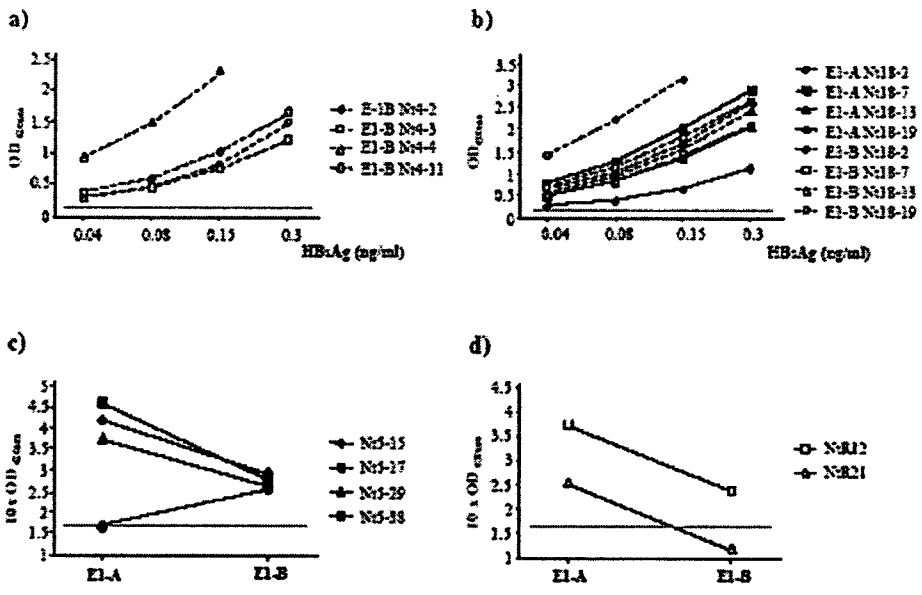


Figure 31

Construct	T0 plant code	T1 plant code	Transgene copy number	Transgene mRNA expression level	ngHBsAg / mgTSPa b
GA1xFlag-Mbasic	Nt4-4	4.4-1	4	1.0	109.7
		4.4-2	3	2.5	46.9
		4.4-3	2	2.6	262.6
		4.4-4	2	1.5	319.5
		4.4-5	4	0.9	159.9
	Nt4-11	4.11-1	3	1.4	809.9
		4.11-2	3	1.3	887.1
		4.11-3	2	1.7	749.0
		4.11-4	2	1.3	877.9
		4.11-5	3	1.3	680.4
GA1xFlag-Mpol.opt	Nt5-15	5.15-1	2	0.6	0.7
		5.15-2	2	0.2	0.4
		5.15-3	2	0.4	HBsAgn.d.c
		5.15-4	2	0.5	HBsAgn.d.
		5.15-5	2	0.5	0.4
	Nt5-17	5.17-1	1	0.5	HBsAgn.d.
		5.17-2	1	0.1	0.5
		5.17-3	1	0.5	0.7
		5.17-4	1	0.2	0.8
		5.17-5	1	0.3	HBsAgn.d.
GA3xFlag-Mbasic	Nt18-2	18.2-1	1	2.0	58.8
		18.2-2	1	4.7	488.2
		18.2-3	1	1.6	401.1
		18.2-4	3	1.5	234.3
		18.2-5	2	2.6	347.7
	Nt18-7	18.7-1	1	1.0	136.7
		18.7-2	1	0.6	69,3
		18.7-3	1	1.3	53.3
		18.7-4	1	2.8	54.0
		18.7-5	1	1.9	292.7
GA3xFlag-Mpol.opt	NtR-12	R.12-1	1	0.5	HBsAgn.d.
		R.12-2	1	0.8	1.2
		R.12-3	1	1.0	0.1
		R.12-4	1	0.4	0.5
		R.12-5	1	0.5	0.7
	NtR-21	R.21-1	2	1.0	0.1
		R.21-2	2	1.1	1.2
		R.21-3	2	0.5	HBsAgn.d.
		R.21-4	2	0.9	0.3
		R.21-5	2	0.7	0.2

