DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
171

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
171

MAARJA ADOJAAN

Molecular variation of HIV-1 and the use of this knowledge in vaccine development

TARTU UNIVERSITY
PRESS

Department of Evolutionary Biology, Institute of Molecular and Cell Biology, Faculty of Science and Technology, University of Tartu, Estonia

Dissertation was accepted for the commencement of the degree of Doctor of Philosophy (in molecular biology) on November 10, 2009 by the Council of the Institute of Molecular and Cell Biology, University of Tartu.

Supervisors: Prof. Richard Villems, DrSci (The Department of Evolutionary Biology, Institute of Molecular and Cell Biology, University of Tartu)

Prof. Mart Ustav, PhD (Centre of Biomedical Technology, Institute of Technology, University of Tartu)

Opponent: Prof. Paul Pumpens, DrSci (Latvian Biomedical Research and Study Centre, Riga, Latvia)

Commencement: Room 218, 23 Riia Street, Tartu, on December 17th, 2009 at 12.00 a.m.

ISSN 1024–6479
ISBN 978–9949–19–272–4 (trükis)
ISBN 978–9949–19–273–1 (PDF)

Tartu Ülikooli Kirjastus
www.tyk.ee
Tellimus nr. 480

CONTENTS

LIST OF THE ORIGINAL PUBLICATIONS

This thesis is based on the following original articles, which are referred to in the text by their Roman numerals.

I. **Adojaan Maarja**, Kivisild Toomas, Männik Andres, Krispin Tõnu, Ustina Valentina, Zilmer Kai, Liebert Elo, Jaroslavtsev Nikolai, Priimägi Ludmilla, Tefanova Valentina, Schmidt Jelena, Krohn Kai, Villems Richard, Salminen Mika, Ustav Mart. Predominance of a Rare Type of HIV-1 in Estonia. (2005) *JAIDS Journal of Acquired Immune Deficiency Syndromes* 39: 598–605.

II. **Adojaan Maarja**, Mölder Tarmo, Männik Andres, Kivisild Toomas, Villems Richard, Krispin Tõnu, Ustav Mart. High Prevalence of the CCR5Δ32 HIV-Resistance Mutation among Estonian HIV Type 1-Infected Individuals. (2007) *AIDS Research and Human Retroviruses* 23:193–197.

III. Mölder Tarmo, **Adojaan Maarja**, Kaldma Katrin, Ustav Mart, Sikut Rein. Elicitation of broad CTL-response against HIV-1 by the DNA vaccine encoding artificial multi-component fusion protein MultiHIV – Study in domestic pigs. (In press), *Vaccine*.

Author's contribution:

I. performed the molecular genetic studies (subtyping, full-genome-sequencing of isolates EE0359 and EE0369), drafted the manuscript
II. performed the molecular genetic studies (determined CCR5Δ32 polymorphism frequency), drafted the manuscript
III. participated in the design and construction of the DNA vaccine, performed SLA-typing, participated in writing the manuscript

ABBREVIATIONS USED IN THE TEXT

APOBEC3G	apolipoprotein beta mRNA editing enzyme, catalytic polypeptide-like 3G
CD4, CD8	cluster of differentiation 4, and 8
CCR5	cysteine-cysteine linked chemokine receptor 5
CTL	cytotoxic T-lymphocyte
CRF	circulating recombinant form
HAART	highly active antiretroviral treatment
HIV-1, HIV-2	human immunodeficiency virus type 1, and 2
HLA-I, HLA-II	human leukocyte antigen class I, and II
IDU	intravenous drug user
IN	HIV-1 integrase
LTR	long terminal repeat
MHC	major histocompatibility complex
M-, O-, N-group	HIV-1 lineages Major, Outlier and non-M/non-O
NHP	non-human primate
PIC	pre-integration complex
PR	HIV-1 protease
RNase H	HIV-1 ribonuclease H
RT	HIV-1 reverse transcriptase
SIV	simian immunodeficiency virus
SLA	swine leukocyte antigen
URF	unique recombinant form

INTRODUCTION

Lentiviruses have been found in many African primate species, and the phylogenetic relationships between the viruses have been used to study the primate origins of human immunodeficiency viruses (HIVs), (Hahn et al., 2000). Evidence exists that simian immunodeficiency virus (SIV) has crossed the species barrier (simians to humans) several times over decades, leading to the various types, groups and possibly clades of HIV. Phylogenetic analysis reveals that closest simian relatives of HIV-1 are SIVcpz in chimpanzees (*Pan troglodytes troglodytes*) and SIVgor in western gorillas (*Gorilla gorilla*); the less pathogenic HIV-2 is most closely related to SIVsmm in sooty mangabeys (*Cercocebus atys*), (Heuverswyn &Peeters 2007). At least three independent transmission events early in the 20th century spawned three HIV-1 groups (named M, O, and N). Strains related to the HIV-1 M and N groups have been found in chimpanzees and although group O-related viruses have been identified in gorillas, chimpanzees are considered the original reservoir of also this SIV infection (Heuverswyn & Peeters 2007, Taylor et al., 2008, Tebit et al., 2007). Recent findings indicate that a new HIV-1 lineage might exist, deriving from gorillas. This novel lineage, represented by a single case currently, will be named group P if additional related isolates will be identified. The discovery of this new lineage highlights the continuing emergence of new HIV variants, particularly in Western Central Africa, the origin of all existing HIV-1 groups (Plantier et al., 2009).

Viral archaeology has enabled to shed light on the origin of the current AIDS pandemic and on the geography and timescale of the early diversification of HIV-1 in humans. It has been hypothesized that HIV-1 was spreading among humans for 60–80 years before AIDS was first recognized (Korber et al., 2000, Worobey et al., 2008). The earliest human samples found to contain HIV-1 were taken in 1959 and 1960 in Leopoldville (now Kinshasa, Democratic Republic of the Congo). The sequence analyses of 50-years-old samples have estimated the date of the common M group ancestor no later than 1933 (Worobey et al., 2008). If the epidemic grew roughly exponentially from only one or a few infected individuals around 1910 to the more than 50 million estimated to have been infected by 2007, there were probably only a few thousand HIV-infected individuals by 1960, all living in central Africa (Sharp & Hahn, 2008).

Today more than 20 million people have died from AIDS and roughly 33 million others are living with HIV. The overwhelming majority of HIV-1-positive individuals live in the developing world. The highest prevalence is seen in the sub-Saharan Africa where the epidemic has become a growing human and economic catastrophe, see Figure 1.

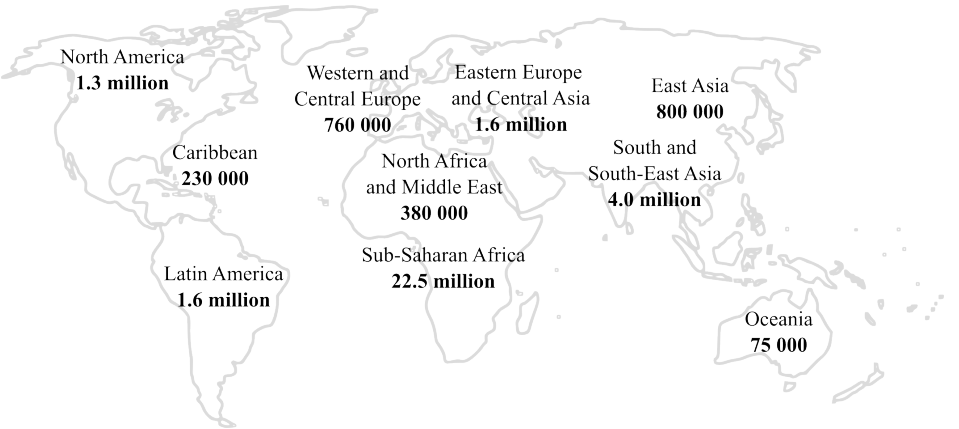

Figure 1. Adults and children estimated to be living with HIV in 2007 was 30.6–36.1 million. Data source: *AIDS epidemic update.* (2008) Joint United Nations Programme on HIV/AIDS (UNAIDS), http://www.unaids.org

Literature overview aims to outline the main characteristic features of HIV-1: its extraordinary genetic variability and mechanisms underlying it; the pandemic spread of the virus and its differential geographical distribution; classification of distinct strains and emergence of new viruses with mosaic genomes; the ability of the virus to evade immune system and implications it poses for vaccine development.

AIMS OF THE STUDY

Objectives of our study were:

- to investigate the genetic epidemiology of HIV-1 in Estonia
- to analyze the role of genetic background in HIV-1 infection and spread
- to develop antigenic sequences for genetic vaccination; to evaluate the swine model for studying the breadth of cellular immune response

Firstly, we characterized HIV-1 genetic diversity in Estonian population since the beginning of the epidemic in August 2000 when HIV-1 incidence grew explosively among Estonian intravenous drug users (IDUs).

Secondly, we explored HIV-1 co-receptor, chemokine receptor 5 (CCR5), polymorphism among Estonian HIV-1 patients. In Estonia the frequency of the HIV-1 restrictive CCR5Δ32 allele has been found to be among the highest in the world; concurrently, Estonia has a very high HIV-1 incidence. These features allowed us to explore the CCR5 polymorphism in a Caucasian population with a high frequency of the CCR5Δ32 allele, on a high background of HIV-1 incidence.

Thirdly, we showed a successful induction of broad cellular anti-HIV immune response with an improved DNA vaccine vector encoding artificial fusion MultiHIV consensus antigen consisting of several viral proteins and selected epitope-rich regions of HIV-1 in domestic pig model. We also demonstrated that pig is an excellent test animal for studying T cell responses in immunological surveys, *e.g.* in vaccine evaluation studies.

REVIEW OF THE LITERATURE

1. General description of HIV-1

HIV-1 is an RNA virus that belongs to the family *Retroviridae* and the genus *Lentiviridae* (Chiu et al., 1985). The viral genome consists of two plus stranded RNA molecules of approximately 9.6 kbp, non-covalently linked near their 5′ end. The RNA genome is surrounded by a protein capsid together with essential viral enzymes. The capsid is enclosed in a matrix layer that in turn is surrounded by a lipid envelope with embedded surface glycoproteins. The lipid bilayer is acquired from the host cell but the glycoproteins that protrude from the membrane are viral (*Immunobiology*, Murphy et. al. 2008), see Figure 2.

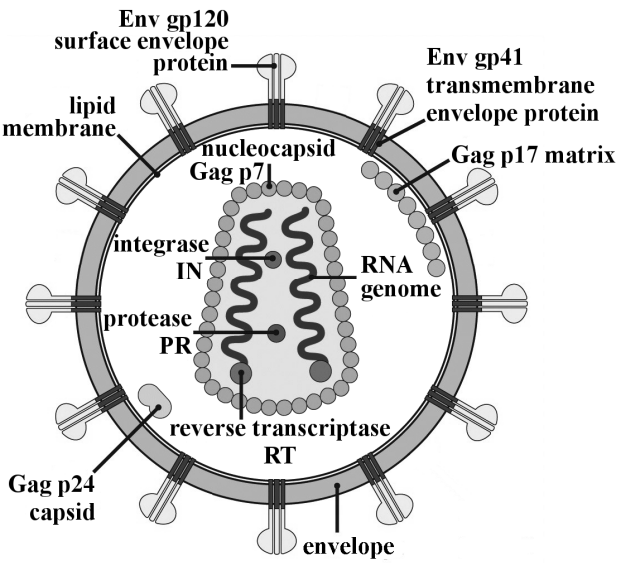

Figure 2. HIV virion. (Figure modified from *Immunobiology*, 7ed. Murphy et. al. 2008).

1.1. Genome structure of HIV-1

The HIV-1 RNA genome encodes the essential retrovirus genes *Gag*, *Pol* and *Env*, as well as the additional accessory/regulatory genes *Vif*, *Vpr*, *Vpu*, *Rev*, *Tat,* and *Nef,* see Figure 3.

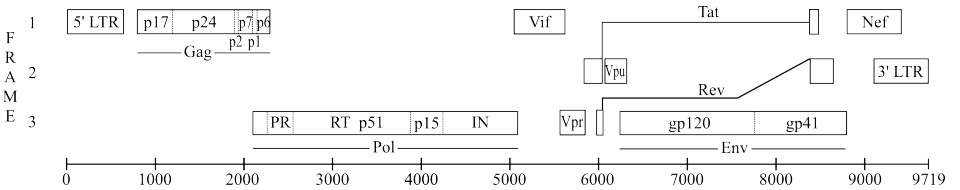

Figure 3. Genomic organization of the HIV-1 proviral genome. Open reading frames are shown as rectangles. Structural and enzymatic proteins are encoded by the *Gag, Pol* and *Env* genes. Regulatory gene products are encoded by the *Tat* and *Rev* genes and the auxiliary proteins are encoded by the *Vif, Vpu, Vpr* and *Nef*. The long terminal repeats (LTRs) are sites for initiation of viral RNA synthesis and necessary for proviral integration into host cell genome. The lower line shows the length of the proviral genome (ref. strain HXB2) in bp.

1.2. Structural genes

The HIV-1 genome has three main coding regions; *Gag* which directs the synthesis of the internal structural proteins, *Pol* which encodes the viral enzymes and *Env* which gives rise to the envelope proteins.

Gag refers to a genomic region encoding the capsid proteins, also known as group specific antigens, reviewed in (Göttlinger, 2001). Gag is often referred to as a precursor, because it is subject to cleavage by the viral protease. The precursor Gag $_{p55}$ protein is processed to Gag $_{p17}$ (matrix), Gag $_{p24}$ (capsid), Gag $_{p7}$ (nucleocapsid), and Gag $_{p6}$ (linker) protein and in addition small spacer peptides p2 and p1. Gag $_{p55}$ is called assemblin to emphasize its role in viral assembly. Gag $_{p17}$ is essential for membrane anchoring and nuclear transport of viral core as well as for the incorporation of the Env glycoprotein spikes during virus assembly. Gag $_{p24}$ has crucial role in particle assembly; in the mature virion, Gag $_{p24}$ forms the shell of the core. Gag $_{p7}$ is a highly basic protein that binds RNA and functions as a nucleic acid chaperone; it is essential for the specific packaging of two copies of the genomic viral RNA into assembling particles; it is also linked to HIV-1 recombination. Gag $_{p6}$ is important for retaining the Pol proteins within the virion once Gag-Pol is cleaved (Göttlinger, 2001, Hill et al., 2005).

