



(43) International Publication Date
31 March 2016 (31.03.2016)

(10) International Publication Number
WO 2016/046635 A1

- (51) **International Patent Classification:**
C12Q 1/68 (2006.01) *C12Q 1/70* (2006.01)
- (21) **International Application Number:**
PCT/IB2015/001973
- (22) **International Filing Date:**
25 September 2015 (25.09.2015)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
62/055217 25 September 2014 (25.09.2014) US
- (71) **Applicant:** INSTITUT PASTEUR [FR/FR]; 25-28, rue du Docteur Roux, F-75724 Paris Cedex 15 (FR).
- (72) **Inventors:** CAVAL, Vincent; Résidence le Clos Simon, Appt D82, Rue Gilbert Becaud, F-37520 La Riche (FR). SUSPENE, Rodolphe; 34, avenue Saint Laurent, Résidence Chevreuse, Bât. 7, F-91400 Orsay (FR). VARTANI-AN, Jean-Pierre; 1, bis rue Mornay, F-75004 Paris (FR). WAIN-HOBSON, Simon; 3, rue Jean de la Fontaine, F-78180 Montigny Le Bretonneux (FR).
- (74) **Agent:** GUTMANN, Ernest; Yves Plasseraud SAS, 3, rue Auber, F-75009 Paris (FR).
- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM,

AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
- *with sequence listing part of description (Rule 5.2(a))*

(54) **Title:** METHODS FOR CHARACTERIZING HUMAN PAPILLOMAVIRUS ASSOCIATED CERVICAL LESIONS

(57) **Abstract:** Methods of monitoring a cervical HPV infection and/or an HPV induced cervical lesion in a subject are provided. Methods of identifying a subject with a cervical HPV infection and/or an HPV induced cervical lesion, with an increased risk of developing cervical squamous cell carcinoma are also provided. The methods comprise performing a genotyping assay on a nucleic acid sample of the subject to determine whether the A3 genotype of the subject comprises at least one copy of an A3A-A3B deletion allele (Δ A3B). In some embodiments, if the A3 genotype of the subject comprises at least one copy of an A3A-A3B deletion allele (Δ A3B) then the subject is advised to receive the next Pap test at a shorter time interval than if the A3 genotype of the subject does not comprise at least one copy of a A3A-A3B deletion allele (Δ A3B). In some embodiments, the subject is diagnosed as having an increased risk of developing cervical squamous cell carcinoma if the A3 genotype of the subject comprises at least one copy of an A3 A- A3 B deletion allele (Δ A3B).



WO 2016/046635 A1

METHODS FOR CHARACTERIZING HUMAN PAPILLOMAVIRUS ASSOCIATED CERVICAL LESIONS

INTRODUCTION

5

Nearly all cases of cervical cancer are caused by infection with oncogenic, or high-risk, types of human papillomavirus (HPV). Infections with these sexually transmitted viruses also cause most anal cancers; many vaginal, vulvar, and penile cancers; and some oropharyngeal cancers. Although HPV infection is very common, most infections will be suppressed by the immune system within 1 to 2 years without causing cancer. These transient infections may cause temporary changes in cervical cells. If a cervical infection with a high-risk HPV type persists, the cellular changes can eventually develop into more severe precancerous lesions. If precancerous lesions are not treated, they can progress to cancer. It can take 10 to 20 years or more for a persistent infection with a high-risk HPV type to develop into cancer.

15

Cervical cancer screening is an essential part of a woman's routine health care. It is a way to detect abnormal cervical cells, including precancerous cervical lesions, as well as early cervical cancers. Both precancerous lesions and early cervical cancers can be treated very successfully. Routine cervical screening has been shown to greatly reduce both the number of new cervical cancers diagnosed each year and deaths from the disease. Cervical cancer screening includes two types of screening tests: cytology-based screening, known as the Pap test or Pap smear, and HPV testing. The main purpose of screening with the Pap test is to detect abnormal cells that may develop into cancer if left untreated. The Pap test can also find noncancerous conditions, such as infections and inflammation. It can also find cancer cells. In regularly screened populations, however, the Pap test identifies most abnormal cells before they become cancer.

20

25

HPV testing is used to look for the presence of high-risk HPV types in cervical cells. These tests can detect HPV infections that cause cell abnormalities, sometimes even before cell abnormalities are evident. Several different HPV tests have been approved for screening. Most approved tests detect the DNA of high-risk HPV, although one approved test detects the RNA of high-risk HPV. Some tests detect any high-risk HPV and do not

30

identify the specific type or types present. Other tests specifically detect infection with HPV 16 or 18, the two types that cause most HPV-associated cancers.

5 Cervical cancer screening can be done in a medical office, a clinic, or a community health center. In a conventional Pap test, the specimen (or smear) is placed on a glass microscope slide and a fixative is added. In an automated liquid-based Pap cytology test, cervical cells collected with a brush or other instrument are placed in a vial of liquid preservative. The slide or vial is then sent to a laboratory for analysis.

10 In the United States, automated liquid-based Pap cytology testing has largely replaced conventional Pap tests. One advantage of liquid-based testing is that the same cell sample can also be tested for the presence of high-risk types of HPV, a process known as "Pap and HPV cotesting." In addition, liquid-based cytology appears to reduce the likelihood of an unsatisfactory specimen. However, conventional and liquid-based Pap tests appear to have a similar ability to detect cellular abnormalities and both are in use today.

15 In March 2012, updated screening guidelines were released by the United States Preventive Services Task Force and jointly by the American Cancer Society, the American Society for Colposcopy and Cervical Pathology, and the American Society for Clinical Pathology. These guidelines recommend that women have their first Pap test at age 21. Although previous guidelines recommended that women have their first Pap test 3 years after they start having sexual intercourse, waiting until age 21 is now recommended
20 because adolescents have a very low risk of cervical cancer and a high likelihood that cervical cell abnormalities will go away on their own.

According to the updated guidelines, women ages 21 through 29 should be screened with a Pap test every 3 years. Women ages 30 through 65 can then be screened every 5 years with Pap and HPV cotesting or every 3 years with a Pap test alone.

25 The guidelines also note that women with certain risk factors may need to have more frequent screening or continue screening beyond age 65. These risk factors include being infected with the human immunodeficiency virus (HIV), being immunosuppressed, having been exposed to diethylstilbestrol before birth, and having been treated for a precancerous cervical lesion or cervical cancer.

30 On April 24, 2014, the Food and Drug Administration (FDA) approved the use of one HPV DNA test (cobas HPV test, Roche Molecular Systems, Inc.) as a first-line primary screening test for use alone for women age 25 and older. This test detects each of HPV

types 16 and 18 and gives pooled results for 12 additional high-risk HPV types. The new approval was based on long-term findings from the ATHENA trial, a clinical trial that included more than 47,000 women. The results showed that the HPV test used in the study performed better than the Pap test at identifying women at risk of developing severe cervical cell abnormalities.

The greater assurance against future cervical cancer risk with HPV testing has also been demonstrated by a cohort study of more than a million women, which found that, after 3 years, women who tested negative on the HPV test had an extremely low risk of developing cervical cancer--about half the already low risk of women who tested negative on the Pap test.

First-line HPV testing has not yet been incorporated into the current professional cervical cancer screening guidelines. Professional societies are developing interim guidance documents, and some medical practices might incorporate primary HPV screening.

Most laboratories in the United States use a standard set of terms, called the Bethesda System, to report Pap test results. Under the Bethesda System, samples that have no cell abnormalities are reported as "negative for intraepithelial lesion or malignancy." A negative Pap test report may also note certain benign (non-neoplastic) findings, such as common infections or inflammation. Pap test results also indicate whether the specimen was satisfactory or unsatisfactory for examination. The Bethesda System considers abnormalities of squamous cells and glandular cells separately. Squamous cell abnormalities are divided into the following categories, ranging from the mildest to the most severe.

Atypical squamous cells (ASC) are the most common abnormal finding in Pap tests. The Bethesda System divides this category into two groups, ASC-US and ASC-H. ASC-US indicates atypical squamous cells of undetermined significance. The squamous cells do not appear completely normal, but doctors are uncertain about what the cell changes mean. The changes may be related to an HPV infection, but they can also be caused by other factors. ASC-H indicates atypical squamous cells, cannot exclude a high-grade squamous intraepithelial lesion. The cells do not appear normal, but doctors are uncertain about what the cell changes mean. ASC-H lesions may be at higher risk of being precancerous compared with ASC-US lesions.

Low-grade squamous intraepithelial lesions (LSILs) are considered mild abnormalities caused by HPV infection. Low-grade means that there are early changes in the size and shape of cells. Intraepithelial refers to the layer of cells that forms the surface of the cervix. When cells from the abnormal area are removed and examined under a microscope (in a procedure called a biopsy), LSILs are usually found to have mild cell changes that may be classified as mild dysplasia or as cervical intraepithelial neoplasia, grade 1 (CIN-1).

High-grade squamous intraepithelial lesions (HSILs) are more severe abnormalities that have a higher likelihood of progressing to cancer if left untreated. High-grade means that there are more evident changes in the size and shape of the abnormal (precancerous) cells and that the cells look very different from normal cells. When examined under a microscope, the cells from HSILs are often found to have more extensive changes that may be classified as moderate or severe dysplasia or as CIN-2, CIN-2/3, or CIN-3 (in order of increasing severity). Microscopic examination of HSILs may also reveal carcinoma in situ (CIS), which is commonly included in the CIN-3 category.

Squamous cell carcinoma is cervical cancer. The abnormal squamous cells have invaded more deeply into the cervix or into other tissues or organs. In a well-screened population, such as that in the United States, a finding of cancer during cervical screening is extremely rare.

Despite the successes of cervical screening, there is a great need for additional methods of detecting subjects with an increased risk of developing cervical cancer following HPV infection and development of an HPV cervical lesion. In western societies additional methods will allow a further fine tuning of monitoring protocols that is expected to provide an even higher success rate at limiting the development of cervical cancer from cervical lesions. In other countries with less robust cervical screening, additional methods of detecting subjects with an increased risk of developing cervical cancer following HPV infection and development of an HPV cervical lesion will allow greater resources to be committed to monitoring those subjects with the greatest risks of developing cervical cancer and thus have significant public health benefits. The possibility of identifying populations of subjects at highest risk of developing cervical cancer would also be very useful. For these and other reasons there is a need for new methods of monitoring a cervical HPV infection and/or an HPV induced cervical lesion in a subject. For these and other reasons there is a need for new methods of identifying a subject with a cervical HPV infection and/or an HPV

induced cervical lesion, with an increased risk of developing cervical squamous cell carcinoma. The inventions provided herein meet these and other needs.

SUMMARY

5

This invention relates to methods for monitoring a cervical HPV infection or HPV induced cervical lesion in a subject and/or respectively detecting a subject's increased risk of progressing to and/or developing cervical squamous cell carcinoma, said subject being a high risk HPV-positive subject or a subject with an HPV induced cervical lesion. The methods may comprise a) providing a nucleic acid sample from a high risk HPV-positive subject or respectively a subject with an HPV-induced cervical lesion; b) performing a genotyping assay on the nucleic acid sample to determine whether the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$); and c) concluding about the follow-up of the subject and/or respectively concluding on the subject's increased risk of progressing to and/or developing cervical squamous cell carcinoma. Particular embodiments of the invention are further detailed herein.

According to a particular embodiment, the methods of the invention as defined herein are *in vitro* methods, performed on a sample removed from the body of a subject.

This invention encompasses methods of monitoring a cervical HPV infection in a subject. The methods may comprise a) providing a nucleic acid sample from a high risk HPV-positive subject; b) performing a genotyping assay on the nucleic acid sample to determine whether the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$); and c) advising the subject to receive a next Pap test based on the A3 genotype of the subject; wherein, if the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$) then the subject is advised to receive the next Pap test at a shorter time interval than if the A3 genotype of the subject does not comprise at least one copy of a A3A-A3B deletion allele ($\Delta A3B$). In some embodiments, the subject is heterozygous for the $\Delta A3B$ allele. In some embodiments, the subject is heterozygous for the $\Delta A3B$ allele and the methods further comprise d) performing a Pap test on the subject at an increased frequency relative to the Pap test frequency for a subject with an A3 genotype not known to comprise at least one copy of a A3A-A3B deletion allele ($\Delta A3B$). In some embodiments, the subject is homozygous for the $\Delta A3B$

allele. In some embodiments, the subject is homozygous for the $\Delta A3B$ allele and the methods further comprise d) performing a Pap test on the subject at an increased frequency relative to the Pap test frequency for a subject with an A3 genotype not known to comprise at least one copy of a A3A-A3B deletion allele ($\Delta A3B$). In some embodiments, if the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$) then the subject is advised to receive the next Pap test after a time interval of less than one year. In some embodiments, the nucleic acid sample is from a subject that has been diagnosed with a low-grade squamous intraepithelial lesion (LSIL) of the cervix. In some embodiments, the subject has been diagnosed with a cervical intraepithelial neoplasia grade 1 (CIN-1). In some embodiments, the nucleic acid sample is from a subject that has been diagnosed with a high-grade squamous intraepithelial lesion (HSIL) of the cervix. In some embodiments, the subject has been diagnosed with a cervical intraepithelial neoplasia grade 2 (CIN-2), cervical intraepithelial neoplasia grade 2/3 (CIN-2/3), or cervical intraepithelial neoplasia grade 3 (CIN-3).

This invention also encompasses methods of monitoring an HPV induced cervical lesion in a subject. The methods may comprise a) providing a nucleic acid sample from a subject with an HPV-induced cervical lesion; b) performing a genotyping assay on the nucleic acid sample to determine whether the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$); and c) advising the subject to receive a next Pap test based on the A3 genotype of the subject; wherein, if the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$) then the subject is advised to receive the next Pap test at a shorter time interval than if the A3 genotype of the subject does not comprise at least one copy of a A3A-A3B deletion allele ($\Delta A3B$). In some embodiments, the subject is heterozygous for the $\Delta A3B$ allele. In some embodiments, the subject is heterozygous for the $\Delta A3B$ allele and the methods further comprise d) performing a Pap test on the subject at an increased frequency relative to the Pap test frequency for a subject with an A3 genotype not known to comprise at least one copy of a A3A-A3B deletion allele ($\Delta A3B$). In some embodiments, the subject is homozygous for the $\Delta A3B$ allele. In some embodiments, the subject is homozygous for the $\Delta A3B$ allele and the methods further comprise d) performing a Pap test on the subject at an increased frequency relative to the Pap test frequency for a subject with an A3 genotype not known to comprise at least one copy of a A3A-A3B deletion allele ($\Delta A3B$). In some embodiments,

if the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$) then the subject is advised to receive the next Pap test after a time interval of less than one year. In some embodiments, the nucleic acid sample is from a subject that has been diagnosed with a low-grade squamous intraepithelial lesion (LSIL) of the cervix. In some embodiments, the subject has been diagnosed with a cervical intraepithelial neoplasia grade 1 (CIN-1). In some embodiments, the nucleic acid sample is from a subject that has been diagnosed with a high-grade squamous intraepithelial lesion (HSIL) of the cervix. In some embodiments, the subject has been diagnosed with a cervical intraepithelial neoplasia grade 2 (CIN-2), cervical intraepithelial neoplasia grade 2/3 (CIN-2/3), or cervical intraepithelial neoplasia grade 3 (CIN-3). In some embodiments, the nucleic acid sample is from a subject that has tested positive for a high risk type of HPV. In some embodiments, the methods further comprise providing a cervical sample from the subject and testing the cervical sample for the presence of a high risk type of HPV.

This invention also encompasses methods of identifying a subject with an HPV induced cervical lesion with an increased risk of developing cervical squamous cell carcinoma. The methods may comprise a) providing a nucleic acid sample from a subject with an HPV induced cervical lesion; b) performing a genotyping assay on the nucleic acid sample to determine whether the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$); and c) diagnosing the subject as having an increased risk of cervical squamous cell carcinoma if the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$). In some embodiments, the subject is heterozygous for the $\Delta A3B$ allele. In some embodiments, the subject is homozygous for the $\Delta A3B$ allele. In some embodiments, the nucleic acid sample is from a subject that has been diagnosed with a low-grade squamous intraepithelial lesion (LSIL) of the cervix. In some embodiments, the subject has been diagnosed with a cervical intraepithelial neoplasia grade 1 (CIN-1). In some embodiments, the nucleic acid sample is from a subject that has been diagnosed with a high-grade squamous intraepithelial lesion (HSIL) of the cervix. In some embodiments, the subject has been diagnosed with a cervical intraepithelial neoplasia grade 2 (CIN-2), cervical intraepithelial neoplasia grade 2/3 (CIN-2/3), or cervical intraepithelial neoplasia grade 3 (CIN-3). In some embodiments, the nucleic acid sample is from a subject that has tested positive for a high risk type of HPV.

In some embodiments, the methods further comprise providing a cervical sample from the subject and testing the cervical sample for the presence of a high risk type of HPV.

This invention also encompasses methods of identifying a high risk HPV-positive subject with an increased risk of developing cervical squamous cell carcinoma. The methods may
5 comprise a) providing a nucleic acid sample from a high risk HPV-positive subject; b) performing a genotyping assay on the nucleic acid sample to determine whether the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$); and c) diagnosing the subject as having an increased risk of developing cervical squamous cell carcinoma if the A3 genotype of the subject comprises at least one copy of a A3A-A3B
10 deletion allele ($\Delta A3B$).

In some embodiments, the subject is heterozygous for the $\Delta A3B$ allele. In some embodiments, the subject is homozygous for the $\Delta A3B$ allele. In some embodiments, the nucleic acid sample is from a subject that has been diagnosed with a low-grade squamous intraepithelial lesion (LSIL) of the cervix. In some embodiments, the subject has been
15 diagnosed with a cervical intraepithelial neoplasia grade 1 (CIN-1). In some embodiments, the nucleic acid sample is from a subject that has been diagnosed with a high-grade squamous intraepithelial lesion (HSIL) of the cervix. In some embodiments, the subject has been diagnosed with a cervical intraepithelial neoplasia grade 2 (CIN-2), cervical intraepithelial neoplasia grade 2/3 (CIN-2/3), or cervical intraepithelial neoplasia grade 3
20 (CIN-3).

BRIEF DESCRIPTION OF THE DRAWINGS

Figs. 1A to 1C show details of A3B molecular clones and cytidine deamination activity.
25 **a)** ClustalW alignment of A3B proteins. Only variable residues are shown. (consecutive SEQ ID NOs : 1 to 4) **b)** Representative Western blot analysis from 3 independent experiments of V5-tagged A3B in human HEK 293T-UGI cells. β -Actin was used as loading control. **c)** In vitro deamination assay performed on TAMRA-FAM coupled oligonucleotide using transfected 293T lysates from two experiments performed in
30 duplicates. Background fluorescence obtained with mock-transfected cells was subtracted, and A3A A3AC106S and A3BE255Q catalytic mutants were used as a negative control.

Figs. 2A to 2H show the role of A3B mutator enzymes in causing nuclear DNA damage.

a) TP53 specific 3DPCR gels after 293T-UGI transfections with A3 proteins. Asterisks denote samples that were cloned and sequenced. **b)** Selection of hypermutated TP53 sequences after A3 transfection in 293T-UGI cells. (consecutive SEQ ID NOs : 5 to 14) **c)** Mutation matrices of hyperedited sequences on TP53 minus strand DNA. The numbers below the matrices indicate the number of bases sequenced. **d)** 5' Dinucleotide analysis of the deamination context performed on minus strand DNA. * denotes significant deviation from expected values (χ^2 test, $p < 0.05$, $n = 35$ for A3A, $n = 64$ for A3Bi7, $n = 65$ for A3Btok). **e)** FACS analysis of γ H2AX-positive HeLa and QT6 cells gated on V5 positive cells after A3 transfection at 48 hours. Error bars represent standard deviation from six independent transfections. **f)** FACS analysis of early apoptosis (Annexin V positive, PI negative cells - white) and late apoptosis/necrosis (Annexin V, PI double positive - red) 36 hours post-transfection. Differences compared to mock transfected cells were calculated using the Mann-Whitney test (* $p < 0.05$). Error bars represent standard deviation from four independent transfections. Top error bars relate to Annexin staining while lower error bars represent standard deviation for PI staining among Annexin + cells. **g)** 3DPCR gels after QT6 transfections with A3 constructs along with 5Me-dC substituted HIV-1 env gene matrix. Asterisks denote the samples that were cloned and sequenced. **h)** 5' Dinucleotide analysis of the deamination context performed on minus strand DNA. * denotes significant deviation from expected values (χ^2 test, $p < 0.05$, $n = 25$ for A3A, $n = 29$ for A3Bi7, $n = 14$ for A3Btok, $n = 21$ for A3Blan).

Figs. 3A to 3H A3B deletion locus and enhanced A3A mutation of nuDNA. **a)** The A3A-A3B locus and 29.5 kb chimeric deletion allele along with the three transcripts. **b)** A3A-3'UTR and Luc-3'UTR expression constructs cloned into pcDNA3.1 expression vector. **c)** Representative Western blot analysis from 4 independent experiments of HA-tagged A3A levels in 293T cells 48 hours post transfection with β -actin as loading control. Average expression level is given by densitometric ratio (DR) normalized on actin loading **d)** Transcriptome analysis of the three constructs using oligonucleotides spanning intron 4 and normalized to the expression levels of the housekeeping gene RPL13A. Error bars represent standard deviation from six independent transfections. Differences compared to Luc- Δ UTR, were calculated using the Mann-Whitney test (** $p < 0.01$). **e)** Quantitation of UTR-tagged firefly luciferase normalized on Renilla luciferase activity in 293T cells 48

hours after transfection. Error bars represent standard deviation from three independent transfections, measured in triplicates. Differences compared to Luc- Δ UTR were calculated using the Mann-Whitney test (**** $p < 0.0001$). **f**) 3DPCR gel gradients gels for TP53 gene after 293T-UGI transfections with HA-A3 proteins. **g**) FACS analysis of γ H2AX-positive HeLa cells 48 hours after transfection gated on the HA positive cells after A3 transfection. Error bars represent standard deviation from four independent transfections. Differences compared to A3A- Δ UTR were calculated using the Mann-Whitney test (* $p < 0.05$; ** $p < 0.01$). **h**) γ H2AX-positive double stranded breaks in the Δ A3B-/- cell line SKBR3 treated with 100 μ M PMA, 100 μ M etoposide and DMSO solvent control (red). Level of endogenous A3A transcript normalized to the expression levels of the housekeeping gene RPL13A. Presented results are from two independent experiments measured in duplicates.

Figs. 4A to 4B show cellular localization of A3B proteins and HBV editing. **a**) Confocal microscopy of V5 tagged A3A proteins performed in HeLa cells 24 hours post transfection. Nuclei are stained using DAPI. **b**) 3D-PCR analysis of HBV DNA editing by A3B proteins.

Figs. 5A to 5E show A3B nuDNA editing and 5-methylcytidine deamination. **a**) Quail cMyc specific 3DPCR gels after QT6 cells transfections with A3 proteins along with UGI coding plasmid. Asterisks denote the samples that were molecularly cloned and sequenced. **b**) Mutation matrices of cMyc hyperedited sequences. The numbers below the matrices indicate the number of bases sequenced. **c**) 5' Dinucleotide analysis of the deamination context performed on DNA minus strand for cloned PCR products. * denote significant deviation from expected values (χ^2 test, $p < 0.05$, $n = 18$ for A3A, $n = 23$ for A3Bi7). **d**) Selection of hypermutated 5Me-dC substituted HIV-1 sequences after co-transfection along with A3 coding plasmids in QT6 cells. (consecutive SEQ ID NOs : 15 to 27) **e**) Mutation matrices of HIV-1 hyperedited sequences. The numbers below the matrices indicate the number of bases sequenced. Data presented correspond to 3DPCR products obtained at 81.3°C.

Fig. 6A shows FACS plots of γ H2AX staining on V5 positive HeLa cells gated on V5 (pink) after A3 (2 μ g) transfection performed at 48 hours. Cells. Blue histograms represent γ H2AX staining for A3AC106S catalytic mutant transfection used as negative control.

Fig. 6B shows FACS plots of γ H2AX staining on V5 positive QT6 cells gated on V5 (pink) after A3 (2 μ g) transfection performed at 48 hours. Cells. Blue histograms represent γ H2AX staining for A3AC106S catalytic mutant transfection used as negative control.

Figs. 7A to 7B show FACS analysis of γ H2AX-positive HeLa cells gated on V5 after A3 (1 μ g) co- transfection with mock (1 μ g) or UGI (1 μ g) coding plasmids at 48 hours. **a)** FACS plots of γ H2AX staining on V5 positive cells. **b)** Red histograms represent γ H2AX staining for transfections performed with 1 μ g of mock plasmid, white histograms represent γ H2AX staining for transfections performed with UGI coding plasmid. Error bars represent standard deviation from four independent transfections. Difference between mock transfected cells and UGI transfected cells for A3A was calculated using student test (*p= 0.000375).

Figs. 8 shows a sequence comparison of the 3'UTR encoded by exon 5 of A3A and exon 8 of A3B genes used in the study (SEQ ID NOs : 28 and 29). An imperfect repeated sequence flanking the Alu element is underlined. A3A and A3B transcripts are identified by their accession numbers and their 3' ends denoted by arrowheads (<).

Figs. 9A to 9B show A3 cytidine deamination. **a)** FACS plots of γ H2AX staining on HA positive HeLa cells gated on HA (pink) after A3 (2 μ g) transfection performed at 48 hours. Cells. Blue histograms represent γ H2AX staining for A3AC106S catalytic mutant transfection used as negative control. **b)** FACS plots of γ H2AX staining of SKBR3 cells after treatment. Blue histograms represent γ H2AX staining for DMSO treated cells used as negative control.

Figs. 10A to 10B show uncropped scans of western blots from **a)** Figure 1B, and **b)** Figure 3C.

Fig. 11 shows an example of a nucleic acid target sequence as discussed herein (SEQ ID NO: 87), corresponding to the nucleic acid sequence upstream (5') to the APOBEC3B gene in the human genome. The numbering refers to the position of the sequence on human chromosome 22. Examples of primers are underlined, the arrowheads indicating the sense of the PCR primer (5' forward or 3' reverse)

DETAILED DESCRIPTION

A. Introduction

The human seven gene *APOBEC3* (*A3*) cluster locus encodes 6 functional polynucleotide cytidine deaminases (*A3A-C*, *A3F-H*) some of which are involved in cellular defense against viruses and retroelements through cytidine deamination of single stranded DNA (ssDNA) (1). *A3A* notably is able to hypermutate nuclear DNA (nuDNA) resulting in the formation of double strand DNA breaks (DSB) and apoptosis (2-4). In addition *A3A* could deaminate DNA molecules harboring a variety of 5-oxidized cytidine bases including 5Me-dC (5-7). These findings parallel the accumulation of cancer genome sequences which revealed that there were far more somatic mutations per genome than hitherto imagined, the majority being, with notable exceptions, GC->AT transitions (8-11). In view of the above and the hallmark 5'TpC editing signature associated with cytidine deamination, it is now accepted that *A3* molecules are involved in the genesis of cancer genomes (9, 12-14).

A3A and the C-terminal domain of *A3B* differ by 9% at the amino acid level. *A3B* has been clearly shown to edit viral genomes such as HIV (15), HTLV (16, 17) and HBV (18, 19). Two recent reports suggested *A3B* editing of nuDNA (20, 21). However, the degree of editing, of the order of a few mutations per kilobase, was unlike that observed for *A3A* (hundreds per kilobase), and probably represents PCR background (22). The role of *A3B* in genome editing is intriguing given that deletion of most of the *A3B* gene results in a higher odds ratio of developing breast, ovarian or liver cancer (23-26). Indeed, complete genome sequencing of $\Delta A3B^{-/-}$ breast cancer genomes revealed a higher than average mutation burden greatly consolidating these genetic studies (27). The 29.5kb deletion leaves a chimeric *A3A-A3B* transcript that does not affect the *A3A* protein sequence but terminates with the *A3B* 3'UTR. Yet it is contra-intuitive that deletion of a nuDNA mutator enzyme increases the risk of cancer, especially as there was a heterozygous effect suggesting a gene dosage effect. As this is a highly prevalent polymorphism with frequencies ~30-40% in south east Asia and China and >90% in Oceania (28), understanding the underlying mechanism is of concern. It turns out that while both *A3A* and *A3B* can generate mutations in nuDNA, *A3A* alone is capable of inducing detectable DSBs. Moreover, it transpires that the *A3A-A3B* deletion polymorphism results in increased chimeric *A3A* mRNA and protein levels which translate into more DNA damage. This may underlie an increased cancer predisposition of individuals harboring this allele.

Taken together those results indicate that A3B is an endogenous mutator of nuDNA with dC and 5Me-dC ssDNA substrate specificities, although it seems to access far less efficiently nuDNA than does A3A resulting in modest damage (Figs. 2a, e, f). As established breast cancer derived cell lines such as HCC1569 and MDA-MB468 express high levels of A3B (20) it would appear that only massive overexpression of A3B leads to experimentally detectable DNA damage.

As demonstrated in the examples presented herein, the naturally occurring 29.5kb deletion between a homology region at the 3' ends of A3A and A3B results in a chimeric A3A mRNA yielding 10-20 fold higher steady state levels of A3A leading to increased DNA damage. If a recent study points to lower A3A expression levels among $\Delta A3B$ cancers (27), interpreting tumor specific A3A and A3B transcription levels needs more work before making correlations with tumour incidence. They can be confounded by the presence of infiltrating hematopoietic cells that generally express much higher levels of A3 genes. A3A levels are generally very low compared to those of A3B, while A3A in effector cells can be massively induced by inflammatory responses (33, 39, 40). In addition, it is not obvious how the accumulation of mutations over a 5-15+ years are correlated to end stage tumour A3A and A3B transcription levels. However, as deletion of the entire A3B gene correlates with a higher risk of developing some cancers and an overall higher mutation burden per cancer genome (27), A3A is clearly an important player in cancer. As the deletion is associated with a higher odds ratio of developing breast, ovarian and liver cancer it constitutes a powerful cancer susceptibility marker. These mechanistic findings above tie in well with the recently identified higher mutation burden in breast cancer genomes (27) indicating that this is causal and not a surrogate marker. They help explain the heterozygous effect between $\Delta A3B$ and cancer (24-27).

Given the sensitivity of A3A to inflammatory environments involving type I and II interferons, while chronic inflammation has long been associated with the onset of cancer (41), it is distinctly possible that repeated A3A induced DNA damage will be a major driving force behind iterative somatic mutation/selection of the human genome, cutting across many cancer types. Cancer incidence will be modulated as a function of the global distribution of the $\Delta A3B$ polymorphism with South East Asia where the prevalence of this allele is 35-40% bearing a particularly heavy burden.

Because the $\Delta A3B$ allele is a causal marker in cancer, the presence of one or two copies of the $\Delta A3B$ allele in a subject's genome increases the risk that an HPV-induced cervical lesion will progress toward cervical cancer more rapidly than is the case in a subject with a genome that does not comprise at least one $\Delta A3B$ allele. On the basis of this discovery, the invention provides methods of monitoring a cervical HPV infection and/or an HPV induced cervical lesion in a subject. The invention also provides methods of identifying a subject with a cervical HPV infection and/or an HPV induced cervical lesion, with an increased risk of developing cervical squamous cell carcinoma are also provided. The methods comprise performing a genotyping assay on a nucleic acid sample of the subject to determine whether the A3 genotype of the subject comprises at least one copy of a $A3A-A3B$ deletion allele ($\Delta A3B$). In some embodiments, if the A3 genotype of the subject comprises at least one copy of a $A3A-A3B$ deletion allele ($\Delta A3B$) then the subject is advised to receive the next Pap test at a shorter time interval than if the A3 genotype of the subject does not comprise at least one copy of a $A3A-A3B$ deletion allele ($\Delta A3B$). In some embodiments, the subject is diagnosed as having an increased risk of developing cervical squamous cell carcinoma if the A3 genotype of the subject comprises at least one copy of a $A3A-A3B$ deletion allele ($\Delta A3B$). These and other aspects and embodiments of the invention are provided herein.

B. Definitions

Unless otherwise defined herein, scientific and technical terms used in connection with the present disclosure shall have the meanings that are commonly understood by those of ordinary skill in the art. Further, unless otherwise required by context, singular terms shall include the plural and plural terms shall include the singular. Generally, nomenclatures used in connection with, and techniques of, biochemistry, enzymology, molecular and cellular biology, microbiology, genetics and protein and nucleic acid chemistry and hybridization described herein are those well-known and commonly used in the art. Certain references and other documents cited herein are expressly incorporated herein by reference. Additionally, all UniProt/SwissProt records cited herein are hereby incorporated herein by reference. In case of conflict, the present specification, including definitions, will control. The materials, methods, and examples are illustrative only and not intended to be limiting.

The methods and techniques of the present disclosure are generally performed according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the present specification unless otherwise indicated. See, e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, 3d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2001); Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing Associates (1992, and Supplements to 2002); Taylor and Drickamer, Introduction to Glycobiology, Oxford Univ. Press (2003); Worthington Enzyme Manual, Worthington Biochemical Corp., Freehold, N.J.; Handbook of Biochemistry: Section A Proteins, Vol I, CRC Press (1976); Handbook of Biochemistry: Section A Proteins, Vol II, CRC Press (1976); Essentials of Glycobiology, Cold Spring Harbor Laboratory Press (1999).

Before the present proteins, compositions, methods, and other embodiments are disclosed and described, it is to be understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting. It must be noted that, as used in the specification and the appended claims, the singular forms "a," "an" and "the" include plural referents unless the context clearly dictates otherwise.

The term "comprising" as used herein is synonymous with "including" or "containing", and is inclusive or open-ended and does not exclude additional, unrecited members, elements or method steps.

As used herein, the term "isolated" refers to a substance or entity that has been (1) separated from at least some of the components with which it was associated when initially produced (whether in nature or in an experimental setting), and/or (2) produced, prepared, and/or manufactured by the hand of man. Isolated substances and/or entities may be separated from at least about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, or more of the other components with which they were initially associated. In some embodiments, isolated agents are more than about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more than about 99% pure. As used herein, a substance is "pure" if it is substantially free of other components.

The term "polynucleotide", "nucleic acid molecule", "nucleic acid", or "nucleic acid sequence" refers to a polymeric form of nucleotides of at least 10 bases in length. The

term includes DNA molecules (e.g., cDNA or genomic or synthetic DNA) and RNA molecules (e.g., mRNA or synthetic RNA), as well as analogs of DNA or RNA containing non-natural nucleotide analogs, non-native internucleoside bonds, or both. The nucleic acid can be in any topological conformation. For instance, the nucleic acid can be single-stranded, double-stranded, triple-stranded, quadruplexed, partially double-stranded, branched, hairpinned, circular, or in a padlocked conformation. The nucleic acid (also referred to as polynucleotides) may include both sense and antisense strands of RNA, cDNA, genomic DNA, and synthetic forms and mixed polymers of the above. They may be modified chemically or biochemically or may contain non-natural or derivatized nucleotide bases, as will be readily appreciated by those of skill in the art. Such modifications include, for example, labels, methylation, substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoramidates, carbamates, etc.), charged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), pendent moieties (e.g., polypeptides), intercalators (e.g., acridine, psoralen, etc.), chelators, alkylators, and modified linkages (e.g., alpha anomeric nucleic acids, etc.) Also included are synthetic molecules that mimic polynucleotides in their ability to bind to a designated sequence via hydrogen bonding and other chemical interactions. Such molecules are known in the art and include, for example, those in which peptide linkages substitute for phosphate linkages in the backbone of the molecule. Other modifications can include, for example, analogs in which the ribose ring contains a bridging moiety or other structure such as the modifications found in "locked" nucleic acids.

In general, "stringent hybridization" is performed at about 25°C below the thermal melting point (T_m) for the specific DNA hybrid under a particular set of conditions. "Stringent washing" is performed at temperatures about 5°C lower than the T_m for the specific DNA hybrid under a particular set of conditions. The T_m is the temperature at which 50% of the target sequence hybridizes to a perfectly matched probe. See Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), page 9.51. For purposes herein, "stringent conditions" are defined for solution phase hybridization as aqueous hybridization (i.e., free of formamide) in 6xSSC (where 20xSSC contains 3.0 M NaCl and 0.3 M sodium citrate), 1% SDS at 65°C for 8-12 hours, followed by two washes in 0.2xSSC, 0.1% SDS at 65°C

for 20 minutes. It will be appreciated by the skilled worker that hybridization at 65°C will occur at different rates depending on a number of factors including the length and percent identity of the sequences which are hybridizing.

As used herein, “a shorter time interval” is used to describe the period of time between successive Pap tests of a subject in relation to a reference period of time between successive Pap tests. Specifically, depending on the genotype of a test subject, the test subject may be advised to receive a next Pap test sooner than if the subject had a reference genotype. That means that if the subject had the reference genotype the subject would be advised to receive a next Pap test after a reference period of time has elapsed, whereas if the subject has a different genotype the subject is advised to receive a next Pap test after a period of time has elapsed that is shorter than the reference period of time. That is, after a “shorter time interval.”

A “high risk type of HPV” may be any type of HPV that induces formation of cervical cancer in a subject. In some embodiments a high risk type of HPV is a type of HPV selected from HPV types 16 and 18. In some embodiments a high risk type of HPV is a type of HPV selected from HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68. In some embodiments a high risk type of HPV is a type of HPV selected from HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 69, 73, and 82. In some embodiments a high risk type of HPV is a type of HPV selected from HPV types 6, 11, 16, 18, 26, 31, 33, 35, 39, 40, 42, 45, 51, 52, 53, 54, 55, 56, 58, 59, 61, 62, 64, 66, 67, 68, 69, 70, 71, 72, 73 (MM9) (novel type related to HPV73), 81, 82 (MM4) (novel type related to HPV82), 83 (MM7) (novel type related to HPV83), 84 (MM8) (novel type related to HPV84), IS39 and CP6108.

Other terms are defined throughout the disclosure.

C. Methods of Identifying Subjects With An Increased Risk of Developing Cervical Squamous Cell Carcinoma

Cervical squamous cell carcinoma is caused by a high risk HPV infection over several years. This process thus relies on several mutation events occurring over time. Accordingly, the combination of infection by a high risk type of HPV and the presence in a subject’s genome of at least one copy of a A3A-A3B deletion allele ($\Delta A3B$) predisposes

the subject to to a higher risk of developing a cervical squamous cell carcinoma. It is also presently believed that infection by a high risk type of HPV and the presence in a subject's genome of at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$) predisposes the subject to to a higher risk of an LSIL progressing into an HSIL, and of a CIN-1 lesion progressing to become a CIN-2 lesion, a CIN-2 lesion progressing to become a CIN-2/3 lesion, and a CIN-2/3 lesion progressing to become a CIN-3 lesion. Accordingly, the invention provides methods for identifying a subject with an HPV induced cervical lesion and/or a high risk HPV infection, wherein the subject has an increased risk of developing cervical squamous cell carcinoma.

In some embodiments the methods comprise a) providing a nucleic acid sample from a subject with an HPV induced cervical lesion and/or a high risk HPV infection; and b) performing a genotyping assay on the nucleic acid sample to determine whether the A3 genotype of the subject comprises at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$). The results of the genotyping assay may be used to characterize the subject with respect to at least one feature of cervical cancer risk, such as an increased risk of developing cervical squamous cell carcinoma.

In some embodiments the methods comprise a) providing a nucleic acid sample from a subject with an HPV induced cervical lesion and/or a high risk HPV infection; b) performing a genotyping assay on the nucleic acid sample to determine whether the A3 genotype of the subject comprises at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$); and c) diagnosing the subject as having an increased risk of cervical squamous cell carcinoma if the A3 genotype of the subject comprises at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$). In some embodiments the subject is heterozygous for the $\Delta A3B$ allele and, therefore, the subject has an increased risk of cervical squamous cell carcinoma compared to a subject with an A3 genotype that does not comprise at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$). In some embodiments the subject is homozygous for the $\Delta A3B$ allele and, therefore, the subject has an increased risk of cervical squamous cell carcinoma compared to a subject with an A3 genotype that does not comprise at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$). In some embodiments the subject also has an increased risk of developing cervical squamous cell carcinoma compared to a subject who is heterozygous for the $\Delta A3B$ allele.

In some embodiments the nucleic acid sample is from a subject that has been diagnosed with a low-grade squamous intraepithelial lesion (LSIL) of the cervix. In some embodiments the LSIL is a cervical intraepithelial neoplasia grade 1 (CIN-1). In some embodiments the subject thus has a higher risk of developing a high-grade squamous intraepithelial lesion (HSIL). In some embodiments the HSIL is a cervical intraepithelial neoplasia grade 2 (CIN-2), cervical intraepithelial neoplasia grade 2/3 (CIN-2/3), or cervical intraepithelial neoplasia grade 3 (CIN-3).

In some embodiments the subject has been diagnosed with an HPV induced cervical lesion and the method further comprises testing the subject for the presence of a high-risk HPV type infection.

In some embodiments the subject has been diagnosed with a high risk HPV infection and the method further comprises performing a Pap test on the subject.

D. Methods of Monitoring a Cervical HPV Infection in a Subject

Because cervical squamous cell carcinoma is caused by a high risk HPV infection over several years, monitoring subjects having a cervical infection by a high risk type of HPV and/or subjects having a cervical lesion characteristic of HPV infection is very important. This invention provides new methods of monitoring such subjects based, in part, on the A3 genotype of the subject, and in particular based on whether the subject's genome comprises at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$). Accordingly, the invention provides methods of monitoring a cervical HPV infection in a subject who is positive for a high risk HPV infection and/or has an HPV-induced cervical lesion.

The methods may comprise a) providing a nucleic acid sample from a subject with an HPV induced cervical lesion and/or a high risk HPV infection; b) performing a genotyping assay on the nucleic acid sample to determine whether the A3 genotype of the subject comprises at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$); and c) advising the subject to receive a next Pap test based on the A3 genotype of the subject; wherein, if the A3 genotype of the subject comprises at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$) then the subject is advised to receive the next Pap test at a shorter time interval than if the A3 genotype of the subject does not comprise at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$). In some embodiments the shorter time interval is one month from

the previous Pap test, two months from the previous Pap test, three months from the previous Pap test, four months from the previous Pap test, five months from the previous Pap test, six months from the previous Pap test, seven months from the previous Pap test, eight months from the previous Pap test, nine months from the previous Pap test, ten months from the previous Pap test, eleven months from the previous Pap test, twelve months (one year) from the previous Pap test, fifteen months from the previous Pap test, eighteen months from the previous Pap test, or two years from the previous Pap test.

In some embodiments the subject is heterozygous for the $\Delta A3B$ allele and, therefore, the subject has an increased risk of cervical squamous cell carcinoma compared to a subject with an A3 genotype that does not comprise at least one copy of a $A3A-A3B$ deletion allele ($\Delta A3B$). In some embodiments the subject is homozygous for the $\Delta A3B$ allele and, therefore, the subject has an increased risk of cervical squamous cell carcinoma compared to a subject with an A3 genotype that does not comprise at least one copy of a $A3A-A3B$ deletion allele ($\Delta A3B$). In some embodiments the subject also has an increased risk of developing cervical squamous cell carcinoma compared to a subject who is heterozygous for the $\Delta A3B$ allele.

In some embodiments the subject comprises at least one copy of a $A3A-A3B$ deletion allele ($\Delta A3B$) and, therefore, the subject is advised to receive the next Pap test at a shorter time interval than if the A3 genotype of the subject does not comprise at least one copy of a $A3A-A3B$ deletion allele ($\Delta A3B$). In some embodiments the subject comprises two copies of a $A3A-A3B$ deletion allele ($\Delta A3B$) and, therefore, the subject is advised to receive the next Pap test at a shorter time interval than if the A3 genotype of the subject does not comprise at least one copy of a $A3A-A3B$ deletion allele ($\Delta A3B$). In some embodiments the subject is also advised to receive the next Pap test at a shorter time interval than if the A3 genotype of the subject is heterozygous for the $A3A-A3B$ deletion allele ($\Delta A3B$).

In some embodiments the nucleic acid sample is from a subject that has been diagnosed with a low-grade squamous intraepithelial lesion (LSIL) of the cervix. In some embodiments the LSIL is a cervical intraepithelial neoplasia grade 1 (CIN-1). In some embodiments the subject thus has a higher risk of developing a high-grade squamous intraepithelial lesion (HSIL). In some embodiments the HSIL is a cervical intraepithelial

neoplasia grade 2 (CIN-2), cervical intraepithelial neoplasia grade 2/3 (CIN-2/3), or cervical intraepithelial neoplasia grade 3 (CIN-3).

In some embodiments the subject has been diagnosed with an HPV induced cervical lesion and the method further comprises testing the subject for the presence of a high-risk
5 HPV type infection.

In some embodiments the subject has been diagnosed with a high risk HPV infection and the method further comprises performing a Pap test on the subject.

In some embodiments the subject is advised to receive the next Pap test after a time interval of less than one year. In some embodiments the subject is advised to receive the
10 next Pap test after a time interval of from six months to less than one year. In some embodiments the subject is advised to receive the next Pap test after a time interval of from six months to nine months. In some embodiments the subject is advised to receive the next Pap test after a time interval of from nine months to less than one year. In some
15 embodiments the subject is advised to receive the next Pap test after a time interval of one, two, three, four, five, six, seven, eight, nine, ten, or eleven months.

The method can comprise treating an HPV positive subject with a treatment comprising administering a composition comprising at least one chimeric recombinant Bordetella sp. adenylate cyclase (CyaA) protein or fragment thereof. In some embodiments the CyaA
20 protein or fragment thereof comprises at least one inserted human papilloma virus (HPV) E6 and/or E7 epitope. Additional treatments are disclosed, for example, in Hung *et al.*, Expert Opin Biol Ther. 2008 April ; 8(4): 421–439.

The method can comprise treating an HPV positive subject, wherein the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$), with a treatment comprising administering a composition comprising a live-vector based vaccine.
25 For example, the live-vector based vaccine may be a bacterial vector-based vaccine, such as *Salmonella typhimurium* or *Listeria monocytogenes*, or a viral vector-based vaccine, such as adenovirus (AdV).

The method can comprise treating an HPV positive subject, wherein the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$), with a
30 treatment comprising administering a composition comprising a peptide and/or protein based vaccine. In preclinical studies, progress has been achieved in augmenting peptide vaccine potency by employing the intranasal route of administration, linking peptides to

immunostimulatory molecules to generate protective immunity and specific CTL responses and using DC-activating agents such as 4'-monophosphoryl lipid A (MPL) and GM-CSF to increase and sustain levels of CTL responses. Combining peptides with CpG oligodeoxynucleotide (CpG ODN), which provides a 'danger signal' for Toll-like receptor 9
5 by mimicking bacterial DNA, has also been shown to enhance the immunogenicity of peptide vaccines. Protein-based vaccination can circumvent the limited specificity of MHC responses associated with some peptide-based vaccines. Various protein vaccines have moved to clinical trials. Fusion proteins containing HPV capsid proteins and HPV early proteins can potentially induce prophylactic and therapeutic immune responses. One
10 example of this experimental fusion vaccine is TA-GW, a fusion of HPV-6 L2 and E7 absorbed onto Alhydrogel. It has been well tolerated by patients in two clinical trials and was effective in clearing HPV genital warts in a subset of patients. A vaccine containing an HPV-16 E6/E7 fusion protein mixed with the ISCOMATRIX adjuvant has also recently been tested in a Phase I study. Immunization with this protein-based vaccine was shown to
15 be safe and immunogenic and resulted in significantly enhanced CD8+ T cell responses to both E6 and E7 in vaccinated patients compared with those observed in placebo recipients. Another protein vaccine, TA-CIN, a fusion of HPV-16 L2, E6 and E7, induced antibodies in all the women tested and induced T cell immunity in a subset of them, proving to be safe. A vaccine termed PD-E7, comprised of mutated HPV-16 E7 fused with
20 a fragment of Haemophilus influenzae protein D and formulated in the GlaxoSmithKline Biologicals adjuvant AS02B, has been evaluated in Phase I/II clinical trials and was shown to induce significant E7-specific CTL responses in patients with CIN-1 and CIN-3 lesions. Recently, a fusion of HPV-16 E7 and M. bovis hsp65 has been shown to be well tolerated in patients with high-grade anal intraepithelial neoplasia (AIN); however, further tests are
25 needed to determine the clinical efficacy of the vaccine. A recent trial employing the same vaccine was conducted in women with CIN III lesions.

The method can comprise treating an HPV positive subject, wherein the A3 genotype of the subject comprises at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$), with a treatment comprising administering a composition comprising a DNA-based vaccine. DNA
30 vaccines have emerged as an attractive and potentially effective strategy for antigen-specific immunotherapy. Naked DNA is safe, stable, relatively easy to manufacture and can be used to sustain the expression of antigen in cells for longer periods of time than

RNA or protein vaccines. Furthermore, unlike live-vector vaccines, DNA vaccines do not elicit neutralizing antibody production in the patient, and thus can be repeatedly administered to the same patient effectively.

5 The method can comprise treating an HPV positive subject, wherein the A3 genotype of the subject comprises at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$), with a treatment comprising administering a composition comprising an RNA replicon-based vaccine. Use of RNA replicons is a relatively new and potentially interesting strategy for HPV vaccination. RNA replicons are naked RNA molecules that replicate within transfected cells. They may be derived from alphaviruses, such as Sindbis virus, Semliki Forest virus,
10 and VEE.

In some embodiments the treated subject is infected with a high risk HPV virus. In some embodiments the treated patient is infected with a human papilloma virus is of a type selected from types types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 69, 73, and 82.

15 **E. Detection of Infection By a High Risk Type of Human Papilloma Virus (HPV)**

The invention encompasses methods comprising detecting the presence or absence of a high risk type of human papilloma virus (HPV). The method can comprise detecting the presence or absence of HPV DNA, RNA, or protein in a sample of a subject. The method
20 can comprise preparing nucleic acids from a cell sample and contacting the nucleic acids with an HPV specific primer or probe. The nucleic acids can be DNA and/or RNA. The method can be performed, for example, by using routine techniques in the art including by the use of commercially available products designed specifically to detect infection by high risk types of HPV.

25 The invention encompasses the use of techniques for detecting HPV based on DNA typing. For example, the COBAS (Roche) and APTIMA (GEN-PROBE) kits are PCR tests of specific targets intended for the qualitative in vitro detection of mRNA of the L1 gene from 17 types of human papillomavirus (HPV) virus considered High risk (HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68, 69, 73, and 82). Six different results are
30 obtainable: HPV16 positive or negative, HPV18 positive or negative, others 12 HPVs positive or negative. LINEAR ARRAY HPV Genotyping Test (Roche) is a qualitative test that detects 37 high- and low-risk human papillomavirus genotypes, including those

considered a significant risk factor for High-grade Squamous Intraepithelial (HSIL) progression to cervical cancer. This test is a qualitative in vitro test for the detection of Human Papillomavirus in clinical specimens. The test utilizes amplification of target DNAs by PCR of the late gene L1 of HPV DNA genotypes 6, 11, 16, 18, 26, 31, 33, 35, 39, 40, 42, 45, 51, 52, 53, 54, 55, 56, 58, 59, 61, 62, 64, 66, 67, 68, 69, 70, 71, 72, 73 (MM9) (novel type related to HPV73), 81, 82 (MM4) (novel type related to HPV82), 83 (MM7) (novel type related to HPV83), 84 (MM8) (novel type related to HPV84), IS39 and CP6108. The digene HC2 HPV DNA Test, developed by Qiagen, is based on Capture Hybridization of HPV DNAs (L1 gene) for the qualitative detection of 18 types (HPV 16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68 [68a], 73, 82MM4 [82IS39]) in cervical specimens.

More recently, NucliSENS EasyQ HPV was made available to qualitative detection of oncogenes E6 / E7 mRNAs of 5 specific High risk HPVs 16, 18, 31, 33 and 45. Detection of HPV E6 and E7 has been proposed as a better correlate of cancer development than HPV DNA.

In addition, WO2011/088573, describes a set of probes to detect and Identify 46 specifically targeted species of mucosal human papillomaviruses (HPV). These probes are used as a multiplex assay based on nested PCR amplification and the Luminex xMAP technology for genotyping DNA of L1 genes of HPV types 6, 11, 13, 16, 18, 26, 30, 31, 32, 33, 35, 39, 40, 42, 43, 44, 45, 51, 52, 53, 54, 56, 58, 59, 61, 62, 66, 67, 68, 69, 70, 71, 72, 73, 74, 81, 82, 83, 84, 85, 86, 87, 89, 90, 91 and 97.

HPV can also be detected by preparing a nucleic acid from cells from a subject and sequencing the HPV nucleic acid. Any sequencing method known in the art can be employed.

HPV can also be detected by methods comprising sequencing HPV nucleic acids present in a sample of a subject. In some embodiments all or part of the E6 and/or E7 region of an HPV nucleic acid is sequenced. As used herein, the term "sequencing" is used in a broad sense and refers to any technique known by the skilled person including but not limited to Sanger dideoxy termination sequencing, whole-genome sequencing, sequencing by hybridization, pyrosequencing, capillary electrophoresis, cycle sequencing, single-base extension sequencing, solid-phase sequencing, high-throughput sequencing, massively parallel signature sequencing (MPSS), sequencing by reversible dye terminator, paired-end sequencing, near-term sequencing, exonuclease sequencing, sequencing by

ligation, short-read sequencing, single-molecule sequencing, sequencing-by-synthesis, real-time sequencing, reverse-terminator sequencing, nanopore sequencing, 454 sequencing, Solexa Genome Analyzer sequencing, SOLiD(R) sequencing, MS-PET sequencing, mass spectrometry, and a combination thereof. In specific embodiments, the method and kit of the invention is adapted to run on ABI PRISM(R) 377 DNA Sequencer, an ABI PRISM(R) 310, 3100, 3100-Avant, 3730, or 3730x1 Genetic Analyzer, an ABI PRISM(R) 3700 DNA Analyzer, or an Applied Biosystems SOLiD(TM) System (all from Applied Biosystems), a Genome Sequencer 20 System (Roche Applied Science), an HiSeq 2500, an HiSeq 2000, a Genome Analyzer IIx, a MiSeq Personal Sequencer, a HiScanSQ (all from Illumina), the Genetic Analysis System, including the Single Molecule Sequencer, Analysis Engine and Sample Loader (all from HeliScope), the Ion Proton™ Sequencer, or the Ion PGM™ Sequencer (both from Ion Torrent).

For all technologies described herein, although in some embodiments the primers are used in solution, in other embodiments the primers are linked to a solid support.

To permit its covalent coupling to the support, the primer is generally functionalized. Thus, it may be modified by a thiol, amine or carboxyl terminal group at the 5' or 3' position. In particular, the addition of a thiol, amine or carboxyl group makes it possible, for example, to couple the oligonucleotide to a support bearing disulphide, maleimide, amine, carboxyl, ester, epoxide, cyanogen bromide or aldehyde functions. These couplings form by establishment of disulphide, thioether, ester, amide or amine links between the primer and the support. Any other method known to a person skilled in the art may be used, such as bifunctional coupling reagents, for example.

Moreover, to improve the hybridization with the coupled oligonucleotide, it can be advantageous for the oligonucleotide to contain an "arm" and a "spacer" sequence of bases. The use of an arm makes it possible, in effect, to bind the primer at a chosen distance from the support, enabling its conditions of interaction with the DNA to be improved. The arm advantageously consists of a linear carbon chain, comprising 1 to 18 and preferably 6 or 12 (CH₂) groups, and an amine which permits binding to the column. The arm is linked to a phosphate of the oligonucleotide or of a "spacer" composed of bases which do not interfere with the hybridization. Thus, the "spacer" can comprise purine bases. As an example, the "spacer" can comprise the sequence GAGG. The arm is advantageously composed of a linear carbon chain comprising 6 or 12 carbon atoms.

For implementation of the present invention, different types of support may be used. These can be functionalized chromatographic supports, in bulk or prepacked in a column, functionalized plastic surfaces or functionalized latex beads, magnetic or otherwise. Chromatographic supports are preferably used. As an example, the chromatographic supports capable of being used are agarose, acrylamide or dextran as well as their derivatives (such as Sephadex, Sepharose, Superose, etc.), polymers such as poly(styrene/divinylbenzene), or grafted or ungrafted silica, for example. The chromatography columns can operate in the diffusion or perfusion mode.

In some embodiments, the invention is aimed at a method for determining a profile of sequences in one or more samples of patients suspected to be infected with or carrying an HPV, comprising detecting HPV sequences in one or more samples comprising: a) amplifying nucleic acid molecules in the sample ; b) spatially isolating individual molecules of said amplified nucleic acid molecules; c) optionally re-amplifying said amplified nucleic acid molecules; d) sequencing said re-amplified nucleic acid molecules; and e) determining the levels of different sequences from said sample to generate said profile of nucleic acid molecules in the sample. Said amplifying and/or re-amplifying comprises PCR, multiplex PCR, TMA, NASBA, or LAMP and spatially isolating individual molecules comprises separating molecules in two dimensions on a solid support, separating said molecules in three dimensions for example in a solution with micelles, or separating molecules using micro-reaction chambers. Said sequencing in step d) comprises dideoxy sequencing, sequencing by synthesis using reversibly terminated labeled nucleotides, pyrophosphate release on nucleotide incorporation, sequencing by synthesis using allele specific hybridization to a library of labeled oligonucleotide probes followed by ligation of said probes, real time monitoring of the incorporation of labeled nucleotides during a polymerization step.

Such amplification techniques include in particular isothermal methods and PCR-based techniques. Isothermal techniques include such methods as e.g. nucleic acid sequence-based amplification (NASBA), loop-mediated isothermal amplification (LAMP), helicase-dependent amplification (HDA), rolling circle amplification (RCA), and strand displacement amplification (SDA), exponential amplification reaction (EXPAR), isothermal and chimeric primer-initiated amplification of nucleic acids (ICANs), signal-mediated amplification of RNA technology (SMART) and others (see e.g. Asiello and Baeumner, Lab Chip; 11(8):

1420-1430, 2011). Preferably, the PCR technique used quantitatively measures starting amounts of DNA, cDNA, or RNA. Examples of PCR-based techniques according to the invention include techniques such as, but not limited to, quantitative PCR (Q-PCR), reverse-transcriptase polymerase chain reaction (RT-PCR), quantitative reverse-transcriptase PCR (QRT-PCR), or digital PCR (for example the Droplet Digital™ PCR technology sold by Bio-Rad, also known as ddPCR™). These techniques are well known and easily available technologies for those skilled in the art and do not need a precise description.

In some embodiments the determination of viral load is performed by quantitative PCR.

In some embodiments the determination of the viral load is performed by digital PCR. Digital PCR involves multiple PCR analyses on extremely dilute nucleic acids such that most positive amplifications reflect the signal from a single template molecule. Digital PCR thereby permits the counting of individual template molecules. The proportion of positive amplifications among the total number of PCRs analyzed allows an estimation of the template concentration in the original or non-diluted sample. This technique has been proposed to allow the detection of a variety of genetic phenomena (Vogelstein et al., Proc Natl Acad Sci USA 96: 9236-924, 1999). Since template molecule quantification by digital PCR does not rely on dose-response relationships between reporter dyes and nucleic acid concentrations, its analytical precision is, at least theoretically, superior to that of real-time PCR. Hence, digital PCR potentially allows the discrimination of finer degrees of quantitative differences between target and reference loci.

The primers are chosen by the person skilled in the art depending on the desired specificity of the PCR amplification step using standard parameters such as the nucleic acid size, GC contents, and temperature reactions.

The invention further encompasses a PCR or other amplified nucleic acid product comprising an HPV nucleic acid sequence. The amplified product can comprise any of the HPV nucleic acid sequences described herein. In some embodiments the amplification products are amplified using a biological sample containing an HPV nucleic acid. These products can be generated using the techniques set forth in the examples or other techniques known to the skilled artisan.

F. Detection of an A3A-A3B Deletion Allele

The human seven gene *APOBEC3* (*A3*) cluster locus encodes 6 functional polynucleotide cytidine deaminases (*A3A-C*, *A3F-H*). A naturally occurring 29.5kb deletion allele occurs between an identical 370 bp segment spanning intron 4/exon5 of *A3A* and intron 7/exon8 of *A3B*. (Figure 3a; see Reference 28, which is hereby incorporated herein by reference.) The deletion allele encodes a chimeric *A3-A3B* transcript that differs from normal *A3A* only by the 3'UTR – the four amino acids of *A3A* exon 5 and *A3B* exon 8 being identical (Fig. 3a and Fig. 8) (28). The major difference is the insertion of an Alu sequence in the *A3A* 3'UTR (Fig. 3a and Fig. 8).

As used herein, a “*A3A-A3B* deletion allele” or “ $\Delta A3B$ ” refers to an *A3* locus allele in which this 29.5 kb fragment is deleted. The presence of one or two copies of the $\Delta A3B$ allele in a subject may be detected using any technique known in the art. Generally, a nucleic acid sample from the subject will be provided and analysed to detect the presence or absence of at least one copy of an $\Delta A3B$ allele, to detect the presence of exactly one copy of an $\Delta A3B$ allele, or to detect the presence of exactly two copies of an $\Delta A3B$ allele. In some embodiments the method comprises directly analyzing the nuclear DNA of the subject. In some embodiments the method comprises analyzing the level and/or structure of an mRNA encoded by the *A3* locus of the subject. As described herein, the $\Delta A3B$ allele results in the presence of a unique mRNA and the mRNA is typically present in a subject at an elevated level compared to the native *A3A* mRNA. In some embodiments the method comprises analyzing the level of *A3A* protein in the subject.

A sequence of the chimeric transcript resulting from the presence of the *A3A-A3B* deletion allele in a human genome, according to the present disclosure, is provided below (SEQ ID NO: 85):

GGCCTCCCTGCCACGGGGATGCTGCCTTTTCTGCTCCGGGTGTTTCCACGAGGCAGGCATGGAATCT
TCCCTGGACAAGCGACATACCGTGGAGAGACAGGCTCCTGAATCCTAAGAGAATGTTGGTGAAGATC
TTAACACCACGCCTTGAGCAAGTCGCAAGAGCGGGAGGACACAGACCAGGAACCGAGAAGGGACAAG
CAC**ATG**GAAGCCAGCCCAGCATCCGGGCCAGACACTTGATGGATCCACACATATTCACCTTCCAAC
TTAACAATGGCATTGGAAGGCATAAGACCTACCTGTGCTACGAAGTGGAGCGCCTGGACAATGGCAC
CTCGGTCAAGATGGACCAGCACAGGGGCTTTCTACACAACCAGGCTAAGAATCTTCTCTGTGGCTTT
TACGGCCGCCATGCGGAGCTGCGCTTCTTGGACCTGGTTCCTTCTTTGCAGTTGGACCCGGCCAG
TCTACAGGGTCACTTGTTTCATCTCCTGGAGCCCCTGCTTCTCCTGGGGCTGTGCCGGGGAAGTGCG

TGCGTTCCTTCAGGAGAACACACACGTGAGACTGCGTATCTTCGCTGCCC GCATCTATGATTACGAC
CCCCATATAAGGAGGCACTGCAAATGCTGCGGGATGCTGGGGCCCAAGTCTCCATCATGACCTACG
ATGAATTTAAGCACTGCTGGGACACCTTTGTGGACCACCAGGGATGTCCCTTCCAGCCCTGGGATGG
ACTAGATGAGCACAGCCAAGCCCTGAGTGGGAGGCTGCGGGCCATTCTCCAGAATCAGGGAAACTGA
5 AGGATGGGCCTCAGTCTCTAAGGAAGGCAGAGACCTGGGTTGAGCAGCAGAATAAAAGATCTTCTTC
CAAGAAATGCAAACAGACCGTTCACCACCATCTCCAGCTGCTCACAGACACCAGCAAAGCAATGTGC
TCCTGATCAAGTAGATTTTTTAAAAATCAGAGTCAATTAATTTTAATTGAAAATTTCTCTTATGTTT
CAAGTGTACAAGAGTAAGATTATGCTCAATATTCCCAGAATAGTTTTCAATGTATTAATGAAGTGAT
TAATTGGCTCCATATTTAGACTAATAAAACATTAAGAATCTTCCATAATTGTTTCCACAAACACTAG
10 CAAATGTGTAGATGTCTTTTCTTGTGTAGCGGACCTGTAGCTGGGAAAGGTCACACAACATCCCTCT
GGATCCAGAAAACCTCAGCTAAACCACACAGGAGAGGAACCTAAATGCAGACCCACCCCTCACTCACA
GAGCCCCGCCACCCCTCACTCACAGAGCCCCGGGCGCTGATTGG

Initiation and Stop codons framing the sequence coding for the corresponding protein
are indicated in bold. The upstream and downstream sequence parts are non-coding
15 sequences.

A sequence coding for the chimeric protein resulting from the presence of the A3A-A3B
deletion allele in a human genome, according to the present disclosure, is therefore (SEQ
ID NO: 86):

ATGGAAGCCAGCCCAGCATCCGGGCCAGACACTTGATGGATCCACACATATTCACCTTCCAACCTTTA
20 ACAATGGCATTGGAAGGCATAAGACCTACCTGTGCTACGAAGTGGAGCGCCTGGACAATGGCACCTC
GGTCAAGATGGACCAGCACAGGGGCTTTCTACACAACCAGGCTAAGAATCTTCTCTGTGGCTTTTAC
GGCCGCCATGCGGAGCTGCGCTTCTTGGACCTGGTTCTTCTTGCAGTTGGACCCGGCCAGATCT
ACAGGGTCACTTGGTTCATCTCCTGGAGCCCCTGCTTCTCCTGGGGCTGTGCCGGGGAAGTGCGTGC
GTTCTTTCAGGAGAACACACACGTGAGACTGCGTATCTTCGCTGCCC GCATCTATGATTACGACCCC
25 CTATATAAGGAGGCACTGCAAATGCTGCGGGATGCTGGGGCCCAAGTCTCCATCATGACCTACGATG
AATTTAAGCACTGCTGGGACACCTTTGTGGACCACCAGGGATGTCCCTTCCAGCCCTGGGATGGACT
AGATGAGCACAGCCAAGCCCTGAGTGGGAGGCTGCGGGCCATTCTCCAGAATCAGGGAAACTGA

In certain methods of this disclosure a nucleic acid sample from a subject is provided.
The term "nucleic acid sample" includes a sample from a subject who is known to be
30 infected with a high risk type of HPV and/or a subject who is known to have an HPV-
associated cervical lesion. The test sample may originate from various sources in the
subject without limitation. If genomic DNA in the sample will be analyzed the tissue
sample can come from any tissue source that comprises genomic DNA of the subject,

including, without limitation, synovial fluid, blood, blood-derived product (such as buffy coat, serum, and plasma), lymph, urine, tear, saliva, hair bulb cells, cerebrospinal fluid, buccal swabs, feces, synovial fluid, synovial cells, sputum, or tissue samples. In addition, one of skill in the art would realize that some samples would be more readily analyzed following a fractionation or purification procedure, for example, isolation of DNA from whole blood. If mRNA in the sample will be analyzed the tissue sample will generally be taken from a tissue of the subject in which the A3A gene is known to be expressed.

In some embodiments the test sample is collected from the subject and then tested with little or no sample processing. In some embodiments the sample is processed, such as for example and without limitation processing to isolate all or a portion of the nucleic acid in the sample, such as genomic DNA in the sample, total RNA in the sample, or mRNA in the sample.

In general, methods for detecting an $\Delta A3B$ allele can be divided into two groups: (1) methods based on hybridization analysis of polynucleotides, and (2) other methods based on biochemical detection or sequencing of polynucleotides. The method used may be based on analysis of a starting nucleic acid that is genomic DNA or total RNA or mRNA obtained from the subject. In some embodiments cDNA is made from the mRNA as part of the method. Alternatively, the method used may be based on analysis of a starting nucleic acid that is genomic DNA obtained from the subject.

Any method known in the art or later developed may be used, in view of the teachings of this disclosure, to detect $\Delta A3B$ allele present in a starting sample that is genomic DNA obtained from a subject. Exemplary methods include, by way of example only, large-scale SNP genotyping, exonuclease-resistant nucleotide detection, solution-based methods, genetic bit analyses, primer guided nucleotide incorporation, allele specific hybridization, and other techniques. In some embodiments a Southern blot assay is used. Any method of detecting a marker may use a labeled oligonucleotide.

Numerous methods and devices are well known to the skilled artisan to identify the presence or absence of an $\Delta A3B$ allele. DNA (genomic and cDNA) for detection can be prepared from a biological sample by methods well known in the art, e.g., phenol/chloroform extraction, PURE GENE DNA® purification system from GentAS Systems (Qiagen, CA). Detection of a DNA sequence may include examining the nucleotide(s) located at either the sense or the anti-sense strand within that region. The

presence or absence of an $\Delta A3B$ allele in a patient may be detected from DNA (genomic or cDNA) obtained from PCR using sequence-specific probes, e.g., hydrolysis probes from Taqman, Beacons, Scorpions, or hybridization probes. For the detection of the $\Delta A3B$ allele, sequence specific probes may be designed such that they specifically hybridize to the genomic DNA for the alleles of interest or, in some cases, an RNA of interest. For example, primers and probes for the $\Delta A3B$ allele may be designed based on context sequences found in genomic DNA databases. These probes may be labeled for direct detection or contacted by a second, detectable molecule that specifically binds to the probe. The PCR products also can be detected by DNA-binding agents. Said PCR products can then be subsequently sequenced by any DNA sequencing method available in the art. Alternatively the presence or absence of an allele can be detected by sequencing using any sequencing methods such as, but not limited to, Sanger-based sequencing, pyrosequencing or next generation sequencing (Shendure J. and Ji, H., Nature Biotechnology (1998), Vol. 26, Nr 10, pages 1135-1145).

Various well-known techniques can be applied, including, e.g., hybridization-based methods, such as dynamic allele-specific hybridization (DASH) genotyping, detection through molecular beacons (Abravaya K., et al. (2003) Clin Chem Lab Med. 41 :468-474), Luminex xMAP technology, Illumina Golden Gate technology and commercially available high-density oligonucleotide arrays, BeadChip kits from Illumina, e.g, Human660W-Quad and Human 1.2M-Duo); enzyme-based methods, such as restriction fragment length polymorphism (RFLP), PCR-based methods (e.g., Tetra-primer ARMS-PCR), Invader assays (Olivier M. (2005) Mutat Res. 573(1-2): 103-10), various primer extension assays (incorporated into detection formats, e.g., MALDI-TOF Mass spectrometry, electrophoresis, blotting, and ELISA-like methods), Taqman assays, and oligonucleotide ligase assays; and other post-amplification methods, e.g., analysis of single strand conformation polymorphism (Costabile et al. (2006) Hum. Mutat. 27(12): 1163-73), temperature gradient gel electrophoresis (TGGE), denaturing high performance liquid chromatography, high-resolution melting analysis, DNA mismatch-binding protein assays (e.g., MutS protein from *Thermus aquaticus* binds different single nucleotide mismatches with different affinities and can be used in capillary electrophoresis to differentiate all six sets of mismatches), SNPLex® (proprietary SNP detecting system available from Applied Biosystems), capillary electrophoresis, mass spectrometry, and various sequencing methods, e.g.,

pyrosequencing and next generation sequencing, etc. Commercial kits for SNP genotyping include, e.g., Fluidigm Dynamic Array® IFCs (Fluidigm), TaqMan® SNP Genotyping Assay (Applied Biosystems), MassARRAY® iPLEX Gold (Sequenom), and Type-it Fast® SNP Probe PCR Kit (Quiagen).

5 According to a particular embodiment, methods are used to determine whether the A3 genotype of the subject comprises, in particular, at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$), but also whether the A3 genotype of the subject comprises either zero copy or one copy or two copies of the *A3A-A3B* deletion allele. To assess the presence of the deleted allele, various well-known techniques can be applied. Similarly to the disclosure
10 provided above with respect to HPV detection, PCR-based methods may be used in the implementation of a genotyping assay on a nucleic acid sample to determine whether the A3 genotype of the subject comprises at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$). Accordingly, examples of such PCR-based techniques include techniques such as, but not limited to, quantitative PCR (Q-PCR), reverse-transcriptase polymerase chain
15 reaction (RT-PCR), quantitative reverse-transcriptase PCR (QRT-PCR), or digital PCR (for example the Droplet Digital™ PCR technology sold by Bio-Rad, also known as ddPCR™). These techniques are well known and easily available technologies for those skilled in the art and do not need a precise description.

According to a particular embodiment, use is made of the Droplet Digital™ PCR
20 technology (ddPCR™) for assessing the presence of the *A3A-A3B* deletion allele ($\Delta A3B$), in particular, to determine the copy number of the *A3A-A3B* deletion allele ($\Delta A3B$) since this technique is suitable for measuring gene copy numbers. The APOBEC3 locus was built up by extensive gene duplication in the human genome: an example of a small region of DNA that is unique in the human genome, i.e., a region not showing any sequence
25 homology with any another part of the human genome, and which can therefore be used as a suitable target sequence for implementation of Droplet Digital™ PCR, is provided under SEQ ID NO: 87 (Figure 11):

CAGCATCTGGCACTGTGGTCGCTTCTTTACTCTGCCTCCCTCCTCTGGAGCTCCCGTGAATGTGAGC
GTCCTCCTCCTGGCTGGACGGCATCTCCTCTCTCTGGGCCTTCTGCCCTGGCCCCATCCTCCCTCGG
30 ACTGGGTGCCGAGTCTCAATGTTGGTCACTTATGACCCCAAATCGCCTTCTCCAGCTGGACACCCCC
AGTCTCCTGGGCTGGACTCTGCAGTCACCTTTAGTGCTCCCAGCTCCCCGTTCCTAAGCCCTGCTG
TTTCTACCTTTGGGAGGACACGCAGGCTGTCA

SEQ ID NO: 87 corresponds to a nucleic acid molecule upstream (5') to the APOBEC3B gene in the human genome. Figure 11 indicates the position of said sequence on chromosome 22 of the human genome.

According to a particular embodiment, the methods of the invention disclosed herein encompasses the use of target sequence SEQ ID NO: 87 for performing a genotyping assay on a nucleic acid sample to determine whether the A3 genotype of a subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$), using any one of the techniques disclosed herein, in particular, the Droplet DigitalTM PCR technology.

Other unique target sequences may be readily identified by the skilled artisan using any technique well-known in the art for determining such sequences.

According to another aspect, the methods of the invention, as disclosed herein, may encompass the use of a suitable target sequence as defined herein, between genes APOBEC3A and APOBEC3B in the human genome, especially SEQ ID NO: 87, for assessing the presence of a A3A-A3B deletion allele ($\Delta A3B$).

On the basis of such target sequences, one can choose suitable primers for implementation of a PCR method, especially the Droplet DigitalTM PCR method, aimed at determining the genotype of the subject comprises, in particular, at least one copy of a A3A-A3B deletion allele ($\Delta A3B$). In addition to the considerations provided in the present disclosure regarding the design of suitable primers, Figure 11 shows an annotated version of SEQ ID NO: 87, on which examples of primers are underlined, the arrowheads indicating the sense of the corresponding PCR primer. Primers can be designed according to standard methods that are commonplace. According to particular embodiments, pair(s) of primers suitable for use in the implementation of the present invention can be chosen amongst any suitable combination(s) of the following primers:

- GCACTGTGGTCGCTTCTTTA (SEQ ID NO : 88) 5' forward primer
- GTGAATGTGAGCGTCCTCCT (SEQ ID NO : 89) 5' forward primer
- GGCTGGACTCTGCAGTCAC (SEQ ID NO : 90) 3' reverse primer
- TGTTCCCTAAGCCCTGCTGTTT (SEQ ID NO : 91) 3' reverse primer

As long as a target sequence is known and identified, obtaining primers such as disclosed above is a matter of routine experimentation. Oligonucleotide (forward and reverse) primers have a length adapted to specifically priming the targeted DNA, in particular, have a length from 10, 15, 17, 18, 19 nucleotides to 25, 26, 27 or 30

nucleotides, especially have a length between 19 to 27 nucleotides. Examples of suitable oligonucleotide primers are polynucleotides comprising or consisting of the sequences disclosed herein.

5 The invention thus also relates to the use of oligonucleotide pairs (primers) either as unique pair or combined primer pairs having the sequences SEQ ID NO: 88, 89, 90 and 91 according to all suitable combinations known to the skilled person, for implementing the methods disclosed herein according to all aspects of the present disclosure.

10 Another object of the invention is to provide a kit suitable for carrying out methods of the invention as defined herein, especially for use in a genotyping assay according to the present disclosure, comprising:

- At least one pair of specific oligonucleotide primers specific for hybridization with a target sequence in human genomic DNA and, optionally, one or several of the following reagents,
- nucleotides (e.g. dATP, dCTP, dGTP, dUTP),
- 15 - a DNA polymerase, in particular a thermostable DNA polymerase,
- optionally, at least one dye for staining nucleic acids,
- optionally, a buffer solution,
- optionally, reagents necessary for the hybridization of the primers to their targets,
- 20 - optionally, a positive control,
- optionally, a negative control,
- optionally, a reference dye and,
- optionally a notice providing instructions for use and expected values for interpretation of results.

25 According to a particular embodiment, a kit of the invention further comprises means for detecting the presence or absence of a high risk type of human papilloma virus (HPV) according to the present disclosure, especially pair(s) of oligonucleotide primers specific for hybridization with HPV nucleic acid(s) sequence(s), as disclosed herein.

30 According to a particular embodiment, a kit of the invention comprises primers suitable for amplifying the target sequence SEQ ID NO: 87 or another target sequence specific for detecting of the presence of a *A3A-A3B* deletion allele ($\Delta A3B$).

According to a particular embodiment, a kit of the invention comprises pair(s) of primers selected amongst the sequences SEQ ID NO: 88, 89, 90 and 91 according to all suitable combinations known to the skilled person.

According to a particular embodiment, a kit of the invention comprises agents
5 necessary for the implementation of a Droplet DigitalTM PCR.

According to another aspect, the invention relates to the use of kit(s) as disclosed herein for performing a genotyping assay on a nucleic acid sample, and/or to determine whether the A3 genotype of a subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$).

10 According to another particular embodiment, the technique used for genotyping is the Droplet DigitalTM PCR technology.

According to particular embodiment, said use is made within a method for monitoring a cervical HPV infection or HPV induced cervical lesion in a subject and/or respectively detecting a subject's increased risk of progressing to and/or developing cervical
15 squamous cell carcinoma, said subject being a high risk HPV-positive subject or a subject with an HPV induced cervical lesion, as disclosed herein.

The invention also relates to the use of agents and kits, as described herein, in particular, when the kits suitable for implementing the invention are described, for the manufacture of a kit suitable for or aimed at performing the methods of the invention as
20 described herein. Instructions for use or guidance for implementing the method of the invention and/or instructions for use or guidance in order to obtain a suitable kit may advantageously be provided. The agents discussed herein can be readily determined by the skilled in the art.

In some embodiments, the presence or absence of an $\Delta A3B$ allele in a patient is
25 detected using a hybridization assay. In a hybridization assay, the presence or absence of the genetic marker is determined based on the ability of the nucleic acid from the sample to hybridize to a complementary nucleic acid molecule, e.g., an oligonucleotide probe. A variety of hybridization assays are available. In some, hybridization of a probe to the sequence of interest is detected directly by visualizing a bound probe, e.g., a Northern or
30 Southern assay. In these assays, DNA (Southern) or RNA (Northern) is isolated. The DNA or RNA is then cleaved with a series of restriction enzymes that cleave infrequently in the genome and not near any of the markers being assayed. The DNA or RNA is then

separated, e.g., on an agarose gel, and transferred to a membrane. A labeled probe or probes, e.g., by incorporating a radionucleotide or binding agent (e.g., SYBR® Green), is allowed to contact the membrane under low-, medium- or high-stringency conditions. Unbound probe is removed and the presence of binding is detected by visualizing the
5 labeled probe. In some embodiments, arrays, e.g., the MassARRAY system (Sequenom, San Diego, California, USA) may be used to genotype a subject.

Sequence-Specific Oligonucleotide (SSO) typing uses PCR target amplification, hybridization of PCR products to a panel of immobilized sequence-specific oligonucleotides on beads, detection of probe-bound amplified product by color formation
10 followed by data analysis. Those skilled in the art would understand that the described Sequence-Specific Oligonucleotide (SSO) hybridization may be performed using various commercially available kits, such as those provided by One Lambda, Inc. (Canoga Park, CA) coupled with Luminex® technology (Luminex, Corporation, TX). LABType® SSO is a reverse SSO (rSSO) DNA typing solution that uses sequence-specific oligonucleotide
15 (SSO) probes and color-coded microspheres to identify alleles. The target DNA is amplified by polymerase chain reactions (PCR) and then hybridized with the bead probe array. The assay takes place in a single well of a 96-well PCR plate; thus, 96 samples can be processed at one time.

Sequence Specific Primers (SSP) typing is a PCR based technique which uses
20 sequence specific primers for DNA based allele typing. The SSP method is based on the principle that only primers with completely matched sequences to the target sequences result in amplified products under controlled PCR conditions. Allele sequence-specific primer pairs are designed to selectively amplify target sequences which are specific to a single allele or group of alleles. PCR products can be visualized on an agarose gel. Control
25 primer pairs that match non-allelic sequences present in all samples act as an internal PCR control to verify the efficiency of the PCR amplification. Those skilled in the art would understand that low, medium and high resolution genotyping with the described sequence-specific primer typing may be performed using various commercially available kits, such as the Olerup SSP™ kits (Olerup, PA) or (Invitrogen) or Allset and™Gold DQA1 Low
30 resolution SSP (Invitrogen).

Sequence Based Typing (SBT) is based on PCR target amplification, followed by sequencing of the PCR products and data analysis. In some cases, RNA, e.g., mature

mRNA or pre-mRNA, can also be used to determine the presence or absence of $\Delta A3B$ alleles. Analysis of the sequence of mRNA transcribed from a given gene can be performed using any known method in the art including, but not limited, to Northern blot analysis, nuclease protection assays (NPA), *in situ* hybridization, reverse transcription-polymerase chain reaction (RT-PCR), RT-PCR ELISA, TaqMan-based quantitative RT-PCR (probe-based quantitative RT-PCR) and SYBR green-based quantitative RT-PCR. In one example, detection of mRNA levels involves contacting the isolated mRNA with an oligonucleotide that can hybridize to mRNA encoded by a coding sequence present in the $\Delta A3B$ allele. The nucleic acid probe can typically be, for example, a full-length cDNA, or a portion thereof, such as an oligonucleotide of at least 7, 15, 30, 50, or 100 nucleotides in length and sufficient to specifically hybridize under stringent conditions to the mRNA. Hybridization of an mRNA with the probe indicates that the marker in question is being expressed. In one format, the RNA is immobilized on a solid surface and contacted with a probe, for example by running the isolated RNA on an agarose gel and transferring the mRNA from the gel to a membrane, such as nitrocellulose. Amplification primers are defined as being a pair of nucleic acid molecules that can anneal to 5' or 3' regions of a gene (plus and minus strands, respectively, or vice-versa) and contain a short region in between. In general, amplification primers are from about 10 to about 30 nucleotides in length and flank a region from about 50 to about 200 nucleotides in length. Under appropriate conditions and with appropriate reagents, such primers permit the amplification of a nucleic acid molecule comprising the nucleotide sequence flanked by the primers. PCR products can be detected by any suitable method including, but not limited to, gel electrophoresis and staining with a DNA-specific stain or hybridization to a labeled probe.

In some embodiments, the presence of a $\Delta A3B$ allele in a subject is determined by measuring RNA levels using, e.g., a PCR-based assay or reverse-transcriptase PCR (RT-PCR). In some embodiments, quantitative RT-PCR with standardized mixtures of competitive templates can be utilized.

A $\Delta A3B$ allele can also be identified by detecting an equivalent genetic marker thereof, which can be, e.g., a SNP allele on the same haplotype as the $\Delta A3B$ allele. Two particular alleles at different loci on the same chromosome are said to be in linkage disequilibrium (LD) if the presence of one of the alleles at one locus tends to predict the presence of the other allele at the other locus. The SNP may be an allele of a polymorphism that is

currently known. Other SNPs may be readily identified by the skilled artisan using any technique well-known in the art for discovering polymorphisms.

EXAMPLES

5

Methods

Plasmids

A3B molecular clones were PCR cloned in pcDNA3.1D/V5-His-TOPO vector (Life Technologies). As subcloning of wild type A3B cDNA into pcDNA3.1D/V5-His-TOPO vector (Life Technologies) proved difficult due to *E. coli* toxicity, a synthetic A3B cDNA A3Bi7 containing the 281 bp intron 7 was synthesized (GeneCust) and subsequently cloned in pcDNA3.1D/V5-His-TOPO vector (Table 1 – consecutive SEQ ID NOs : 30 to 53). Derivatives were then generated by site directed mutagenesis (GeneArt® Site-Directed Mutagenesis System, Life Technologies) (Table 2 – consecutive SEQ ID NOs : 54 to 61). HA tagged A3As fused to UTRs were constructed by assembly PCR. A3A coding sequence retaining the 281 bp intron 4 (GeneCust) was amplified using primers designed to add a N-terminal HA tag and A3A/A3B 3'UTRs sequences were amplified from 293T genomic DNA extracted using MasterPure™ complete DNA and RNA purification kit (Epicentre). Overlapping amplicons were purified and mixed for PCR elongation. Full-length constructs were then amplified using external primers (Table 1) and cloned in pcDNA3.1D/V5-His-TOPO vector. Luciferase plasmids were constructed the same way using Firefly Luciferase sequence amplified from pGL4.50[luc2/CMV/Hygro] Vector (Promega). All constructs were verified by sequencing.

25

Cells

Quail QT6 embryo fibroblast cells were maintained in HAM's F40 medium (Eurobio), supplemented with 1% chicken serum, 10% FCS, 5% tryptose phosphate, 2 mM L-glutamine, 50 U/ml penicillin and 50 mg/ml streptomycin. Human HeLa cells, SKBR3 cells and 293T-UGI cells stably expressing *Bacillus subtilis* phage uracil-DNA glycosylase inhibitor (UGI) were maintained in DMEM glutamax medium (Life Technologies) supplemented with 10% FCS, 50 U/ml penicillin and 50 mg/ml streptomycin.

30

Transfections

One million QT6 cells were co-transfected with 0.5 µg of pCayw HBV coding plasmid and 1.5 µg A3 expression plasmids using JetPrime (Polyplus) following manufacturer's recommendations and harvested 48 hours post-transfection. QT6 nuDNA editing was assessed after transfection of 1 µg A3 expression plasmids along with 1 µg of the UGI expression plasmid using the same transfection procedure. For single plasmid transfections, 8×10^5 of HeLa, 293T-UGI, cells were transfected using 2 µg APOBEC3 expression plasmids using JetPrime (Polyplus) following manufacturer's recommendations and harvested 48 hours post-transfection. For the 5Me-dC deamination assay, a 679 bp fragment of HIV-1 pNL4.3 *env* gene was amplified using total substitution of dCTP by 5Me-dCTP (Trilink) using the primer pair 5'-TTGATGATCTGTAGTGCTACAGCA (SEQ ID NO : 62) and 5'-GCCTAATTCCATGTGTACATTGTA (SEQ ID NO : 63). The 5Me-dC containing DNA was heat denatured, chilled on ice and 200 ng of synthesized DNA was transfected using JetPrime 24 hours following initial transfection of APOBEC3 coding plasmids in QT6 cells as described above. For immunofluorescence labeling, 5×10^4 HeLa cells grown on chamber slides (LabTek) were transfected with 1 µg APOBEC3 expression plasmids using Fugene HD (Roche) following manufacturer's recommendations. For Luciferase activity assay, 8×10^5 of 293T cells were co-transfected with 1.8 µg of Firefly Luciferase plasmids along with 0.2 µg of *Renilla* Luciferase control plasmid (Promega).

Western blots

After blocking, membranes were probed with 1:5000 diluted mouse monoclonal antibody specific for the V5 epitope (Life Technologies), or 1:2000 diluted rabbit monoclonal antibody specific HA tag (Sigma) in PSB-0.01% Tween 5% dry milk applied overnight. After PBS-Tween washings and incubation with an anti-mouse or anti-rabbit IgG horseradish peroxidase-coupled secondary antibody (Amersham), the membrane was revealed by enhanced chemiluminescence (Pierce). β -Actin was used as a loading control using 1:50000 diluted mouse monoclonal antibody specific for β -Actin (Sigma). Signal was quantified using ImageJ Software. Uncropped scans of western blots are presented in Figures 10A and 10B.

Immunofluorescence

After PBS washings, transfected HeLa cells grown on chamber slides were fixed with 4% PFA for 15 min. After PBS washing cells were incubated in 50/50 acetone/methanol for 20 minutes. Mouse monoclonal anti-V5 antibody (Life Technologies) was then incubated at 1:200 for 1 h at room temperature, followed by incubation with 1:500 diluted mouse specific Alexa-488 conjugated goat antibody (Life Technologies) for 1 h at room temperature in the dark. After washing, slides were mounted with Vectashield imaging medium containing DAPI (Vector Laboratories). Imaging was performed using Leica SP5 confocal microscope.

10

In vitro deamination assay

At 72 hours post transfection, APOBEC3 transfected 293T cells were extensively wash with PBS and mechanically harvested. Total proteins were extracted using specific lysis buffer (25 mM HEPES (pH 7.4), 10% glycerol, 150 mM NaCl, 0.5% Triton X-100, 1 mM EDTA, 1 mM MgCl₂, 1 mM ZnCl₂) supplemented with protease inhibitors, and submitted to sonication. Deaminase activity was assessed by incubating whole cell lysates with 1 pmole DNA oligonucleotide 5'-(6-FAM)-AAATTCTAATAGATAATGTGA-(TAMRA) (SEQ ID NO : 64) in presence of 0.4 unit UDG (NEB) in a 20 mM Tris-HCl, 1 mM DTT, 1 mM EDTA reaction buffer. After 2 hours incubation at 37°C, generated abasic sites were cleaved by heating 2 minutes at 95°C, and endpoint fluorescence were measured using realplex (2) Mastercycler (Biorad) with FAM setting and background fluorescence obtained with mock-transfected cells as negative control. Results are normalized to the quantity of protein using Pierce™ BCA Protein Assay Kit (Thermo Scientific).

20

DNA extraction and 3DPCR amplification

Total DNA from transfected cells was extracted using the MasterPure™ complete DNA and RNA purification kit (Epicentre) and resuspended in 30 µL sterile water. All amplifications were performed using first-round standard PCR followed by nested 3DPCR (Table 3 – consecutive SEQ ID NOs : 65 to 80). PCR was performed with 1 u Taq DNA polymerase (Eurobio) per reaction. After purification, PCR products were cloned into TOPO 2.1 vector (Life Technologies) and sequencing was outsourced to GATC. Expected values are derived from the base composition of the target sequence assuming no dinucleotide bias

30

(% of NpC= numbers of NpC / numbers of Cs) x100). Statistical significance of observed percentages compared to expected values was assessed using χ^2 test.

A3A mRNA quantification

5 Transfected 293T cells mRNAs were extracted using RNA EXTRACT RNeasyR Plus Mini Kit (Qiagen). Corresponding cDNAs were synthesized using QuantiTect Reverse Transcription Kit (Qiagen). Quantification was performed by TaqMan using Takyon™ Rox Probe MasterMix dTTP blue (Eurogentec). A3A mRNA from transfected plasmids was quantified with primers overlapping retained intron4, qA3Afor: 5'-
10 CTGAGGCCCATCCTTCAGTTTCCCT (SEQ ID NO : 81), qA3Arev: 5'-GTGGACCACCAGGGATGT (SEQ ID NO : 82) and molecular probe #11 (Roche). Results were normalized to RPL13 reference gene qRPL13for: 5'-CTGGACCGTCTCAAGGTGTT (SEQ ID NO : 83), qRPL13rev: 5'-GCCCCAGATAGGCAAACCTT (SEQ ID NO : 84), probe #74. Endogenous level of A3A transcripts in SKBR3 cells was measured a previously
15 described (4).

Luciferase activity

3.10⁵ of transfected 293T cells were split in 96well plates 24 hours after luciferase plasmids transfection. Luciferase activity was measured 24 hours later using Dual-Glo
20 luciferase assay (Promega) with 30 minutes incubation times. Presented results represent luciferase activity from 3 independent transfections measured in triplicates.

FACS analysis of apoptosis

Transfected HeLa cells were resuspended in binding buffer (BD Pharmingen) and stained
25 with FITC-labeled Annexin V antibody (1 µg/ml) (BD Pharmingen). Cells were counterstained 5 µg/ml PI (BD Pharmingen) to distinguish between early apoptotic and late apoptotic or necrotic events. Treatment by 100 µM etoposide in DMSO was used as positive control.

30 FACS analysis of double strand breaks

At 48 hours after transfection, cells were washed with PBS, fixed in 2-4% ice-cold paraformaldehyde (Electron Microscopy Sciences) for 10 minutes and permeabilized in

90% ice-cold methanol (Sigma) for 30 minutes. After washing with PBS, cells were incubated with 1:200 diluted mouse anti-V5 antibody (Life Technologies) in PBS-BSA 0.5% for 1 hour. After PBS washings incubation with 1:500 diluted Alexa Fluor 633 F fragment of goat anti-mouse IgG (H+L) (Life Technologies) was performed for 45 minutes. HA tag
5 detection was performed in the same conditions using 1:100 anti-HA antibody directly linked to Alexa Fluor 633 (Cell Signaling). DNA double strand breaks were analyzed by staining for 1 hour with 1:50 diluted Alexa Fluor 488-conjugated rabbit monoclonal anti- γ H2AX (20E3) antibody (Cell Signaling). All incubation steps were performed on ice. Stained samples were acquired on a MACSQuant Analyser (Miltenyi Biotec) and data
10 were analyzed with FlowJo software (Tree Star Inc. version 8.7.1). DSB probing in SKBR3 cells was performed 24 hours after 100 μ M phorbolmyristic acetate (PMA) (Sigma) induction with 1:50 diluted Alexa Fluor 647-conjugated rabbit monoclonal anti- γ H2AX (20E3) antibody (Cell Signaling).

15 **Example 1: Functional attenuation of APOBEC3B in *E. coli***

The starting point was the observation that human A3B cDNA expression plasmids could occasionally accumulate somatic mutations. Our original A3B construct derived from an Image clone encoding a known but minor T146K polymorphism (A3Bwh)⁴ had recently acquired two somatic mutations, notably F308L and W359L (Fig. 1a). Another available
20 plasmid A3Blan (29), encoded 8 amino acid substitutions K62E, L80P, F107L, T146K, M193V, D205G, T337A, R372K, while A3Btok (21, 30) encoded the T146K polymorphism (Fig. 1a). As none of the other mutations are known polymorphisms, they were probably selected to attenuate toxic A3B expression via leaky T7 promoters in *E. coli*, as has been reported for some A3 constructs (31, 32). V5 tagged versions of germ line A3B and A3Btok
25 proved to be genetically unstable while A3Blan posed no problem. To overcome this issue, a germ line A3B cDNA retaining the entire intron 7 was synthesized and cloned into pcDNA3.1 (plasmid pA3Bi7). This non-toxic and stable construct was subsequently used to generate A3B derivatives, notably the T146K polymorphism, an E255Q null mutant as well as individual mutations F308L and W359L found in A3Bwh. Western blot analysis
30 confirmed that A3B proteins were produced in comparable amounts (Fig. 1b), all of them exhibiting A3B classical nuclear localization (Figure 4a).

A3B plasmids were transfected into 293T cells and cellular lysates were used in a FRET based *in vitro* deamination assay where C to U conversion in a TAM-FAM labelled DNA oligonucleotide, allows fluorescence detection following cleavage by uracil-DNA glycosylase (UNG) activity (31, 33). All A3B constructs showed cytidine deamination activity while A3AC106S and A3BE255Q null mutants did not (Fig. 1c). If A3Bi7 and A3Btok displayed cytidine deaminase activity comparable to A3A, A3Bwh and A3Blan showed diminished activity, while F308L and W359L individually attenuated activity. Turning to cell based systems, A3 activity was assessed by HBV genome editing which has proven to be the most sensitive *in vivo* assay for A3 activity (19, 34). Following co-transfection of A3 plasmids along with the pCayw HBV infectious molecular clone, total cellular DNA was extracted and HBV editing was analyzed by 3DPCR, a technique that allows recovery of AT rich DNA (35). For all A3B constructs, HBV DNA was recovered at temperatures below the limiting denaturation temperature $T_d=89.0^{\circ}\text{C}$ for unedited DNA obtained with inactive A3BE255Q, A3AC106S mutants or mock transfected cells (Figure 4b). While A3Bi7 and A3Btok transfections allowed recovery of edited HBV DNA down to 81.3°C , again A3Blan and A3Bwh were less active, DNA being recovered at higher T_d s. Once again F308L and W359L proved to be attenuating mutations (Figure 4b).

Example 2: APOBEC3B can deaminate nuDNA and 5-methylcytidine.

The ability of A3Bi7 to edit nuDNA and 5-methylcytidine (5Me-dC), hallmark activities of A3A, was investigated. A3B plasmids were transfected into 293T-UGI cells, which stably express the UNG inhibitor UGI (33), allowing detection of nuDNA editing as in the case of A3A⁴. Edited TP53 nuDNA was recovered using 3DPCR, from A3Bi7 and A3Btok transfections (Fig. 2a). Sequences were peppered with C to T mutations preferentially on the minus strand (Figs. 2b, c), reflecting an A3 deamination process occurring on the non-transcribed template strand (36). Dinucleotide analysis of edited bases showed a strong 5'TpC preference for cytidine deamination (Fig. 2d). By contrast A3Bwh and A3Blan failed to edit nuDNA (Fig. 2a). Once again F308L and W359L both abrogated nuDNA editing (Fig. 2a). To confirm this in another cell type, we co-transfected QT6 cells, which are devoid of any APOBEC editing background, with A3B plasmids along with a UGI coding plasmid. 3DPCR amplification of a segment of quail CMYC recovered hyperedited sequences with C->T transitions in the 5'TpC dinucleotide context for A3Bi7 and A3Btok

constructs to 89.6°C and 88.7°C respectively, below the restrictive temperature $T_d=91.7^\circ\text{C}$ obtained for controls (Figure 5a-c).

Although 3DPCR is not a quantitative technique, A3A consistently outperformed A3B regarding editing activity (Fig. 2a, Fig. 4a, Fig. 5a), suggesting that *in vivo* editing by A3A is more efficient than A3B. Given that DSB formation is a macroscopic and quantitative method to assess the efficiency of A3A or A3B editing on nuclear DNA, γH2AX histone phosphorylation gated on A3-V5 tagged positive cells was probed 48 hours post transfection. Where A3A expression resulted in numerous γH2AX foci in HeLa or QT6 cells, transfection with A3Bi7 or A3Btok consistently failed to make detectable DSBs (Fig. 2e, Figs 6a and 6b). As DSBs are a direct consequence of A3A mediated catastrophic mutations on cellular DNA (Fig. 7), it is possible that A3B failed to induce the number of mutations required for detectable DSB formation.

Nonetheless, A3B activity results in genotoxicity (Fig. 2) and apoptosis just as A3A (Fig. 2f) (3). As orthologous mammalian A3A enzymes are capable of deaminating 5Me-dC residues in single stranded DNA (31), 5Me-dC deamination being involved in 5Me-CpG mutation hot-spots associated with cancer related genes (37), was analyzed as previously described⁶. Both A3Bi7, A3Btok could edit 5Me-dC-substituted HIV-1 *env* gene DNA like the A3A positive control compared to A3C used as negative control (Fig. 2g). Sequencing of 3DPCR products revealed hypermutated sequences with the same 5'TpC dinucleotide editing preference (Fig. 2h, Fig. 5d, e). The above findings indicate that A3B is a *bona fide* mutator of chromosomal DNA although A3A would appear to be the more powerful of the two *in vivo* (Fig. 2a, e).

Example 3: Increased DNA damage from the chimeric A3A transcript.

A natural 29.5kb deletion allele occurs between an identical 370 bp segment spanning intron 4/exon5 of A3A and intron 7/exon8 of A3B. It leaves a chimeric A3-A3B transcript that differs from normal A3A only by the 3'UTR – the four amino acids of A3A exon 5 and A3B exon 8 being identical (Fig. 3a and Fig. 8) (28). The major difference is the insertion of an Alu sequence in the A3A 3'UTR (Fig. 3a and Fig. 8). As previous constructs were C-terminal V5 tagged, a new set of A3A intron 4 constructs were generated, harboring a N-terminal HA-tag and the 3' untranslated regions from A3A or A3B (UTR_{A3A} and UTR_{A3B}) (Fig. 3b). As can be seen from Fig. 3c transfection of control A3A- ΔUTR , or the chimeric

A3A-UTR_{A3B} construct produced much higher levels of A3A compared to the natural A3A transcript (A3A-UTR_{A3A}). RTqPCR across A3A intron 4 showed a significant two fold reduction in mRNA levels (Fig. 3d). To quantify accurately the difference, the UTR sequences were cloned at the 3' end of the firefly luciferase (Fig. 3b) reporter gene and relative luminescence was measured using the *Renilla* luciferase gene as transfection control. The activity from Luc-ΔUTR and Luc-UTR_{A3B} constructs proved to be ~20 fold greater than that for and Luc-UTR_{A3A} (Fig. 3e) demonstrating that A3A expression levels of the natural and chimeric transcripts are primarily controlled by the UTRs.

To assess A3A cytidine deamination, transfection of 293T-UGI cells showed that the A3A-UTR_{A3B} transcript was capable of hyperediting nuDNA, just as A3A-ΔUTR, while the natural A3A-UTR_{A3A} construct proved to be much less efficient (Fig. 3f), a finding paralleled by the frequency of DSBs in HeLa transfected cells (Fig. 3g, Fig. 9a). Of note, N-terminal HA tagged A3A proteins proved somewhat less active than C-terminally V5 tagged constructs (Fig. 3f, g compared to Fig. 2a, e). The breast cancer cell line SKBR3 harbors the ΔA3B^{-/-} deletion (23) and sequencing of chimeric A3A-UTR_{A3B} mRNA transcripts confirmed the deletion between A3A and A3B. As A3A was originally identified following induction by phorbol myristic acetate (PMA) in normal human keratinocytes (38), SKBR3 cells were treated with 100 μM PMA and DSB formation measured 24 hours after induction (Fig. 3h, Fig. 9). A3A upregulation was observed in treated SKBR3 cells, accompanied by DSB formation (Fig. 3h), confirming that somatic DNA damage can occur in the absence of A3B.

Table 1 (consecutive SEQ ID NOs : 30 to 53)

Plasmid	Matrix	Primers
A3B1an	pCDNA3.1-A3B-HA	For : 5'-CACCATGAATCCACAGATCAGAAATCCGATGGAGC
		Rev : 5'-GTTTCCCTGATTCCTGGAGAAAT
A3B17	Puc57-A3B1	For : 5'-CACCATGAATCCACAGATCAGAAATCCGATGGAGC
		Rev : 5'-GTTTCCCTGATTCCTGGAGAAATCCGATGGAGC
A3A ΔUTR	A3A14	For : 5'-CACCATGGCTTATCCTTACAGACCTGCTGACTACGCCATGGAAGCCAGCCAGCATCCGGGCCAGACACTTG
		Rev : 5'-ACTGAGGCCCATCCTTCAGTTCCCTGATTCCTGGAAGAGCAG
A3AC106S	A3AC106S	For : 5'-CACCATGGCTTATCCTTACAGACCTGCTGACTACGCCATGGAAGCCAGCCAGCATCCGGGCCAGACACTTG
		Rev : 5'-ACTGAGGCCCATCCTTCAGTTCCCTGATTCCTGGAAGAGCAG
A3A UTRA	A3A14	For : 5'-CACCATGGCTTATCCTTACAGACCTGCTGACTACGCCATGGAAGCCAGCCAGCATCCGGGCCAGACACTTG
		Rev : 5'-ACTGAGGCCCATCCTTCAGTTCCCTGATTCCTGGAAGAGCAG
293T gDNA	293T gDNA	For : 5'-AATCAGGGAACCTGAGGATGGCTCAGTCTCTAAGGAAG
		Rev : 5'-CCATCAGCGCCCGGGCTCTCTAGTGAGGGTGGCGGGGC
A3A UTRB	A3A14	For : 5'-CACCATGGCTTATCCTTACAGACCTGCTGACTACGCCATGGAAGCCAGCCAGCATCCGGGCCAGACACTTG
		Rev : 5'-ACTGAGGCCCATCCTTCAGTTCCCTGATTCCTGGAAGAGCAG
293T gDNA	293T gDNA	For : 5'-AATCAGGGAACCTGAGGATGGCTCAGTCTCTAAGGAAG
		Rev : 5'-CCATCAGCGCCCGGGCTCTCTAGTGAGGGTGGCGGGGC
Luc ΔUTR	pGL4.50 [luc2/CMV/Hygro]	For : 5'-CACCATGGAGATGCCAAAACATTAAAG
		Rev : 5'-TTACACGGCGATCTTGGCGCCCTTCTTGGCCCTT
Luc UTRA	pGL4.50 [luc2/CMV/Hygro]	For : 5'-CACCATGGAGATGCCAAAACATTAAAG
		Rev : 5'-CCTTCTTACAGACTGAGGCCCATCTTTACACGGCGATCTTGGCCGCCCTTCTTGGCCCTT
A3A UTRA	A3A UTRA	For : 5'-AGGATGGGCTCAGTCTCTAAGGAAG
		Rev : 5'-GTTTCCCTGATTCCTGGAAGAGCAGGGGGTGG
pGL4.50	pGL4.50	For : 5'-CACCATGGAGATGCCAAAACATTAAAG
		Rev : 5'-CCTTCTTACAGACTGAGGCCCATCTTTACACGGCGATCTTGGCCGCCCTTCTTGGCCCTT
Luc UTRB	[luc2/CMV/Hygro]	For : 5'-AGGATGGGCTCAGTCTCTAAGGAAG
		Rev : 5'-GTTTCCCTGATTCCTGGAAGAGCAGGGGGTGG
A3A UTRB	A3A UTRB	For : 5'-AGGATGGGCTCAGTCTCTAAGGAAG
		Rev : 5'-CCTTCTTACAGACTGAGGCCCATCTTTACACGGCGATCTTGGCCGCCCTTCTTGGCCCTT

Table 2 (consecutive SEQ ID NOs : 54 to 61)

Plasmid	Matrix	Primers
A3Btok (T146K)	A3Bi7	For : 5'-GCAGGAGCCCGCGTGAAGATCATGGACTATGAA Rev : 5'-TTCATAGTCCATGATCTTCACGGGGCTCCTGC
A3BiE255Q	A3Bi7	For : 5'-TACGGCCGCCATGCGCGCAGCTGGCGTTCTTTGGAC Rev : 5'-GTCCAAGAGCGCGCAGCTGCGGCATGGCGGCGGA
A3BF308L	A3Bi7	For : 5'-GTGAGACATGGGCATCTTAACTGCCCCGCATCTAT Rev : 5'-ATAGATGCGGGCAGCTAAGATGCGGCAGTCTCAC
A3BW359L	A3Bi7	For : 5'-TGTCCCTTCCAGCCCTTTGGATGGACTAGAGGAG Rev : 5'-CTCCTCTAGTCCATCCAAAGGGCTGGAAGGGACA

Table 3 (consecutive SEQ ID NOs : 65 to 80)

Primer	Application	Sequence	Cycling conditions
HBV out fwd	HBV PCR	5'-AAGAGTYTYTTAATGTAAGACYTT	95°C 5 min 35x (95°C 1 min, 50°C 30 sec, 72°C 5 min), 72°C 10 min
HBV out rev	HBV PCR	5'-CGCAATATACATCGTATCCAT	
HBV in fwd	HBV 3DPCR	5'-AAGTGCACACGGTYGGCAGAT	80-90°C 5 min 35x (80-90°C 1 min, 50°C 30 sec, 72°C 5 min), 72°C 10 min
HBV in rev	HBV 3DPCR	5'-ATGGCTGCTARGCTGTGCTGCCAA	
TP53 out fwd	TP53 PCR	5' GAGCTGGACCTTAGGCTCCAGAAAGGACAA	95°C 5 min 42x (95°C 30 sec, 60°C 30 sec, 72°C 2 min), 72°C 10 min
TP53 out rev	TP53 PCR	5'-GCTGGTGTGTGGGCGAGTCTAGGAA	
TP53 in fwd	TP53 3DPCR	5'-TTCTCTTTTCCATACCTGAGTAGTGTAA	84-92°C 5 min 40x (84-92°C 1 min, 58°C 30 sec, 72°C 2 min), 72°C 10 min
TP53 in rev	TP53 3DPCR	5'-AAAGGTGATAAAAGTGAATCTGAGGCATAA	
HIV-1 out fwd	HIV-1 PCR	5'-TGTACCCACAGACCCCAACCCACAA	95°C 5 min 30x (95°C 45 sec, 54°C 45 sec, 72°C 2 min), 72°C 10 min
HIV-1 out rev	HIV-1 PCR	5'-TTCCATTGAACGCTCTTATTATTACA	
HIV-1 in fwd	HIV-1 3DPCR	5'-ATCAAAGCCTAAGCCCATGTGTAA	78-90°C 5 min 45x (78-90°C 1 min, 56°C 30 sec, 72°C 5 min), 72°C 10 min
HIV-1 in rev	HIV-1 3DPCR	5'-CAATAATGTATGGGAATGGCTCAA	
QT6 cmyc out fwd	cmyc PCR	5'-GATGCCGCTCAGCCCGAGCCTCCCCAGCAA	95°C 5 min 40x (95°C 1 min, 58°C 45 sec, 72°C 2 min), 72°C 10 min
QT6 cmyc out rev	cmyc PCR	5'-GCAGTCTTGGATGATGATGATTGACGAA	
QT6 cmyc in fwd	cmyc 3DPCR	5'-TTCGAGGAGGAGGAGGAGAACTTCTA	88-95°C 5 min 42x (88-95°C 1 min, 60°C 1 min, 72°C 2 min), 72°C 10 min
QT6 cmyc in rev	cmyc 3DPCR	5'-GATTCTGCTCCGGGTCGAGATGAA	

REFERENCES

1. Jarmuz, A. *et al.* An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics* **79**, 285-296 (2002).
2. Landry, S., Narvaiza, I., Linfesty, D.C. & Weitzman, M.D. APOBEC3A can activate the
5 DNA damage response and cause cell-cycle arrest. *EMBO reports* **12**, 444-450 (2011).
3. Mussil, B., Suspène, R., Aynaud, M.M., Gauthier, A., Vartanian, J.P. & Wain-Hobson, S. Human APOBEC3A Isoforms Translocate to the Nucleus and Induce DNA Double Strand Breaks Leading to Cell Stress and Death. *PLoS ONE* **8**, e73641 (2013).
4. Suspène, R. *et al.* Somatic hypermutation of human mitochondrial and nuclear DNA by
10 APOBEC3 cytidine deaminases, a pathway for DNA catabolism. *Proc. Natl. Acad. Sci. USA* **108**, 4858-4863 (2011).
5. Carpenter, M.A. *et al.* Methylcytosine and Normal Cytosine Deamination by the Foreign DNA Restriction Enzyme APOBEC3A. *J. Biol. Chem.* **287**, 34801-34808 (2012).
6. Suspène, R., Aynaud, M.M., Vartanian, J.P. & Wain-Hobson, S. Efficient deamination of
15 5-methylcytidine and 5-substituted cytidine residues in DNA by human APOBEC3A cytidine deaminase. *PLoS ONE* **8**, e63461 (2013).
7. Wijesinghe, P. & Bhagwat, A.S. Efficient deamination of 5-methylcytosines in DNA by human APOBEC3A, but not by AID or APOBEC3G. *Nucleic Acids Res.* **40**, 9206-9217 (2012).
8. Greenman, C. *et al.* Patterns of somatic mutation in human cancer genomes. *Nature*
20 **446**, 153-158 (2007).
9. Nik-Zainal, S. *et al.* Mutational Processes Molding the Genomes of 21 Breast Cancers. *Cell* **149**, 979-993 (2012).
10. Stephens, P.J. *et al.* Complex landscapes of somatic rearrangement in human
25 breast cancer genomes. *Nature* **462**, 1005-1010 (2009).
11. Stephens, P.J. *et al.* The landscape of cancer genes and mutational processes in breast cancer. *Nature* **486**, 400-404 (2012).
12. Alexandrov, L.B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415-421 (2013).
13. Roberts, S.A. *et al.* An APOBEC cytidine deaminase mutagenesis pattern is
30 widespread in human cancers. *Nat. Genet.* **45**, 970-976 (2013).

14. Roberts, S.A. & Gordenin, D.A. Clustered and genome-wide transient mutagenesis in human cancers: Hypermutation without permanent mutators or loss of fitness. *Bioessays* doi: 10.1002/bies.201300140, (2014).
15. Doehle, B.P., Schafer, A., Wiegand, H.L., Bogerd, H.P. & Cullen, B.R. Differential sensitivity of murine leukemia virus to APOBEC3-mediated inhibition is governed by virion exclusion. *J. Virol.* **79**, 8201-8207 (2005).
16. Mahieux, R. *et al.* Extensive editing of a small fraction of human T-cell leukemia virus type 1 genomes by four APOBEC3 cytidine deaminases. *J. Gen. Virol.* **86**, 2489-2494 (2005).
- 10 17. Ooms, M., Krikoni, A., Kress, A.K., Simon, V. & Munk, C. APOBEC3A, APOBEC3B, and APOBEC3H haplotype 2 restrict human T-lymphotropic virus type 1. *J. Virol.* **86**, 6097-6108 (2012).
18. Bonvin, M. & Greeve, J. Effects of point mutations in the cytidine deaminase domains of APOBEC3B on replication and hypermutation of hepatitis B virus in vitro. *J. Gen. Virol.* **88**, 3270-3274 (2007).
- 15 19. Suspène, R., Guétard, D., Henry, M., Sommer, P., Wain-Hobson, S. & Vartanian, J.P. Extensive editing of both hepatitis B virus DNA strands by APOBEC3 cytidine deaminases in vitro and in vivo. *Proc. Natl. Acad. Sci. USA* **102**, 8321-8326 (2005).
20. Burns, M.B. *et al.* APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* **494**, 366-370 (2013).
- 20 21. Shinohara, M. *et al.* APOBEC3B can impair genomic stability by inducing base substitutions in genomic DNA in human cells. *Sci. Rep.* **2**, 806 (2012).
22. Suspène, R., Caval, V., Henry, M., Bouzidi, M.S., Wain-Hobson, S. & Vartanian, J.P. Erroneous identification of APOBEC3-edited chromosomal DNA in cancer genomics. *Br. J. Cancer* **110**, 2615-2622 (2014).
- 25 23. Komatsu, A., Nagasaki, K., Fujimori, M., Amano, J. & Miki, Y. Identification of novel deletion polymorphisms in breast cancer. *Int J. Oncol.* **33**, 261-270 (2008).
24. Long, J. *et al.* A Common Deletion in the APOBEC3 Genes and Breast Cancer Risk. *J. Natl. Cancer Inst.* **105**, 573-579 (2013).
- 30 25. Xuan, D. *et al.* APOBEC3 deletion polymorphism is associated with breast cancer risk among women of European ancestry. *Carcinogenesis* **34**, 2240-2243 (2013).

26. Zhang, T. *et al.* Evidence of associations of APOBEC3B gene deletion with susceptibility to persistent HBV infection and hepatocellular carcinoma. *Hum. Mol. Genet.* **22**, 1262-1269 (2012).
27. Nik-Zainal, S. *et al.* Association of a germline copy number polymorphism of
5 APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. *Nat. Genet.* **46**, 487-491 (2014).
28. Kidd, J.M., Newman, T.L., Tuzun, E., Kaul, R. & Eichler, E.E. Population stratification of a common APOBEC gene deletion polymorphism. *PLoS Genet.* **3**, e63 (2007).
- 10 29. Yu, Q., Chen, D., Konig, R., Mariani, R., Unutmaz, D. & Landau, N.R. APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. *J. Biol. Chem.* **279**, 53379-53386 (2004).
30. Kinomoto, M. *et al.* All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition. *Nucleic Acids Res.* **35**, 2955-2964 (2007).
- 15 31. Caval, V., Suspène R., Vartanian, J.P. & Wain-Hobson, S. Orthologous mammalian APOBEC3A cytidine deaminases hypermutate nuclear DNA. *Mol. Biol. Evol.* **31**, 330-340 (2014).
32. McDougale, R.M., Hultquist, J.F., Stabell, A.C., Sawyer, S.L. & Harris, R.S. D316 is critical for the enzymatic activity and HIV-1 restriction potential of human and rhesus
20 APOBEC3B. *Virology* **441**, 31-39 (2013).
33. Stenglein, M.D., Burns, M.B., Li, M., Lengyel, J. & Harris, R.S. APOBEC3 proteins mediate the clearance of foreign DNA from human cells. *Nat. Struct. Mol. Biol.* **17**, 222-229 (2010).
34. Henry, M., Guétard, D., Suspène, R., Rusniok, C., Wain-Hobson, S. & Vartanian, J.P. Genetic editing of HBV DNA by monodomain human APOBEC3 cytidine deaminases
25 and the recombinant nature of APOBEC3G. *PLoS ONE* **4**, e4277 (2009).
35. Suspène, R., Henry, M., Guillot, S., Wain-Hobson, S. & Vartanian, J.P. Recovery of APOBEC3-edited human immunodeficiency virus G->A hypermutants by differential DNA denaturation PCR. *J. Gen. Virol.* **86**, 125-129 (2005).
- 30 36. Pham, P., Landolph, A., Mendez, C., Li, N. & Goodman, M.F. A biochemical analysis linking APOBEC3A to disparate HIV-1 restriction and skin cancer. *J. Biol. Chem.* **288**, 29294-29304 (2013).

37. Jones, P.A. & Baylin, S.B. The epigenomics of cancer. *Cell* **128**, 683-692 (2007).
38. Madsen, P. *et al.* Psoriasis upregulated phorbolin-1 shares structural but not functional similarity to the mRNA-editing protein apobec-1. *J. Invest. Dermatol.* **113**, 162-169 (1999).
- 5 39. Koning, F.A., Newman, E.N., Kim, E.Y., Kunstman, K.J., Wolinsky, S.M. & Malim, M.H. Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. *J. Virol.* **83**, 9474-9485 (2009).
40. Refsland, E.W., Stenglein, M.D., Shindo, K., Albin, J.S., Brown, W.L. & Harris, R.S. Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues:
10 implications for HIV-1 restriction. *Nucleic Acids Res.* **38**, 4274-4284 (2010).
41. Balkwill, F. & Mantovani, A. Inflammation and cancer: back to Virchow? *Lancet* **357**, 539-545 (2001).

CLAIMS

1. A method for monitoring a cervical HPV infection or HPV induced cervical lesion in a subject and/or respectively detecting a subject's increased risk of progressing to and/or developing cervical squamous cell carcinoma, said subject being a high risk HPV-positive subject or a subject with an HPV induced cervical lesion, comprising the steps of:
- a) providing a nucleic acid sample from a high risk HPV-positive subject or respectively a subject with an HPV-induced cervical lesion;
 - b) performing a genotyping assay on the nucleic acid sample to determine whether the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$); and
 - c) concluding about the follow-up of the subject and/or respectively concluding on the subject's increased risk of progressing to and/or developing cervical squamous cell carcinoma.
2. The method according to claim 1, which is for monitoring a cervical HPV infection in a subject, comprising:
- a) providing a nucleic acid sample from a high risk HPV-positive subject;
 - b) performing a genotyping assay on the nucleic acid sample to determine whether the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$); and
 - c) advising the subject to receive a next Pap test based on the A3 genotype of the subject;
- wherein, if the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$) then the subject is advised to receive the next Pap test at a shorter time interval than if the A3 genotype of the subject does not comprise at least one copy of a A3A-A3B deletion allele ($\Delta A3B$).
3. The method of claim 2, wherein the subject is heterozygous for the $\Delta A3B$ allele.
4. The method of claim 2, wherein the subject is homozygous for the $\Delta A3B$ allele.

5. The method of any one of claims 1 to 4, further comprising d) performing a Pap test on the subject at an increased frequency relative to the Pap test frequency for a subject with an A3 genotype not known to comprise at least one copy of a A3A-A3B deletion allele ($\Delta A3B$).

5

6. The method of any one of claims 1 to 5, wherein if the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$) then the subject is advised to receive the next Pap test after a time interval of less than one year.

10 7. The method of any one of claims 1 to 6, wherein the nucleic acid sample is from a subject that has been diagnosed with a low-grade squamous intraepithelial lesion (LSIL) of the cervix.

15 8. The method of claim 7, wherein the subject has been diagnosed with a cervical intraepithelial neoplasia grade 1 (CIN-1).

9. The method of any one of claims 1 to 6, wherein the nucleic acid sample is from a subject that has been diagnosed with a high-grade squamous intraepithelial lesion (HSIL) of the cervix.

20

10. The method of claim 9, wherein the subject has been diagnosed with a cervical intraepithelial neoplasia grade 2 (CIN-2), cervical intraepithelial neoplasia grade 2/3 (CIN-2/3), or cervical intraepithelial neoplasia grade 3 (CIN-3).

25 11. The method of claim 1, which is for monitoring an HPV induced cervical lesion in a subject, comprising:

a) providing a nucleic acid sample from a subject with an HPV-induced cervical lesion;
b) performing a genotyping assay on the nucleic acid sample to determine whether the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$);

30 and

c) advising the subject to receive a next Pap test based on the A3 genotype of the subject;

wherein, if the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$) then the subject is advised to receive the next Pap test at a shorter time interval than if the A3 genotype of the subject does not comprise at least one copy of a A3A-A3B deletion allele ($\Delta A3B$).

5

12. The method of claim 11, wherein the subject is heterozygous for the $\Delta A3B$ allele.

13. The method of claim 11, wherein the subject is homozygous for the $\Delta A3B$ allele.

10 14. The method of any one of claims 11 to 13, further comprising d) performing a Pap test on the subject at an increased frequency relative to the Pap test frequency for a subject with an A3 genotype not known to comprise at least one copy of a A3A-A3B deletion allele ($\Delta A3B$).

15 15. The method of any one of claims 11 to 14, wherein if the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$) then the subject is advised to receive the next Pap test after a time interval of less than one year.

20 16. The method of any one of claims 11 to 15, wherein the nucleic acid sample is from a subject that has been diagnosed with a low-grade squamous intraepithelial lesion (LSIL) of the cervix.

17. The method of claim 16, wherein the subject has been diagnosed with a cervical intraepithelial neoplasia grade 1 (CIN-1).

25

18. The method of any one of claims 11 to 15, wherein the nucleic acid sample is from a subject that has been diagnosed with a high-grade squamous intraepithelial lesion (HSIL) of the cervix.

30 19. The method of claim 18, wherein the subject has been diagnosed with a cervical intraepithelial neoplasia grade 2 (CIN-2), cervical intraepithelial neoplasia grade 2/3 (CIN-2/3), or cervical intraepithelial neoplasia grade 3 (CIN-3).

20. The method of any one of claims 11 to 17, wherein the nucleic acid sample is from a subject that has tested positive for a high risk type of HPV.

5 21. The method of any one of claims 11 to 20, further comprising providing a cervical sample from the subject and testing the cervical sample for the presence of a high risk type of HPV.

10 22. The method of claim 1, which is for identifying a subject with an HPV induced cervical lesion with an increased risk of developing cervical squamous cell carcinoma, comprising:

a) providing a nucleic acid sample from a subject with an HPV induced cervical lesion;

b) performing a genotyping assay on the nucleic acid sample to determine whether the A3 genotype of the subject comprises at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$);

15 and

c) diagnosing the subject as having an increased risk of cervical squamous cell carcinoma if the A3 genotype of the subject comprises at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$).

20 23. The method of claim 22, wherein the subject is heterozygous for the $\Delta A3B$ allele.

24. The method of claim 22, wherein the subject is homozygous for the $\Delta A3B$ allele.

25 25. The method of any one of claims 22 to 24, wherein the nucleic acid sample is from a subject that has been diagnosed with a low-grade squamous intraepithelial lesion (LSIL) of the cervix.

26. The method of claim 25, wherein the subject has been diagnosed with a cervical intraepithelial neoplasia grade 1 (CIN-1).

30

27. The method of any one of claims 22 to 24, wherein the nucleic acid sample is from a subject that has been diagnosed with a high-grade squamous intraepithelial lesion (HSIL) of the cervix.

5 28. The method of claim 27, wherein the subject has been diagnosed with a cervical intraepithelial neoplasia grade 2 (CIN-2), cervical intraepithelial neoplasia grade 2/3 (CIN-2/3), or cervical intraepithelial neoplasia grade 3 (CIN-3).

10 29. The method of any one of claims 22 to 28, wherein the nucleic acid sample is from a subject that has tested positive for a high risk type of HPV.

30. The method of any one of claims 22 to 29, further comprising providing a cervical sample from the subject and testing the cervical sample for the presence of a high risk type of HPV.

15

31. The method of claim 1, which is for identifying a high risk HPV-positive subject with an increased risk of developing cervical squamous cell carcinoma, comprising:

a) providing a nucleic acid sample from a high risk HPV-positive subject;

20 b) performing a genotyping assay on the nucleic acid sample to determine whether the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$); and

c) diagnosing the subject as having an increased risk of developing cervical squamous cell carcinoma if the A3 genotype of the subject comprises at least one copy of a A3A-A3B deletion allele ($\Delta A3B$).

25

32. The method of claim 31, wherein the subject is heterozygous for the $\Delta A3B$ allele.

33. The method of claim 31, wherein the subject is homozygous for the $\Delta A3B$ allele.

30 34. The method of any one of claims 31 to 33, wherein the nucleic acid sample is from a subject that has been diagnosed with a low-grade squamous intraepithelial lesion (LSIL) of the cervix.

35. The method of claim 34, wherein the subject has been diagnosed with a cervical intraepithelial neoplasia grade 1 (CIN-1).

5 36. The method of any one of claims 31 to 33, wherein the nucleic acid sample is from a subject that has been diagnosed with a high-grade squamous intraepithelial lesion (HSIL) of the cervix.

10 37. The method of claim 36, wherein the subject has been diagnosed with a cervical intraepithelial neoplasia grade 2 (CIN-2), cervical intraepithelial neoplasia grade 2/3 (CIN-2/3), or cervical intraepithelial neoplasia grade 3 (CIN-3).

38. Kit suitable for carrying out a method as defined in any one of claims 1 to 37, comprising:

- 15
- at least one pair of specific oligonucleotide primers specific for hybridization with a target sequence in human genomic DNA, especially SEQ ID NO: 87, and, optionally, one or several of the following reagents,
 - nucleotides (e.g. dATP, dCTP, dGTP, dUTP),
 - a DNA polymerase, in particular a thermostable DNA polymerase,

20

 - optionally, at least one dye for staining nucleic acids,
 - optionally, a buffer solution,
 - optionally, reagents necessary for the hybridation of the primers to their targets,
 - optionally, a positive control,

25

 - optionally, a negative control,
 - optionally, a reference dye and,
 - optionally, a notice providing instructions for use and expected values for interpretation of results.

30 39. Use of a kit according to claim 38, for performing a genotyping assay on a nucleic acid sample of a subject being a high risk HPV-positive subject or a subject with an HPV

induced cervical lesion, to determine whether the A3 genotype of the subject comprises at least one copy of a *A3A-A3B* deletion allele ($\Delta A3B$)

40. Use according to claim 39, in a method for monitoring a cervical HPV infection or HPV
5 induced cervical lesion in a subject and/or respectively detecting a subject's increased risk of progressing to and/or developing cervical squamous cell carcinoma.

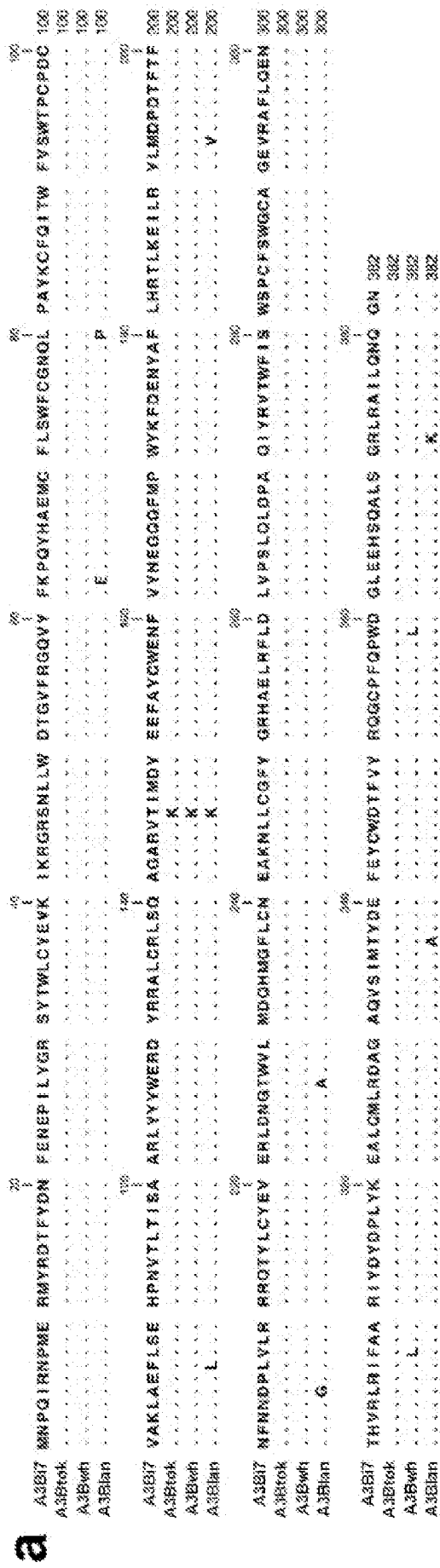


Fig. 1A

b

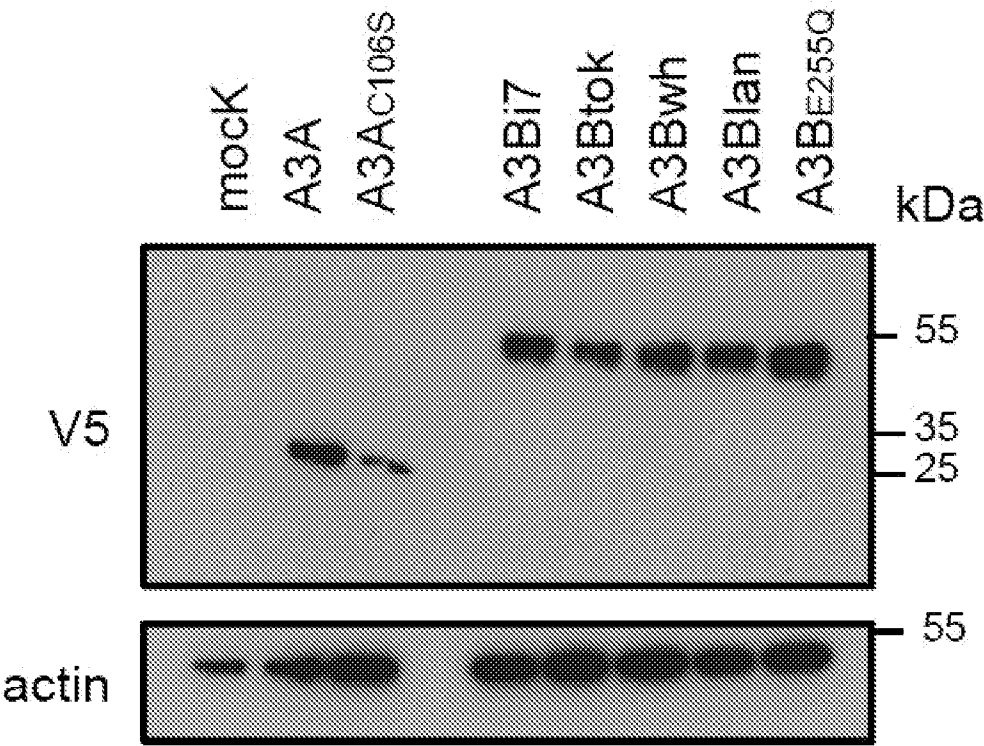


Fig. 1B

c

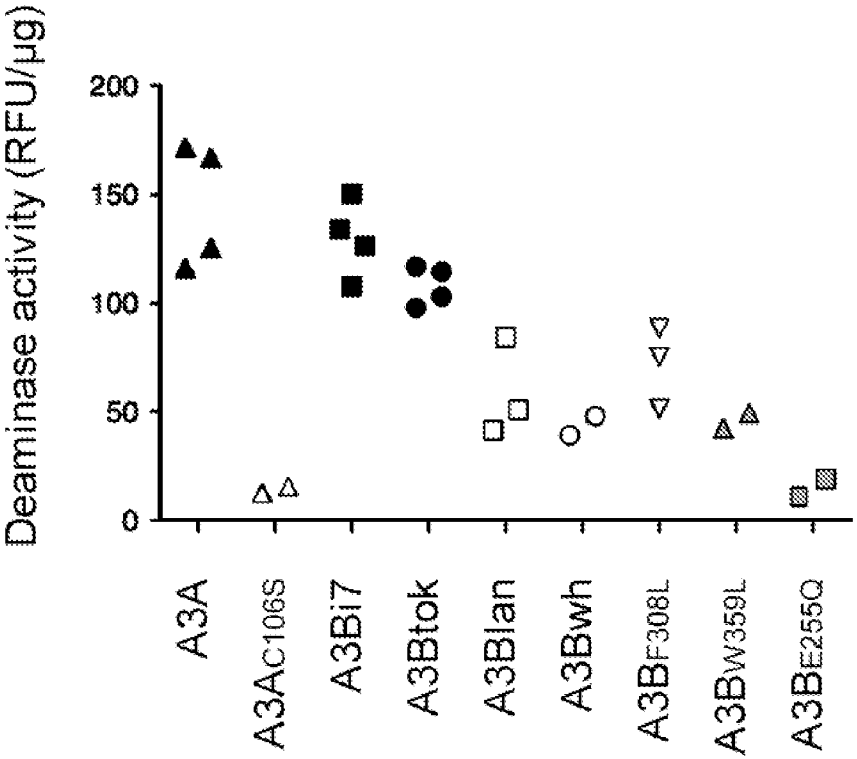
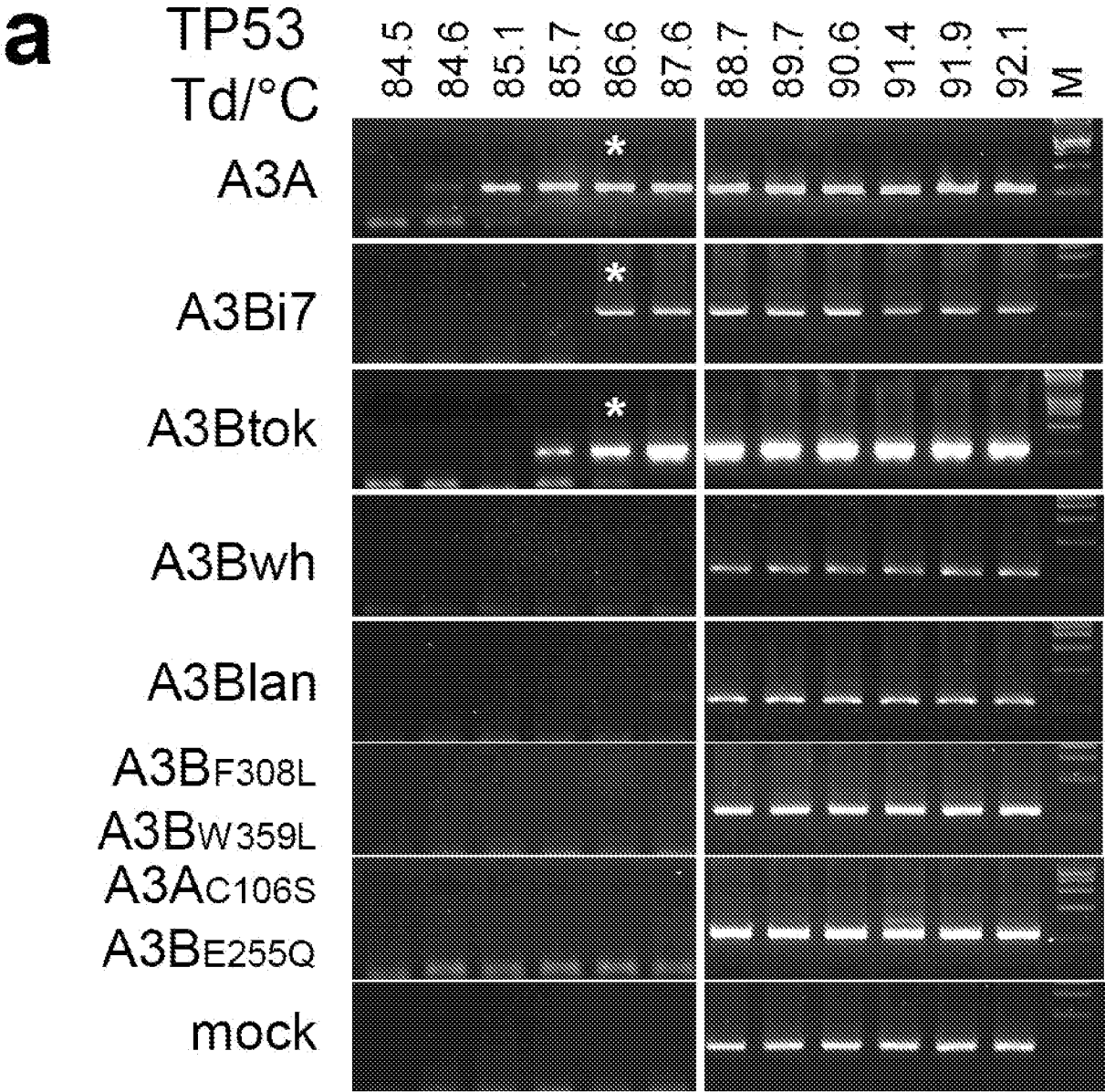


Fig. 1C



TP53 : 100 bp from 178 bp

Fig. 2A

b

	TGCACCCCTTGGTCTCTCCCTCCACCGCTTCTTGTGCTGCTTGTCTTACCTCGCTTAGTGTCTCCCTGGGGGCGACCTCGTGGTGAGGCTCCCCCTTCTTGGGAGA
A3A	T.TT.TT.....T..T.....TTT.....TTT.....T.....TTT.....T.....T.TT.TT.....T.....TT.....TTT.....T.....TTTT.....T.....T.T.TT.....T.....TT.....TTT.....T.....TTTT.....T.....
A3B ⁱ 7	T.....T.....T.....T.....TTTT.....T.....T.TT.T.....T.....T.....TTTT.....T.....T.....T.....T.....TTTT.....T.....
A3Btok	T.....T.TT.TT.....T.....T.....T.....T.TTTT.....T.....T.TT.TT.....T.....T.....T.....TTTT.....T.....T.....T.TT.TT.....T.....T.....T.....TT.....T.....

Fig. 2B

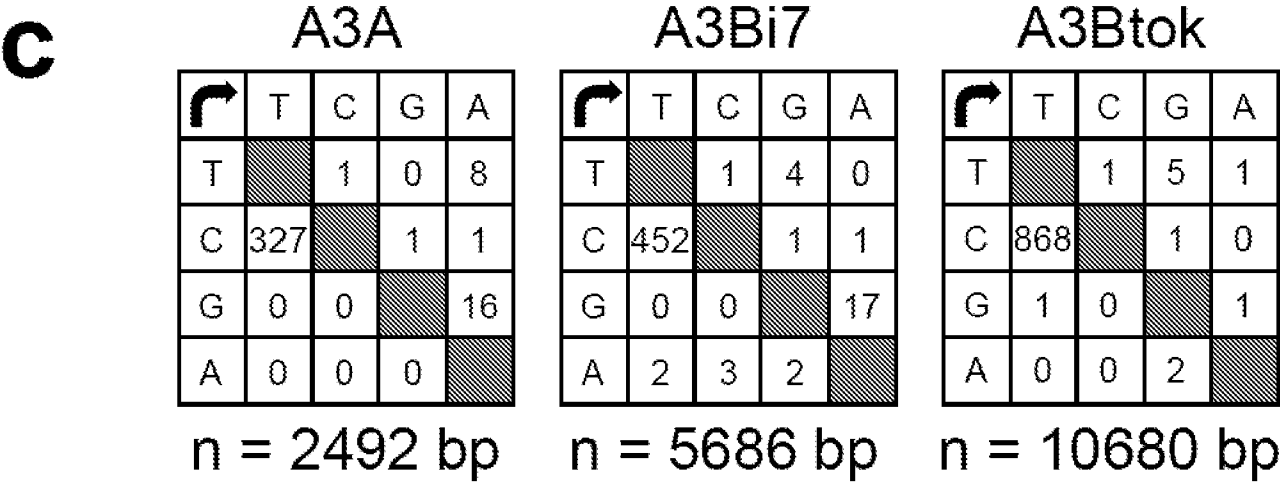


Fig. 2C

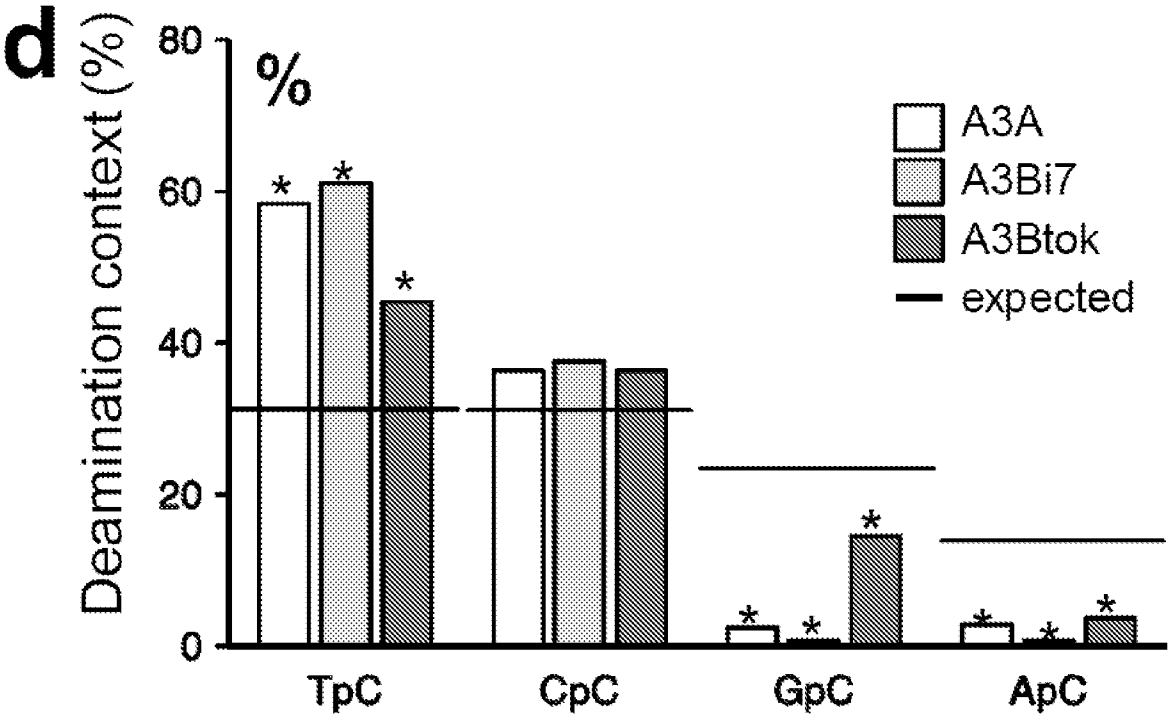


Fig. 2D

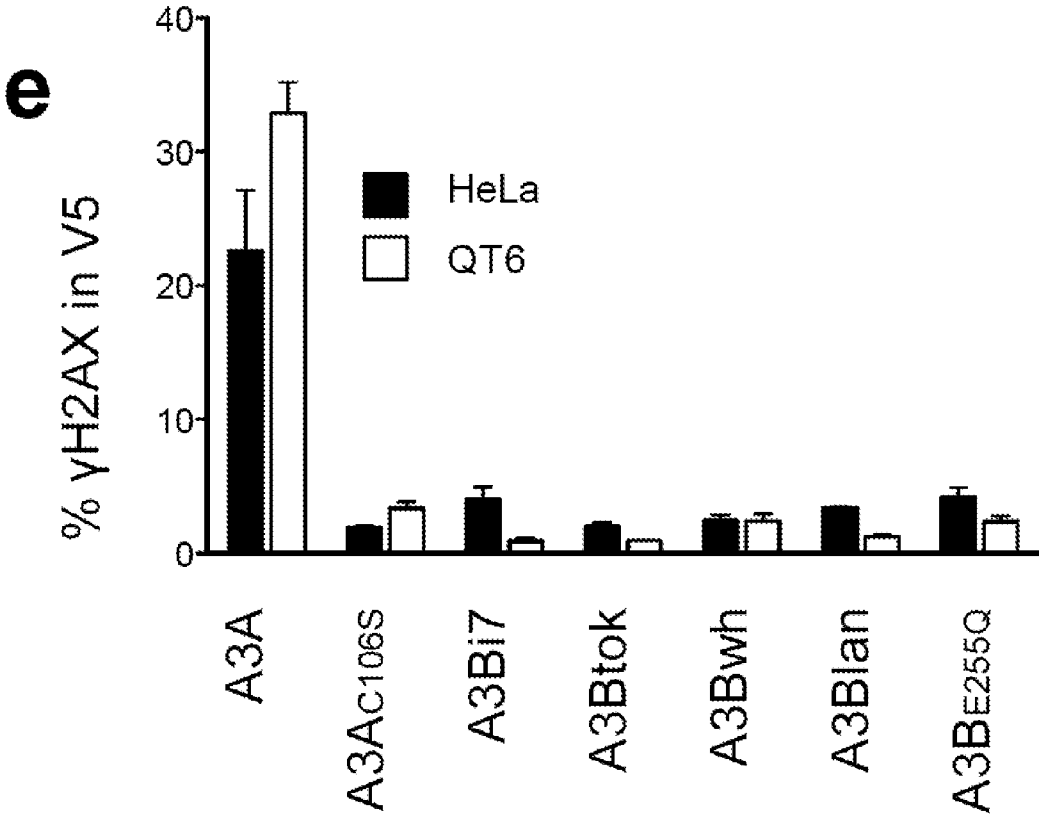


Fig. 2E

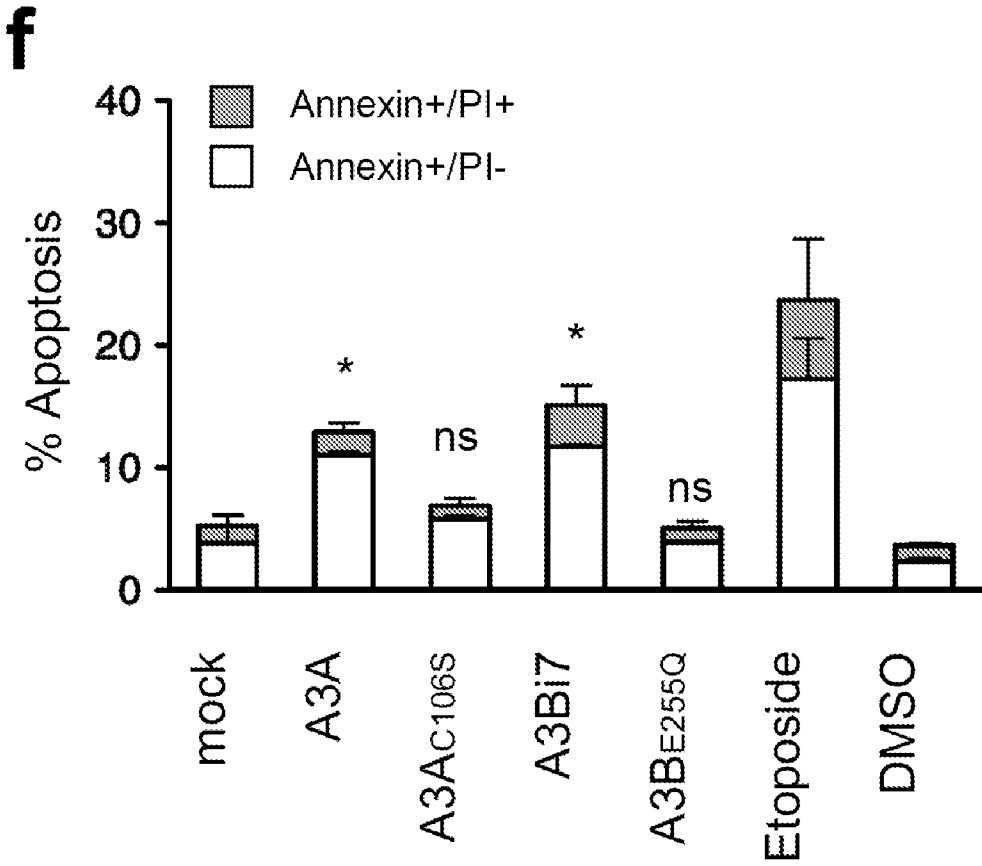


Fig. 2F

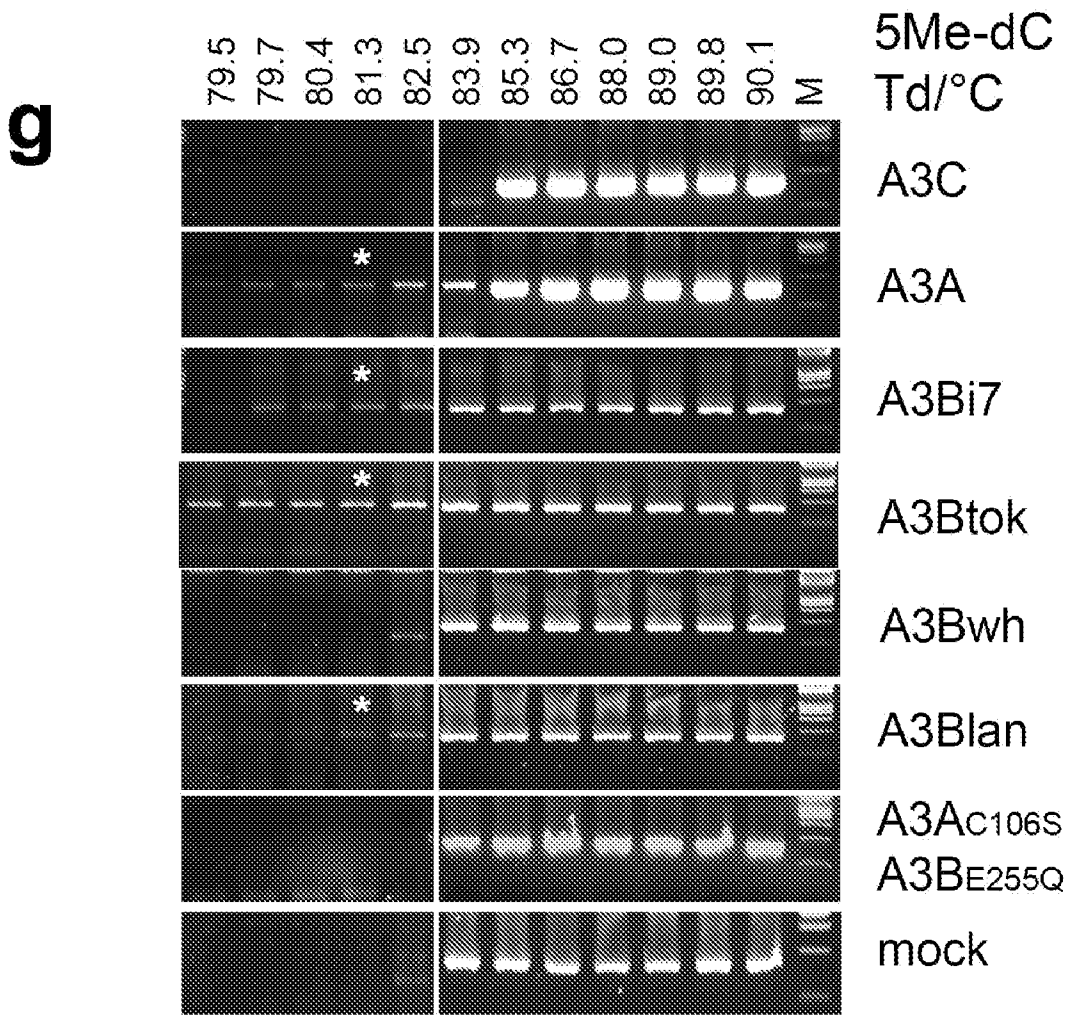


Fig. 2G

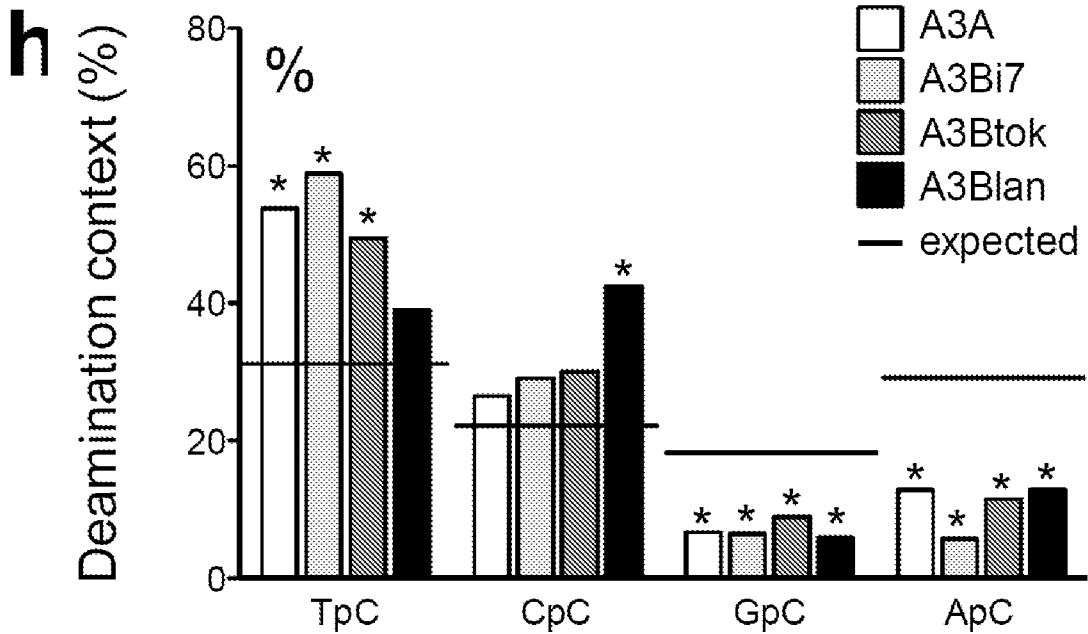


Fig. 2H

a

Normal A3 locus

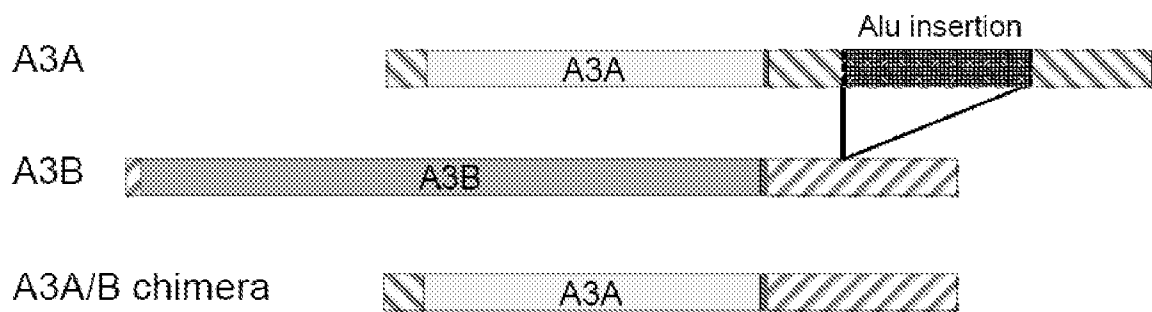
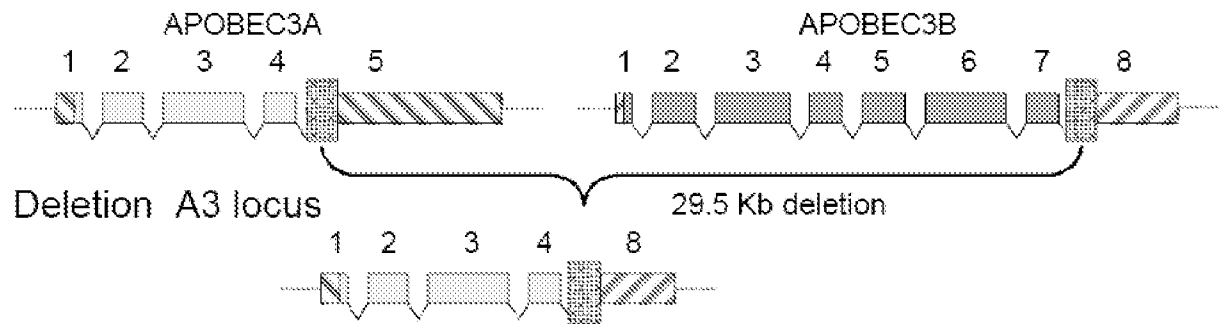


Fig. 3A

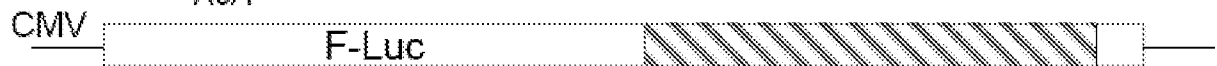
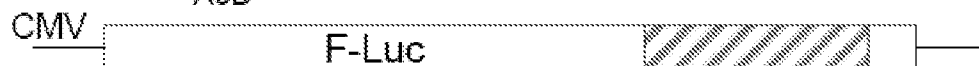
bA3A- Δ UTRA3A-UTR_{A3A}A3A-UTR_{A3B}Luc- Δ UTRLuc-UTR_{A3A}Luc-UTR_{A3B}

Fig. 3B

c

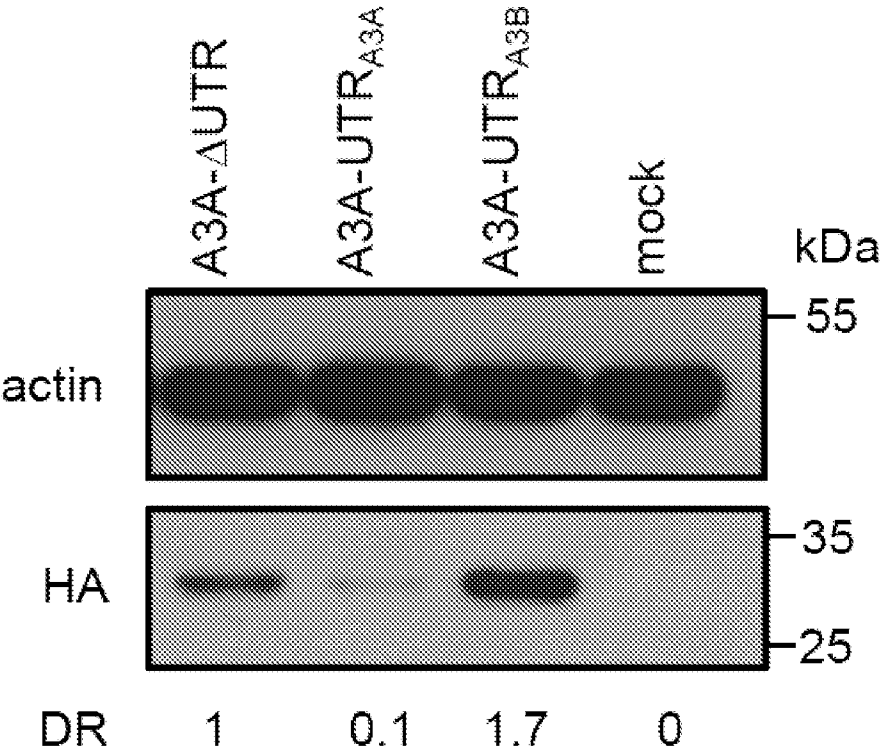


Fig. 3C

d

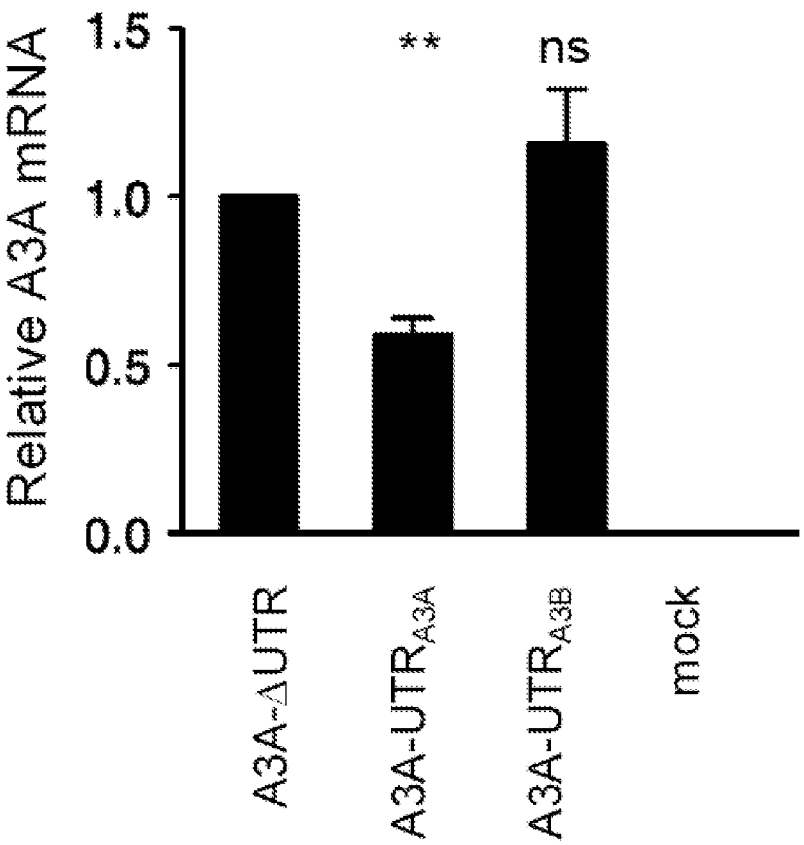


Fig. 3D

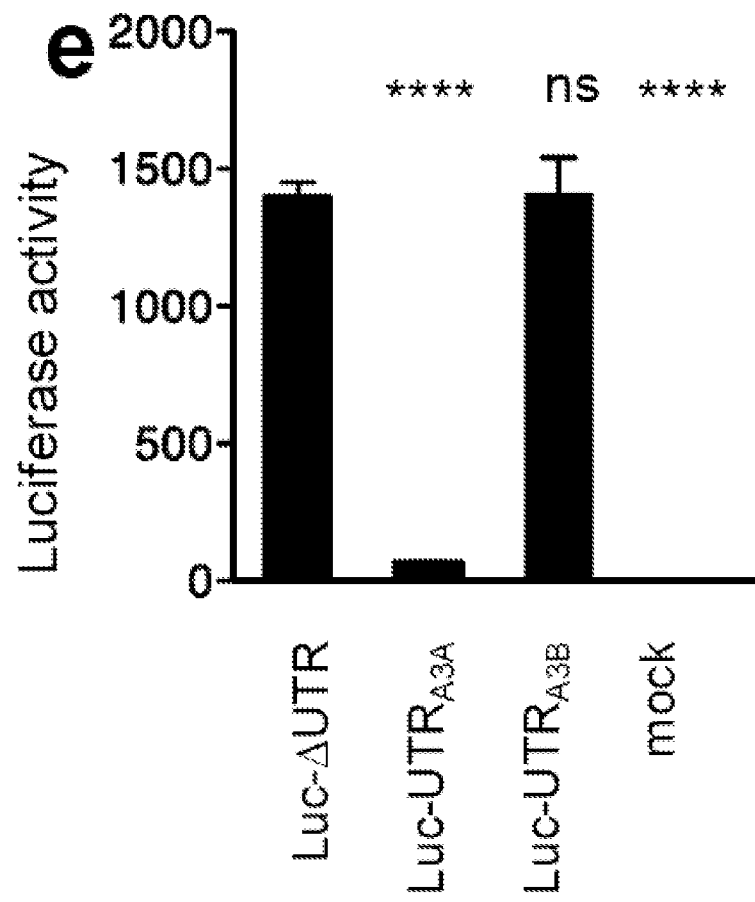


Fig. 3E

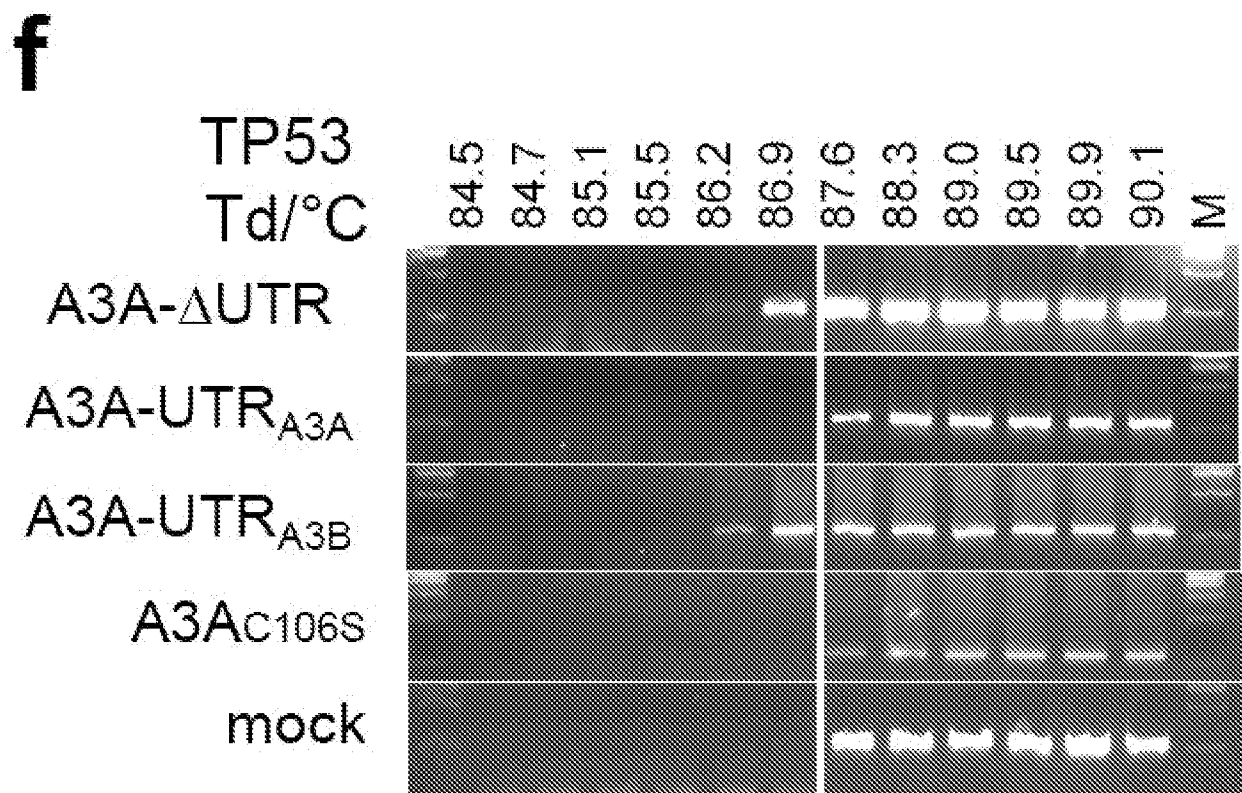


Fig. 3F

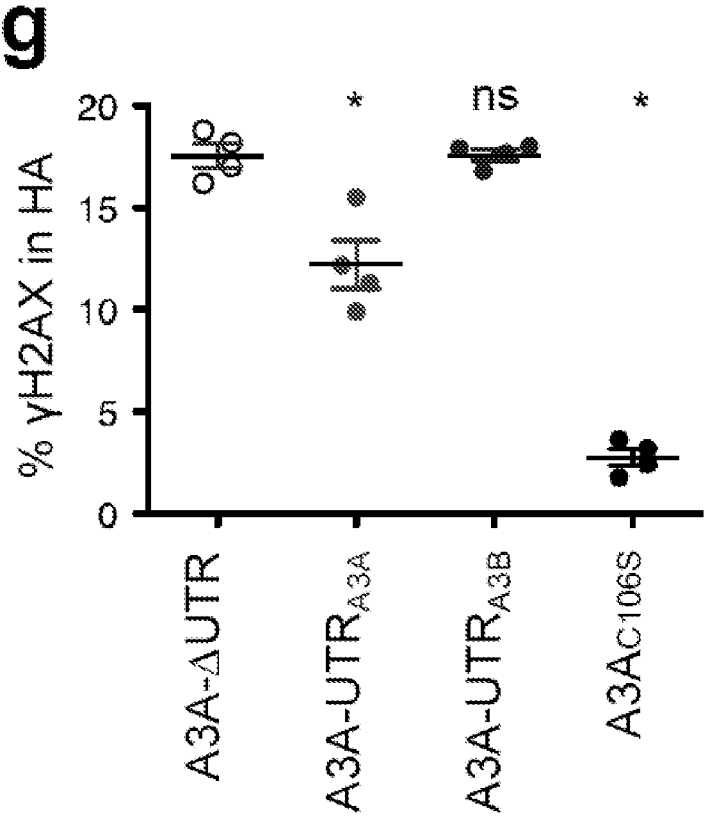


Fig. 3G

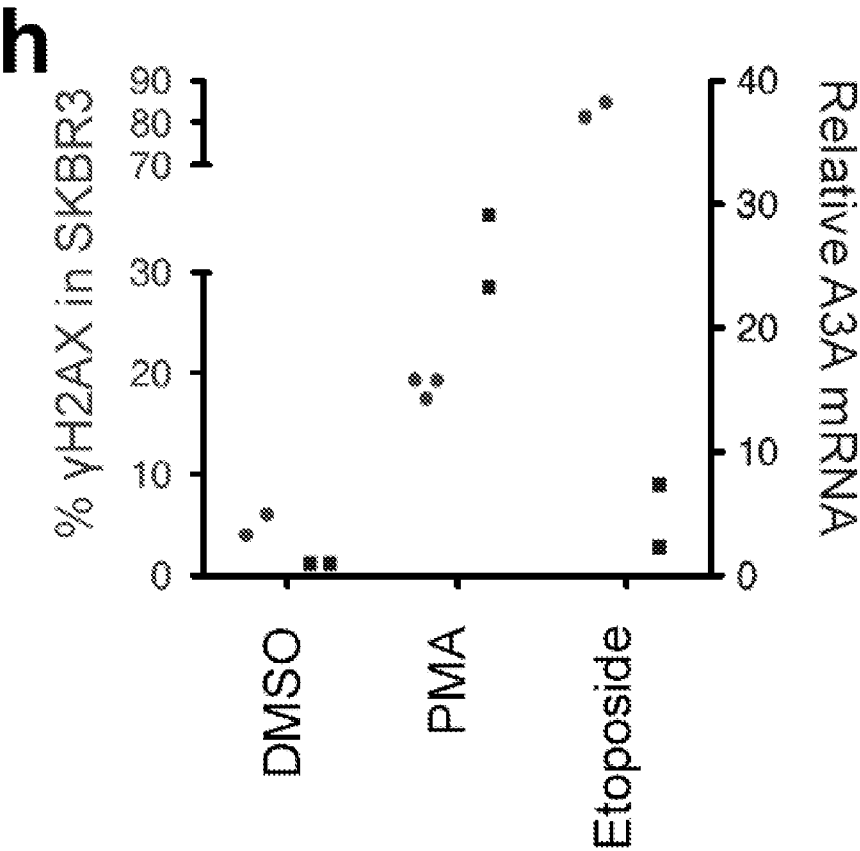


Fig. 3H

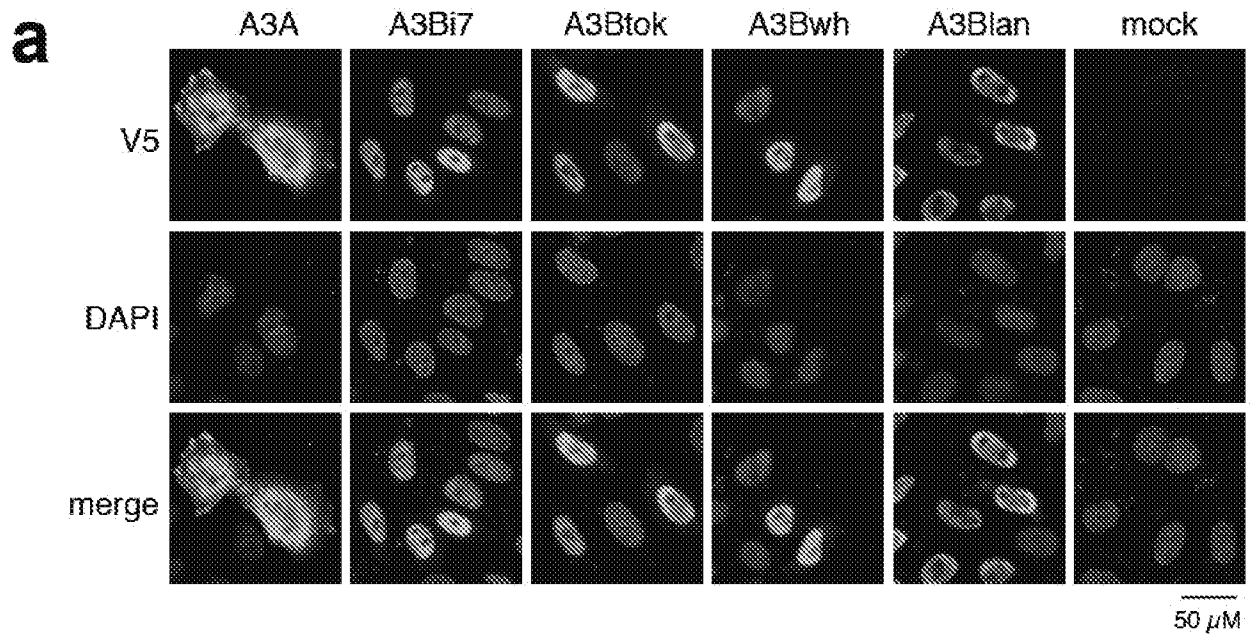


Fig. 4A

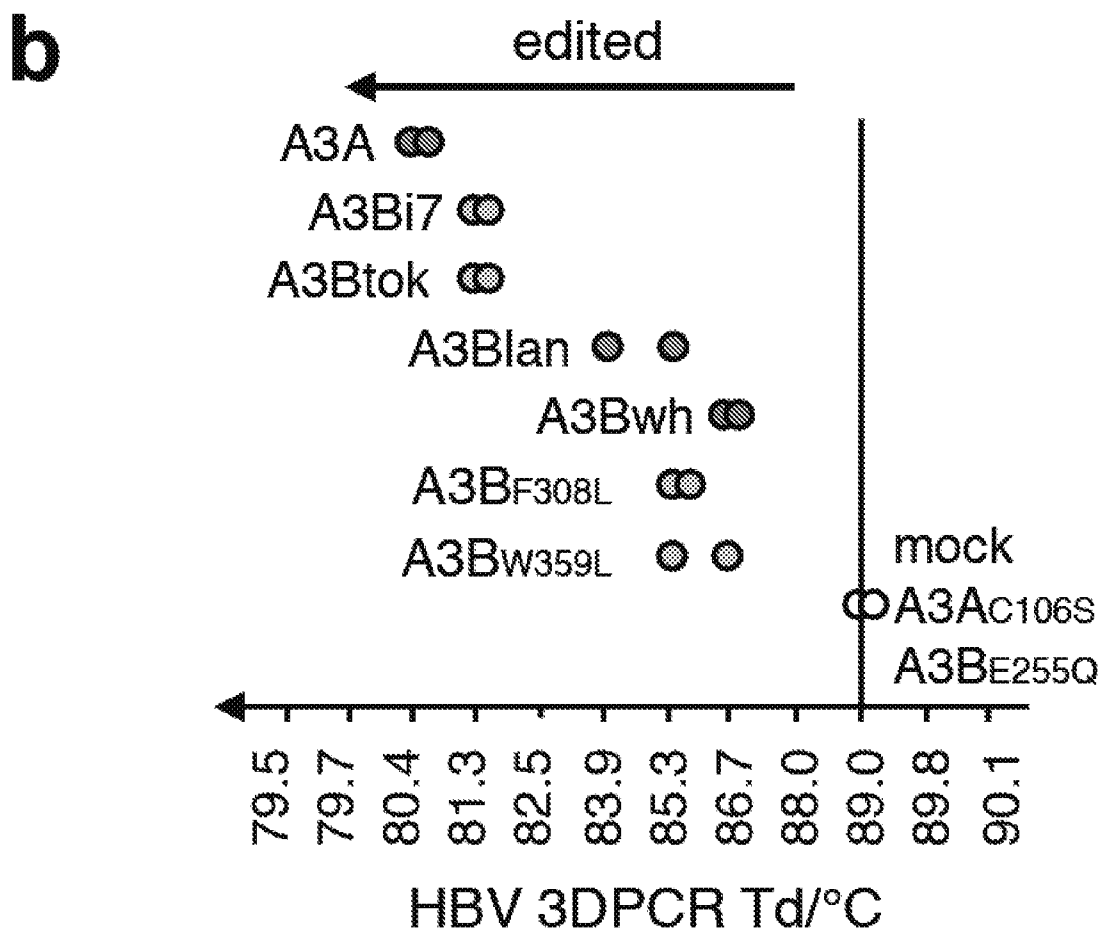
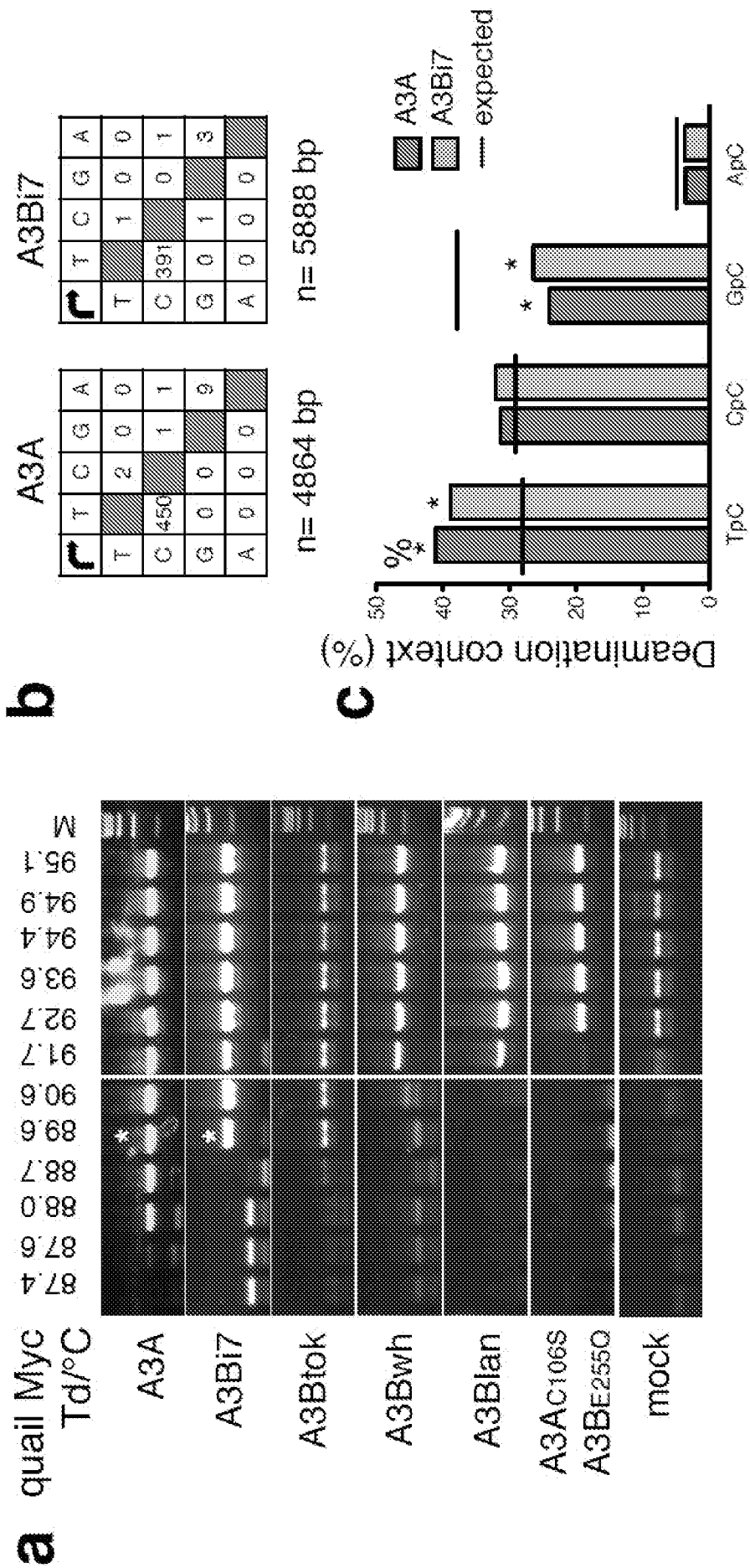


Fig. 4B



5Me-dC HIV-1 : 100 bp from 276 bp

Fig. 5A, 5B, 5C

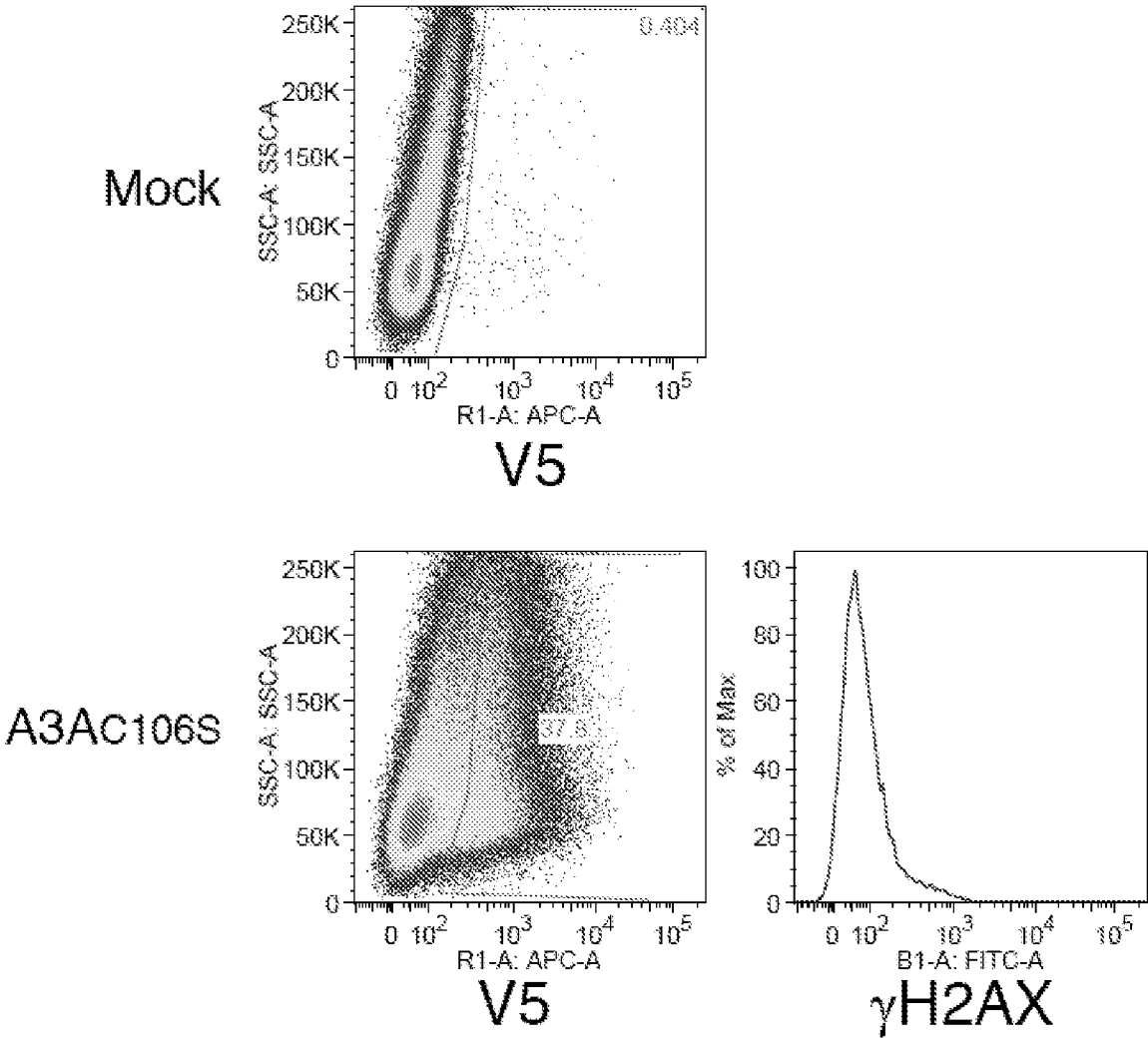


Fig. 6A.1

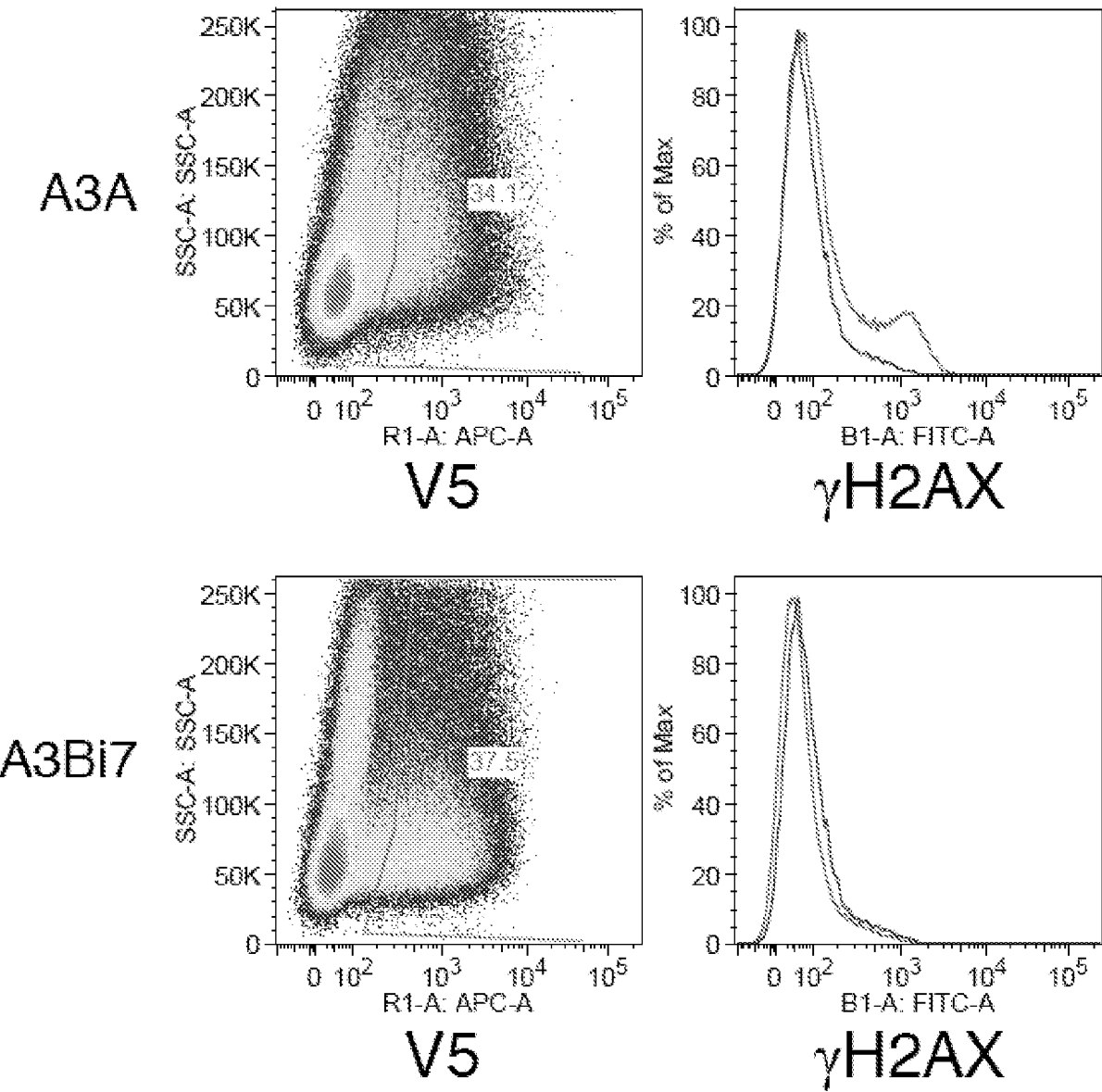


Fig. 6A.2

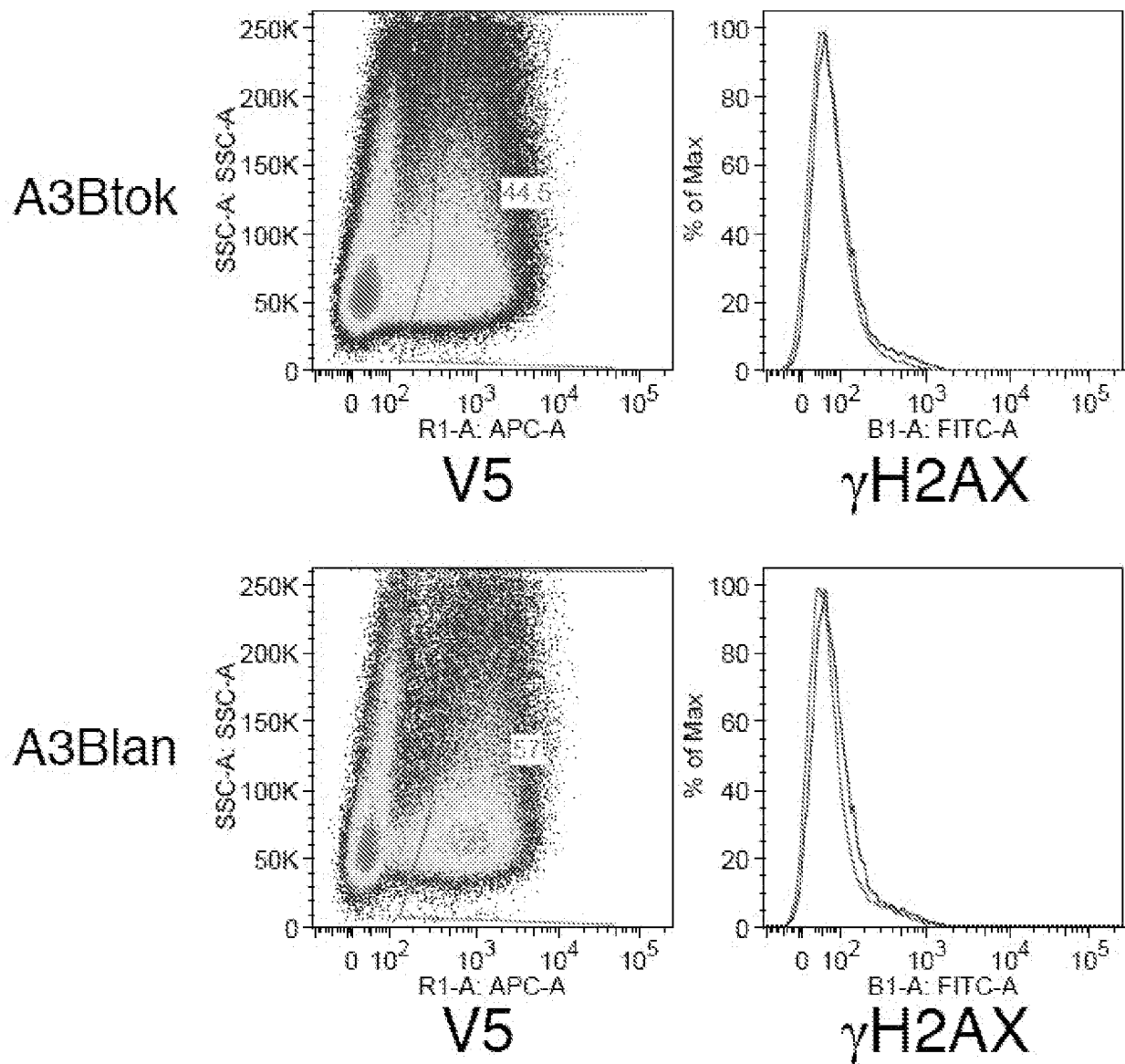


Fig. 6A.3

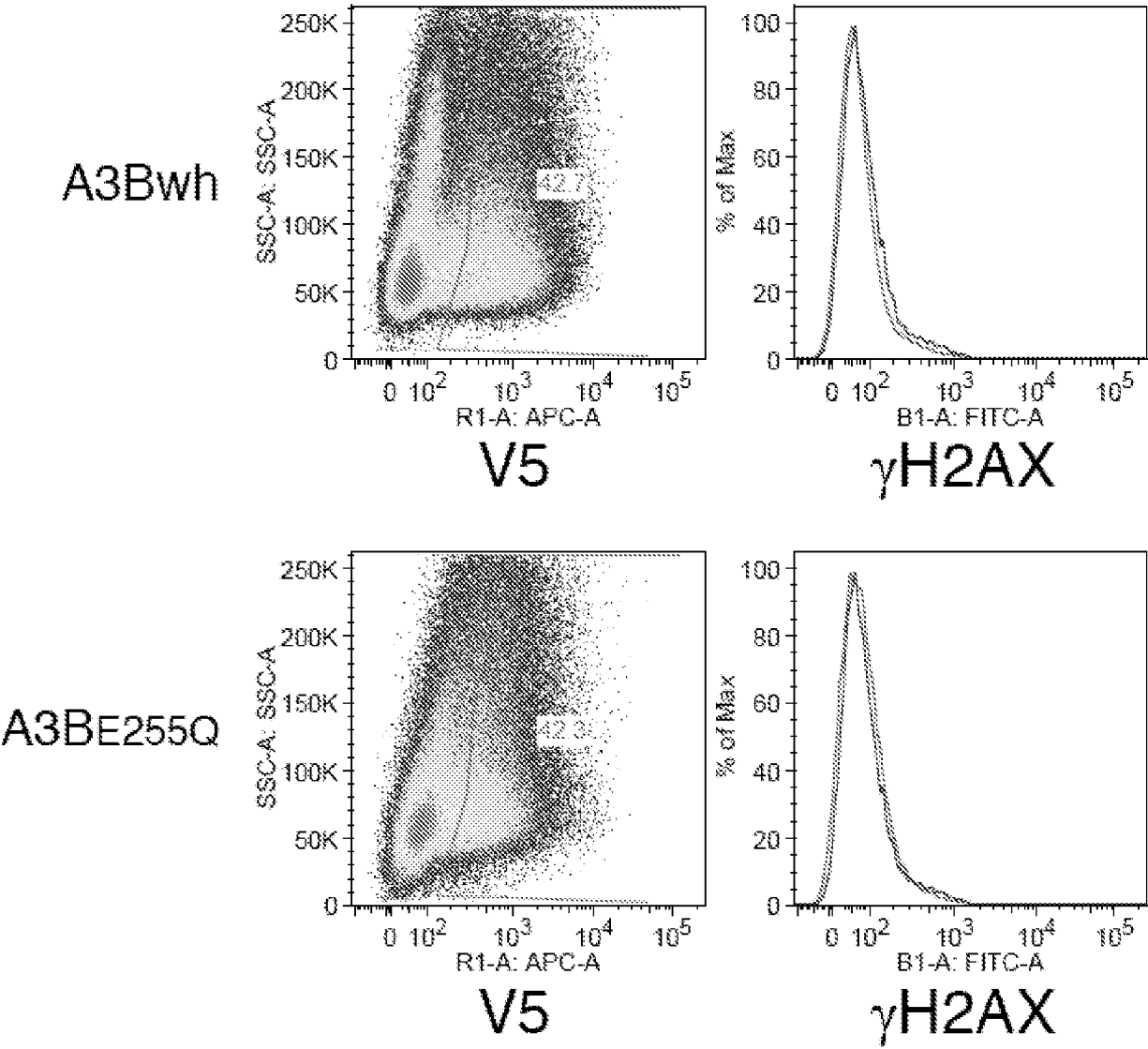


Fig. 6A.4

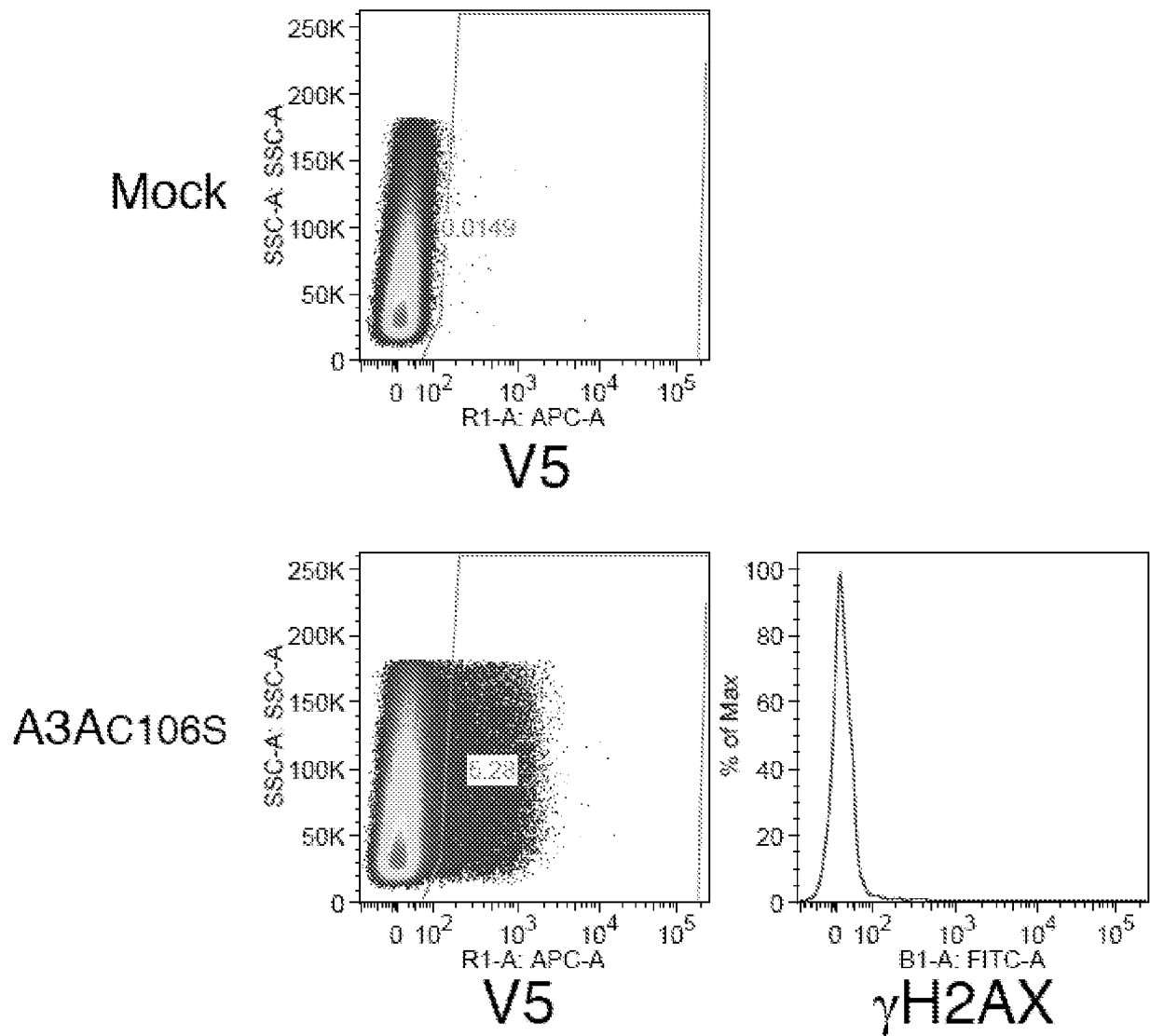


Fig. 6B.1

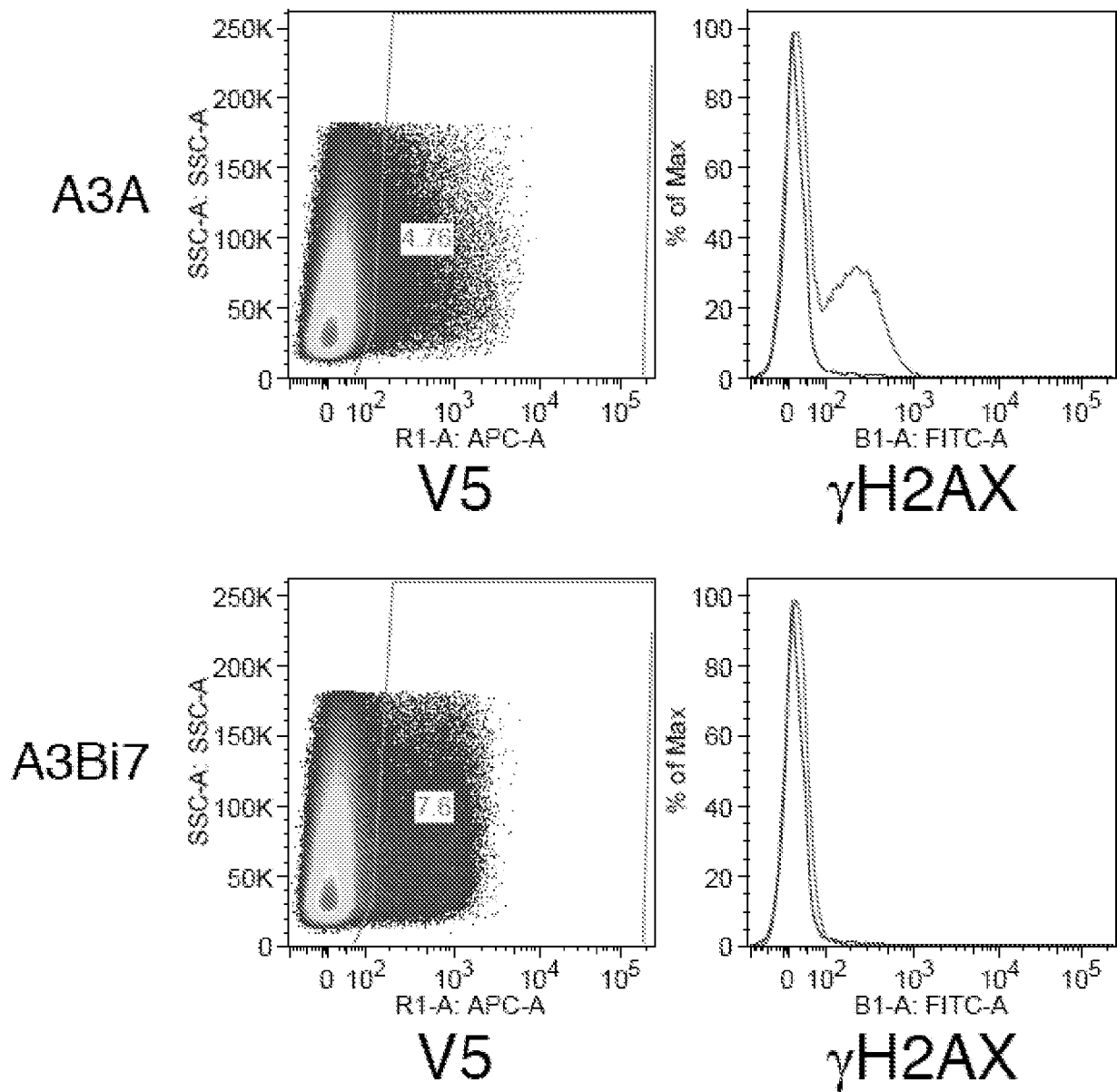


Fig. 6B.2

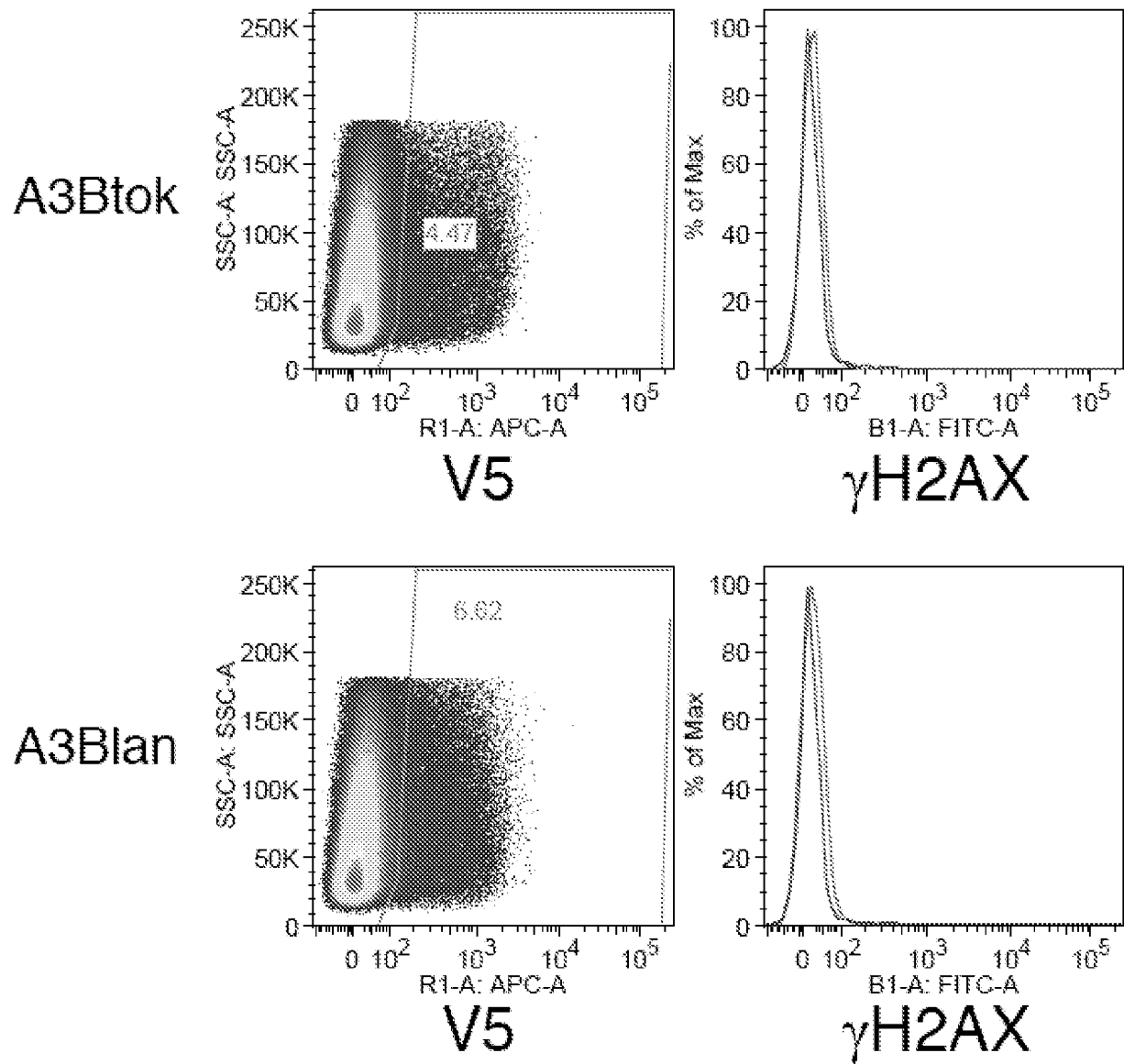


Fig. 6B.3

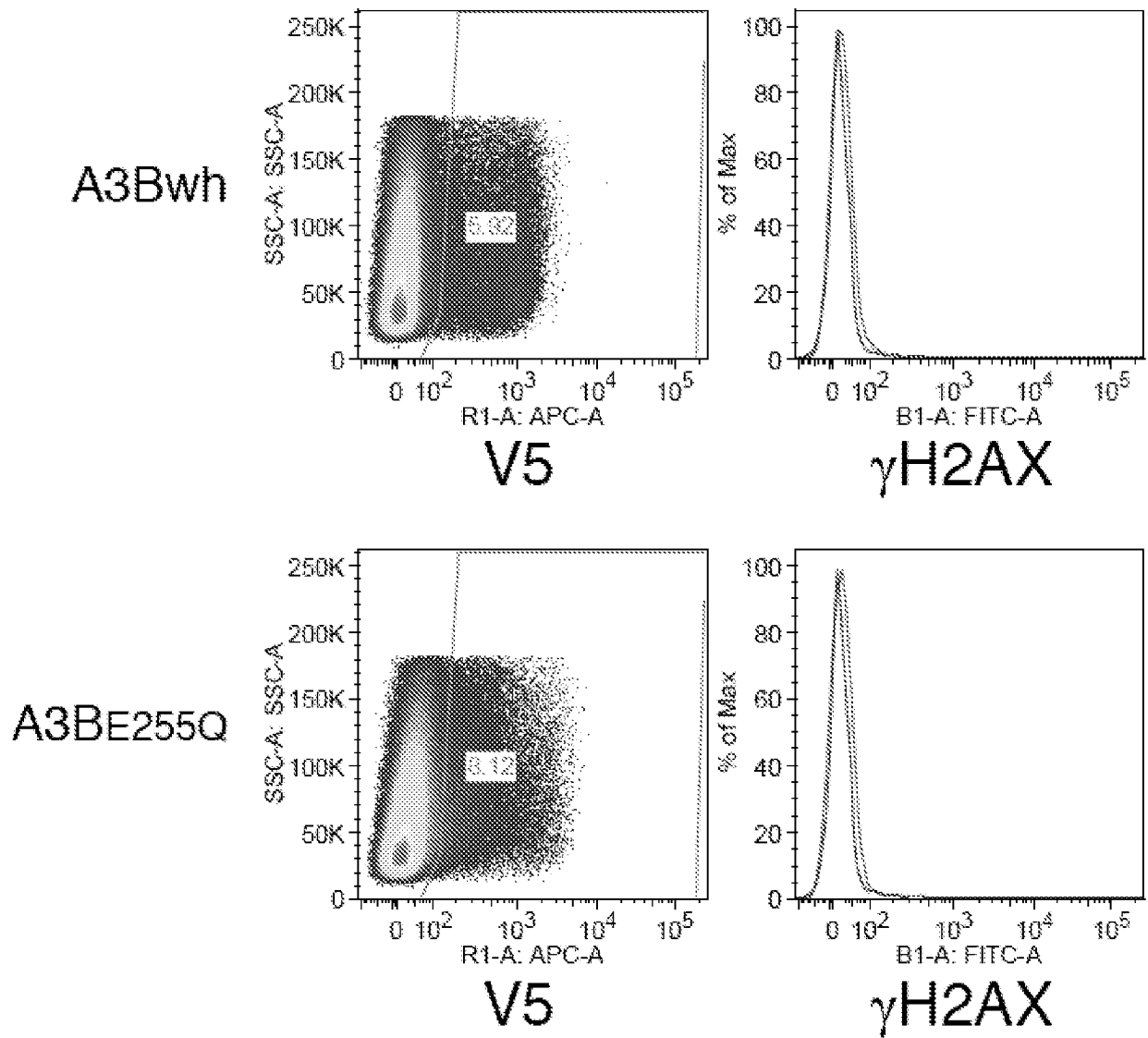


Fig. 6B.4

a

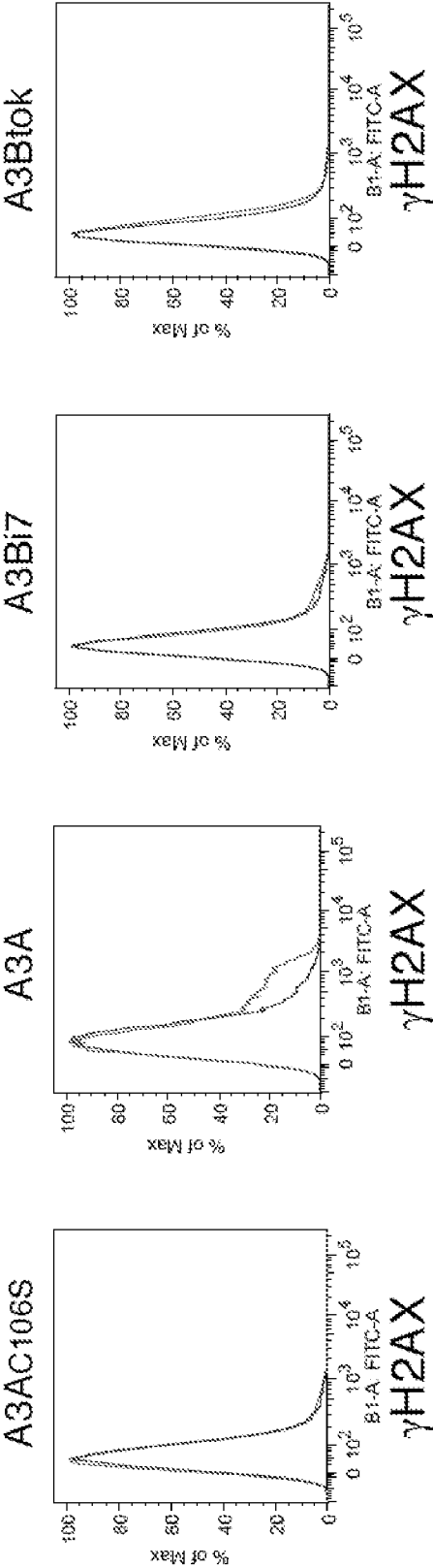


Fig. 7A.1

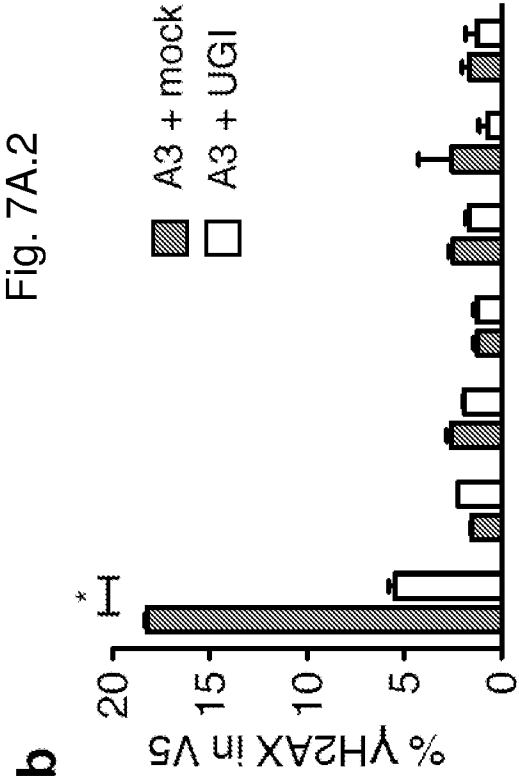
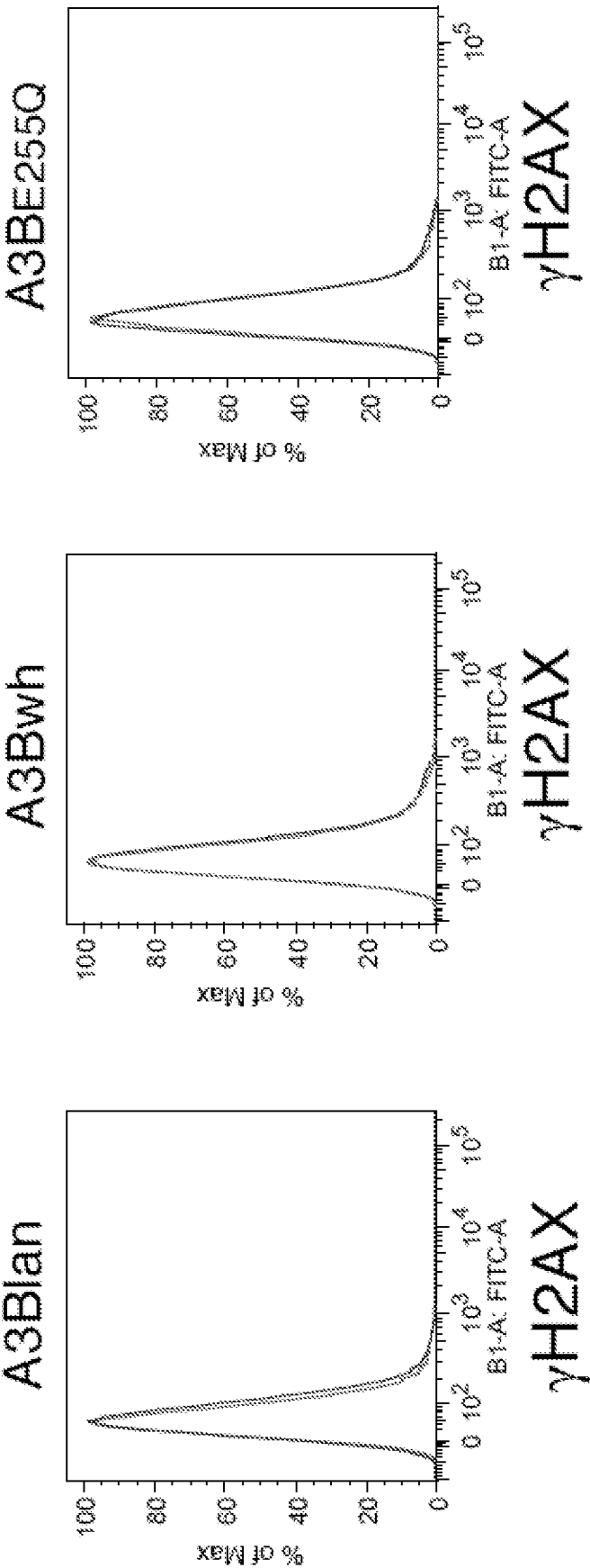


Fig. 7A.2

Fig. 7B

```

exon 5/8
AsnGlnGlyAsn***
A3A-UTR AATCAGGGAAACTGAAAGGATGGGCTCAGTCTCTAAGGAAGGCAGAGACCTGGGTGAGCAGAAATAAAGATCTTCTTCCAAGAAATGCAACAGAC 100
A3B-UTR AATCAGGGAAACTGAAAGGATGGGCTCAGTCTCTAAGGAAGGCAGAGACCTGGGTGAGCAGAAATAAAGATCTTCTTCCAAGAAATGCAACAGAC 100
*****
A3A-UTR CGTTCACCAACCATCTCCAGCTGCTCACAGACGCCAGCAAGCAGTATGCTCCGATCAAGTAGATTTTTAAAAATCAGAGTGGCCGGCGGTGGCT 200
A3B-UTR CGTTCACCAACCATCTCCAGCTGCTCACAGACGCCAGCAAGCAGTATGCTCCGATCAAGTAGATTTTTAAAAATCAGAGT----- 182
*****>>> Alu element
A3A-UTR CACGCCCTGTATCCAGCAGCACTTTGGAGGCCAAGCGGGTGGATCAGGAGTCAGAGATCGAGACCATCTCGGTAAACCGGTGAACCCCTGTCTCTACT 300
A3B-UTR ----- 182
A3A-UTR AAAAAATACAAAAATTAGCCAGGCGTGGTGGGCCCTGTAGTCCAGCTACTCTGGAGGCTGAGGCAGGAGAGTAGCGTGAACCCGGGAGGCAGAGC 400
A3B-UTR ----- 182
A3A-UTR TTGGCGGTGAGCCGAGATTGGCGTACTGCACTCCAGCTGGCCACAGTACAGACTCCATCTCAAAAAAACCAGACTGAATTAATTTAACTGA 500
A3B-UTR ----- 199
*****
A3A-UTR ABAATTTCTCTTATGTTCBAAGTACACAATAGTAAGATTATGCTCAATATTTTCAGAAATAATTTCAATGTATTAATGAATGATAATTTGGCTTC 600
A3B-UTR ABAATTTCTCTTATGTTCBAAGTACACAATAGTAAGATTATGCTCAATATTTTCAGAAATAATTTCAATGTATTAATGAAG-----TGATTAAATTTGGCTCC 294
*****
A3A-UTR ATATCTAGACTACACAAAATTAAGAAATCTTCCATAATTTGCTTTTCTCAGTAACTGTGTCATGAATTCAGAGATTTCACAAACACTAGCAAAATGTGT 700
A3B-UTR BC031803, AK024854 <
BC053859 <
A3A-UTR AGATGTCTTTTCTTGTGTAGCCGGACCTGTAGCCGGGAAAGGTCACACACATCCCTCTGGATCCAGAAAACCTCAGCTAAACCTCAGAGGAGGAACCTA 800
A3B-UTR AGATGTCTTTTCTTGTGTAGCCGGACCTGTAGCTGGGAAAGGTCACACACATCCCTCTGGATCCAGAAAACCTCAGCTAAACCTCAGAGGAGGAACCTA 459
*****
A3A-UTR AATGCAGACCCACCCCTCAGTCAACAGAGCCCGCCCGCCCTCAGTCAAGAGCCCGGGCCGCTGATTGG 869
A3B-UTR AATGCAGACCCACCCCTCAGTCAACAGAGCCCGCCCGCCCTCAGTCAAGAGCCCGGGCCGCTGATTGG 528
*****

```

Fig. 8

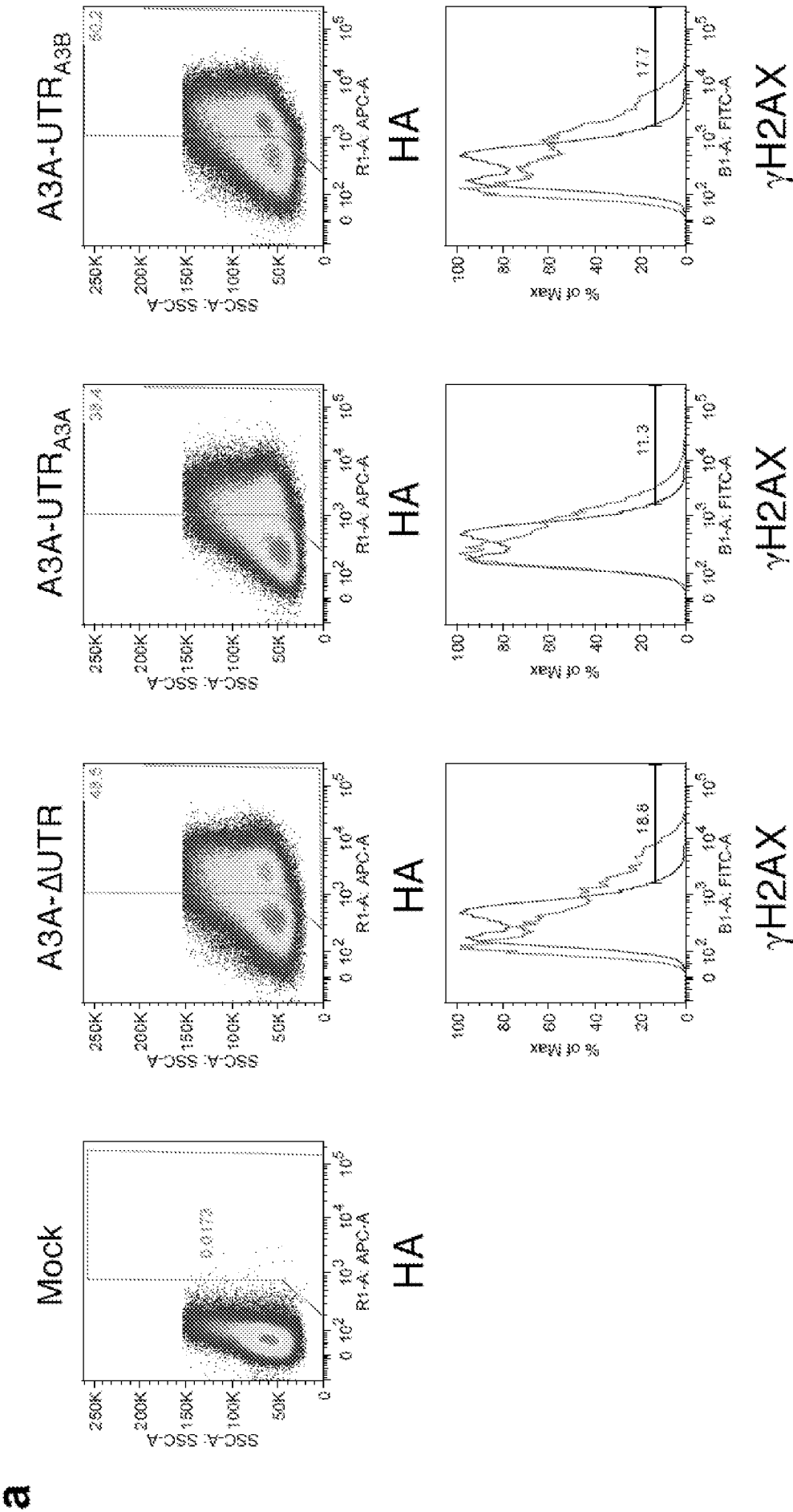


Fig. 9A

b

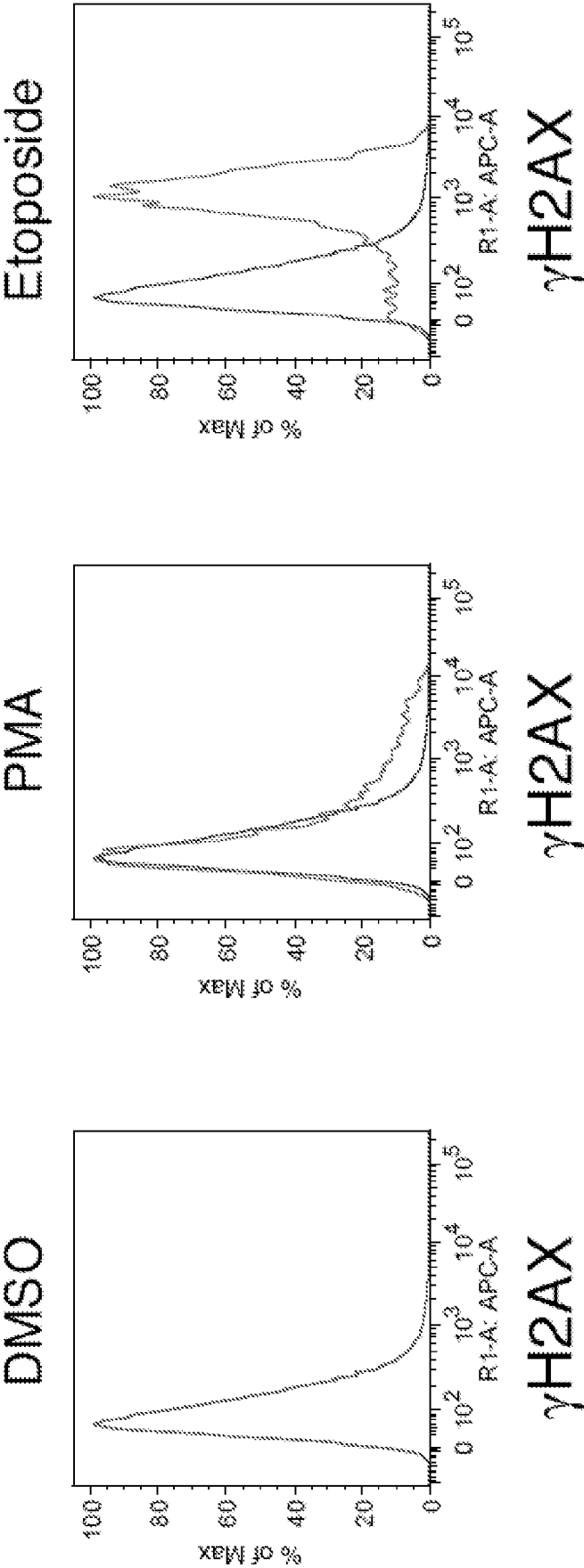


Fig. 9B

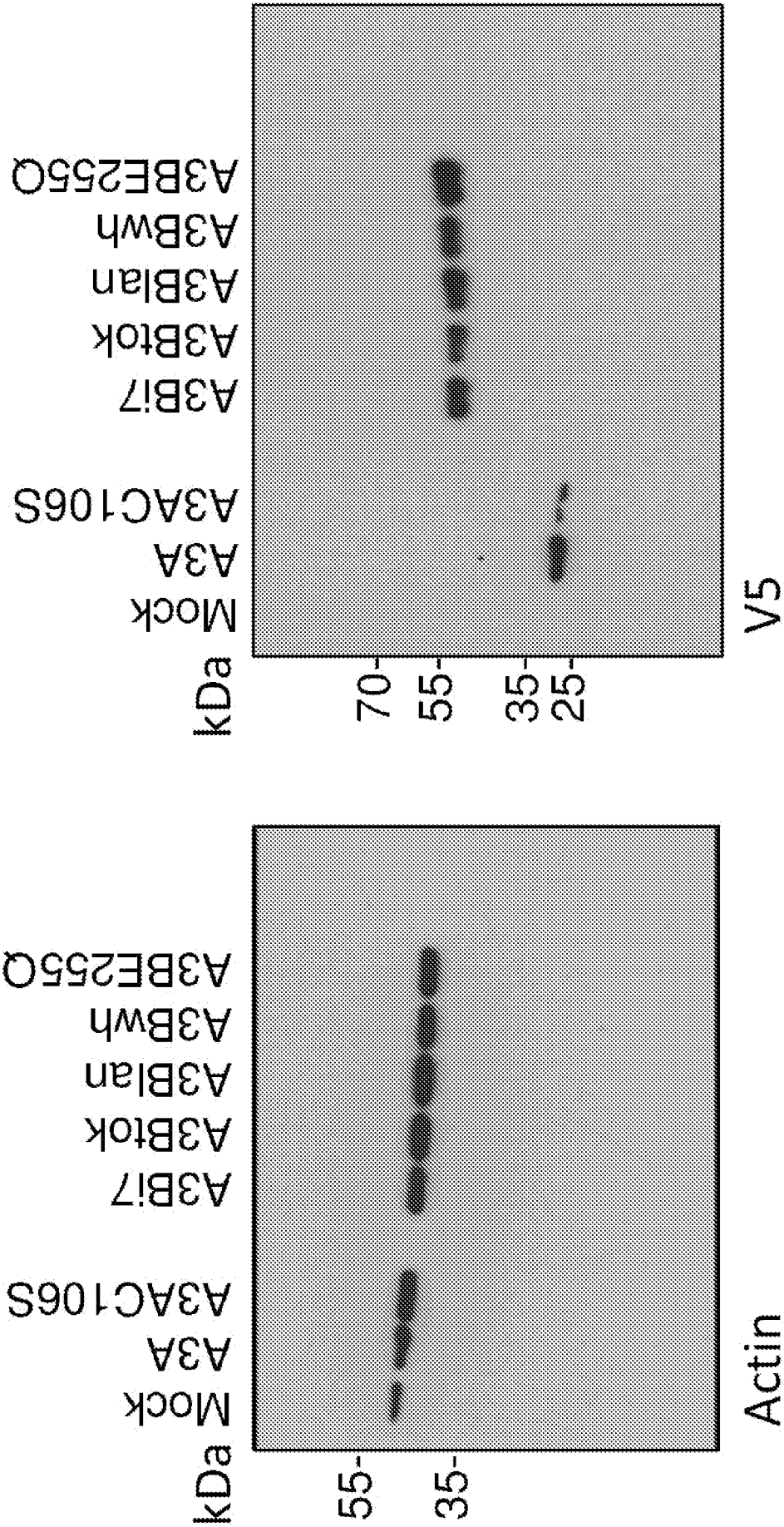


Fig. 10A

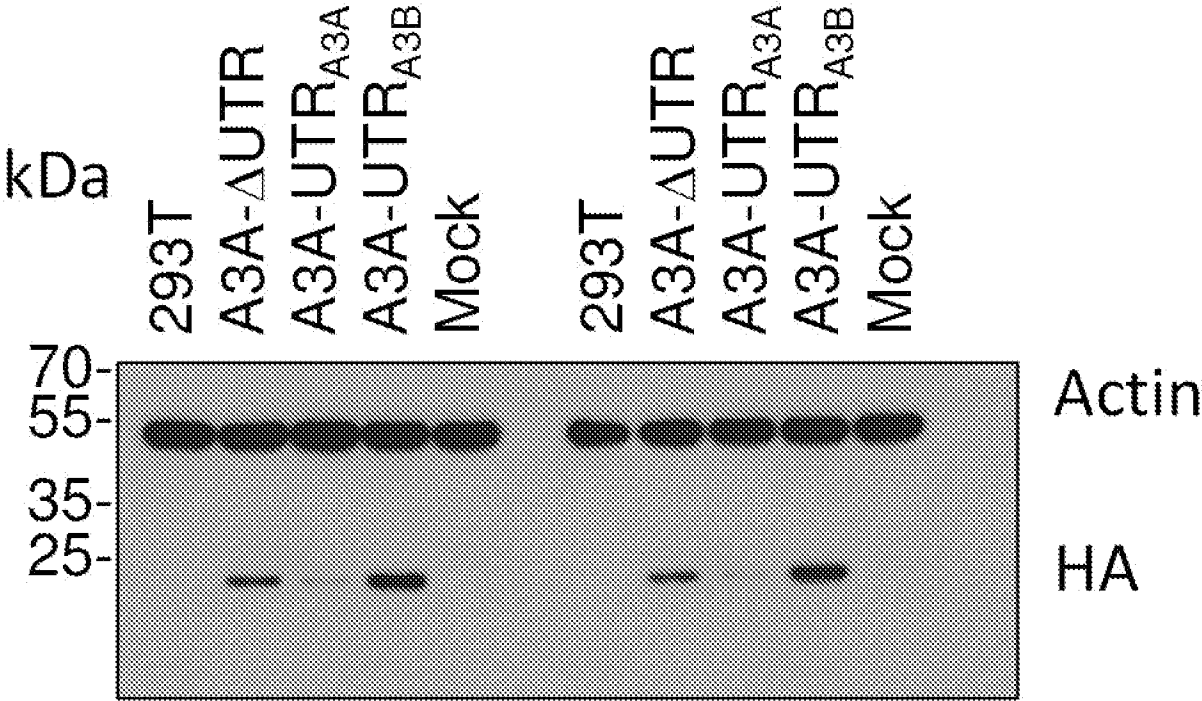


Fig. 10B

```
CAGCATCTGG CACTGTGGTC GCTTCTTTAC TCTGCCTCCC TCCTCTGGAG 39370703
> >>>>>>>>> >>>>>>>>>
CTCCCGTGAA TGTGAGCGTC CTCCTCCTGG CTGGACGGCA TCTCCTCTCT 39370753
>>>>> >>>>>>>>>> >>>>>
CTGGGCCTTC TGCCCTGGCC CCATCCTCCC TCGGACTGGG TGCCGAGTCT 39370803
CAATGTTGGT CACTTATGAC CCCAAATCGC CTTCTCCAGC TGGACACCCC 39370853
CAGTCTCCTG GGCTGGACTC TGCAGTCACC TTTAGTGCTC CCAGCTCCCC 39370903
<<<<<<<<<< <<<<<<<<<<
TGTTCCCTAAG CCCTGCTGTT TCTACCTTTG GGAGGACACG CAGGCTGTCA 39370953
<<<<<<<<<< <<<<<<<<<< <
```

Fig. 11

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2015/001973

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/68 C12Q1/70
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2004/018711 A2 (UNIV LONDON [GB]; MING-QING DU [GB]) 4 March 2004 (2004-03-04) claims 1-12	1-37, 39, 40
Y	----- SERENA NIK-ZAINAL ET AL: "Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer", NATURE GENETICS., vol. 46, no. 5, 13 April 2014 (2014-04-13) , pages 487-491, XP055242124, NEW YORK, US ISSN: 1061-4036, DOI: 10.1038/ng.2955 cited in the application the whole document ----- -/--	1-37, 39, 40

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

20 January 2016

Date of mailing of the international search report

02/02/2016

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Bradbrook, Derek

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2015/001973

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JEFFREY M. KIDD ET AL: "Population Stratification of a Common APOBEC Gene Deletion Polymorphism", PLOS GENETICS, vol. 296, no. 4, 1 April 2007 (2007-04-01) , page 2225, XP055242128, ISSN: 0193-4511, DOI: 0193-4511(2002)296[2225:TSOHBI]2.0.CO;2 cited in the application page 590, column 1, paragraph 4 -----	38
X	D. XUAN ET AL: "APOBEC3 deletion polymorphism is associated with breast cancer risk among women of European ancestry", CARCINOGENESIS., vol. 34, no. 10, 28 May 2013 (2013-05-28), pages 2240-2243, XP055242139, GB ISSN: 0143-3334, DOI: 10.1093/carcin/bgt185 cited in the application abstract	38
A	-----	1-37,39, 40
Y	STEVEN A ROBERTS ET AL: "An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers", NATURE GENETICS., vol. 45, no. 9, 14 July 2013 (2013-07-14), pages 970-976, XP055242141, NEW YORK, US ISSN: 1061-4036, DOI: 10.1038/ng.2702 cited in the application the whole document -----	1-37,39, 40
A,P	MORI SEIICHIRO ET AL: "Identification ofAPOBEC3Bpromoter elements responsible for activation by human papillomavirus type 16 E6", BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 460, no. 3, 20 March 2015 (2015-03-20), pages 555-560, XP029156526, ISSN: 0006-291X, DOI: 10.1016/J.BBRC.2015.03.068 abstract ----- -/--	1-40

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2015/001973

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A,P	Shan Chen ET AL: "APOBEC3A possesses anticancer and antiviral effects by differential inhibition of HPV E6 and E7 expression on cervical cancer", Int J Clin Exp Med., Vol. 8, 15 July 2015 (2015-07-15), pages 10548-10557, XP055242158, Retrieved from the Internet: URL:http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4565227/pdf/ijcem0008-10548.pdf [retrieved on 2016-01-18] the whole document -----	1-40
A,P	VINCENT CAVAL ET AL: "A prevalent cancer susceptibility APOBEC3A hybrid allele bearing APOBEC3B 3'UTR enhances chromosomal DNA damage", NATURE COMMUNICATIONS, vol. 5, 9 October 2014 (2014-10-09), page 5129, XP055242177, United Kingdom ISSN: 2041-1723, DOI: 10.1038/ncomms6129 the whole document -----	1-40

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB2015/001973

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☒ forming part of the international application as filed:
 - ☒ in the form of an Annex C/ST.25 text file.
 - ☐ on paper or in the form of an image file.
 - b. ☐ furnished together with the international application under PCT Rule 13~~ter~~.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
 - c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
 - ☐ in the form of an Annex C/ST.25 text file (Rule 13~~ter~~.1(a)).
 - ☐ on paper or in the form of an image file (Rule 13~~ter~~.1(b) and Administrative Instructions, Section 713).
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/IB2015/001973

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
W0 2004018711 A2	04-03-2004	AU 2003255823 A1	11-03-2004
		W0 2004018711 A2	04-03-2004