——— p>0.05d ———

Figure 32

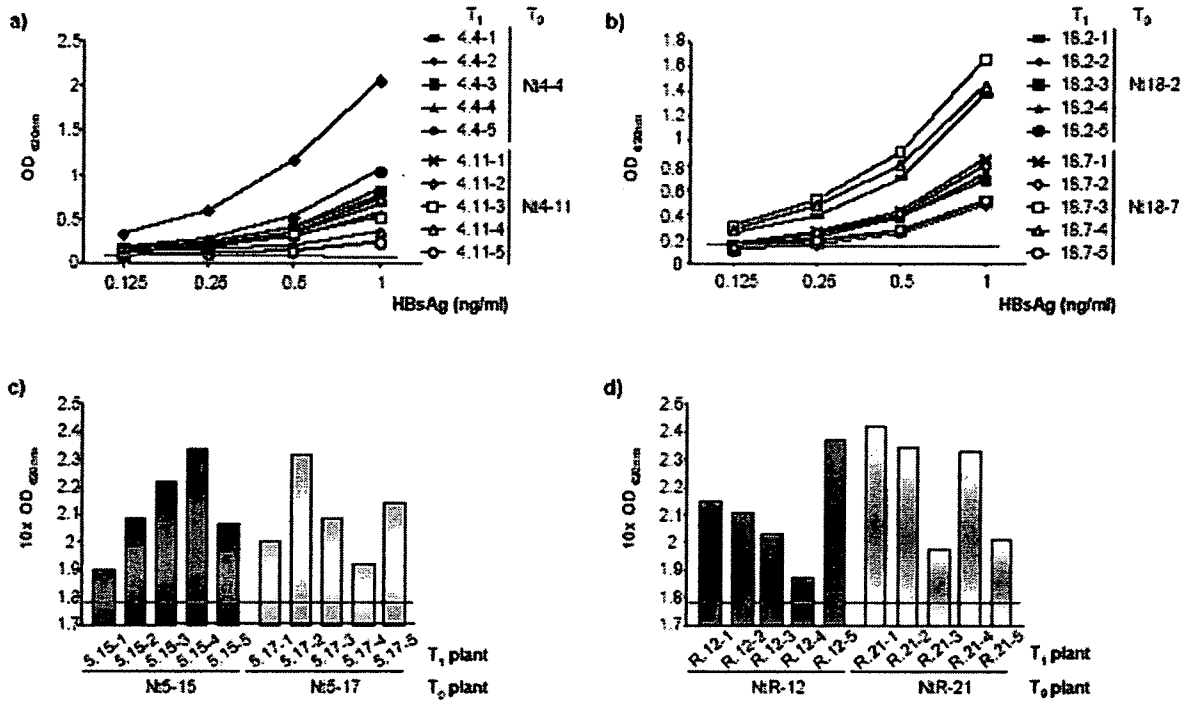


Figure 33

Characterization of 15 selected transgenic Arabidopsis plants

Construct	Plantcode	Transgene copy number	Transgene mRNA expression[a]	ng HBsAg / mg TSP[b,c]
1xFlag-Mbasic	At4-13	2	1.8	119.9
	At4-13	2	3.4	257.8
	At4-14	1	0.4	42.6
	At4-23	2	0.3	41.1
1xFlag-Mpolopt	At5-13	2	1.1	13.6
	At5-19	1	0.5	8.3
	At5-25	4	1.0	6.0
	At5-40	2	0.1	6.7
3xFlag-Mbasic	At18-5	2	0.4	29.3
	At18-8	1	1.1	37.3
	At18-12	4	2.9	33.2
	At18-14	3	2.3	36.9
3xFlag-Mpolopt	AtR-19	2	0.5	6.2
	AtR-23	6	0.1	7.7
	AtR-43	4	0.1	10.0
	AtR-45	1	0.7	12.3

$p = 0.05$[d]

[a] Normalised values.

[b] TSP: total soluble protein.

[c] Reported values are the mean among three independent measurements.

[d] Z-correlation test.