Protease and other essential viral enzymes are brought into the virion as components of the Gag-Pol polyprotein, which is produced by ribosomal frameshifting between the overlapping *Gag* and *Pol* genes. Pol is expressed in the form of a 160 kDa Gag-Pol fusion protein and is cleaved to produce Gag $_{p17}$, Gag $_{p24}$, Gag $_{p7}$, the transframe protein (TF), and the viral enzymes protease (PR), reverse transcriptase (RT) and integrase (IN). PR catalyses the proteolytic processing of the Gag $_{p55}$ and Gag-Pol precursor-proteins. RT is a multifunctional enzyme with RNA-directed DNA polymerase, DNA-directed DNA polymerase and ribonuclease hybrid activities (RNase H). It catalyses the conversion of single stranded RNA into double stranded viral DNA for integration

into the host chromosomes. PR, RT and IN are not functional in their monomeric forms and must come together as either dimers (PR), heterodimers (RT: p_{66} [consisting of p_{51} and p_{15}] + p_{51}) or tetramers (IN) to be catalytically active, reviewed in (Hill et al., 2005). Gag and Pol are the most highly conserved HIV proteins and they contain the largest number of defined epitopes (Korber et al., 2001, Yusim et al., 2002).

HIV-1 interacts with target cells using envelope glycoproteins (Env_{gp120} and Env_{gp41}). Env_{gp120} and Env_{gp41} are formed by proteolysis of Env_{gp160} and remain non-covalently associated in a spike on the outside of the viral membrane. The Env_{gp120} surface subunit binds cell receptors to attach virus to target cells and regulate viral entry. Env_{gp120} is a polyclonal activator of B cells, causing them to differentiate into immunoglobulin producing cells while activating the complement cascade. The Env_{gp41} transmembrane subunit fuses host-cell and viral membranes to deliver the viral core into the cell cytoplasm. On the surface of the virion, these Env proteins form trimers, with Env_{gp120} on the outer surface bound to Env_{gp41} spanning the cellular membrane. The uncleaved envelope protein, which contains five regions of hyper-variability, differs between isolates not only in terms of amino acid sequences, but also in terms of length and sites of glycosylation, the latter accounting for more than half of the molecular weight of the protein. While variable parts of Env are frequent targets of immune activity, the conserved regions are often masked or hidden within the three-dimensional structure of HIV-1, reviewed in (McBurney & Ross, 2008, Stevceva et al., 2007, Surman et al., 2009, Wyatt R, 1998).

1.3. Regulatory and accessory genes

Tat and Rev proteins modulate transcriptional and posttranscriptional steps of virus gene expression and are essential for virus propagation.

HIV transcription is controlled primarily by the *trans*-activator protein (Tat). It acts by binding to a short nascent leader RNA, termed the *trans*-activation responsive region, or TAR, that is present at the 5' extremity of all viral transcripts, and by activating transcription initiation and elongation from the long terminal repeat (LTR) promoter (Campbell & Loret, 2009, Karn, 2000). Tat is also secreted by the infected cells and taken up by neighboring cells. This exogenous Tat is able to enter both uninfected and latently infected cells, inducing apoptosis in the former and activating the transcription of the viral genome in the latter. Thus, Tat fulfills a role in HIV-1 pathogenesis not only as an essential protein for HIV-1 transcription, but also as an extra-cellular toxin, reviewed in (Campbell & Loret, 2009).

Rev (regulator of expression of the virion) transports to cytoplasm single-spliced and un-spliced viral mRNAs that are required for expression of HIV structural proteins and production of genomic RNA. In addition to facilitating nuclear export, Rev increases stability and translation of HIV RNA. Rev is

considered the most functionally conserved regulatory protein of lentiviruses, reviewed in (Li et al., 2005).

Vif (virus infectivity factor) binds to an innate host defense protein APOBEC3G and counteracts its anti-HIV activity by promoting its degradation. APOBEC3G is a member of the cytidine deaminase family, which prevents viral cDNA synthesis via deaminating deoxycytidines in the minus-strand retroviral cDNA replication intermediate. As a result, it creates stop codons or G-A transitions in the newly synthesized viral cDNA which is then subjective to elimination by host DNA repair machinery. However, the virus has also developed an offensive strategy to suppress the antiviral effect of APOBEC3G through Vif, reviewed in (Romani et al., 2009).

Vpr (viral protein R) has been shown to have multiple activities during virus replication, including effects on the nuclear import of the proviral DNA as a component of the pre-integration complex (PIC). Vpr also alters cell cycle and induces apoptosis in infected cells (Romani et al., 2009).

Vpu (viral protein U) enhances the release of progeny virions from infected cells and induces the degradation of the CD4 receptor (Nomaguchi et al., 2008).

The HIV-1 Nef protein is abundantly produced during the early phase of viral replication cycle (Arien & Verhasselt, 2008). The role of Nef in HIV-1 replication and disease pathogenesis is determined by at least four independent activities. First, Nef affects the cell surface expression of several cellular proteins: it down-regulates CD4, CD8, major histocompatibility complex (MHC) class I and class II proteins, but up-regulates the invariant chain of MHC II (Trono, 1999). Secondly, it interferes with cellular signal transduction pathways. Thirdly, Nef enhances virion infectivity and viral replication. Fourthly, Nef regulates cholesterol trafficking in HIV-infected cells, reviewed in (Arien & Verhasselt, 2008).

1.4. Viral life-cycle

The main cellular receptor for HIV-1 is the CD4 molecule which is expressed in monocytes, macrophages, dendritic cells, $CD4^+$ T lymphocytes and microglial cells in brain (Klatzmann et al., 1984). Upon binding to the HIV receptor CD4, Env_{gp120} undergoes conformational changes, leading to the exposure of Env_{gp120}-buried domains that interact with chemokine receptors, which behave as HIV co-receptors (Deng et al., 1996, Moore et al., 2004). Co-receptor binding further induces conformational change in the transmembrane part of Env_{gp41} which triggers the fusion of the viral envelope to the cell membrane.

After entering the host cell the viral genome is reverse transcribed by the viral RT and the single stranded RNA genome of the virus is converted to a double strand DNA copy, termed provirus. Provirus is transported to the nucleus and subsequently integrated by the viral enzyme IN into the host cell's genome, where it is further transcribed by the host polymerases. By transcription from this proviral DNA, viral genomic RNA as well as series of

fully or partially spliced mRNAs is generated to serve for viral protein synthesis. In the early phase of transcription, the regulatory genes *Tat*, *Rev*, and *Nef* are expressed. Transcription of late genes results in production of structural proteins for the viral particle, as well as viral enzymes. Assembly of the viral proteins and genome takes place in the cell membrane where the new immature viral particle buds out, thereby receiving its envelope. Upon budding, immature particles rearrange to form mature, infectious virions. The viral protease is activated during assembly, and it cleaves Gag to generate a set of new proteins and spacer peptides. These newly processed proteins then reassemble to form the distinct layers of the mature virion: Gag $_{p17}$ remains associated with the inner viral membrane (the 'matrix' layer), Gag $_{p7}$ coats the viral RNA genome (the 'nucleocapsid' layer), and Gag $_{p24}$ assembles into the conical capsid that surrounds the nucleocapsid and its associated enzymes (*Molecular Cell Biology,* fifth ed. Lodish et al. 2004).

1.5. Co-receptor usage

The interaction of Env_{gp120}/CD4 complex with co-receptors stabilizes virus binding and triggers conformational changes in Env_{gp41}, which lead to the insertion of the hydrophobic Env_{gp41} amino-terminus into the host cell membrane, and eventually to the formation of a fusion pore and viral entry. Several chemokine receptors can serve as co-receptors for HIV-1 entry, including CCR2, CCR3, CCR8, CCR9, CXCR6, CXCR1, APJ, GRP1, and RDC1. However, based on the pathophysiological data concerning the viral life cycle *in vivo*, the major co-receptors are CCR5 which is predominantly expressed on dendritic cells, macrophages, and CD4 T cells, and CXCR4, which is expressed on activated T cells (*Immunobiology*, 7ed. Murphy et. al. 2008). As CCR5 is the primary co-receptor in the initial phases of infection, transmission of HIV-1 is in approximately 90% of cases associated with viruses that utilize CCR5, reviewed in (Moore et al., 2004, Taylor et al., 2008). It has been shown that R5 strains are preferentially transmitted from one infected individual to another as well as via sexual contact, vertical transmission, and direct blood transmission (Dragic et al., 1996).

CCR5 is a 352-aa cell surface receptor protein with seven transmembrane segments encoded by CMKBR5 gene. The 32-bp deletion in the gene segment encoding the second extracellular loop appears to be the most prevalent as well as most thoroughly studied mutation in CMKBR5 gene (Galvani & Novembre, 2005, Martinson et al., 1997, Novembre et al., 2005, O'Brien & Moore, 2000). Homozygous 32-bp deletion produces a truncated receptor, which is retained and degraded in endoplasmatic reticulum and is not expressed on the cell surface, neither in natural leukocyte populations, nor on transfected cell lines (Liu et al., 1996). Cells with heterozygous CCR5Δ32 genotype have reduced number of CCR5 receptors exposed on the surface (Liu et al., 1996, Venkatesan et al., 2002).

Among Caucasians in North America and Europe, the prevalence of CCR5Δ32 is approximately 1% for homozygotes and 10–20% to heterozygotes, with the highest frequencies in Nordic countries. Across Eurasia, the allele frequency of CCR5Δ32 forms a north-to-south cline, with frequencies above 15% in populations of Sweden, Estonia, Finland, and Northern Russia, to 4% in Sardinia and 2% in Greece. The Δ32 allele is virtually absent among non-Caucasians, as several studies of African, Asian and South-American populations have revealed (Balanovsky et al., 2005, Limborska et al., 2002, Martinson et al., 1997, McNicholl et al., 1997, Novembre et al., 2005, Voevodin et al., 1998). Very low frequencies are found in the Middle East and India and this allele is absent in China and Japan (Lu et al., 1999, Yudin et al., 1998).

Epidemiological studies have linked CCR5Δ32 heterozygosity with delayed onset of AIDS and reduced risk of initial transmission. This finding correlates with reduced levels of CCR5 and reduced infectability of $CD4^+$T cells (Mulherin et al., 2003, Reynes et al., 2000, Visco-Comandini et al., 1998). Individuals homozygous for the CCR5Δ32 allele do not express any of the CCR5 chemokine receptors on their cell surfaces, and in turn, they are largely resistant to infection by HIV-1 (Liu et al., 1996). Heterozygous individuals exhibit elevated resistance relative to wild-type individuals, and if heterozygotes do become infected, they have reduced HIV-1 viral loads with slowed progression to AIDS. Besides, heterozygosity could limit the epidemic by decreasing the probability of infection, as transmission rates are reduced for individuals with lower viral loads (Sullivan et al., 2001).

In contrast to humans, who are only recent hosts to HIV, natural SIV host species express remarkably low levels of CCR5 on $CD4^+$ T cells, especially in mucosal tissues where the fraction of $CD4^+$ T cells expressing CCR5 is >50% in humans, but only 1%–5% in African non-human primates (NHPs). As natural SIV infection in these animals is asymptomatic and usually does not induce significant lymphocyte depletion despite high levels of virus replication, this down-modulation of CCR5 expression on $CD4^+$ T cells is considered to be a key feature of SIV infection as opposed to HIV infection (Paiardini et al., 2009). CCR5 is an attractive antiviral target and therefore a number of gene therapy strategies have been developed to inhibit co-receptor synthesis or surface expression. Most of these strategies have not conferred complete down-regulation of cell surface CCR5 expression, however, and have failed to confer complete inhibition of HIV-1 replication for a sustained period of time, reviewed in (Nazari & Joshi, 2008).

2. Genetic variation of HIV-1

A central feature of *Lentiviridae* in general, and HIV-1 in particular, is the propensity to rapidly mutate and evolve in response to immunological or other selective pressures.

2.1. Classification

HIV is divided into HIV-1 and HIV-2, with sequence differentiation reaching up to 50%. HIV-2 is (supposedly due to its limited transmissibility) responsible for less than 1 million infections worldwide contrary to HIV-1, which has caused more than ~ 33 million infections (Tebit et al., 2007).

Phylogenetic analysis of numerous HIV-1 strains, isolated from diverse geographic origins, have revealed that they can be subdivided into groups, subtypes, sub-subtypes, circulating recombinant forms (CRFs) and unique recombinant forms (URFs), (Buonaguro et al., 2007, Korber et al., 2001, Peeters, 2000, Peeters & Sharp, 2000, Riva et al., 2008, Robertson et al., 2000, Taylor et al., 2008). The cumulative genetic variability of HIV-1 is managed on paper by the Los Alamos National Laboratory (LANL) HIV database (http://www.hiv.lanl.gov/). Figure 4 shows a phylogenetic tree demonstrating the known groups, subtypes, and sub-subtypes of HIV-1.

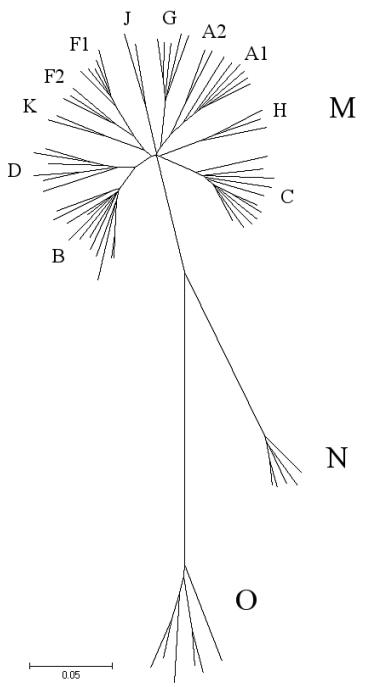

Figure 4. The phylogenetic tree showing all groups (M, N, and O), subtypes, and sub-subtypes (A1, A2, B, C, D, F1, F2, G, H, J, and K) of HIV-1. The tree was inferred from near-full-length non-recombinant sequences in the HIV sequence database of Los Alamos National Laboratory (http://www.hiv.lanl.gov/) using the Neighbor-Joining method (Tamura et al. 2007). A genetic distance scale of 0.05 is indicated.

HIV classification:

Groups – Distinctive HIV-1 lineages Major (M), Outlier (O) and non-M/non-O (N). While HIV-1 and -2 nucleotide variation might reach 50%, the differences between groups are ~ 30%.

- HIV-1 group M viruses are found globally and are largely responsible for the AIDS pandemic, while group O and N viruses are restricted to West Central Africa. Within group M, HIV-1 isolates have been further divided into 13 currently recognized distinct subtypes (also called clades) and sub-subtypes (Taylor et al., 2008).
- HIV-1 group O viruses were named "outlier" because of their distinct clustering from members of the group M viruses. Phylogenetic analyses of group O strains have not revealed the same substructure as found within the evolutionary tree of group M; they do exhibit a high diversity, but do not show the same radiation like group M viruses.
- HIV-1 group N viruses have only been identified in few patients and within last years. Phylogenetically they form an independent clade related to group M, whereas sequences from the 3' end cluster more closely with the SIVcpz chimpanzee virus, suggesting a possible ancient recombination event.

Subtypes/clades – Genetically related HIV-1 strains from group M that are phylogenetically approximately equidistant from each other (subtypes A, B, C, D, F, G, H, J, and K). Sequence similarity between subtypes is 70–90%, with the greatest genetic differences observed in the *Env* gene (up to 30% nucleotide diversity) followed by *Gag* (20%) and *Pol* (15%). Genetic variation within a subtype can be 15–20%.

Sub-subtypes – Within some subtypes (A and F), further phylogenetic structure can be identified (A1 and A2; F1 and F2). Also B and D clades are more closely related to each other than to other subtypes and currently subtype D is considered to be the early B-African variant, but their original designation as subtypes is retained for consistency with earlier published work. Also sub-subtypes A3 and A4 have been suggested to classification (Vidal et al., 2006).

Circulating recombinant forms (CRFs) – CRFs represent viral isolates that are the combination of two or more viral isolates from different subtypes that have been found in at least three epidemiologically independent individuals and play an important role in the HIV-1 pandemic. The assigned name of the CRF reflects sequence of discovery and subtype composition. If the genome contains sequences originating from more than two subtypes, the letters are replaced by "cpx", denoting "complex". It is estimated that CRFs account for about 10–20% of all new infections. The majority of CRFs have been found in areas with a high prevalence of different genetic subtypes. In regions where multiple subtypes are co-circulating with a high prevalence, recombination between recombinants has also been observed (so called second-generation recombinants) 43 CRFs are currently characterized, eight of which are mosaic genomes con-

taining gene regions of more than two subtypes and nine of which are second-generation recombinants (http://www.hiv.lanl.gov/content/index). More than 20 CRFs whose origins can be tracked to areas where the parental strains are co-circulating have been reported.

Unique recombinant forms (URFs) – URFs are generated in patients after superinfection or co-infection with different subtypes. URFs have heterogenous recombination breakpoints and have limited transmission in the human population. URFs are responsible for over 30% of the infections where several HIV-1 subtypes co-circulate.

During the early phases of the epidemic, HIV subtype classification was based on a sub-genomic region of an individual gene, mostly Env_{gp120}. This led to mistakes in the classification, for instance re-analysis has indicated that CRF02_AG is the parent of the recombinant subtype G, rather than the two having the opposite evolutionary relationship, as is currently proposed (Abecasis et al., 2007). At present, nucleotide sequences derived from multiple sub-genomic regions of the same isolates or near-full-length sequences are used for classification.

2.2. Mechanisms for producing viral variation

HIV has several intrinsic mechanisms that ensure rapid viral evolution (Gao et al., 2004).

Firstly, HIV-1 undergoes continuous genetic differentiation within individual patients due the error-prone nature of the viral RT ($3.4x10^{-5}$ mutations per base per cycle) combined with rapid viral turnover (~2.6 days/replication cycle) and high levels of persistent virus replication (10^{10} viral particles/day). This corresponds to approximately one new nucleotide substitution per genome per replication cycle (Gao et al., 2004). Beyond base-substitutions, HIV-1 is subject to relatively large deletions, insertions, and duplications. In addition, selective forces of the host environment, the pressure for change coupled with tolerance to change further contributes to the genetic variation. While immune escape clearly can drive positive selection in specific HIV epitopes, neutral mutations and genetic drift can also contribute to the overall diversity of the virus, reviewed in (Rambaut et al., 2004).

Secondly, genomic recombination between two different HIV-1 populations frequently occurs *in vivo*, resulting in biologically viable viruses with mosaic genomes. In regions where two or more different strains are in co-circulation, intersubtype recombination is common. Co-infection or super-infection with different genetic variants of HIV-1 does happen in nature; active infection with one virus strain does not confer complete protection against infection with another strain (Jung et al., 2002).

As a consequence, HIV-1 populations at the individual and global level exist as a heterogeneous set of highly related viral variants, referred to as quasi-species, unique to a host or, on a larger scale, to a geographical area (Tebit et al., 2007).

Recent studies continue to reveal evidence that the viral subtypes have different phenotypic properties, such as co-receptor utilization, *ex vivo* replicative fitness, rate of disease progression, biology of transmission, antigenicity, and mutational patterns (Lynch et al., 2009). Namely, the subtype-specific differences in frequency of usage of certain co-receptors have been found among subtype C and D isolates, as well as among Thailand B and E viruses (Subbarao et al., 2000). Also, it has been shown that subtype C HIV-1 isolates might have reduced *ex vivo* pathogenic fitness relative to other group M HIV-1 viruses (Abraha et al., 2009). *Ex vivo* studies on pathogenic fitness as well as on transmission fitness suggest that HIV-1 group M viruses are more fit than HIV-2, which in turn are more fit than group O viruses; this order in replicative fitness is reflected also in their prevalence in the epidemic (Arien et al., 2007).

2.3. Recombination in HIV-1

Genetic recombination is an integral part of the HIV lifecycle and it appears to be a common feature of retroviral replication in general. By this means the virus is provided with far more adaptive potential than is available from nucleotide substitution alone. The high capacity of the HIV-1 RT to generate recombinant progeny viruses has lead to the hypothesis that the reason retroviruses are diploid, is to provide a recombination partner. Two fundamental properties of retroviruses are critical to their high frequency of recombination. Firstly, retroviral genomic RNAs are encapsidated in pairs and upon infection of a new cell, the proximity of the two RNA molecules facilitates template switching that is in orders of magnitude more frequent than that of other viruses. Secondly, retroviruses are characterized by error-prone replication machinery, reviewed in (Clavel et al., 1989, Onafuwa-Nuga & Telesnitsky, 2009, Rambaut et al., 2004).

Retroviral genomes are composed of single-stranded RNAs, designated as "plus-strand", as they contain open reading frames that are recognizable by host ribosomes. Thus, upon infection of a target cell, the virally encoded RT is responsible for synthesis of the first DNA strand, designated as "minus single stranded" (-)ssDNA complementary to the genomic RNA, see Figure 5.

Once the genomic RNA is copied, it is degraded by the RNase H activity of the RT. Subsequently, using (-)ssDNA as template, the second DNA strand is synthesized (Ramirez et al., 2008). Following reverse transcription, the viral cDNA is primed for integration in the cytoplasm by integrase-mediated trimming of the 3'-ends of the viral DNA. Following 3'-processing, integrase remains bound to the viral cDNA as a multimeric complex that bridges both ends of the viral DNA within intracellular particles called pre-integration

complexes (PICs). The viral RT, Gag $_{p17}$, Gag $_{p7,}$ and Vpr contribute to the transport of PICs through the nuclear envelope. Once in the nucleus, IN catalyses the insertion of the viral cDNA ends into host chromosomes. This strand transfer reaction consists of the ligation of the viral 3'-OH DNA ends to the 5'-DNA phosphate of a host chromosome (Pommier et al. 2005).

During synthesis of the (-)ssDNA strand, the RT frequently switches from one copy of genomic RNA onto the other (a process known as copy choice) and if the two RNAs contain divergent sequences, copy choice results in genetic recombination (Ramirez et al., 2008). For recombination to take place, it is necessary that two divergent viruses establish successful infection of the same cell, and that the two types of genomic RNA have the capacity to form heterodimers. If these requirements are fulfilled, the nascent viral progeny will be constituted by homozygous as well as heterozygous particles and, during reverse transcription in a heterozygous virus, copy choice can eventually lead to synthesis of a recombinant proviral DNA (Ramirez et al., 2008). The fact that mosaic viruses are indeed recombinants is supported by the existence of discrete breakpoints between the genomic regions with different phylogenetic associations (Peeters, 2000).

The current estimates are that HIV-1 undergoes four to five crossovers per genome. Although recombination is observed throughout HIV-1 genomes, frequencies are not uniform and recombination hot spots have been identified. It has been suggested though, that most common crossovers observed *in vivo* are more likely the result of selection rather than template-switching hot spots (Onafuwa-Nuga & Telesnitsky, 2009, Zhuang et al., 2002).

Recombination contributes to HIV-1 genetics on several levels: from origins of the virus to its adaptability in individual patients. The chimpanzee virus that gave rise to HIV-1 was probably a recombinant of other SIVs; HIV-1 group N arose likely *via* recombination between an SIV and an early form of group M. Within patients, recombination is an extensive and ongoing source of viral diversity (Onafuwa-Nuga & Telesnitsky, 2009). The efficient spread of recombinant viruses has generated several CRFs, reviewed in (Buonaguro et al., 2007).

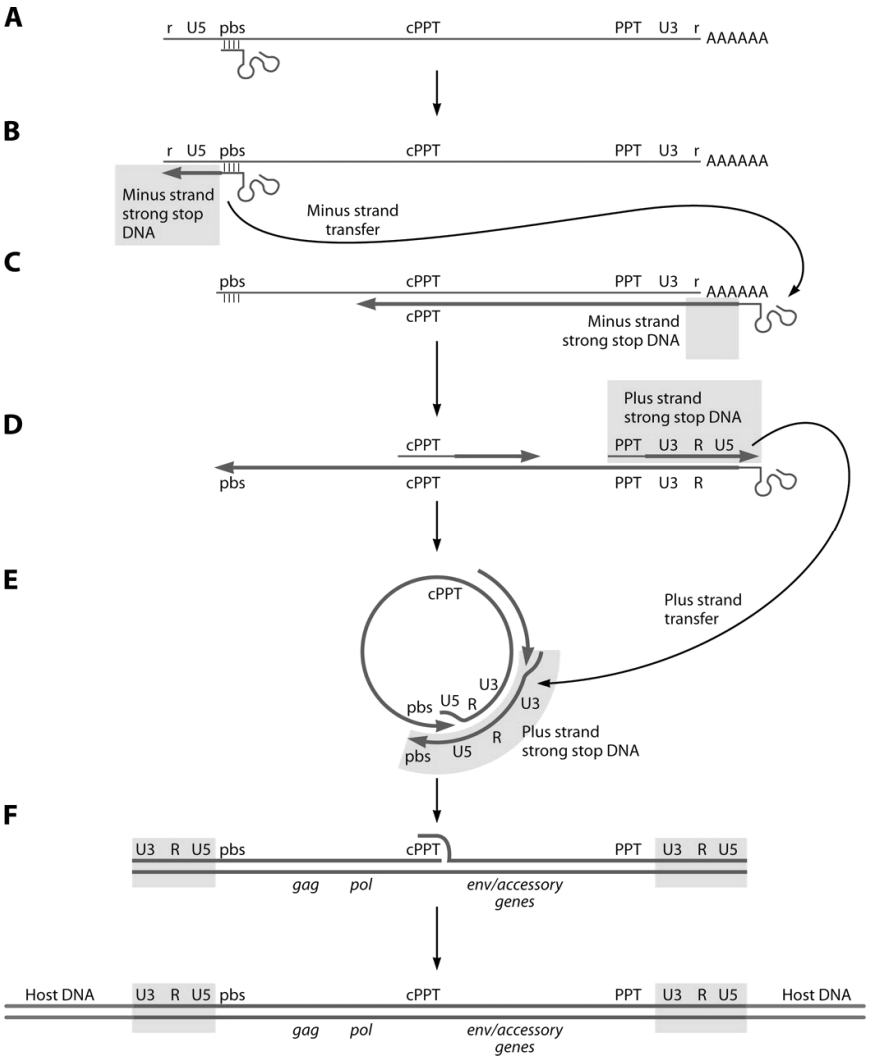

Figure 5. Process of HIV-1 reverse transcription. (A) A tRNA primer is bound to complementary sequences at the primer binding site (pbs) on encapsidated HIV-1 RNAs. (B) Minus-strand DNA synthesis initiates from primer tRNA and halts when it reaches the 5' end of RNA, thus creating a short DNA stretch comprising the 5' LTR – i.e. minus strong stop DNA. (C) The RNase H activity of RT degrades the copied region of the template RNA, a replicative template switch occurs and the minus-strand strong-stop DNA hybridizes with the 3' end of the RNA. (D) As minus-strand synthesis proceeds, plus-strand DNA synthesis is initiated from the polypurine tracks (PPTs). (E) After the RNaseH-mediated removal of the tRNA primer, plus-strand strong-stop strand transfer results from the annealing of the 3' end of plus-strand strong-stop DNA to complementary sequences at the end of the minus-strand DNA intermediate. (F) DNA synthesis is completed, generating double-stranded DNA with LTRs. Integrase catalyzes the establishment of the provirus, and host repair enzymes remove flaps and gaps. Reproduced from (Onafuwa-Nuga & Telesnitsky, 2009).

3. Molecular epidemiology of HIV-1

3.1. Geographical distribution

Throughout the world, individuals are infected with a diverse range of HIV isolates. The expansion of different HIV strains has been largely uneven. Founder effects, host restrictive factors, as well as social and behavioral facts all have contributed to this differential spread.

Today, the highest prevalence of HIV-1 is found in Southern Africa, where HIV-1 subtype C is dominant. In fact, among all known subtypes and CRFs, subtype C is the most prevalent and is accounting for more than 50% of group M infections worldwide and representing the predominant HIV-1 lineage in addition to Southern Africa, also in China and India (Osmanov et al., 2002, Tebit et al., 2007). In East Africa, subtypes A, D, and C as well as unique inter-subtype recombinants combining these three subtypes predominate. Subtype A and related CRFs account for roughly 30% of group M infections, and are primarily found in Western and Central Africa, which is described as an „HIV diversity hotspot" and where the greatest genetic variability of HIV-1 is found. For example, from Democratic Republic of Congo as well as from Cameroon, nearly every HIV type, group, subtype and circulating recombinant form can be identified in the infected population. Subtype B comprises about 15% of group M infections and is the predominant subtype in Europe, Australia and the Americas. Subtype B and related recombinants are also common in Asia (Kothe et al., 2007).

Sub-subtype A1 variants are spread globally and are differentiated into two distinct evolutionary lineages, one mainly restricted to Central Africa and the other responsible for the majority of epidemics in Eastern European countries formerly constituting the Soviet Union (Riva et al., 2008).

In the global HIV-1 pandemic the relevance of CRFs is increasingly rising, representing the predominant local form in South-East Asia (CRF01_AE) and in West and West Central Africa (CRF02_AG), (Buonaguro et al., 2007).

3.2. HIV-1 epidemiology in Estonian population

The spread of HIV-1 in eastern European countries has been closely linked with the rise of injecting drug use, which increased promptly after the collapse of Soviet Union in the 1990s. The HIV-1 epidemic among intravenous drug users (IDUs) started in 1996 in southern Ukraine with subtype A, probably derived from a single introduction event in Odessa, and later spread to other Former Soviet Union regions (Hamers & Downs, 2003, Riva et al., 2008), see Figure 6.

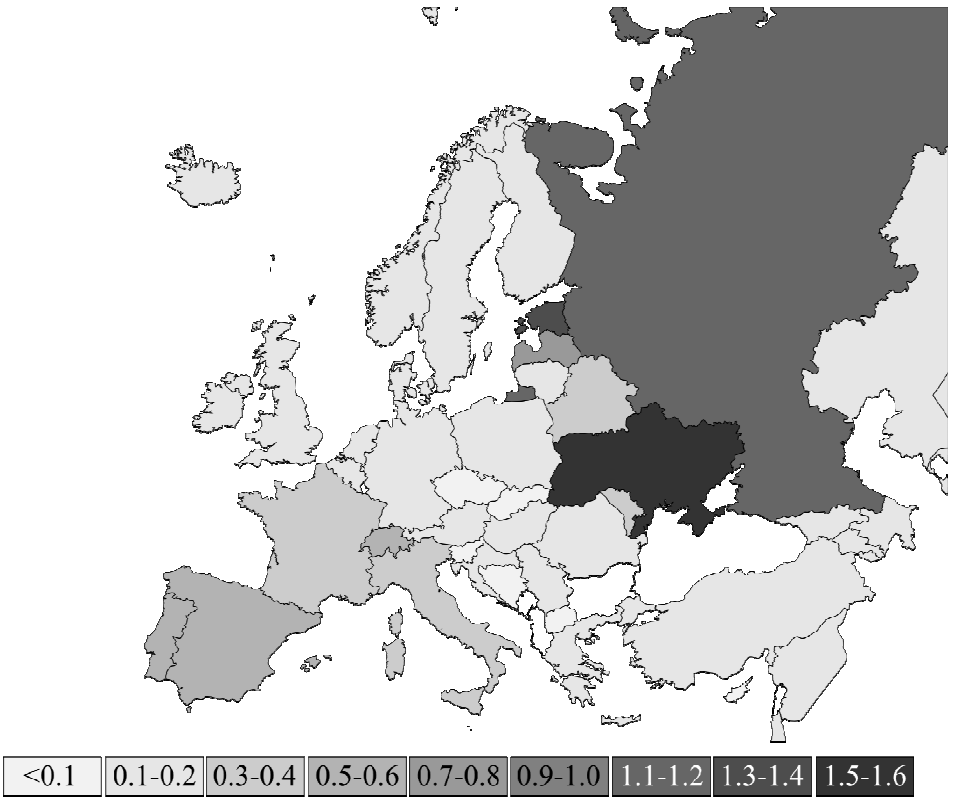

Figure 6. Estimated adult prevalence percentage of HIV in Central and Eastern Europe in 2007. Data source: *Report on the global AIDS epidemic* (2008) UNAIDS, http://www.unaids.org

In Estonia, during the 1990s, the reported number of HIV-1 cases was relatively low and the predominant HIV-1 type was B (likewise in other Baltic countries), with a limited distribution of subtypes A, C, D, F, and G as well as CRF02_AG (Liitsola et al., 1996, Naganawa et al., 2002). Before the epidemic outbreak, in 1999, Estonia reported 96 diagnosed HIV cases, only 4% of which were detected among IDUs. This situation changed drastically in August 2000 when the number of IDU associated HIV-positive cases started to increase rapidly and comprised nearly 90% of all new HIV cases reported in 2000 (Uusküla et al., 2002). Based on studies conducted among IDUs and other risk groups, the total number of Estonian inhabitants living with HIV was estimated to be approximately 11 000 in 2008 (Lai et al., 2009).

4. Medical interventions against HIV

The introduction of highly active antiretroviral therapy (HAART) has been successful in reducing the rates of vertical transmission as well as HIV-related morbidity and mortality among HIV-infected individuals. Antiretroviral therapy has moved from almost ineffective monotherapy to combination multidrug regimens which are able to effectively suppress viral replication in most HIV-infected patients. Currently, five different classes of anti-HIV drugs are available, classified by the phase of the retrovirus life-cycle that the drug inhibits: nucleoside and nucleotide RT inhibitors (inhibit reverse transcription by being incorporated into the newly synthesized viral DNA and preventing its further elongation); non-nucleoside inhibitors (inhibit the reverse transcription by binding to the p_{66} subunit of RT and altering its ability to function); protease inhibitors (inhibit the PR cleavage of polyproteins to mature proteins in budding virion); fusion inhibitors (interfere with binding, fusion and entry of HIV-1 to the host cell by binding to Env_{gp41}); and integrase inhibitors (inhibit the strand transfer process during the integration of viral DNA into the DNA of the infected cell).

The reality is however, that most people living with HIV in the developing world still have very limited access to antiretroviral treatment. In addition, response to HAART is a complex phenomenon and is often limited by the occurrence of acute or chronic toxicities or by the emergence or drug resistance (Tozzi, 2009).

It is estimated that approximately 60% of the HIV-infected individuals in Estonian population who could benefit from antiretroviral therapy, do not receive it, either because they are not aware of their HIV status, or because for various reasons they have not accessed health care services. This suboptimal coverage greatly increases the risk of HIV transmission in the general population (Lai et al., 2009).

The common notion is that the development of safe and effective HIV-1 vaccine would be better solution for the ultimate control of the worldwide HIV pandemic than expensive and life-long HAART. Unfortunately, regardless of two decades of research, this goal has not yet been achieved.

Recently the first tiny morsel of good news in HIV vaccine field came with the announcement that, for a first time, an HIV vaccine was shown to provide a modest degree of efficacy in phase 3 clinical trial (Lancet, 2009). A combination of ALVAC HIV, which consists of an attenuated canarypox virus vector encoding HIV *gag*, *pol,* and *env* genes and AIDSVAX B/E recombinant Env_{gp120} *versus* placebo were tested in more than 16 000 HIV-negative volunteers in Thailand. During 3 years of the study, HIV infection was detected in 51 (0.62%) of 8197 volunteers given the vaccine combination and 74 (0.90%) of 8198 given placebo, leading to the absolute difference of 0.28% (Lancet, 2009).

4.1. Vaccines

HIV-1 has escaped from the attempts of scientists to develop an HIV vaccine for 25 years due to its unique virologic and immunologic characteristics: its extraordinarily high diversification rate, which enables immune evasion and its capacity to evade adaptive immune responses, the inability to induce broadly reactive antibody responses and the early establishment of latent viral reservoirs, reviewed in (Barouch, 2008, Taylor et al., 2008).

An ideal long-term solution to the HIV-1 pandemic is a vaccine that blocks infection completely – confers sterilizing immunity. Neutralizing antibodies, $CD4^+$ and $CD8^+$ T-cell responses as well as innate immunity have all been linked to protective efficacy; however none of these parameters are so far shown as universal (McBurney & Ross, 2008). A less efficient, but still valuable alternative would be a vaccine that considerably diminishes the risk of transmitting infection to a new host and reduces manifestation of HIV-1 induced disease phenotype by providing partial immune control of viral replication after infection. The latter is realistic goal of human leukocyte antigen (HLA) -mediated cytotoxic T lymphocyte (CTL)-inducing vaccines that do not trigger neutralizing antibodies (Barouch, 2008, Letvin, 2006).

4.1.1. Current HIV-1 vaccine strategies

Vaccine strategies for HIV-1 can be divided into traditional and novel approaches, reviewed in (Barouch, 2008, Letvin, 2006, Mooij & Heeney, 2001). The traditional technologies include live attenuated viruses, whole killed viruses, and protein subunits. These strategies have been most successful with vaccines for other viruses; however they are not exploitable for HIV-1. Although live attenuated viruses provide substantial protective efficacy against SIV challenges in rhesus monkeys, they can not be used in humans due to the significant safety concerns (possibility of reversion). The use of whole killed viruses and protein subunits, on the contrary, have been insufficient to induce broadly active neutralizing antibody responses as well as $CD8^+$ T lymphocyte responses, reviewed in (Barouch, 2008, Mooij & Heeney, 2001).

Novel strategies include plasmid DNA vaccines and live recombinant vectors that are engineered to express HIV-1 antigens. Recombinant vectors include attenuated or replication-incompetent viruses, *e.g.* adenoviruses and poxviruses with replication potential to express HIV-1 antigens. However, the magnitude of the cellular immune responses of live recombinant virus vectors in human volunteers immunized with such vaccines has been substantially lower than that seen in non-human primates. This reduced immunogenicity might be a consequence of pre-existing vector-specific immunity – as with most viral vectors derive from current widespread human infections. Plasmid DNA vaccine constructs are non-live, non-replicating, non-spreading, highly flexible and have proved to be effective immunogens in mice for eliciting cellular

immune responses and for priming antibody responses. However, they are less immunogenic in non-human primates than they are in mice, and even less immunogenic in humans than in non-human primates. As a result, with plasmid DNA vaccines, multiple injections of high doses of DNA are usually required to elicit detectable immune responses. To overcome this, substantial research effort is put on the molecular adjuvants, such as plasmid DNA immunogens encoding cytokines and delivery technologies, such as *in vivo* electroporation, to increase the efficiency of cellular transfection with DNA. Also, gene optimization strategies (modification of the Kozak sequence, species-specific codon optimization) and RNA structural design (removal of instability elements that lower expression) have been introduced to increase protein production on a per-cell basis, leading to enhanced T-cell responses and antibody induction, reviewed in (Barouch, 2008, Kutzler & Weiner, 2008, Letvin, 2006, Mooij & Heeney, 2001).

Among all known subtypes, subtype C is the most prevalent, accounting for more than 50% of group M infections worldwide; subtype A and related CRFs account for roughly 30% and subtype B comprises about 15% of group M infections (Osmanov et al., 2002). Since all other subtypes and CFRs are less prevalent, candidate vaccines have historically been selected from members of subtypes A, B, and C (Douek et al., 2006).

4.1.2. Immunologic challenges for vaccine development

4.1.2.1. HIV-1-specific humoral immunity

The earliest vaccine efforts attempted to block HIV-1 during the initial steps of infection, with the goal of achieving sterilizing immunity. After infection, antibodies bind to and interfere with the HIV envelope protein, and these 'neutralizing antibodies' were seen as promising key components of a protective vaccine. Although antibodies cannot fully control an established infection, they can reduce viral infectivity and modulate infection if present early. Also, virus-specific neutralizing antibody titers represent key immune correlates of protection for most licensed viral vaccines, and thus first vaccine developmental strategies have been focused on developing HIV-1 Env subunit immunogens (Barouch, 2008, Haigwood & Hirsch, 2009). Proof-of-concept passive transfer studies in NHPs have shown that administration of high doses of broadly reactive monoclonal antibodies can afford sterilizing protection from infection, thus demonstrating the potential of virus-specific humoral immunity (Mascola et al., 2000). The first vaccine concept that completed clinical efficacy studies used monomeric HIV-1 Env_{gp120} protein with the aim to induce Env-specific humoral immune responses. Phase 3 efficacy trials indicated though, that the type-specific antibody responses that were generated were insufficient to protect against HIV-1 (Flynn et al., 2005, Gilbert et al., 2005). Development of a vaccine that induces neutralizing antibodies that bind to the trimeric envelope

on the surface of the virion remains a great challenge. There are currently no vaccine candidates that are aimed at eliciting broadly reactive Env-specific neutralizing antibodies in clinical trial (Barouch, 2008).

Instead a markedly different approach for achieving humoral immunity was recently introduced by two different groups in NHPs, reviewed in (Haigwood & Hirsch, 2009). The study of Hessell *et al.* indicated that moderate amounts of repeatedly transferred neutralizing antibodies can protect from multiple viral challenges (Hessell et al., 2009). Johnson *et al*. showed that long-lasting neutralizing antibodies can be produced by gene transfer *in vivo* — thus providing continuous protection against the virus (Johnson et al., 2009). Both of these studies, however, used viruses for challenge that were specifically matched to the neutralizing antibody tested.

4.1.2.2. HIV-1-specific cellular immunity

The HIV-1 vaccine field has also been developing immunogens for induction of T-cell response, encouraged by data from natural history studies and NHP models: $CD8^+$ T cell responses emerge during acute infection coincident with initial control of primary viremia; potent cellular responses have been reported in long-term non-progressors; specific HLA alleles have been correlated with control of viral replication; experimental depletion of $CD8^+$ lymphocytes has been shown to abrogate immune control of SIV replication in NHPs, reviewed in (Barouch, 2008, Korber et al., 2001, Mooij & Heeney, 2001).

Recently, results of a vaccine concept evaluated in phase 3 efficacy trials with the aim to elicit HIV-1-specific cellular immune responses were announced (Buchbinder et al., 2008). This approach used replication-incompetent recombinant adenovirus serotype 5 (rAd5) vectors expressing HIV-1 Gag, Pol, and Nef. It failed to protect against infection or to reduce viral loads after infection and, furthermore, the vaccinees with pre-existing Ad5-specific neutralizing antibodies exhibited an enhanced rate of HIV-1 acquisition (Watkins et al., 2008).

HIV-specific $CD4^+$ and $CD8^+$ T cell responses have been shown to recognize peptides from every HIV-1 protein, although vast majority of defined epitopes are found in HIV Gag, Pol, Env, and Nef (Korber et al., 2001). The breadth of epitope-specific T lymphocyte responses is considered to be critical for an HIV-1 vaccine, for maximizing immunologic coverage of HIV-1 diversity on one hand and for minimizing the potential for viral escape from recognition by T lymphocytes on the other. One basic approach is the inclusion of multiple whole HIV-1 protein coding regions to maximize the potential breadth of the response. Another strategy is to concatenate conserved DNA encoding epitopes into a single contiguous DNA vaccine (Korber et al., 2001). It has been suggested that HIV-1 vaccines including highly conserved, cross-clade or multi-clade T cell epitopes would be the most effective types of vaccine in the global context of

HIV epidemic, especially if supertype epitopes and promiscuous epitopes are included (De Groot et al., 2003).

4.1.2.3. MHC-I role in HIV-1 infection

Viral proteins are not recognized by CTL in their native form, but are cleaved into peptides in the cytosol of infected cells, transported to the endoplasmatic reticulum and loaded into a groove on major histocompatibility complex class I (MHC-I) molecules, *i.e.* HLA-I in humans (*Immunobiology*, 7ed. Murphy et. al. 2008). HLA-I molecules present peptide antigens on the surface of infected cells, targeting them for removal by the CTLs. HLA-I genes are highly polymorphic loci and the efficiency of this process differs amongst individuals because of the high number of allelic variants of HLA-I genes. Almost all of the extensive HLA-I polymorphism is concentrated to those residues that are located in the peptide-binding groove of the molecules, thereby defining the peptides that bind to each HLA molecule, reviewed in (Goulder & Watkins, 2008). However, HIV-1 can escape the HLA-mediated CTL response through the evolution of variants within the targeted epitopes, abolishing recognition by the CTL. Virus variability may influence any step in epitope processing (Paradela et al., 2000), however, escape from a CTL response is balanced by detrimental impacts on viral replication (Moore et al., 2002). Higher HLA polymorphism is suggested to be associated with elevated immune response efficiency towards HIV-1. This effect rises from the ability to present a broader repertoire of antigens to T cells, thus complicating the emergence of HIV-1 escape mutants (Piacentini et al., 2008). Individuals with HLA alleles that specifically select regions of HIV proteins that infrequently mutate, due to the effects on viral fitness, could therefore have better outcomes when infected with HIV, as the potential for the appearance of escape mutants is reduced (Blackwell et al., 2009). An example is HLA-B*27 allele, which recognizes a conserved epitope in the Gag $_{p24}$ protein and is associated with significantly improved survival in HIV-infected individuals. As an opposite, certain other HLA-I alleles are associated with rapid disease progression of HIV-1, namely, HLA-B*35-restricted escape mutants affect the recognition of CTL by reducing both peptide binding and T cell receptor recognition, leading to rapid HIV disease progression (Blackwell et al., 2009). However, by investigating numerous cohorts from North America, the Caribbean, Europe, sub-Saharan Africa, Australia and Japan, it was shown recently that the frequency of escape mutations to CTL pressure in a given population is correlated with the prevalence of the relevant HLA allele in that population. Therefore, the accumulation and fixation of escape mutations at a population level suggests that viruses are becoming resistant to CTL responses and that previously identified protective HLA alleles (*e.g.* B*57) might lose their advantage (Kawashima et al., 2009).

SLA-I

The process of evolution of MHC-I genes in mammals is highly complicated, comprising many types of species-specific expansions of class I genes. Consequently, the structure of the gene family of MHC-I varies among species, and it is often difficult to find counterparts of a particular class I gene even in closely related species (Tanaka-Matsuda et al., 2009). Surprisingly, the swine MHC-I or swine leukocyte antigen (SLA-I) complex has been shown to be remarkably similar to human MHC-I (HLA-I), (Renard et al., 2006). SLA-I differs by only one rather extensive deletion and two moderate gene expansions from the HLA-I region, opposite to other mammalian species, *e.g*. rhesus macaques, among whom multiple blocks of expansion have been observed (Renard et al., 2003).

The polymorphic amino acid positions of the classical SLA-I sequences have been shown to be almost perfectly superimposable on those of the HLA class I sequences. However, the overall genomic organization of the SLA-I region is quite different from that of the HLA-I region (Ando et al., 2003, Lunney et al., 2009). It has been suggested that differences in gene organization of the MHC-I region in mammalian species arose due to gene duplications after speciation (Lunney et al., 2009). In the pig genome, three phases of the SLA class I region evolution have been distinguished, the most recent corresponding to the evolution of the SLA-I genes, which arose through a series of five duplications of a class I containing elementary unit. The origin of SLA-1, SLA-2, and SLA-3 genes from the common SLA-I unit ancestor explains the high similarity present among different SLA loci: SLA-1 and SLA-2 coding sequences are nearly 98% similar to each other on the nucleotide level and SLA-3 about 97% (Renard et al., 2003).

Compared with a large number of alleles of the HLA-A, -B and -C loci, a relatively limited number of alleles in the three SLA-I loci have been reported so far. According to the Immuno Polymorphism Database (IPD)-MHC SLA sequence database (http://www.ebi.ac.uk/ipd/mhc/sla/), there are 125 classical SLA-I (SLA-1, SLA-3 and SLA-2) alleles, designated by the SLA Nomenclature Committee of the International Society for Animal Genetics (ISAG), (Ho et al., 2009). With numerous swine breeds worldwide, the actual extent of SLA diversity in outbred pig populations is still not known (Lunney et al., 2009). The polymorphisms of the SLA-I genes are concentrated in exons 2 and 3 of the coding regions which form the class I protein peptide-binding groove (Lunney et al., 2009).

Unlike the HLA-I molecules, the foreign antigen binding sites and T cell recognition sites in the SLA-I have not been precisely clarified yet. Some putative antigen recognition sites in the SLA-I in relation to those in the HLA-I have been determined by analyzing selective forces operating at the amino acid sequence level. This analysis suggested that selective forces operating at single amino acid sites of the SLA-I molecules are almost similar to those of the HLA-I, although some functional sites for antigen and CTL recognition might differ (Ando et al., 2003).

4.1.3. Implications of genetic subtypes for vaccine development

Another challenge facing AIDS vaccine development is the vast variability among circulating HIV-1 populations in different geographical locations and the need to develop vaccines that can elicit enduring protective immunity to various HIV-1 strains.

Due to the huge diversity of HIV-1 isolates circulating in the population, it would be overly optimistic to expect that a vaccine based on a single viral isolate will be able to protect against the infection, reviewed in (Gaschen et al., 2002, McBurney & Ross, 2008). If any currently circulating HIV-1 strain would be selected as a candidate immunogen, then one should bear in mind that it might be as distant from other present-day viruses as these are from each other. To maximize the degree of sequence similarity, the use of "centralized" HIV-1 immunogens, expressed from consensus or reconstructed ancestor gene sequences have been proposed (Ellenberger et al., 2002, Gaschen et al., 2002, Korber et al., 2001, McBurney & Ross, 2008, Nickle et al., 2003, Novitsky et al., 2002). Because of their "central" position within an evolutionary tree, these inferred sequences are almost half as distant from extant HIV-1 strains, as the latter are from each other and should thus contain a larger number of conserved epitopes and may have enhanced potential for eliciting cross-reactive responses (Gaschen et al., 2002).

Three methods of developing centralized vaccines are ancestral, center-of-the-tree and consensus, see Figure 7.

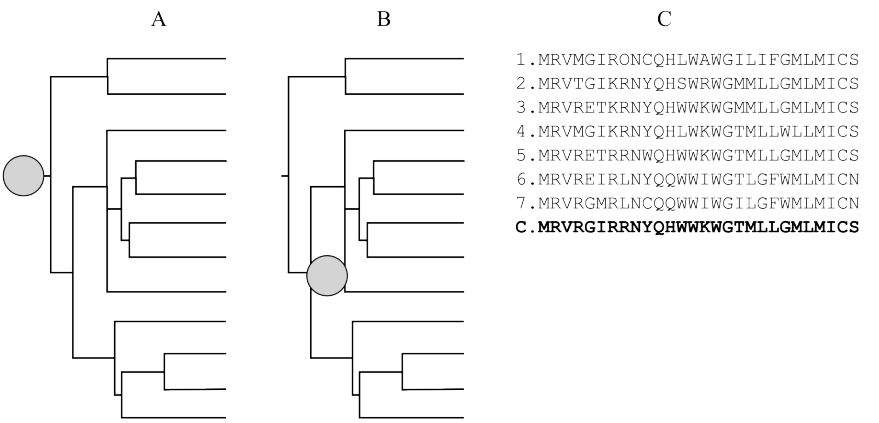

Figure 7. Three methods of developing a centralized vaccine. (A) Ancestral vaccine. This method of centralization utilizes the theoretical ancestor that gave rise to the phylogenetic tree. The location of the ancestral vaccine is shown with a grey dot. (B) Center of the tree. This method of centralization generates a sequence that is equidistant to all points of the phylogenetic tree. The approximate location of the vaccine sequence is shown with a grey dot. (C) Consensus vaccine. A consensus sequence is generated from the seven sample *Env* sequences by assigning the most common amino acid at each position in the amino acid sequence (McBurney & Ross, 2008).

A phylogenetic tree, for example, by means of maximum likelihood approach is developed using primary viral protein sequences to develop center-of-the-tree or ancestral sequences. Center-of-the-tree sequences are based upon generating a sequence that is equidistant to all points of the tree. The ancestral approach is based upon the theoretical ancestor sequence that gave rise to all sequences of a particular tree. In addition, the population of these sequences can be aligned to make a consensus sequence based on the most common amino acid in each position in an alignment (Gaschen et al., 2002, Nickle et al., 2003).

Centralized vaccine strategy has been used to construct a number of HIV-1 gene sequences (Burgers et al., 2006, Malm et al., 2005, Santra et al., 2008, Thomson et al., 2005). It remains to be determined weather vaccines that are based on such artificial genes elicit immune responses with greater width than vaccines that incorporate single, naturally occurring viral genes.

RESULTS AND DISCUSSION

Genetic structure of the HIV-1 strains causing the epidemic outbreak of HIV-1 infection in Estonia (Ref I)

The spread of HIV-1 in eastern European countries has been closely linked with the rise of intravenous drug use, which increased promptly after the collapse of Soviet Union in 1990s. The drug-related epidemic started in 1996 in Ukraine, rapidly followed by neighboring countries: Belarus, Russia, Moldova, and Kazakhstan (Hamers & Downs, 2003). In all these regions, sub-subtype A1 viruses predominated; in fact, due to the distinctive way of transmission, the lineage was designated as IDU-A (Riva et al., 2008). Eventually, in year 2000 the HIV-1 epidemic reached to Estonia.

Studies carried out in the 1990s, before the epidemic outbreak, had determined subtype B to be a predominant form of HIV-1 in the Estonian population, like elsewhere in Europe (Liitsola et al., 1996).

The study was initiated to establish the exact genetic structure of the HIV-1 strains causing the IDU-associated epidemic outbreak of HIV-1 infection in Estonia after August 2000.

Two proviral genomic regions were selected to identify existing strains and to detect possible recombination events: a 430–base pair (bp) fragment of the *gag/pol* region and a 387-bp fragment of env_{gp41}. For subtyping, proviral genomic DNA samples from 141 HIV-1 positive individuals were amplified by PCR reaction and sequenced. Detailed descriptions of methods and materials used are given in *reference I*.

Phylogenetic analysis revealed that a rare circulating recombinant form, CRF06_cpx is predominant in population of Estonian HIV-1 positive individuals. CRF06_cpx is a complex mosaic form of the virus composed of successive fragments of subtype A, G, K and J and shows similarity to G sequences in both selected env_{gp41} and *gag/pol* regions (Montavon et al., 1999, Oelrichs et al., 1998), see Figure 8.

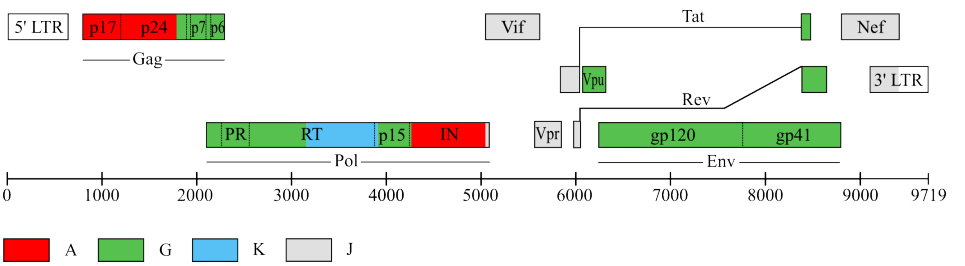

Figure 8. The mosaic structure of CRF06_cpx composed of subtypes A, G, K, and J. Figure adopted from (Montavon et al., 1999, Oelrichs et al., 1998).

While majority of the samples from recently infected individuals belonged to type CRF06_cpx, the remaining demonstrated remarkable similarity with a subtype A sequence isolate derived from the Ukraine. The samples of patients with HIV-positive status identified before 2000 were classified as subtype B.

In addition to CRF06_cpx, subtype A, and subtype B sequences, the phylogenetic analysis exposed yet another type of viruses with mosaic structure, resembling subtype A in the first part of the *gag/pol* segment and CRF06_cpx in the second. Moreover, in cases of many individuals, discordant results of subtype determination became obvious between *gag/pol* or env_{gp41} sequence data, see Table 1.

Table 1. Genetic subtypes of HIV-1 in Estonian samples

		gag/pol					
		A	B	CRF06	A/CRF06	Total	%
env_{gp41}	A	1	0	1	2	4	3
	B	0	6	0	0	6	4
	CRF06	3	0	107	21	131	93
	Total	4	6	108	23	141	100
	%	3	4	77	16	100	

Detailed comparison of the mosaic *gag/pol* sequences revealed the presence of multiple different types of recombinant forms, each type showing different pattern of potential cross-over points. Sequences that differed from each other by less than three consecutive character changes were grouped together, defining 9 different groups. It is noteworthy, that the crossover sites of Estonian *gag/pol* URFs correspond to the genomic region (5'end of *Pol*) of HIV-1 that has been characterized as a region of recombination hot spot (Jetzt et al., 2000). A high incidence of different recombinant forms in the Estonian IDU population indicates that co-infection with different strains of HIV-1 occurs frequently among drug users.

The CRF06_cpx isolates that are circulating in Estonia show remarkably low nucleotide diversity which suggests a recent introduction of the strain to Estonia. Due to the small number of available reference sequences from geographically remote areas and the significant difference between our sequences compared to other CRF06_cpx isolates, the origin of the CRF06_cpx lineage in Estonia could not be traced. There is evidence, though that this lineage was also present in neighboring countries Russia and Sweden – already in 2002 (Skar et al., 2008).

Identification of a novel circulating recombinant form CRF32_06A1 (Ref I)

Currently, 43 intersubtype recombinants fulfill the requirement necessary to be circulating recombinant forms (CRFs), meaning they have been identified in at least three individuals who are epidemiologically unlinked and at least two full-length viruses with identical recombination breakpoints must be sequenced (Robertson et al., 2000). The impact of these CRFs on the evolution of the pandemic is evident, as already some have spread more widely than their parental subtypes. For example, CRF01_AE is a highly prevalent strain in Southeast Asia whose parental subtype E has never been detected (Peeters, 2000).

To study the prevalent Estonian HIV-1 forms in more detail, four representative HIV-1 near-full-length proviral genomes were sequenced. For detecting HIV-1 intersubtype recombination the bootscanning-based method Simplot (Lole et al., 1999) was used. SimPlot calculates and graphically plots the phylogenetic identity (bootscanning) of the query sequence to a group of reference sequences. A detailed description of parameters used is given in *reference I*.

The results of the bootscanning analyses confirmed the mosaic structure of the isolates EE0369 and EST2002-1169, consisting of stretches of subtypes A and CRF06_cpx.

The data on 2 full-length and 6 partial sequences with the same recombinant structure allowed us to designate this as the emergence of a new CRF, namely CRF32_06A1, see Figure 9.

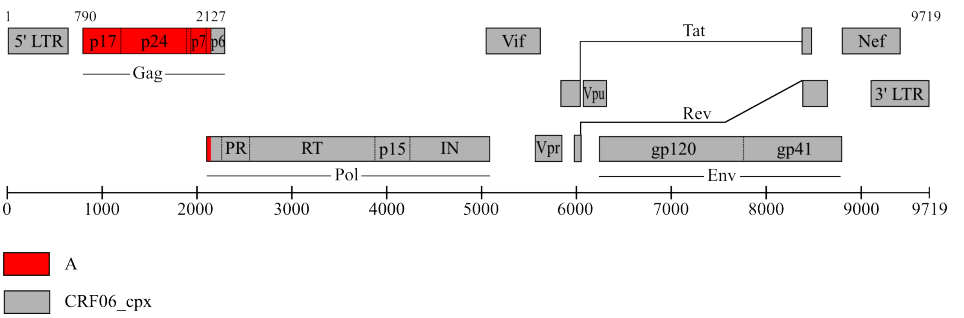

Figure 9. Novel CRF CRF32_06A1, breakpoints: 1-789: 06_cpx; 790-2126: A1; 2127-9719: 06_cpx; reference sequence EE0369 (AY535660). Breakpoint locations are based on the HxB2 numbering. The illustration was adopted from the Los Alamos HIV Sequence Database Website.

As reviewed in (Thomson & Najera, 2007), apart from subtype B and the IDU-A, HIV-1 clades reported to be circulating in Europe include subtype G and CRF14_BG in Portugal and Spain, CRF11_cpx in Switzerland, CRF01_AE in

Finland, CRF04_cpx in Greece, subtype F in Romania, CRF03_AB in Russia, and a subtype G variant (unrelated to the Iberian strain) in Italy, demonstrating the increasing diversity of HIV-1 in Europe. Thus, the emergence of a rare CRF06_cpx as well as CRF32_06A1 in Estonia adds to the findings of other studies indicating a trend toward an increasing genetic complexity of HIV-1 epidemics in Europe.

Frequency of CCR5Δ32 polymorphism among HIV-1-positive individuals in Estonian population (Ref II)

The Δ32 deletion mutant of CCR5, resulting in a non-functional receptor not reaching the cell surface, is clearly associated with strong, although incomplete, resistance to HIV infection for homozygotes (Dean et al., 1996, Liu et al., 1996, Samson et al., 1996). Several studies have investigated whether heterozygosity for a CCR5Δ32 polymorphism affects susceptibility towards infection of the HIV-1, but the results obtained so far appear to be contradictory, reviewed in (Arenzana-Seisdedos & Parmentier, 2006). In many cases, the interpretation of observed differences in genotype and allele frequencies in terms of susceptibility is often complicated, as in populations with higher frequency of CCR5Δ32 allele (Caucasians in North America and Europe) the HIV-1 incidence is relatively low. Additionally, heterogeneous ethnical/genetic backgrounds as well as differences in route of transmission likely contribute significantly to the end result.

We analyzed the CCR5Δ32 polymorphism among HIV-seropositive individuals in a Caucasian population simultaneously characterized by a high frequency of CCR5Δ32 allele (Kalev et al., 2000) and high HIV-1 incidence – a combination of features that only seldom coexist. A detailed description of methods and materials used is given in *reference II*. We found that CCR5Δ32 heterozygosity occurs with the same frequency among HIV-1 seropositives as in healthy control population, thus suggesting that in case of IDU associated epidemics, CCR5Δ32 heterozygosity does not confer protection from HIV-1 infection, see Table 2.

It has been hypothesized, that difference in transmission route or characteristics of the HIV-1 variants that circulate in IDUs could explain why different results are obtained when IDU-associated epidemics are involved in such surveys (Lukashov et al., 1996, Op de Coul et al., 2001, Schinkel et al., 1999). The distinct transmission route in IDUs, among whom the initial virus replication takes place in lymph nodes (not in mucosal macrophages like in sexual transmission), could explain why our results are different from several other studies, mainly involving homosexuals, where protective effect of the CCR5Δ32 heterozygosity was postulated (Marmor et al., 2001, Meyer et al., 1997, Papa et al., 2000, Samson et al., 1996). However, if this were a risk-group specific phenomenon, we would see elevated frequency of CCR5Δ32 heterozygotes among

IDUs compared to sexually infected patients. Our results obtained by estimating CCR5Δ32 allele frequencies among sexually infected patients and IDUs did not reveal statistically significant differences, though.

Table 2. CCR5Δ32 genotypes and allele frequencies in controls and in HIV-1 patients.

	HIV-1 patients	Controls
Total	300	504
ccr5 wt/wt	230 (76.7%)	371 (73.6%)
ccr5 wt/Δ32	70 (23.3%)	117 (23.2%)
ccr5 Δ32/Δ32	0	16 (3.2%)
Allelic frequency of Δ32	0.117	0.148

HIV-1 cellular tropism and co-receptor specificity are largely determined by the sequence of the third hypervariable loop (V3) of the viral Env_{gp120} glycoprotein, and distinct changes in this region have been associated with the NSI/SI phenotype (Deng et al., 1996). However, the V3 regions of the Estonian isolates are highly conserved and have a low overall positive charge, which is consistent with the NSI phenotypes and using CCR5 as a co receptor (Cilliers et al., 2003).

It has been demonstrated that the concentration of the CCR5 co-receptor on the cell surface correlates significantly with cell permissiveness and thus modulates primary HIV infection. However, it has remained unknown, how the reduced cell surface expression of CCR5 affects dual infection (Ciuffi et al., 2004, Reynes et al., 2000). We analyzed, whether CCR5Δ32 heterozygotes harbor URFs on an equal frequency to CCR5WT homozygotes. The statistical analysis confirmed equal distribution of URFs both among CCR5Δ32 heterozygotes as well as among CCR5WT homozygotes. Thus, our data does not reveal any association between CCR5Δ32 genotype status and frequency of dual infection.

Heterozygotes of CCR5Δ32 allele have been found to display slower progression to clinical stages of AIDS, reviewed in (Arenzana-Seisdedos & Parmentier, 2006). As during the current survey the HIV-1 epidemic was in its early stage, the impact of the CCR5Δ32 polymorphism on the progression of the HIV infection towards AIDS, both among IDUs and among sexually infected individuals, remains to be monitored in future.

MultiHIV antigen design (Ref III)

The leading questions in selecting the immunogens for effective and broadly active HIV-1 vaccine concern the choice of antigen sequences and the HIV genomes from which the antigen sequences ought to be selected.

For a vaccine designed to elicit cellular immunity, the broadest responses would be expected from immunogens that encode the largest number of conserved T cell epitopes, both to maximize the strength of response and to minimize the probability of viral escape. Incorporating highly conserved proteins in a vaccine considerably improves the likelihood of cross-clade efficacy, and as Gag and Pol are the most highly conserved HIV proteins and as they contain the largest number of defined epitopes, they have believed being the best candidate antigens (De Groot et al., 2003, Finnefrock et al., 2007, Korber et al., 2001). In addition, more variable Nef, Tat, and Rev, have been suggested as vaccine antigen candidates, as they contribute substantially to the immune response in many HIV-infected patients (especially Nef) and, importantly, they are early expressed within the viral lifecycle (Betts et al., 2002). On the other hand, among all HIV vaccine candidates evaluated, live attenuated vaccines which express the majority of viral proteins have been the most successful, indicating that multiple viral proteins might still be required as antigens for effective HIV vaccine (Mooij & Heeney, 2001). Furthermore, HIV-1-specific $CD4^+$ and $CD8^+$ T cell responses have been shown to recognize peptides from every HIV protein, thus all HIV proteins may contribute to control of HIV replication (Betts et al., 2002, Korber et al., 2001).

In the current study, four different prototype clade/ancestor MultiHIV antigens were designed based on sequence data of HIV-1 A, B, C, F, G, and H subtypes. These antigens were named as MultiHIV-A (based on subtype A consensus sequence), MultiHIV-B (based on subtype B consensus sequence), MultiHIV-C (based on subtype C consensus sequence) and MultiHIV-FGH (based on ancestral sequence for subtypes F, G and H), respectively. All four MultiHIV antigens were designed by fusing regulatory early viral proteins Rev, Nef, and Tat, as well as structural proteins Gag_{p17} and Gag_{p24} (see Figure 10) with similar structure and size of 1053–1079 aa.

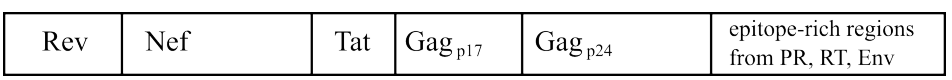

Figure 10. Structure of the MultiHIV antigens.

In addition, with the aim to maximize the coverage of potential CTL epitopes, nine human T cell epitope-rich regions from PR (1 region), RT (6 regions) and envelope proteins Env_{gp120} (1 region) and Env_{gp41} (1 region) with length 17–46 aa were selected from *The 2001 Immunology Compendium*, Brander & Goulder. It is

important to note, that all epitopes included in this database represent experimentally mapped, not bioinformatically predicted ones. Thereafter, artificial protein sequence was generated by consecutive incorporation of the selected regions and this fragment was incorporated into the C-terminal end of the MultiHIV antigen.

In addition, two dominant CTL epitopes were included into this fragment for potency assays in the mouse and macaque systems, respectively, see Figure 11.

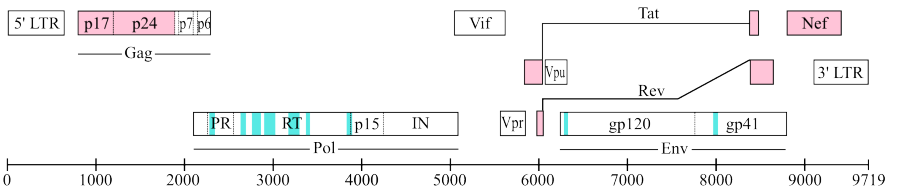

Figure 11. Genomic regions of the HIV-1 genome included into the artificial multigen are colored pink (Rev, Nef, Tat, p17, and p24); T-cell epitope rich regions included into the multigen are coloured blue

After designing general constitution of the MultiHIV antigen, bioinformatical manipulations were performed to confront the extensive variability present in global HIV-1 population:

To reduce the degree of sequence variation between a vaccine strain and circulating viruses, all HIV-1 isolate sequences of particular subtypes listed in HIV database http://www.hiv.lanl.gov in August 2002, were taken into account and synthetic viral protein sequences were created using either subtype consensus or ancestral lineages. The consensus sequence of a subtype is generally biased towards a subset of lineages that are sampled from the region where the subtype predominates. A, B, and C are all major subtypes with distinct subsets and geographical distribution patterns. For any one of these, the consensus sequence has less mutational differences from an average individual isolate within the group than the ancestral sequence of the subtype. However, in case of minor subtypes, like F, G and H with relatively small number of representatives, each deriving from different geographic location, consensus approach could introduce a bias towards specific subtype consensus that is represented by more than 50% of the sequences. For that reason, the FGH ancestral sequence was constructed separately for every amino acid sequence segment of HIV-1 proteins from all available isolates belonging to F, G, and H subtypes using maximum likelihood program *proml* from the software package PHYLIP.

As a next step, the epitope rich areas of created consensus/ancestor sequences were compared with the isolates of particular subtypes. It appeared that the constructed sequences were complementing each other in a window of the assumed size of the T cell epitopes (20 aa). The average of maximum hits from the four sequences was always higher than the hit rate from one particular

sequence. However, over the stretch of many 20 aa windows one or two consensus/ancestral sequences showed identity with each other. Trying to economize this redundancy, an additional step of optimization was taken: in one of the two subtype consensus sequences identical to each other in the sliding window of 20 aa, the redundant amino acid was replaced with second frequent one. The change was introduced into the subtype consensus, which showed higher incidence of the alternative variant in a given position in the HIV-1 database. By the end of this optimization and compilation every 20 aa window showed on average a 94.8% identity with the isolates from the database. Subtype B showed the highest average identity (96.1%), followed by subtypes C (95.2%), A (94.7%), and FGH (93.2%).

After designing the protein sequences of MultiHIV antigens, the presence of known CTL epitopes in the MultHIV antigens was analyzed. First, the antigen sequences were compared manually with the list of the best-defined HIV-derived CTL epitopes taken from *The 2001 Immunology Compendium*, Brander & Goulder; also the status of these 105 epitopes included into the MultiHIV antigens was checked. It was concluded that in most cases the defined epitopes were identical or very similar to motifs present in our antigens and the anchor amino acids were only rarely affected.

For viral proteins to be recognized by CTLs they must first be cleaved into short peptides; this step occurs in the cytosol and is due to the immuno-proteasome (Murphy, 2008). Cleavage sites generated by immunoproteasome are sensitive to the surrounding amino acid sequences and predictable by neural network approach NetChop (Kesmir et al., 2002). Thus, the created MultiHIV antigen sequences were also analyzed by NetChop prediction server (http://www.cbs.dtu.dk/services/NetChop). When analyzed separately, approxi-mately 60% of C-terminal ends of the best defined epitopes were predicted as immunoproteasome cleavage sites. However, as the processing of an epitope was often predicted for only one or two antigens out of four, the simultaneous inclusion of all antigens resulted in predicting even more than 70% C-terminal ends of 105 best defined epitopes.

Generation and analysis of the MultiHIV-encoding GTU DNA vaccine vectors (Ref III)

The DNA coding sequences of all four MultiHIV antigens were designed by backtranslation of the amino acid sequences, followed by optimization steps resulting in human codon-optimized cDNAs with appropriate start and stop codons and Kozak consensus motifs. Unwanted RNA secondary structure and instability elements as well as cryptic splice sites were eliminated by synony-mous replacements. Finally, the cDNAs were synthesized by combining chemical oligonucleotide synthesis, PCR and recombinant DNA techniques.

Subsequently, the created cDNAs were cloned into GTU® DNA vaccine vector. The GTU® vector has several unique features while compared to

ordinary DNA plasmid vectors. First, the vector expresses nuclear anchoring E2 protein of bovine papillomavirus type 1 that together with E2 binding sites present in the GTU plasmid provide segregation/partitioning function to the GTU®, therefore assuring the maintenance of the plasmid in the nuclei of dividing cell population (Abroi et al., 2004, Baars et al., 2003, McPhillips et al., 2005, You et al., 2004, You et al., 2005). Second, the E2 protein and E2 binding sites are strongly acting as transcription enhancer of the promoter driving the expression of the gene antigen (Baars et al., 2003, Kurg et al., 2006). As an outcome, different promoters become up to 100 times stronger in GTU® context compared to the regular vectors (Blazevic et al., 2006).

The expression studies in cell culture system showed that MultiHIV fusion proteins were all expressed as full-length proteins at high expression levels in the cytoplasm of the transfected cells (Martinon et al., 2009). In addition, the MultiHIV-B expression was studied *in vivo* in pig skin, which is standard biomedical animal model of human skin, especially suitable for studies on DNA uptake and gene expression (Hengge et al., 1996). The intra-dermal injection resulted in significant expression of the MultiHIV antigen, mainly in cytoplasm of epidermal keratinocytes. Furthermore, similar results were also obtained in macaque skin. However, intra-dermal injection of macaques with the plasmid without electroporation resulted in antigen expression only in the dermis; electroporation substantially increased antigen expression level in the epidermis (Martinon et al., 2009).

GTU-MultiHIV DNA vaccine-induced CTL-response in a heterogenous domestic pig population (Ref III)

The immunogenicity of GTU-MultiHIV antigen has previously been shown in mice. However, homozygous inbred animals are not suitable models for analyzing the breadth of the T-cell response due to limited number of different MHC class I antigens on the cells, leading to a limited number of presented epitopes to T-cells. As a result, mouse immunisation with vectors encoding MultiHIV antigen elicited strong CTL response mainly against only 2 epitopes, both described as potent epitopes in Balb/c mice (Blazevic et al., 2006). In order to be able to analyze the breadth of the induced T-cell response, we started to use domestic pigs as model animals, instead of mice. Domestic pig is considered to be a useful test animal for biomedical research (mainly transplantation) due to the similarity of its organ size and body weight as well as cardiovascular system structure, to those in humans (Tanaka-Matsuda et al., 2009). In addition, domestic swine are usually outbred animals, and the swine MHC (SLA) complex, which plays a central role in the presentation of antigenic peptides to T cells, has been shown to be remarkably similar to human MHC (HLA). We took advantage of these features and analyzed the immunogenicity profile of our DNA vaccine using domestic swine as test animal.

To assess the genome diversity of the SLA class I loci, all five immunized study animals were SLA-I- genotyped, see Table 3.

SLA typing revealed considerably high heterogeneity among study animals: all swine had a unique SLA class I allelic combination, with 16 different alleles. Five of the SLA-I alleles found in the local pigs were novel, while eleven of the alleles have been previously described in other pig breeds. Two alleles were shared by two animals; all other alleles were present in a single animal. All of the putative new alleles were different from each other. Among five animals, three were heterozygous for one or two loci; remaining two were homozygous for all three loci. However, in a case of only one allele detected, a possibility of the existence of another allele on the other chromosome cannot be excluded. Five alleles (SLA-1*1501, SLA-1*0401, SLA-1*1201, SLA-2*1001, and SLA-3*0401) have been described in Large White/Landrace before; three alleles have been identified in porcine cell line ESK4 (SLA-1*es11, SLA-1*es12, SLA-2*es22) and three alleles (SLA-3*03an04 and SLA-3*0501, and SLA-2*w06sn01) have been found in other pig breeds.

Table 3. SLA-I-typing results in five vaccinated animals.

Pig ID	Locus	(Allelic) designation
53–14	SLA-1	SLA-1*1501
	SLA-2	SLA-2*1001
	SLA-3	SLA-3*03an04, new allele
42–09	SLA-1	SLA-1*0401
	SLA-2	new allele
	SLA-3	SLA-3*0401
38–07	SLA-1	SLA-1*es12; SLA-1*es11
	SLA-2	SLA-2*es22; SLA-2*w06sn01
	SLA-3	not determined
37–14	SLA-1	SLA-1*1201
	SLA-2	SLA-2*1001
	SLA-3	new allele
37–11	SLA-1	new allele
	SLA-2	new allele
	SLA-3	SLA-3*0501; SLA-3*03an04

Using swine as test animals we were able to show that the CTL response against MultiHIV antigen is directed against all constituents of the MultiHIV antigen, in which detailed epitope mapping was done. The results show that the epitopes recognized in pigs lie in the same antigenic regions as recognized in humans, indicating that immunogenicity profile of vaccines in domestic swine may predict the outcome of human immunisation.

SUMMARY

Firstly, we characterized HIV-1 genetic diversity in Estonian population since the beginning of the epidemic in August 2000 when HIV-1 incidence grew explosively among Estonian intravenous drug users (IDUs). Phylogenetic analysis revealed that a rare circulating recombinant form (CRF), CRF06_cpx was predominant in population of Estonian HIV-1 positive individuals, comprising 76% of all circulating viruses. Additionally, high incidence (17%) of different unique recombinant forms (URFs) was detected, consisting of stretches of subtypes A and CRF06_cpx, indicating that co-infection with different strains of HIV-1 occurs frequently among drug users. The fact that one particular type comprises one third of all URFs in Estonia, allowed us to designate this as the emergence of a new CRF, namely CRF32_06A1.

Secondly, we studied HIV-1 co-receptor, chemokine receptor 5 (CCR5), polymorphism among Estonian HIV-1 patients. In literature, CCR5Δ32 heterozygosity has been linked with delayed onset of AIDS and reduced risk of initial transmission. In Estonia the frequency of the HIV-1 restrictive CCR5Δ32 allele has been found to be among the highest in the world. Nevertheless, Estonia has a very high HIV-1 incidence, allowing us, therefore, to follow the CCR5 polymorphism in a Caucasian population with a high frequency of the CCR5Δ32 allele, at the high background of HIV-1 incidence. We found that CCR5Δ32 heterozygosity occurs with the same frequency among HIV-1 seropositives as in healthy control population. It suggests that in case of IDU associated epidemics, CCR5Δ32 heterozygosity does not confer protection from HIV-1 infection. CCR5Δ32 allele frequencies among sexually infected patients *versus* IDUs did not reveal statistically significant differences, thus indicating that this was not a risk-group specific phenomenon. We also analyzed, whether CCR5Δ32 heterozygotes harbor URFs of HIV-1 (indicative of dual infection) with the same frequency as CCR5WT homozygotes. The statistical analysis confirmed equal distribution of URFs both among CCR5Δ32 heterozygotes as well as among CCR5WT homozygotes. Thus, our data did not reveal any association between CCR5Δ32 genotype status and frequency of dual infection.

Thirdly, four different prototype clade/ancestor MultiHIV antigens were designed based on sequence data of HIV-1 A, B, C, F, G, and H subtypes. All four MultiHIV antigens were designed by fusing regulatory early viral proteins Rev, Nef, and Tat; structural proteins Gag $_{p17}$ and Gag $_{p24}$, as well as human T cell epitope-rich regions from PR, RT, and envelope proteins Env_{gp120} and $Env_{gp41,}$ with similar structure and size. We showed a successful induction of broad cellular anti-HIV immune response with a MultiHIV-B consensus antigen and demonstrated that the CTL response was directed against all constituents of the MultiHIV antigen in domestic pig. To assess the genome diversity of the SLA class I loci, all immunized study animals were SLA-I- genotyped. SLA typing revealed considerably high heterogeneity among study animals: all swine had a unique SLA class I allelic combination. We thus demonstrated that pig is an excellent test animal for studying T cell responses in immunological surveys, *e.g.* in vaccine evaluation studies.

REFERENCES

Abecasis, A. B., Lemey, P., Vidal, N., Oliveira, T., Peeters, M., Camacho, R., Shapiro, B., Rambaut, A., Vandamme, A.-M. (2007). Recombination Confounds the Early Evolutionary History of Human Immunodeficiency Virus Type 1: Subtype G Is a Circulating Recombinant Form. *J Virol* **81**, 8543–8551.

Abraha, A., Nankya, I. L., Gibson, R., Demers, K., Tebit, D. M., Johnston, E., Katzenstein, D., Siddiqui, A., Herrera, C., Fischetti, L., Shattock, R. J. & Arts, E. J. (2009). CCR5- and CXCR4-tropic subtype C human immunodeficiency virus type 1 isolates have a lower level of pathogenic fitness than other dominant group M subtypes: implications for the epidemic. *J Virol* **83**, 5592–605.

Abroi, A., Ilves, I., Kivi, S. & Ustav, M. (2004). Analysis of chromatin attachment and partitioning functions of bovine papillomavirus type 1 E2 protein. *J Virol* **78**, 2100–13.

Ando, A., Kawata, H., Shigenari, A., Anzai, T., Ota, M., Katsuyama, Y., Sada, M., Goto, R., Takeshima, S. N., Aida, Y., Iwanaga, T., Fujimura, N., Suzuki, Y., Gojobori, T. & Inoko, H. (2003). Genetic polymorphism of the swine major histocompatibility complex (SLA) class I genes, SLA-1, -2 and -3. *Immunogenetics* **55**, 583–93.

Arenzana-Seisdedos, F. & Parmentier, M. (2006). Genetics of resistance to HIV infection: Role of co-receptors and co-receptor ligands. *Semin Immunol* **18**, 387–403.

Ariën, K. K., Vanham, G. & Arts, E. J. (2007). Is HIV-1 evolving to a less virulent form in humans? *Nat Rev Microbiol* **5**, 141–51.

Ariën, K. K. & Verhasselt, B. (2008). HIV Nef: role in pathogenesis and viral fitness. *Curr HIV Res* **6**, 200–8.

Baars, S., Bachmann, A., Levitzki, A. & Rosl, F. (2003). Tyrphostin AG 555 inhibits bovine papillomavirus transcription by changing the ratio between E2 transactivator/repressor function. *J Biol Chem* **278**, 37306–13.

Balanovsky, O., Pocheshkhova, E., Pshenichnov, A., Solovieva, D., Kuznetsova, M., Voronko, O., Churnosov, M., Tegako, O., Atramentova, L., Lavryashina, M., Evseeva, I., Borinska, S., Boldyreva, M., Dubova, N. & Balanovska, E. (2005). Is spatial distribution of the HIV-1-resistant CCR5Delta32 allele formed by ecological factors? *J Physiol Anthropol Appl Human Sci* **24**, 375–82.

Barouch, D. H. (2008). Challenges in the development of an HIV-1 vaccine. *Nature* **455**, 613–9.

Betts, M. R., Yusim, K. & Koup, R. A. (2002). Optimal antigens for HIV vaccines based on CD8+ T response, protein length, and sequence variability. *DNA Cell Biol* **21**, 665–70.

Blackwell, J. M., Jamieson, S. E. & Burgner, D. (2009). HLA and infectious diseases. *Clin Microbiol Rev* **22**, 370–85, Table of Contents.

Blazevic, V., Männik, A., Malm, M., Sikut, R., Valtavaara, M., Toots, U., Ustav, M. & Krohn, K. (2006). Induction of human immunodeficiency virus type-1-specific immunity with a novel gene transport unit (GTU)-MultiHIV DNA vaccine. *AIDS Res Hum Retroviruses* **22**, 667–77.

Brander, C. & Goulder, P. (2001). The Identification of Optimal HIV-Derived CTL Epitopes in Diverse Populations Using HIV Clade-Specific Consensus. In *HIV Molecular Immunology 2001*. Edited by B. C. Korber B, Haynes B, Koup R, Kuiken C, Moore J, Walker B, and Watkins D. Los Alamos: Los Alamos National Laboratory, Theoretical Biology and Biophysics.

Buchbinder, S. P., Mehrotra, D. V., Duerr, A., Fitzgerald, D. W., Mogg, R., Li, D., Gilbert, P. B., Lama, J. R., Marmor, M., Del Rio, C., McElrath, M. J., Casimiro, D. R., Gottesdiener, K. M., Chodakewitz, J. A., Corey, L. & Robertson, M. N. (2008). Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. *Lancet* **372**, 1881–93.

Buonaguro, L., Tornesello, M. L. & Buonaguro, F. M. (2007). Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. *J Virol* **81**, 10209–19.

Burgers, W. A., van Harmelen, J. H., Shephard, E., Adams, C., Mgwebi, T., Bourn, W., Hanke, T., Williamson, A. L. & Williamson, C. (2006). Design and preclinical evaluation of a multigene human immunodeficiency virus type 1 subtype C DNA vaccine for clinical trial. *J Gen Virol* **87**, 399–410.

Campbell, G. R. & Loret, E. P. (2009). What does the structure-function relationship of the HIV-1 Tat protein teach us about developing an AIDS vaccine? *Retrovirology* **6**, 50.

Chiu, I. M., Yaniv, A., Dahlberg, J. E., Gazit, A., Skuntz, S. F., Tronick, S. R. & Aaronson, S. A. (1985). Nucleotide sequence evidence for relationship of AIDS retrovirus to lentiviruses. *Nature* **317**, 366–8.

Cilliers, T., Nhlapo, J., Coetzer, M., Orlovic, D., Ketas, T., Olson, W. C., Moore, J. P., Trkola, A. & Morris, L. (2003). The CCR5 and CXCR4 coreceptors are both used by human immunodeficiency virus type 1 primary isolates from subtype C. *J Virol* **77**, 4449–56.

Ciuffi, A., Bleiber, G., Munoz, M., Martinez, R., Loeuillet, C., Rehr, M., Fischer, M., Gunthard, H. F., Oxenius, A., Meylan, P., Bonhoeffer, S., Trono, D. & Telenti, A. (2004). Entry and transcription as key determinants of differences in CD4 T-cell permissiveness to human immunodeficiency virus type 1 infection. *J Virol* **78**, 10747–54.

Clavel, F., Hoggan, M. D., Willey, R. L., Strebel, K., Martin, M. A. & Repaske, R. (1989). Genetic recombination of human immunodeficiency virus. *J Virol* **63**, 1455–9.

De Groot, A. S., Jesdale, B., Martin, W., Saint Aubin, C., Sbai, H., Bosma, A., Lieberman, J., Skowron, G., Mansourati, F. & Mayer, K. H. (2003). Mapping cross-clade HIV-1 vaccine epitopes using a bioinformatics approach. *Vaccine* **21**, 4486–504.

Dean, M., Carrington, M., Winkler, C., Huttley, G. A., Smith, M. W., Allikmets, R., Goedert, J. J., Buchbinder, S. P., Vittinghoff, E., Gomperts, E., Donfield, S., Vlahov, D., Kaslow, R., Saah, A., Rinaldo, C., Detels, R. & O'Brien, S. J. (1996). Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science* **273**, 1856–62.

Deng, H., Liu, R., Ellmeier, W., Choe, S., Unutmaz, D., Burkhart, M., Di Marzio, P., Marmon, S., Sutton, R. E., Hill, C. M., Davis, C. B., Peiper, S. C., Schall, T. J., Littman, D. R. & Landau, N. R. (1996). Identification of a major co-receptor for primary isolates of HIV-1. *Nature* **381**, 661–6.

Douek, D. C., Kwong, P. D. & Nabel, G. J. (2006). The rational design of an AIDS vaccine. *Cell* **124**, 677–81.

Dragic, T., Litwin, V., Allaway, G. P., Martin, S. R., Huang, Y., Nagashima, K. A., Cayanan, C., Maddon, P. J., Koup, R. A., Moore, J. P. & Paxton, W. A. (1996).

HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* **381**, 667–73.

Ellenberger, D. L., Li, B., Lupo, L. D., Owen, S. M., Nkengasong, J., Kadio-Morokro, M. S., Smith, J., Robinson, H., Ackers, M., Greenberg, A., Folks, T. & Butera, S. (2002). Generation of a consensus sequence from prevalent and incident HIV-1 infections in West Africa to guide AIDS vaccine development. *Virology* **302**, 155–63.

Finnefrock, A. C., Liu, X., Opalka, D. W., Shiver, J. W., Casimiro, D. R. & Condra, J. H. (2007). HIV type 1 vaccines for worldwide use: predicting in-clade and cross-clade breadth of immune responses. *AIDS Res Hum Retroviruses* **23**, 1283–92.

Flynn, N. M., Forthal, D. N., Harro, C. D., Judson, F. N., Mayer, K. H. & Para, M. F. (2005). Placebo-controlled phase 3 trial of a recombinant glycoprotein 120 vaccine to prevent HIV-1 infection. *J Infect Dis* **191**, 654–65.

Galvani, A. P. & Novembre, J. (2005). The evolutionary history of the CCR5-Delta32 HIV-resistance mutation. *Microbes Infect* **7**, 302–9.

Gao, F., Chen, Y., Levy, D. N., Conway, J. A., Kepler, T. B. & Hui, H. (2004). Unselected mutations in the human immunodeficiency virus type 1 genome are mostly nonsynonymous and often deleterious. *J Virol* **78**, 2426–33.

Gaschen, B., Taylor, J., Yusim, K., Foley, B., Gao, F., Lang, D., Novitsky, V., Haynes, B., Hahn, B. H., Bhattacharya, T. & Korber, B. (2002). Diversity considerations in HIV-1 vaccine selection. *Science* **296**, 2354–60.

Gilbert, P. B., Peterson, M. L., Follmann, D., Hudgens, M. G., Francis, D. P., Gurwith, M., Heyward, W. L., Jobes, D. V., Popovic, V., Self, S. G., Sinangil, F., Burke, D. & Berman, P. W. (2005). Correlation between immunologic responses to a recombinant glycoprotein 120 vaccine and incidence of HIV-1 infection in a phase 3 HIV-1 preventive vaccine trial. *J Infect Dis* **191**, 666–77.

Goulder, P. J. & Watkins, D. I. (2008). Impact of MHC class I diversity on immune control of immunodeficiency virus replication. *Nat Rev Immunol* **8**, 619–30.

Göttlinger, H. (2001). HIV-1 Gag: a Molecular Machine Driving Viral Particle Assembly and Release. In *HIV Sequence Compendium 2001*, pp. pp2–28. Edited by F. B. Kuiken C, Hahn B, Marx P, McCutchan F, Mellors JW, Wolinsky S, Korber B. Los Alamos: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory.

Hahn, B. H., Shaw, G. M., De Cock, K. M. & Sharp, P. M. (2000). AIDS as a zoonosis: scientific and public health implications. *Science* **287**, 607–14.

Haigwood, N. L. & Hirsch, V. M. (2009). Blocking and tackling HIV. *Nat Med* **15**, 841–2.

Hamers, F. F. & Downs, A. M. (2003). HIV in central and eastern Europe. *Lancet* **361**, 1035–44.

Hengge, U. R., Walker, P. S. & Vogel, J. C. (1996). Expression of naked DNA in human, pig, and mouse skin. *J Clin Invest* **97**, 2911–6.

Hessell, A. J., Poignard, P., Hunter, M., Hangartner, L., Tehrani, D. M., Bleeker, W. K., Parren, P. W., Marx, P. A. & Burton, D. R. (2009). Effective, low-titer antibody protection against low-dose repeated mucosal SHIV challenge in macaques. *Nat Med* **15**, 951–4.

Heuverswyn, F., Peeters, M. (2007) The Origins of HIV and Implications for the Global Epidemic. *Curr Inf Dis Rep* **9**, 338–346.

Hill, M., Tachedjian, G. & Mak, J. (2005). The packaging and maturation of the HIV-1 Pol proteins. *Curr HIV Res* **3**, 73–85.

Ho, C. S., Lunney, J. K., Franzo-Romain, M. H., Martens, G. W., Lee, Y. J., Lee, J. H., Wysocki, M., Rowland, R. R. & Smith, D. M. (2009). Molecular characterization of swine leucocyte antigen class I genes in outbred pig populations. *Anim Genet* **40**, 468–78.

Jetzt, A. E., Yu, H., Klarmann, G. J., Ron, Y., Preston, B. D. & Dougherty, J. P. (2000). High rate of recombination throughout the human immunodeficiency virus type 1 genome. *J Virol* **74**, 1234–40.

Johnson, P. R., Schnepp, B. C., Zhang, J., Connell, M. J., Greene, S. M., Yuste, E., Desrosiers, R. C. & Reed Clark, K. (2009). Vector-mediated gene transfer engenders long-lived neutralizing activity and protection against SIV infection in monkeys. *Nat Med* **15**, 901–6.

Jung, A., Maier, R., Vartanian, J. P., Bocharov, G., Jung, V., Fischer, U., Meese, E., Wain-Hobson, S. & Meyerhans, A. (2002). Multiply infected spleen cells in HIV patients. *Nature* **418**, 144.

Kalev, I., Mikelsaar, A. V., Beckman, L., Tasa, G. & Pärlist, P. (2000). High frequency of the HIV-1 protective CCR5 delta32 deletion in native Estonians. *Eur J Epidemiol* **16**, 1107–9.

Karn, J. (2000). Tat, a novel regulator of HIV transcription and latency. In *HIV Sequence Compendium 2000*, pp. pp. 2–18. Edited by M. F. Kuiken C, Foley B, Mellors JW, Hahn B, Mullins J, Marx P, Wolinsky S. Los Alamos: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory.

Kawashima, Y., Pfafferott, K., Frater, J., Matthews, P., Payne, R., Addo, M., Gatanaga, H., Fujiwara, M., Hachiya, A., Koizumi, H., Kuse, N., Oka, S., Duda, A., Prendergast, A., Crawford, H., Leslie, A., Brumme, Z., Brumme, C., Allen, T., Brander, C., Kaslow, R., Tang, J., Hunter, E., Allen, S., Mulenga, J., Branch, S., Roach, T., John, M., Mallal, S., Ogwu, A., Shapiro, R., Prado, J. G., Fidler, S., Weber, J., Pybus, O. G., Klenerman, P., Ndung'u, T., Phillips, R., Heckerman, D., Harrigan, P. R., Walker, B. D., Takiguchi, M. & Goulder, P. (2009). Adaptation of HIV-1 to human leukocyte antigen class I. *Nature* **458**, 641–5.

Kesmir, C., Nussbaum, A. K., Schild, H., Detours, V. & Brunak, S. (2002). Prediction of proteasome cleavage motifs by neural networks. *Protein Eng* **15**, 287–96.

Klatzmann, D., Champagne, E., Chamaret, S., Gruest, J., Guetard, D., Hercend, T., Gluckman, J. C. & Montagnier, L. (1984). T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* **312**, 767–8.

Korber, B., Gaschen, B., Yusim, K., Thakallapally, R., Kesmir, C. & Detours, V. (2001). Evolutionary and immunological implications of contemporary HIV-1 variation. *Br Med Bull* **58**, 19–42.

Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B. H., Wolinsky, S. & Bhattacharya, T. (2000). