*G00001*

# METHODS AND COMPOSITIONS FOR ENGINEERING OF ATTENUATED VACCINES

## STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH

Not Applicable.

## CROSS-REFERENCE TO RELATED APPLICATIONS

Not Applicable.

## BACKGROUND OF THE INVENTION

### Field of the Invention

This invention pertains to the field of vaccine development. Methods are provided for obtaining attenuated vaccines that exhibit improvements compared to previously available attenuated vaccines.

### Background

Edward Jenner demonstrated in 1796 that inoculation of a person with the cowpox virus would confer protection against its deadly relative, smallpox. Jenner's discovery was followed by Louis Pasteur's development in 1879-1881 of attenuated vaccines for chicken cholera, anthrax, and rabies. Since these early discoveries, attenuated vaccines have provided a significant addition to medicine's arsenal of weapons against a wide variety of infectious and other diseases.

Attenuated vaccines are otherwise pathogenic organisms that lack certain characteristics that are necessary to produce disease. Both bacteria and viruses are suitable for use as attenuated vaccines. Attenuated virus vaccines include, for example, vaccinia virus and other attenuated poxvirus vectors. One such poxvirus attenuated vaccine, the NYVAC vaccine, was obtained by attenuating the Copenhagen strain of vaccinia by complete deletion of the reading frames of 18 genes involved in virulence, tissue tropism and host range (Tartaglia *et al.*, *Virology* 188: 217-32 (1992)). The resultant vector, which is highly



*159159*

attenuated but nevertheless has the immunogenic potential of the original Copenhagen strain, is used in clinical trials as a human vaccine vector (Lanar *et al.*, *Infect. Immun.* 64: 1666-71 (1996); Limbach and Paoletti, *Epidemiol. Infect.* 116: 241-56 (1996); Paoletti *et al.*, *Dev. Biol. Stand.* 84:159-63 (1995)). Another poxvirus vector, the ALVAC vector, is a highly attenuated strain derived from canarypox and is licensed as a veterinary vaccine for canaries. It is a safe and efficacious vector in several mammalian species, including man. The TROVAC vector is based on a highly attenuated fowlpox vaccine strain and is used to immunize day-old chicks against fowlpox disease. TROVAC can also be used as a vector for vaccination of chicks against avian influenza virus and Newcastle disease virus (Paoletti *et al.*, *supra.*). Other attenuated viral vaccines include those that are useful against polio, measles, mumps, rubella, yellow fever and varicella. Examples of attenuated bacteria that are useful as vaccines include, for example, BCG (Bacillus Calmette-Guerin), *Salmonella* and *E. coli*.

Attenuated viruses and bacteria have been produced by the various methods, including UV irradiation (Hristov and Karadjov, *Vet. Med. Nauki*, 13: 8 (1975)), chemical attenuation by formalin or ethanol treatment (Zuschek *et al.*, *J. Am. Vet. Med. Assoc.* 139: 236 (1961); Haralambiev, *Acta Vet. Acad. Sci. Hung.* 26: 215 (1976)), and passage under conditions of stress *in vitro* or *in vivo*. An example of *in vitro* passage under stress was the development of the first attenuated virus licensed as a smallpox vaccine, which was obtained by conducting 36 passages of the original Lister strain (LO) of vaccinia in primary rabbit kidney cells at 30°C, followed by another 6 passages. Selection then yielded LC16m0 as a temperature sensitive and medium pock-forming virus. LC16m8 was cloned from LC16m0 as a small-pock forming variant (*see, e.g.*, Sugimoto, *Vaccine* 12: 675-681 (1994); Morita *et al.*, *Vaccine* 5: 65-70 (1987)). Another example of the development of a vaccine by serial passage *in vitro* is the passage of vaccinia virus through >570 passages in chick embryo fibroblasts to generate the modified vaccinia virus (MVA) that is host restricted, unable to replicate in human and other mammalian cells. MVA is avirulent in normal and immunosuppressed animals and produced no side effects, unlike the conventional smallpox vaccine strains. MVA had at least 30,000 bp of DNA deleted with loss of at least two host range genes.

Another approach to development of attenuated vaccines involves the specific deletion of genes known to confer virulence. For example, in the case of vaccinia virus the following genes have been deleted to alter the viral phenotype: TK, growth factor, hemagglutinin, 13.8 kb secreted protein, ribonucleotide reductase, envelope proteins, steroid dehydrogenase, complement control protein, host range genes (Moss, *Dev. Biol. Stand.* 82: 55-63 (1994); Blanchard *et al.*, *J. Gen. Virol.* 79: 1159-1167 (1998); Carroll and Moss, *Virology* 238: 198-211 (1997)). The example of development of the NYVAC vector, discussed above, is notable in that the entire genome was sequenced and relevant genes were precisely deleted. Thus, this is an example of the construction of an attenuated vaccine wherein the exact alterations are known and distance from the wild type understood. Such precise delineations of the genetic alterations of an attenuated strain are obviously more accessible technically with a virus than with the larger genomes of bacteria. Similarly, HSV lacking an essential glycoprotein (gH gene) can undergo a single round of replication in normal cells, but the virus particles derived from this infection are noninfectious. They are called DISC viruses (disabled infectious single cycle) and in the case of an HSV-2 lacking gH sequences have been shown in a guinea pig model of genital HSV-2 infection to protect against infection and against primary and recurrent disease (Boursnell *et al.*, *J. Infect. Dis.* 175: 16-25 (1997)).

The insertion of specific genes into the viral genome provides another approach to developing attenuated vaccines. For example, lymphokine genes have been inserted into the genome of vaccinia in order to decrease virulence without affecting immunogenicity. Murine or human IL-2 or interferon gamma have been inserted into the genome of vaccinia to produce virus of much lower pathogenicity yet unaltered immunogenicity (Moss, *Dev. Biol. Stand.*, 82: 55-63 (1994)). A similar approach involved the insertion of the B5R gene of the LO strain of vaccinia virus into LC16m8 infected RK13 cells with derivation of the LOTC virus strains (LOTC-1 through 5). The BR5 gene is responsible for plaque and pock size and host range and corresponds to the *ps/hr* gene (Sugimoto, *Vaccine* 12: 675-678 (1994)).

In some of the cases described above, defined genetic alterations have been performed with an already attenuated virus while in others, particularly with bacterial vaccines, there is little or no knowledge of the genetic lesions that confer attenuation.

Diverse types of genetic alterations are presumed to have been generated in the course of attenuation, including point mutations, DNA deletion and rearrangement. These methods of attenuation do not, in themselves, dictate the precise nature of the genetic alterations that confer the attenuated phenotype. Even though mutations may be chemically or UV irradiation induced via characterized chemical mechanisms, the positions of mutation are not controllable by current technologies. Nor are the number of sites in the genome that have been altered controlled or characterized. Although there may be a dose response relationship between concentration of mutagen, for example, and degree of phenotypic alteration, whether there are 10 or 100 genes whose function has been disrupted in the process of attenuation is unknown. Consequently, these mutagenesis methods do not control which genes or control elements have been modified. In the absence of sequencing the entire genome of the bacterium or virus (until recently not a practical or technical feasibility), the positions of mutations/deletions/rearrangements in the genome are unknown. An attenuated phenotype could therefore reflect modification in one virulence gene and many other genes not relevant to virulence, or modification of 10 virulence genes as well as many other genes unrelated to virulence. In such cases, the degree of difference between the attenuated genotypes is very different genetically but may appear little different phenotypically.

Other problems that are associated with current methods of attenuation include, for example, the reversion of attenuated organisms to wild type phenotype in vaccinees, with consequent disease pathology including severe morbidity and death. Vaccine strains of infectious bronchitis virus (IBV) are one example of attenuated vaccines that easily revert to more virulent strains *in vivo* (Hopkins and Yoder, *Avian Dis.*, 30: 221-3 (1986)). A related problem is the inability to monitor the stability of the genetic alterations in the attenuated vaccine organism during the process of developing, manufacturing, and distributing the vaccine over many years or decades. If one has not defined what is altered, it is impossible to identify whether the genotype is in flux and whether the organism is reverting closer to wild type phenotype. Vaccine manufacturers monitor the phenotype of the attenuated organism by standard protocols to ensure that the organism appears stable under standard conditions, but such methods are imprecise and do not allow any understanding of whether the number of mutations that separate the attenuated organism from wild type phenotype may have decreased to such a number that a very small number of additional

reversions in genotype might confer wild type phenotype). An additional problem with previously available methods for developing attenuated vaccines is a loss of immunogenicity due alteration of genes, by loss or altered sequence, that are important for elicitation of desired immune responses in vaccinees.

Another problem, the retention of replication competence in an attenuated vaccine or vaccine vector raises safety concerns, particularly in immunocompromised persons or animals in whom even a substantially attenuated virus or bacterium may cause disease. The retention of replication competence, on the other hand, can be an advantage for the stimulation of broad based and long lasting immunity. One approach to solving this dilemma is to engineer the virus to undergo only one round of replication such that its capacity to cause disease is eliminated yet capacity to immunize effectively enhanced compared to replication incompetent virus or inactivated virus. An example is the DISC virus vaccine approach, as exemplified with HSV-1 or HSV-2 viruses in which the essential gH gene is deleted. Reactivation of latent Herpes zoster virus elicited by immunization with a live attenuated varicella virus vaccine is another drawback to the use of attenuated vaccines (Garnett and Grenfell, *Epidemiol. Infect.* 108: 513-528 (1992)).

Although the potential problems associated with attenuated vaccines are significant, the attenuated vaccines have shown promise against infectious diseases for which other vaccines are not yet available. For example, an International AIDS Vaccine Initiative report indicates that some people infected with a weak strain of HIV have remained healthy for more than a dozen years, and at least one person with a weakened HIV strain may have successfully warded off multiple exposures to other HIV strains. This finding is supported by some primate studies. Accordingly, given the great need for an AIDS vaccine, attenuated vaccines are currently receiving a great deal of attention. However, a recent study which used attenuated versions of the HIV analog SIV, which infects monkeys, found that some monkeys may have acquired AIDS from the attenuated vaccine. Therefore, a need exists for improved attenuated vaccines, and for methods of developing such vaccines. The present invention fulfills these and other needs.

## SUMMARY OF THE INVENTION

The present invention provides methods for obtaining attenuated vaccines. The vaccines are useful for therapeutic and prophylactic purposes, and are effective against

pathogenic agents such as viruses, bacteria, parasites, and others. In some embodiments, the methods involve recombining a first set of one or more nucleic acid segments that comprise a complete or partial genomic library of a virus or a cell with at least a second set of one or more nucleic acid segments. Viruses or cells that contain members of the resulting library of recombinant nucleic acid fragments are then screened to identify those that are attenuated under physiological conditions that exist in a host organism. For example, the viruses or cells can be "attenuated" in that they are less able to propagate and/or cause disease in the host organism than a naturally occurring isolate of the viruses or cells.

In presently preferred embodiments, the attenuated viruses or cells are screened to identify those that can induce an immune response against a pathogenic agent that displays an immunogenic determinant that is also displayed by the attenuated viruses or cells. Attenuated viruses or cells that can induce the immune response are useful as attenuated vaccines against the pathogenic agent.

In some embodiments, the methods involve performing recursive recombination and screening/selection. This involves recombining polynucleotides from the attenuated cells or viruses obtained in a first round of recombination and screening/selection with a further set of one or more forms of a nucleic acid, to form a further library of recombinant nucleic acids. Viruses or cells that contain members of the further library of recombinant nucleic acid fragments are screened to identify those viruses or cells that are further attenuated under physiological conditions that exist in a host organism. The recombination and selection/screening can be repeated until the resulting attenuated viruses or cells have lost the ability to replicate or cause disease under physiological conditions that exist in the host organism.

The nucleic acid segments of at least one of the sets are, in some embodiments, obtained from a non-pathogenic strain of a virus or cell. In such cases, at least one of the other sets of nucleic acid segments are typically derived from a pathogenic agent, which can be of the same species as the nonpathogenic strain, or of a different species. One or more of the sets of nucleic acid segments can comprise a complete or substantially complete genomic library of the cell or virus from which the segments are derived, or can be a partial genomic library of the cell or virus. The nucleic acid segments can also include those that encode all or part of an immunogenic polypeptide that is displayed on a

pathogenic agent, or that encode a polypeptide, such as an immunomodulatory molecule or therapeutic protein, that has a desirable effect on an immune response induced by the vaccine.

In presently preferred embodiments of the invention, the attenuated viruses or cells are backcrossed to remove superfluous mutations. Backcrossing, according to the invention, involves recombining nucleic acids from the attenuated viruses or cells with a library of nucleic acids from a wild-type or naturally occurring strain of the virus or cell to form a further library of recombinant nucleic acids. Viruses or cells that contain members of the further library of recombinant nucleic acids are then screened to identify backcrossed viruses or cells that remain attenuated under physiological conditions present in an inoculated host organism. The library of nucleic acids from the wild-type strain is, in some embodiments, a partial or complete genomic library of the naturally occurring strain. The backcrossed attenuated viruses or cells can also be screened to identify those that can induce an immune response against a pathogenic agent that displays an immunogenic determinant that is also displayed by the attenuated viruses or cells. The backcrossing can be repeated one or more times, as desired.

The methods can also involve screening the attenuated viruses or cells to identify those that propagate under permissive conditions used for production of the attenuated viruses or cells, but do not propagate significantly in an inoculated host organism. The permissive conditions used for production can differ from the physiological conditions in the host in, for example, temperature, pH, sugar content, a compromised immune system, absence of complement or complement components, and presence or absence of serum proteins.

In some embodiments, the permissive condition used for the production of the attenuated vaccine is the presence of a suppressor tRNA molecule that can suppress termination of translation at non-naturally occurring stop codons that are introduced into the genome of the attenuated virus or cell. The nucleic acid segments subjected to recombination include one or more polynucleotides that encode all or part of a polypeptide that is involved in replication or pathogenicity of the virus or cells. The polypeptide-encoding polynucleotides can be included within a complete or partial genomic library of a virus or cell. Also included in the recombination is a population of oligonucleotides that have one or

more stop codons interspersed within the coding sequences for the polypeptide. The oligonucleotides undergo recombination with the polypeptide-encoding polynucleotides to form a library of recombinant nucleic acids in which at least one nonnaturally occurring stop codon is interspersed within the coding sequence of the replication polypeptide. The attenuated viruses or cells are obtained by contacting the library of recombinant nucleic acid fragments with suppressor tRNA molecules that suppress the termination of translation at the nonnaturally occurring stop codons and collecting progeny viruses or cells that propagate in the presence of the suppressor tRNA molecules but not in the absence of the suppressor tRNA molecules.

The invention also provides methods of obtaining an attenuated vaccine by introducing a library of nucleic acid fragments into a plurality of cells, whereby at least one of the fragments undergoes recombination with a segment in the genome or an episome of the cells to produce modified cells. The modified cells are screened to identify conditionally defective cells that have evolved toward loss of the ability to proliferate under physiological conditions as found in a host organism. The conditionally defective cells are, in turn, screened to identify those modified cells that have maintained the ability to replicate under permissive conditions used for production of the attenuated vaccine. The conditionally defective cells that replicate under permissive conditions but not in a host mammal are suitable for use as an attenuated vaccine organism.

In other embodiments, the invention provides methods of obtaining a chimeric attenuated vaccine. These methods generally involve recombining a first set of one or more nucleic acid segments from a virus or cell with at least a second set of one or more nucleic acid segments. The nucleic acid segments of the second set generally confer upon viruses or cells that contain the nucleic acid segments a property that is desirable for vaccination. A library of recombinant DNA fragments is thus formed. Attenuated viruses or cells are then identified by screening viruses or cells that contain members of the library of recombinant DNA fragments to identify those viruses or cells that are attenuated under physiological conditions present in a host organism inoculated with the viruses or cells. The attenuated viruses or cells are then screened to identify those that exhibit an improvement in the property that is desirable for vaccination. The screening can be conducted in any order.

In some embodiments, at least one of the sets of nucleic acid segments is a partial or substantially complete genomic library of a virus or cell. For example, one set can be from a pathogenic virus or cell, while another set is from a non-pathogenic isolate of virus or cell. The pathogenic and non-pathogenic isolates can be of the same or different species.

The recombination is performed, in some embodiments, by introducing the second set of nucleic acid segments into a plurality of nonpathogenic cells. At least one member of the second set of nucleic acid segments undergoes recombination with a segment in the genome or an episome of the nonpathogenic cells to produce modified cells. The modified cells are then screened to identify attenuated chimeric cells that are nonpathogenic and exhibit an improvement in the property that is desirable for vaccination.

In some embodiments, the methods of obtaining chimeric vaccines involve recursive recombination and screening. For example, one can recombine nucleic acids from the attenuated chimeric cells or viruses with a further set of nucleic acid segments to form a further library of recombinant nucleic acids. Attenuated chimeric viruses or cells that exhibit further improvement in attenuation or in the property that is desirable for vaccination are identified by screening viruses or cells that contain members of the further library of recombinant DNA fragments to identify those that exhibit further improved attenuation or desirable property. The recombination and screening can be repeated one or more times as desired until the attenuated chimeric viruses or cells have achieved a desired level of pathogenicity loss or improvement in the property that is desirable for vaccination.

The invention also provides attenuated and chimeric viruses and cells that are produced using the methods described herein. Also provided are vaccine compositions and methods of vaccinating using the vaccine compositions of the invention.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a schematic representation of recombinatorial shuffling of a collection of families of viral genomes having a variety of mutations or distinct genome portions; distinct genome segments (*e.g.*, obtained from the genomes of different virus isolates) are indicated by shaded boxes.

Figure 2 shows a schematic of a protocol for screening chimeric viral vaccines that are obtained using the recombination methods of the invention.

Figure 3 shows a summary of a strategy for using the methods of the invention to evolve a multivalent viral vaccine. HPV is used as an illustrative example. The arrows in the right lane of the figure point to each subsequent task after successful accomplishment of the previous. The broken arrows in the left lane outline steps of alternative methods if the previous task fails to lead to the next step.

Figure 4 shows a phylogenetic tree of papillomavirus (source: *Human Papillomavirus comp.* 1997).

Figure 5 shows a high throughput (HTP) *in vitro* screening assay for identifying recombinant nucleic acids that encode an improved antigen.

Figure 6 shows a schematic of a protocol for an antigen library immunization and cross-neutralization assay.

## DETAILED DESCRIPTION

### Definitions

A "pathogenic agent" refers to an organism or virus that is capable of infecting a host cell. Pathogenic agents are typically capable of causing a disease or other adverse effect on an infected cell or organism. Pathogenic agents include, for example, viruses, bacteria, fungi, parasites, and the like. The term "virus" includes not only complete virus particles, but also virus-like particles (VLPs) that include one or more viral polypeptides.

The term "attenuated," when used with respect to a virus or cell, means that the virus or cell has lost some or all of its ability to proliferate and/or cause disease or other adverse effect when the virus or cell infects an organism. For example, an "attenuated" virus or cell can be unable to replicate at all, or be limited to one or a few rounds of replication, when present in an organism in which a wild-type or other pathogenic version of the attenuated virus or cell can replicate. Alternatively or additionally, an "attenuated" virus or cell might have one or more mutations in a gene or genes that are involved in pathogenicity of the viruses or cells.

A "host organism" is an animal that is a target of vaccination with the attenuated and chimeric vaccines of the invention. Such host organisms have an immune system that is responsive to inoculation with an immunogen. Suitable host organisms

include, for example, humans, livestock, birds, and other animals in which it is desirable to vaccinate for either therapeutic or prophylactic purposes.

A "vaccine," as used herein, refers to an immunogen that, upon inoculation into a host organim, can induce complete or partial immunity to pathogenic agents, or can reduce the effects of diseases associated with pathogenic agents. Vaccines are also useful to alleviate immune system disorders other than those associated with pathogenic agents, such as autoimmune conditions.

The term "screening" describes, in general, a process that identifies vaccines that have optimal properties, such as attenuation. Selection is a form of screening in which identification and physical separation of attenuated vaccines are achieved simultaneously. For example, expression of a selection marker, which, in some genetic circumstances, allows cells expressing the marker to survive while other cells die (or *vice versa*) can be used as a selection method. Screening markers include, for example, genes that express luciferase, β-galactosidase and green fluorescent protein, or other gene products that are readily detected upon expression. Selection markers include, for example, drug and toxin resistance genes. Because of limitations in studying primary immune responses *in vitro*, *in vivo* studies are particularly useful screening methods. In these studies, the putative vaccines are introduced into test animals, and the immune responses are subsequently studied by analyzing protective immune responses or by studying the quality or strength of the induced immune response using, for example, lymphoid cells derived from the immunized animal. Although spontaneous selection can and does occur in the course of natural evolution, in the present methods selection is performed by man.

A "exogenous DNA segment", "heterologous sequence" or a "heterologous nucleic acid", as used herein, is one that originates from a source foreign to the particular host cell, or, if from the same source, is modified from its original form. Thus, a heterologous gene in a host cell includes a gene that is endogenous to the particular host cell, but has been modified. Modification of a heterologous sequence in the applications described herein typically occurs through the use of DNA shuffling. Thus, the terms refer to a DNA segment which is foreign or heterologous to the cell, or homologous to the cell but in a position within the host cell nucleic acid in which the element is not ordinarily found. Exogenous DNA segments are expressed to yield exogenous polypeptides.

The term "gene" is used broadly to refer to any segment of DNA associated with a biological function. Thus, genes include coding sequences and/or the regulatory sequences required for their expression. Genes also include nonexpressed DNA segments that, for example, form recognition sequences for other proteins. Genes can be obtained from a variety of sources, including cloning from a source of interest or synthesizing from known or predicted sequence information, and may include sequences designed to have desired parameters.

The term "isolated", when applied to a nucleic acid or protein, denotes that the nucleic acid or protein is essentially free of other cellular components with which it is associated in the natural state. It is preferably in a homogeneous state although it can be in either a dry or aqueous solution. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein which is the predominant species present in a preparation is substantially purified. In particular, an isolated gene is separated from open reading frames which flank the gene and encode a protein other than the gene of interest. The term "purified" denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Particularly, it means that the nucleic acid or protein is at least about 50% pure, more preferably at least about 85% pure, and most preferably at least about 99% pure.

The term "naturally-occurring" is used to describe an object that can be found in nature as distinct from being artificially produced by man. For example, an organism, or a polypeptide or polynucleotide sequence that is present in an organism (including viruses, bacteria, protozoa, insects, plants or mammalian tissue) that can be isolated from a source in nature and which has not been intentionally modified by man in the laboratory is naturally-occurring.

The term "nucleic acid" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form. Unless specifically limited, the term encompasses nucleic acids containing known analogues of natural nucleotides which have similar binding properties as the reference nucleic acid and are metabolized in a manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (*e.g.*

degenerate codon substitutions) and complementary sequences and as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer *et al.* (1991) *Nucleic Acid Res. 19*: 5081; Ohtsuka *et al.* (1985) *J. Biol. Chem. 260.* 2605-2608; Cassol *et al.* (1992) ; Rossolini *et al.* (1994) *Mol. Cell. Probes 8*: 91-98). The term nucleic acid is used interchangeably with gene, cDNA, and mRNA encoded by a gene.

"Nucleic acid derived from a gene" refers to a nucleic acid for whose synthesis the gene, or a subsequence thereof, has ultimately served as a template. Thus, an mRNA, a cDNA reverse transcribed from an mRNA, an RNA transcribed from that cDNA, a DNA amplified from the cDNA, an RNA transcribed from the amplified DNA, *etc.*, are all derived from the gene and detection of such derived products is indicative of the presence and/or abundance of the original gene and/or gene transcript in a sample.

A nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For instance, a promoter or enhancer is operably linked to a coding sequence if it increases the transcription of the coding sequence. Operably linked means that the DNA sequences being linked are typically contiguous and, where necessary to join two protein coding regions, contiguous and in reading frame. However, since enhancers generally function when separated from the promoter by several kilobases and intronic sequences may be of variable lengths, some polynucleotide elements may be operably linked but not contiguous.

A specific binding affinity between two molecules, for example, a ligand and a receptor, means a preferential binding of one molecule for another in a mixture of molecules. The binding of the molecules can be considered specific if the binding affinity is about $1 \times 10^4\ M^{-1}$ to about $1 \times 10^6\ M^{-1}$ or greater.

The term "recombinant" when used with reference to a cell or virus indicates that the cell or a cell infected by the virus, replicates a heterologous nucleic acid, or expresses a peptide or protein encoded by a heterologous nucleic acid. Recombinant cells and viruses can contain genes that are not found within the native (non-recombinant) form of the cell or virus. Recombinant cells and viruses can also contain genes found in the native form of the cell or virus wherein the genes are modified and re-introduced into the cell by

artificial means. The term also encompasses cells and viruses that contain a nucleic acid endogenous to the cell that has been modified without removing the nucleic acid from the cell; such modifications include those obtained by gene replacement, site-specific mutation, and related techniques.

A "recombinant expression cassette" or simply an "expression cassette" is a nucleic acid construct, generated recombinantly or synthetically, with nucleic acid elements that are capable of effecting expression of a structural gene in hosts compatible with such sequences. Expression cassettes include at least promoters and optionally, transcription termination signals. Typically, the recombinant expression cassette includes a nucleic acid to be transcribed (*e.g.*, a nucleic acid encoding a desired polypeptide), and a promoter. Additional factors necessary or helpful in effecting expression may also be used as described herein. For example, an expression cassette can also include nucleotide sequences that encode a signal sequence that directs secretion of an expressed protein from the host cell. Transcription termination signals, enhancers, and other nucleic acid sequences that influence gene expression, can also be included in an expression cassette.

The terms "identical" or percent "identity," in the context of two or more nucleic acid or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence, as measured using one of the following sequence comparison algorithms or by visual inspection.

The phrase "substantially identical," in the context of two nucleic acids or polypeptides, refers to two or more sequences or subsequences that have at least 60%, preferably 80%, most preferably 90-95% nucleotide or amino acid residue identity, when compared and aligned for maximum correspondence, as measured using one of the following sequence comparison algorithms or by visual inspection. Preferably, the substantial identity exists over a region of the sequences that is at least about 50 residues in length, more preferably over a region of at least about 100 residues, and most preferably the sequences are substantially identical over at least about 150 residues. In a most preferred embodiment, the sequences are substantially identical over the entire length of the coding regions.

For sequence comparison, typically one sequence acts as a reference sequence to which test sequences are compared. When using a sequence comparison algorithm, test

and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

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 computerized 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 visual inspection (*see generally* Ausubel *et al., infra*).

One example of an algorithm that is suitable for determining percent sequence identity and sequence similarity is the BLAST algorithm, which is described in Altschul *et al., J. Mol. Biol.* 215:403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://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 neighborhood word score threshold (Altschul *et al., supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then 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, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915).

In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (*see, e.g.*, Karlin & Altschul (1993) *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787). 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 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

Another indication that two nucleic acid sequences are substantially identical is that the two molecules hybridize to each other under stringent conditions. The phrase "hybridizing specifically to", refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (*e.g.*, total cellular) DNA or RNA. "Bind(s) substantially" refers to complementary hybridization between a probe nucleic acid and a target nucleic acid and embraces minor mismatches that can be accommodated by reducing the stringency of the hybridization media to achieve the desired detection of the target polynucleotide sequence.

"Stringent hybridization conditions" and "stringent hybridization wash conditions" in the context of nucleic acid hybridization experiments such as Southern and northern hybridizations are sequence dependent, and are different under different environmental parameters. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes* part I chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York. Generally, highly stringent hybridization and wash conditions are selected to be about 5° C lower than the thermal melting point ($T_m$)

for the specific sequence at a defined ionic strength and pH. Typically, under "stringent conditions" a probe will hybridize to its target subsequence, but to no other sequences.

The $T_m$ is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Very stringent conditions are selected to be equal to the $T_m$ for a particular probe. An example of stringent hybridization conditions for hybridization of complementary nucleic acids which have more than 100 complementary residues on a filter in a Southern or northern blot is 50% formamide with 1 mg of heparin at 42°C, with the hybridization being carried out overnight. An example of highly stringent wash conditions is 0.15M NaCl at 72°C for about 15 minutes. An example of stringent wash conditions is a 0.2x SSC wash at 65°C for 15 minutes (*see*, Sambrook, *infra.*, for a description of SSC buffer). Often, a high stringency wash is preceded by a low stringency wash to remove background probe signal. An example medium stringency wash for a duplex of, *e.g.*, more than 100 nucleotides, is 1x SSC at 45°C for 15 minutes. An example low stringency wash for a duplex of, *e.g.*, more than 100 nucleotides, is 4-6x SSC at 40°C for 15 minutes. For short probes (*e.g.*, about 10 to 50 nucleotides), stringent conditions typically involve salt concentrations of less than about 1.0 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3, and the temperature is typically at least about 30°C. Stringent conditions can also be achieved with the addition of destabilizing agents such as formamide. In general, a signal to noise ratio of 2x (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization. Nucleic acids which do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, *e.g.*, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code.

A further indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with, or specifically binds to, the polypeptide encoded by the second nucleic acid. Thus, a polypeptide is typically substantially identical to a second polypeptide, for example, where the two peptides differ only by conservative substitutions.

The phrase "specifically (or selectively) binds to an antibody" or "specifically (or selectively) immunoreactive with", when referring to a protein or peptide, refers to a binding reaction which is determinative of the presence of the protein in the presence of a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein and do not bind in a significant amount to other proteins present in the sample. Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular protein. For example, antibodies raised to the protein with the amino acid sequence encoded by any of the polynucleotides of the invention can be selected to obtain antibodies specifically immunoreactive with that protein and not with other proteins except for polymorphic variants. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays, Western blots, or immunohistochemistry are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. *See* Harlow and Lane (1988) *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, New York "Harlow and Lane"), for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity. Typically a specific or selective reaction will be at least twice background signal or noise and more typically more than 10 to 100 times background.

"Conservatively modified variations" of a particular polynucleotide sequence refers to those polynucleotides that encode identical or essentially identical amino acid sequences, or where the polynucleotide does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given polypeptide. For instance, the codons CGU, CGC, CGA, CGG, AGA, and AGG all encode the amino acid arginine. Thus, at every position where an arginine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of "conservatively modified variations." Every polynucleotide sequence described herein which encodes a polypeptide also describes every possible silent variation, except where otherwise noted. One of skill will recognize that each codon in a nucleic acid (except AUG, which is

ordinarily the only codon for methionine) can be modified to yield a functionally identical molecule by standard techniques. Accordingly, each "silent variation" of a nucleic acid which encodes a polypeptide is implicit in each described sequence.

Furthermore, one of skill will recognize that individual substitutions, deletions or additions which alter, add or delete a single amino acid or a small percentage of amino acids (typically less than 5%, more typically less than 1%) in an encoded sequence are "conservatively modified variations" where the alterations result in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. *See, e.g.*, Creighton (1984) *Proteins*, W.H. Freeman and Company. In addition, individual substitutions, deletions or additions which alter, add or delete a single amino acid or a small percentage of amino acids in an encoded sequence are also "conservatively modified variations".

A "subsequence" refers to a sequence of nucleic acids or amino acids that comprise a part of a longer sequence of nucleic acids or amino acids (*e.g.*, polypeptide) respectively.

**Description of the Preferred Embodiments**

The present invention provides a new approach to the development of attenuated vaccines. This approach is evolution-based, using methods such as DNA shuffling in particular, and is optimal for developing attenuated vaccines. DNA shuffling is a process for recursive recombination and mutation and is performed by random fragmentation of related DNA sequences followed by reassembly of the fragments by primerless PCR. As in natural evolution, the technique takes advantage of deletions, insertions, inversions and point mutations in the DNA sequences to generate large pools of recombinant sequences, from which the best for a particular purpose are identified by screening or selection that is based on the improved function sought. If further improvement is desired, the optimized nucleic acids can be subjected to new rounds of shuffling and selection. DNA shuffling is the most efficient known method to generate large libraries of homologous DNA sequences and, of particular relevance to the development of attenuated vaccines, it can also be applied to whole organisms, such as viruses, bacteria and other pathogens.

The use of such recombination/screening methods to generate attenuated vaccines provides significant advantages over previously available methods. Generally, little

information is available on particular mutations that might reduce pathogenicity while not hampering the replication of the vaccine viruses or other organisms in manufacturing cells or media or the capacity of the vaccine organisms to induce protective immune responses in host cells. The methods of the invention overcome this obstacle by eliminating any reliance on *a priori* assumptions regarding attenuation or the mechanisms that regulate replication and pathogenicity. One can simply carry out the recombination and screening/selection methods of the invention and obtain an attenuated vaccine that has the desired properties.

Moreover, attenuated vaccines obtained using previously available methods, *e.g.*, those that involve mutagenesis, generally have many diverse types of mutation both in genes that are involved in virulence and in other genes. For example, the use of chemical or ultraviolet irradiation-induced mutagenesis to develop attenuated vaccines does not provide any information as to how many genes that are not involved in pathogenesis are mutated in order to obtain a mutation in a pathogenicity gene. This is particularly true in the case of bacterial vaccines. Because the efficacy of an attenuated vaccine depends on the presence of particular antigens that remain immunogenic in the vaccinated host, it is desirable to maintain as much of the wild-type amino acid sequence as possible. Thus, the introduction of mutations into genes that encode these antigens that can result from previously available attenuation methods can decrease the immunogenicity of the vaccine.

This problem is avoided by the present invention, which provides attenuated vaccines that generally have only one or a few mutations that result in the attenuation. In presently preferred embodiments, the recombinant vaccines that are obtained using the methods of the invention are subjected to molecular backcrossing. This allows one to move a mutant gene or genes that are responsible for conferring attenuation back to a parental or wildtype genome, thus retaining those few mutations that are critical to the desired evolved attenuation phenotype while eliminating at least some of the mutations that are not involved in attenuation. Thus, backcrossing can be used to retain the sequences from the wild-type organism that are critical for induction of protective immune responses, while also retaining those mutant genes that are responsible for the attenuated phenotype of the vaccine organism. Backcrossing can also be used to identify the mutations that are critical to the desired phenotype.

Other problems that are associated with current methods of attenuation are also avoided by the methods and attenuated vaccines of the invention. For example, the methods of the invention can greatly reduce or eliminate the reversion of attenuated organisms to the wild type, disease-causing phenotype in vaccinees that often occurs with attenuated vaccines prepared using previously available methods. Moreover, because the attenuated vaccines obtained using the methods of the invention can have well-characterized mutations, one can monitor the stability of the genetic alterations in the attenuated vaccine organism during the process of developing, manufacturing, and distributing the vaccine over many years or decades. Thus, the invention provides a means by which one can obtain an attenuated vaccine that has improved ability to induce an immune response to a pathogenic agent without causing the potential problems associated with previously available attenuated vaccines.

### *I. General Approach to Attenuated Vaccine Evolution*

Attenuated vaccines of the invention are created by first creating a library of recombinant nucleic acids. The library is created by recombining two or more variant forms of a nucleic acid are recombined to produce a library of recombinant nucleic acids. The library is then screened to identify those recombinant nucleic acids that include mutations that result in attenuation of a potential vaccine virus or other organism. For example, the recombinant nucleic acid can include one or more mutations that render a normally pathogenic organism non-pathogenic. Importantly, complete or partial genomes of viruses, bacteria, fungi, parasites or other pathogens can be fragmented and subjected to the recombination and screening methods of the invention. Single genes or other nucleic acid fragments can also be subjected to the recombination and screening methods, either alone or in combination with a complete or partial genome of a virus or organism. Recombination and selection of single pathogen gene is useful in cases when the critical genes that regulate pathogenicity are known. For example, single gene shuffling is useful when the protein responsible for binding to the natural host cell is known. Whole genome shuffling is particularly useful when, for example, the genome is relatively small and little is known of the critical sequences affecting attenuation.

A number of different formats are available by which one can create a library of recombinant nucleic acids for screening or selection. In some embodiments, the methods

of the invention entail performing recombination ("shuffling" or "sequence recombination") and screening or selection to "evolve" individual genes, whole plasmids or viruses, multigene clusters, or even whole genomes (Stemmer (1995) *Bio/Technology* 13:549-553; PCT US98/00852; US Patent Appl. No. 09/116,188, filed July 15, 1998). Reiterative cycles of recombination and screening/selection can be performed to further evolve the nucleic acids of interest. Such techniques do not require the extensive analysis and computation required by conventional methods for polypeptide engineering. Shuffling allows the recombination of large numbers of mutations in a minimum number of selection cycles, in contrast to traditional, pairwise recombination events (*e g.*, as occur during sexual replication). Thus, the sequence recombination techniques described herein provide particular advantages in that they provide recombination between any or all of the mutations, thereby providing a very fast way of exploring the manner in which different combinations of mutations can affect a desired result. In some instances, however, structural and/or functional information is available which, although not required for sequence recombination, provides opportunities for modification of the technique.

Sequence recombination can be achieved in many different formats and permutations of formats, as described in further detail below. These formats share some common principles. A group of nucleic acid molecules that have some sequence identity to each other, but differ in the presence of mutations, is typically used as a substrate for recombination. In any given cycle, recombination can occur *in vivo* or *in vitro*, intracellularly or extracellularly. Furthermore, diversity resulting from recombination can be augmented in any cycle by applying other methods of mutagenesis (*e.g.*, error-prone PCR or cassette mutagenesis) to either the substrates or products of recombination. In some instances, a new or improved property or characteristic can be achieved after only a single cycle of *in vivo* or *in vitro* recombination, as when using different, variant forms of the sequence, as homologs from different individuals or strains of an organism, or related sequences from the same organism, as allelic variations. However, recursive sequence recombination, which entails successive cycles of recombination and selection/screening, can generate further improvement.

The DNA shuffling methods can involve one or more of at least four different approaches to improve attenuation of an otherwise pathogenic vaccine candidate, as well as

improve other properties that are of interest for a vaccine (*e.g.*, increased immunogenicity). First, DNA-shuffling can be performed on a single gene. Secondly, several highly homologous genes can be identified by sequence comparison with known homologous genes. These genes can be synthesized and shuffled as a family of homologs, to select recombinants with the desired activity. This "family shuffling" procedure is shown schematically in Figure 1. The shuffled genes can be introduced into appropriate host cells, which can include *E. coli*, yeast, plants, fungi, animal cells, and the like, and those having the desired properties can be identified by the methods described herein. Third, whole genome shuffling can be performed to shuffle genes that are involved in pathogenicity (along with other genomic nucleic acids), thus obtaining mutated pathogenicity genes that reduce or eliminate pathogenicity of the organism. For whole genome shuffling approaches, it is not even necessary to identify which genes are being shuffled. Instead, *e.g.*, bacterial cell or viral genomes are combined and shuffled to acquire recombinant nucleic acids that, either itself or through encoding a polypeptide, have enhanced ability to induce an immune response, as measured in any of the assays described herein. Fourth, polypeptide-encoding genes can be codon modified to access mutational diversity not present in any naturally occurring gene.

Exemplary formats and examples for sequence recombination, sometimes referred to as DNA shuffling, evolution, or molecular breeding, have been described by the present inventors and co-workers in co-pending applications U.S. Patent Application Serial No. 08/198,431, filed February 17, 1994; Serial No. PCT/US95/02126, filed February 17, 1995; Serial No. 08/425,684, filed April 18, 1995; Serial No. 08/537,874, filed October 30, 1995; Serial No. 08/564,955, filed November 30, 1995; Serial No. 08/621,859, filed March 25, 1996; Serial No. 08/621,430, filed March 25, 1996; Serial No. PCT/US96/05480, filed April 18, 1996; Serial No. 08/650,400, filed May 20, 1996; Serial No. 08/675,502, filed July 3, 1996; Serial No. 08/721, 824, filed September 27, 1996; Serial No. PCT/US97/17300, filed September 26, 1997; and Serial No. PCT/US97/24239, filed December 17, 1997. *See also*, Stemmer, *Science* 270: 1510 (1995); Stemmer *et al.*, *Gene* 164: 49-53 (1995); Stemmer, *Bio/Technology* 13: 549-553 (1995); Stemmer, *Proc. Natl. Acad. Sci. U.S.A.* 91:10747-10751 (1994); Stemmer, *Nature* 370: 389-391 (1994); Crameri *et al.*, *Nature*

*Medicine* 2(1):1-3 (1996); Crameri *et al.*, *Nature Biotechnology* 14: 315-319 (1996). Each of these references is incorporated herein by reference in its entirety for all purposes.

Other methods for obtaining libraries of recombinant polynucleotides and/or for obtaining diversity in nucleic acids used as the substrates for shuffling include, for example, homologous recombination (PCT/US98/05223; Publ. No. WO98/42727); oligonucleotide-directed mutagenesis (for review *see*, Smith, *Ann. Rev. Genet.* 19: 423-462 (1985); Botstein and Shortle, *Science* 229: 1193-1201 (1985); Carter, *Biochem. J.* 237: 1-7 (1986); Kunkel, "The efficiency of oligonucleotide directed mutagenesis" *in Nucleic acids & Molecular Biology*, Eckstein and Lilley, eds., Springer Verlag, Berlin (1987)). Included among these methods are oligonucleotide-directed mutagenesis (Zoller and Smith, *Nucl. Acids Res.* 10: 6487-6500 (1982), *Methods in Enzymol.* 100: 468-500 (1983), and *Methods in Enzymol.* 154: 329-350 (1987)) phosphothioate-modified DNA mutagenesis (Taylor *et al.*, *Nucl. Acids Res.* 13: 8749-8764 (1985); Taylor *et al.*, *Nucl. Acids Res.* 13: 8765-8787 (1985); Nakamaye and Eckstein, *Nucl. Acids Res.* 14: 9679-9698 (1986); Sayers *et al.*, *Nucl. Acids Res.* 16: 791-802 (1988); Sayers *et al.*, *Nucl. Acids Res.* 16: 803-814 (1988)), mutagenesis using uracil-containing templates (Kunkel, *Proc. Nat'l. Acad. Sci. USA* 82: 488-492 (1985) and Kunkel *et al.*, *Methods in Enzymol.* 154: 367-382)); mutagenesis using gapped duplex DNA (Kramer *et al.*, *Nucl. Acids Res.* 12: 9441-9456 (1984); Kramer and Fritz, *Methods in Enzymol.* 154: 350-367 (1987); Kramer *et al.*, *Nucl. Acids Res.* 16: 7207 (1988)); and Fritz *et al.*, *Nucl. Acids Res.* 16: 6987-6999 (1988)). Additional suitable methods include point mismatch repair (Kramer *et al.*, *Cell* 38: 879-887 (1984)), mutagenesis using repair-deficient host strains (Carter *et al.*, *Nucl. Acids Res.* 13: 4431-4443 (1985); Carter, *Methods in Enzymol.* 154: 382-403 (1987)), deletion mutagenesis (Eghtedarzadeh and Henikoff, *Nucl. Acids Res.* 14: 5115 (1986)), restriction-selection and restriction-purification (Wells *et al.*, *Phil. Trans. R. Soc. Lond.* A 317: 415-423 (1986)), mutagenesis by total gene synthesis (Nambiar *et al.*, *Science* 223: 1299-1301 (1984); Sakamar and Khorana, *Nucl. Acids Res.* 14: 6361-6372 (1988); Wells *et al.*, *Gene* 34: 315-323 (1985); and Grundström *et al.*, *Nucl. Acids Res.* 13: 3305-3316 (1985). Kits for mutagenesis are commercially available (*e.g.*, Bio-Rad, Amersham International, Anglian Biotechnology).

The recombination procedure starts with at least two nucleic acid substrates that generally show substantial sequence identity to each other (*i.e.*, at least about 30%, 50%, 70%, 80% or 90% sequence identity), but differ from each other at certain positions. The difference can be any type of mutation, for example, substitutions, insertions and deletions. Often, different segments differ from each other in about 5-20 positions. For recombination to generate increased diversity relative to the starting materials, the starting materials must differ from each other in at least two nucleotide positions. That is, if there are only two substrates, there should be at least two divergent positions. If there are three substrates, for example, one substrate can differ from the second at a single position, and the second can differ from the third at a different single position. The starting DNA segments can be natural variants of each other, for example, allelic or species variants. The segments can also be from nonallelic genes showing some degree of structural and usually functional relatedness (*e.g.*, different genes within a superfamily, such as the family of human papillomavirus L1 and L2-encoding genes, for example). The starting DNA segments can also be induced variants of each other. For example, one DNA segment can be produced by error-prone PCR replication of the other, the nucleic acid can be treated with a chemical or other mutagen, or by substitution of a mutagenic cassette. Induced mutants can also be prepared by propagating one (or both) of the segments in a mutagenic strain, or by inducing an error-prone repair system in the cells. In these situations, strictly speaking, the second DNA segment is not a single segment but a large family of related segments. The different segments forming the starting materials are often the same length or substantially the same length. However, this need not be the case; for example; one segment can be a subsequence of another. The segments can be present as part of larger molecules, such as vectors, or can be in isolated form.

The starting DNA segments are recombined by any of the sequence recombination formats provided herein to generate a diverse library of recombinant DNA segments. Such a library can vary widely in size from having fewer than 10 to more than $10^5$, $10^9$, $10^{12}$ or more members. In some embodiments, the starting segments and the recombinant libraries generated will include full-length coding sequences and any essential regulatory sequences required for expression, such as a promoter and polyadenylation sequence. In other embodiments, the recombinant DNA segments in the library can be

inserted into a common vector providing sequences necessary for expression before performing screening/selection.

A further technique for recombining mutations in a nucleic acid sequence utilizes "reassembly PCR." This method can be used to assemble multiple segments that have been separately evolved into a full length nucleic acid template such as a gene. This technique is performed when a pool of advantageous mutants is known from previous work or has been identified by screening mutants that may have been created by any mutagenesis technique known in the art, such as PCR mutagenesis, cassette mutagenesis, doped oligo mutagenesis, chemical mutagenesis, or propagation of the DNA template *in vivo* in mutator strains. Boundaries defining segments of a nucleic acid sequence of interest preferably lie in intergenic regions, introns, or areas of a gene not likely to have mutations of interest. Preferably, oligonucleotide primers (oligos) are synthesized for PCR amplification of segments of the nucleic acid sequence of interest, such that the sequences of the oligonucleotides overlap the junctions of two segments. The overlap region is typically about 10 to 100 nucleotides in length. Each of the segments is amplified with a set of such primers. The PCR products are then "reassembled" according to assembly protocols such as those discussed herein to assemble randomly fragmented genes. In brief, in an assembly protocol the PCR products are first purified away from the primers, by, for example, gel electrophoresis or size exclusion chromatography. Purified products are mixed together and subjected to about 1-10 cycles of denaturing, reannealing, and extension in the presence of polymerase and deoxynucleoside triphosphates (dNTP's) and appropriate buffer salts in the absence of additional primers ("self-priming"). Subsequent PCR with primers flanking the gene are used to amplify the yield of the fully reassembled and shuffled genes.

In a further embodiment, PCR primers for amplification of segments of the nucleic acid sequence of interest are used to introduce variation into the gene of interest as follows. Mutations at sites of interest in a nucleic acid sequence are identified by screening or selection, by sequencing homologues of the nucleic acid sequence, and so on. Oligonucleotide PCR primers are then synthesized which encode wild type or mutant information at sites of interest. These primers are then used in PCR mutagenesis to generate libraries of full length genes encoding permutations of wild type and mutant information at the designated positions. This technique is typically advantageous in cases where the

screening or selection process is expensive, cumbersome, or impractical relative to the cost of sequencing the genes of mutants of interest and synthesizing mutagenic oligonucleotides.

In a presently preferred embodiment, DNA shuffling is used to obtain the library of recombinant nucleic acids. DNA shuffling can result in attenuation of a pathogen even in the absence of a detailed understanding of the mechanism by which the pathogenicity is mediated. Examples of candidate substrates for acquisition of a property or improvement in a property include bacterial, viral and nonviral vectors used in genetic and classical types of vaccination, as well as nucleic acids that are involved in mediating a particular aspect of an immune response (*e.g.*, a nucleic acid that encodes an antigen). The methods require at least two variant forms of a starting substrate. The variant forms of candidate components can have substantial sequence or secondary structural similarity with each other, but they should also differ in at least two positions. The initial diversity between forms can be the result of natural variation, *e.g.*, the different variant forms (homologs) are obtained from different individuals or strains of an organism (including geographic variants; termed "family shuffling" (Figure 1)) or constitute related sequences from the same organism (*e.g.*, allelic variations). Alternatively, the initial diversity can be induced, *e.g.*, the second variant form can be generated by error-prone transcription, such as an error-prone PCR or use of a polymerase which lacks proof-reading activity (*see*, Liao (1990) *Gene* 88:107-111), of the first variant form, or, by replication of the first form in a mutator strain.

**1. Attenuated viral vaccines**

In some embodiments, the invention provides attenuated viral vaccines and methods for obtaining the attenuated viral vaccines. By using the methods of the invention, one can generate novel variant viruses having genotypes and phenotypes that do not naturally occur or would not otherwise be anticipated to occur at a substantial frequency. A preferred aspect of the method employs recursive nucleotide sequence recombination, termed "DNA shuffling," which enables the rapid generation of a collection of broadly diverse viral phenotypes that can be selectively bred for a broader range of novel phenotypes or more extreme phenotypes than would otherwise occur by natural evolution in the same time period. The method typically involves: (1) shuffling of a plurality of viral genomes, and (2) selection of the resultant shuffled viral genomes to isolate or enrich a plurality of shuffled viral genomes having a desired phenotype(s) (*e.g.*, attenuation), and optionally (3) repeating

steps (1) and (2) on the plurality of shuffled viral genomes conferring on a virus the desired phenotype(s) until one or more variant viral genomes conferring a sufficiently optimized desired phenotype is obtained. In this general manner, the method facilitates the "forced evolution" of a viral genome to encode an attenuated virus which natural selection and evolution has heretofore not generated. Figure 2 shows a block diagram of a basic method for viral genome shuffling and selection for a desired phenotype; the recursion option is generally selected each cycle until one or more viral genomes conferring a satisfactory optimization for the desired phenotype(s) are obtained.

Typically, a plurality of viral genomes of the same taxonomic classification are shuffled and selected by the present method. It is believed that a common use of the method will be to shuffle mutant variants of a clinical isolate(s) or of a laboratory strain of a virus to obtain a variant of the clinical isolate or laboratory strain that possesses a novel desired phenotype (*e.g.*, attenuation). However, the method can be used with a plurality of strains (or clades) of a virus, or even with a plurality of related viruses (*e.g.*, lentiviruses, herpesviruses, adenoviruses, *etc.*), and in some instances with unrelated viruses or portions thereof which have recombinogenic portions (either naturally or generated via genetic engineering). The method can be used to shuffle xenogeneic viral sequences into a viral genome (*e.g.*, incorporating and evolving a gene of a first virus in the genome of a second virus so as to confer a desired phenotype to the evolved genome of the second virus). Furthermore, the method can be used to evolve a heterologous nucleic acid (*e.g.*, a non-naturally occurring mutant viral gene) to optimize its phenotypic expression (*e.g.*, immunogenicity) in a viral genome, and/or in a particular host cell or expression system (*e.g.*, an expression cassette or expression replicon). Figure 1 shows a schematic representation of recombinatorial shuffling of a collection of families of viral genomes having a variety of mutations or distinct genome portions; distinct genome segments (*e.g.*, obtained from the genomes of different virus isolates) are indicated by shaded boxes.

Availability of infectious cDNA clones of RNA viruses is useful, but not necessary, for the development of improved strains of attenuated viral vaccines. Infectious cDNA clones have been established, for example, from porcine reproductive and respiratory syndrome virus (Meulenberg *et al.*, *Adv. Exp. Med. Biol.* (1998) 440:199-206), hepatitis C virus (Yanagi *et al.*, *Proc. Nat'l. Acad. Sci. USA* (1999) 96:2291-5), tick-borne encephalitis

virus (Gritsun *et al.*, *J. Virol. Methods* (1998) 76:109-20; Mandl *et al.*, *J. Gen. Virol.* (1997) 78:1049-57), plum pox potyvirus (Guo *et al.*, *Virus Res.* (1998) 57:183-95), respiratory syncytial virus (Jin *et al.*, *Virology* (1998) 251:206-14), paramyxovirus (He *et al.*, *Virology* (1998) 250:30-40), bovine viral diarrhea virus (Zhong *et al.*, *J. Virol.* (1998) 72:9365-9), feline calicivirus (Sosnovtsev *et al.*, *J. Virol.* (1998) 72:3051-9), infectious bursal disease virus (Yao *et al.*, *J. Virol.* (1998) 72:2647-54), dengue virus type 2 (Gualano *et al.*, *J. Gen. Virol.* (1998) 79:437-46), swine fever virus (Mittelholzer *et al.*, *Virus. Res.* (1997) 51:125-37), coxsackievirus B3 (Lee *et al.*, *Virus. Res.* (1997) 50:225-35), Hoffman and Banerjee, *J. Virol.* (1997) 71:4272-7), equine arteritis virus (van Dinten *et al.*, *Proc. Nat'l. Acad. Sci. USA* (1997) 94: 991-6), yellow fever virus (Galler *et al.*, *Braz. J. Med. Biol. Res.* (1997) 30:157-68), human astrovirus serotype 1 (Geigenmuller *et al.*, *J. Virol.* (1997) 71: 1713-7). The fact that infectious cDNA clones of these viruses have been established indicates that same approaches can be used to generate infectious cDNA clones of other viruses belonging to the same families and their shuffled variants. Therefore, family shuffling of these and related viruses provides an excellent starting point for development of attenuated vaccine strains.

The methods of the invention are applicable to generation of attenuated versions of many different viruses. Examples of viruses that are of particular interest include, but are not limited to, rotavirus, parvovirus B19, herpes simplex-1 and -2, CMV, RSV, varicella zoster virus, influenza viruses, HPV, HIV, EBV, hepatitis A, B, C, D & E virus. Also of particular interest are viruses of the picornavirus family, which includes the following genera: Rhinoviruses, which are responsible for approximately 50% cases of the common cold, and thus are of interest for medical applications; Enteroviruses, including polioviruses and coxsackieviruses; echoviruses and human enteroviruses such as hepatitis A virus; and the Apthoviruses, which are the foot and mouth disease viruses, and thus of interest particularly for veterinary uses (target antigens include VP1, VP2, VP3, VP4 and VPG. The Calcivirus family, which includes the Norwalk group of viruses which are an important causative agent of epidemic gastroenteritis, are also of particular interest for use as attenuated vaccines. Other viruses for which the methods of the invention are useful for generating attenuated vaccines include, for example , bovine viral diarrhea virus, Marek's Disease Virus (MDV), bovine herpes virus type-1 (BHV-1), infectious bronchitis virus,

infectious bursal disease virus (IBDV), porcine reproductive and respiratory syndrome virus, canine cistemper virus (CDV).

