



(12) **DEMANDE DE BREVET CANADIEN**
CANADIAN PATENT APPLICATION
(13) **A1**

(86) Date de dépôt PCT/PCT Filing Date: 2016/07/25
(87) Date publication PCT/PCT Publication Date: 2017/02/02
(85) Entrée phase nationale/National Entry: 2018/01/23
(86) N° demande PCT/PCT Application No.: EP 2016/067621
(87) N° publication PCT/PCT Publication No.: 2017/017049
(30) Priorité/Priority: 2015/07/27 (GB1513176.6)

(51) Cl.Int./Int.Cl. *A61K 39/04* (2006.01),
C07K 14/075 (2006.01), *C12N 15/861* (2006.01),
C12N 7/00 (2006.01)

(71) Demandeur/Applicant:
GLAXOSMITHKLINE BIOLOGICALS S.A., BE

(72) Inventeurs/Inventors:
CORTESE, RICCARDO (DECEASED), CH;
COLLOCA, STEFANO, IT;
VITELLI, ALESSANDRA, IT;
NICOSIA, ALFREDO, IT;
GRAZIOLI, FABIANA, IT;
AMMENDOLA, VIRGINIA, IT

(74) Agent: NORTON ROSE FULBRIGHT CANADA
LLP/S.E.N.C.R.L., S.R.L.

(54) Titre : NOUVEL ADENOVIRUS
(54) Title: NOVEL ADENOVIRUS

(57) **Abrégé/Abstract:**

There is provided adenoviral vectors encoding a mycobacterial antigen derived from a chimp adenovirus, and to related aspects.

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

(19) World Intellectual Property Organization

International Bureau

(43) International Publication Date

2 February 2017 (02.02.2017)





(10) International Publication Number

WO 2017/017049 A1

(51) International Patent Classification:
A61K 39/04 (2006.01) C12N 15/861 (2006.01)
C12N 7/00 (2006.01) C07K 14/075 (2006.01)

(21) International Application Number:
PCT/EP2016/067621

(22) International Filing Date:
25 July 2016 (25.07.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1513176.6 27 July 2015 (27.07.2015) GB

(71) Applicant: GLAXOSMITHKLINE BIOLOGICALS S.A. [BE/BE]; rue de l'Institut 89, B-1330 Rixensart (BE).

(72) Inventor: COLLOCA, Stefano; GlaxoSmithKline, Avenue Fleming 20, B-1300 Wavre (BE).

(74) Agent: BAKER, Suzanne Jane; GlaxoSmithKline, Global Patents CN925.1, 980 Great West Road, Brentford Middlesex TW8 9GS (GB).

(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).

Declarations under Rule 4.17:

— of inventorship (Rule 4.17(iv))

Published:

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

— with sequence listing part of description (Rule 5.2(a))

(54) Title: NOVEL ADENOVIRUS

(57) Abstract: There is provided adenoviral vectors encoding a mycobacterial antigen derived from a chimp adenovirus, and to related aspects.

WO 2017/017049 A1

NOVEL ADENOVIRUS

FIELD OF THE INVENTION

5 The present invention relates to adenoviral vectors encoding a mycobacterial antigen, more particular to adenoviral vectors derived from chimp adenovirus ChAd155, and to related aspects.

BACKGROUND OF THE INVENTION

10 Tuberculosis (TB) is a chronic infectious disease caused by infection with *Mycobacterium tuberculosis* and other *Mycobacterium* species. It is a major disease in developing countries, as well as an increasing problem in developed areas of the world.

15 Vaccination is one of the most effective methods for preventing or treating infectious diseases. Adenovirus has been widely used for gene transfer applications due to its ability to achieve highly efficient gene transfer in a variety of target tissues and large transgene capacity. Conventionally, E1 genes of adenovirus are deleted and replaced with a transgene cassette consisting of the promoter of choice, cDNA sequence of the gene of interest and a poly A
20 signal, resulting in a replication defective recombinant virus.

Recombinant adenoviruses are useful in gene therapy and as vaccines. Viral vectors based on chimpanzee adenovirus represent an alternative to the use of human derived adenoviral (Ad) vectors for the development of genetic vaccines. Adenoviruses isolated from
25 chimpanzees are closely related to adenoviruses isolated from humans as demonstrated by their efficient propagation in cells of human origin. However, since human and chimp adenoviruses are close relatives, serologic cross reactivity between the two virus species is possible.

30 There is a demand for vectors which effectively deliver molecules to a target and minimize the effect of pre-existing immunity to selected adenovirus serotypes in the population. One aspect of pre-existing immunity that is observed in humans is humoral immunity, which can result in the production and persistence of antibodies that are specific for adenoviral proteins. The humoral response elicited by adenovirus is mainly directed against the three major structural
35 capsid proteins: fiber, penton and hexon.

There remains a need for novel methods of immunising against tuberculosis, which are highly efficacious, safe, convenient, cost-effective, long-lasting and induce a broad spectrum of immune responses.

40 Vectors, compositions and methods of the present invention may have one or more following improved characteristics over the prior art, including but not limited to higher productivity, improved immunogenicity and increased transgene expression.

SUMMARY OF THE INVENTION

- 5 The present invention provides a recombinant adenovirus comprising at least one polynucleotide or polypeptide selected from the group consisting of:
- (a) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 1,
 - 10 (b) a polynucleotide which encodes a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1,
 - (c) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 3,
 - 15 (d) a polypeptide having the amino acid sequence according to SEQ ID NO: 1,
 - (e) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1, and
 - 20 (f) a polypeptide having the amino acid sequence according to SEQ ID NO: 3:
- wherein the adenovirus comprises a nucleic acid sequence encoding a mycobacterial antigen, wherein the nucleic acid sequence is operatively linked to one or more sequences which direct expression of said mycobacterial antigen in a host cell.
- 25 Also provided is a composition comprising the recombinant adenovirus and a pharmaceutically acceptable excipient.

The recombinant adenovirus and compositions may be used as medicaments, in particular for the stimulation of an immune response against mycobacterial infection, such as

30 *Mycobacterium tuberculosis* infection.

DESCRIPTION OF THE FIGURES

- Figure 1A-C - Alignment of fiber protein sequences from the indicated simian adenoviruses.
- 35 ChAd3 (SEQ ID NO:27)
 - PanAd3 (SEQ ID NO:28)
 - ChAd17 (SEQ ID NO:29)
 - ChAd19 (SEQ ID NO:30)
 - ChAd24 (SEQ ID NO:31)
 - 40 ChAd155 (SEQ ID NO:1)
 - ChAd11 (SEQ ID NO:32)
 - ChAd20 (SEQ ID NO:33)
 - ChAd31 (SEQ ID NO:34)

	PanAd1 (SEQ ID NO:35)
	PanAd2 (SEQ ID NO:36)
Figure 2 –	Flow diagram for production of specific ChAd155 BAC and plasmid vectors
Figure 3 –	Species C BAC Shuttle #1365 schematic
5 Figure 4 –	pArsChAd155 Ad5E4orf6-2 (#1490) schematic
Figure 5 –	pChAd155/RSV schematic
Figure 6 –	BAC ChAd155/RSV schematic
Figure 7 –	Productivity of ChAd3 and ChAd155 vectors expressing an HIV Gag transgene (Experiment 1)
10 Figure 8 –	Productivity of ChAd3 and ChAd155 vectors expressing an HIV Gag transgene (Experiment 2)
Figure 9 –	Productivity of PanAd3 and ChAd155 vectors expressing RSV transgene
Figure 10 –	Expression levels of ChAd3 and ChAd155 vectors expressing an HIV Gag transgene
15 Figure 11 –	Expression levels of PanAd3 and ChAd155 vectors expressing an HIV Gag transgene – Western Blot
Figure 12 –	Immunogenicity of ChAd3 and ChAd155 vectors expressing an HIV Gag transgene – IFN-gamma ELISpot
Figure 13 –	Immunogenicity of PanAd3 and ChAd155 vectors expressing an HIV Gag transgene – IFN-gamma ELISpot
20	

DESCRIPTION OF THE SEQUENCES

SEQ ID NO: 1 -	Polypeptide sequence of ChAd155 fiber
25 SEQ ID NO: 2 -	Polynucleotide sequence encoding ChAd155 fiber
SEQ ID NO: 3 -	Polypeptide sequence of ChAd155 penton
SEQ ID NO: 4 -	Polynucleotide sequence encoding ChAd155 penton
SEQ ID NO: 5 -	Polypeptide sequence of ChAd155 hexon
SEQ ID NO: 6 -	Polynucleotide sequence encoding ChAd155 hexon
30 SEQ ID NO: 7 -	Polynucleotide sequence encoding ChAd155#1434
SEQ ID NO: 8 -	Polynucleotide sequence encoding ChAd155#1390
SEQ ID NO: 9 -	Polynucleotide sequence encoding ChAd155#1375
SEQ ID NO: 10 -	Polynucleotide sequence encoding wild type ChAd155
SEQ ID NO: 11 -	Polynucleotide sequence encoding ChAd155/RSV
35 SEQ ID NO: 12 -	Polynucleotide sequence encoding the CASI promoter
SEQ ID NO: 13 -	Ad5orf6 primer 1 polynucleotide sequence
SEQ ID NO: 14 -	Ad5orf6 primer 2 polynucleotide sequence
SEQ ID NO: 15 -	BAC/CHAd155 Δ E1_TetO hCMV RpsL-Kana primer 1 polynucleotide sequence
40 SEQ ID NO: 16 -	BAC/CHAd155 Δ E1_TetO hCMV RpsL-Kana (#1375) primer 2 polynucleotide sequence
SEQ ID NO: 17 -	1021-FW E4 Del Step1 primer polynucleotide sequence
SEQ ID NO: 18 -	1022-RW E4 Del Step1 primer polynucleotide sequence

	SEQ ID NO: 19 -	1025-FW E4 Del Step2 primer polynucleotide sequence
	SEQ ID NO: 20 -	1026-RW E4 Del Step2 primer polynucleotide sequence
	SEQ ID NO: 21 -	91-SubMonte FW primer polynucleotide sequence
	SEQ ID NO: 22 -	890-BghPolyA RW primer polynucleotide sequence
5	SEQ ID NO: 23 -	CMVfor primer polynucleotide sequence
	SEQ ID NO: 24 -	CMVrev primer polynucleotide sequence
	SEQ ID NO: 25 -	CMVFAM-TAMRA qPCR probe polynucleotide sequence
	SEQ ID NO: 26 -	Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE) polynucleotide sequence
10	SEQ ID NO: 27 -	Amino acid sequence for the fiber protein of ChAd3
	SEQ ID NO: 28 -	Amino acid sequence for the fiber protein of PanAd3
	SEQ ID NO: 29 -	Amino acid sequence for the fiber protein of ChAd17
	SEQ ID NO: 30 -	Amino acid sequence for the fiber protein of ChAd19
	SEQ ID NO: 31 -	Amino acid sequence for the fiber protein of ChAd24
15	SEQ ID NO: 32 -	Amino acid sequence for the fiber protein of ChAd11
	SEQ ID NO: 33 -	Amino acid sequence for the fiber protein of ChAd20
	SEQ ID NO: 34 -	Amino acid sequence for the fiber protein of ChAd31
	SEQ ID NO: 35 -	Amino acid sequence for the fiber protein of PanAd1
	SEQ ID NO: 36 -	Amino acid sequence for the fiber protein of PanAd2
20	SEQ ID NO: 37 -	RSV FΔTM amino acid sequence
	SEQ ID NO: 38 -	HIV Gag polynucleotide sequence
	SEQ ID NO: 39 -	polypeptide sequence of Rv1174 (Mtb8.4)
	SEQ ID NO: 40 -	polypeptide sequence of Rv0287 (Mtb9.8)
	SEQ ID NO: 41 -	polypeptide sequence of Rv1793 (Mtb9.9)
25	SEQ ID NO: 42 -	polypeptide sequence of Rv0915 (Mtb41)
	SEQ ID NO: 43 -	polypeptide sequence of Rv3875 (ESAT-6)
	SEQ ID NO: 44 -	polypeptide sequence of Rv3804 (Ag85A)
	SEQ ID NO: 45 -	polypeptide sequence of Rv1886 (Ag85B)
	SEQ ID NO: 46 -	polypeptide sequence of Rv2031 (alpha-crystallin)
30	SEQ ID NO: 47 -	polypeptide sequence of Rv1980 (Mpt64)
	SEQ ID NO: 48 -	polypeptide sequence of Rv0288 (TB10.4)
	SEQ ID NO: 49 -	polypeptide sequence of Rv1753
	SEQ ID NO: 50 -	polypeptide sequence of Rv2386
	SEQ ID NO: 51 -	polypeptide sequence of Rv3616 (Mtb40)
35	SEQ ID NO: 52 -	polypeptide sequence of Rv3407
	SEQ ID NO: 53 -	polypeptide sequence of Rv2660
	SEQ ID NO: 54 -	polypeptide sequence of Rv2608
	SEQ ID NO: 55 -	polypeptide sequence of Rv3619
	SEQ ID NO: 56 -	polypeptide sequence of Rv3620
40	SEQ ID NO: 57 -	polypeptide sequence of Rv1813
	SEQ ID NO: 58 -	polypeptide sequence of Rv1009 (RpfB)
	SEQ ID NO: 59 -	polypeptide sequence of Rv2389 (RpfD)
	SEQ ID NO: 60 -	polypeptide sequence of Rv2626

- SEQ ID NO: 61 - polypeptide sequence of Rv1733
- SEQ ID NO: 62 - polypeptide sequence of Rv3136 (PPE51)
- SEQ ID NO: 63 - polypeptide sequence of Rv0475 (HBHA)
- SEQ ID NO: 64 - polypeptide sequence of Rv0125
- 5 SEQ ID NO: 65 - polypeptide sequence of Ser/Ala mutated mature Rv0125
- SEQ ID NO: 66 - polypeptide sequence of Ra12
- SEQ ID NO: 67 - polypeptide sequence of Ra35
- SEQ ID NO: 68 - polypeptide sequence of Rv1196
- SEQ ID NO: 69 - polypeptide sequence of Mtb72f
- 10 SEQ ID NO: 70 - polypeptide sequence of M72

DETAILED DESCRIPTION OF THE INVENTION

Tuberculosis

15 Tuberculosis (TB) is a chronic infectious disease caused by infection with *Mycobacterium tuberculosis* and other *Mycobacterium* species. It is a major disease in developing countries, as well as an increasing problem in developed areas of the world. About one third of the world's population are believed to be latently infected with TB bacilli, with about
20 9 million new cases of active TB and 1.5 million deaths each year. Around 10% of those infected with TB bacilli will develop active TB, each person with active TB infecting an average of 10 to 15 others per year. (World Health Organisation *Tuberculosis Facts* 2014).

25 *Mycobacterium tuberculosis* infects individuals through the respiratory route. Alveolar macrophages engulf the bacterium, but it is able to survive and proliferate by inhibiting phagosome fusion with acidic lysosomes. A complex immune response involving CD4+ and CD8+ T cells ensues, ultimately resulting in the formation of a granuloma. Central to the success of *Mycobacterium tuberculosis* as a pathogen is the fact that the isolated, but not eradicated, bacterium may persist for long periods, leaving an individual vulnerable to the later
30 development of active TB.

Fewer than 5% of infected individuals develop active TB in the first years after infection. The granuloma can persist for decades and is believed to contain live *Mycobacterium tuberculosis* in a state of dormancy, deprived of oxygen and nutrients. However, it has been suggested that
35 the majority of the bacteria in the dormancy state are located in non-macrophage cell types spread throughout the body (Locht et al, *Expert Opin. Biol. Ther.* 2007 7(11):1665-1677). The development of active TB occurs when the balance between the host's natural immunity and the pathogen changes, for example as a result of an immunosuppressive event (Anderson P *Trends in Microbiology* 2007 15(1):7-13; Ehlers S *Infection* 2009 37(2):87-95).

40 A dynamic hypothesis describing the balance between latent TB and active TB has also been proposed (Cardona P-J *Inflammation & Allergy – Drug Targets* 2006 6:27-39; Cardona P-J *Infection* 2009 37(2):80-86).

Although an infection may be asymptomatic for a considerable period of time, the active disease is most commonly manifested as an acute inflammation of the lungs, resulting in tiredness, weight loss, fever and a persistent cough. If untreated, serious complications and death typically result.

Tuberculosis can generally be controlled using extended antibiotic therapy, although such treatment is not sufficient to prevent the spread of the disease. Actively infected individuals may be largely asymptomatic, but contagious, for some time. In addition, although compliance with the treatment regimen (which typically lasts 6 months or more) is critical, patient behaviour is difficult to monitor. Some patients do not complete the course of treatment, which can lead to ineffective treatment and the development of drug resistance.

Multidrug-resistant TB (MDR-TB) is a form which fails to respond to first line medications. An estimated 480,000 people developed MDR-TB in 2013. MDR-TB is treatable by using second-line drugs. However, second-line treatment options are limited and recommended medicines are not always available. The extensive chemotherapy required (up to two years of treatment) is costly and can produce severe adverse drug reactions in patients.

Extensively drug-resistant TB (XDR-TB) occurs when resistance to second line medications develops on top of resistance to first line medications. It is estimated that about 9.0% of MDR-TB cases had XDR-TB (World Health Organisation *Tuberculosis Facts* 2014).

Even if a full course of antibiotic treatment is completed, infection with *M. tuberculosis* may not be eradicated from the infected individual and may remain as a latent infection that can be reactivated. Consequently, accurate and early diagnosis of the disease are of utmost importance.

Currently, vaccination with attenuated live bacteria is the most widely used method for inducing protective immunity. The most common *Mycobacterium* employed for this purpose is *Bacillus Calmette-Guerin* (BCG), an avirulent strain of *M. bovis* which was first developed over 60 years ago. It is administered at birth in TB endemic regions. However, the safety and efficacy of BCG is a source of controversy - while protecting against severe disease manifestation in children, the efficacy of BCG against disease in adults is variable. Additionally, some countries, such as the United States, do not vaccinate the general public with this agent.

Several of the proteins which are strongly expressed during the early stages of *Mycobacterium* infection have been shown to provide protective efficacy in animal vaccination models. However, vaccination with antigens which are highly expressed during the early stages of infection may not provide an optimal immune response for dealing with later stages of infection. Adequate control during latent infection may require T cells which are specific for the particular antigens which are expressed at that time. Post-exposure vaccines which directly target the dormant persistent bacteria may aid in protecting against TB reactivation, thereby

enhancing TB control, or even enabling clearance of the infection. A vaccine targeting latent TB could therefore significantly and economically reduce global TB infection rates.

5 Vaccines based on late stage antigens could also be utilised in combination with early stage antigens to provide a multiphase vaccine. Alternatively, early and/or late stage antigens could be used to complement and improve BCG vaccination.

10 Typically, the aim of the methods of the invention is to induce a protective immune response, i.e. immunise or vaccinate the subject against a related pathogen. The invention may therefore be applied for the prophylaxis, treatment or amelioration of infection by mycobacteria, such as infection by *Mycobacterium bovis* or *Mycobacterium tuberculosis*, in particular *Mycobacterium tuberculosis*.

The invention may be provided for the purpose of:

- 15 - prophylaxis of active tuberculosis due to infection (i.e. primary tuberculosis) or reactivation (i.e. secondary tuberculosis), such as by administering to a subject who is uninfected, or alternatively a subject who has a latent infection;
- prophylaxis of latent tuberculosis, such as by administering to a subject who is uninfected;
- 20 - treating latent tuberculosis;
- preventing or delaying reactivation of tuberculosis, especially the delay of reactivation, for example by a period of months, years or indefinitely; or
- treating active tuberculosis (such as to reduce the need for chemotherapeutic treatment: such as reduced term of chemotherapeutic treatment, complexity of drug
- 25 regimen or dosage of chemotherapeutic treatment; alternatively, to reduce the risk of a later relapse following chemotherapeutic treatment).

30 The elicited immune response may be an antigen specific T cell response (which may be a systemic and/or a local response). Systemic responses may be detected, for example, from a sample of whole blood. Local responses (for example, the local response in the lung) may be detected from an appropriate sample of tissue (for example, lung tissue) or other locally focused sample method (e.g. bronchoalveolar lavage). The antigen specific T cell response may comprise a CD4+ T cell response, such as a response involving CD4+ T cells expressing a plurality of cytokines (e.g. IFNgamma, TNFalpha or IL2, especially IFNgamma, TNFalpha and IL2).

35 Alternatively, or additionally, the antigen specific T cell response comprises a CD8+ T cell response, such as a response involving CD8+ T cells expressing a plurality of cytokines (e.g. IFNgamma, TNFalpha or IL2, especially IFNgamma, TNFalpha and IL2).

40 The term “active infection” refers to an infection, e.g. infection by *M. tuberculosis*, with manifested disease symptoms and/or lesions, suitably with manifested disease symptoms.

The terms “inactive infection”, “dormant infection” or “latent infection” or “latent tuberculosis” refer to an infection, e.g. infection by *M. tuberculosis*, without manifested disease symptoms

and/or lesions, suitably without manifested disease symptoms. A subject with latent infection will suitably be one which tests positive for infection, e.g. by Tuberculin skin test (TST) or Interferon–Gamma Release Assays (IGRAs), but which has not demonstrated the disease symptoms and/or lesions which are associated with an active infection.

5

The term “primary tuberculosis” refers to clinical illness, e.g., manifestation of disease symptoms, directly following infection, e.g. infection by *M. tuberculosis*. See, *Harrison’s Principles of Internal Medicine*, Chapter 150, pp. 953-966 (16th ed., Braunwald, *et al.*, eds., 2005).

10

The terms “secondary tuberculosis” or “postprimary tuberculosis” refer to the reactivation of a dormant, inactive or latent infection, e.g. infection by *M. tuberculosis*. See, *Harrison’s Principles of Internal Medicine*, Chapter 150, pp. 953-966 (16th ed., Braunwald, *et al.*, eds., 2005).

15

The term “tuberculosis reactivation” refers to the later manifestation of disease symptoms in an individual that tests positive for infection (e.g. by Tuberculin skin test (TST) or Interferon–Gamma Release Assays (IGRAs)) but does not have apparent disease symptoms. Suitably the individual will not have been re-exposed to infection. The positive diagnostic test indicates that the individual is infected, however, the individual may or may not have previously manifested active disease symptoms that had been treated sufficiently to bring the tuberculosis into an inactive or latent state.

20

Suitability the methods are applied to a subject who is uninfected or who has a latent infection by mycobacteria, such as infection by *Mycobacterium tuberculosis*. In one embodiment the methods are applied to a subject who does not have an infection by *Mycobacterium tuberculosis* (in the context of human subjects) or *Mycobacterium bovis* (in the context of bovine subjects). In another embodiment the methods are applied to a subject who has a latent infection by mycobacteria, such as *Mycobacterium tuberculosis* (in the context of human subjects) or *Mycobacterium bovis* (in the context of bovine subjects).

25

30

In some embodiments, the subject has previously been vaccinated with BCG. The approaches of the present invention may, for example, be utilised for a subject at least one year after BCG vaccination, for example at least two years after BCG vaccination such as at least at least five years after BCG vaccination.

35

In some embodiments, the subject has previously been infected with *M. tuberculosis*.

Antigens of use in the invention

40

The mycobacterial antigen is an antigenic sequence (i.e. a sequence from a mycobacterial protein which comprises at least one B or T cell epitope). Suitably the mycobacterial antigen comprises at least one T cell epitope.

Mycobacterial antigens of particular interest in the present invention are those derived from:

(i) Mtb8.4 (also known as DPV and Rv1174), the polypeptide sequence of which is described in SEQ ID NO: 102 of WO97/09428 (cDNA in SEQ ID NO: 101) and in Coler et al *Journal of Immunology* 1998 161:2356-2364. Of particular interest is the mature Mtb8.4 sequence which is absent the leading signal peptide (i.e. amino acid residues 15-96 from SEQ ID NO: 102 of WO97/09428). The full-length polypeptide sequence of Rv1174 is shown in SEQ ID NO: 39;

(ii) Mtb9.8 (also known as MSL and Rv0287), the polypeptide sequence of which is described in SEQ ID NO: 109 of WO98/53075 (fragments of MSL are disclosed in SEQ ID NOs: 110-124 of WO98/53075, SEQ ID NOs: 119 and 120 being of particular interest) and also in Coler et al *Vaccine* 2009 27:223-233 (in particular the reactive fragments shown in Figure 2 therein). The full-length polypeptide sequence for Rv0287 is shown in SEQ ID NO: 41;

(iii) Mtb9.9 (also known as Mtb9.9A, MTI, MTI-A and Rv1793) the polypeptide sequence of which is described in SEQ ID NO: 19 of WO98/53075 and in Alderson et al *Journal of Experimental Medicine* 2000 7:551-559 (fragments of MTI are disclosed in SEQ ID NOs: 17 and 51-66 of WO98/53075, SEQ ID NOs: 17, 51, 52, 53, 56 and 62-65 being of particular interest). A number of polypeptide variants of MTI are described in SEQ ID NOs: 21, 23, 25, 27, 29 and 31 of WO98/53075 and in Alderson et al *Journal of Experimental Medicine* 2000 7:551-559. The full-length polypeptide sequence for Rv1793 is shown in SEQ ID NO: 41;

(iv) Mtb41 (also known as MTCC2 and Rv0915) the polypeptide sequence of which is described in SEQ ID NO: 142 of WO98/53075 (cDNA in SEQ ID NO: 140) and in Skeiky et al *Journal of Immunology* 2000 165:7140-7149. The full-length polypeptide sequence for Rv0915 is shown in SEQ ID NO: 42;

(v) ESAT-6 (also known as esxA and Rv3875) the polypeptide sequence of which is described in SEQ ID NO: 103 of WO97/09428 (cDNA in SEQ ID NO: 104) and in Sorensen et al *Infection and Immunity* 1995 63(5):1710-1717. The full-length polypeptide sequence for Rv3875 is shown in SEQ ID NO: 43;

(vi) Ag85 complex antigens (e.g. Ag85A, also known as fbpA and Rv3804; or Ag85B, also known as fbpB and Rv1886) which are discussed, for example, in Content et al *Infection and Immunity* 1991 59:3205-3212 and in Huygen et al *Nature Medicine* 1996 2(8):893-898. The full-length polypeptide sequence for Rv3804/Ag85A is shown in SEQ ID NO: 44 (the mature protein of residues 43-338, i.e. lacking the signal peptide, being of particular interest). The full-length polypeptide sequence for Ag85B is shown in

SEQ ID NO: 45 (the mature protein of residues 41-325, i.e. lacking the signal peptide, being of particular interest);

5 (vii) Alpha-crystallin (also known as hspX and Rv2031) which is described in Verbon et al *Journal of Bacteriology* 1992 174:1352-1359 and Friscia et al *Clinical and Experimental Immunology* 1995 102:53-57 (of particular interest are the fragments corresponding to residues 71-91, 21-40, 91-110 and 111-130). The full-length polypeptide sequence for Rv2031 is shown in SEQ ID NO: 46;

10 (viii) Mpt64 (also known as Rv1980) which is described in Roche et al *Scandinavian Journal of Immunology* 1996 43:662-670. The full-length polypeptide sequence for Mpt64 is shown in SEQ ID NO: 47 (the mature protein of residues 24-228, i.e. lacking the signal peptide, being of particular interest);

15 (ix) TB10.4 (also known as cfp7 and Rv0288), described for example in Skjot et al. *Infect Immun* 2002 70: 5446-5453, Dietrich et al. *J Immunol* 2005 174:6332-6339 and Elvang et al. *PLoS One* 2009;4(4):e5139. The full-length polypeptide sequence for Rv0288 is shown in SEQ ID NO: 48;

20 (x) RV1753, such as described in Seq ID NOs: 1 and 2-7 of WO2010010180. The full-length polypeptide sequence for Rv1753 from *Mycobacterium tuberculosis* strain C is shown in SEQ ID NO: 49;

25 (xi) Rv2386c, Seq ID NOs: 1 and 2-7 of WO2010010179. The full-length polypeptide sequence for Rv2386c from *Mycobacterium tuberculosis* H37Rv is shown in SEQ ID NO: 50; and/or

30 (xii) Mtb40 (also known as HTCC1 and Rv3616) such as described in WO2011092253, for example a natural Rv3616 sequence selected from Seq ID NOs: 1 and 2-7 of WO2011092253 or a modified Rv3616 sequence such as those selected from Seq ID NOs: 161 to 169, 179 and 180 of WO2011092253. Rv3616 fragments selected from SEQ ID NO: 127, 128, 130, 131-133, 135, 143-148 or 150-156 of WO2011092253 are of particular interest. The full-length polypeptide sequence for Rv3616 is shown in SEQ ID NO: 51;

35 (xiii) Rv3407, the full-length polypeptide sequence for Rv3407 is shown in SEQ ID NO: 52;

40 (xiv) Rv2660, the full-length polypeptide sequence for Rv2660 is shown in SEQ ID NO: 53;

(xv) Rv2608, the full-length polypeptide sequence for Rv2608 is shown in SEQ ID NO: 54;

(xvi) Rv3619, the full-length polypeptide sequence for Rv3619 is shown in SEQ ID NO: 55;

5 (xvii) Rv3620, the full-length polypeptide sequence for Rv3620 is shown in SEQ ID NO: 56;

(xviii) Rv1813, the full-length polypeptide sequence for Rv1813 is shown in SEQ ID NO: 57;

10 (xix) Rv1009, also known as RpfB, the full-length polypeptide sequence for Rv1009 is shown in SEQ ID NO: 58;

15 (xx) Rv2389, also known as RpfD, the full-length polypeptide sequence for Rv2389 is shown in SEQ ID NO: 59;

(xxi) Rv2626, the full-length polypeptide sequence for Rv2626 is shown in SEQ ID NO: 60;

20 (xxii) Rv1733, the full-length polypeptide sequence for Rv1733 is shown in SEQ ID NO: 61;

(xxiii) Rv3136, also known as PPE51, the full-length polypeptide sequence for Rv3136 is shown in SEQ ID NO: 62;

25 (xxiv) Rv0475, also known as HBHA and described in WO97044463, WO03044048 and WO2010149657, the full-length polypeptide sequence for Rv0475 is shown in SEQ ID NO: 63;

30 (xxv) Mtb32A (also known as Rv0125), the polypeptide sequence of which is described in SEQ ID NO: 2 (full-length) and residues 8-330 of SEQ ID NO: 4 (mature) of WO01/98460, especially variants having at least one of the catalytic triad mutated (e.g. the catalytic serine residue, which may for example be mutated to alanine). The mature polypeptide sequence for Rv0125 is shown in SEQ ID NO: 64. The mature form of Mtb32A having a Ser/Ala mutation is shown in SEQ ID NO: 65. Fragments of Rv0125 of particular interest include Ra12 (also known as Mtb32A C-terminal antigen) the polypeptide sequence of which is described in SEQ ID NO: 10 of WO01/98460 and in Skeiky et al *Journal of Immunology* 2004 172:7618-7682. The full-length polypeptide sequence for Ra12 is shown in SEQ ID NO: 66. Another fragment of Rv0125 of particular interest is Ra35 (also known as Mtb32A N-terminal antigen) the polypeptide sequence of which is described in SEQ ID NO: 8 of WO01/98460 and in Skeiky et al *Journal of Immunology* 2004 172:7618-7682. The full-length polypeptide sequence for Ra35 is shown in SEQ ID NO: 67; and

35

40

(xxvi) TbH9 (also known as Mtb39, Mtb39A, TbH9FL and Rv1196) the polypeptide sequence of which is described in SEQ ID NO: 107 of WO97/09428, and also in Dillon et al *Infection and Immunity* 1999 67(6):2941-2950 and Skeiky et al *Journal of Immunology* 2004 172:7618-7682. The full-length polypeptide sequence for Rv1196 is shown in SEQ ID NO: 68.

Combinations of mycobacterial antigens may also be utilised. In such cases the mycobacterial antigens may be encoded individually or as part of one or more fusion proteins. "Fusion polypeptide" or "fusion protein" refers to a protein having at least two heterologous polypeptides (e.g. at least two *Mycobacterium* sp. polypeptides) covalently linked, either directly or via an amino acid linker. The polypeptides of the fusion protein can be in any order.

Combinations of antigens which may be encoded (suitably in the form of a single fusion protein) include:

- (a) an Ag85B sequence and an ESAT-6 sequence (such as the H1 fusion protein antigen);
- (b) an Ag85B sequence and a TB10.4 sequence (such as the H4 fusion protein antigen);
- (c) an Ag85B sequence, an ESAT-6 sequence and an Rv2660 sequence (such as the H56 fusion protein antigen);
- (d) an Rv2608 sequence, an Rv3619 sequence, an Rv3620 sequence and an Rv1813 sequence (such as the ID93 fusion protein antigen);
- (e) an Ag85A sequence, an Ag85B sequence and an Rv3407 sequence;
- (f) an Rv1733 sequence and an Rv2626 sequence; such as:
 - (f-i) an Rv1733 sequence, an Rv2626 sequence and an ESAT-6 sequence;
 - (f-i-a) an Rv1733 sequence, an Rv2626 sequence, an ESAT-6 sequence and an RpfD sequence; such as:
 - (f-i-a-i) an Rv1733 sequence, an Rv2626 sequence, an ESAT-6 sequence, an RpfD sequence and an Ag85B sequence;
 - (f-i-b) an Rv1733 sequence, an Rv2626 sequence, an ESAT-6 sequence and an RpfB sequence; such as:
 - (f-i-b-i) an Rv1733 sequence, an Rv2626 sequence, an ESAT-6 sequence, an RpfB sequence and an Ag85B sequence;
 - (f-ii) an Rv1733 sequence, an Rv2626 sequence and an Rv3407 sequence; such as:
 - (f-ii-a) an Rv1733 sequence, an Rv2626 sequence, an Rv3407 sequence and an RpfD sequence;
 - (f-ii-b) an Rv1733 sequence, an Rv2626 sequence, an Rv3407 sequence and an RpfB sequence;
 - (f-iii) an Rv1733 sequence, an Rv2626 sequence and a PPE51 sequence; such as:
 - (f-iii-a) an Rv1733 sequence, an Rv2626 sequence, a PPE51 sequence and an RpfD sequence;
 - (f-iii-b) an Rv1733 sequence, an Rv2626 sequence, a PPE51 sequence and an RpfB sequence.

The skilled person will recognise that combinations need not rely upon the specific sequences described above in (i)-(xxvi), and that variants having at least 80% identity, such as at least 90% identity, in particular at least 95% identity and especially at least 98% identity) or immunogenic fragments (e.g. at least 20% of the full length antigen, such as at least 40% of the antigen, in particular at least 50% and especially at least 75%) of the described sequences can alternatively be used.

The adenovirus may therefore encode any of combinations (a) to (f), such as in a single fusion protein, wherein each of the components has at least 80% identity, such as at least 90% identity, in particular at least 95% identity and especially at least 98% identity) or is an immunogenic fragments (e.g. at least 20% of the full length antigen, such as at least 40% of the antigen, in particular at least 50% and especially at least 75%) of the described sequences (which may be in any order).

Particularly suitable combinations are those comprising an Rv1196 sequence and an Rv0125 sequence, such as the Mtb72f fusion protein antigen or M72 fusion protein antigen.

Rv1196 (described, for example, by the name Mtb39a in Dillon et al *Infection and Immunity* 1999 67(6): 2941-2950) is highly conserved, with 100% sequence identity across H37Rv, C, Haarlem, CDC1551, 94-M4241A, 98-R604INH-RIF-EM, KZN605, KZN1435, KZN4207, KZNR506 strains, the F11 strain having a single point mutation Q30K (most other clinical isolates have in excess of 90% identity to H37Rv). An adenovirus encoding an Rv1196 related antigen is described in Lewinsohn et al *Am J Respir Crit Care Med* 2002 116:843-848.

Rv0125 (described, for example, by the name Mtb32a in Skeiky et al *Infection and Immunity* 1999 67(8): 3998-4007) is also highly conserved, with 100% sequence identity across many strains. An adenovirus (human Ad5) encoding an Rv0125 related antigen is described in Zhang et al *Human Vaccines & Therapeutics* 2015 11(7):1803-1813 doi: 10.1080/21645515.2015.1042193. Full length Rv0125 includes an N-terminal signal sequence which is cleaved to provide the mature protein.

Mtb72f has been shown to provide protection in a number of animal models (see, for example: Brandt et al *Infect. Immun.* 2004 72(11):6622-6632; Skeiky et al *J. Immunol.* 2004 172:7618-7628; Tsenova et al *Infect. Immun.* 2006 74(4):2392-2401). Mtb72f has also been the subject of clinical investigations (Von Eschen et al 2009 *Human Vaccines* 5(7):475-482). M72 is an improved antigen which incorporates a single serine to alanine mutation relative to Mtb72f, resulting in improved stability characteristics. M72 related antigens have also been shown to be of value in a latent TB model (international patent application WO2006/117240, incorporated herein by reference). Previous pre-clinical and clinical investigations have led to M72 being administered in humans in conjunction with the immunostimulants 3-O-deacylated monophosphoryl lipid A (3D-MPL) and QS21 in a liposomal formulation and in a 0,1 month schedule using 10 ug M72 polypeptide, 25 ug 3D-MPL and 25 ug QS21 (see, for example, Leroux-Roels et al *Vaccine* 2013 31 2196-2206, Montoya et al *J. Clin. Immunol.* 2013 33(8):

1360-1375; Thacher EG et al *AIDS* 2014 28(12):1769-1781; Idoko OT et al *Tuberculosis (Edinb)* 2014 94(6):564-578; Penn-Nicholson A, et al *Vaccine* 2015 33(32):4025-4034 doi:10.1016/j.vaccine.2015.05.088).

- 5 In an embodiment of the invention the adenovirus comprises a nucleic acid sequence encoding a mycobacterial antigen derived from at least one of Rv0125, Rv0287, Rv0288, Rv0475, Rv0915, Rv1009, Rv1174, Rv1196, Rv1733, Rv1753, Rv1793, Rv1813, Rv1886, Rv1980, Rv2031, Rv2386, Rv2389, Rv2608, Rv2626, Rv2660, Rv3136, Rv3407, Rv3616, Rv3619, Rv3620, Rv3804 and Rv3875, in particular at least one of Rv0125, Rv0287, Rv0915,
10 Rv1174, Rv1196, Rv1753, Rv1793, Rv2386 and Rv3616, especially at least one of Rv0125 and Rv1196.

By the term RvNNNN, means the protein encoded by the gene number NNNN identified from the H37Rv strain of *Mycobacterium tuberculosis* or a homologue thereof from another mycobacterium, such as *Mycobacterium bovis*, or in particular from another strain of *Mycobacterium tuberculosis*. Sequences for proteins from H37Rv are known in the art and
15 may be obtained, for example, from Tuberculist (<http://tuberculist.epfl.ch/>).

The adenovirus may comprise a nucleic acid sequence encoding a mycobacterial antigen which comprises at least one of Rv0125, Rv0287, Rv0288, Rv0475, Rv0915, Rv1009, Rv1174, Rv1196, Rv1733, Rv1753, Rv1793, Rv1813, Rv1886, Rv1980, Rv2031, Rv2386, Rv2389, Rv2608, Rv2626, Rv2660, Rv3136, Rv3407, Rv3616, Rv3619, Rv3620, Rv3804 and Rv3875, in particular at least one of Rv0125, Rv0287, Rv0915, Rv1174, Rv1196, Rv1753, Rv1793, Rv2386 and Rv3616, especially at least one of Rv0125 and Rv1196.
20

25 The adenovirus may comprise a nucleic acid sequence encoding a mycobacterial antigen which comprises a sequence which has at least 80% identity to Rv0125, Rv0287, Rv0288, Rv0475, Rv0915, Rv1009, Rv1174, Rv1196, Rv1733, Rv1753, Rv1793, Rv1813, Rv1886, Rv1980, Rv2031, Rv2386, Rv2389, Rv2608, Rv2626, Rv2660, Rv3136, Rv3407, Rv3616, Rv3619, Rv3620, Rv3804 or Rv3875, in particular Rv0125, Rv0287, Rv0915, Rv1174, Rv1196, Rv1753, Rv1793, Rv2386 or Rv3616, especially at least one of Rv0125 or Rv1196. Suitably the mycobacterial antigen which comprises a sequence which has at least 90% identity to the reference sequence, in particular at least 95%, such as at least 98%.

35 The adenovirus may comprise a nucleic acid sequence encoding a mycobacterial antigen which comprises a sequence which is an immunogenic fragment (such as comprising one or more T cell epitopes) of Rv0125, Rv0287, Rv0288, Rv0475, Rv0915, Rv1009, Rv1174, Rv1196, Rv1733, Rv1753, Rv1793, Rv1813, Rv1886, Rv1980, Rv2031, Rv2386, Rv2389, Rv2608, Rv2626, Rv2660, Rv3136, Rv3407, Rv3616, Rv3619, Rv3620, Rv3804 or Rv3875, in particular Rv0125, Rv0287, Rv0915, Rv1174, Rv1196, Rv1753, Rv1793, Rv2386 or Rv3616, especially at least one of Rv0125 or Rv1196. Suitably the mycobacterial antigen which comprises a sequence which is an immunogenic fragment of at least 20 amino acid residues, such as at least 50 amino acid residues from the reference sequence. Alternatively, the
40

mycobacterial antigen which comprises a sequence which is an immunogenic fragment of at least 20% of the total length of the reference sequence, such as at least 30% of the total length of the reference sequence residues, in particular at least 40% of the total length of the reference sequence residues, especially at least 50% of the total length of the reference sequence residues such as at least 75% of the total length of the reference sequence.

In an embodiment of the invention the adenovirus comprises a nucleic acid sequence encoding a mycobacterial antigen derived from at least one of SEQ ID NOs: 39 to 68, in particular SEQ ID NOs: 65 to 68 and especially SEQ ID NO: 70.

Each of the above individual antigen sequences is also disclosed in Cole et al *Nature* 1998 393:537-544 and Camus *Microbiology* 2002 148:2967-2973. The genome of *M. tuberculosis* H37Rv is publicly available, for example at the Wellcome Trust Sanger Institute website (www.sanger.ac.uk/Projects/M_tuberculosis/) and elsewhere.

Suitably the mycobacterial antigen contains 2500 amino acid residues or fewer, such 1500 amino acid residues or fewer, in particular 1200 amino acid residues or fewer, especially 1000 amino acid residues or fewer, typically 800 amino acid residues or fewer.

T cell epitopes are short contiguous stretches of amino acids which are recognised by T cells (e.g. CD4+ or CD8+ T cells). Identification of T cell epitopes may be achieved through epitope mapping experiments which are known to the person skilled in the art (see, for example, Paul, *Fundamental Immunology*, 3rd ed., 243-247 (1993); Beißbarth et al *Bioinformatics* 2005 21(Suppl. 1):i29-i37). In a diverse out-bred population, such as humans, different HLA types mean that particular epitopes may not be recognised by all members of the population. As a result of the crucial involvement of the T cell response in tuberculosis, to maximise the level of recognition and scale of immune response, an immunogenic derivative of a reference sequence is desirably one which contains the majority (or suitably all) T cell epitopes intact. Mortier et al *BMC Immunology* 2015 16:63 undertake sequence conservation analysis and *in silico* human leukocyte antigen-peptide binding predictions for Mtb72f and M72 tuberculosis candidate vaccine antigens.

The skilled person will recognise that individual substitutions, deletions or additions to a protein which alters, adds or deletes a single amino acid or a small percentage of amino acids is an “immunogenic derivative” where the alteration(s) results in the substitution of an amino acid with a functionally similar amino acid or the substitution/deletion/addition of residues which do not substantially impact the immunogenic function.

Conservative substitution tables providing functionally similar amino acids are well known in the art. In general, such conservative substitutions will fall within one of the amino-acid groupings specified below, though in some circumstances other substitutions may be possible without substantially affecting the immunogenic properties of the antigen. The following eight groups each contain amino acids that are typically conservative substitutions for one another:

- 1) Alanine (A), Glycine (G);
 - 2) Aspartic acid (D), Glutamic acid (E);
 - 3) Asparagine (N), Glutamine (Q);
 - 4) Arginine (R), Lysine (K);
 - 5 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V);
 - 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W);
 - 7) Serine (S), Threonine (T); and
 - 8) Cysteine (C), Methionine (M)
- (see, e.g., Creighton, *Proteins* 1984).

10

Suitably such substitutions do not occur in the region of an epitope, and do not therefore have a significant impact on the immunogenic properties of the antigen.

15

Immunogenic derivatives may also include those wherein additional amino acids are inserted compared to the reference sequence. Suitably such insertions do not occur in the region of an epitope, and do not therefore have a significant impact on the immunogenic properties of the antigen. One example of insertions includes a short stretch of histidine residues (e.g. 2-6 residues) to aid expression and/or purification of the antigen in question.

20

Immunogenic derivatives include those wherein amino acids have been deleted compared to the reference sequence. Suitably such deletions do not occur in the region of an epitope, and do not therefore have a significant impact on the immunogenic properties of the antigen. The skilled person will recognise that a particular immunogenic derivative may comprise substitutions, deletions and additions (or any combination thereof).

25

30

The terms "identical" or percentage "identity," in the context of two or more antigen sequences, refer to two or more sequences or sub-sequences that are the same or have a specified percentage of amino acid residues that are the same, when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. This definition also refers to the complement of a test sequence. Optionally, the identity exists over a region that is at least 200 amino acids in length, such as at least 300 amino acids or at least 400 amino acids. Suitably, the comparison is performed over a window corresponding to the entire length of the reference sequence (as opposed to the derivative sequence).

35

40

For sequence comparison, one sequence acts as the reference sequence, to which the test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percentage sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A “comparison window”, as used herein, refers to a segment in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerised implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (see, e.g., *Current Protocols in Molecular Biology* (Ausubel *et al.*, eds. 1995 supplement)).

One example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments to show relationship and percent sequence identity. It also plots a tree or dendrogram showing the clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng & Doolittle, *J. Mol. Evol.* 35:351-360 (1987). The method used is similar to the method described by Higgins & Sharp, *CABIOS* 5:151-153 (1989). The program can align up to 300 sequences, each of a maximum length of 5,000 nucleotides or amino acids. The multiple alignment procedure begins with the pairwise alignment of the two most similar sequences, producing a cluster of two aligned sequences. This cluster is then aligned to the next most related sequence or cluster of aligned sequences. Two clusters of sequences are aligned by a simple extension of the pairwise alignment of two individual sequences. The final alignment is achieved by a series of progressive, pairwise alignments. The program is run by designating specific sequences and their amino acid coordinates for regions of sequence comparison and by designating the program parameters. Using PILEUP, a reference sequence is compared to other test sequences to determine the percent sequence identity relationship using the following parameters: default gap weight (3.00), default gap length weight (0.10), and weighted end gaps. PILEUP can be obtained from the GCG sequence analysis software package, e.g., version 7.0 (Devereaux *et al.*, *Nuc. Acids Res.* 12:387-395 (1984)).

Another example of algorithm that is suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.*, *Nuc. Acids Res.* 25:3389-3402 (1997) and Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990), respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (website at www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighbourhood word score threshold (Altschul *et al.*, *supra*). These initial neighbourhood word hits act as seeds for initiating searches to find longer HSPs containing

them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001.

In any event, immunogenic derivatives of a polypeptide sequence will usually have essentially the same activity as the reference sequence. By essentially the same activity is meant at least 50%, suitably at least 75% and especially at least 90% activity of the reference sequence in an *in vitro* restimulation assay of PBMC, whole blood, lung tissue or bronchoalveolar lavage with specific antigens (e.g. restimulation for a period of between several hours to up to two weeks, such as up to one day, 1 day to 1 week or 1 to 2 weeks) that measures the activation of the cells via lymphoproliferation, production of cytokines in the supernatant of culture (measured by ELISA, CBA etc) or characterisation of T and B cell responses by intra and extracellular staining (e.g. using antibodies specific to immune markers, such as CD3, CD4, CD8, IL2, TNF-alpha, IFN-gamma, IL-17, CD40L, CD69 etc) followed by analysis with a flow cytometer. Suitably, by essentially the same activity is meant at least 50%, suitably at least 75% and especially at least 90% activity of the reference sequence in a T cell proliferation and/or IFN-gamma production assay.

In one embodiment the encoded antigen is a Rv1196 related antigen. The term 'Rv1196 related antigen' refers to the Rv1196 protein provided in SEQ ID NO: 68 or an immunogenic derivative thereof. As used herein the term "derivative" refers to an antigen that is modified relative to the reference sequence. Immunogenic derivatives are sufficiently similar to the

reference sequence to substantially retain the immunogenic properties of the reference sequence and remain capable of allowing an immune response to be raised against the reference sequence. An immunogenic derivative may, for example, comprise a modified version of the reference sequence or alternatively may consist of a modified version of the reference sequence.

The Rv1196 related antigen may for example contain 2500 amino acid residues or fewer, such as 1500 amino acid residues or fewer, in particular 1200 amino acid residues or fewer, especially 1000 amino acid residues or fewer, typically 800 amino acid residues or fewer.

Suitably the Rv1196 related antigen will comprise, such as consist of, a sequence having at least 70% identity to SEQ ID NO: 68, such as at least 80%, in particular at least 90%, especially at least 95%, for example at least 98%, such as at least 99%.

A specific example of an Rv1196 related antigen is Rv1196 from *Mycobacterium tuberculosis* strain H37Rv, as provided in SEQ ID NO: 68. Consequently, in one embodiment of the invention the Rv1196 related antigen is a protein comprising SEQ ID NO: 68. In a second embodiment of the invention the Rv1196 related antigen is a protein consisting of SEQ ID NO: 68.

Typical Rv1196 related antigens will comprise (such as consist of) an immunogenic derivative of SEQ ID NO: 68 having a small number of deletions, insertions and/or substitutions. Examples are those having deletions of up to 5 residues at 0-5 locations, insertions of up to 5 residues at 0-5 five locations and substitution of up to 20 residues.

Other immunogenic derivatives of Rv1196 are those comprising (such as consisting of) a fragment of SEQ ID NO: 68 which is at least 200 amino acids in length, such as at least 250 amino acids in length, in particular at least 300 amino acids in length, especially at least 350 amino acids in length.

In one embodiment the polypeptide antigen and the encoded antigen are Rv0125 related antigens. The term 'Rv0125 related antigen' refers to the Rv0125 protein provided in SEQ ID NO: 64 or an immunogenic derivative thereof. As used herein the term "derivative" refers to an antigen that is modified relative to the reference sequence. Immunogenic derivatives are sufficiently similar to the reference sequence to substantially retain the immunogenic properties of the reference sequence and remain capable of allowing an immune response to be raised against the reference sequence. An immunogenic derivative may, for example, comprise a modified version of the reference sequence or alternatively may consist of a modified version of the reference sequence.

The Rv0125 related antigen may for example contain 2500 amino acid residues or fewer, such as 1500 amino acid residues or fewer, in particular 1200 amino acid residues or fewer, especially 1000 amino acid residues or fewer, typically 800 amino acid residues or fewer.

Suitably the Rv0125 related antigen will comprise, such as consist of, a sequence having at least 70% identity to SEQ ID NO: 64, such as at least 80%, in particular at least 90%, especially at least 95%, for example at least 98%, such as at least 99%.

5

A specific example of an Rv0125 related antigen is Rv0125 from *Mycobacterium tuberculosis* strain H37Rv, as provided in SEQ ID NO: 64. Consequently, in one embodiment of the invention the Rv0125 related antigen is a protein comprising SEQ ID NO: 64. In a second embodiment of the invention the Rv0125 related antigen is a protein consisting of SEQ ID NO:

10

64.

Typical Rv0125 related antigens will comprise (such as consist of) an immunogenic derivative of SEQ ID NO: 64 having a small number of deletions, insertions and/or substitutions. Examples are those having deletions of up to 5 residues at 0-5 locations, insertions of up to 5 residues at 0-5 five locations and substitution of up to 20 residues.

15

Other immunogenic derivatives of Rv0125 are those comprising (such as consisting of) a fragment of SEQ ID NO: 64 which is at least 150 amino acids in length, such as at least 200 amino acids in length, in particular at least 250 amino acids in length, especially at least 300 amino acids in length. Particular immunogenic derivatives of Rv0125 are those comprising (such as consisting of) the fragment of SEQ ID NO: 64 corresponding to residues 1-195 of SEQ ID NO: 3. Further immunogenic derivatives of Rv1196 are those comprising (such as consisting of) the fragment of SEQ ID NO: 64 corresponding to residues 192-323 of SEQ ID NO: 64.

20

25

Particularly preferred Rv0125 related antigens are derivatives of SEQ ID NO: 64 wherein at least one (for example one, two or even all three) of the catalytic triad have been substituted or deleted, such that the protease activity has been reduced and the protein more easily produced – the catalytic serine residue may be deleted or substituted (e.g. substituted with alanine) and/or the catalytic histidine residue may be deleted or substituted and/or substituted the catalytic aspartic acid residue may be deleted or substituted. Especially of interest are derivatives of SEQ ID NO: 64 wherein the catalytic serine residue has been substituted (e.g. substituted with alanine). Also of interest are Rv0125 related antigens which comprise, such as consist of, a sequence having at least 70% identity to SEQ ID NO: 64, such as at least 80%, in particular at least 90%, especially at least 95%, for example at least 98%, such as at least 99% and wherein at least one of the catalytic triad have been substituted or deleted or those comprising, such as consisting of, a fragment of SEQ ID NO: 64 which is at least 150 amino acids in length, such as at least 200 amino acids in length, in particular at least 250 amino acids in length, especially at least 300 amino acids in length and wherein at least one of the catalytic triad have been substituted or deleted. Further immunogenic derivatives of Rv0125 are those comprising (such as consisting of) the fragment of SEQ ID NO: 64 corresponding to residues 192-323 of SEQ ID NO: 64 wherein at least one (for example one, two or even all three) of the catalytic triad have been substituted or deleted. Particular

30

35

40

immunogenic derivatives of Rv1196 are those comprising (such as consisting of) the fragment of SEQ ID NO: 64 corresponding to residues 1-195 of SEQ ID NO: 64 wherein the catalytic serine residue (position 176 of SEQ ID NO: 64) has been substituted (e.g. substituted with alanine).

5

In certain embodiments the mycobacterial antigen is an Rv1196 and Rv0125 related antigen, such as M72 related antigens. Particular derivatives of the M72 protein include those with additional His residues at the N-terminus (e.g. two His residues; or a polyhistidine tag of five or particularly six His residues, which may be used for nickel affinity purification). Mtb72f which contains the original serine residue that has been mutated in M72, is a further derivative of M72, as are Mtb72f proteins with additional His residues at the N-terminus (e.g. two His residues; or a polyhistidine tag of five or particularly six His residues, which may be used for nickel affinity purification).

10

15

In some embodiments a single adenovirus may encode two distinct polypeptides, one being a Rv1196 related antigen and one being a Rv0125 related antigen.

20

Suitably an M72 related antigen will comprise, such as consist of, a sequence having at least 70% identity to SEQ ID NO. 70, such as at least 80%, in particular at least 90%, especially at least 95%, such as at least 98%, for example at least 99%.

25

Typical M72 related antigens will comprise, such as consist of, a derivative of SEQ ID NO: 70 having a small number of deletions, insertions and/or substitutions. Examples are those having deletions of up to 5 residues at 0-5 locations, insertions of up to 5 residues at 0-5 five locations and substitution of up to 20 residues.

30

Other derivatives of M72 are those comprising, such as consisting of, a fragment of SEQ ID NO: 70 which is at least 450 amino acids in length, such as at least 500 amino acids in length, such as at least 550 amino acids in length, such as at least 600 amino acids in length, such as at least 650 amino acids in length or at least 700 amino acids in length. As M72 is a fusion protein derived from two individual antigens, any fragment of at least 450 residues will comprise a plurality of epitopes from the full length sequence (Skeiky et al *J. Immunol.* 2004 172:7618-7628; Skeiky *Infect. Immun.* 1999 67(8):3998-4007; Dillon *Infect. Immun.* 1999 67(6):2941-2950).

35

In particular embodiments the M72 related antigen will comprise residues 2-723 of SEQ ID NO. 70, for example comprise (or consist of) SEQ ID NO. 70.

Adenovirus

40

Adenoviruses have a characteristic morphology with an icosahedral capsid comprising three major proteins, hexon (II), penton base (III) and a knobbed fiber (IV), along with a number of other minor proteins, VI, VIII, IX, IIIa and IVa2. The virus genome is a linear, double-stranded

DNA. The virus DNA is intimately associated with the highly basic protein VII and a small peptide pX (formerly termed mu). Another protein, V, is packaged with this DNA-protein complex and provides a structural link to the capsid via protein VI. The virus also contains a virus-encoded protease, which is necessary for processing of some of the structural proteins to produce mature infectious virus.

The adenoviral genome is well characterized. There is general conservation in the overall organization of the adenoviral genome with respect to specific open reading frames being similarly positioned, e.g. the location of the E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4 and L5 genes of each virus. Each extremity of the adenoviral genome comprises a sequence known as an inverted terminal repeat (ITR), which is necessary for viral replication. The virus also comprises a virus-encoded protease, which is necessary for processing some of the structural proteins required to produce infectious virions. The structure of the adenoviral genome is described on the basis of the order in which the viral genes are expressed following host cell transduction. More specifically, the viral genes are referred to as early (E) or late (L) genes according to whether transcription occurs prior to or after onset of DNA replication. In the early phase of transduction, the E1A, E1B, E2A, E2B, E3 and E4 genes of adenovirus are expressed to prepare the host cell for viral replication. During the late phase of infection, expression of the late genes L1-L5, which encode the structural components of the virus particles, is activated.

Adenoviruses are species-specific and different serotypes, i.e., types of viruses that are not cross-neutralized by antibodies, have been isolated from a variety of mammalian species. For example, more than 50 serotypes have been isolated from humans which are divided into six subgroups (A–F; B is subdivided into B1 and B2) based on sequence homology and on their ability to agglutinate red blood cells (Tatsis and Ertl *Molecular Therapy* (2004) 10:616–629). Numerous adenoviruses have been isolated from nonhuman simians such as chimpanzees, bonobos, rhesus macaques and gorillas, and they are classified into the same human groups based on phylogenetic relationships based on hexon or fiber sequences (Colloca *et al.* (2012) *ScienceTranslational Medicine* 4:1-9; Roy *et al.* (2004) *Virology* 324: 361-372; Roy *et al.* (2010) *Journal of Gene Medicine* 13:17-25).

Adenovirus Capsid Proteins Including the Fiber Protein and Polynucleotides Encoding These Proteins

As outlined above, the adenoviral capsid comprises three major proteins, hexon, penton and fiber. The hexon accounts for the majority of the structural components of the capsid, which consists of 240 trimeric hexon capsomeres and 12 penton bases. The hexon has three conserved double barrels, while the top has three towers, each tower containing a loop from each subunit that forms most of the capsid. The base of hexon is highly conserved between adenoviral serotypes, while the surface loops are variable (Tatsis and Ertl *Molecular Therapy* (2004) 10:616–629).

Penton is another adenoviral capsid protein that forms a pentameric base to which fiber attaches. The trimeric fiber protein protrudes from the penton base at each of the 12 vertices of the capsid and is a knobbed rod-like structure. A remarkable difference in the surface of adenovirus capsids compared to that of most other icosahedral viruses is the presence of the long, thin fiber protein. The primary role of the fiber protein is the tethering of the viral capsid to the cell surface via its interaction with a cellular receptor.

The fiber proteins of many adenovirus serotypes share a common architecture: an N-terminal tail, a central shaft made of repeating sequences, and a C-terminal globular knob domain (or “head”). The central shaft domain consists of a variable number of beta-repeats. The beta-repeats connect to form an elongated structure of three intertwined spiralling strands that is highly rigid and stable. The shaft connects the N-terminal tail with the globular knob structure, which is responsible for interaction with the target cellular receptor. The globular nature of the adenovirus knob domain presents large surfaces for binding the receptor laterally and apically. The effect of this architecture is to project the receptor-binding site far from the virus capsid, thus freeing the virus from steric constraints presented by the relatively flat capsid surface.

Although fibers of many adenovirus serotypes have the same overall architecture, they have variable amino acid sequences that influence their function as well as structure. For example, a number of exposed regions on the surface of the fiber knob present an easily adaptable receptor binding site. The globular shape of the fiber knob allows receptors to bind at the sides of the knob or on top of the fiber knob. These binding sites typically lie on surface-exposed loops connecting beta-strands that are poorly conserved among human adenoviruses. The exposed side chains on these loops give the knob a variety of surface features while preserving the tertiary and quaternary structure. For example, the electrostatic potential and charge distributions at the knob surfaces can vary due to the wide range of isoelectric points in the fiber knob sequences, from pI approximately 9 for Ad 8, Ad 19, and Ad 37 to approximately 5 for subgroup B adenoviruses. As a structurally complex virus ligand, the fiber protein allows the presentation of a variety of binding surfaces (knob) in a number of orientations and distances (shaft) from the viral capsid.

One of the most obvious variations between some serotypes is fiber length. Studies have shown that the length of the fiber shaft strongly influences the interaction of the knob and the virus with its target receptors. Further, fiber proteins between serotypes can also vary in their ability to bend. Although beta-repeats in the shaft form a highly stable and regular structure, electron microscopy (EM) studies have shown distinct hinges in the fiber. Analysis of the protein sequence from several adenovirus serotype fibers pinpoints a disruption in the repeating sequences of the shaft at the third beta-repeat from the N-terminal tail, which correlates strongly with one of the hinges in the shaft, as seen by EM. The hinges in the fiber allow the knob to adopt a variety of orientations relative to the virus capsid, which may circumvent steric hindrances to receptor engagement requiring the correct presentation of the receptor binding site on the knob. For example, the rigid fibers of subgroup D Ads thus require

a flexible receptor or one prepositioned for virus attachment, as they are unable to bend themselves. (Nicklin et al *Molecular Therapy* 2005 12:384–393)

The identification of specific cell receptors for different Ad serotypes and the knowledge of how they contribute to tissue tropism have been achieved through the use of fiber pseudotyping technology. Although Ads of some subgroups use CAR as a primary receptor, it is becoming clear that many Ads use alternate primary receptors, leading to vastly different tropism in vitro and in vivo. The fibers of these serotypes show clear differences in their primary and tertiary structures, such as fiber shaft rigidity, the length of the fiber shaft, and the lack of a CAR binding site and/or the putative HSPG binding motif, together with the differences in net charge within the fiber knob. Pseudotyping Ad 5 particles with an alternate fiber shaft and knob therefore provides an opportunity to remove important cell binding domains and, in addition, may allow more efficient (and potentially more cell-selective) transgene delivery to defined cell types compared to that achieved with Ad 5. Neutralization of fiber-pseudotyped Ad particles may also be reduced if the fibers used are from Ads with lower seroprevalence in humans or experimental models, a situation that favours successful administration of the vector (Nicklin et al *Molecular Therapy* (2005) 12:384–393). Furthermore, full length fiber as well as isolated fiber knob regions, but not hexon or penton alone, are capable of inducing dendritic cell maturation and are associated with induction of a potent CD8+ T cell response (Molinier-Frenkel et al. *J. Biol. Chem.* (2003) 278:37175-37182). Taken together, adenoviral fiber plays an important role in at least receptor-binding and immunogenicity of adenoviral vectors.

Illustrating the differences between the fiber proteins of Group C simian adenoviruses is the alignment provided in Figure 1. A striking feature is that the fiber sequences of these adenoviruses can be broadly grouped into having a long fiber, such as ChAd155, or a short fiber, such as ChAd3. This length differential is due to a 36 amino acid deletion at approximately position 321 in the short fiber relative to the long fiber. In addition, there are a number of amino acid substitutions that differ between the short versus long fiber subgroup yet are consistent within each subgroup. While the exact function of these differences have not yet been elucidated, given the function and immunogenicity of fiber, they are likely to be significant. It has been shown that one of the determinants of viral tropism is the length of the fiber shaft. It has been demonstrated that an Ad5 vector with a shorter shaft has a lower efficiency of binding to CAR receptor and a lower infectivity (Ambriović-Ristov A. et al.: *Virology*. (2003) 312(2):425-33): It has been speculated that this impairment is the results of an increased rigidity of the shorter fiber leading to a less efficient attachment to the cell receptor (Wu, E et al.: *J Virol.* (2003) 77(13): 7225–7235). These studies may explain the improved properties of ChAd155 carrying a longer and more flexible fiber in comparison with the previously described ChAd3 and PanAd3 carrying a fiber with a shorter shaft.

In one aspect of the invention there is provided isolated fiber, penton and hexon capsid polypeptides of chimp adenovirus ChAd155 and isolated polynucleotides encoding the fiber, penton and hexon capsid polypeptides of chimp adenovirus ChAd155.

All three capsid proteins are expected to contribute to low seroprevalence and can, thus, be used independently from each other or in combination to suppress the affinity of an adenovirus to preexisting neutralizing antibodies, e.g. to manufacture a recombinant adenovirus with a reduced seroprevalence. Such a recombinant adenovirus may be a chimeric adenovirus with capsid proteins from different serotypes with at least a fiber protein from ChAd155.

The ChAd155 fiber polypeptide sequence is provided in SEQ ID NO: 1.

The ChAd155 penton polypeptide sequence is provided in SEQ ID NO: 3.

The ChAd155 hexon polypeptide sequence is provided in SEQ ID NO: 5.

Recombinant Adenoviruses or Compositions Comprising Polypeptide Sequences of ChAd155 Fiber or a Functional Derivative Thereof

Suitably the recombinant adenovirus or composition of the invention comprises a polypeptide having the amino acid sequence according to SEQ ID NO: 1.

Suitably the recombinant adenovirus or composition of the invention comprises a polypeptide which is a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1. Suitably the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1 has an amino acid sequence which is at least 80% identical, such as at least 85.0% identical, , such as at least 90% identical, such as at least 91.0% identical, such as at least 93.0% identical, such as at least 95.0% identical, such as at least 97.0% identical, such as at least 98.0% identical, such as at least 99.0% identical, such as at least 99.2% identical, such as at least 99.4% identical, such as 99.5% identical, such as at least 99.6% identical, such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO: 1. Alternatively the functional derivative has no more than 130, more suitably no more than 120, more suitably no more than 110, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 1.

Suitably the recombinant adenovirus or composition according to the invention further comprises:

(a) a polypeptide having the amino acid sequence according to SEQ ID NO: 3; or

(b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 3, wherein the functional derivative has an amino acid sequence which is at least 50.0% identical over its entire length to the amino acid sequence of SEQ ID NO: 3, and/or

(a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or
 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.

5

Suitably the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 3 has an amino acid sequence which is at least 60.0%, such as at least 70.0%, such as at least 80.0%, such as at least 85.0%, such as at least 90.0%, such as at least 91.0% identical, such as at least 93.0% identical, such as at least 95.0% identical, such as at least 97.0% identical, such as at least 98.0% identical, such as at least 99.0%, such as at least 99.2%, such as at least 99.4%, such as 99.5% identical, such as at least 99.6%, such as 99.7% identical such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO: 3. Alternatively the functional derivative has no more than 300, more suitably no more than 250, more suitably no more than 200, more suitably no more than 150, more suitably no more than 125, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 3.

Suitably the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5 has an amino acid sequence which is at least 60.0%, such as at least 70.0%, such as at least 80.0%, such as at least 85.0%, such as at least 90.0%, such as at least 91.0% identical, such as at least 93.0% identical, such as at least 95.0% identical, such as at least 97.0% identical, such as at least 98.0% identical, such as at least 99.0%, such as at least 99.2%, such as at least 99.4%, such as 99.5% identical, such as at least 99.6%, such as 99.7% identical such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO: 5. Alternatively the functional derivative has no more than 500, more suitably no more than 400, more suitably no more than 450, more suitably no more than 300, more suitably no more than 250, more suitably no more than 200, more suitably no more than 150, more suitably no more than 125, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 5.

40

Recombinant Adenoviruses or Compositions Comprising Polypeptide Sequences of ChAd155 Penton

Suitably the recombinant adenovirus or composition of the invention comprises a polypeptide having the amino acid sequence according to SEQ ID NO: 3.

Suitably the recombinant adenovirus or composition of the invention further comprises:

- 5 (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 1; or
(b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1
and/or
- 10 (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or
(b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which is at least 60% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.
- 15 Suitably the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1 has an amino acid sequence which is at least 60.0% identical, such as at least 70.0% identical, such as at least 80.0% identical, such as at least 85.0% identical, such as at least 87.0% identical, such as at least 89.0% identical, such as at least 91.0% identical, such as at least 93.0% identical, such as at least 95.0% identical, such as at least 97.0%
20 identical, such as at least 98.0% identical, such as at least 99.0% identical, such as at least 99.2%, such as at least 99.4%, such as 99.5% identical, such as at least 99.6%, such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO: 1. Alternatively the functional derivative has no more than 130, more suitably no more than 120, more suitably no more than 110, more suitably no more than 100,
25 more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or
30 substitutions(s) compared to SEQ ID NO: 1.

- Suitably the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5 has an amino acid sequence which is at least 60.0%, such as at least 70.0%, such as at least 80.0%, such as at least 85.0%, such as at least 90.0%, such as at least
35 95.0%, such as at least 97.0%, such as at least 99.0%, such as at least 99.0%, such as at least 99.2%, such as at least 99.4%, such as 99.5% identical, such as at least 99.6%, such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO:5. Alternatively the functional derivative has no more than 500, more suitably no more than 400, more suitably no more than 450, more suitably no more than 300,
40 more suitably no more than 250, more suitably no more than 200, more suitably no more than 150, more suitably no more than 125, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no

more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 5.

5

Recombinant Adenoviruses or Compositions comprising Polynucleotides Encoding ChAd155 Fiber or a Functional Derivative Thereof

10 Suitably the recombinant adenovirus or composition of the invention comprises a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 1. Suitably the polynucleotide has a sequence according to SEQ ID NO: 2.

15 Alternatively, the recombinant adenovirus or composition of the invention comprises a polynucleotide which encodes a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1. Suitably the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1 has an amino acid sequence which is at least 80% identical, such as at least 85.0% identical, such as at least 90% identical, such as at least 91.0% identical, such as at least 93.0% identical, such as at least 95.0% identical, such as at least 97.0% identical, such as at least 98.0% identical, such as at least 99.0% identical, such as at least 99% identical, such as at least 99.4% identical, such as at least 99.6% identical, such as at least 99.8% identical over its entire length to the amino acid sequence of SEQ ID NO: 1. Alternatively the functional derivative has no more than 130, more suitably no more than 120, more suitably no more than 110, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 1.

Suitably the recombinant adenovirus or composition of the invention further comprises a polynucleotide encoding:

- 35 (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 3; or
 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 3, wherein the functional derivative has an amino acid sequence which is at least 50.0% identical over its entire length to the amino acid sequence of SEQ ID NO: 3, and/or
- 40 (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or
 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.

Suitably the functional derivative of the polypeptide having the amino acid sequence according to SEQ ID NO: 3 has an amino acid sequence which is at least 60.0%, such as at least 70.0%, such as at least 80.0%, such as at least 85.0%, such as at least 90.0%, such as at least 91.0% identical, such as at least 93.0% identical, such as at least 95.0% identical, such as at least 97.0% identical, such as at least 98.0% identical, such as at least 99.0%, such as at least 99%, such as at least 99.4%, such as at least 99.6%, such as at least 99.8% identical over its entire length to the amino acid sequence of SEQ ID NO: 3. Alternatively the functional derivative has no more than 300, more suitably no more than 250, more suitably no more than 200, more suitably no more than 150, more suitably no more than 125, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 3.

Suitably the functional derivative of the polypeptide having the amino acid sequence according to SEQ ID NO: 5 has an amino acid sequence which is at least 60.0%, such as at least 70.0%, such as at least 80.0%, such as at least 85.0%, such as at least 90.0%, such as at least 95.0%, such as at least 97.0%, such as at least 98.0%, such as at least 99.0%, such as at least 99.2%, such as at least 99.4%, such as 99.5% identical, such as at least 99.6%, such as 99.7% identical such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO: 5. Alternatively the functional derivative has no more than 500, more suitably no more than 400, more suitably no more than 450, more suitably no more than 300, more suitably no more than 250, more suitably no more than 200, more suitably no more than 150, more suitably no more than 125, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 5.

35 Recombinant Adenoviruses or Compositions comprising Polynucleotides Encoding ChAd155 Penton

Suitably the recombinant adenovirus or composition of the invention comprises a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 3. Suitably the polynucleotide has a sequence according to SEQ ID NO: 4.

Suitably the recombinant adenovirus or composition of the invention further comprises a polynucleotide encoding:

(a) a polypeptide having the amino acid sequence according to SEQ ID NO: 1; or
 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 1

5 and/or

(a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or
 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.

10

Suitably the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1 has an amino acid sequence which is at least 60.0% identical, such as at least 70.0% identical, such as at least 80.0% identical, such as at least 85.0% identical, such as at least 87.0% identical, such as at least 89.0% identical, such as at least 91.0% identical, such as at least 93.0% identical, such as at least 95.0% identical, such as at least 97.0% identical, such as at least 98.0% identical, such as at least 99.0%, such as at least 99.2%, such as at least 99.4%, such as 99.5% identical, such as at least 99.6%, such as 99.7% identical such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO: 1. Alternatively the functional derivative has no more than 130, more suitably no more than 120, more suitably no more than 110, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 1.

20

25

30

Suitably the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5 has an amino acid sequence which is at least 60.0%, such as at least 70.0%, such as at least 80.0%, such as at least 85.0%, such as at least 90.0%, such as at least 95.0%, such as at least 97.0%, such as at least 98.0%, such as at least 99.0%, such as at least 99.2%, such as at least 99.4%, such as 99.5% identical, such as at least 99.6%, such as 99.7% identical such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO: 5. Alternatively the functional derivative has no more than 500, more suitably no more than 400, more suitably no more than 450, more suitably no more than 300, more suitably no more than 250, more suitably no more than 200, more suitably no more than 150, more suitably no more than 125, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 5.

35

40

ChAd155 Backbones

The present application describes isolated polynucleotide sequences of chimp adenovirus ChAd155, including that of wild type, unmodified ChAd155 (SEQ ID NO: 10) and modified backbone constructs of ChAd155. These modified backbone constructs include ChAd155#1434 (SEQ ID NO: 7), ChAd155#1390 (SEQ ID NO: 8) and ChAd155#1375 (SEQ ID NO: 9). ChAd155 backbones may be used in the construction of recombinant replication-competent or replication-incompetent adenoviruses for the delivery of transgenes.

Annotation of the ChAd155 wild type sequence (SEQ ID NO: 10) sequence is provided below.

```

LOCUS      ChAd155                      37830 bp    DNA        linear        10-JUN-2015
DEFINITION Chimp adenovirus 155, complete genome.
COMMENT    Annotation according to alignment of ChAd155 against the human
           Adenovirus 2 reference strain NC_001405
           Two putative ORFs in the E3 region added manually
FEATURES             Location/Qualifiers
     source           1..37830
                     /organism="Chimpanzee adenovirus 155"
                     /mol_type="genomic DNA"
                     /acronym="ChAd155"
     repeat_region    1..101
                     /standard_name="ITR"
                     /rpt_type=inverted
     gene             466..1622
                     /gene="E1A"
     TATA_signal       466..471
                     /gene="E1A"
     prim_transcript   497..1622
                     /gene="E1A"
     CDS               join(577..1117,1231..1532)
                     /gene="E1A"
                     /product="E1A_280R"
     CDS               join(577..979,1231..1532)
                     /gene="E1A"
                     /product="E1A_243R"
     polyA_signal      1600..1605
                     /gene="E1A"
     gene             1662..4131
                     /gene="E1B"
     TATA_signal       1662..1667
                     /gene="E1B"

```

	prim_transcript	1692..4131
		/gene="E1B"
	CDS	1704..2267
		/gene="E1B"
5		/product="E1B_19K"
	CDS	2009..3532
		/gene="E1B"
		/product="E1B_55K"
	gene	3571..4131
10		/gene="IX"
	TATA_signal	3571..3576
		/gene="IX"
	prim_transcript	3601..4131
		/gene="IX"
15	CDS	3628..4092
		/gene="IX"
		/product="IX"
	polyA_signal	4097..4102
		/note="E1B, IX"
20	gene	complement(4117..27523)
		/gene="E2B"
	prim_transcript	complement(4117..27494)
		/gene="E2B"
	gene	complement(4117..5896)
25		/gene="IVa2"
	prim_transcript	complement(4117..5896)
		/gene="IVa2"
	CDS	complement(join(4151..5487,5766..5778))
		/gene="IVa2"
30		/product="E2B_IVa2"
	polyA_signal	complement(4150..4155)
		/note="IVa2, E2B"
	CDS	complement(join(5257..8838,14209..14217))
		/gene="E2B"
35		/product="E2B_polymerase"
	gene	6078..34605
		/gene="L5"
	gene	6078..28612
		/gene="L4"
40	gene	6078..22658
		/gene="L3"
	gene	6078..18164
		/gene="L2"

	gene	6078..14216
		/gene="L1"
	TATA_signal	6078..6083
		/note="L"
5	prim_transcript	6109..34605
		/gene="L5"
	prim_transcript	6109..28612
		/gene="L4"
10	prim_transcript	6109..22658
		/gene="L3"
	prim_transcript	6109..18164
		/gene="L2"
	prim_transcript	6109..14216
		/gene="L1"
15	CDS	join(8038..8457,9722..9742)
		/gene="L1"
		/product="L1_13.6K"
	CDS	complement(join(8637..10640,14209..14217))
		/gene="E2B"
20		/product="E2B_pTP"
	gene	10671..10832
		/gene="VAI"
	misc_RNA	10671..10832
		/gene="VAI"
25		/product="VAI"
	gene	10902..11072
		/gene="VAII"
	misc_RNA	10902..11072
		/gene="VAII"
30		/product="VAII"
	CDS	11093..12352
		/gene="L1"
		/product="L1_52K"
35	CDS	12376..14157
		/gene="L1"
		/product="L1_pIIIIa"
	polyA_signal	14197..14202
		/gene="L1"
40	CDS	14254..16035
		/gene="L2"
		/product="L2_penton"
	CDS	16050..16646
		/gene="L2"

		/product="L2_pVII"
	CDS	16719..17834
		/gene="L2"
		/product="L2_v"
5	CDS	17859..18104
		/gene="L2"
		/product="L2_pX"
	polyA_signal	18143..18148
		/gene="L2"
10	CDS	18196..18951
		/gene="L3"
		/product="L3_pVI"
	CDS	19063..21945
		/gene="L3"
15		/product="L3_hexon"
	CDS	21975..22604
		/gene="L3"
		/product="L3_protease"
	polyA_signal	22630..22635
20		/gene="L3"
	gene	complement(22632..27523)
		/gene="E2A"
	prim_transcript	complement(22632..27494)
		/gene="E2A"
25	gene	complement(22632..26357)
		/gene="E2A-L"
	prim_transcript	complement(22632..26328)
		/gene="E2A-L"
	polyA_signal	complement(22649..22654)
30		/note="E2A, E2A-L"
	CDS	complement(22715..24367)
		/gene="E2A"
		/note="DBP; genus-common; DBP family"
		/codon_start=1
35		/product="E2A"
	CDS	24405..26915
		/gene="L4"
		/product="L4_100k"
	TATA_signal	complement(26352..26357)
40		/gene="E2A-L"
	CDS	join(26602..26941,27147..27529)
		/gene="L4"
		/product="L4_33K"

	CDS	26602..27207
		/gene="L4"
		/product="L4_22K"
5	TATA_signal	complement(27518..27523)
		/note="E2A, E2B; nominal"
	CDS	27604..28287
		/gene="L4"
		/product="L4_pVIII"
10	gene	27969..32686
		/gene="E3B"
	gene	27969..31611
		/gene="E3A"
	TATA_signal	27969..27974
		/note="E3A, E3B"
15	prim_transcript	27998..32686
		/gene="E3B"
	prim_transcript	27998..31611
		/gene="E3A"
20	CDS	28288..28605
		/gene="E3A"
		/product="E3 ORF1"
	polyA_signal	28594..28599
		/gene="L4"
25	CDS	29103..29303
		/gene="E3A"
		/product="E3 ORF2"
	CDS	29300..29797
		/gene="E3A"
		/product="E3 ORF3"
30	CDS	29826..30731
		/gene="E3A"
		/product="E3 ORF4"
	CDS	30728..31579
		/gene="E3A"
35		/product="E3 ORF5"
	CDS	31283..31579
		/gene="E3A"
		/product="E3 ORF6"
40	polyA_signal	31578..31584
		/gene="E3A"
	CDS	31591..31863
		/gene="E3B"
		/product="E3 ORF7"

5	CDS	31866..32264 /gene="E3B" /product="E3 ORF8"
	CDS	32257..32643 /gene="E3B" /product="E3 ORF9"
	polyA_signal	32659..32664 /gene="E3B"
	gene	complement(<32678..32838) /gene="U"
	CDS	complement(<32678..32838) /gene="U" /note="exon encoding C terminus unidentified; genus-common" /product="protein U"
15	CDS	32849..34585 /gene="L5" /product="L5_fiber"
	polyA_signal	34581..34586 /gene="L5"
	gene	complement(34611..37520) /gene="E4"
	prim_transcript	complement(34611..37490) /gene="E4"
	polyA_signal	complement(34625..34630) /gene="E4"
25	CDS	complement(join(34794..35069,35781..35954)) /gene="E4" /product="E4 ORF7"
	CDS	complement(35070..35954) /gene="E4" /product="E4 ORF6"
	CDS	complement(35875..36219) /gene="E4" /product="E4 ORF4"
	CDS	complement(36235..36582) /gene="E4" /product="E4 ORF3"
	CDS	complement(36579..36971) /gene="E4" /product="E4 ORF2"
35	CDS	complement(37029..37415) /gene="E4"
	CDS	complement(37029..37415) /gene="E4"
	CDS	complement(37029..37415) /gene="E4"
	CDS	complement(37029..37415) /gene="E4"
	CDS	complement(37029..37415) /gene="E4"
40	CDS	complement(37029..37415) /gene="E4"
	CDS	complement(37029..37415) /gene="E4"
	CDS	complement(37029..37415) /gene="E4"
	CDS	complement(37029..37415) /gene="E4"
	CDS	complement(37029..37415) /gene="E4"

/product="E4 ORF1"
 TATA_signal complement(37515..37520)
 /gene="E4"
 repeat_region 37740..37830
 5 /standard_name="ITR"
 /rpt_type=inverted

Definitions

10 Recombinant means that the polynucleotide is the product of at least one of cloning, restriction or ligation steps, or other procedures that result in a polynucleotide that is distinct from a polynucleotide found in nature. A recombinant adenovirus is an adenovirus comprising a recombinant polynucleotide.

15 Typically, "heterologous" means derived from a genotypically distinct entity from that of the rest of the entity to which it is being compared. A heterologous nucleic acid sequence refers to any nucleic acid sequence that is not isolated from, derived from, or based upon a naturally occurring nucleic acid sequence of the adenoviral vector. "Naturally occurring" means a sequence found in nature and not synthetically prepared or modified. A sequence is "derived"
 20 from a source when it is isolated from a source but modified (e.g., by deletion, substitution (mutation), insertion, or other modification), suitably so as not to disrupt the normal function of the source gene.

A "functional derivative" of a polypeptide suitably refers to a modified version of a polypeptide,
 25 e.g. wherein one or more amino acids of the polypeptide may be deleted, inserted, modified and/or substituted. A derivative of an unmodified adenoviral capsid protein is considered functional if, for example:

- 30 (a) an adenovirus comprising the derivative capsid protein within its capsid retains substantially the same or a lower seroprevalence compared to an adenovirus comprising the unmodified capsid protein and/or
- (b) an adenovirus comprising the derivative capsid protein within its capsid retains substantially the same or a higher host cell infectivity compared to an adenovirus comprising the unmodified capsid protein and/or
- 35 (c) an adenovirus comprising the derivative capsid protein within its capsid retains substantially the same or a higher immunogenicity compared to an adenovirus comprising the unmodified capsid protein and or
- (d) an adenovirus comprising the derivative capsid protein within its capsid retains substantially the same or a higher level of transgene productivity compared to an
 40 adenovirus comprising the unmodified capsid protein.

Properties (a)-(d) above may suitably be measured using the methods described in the Examples section below.

Suitably, the recombinant adenovirus has a low seroprevalence in a human population. "Low seroprevalence" may mean having a reduced pre-existing neutralizing antibody level as compared to human adenovirus 5 (Ad5). Similarly or alternatively, "low seroprevalence" may mean less than about 20% seroprevalence, less than about 15% seroprevalence, less than about 10% seroprevalence, less than about 5% seroprevalence, less than about 4% seroprevalence, less than about 3% seroprevalence, less than about 2% seroprevalence, less than about 1% seroprevalence or no detectable seroprevalence. Seroprevalence can be measured as the percentage of individuals having a clinically relevant neutralizing titre (defined as a 50% neutralisation titer >200) using methods as described in Aste-Amézaga *et al.*, Hum. Gene Ther. (2004) 15(3):293-304.

The terms polypeptide, peptide and protein are used interchangeably herein.

The term "simian" is typically meant to encompass nonhuman primates, for example Old World monkeys, New World monkeys, apes and gibbons. In particular, simian may refer to nonhuman apes such as chimpanzees (*Pan troglodyte*), bonobos (*Pan paniscus*) and gorillas (genus *Gorilla*). Non-ape simians may include rhesus macaques (*Macaca mulatta*).

Adenovirus Sequence Comparison

For the purposes of comparing two closely-related polynucleotide or polypeptide sequences, the "% identity" between a first sequence and a second sequence may be calculated using an alignment program, such as BLAST® (available at blast.ncbi.nlm.nih.gov, last accessed 09 March 2015) using standard settings. The % identity is the number of identical residues divided by the number of residues in the reference sequence, multiplied by 100. The % identity figures referred to above and in the claims are percentages calculated by this methodology. An alternative definition of % identity is the number of identical residues divided by the number of aligned residues, multiplied by 100. Alternative methods include using a gapped method in which gaps in the alignment, for example deletions in one sequence relative to the other sequence, are accounted for in a gap score or a gap cost in the scoring parameter. For more information, see the BLAST® fact sheet available at ftp.ncbi.nlm.nih.gov/pub/factsheets/HowTo_BLASTGuide.pdf, last accessed on 09 March 2015.

Sequences that preserve the functionality of the polynucleotide or a polypeptide encoded thereby are likely to be more closely identical. Polypeptide or polynucleotide sequences are said to be the same as or identical to other polypeptide or polynucleotide sequences, if they share 100% sequence identity over their entire length.

A "difference" between sequences refers to an insertion, deletion or substitution of a single amino acid residue in a position of the second sequence, compared to the first sequence. Two polypeptide sequences can contain one, two or more such amino acid differences. Insertions,

deletions or substitutions in a second sequence which is otherwise identical (100% sequence identity) to a first sequence result in reduced percent sequence identity. For example, if the identical sequences are 9 amino acid residues long, one substitution in the second sequence results in a sequence identity of 88.9%. If the identical sequences are 17 amino acid residues long, two substitutions in the second sequence results in a sequence identity of 88.2%. If the identical sequences are 7 amino acid residues long, three substitutions in the second sequence results in a sequence identity of 57.1%. If first and second polypeptide sequences are 9 amino acid residues long and share 6 identical residues, the first and second polypeptide sequences share greater than 66% identity (the first and second polypeptide sequences share 66.7% identity). If first and second polypeptide sequences are 17 amino acid residues long and share 16 identical residues, the first and second polypeptide sequences share greater than 94% identity (the first and second polypeptide sequences share 94.1% identity). If first and second polypeptide sequences are 7 amino acid residues long and share 3 identical residues, the first and second polypeptide sequences share greater than 42% identity (the first and second polypeptide sequences share 42.9% identity).

Alternatively, for the purposes of comparing a first, reference polypeptide sequence to a second, comparison polypeptide sequence, the number of additions, substitutions and/or deletions made to the first sequence to produce the second sequence may be ascertained. An addition is the addition of one amino acid residue into the sequence of the first polypeptide (including addition at either terminus of the first polypeptide). A substitution is the substitution of one amino acid residue in the sequence of the first polypeptide with one different amino acid residue. A deletion is the deletion of one amino acid residue from the sequence of the first polypeptide (including deletion at either terminus of the first polypeptide).

For the purposes of comparing a first, reference polynucleotide sequence to a second, comparison polynucleotide sequence, the number of additions, substitutions and/or deletions made to the first sequence to produce the second sequence may be ascertained. An addition is the addition of one nucleotide residue into the sequence of the first polynucleotide (including addition at either terminus of the first polynucleotide). A substitution is the substitution of one nucleotide residue in the sequence of the first polynucleotide with one different nucleotide residue. A deletion is the deletion of one nucleotide residue from the sequence of the first polynucleotide (including deletion at either terminus of the first polynucleotide).

Suitably substitutions in the sequences of the present invention may be conservative substitutions. A conservative substitution comprises the substitution of an amino acid with another amino acid having a chemical property similar to the amino acid that is substituted (see, for example, Stryer et al, *Biochemistry*, 5th Edition 2002, pages 44-49). Preferably, the conservative substitution is a substitution selected from the group consisting of: (i) a substitution of a basic amino acid with another, different basic amino acid; (ii) a substitution of an acidic amino acid with another, different acidic amino acid; (iii) a substitution of an aromatic amino acid with another, different aromatic amino acid; (iv) a substitution of a non-polar, aliphatic amino acid with another, different non-polar, aliphatic amino acid; and (v) a

substitution of a polar, uncharged amino acid with another, different polar, uncharged amino acid. A basic amino acid is preferably selected from the group consisting of arginine, histidine, and lysine. An acidic amino acid is preferably aspartate or glutamate. An aromatic amino acid is preferably selected from the group consisting of phenylalanine, tyrosine and tryptophane. A non-polar, aliphatic amino acid is preferably selected from the group consisting of glycine, alanine, valine, leucine, methionine and isoleucine. A polar, uncharged amino acid is preferably selected from the group consisting of serine, threonine, cysteine, proline, asparagine and glutamine. In contrast to a conservative amino acid substitution, a non-conservative amino acid substitution is the exchange of one amino acid with any amino acid that does not fall under the above-outlined conservative substitutions (i) through (v).

Recombinant Adenovirus

The ChAd155 sequences are useful as therapeutic agents and in construction of a variety of vector systems, recombinant adenovirus and host cells. Suitably the term "vector" refers to a nucleic acid that has been substantially altered (e.g., a gene or functional region that has been deleted and/or inactivated) relative to a wild type sequence and/or incorporates a heterologous sequence, *i.e.* nucleic acid obtained from a different source (also called an "insert"), and replicating and/or expressing the inserted polynucleotide sequence, when introduced into a cell (e.g., a host cell). For example, the insert may be all or part of the ChAd155 sequences described herein. In addition or alternatively, a ChAd155 vector may be a ChAd155 adenovirus comprising one or more deletions or inactivations of viral genes, such as E1 or other viral gene or functional region described herein. Such a ChAd155, which may or may not comprise a heterologous sequence, is often called a "backbone" and may be used as is or as a starting point for additional modifications to the vector.

A vector may be any suitable nucleic acid molecule including naked DNA, a plasmid, a virus, a cosmid, phage vector such as lambda vector, an artificial chromosome such as a BAC (bacterial artificial chromosome), or an episome. Alternatively, a vector may be a transcription and/or expression unit for cell-free *in vitro* transcription or expression, such as a T7-compatible system. The vectors may be used alone or in combination with other adenoviral sequences or fragments, or in combination with elements from non-adenoviral sequences. The ChAd155 sequences are also useful in antisense delivery vectors, gene therapy vectors, or vaccine vectors. Thus, further provided are gene delivery vectors, and host cells which contain the ChAd155 sequences.

The term "replication-competent" adenovirus refers to an adenovirus which can replicate in a host cell in the absence of any recombinant helper proteins comprised in the cell. Suitably, a "replication-competent" adenovirus comprises the following intact or functional essential early genes: E1A, E1B, E2A, E2B, E3 and E4. Wild type adenoviruses isolated from a particular animal will be replication competent in that animal.

The term "replication-incompetent" or "replication-defective" adenovirus refers to an adenovirus which is incapable of replication because it has been engineered to comprise at least a functional deletion (or "loss-of-function" mutation), i.e. a deletion or mutation which impairs the function of a gene without removing it entirely, e.g. introduction of artificial stop codons, deletion or mutation of active sites or interaction domains, mutation or deletion of a regulatory sequence of a gene etc, or a complete removal of a gene encoding a gene product that is essential for viral replication, such as one or more of the adenoviral genes selected from E1A, E1B, E2A, E2B, E3 and E4 (such as E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8, E3 ORF9, E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2 and/or E4 ORF1). Particularly suitably E1 and optionally E3 and/or E4 are deleted. If deleted, the aforementioned deleted gene region will suitably not be considered in the alignment when determining % identity with respect to another sequence.

The present invention provides recombinant adenovirus that deliver a mycobacterial antigen, to cells, either for therapeutic or vaccine purposes. A vector may include any genetic element including naked DNA, a phage, transposon, cosmid, episome, plasmid, or a virus. Such vectors contain DNA of ChAd155 as disclosed herein and a minigene. By "minigene" (or "expression cassette") is meant the combination of a selected heterologous gene (transgene) and the other regulatory elements necessary to drive translation, transcription and/or expression of the gene product in a host cell.

Typically, a ChAd155-derived adenoviral vector is designed such that the minigene is located in a nucleic acid molecule which contains other adenoviral sequences in the region native to a selected adenoviral gene. The minigene may be inserted into an existing gene region to disrupt the function of that region, if desired. Alternatively, the minigene may be inserted into the site of a partially or fully deleted adenoviral gene. For example, the minigene may be located in the site of a mutation, insertion or deletion which renders non-functional at least one gene of a genomic region selected from the group consisting of E1A, E1B, E2A, E2B, E3 and E4. The term "renders non-functional" means that a sufficient amount of the gene region is removed or otherwise disrupted, so that the gene region is no longer capable of producing functional products of gene expression. If desired, the entire gene region may be removed (and suitably replaced with the minigene).

For example, for a production vector useful for generation of a recombinant virus, the vector may contain the minigene and either the 5' end of the adenoviral genome or the 3' end of the adenoviral genome, or both the 5' and 3' ends of the adenoviral genome. The 5' end of the adenoviral genome contains the 5' cis-elements necessary for packaging and replication; i.e., the 5' ITR sequences (which function as origins of replication) and the native 5' packaging enhancer domains (that contain sequences necessary for packaging linear Ad genomes and enhancer elements for the E1 promoter). The 3' end of the adenoviral genome includes the 3' cis-elements (including the ITRs) necessary for packaging and encapsidation. Suitably, a recombinant adenovirus contains both 5' and 3' adenoviral cis-elements and the minigene

(suitably containing a transgene) is located between the 5' and 3' adenoviral sequences. A ChAd155-based adenoviral vector may also contain additional adenoviral sequences.

5 Suitably, ChAd155-based vectors contain one or more adenoviral elements derived from the adenoviral ChAd155 genome. In one embodiment, the vectors contain adenoviral ITRs from ChAd155 and additional adenoviral sequences from the same adenoviral serotype. In another embodiment, the vectors contain adenoviral sequences that are derived from a different adenoviral serotype than that which provides the ITRs.

10 As defined herein, a pseudotyped adenovirus refers to an adenovirus in which the capsid proteins of the adenovirus are from a different adenovirus than the adenovirus which provides the ITRs.

15 Further, chimeric or hybrid adenoviruses may be constructed using the adenoviruses described herein using techniques known to those of skill in the art (e.g., US 7,291,498).

ITRs and any other adenoviral sequences present in the vector may be obtained from many sources. A variety of adenovirus strains are available from the American Type Culture Collection, Manassas, Virginia, or available by request from a variety of commercial and
20 institutional sources. Further, the sequences of many such strains are available from a variety of databases including, e.g., PubMed and GenBank. Homologous adenovirus vectors prepared from other chimp or from human adenoviruses are described in the published literature (for example, US 5,240,846). The DNA sequences of a number of adenovirus types are available from GenBank, including type Ad5 (GenBank Accession Number M73370). The
25 adenovirus sequences may be obtained from any known adenovirus serotype, such as serotypes 2, 3, 4, 7, 12 and 40, and further including any of the presently identified human types. Similarly adenoviruses known to infect nonhuman animals (e.g., simians) may also be employed in the vector constructs of this invention (e.g., US 6,083,716). The viral sequences, helper viruses (if needed), and recombinant viral particles, and other vector components and
30 sequences employed in the construction of the vectors described herein may be obtained as described below.

Sequence, Vector and Adenovirus Production

35 The sequences may be produced by any suitable means, including recombinant production, chemical synthesis, or other synthetic means. Suitable production techniques are well known to those of skill in the art. Alternatively, peptides can also be synthesized by well known solid phase peptide synthesis methods.

40 The adenoviral plasmids (or other vectors) may be used to produce adenoviral vectors. In one embodiment, the adenoviral vectors are adenoviral particles which are replication-incompetent. In one embodiment, the adenoviral particles are rendered replication-incompetent by deletions in the E1A and/or E1B genes. Alternatively, the adenoviruses are rendered replication-

incompetent by another means, optionally while retaining the E1A and/or E1B genes. Similarly, in some embodiments, reduction of an immune response to the vector may be accomplished by deletions in the E2B and/or DNA polymerase genes. The adenoviral vectors can also contain other mutations to the adenoviral genome, e.g., temperature-sensitive mutations or deletions in other genes. In other embodiments, it is desirable to retain an intact E1A and/or E1B region in the adenoviral vectors. Such an intact E1 region may be located in its native location in the adenoviral genome or placed in the site of a deletion in the native adenoviral genome (e.g., in the E3 region).

In the construction of adenovirus vectors for delivery of a gene to a mammalian (such as human) cell, a range of modified adenovirus nucleic acid sequences can be employed in the vectors. For example, all or a portion of the adenovirus delayed early gene E3 may be eliminated from the adenovirus sequence which forms a part of the recombinant virus. The function of E3 is believed to be irrelevant to the function and production of the recombinant virus particle. Adenovirus vectors may also be constructed having a deletion of at least the ORF6 region of the E4 gene, and more desirably because of the redundancy in the function of this region, the entire E4 region. Still another vector of the invention contains a deletion in the delayed early gene E2A. Deletions may also be made in any of the late genes L1 to L5 of the adenovirus genome. Similarly, deletions in the intermediate genes IX and IVa2 may be useful for some purposes. Other deletions may be made in the other structural or non-structural adenovirus genes. The above discussed deletions may be used individually, i.e., an adenovirus sequence for use as described herein may contain deletions in only a single region. Alternatively, deletions of entire genes or portions thereof effective to destroy their biological activity may be used in any combination. For example, in one exemplary vector, the adenovirus sequence may have deletions of the E1 genes and the E4 gene, or of the E1, E2A and E3 genes, or of the E1 and E3 genes, or of E1, E2A and E4 genes, with or without deletion of E3, and so on. Any one or more of the E genes may suitably be replaced with an E gene (or one or more E gene open reading frames) sourced from a different strain of adenovirus. Particularly suitably the ChAd155 E1 and E3 genes are deleted and the ChAd155E4 gene is replaced with E4Ad5orf6. As discussed above, such deletions and/or substitutions may be used in combination with other mutations, such as temperature-sensitive mutations, to achieve a desired result.

An adenoviral vector lacking one or more essential adenoviral sequences (e.g., E1A, E1B, E2A, E2B, E4 ORF6, L1, L2, L3, L4 and L5) may be cultured in the presence of the missing adenoviral gene products which are required for viral infectivity and propagation of an adenoviral particle. These helper functions may be provided by culturing the adenoviral vector in the presence of one or more helper constructs (e.g., a plasmid or virus) or a packaging host cell.

Complementation of Replication-Incompetent Vectors

To generate recombinant adenoviruses deleted in any of the genes described above, the function of the deleted gene region, if essential to the replication and infectivity of the virus, must be supplied to the recombinant virus by a helper virus or cell line, i.e., a complementation or packaging cell line.

5

Helper Viruses

Depending upon the adenovirus gene content of the viral vectors employed to carry the minigene, a helper adenovirus or non-replicating virus fragment may be used to provide
10 sufficient adenovirus gene sequences necessary to produce an infective recombinant viral particle containing the minigene. Useful helper viruses contain selected adenovirus gene sequences not present in the adenovirus vector construct and/or not expressed by the packaging cell line in which the vector is transfected. In one embodiment, the helper virus is replication-defective and contains adenovirus genes in addition, suitably, to one or more of the
15 sequences described herein. Such a helper virus is suitably used in combination with an E1 expressing (and optionally additionally E3 expressing) cell line.

A helper virus may optionally contain a reporter gene. A number of such reporter genes are known to the art as well as described herein. The presence of a reporter gene on the helper
20 virus which is different from the transgene on the adenovirus vector allows both the adenoviral vector and the helper virus to be independently monitored. This reporter is used to enable separation between the resulting recombinant virus and the helper virus upon purification.

Complementation Cell Lines

25

In many circumstances, a cell line expressing the one or more missing genes which are essential to the replication and infectivity of the virus, such as human E1, can be used to transcomplement a chimp adenoviral vector. This is particularly advantageous because, due to the diversity between the chimp adenovirus sequences of the invention and the human
30 adenovirus sequences found in currently available packaging cells, the use of the current human E1-containing cells prevents the generation of replication-competent adenoviruses during the replication and production process.

Alternatively, if desired, one may utilize the sequences provided herein to generate a
35 packaging cell or cell line that expresses, at a minimum, the E1 gene from ChAd155 under the transcriptional control of a promoter for expression in a selected parent cell line. Inducible or constitutive promoters may be employed for this purpose. Examples of such promoters are described in detail elsewhere in this document. A parent cell is selected for the generation of a novel cell line expressing any desired ChAd155 gene. Without limitation, such a parent cell
40 line may be HeLa [ATCC Accession No. CCL 2], A549 [ATCC Accession No. CCL 185], HEK 293, KB [CCL 17], Detroit [e.g., Detroit 510, CCL 72] and WI-38 [CCL 75] cells, among others. These cell lines are all available from the American Type Culture Collection, 10801 University Boulevard, Manassas, Virginia 20110-2209.

Such E1-expressing cell lines are useful in the generation of recombinant adenovirus E1 deleted vectors. Additionally, or alternatively, cell lines that express one or more adenoviral gene products, e.g., E1A, E1B, E2A, E3 and/or E4, can be constructed using essentially the same procedures as used in the generation of recombinant viral vectors. Such cell lines can be utilised to transcomplement adenovirus vectors deleted in the essential genes that encode those products, or to provide helper functions necessary for packaging of a helper-dependent virus (e.g., adeno-associated virus). The preparation of a host cell involves techniques such as assembly of selected DNA sequences.

In another alternative, the essential adenoviral gene products are provided in trans by the adenoviral vector and/or helper virus. In such an instance, a suitable host cell can be selected from any biological organism, including prokaryotic (e.g., bacterial) cells, and eukaryotic cells, including, insect cells, yeast cells and mammalian cells.

Host cells may be selected from among any mammalian species, including, without limitation, cells such as A549, WEHI, 3T3, 10T1/2, HEK 293 cells or Per.C6 (both of which express functional adenoviral E1) [Fallaux, FJ et al, (1998), Hum Gene Ther, 9:1909-1917], Saos, C2C12, L cells, HT1080, HepG2 and primary fibroblast, hepatocyte and myoblast cells derived from mammals including human, monkey, mouse, rat, rabbit, and hamster.

A particularly suitable complementation cell line is the Procell92 cell line. The Procell92 cell line is based on HEK 293 cells which express adenoviral E1 genes, transfected with the Tet repressor under control of the human phosphoglycerate kinase-1 (PGK) promoter, and the G418-resistance gene (Vitelli et al. *PLOS One* (2013) 8(e55435):1-9). Procell92.S is adapted for growth in suspension conditions and is useful for producing adenoviral vectors expressing toxic proteins (www.okairos.com/e/inner.php?m=00084, last accessed 13 April 2015).

Assembly of a Viral Particle and Transfection of a Cell Line

Generally, when delivering the vector comprising the minigene by transfection, the vector is delivered in an amount from about 5 µg to about 100 µg DNA, and preferably about 10 to about 50 µg DNA to about 1×10^4 cells to about 1×10^{13} cells, and preferably about 10^5 cells. However, the relative amounts of vector DNA to host cells may be adjusted, taking into consideration such factors as the selected vector, the delivery method and the host cells selected.

Introduction into the host cell of the vector may be achieved by any means known in the art, including transfection, and infection. One or more of the adenoviral genes may be stably integrated into the genome of the host cell, stably expressed as episomes, or expressed transiently. The gene products may all be expressed transiently, on an episome or stably integrated, or some of the gene products may be expressed stably while others are expressed transiently.

Introduction of vectors into the host cell may also be accomplished using techniques known to the skilled person. Suitably, standard transfection techniques are used, e.g., CaPC transfection or electroporation.

5

Assembly of the selected DNA sequences of the adenovirus (as well as the transgene and other vector elements) into various intermediate plasmids, and the use of the plasmids and vectors to produce a recombinant viral particle are all achieved using conventional techniques. Such techniques include conventional cloning techniques of cDNA, use of overlapping
10 oligonucleotide sequences of the adenovirus genomes, polymerase chain reaction, and any suitable method which provides the desired nucleotide sequence. Standard transfection and co-transfection techniques are employed, e.g., CaPC precipitation techniques. Other conventional methods employed include homologous recombination of the viral genomes, plaquing of viruses in agar overlay, methods of measuring signal generation, and the like.

15

For example, following the construction and assembly of the desired minigene-containing viral vector, the vector is transfected in vitro in the presence of a helper virus into the packaging cell line. Homologous recombination occurs between the helper and the vector sequences, which permits the adenovirus-transgene sequences in the vector to be replicated and packaged into
20 virion capsids, resulting in the recombinant viral vector particles. The resulting recombinant adenoviruses are useful in transferring a selected transgene to a selected cell. In in vivo experiments with the recombinant virus grown in the packaging cell lines, the E1-deleted recombinant adenoviral vectors of the invention demonstrate utility in transferring a transgene to a non-simian mammal, preferably a human, cell.

25

Transgenes

The transgene is a nucleic acid sequence, heterologous to the vector sequences flanking the transgene, which encodes a protein of interest. The nucleic acid coding sequence is
30 operatively linked to regulatory components in a manner which permits transgene transcription, translation, and/or expression in a host cell.

The composition of the transgene sequence will depend upon the use to which the resulting vector will be put. For example, the transgene may be a therapeutic transgene or an immunogenic transgene. Alternatively, a transgene sequence may include a reporter
35 sequence, which upon expression produces a detectable signal. Such reporter sequences include, without limitation, DNA sequences encoding β -lactamase, β -galactosidase (LacZ), alkaline phosphatase, thymidine kinase, green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), luciferase, membrane bound proteins including, for example, CD2,
40 CD4, CD8, the influenza hemagglutinin protein, and others well known in the art, to which high affinity antibodies directed thereto exist or can be produced by conventional means, and fusion proteins comprising a membrane bound protein appropriately fused to an antigen tag domain from, among others, hemagglutinin or Myc. These coding sequences, when associated with

regulatory elements which drive their expression, provide signals detectable by conventional means, including enzymatic, radiographic, colorimetric, fluorescence or other spectrographic assays, fluorescent activating cell sorting assays and immunological assays, including enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and immunohistochemistry.

5

In one embodiment, the transgene is a non-marker sequence encoding a product which is useful in biology and medicine, such as a therapeutic transgene or an immunogenic transgene such as proteins, RNA, enzymes, or catalytic RNAs. Desirable RNA molecules include tRNA, dsRNA, ribosomal RNA, catalytic RNAs, and antisense RNAs. One example of a useful RNA sequence is a sequence which extinguishes expression of a targeted nucleic acid sequence in the treated animal.

10

The transgene may be used for treatment, e.g., as a vaccine, for induction of an immune response, and/or for prophylactic vaccine purposes. As used herein, induction of an immune response refers to the ability of a protein to induce a T cell and/or a humoral immune response to the protein.

15

Regulatory Elements

20

In addition to the transgene the vector also includes conventional control elements which are operably linked to the transgene in a manner that permits its transcription, translation and/or expression in a cell transfected with the plasmid vector or infected with the virus produced by the invention. As used herein, "operably linked" sequences include both expression control sequences that are contiguous with the gene of interest and expression control sequences that act in trans or at a distance to control the gene of interest.

25

Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation (poly A) signals including rabbit beta-globin polyA; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (e.g., Kozak consensus sequence); sequences that enhance protein stability; and when desired, sequences that enhance secretion of the encoded product. Among other sequences, chimeric introns may be used.

30

35

In some embodiments, the Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE) (Zuffrey et al. (1999) J Virol; 73(4):2886-9) may be operably linked to the transgene. An exemplary WPRE is provided in SEQ ID NO: 26.

40

A "promoter" is a nucleotide sequence that permits binding of RNA polymerase and directs the transcription of a gene. Typically, a promoter is located in the 5' non-coding region of a gene, proximal to the transcriptional start site of the gene. Sequence elements within promoters that function in the initiation of transcription are often characterized by consensus nucleotide sequences. Examples of promoters include, but are not limited to, promoters from bacteria,

yeast, plants, viruses, and mammals (including humans). A great number of expression control sequences, including promoters which are internal, native, constitutive, inducible and/or tissue-specific, are known in the art and may be utilized.

5 Examples of constitutive promoters include, without limitation, the TBG promoter, the retroviral Rous sarcoma virus LTR promoter (optionally with the enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer, see, e.g., Boshart et al, Cell, 41:521-530 (1985)), the CASI promoter, the SV40 promoter, the dihydrofolate reductase promoter, the β -actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1a promoter (Invitrogen).

10 In some embodiments, the promoter is a CASI promoter (see, for example, WO2012/115980). The CASI promoter is a synthetic promoter which contains a portion of the CMV enhancer, a portion of the chicken beta-actin promoter, and a portion of the UBC enhancer. In some embodiments, the CASI promoter can include a nucleic acid sequence having at least about
15 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or more, sequence identity to SEQ ID NO: 12. In some embodiments, the promoter comprises or consists of a nucleic acid sequence of SEQ ID NO: 12.

Inducible promoters allow regulation of gene expression and can be regulated by exogenously
20 supplied compounds, environmental factors such as temperature, or the presence of a specific physiological state, e.g., acute phase, a particular differentiation state of the cell, or in replicating cells only. Inducible promoters and inducible systems are available from a variety of commercial sources, including, without limitation, Invitrogen, Clontech and Ariad. Many other systems have been described and can be readily selected by one of skill in the art. For
25 example, inducible promoters include the zinc-inducible sheep metallothionine (MT) promoter and the dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter. Other inducible systems include the T7 polymerase promoter system (WO 98/10088); the ecdysone insect promoter (No et al, Proc. Natl. Acad. Sci. USA, 93:3346-3351 (1996)), the tetracycline-repressible system (Gossen et al, Proc. Natl. Acad. Sci. USA, 89:5547-5551
30 (1992)), the tetracycline-inducible system (Gossen et al, Science, 378:1766-1769 (1995), see also Harvey et al, Curr. Opin. Chem. Biol, 2:512-518 (1998)). Other systems include the FK506 dimer, VP16 or p65 using castradiol, diphenol murislerone, the RU486-inducible system (Wang et al, Nat. Biotech., 15:239-243 (1997) and Wang et al, Gene Ther., 4:432-441 (1997)) and the rapamycin-inducible system (Magari et al, J. Clin. Invest., 100:2865-2872 (1997)). The
35 effectiveness of some inducible promoters increases over time. In such cases one can enhance the effectiveness of such systems by inserting multiple repressors in tandem, e.g., TetR linked to a TetR by an IRES.

40 In another embodiment, the native promoter for the transgene will be used. The native promoter may be preferred when it is desired that expression of the transgene should mimic the native expression. The native promoter may be used when expression of the transgene must be regulated temporally or developmentally, or in a tissue-specific manner, or in response to specific transcriptional stimuli. In a further embodiment, other native expression

control elements, such as enhancer elements, polyadenylation sites or Kozak consensus sequences may also be used to mimic the native expression.

The transgene may be operably linked to a tissue-specific promoter. For instance, if expression in skeletal muscle is desired, a promoter active in muscle should be used. These include the promoters from genes encoding skeletal β -actin, myosin light chain 2A, dystrophin, muscle creatine kinase, as well as synthetic muscle promoters with activities higher than naturally occurring promoters (see Li et al, Nat. Biotech., 17:241-245 (1999)). Examples of promoters that are tissue-specific are known for liver (albumin, Miyatake et al, J. Virol, 71:5124-32 (1997); hepatitis B virus core promoter, Sandig et al, Gene Ther., 3:1002-9 (1996); alpha-fetoprotein (AFP), Arbuthnot et al., Hum. Gene Ther., 7: 1503-14 (1996)), bone osteocalcin (Stein et al, Mol. Biol. Rep., 24:185-96 (1997)); bone sialoprotein (Chen et al., J. Bone Miner. Res., 11:654-64 (1996)), lymphocytes (CD2, Hansal et al, J. Immunol, 161:1063-8 (1998); immunoglobulin heavy chain; T cell receptor chain), neuronal such as neuron-specific enolase (NSE) promoter (Andersen et al, Cell. Mol. Neurobiol, 13:503-15 (1993)), neurofilament light-chain gene (Piccioli et al, Proc. Natl. Acad. Sci. USA, 88:5611-5 (1991)), and the neuron-specific vgf gene (Piccioli et al, Neuron, 15:373-84 (1995)), among others.

Optionally, vectors carrying transgenes encoding therapeutically useful or immunogenic products may also include selectable markers or reporter genes which may include sequences encoding geneticin, hygromycin or purimycin resistance, among others. Such selectable reporters or marker genes (preferably located outside the viral genome to be packaged into a viral particle) can be used to signal the presence of the plasmids in bacterial cells, such as ampicillin resistance. Other components of the vector may include an origin of replication.

These vectors are generated using the techniques and sequences provided herein, in conjunction with techniques known to those of skill in the art. Such techniques include conventional cloning techniques of cDNA such as those described in texts, use of overlapping oligonucleotide sequences of the adenovirus genomes, polymerase chain reaction, and any suitable method which provides the desired nucleotide sequence.

Therapeutics and Prophylaxis

The recombinant ChAd155-based vectors are useful for gene transfer to a human or non-simian mammal in vitro, ex vivo, and in vivo.

The recombinant adenovirus vectors described herein can be used as expression vectors for the production of the products encoded by the heterologous transgenes in vitro. For example, the recombinant replication-incompetent adenovirus containing a transgene may be transfected into a complementation cell line as described above.

A ChAd155-derived recombinant adenoviral vector provides an efficient gene transfer vehicle that can deliver a selected transgene to a selected host cell in vivo or ex vivo even where the

organism has neutralizing antibodies to one or more adenovirus serotypes. In one embodiment, the vector and the cells are mixed ex vivo; the infected cells are cultured using conventional methodologies; and the transduced cells are re-infused into the patient. These techniques are particularly well suited to gene delivery for therapeutic purposes and for immunisation, including inducing protective immunity.

Immunogenic Transgenes

The recombinant ChAd155 vectors may also be administered in immunogenic compositions. An immunogenic composition as described herein is a composition comprising one or more recombinant ChAd155 vector capable of inducing an immune response, for example a humoral (e.g., antibody) and/or cell-mediated (e.g., a cytotoxic T cell) response, against a transgene product delivered by the vector following delivery to a mammal, suitably a human. A recombinant adenovirus may comprise (suitably in any of its gene deletions) a gene encoding a desired immunogen and may therefore be used in a vaccine.

Such vaccine or other immunogenic compositions may be formulated in a suitable delivery vehicle. Generally, doses for the immunogenic compositions are in the range defined below under 'Delivery Methods and Dosage'. The levels of immunity of the selected gene can be monitored to determine the need, if any, for boosters. Following an assessment of antibody titers in the serum, optional booster immunizations may be desired.

Optionally, a vaccine or immunogenic composition of the invention may be formulated to contain other components, including, e.g., adjuvants, stabilizers, pH adjusters, preservatives and the like. Examples of suitable adjuvants are provided below under 'Adjuvants'. Such an adjuvant can be administered with a priming DNA vaccine encoding an antigen to enhance the antigen-specific immune response compared with the immune response generated upon priming with a DNA vaccine encoding the antigen only. Alternatively, such an adjuvant can be administered with a polypeptide antigen which is administered in an administration regimen involving the ChAd155 vectors of the invention (as described below under 'Administration Regimens').

The recombinant adenoviruses are administered in an immunogenic amount, that is, an amount of recombinant adenovirus that is effective in a route of administration to transfect the desired target cells and provide sufficient levels of expression of the selected gene to induce an immune response. Where protective immunity is provided, the recombinant adenoviruses are considered to be vaccine compositions useful in preventing infection and/or recurrent disease.

The recombinant vectors described herein are expected to be highly efficacious at inducing cytolytic T cells and antibodies directed to the inserted heterologous antigenic protein expressed by the vector.

Adjuvants

An "adjuvant" as used herein refers to a composition that enhances the immune response to an immunogen. Examples of such adjuvants include but are not limited to inorganic adjuvants (e.g. inorganic metal salts such as aluminium phosphate or aluminium hydroxide), organic adjuvants (e.g. saponins, such as QS21, or squalene), oil-based adjuvants (e.g. Freund's complete adjuvant and Freund's incomplete adjuvant), cytokines (e.g. IL-1 β , IL-2, IL-7, IL-12, IL-18, GM-CSF, and INF- γ) particulate adjuvants (e.g. immuno-stimulatory complexes (ISCOMS), liposomes, or biodegradable microspheres), virosomes, bacterial adjuvants (e.g. monophosphoryl lipid A, such as 3-de-O-acylated monophosphoryl lipid A (3D-MPL), or muramyl peptides), synthetic adjuvants (e.g. non-ionic block copolymers, muramyl peptide analogues, or synthetic lipid A), synthetic polynucleotides adjuvants (e.g. polyarginine or polylysine) and immunostimulatory oligonucleotides containing unmethylated CpG dinucleotides ("CpG").

One suitable adjuvant is monophosphoryl lipid A (MPL), in particular 3-de-O-acylated monophosphoryl lipid A (3D-MPL). Chemically it is often supplied as a mixture of 3-de-O-acylated monophosphoryl lipid A with either 4, 5, or 6 acylated chains. It can be purified and prepared by the methods taught in GB 2122204B, which reference also discloses the preparation of diphosphoryl lipid A, and 3-O-deacylated variants thereof. Other purified and synthetic lipopolysaccharides have been described (U.S. Pat. No. 6,005,099 and EP 0 729 473 B1; Hilgers et al., 1986, *Int.Arch.Allergy.Immunol.*, 79(4):392-6; Hilgers et al., 1987, *Immunology*, 60(1):141-6; and EP 0 549 074 B1).

Saponins are also suitable adjuvants (see Lacaille-Dubois, M and Wagner H, A review of the biological and pharmacological activities of saponins. *Phytomedicine* vol 2 pp 363-386 (1996)). For example, the saponin Quil A (derived from the bark of the South American tree *Quillaja Saponaria* Molina), and fractions thereof, are described in U.S. Pat. No. 5,057,540 and Kensil, *Crit. Rev. Ther. Drug Carrier Syst.*, 1996, 12:1-55; and EP 0 362 279 B1. Purified fractions of Quil A are also known as immunostimulants, such as QS21 and QS17; methods of their production is disclosed in U.S. Pat. No. 5,057,540 and EP 0 362 279 B1. Also described in these references is QS7 (a non-haemolytic fraction of Quil-A). Use of QS21 is further described in Kensil et al. (1991, *J. Immunology*, 146: 431-437). Combinations of QS21 and polysorbate or cyclodextrin are also known (WO 99/10008). Particulate adjuvant systems comprising fractions of QuilA, such as QS21 and QS7 are described in WO 96/33739 and WO 96/11711.

Another adjuvant is an immunostimulatory oligonucleotide containing unmethylated CpG dinucleotides ("CpG") (Krieg, *Nature* 374:546 (1995)). CpG is an abbreviation for cytosine-guanosine dinucleotide motifs present in DNA. CpG is known as an adjuvant when administered by both systemic and mucosal routes (WO 96/02555, EP 468520, Davis et al, *J.Immunol*, 1998, 160:870-876; McCluskie and Davis, *J.Immunol.*, 1998, 161:4463-6). CpG, when formulated into vaccines, may be administered in free solution together with free antigen

(WO 96/02555) or covalently conjugated to an antigen (WO 98/16247), or formulated with a carrier such as aluminium hydroxide (Brazolot-Millan et al., Proc. Natl. Acad. Sci., USA, 1998, 95:15553-8).

- 5 Adjuvants such as those described above may be formulated together with carriers, such as liposomes, oil in water emulsions, and/or metallic salts (including aluminum salts such as aluminum hydroxide). For example, 3D-MPL may be formulated with aluminum hydroxide (EP 0 689 454) or oil in water emulsions (WO 95/17210); QS21 may be formulated with cholesterol containing liposomes (WO 96/33739), oil in water emulsions (WO 95/17210) or alum (WO 10 98/15287); CpG may be formulated with alum (Brazolot-Millan, supra) or with other cationic carriers.

Combinations of adjuvants may be utilized in the present invention, in particular a combination of a monophosphoryl lipid A and a saponin derivative (see, e.g., WO 94/00153; WO 95/17210; 15 WO 96/33739; WO 98/56414; WO 99/12565; WO 99/11241), more particularly the combination of QS21 and 3D-MPL as disclosed in WO 94/00153, or a composition where the QS21 is quenched in cholesterol-containing liposomes (DQ) as disclosed in WO 96/33739. Alternatively, a combination of CpG plus a saponin such as QS21 is an adjuvant suitable for use in the present invention. A potent adjuvant formulation involving QS21, 3D-MPL & 20 tocopherol in an oil in water emulsion is described in WO 95/17210 and is another formulation for use in the present invention. Saponin adjuvants may be formulated in a liposome and combined with an immunostimulatory oligonucleotide. Thus, suitable adjuvant systems include, for example, a combination of monophosphoryl lipid A, preferably 3D-MPL, together with an aluminium salt (e.g. as described in WO00/23105). A further exemplary adjuvant 25 comprises comprises QS21 and/or MPL and/or CpG. QS21 may be quenched in cholesterol-containing liposomes as disclosed in WO 96/33739.

Other suitable adjuvants include alkyl Glucosaminide phosphates (AGPs) such as those disclosed in WO9850399 or U.S. Pat. No. 6,303,347 (processes for preparation of AGPs are 30 also disclosed), or pharmaceutically acceptable salts of AGPs as disclosed in U.S. Pat. No. 6,764,840. Some AGPs are TLR4 agonists, and some are TLR4 antagonists. Both are thought to be useful as adjuvants.

It has been found (WO 2007/062656, which published as US 2011/0293704 and is 35 incorporated by reference for the purpose of disclosing invariant chain sequences) that the fusion of the invariant chain to an antigen which is comprised by an expression system used for vaccination increases the immune response against said antigen, if it is administered with an adenovirus. Accordingly, in one embodiment of the invention, the immunogenic transgene may be co-expressed with invariant chain in a recombinant ChAd155 viral vector.

40 In another embodiment, the invention provides the use of the capsid of ChAd155 (optionally an intact or recombinant viral particle or an empty capsid is used) to induce an immunomodulatory effect response, or to enhance or adjuvant a cytotoxic T cell response to another active agent

by delivering a ChAd155 capsid to a subject. The ChAd155 capsid can be delivered alone or in a combination regimen with an active agent to enhance the immune response thereto. Advantageously, the desired effect can be accomplished without infecting the host with an adenovirus.

5

Administration Regimens

Commonly, the ChAd155 recombinant adenoviral vectors will be utilized for delivery of therapeutic or immunogenic molecules (such as proteins). It will be readily understood for both applications, that the recombinant adenoviral vectors of the invention are particularly well suited for use in regimens involving repeat delivery of recombinant adenoviral vectors. Such regimens typically involve delivery of a series of viral vectors in which the viral capsids are alternated. The viral capsids may be changed for each subsequent administration, or after a pre-selected number of administrations of a particular serotype capsid (e.g. one, two, three, four or more). Thus, a regimen may involve delivery of a recombinant adenovirus with a first capsid, delivery with a recombinant adenovirus with a second capsid, and delivery with a recombinant adenovirus with a third capsid. A variety of other regimens which use the adenovirus capsids of the invention alone, in combination with one another, or in combination with other adenoviruses (which are preferably immunologically non-crossreactive) will be apparent to those of skill in the art. Optionally, such a regimen may involve administration of recombinant adenovirus with capsids of other non-human primate adenoviruses, human adenoviruses, or artificial sequences such as are described herein.

The adenoviral vectors of the invention are particularly well suited for therapeutic regimens in which multiple adenoviral-mediated deliveries of transgenes are desired, e.g., in regimens involving redelivery of the same transgene or in combination regimens involving delivery of other transgenes. Such regimens may involve administration of a ChAd155 adenoviral vector, followed by re-administration with a vector from the same serotype adenovirus. Particularly desirable regimens involve administration of a ChAd155 adenoviral vector, in which the source of the adenoviral capsid sequences of the vector delivered in the first administration differs from the source of adenoviral capsid sequences of the viral vector utilized in one or more of the subsequent administrations. For example, a therapeutic regimen involves administration of a ChAd155 vector and repeat administration with one or more adenoviral vectors of the same or different serotypes.

35

In another example, a therapeutic regimen involves administration of an adenoviral vector followed by repeat administration with a ChAd155 vector which has a capsid which differs from the source of the capsid in the first delivered adenoviral vector, and optionally further administration with another vector which is the same or, preferably, differs from the source of the adenoviral capsid of the vector in the prior administration steps. These regimens are not limited to delivery of adenoviral vectors constructed using the ChAd155 sequences. Rather, these regimens can readily utilize other adenoviral sequences, including, without limitation,

40

other adenoviral sequences including other non-human primate adenoviral sequences, or human adenoviral sequences, in combination with the ChAd155 vectors.

In a further example, a therapeutic regimen may involve either simultaneous (such as co-administration) or sequential (such as a prime-boost) delivery of (i) one or more ChAd155 adenoviral vectors and (ii) a further component such as non-adenoviral vectors, non-viral vectors, and/or a variety of other therapeutically useful compounds or molecules such as antigenic proteins optionally simultaneously administered with adjuvant. Examples of co-administration include homo-lateral co-administration and contra-lateral co-administration (further described below under 'Delivery Methods and Dosage').

Suitable non-adenoviral vectors for use in simultaneous or particularly in sequential delivery (such as prime-boost) with one or more ChAd155 adenoviral vectors include one or more poxviral vectors. Suitably, the poxviral vector belongs to the subfamily chordopoxvirinae, more suitably to a genus in said subfamily selected from the group consisting of orthopox, parapox, yatapox, avipox (suitably canarypox (ALVAC) or fowlpox (FPV)) and molluscipox. Even more suitably, the poxviral vector belongs to the orthopox and is selected from the group consisting of vaccinia virus, NYVAC (derived from the Copenhagen strain of vaccinia), Modified Vaccinia Ankara (MVA), cowpoxvirus and monkeypox virus. Most suitably, the poxviral vector is MVA.

"Simultaneous" administration suitably refers to the same ongoing immune response. Preferably both components are administered at the same time (such as simultaneous administration of both DNA and protein), however, one component could be administered within a few minutes (for example, at the same medical appointment or doctor's visit), within a few hours. Such administration is also referred to as co-administration. In some embodiments, co-administration may refer to the administration of an adenoviral vector, an adjuvant and a protein component. In other embodiments, co-administration refers to the administration of an adenoviral vector and another viral vector, for example a second adenoviral vector or a poxvirus such as MVA. In other embodiments, co-administration refers to the administration of an adenoviral vector and a protein component, which is optionally adjuvanted.

A prime-boost regimen may be used. Prime-boost refers to two separate immune responses: (i) an initial priming of the immune system followed by (ii) a secondary or boosting of the immune system many weeks or months after the primary immune response has been established.

Such a regimen may involve the administration of a recombinant ChAd155 vector to prime the immune system to second, booster, administration with a traditional antigen, such as a protein (optionally co-administered with adjuvant), or a recombinant virus carrying the sequences encoding such an antigen (e.g., WO 00/11140). Alternatively, an immunization regimen may involve the administration of a recombinant ChAd155 vector to boost the immune response to a vector (either viral or DNA-based) encoding an antigen. In another alternative, an immunization regimen involves administration of a protein followed by booster with a

recombinant ChAd155 vector encoding the antigen. In one example, the prime-boost regimen can provide a protective immune response to the virus, bacteria or other organism from which the antigen is derived. In another embodiment, the prime-boost regimen provides a therapeutic effect that can be measured using conventional assays for detection of the presence of the condition for which therapy is being administered.

Preferably, a boosting composition is administered about 2 to about 27 weeks after administering the priming composition to the subject. The administration of the boosting composition is accomplished using an effective amount of a boosting composition containing or capable of delivering the same antigen or a different antigen as administered by the priming vaccine. The boosting composition may be composed of a recombinant viral vector derived from the same viral source or from another source. Alternatively, the boosting composition can be a composition containing the same antigen as encoded in the priming vaccine, but in the form of a protein, which composition induces an immune response in the host. The primary requirements of the boosting composition are that the antigen of the composition is the same antigen, or a cross-reactive antigen, as that encoded by the priming composition.

Delivery Methods and Dosage

The vector may be prepared for administration by being suspended or dissolved in a pharmaceutically or physiologically acceptable carrier such as isotonic saline; isotonic salts solution or other formulations that will be apparent to those skilled in the art. The appropriate carrier will be evident to those skilled in the art and will depend in large part upon the route of administration. The compositions described herein may be administered to a mammal in a sustained release formulation using a biodegradable biocompatible polymer, or by on-site delivery using micelles, gels and liposomes.

In some embodiments, the recombinant adenovirus of the invention is administered to a subject by intramuscular injection, intravaginal injection, intravenous injection, intraperitoneal injection, subcutaneous injection, epicutaneous administration, intradermal administration, nasal administration or oral administration. Of particular interest in the context of tuberculosis are intramuscular injection, subcutaneous injection, intradermal administration, nasal administration or aerosol administration, especially intramuscular injection, nasal administration or aerosol administration.

If the therapeutic regimen involves co-administration of one or more ChAd155 adenoviral vectors and a further component, each formulated in different compositions, they are favourably administered co-locally at or near the same site. For example, the components can be administered (e.g. via an administration route selected from intramuscular, transdermal, intradermal, sub-cutaneous) to the same side or extremity ("co-lateral" administration) or to opposite sides or extremities ("contra-lateral" administration).

Dosages of the viral vector will depend primarily on factors such as the condition being treated, the age, weight and health of the patient, and may thus vary among patients. For example, a therapeutically effective adult human or veterinary dosage of the viral vector generally contains 1×10^5 to 1×10^{15} viral particles, such as from 1×10^8 to 1×10^{12} (e.g., 1×10^8 , 2.5×10^8 , 5×10^8 , 1×10^9 , 1.5×10^9 , 2.5×10^9 , 5×10^9 , 1×10^{10} , 1.5×10^{10} , 2.5×10^{10} , 5×10^{10} , 1×10^{11} , 1.5×10^{11} , 2.5×10^{11} , 5×10^{11} , 1×10^{12} particles). Alternatively, a viral vector can be administered at a dose that is typically from 1×10^5 to 1×10^{10} plaque forming units (PFU), such as 1×10^5 PFU, 2.5×10^5 PFU, 5×10^5 PFU, 1×10^6 PFU, 2.5×10^6 PFU, 5×10^6 PFU, 1×10^7 PFU, 2.5×10^7 PFU, 5×10^7 PFU, 1×10^8 PFU, 2.5×10^8 PFU, 5×10^8 PFU, 1×10^9 PFU, 2.5×10^9 PFU, 5×10^9 PFU, or 1×10^{10} PFU. Dosages will vary depending upon the size of the animal and the route of administration. For example, a suitable human or veterinary dosage (for about an 80 kg animal) for intramuscular injection is in the range of about 1×10^9 to about 5×10^{12} particles per mL, for a single site. Optionally, multiple sites of administration may be used. In another example, a suitable human or veterinary dosage may be in the range of about 1×10^{11} to about 1×10^{15} particles for an oral formulation.

The viral vector can be quantified by Quantitative PCR Analysis (Q-PCR), for example with primers and probe designed on CMV promoter region using as standard curve serial dilution of plasmid DNA containing the vector genome with expression cassette including HCMV promoter. The copy number in the test sample is determined by the parallel line analysis method. Alternative methods for vector particle quantification can be analytical HPLC or spectrophotometric method based on A_{260} nm.

An immunologically effective amount of a nucleic acid may suitably be between 1 ng and 100 mg. For example, a suitable amount can be from 1 μ g to 100 mg. An appropriate amount of the particular nucleic acid (e.g., vector) can readily be determined by those of skill in the art. Exemplary effective amounts of a nucleic acid component can be between 1 ng and 100 μ g, such as between 1 ng and 1 μ g (e.g., 100 ng-1 μ g), or between 1 μ g and 100 μ g, such as 10 ng, 50 ng, 100 ng, 150 ng, 200 ng, 250 ng, 500 ng, 750 ng, or 1 μ g. Effective amounts of a nucleic acid can also include from 1 μ g to 500 μ g, such as between 1 μ g and 200 μ g, such as between 10 and 100 μ g, for example 1 μ g, 2 μ g, 5 μ g, 10 μ g, 20 μ g, 50 μ g, 75 μ g, 100 μ g, 150 μ g, or 200 μ g. Alternatively, an exemplary effective amount of a nucleic acid can be between 100 μ g and 1 mg, such as from 100 μ g to 500 μ g, for example, 100 μ g, 150 μ g, 200 μ g, 250 μ g, 300 μ g, 400 μ g, 500 μ g, 600 μ g, 700 μ g, 800 μ g, 900 μ g or 1 mg.

Generally a human dose will be in a volume of between 0.1 ml and 2 ml. Thus the composition described herein can be formulated in a volume of, for example 0.1, 0.15, 0.2, 0.5, 1.0, 1.5 or 2.0 ml human dose per individual or combined immunogenic components.

One of skill in the art may adjust these doses, depending on the route of administration and the therapeutic or vaccine application for which the recombinant vector is employed. The levels of expression of the transgene, or for an adjuvant, the level of circulating antibody, can be monitored to determine the frequency of dosage administration.

If one or more priming and/or boosting steps are used, this step may include a single dose that is administered hourly, daily, weekly or monthly, or yearly. As an example, mammals may receive one or two doses containing between about 10 µg to about 50 µg of plasmid in carrier.

5 The amount or site of delivery is desirably selected based upon the identity and condition of the mammal.

The therapeutic levels of, or level of immune response against, the protein encoded by the selected transgene can be monitored to determine the need, if any, for boosters. Following an
10 assessment of CD8+ T cell response, or optionally, antibody titers, in the serum, optional booster immunizations may be desired. Optionally, the recombinant ChAd155 vectors may be delivered in a single administration or in various combination regimens, e.g., in combination with a regimen or course of treatment involving other active ingredients or in a prime-boost regimen.

15 The present invention will now be further described by means of the following non-limiting examples.

20 EXAMPLES

Example 1: Isolation of ChAd155

Wild type chimpanzee adenovirus type 155 (ChAd155) was isolated from a healthy young chimpanzee housed at the New Iberia Research Center facility (New Iberia Research Center;
25 The University of Louisiana at Lafayette) using standard procedures as described in Colloca *et al.* (2012) and WO2010086189, which is hereby incorporated by reference for the purpose of describing adenoviral isolation and characterization techniques

Example 2: ChAd155 vector construction

30 The ChAd155 viral genome was then cloned in a plasmid or in a BAC vector and subsequently modified (Figure 2) to carry the following modifications in different regions of the ChAd155 viral genome:

- 35 a) deletion of the E1 region (from bp 449 to bp 3529) of the viral genome;
b) deletion of the E4 region (from bp 34731 to bp 37449) of the viral genome;
c) insertion of the E4orf6 derived from human Ad5.

2.1: Deletion of E1 region: Construction of BAC/ChAd155 ΔE1 TetO hCMV RpsL-Kana#1375

40 The ChAd155 viral genome was cloned into a BAC vector by homologous recombination in *E. coli* strain BJ5183 electroporation competent cells (Stratagene catalog no. 2000154) co-transformed with ChAd155 viral DNA and Subgroup C BAC Shuttle (#1365). As shown in the

schematic of Figure 3, the Subgroup C Shuttle is a BAC vector derived from pBeloBAC11 (GenBank U51113, NEB) and which is dedicated to the cloning of ChAd belonging to species C and therefore contains the pIX gene and DNA fragments derived from right and left ends (including right and left ITRs) of species C ChAd viruses.

5

The Species C BAC Shuttle also contains a RpsL-Kana cassette inserted between left end and the pIX gene. In addition, an Amp-LacZ-SacB selection cassette, flanked by IScel restriction sites, is present between the pIX gene and right end of the viral genome. In particular, the BAC Shuttle comprised the following features: Left ITR: bp 27 to 139, hCMV(tetO) RpsL-Kana cassette: bp 493 to 3396, pIX gene: bp 3508 to 3972, IScel restriction sites: bp 3990 and 7481, Amp-LacZ-SacB selection cassette: bp 4000 to 7471, Right ITR: bp 7805 to 7917.

10

15

BJ5183 cells were co-transformed by electroporation with ChAd155 purified viral DNA and Subgroup C BAC Shuttle vector digested with IScel restriction enzyme and then purified from gel. Homologous recombination occurring between pIX gene and right ITR sequences (present at the ends of Species C BAC Shuttle linearized DNA) and homologous sequences present in ChAd155 viral DNA lead to the insertion of ChAd155 viral genomic DNA in the BAC shuttle vector. At the same time, the viral E1 region was deleted and substituted by the RpsL-Kana cassette, generating BAC/ChAd155 Δ E1/ TetO hCMV RpsL-Kana#1375.

20

2.2: Plasmid construction by homologous recombination in E.coli BJ5183

2.2.1: Deletion of E4 region - Construction of pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV RpsL-Kana (#1434)

25

30

To improve propagation of the vector, a deletion of the E4 region spanning from nucleotide 34731-37449 (ChAd155 wild type sequence) was introduced in the vector backbone by replacing the native E4 region with Ad5 E4orf6 coding sequence using a strategy involving several steps of cloning and homologous recombination in *E.coli*. The E4 coding region was completely deleted while the E4 native promoter and polyadenylation signal were conserved. To this end, a shuttle vector was constructed to allow the insertion of Ad5orf6 by replacing the ChAd155 native E4 region by homologous recombination in *E.coli* BJ5183 as detailed below.

Construction of pARS SpeciesC Ad5E4orf6-1

35

40

A DNA fragment containing Ad5orf6 was obtained by PCR using Ad5 DNA as template, with the oligonucleotides 5'-ATACGGACTAGTGGAGAAGTACTCGCCTACATG-3' (SEQ ID NO: 13) and 5'-ATACGGAAGATCTAAGACTTCAGGAAATATGACTAC-3' (SEQ ID NO: 14). The PCR fragment was digested with BglII and SpeI and cloned into Species C RLD-EGFP shuttle digested with BglII and SpeI, generating the plasmid pARS Species C Ad5orf6-1. Details regarding the shuttle can be found in Colloca *et al*, Sci. Transl. Med. (2012) 4:115ra.

Construction of pARS SpeciesC Ad5E4orf6-2

To delete the E4 region, a 177 bp DNA fragment spanning bp 34586 to bp 34730 of the ChAd155 wt sequence (SEQ ID NO: 10) was amplified by PCR using the plasmid BAC/ChAd155 Δ E1_TetO hCMV RpsL-Kana (#1375) as a template with the following
 5 oligonucleotides: 5'-ATTCAGTGTACAGGCGCGCCAAAGCATGACGCTGTTGATTTGATTC-3' (SEQ ID NO: 15) and 5'-ACTAGGACTAGTTATAAGCTAGAATGGGGCTTTGC-3' (SEQ ID NO: 16). The PCR fragment was digested with BsrGI and SpeI and cloned into pARS SubGroupC Ad5orf6-1 digested with BsrGI and SpeI, generating the plasmid pARS SpeciesC Ad5orf6-2 (#1490). A schematic diagram of this shuttle plasmid is provided in Figure 4. In
 10 particular, the shuttle plasmid comprised the following features: Left ITR: bp 1 to 113, Species C first 460bp: bp 1 to 460, ChAd155 wt (bp 34587 to bp 34724 of SEQ ID NO:10) : bp 516 to 650, Ad5 orf6: bp 680 and 1561, Species C last 393 bp: bp 1567 to 1969, Right ITR: bp 1857 to 1969.

15 *Construction of pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV RpsL-Kana (#1434)*

The resulting plasmid pARS SubGroupC Ad5orf6-2 was then used to replace the E4 region within the ChAd155 backbone with Ad5orf6. To this end the plasmid BAC/ChAd155 Δ E1_TetO hCMV RpsL-Kana (#1375) was digested with PacI/PmeI and co-transformed into BJ5183 cells
 20 with the digested plasmid pARS SubGroupC Ad5orf6-2 BsrGI/Ascl, to obtain the pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV RpsL-Kana (#1434) pre-adeno plasmid.

25 *2.2.2: Insertion of RSV Expression Cassette - Construction of pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV RSV*

An RSV cassette was cloned into a linearised pre-adeno acceptor vector via homologous recombination in *E.coli* by exploiting the homology existing between HCMV promoter and BGH polyA sequences. The plasmid pvjTetOhCMV-bghpolyA_RSV was cleaved with SfiI and SpeI to excise the 4,65 Kb fragment containing the HCMV promoter with tetO, RSV and BGHpolyA
 30 sequence. The resulting RSV 4,65 Kb fragment was cloned by homologous recombination into the pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV RpsL-Kana (#1434) acceptor vector carrying the RpsL-Kana selection cassette under control of HCMV and BGHpA. The acceptor pre-adeno plasmid was linearized with the restriction endonuclease SnaBI. The resulting construct was the pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV RSV vector (Figure 5).

35 2.3: BAC vector construction by recombineering

40 *2.3.1: Deletion of E4 region - Construction of BAC/ChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV RpsL-Kana#1390*

A deletion of the E4 region spanning from nucleotide 34731-37449 of the ChAd155 wt sequence was introduced in the vector backbone by replacing this native E4 region with the

Ad5 E4orf6 coding sequence using a strategy involving two different steps of recombineering in *E. Coli* SW102 competent cells.

5 The first step resulted in insertion of a selection cassette including the suicide gene SacB, ampicillin–R gene and lacZ (Amp-LacZ-SacB selection cassette) in the E4 region of ChAd155, for the purpose of positive/negative selection of recombinants.

First step – Substitution of ChAd155 native E4 region with Amp-LacZ-SacB selection cassette

10 The Amp-LacZ-SacB selection cassette was amplified by PCR using the oligonucleotides provided below containing E4 flanking sequences to allow homologous recombination: 1021-FW E4 Del Step1 (5'-TTAATAGACACAGTAGCTTAATAGACCCAGTAGTGCAAAGCCCCATTCTAGCTTATAA
CCCCTATTTGTTTATTTTCT-3') (SEQ ID NO: 17) and 1022-RW E4 Del Step1
15 (5'-ATATATACTCTCTCGGCACTTGGCCTTTTACACTGCGAAGTGTTGGTGCTGGTGCTGCGTTGAGAGATCTTTATTTGTAACTGTAAATTGTC-3') (SEQ ID NO: 18).

20 The PCR product was used to transform *E. Coli* SW102 competent cells containing the pAdeno plasmid BAC/ChAd155 (DE1) tetO hCMV - RpsLKana#1375. The transformation of SW102 cells allowed the insertion of the selection cassette in the E4 region of ChAd155 via lambda (λ) Red-mediated homologous recombination, thus obtaining BAC/ChAd155 (DE1) TetOhCMV - RpsL Kana #1379 (including Amp-LacZ-SacB cassette by substituting ChAd155 native E4 region).

25 *Second step – Substitution of Amp-lacZ-SacB selection cassette with Ad5E4orf6 region*

30 The resulting plasmid BAC/ChAd155 (DE1) TetOhCMV - RpsL Kana #1379 (with Amp-LacZ-SacB cassette in place of ChAd155 E4 region) was then manipulated to replace the Amp-lacZ-SacB selection cassette with Ad5orf6 within the ChAd155 backbone. To this end, a DNA fragment containing the Ad5orf6 region was obtained by PCR, using the oligonucleotides 1025-FW E4 Del Step2 (5'-TTAATAGACACAGTAGCTTAATA-3') (SEQ ID NO: 19) and 1026-RW E4 Del Step2 (5'-GGAAGGGAGTGTCTAGTGTT-3') (SEQ ID NO: 20). The resulting DNA fragment was introduced into *E. coli* SW102 competent cells containing the pAdeno plasmid BAC/ChAd155 (DE1) TetOhCMV - RpsL Kana)#1379, resulting in a final plasmid
35 BAC/ChAd155 (Δ E1, E4 Ad5E4orf6) TetOhCMV - RpsL Kana#1390 containing Ad5orf6 substituting the native ChAd155 E4 region.

2.3.2: Insertion of RSV Expression Cassette: Construction of BAC/ChAd155 Δ E1, E4_Ad5E4orf6 /TetOhCMV RSV#1393

40

An RSV transgene was cloned into the BAC/ChAd155 Δ E1, E4_Ad5E4orf6 /TetOhCMV RSV#1393 vector by substituting the RpsL-Kana selection cassette. The construction strategy was based on two different steps of recombineering in *E. Coli* SW102 competent cells.

First step – Substitution of RpsL-Kana cassette with Amp-LacZ-SacB selection cassette

The Amp-LacZ-SacB selection cassette was obtained from plasmid BAC/ChAd155 (DE1) TetO
 5 hCMV Amp-LacZ-SacB#1342 by PCR using the oligonucleotides 91-SubMonte FW (5'-
 CAATGGGCGTGGATAGCGGTTTGAC-3') (SEQ ID NO: 21) and 890-BghPolyA RW (5'-
 CAGCATGCCTGCTATTGTC-3') (SEQ ID NO: 22). The product was transformed into *E. Coli*
 SW102 competent cells containing the pAdeno plasmid BAC/ChAd155 (DE1, E4 Ad5E4orf6)
 TetOhCMV - RpsL Kana#1390, resulting in BAC/ChAd155 (DE1, E4 Ad5E4orf6) TetOhCMV –
 10 Amp-LacZ-SacB#1386.

Second step – Substitution of Amp-lacZ-SacB selection cassette with RSV transgene

The RSV transgene was inserted in plasmid BAC/ChAd155 (DE1, E4 Ad5E4orf6) TetOhCMV
 15 – Amp-LacZ-SacB#1386 by replacing the Amp-lacZ-SacB selection cassette by homologous
 recombination. To this end, the plasmid pvjTetOhCMV-bghpolyA_RSV#1080 (containing an
 RSV expression cassette) was cleaved with *SpeI* and *SfiI* to excise the 4.4 Kb fragment
 including the HCMV promoter, RSV and BGHpolyA. The resulting RSV 4.4 Kb fragment was
 transformed into *E. Coli* SW102 competent cells containing the pAdeno plasmid BAC/ChAd155
 20 (DE1, E4 Ad5E4orf6) TetOhCMV – Amp-LacZ-SacB#1386, resulting in the final plasmid
 BAC/ChAd155 ΔE1, E4_Ad5E4orf6 / TetO hCMV Kana#1390. The structure of the BAC
 carrying ChAd155/RSV (SEQ ID NO: 11) is illustrated in Figure 6. In particular, ChAd155/RSV
 comprised the following features: Species C Left ITR: bp 1 to 113, hCMV(tetO) bp 467 to
 1311, RSV gene: bp 1348 to 4785, bghpolyA: bp 4815 to 5032, Ad5E4orf6: bp 36270 to
 25 37151, Species C Right ITR: bp 37447 to 37559.

Example 3: Vector production

The productivity of ChAd155 was evaluated in comparison to ChAd3 and PanAd3 in the
 30 Procell 92 cell line.

3.1: Production of vectors comprising an HIV Gag transgene

Vectors expressing the HIV Gag protein were prepared as described above (ChAd155/GAG)
 35 or previously (ChAd3/GAG Colloca *et al*, Sci. Transl. Med. (2012) 4:115ra). ChAd3/GAG and
 ChAd155/GAG were rescued and amplified in Procell 92 until passages 3 (P3); P3 lysates
 were used to infect 2 T75 flasks of Procell 92 cultivated in monolayer with each vector. A
 multiplicity of infection (MOI) of 100 vp/cell was used for both infection experiments. The
 infected cells were harvested when full CPE was evident (72 hours post-infection) and pooled;
 40 the viruses were released from the infected cells by 3 cycles of freeze/thaw (-70°/37°C) then
 the lysate was clarified by centrifugation. The clarified lysates were quantified by Quantitative
 PCR Analysis with primers and probe complementary to the CMV promoter region. The
 oligonucleotide sequences are the following: CMVfor 5'-

CATCTACGTATTAGTCATCGCTATTACCA-3' (SEQ ID NO: 23), CMVrev 5'-
GACTTGGAAATCCCCGTGAGT-3' (SEQ ID NO: 24), CMVFAM-TAMRA probe 5'-
ACATCAATGGGCGTGGATAGCGGTT-3' (SEQ ID NO: 25) (QPCRs were run on ABI Prism
7900 Sequence detector – Applied Biosystem). The resulting volumetric titers (vp/ml)
5 measured on clarified lysates and the specific productivity expressed in virus particles per cell
(vp/cell) are provided in Table 1 below and illustrated in Figure 7.

Table 1. Vector productivity from P3 lysates.

Vector	vp/ml	Total vp (20 ml conc.)	vp/cell
ChAd3/GAG	9,82E+09	1,96E+11	6,61E+03
ChAd155/GAG	1,11E+10	2,22E+11	7,46E+03

To confirm the higher productivity of the ChAd155 vector expressing HIV Gag transgene, a
second experiment was performed by using purified viruses as inoculum. To this end, Procell
92 cells were seeded in a T25 Flask and infected with ChAd3/GAG and ChAd155/GAG when
the confluence of the cells was about 80%, using a MOI=100 vp/cell of infection. The infected
15 cells were harvested when full CPE was evident; the viruses were released from the infected
cells by freeze/thaw and clarified by centrifugation. The clarified lysates were quantified by
Quantitative PCR Analysis by using following primers and probe: CMVfor 5'-
CATCTACGTATTAGTCATCGCTATTACCA-3' (SEQ ID NO: 23), CMV rev
GACTTGGAAATCCCCGTGAGT (SEQ ID NO: 24), CMV FAM-TAMRA probe 5'-
20 ACATCAATGGGCGTGGATAGCGGTT-3' (SEQ ID NO: 25) complementary to the CMV
promoter region (samples were analysed on an ABI Prism 7900 Sequence detector- Applied
Biosystems). The resulting volumetric titers (vp/ml) measured on clarified lysates and the
specific productivity expressed in virus particles per cell (vp/cell) are provided in Table 2 below
and illustrated in Figure 8.

Table 2. Vector productivity from purified viruses.

Vector	vp/ml	Total vp/T25 flask (5ml of lysate)	vp/cell
ChAd3/GAG	1,00E+10	5,00E+10	1,67E+04
ChAd155/GAG	1,21E+10	6,05E+10	2,02E+04

3.2: Production of vectors comprising an RSV transgene

A different set of experiments were performed to evaluate the productivity of RSV vaccine
vectors in Procell 92.S cultivated in suspension. The experiment compared PanAd3/RSV

(described in WO2012/089833) and ChAd155/RSV in parallel by infecting Procell 92.S at a cell density of 5×10^5 cells/ml. The infected cells were harvested 3 days post infection; the virus was released from the infected cells by 3 cycles of freeze/thaw and the lysate was clarified by centrifugation. The clarified lysates were then quantified by Quantitative PCR Analysis as reported above. The volumetric productivity and the cell specific productivity are provided in Table 3 below and illustrated in Figure 9.

Table 3

Virus	Volumetric productivity (Vp/ml)	Total vp	Cell specific productivity (vp/cell)
PanAd3/RSV	5,82E+09	2,91E+11	1,16E+4
ChAd155/RSV	3,16E+10	1,58E+12	6,31E+04

Example 4: Transgene Expression Levels

4.1: Expression level of HIV Gag transgene

Expression levels were compared in parallel experiments by infecting HeLa cells with ChAd3 and ChAd155 vectors comprising an HIV Gag transgene. HeLa cells were seeded in 24 well plates and infected in duplicate with ChAd3/GAG and ChAd155/GAG purified viruses using a MOI=250 vp/cell. The supernatants of HeLa infected cells were harvested 48 hours post-infection, and the production of secreted HIV GAG protein was quantified by using a commercial ELISA Kit (HIV-1 p24 ELISA Kit, PerkinElmer Life Science). The quantification was performed according to the manufacturer's instruction by using an HIV-1 p24 antigen standard curve. The results, expressed in pg/ml of GAG protein, are illustrated in Figure 10.

4.1: Expression level of RSV F transgene

Expression levels were compared in parallel experiments by infecting HeLa cells with the above-described PanAd3 and ChAd155 vectors comprising an RSV F transgene. To this end, HeLa cells were seeded in 6 well plates and infected in duplicate with PanAd3/RSV and ChAd155/RSV purified viruses using a MOI=500 vp/cell. The supernatants were harvested 48 hours post-infection, and the production of secreted RSV F protein was quantified by ELISA. Five different dilutions of the supernatants were transferred to microplate wells which are coated with a commercial mouse anti-RSV F monoclonal antibody. The captured antigen was revealed using a secondary anti-RSV F rabbit antiserum followed by Biotin-conjugated anti-rabbit IgG, then by adding Streptavidin-AP conjugate (BD Pharmingen cat. 554065). The quantification was performed by using an RSV F protein (Sino Biological cat. 11049-V08B) standard curve. The results obtained, expressed as ug/ml of RSV F protein, are provided in Table 4 below.

Table 4

Sample	µg/ml RSV F protein
ChAd155/RSV	5,9
PanAd3/RSV	4

A western blot analysis was also performed to confirm the higher level of transgene expression provided by the ChAd155 RSV vector relative to the PanAd3 RSV vector. HeLa cells plated in 6 well plates were infected with PanAd3/RSV and ChAd155/RSV purified viruses using MOI=250 and 500 vp/cell. The supernatants of HeLa infected cells were harvested and the production of secreted RSV F protein were analysed by non-reducing SDS gel followed by Western Blot analysis. Equivalent quantities of supernatants were loaded on non-reducing SDS gel; after electrophoresis separation, the proteins were transferred to a nitrocellulose membrane to be probed with an anti-RSV F mouse monoclonal antibody (clone RSV-F-3 catalog no: ABIN308230 available at antibodies-online.com (last accessed 13 April 2015). After the incubation with primary antibody, the membrane was washed and then incubated with anti-mouse HRP conjugate secondary antibody. Finally, the assay was developed by electrochemiluminescence using standard techniques (ECL detection reagents Pierce catalog no W3252282). The Western Blot results are shown in Figure 11. A band of about 170 kD indicated by the arrow was revealed by monoclonal antibody mAb 13 raised against the F protein, which corresponds to the expected weight of trimeric F protein. It can be seen that the ChAd155 RSV vector produced a darker band at both MOI=250 and 500vp/cell.

Example 5: Evaluation of immunological potency by mouse immunization experiments

5.1: Immunogenicity of vectors comprising the HIV Gag transgene

The immunogenicity of the ChAd155/GAG vector was evaluated in parallel with the ChAd3/GAG vector in BALB/c mice (5 per group). The experiment was performed by injecting 10^6 viral particles intramuscularly. T-cell response was measured 3 weeks after the immunization by ex vivo IFN-gamma enzyme-linked immunospot (ELISpot) using a GAG CD8+ T cell epitope mapped in BALB/c mice. The results are shown in Figure 12, expressed as IFN-gamma Spot Forming Cells (SFC) per million of splenocytes. Each dot represents the response in a single mouse, and the line corresponds to the mean for each dose group. Injected dose in number of virus particles and frequency of positive mice to the CD8 immunodominant peptide are shown on the x axis.

5.2 Immunogenicity of vectors comprising the RSV transgene

The immunological potency of the PanAd3/RSV and ChAd155/RSV vectors was evaluated in BALB/c mice. Both vectors were injected intramuscularly at doses of 10^8 , 10^7 and 3×10^6 vp. Three weeks after vaccination the splenocytes of immunized mice were isolated and analyzed

by IFN-gamma-ELISpot using as antigens immunodominant peptide F and M epitopes mapped in BALB/c mice. The levels of immune-responses were reduced in line with decreasing dosage (as expected) but immune responses were clearly higher in the groups of mice immunized with ChAd155/RSV vector compared to the equivalent groups of mice immunized with PanAd3/RSV vaccine (Figure 13). In Figure 13, symbols show individual mouse data, expressed as IFN-gamma Spot Forming Cells (SFC)/million splenocytes, calculated as the sum of responses to the three immunodominant epitopes (F_{51-66} F_{85-93} and $M2-1_{282-290}$) and corrected for background. Horizontal lines represent the mean number of IFN-gamma SFC/million splenocytes for each dose group.

Conclusion

Taken together the results reported above demonstrated that ChAd155 is an improved adenoviral vector in comparison to ChAd3 and PanAd3 vectors. ChAd155 was shown to be more productive therefore facilitating the manufacture process, able to express higher level of transgene in vitro and also in vivo providing a stronger T-cell response against the antigens expressed in animal models.

CLAIMS

1. A recombinant adenovirus comprising at least one polynucleotide or polypeptide selected from the group consisting of:
 - 5 (a) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 1,
 - (b) a polynucleotide which encodes a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1,
 - 10 (c) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 3,
 - (d) a polypeptide having the amino acid sequence according to SEQ ID NO: 1,
 - (e) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1, and
 - 15 (f) a polypeptide having the amino acid sequence according to SEQ ID NO: 3;wherein the adenovirus comprises a nucleic acid sequence encoding a mycobacterial antigen, wherein the nucleic acid sequence is operatively linked to one or more sequences which direct expression of said mycobacterial antigen in a host cell.
- 20 2. A composition comprising the recombinant adenovirus according to claim 1 and a pharmaceutically acceptable excipient.
- 25 3. The adenovirus or composition according to either claim 1 or 2, wherein the polynucleotide encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 1.
- 30 4. The adenovirus or composition according to claim 3, wherein the polynucleotide has a sequence according to SEQ ID NO: 2.
5. The adenovirus or composition according to any one of claims 1 to 4, wherein the polynucleotide encodes a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1.
- 35 6. The adenovirus or composition according to claim 5, wherein the functional derivative has an amino acid sequence which is at least 89.0% identical, such as at least 99.0% identical, such as at least 99.6% identical over its entire length to the amino acid sequence of SEQ ID NO: 1.
- 40

7. The adenovirus or composition according to any one of claims 1 to 6, further comprising a polynucleotide encoding:
- 5 (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 3; or
 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 3, wherein the functional derivative has an amino acid sequence which is at least 50.0% identical over its entire length to the amino acid sequence of SEQ ID NO: 3,
 or
 (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or
 10 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.
- 15 8. The adenovirus or composition according to any one of claims 1 to 6, further comprising a polynucleotide encoding:
- (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 3; or
 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 3, wherein the functional derivative has an amino acid sequence which
 20 is at least 50.0% identical over its entire length to the amino acid sequence of SEQ ID NO: 3,
 and
 (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or
 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which
 25 is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.
- 30 9. The adenovirus or composition according to either claim 7 or 8, wherein the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 3 has an amino acid sequence which is at least 70.0%, such as at least 90.0%, such as at least 99% identical over its entire length to the amino acid sequence of SEQ ID NO: 3.
- 35 10. The adenovirus or composition according to any one of claims 7 to 9, wherein the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5 has an amino acid sequence which is at least 70.0%, such as at least 90.0%, such as at least 99% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.
 40
11. The adenovirus or composition according to either claim 1 or 2, wherein the polynucleotide encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 3.

12. The adenovirus or composition according to claim 11, wherein the polynucleotide has a polynucleotide sequence according to SEQ ID NO: 5.
- 5 13. The adenovirus or composition according to either claim 11 or 12, further comprising a polynucleotide encoding:
- 10 (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 1; or
(b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 1
or
(a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or
15 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.
- 20 14. The adenovirus or composition according to either claim 11 or 12, further comprising a polynucleotide encoding:
- 25 (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 1; or
(b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 1
and
(a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or
30 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.
- 35 15. The adenovirus or composition according to either claim 13 or 14, wherein the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1 has an amino acid sequence which is at least 70.0% identical, such as at least 99.0% identical, such as at least 99.4% identical over its entire length to the amino acid sequence of SEQ ID NO: 1.
- 40 16. The adenovirus or composition according to any one of claims 13 to 15, wherein the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5 has an amino acid sequence which is at least 70.0%, such as at least 90.0%, such as at least 99% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.

17. The adenovirus or composition according to either claim 1 or 2, wherein the polypeptide has the amino acid sequence according to SEQ ID NO: 1.
- 5 18. The adenovirus or composition according to either claim 1 or 2, wherein the polypeptide is a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1.
- 10 19. The adenovirus or composition according to claim 18, wherein the functional derivative has an amino acid sequence which is at least 89.0% identical, such as at least 99.0% identical, such as at least 99.6% identical over its entire length to the amino acid sequence of SEQ ID NO: 1.
- 15 20. The adenovirus or composition according to any one of claims 17 to 19, further comprising:
(a) a polypeptide having the amino acid sequence according to SEQ ID NO: 3; or
(b) a functional derivative of a polypeptide having the amino acid sequence according
20 to SEQ ID NO: 3, wherein the functional derivative has an amino acid sequence which is at least 50.0% identical over its entire length to the amino acid sequence of SEQ ID NO: 3,
or
(a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or
25 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.
- 30 21. The adenovirus or composition according to any one of claims 17 to 19, further comprising:
(a) a polypeptide having the amino acid sequence according to SEQ ID NO: 3; or
(b) a functional derivative of a polypeptide having the amino acid sequence according
35 to SEQ ID NO: 3, wherein the functional derivative has an amino acid sequence which is at least 50.0% identical over its entire length to the amino acid sequence of SEQ ID NO: 3,
and
(a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or
40 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.

22. The adenovirus or composition according to either claim 20 or 21, wherein the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 3 has an amino acid sequence which is at least 70.0%, such as at least 90.0%, such as at least 99% identical over its entire length to the amino acid sequence of SEQ ID NO: 3.
23. The adenovirus or composition according to any one of claims 20 to 22, wherein the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5 has an amino acid sequence which is at least 70.0%, such as at least 90.0%, such as at least 99% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.
24. The adenovirus or composition according to either claim 1 or 2, wherein the polypeptide has the amino acid sequence according to SEQ ID NO: 3.
25. The adenovirus or composition according claim 24, further comprising:
 (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 1; or
 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 1
 or
 (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or
 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.
26. The adenovirus or composition according claim 24, further comprising:
 (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 1; or
 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 1
 and
 (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or
 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.
27. The adenovirus or composition according to either claim 25 or 26, wherein the functional derivative of a polypeptide having the amino acid sequence according to

SEQ ID NO: 1 has an amino acid sequence which is at least 70.0% identical, such as at least 99.0% identical, such as at least 99.4% identical over its entire length to the amino acid sequence of SEQ ID NO: 1.

- 5 28. The adenovirus or composition according to any one of claims 25 to 27, wherein the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5 has an amino acid sequence which is at least 70.0%, such as at least 90.0%, such as at least 99% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.
- 10 29. The adenovirus or composition according to any one of claims 1-28, wherein the polynucleotide comprises at least one of the following:
- (a) an adenoviral 5'-end, preferably an adenoviral 5' inverted terminal repeat;
 - (b) an adenoviral E1A region, or a fragment thereof selected from among the E1A_280R and E1A_243R regions;
 - (c) an adenoviral E1B or IX region, or a fragment thereof selected from among the group consisting of the E1B_19K, E1B_55K or IX regions;
 - (d) an adenoviral E2b region; or a fragment thereof selected from among the group consisting of the E2B_pTP, E2B_Polymerase and E2B_IVa2 regions;
 - 20 (e) an adenoviral L1 region, or a fragment thereof, said fragment encoding an adenoviral protein selected from the group consisting of the L1_13.6k protein, L1_52k and L1_Illa protein;
 - (f) an adenoviral L2 region, or a fragment thereof, said fragment encoding an adenoviral protein selected from the group consisting of the L2_penton protein, L2_pVII, L2_V, and L2_pX protein;
 - 25 (g) an adenoviral L3 region, or a fragment thereof, said fragment encoding an adenoviral protein selected from the group consisting of the L3_pVI protein, L3_hexon protein and L3_protease;
 - (h) an adenoviral E2A region;
 - 30 (i) an adenoviral L4 region, or a fragment thereof said fragment encoding an adenoviral protein selected from the group consisting of the L4_100k protein, the L4_33k protein and protein L4_VIII;
 - (j) an adenoviral E3 region, or a fragment thereof selected from the group consisting of E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8, and E3 ORF9;
 - 35 (k) an adenoviral L5 region, or a fragment thereof said fragment encoding the L5_fiber fiber protein;
 - (l) an adenoviral E4 region, or a fragment thereof selected from the group consisting of E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2, and E4 ORF1;
 - 40 (m) an adenoviral 3'-end, preferably an adenoviral 3' inverted terminal repeat; and/or

- (n) an adenoviral VAI or VAII RNA region, preferably an adenoviral VAI or VAII RNA region from an adenovirus other than ChAd155, more preferably from Ad5.

30. The adenovirus or composition according to claim 29, wherein the polynucleotide comprises at least one of the following:

- (a) an adenoviral 5'-end, preferably an adenoviral 5' inverted terminal repeat;
- (e) an adenoviral L1 region, or a fragment thereof, said fragment encoding an adenoviral protein selected from the group consisting of the L1_13.6k protein, L1_52k and L1_IllA protein;
- (f) an adenoviral L2 region, or a fragment thereof, said fragment encoding an adenoviral protein selected from the group consisting of the L2_penton protein, L2_pVII, L2_V, and L2_pX protein;
- (g) an adenoviral L3 region, or a fragment thereof, said fragment encoding an adenoviral protein selected from the group consisting of the L3_pVI protein, L3_hexon hexon protein and L3_protease;
- (i) an adenoviral L4 region, or a fragment thereof said fragment encoding an adenoviral protein selected from the group consisting of the L4_100k protein, the L4_33k protein and protein L4_VIII;
- (k) an adenoviral L5 region, or a fragment thereof said fragment encoding the L5_fiber fiber protein;
- (m) an adenoviral 3'-end, preferably an adenoviral 3' inverted terminal repeat.

31. The adenovirus or composition according to any one of claims 1-28, wherein the polynucleotide comprises an adenoviral VAI or VAII RNA region.

32. The adenovirus or composition according to claim 31, wherein the VAI or VAII RNA region is from an adenovirus other than ChAd155.

33. The adenovirus or composition according to claim 32, wherein the VAI or VAII RNA region is from Ad5.

34. The adenovirus or composition according to any one of claims 1-28, wherein the polynucleotide comprises or consists of a polynucleotide which is at least 98.6% identical over its entire length to a reference sequence that consists essentially of SEQ ID NO: 6, 7 or 8.

35. The adenovirus or composition according to claim 34, wherein the polynucleotide comprises or consists of a polynucleotide which is at least 99% identical over its entire length to the reference sequence.

36. The adenovirus or composition according to claim 35, wherein the polynucleotide comprises or consists of a polynucleotide which is at least 99.5% identical over its entire length to the reference sequence.

37. The adenovirus or composition according to any one of claims 34-46, wherein the polynucleotide comprises or consists of a polynucleotide which is identical over its entire length to the reference sequence.
- 5 38. The adenovirus or composition according to claim 37, wherein the reference sequence is SEQ ID NO: 7.
- 10 39. The adenovirus or composition according to claim 37, wherein the reference sequence is SEQ ID NO: 8.
40. The adenovirus or composition according to claim 37, wherein reference sequence is SEQ ID NO: 8 or 9.
- 15 41. The adenovirus or composition according to any one of claims 1 to 28, wherein the polynucleotide comprises a mutation or deletion which renders non-functional at least one gene of a genomic region selected from the group consisting of E1A, E1B, E2A, E2B, E3 and E4.
- 20 42. The adenovirus or composition according to claim 41, wherein the polynucleotide lacks at least one gene of a genomic region selected from the group consisting of E1A, E1B, E2A, E2B, E3 and/or E4.
- 25 43. The adenovirus or composition according to either claim 41 or 42, wherein the genomic regions are E1A and/or E1B.
44. The adenovirus or composition according to any one of claims 1 to 43, wherein the recombinant adenovirus is replication-competent.
- 30 45. The adenovirus or composition according to any one of claims 1 to 43, wherein the recombinant adenovirus is replication-incompetent.
- 35 46. The adenovirus or composition according to any one of claims 1 to 45, wherein the nucleic acid sequence encodes a mycobacterial antigen derived from at least one of Rv0125, Rv0287, Rv0288, Rv0475, Rv0915, Rv1009, Rv1174, Rv1196, Rv1733, Rv1753, Rv1793, Rv1813, Rv1886, Rv1980, Rv2031, Rv2386, Rv2389, Rv2608, Rv2626, Rv2660, Rv3136, Rv3407, Rv3616, Rv3619, Rv3620, Rv3804 and Rv3875 from H37Rv or a homologue thereof from another mycobacterium or another strain of *Mycobacterium tuberculosis*.
- 40 47. The adenovirus or composition according to any one of claims 1 to 45, wherein the nucleic acid sequence encodes a mycobacterial antigen derived from at least one of Rv0125, Rv0287, Rv0288, Rv0475, Rv0915, Rv1009, Rv1174, Rv1196, Rv1733,

Rv1753, Rv1793, Rv1813, Rv1886, Rv1980, Rv2031, Rv2386, Rv2389, Rv2608, Rv2626, Rv2660, Rv3136, Rv3407, Rv3616, Rv3619, Rv3620, Rv3804 and Rv3875 from H37Rv or a homologue thereof from another strain of *Mycobacterium tuberculosis*.

- 5 48. The adenovirus or composition according to any one of claims 1 to 45, wherein the nucleic acid sequence encodes a mycobacterial antigen derived from at least one of Rv0125, Rv0287, Rv0288, Rv0475, Rv0915, Rv1009, Rv1174, Rv1196, Rv1733, Rv1753, Rv1793, Rv1813, Rv1886, Rv1980, Rv2031, Rv2386, Rv2389, Rv2608, Rv2626, Rv2660, Rv3136, Rv3407, Rv3616, Rv3619, Rv3620, Rv3804 and Rv3875 from H37Rv.
10
49. The adenovirus or composition according to any one of claims 1 to 48, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises at least one of Rv0125, Rv0287, Rv0288, Rv0475, Rv0915, Rv1009, Rv1174, Rv1196, Rv1733, Rv1753, Rv1793, Rv1813, Rv1886, Rv1980, Rv2031, Rv2386, Rv2389, Rv2608, Rv2626, Rv2660, Rv3136, Rv3407, Rv3616, Rv3619, Rv3620, Rv3804 and Rv3875.
15
50. The adenovirus or composition according to claim 49, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises at least one of Rv0125, Rv0287, Rv0915, Rv1174, Rv1196, Rv1753, Rv1793, Rv2386 and Rv3616.
20
51. The adenovirus or composition according to claim 50, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises at least one of Rv0125 and Rv1196.
25
52. The adenovirus or composition according to any one of claims 1 to 48, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises a sequence which has at least 80% identity to Rv0125, Rv0287, Rv0288, Rv0475, Rv0915, Rv1009, Rv1174, Rv1196, Rv1733, Rv1753, Rv1793, Rv1813, Rv1886, Rv1980, Rv2031, Rv2386, Rv2389, Rv2608, Rv2626, Rv2660, Rv3136, Rv3407, Rv3616, Rv3619, Rv3620, Rv3804 or Rv3875.
30
53. The adenovirus or composition according to claim 52, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises a sequence which has at least 80% identity to Rv0125, Rv0287, Rv0915, Rv1174, Rv1196, Rv1753, Rv1793, Rv2386 or Rv3616.
35
54. The adenovirus or composition according to claim 53, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises a sequence which has at least 80% identity to Rv0125 or Rv1196.
40

55. The adenovirus or composition according to any one of claims 52 to 54, wherein the mycobacterial antigen which comprises a sequence which has at least 90% identity to the reference sequence, in particular at least 95%, such as at least 98%.
- 5 56. The adenovirus or composition according to any one of claims 1 to 48, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises a sequence which is an immunogenic fragment (such as comprising one or more T cell epitopes) of Rv0125, Rv0287, Rv0288, Rv0475, Rv0915, Rv1009, Rv1174, Rv1196, Rv1733, Rv1753, Rv1793, Rv1813, Rv1886, Rv1980, Rv2031, Rv2386, Rv2389, Rv2608, 10 Rv2626, Rv2660, Rv3136, Rv3407, Rv3616, Rv3619, Rv3620, Rv3804 or Rv3875.
57. The adenovirus or composition according to claim 52, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises a sequence which is an immunogenic fragment (such as comprising one or more T cell epitopes) of Rv0125, 15 Rv0287, Rv0915, Rv1174, Rv1196, Rv1753, Rv1793, Rv2386 or Rv3616.
58. The adenovirus or composition according to claim 53, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises a sequence which is an immunogenic fragment (such as comprising one or more T cell epitopes) of Rv0125 or 20 Rv1196.
59. The adenovirus or composition according to any one of claims 56 to 58, wherein the mycobacterial antigen comprises a sequence which is an immunogenic fragment of at least 20 amino acid residues, such as at least 50 amino acid residues from the 25 reference sequence.
60. The adenovirus or composition according to any one of claims 56 to 58, wherein the mycobacterial antigen comprises a sequence which is an immunogenic fragment least 20% of the total length of the reference sequence, such as at least 30% of the total 30 length of the reference sequence residues, in particular at least 40% of the total length of the reference sequence residues, especially at least 50% of the total length of the reference sequence residues such as at least 75% of the total length of the reference sequence.
- 35 61. The adenovirus or composition according to any one of claims 1 to 48, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises at least one of SEQ ID NOs: 39 to 68.
- 40 62. The adenovirus or composition according to claim 61, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises at least one of SEQ ID NOs: 65 to 68.

63. The adenovirus or composition according to claim 61, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises SEQ ID NO: 70.
- 5 64. The adenovirus or composition according to any one of claims 1 to 48, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises a sequence which has at least 80% identity to any of SEQ ID NOs: 39 to 68.
- 10 65. The adenovirus or composition according to claim 64, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises a sequence which has at least 80% identity to any of SEQ ID NOs: 65 to 68.
- 15 66. The adenovirus or composition according to claim 64, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises a sequence which has at least 80% identity to SEQ ID NO: 70.
- 20 67. The adenovirus or composition according to any one of claims 64 to 66, wherein the mycobacterial antigen which comprises a sequence which has at least 90% identity to the reference sequence, in particular at least 95%, such as at least 98%.
- 25 68. The adenovirus or composition according to any one of claims 1 to 48, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises a sequence which is an immunogenic fragment (such as comprising one or more T cell epitopes) of any of SEQ ID NOs: 39 to 68.
- 30 69. The adenovirus or composition according to claim 68, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises a sequence which is an immunogenic fragment (such as comprising one or more T cell epitopes) of any of SEQ ID NOs: 65 to 68.
- 35 70. The adenovirus or composition according to claim 68, wherein the nucleic acid sequence encodes a mycobacterial antigen which comprises a sequence which is an immunogenic fragment (such as comprising one or more T cell epitopes) of SEQ ID NO: 70.
- 40 71. The adenovirus or composition according to any one of claims 68 to 70, wherein the mycobacterial antigen comprises a sequence which is an immunogenic fragment of at least 20 amino acid residues, such as at least 50 amino acid residues from the reference sequence.
72. The adenovirus or composition according to any one of claims 68 to 70, wherein the mycobacterial antigen comprises a sequence which is an immunogenic fragment least 20% of the total length of the reference sequence, such as at least 30% of the total length of the reference sequence residues, in particular at least 40% of the total length

of the reference sequence residues, especially at least 50% of the total length of the reference sequence residues such as at least 75% of the total length of the reference sequence.

- 5 73. The adenovirus or composition of any one of claims 1 to 72, wherein the one or more sequences which direct expression of said product in a host cell includes a sequence selected from one or more of the group consisting of: transcription initiation, transcription termination, promoter and enhancer sequences.
- 10 74. The adenovirus or composition of claim 73, wherein the one or more sequences which direct expression of said product in a host cell includes a promoter sequence.
- 15 75. The adenovirus or composition of claim 74, wherein the promoter sequence is selected from the group consisting of an internal promoter, a native promoter, RSV LTR promoter, CMV promoter, SV40 promoter, dihydrofolate reductase promoter, β -actin promoter, PGK promoter, EF1a promoter and CASI promoter.
- 20 76. The adenovirus or composition according to any one of claims 1 to 75, wherein the adenovirus has a seroprevalence of less than 10% in human subjects and preferably no seroprevalence in human subjects.
- 25 77. The adenovirus or composition according to any one of claims 1 to 76, wherein the adenovirus is capable of infecting a mammalian cell.
- 30 78. The composition according to any one of claims 2 to 77, comprising an adjuvant selected from the list consisting of: inorganic adjuvants (e.g. inorganic metal salts such as aluminium phosphate or aluminium hydroxide), organic adjuvants (e.g. saponins, such as QS21, or squalene), oil-based adjuvants (e.g. Freund's complete adjuvant and Freund's incomplete adjuvant), cytokines (e.g. IL-1 β , IL-2, IL-7, IL-12, IL-18, GM-CSF, and INF- γ), particulate adjuvants (e.g. immuno-stimulatory complexes (ISCOMS), liposomes, or biodegradable microspheres), virosomes, bacterial adjuvants (e.g. monophosphoryl lipid A, such as 3-de-O-acylated monophosphoryl lipid A (3D-MPL), or muramyl peptides), synthetic adjuvants (e.g. non-ionic block copolymers, muramyl peptide analogues, or synthetic lipid A), synthetic polynucleotide adjuvants (e.g. polyarginine or polylysine) and immunostimulatory oligonucleotides containing unmethylated CpG dinucleotides ("CpG") or mixtures of such adjuvants.
- 35 79. The composition according to claim 78, wherein the adjuvant is a 3D-MPL and/or QS21.
- 40 80. The adenovirus or composition according to any one of claims 1 to 79, for use as a medicament.

81. The adenovirus or composition according to claim 80, for use in the stimulation of an immune response.
82. The adenovirus or composition according to claim 80, for use as a vaccine.
83. The adenovirus or composition according to claim 80, for use in the prophylaxis of mycobacterial, such as *Mycobacterium tuberculosis*, infection.
84. The adenovirus or composition according to claim 80, for use in the treatment of mycobacterial, such as *Mycobacterium tuberculosis*, infection.
85. Use of the adenovirus or composition according to any one of claims 1 to 79 in the manufacture of a medicament for the treatment or prophylaxis of mycobacterial infection.
86. A method of inducing an immune response in a subject comprising administering the adenovirus or composition according to any one of claims 1 to 79 to the subject.
87. A non-human simian adenovirus containing at least the penton of SEQ ID NO: 3, the hexon of SEQ ID NO: 5 or the fiber of SEQ ID NO: 1 and comprising a transgene encoding a Rv1196 or Rv0125 related antigen.
88. The non-human simian adenovirus according to claim 87, comprising the penton (SEQ ID NO: 3), hexon (SEQ ID NO: 5) and fiber (SEQ ID NO: 1) protein from ChAd155.
89. The non-human simian adenovirus according to claim 87 or 88 wherein the encoded antigen comprises a sequence having at least 90% identity to SEQ ID NO. 70.
90. The non-human simian adenovirus according to claim 87 or 88 wherein the encoded antigen comprises a fragment of SEQ ID NO: 70 which is at least 450 amino acids in length, such as 2-723 of SEQ ID NO. 70.
91. The non-human simian adenovirus according to any one of claims 87 to 90, which is a replication deficient adenovirus.
92. The non-human simian adenovirus according to any one of claims 87 to 91, wherein the adenovirus comprises functional inactivation (such as deletion) of the E1 gene.
93. The non-human simian adenovirus according to any one of claims 97 to 92, wherein the adenovirus comprises functional inactivation (such as deletion) of the E4 gene.
94. The non-human simian adenovirus according to any one of claims 87 to 93, wherein the adenovirus comprises functional inactivation (such as deletion) of the E3 gene.

95. The non-human simian adenovirus according to any one of claims 87 to 94, wherein the adenovirus comprises an Ad5E4orf6 gene substitution.

1/15
Figure 1A

ChAd3	(1)	MKR	TK	T	S	D	E	S	F	N	P	V	Y	P	D	T	E	S	G	P	P	S	V	P	F	L	T	P	P	F	V	S	P	D	G	F	Q	E	S	P	P	G	V	L	S	L	N	L	A	E	P	L	V	T	S	H	G	M	L	A	K	M	G	S	G	L	S	L	D	D	A	G	N	L	T	S	Q	D	I	T	T	A	S	P	P	L	K	K	T	K	T	N	L	S	L	E					
PanAd3	(1)	MK	RA	K	T	S	D	E	T	F	N	P	V	Y	P	D	T	E	N	G	P	P	S	V	P	F	L	T	P	P	F	V	S	P	D	G	F	Q	E	S	P	P	G	V	L	S	L	R	L	S	E	P	L	V	T	S	H	G	M	L	A	K	M	G	N	G	L	S	L	D	D	A	G	N	L	T	S	Q	D	V	T	T	V	T	P	P	L	K	K	T	K	T	N	L	S	L	Q				
ChAd17	(1)	MKR	TK	T	S	D	E	S	F	N	P	V	Y	P	D	T	E	S	G	P	P	S	V	P	F	L	T	P	P	F	V	S	P	D	G	F	Q	E	S	P	P	G	V	L	S	L	N	L	A	E	P	L	V	T	S	H	G	M	L	A	K	M	G	S	G	L	S	L	D	D	A	G	N	L	T	S	Q	D	I	T	S	T	T	P	P	L	K	K	T	K	T	N	L	S	L	E					
ChAd19	(1)	MKR	TK	T	S	D	K	S	F	N	P	V	Y	P	D	T	E	N	G	P	P	S	V	P	F	L	T	P	P	F	V	S	P	D	G	F	Q	E	S	P	P	G	V	L	S	L	N	L	A	E	P	L	V	T	S	H	G	M	L	A	K	M	G	S	G	L	S	L	D	D	A	G	N	L	T	S	Q	D	V	T	T	T	P	P	L	K	K	T	K	T	N	L	S	L	E						
ChAd24	(1)	MKR	TK	T	S	D	E	S	F	N	P	V	Y	P	D	T	E	N	G	P	P	S	V	P	F	L	T	P	P	F	V	S	P	D	G	F	Q	E	S	P	P	G	V	L	S	L	N	L	A	E	P	L	V	T	S	H	G	M	L	A	K	M	G	S	G	L	S	L	D	D	A	G	N	L	T	S	Q	D	V	T	T	T	P	P	L	K	K	T	K	T	N	L	S	L	E						
ChAd155	(1)	MKR	TK	T	S	D	E	S	F	N	P	V	Y	P	D	T	E	S	G	P	P	S	V	P	F	L	T	P	P	F	V	S	P	D	G	F	Q	E	S	P	P	G	V	L	S	L	N	L	A	E	P	L	V	T	S	H	G	M	L	A	K	M	G	S	G	L	S	L	D	D	A	G	N	L	T	S	Q	D	I	T	T	A	S	P	P	L	K	K	T	K	T	N	L	S	L	E					
ChAd11	(1)	MKR	TK	T	S	D	E	S	F	N	P	V	Y	P	D	T	E	N	G	P	P	S	V	P	F	L	T	P	P	F	V	S	P	D	G	F	Q	E	S	P	P	G	V	L	S	L	N	L	A	E	P	L	V	T	S	H	G	M	L	A	K	M	G	S	G	L	S	L	D	D	A	G	N	L	T	S	Q	D	V	T	T	T	P	P	L	K	K	T	K	T	N	L	S	L	E						
ChAd20	(1)	MKR	TK	T	S	D	E	S	F	N	P	V	Y	P	D	T	E	S	G	P	P	S	V	P	F	L	T	P	P	F	V	S	P	D	G	F	Q	E	S	P	P	G	V	L	S	L	N	L	A	E	P	L	V	T	S	H	G	M	L	A	K	M	G	S	G	L	S	L	D	D	A	G	N	L	T	S	Q	D	I	T	T	A	S	P	P	L	K	K	T	K	T	N	L	S	L	E					
ChAd31	(1)	MKR	TK	T	S	D	E	S	F	N	P	V	Y	P	D	T	E	S	G	P	P	S	V	P	F	L	T	P	P	F	V	S	P	D	G	F	Q	E	S	P	P	G	V	L	S	L	N	L	A	E	P	L	V	T	S	H	G	M	L	A	K	M	G	S	G	L	S	L	D	D	A	G	N	L	T	S	Q	D	I	T	T	A	S	P	P	L	K	K	T	K	T	N	L	S	L	E					
PanAd1	(1)	MK	RA	K	T	S	D	E	T	F	N	P	V	Y	P	D	T	E	N	G	P	P	S	V	P	F	L	T	P	P	F	V	S	P	D	G	F	Q	E	S	P	P	G	V	L	S	L	R	L	S	E	P	L	V	T	S	H	G	M	L	A	K	M	G	N	G	L	S	L	D	D	A	G	N	L	T	S	Q	D	V	T	T	V	T	P	P	L	K	K	T	K	T	N	L	S	L	Q				
PanAd2	(1)	MK	RA	K	T	S	D	E	T	F	N	P	V	Y	P	D	T	E	N	G	P	P	S	V	P	F	L	T	P	P	F	V	S	P	D	G	F	Q	E	S	P	P	G	V	L	S	L	R	L	S	E	P	L	V	T	S	H	G	M	L	A	K	M	G	N	G	L	S	L	D	D	A	G	N	L	T	S	Q	D	V	T	T	V	T	P	P	L	K	K	T	K	T	N	L	S	L	Q				
ChAd3	(101)	T	S	S	P	L	T	V	S	T	S	G	A	L	T	V	A	A	A	A	P	L	A	V	A	G	T	S	L	T	M	Q	S	E	A	P	L	T	V	Q	D	A	K	L	T	L	A	T	K	G	P	L	T	V	S	E	G	K	L	A	L	Q	T	S	A	P	L	T	A	A	D	S	S	T	L	T	V	S	A	T	P	P	I	N	V	S	S	G	S	L	G	L	D	M	E	D	P	M	Y	T	H
PanAd3	(101)	T	S	A	P	L	T	V	S	-	S	G	S	L	T	V	A	A	A	A	P	L	A	V	A	G	T	S	L	T	M	Q	S	E	A	P	L	T	V	Q	D	A	K	L	G	L	A	T	Q	G	P	L	T	V	S	E	G	K	L	T	L	Q	T	S	A	P	L	T	A	A	D	S	S	T	L	T	V	G	T	T	P	P	I	S	V	S	S	G	S	L	G	L	D	M	E	D	P	M	Y	T	H
ChAd17	(101)	T	S	S	P	L	T	V	S	T	S	G	A	L	T	V	A	A	A	A	P	L	A	V	A	G	T	S	L	T	M	Q	S	E	A	P	L	A	V	Q	D	A	K	L	T	L	A	T	K	G	P	L	T	V	S	E	G	K	L	A	L	Q	T	S	A	P	L	T	A	A	D	S	S	T	L	T	V	S	S	T	P	P	I	S	V	S	S	G	S	L	G	L	D	M	E	D	P	M	Y	T	H
ChAd19	(101)	T	S	A	P	L	T	V	S	T	S	G	A	L	T	L	A	A	A	A	P	L	A	V	A	G	T	S	L	T	M	Q	S	E	A	P	L	T	V	Q	D	A	K	L	T	L	A	T	K	G	P	L	T	V	S	E	G	K	L	A	L	Q	T	S	A	P	L	T	A	A	D	S	S	T	L	T	V	S	A	T	P	P	I	S	V	S	S	G	S	L	G	L	D	M	E	D	P	M	Y	T	H
ChAd24	(101)	T	S	A	P	L	T	V	S	T	S	G	A	L	T	L	A	A	A	A	P	L	A	V	A	G	T	S	L	T	M	Q	S	E	A	P	L	T	V	Q	D	A	K	L	T	L	A	T	K	G	P	L	T	V	S	E	G	K	L	A	L	Q	T	S	A	P	L	T	A	A	D	S	S	T	L	T	V	S	A	T	P	P	I	N	V	S	S	G	S	L	G	L	D	M	E	N	P	M	Y	T	H
ChAd155	(101)	T	S	S	P	L	T	V	S	T	S	G	A	L	T	V	A	A	A	A	P	L	A	V	A	G	T	S	L	T	M	Q	S	E	A	P	L	T	V	Q	D	A	K	L	T	L	A	T	K	G	P	L	T	V	S	E	G	K	L	A	L	Q	T	S	A	P	L	T	A	A	D	S	S	T	L	T	V	S	A	T	P	P	L	S	T	S	N	G	S	L	G	I	D	M	Q	A	P	I	Y	T	T
ChAd11	(101)	T	S	A	P	L	T	V	S	T	S	G	A	L	T	L	A	A	A	V	P	L	A	V	A	G	T	S	L	T	M	Q	S	E	A	P	L	T	V	Q	D	A	K	L	T	L	A	T	K	G	P	L	T	V	S	E	G	K	L	A	L	Q	T	S	A	P	L	T	A	A	D	S	S	T	L	T	I	S	A	T	P	P	L	S	T	S	N	G	S	L	G	I	D	M	Q	A	P	I	Y	T	T
ChAd20	(101)	T	S	S	P	L	T	V	S	T	S	G	A	L	T	V	A	A	A	A	P	L	A	V	A	G	T	S	L	T	M	Q	S	E	A	P	L	T	V	Q	D	A	K	L	T	L	A	T	K	G	P	L	T	V	S	E	G	K	L	A	L	Q	T	S	A	P	L	T	A	A	D	S	S	T	L	T	V	S	A	T	P	P	L	S	T	S	N	G	S	L	G	I	D	M	Q	A	P	I	Y	T	T
ChAd31	(101)	T	S	S	P	L	T	V	S	T	S	G	A	L	T	V	A	A	A	A	P	L	A	V	A	G	T	S	L	T	M	Q	S	E	A	P	L	T	V	Q	D	A	K	L	T	L	A	T	K	G	P	L	T	V	S	E	G	K	L	A	L	Q	T	S	A	P	L	T	A	A	D	S	S	T	L	T	V	S	A	T	P	P	L	S	T	S	N	G	S	L	G	I	D	M	Q	A	P	I	Y	T	T
PanAd1	(101)	T	S	A	P	L	T	V	S	-	S	G</																																																																																									

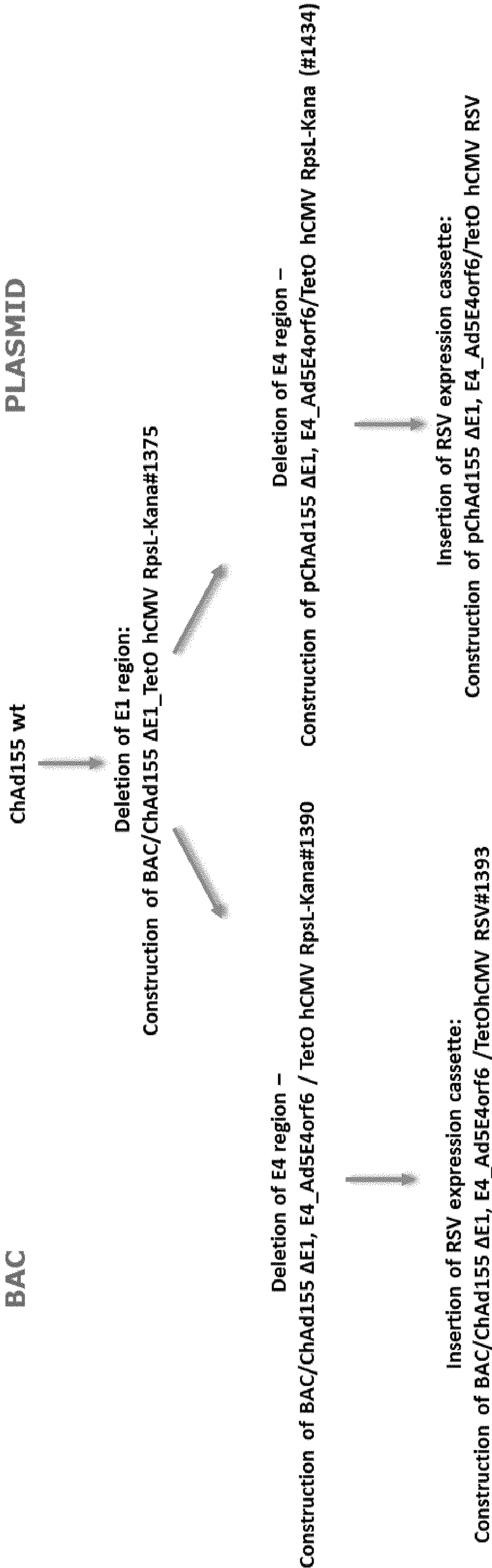
2/15
Figure 1B

ChAd3	(201)	DGKLGIRIGGPLRVVDSLHTLTVVVTGNGLTVDNNAIQTRVTGALGYDTSGNLQLRAGGMRIDANGQLILNVAYPFDAQNNLSRLGGQPLYINTDHNLD
PanAd3	(200)	DGKLGIRIGGPLQVVDSLHTLTVVVTGNGITVANNALQTKVAGALGYDSSGNLELRAGGMRINTGGQLILDVAYPFDAQNNLSRLGGQPLYVNTNHNLD
ChAd17	(201)	DGKLGIRIGGPLRVVDSLHTLTVVVTGNGLTVDNNAIQTRVTGALGYDTSGNLQLRAGGMRIDANGQLILDVAYPFDAQNNLSRLGGQPLYVNTDHNLD
ChAd19	(201)	DGKLGIRIGGPLRVVDSLHTLTVVVTGNGIAVDNNAIQTRVTGALGYDTSGNLQLRAGGMRIDANGQLILDVAYPFDAQNNLSRLGGQPLYVNTDHNLD
ChAd24	(201)	DGKLGIRIGGPLRVVDSLHTLTVVVTGNGIAVDNNAIQTRVTGALGYDTSGNLQLRAGGMRIDANGQLILDVAYPFDAQNNLSRLGGQPLYVNTDHNLD
ChAd155	(201)	NGKGLNFGAPLHVVDLSNALTVVVTGQGLTINGTALQTRVSGALNYDTSGNLELRAGGMRVDANGQLILDVAYPFDAQNNLSRLGGQPLFVNSAHNLD
ChAd11	(201)	NGKGLNFGAPLHVVDLSNALTVVVTGQGLTINGTALQTRVSGALNYDSSGNLELRAGGMRVDANGKLLILDVAYPFDAQNNLSRLGGQPLFVNSAHNLD
ChAd20	(201)	NGKGLNFGAPLHVVDLSNALTVVVTGQGLTINGTALQTRVSGALNYDTSGNLELRAGGMRVDANGQLILDVAYPFDAQNNLSRLGGQPLFVNSAHNLD
ChAd31	(201)	NGKGLNFGAPLHVVDLSNALTVVVTGQGLTINGTALQTRVSGALNYDTSGNLELRAGGMRVDANGQLILDVAYPFDAQNNLSRLGGQPLFVNSAHNLD
PanAd 1	(200)	NGKLALNIGAPLHVVDTLNALTVVVTGQGLTINGRALQTRVTGALS YDTEGNIQLQAGGMRIDNNGQLILNVAYPFDAQNNLSRLGGQPLI VNSAHNLD
PanAd2	(200)	NGKLALNIGAPLHVVDTLNALTVVVTGQGLTINGRALQTRVTGALS YDTEGNIQLQAGGMRIDNNGQLILNVAYPFDAQNNLSRLGGQPLI VNSAHNLD
ChAd3	(301)	LNCNRGLTTTTNNTKKLET-----KISSGLDYDTNGAVIIKLTGGLSFDNTGALT VGNTGDDDKLTLWT
PanAd3	(300)	LNCNRGLTTTTSSNTTKLET-----KIDSGLDYNANGAIIAKLTGGLTFDNTGAITVGNTGDDDKLTLWT
ChAd17	(301)	LNCNRGLTTTTNNTKKLET-----KISSGLDYDTNGAVIIKLTGGLSFDNTGALT VGNTGDDDKLTLWT
ChAd19	(301)	LNCNRGLTTTTNNTKKLET-----KIGSGLDYDTNGAVIIKLTGVSFSDSTGALS VGNTGDDDKLTLWT
ChAd24	(301)	LNCNRGLTTTTNNTKKLET-----KIGSGLDYDTNGAVIIKLTGVSFSDSTGALS VGNTGDDDKLTLWT
ChAd155	(301)	VNYNRGLYLFTSGNTKKLEVNIKTAKGLIYDDTAIAINAGDGLQFDSGSDTNPLKTKLGLGLDYDSSRAIIAKLTGGLSFDNTGAITVGKNKND DKLTLWT
ChAd11	(301)	VNYNRGLYLFTSGNTKKLEVNIKTAKGLIYDDTAIAINPGDGLEFCGSDTNPLKTKLGLGLEIDSSRAIIAKLTGGLSFDNTGAITVGKNKND DKLTLWT
ChAd20	(301)	VNYNRGLYLFTSGNTKKLEVNIKTAKGLIYDDTAIAINAGDGLQFDSGSDTNPLKTKLGLGLDYDSSRAIIAKLTGGLSFDNTGAITVGKNKND DKLTLWT
ChAd31	(301)	VNYNRGLYLFTSGNTKKLEVNIKTAKGLIYDDTAIAINAGDGLQFDSGSDTNPLKTKLGLGLDYDSSRAIIAKLTGGLSFDNTGAITVGKNKND DKLTLWT
PanAd1	(300)	LNLNRGLYLFTSGNTKKLEVNIKTAKGLFYDGTAAIAINAGDGLQFDSGSDTNPLQTKLGLGLEIDYSNKAIIKTGGLSFDNTGAITVGKNKND DKLTLWT
PanAd2	(300)	LNLNRGLYLFTSGNTKKLEVNIKTAKGLFYDGTAAIAINAGDGLQFDSGSDTNPLQTKLGLGLEIDYSNKAIIKTGGLSFDNTGAITVGKNKND DKLTLWT

3/15
Figure 1C

ChAd3	(365)	TPDPSPNCRIHSDKDKCKFTLVLTCKGSQILASVAALAVSGNLASITGTVASVTIFLRFDQNGVLMENSSLDQYWNFRNGNSTNAAPYTNAVGFMPNLAA
PanAd3	(364)	TPDPSPNCRIHADKDKCKFTLVLTCKGSQILASVAALAVSGNLSSMTGTVSSVTIFLRFDQNGVLMENSSLDKEYWNFRNGNSTNATPYTNAVGFMPNLISA
ChAd17	(365)	TPDPSPNCRIHSDKDKCKFTLVLTCKGSQILASVAALAVSGNLASITGTVASVTIFLRFDQNGVLMENSSLDKQYWNFRNGNSTNAAPYTNAVGFMPNLAA
ChAd19	(365)	TPDPSPNCRIHSDKDKCKFTLVLTCKGSQILASVAALAVSGNLASITGTVSSVTIFLRFDQNGVLMENSSLDKQYWNFRNGNSTNATPYTNAVGFMPNLAA
ChAd24	(365)	TPDPSPNCRIHSDKDKCKFTLVLTCKGSQILASVAALAVSGNLASITGTVSSVTIFLRFDQNGVLMENSSLDKQYWNFRNGNSTNATPYTNAVGFMPNLAA
ChAd155	(401)	TPDPSPNCRIYSEKDAKFTLVLTCKGSQVLASVLSVKGSLAPISGTVTSAQIVLRFDENGVLNSSLDPQYWNYRKGDLTEGTAYTNAVGFMPNLTA
ChAd11	(401)	TPDPSPNCRIYSEKDAKFTLVLTCKGSQVLASVLSVKGSLAPISGTVTSAQII LRFDENGVLNSSLDPQYWNYRKGDLTEGTAYTNAVGFMPNLTA
ChAd20	(401)	TPDPSPNCRIYSEKDAKFTLVLTCKGSQVLASVLSVKGSLAPISGTVTSAQIVLRFDENGVLNSSLDPQYWNYRKGDLTEGTAYTNAVGFMPNLTA
ChAd31	(401)	TPDPSPNCRIYSEKDAKFTLVLTCKGSQVLASVLSVKGSLAPISGTVTSAQIVLRFDENGVLNSSLDPQYWNYRKGDLTEGTAYTNAVGFMPNLTA
PanAd1	(400)	TPDPSPNCRINSEKDAKLTLVLTCKGSQVLASVLSVKGSLAPISGTVTSAQIVLRFDENGVLNSSLDPQYWNYRKGDSTEGTAYTNAVGFMPNLTA
PanAd2	(400)	TPDPSPNCRINSEKDAKLTLVLTCKGSQVLASVLSVKGSLAPISGTVTSAQIVLRFDENGVLNSSLDPQYWNYRKGDSTEGTAYTNAVGFMPNLTA
ChAd3	(465)	YPKTQSQTAKNNIVSQVYLNQDKSKPMTLTI TLNGTNESSETSQVSHYSMSFTWAWESGQYATETFATNSFTFSYIAEQ
PanAd3	(464)	YPKTQSQTAKNNIVSEVYLHGDKSKPMILTI TLNGTNESSETSQVSHYSMSFTWSWDSGKYATETFATNSFTFSYIAEQ
ChAd17	(465)	YPKTQSQTAKNNIVSQVYLNQDKSKPMTLTI TLNGTNESSETSQVSHYSMSFTWAWESGQYATETFATNSFTFSYIAEQ
ChAd19	(465)	YPKTQSQTAKNNIVSQVYLNQDKSKPMTLTI TLNGTNESSETSQVSHYSMSFTWAWESGQYATETFATNSFTFSYIAEQ
ChAd24	(465)	YPKTQSQTAKNNIVSQVYLNQDKSKPMTLTI TLNGTNESSETSQVSHYSMSFTWAWESGQYATETFATNSFTFSYIAEQ
ChAd155	(501)	YPKTQSQTAKSNIVSQVYLNQDKSKPMTLTI TLNGTNETG- DATVSTYSMSFSWNWNGSNYINETFQTNSTFTFSYIAQE
ChAd11	(501)	YPKTQSQTAKSNIVSQVYLNQDKSKPMTLTI TLNGTNETG- DATVSTYSMSFSWNWNGSNYINETFQTNSTFTFSYIAQE
ChAd20	(501)	YPKTQSQTAKSNIVSQVYLNQDKSKPMTLTI TLNGTNETG- DATVSTYSMSFSWNWNGSNYINETFQTNSTFTFSYIAQE
ChAd31	(501)	YPKTQSQTAKSNIVSQVYLNQDKSKPMTLTI TLNGTNETG- DATVSTYSMSFSWNWNGSNYINETFQTNSTFTFSYIAQE
PanAd1	(500)	YPKTQSQTAKSNIVSQVYLNQDKTKPMTLTI TLNGTNETG- DATVSTYSMSFSWNWNGSNYINDTFQTNSTFTFSYIAQE
PanAd2 (	500)	YPKTQSQTAKSNIVSQVYLNQDKTKPMTLTI TLNGTNETG- DATVSTYSMSFSWNWNGSNYINDTFQTNSTFTFSYIAQE

4/15
Figure 2



5/15
Figure 3

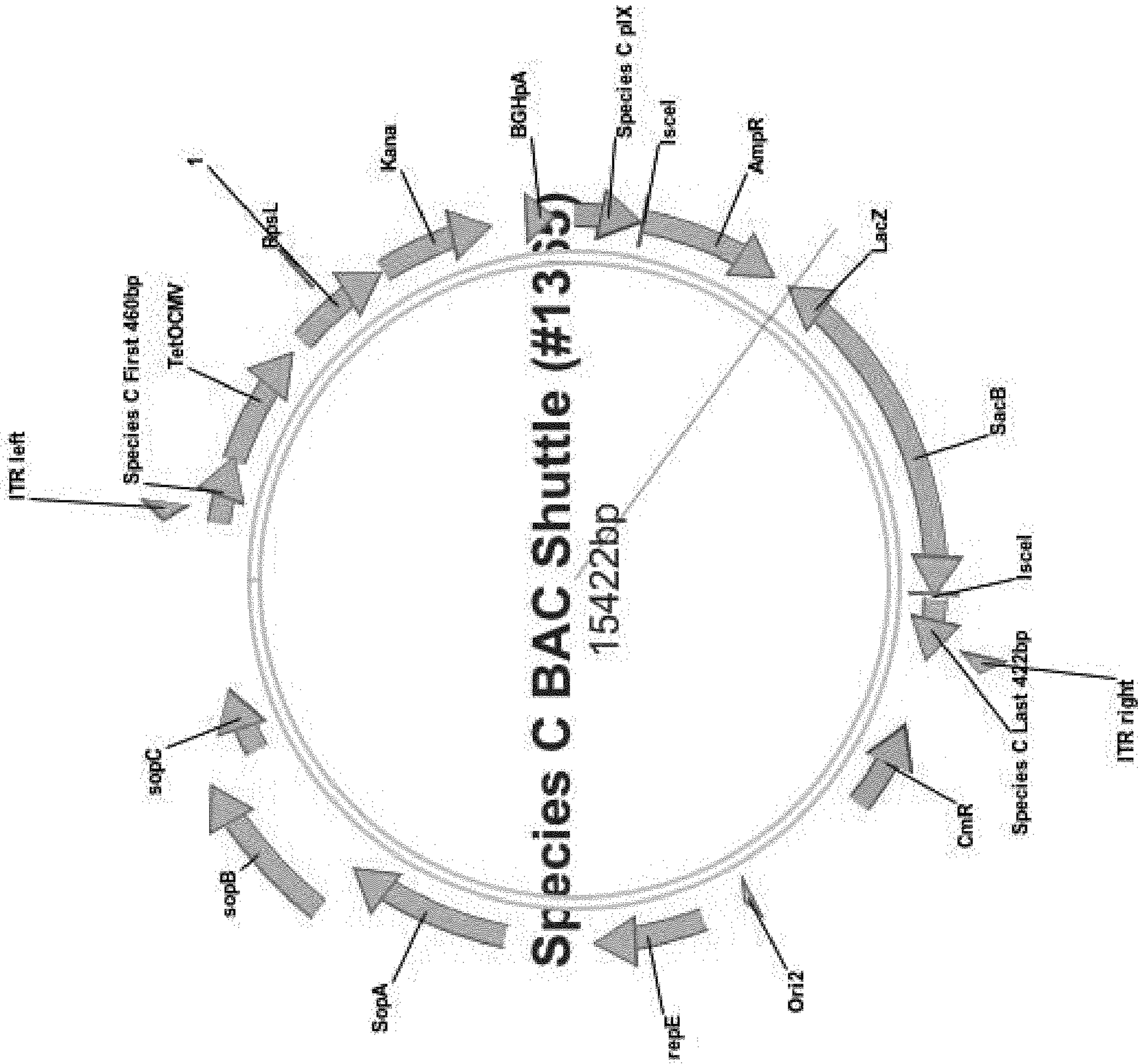
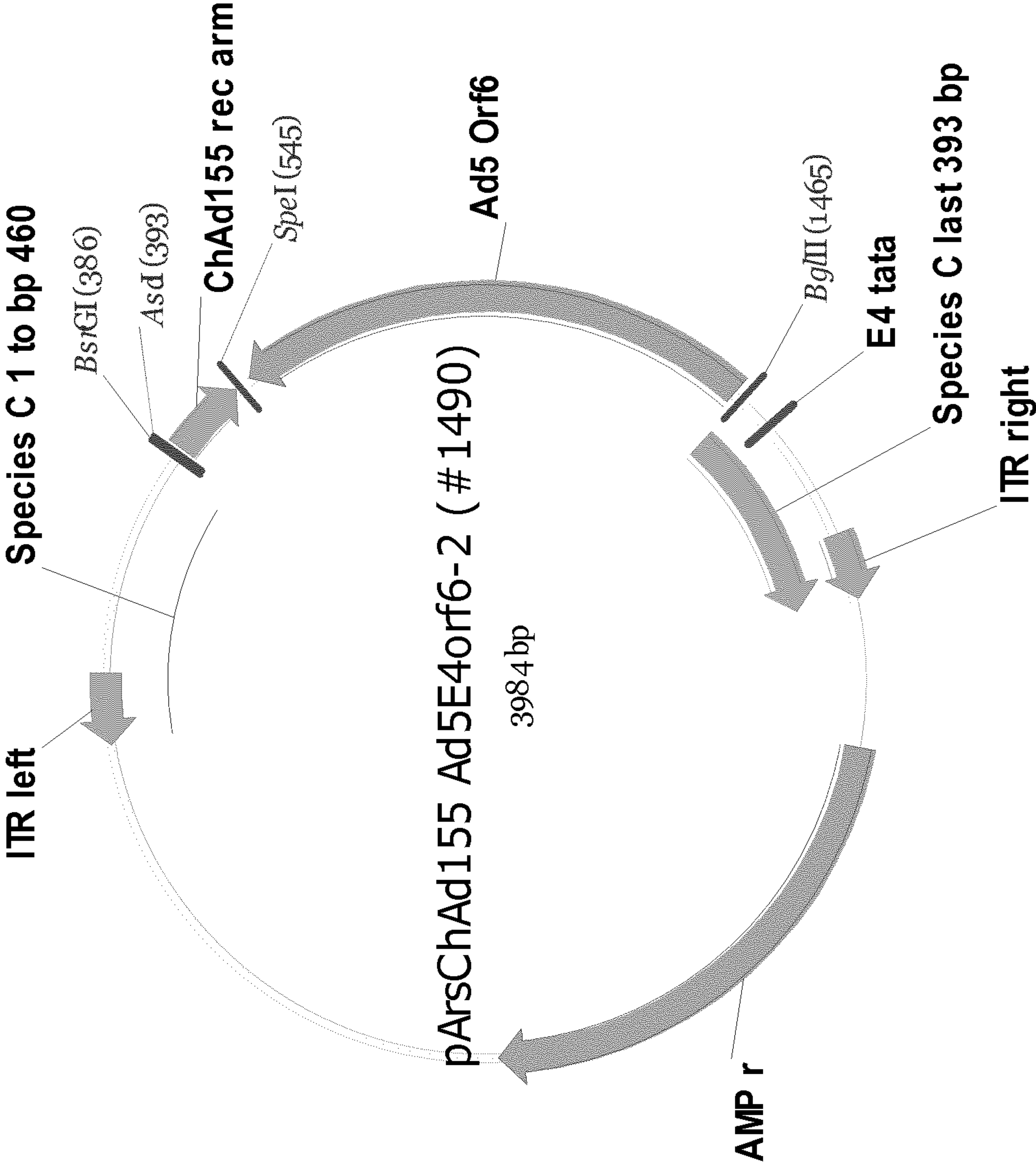


Figure 4



7/15
Figure 5

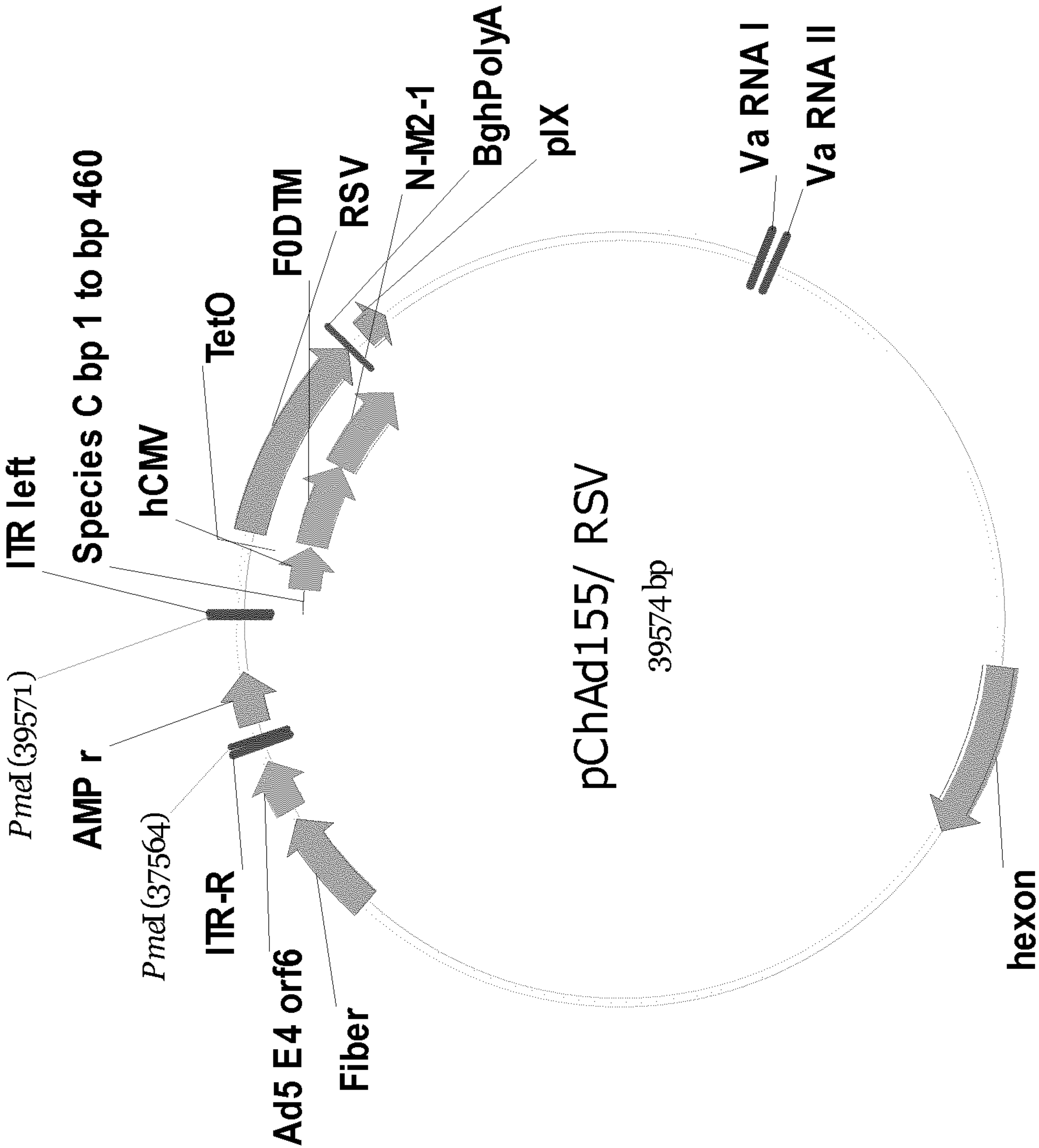
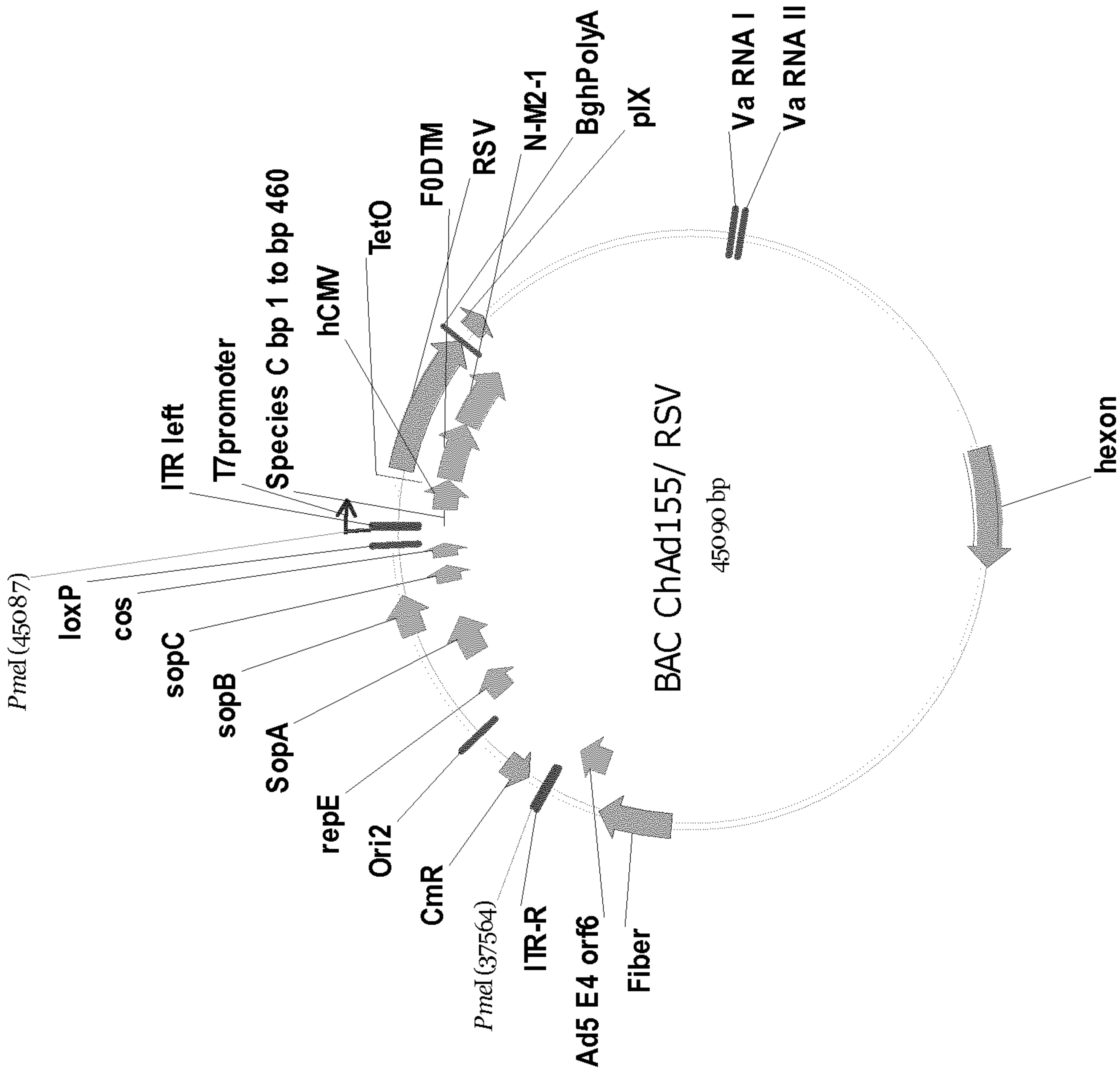
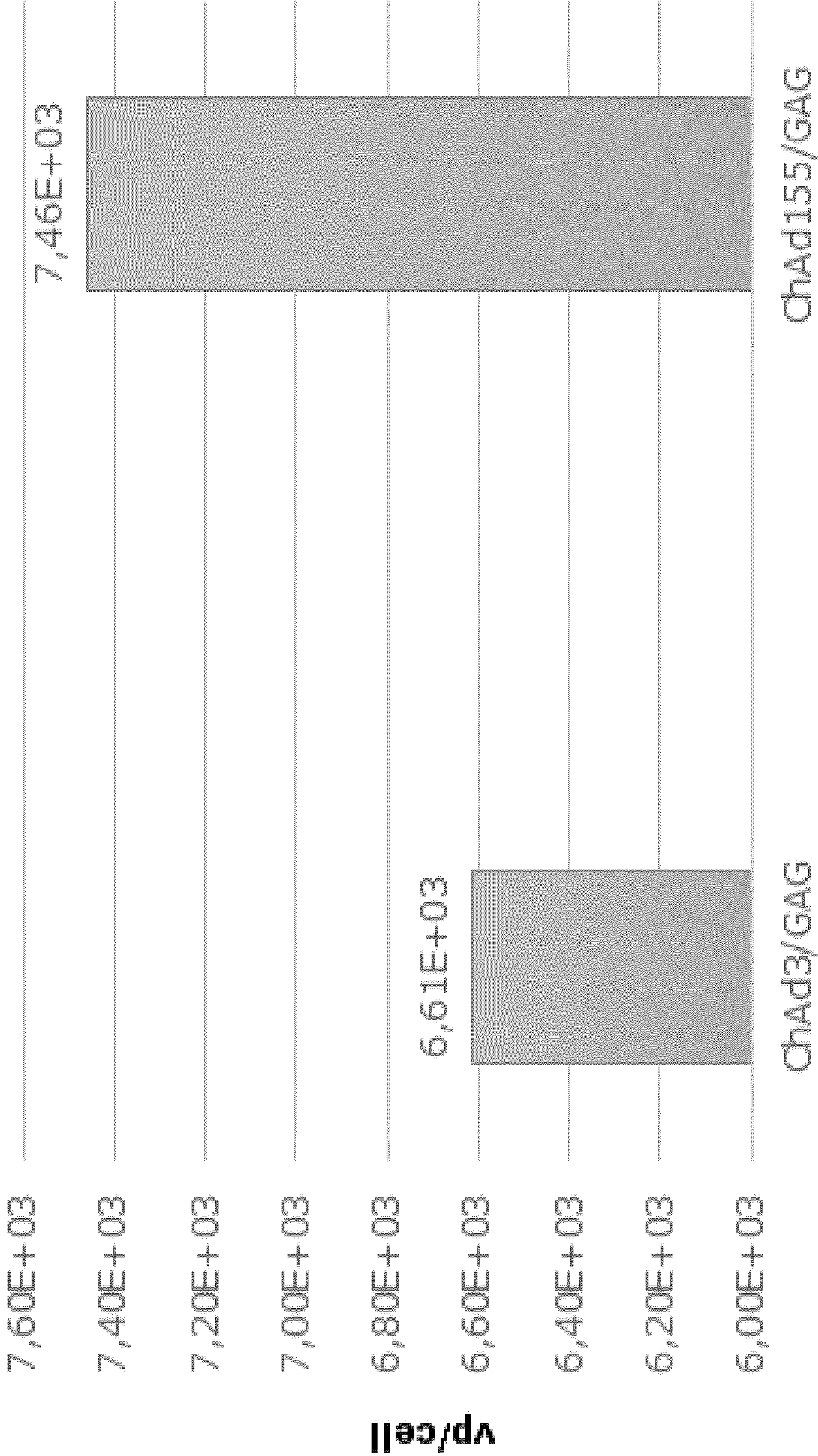


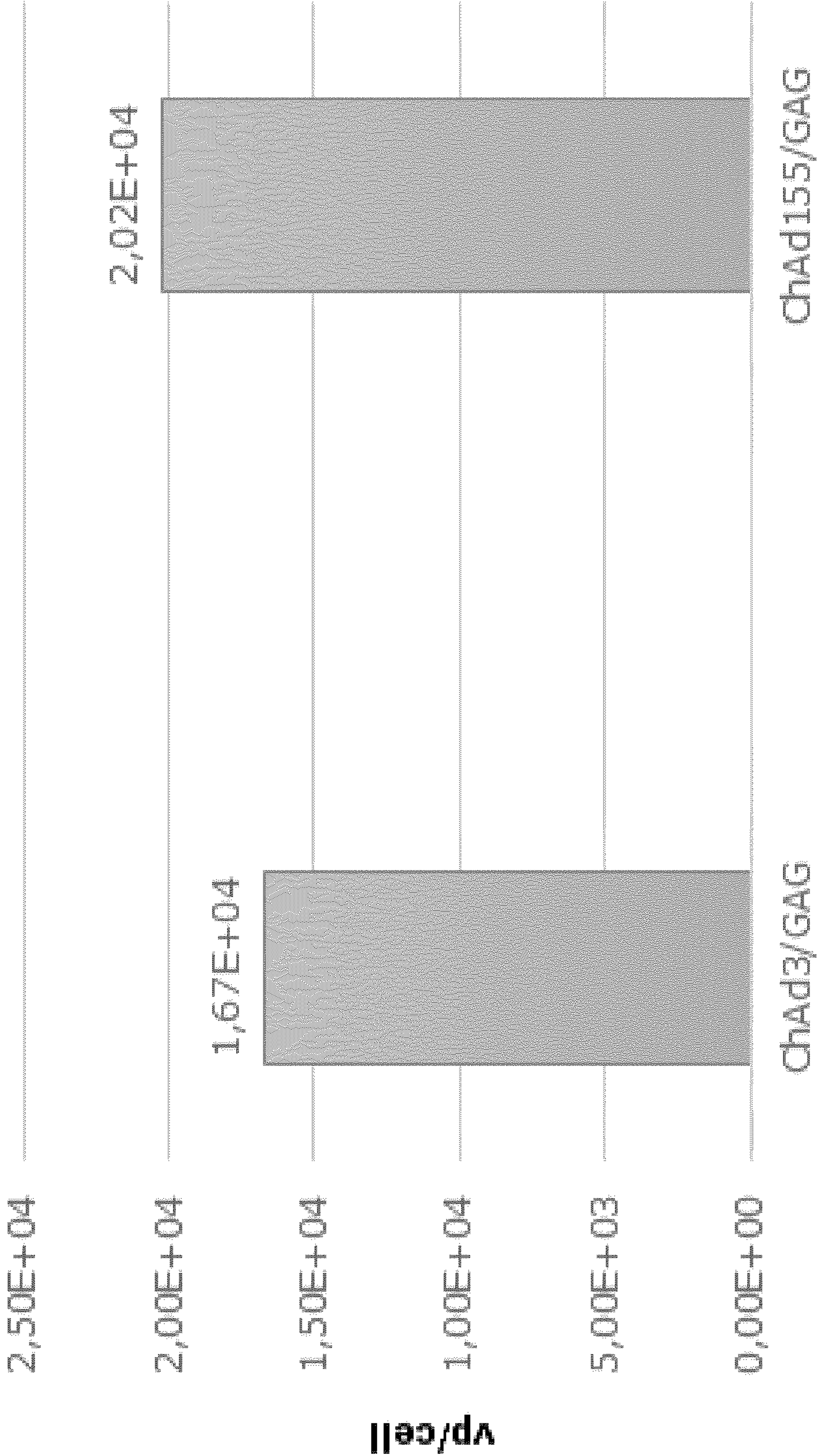
Figure 6



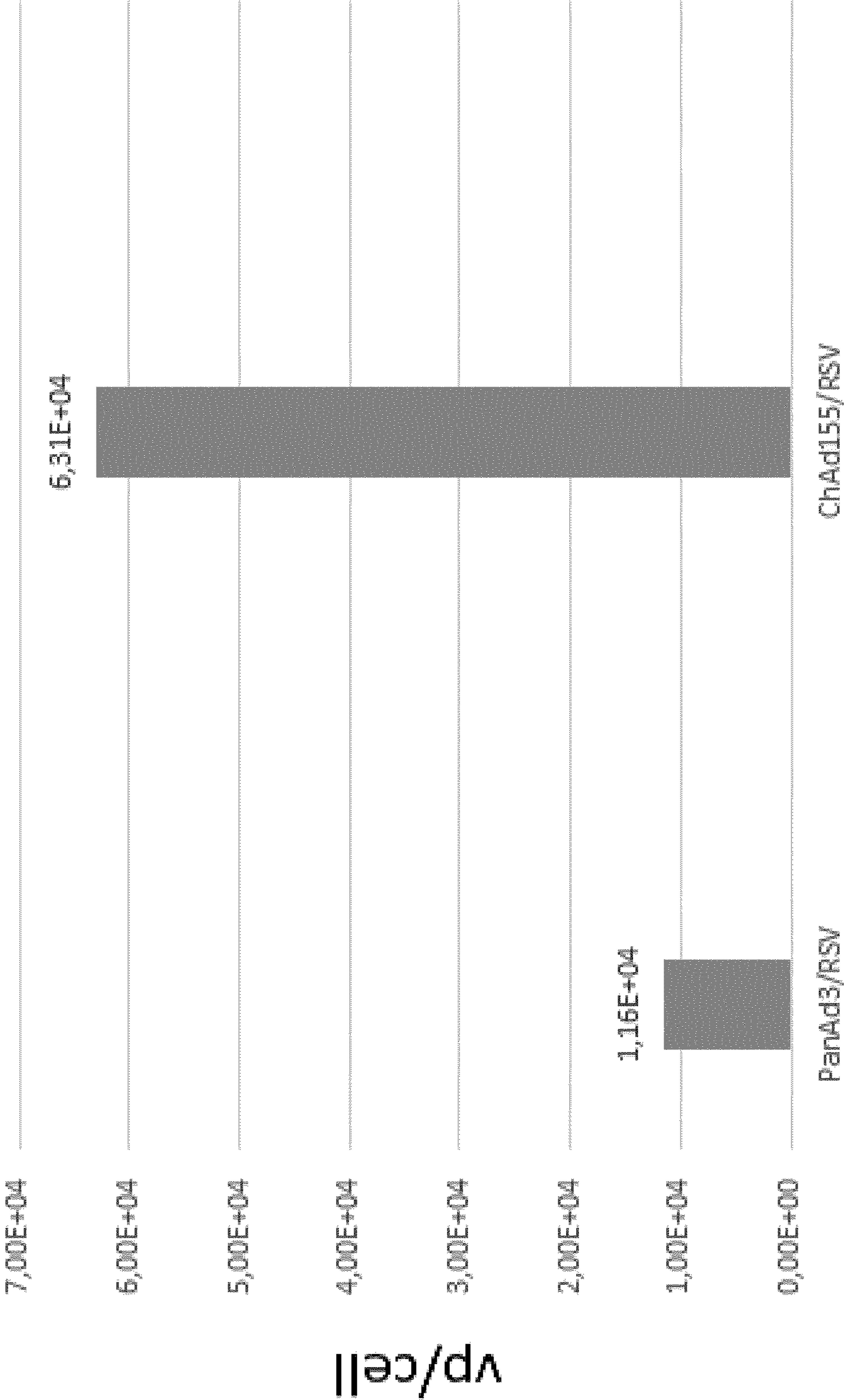
9/15
Figure 7



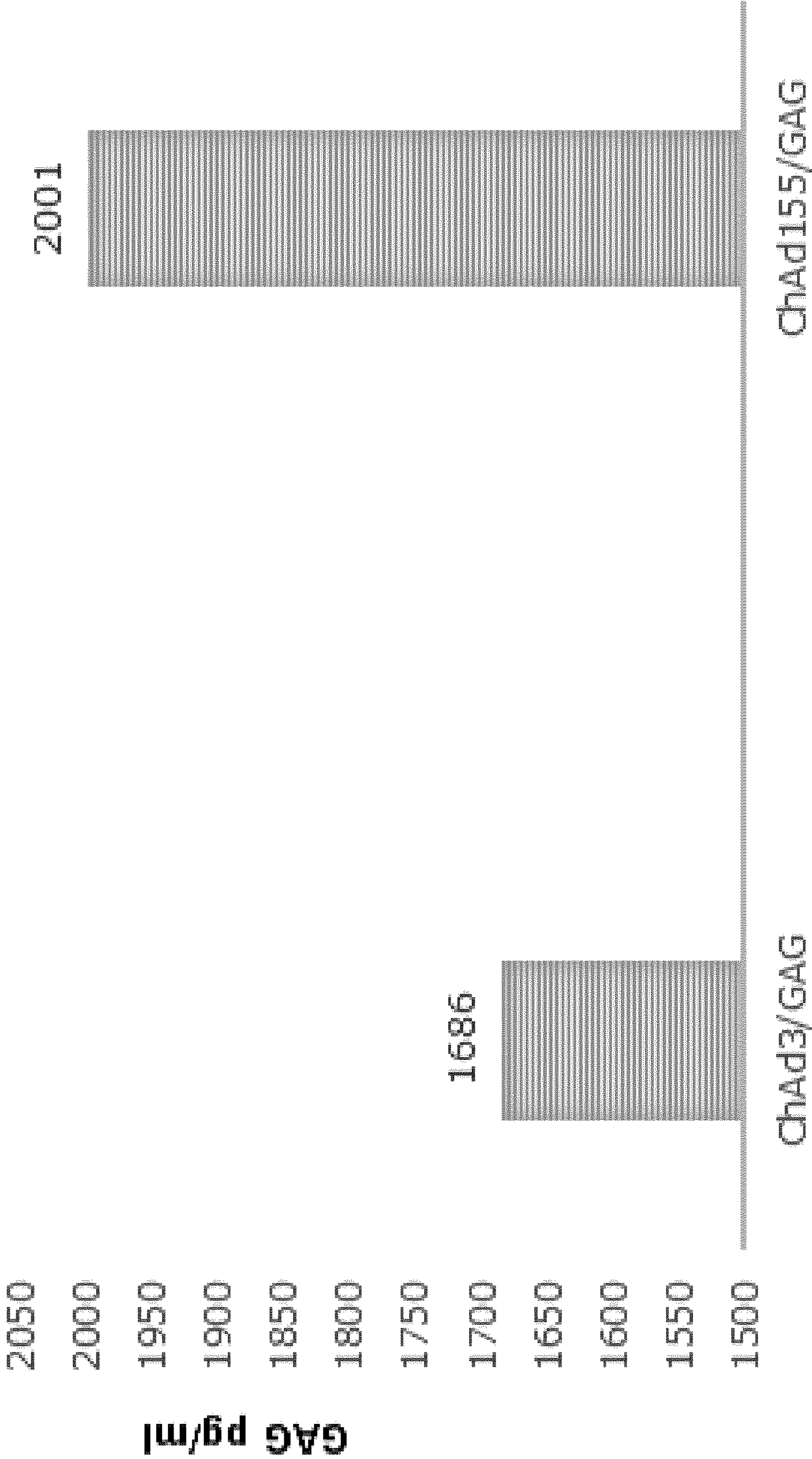
10/15
Figure 8



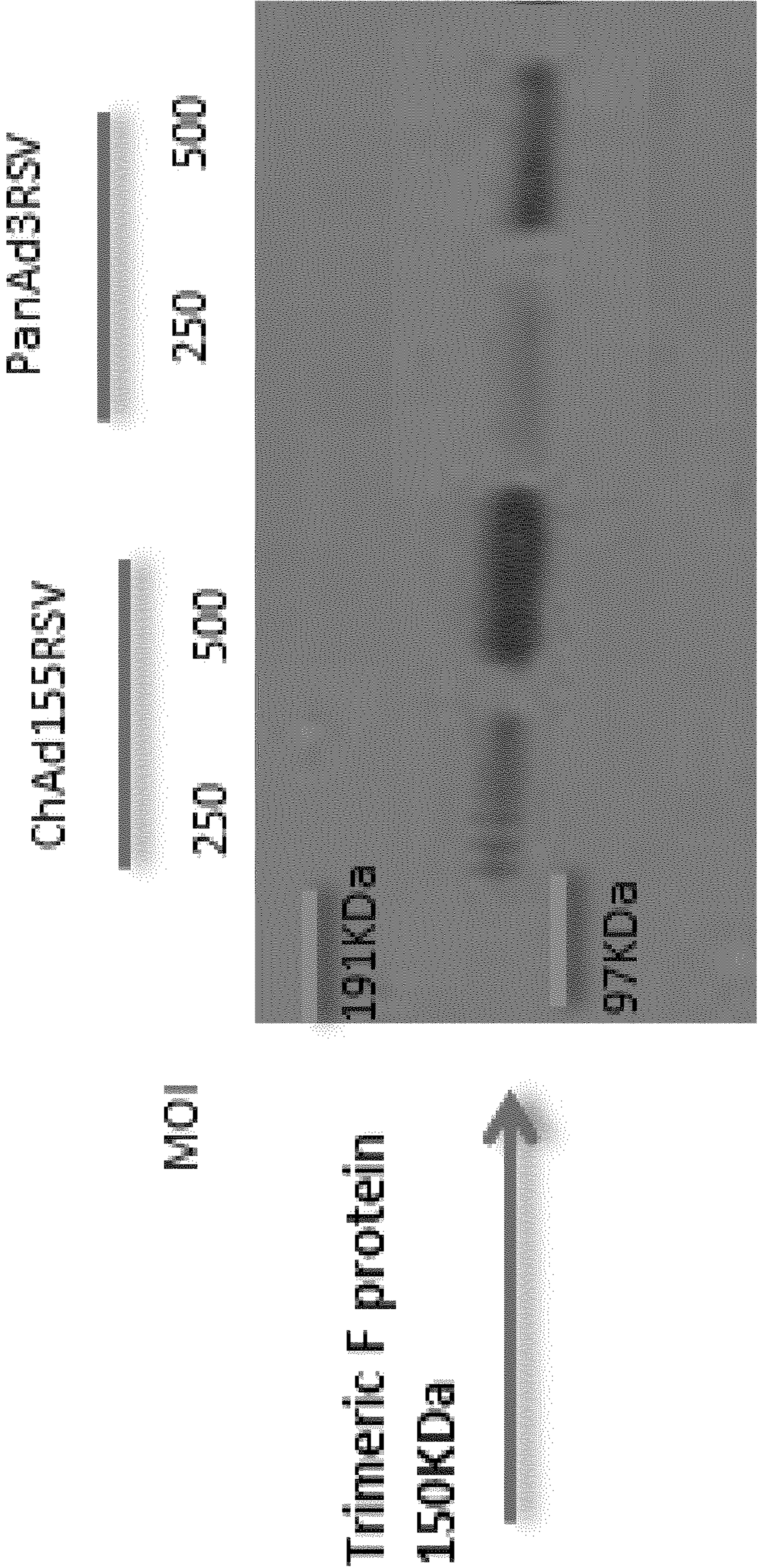
11/15
Figure 9



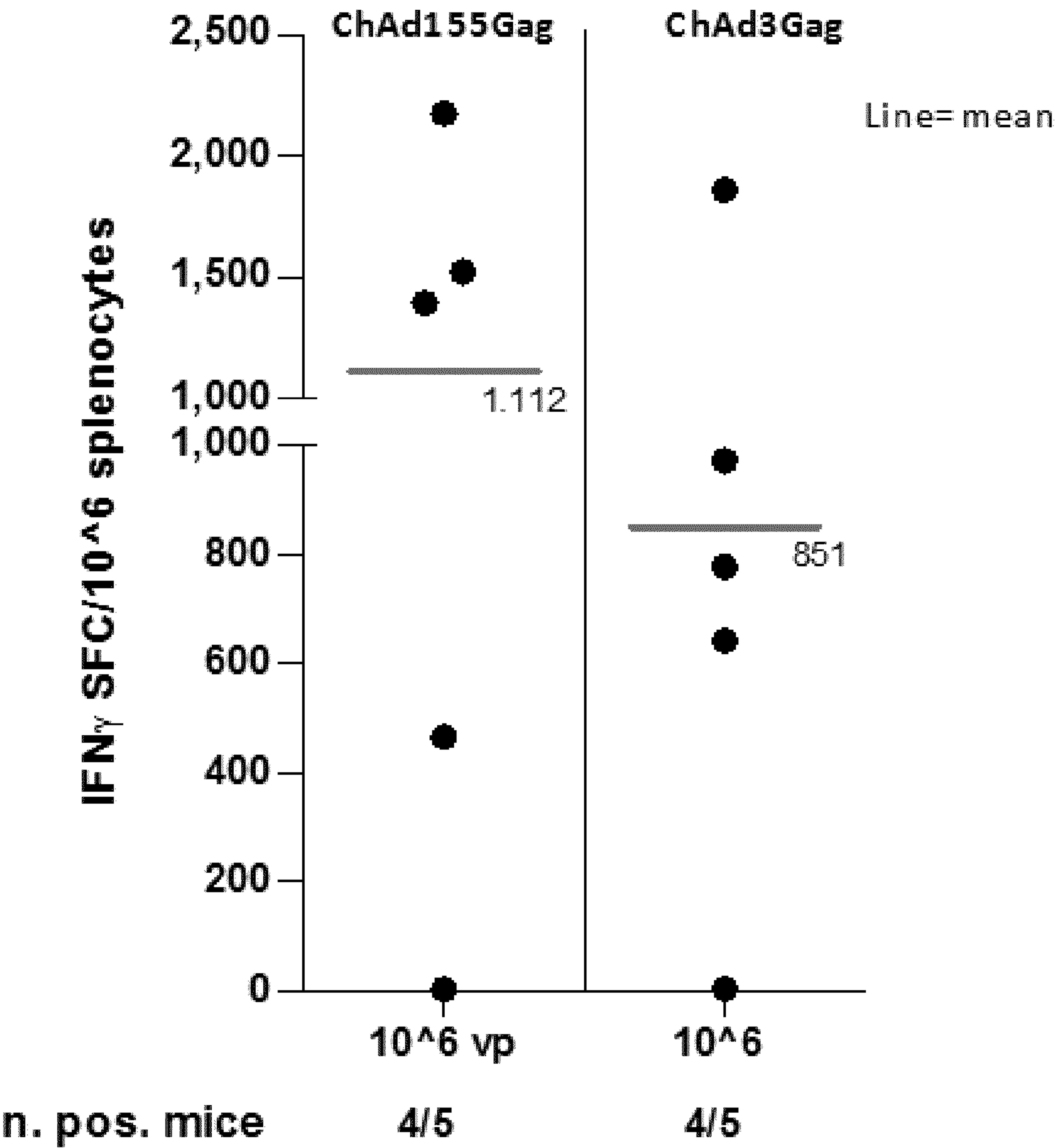
12/15
Figure 10



13/15
Figure 11



14/15
Figure 12



15/15

Figure 13