Figure 34

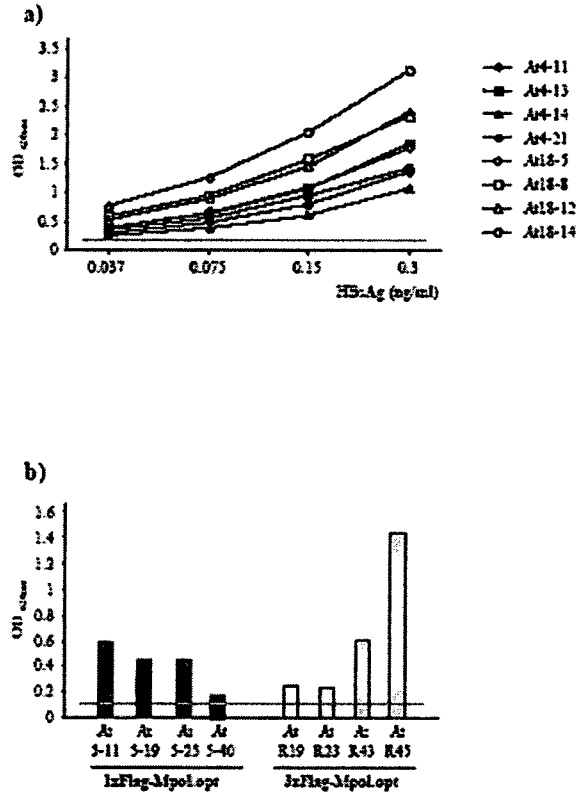
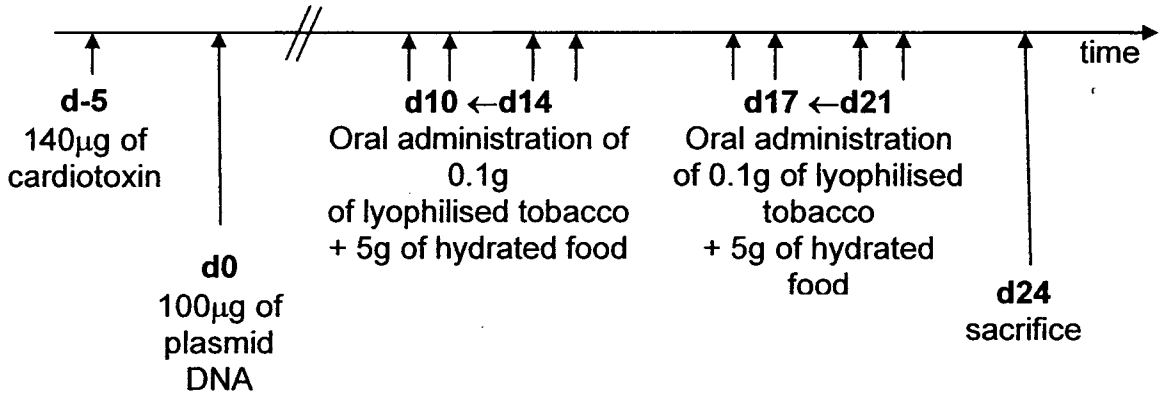

a)
3.5
3
2.5
2
1.5
1
0.5
0
0.037
0.075
0.15
0.3
At4-11
At4-13
At4-14
At4-21
At18-5
At18-8
At18-12
At18-14
b)
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
At 5-11
At 5-19
At 5-25
At 5-40
At R19
At R23
At R43
At R45
1xFlag-hIpoLopt
3xFlag-hIpoLopt

Figure 35

A


time
d-5
140µg of cardiotoxin
d0
100µg of plasmid DNA
d10 ←d14
Oral administration of 0.1g of lyophilised tobacco + 5g of hydrated food
d17 ←d21
Oral administration of 0.1g of lyophilised tobacco + 5g of hydrated food
d24
sacrifice

B

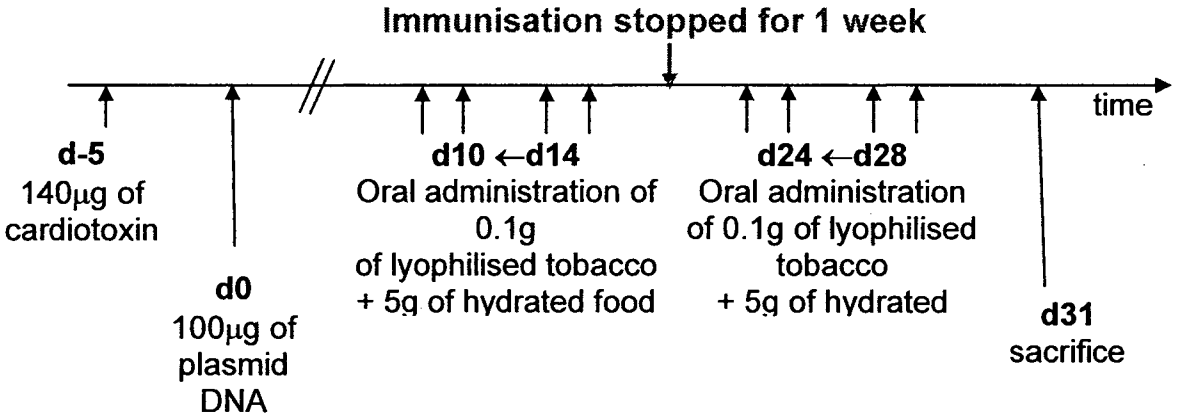

Immunisation stopped for 1 week
time
d-5
140µg of cardiotoxin
d0
100µg of plasmid DNA
d10 ←d14
Oral administration of 0.1g of lyophilised tobacco + 5g of hydrated food
d24 ←d28
Oral administration of 0.1g of lyophilised tobacco + 5g of hydrated
d31
sacrifice

Figure 36

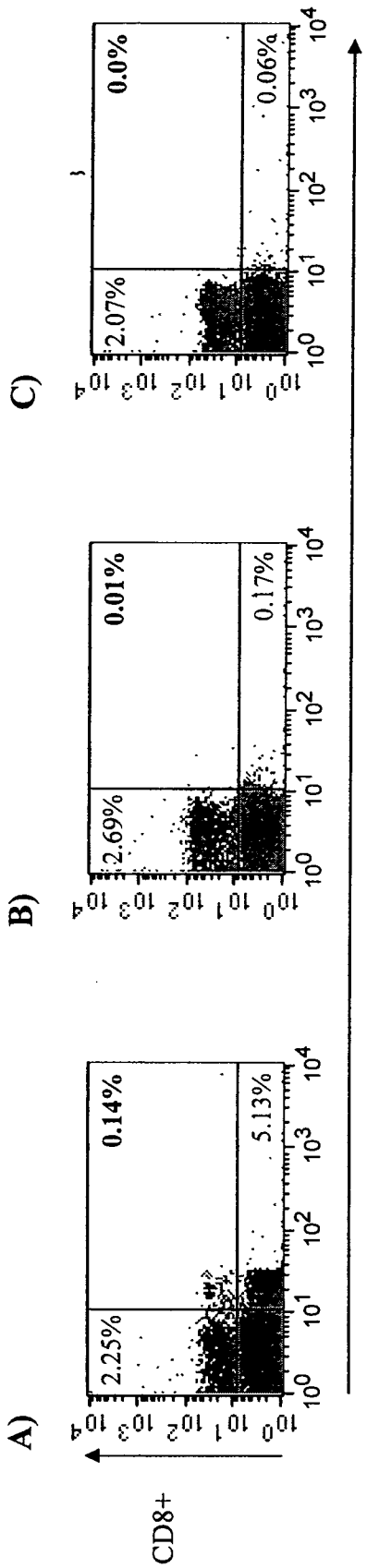
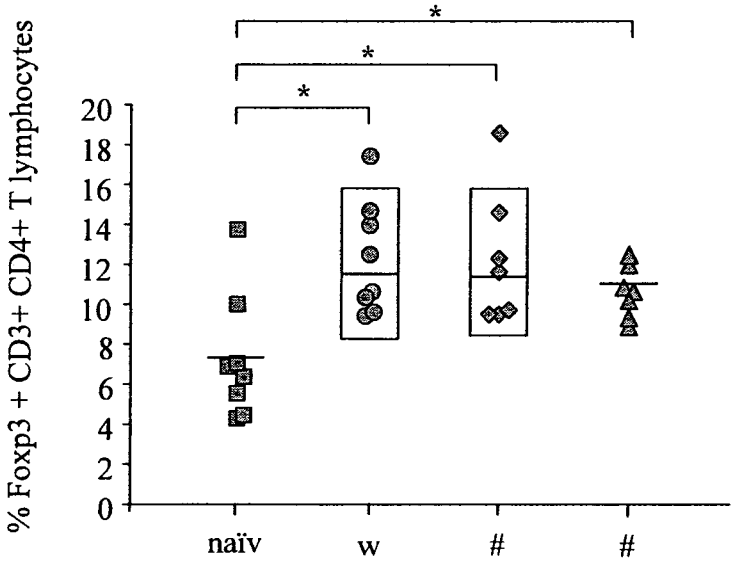


Figure 37

A)


% Foxp3 + CD3+ CD4+ T lymphocytes
20
18
16
14
12
10
8
6
4
2
0
naïv
w
#
#
*
*
*

B)

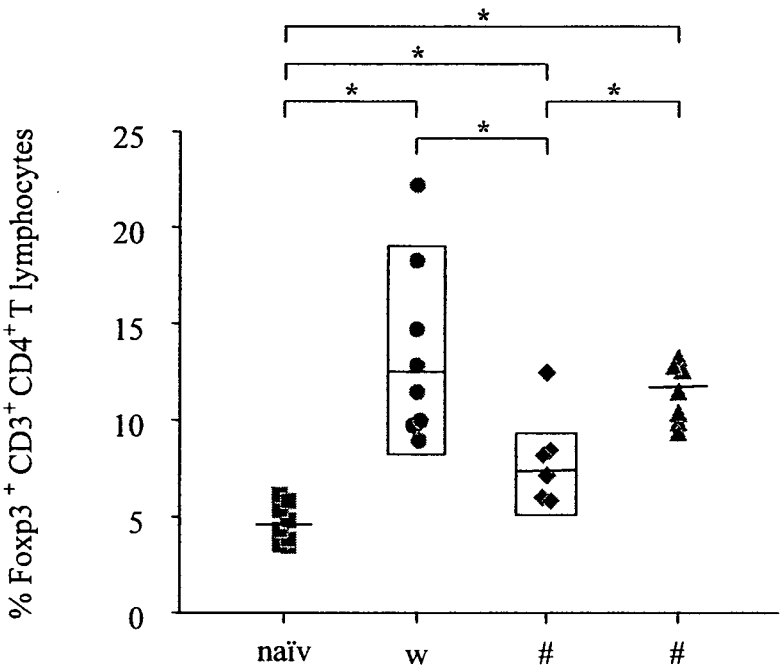

% Foxp3 + CD3+ CD4+ T lymphocytes
25
20
15
10
5
0
naïv
w
#
#
*
*
*
*
*

Figure 38

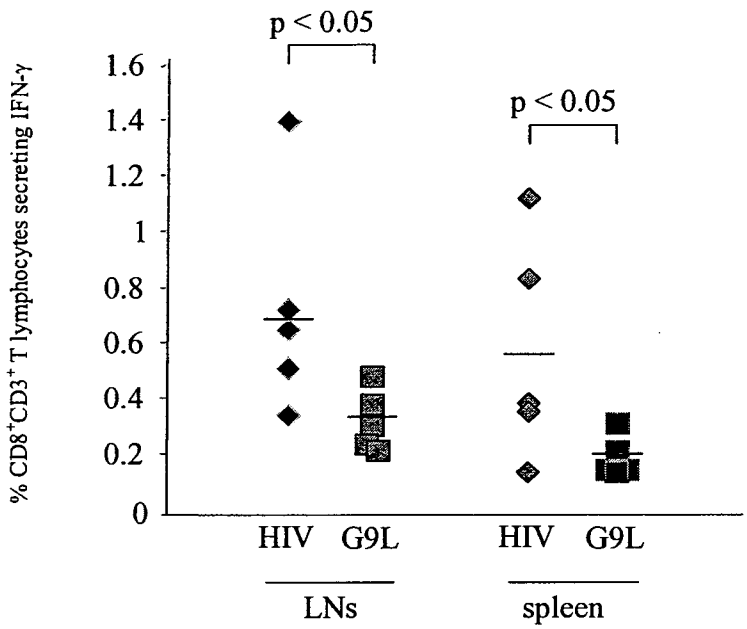

p < 0.05
p < 0.05
% CD8+CD3+ T lymphocytes secreting IFN-γ
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
HIV
G9L
HIV
G9L
LNs
spleen

Figure 39