Also of interest for use as attenuated vaccines are the Togavirus family, including the following genera: Alphaviruses, which are of interest for both medical and veterinary use and include, for example, Senilis viruses, the RossRiver virus and Eastern & Western equine virus, as well as the Reovirus family, which includes Rubella virus. The Flariviridue family of viruses are also of particular interest for the development of attenuated vaccines using the methods of the invention. Examples include: for example, dengue, yellow fever, Japanese encephalitis, St. Louis encephalitis and tick borne encephalitis viruses. The hepatitis C virus, when attenuated using the methods of the invention, is also of particular interest for medical use.

Attenuated viral vaccines of interest also include those of the Coronavirus family, which find use for both medical and veterinary applications. Examples of coronaviruses that are useful for veterinary applications include, but are not limited to, infectious bronchitis virus (poultry), porcine transmissible gastroenteric virus (pig), porcine hemagglutinatiny encephalomyelitis virus (pig), feline infectious peritonitis virus (cats), feline enteric coronavirus (cat) and canine coronavirus (dog). For medical use, coronavirus family members of particular interest include, for example, the human respiratory coronaviruses, which cause about 40 percent of cases of the common cold (*see, e.g.*, Winther *et al.*, *Am. J. Rhinol.* 12: 17-20 (1998)). Coronaviruses may also cause non-A, B or C hepatitis. Target antigens include, for example, E1 (also called M or matrix protein), E2 (also called S or Spike protein), E3 (also called HE or hemagglutinelterose glycoprotein (not present in all coronaviruses), and N-nucleocapsid.

The Rhabdovirus family is another family of viruses that, when attenuated using the methods of the invention, are useful as vaccines. Genera within this family that are of particular use include, for example, vesiliovirus, Lyssavirus (rabies; finds both medical and veterinary use. Target antigens for this family of viruses include, for example, G protein and N protein. Also of interest are viruses of the Filoviridue famlily, which includes the hemorrhagic fever viruses such as Marburg and Ebola viruses.

Viruses of the Paramyxovirus family, when attenuated using the methods of the invention, also provide vaccines that find both medical and veterinary use. Examples of

genera of interest include, for example, paramyxovirus (both medical and veterinary use), the mumps virus, New Castle disease virus (important pathogen in chickens), morbillivirus: (medical and veterinary use), measles, canine distemper, pneumonia viruses, and respiratory syncytial virus. The Orthomyxovirus family of viruses are also of interest for medical use when attenuated using the methods of the invention. These include, for example, the influenza virus.

The Bungavirus family is also of interest, including the following genera: bungavirus (including California encephalitis virus, LA Crosse virus), Phlebovirus (including Rift Valley Fever virus), Hantavirus (Puumala is a hemorrhagic fever virus), Nairovirus (causes Nairobi sheep disease, and thus vaccines find use in veterinary applications). Many unassigned bungaviruses are known and are also useful as vaccines when attenuated using the methods of the invention.

Also of interest for use as attenuated vaccines are viruses of the arenavirus family, which includes the LCM and Lassa fever virus. The Reovirus family, when attenuated, also provides a vaccine of interest. Genera of particular interest include, for example, reovirus (a possible human pathogen), rotavirus (causes acute gastroenteritis in children), orbiviruses: (which find both medical and veterinary use and include Colorado Tick fever, Lebombo (humans), equine encephalosis, and blue tongue.

The Retrovirus family includes many viral pathogens that cause significant diseases that are recalcitrant to existing treatment methods. Attenuated vaccines derived from these viruses find both veterinary and medical use. Sub-families of the retrovirus family include, for example, the oncorivirinal retroviruses (*e.g.*, feline leukemia virus, HTLVI and HTLVII), the lentivirinal retroviruses (*e.g.*, HIV, feline immunodeficiency virus, equine infections and anemia virus), and the spumavirinal retrovirus family.

Also of interest for use as vaccines are attenuated viruses of the papovavirus family. This family includes the sub-families: polyomaviruses (including BKU and JCU viruses), papillomavirus (which includes many viral types associated with cancers or malignant progression of papilloma), adenovirus (useful for medical applications, including AD7, ARD., O.B; some adenoviruses cause respiratory disease, while others (*e.g.*, 275) can cause enteritis).

Attenuated viruses of the parvovirus family find use for veterinary applications in particular. For example, attenuated virus vaccines can be obtained using the methods of the invention for feline parvovirus (causes feline enteritis), feline panleucopeniavirus, canine parvovirus and porcine parvovirus.

Viruses of the herpesvirus family can also be subjected to the attenuation methods of the invention to obtain attenuated vaccines. Herpesvirus sub-families of particular interest include the alphaherpesviridue subfamily (including the simplexvirus genus (*e.g.*, HSVI, HSVII, both of which are suitable for medical use; Varicellovirus (useful for both medical and veterinary use), and pseudorabies (varicella zoster)), the betaherpesviridue (which includes the cytomegalovirus (*e.g.*, HCMV) and muromegalovirus genera), and the gammaherpesvirdiue sub-family (including the genera lymphocryptovirus, EBV (Burkitts lymphoma), and rhadinovirus.

The poxvirus family is also of particular interest for the development of attenuated vaccines. The poxvirus family includes the Chordopoxviridue subfamily (includes viruses that, when attenuated using the methods of the invention, are useful for both medical and veterinary applications; genera include variola (smallpox), vaccinia (cowpox), parapoxvirus (veterinary), auipoxvirus (veterinary), capripoxvirus, leporipoxvirus and suipoxvirus) and the entemopoxviridue subfamily.

Another viral family of interest is the hepadnavirus family, which includes, for example, hepatitis B virus. Unclassified viruses of interest for development of attenuated vaccines include, for example, hepatitis delta virus. A list of viruses of interest is presented in Table 1.

## Table 1: Viral Targets for Development of Attenuated Vaccines

Deltavirus
- Hepatitis D virus
- Hepatitis D virus (ISOLATE 7/18/83)
- Hepatitis D virus (ISOLATE AMERICAN)
- Hepatitis D virus (isolate D380)
- Hepatitis D virus (ISOLATE ITALIAN)
- Hepatitis D virus (ISOLATE JAPANESE M-1)
- Hepatitis D virus (ISOLATE JAPANESE M-2)
- Hepatitis D virus (ISOLATE JAPANESE S-1)
- Hepatitis D virus (ISOLATE JAPANESE S-2)
- Hepatitis D virus (isolate Lebanon-1)
- Hepatitis D virus (ISOLATE NAURU)
- Hepatitis D virus (isolate woodchuck)

dsDNA viruses, no RNA stage
- Adenoviridae
  - Aviadenovirus
  - Mastadenovirus
  - unclassified Adenoviridae
- African swine fever-like viruses
  - African swine fever virus
- Ascoviridae
  - ascovirus
- Baculoviridae
  - Granulovirus
  - Nucleopolyhedrovirus
  - unclassified Baculoviridae
- Corticoviridae
  - Corticovirus
- Fuselloviridae
  - Fusellovirus
- Herpesviridae
  - Alphaherpesvirinae
  - Betaherpesvirinae
  - Gammaherpes-virinae(lymphoproliferative virus group)
  - Unassigned Herpesviridae
  - unclassified Herpesviridae
- Iridoviridae
  - Iridovirus(small iridescent insect viruses)
  - Lymphocystivirus
  - Ranavirus
  - Red sea bream iridovirus
- Lipothrixviridae
  - Lipothrixvirus
- Papovaviridae
  - Papillomavirus
  - Polyomavirus
- Phycodnaviridae
  - Phycodnavirus
  - Feldmannia sp. virus
- Plasmaviridae
  - Acholeplasma phage L2
  - Plasmavirus
- Polydnaviridae
  - Bracovirus
  - Ichnovirus
- Poxviridae
  - Chordopoxvirinae
  - Entomopoxvirinae
  - unclassified Poxviridae
- Tailed phages
  - Myoviridae(phages with contractile tails)
  - Podoviridae(phages with short tails)
  - Siphoviridae(phages with long non-contractile tails)
- Tectiviridae
  - Tectivirus
- unclassified dsDNA phages
  - Bacteriophage 933W
  - Lactobacillus casei bacteriophage A2
  - Mycoplasma arthritidi bacteriophage MAV1
  - virus of Serpulina hyodysenteriae 1

dsRNA viruses

Table 1(cont'd)

Birnaviridae
- Aquabirnavirus
- Avibirnavirus
- Birnavirus
- Entomobirnavirus
- yellowtail ascites virus

Cystoviridae
- Cystovirus

Hypoviridae
- Hypovirus

Partitiviridae
- Alphacryptovirus
- Partitivirus
- unclassified Partitiviridae

Reoviridae
- Aquareovirus
- Coltivirus
- Cypovirus(cytoplasmic polyhedrosis viruses)
- Fijivirus(plant reovirus 2)
- Orbivirus
- Orthoreovirus
- Oryzavirus
- Phytoreovirus(plant reovirus 1)
- Rotavirus
- unclassified Reoviridae

Totiviridae(monopartite dsRNA genome mycoviruses)
- Giardiavirus
- Leishmaniavirus
- Totivirus
- unclassified Totiviridae

unclassified dsRNA viruses
- Eimeria nieschulzi virus
- Non-encapsidated dsRNA virus

Retroid viruses

Badnavirus
- banana streak virus
- Cacao swollen shoot virus
- Commelina yellow mottle virus
- Rice tungro bacilliform virus
- Sugarcane bacilliform virus

Caulimovirus
- Carnation etched ring virus
- Cauliflower mosaic virus
- Figwort mosaic virus
- Peanut chlorotic streak virus
- petunia vein clearing virus
- Soybean chlorotic mottle virus
- Strawberry vein banding virus

Hepadnaviridae(hepatitis B-type viruses)
- Avihepadnavirus(avian hepatitis B-type viruses)
- Orthohepadnavirus (mammalian hepatitis B-type viruses)

Retroviridae
- Avian type C retroviruses
- BLV-HTLV retroviruses
- Intracisternal A-particles
- Lentivirus
- Mammalian type B retroviruses
- Mammalian type C retroviruses
- Spumavirus
- Type D retroviruses
- unclassified Retroviridae

unclassified retroid viruses
- Cassava vein mosaic virus
- Endogenous sheep retrovirus
- Exogenous Jaagsiekte sheep retrovirus
- porcine endogenous type C retrovirus

Satellites
- RYMV small circular viroid-like RNA(Smal circular viroid-like RNA associated with rice yellow mottle virus)
- Satellite cucumber mosaic virus
- Satellite maize white line mosaic virus
- Satellite panicum mosaic virus

**Table 1**(cont'd)

Satellite phage P4
Satellite St. Augustine decline virus
Satellite tobacco mosaic virus
Satellite tobacco necrosis virus
Satellite tobacco necrosis virus 1
Satellite tobacco necrosis virus 2
Satellite tobacco ringspot virus
Satellite virus of maize white line mosaic virus
ssRNA Satellites
Arabis mosaic virus satellite
Tomato leaf curl virus satellite
ssDNA viruses
Circoviridae
Circovirus
unclassified Circoviridae
Geminiviridae
Begomovirus
Curtovirus
Mastrevirus
Unclassified Geminiviridae
Inoviridae
Inovirus
Plectrovirus
Microviridae(isometric ssDNA phages)
Chlamydiamicrovirus
Microvirus
Spiromicrovirus
Parvoviridae
Densovirinae
Parvovirinae
unclassified ssDNA viruses
TT virus
Xanthomonas campestris phage phi-Lf
ssRNA negative-strand viruses
Arenaviridae
Arenavirus
Bunyaviridae
Bunyavirus
Hantavirus
Nairovirus
Phlebovirus
Tospovirus
unclassified Bunyaviridae
Mononegavirales(negative-sense genome single-stranded RNA viruses)
Borna disease virus
Filoviridae
Paramyxoviridae
Rhabdoviridae
Orthomyxoviridae
Influenza virus A and B group
Influenza virus C group
Thogoto-like viruses
unclassified Orthomyxoviridae
Tenuivirus
Echinochloa hoja blanca virus
maize stripe virus
rice grassy stunt virus
rice hoja blanca virus
rice stripe virus
Urochloa hoja blanca virus
ssRNA positive-strand viruses, no DNA stage
Astroviridae
Astrovirus
Barnaviridae
Barnavirus
beet soil-borne mosaic virus(BSBMV)
Beet virus Q
Bromoviridae
Alfamovirus
Bromovirus
Cucumovirus
Ilarvirus
Caliciviridae
Calicivirus
unclassified Caliciviridae
Capillovirus
Apple stem grooving virus
Cherry capillovirus
Citrus tatter leaf virus
Carlavirus
apple stem pitting virus
blueberry scorch virus
Carnation latent virus
Chrysanthemum virus B

**Table 1(cont'd)**

cowpea mild mottle virus

garlic common latent virus

garlic latent virus

garlic virus 1

Helenium virus S

Lily symptomless virus(LSV)

Poplar mosaic virus

Potato virus M

Potato virus S

Potato virus V

Rupestris stem pitting associated virus-1

Shallot latent virus

Shallot virus X

sour cherry green ring mottle virus

sugarcane striate mosaic virus

Closteroviridae

Closterovirus

Crinivirus

Comoviridae

Comovirus

Fabavirus

navel orange infectious mottling virus

Nepovirus

Dianthovirus

Carnation ringspot virus

Red clover necrotic mosaic virus

Sweet clover necrotic mosaic virus

Enamovirus

Pea enation mosaic virus

Flaviviridae

Flavivirus(arboviruses group B)

Hepatitis C-like viruses

Pestivirus

unclassified Flaviviridae

Furovirus

Beet necrotic yellow vein virus

beet soil-borne virus

broad bean necrosis virus

Indian peanut clump virus

Nicotiana velutina mosaic virus

Oat golden stripe virus

Peanut clump virus

Potato mop-top virus

soil-borne wheat mosaic virus

Hordeivirus

Barley stripe mosaic virus

Lychnis ringspot virus

Poa semilatent virus

Idaeovirus

Raspberry bushy dwarf virus

Leviviridae

Allolevivirus

Levivirus

serogroup III

serumgroup IV

unclassified Leviviridae

Luteovirus

Barley yellow dwarf virus

Bean leaf roll virus

Beet western yellows virus

chickpea stunt disease associated virus

Groundnut rosette assistor virus

Potato leaf roll virus

Soybean dwarf virus

Machlomovirus

Maize chlorotic mottle virus

Marafivirus

maize rayado fino virus

oat blue dwarf virus

Necrovirus

leek white stripe virus

Tobacco necrosis virus

Nidovirales

Arteriviridae

Coronaviridae

Nodaviridae

Isometric bipartite viruses

Nodavirus

Picornaviridae

Aphthovirus

Cardiovirus

Enterovirus

equine rhinovirus 1

Hepatovirus

Rhinovirus(common

Table 1(cont'd)

| | | | |
|---|---|---|---|
| cold viruses) | virus | Solanum nodiflorum mottle virus | Pepper mild mottle virus |
| unclassified Picornaviridae | White clover mosaic virus | Southern bean mosaic virus | Ribgrass mosaic virus |
| Potexvirus | Potyviridae | Bymovirus | Subterranean clover mottle virus |
| Alternanthera potexvirus | Maclura mosaic virus | Velvet tobacco mottle virus | Tobacco mosaic virus |
| Bamboo mosaic virus | Narcissus latent virus | Tetraviridae | Tobamovirus Ob |
| Cassava common mosaic virus | Potyvirus | Betatetravirus | Tomato mosaic virus |
| Clover yellow mosaic virus | Rymovirus | Omegatetravirus | Turnip vein-clearing virus |
| Cymbidium mosaic virus | Wheat streak mosaic virus | unclassified Tetraviridae | Tobravirus |
| Foxtail mosaic virus | wheat yellow mosaic virus | Tobamovirus | Pea early browning virus |
| Lily virus X | Sequiviridae | Beet necrotic yellow vein mosaic virus | Pepper ringspot virus |
| Narcissus mosaic virus | Sequivirus | Chinese Rape Mosaic Virus | Tobacco rattle virus |
| Papaya mosaic virus | Waikavirus | crucifer tobamovirus | Togaviridae |
| Plantago asiatica mosaic potexvirus | Sobemovirus | Cucumber green mottle mosaic virus | Alphavirus(arbovir uses group A) |
| Potato aucuba mosaic virus | cocksfoot mottle virus | Odontoglossum ringspot virus | Rubivirus |
| Potato virus X | Lucerne transient streak virus | Paprika mild mottle virus | Tombusviridae |
| Strawberry mild yellow edge-associated | Olive latent virus 1 | | Carmovirus |
| | Rice yellow mottle virus | | Panicum mosaic virus |
| | | | Tombusvirus |
| | | | Trichovirus |
| | | | Apple chlorotic leaf spot virus |
| | | | Grapevine virus A |
| | | | Grapevine virus B |
| | | | Grapevine virus D |
| | | | Heracleum latent virus |
| | | | potato trichovirus T |
| | | | Tymovirus |
| | | | Andean potato latent virus |
| | | | Belladonna mottle virus |
| | | | Cacao yellow mosaic virus |
| | | | Clitoria yellow vein virus |
| | | | Desmodium yellow mottle tymovirus |
| | | | Dulcamara mottle |

Table 1(cont'd)

tymovirus

Eggplant mosaic virus

Erysimum latent virus

Kennedya yellow mosaic virus

okra mosaic tymovirus

Ononis yellow mosaic virus

Physalis mottle tymovirus

Turnip yellow mosaic virus

wild cucumber mosaic virus

Umbravirus

carrot mottle mimic virus

groundnut rosette virus

unclassified ssRNA positive-strand viruses, no DNA stage

Acyrthosiphon pisum virus

apricot latent virus

garlic virus B

Pear vein yellows-associated virus

Simian hemorrhagic fever virus

SIRS virus ATCC VR-2332

sweet potato chlorotic stunt virus

Swine hepatitis E virus

unclassified bacteriophages

Actinophage JHJ-1

Actinophage JHJ-3

bacteriophage #D

Bacteriophage 12826

Bacteriophage 13

Bacteriophage 16

Bacteriophage 3/14

Bacteriophage 5

bacteriophage 80 alpha

bacteriophage 85

Bacteriophage AR1

Bacteriophage Bastille

Bacteriophage BFK20

Bacteriophage biL66

Bacteriophage c-st

Bacteriophage chp1

Bacteriophage Cp-5

Bacteriophage Cz

Bacteriophage d-16 phi

Bacteriophage DD VI

bacteriophage Dp-1

Bacteriophage FAAU2

bacteriophage Felix 01

bacteriophage fuse 5

Bacteriophage h30

Bacteriophage K139

bacteriophage K1E

Bacteriophage KVP20

Bacteriophage KVP40

Bacteriophage L

Bacteriophage L10

Bacteriophage L17

Bacteriophage L5

Bacteriophage l54a

Bacteriophage LL-H

Bacteriophage LM4

Bacteriophage LZ1

Bacteriophage LZ4

Bacteriophage LZ6

Bacteriophage LZ7

Bacteriophage LZ8

Bacteriophage M13mp18

Bacteriophage M13mp7

Bacteriophage M2Y

Bacteriophage MB78

Bacteriophage mv4

Bacteriophage Mx8

Bacteriophage N 209

Bacteriophage N15

Bacteriophage NJL

Bacteriophage P001

Bacteriophage P32

Bacteriophage P37

Bacteriophage P54

Bacteriophage PAt5-12

Bacteriophage phi 7-9

bacteriophage phi AR29

Bacteriophage phi LC3

bacteriophage phi PVL

Bacteriophage phi-LC3

Bacteriophage phi-PLS27

Bacteriophage phi-vml3

bacteriophage phiFSW

Bacteriophage phig1e

bacteriophage phiU

Bacteriophage PR4

Bacteriophage PR5

Bacteriophage PR722

Bacteriophage PsiM1

Bacteriophage PSP3

Bacteriophage PZt11-15

Bacteriophage r1t

Bacteriophage Rho11s

Bacteriophage rho15

**Table 1**(cont'd)

(ISOLATE 7-9)

Bacteriophage SC1

bacteriophage SfII

Bacteriophage T1

Bacteriophage T270

Bacteriophage TP21

Bacteriophage TP901-1

Bacteriophage US3

Bacteriophage VWB

Bacteriophage Z

Bartonella henselae phage 60457

Clostridium botulinum C phage

Clostridium botulinum D phage

Clostridium botulinum phage 1C

Clostridium botulinum phage CN

Clostridium botulinum phage d-sA

Coliphage 933J

Coliphage K1

Coliphage K1F

coliphage K5

Colitis phage

Corynebacterium diphtheriae virus

corynephage phi16

filamentous bacteriophage

Halophage HF2

Klebsiella phage No. 11

Lactobacillus bacteriophage phi adh

Lactobacillus casei bacteriophage J1

Lactobacillus delbrueckii bacteriophage mv1

Lactococcal bacteriophage P6

Lactococcus 1 bacteriophage

Lactococcus bacteriophage

Lactococcus bacteriophage

*Lactococcus lactis* bacteriophage F4-1

Lactococcus lactis phage phi 197

Lactococcus lactis phage US3

Lactococcus phage mi7-9

Leuconostoc oenos bacteriophage 10MC

Methanobacterium phage psiM1

Mycobacteriophage D29

Mycobacteriophage I3

Mycobacteriophage I5

Mycobacteriophage TM4

Mycobacterium phage DS6A

Mycobacterium phage FRAT1

Mycobacterium phage L5

Mycobacterium phage Ms6

*Oenococcus oeni* temperate bacteriophage fOg44

phage E

phage SV40

Pseudomonas aeruginosa phage D3

Pseudomonas aeruginosa phage phi CTX

Pseudomonas aeruginosa phage PS17

Staphylococcus aureus bacteriophage phi-42

Staphylococcus bacteriophage phi 11

Streptococcus phage phi7201

Streptococcus pyogenes phage

Streptococcus pyogenes phage H10403

Streptococcus pyogenes phage H4489A

Streptococcus pyogenes phage T12

Streptococcus thermophilus bacteriophage DT1

Vibrio cholerae bacteriophage K139

Vibrio cholerae O139 fs1 phage

Vibrio cholerae V86 phage

virus-like particle CAK1

unidentified bacteriophage

unclassified viruses

20S RNA replicon

acute paralysis virus

Agaricus bisporus virus 1

Archaeal virus SIRV

carrot red leaf luteovirus associated RNA

Citrus psorosis virus

cloudy wing virus

East African cassava

**Table 1**(cont'd)

mosaic virus

ecotropic provirus

Ectocarpus siliculosus virus

Ectocarpus virus 63124

Ectocarpus virus 63125

Ectocarpus virus 63126

Ectocarpus virus 63127

flame chlorosis virus-like agent

Frog erythrocytic virus

garlic mite-borne filamentous virus

garlic virus A

garlic virus C

garlic virus D

grapevine Rupestris stem pitting associated virus

Grapevine trichovirus B

Halobacterium halobium phage phi-H

Halobacterium salinarium phage phi H

Heliothis zea virus 1

High Plains Virus

Japanese garlic virus

JDV virus

K28 killer virus

Kashmir bee virus

Killer virus of S.cerevisiae

La France disease virus

M28 virus

non-A, non-B hepatitis virus

Olive latent virus 2

peanut chlorotic fan-spot virus

pineapple bacilliform virus

Saccharomyces cerevisiae killer particle M2

Saccharomyces cerevisiae killer virus M1

Sesbania mosaic virus

Spiroplasma virus

stealth virus 1

stealth virus 4

## 2. Attenuated bacterial, fungal and parasite vaccines

The invention also provides attenuated vaccines against bacterial, fungal and other pathogens, such as parasites. Methods for obtaining these attenuated vaccines are also provided. The methods of the invention provides a means by which one can generate novel variant bacteria or parasites that have genotypes and phenotypes that do not naturally occur or would not otherwise be anticipated to occur at a substantial frequency. A preferred aspect of the method employs recursive nucleotide sequence recombination, termed "sequence shuffling" or "DNA shuffling," which enables the rapid generation of a collection of broadly diverse bacterial or parasite phenotypes that can be selectively bred for a broader range of novel phenotypes or more extreme phenotypes than would otherwise occur by natural evolution in the same time period.

Similarly to the case for viral vaccines described above, the presently preferred methods involve (1) sequence shuffling of a plurality of whole or partial bacterial or parasite genomes, and (2) selection of the resultant shuffled bacterial or parasite genomes to isolate or enrich a plurality of shuffled genomes that result in an organism that has a desired phenotype(s) (*e.g.*, attenuation). In preferred embodiments, the method includes (3) repeating steps (1) and (2) on the plurality of shuffled genomes that confer the desired phenotype(s) until one or more variant genomes that confer a sufficiently optimized desired phenotype is obtained. In this general manner, the method facilitates the "forced evolution" of a bacterial or other pathogen genome to encode an attenuated organism which natural selection and evolution has heretofore not generated. The recursion option is generally selected each cycle until one or more genomes that confer a satisfactory optimization for the desired phenotype(s) are obtained.

Typically, a plurality of bacterial, parasite, or other genomes of the same taxonomic classification are shuffled and selected by the present method. A common use of the method will be to shuffle mutant variants of a clinical isolate(s) or of a laboratory strain of an organism to obtain a variant of the clinical isolate or laboratory strain that possesses a novel desired phenotype (*e.g.*, attenuation). However, the method can be used with a plurality of strains (or clades) of a pathogenic organism, or even with a plurality of related organisms (*e.g.*, *Mycobacterium tuberculosis*, *Mycobacterium vaccae* and *Mycobacterium bovis* (BCG)), and in some instances with unrelated pathogens or portions thereof which

have recombinogenic portions (either naturally or generated via genetic engineering). The method can be used to shuffle xenogeneic sequences into a pathogen's genome (*e.g.*, incorporating and evolving a gene of a first pathogenic organism in the genome of a second organism so as to confer a desired phenotype (such as immunogenicity) to the evolved genome of the second organism). Furthermore, the method can be used to evolve a heterologous nucleic acid (*e.g.*, a non-naturally occurring mutant gene) to optimize its phenotypic expression (*e.g.*, immunogenicity) when present in a bacterial or parasite genome, and/or in a particular host cell or expression system (*e.g.*, an expression cassette or expression replicon).

In some embodiments, the methods of the invention are used to create chimeras of pathogenic and non-pathogenic bacteria, fungi or parasites. In these applications, whole genome or partial genome shuffling is preferred for generating the libraries of recombinant nucleic acids. For example, a specific and broad-spectrum bacterial vaccine against nosocomical infections can be obtained by whole genome shuffling of pathogenic bacteria with, for example, *Lactococcus lactis*. Protocols for whole genome shuffling are described in, for example, PCT patent application No. US98/00852 (Publ. No. WO 98/31837).

The methods of the invention, including family shuffling of single genes and whole genomes, can also be used to generate attenuated strains of bacteria that are useful as vaccines or vaccine antigen delivery vehicles. In addition, these methods can be used to improve expression levels of vaccine antigens in bacterial strains used as vaccines. For example, *Mycobacterium bovis bacillus* Calmette-Guerin (BCG) has been widely used as human tuberculosis vaccine, and it has several features that make it a particularly attractive live recombinant vaccine vehicle. BCG, like other mycobacteria, are potent adjuvants, and the immune response to mycobacteria has been studied extensively (Orme, *Int. J. Tuberc. Lung Dis.* (1997) 1: 95-100). More than two billion immunizations with BCG have been performed with a long record of safe use in man. It is one of the few vaccines that can be given at birth, and it provides long- lived immune responses after a single dose. Foreign genes have been successfully introduced into BCG enabling the generation of BCG-based vaccines against non-mycobacterial diseases, including HIV (Aldovini and Young, *Nature* (1991) 351: 479-82). Another useful bacterial strain for evolution by DNA shuffling is

*Mycobacterium vaccae* (*M. vaccae*), which has previously been implicated in the treatment of psoriasis (Lehrer *et al.*, *FEMS Immunol. Med. Microbiol.* (1998) 21: 71-7).

The methods of the invention enable the generation of chimeric bacteria or other pathogenic organisms that have antigenic determinants from other bacteria or pathogens. For example, whole genome shuffling can be used to generate BCG-like strains that have multiple antigenic determinants derived from *Mycobacterium tuberculosis* (Mt). More specifically, BCG and Mt can be crossed by whole genome shuffling and the optimal vaccine strains selected by Mt-specific antibodies. Alternatively, chimeras of *M. vaccae* and Mt can be generated using similar type of approach. The attenuated phenotype of the new shuffled strains can be confirmed in animals models, which will simultaneously allow the analysis of the immunogenicity of such strains. In fact, the optimal vaccine strains can be selected by using *in vivo* immunizations. The strains that induce potent Mt-specific antibody responses *in vivo*, while retaining their attenuated phenotype, can also be selected for new rounds of whole genome shuffling and selection. Challenge of the immunized animals with live Mt will enable the analysis of the quality of the protective immune response.

Further examples of useful targets include, but are not limited to, whole genome shuffling of *Bacillus subtilis* and *Bacillus anthracis* to generate nonpathogenic bacillus strains that have antigenic determinants from *Bacillus anthracis*, which can provide protective immune responses against anthrax. Moreover, shuffling attenuated and pathogenic strains of *Salmonella* species can be used to generate strains that have attenuated phenotype, while expressing immunogenic determinants from pathogenic *Salmonellae*, providing protective immune responses against *Salmonella* infection.

DNA shuffling can also be used to improve expression levels of antigens to be expressed in the attenuated bacteria. Because pathogens infecting mammalian cells have generally not coevolved with bacteria, expression of viral antigens in bacteria is problematic often resulting in poor expression levels. Expression, solubility and folding of pathogens antigens, viral antigens in particular, are also often impaired in BCG, reducing the efficacy of immunizations. DNA shuffling can be used to improve solubility of proteins in bacteria. For example, one can generate libraries of pathogen antigens and select the most efficiently expressed variants in bacteria. For example, HIV antigen gp120 can be fused to GFP, and the fusion proteins expressed in BCG or *M. vaccae*. Expression of GFP is an indication of

gp120 expression, and the brightest cells can be selected for example by flow cytometry based cell sorting.

These approaches using *Mycobacteria* and DNA shuffling also provide opportunities to improve orally or intranasally delivered vaccines. BCG has also been shown to provide protective immune responses via aerogenic vaccination (Lagranderie *et al.*, *Tubercle and Lung Disease* (1993) 74: 38-46). Improved expression of foreign antigens in BCG by DNA shuffling can substantially improve the efficacy of BCG as an oral or inhaled vaccine delivery strain. Further improvements can be obtained by fusing the antigen of interest to adjuvant enterotoxins, such as cholera toxin (CT) or heat-labile enterotoxin of *E. coli* (LT), which can then be secreted from the cells. In the most desired approach, a library of enterotoxins are generated by DNA shuffling (for example by shuffling CT and LT), and these shuffled enterotoxins are fused to shuffled vaccine antigens of interest. These libraries can be screened as expressed purified proteins or they can be expressed in BCG or *M. vaccae*, and these strains will subsequently be screened in animals for immunogenicity. Because several enterotoxins, such as CT and LT, have been shown to act as adjuvants, particularly in the skin and mucosal membranes, this approach is expected to further improve the efficacy of oral, intranasal, transdermal and inhaled vaccines.

Screening for ability of the attenuated vaccines to induce an immune response can be performed using methods known to those of skill in the art. In a presently preferred embodiment, an *in vitro* screen is employed to test for attenuation. The *in vitro* screen will be followed up by further testing *in vivo*. Moreover, the ability of the modified cells to induce protective immunity upon inoculation into a mammal will be studied.

The methods of the invention are useful for producing attenuated vaccines against a wide range of bacterial and other pathogenic cells. For example, one can obtain vaccines against pathogenic gram-positive cocci, including pneumococcal, staphylococcal and streptococcal bacteria. Pathogenic gram-negative cocci are also suitable targets. Of particular interest are the meningococcal and gonococcal bacteria.

Also of interest are vaccines against the pathogenic enteric gram-negative bacilli. Examples include, but are not limited to, enterobacteriaceae (pseudomonas, acinetobacteria and eikenella), melioidosis, salmonella, shigellosis, hemophilus, chancroid, brucellosis, tularemia, yersinia (pasteurella), streptobacillus moniliformis and spirillum,

listeria monocytogenes, erysipelothrix rhusiopathiae, diphtheria, cholera, anthrax, donovanosis (granuloma inguinale), and bartonellosis.

The pathogenic anaerobic bacteria are also suitable targets for the development of attenuated vaccines. Those of particular interest include, for example, tetanus, botulism, other clostridia, tuberculosis, leprosy, and other mycobacteria. Pathogenic spirochetal diseases include syphilis, treponematoses (yaws, pinta and endemic syphilis), and leptospirosis.

Vaccine targets also include other infections caused by higher pathogen bacteria and pathogenic fungi, including, for example, actinomycosis, nocardiosis, cryptococcosis, blastomycosis, histoplasmosis and coccidioidomycosis; candidiasis, aspergillosis, and mucormycosis; sporotrichosis, paracoccidiodomycosis, petriellidiosis, torulopsosis, mycetoma and chromomycosis; and dermatophytosis. The methods of the invention are also useful for obtaining attenuated vaccines against rickettsial infections (*e.g.*, rickettsial and rickettsioses) and mycoplasma and chlamydial infections (*e.g.*, mycoplasma pneumoniae, lymphogranuloma venereum, psittacosis and perinatal chlamydial infections).

Other targets include parasites, including but not limited to, amebiasis, malaria, leishmaniasis, trypanosomiasis, toxoplasmosis, pneumocystis carinii, babesiosis, giardiasis, trichinosis, filariasis, schistosomiasis, nematodes, trematodes or flukes, and cestode (tapeworm) infections.

### *II. Methods to Screen for Attenuated Organisms*

A recombination cycle is usually followed by at least one cycle of screening or selection for recombinant nucleic acid molecules having a desired property or characteristic that is of interest for vaccination. The nature of screening or selection depends on what property or characteristic is to be acquired or the property or characteristic for which improvement is sought, and many examples are discussed below. It is not usually necessary to understand the molecular basis by which particular products of recombination (recombinant segments) have acquired new or improved properties or characteristics relative to the starting substrates. For example, an attenuated vaccine of the invention can have many component sequences each having a different intended role (*e.g.*, coding sequence, regulatory sequences, targeting sequences, stability-conferring sequences,

immunomodulatory sequences and sequences affecting antigen presentation). Each of these component sequences can be varied and recombined simultaneously. Screening/selection can then be performed to identify recombinant segments that have, for example, increased attenuation and/or immunogenicity without the need to attribute such improvement to any of the individual component sequences of the vector.

If a recombination cycle is performed *in vitro*, the products of recombination, *i.e.*, recombinant segments, are sometimes introduced into cells before the screening step. Recombinant segments can also be linked to an appropriate vector or other regulatory sequences before screening. The products of *in vitro* recombination are sometimes packaged as viruses before screening, or as part of the screening. For example, an attenuated viral vaccine can be identified by the inability of the recombinant nucleic acids to direct synthesis of viruses upon introduction into a non-permissive host cell (*e.g.*, a cell from the species that is to be vaccinated). If recombination is performed *in vivo*, recombination products can sometimes be screened in the cells in which recombination occurred. In other applications, recombinant segments are extracted from the cells, and optionally packaged as viruses, before screening.

The introduction of the recombinant nucleic acids into cells for screening can introduce multiple copies of the recombinant nucleic acids. For some applications, it is desirable to insert only a single copy of the modified gene into each cell. Another preferred variation of this assay involves reducing the amount of variability in transcription of a recombinant nucleic acid that can result from differences in chromosomal location of integration sites. This requires a means for defined, site-specific integration of the recombinant nucleic acids. These methods can also be used to evolve an episomal vector (which can replicate inside the cell) that can site-specifically integrate into a chromosome. One way to obtain single copy integrations of recombinant nucleic acids is to use retroviruses as a shuttle vector. Retroviruses integrate as a single copy. However, this insertion is not site-specific, *i.e.*, the retrovirus inserts in a random location in the chromosome. Adenoviruses and ars-plasmids are also used to shuttle modified transgenes, however, they integrate as multiple copies. While wild type AAV integrates as a single copy in chromosome q19, commonly used modified versions of AAV do not. Homologous recombination is also used to insert a modified recombinant segment into a chromosome, but

this method can be inefficient and may result in the integration of two copies in the pair of chromosomes.

To solve these problems, some embodiments of the invention utilize site-specific integration systems to target the transgene to a specific, constant location in the genome. A preferred embodiment uses the Cre/LoxP or the related FLP/FRT site-specific integration system. The Cre/LoxP system uses a Cre recombinase enzyme to mediate site-specific insertion and excision of viral or phage vectors into a specific palindromic 34 base pair sequence called a "*Lox*P site." LoxP sites can be inserted to a mammalian genome of choice, to create, for example, a transgenic animal containing the LoxP site, by homologous recombination (*see* Rohlmann (1996) *Nature Biotech.* 14:1562-1565). If a cell's genome is engineered to contain a *Lox*P site in a desired location, infection of such cells with recombinant nucleic acids that are flanked by *LoxP* sites, in the presence of Cre recombinase (*e.g.*, expressed by a vector that expresses a gene for the Cre recombinase) results in the efficient, site-specific integration of the recombinant nucleic acids into the *Lox*P site. This approach is reproducible from cycle to cycle and provides a single copy of the recombinant sequence at a constant, defined location. Thus, a recombinant nucleotide obtained using the methods of the invention *in vitro* can be reinserted into the cell for *in vivo/in situ* selection for the new or improved property in the optimal way with minimal noise. This technique can also be used *in vivo*. *See*, for example, Agah (1997) *J. Clin. Invest.* 100:169-179; Akagi (1997) *Nucleic Acids Res.* 25:1766-1773; Xiao (1997) *Nucleic Acids Res* 25:2985-2991; Jiang (1997) *Curr Biol* 7:321-R323, Rohlmann (1996) *Nature Biotech.* 14:1562-1565; Siegal (1996) *Genetics* 144: 715-726; Wild (1996) *Gene* 179:181-188. The evolution of *Cre* is discussed in further detail in PCT patent application US97/17300 (Publ. No. WO98/13487), filed September 26, 1997.

In presently preferred embodiments, the attenuated vaccines of the invention are screened in mammalian cells or organisms. Once a group of attenuated strains have been identified, these vaccine strains are subsequently analyzed for their immunogenicity *in vivo*. Useful animal species for such studies include, but are not limited to, mice, rats, guinea pigs, cats, dogs, cows, pigs, horses, chicken. These experiments are useful in identifying improved veterinary vaccines, and they also provide information about their safety and efficacy for use as human vaccines. The attenuated vaccines that are intended for use in humans are often

subjected to further testing in humans. In some instances, cells used for screening can be obtained from a patient to be treated with a view, for example, to minimizing problems of immunogenicity in this patient. Use of an attenuated vaccine in treatment can itself be used as a round of screening. That is, attenuated vaccines that are successively taken up and/or expressed by the intended target cells in one patient are recovered from those target cells and used to treat another patient. The attenuated vaccines that are recovered from the intended target cells in one patient are enriched for vectors that have evolved, *i.e.*, have been modified by recursive recombination, toward improved or new properties or characteristics for attenuation, specific uptake, immunogenicity, stability, and the like.

The screening or selection step identifies a subpopulation of recombinant segments (*e.g.*, viral or bacterial whole or partial genomes, or other nucleic acid segments) that have evolved toward acquisition of improved attenuation, and/or other new or improved desired properties useful in vaccination. Depending on the screen, the recombinant segments can be screened as components of cells, components of viruses or other vectors, or in free form. More than one round of screening or selection can be performed after each round of recombination.

If further improvement in a property is desired, at least one and usually a collection of recombinant segments surviving a first round of screening/selection are subject to a further round of recombination. These recombinant segments can be recombined with each other or with exogenous segments representing the original substrates or further variants thereof. Again, recombination can proceed *in vitro* or *in vivo*. If the previous screening step identifies desired recombinant segments as components of cells, the components can be subjected to further recombination *in vivo*, or can be subjected to further recombination *in vitro*, or can be isolated before performing a round of *in vitro* recombination. Conversely, if the previous screening step identifies desired recombinant segments in naked form or as components of viruses or other vectors, these segments can be introduced into cells to perform a round of *in vivo* recombination. The second round of recombination, irrespective how performed, generates further recombinant segments which encompass additional diversity compared to recombinant segments resulting from previous rounds.

The second round of recombination can be followed by a further round of screening/selection according to the principles discussed above for the first round. The stringency of screening/selection can be increased between rounds. Also, the nature of the screen and the property being screened for can vary between rounds if improvement in more than one property is desired or if acquiring more than one new property is desired. Additional rounds of recombination and screening can then be performed until the recombinant segments have sufficiently evolved to acquire the desired new or improved property or function.

After a desired phenotype is acquired to a satisfactory extent by a selected shuffled viral or other pathogen genome or portion thereof, it is often desirable to remove mutations which are not essential or substantially important to retention of the desired phenotype ("superfluous mutations"). Superfluous mutations can be removed by backcrossing, which involves shuffling the selected shuffled genome(s) with one or more parental genome and/or naturally-occurring genome(s) (or portions thereof) and selecting or screening the resultant collection of shufflants to identify those that retain the desired phenotype. By employing this method, typically in one or more recursive cycles of shuffling against parental or naturally-occurring genome(s) (or portions thereof) and selection or screening for retention of the desired phenotype, it is possible to generate and isolate selected shufflants that incorporate substantially only those mutations necessary to confer the desired phenotype (*e.g.*, attenuation), while having the remainder of the genome (or portion thereof) consist of sequence which is substantially identical to the parental (or wild-type) sequence(s). As one example of backcrossing, a viral genome can be shuffled and selected for attenuation in target host cells; the resultant selected shufflants can be backcrossed with one or more genomes of clinical isolates of the virus and selected for retention of the attenuation. After several cycles of such backcrossing, the backcrossing will yield viral genome(s) that contain the mutations necessary for attenuation, and will otherwise have a genomic sequence substantially identical to the genome(s) of the clinical isolate(s) of the virus.

Examples of the types of approaches that are useful for obtaining attenuated vaccines, and screening/selection techniques that are suitable for identifying vaccines having the desired properties, are described in the following section.

### *III. Illustrative Examples of Attenuated Vaccines*

The invention provides attenuated vaccines, and methods for obtaining attenuated vaccines, that have a wide variety of properties. To obtain attenuated vaccines that have these and other desired properties, a suitable screening and/or selection method is used which is specific for the particular properties desired. The screening and/or selection methods can be used in combination to obtain attenuated vaccines that have more than one desired improvement. The following are illustrative examples of types of attenuated viral, and bacterial vaccines, and methods for obtaining such vaccines. Different selection/ screening methods can also be used as is appropriate to identify attenuated vaccines that have other desirable properties. Analogous methods are useful for developing attenuated fungal and parasite vaccines.

#### 1. Non-pathogenic chimeric vaccines

In some embodiments, the invention provides chimeric viruses, bacteria or other organisms into which are introduced nucleic acids that encode one or more immunogenic polypeptides from a pathogenic virus or other organism. In presently preferred embodiments, the chimeric vaccine is non-pathogenic. Both the pathogenic and non-pathogenic virus or organism can be of the same species (*e.g.*, a coding region for an immunogenic polypeptide from a pathogenic strain is introduced into a non-pathogenic strain). These methods are useful, for example, in the development of polyvalent vaccines that express immunogenic polypeptides from multiple strains of an organism or virus. As one example, vaccines derived from human papillomavirus strains are typically not cross-protective against other HPV strains. By using the methods of the invention, one can obtain polyvalent, cross-protective HPV strains. Alternatively, the pathogenic and non-pathogenic virus or organism can be of different species. As one example, a vaccinia virus can function as a non-pathogenic carrier for a nucleic acid that encodes an immunogenic polypeptide such as, for example, gp120 of HIV.

These methods typically involve recombining a first set of one or more nucleic acid segments from a virus or cell with a second set of one or more nucleic acid segments. The nucleic acid segments of the second set typically encode one or more polypeptides, or portions thereof, that confer upon a viruses or cells that include the polypeptide a property that is desirable for vaccination. For example, the second set of

nucleic acid segments can encode an immunogenic polypeptide from a pathogenic strain of a virus or cell, an adjuvant or immunomodulatory molecule, and the like.

The resulting library of recombinant DNA fragments is then screened to identify those that confer upon a virus or cell an improvement in the desired property. In presently preferred embodiments, the viruses or cells that contain the recombinant fragments are screened to identify those viruses or cells that have become, or remain, attenuated (*i.e.*, nonpathogenic) under physiological conditions present in a host organism inoculated with the virus or cell. The screening for attenuation can be conducted before or after, or simultaneously with, the screening for the improvement in the other desired property.

In some embodiments, the recombination is performed *in vivo*. For example, one can introduce a library of DNA fragments that comprises at least a partial genomic library of a pathogenic cell into a plurality of nonpathogenic cells. At least one of the fragments from the pathogenic cell undergoes recombination with a segment in the genome or an episome of the non-pathogenic cells to produce modified cells. The modified cells are screened to identify those that are nonpathogenic but have evolved towards an ability to induce an immune response against the pathogenic cells. The resulting nonpathogenic cells that have evolved towards an ability to induce an immune response against the pathogenic cells are suitable for use as an attenuated vaccine.

If desired, further improvement can obtained by subjecting the DNA from the modified cells that are nonpathogenic and have evolved an ability to induce an immune response against the pathogenic cells to recombination with a further library of DNA fragments from a pathogenic organism. At least one of the fragments from the pathogenic organism undergoes recombination with a segment in the genome or the episome of the modified cells to produce further modified cells. Alternatively, one can recombine DNA from the modified cells that are nonpathogenic and have evolved an ability to induce an immune response with DNA from the pathogenic cells to produce further modified cells. The further modified cells are then screened to identify further modified cells that are nonpathogenic and have evolved a further ability to induce an immune response against the pathogenic cells. The recombination and selection/screening steps can be repeated as required until the further modified cells are nonpathogenic and have acquired the ability to induce an immune response against the pathogenic cells.

The recombination and selection/screening methods of the invention provide a means not only for obtaining attenuated viruses or cells for use as the carrier, but can be used to obtain chimeric viruses or organisms that exhibit improvements in properties such as enhanced expression of the antigen and improved immunogenicity of the antigen. The genes that encode the antigen can be subjected to recombination separately from the non-pathogenic virus or other vaccine organism; alternatively, one can perform the recombination on whole or partial viral, bacterial or parasite genomes. Methods for improving antigen expression and immunogenicity are described in co-pending, commonly assigned US patent application Ser. No. 09/247,890 (filed February 10, 1999). A polynucleotide that encodes a recombinant antigenic polypeptide can be placed under the control of a promoter, *e.g.*, a high activity or tissue-specific promoter. The promoter used to express the antigenic polypeptide can itself be optimized using recombination and selection methods analogous to those described herein (*see, e.g.,* US Ser. No. 09/247,888, filed February 10, 1999).

In some embodiments, the methods of the invention are used to obtain virus-like particles (VLP's) that have desired characteristics. VLPs lack the viral components that are required for virus replication and, therefore, represent a highly attenuated form of a virus. The VLPs can display antigens from multiple viral strains, and thus are useful as a polyvalent vaccine. Viral proteins from several viruses are known to form VLPs, including human papillomavirus, HIV (Kang *et al.*, *Biol. Chem.* 380: 353-64 (1999)), Semliki-Forest virus (Notka *et al.*, *Biol. Chem.* 380: 341-52 (1999)), human polyomavirus (Goldmann *et al.*, *J. Virol.* 73: 4465-9 (1999)), rotavirus (Jiang *et al.*, *Vaccine* 17: 1005-13 (1999)), parvovirus (Casal, *Biotechnology and Applied Biochemistry*, Vol 29, Part 2, pp 141-150 (1999)), canine parvovirus (Hurtado *et al.*, *J. Virol.* 70: 5422-9 (1996)), and hepatitis E virus (Li *et al.*, *J. Virol.* 71: 7207-13 (1997)).

Screening for ability to induce an immune response against the pathogenic cells can be performed using methods known to those of skill in the art. In a presently preferred embodiment, the screening is performed by testing for ability of the modified cells to induce protective immunity upon inoculation into a mammal.

### 2. Replication-deficient viral vaccines that are evolved for high efficiency infectivity and protectivity

In other embodiments, the invention provides replication-deficient viruses that are evolved to infect target cells with high efficiency, but with only one round of replication or no replication once the cells are infected. Replication-deficient viruses can be obtained either by rational design (*e.g.*, targeted disruption of a gene that is involved in viral replication) or by the recombination and screening/selection methods of the invention. The nucleic acids of the replication deficient viruses thus obtained are then subjected to recombination and selection for those viruses that exhibit improved entry into host cells. The screening can be accomplished by, for example, fluorescence-activated cell sorting of cells that contain the virus (based on expression of a gene such as Mx). Viruses can be recovered from these cells, re-infected into host cells, re-sorted. After one or more repetitions of the screening/selection, individual colonies analyzed for their capacity to replicate in cell culture. Those that do not replicate in host cells are suitable for further testing as vaccines, *e.g.*, for ability to induce an "antiviral state" in infected cells. Again, backcrossing can be used to obtain viral vaccines that include the mutations that prevent replication but lack other mutations that reduce immunogenicity.

### 3. Vaccines attenuated by insertions of stop codons in polypeptide coding regions

In another embodiment, the invention provides attenuated vaccines, and methods of obtaining attenuated vaccines, in which numerous stop codons are introduced into the nucleic acids of the vaccine virus or other organism. These attenuated vaccines thus require the presence of a suppressor tRNA in a host cell in order to replicate (Drabkin *et al.*, *Mol. Cell. Biol.* 16, 907-13 (1996); Park and RajBhandary, *Mol. Cell. Biol.* 18, 4418-25 (1998)). Accordingly, a production cell is used that contains the appropriate suppressor tRNA (*e.g.*, amber, ochre, frameshift or other suppressor) that corresponds to the stop codons that are introduced into the genome of the vaccine virus or other organism. Preferably, the stop codons are of a type that is not frequently used in the respective naturally occurring virus or organism.

The stop codon-containing attenuated vaccines are obtained by recombining a nucleic acid segment, or mixture of fragments, from a virus or other vaccine organism with a population of oligonucleotides that include one or more stop codons interspersed within one

or more polynucleotide sequences that code for at least a portion of a polypeptide necessary for replication of the virus or organism. For example, one can use a library of stop codon-containing oligonucleotides, wherein the sequences of the oligos are determined based on known sequence information for the viral proteins. The mixture of fragments can be a full or partial fragmented genome of the organism. A library of recombinant nucleic acid segments is produced by subjecting the oligonucleotides to recombination together with nucleic acids of the virus or other organism. This results in the non-naturally occurring stop codons becoming incorporated into the coding regions of the cell or virus, thus causing premature translational termination in the absence of a corresponding suppressor tRNA.

The library of recombinant nucleic acid segments are then screened by contacting introduced into production cells, for example, that contain suppressor tRNA which suppresses the translational termination that would otherwise occur at the introduced stop codons (*see, e.g.,* (Drabkin *et al.*, *Mol. Cell. Biol.* 16, 907-13 (1996); Park and RajBhandary, *Mol. Cell. Biol.* 18, 4418-25 (1998)). Those viruses or other organisms that reproduce in the production cells, and resulting progeny, are then collected. In presently preferred embodiments, those that replicate are then tested for ability to replicate in non-suppressor cells. Preferably, cells of the mammal that is to be inoculated are tested. Progeny viruses or other organisms that do not replicate in the non-suppressor cells are suitable for use as attenuated vaccine organisms.

**4. Selection of conditionally replicating mutant strains**

In some embodiments, the invention provides attenuated vaccines that are unable to proliferate under physiological conditions in an inoculated host mammal, but are capable of replication under permissive conditions such as those used for production of the vaccine. Permissive conditions can include a property of a production cell or growth conditions, that differs from the corresponding physiological condition in the cells of an inoculated host mammal. For example, a permissive condition can be a temperature, pH, sugar content, the presence or absence of complement components and/or serum proteins, and the like, that differs from the physiological condition found in the inoculated mammal. A permissive condition can also be the presence of an essential nutrient that is absent in the inoculated host mammal.

Methods of obtaining such vaccines are also provided by the invention. For bacterial, parasite, and other whole-cell vaccines, these methods typically involve introducing a library of recombinant DNA fragments into population of the bacterial cells. The recombinant fragments can be a whole or partial bacterial genome, or a recombinant gene or genes, that can become incorporated into the genome or an episome of the cells. The modified cells are then screened to identify conditionally defective cells that have evolved toward loss of the ability to proliferate under physiological conditions as found in a host organism. The conditionally defective cells are then screened to identify those modified cells that have evolved toward ability to replicate under the permissive conditions. Modified cells that replicate under permissive conditions but not in a host mammal are suitable for testing as an attenuated vaccine organism.

If further improvement in attenuation is desired, additional rounds of recombination and screening can be performed. DNA from the modified cells that have evolved toward inability to replicate under physiological conditions and ability to replicate under permissive conditions is recombined with a further library of DNA fragments, genomes, or partial genomes. The recombined DNA is introduced into the modified cells to produce further modified cells. Alternatively, one can recombine DNA among the modified cells that have evolved toward the desired function to produce the further modified cells. The further modified cells are then screened to identify those cells that have further evolved toward loss of ability to replicate under physiological conditions and toward ability to replicate under permissive conditions. These steps can be repeated as required until the further modified cells have lost the ability to replicate under physiological conditions in a host mammal and have acquired the ability to replicate under permissive conditions.

To obtain conditional-sensitive attenuated viral vaccines, the recombinant libraries of viral genomes can be introduced into suitable test cells which are similar to those of the target mammal in terms of the conditions that can affect viral replication. The virus-containing cells are cultured in increasing or decreasing temperatures, pHs, sugar content, or other condition, and the surviving viruses are chosen for new rounds of recombination and selection. The viruses that grow in altered temperatures or other condition are further analyzed to identify those that do not replicate at the temperature or other condition found in the mammal to be inoculated. In addition, it is generally desirable to analyze the conditional-

sensitive viruses for their capacity to induce pathology and protective immune response in the natural host. Backcrossing can be employed as discussed herein to obtain an attenuated vaccine in which the mutated gene or genes that are responsible for the conditional sensitivity are found in a virus that is otherwise unmodified in terms of its immunogenicity or other properties related to effectiveness as a vaccine.

In some embodiments, viral vaccines are screened by introducing the recombinant nucleic acids into allantoid cavities of embryonal eggs. Alternatively, *in vitro* tissue culture can be used. This selection scheme can employ as the culture cells a mixture of cells from various species that have different requirements (*e.g.*, for temperature or other condition) in tissue culture. This can overcome potential loss of survival and growth of the host cells that could occur when cultured under changing conditions *in vitro*.

**5. Altered host cell specificity**

Also provided by the invention are attenuated vaccines that are evolved to exhibit altered host specificity. One aim is to evolve viral, bacterial, or parasite strains that can specifically grow in cell types and/or organisms that allow efficient production of the vaccine strain, but cannot grow in the natural host cells or organisms in which they could cause a disease.

As one example, one can gradually change the selection pressure by using cell lines from different species. One can start by adapting the virus or other organism to simultaneously grow in the natural host and in phylogenetically related species. When generating attenuated human viruses, for example, one can start by adapting the virus to grow in both monkey cells and human cells. Thereafter, one would start selecting mutants that also grow in bovine cell lines, and the human cells are removed from the culture system. After recursive rounds of shuffling and selection, one is likely to be able to find a mutant strain that specifically grows in nonnatural host cell. In presently preferred embodiments, the mutant strain will not replicate in human cells. These screening methods can be done using pooled whole libraries of shuffled viruses, significantly reducing the numbers of samples that are handled.

In addition, the attenuated strains can be selected by screening for virus growth only in selected host cells rather than in many cell types (*i.e.* restriction of host cell specificity). This selection system allows generation of viruses that can replicate in certain

cells of the body, sufficient to elicit an immune response, but the restricted cell specificity will reduce the pathogenicity of the virus, thus preventing clinical symptoms.

Because the genes that regulate replication in host cells and the important antigenic determinants are likely to be encoded by different genes, backcrossing provides a means to retain the maximal number of the epitopes that are important for induction of protective immune responses.

**6. Rapid growth in manufacturing cells but reduced proliferation in host cells**

The invention also provides attenuated vaccines that exhibit rapid growth in manufacturing cells, but reduced proliferation in cells of the inoculated host. For example, one can first select for growth in manufacturing cell lines or culture conditions, and then generate a library of recombinant viruses and test those individually for growth in natural host cells. The clones with the slowest growth rate in the natural host cells are selected and subjected to new rounds of shuffling and selection. The selection for high growth in manufacturing cells and slow growth in host cells can be repeated. An advantage of this method is that simultaneous selection and screening for attenuation and high yield manufacturing is performed. Alternatively, one can first select individual clones that exhibit slow growth in host cells, and then select for growth in manufacturing cells. Again, selected mutants/chimeras are selected for new rounds of shuffling and selection.

**7. Screening based on adherence to target cells or target cell receptors**

The invention also provides methods for screening to identify viral vaccines that exhibit reduced adherence to target host cells. A library of mutant/chimeric viruses is incubated in the presence of cells of the type in which virus entry and replication is not desired. The viruses that do not bind or enter the cells are harvested from the supernatant. Kinetic studies can be performed to identify the optimal incubation time to most efficiently remove viruses which demonstrate specific binding to the cells. Also, several rounds of cell panning may be required to achieve optimal removal of the mutants/chimeras that have retained their capacity to bind to their specific cell surface receptors.

In addition to the intact cells, one can also use purified virus receptors in the panning in cases when the cell surface receptors for the given virus have been identified. In this system the viruses are mixed with purified (*e.g.*, recombinant) receptor in solution or

crosslinked to a plate. The viruses that bind to the receptors are captured by using, for example, monoclonal antibodies that are specific for the receptor or simply by allowing them to bind to the receptor crosslinked to the plate. The viruses are subsequently selected for growth in manufacturing cells, and the shuffling and selection is repeated as desired for further optimization. The selection is oscillated between binding to the specific receptors and growth in manufacturing cells.

**8. Selection based on sensitivity to complement**

Attenuated viruses can also be selected based on their sensitivity to complement or complement components. The library of recombinant viruses is generated by a method such as DNA shuffling, preferably family shuffling. Individuals clones of virus are expanded in manufacturing cells, and subsequently cultured in the presence of complement or complement components. Clones that demonstrate decreased virus titer upon exposure to complement are selected for new rounds of shuffling and selection. Viruses that are susceptible to killing by complement are likely to have strongly reduced capacity to induce pathology *in vivo*, yet they are likely to elicit immune responses that protect from future infections.

In addition, one can select virus mutants/chimeras that bind purified complement components (*e.g.*, C3 or components thereof). Binding of complement components directly to the virus may induce the cascade of complement mediated killing, and it may also cause opsonization of the virus rendering them more susceptible for killing by phagocytic cells, such as monocytes. The selection of mutants/chimeras that bind complement components can be done for example by panning or affinity column chromatography.

**9. Selection for growth in immunocompromised animals only**

Additional selection systems for attenuated vaccine strains include screening for growth in immunocompromised hosts only. Genetically immunocompromised strains of certain species are available, such immunodeficient mouse strains. Such strains include, but are not limited to, SCID mice, nude mice and mice rendered deficient in their genes encoding RAG1 or RAG2 genes. Importantly, however, practically any host species can be rendered transiently immunodeficient by drug treatment or irradiation. Immunosuppressive

drugs include, but are not limited to cyclosporin A, FK506, IL-10, soluble IL-2 receptor, steroids and anti-proliferative cancer drugs, such as methotrexate.

Individual virus, bacterial, or parasite clones are tested for their capacity to propagate in either genetically immunodeficient or transiently immunocompromised hosts. The individual clones that grow well in immunocompromised hosts are then tested for growth in normal hosts. Clones that grow in immunocompromised hosts but demonstrate slow growth in immunocompetent hosts are selected for new rounds of shuffling and selection.

In addition, viruses or other pathogens may grow, in immunocompromised hosts, in tissues where they normally do not grow. This allows selection of virus, bacterial, or parasite mutants/chimeras that have altered tissue specificity in immunocompromised hosts. In this selection system, one can infect the host with a library of viruses or other organisms and harvest viruses, bacteria, or parasites from novel target tissues. Individual clones from these tissues can then be tested for growth in immunocompetent hosts and manufacturing cells. Like in other shuffling formats, the selected mutants/chimeras can be subjected to additional rounds of shuffling and selection.

**10. Selection of pathogen variants that are killed by Ab response alone or by components of normal serum rendering them less virulent**

Immune defense against viral and other infections generally includes both cell mediated and humoral immune responses. Attenuated virus or cell strains can be selected by screening mutants/chimeras that are killed by either one of these mechanisms alone. For example, virus or cell mutants can be selected that are killed by specific antibodies alone, in the absence of T cell mediated immune responses. First, antisera are generated against the wild-type virus or cell by immunization. Thereafter, the shuffled library of virus or cell mutants/chimeras is mixed with the antisera and viruses or cells that are recognized by the antisera are selected for further analysis. These clones are then tested individually whether the antiserum neutralizes the function of the virus or cell. These studies can be done either *in vitro* or *in vivo*, or both. For example, one can first analyze whether the antisera, or pools of antisera, neutralize the virus or cell variants *in vitro* and then test these individual clones for their capacity to induce pathology and immunity *in vivo*.

In a variant of this selection system, one can also use sera from non-immunized hosts in order to identify virus or cell variants that are neutralized by serum components normally present in the host. As described above, the serum components that limit virus growth can include complement and complement components. However, the use of whole serum can allow evolution for killing by additional agents present in normal serum. Such mutants/chimeras are also likely to have an attenuated phenotype *in vivo*. As in other recombination and screening/selection formats, the selected mutants/chimeras can be subjected to new rounds of shuffling and selection if further optimization is desired.

**11. Combinations of selection systems to obtain optimal attenuation**

The optimal attenuation may often be obtained by combining several different selection mechanisms. For example, optimal attenuation can be achieved by simultaneously selecting mutants that grow at altered temperatures in a nonconventional host cell, *e.g.*, a cell line from a species other than the normal host for the virus or other pathogen. To achieve such a large change in the function of the virus or other organism, gradual selection pressure can be important. The temperature is readily increased or decreased gradually, and mixtures of cells derived from different species can be used to allow gradual adaptation to grow in nonhost tissues.

Furthermore, *in vitro* and *in vivo* selection systems can be used in combination. One can first select mutants that do not bind cells that normally are target cells for the given virus *in vitro*. In this format the natural host cells and the virus library are incubated as a mixture, and the mutants that do dot bind are then harvested, and the process is repeated until minimal or no viruses are removed from the pool by the host cells. Thereafter, the remaining population is analyzed *in vivo* either as pools or as individual recombinant viruses. Both induction of pathology and immune responses are measured. For example, one can inject the mutant viruses into animals, observe pathology, omit the ones that become sick, and then analyze the remaining animals by challenging them by the wild-type virus. Animals that have developed protective immunity after inoculation of the attenuated mutant strain will survive the challenge with the wild-type infectious virus. These can be selected for new rounds of shuffling and selection. Again, backcrossing with the wild-type virus can be used to retain maximal number of immunologically important residues.

### 12. Evolution of vaccine strains that can escape recognition by maternal antibodies

The invention is also useful for generating vaccines that are not recognized by maternal antibodies. Vaccines that exhibit reduced maternal antibody binding are expected to have improved efficacy. To obtain such vaccines, nucleic acids that encode either a whole or partial genome of a potential vaccine organism, or that encode a particular immunogenic polypeptide, are subjected to recombination as described herein. The resulting recombinant nucleic acids are then introduced into cells or viruses, which are then negatively selected so that the vaccines recognized by maternal antibodies are removed from the libraries. Typical methods to negatively select vaccines that are not recognized by maternal antibodies include, but are not limited to, affinity selection using flasks or columns coated with the maternal antibodies.

The family shuffling approach has an advantage in that one will simultaneously generate chimeras of the different strains. These chimeras can then be screened for optimal immunogenicity and crossprotection *in vivo*. These multivalent strains are likely to be more potent in inducing crossprotection against all different existing and emerging variants of the parent virus or cell. In some embodiments, libraries of the virus or cells that contain the recombinant nucleic acids are generated, and these libraries are initially screened for lack of binding to antibodies derived from animals that were previously immunized against the respective cell or virus. The immunogenicity of the selected vaccines can be verified by immunizing animals, and subsequently challenging the vaccinated animals with different strains of live virus or cells.

### 13. Evolution of more stable attenuated vaccine strains

The inherent lability of live organisms presents a challenge in terms of stabilizing and preserving viability of attenuated vaccine strains during manufacturing, storage, and administration (Burke *et al.*, *Crit. Rev. Ther. Drug. Carrier Syst.* (1999) 16: 1-83). Due to instability of several attenuated vaccine strains, shelf life of the vaccines is often limited, reducing the practicality of the vaccines for veterinary application or limiting their usage in undeveloped or tropical areas. Such examples include, but are not limited to, rinderpest vaccine for cattle (House and Mariner, *Dev. Biol. Stand.* (1996) 87: 235-44) and Varicella-Zoster virus (VZV) vaccines (Fanget and Francon, *Dev. Biol. Stand.* (1996) 87:

167-71). The stability of these vaccine strains can be improved by the methods of the invention.

Libraries of vaccine strains are generated by the recombination methods described herein, *e.g.*, DNA shuffling, and the resulting libraries are analyzed for stability. Different temperatures, formulations and time periods can be used to generate a selection pressure that only allows propagation of viruses or cells that have the desired properties. In addition, the virus and cell libraries can be freeze dried, reconstituted, and the most stable viruses or cells selected varying periods after reconstitution. The viruses or cells that demonstrate improved stability can be subjected to new rounds of shuffling and selection. Subsequent immunizations and challenge studies can be used to further evaluate the degree of attenuation, immunogenicity and stability.

### *IV. Use of the Attenuated Vaccines*

The attenuated vaccines of the invention are useful for treating and/or preventing the various diseases and conditions that are caused by viral or cellular pathogens. The attenuated vaccines obtained using the methods of the invention can be further modified to enhance their effectiveness in vaccination. For example, one can incorporate into the attenuated vaccines immunostimulatory sequences such as are described in copending, commonly assigned US Patent Application Serial No. 09/248,716, filed February 10, 1999. The vaccine vector can be modified to direct a particular type of immune response, *e.g.*, a $T_H1$ or a $T_H2$ response, as described in US Patent Application Serial No. 09/247,888, filed February 10, 1999. It is sometimes advantageous to employ a vaccine that is targeted for a particular target cell type (*e.g.*, an antigen presenting cell or an antigen processing cell); suitable targeting methods are described in copending, commonly assigned US patent application Serial No. 09/247,886, filed February 10, 1999.

The attenuated vaccines obtained using the methods of the invention find use not only for inducing a prophylactic or therapeutic immune response against the vaccine itself, but the backbone of the vaccines can be used to carry other pharmaceutically useful proteins into a cell. Such molecules include, for example, vaccine antigens, immunomodulatory molecules, therapeutic proteins, and the like.

Suitable formulations and dosage regimes for vaccine delivery are well known to those of skill in the art. The vaccines of the invention can be delivered to a mammal (including humans) to induce a therapeutic or prophylactic immune response. Vaccine delivery vehicles can be delivered *in vivo* by administration to an individual patient, typically by systemic administration (*e.g.*, intravenous, intraperitoneal, intramuscular, subdermal, intracranial, anal, vaginal, oral, buccal route or they can be inhaled) or they can be administered by topical application. Alternatively, vaccines can be delivered to cells *ex vivo*, such as cells explanted from an individual patient (*e.g.*, lymphocytes, bone marrow aspirates, tissue biopsy) or universal donor hematopoietic stem cells, followed by reimplantation of the cells into a patient, usually after selection for cells which have incorporated the vector.

A large number of delivery methods are well known to those of skill in the art. Such methods include, for example liposome-based gene delivery (Debs and Zhu (1993) WO 93/24640; Mannino and Gould-Fogerite (1988) *BioTechniques* 6(7): 682-691; Rose U.S. Pat No. 5,279,833; Brigham (1991) WO 91/06309; and Felgner *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84: 7413-7414), as well as use of viral vectors (*e.g.*, adenoviral (*see, e.g.*, Berns *et al.* (1995) *Ann. NY Acad. Sci.* 772: 95-104; Ali *et al.* (1994) *Gene Ther.* 1: 367-384; and Haddada *et al.* (1995) *Curr. Top. Microbiol. Immunol.* 199 ( Pt 3): 297-306 for review), papillomaviral, retroviral (*see, e.g.,* Buchscher *et al.* (1992) *J. Virol.* 66(5) 2731-2739; Johann *et al.* (1992) *J. Virol.* 66 (5):1635-1640 (1992); Sommerfelt *et al.*, (1990) *Virol.* 176:58-59; Wilson *et al.* (1989) *J. Virol.* 63:2374-2378; Miller et al., *J. Virol.* 65:2220-2224 (1991); Wong-Staal *et al.,* PCT/US94/05700, and Rosenburg and Fauci (1993) in *Fundamental Immunology, Third Edition* Paul (ed) Raven Press, Ltd., New York and the references therein, and Yu *et al., Gene Therapy* (1994) *supra.*), and adeno-associated viral vectors (*see*, West *et al.* (1987) *Virology* 160:38-47; Carter *et al.* (1989) U.S. Patent No. 4,797,368; Carter *et al.* WO 93/24641 (1993); Kotin (1994) *Human Gene Therapy* 5:793-801; Muzyczka (1994) *J. Clin. Invst.* 94:1351 and Samulski (*supra*) for an overview of AAV vectors; *see also*, Lebkowski, U.S. Pat. No. 5,173,414; Tratschin *et al.* (1985) *Mol. Cell. Biol.* 5(11):3251-3260; Tratschin, *et al.* (1984) *Mol. Cell. Biol.*, 4:2072-2081; Hermonat and Muzyczka (1984) *Proc. Natl. Acad. Sci. USA*, 81:6466-6470; McLaughlin *et al.* (1988) and Samulski *et al.* (1989) *J. Virol.*, 63:03822-3828), and the like.

"Naked" DNA and/or RNA that comprises a genome of an attenuated vaccine can be introduced directly into a tissue, such as muscle. *See, e.g.*, USPN 5,580,859. Other methods such as "biolistic" or particle-mediated transformation (*see, e.g.*, Sanford *et al.*, USPN 4,945,050; USPN 5,036,006) are also suitable for introduction of vaccines into cells of a mammal according to the invention. These methods are useful not only for *in vivo* introduction of DNA into a mammal, but also for *ex vivo* modification of cells for reintroduction into a mammal. As is the case for other methods of delivering vaccines, vaccine administration is repeated, if necessary, in order to maintain the desired level of immunomodulation.

Attenuated vaccines can be administered directly to the mammal. The vaccines obtained using the methods of the invention can be formulated as pharmaceutical compositions for administration in any suitable manner, including parenteral (*e.g.*, subcutaneous, intramuscular, intradermal, or intravenous), topical, oral, rectal, intrathecal, buccal (*e.g.*, sublingual), or local administration, such as by aerosol or transdermally, for prophylactic and/or therapeutic treatment. Pretreatment of skin, for example, by use of hair-removing agents, may be useful in transdermal delivery. Although more than one route can be used to administer a particular composition, a particular route can often provide a more immediate and more effective reaction than another route.

Pharmaceutically acceptable carriers are determined in part by the particular composition being administered, as well as by the particular method used to administer the composition. Accordingly, there is a wide variety of suitable formulations of pharmaceutical compositions of the present invention. *See, e.g.*, Lieberman, *Pharmaceutical Dosage Forms,*, Marcel Dekker, Vols. 1-3 (1998); *Remington's Pharmaceutical Science*, 15th ed., Mack Publishing Company, Easton, Pennsylvania (1980) and similar publications. A variety of aqueous carriers can be used, *e.g.*, buffered saline and the like. These solutions are sterile and generally free of undesirable matter. These compositions can be sterilized by conventional, well known sterilization techniques. The compositions can contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, toxicity adjusting agents and the like, for example, sodium acetate, sodium chloride, potassium chloride, calcium chloride, sodium lactate and the like. The concentration of attenuated vaccine in these formulations can vary

widely, and will be selected primarily based on fluid volumes, viscosities, body weight and the like in accordance with the particular mode of administration selected and the patient's needs.

Formulations suitable for oral administration can consist of (a) liquid solutions, such as an effective amount of the packaged nucleic acid suspended in diluents, such as water, saline or PEG 400; (b) capsules, sachets or tablets, each containing a predetermined amount of the active ingredient, as liquids, solids, granules or gelatin; (c) suspensions in an appropriate liquid; and (d) suitable emulsions. Tablet forms can include one or more of lactose, sucrose, mannitol, sorbitol, calcium phosphates, corn starch, potato starch, tragacanth, microcrystalline cellulose, acacia, gelatin, colloidal silicon dioxide, croscarmellose sodium, talc, magnesium stearate, stearic acid, and other excipients, colorants, fillers, binders, diluents, buffering agents, moistening agents, preservatives, flavoring agents, dyes, disintegrating agents, and pharmaceutically compatible carriers. Lozenge forms can comprise the active ingredient in a flavor, usually sucrose and acacia or tragacanth, as well as pastilles comprising the active ingredient in an inert base, such as gelatin and glycerin or sucrose and acacia emulsions, gels, and the like containing, in addition to the active ingredient, carriers known in the art. It is recognized that the attenuated vaccines, when administered orally, must be protected from digestion. This is typically accomplished either by complexing the vaccines with a composition to render it resistant to acidic and enzymatic hydrolysis or by packaging the vaccines in an appropriately resistant carrier such as a liposome. Means of protecting vectors from digestion are well known in the art. The pharmaceutical compositions can be encapsulated, *e.g.*, in liposomes, or in a formulation that provides for slow release of the active ingredient.

The attenuated vaccines, alone or in combination with other suitable components, can be made into aerosol formulations (*e.g.*, they can be "nebulized") to be administered via inhalation. Aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like.

Suitable formulations for rectal administration include, for example, suppositories, which consist of the packaged nucleic acid with a suppository base. Suitable suppository bases include natural or synthetic triglycerides or paraffin hydrocarbons. In addition, it is also possible to use gelatin rectal capsules which consist of a combination of

the packaged nucleic acid with a base, including, for example, liquid triglycerides, polyethylene glycols, and paraffin hydrocarbons.

Formulations suitable for parenteral administration, such as, for example, by intraarticular (in the joints), intravenous, intramuscular, intradermal, intraperitoneal, and subcutaneous routes, include aqueous and non-aqueous, isotonic sterile injection solutions, which can contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the blood of the intended recipient, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. In the practice of this invention, compositions can be administered, for example, by intravenous infusion, orally, topically, intraperitoneally, intravesically or intrathecally. Parenteral administration and intravenous administration are the preferred methods of administration. The formulations of attenuated vaccines can be presented in unit-dose or multi-dose sealed containers, such as ampoules and vials.

The dose administered to a patient, in the context of the present invention should be sufficient to effect a beneficial therapeutic and/or prophylactic response in the patient over time. The dose will be determined by the efficacy of the particular attenuated vaccine employed and the condition of the patient, as well as the body weight or vascular surface area of the patient to be treated. The size of the dose also will be determined by the existence, nature, and extent of any adverse side-effects that accompany the administration of a particular vaccine in a particular patient.

In determining the effective amount of the vaccine to be administered in the treatment or prophylaxis of an infection or other condition, the physician evaluates vaccine toxicities, progression of the disease, and the production of anti-vaccine vector antibodies, if any. In general, the dose equivalent of a naked nucleic acid from a vector is from about 1 μg to 1 mg for a typical 70 kilogram patient, and doses of vectors used to deliver the nucleic acid are calculated to yield an equivalent amount of therapeutic nucleic acid. Administration can be accomplished via single or divided doses.

In therapeutic applications, compositions are administered to a patient suffering from a disease (*e.g.*, an infectious disease or autoimmune disorder) in an amount sufficient to cure or at least partially arrest the disease and its complications. An amount adequate to accomplish this is defined as a "therapeutically effective dose." Amounts

effective for this use will depend upon the severity of the disease and the general state of the patient's health. Single or multiple administrations of the compositions may be administered depending on the dosage and frequency as required and tolerated by the patient. In any event, the composition should provide a sufficient quantity of the proteins of this invention to effectively treat the patient.

In prophylactic applications, compositions are administered to a human or other mammal to induce an immune response that can help protect against the establishment of an infectious disease or other condition. Subsequent challenge by the corresponding pathogen will trigger the immune response that has been primed by pre-exposure to the vaccine.

The toxicity and therapeutic efficacy of the attenuated vaccines provided by the invention are determined using standard pharmaceutical procedures in cell cultures or experimental animals. One can determine the $LD_{50}$ (the dose lethal to 50% of the population) and the $ED_{50}$ (the dose therapeutically effective in 50% of the population) using procedures presented herein and those otherwise known to those of skill in the art.

The attenuated vaccines of the invention can be packaged in packs, dispenser devices, and kits for administering genetic vaccines to a mammal. For example, packs or dispenser devices that contain one or more unit dosage forms are provided. Typically, instructions for administration of the compounds will be provided with the packaging, along with a suitable indication on the label that the compound is suitable for treatment of an indicated condition. For example, the label may state that the active compound within the packaging is useful for treating a particular infectious disease, autoimmune disorder, tumor, or for preventing or treating other diseases or conditions that are mediated by, or potentially susceptible to, a mammalian immune response.

## EXAMPLES

The following examples are offered to illustrate, but not to limit the present invention.

## Example 1

### Specific examples of using DNA shuffling to generate attenuated viruses to be used as vaccines or vaccine vectors

This Example describes several illustrative methods for using the methods of the invention to generate attenuated viral vaccines or vaccine vectors.

### A. Bovine viral diarrhea virus

Bovine viral diarrhea virus (BVDV) is a togavirus that is the most insidious and devastating viral pathogen of cattle in the United States (Vassilev *et al.*, *J. Virol.* (1997) 71:471-8). The virus causes immunosuppression, diarrhea, respiratory distress, abortion and persistent infection in calves. There are at least two serotypes of BVDV, as well as two biotypes (cytopathic and non-cytopathic). Because of the existing natural diversity in the BVDV strains, the virus offers an excellent starting point for evolution by family shuffling. The approach is particularly feasible, because stable full-length cDNA copies of BVDV have been established (Mendez *et al.*, *J. Virol.* 1998;72: 4737-45; Vassilev *et al.*, *J. Virol.* (1997) 71:471-8). As assayed by transfection of MDBK cells, uncapped RNAs transcribed from these cDNA clones were highly infectious (>$10^5$ PFU/μg). The recovered virus was similar in plaque morphology, growth properties, polyprotein processing, and cytopathogenicity to the parental BVDV strain (Mendez *et al.*, *J. Virol.* (1998) 72: 4737-45).

In addition, the principle of generating chimeras of infectious BVDV and antigenic determinants from other viruses has been demonstrated using rational design. More specifically, the coding region for the major envelope glycoprotein E2/gp53 in the molecular genomic clone of BVDV was substituted with that of the Singer strain, giving rise to a chimeric virus (Vassilev *et al.*, *J. Virol.* (1997) 71:471-8). However, such approaches to design of chimeric viruses suffer a significant drawback in that immunogenicity against the original virus is lost when the immunogenic envelope proteins are replaced by those derived from other pathogens.

To obtain chimeric viruses that have maintained the immunogenicity of the parental virus, DNA is used shuffling to generate chimeras between different viruses or their immunogenic fragments. Nucleic acids of different serotypes of BVDV are shuffled using family shuffling approach, for example. Either the entire infectious cDNA clones or the nucleic acids that encode the envelope proteins are shuffled, and a library of chimeric viruses

is generated. Shuffling the entire viral genome has the advantage that one is likely to simultaneously find solutions to attenuation and immunogenicity. MDBK cells, for example, are suitable for use in screening for attenuation *in vitro*, because the wild-type BVDV is highly infectious in these cells, and some strains cause a cytopathic effect (Mendez *et al.*, *J. Virol.* (1998) 72: 4737-45). Other ways by which one can screen for attenuation include, for example, analysis of temperature sensitivity, altered host-cell specificity, selection based upon sensitivity to complement, and selection for growth in immunocompromised animals only. Alternatively, one can also choose to shuffle only one virus strain to attenuate the virus first, and thereafter use family DNA shuffling of the immunogenic regions to generate chimeras that provide potent crossprotective immunity.

B. Marek's Disease Virus (MDV); improved manufacturing of viral vaccines by DNA shuffling

Marek's disease (MD) is a lymphoproliferative disease of chicken, which is characterized by malignant T cell-lymphoma formation (Morimura *et al.*, *J. Vet. Med. Sci.* 1998; 60:1-8). Relatively efficient vaccines are available to prevent the disease, but methods for manufacturing of the vaccine in particular need major improvements. The previously available attenuated MDV-vaccine is propagated in primary chick embryo fibroblasts, and the vaccine is the frozen, virus-infected cell preparation. Cell-free vaccines have also been tested, but their immunogenicity is inferior as compared to the cell-associated vaccines. In addition, MDV has been significantly evolving over the past 40 years to gain greater oncogenicity, and some of these viruses are not adequately controlled by the vaccines that are currently available (Biggs, *Philos. Trans. R. Soc. Lond. B. Biol Sci.* (1997) 352:1951-62). The development of vaccines against MDV is also hampered by the existence of multiple serotypes. Mixtures of different serotypes have been used in the vaccine preparations; such as vaccines based on the attenuated strains of serotype 1, 2 and 3, but failures resulting in disease progress have been reported indicating a need for production of new, more effective vaccines (Zelnik, *Acta Virol.* (1995) 39:53-63).

Molecular evolution of MDV by DNA shuffling provides solutions to the problems otherwise associated with the manufacturing, attenuation and immunogenicity of the vaccine. Generation of cell-free vaccine preparations provides a major improvement in the manufacturing process. Libraries of recombinant MDV nucleic acids, or fragments

thereof, are generated and screened for efficient propagation in cell culture in a cell-free manner. Cell lines from other species can be used to simultaneously achieve proper attenuation. The cell-free viruses that are then analyzed for their immunogenicity *in ovo* or in chicks. These cell-free preparations will provide major improvements to the manufacturing and storage of the vaccine, because the current vaccine preparations have to be shipped in liquid nitrogen containers to ensure the stability of the cells.

Because family DNA shuffling allows one to generate chimeric antigens and viruses, the approach is useful in generating vaccine strains that provide efficient cross-protection against all or most different serotypes of MDV. Furthermore, because maternal antibodies interfere with the vaccine, thus reducing its efficacy (Sharma and Graham, *Avian Dis.* (1982) 26: 860-70; Nazerian *et al.*, *Avian Dis.* (1996) 40: 368-76), and because DNA shuffling can generate new antigenic variants, this approach allows one to generate vaccine strains that are not recognized ~~by~~ maternal antibodies of previously vaccinated animals. By generating large libraries of vaccine strains, several different immunogenic variants are found. This enables vaccinations of different generations with antigenically variable vaccines, reducing the interference by maternal antibodies induced by immunization. Screening for reduced interference by maternal antibodies can be done *in vitro* using sera derived from vaccinated animals. For example, negative selection techniques, such as panning, are used to remove all strains that are recognized by antibodies from immunized animals. The antigenicity and immunogenicity of the remaining strains can subsequently be verified in *in vivo* studies. Protective immunity will be analyzed by challenging the chicken by different live MDV serotypes.

### C. Bovine herpes virus type-1 (BHV-1), also known as Infectious Bovine Rhinotracheatis virus (IBRV)

Bovine herpesvirus 1 (BHV-1), also called Infectious bovine rhinotracheatis virus (IBRV), replicates in a wide range of cell types and the disease manifestations include respiratory tract disease, conjunctivitis, vulvovaginitis, abortion, balanoposthitis, meningoencephalitis, alimentary tract disease and fatal systemic infection (Lupton and Reed, *Am. J. Vet. Res.* (1980) 41: 383). Immune responses to BHV-1 have been observed after exposure of animals to virulent virus, conventional live or killed vaccines, genetically engineered live virus vaccines, subunit vaccines and, more recently, following immunization

with plasmids encoding immunogenic antigens (Babiuk *et al.*, *Vet. Microbiol.* (1996) 53:31-42). Exposure to BHV-1 or its glycoproteins induced specific responses to the virus which are capable of neutralizing virus and killing virus infected cells. Killing of virus infected cells occurs after the expression of viral antigens on the cell surface of infected cells (Babiuk *et al.*, *Vet. Microbiol.* (1996) 53:31-42). BHV-1 may spread in the infected host by viremia, gaining access to a broader range of tissues and organs, and it may cause a variety of symptoms (Engels and Ackermann, *Vet. Microbiol.* (1996) 53: 3-15). Herpesviruses may also establish latency in neuronal or lymphoid cells, and during latency few viral antigens are synthesized. Upon reactivation, the viruses re-establish the lytic cycle of replication. Although a vigorous immune response is often induced during the primary viral infection, these mechanisms help the herpesviruses to escape the host immune system during latency and to a lesser degree also during reactivation (*Id.*).

Evolution of attenuated BHV-1 strains is achieved by random DNA shuffling of the virus or targeted evolution of virus components that are critical to the penetration and propagation of the virus. One example of such virus component is the glycoprotein H (Meyer *et al.*, *J. Gen. Virol.* (1998) 79: 1983-7). gH is a structural component of the virus and forms a complex with glycoprotein gL. Experiments with gH-deficient BHV-1 demonstrated that gH is crucial in the infectious cycle of the virus and is involved in virus entry and cell-to-cell spread, but not in the attachment of the virus (*Id.*). Another example of useful component of BHV-1 for molecular evolution by DNA shuffling is the glycoprotein D (gD), which has also been shown to be an essential component involved in virus entry (Hanon *et al.*, *Virology* (1999) 257: 191-197). BHV-1 viruses devoid of gD (BHV-1 gD-/-) are able to bind to BL-3 cells, but they are no longer able to induce apoptosis (*Id.*). Furthermore, immunity against gD has been shown to confer resistance to BHV-1 replication in cattle (Zhu and Letchworth, *Vaccine* (1996) 14: 61-9).

DNA shuffling is used to cause molecular evolution of gH or gD so as to generate BHV-1 variants that have altered capacity to spread from cell to cell, a crucial event in pathogenesis. Taken together, family DNA shuffling of gD and or gH provides means to simultaneously attenuate the virus and to generate chimeric viruses that provide protection against multiple serotypes.

More specifically, in this approach the virus, or fragments thereof, such as gH or gD, will be shuffled and a library are generated. The shuffled fragments can be incorporated into the virus backbone using conventional techniques known to those skilled in the art. The library of viruses are then selected for attenuation. A number of different approaches for selection can be taken, some of which have been previously described during attempts to attenuate the wild-type virus. Such selection techniques can also be applied to selection of attenuated, DNA shuffled vaccine strains. These methods include rapid passage in bovine cell culture (Schwartz *et al.*, *Proc. Soc. Exp. Biol. Med.* (1957) 96: 453) or by adaptation to porcine or canine cell cultures (Schwartz *et al.*, *Proc. Soc. Exp. Biol. Med.* (1958) 97: 680; Zuschek *et al.*, *J. Am. Vet. Med. Assoc.* (1961) 139: 236). In addition, the virus can be adapted to grow in cell culture in reduced temperature (30°C), or by selection of temperature sensitive mutants (56 °C for 40 minutes) (Inaba, *J. Jpn. Vet. Med. Assoc.* (1975) 28: 410; Bartha, *Dev. Biol. Stand.* (1974) 26:5).

Intranasally administered BHV-1 has also been attenuated by serial passage in rabbit cells cultured *in vitro*, or were modified by treatment with $HNO_2$ followed by selection of temperature sensitive mutants (Todd, *Can. Vet. J.* (1974) 15: 257; Zygraich *et al.*, *Res. Vet. Sci.* (1974) 16: 328). These selection techniques are also useful in identifying attenuated strains of shuffled viruses; the library of shuffled viruses can contain viruses that have additional improvements in addition to attenuation. A crucial advantage as compared to previously described attenuation techniques is the fact that efficient chimerism between the different serotypes can be achieved, which enables one to generate crossprotective strains that provide improved immune responses *in vivo*. The analysis of the efficacy of the immune response in animals can be done by analyzing immune parameters in the sera and circulating lymphocytes in immunized animals. Moreover, the protective and crossprotective immune responses can be studied by challenging the immunized animals with the wild-type viruses of different serotypes. The viruses that demonstrate the best attenuation combined with the most potent crossprotective immune response *in vivo* are selected for further rounds of shuffling and selection, when desired.

D. Infectious bronchitis virus

Infectious bronchitis virus (IBV) is a member of the family Coronaviridae and causes highly contagious respiratory and reproductive disease in chickens. IBV has a single-stranded, positive sense RNA genome of 27.6 kb. The construction of a full-length clone of IBV downstream of the bacteriophage T7 promoter has been described (Penzes *et al.*, *J. Virol.* (1996) 70: 8660-8). Electroporation of *in vitro* T7-transcribed RNA from the two different constructs into IBV helper virus-infected cells resulted in the replication and packaging of the RNA (Penzes *et al.*, *J. Virol.* (1996) 70:8660-8).

The three structural proteins of IBV are the spike glycoprotein (S protein), the membrane glycoprotein (M protein) and the nucleocapsid protein (N protein) (Jia *et al.*, *Arch. Virol.* (1995) 140: 259). There are at least ten serotypes of IBV. Massachusetts, Connecticut, Arkansas and California serotypes are the most persistent in the US. Mutations and crossovers are common mechanisms for the generation of new serotypes and provide means for the virus to escape naturally existing immunity or that induced by vaccinations (Keck *et al.*, *J. Virol.* (1988) 62: 1810). More immunogenic and crossprotective vaccine strains are needed, and molecular evolution technologies provide improved and faster means to generate novel chimeras that can be screened for the desired properties using *in vitro* and *in vivo* screenings.

Libraries of IBV strains are generated by shuffling, and the desired clones are selected from the library by analyzing the entire library, pools thereof, or individual clones. More specifically, the degree of attenuation can be studied for example by using tracheal organ cultures (TOC) and oviduct organ cultures (OOC) (Raj and Jones, *Vaccine* (1997) 15: 163-8). Ciliostasis (CD50), immunofluorescence staining (IFID50) and organ culture infectivity (OCID50) have been shown to associate with attenuation and are useful methods for screening candidate live respiratory viral vaccines for attenuation (*Id.*). These methods are also useful when selecting shuffled, attenuated strains of IBV. Subsequent immunizations and challenge studies in animals can be used to further evaluate the degree of attenuation, immunogenicity and cross-protection.

E. Evolution of stable Yellow Fever vaccine strains by DNA shuffling

The stability of these vaccine strains can be improved by DNA shuffling technology. This Example describes the evolution of yellow fever (YF) virus, which is an

example of a vaccine strain, the stability of which can be improved by the methods described here. YF is an acute mosquito-borne viral haemorrhagic fever that has reemerged across Africa and in South America. A total of 18,735 yellow fever cases and 4,522 deaths were reported from 1987 to 1991. This represents the greatest reported amount of yellow fever activity for any 5-year period since 1948 (Robertson *et al.*, *JAMA* (1996) 276: 1157-62). In Africa, a large proportion of cases have occurred in children. There is an efficient vaccine against yellow fever available, but financing the vaccine has been difficult for the poorest in the world (Robertson *et al.*, *JAMA* (1996) 276: 1157-62).

The stability of YF vaccine is a major problem in undeveloped countries and tropical areas. The lyophilized vaccine strain without stabilizers deteriorates rapidly when exposed to temperatures above -20 °C (Monath, *Dev. Biol. Stand.* (1996) 87: 219-25). Additives, such as sugars, amino acids, and divalent cations have improved the stability of the vaccine preparations. However, despite the relatively good stability of these vaccine formulations stability when freeze dried, the vaccine is unstable after reconstitution and must be discarded after one hour (Monath, *Dev. Biol. Stand.* (1996) 87: 219-25). Improvements in vaccine stability after reconstitution would significantly reduce cost, stretch supplies of the vaccine, and would also reduce the frequency of vaccine failures due to use of degraded vaccine.

In this Example, infectious cDNA clones of YF are subjected to recombination by, for example, DNA shuffling. The resulting virus libraries are analyzed for stability. Selection is conducted under different temperatures, formulations and time periods, as desired, to obtain suitable YF viruses that are stable under the conditions of interest. Only those viruses that can propagate under such conditions survive the selection. In addition, the virus libraries can be freeze dried, reconstituted, and the most stable viruses selected varying periods after reconstitution. The viruses that demonstrate improved stability can be subjected to new rounds of shuffling and selection. Subsequent immunizations and challenge studies will further evaluate the degree of attenuation, immunogenicity and stability.

**F. Influenza A virus**

Epidemic infections with influenza A continue to associate with significant morbidity and mortality in the general population, particularly among the elderly and other

high risk patients (Calfee and Hayden, *Drugs* (1998) 56: 537-53). Tens of thousands of deaths occur each year despite the availability of relatively efficient vaccines. Efficient control of the disease has not been achieved through immunization programs because of incomplete protective efficacy and antigenic variations of the virus. Vaccinations must be given annually because of the antigenic changes that the virus undergoes, and because the antibody responses decrease significantly over time (Rimmelzwaan *et al.*, *Vaccine* (1999) 17: 1355-8).

Influenza A virus belongs to family Orthomyxoviridae, which are segmented, negative-stranded viruses. Additional members of the family are Influenza B and Influenza C viruses. Viral replication occurs after synthesis of the mRNAs and requires synthesis of the viral proteins. Complete infectious segmented negative-strand viruses have been successfully recovered from cloned cDNA (Bridgen and Elliott, *Proc. Nat'l. Acad. Sci. USA* (1996) 93: 15400-4). Plasmids encoding full-length cDNA copies of three *Bunyamwera bunyavirus* RNA genome segments flanked by bacteriophage T7 promoter and hepatitis delta virus ribozyme sequences were capable of encoding infectious virus with the characteristic of the parental cDNA clones (Bridgen and Elliott, *Proc. Nat'l. Acad. Sci. USA* (1996) 93: 15400-4). Similarly, full-length, infectious vesicular stomatitis virus (VSV), the prototypic nonsegmented negative-strand RNA virus, has been recovered from a cDNA clone (Whelan *et al.*, *Proc. Nat'l. Acad. Sci. USA* (1995) 92: 8388-92). These data illustrate the feasibility of shuffling the entire genomes of negative-strand RNA viruses. In addition, foreign genes have been successfully introduced into Influenza A genome (Luytjes *et al.*, *Cell* (1989) 59: 1107-13), indicating that Influenza A virus genomes can be successfully engineered. Infectious Influenza A virus has been recovered after transfection of cDNA encoding the PB2 polymerase gene, followed by transfection of the RNA transcripts (Subbarao *et al.*, *J. Virol.* (1993) 67, 7223-8). Therefore, as an alternative to whole viral genome shuffling, novel attenuated vaccine strains of Influenza A can be generated by shuffling individual segments of the virus, further illustrating the feasibility of DNA shuffling approach in evolution of Influenza A viruses.

Libraries of Influenza A viruses are generated by shuffling the entire genome, or segments thereof (such as, for example, the PB2 polymerase gene). Because of the large antigenic diversity of Influenza viruses, the segments encoding the immunogenic proteins,

such as the nucleocapsid, matrix proteins, hemagglutinin, nucleoprotein or neuramididase, provide additional targets for molecular evolution by DNA shuffling. Because live attenuated influenza A virus vaccines have been widely produced by the transfer of attenuating genes from a donor virus to new epidemic variants of influenza A virus (Subbarao *et al.*, *J. Virol.* (1993) 67: 7223-8), the shuffled segments can be introduced back to other Influenza strains of interest. The shuffled viruses can be selected using multi-tiered screening process including selection for growth in manufacturing cells, identification of temperature sensitive mutants, screening for presence of multiple epitopes by polyclonal antibodies, analysis of potent crossprotective immune response *in vivo*, or all of the above. The best variants can be selected for new rounds of shuffling and screening when desired.

G. Respiratory syncytial virus (RSV)

Respiratory syncytial virus (RSV) is the most important cause of lower respiratory tract infection during infancy and early childhood (Domachowske and Rosenberg, *Clin. Microbiol. Rev.* (1999) 12: 298-309). RSV infection can be devastating in elderly and immunosuppressed individuals (Wyde, *Antiviral Res.* (1998) 39: 63-79). The infection generally results in the development of anti-RSV neutralizing-antibodies, but these are often suboptimal during an infant's initial infection. Reinfection during subsequent exposures is common, and efficient vaccines are highly desired.

Functional, infectious RSV has been recovered from expressed, cloned cDNAs. RSV was expressed in a functional form by coexpressing the viral polymerase protein, phosphoprotein, and nucleocapsid protein from cDNA clones (Yu *et al.*, *J. Virol.* (1995) 69: 2412-9). Such cDNA clones provide an excellent starting point for molecular evolution by DNA shuffling. Several different antigenic groups of RSV have been identified (Sanz *et al.*, *Virus Res.* (1994) 33: 203-17), and very high mortality rates, up to 78%, have been observed in immunocompromised patients (Harrington *et al.*, *J. Infect. Dis.* (1992) 165: 987-93). Therefore, efficient vaccines that provide protection against multiple different variants of RSV are highly desired.

Although the antigenic heterogeneity of RSV is a challenge for vaccine development, these naturally existing variants of the pathogen provide a pool of existing sequences that can be used to generate a family shuffled library of RSV. Libraries of RSV

viruses are generated by shuffling infectious cDNA clones derived from various RSV isolates. The resulting RSV variants are screened for attenuation and for their properties as vaccines. The stability of the viruses can be selected *in vitro* by storing the vaccine strains for prolonged periods of time. Moreover, the attenuation will be evaluated in animal models for lack of disease or for reduced levels of symptoms. These attenuated strains are be further analyzed for their capacity to induce protective immune responses *in vivo*. This can be achieved by challenging the immunized animals by live wild-type pathogens and scoring the different strains for their level of attenuation and efficacy in inducing protective immune responses. The optimal strains with desired properties can be selected for new rounds of shuffling and screening.

H. Canine Distemper virus (CDV)

Canine distemper virus (CDV) is a morbillivirus that affects the neurologic system and causes a frequently fatal systemic disease in a wide range of carnivore species, including domestic dogs. Classical serology provides data of diagnostic and prognostic values and is also used to predict the optimal vaccination age of pups, because maternal antibodies can interfere with the vaccines (Appel and Harris, *J. Am. Vet. Med. Assoc.* (1988) 193: 332-3). Several antigenically different strains of CDV have been identified (Ohashi *et al.*, *J. Vet. Med. Sci.* (1998) 60: 1209-12; Carpenter *et al.*, *Vet. Immunol. Immunopathol.* (1998) 65(2-4): 259-66), and the virus appears to frequently cross host species among carnivores (*Id.*). The antigenic heterogeneity of the different strains is a challenge for vaccine development, but it also provides and excellent genetic diversity that enables further evolution *in vitro* using the methods of the invention.

Infectious morbilliviruses have been reconstituted from cDNA (Cathomen *et al.*, *EMBO J.* (1998) 17: 3899-908), indicating that the use of DNA shuffling to generate attenuated morbilliviruses, such as attenuated vaccine strains of CDV, is feasible. More stable and immunogenic viruses are highly desired. In addition, because maternal antibodies can interfere with CDV vaccines, vaccines not recognized by such antibodies are expected have improved efficacy (Appel and Harris, *J. Am. Vet. Med. Assoc.* (1988) 193: 332-3). Libraries of CDV viruses are generated by DNA shuffling of infectious cDNA, and the resulting viruses are screened for their properties as vaccines. Studies *in vivo* in dogs are

used to identify attenuated strains (strains that cause inhibited or no clinical disease) that provide efficient immune response upon challenge of the immunized animals by wild-type viruses. In addition, the viruses can be selected for increased stability *in vitro* by storing the vaccine strains for prolonged periods of time. Strains with desired properties can be subjected to new rounds of shuffling and screening if further improvement is desired.

## Example 2

### Evolution of attenuated alphaviruses; VEE as a vehicle for airborne vaccinations

This Example describes the use of DNA shuffling to evolve attenuated alphaviruses, which are useful as a vehicle for vaccines that are suitable for airborne administration.

The alphaviruses are a genus of 26 enveloped viruses that cause disease in several species, including humans and domestic animals. Mosquitoes and other hematophagous arthropods serve as vectors (Strauss and Strauss, *Microbiol. Rev.* (1994) 58: 491-562). Alphaviruses include Venezuelan Equine Encephalitis virus (VEE), Semliki Forest virus (SFV) and Sindbis virus (SIN), which have also been targets of interest as vaccine vectors, because of the broad host range and superior infectivity of these viruses. The generation of high-titer recombinant alphavirus stocks has enabled high-level expression of several nuclear, cytoplasmic, membrane-associated and secreted proteins in a variety of cell lines and primary cell cultures (Lundstrom, *J. Recept. Signal. Transduct. Res.* (1999) 19: 673-86).

The complete sequences of the positive stranded RNA genomes of at least eight alphaviruses have been determined, and partial sequences are known for several others (Strauss and Strauss, *Microbiol. Rev.* (1994) 58: 491-562). Importantly, full-length cDNA clones from which infectious RNA can be recovered have been constructed for four alphaviruses, including VEE, SFV and SIN (Davis *et al.*, *Virology* (1989) 171:189-204; Polo *et al.*, *Proc. Nat'l. Acad. Sci. USA* (1999) 96: 4598-603; Atkins *et al.*, *Mol. Biotechnol.* (1996) 5: 33-8). Therefore, VEE, SIN and SFV are particularly good examples of alphaviruses that can be attenuated by DNA shuffling. In the present Example, although VEE is described in detail, the high degree of structural and functional relatedness among alphaviruses allows one to use a similar approach for other alphaviruses.

VEE is an unusual alphavirus in that it is also highly infectious for both humans and rodents by aerosol inhalation. Therefore, attenuated strains of VEE provide vehicles to deliver the vaccines in an aerosol formulation, which enables rapid vaccinations of large populations of humans or animals at the same time. Aerosol vaccination with inactivated or attenuated recombinant pathogens has been shown to be an efficient way to induce local protection against lung diseases, and aerosol vaccinations have also been shown to protect against infectious diseases (Hensel and Lubitz, *Behring. Inst. Mitt.* (1997) 98: 212-9). Because VEE can also be used as a vector to deliver antigens from other pathogenic organisms, aerosol mediated vaccinations with attenuated strains of VEE are expected to provide very efficient and rapid vaccination protocols against a variety of diseases. VEE is an unusual virus also because its primary target outside the central nervous system is the lymphoid tissue, and therefore, attenuated variants may provide means to target vaccines or pharmaceutically useful proteins to the immune system.

There are at least seven subtypes of VEE that can be identified genetically and serologically. Based on epidemiological data the virus isolates fall into two main categories: I-AB and I-C strains, which are associated with VEE epizootics/epidemics, and the remaining serotypes, which are associated primarily with enzootic vertebrate-mosquito cycles and circulate in specific ecological zones (Johnston and Peters, In *Fields Virology*, Third Edition, eds. B.N. Fields *et al.*, Lippincott-Raven Publishers, Philadelphia, 1996).

Libraries of recombinant VEE viruses are created and screened to identify those recombinant viruses that exhibit an attenuated phenotype. Libraries are generated by subjecting to DNA shuffling either the entire infectious cDNA clones or the regions known to play a role in pathogenesis and protection, such as the envelope protein. These libraries and individuals chimeras/mutants thereof are subsequently screened for their capacity to induce widely crossreacting and protective antibody responses. Alternatively or additionally, the libraries are screened to identify those recombinant viruses that have the capacity to immunize mammals through aerosol delivery. The libraries are delivered in aerosol formulations and the optimal viruses are subsequently identified in animals that develop specific antibodies and survive the infection, as an indication that the strain was sufficiently attenuated. When the libraries are generated by family shuffling of the different serotypes of

VEE, one also is able to identify the most crossprotective strains by challenging the immunized animals with various serotypes of the live, wild-type pathogen.

Furthermore, because VEE can act as a vehicle to deliver foreign antigens, attenuated VEE strains encoding antigens from other pathogens can be useful in vaccinations against a variety of diseases. Packaging cell lines can be used to enable production of biologically active vector particles even when the structural proteins of the virus have been replaced by foreign antigens (Polo *et al.*, *supra.*). These packaging cells can be engineered to encode those structural proteins that were shown to mediate efficient immunization, for example through aerosol mediated immunization. As an example, the gene that encodes VEE envelope protein is replaced with that which encodes hantavirus glycoproteins. Because VEE is known to infect mice by airborne challenge, the approach provides a method for aerosol mediated vaccinations of wild mice against hantavirus infections, which should greatly reduce the risk to humans of encountering hantavirus in endemic areas. A significant advantage of the method is that the laboratory setting is useful in the initial screening, because mice can be infected with VEE by airborne challenge (Wright and Phillpotts, *Arch. Virol.* (1998) 143: 1155-62). The best clones are subjected to new rounds of shuffling and selection, when desired.

## Example 3

### Infectious bursal disease virus (IBDV)

Infectious bursal disease (IBD) emerged in 1957, spread rapidly, and became recognized throughout the U.S. broiler and commercial egg production areas by 1965 (Lasher and Davis, *Avian Dis.* (1997) 41: 11-9). Infectious bursal disease virus (IBDV) attacks the Bursa of Fabritius causing immunosuppression and death. The acute stage of IBD, the immunosuppression that follows, and the widespread distribution of IBD virus (IBDV), contribute to the major economic significance of the disease (Saif, *Poult. Sci.* (1998) 77:1186-9). A live attenuated vaccine was developed and demonstrated a rather good efficacy (Lasher and Davis, *supra.*). However, in the mid-1980s novel variants emerged, such as the Delaware variants, and variants with increased virulence were identified in Europe and Asia in 1989 (*Id.*). In addition, maternal immunity significantly reduces the

efficacy of the vaccines (Bayyari *et al.*, *Avian Dis.* (1996) 40: 588-99), and vaccines with improved efficacy are desired.

IBDV is a double stranded RNA virus and belongs to the family Birnaviridae of the genus Avibirnavirus (Nagarajan and Kibenge, *Can. J. Vet. Res.* (1997) 61: 81-8). The genome consists of two segments, designated A and B. cDNA clones encoding the segments A and B of IBDV have been shown to generate viable virus progeny. Independent full-length cDNA clones were constructed that contained the entire coding and noncoding regions of RNA segments A and B (Mundt and Vakharia, *Proc. Nat'l. Acad. Sci. USA* (1996) 93: 11131-6). These cDNAs provide an excellent starting point for DNA shuffling. In addition, because the NS protein is dispensable for viral replication *in vitro* and *in vivo* plays an important role in viral pathogenesis (Yao *et al.*, *J. Virol.* (1998) 72: 2647-54), the segment encoding the NS protein is a particularly useful target for evolution by DNA shuffling. Furthermore, the existence of several natural variants of IBDV (Nagarajan and Kibenge, *supra.*), provides a source of natural diversity for family DNA shuffling to generate attenuated strains and chimeras of the different serotypes.

Libraries of recombinant IBDV nucleic acids are generated by shuffling both A and B segments. In addition, libraries of viruses that contain shuffled NS genes are generated by shuffling NS genes that are then incorporated back into the viral genome. The libraries of viruses is initially screened for attenuation *in vitro*. However, because chicken are accessible in large numbers, the primary screen is for attenuation and immunogenicity *in vivo*. More specifically the viruses or pools of viruses are injected *in ovo*, at one day of age or applied to the drinking water at 7-14 days of age. Attenuated viruses not causing clinical disease are identified, and the immunized animals are subsequently screened for protective immune responses by challenging them with the live wild-type viruses. The best viruses are subjected to new rounds of shuffling and selection, when desired.

## Example 4

### Canine parvovirus; evolution of vaccine strains that can escape interference by maternal antibodies

Canine parvovirus is a newly emerged pathogen of dogs that was identified in the late 1970s (Pollock and Coyne, *Vet. Clin. North Am. Small. Anim. Pract.* (1993) 23:555-

68). Although parvoviral enteritis was initially seen as epidemic disease in all dogs, the disease now primarily occurs in 1- to 6-month-old dogs. Interference by maternal antibodies accounts for the vast majority of vaccine failures (*Id.*; Waner *et al.*, J. Vet. Diagn. Invest. (1996) 8: 427-32). Maternally derived hemagglutination inhibition (HI) titers have also been shown to correlate to the efficacy of attenuated CPV vaccines (Hoare *et al., Vaccine* (1997) 15: 273-5). Intranasal vaccination of pups with maternal antibodies has demonstrated some success in avoiding the interference, but vaccines that have improved immunogenicity in the presence of such antibodies are desired (Buonavoglia *et al., Zentralbl. Veterinarmed.* (1994) 41: 3-8). Molecular virologic methods have also revealed continued evolution of the virus adding to the antigenic heterogeneity of the virus and to the challenges in vaccine development.

In this Example, canine parvovirus strains are evolved and screened to identify those that, when used as vaccines, do not interfere with maternal antibodies. Either the entire vaccine strain, or the immunogenic antigens of the vaccine alone, is shuffled and negatively selected so that the vaccines that are recognized by maternal antibodies are removed from the libraries. For example, affinity selection using flasks or columns coated with the maternal antibodies are used. To identify strains that have epitopes derived form multiple different strains of CPV, monoclonal antibodies that discriminate among different strains of CPV (Sagazio *et al., J. Virol. Methods* (1998) 73: 197-200) can be used, provided these antibody specificities are not a major component of the maternal antibodies. These multivalent strains are likely to be more potent in inducing crossprotection against all different existing and emerging variants of CPV.

Molecular infectious clones of parvoviruses have been generated (Bloom *et al., J. Virol.* (1993) 67: 5976-88). Such clones provide a suitable substrate for family DNA shuffling. The family shuffling approach has the advantage that one will simultaneously generate chimeras of the different strains. These chimeras can then be screened for optimal immunogenicity and crossprotection *in vivo*. In one approach, libraries of canine parvoviruses are generated, and these libraries are initially screened for lack of binding to antibodies derived from dogs previously immunized against CPV. The immunogenicity of the selected vaccines can be verified by immunizing dogs, and subsequently challenging the vaccinated animals with different strains of live CPV. Vaccines that have reduced binding to

maternal antibodies are expected have improved efficacy, and chimeric vaccines are expected to provide crossprotective immune responses.

## Example 5

### Flaviviruses; evolution of chimeric attenuated vaccine strains of Dengue viruses

Dengue viruses are transmitted through mosquito bites and 50-100 million people each year. The virus causes serious clinical manifestations, such as dengue hemorrhagic fever (DHF). There are four major serotypes of Dengue virus, namely Dengue 1, 2, 3 and 4. The spread of the four dengue virus serotypes had led to increased incidence of DHF and estimated 2.5 billion people at risk of the infection (Cardosa, *Br. Med. Bull.* (1998) 54: 395-405). No efficient vaccine against dengue infections are currently available.

The envelope protein of Dengue virus has been shown to provide an immune response that protects from a future challenge with the same strain of virus. However, the levels of neutralizing antibodies produced in response to subunit vaccines are relatively low and protection from live virus challenge is not always observed. For example, mice injected with genetic vaccines encoding envelope protein of Dengue-2 virus develop neutralizing antibodies when analyzed by *in vitro* neutralization assays, but the mice did not survive the challenge with live Dengue-2 virus (Kochel *et al.*, *Vaccine* (1997) 15: 547-552). However, protective immune responses were observed in mice immunized with recombinant vaccinia virus expressing Dengue 4 virus structural proteins (Bray *et al.*, (1989) *J. Virol.* 63: 2853. Furthermore, live attenuated Dengue-2 vaccines protected monkeys against homologous challenge (Velzing *et al.*, *Vaccine* (1999) 17: 1312-20). These data suggest that live attenuated Dengue vaccines will be an efficient approach to Dengue vaccine development. However, a tetravalent vaccine that induces neutralizing antibodies against all four strains of Dengue is required to avoid antibody-mediated enhancement of the disease when the individual encounters Dengue virus of different serotype. Family DNA shuffling provides a technology that simultaneously enables attenuation and generation of chimeric viruses that protect against all four serotypes.

Infectious cDNA clones of Dengue viruses have been generated and the immunogenic envelope genes sequenced, providing a good starting point for shuffling (Lai *et al.*, *Proc. Nat'l. Acad. Sci. USA* (1991) 88: 5139-43; Lanciotti *et al.*, *J. Gen. Virol.* (1994)

75: 65-75; Kinney *et al.*, *Virology* (1997) 230: 300-8; Puri *et al.*, *Virus Genes* (1998) 17: 85-8).

In this Example, the entire cDNAs encoding dengue viruses are shuffled to obtain proper attenuation. Alternatively, selected regions of the virus are be shuffled to generate attenuated strains. Such viral genes that are useful targets for shuffling include the prM, E, NS1 and NS3 genes (Pryor *et al.*, *J. Gen. Virol.* (1998) 79: 2631-9; Valle and Falgout, *J. Virol.* (1998) 72: 624-32). Additionally, the envelope genes of all four serotypes can be shuffled separately to generate crossprotective antigens, which can subsequently be incorporated back into the attenuated clones. Efficient chimerism is expected to be achieved also by family shuffling the entire infectious cDNA clones. The E proteins of the different dengue viruses share 62% to 77% of their amino acids. Dengue 1 and Dengue 3 are most closely related (77% identical), followed by Dengue 2 (69%) and Dengue 4 (62%). These identities are well in the range that allows efficient family shuffling.

Initially, screening for proper attenuation can be performed *in vitro*. For example, passaging in primary dog kidney cells *can be* used (Puri *et al.*, *J. Gen. Virol.* (1997) 78: 2287-91). One can also select temperature sensitive mutants. The immunogenicity and crossprotection provided by the attenuated vaccine strains can be further studied in mice challenged with the various live wild-type pathogens. If desired, further studies are performed in large non-human primates to verify the level of attenuation. This enables simultaneous studies of the immune responses. Infectivity in monkeys has been shown to correlate with that in humans, indicating that the monkey model is useful in selecting the properly attenuated strains (Marchette *et al.*, *Am. J. Trop. Med. Hyg.* (1990) 43: 212-8). Eventually, the attenuated, shuffled vaccines are tested in humans for their capacity to protect against infections with the different dengue serotypes.

## Example 6

### Flaviviruses; Hepatitis C virus (HCV)

Hepatitis C virus (HCV) infection is a major health problem that leads to cirrhosis and hepatocellular carcinoma in a substantial number of infected individuals, estimated to be 100-200 million worldwide (Martin *et al.*, *Biochemistry* (1998) 37: 11459-68). Vaccines or effective treatments for HCV infection are not available. Antigenic

heterogeneity of different strains of hepatitis C virus (HCV) is a major problem in development of efficient vaccines against HCV. Antibodies or CTLs specific for one strain of HCV typically do not protects against other strains. Multivalent vaccine antigens that simultaneously protect against several strains of HCV would be of major importance when developing efficient vaccines against HCV.

HCV is a particularly suitable target for attenuation by DNA shuffling because full-length cDNAs have been constructed (Yanagi *et al.*, *Proc. Nat'l. Acad. Sci. USA* (1997) 94: 8738-43) and infectious HCV cDNA can be obtained from infectious blood samples (Aizaki *et al.*, *Hepatology* (1998) 27:621-7). Thus, the naturally existing diversity of HCV can be directly isolated from the blood of infected individuals providing an excellent starting point for family DNA shuffling.

Several approaches to attenuation of HCV by DNA shuffling are available. First, the entire cDNA clones may shuffled and the libraries screened using the methods described below. Secondly, the polymerase gene may be replaced by that derived from other related viruses, such as other flaviviruses. Polymerase genes from flaviviruses, such as Dengue virus, may allow replication of the HCV in tissue culture cells, which is not possible with the wild-type HCV. The NS5B protein of HCV is an RNA-dependent RNA polymerase that is required for virus replication (Behrens *et al.*, *EMBO J.* (1996) 15: 12-22). Because other flaviviruses readily replicate *in vitro*, RNA-polymerase genes derived from other flavivirus family members may allow HCV to replicate *in vitro*. Importantly, family DNA shuffling of flavivirus derived polymerase genes can be used to optimize the genes to enable the most potent replication *in vitro*. Generation of a library of HCV viruses with shuffled polymerase genes improves the chance to identify the viruses that efficiently grow *in vitro*. Another useful target for DNA shuffling is the internal ribosome entry site of HCV, which is required for viral translation (Honda *et al.*, *J. Virol.* (1999) 73: 1165-74). The region has been shown to be relatively conserved in the different flaviviruses, indicating that family shuffling is a particularly attractive approach (*Id.*).

The libraries of shuffled HCV, generated by shuffling the entire genomes, or fragments thereof, are generally screened for growth *in vitro* and *in vivo*. Family DNA shuffling allows the simultaneous generation of chimeric viruses and viral proteins. This is expected to improve the crossprotection against the different strains of HCV. One important

approach to attenuation is the selection of variants that grow in tissue culture cells as described above. A particularly significant selection is that which occurs *in vivo*. Virus clones isolated during the acute phase from patients and chimpanzees had identical sequences in the hypervariable region, and the infectious clones derived from patients were infectious in chimpanzees (Aizaki *et al.*, *Hepatology* (1998) 27:621-7). Therefore, the chimpanzee model will be useful when analyzing the level of attenuation and immune response in animals. The degree of attenuation of the selected HCV variants can thus be readily analyzed in chimpanzees. The animals are subsequently challenged with various wild-type HCV isolates to detect the most protective and crossprotective vaccine strains. The best variants can be selected for new rounds of shuffling and screening, when desired.

## Example 7

### Porcine reproductive and respiratory syndrome virus

Porcine reproductive and respiratory syndrome virus (PRRSV) is a recently identified virus that continues to challenge swine producers, veterinary practitioners, and animal health researchers. The prevalence of infection is high, approximately 60% to 80% of herds, and the clinical effects of infection vary widely among the farms (Zimmerman *et al.*, *Vet. Microbiol.* (1997) 55: 187-96). In many herds, infection is subclinical and productivity seemingly unaffected. However, some infected herds report occasional respiratory disease outbreaks in young pigs, periodic outbreaks of reproductive disease, or severe, chronic disease problems, particularly in young pigs. In these herds, secondary infections with viral or bacterial pathogens, such as *Salmonella choleraesuis*, *Streptococcus suis*, or *Haemophilus parasuis* typically occur concurrently with PRRSV infections (Zimmerman *et al.*, *Vet. Microbiol.* (1997) 55: 187-96).

There is evidence that existing candidate vaccine strains may persist and mutate to a less attenuated form *in vivo* (Mengeling *et al.*, *Am. J. Vet. Res.* (1999) 60: 334-4). Moreover, different isolates of PRRSV have been shown to differ substantially in their antigenic properties increasing the challenges in vaccine development (Pirzadeh *et al.*, *Can. J. Vet. Res.* (1998) 62: 170-7). Protection against heterologous strains is poor (van Woensel *et al.*, *Vet. Rec.* (1998) 142: 510-2), and need for improved vaccines is evident.

DNA shuffling is used to generate improved, attenuated vaccine strains of PRRSV. Infectious transcripts from cloned genome-length cDNA of PRRSV have been generated (Meulenberg *et al., J. Virol.* (1998) 72: 380-7). Thus, family DNA shuffling of such infectious cDNA clones derived from various isolates provides a means to generate chimeras that simultaneously protect against the different serotypes. Furthermore, the genetic heterogeneity of the different isolates provides an excellent starting pool to generate optimally attenuated strains that provide efficient protection without clinical disease. As an example, porcine alveolar lung macrophages or CL2621 cells can be used to grow the viruses to address the level of attenuation (Meulenberg *et al., J. Virol.* (1998) 72: 380-7). Furthermore, BHK-21 cells are useful in the initial screening, because they are readily transfectable by the cDNA clones (Meulenberg *et al., J. Virol.* (1998) 72: 380-7). The subsequent analysis of the attenuated strains can be done in pigs or other suitable test animal. For example, pigs are immunized with the attenuated strains, and the animals that do not become ill are challenged with the wild-type viruses to assess the efficacy of the vaccines. The attenuated viruses that provided the most efficient protection can be subjected to new rounds of shuffling and selection, when desired.

## Example 8

### Human immunodeficiency virus (HIV)

The development of a safe and effective vaccine for the prevention of HIV infections has proven to be extremely difficult, at least in part because of the complexity associated with HIV-1 and its pathogenesis (Hulskotte *et al., Vaccine* (1998) 16: 904-15). Previous studies suggest that that HIV infections in human may be abrogated by the host immune system supporting the conclusion that it is possible to generate a weakened virus that induces a protective immune response, but does not cause a disease. In addition, some individuals survive the infection for more than 15 years, further suggesting that the immune response can control HIV-1 infection at least in some individuals (*Id.*). Live, attenuated viruses have been the most successful vaccines in monkey models of HIV-1 infection (Berkhout *et al., J. Virol.* (1999) 73: 1138-45).

A number of infectious molecular clones from various HIV isolates have been constructed (Srinivasan *et al., Gene* (1987); 52: 71-82; Sauermann *et al., AIDS Res. Hum.*

*Retroviruses* (1990) 6: 813-23; Collman *et al.*, *J. Virol.* (1992) 66: 7517-21), thus providing suitable substrates for using family DNA shuffling to achieve evolution of attenuated vaccine strains of HIV. Libraries of HIV viruses are generated by family shuffling the entire molecular clones or fragments thereof. Useful HIV regions for evolution by DNA shuffling are the genes encoding the Gag and Env structural proteins MA (matrix), CA (capsid), NC (nucleocapsid), p6, SU (surface), and TM (transmembrane); the Pol enzymes PR (protease), RT (reverse transcriptase), and IN (integrase); the gene regulatory proteins Tat and Rev; and the accessory proteins Nef, Vif, Vpr, and Vpu (Frankel and Young, *Annu. Rev. Biochem.* (1998) 67: 1-25; Turner and Summers, *J. Mol. Biol.* (1999) 285: 1-32). Nucleic acids that encode any of these proteins can be targeted by DNA shuffling. In addition, combination libraries can be generated by combining shuffled libraries of different viral genes.

The shuffled libraries are screened for attenuation using *in vitro* and *in vivo* methods. *In vitro* methods include, but are not limited to, selection of strains with altered host cell specificity, selection of temperature sensitive mutants and selection of variants that can enter the human host cells but do not replicate. *In vivo* methods include selection of variants that grow in animals that cannot be infected with the wild-type virus. The chimpanzee model provides an excellent *in vivo* system to address the degree of attenuation of the selected variants (Murthy *et al.*, *AIDS Res. Hum. Retroviruses* (1998) Suppl 3:S271-6). Because chimpanzees can also be challenged with the live viruses resulting in AIDS, the model simultaneously provides an opportunity to address the efficacy of the vaccines. The chimpanzees that are immunized with the attenuated, shuffled vaccine strains of HIV can be challenged with various wild-type strains to analyze the level of protection and crossprotection. The clones demonstrating optimal level of attenuation with efficient protection against subsequent challenge can be chosen for additional rounds of shuffling and selection, if so desired.

## Example 9

## Evolution of Multivalent HPV Vaccines

### Background

Mucosal/genital HPV is one of the most common sexually transmitted diseases worldwide, with approximately 5.5 million new cases diagnosed per year in the US

alone. Estimates of up to 40 million infected individuals have been reported in the United States alone (Cancer Weekly Plus, June 29, 1998). Approximately 20 different HPV types infect the mucosal area, the majority of which are classified as "high risk" HPV types because of their association with over 90 % of cervical carcinomas (Bosch *et al.*, *J. Natl. Cancer Inst.* 87:796-802 (1995)) and other ano-genital and oral malignancies. Human papillomavirus infections are the primary cause of cervical cancers (Bosch et al, J. Natl. Cancer Inst. 87:796-802, 1995) and have been also linked to anal, vaginal, vulvar, oral and cutaneous cancers. In the United States alone, 15,000 women a year are diagnosed with cervical cancer, resulting in 5,000 annual deaths (U.S. CDC). A summary of the association of different HPV types with human diseases is outlined in Table 2 (source: M. Stanley, *Antiviral Research* 24:1-15, 1994).. It has been estimated that up to 50% of sexually active women are infected with these so-called 'high risk' HPVs. There are no curative treatments to date.

Table 2

| DISEASE | HPV TYPE |
|---|---|
| Skin | |
| Common warts, hands and feet etc. | 1. 2. 4. (26), (27), 29. 57 |
| Plane warts | 3. 10, 28, (49) |
| Butchers' warts | 7 |
| Epidermodysplasia verruciformis, benign | 5, 8, 9, 12, 14, 15, 17, 19, 20, 21, 22, 23, 24,25,36,46,47,50 |
| Epidermodysplasia verruciformis, SCC | 5, 8, 14, 17, 20 |
| Keiatocanthoma | 37 |
| Malignant melanoma | 38 |
| Actinic keratosis | 5,8 |
| SCC - | 41 , (48) |
| Epidermoid cyst | 60 |
| **Genitalia and mucous membranes** | |
| Normal cervix | 16,53 |
| Genital warts | 6,11,44,54 |
| Buschke Lowenstein tumours | 6,11 |
| Cervical intraepithelial neoplasia | 6,11, 16, 18, 30, 31, 33, 34, 35, 39, 40, 42, (43), (44), 45, 51, 52, 56, 57, 58, 66* |
| Cervical carcinoma | 16, 18, 31, 33, 35, 39, 45, 51, 52, 56,.66* |
| Vulvar intraepithelial neoplasia | 16,18,43,59 |
| Penile intraepithelial neoplasia | 16,18,39,40 |
| Bowen's disease | 34 |
| Bowenoid papulosis | 16,39,55 |
| Laryngeal papillomas | 6,11 |
| Laryngeal carcinoma | 30 |
| Focal epithelial hyperplasia (Heck's) | 13,32 |
| Oral papillomas | 32 |

Papillomavirus are small, non-enveloped DNA virus with a circular genome of 7.8 kb in size. Expression of the viral early and late proteins as well as episomal replication are regulated by a complex interaction of viral and host transcription and replication factors (Bernard and Apt, *Arch. Dermatol.* 130:210 (1994)). The viral early proteins E6 and E7 of "high risk" HPVs are powerful oncogenes and are constitutively expressed in cancer cells. They can thus serve as tumor antigens for therapeutic vaccine development. Vaccination of mice with recombinant E6 and E7 proteins has been shown to elicit a CTL mediated protective immune response against challenge with E6/E7 expressing tumors (Hariharan *et al, Int. J. Oncol.*, 12:1229-35 (1998)).

Two structural proteins, L1 and L2, form papillomavirus viruses, which are expressed in the late viral life cycle. A virus is composed of 72 capsomers, each of which is

formed by 5 L1 molecules. The ratio of the major capsid protein L1 to the minor capsid protein L2 is estimated as 30:1. The inability of HPVs to productively grow in cell culture has severely hampered attempts to generate sufficient amounts of viruses for experimental vaccination. The discovery that papillomavirus L1 proteins have the intrinsic capacity to self-assemble into viral like-particles (VLPs) when expressed in the absence of other viral gene products and epithelial differentiation (Kirnbauer *et al., Proc. Nat'l. Acad. Sci. USA* 89: 12180-84 (1992)) was the technological breakthrough that has driven the recent flurry of prophylactic vaccine development. Papillomavirus VLPs were used in animal models as effective virus surrogates to induce protective immunity. They provide essential conformational epitopes without being infectious. There is no experimental system for testing prophylactic anti-HPV vaccines due to the species specificity of all papillomaviruses. However, experimental animal vaccinations have resulted in antibody-mediated protection of domestic rabbits, cows and dogs following infections with cotton-tail rabbit papillomavirus, bovine papillomavirus and canine oral papillomavirus, respectively (Lowy and Schiller, *Biochim. Biophys. Acta*, 1423: M1-8, 1998, and references herein). Post-attachment neutralization of PVs by antibodies could also be demonstrated.

Effective immune surveillance of human papillomavirus may be difficult to achieve. Natural HPV infections at the genital mucosal surface are poorly immunogenic, presumably reflecting the non-lytic viral life cycle and the co-evolution of the viruses with their natural hosts. Individuals infected with HPVs are usually seropositive both for the late viral proteins and for the viral early oncogenes, but antibody titers are low. For epitheliotropic genital HPVs, which have no blood-borne phase in the viral cycle, the relevance of circulating antibodies will be restricted to their availability on the mucosal surface. The demonstration that systemic immunization with viral-like particles (VLPs) in African green monkeys induced neutralizing IgG in both sera and cervical-vaginal secretions makes it likely that VLP immunization will give similar results in humans (Lowe *et al, J. Infect. Dis.* 176: 1141-1145 (1997)). However, any enhancement of the elicited immune response will result in greater clinical benefit.

HPV pseudotype assays have shown that VLP based prophylactic vaccines are type specific and do not mediate cross-protection (Roden *et al., J. Virol.*, 70: 5875-3383 (1996)). Efficient prophylactic vaccines should therefore include antigenic determinants

from several HPV types. Recent experimental studies point to the minor capsid protein L2 as an alternative target for vaccine development. Immunization of mice with L2 proteins led to the induction of neutralizing antibodies with some degree of cross-neutralization. The antibody titers were, however, very low compared to the titers induced by VLP immunization.

Immunotherapy may offer a novel and more effective means for both prevention and treatment of HPV infection and associated diseases. The majority of current efforts in vaccine development are directed against two of the more prevalent "high risk" types, HPV-16 and HPV-18, which are most commonly found in malignant cancers. However, *in vitro* neutralization assays have demonstrated that the immune surveillance of HPV is type specific (Roden *et al*, *J. Virol.* 70:5875-5883 (1996)). Eradication of the two major HPV types, HPV-16 and HPV-18 could, therefore, drive the evolution of the large number of serological distinct genotypes. Additionally, different HPV-type specific variants, co-evolved with human races, need to be considered in effective prophylactic vaccine development. There is epidemiological evidence that minor variations within HPV-types may be more strongly associated with the risk of developing cancer (Xi *et al*, *J. Nat'l. Cancer Inst.* 89: 796-802 (1997)). Broadly protective HPV vaccines therefore must be multivalent.

Rational design of cross protective, multivalent VLP vaccines is extremely difficult given the lack of knowledge concerning the structure, localization and sequences involved in the antigenic epitopes exposed on the virus surfaces. Furthermore, direct approaches to study HPV immune surveillance are not possible due to the lack of animal models. Papillomavirus evolved with their hosts and are strictly species-specific.

**Evolution of Polyvalent HPV Vaccines**

In this Example, the challenges that have hampered development of HPV vaccines are addressed by the use of molecular breeding by recursive rounds of DNA family shuffling and screening. Since shuffling does not require an understanding of the mode or mechanism of infection, but simply relies on a functional screen for desired improvements, it is the tool most likely to quickly yield a product of clinical and commercial relevance. Recombination of antigenic sequences from related "high risk" HPVs are used to generate

large pools of functionally diverse chimeric sequences from which the best are selected based on improved immunogenic and cross-reactive properties. For example, one can generate potent multivalent VLP vaccines by shuffling nucleic acids that encode antigen epitopes from different L1 proteins and/or L2 genes to improve cross-neutralizing epitopes and antibody titers.

Nucleic acids that encode the antigens are used to generate complex, high-quality antigen libraries that are screened with high throughput (HTP) screening assays *in vitro* and *in vivo* for the selection of superior cross protective antigens against the major "high risk" HPV types. An example of a suitable strategy is summarized in Figure 3. The naturally existing diversity of HPV virus antigens is combined to generate complex antigen libraries by DNA shuffling. HTP *in vitro* assay systems are then used for the production of VLPs and subsequent screening to enrich the libraries for antigenic epitope display. An *in vivo* antigen library screen with subsequent neutralization assays then allows one to select for broad-spectrum VLP vaccines.

### *A. Generation of chimeric L1 antigen libraries*

First, different L1 genes are isolated from "high risk" HPVs and associated variants to generate complex libraries. Papillomaviruses are a large family of related viruses with specific tropism for different epithelia. Based on sequence alignment of the viral genomes, papillomaviruses can be divided into several distinct groups. A phylogenetic tree, computed for 108 different papillomavirus L1 genes, contains three supergroups (mucosal/genital, cutaneous/EV, and certain animal PVs) and 24 subgroups (Figure 4). The phylogenetic relationship is reflected by a similar tissue tropism (cutaneous, mucosal/oral, ano-genital) of the virus and the pathogenic lesions they induce (benign or malignant tumors). Eight "high risk" HPV types (HPV-18, 39, 45 and HPV-16, 31, 33, 35, 52) are found in the majority of malignant cancers and cluster in two distinct subgroups of the phylogenetic tree (A7 and A9, Figure 4).

The phylogenetic distance between the two subgroups is, however, greater than desired for successful DNA family shuffling. Therefore, two different libraries are generated, one for each subgroup. HPV-16 and HPV-18, which are associated with 80 % of

the HPV-related cancers, are used as major templates and different amounts of sequences from related types are added to the shuffling reactions.

HPV-16 L1 genes are pooled with the closely related variants HPV-31, 33, 35, 52 as well as different HPV-16 variants, and HPV-18 genes are pooled with HPV-45 and HPV-39. Pools of related L1 genes are subjected to random fragmentation and subsequent reassembly in a primerless PCR reaction according to established DNA family shuffling protocols as described herein (*see also*, Crameri *et al.*, *Nature* 391: 288 (1998)). Additional sequence heterogeneity can be added by spiking homologous sequences from more distant "high risk" HPV types (*e.g.*, HPV-51, 56 and 66, subgroup A5 and A6, Figure 4) into the assembly reaction, in the form of short oligonucleotides with homologous ends. Reassembled L1 chimeras are amplified by PCR with primers flanking the L1 genes and subcloned into shuttle vectors, which allow for high throughput DNA amplification in *E. coli* and protein expression in mammalian cells. The complexity of the libraries is estimated by restriction analyses and sequencing of randomly selected clones. The quantitative goal is to gain large libraries ($>10^5$), with 90 – 100 % chimeric sequences. The same experimental strategy can be applied to the L2 genes.

### *B. Development of high throughput* in vitro *screening assays*

The next step is to establish HTP assay systems for selection of L1 and L2 chimeras displaying antigenic epitopes. The ultimate goal is to select the best L1 and L2 chimeras for their ability to induce broadly reactive antibodies *in vivo*. DNA family shuffling, however, can generate "wrongly" assembled genes, which give rise to abortive protein expression. Therefore, pre-screening the libraries *in vitro* for L1 and L2 protein expression and for ability to display immunogenic epitopes will help to avoid unnecessary animal studies during subsequent *in vivo* screening.

A general strategy of a high throughput in vitro assay is outlined in Figure 5. Chimeric L1 and L2 genes are transfected into mammalian cells for LI, L2 protein expression and VLP assembly. Cell lysates are prepared from random clones to check the library quality by immunoblotting or simple plate ELISA assays. If the libraries show a higher than desired "knock out" rate, one can apply milder shuffling conditions by using

lower percentages of L1 genes which are most distant from the main types, HPV-16 and HPV-18.

The *in vitro* screening assays generally involve the following components:

(i) For expression in mammalian cells and amplification in *E. coli*, the shuffled L1 and L2 genes are linked to a strong eukaryotic promoter and cloned into plasmid vectors containing a bacterial origin of replication and a drug selection marker. For L1, it is particularly preferred to use a strong promoter, *e.g.*, a CMV promoter or a promoter that has been improved using DNA shuffling. High expression of L1 is desirable to obtain expression levels that are sufficiently high for efficient VLP assembly.

(ii) Different mammalian cell lines are tested with different transfection agents to optimize transfection efficiencies. An efficiency of 80–90 % is desirable. Human 293 cells are often suitable. Transfection efficiency levels and promoter strength can be tested using a control plasmid (*e.g.*, one that expresses GFP), which allows rapid fluorescence read outs. Random clones of the library are transfected into the selected cell line. ELISA or Western blotting is used to examine the cell lysates for L1 and L2 protein expression.

(iii) Preferred screening assays are based on immune recognition, for which specific antibodies are needed. For some embodiments, it is sufficient for the *in vitro* screening assays to use antibodies against HPV-16 and HPV-18 L1 and L2 proteins, since they are most prevalent in malignant cancers and selected chimeras should have the property to induce high levels of antibodies against these two types. Cross protection of the antigens against the related HPV types is tested in the final *in vivo* screening assays.

For the generation of polyclonal antisera against L1/L2 proteins and conformational L1 VLP epitopes, the L1 and L2 genes of HPV-16 and HPV-18 are cloned into bacterial expression vectors (*e.g.*, pET-3a) from which proteins can be expressed in quantitative amounts in appropriate bacterial strains (BL21/DE3, HMS174/DE3). It has been shown that HPV-L1 proteins can reassemble into VLPs during the subsequent protein purification steps (Cripe *et al.*, *J. Virol.* 71: 2988-2995 (1997)). Purified L1 and L2 proteins and sucrose gradient purified L1 VLPs can be used to inject rabbits or mice for induction antibodies against conformational antigen epitopes.

*C. HTP* in vitro *screening assay*

To select and enrich for chimeric clones that give rise to L1/L2 proteins that expose conformational antigenic epitopes, the libraries are subjected to subsequent rounds of screening *in vitro*. A schematic overview of the *in vitro* screening assay is illustrated in Figure 5. The libraries are transfected into mammalian cells for protein expression and VLP formation. For subsequent screening, L1/L2 proteins and VLPs are preferably purified from the crude cell lysates. To achieve high throughput purification for the L1 VLPs, one can use expression vectors that direct the expression of the chimeric L1/L2 proteins as fusions with a heterologous antigenic epitope (*e.g.*, a hexahistidine tag). The presence of such heterologous amino acid sequences does not hamper the self-assembly of L1 proteins into VLPs and the display of conformational antigenic epitopes (Peng *et al.*, *Virology* 240:1800-1805 (1998)). Fusion of a hexahistidine tag to the C-terminus of the protein chimeras provides an efficient and fast way of protein purification in HTP plate assays. The hexahistidine tag is uncharged at physiological pH and rarely interferes with protein structure and function. The His tag can be linked to the C-terminal part of the shuffled chimeras by simply adding a short sequence coding for 6 histidine residues to the 3'-PCR primers used for the final amplification of the shuffled products.

Following amplification of the shuffled libraries in *E. coli*, plasmid DNA from individual clones can be robotically prepared in a high throughput 96 well format. Robotic plasmid purification protocols that allow purification of 600-800 plasmids per day or more are feasible. The quantity and purity of the DNA can also be analyzed on the plates. The libraries are transfected into mammalian cells seeded in 96 well formats, allowing for up to 1000 individual transfections at a time. Crude lysates are prepared after culturing the cells for two days. The lysates are transferred to new 96 well plates coated with nickel-nitrilotriacetic acid (Ni-NTA HisSorb plates, Qiagen), to efficiently immobilize the 6xHis tagged L1 / L2 proteins and VLPs on the plates. The plates are incubated with the anti-HPV-16 and HPV-18 L1/L2 antibodies and the detection conjugate, and analyzed by automated plate read out. The HPV-16 and 18 wild type genes will serve as positive controls for the assay. HPV-16 and HPV-18 do not display cross-reactive epitopes and can be used as background control. The quantitative goal of the *in vitro* assay is to select 1000 or more chimeras from each library for subsequent immune stimulation in mice.

*D. Alternative assays for VLP production and screening:*

An alternative HTP assay can be set up by making use of the ability of VLPs to package plasmid DNA of up to 10 kb in size (Stauffer *et al.*, *J. Mol. Biol.* 263: 529-536 (1998)). A marker gene expressing a photon emitting protein (*e.g.*, GFP, LacZ, luciferase) is cloned into the L1 expression plasmids, which are packaged during the VLP assembly. After VLP production in one cell line, all cells from a 96 well plate are pooled, and VLPs can be purified in a single reaction. Subsequent incubation of the VLPs with mammalian cells leads to marker transfer for L1 chimeras, which have the capacity to assemble into infectious VLPs. Cells expressing the marker gene can be easily monitored by fluorescence microscopy and plate ELISA, and selected by FACS sorting. The plasmid DNA can be purified by Hirt preparation, followed by amplification in *E. coli*. This direct screening assay is more stringent, but selection is only for VLPs which are able to package DNA, so variants that have lost the ability to package DNA but express strong immunogenic epitopes might be missed.

Other alternative transfection protocols and low-throughput (LTP) chromatographic VLP purification steps (affinity chromatography, capillary electrophoresis or sucrose centrifugation) can also be used.

*E. Analyses of the shuffled library* in vivo

Pre-selected chimeric library clones from *in vitro* assays are used to immunize mice. Two different routes of application can be envisioned: (1) Injection of L2 proteins and purified L1 VLPs, which has been successfully used in other experimental studies, and (2) naked DNA delivery, which offers the advantage of easy commercial scale vaccine manufacturing and non-invasive dermal application, specifically for future clinical applications. Naked DNA vaccinations have resulted in sufficient conformational L1 epitope delivery and immune protection in an experimental rabbit model (Sundaram *et al.*, *Vaccine* 15: 664-671 (1997)).

One can initially conduct the *in vivo* screening experiments using dermal naked DNA application. If this assay is not sufficiently sensitive, one can express and purify VLPs. L1 proteins can be quantitatively expressed in, for example, *E. coli* (or yeast), assembled into VLPs *in vitro* and purified by sucrose/CsCl gradients for injection. Using the naked DNA delivery approach in the experimental system has the additional advantage of

enabling one to select concomitantly for L1 chimeras with the highest ability to assemble into VLPs *in vivo*.

Pooling plasmids or VLPs and deconvoluting in subsequent screening rounds can reduce the number of small animals required to identify potent immunogens. The lowest concentration of plasmids or VLPs, which leads to induction of neutralizing antibodies in mice, are evaluated with the HPV-16 wild-type L1 and L2 plasmids or proteins. Pools of 10 or 20 clones of the library can be used in a small number of experiments to examine whether the strategy of pooling and deconvolution in subsequent screening rounds is feasible. If no significant differences can be detected between different pools and the wild type control, single clones can be used for library immunization.

### *F. Neutralization assay and analyses of cross protective immunity*

Sera from immunized mice is collected and tested for cross-neutralization efficiency of the wild type VLPs in 96 well plate assays. Papillomavirus L1/VLPs retain the ability of natural viruses to agglutinate mouse erythrocytes in culture and antibodies raised against VLPs can inhibit agglutination (Roden *et al.*, *J. Virol.* 70: 3298-3201 (1996)). Hemagglutination inhibition assays (HIA) provide therefore reliable and sensitive surrogate neutralization assays. VLPs from all wild-type high risk HPVs can be prepared by expression in *E. coli* and seeded together with the mouse erythrocytes in 96 well plates. Collected sera will be added in serial dilutions to evaluate the neutralization titers and cross-neutralization ability. L1 and L2 chimeras from pooling experiments, which induced antibodies with improved neutralization titers compared to anti-wild type antigens and are cross reactive with other related wild type VLPs, can then be deconvoluted in the next round of *in vivo* screening. The improved characteristics of selected clones can be confirmed in direct neutralization assays using marker gene transfer as described above.

To further improve the quality of the novel antigens, a second round of shuffling and screening is preferably applied to obtain the best variants. Improved variants--defined as those inducing potent cross-reactive immunity against a broad range of related HPV--are reshuffled and screened. Shuffled chimeras can be backcrossed with the wild type genes to further improve the antibody titers. Backcrossing is performed by shuffling the

improved sequence with a large molar excess of the parental sequence and provides a means to breed the shuffled chimeras/mutants back to a parental or wild-type sequence, while retaining the mutations that are critical to the phenotype that provides cross-protective antibody responses. In addition to removing the neutral mutations, molecular backcrossing can also be used to characterize which of the many mutations in an improved variant contribute most to the phenotype. This cannot be accomplished in an efficient library fashion by any other method.

It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference for all purposes.

WHAT IS CLAIMED IS:

1. A method of obtaining an attenuated vaccine, the method comprising:

(1) recombining a first set of one or more nucleic acid segments that comprise a complete or partial genomic library of a virus or a cell with at least a second set of one or more nucleic acid segments to form a library of recombinant nucleic acid fragments;

(2) screening viruses or cells that contain members of the library of recombinant nucleic acid fragments to identify those viruses or cells that are attenuated under physiological conditions that exist in a host organism; and

(3) screening the attenuated viruses or cells to identify those that can induce an immune response against a pathogenic agent that displays an immunogenic determinant that is also displayed by the attenuated viruses or cells;

wherein the attenuated viruses or cells that can induce the immune response are useful as attenuated vaccines against the pathogenic agent.

2. The method of claim 1, wherein the screening of (2) is performed simultaneously with or after the screening of (3).

3. The method of claim 1, wherein the attenuated vaccine is a cell and the library of recombinant DNA fragments is created by introducing the nucleic acids of the second set into a plurality of cells, whereby at least one of the nucleic acids undergoes recombination with a nucleic acid segment in the genome or an episome of the cells.

4. The method of claim 1, wherein the attenuated vaccine is a bacterium, fungus or parasite.

5. The method of claim 1, wherein the attenuated vaccine is a virus or a virus-like particle.

6. The method of claim 5, wherein the attenuated vaccine is a virus which is selected from the group consisting of influenza virus, human immunodeficiency virus,

hepatitis A, B, C, D and E viruses, rotavirus, parvovirus B19, herpes simplex virus 1, herpes simplex virus 2, cytomegalovirus, varicella zoster virus, Epstein-Barr virus, encephalitis virus, respiratory syncytial virus, feline calicivirus, infectious bursal disease, virus, dengue virus type 2, swine fever virus, coxcackievirus B3, equine arteritis virus, yellow fever virus, human astrovirus, and porcine reproductive and respiratory syndrome virus.

7. The method of claim 5, wherein the attenuated vaccine is a virus which is a member of a family selected from the group consisting of picornavirus, togavirus, coronavirus, rhabdovirus, paramyxovirus, bungavirus, arenavirus, retrovirus, papovirus, parvovirus, herpesvirus, poxvirus, and hepadnavirus.

8. The method of claim 5, wherein the attenuated vaccine is a virus-like particle comprises one or more polypeptides of a virus selected from the group consisting of human papillomavirus, human immunodeficiency virus, Semliki-Forest virus, human polyomavirus, rotavirus, parvovirus, and hepatitis E virus.

9. The method of claim 1, wherein the method further comprises:

(4) recombining DNA from the attenuated cells or viruses with a further set of one or more forms of a nucleic acid, to form a further library of recombinant nucleic acids;

(5) screening viruses or cells that contain members of the further library of recombinant nucleic acid fragments to identify those viruses or cells that are further attenuated under physiological conditions that exist in a host organism; and

(6) repeating (4) and (5) as required until the further attenuated viruses or cells have lost the ability to replicate or cause disease under physiological conditions that exist in the host organism.

10. The method of claim 9, wherein the further set of nucleic acid segments comprises nucleic acids from the attenuated cells or viruses identified in (3).

11. The method of claim 1, wherein the nucleic acid segments of the second set are derived from a pathogenic agent.

12. The method of claim 1, wherein the second set of nucleic acid segments comprises a substantially complete genomic library from at least one heterologous cell or virus type.

13. The method of claim 1, wherein the second set of nucleic acid segments comprises a partial genomic library from at least one heterologous cell or virus type.

14. The method of claim 1, wherein the second set of nucleic acid segments encode an immunogenic determinant from a heterologous organism.

15. The method of claim 1, wherein the nucleic acid segments of either or both of the first or second sets comprise natural variants of a gene from different individual cells or viruses.

16. The method of claim 1, wherein the attenuated vaccines are unable to replicate or cause a disease or other adverse effect when present in an inoculated host organism.

17. The method of claim 16, wherein a naturally occurring isolate of the attenuated viruses or cells is capable of replicating or causing a disease or other adverse effect when present in a host organism infected by the naturally occurring isolate.

18. The method of claim 1, wherein the method further comprises:

backcrossing nucleic acids of the attenuated viruses or cells by recombining the nucleic acids with a library of nucleic acids from a wild-type strain of the virus or cell to form a further library of recombinant nucleic acids; and

screening viruses or cells that contain members of the further library of recombinant nucleic acids to identify backcrossed viruses or cells that are attenuated under physiological conditions present in an inoculated host organism.

19. The method of claim 18, wherein the library of nucleic acids from the wild-type strain comprises a complete or partial genomic library of the virus or cell.

20. The method of claim 18, wherein the method further comprises screening the backcrossed attenuated viruses or cells to identify those that can induce an immune response against a pathogenic agent that displays an immunogenic determinant that is also displayed by the attenuated viruses or cells.

21. The method of claim 1, wherein (2) comprises screening viruses or cells that contain members of the library of recombinant nucleic acid fragments to identify those viruses or cells that exhibit reduced ability to bind to a host cell or host tissue.

22. The method of claim 21, wherein the viruses are screened by panning using purified or recombinantly produced receptors of the host cell.

23. The method of claim 1, wherein (2) comprises screening viruses or cells that contain members of the library of recombinant nucleic acid fragments to identify those viruses or cells that exhibit reduced ability to bind to a complement component.

24. The method of claim 1, wherein (2) comprises screening viruses or cells that contain members of the library of recombinant nucleic acid fragments to identify those viruses or cells that exhibit enhanced susceptibility to a humoral or a cell mediated immune response.

25. The method of claim 24, wherein the viruses or cells are screened by contacting the viruses or cells with antibodies that specifically bind to a naturally occurring isolate of the viruses or cells.

26. The method of claim 25, wherein the method further comprises testing the viruses or cells that bind to the antibodies to identify those that are inactivated by the antibodies.

27. The method of claim 1, wherein (2) comprises screening viruses or cells that contain members of the library of recombinant nucleic acid fragments to identify those viruses or cells that do not bind to maternal antibodies.

28. The method of claim 1, wherein the method further comprises screening the attenuated viruses or cells to identify those that propagate under permissive conditions used for production of the attenuated viruses or cells.

29. The method of claim 28, wherein the permissive conditions differ from the physiological conditions in a property selected from the group consisting of temperature, pH, sugar content, a compromised immune system, absence of complement or complement components, and presence or absence of serum proteins.

30. The method of claim 29, wherein the permissive conditions comprise an immunocompromised host organism.

31. The method of claim 28, wherein the permissive conditions comprise a medium that contains a nutrient which is absent under the physiological conditions.

32. The method of claim 28, wherein the method comprises identifying cells or viruses that propagate rapidly in producer cells or culture conditions prior to (1).

33. The method of claim 28, wherein:

the permissive conditions comprise the presence of a suppressor tRNA molecule;

the sets of nucleic acid segments comprise:

a) a first set of one or more polynucleotides that encode all or part of a polypeptide that is involved in replication of the viruses or cells; and

b) a second set of one or more oligonucleotides that comprise one or more stop codons interspersed within a polynucleotide sequence that encodes the polypeptide;

wherein the oligonucleotides undergo recombination with the polypeptide-encoding polynucleotides to form a library of recombinant nucleic acids that comprise one or more recombinant nucleic acids in which at least one nonnaturally occurring stop codon is interspersed within the coding sequence of the polypeptide; and

the attenuated viruses or cells are obtained by contacting the library of recombinant nucleic acid fragments with suppressor tRNA molecules that suppress the termination of translation at the nonnaturally occurring stop codons and collecting progeny viruses or cells that propagate in the presence of the suppressor tRNA molecules but not in the absence of the suppressor tRNA molecules.

34. The method of claim 33, wherein the suppressor tRNA molecules are present in a suppressor cell that is used for production of the attenuated viruses or cells.

35. The method of claim 33, wherein the first set comprises a partial or complete genomic library of the viruses or cells.

36. The method of claim 33, wherein the suppressor tRNA molecules are absent in cells of a mammal that is to be inoculated with the attenuated viruses or cells.

37. The method of claim 33, wherein the method further comprises:

(4) recombining one or more recombinant nucleic acid segments from the attenuated viruses or cells with a further population of oligonucleotides to produce a further library of recombinant nucleic acid segments;

(5) contacting the further library of recombinant nucleic acid segments with the suppressor tRNA molecules, and collecting progeny viruses or cells;

(6) contacting non-suppressor cells with the progeny viruses or cells to identify attenuated viruses or cells that are incapable of replicating in the non-suppressor cells; and

(7) repeating (4) to (6), as necessary, until a further progeny virus or cell has acquired a desired degree of ability to replicate in suppressor cells and inability to replicate in non-suppressor cells.

38. The method of claim 28, wherein the permissive conditions comprise a producer cell or organism in which a naturally occurring isolate of the virus or cell does not propagate, and the method comprises introducing viruses or cells that contain members of the library of recombinant nucleic acid fragments into the producer cell or organism to identify those viruses or cells that can propagate in the producer cell or organism.

39. The method of claim 38, wherein the producer cells are monkey cells and the attenuated viruses or cells propagate in the monkey cells but not in human cells.

40. The method of claim 38, wherein the viruses or cells are introduced into a mixed population of producer cells and cells of the host organism and the method further comprises performing a subsequent recombination and screening in a population of producer cells in the absence of host cells.

41. The method of claim 38, wherein the method further comprises performing a subsequent recombination and screening in a population of a second type of producer cells.

42. An attenuated virus or cell obtained by the method of claim 1.

43. A vaccine composition which comprises an attenuated virus or cell of claim 42, or a polynucleotide obtained from the attenuated virus or cell, and a carrier.

44. A method of vaccinating an animal, the method comprising administering to the animal a vaccine composition of claim 43.

45. The method of claim 44, wherein the vaccination is performed for a prophylactic purpose.

46. The method of claim 44, wherein the vaccination is performed for a therapeutic purpose.

47. A method of obtaining an attenuated vaccine, the method comprising:

(1) introducing a library of DNA fragments into a plurality of cells, whereby at least one of the fragments undergoes recombination with a segment in the genome or an episome of the cells to produce modified cells;

(2) screening the modified cells to identify conditionally defective cells that have evolved toward loss of the ability to proliferate under physiological conditions as found in a host organism; and

(3) screening the conditionally defective cells to identify those modified cells that have maintained the ability to replicate under the permissive conditions;

wherein conditionally defective cells that replicate under permissive conditions but not in a host mammal are suitable for use as an attenuated vaccine organism.

48. The method of claim 47, wherein the method further comprises:

(4) recombining DNA from the modified cells that have evolved toward inability to replicate under physiological conditions with a further library of DNA fragments, at least one of which undergoes recombination with a segment in the genome or the episome of the modified cells to produce further modified cells, or recombining DNA between the modified cells that have evolved toward the desired function to produce further modified cells;

(5) screening the further modified cells for further modified cells that have further evolved toward loss of ability to replicate under physiological conditions and have maintained the ability to replicate under permissive conditions;

(6) repeating (4) and (5) as required until the further modified cells have lost the ability to replicate under physiological conditions in a host mammal and have maintained the ability to replicate under permissive conditions.

49. A method of obtaining a chimeric attenuated vaccine, the method comprising:

(1) recombining a first set of one or more nucleic acid segments from a virus or cell with at least a second set of one or more nucleic acid segments, wherein the

nucleic acid segments of the second set confer upon viruses or cells that contain the nucleic acid segments a property that is desirable for vaccination, to form a library of recombinant DNA fragments;

(2) identifying attenuated viruses or cells by screening viruses or cells that contain members of the library of recombinant DNA fragments to identify those viruses or cells that are attenuated under physiological conditions present in a host organism inoculated with the viruses or cells; and

(3) screening the attenuated viruses or cells to identify those that exhibit an improvement in the property that is desirable for vaccination.

50. A chimeric attenuated vaccine that comprises an attenuated virus or cell obtained by the method of claim 49.

51. The method of claim 49, wherein the screening of (2) is performed simultaneously with or after the screening of (3).

52. The method of claim 49, wherein the property that is desirable for vaccination is enhanced stability of the attenuated chimeric vaccine *in vitro* and (3) comprises exposing viruses or cells that contain recombinant DNA fragments to desired storage conditions.

53. The method of claim 49, wherein the first set of nucleic acid segments comprises a full or a partial genomic library of a nonpathogenic cell or virus.

54. The method of claim 53, wherein the recombination is performed by introducing the second set of nucleic acid segments into a plurality of nonpathogenic cells, whereby at least one member of the second set of nucleic acid segments undergoes recombination with a segment in the genome or an episome of the nonpathogenic cells to produce modified cells; and the modified cells are screened to identify attenuated chimeric cells, which are nonpathogenic and exhibit an improvement in the property that is desirable for vaccination.

55. The method of claim 53, wherein the nonpathogenic cells are selected from the group consisting of *Lactococcus lactis*, *Mycobacterium bovis* (BCG), *Mycobacterium vaccae*, nonpathogenic *Salmonella* species, and nonpathogenic *Bacillus* species.

56. The method of claim 55, wherein the nonpathogenic cells are *Mycobacterium bovis* or *M. vaccae*, and the second set of nucleic acid segments comprises a full or partial genomic library of *M. tuberculosis*.

57. The method of claim 49, wherein the method further comprises:

(4) recombining DNA from the attenuated chimeric cells or viruses with a further set of nucleic acid segments to form a further library of recombinant nucleic acids;

(5) obtaining improved attenuated chimeric viruses or cells that exhibit further improvement in attenuation or in the property that is desirable for vaccination by screening viruses or cells that contain members of the further library of recombinant DNA fragments to identify those that exhibit further improved attenuation or desirable property; and

(6) repeating (4) and (5) as required until the attenuated chimeric viruses or cells have achieved a desired level of pathogenicity loss or improvement in the property that is desirable for vaccination.

58. The method of claim 57, wherein the property that is desirable for vaccination in (5) is different from the property that is desirable for vaccination in (3).

59. The method of claim 57, wherein the further set of nucleic acid segments comprises polynucleotides that encode one or more polypeptides that confer upon a viruses or cells that include the polypeptide an improvement in a property that is desirable for vaccination.

60. The method of claim 57, wherein the further set of nucleic acid segments comprises a full or partial genomic library of a wild-type strain of the cell or virus, and the improvement in a property that is desirable for vaccination is removal from the attenuated chimeric virus or cell one or more superfluous mutations.

61. The method of claim 49, wherein the second set of nucleic acid segments comprises polynucleotides that encode an immunogenic polypeptide, or a portion thereof, of a pathogenic cell or virus and the desired property is induction of a prophylactic or therapeutic immune response against the pathogenic cell or virus in an animal that has been inoculated with the attenuated chimeric cell or virus.

62. The method of claim 61, wherein the screening for ability to induce an immune response against the pathogenic cells or viruses is performed by testing for ability of the attenuated chimeric cells or viruses to induce protective immunity *in vivo*.

63. The method of claim 61, wherein the second set of nucleic acid segments comprises a substantially complete or partial genomic library of a pathogenic cell or virus.

64. The method of claim 61, wherein the pathogenic cells are selected from the group consisting of gram-positive cocci, gram-negative cocci, enteric gram-negative bacilli, anaerobic bacteria, pathogenic fungi, and parasites.

65. The method of claim 64, wherein the pathogenic cells are bacterial cells selected from the group consisting of pneumococci, staphylococci, streptococci, meningococci, gonococci, enterobacteriaceae, melioidosis, salmonella, shigellosis, hemophilus, chancroid, brucellosis, tularemia, yersinia, streptobacilli, listeria monocytogenes, and erysipelothrix rhusiopathiae, and causative agents of conditions selected from the group consisting of diptheria, cholera, anthrax, donovanosis, bartonellosis, tetanus, botulism, tuberculosis, leprosy, syphilis, treponematoses, and leptospirosis.

66. The method of claim 61, wherein the pathogenic cells are causative agents of a condition selected from the group of actinomycosis, nocardiosis, cryptococcosis, blastomycosis, histoplasmosis, coccidioidomycosis, candidiasis, aspergillosis, mucormycosis, sporotrichosis, paracoccidiodomycosis, petriellidiosis, torulopsosis, mycetoma, chromomycosis, dermatophytosis, rickettsial infections, mycoplasma and chlamydial infections, amebiasis, malaria, leishmaniasis, trypanosomiasis, toxoplasmosis, pneumocystis carinii, babesiosis, giardiasis, trichinosis, filariasis, schistosomiasis, nematodes, trematodes or flukes, and cestode (tapeworm) infections.

67. The method of claim 49, wherein the polypeptide is selected from the group consisting of an immunomodulatory molecule and a therapeutic protein.

68. The method of claim 49, wherein the property that is desirable for vaccination is improved expression of an immunogenic polypeptide, specific uptake of the attenuated vaccine, enhanced stability, and enhanced immunogenicity.

69. The method of claim 49, wherein either or both of the first and second sets of nucleic acid segments comprise polynucleotides that encode a viral polypeptide that can associate into a virus-like particle (VLP).

70. The method of claim 69, wherein the viral polypeptide encoded by the polynucleotides of the first set of nucleic acid segments is of a nonpathogenic viral strain and the viral polypeptide encoded by the polynucleotides of the second set of nucleic acid segments is of one or more pathogenic viral strains.

71. The method of claim 70, wherein the nonpathogenic viral strain and the pathogenic viral strain are of different species.

72. An attenuated virus or cell obtained by the method of claim 49.

73. A vaccine composition which comprises an attenuated virus or cell of claim 72, or a polynucleotide obtained from the attenuated virus or cell, and a carrier.

74. A method of vaccinating an animal, the method comprising administering to the animal a vaccine composition of claim 73.

75. The method of claim 74, wherein the vaccination is performed for a prophylactic purpose.

76. The method of claim 74, wherein the vaccination is performed for a therapeutic purpose.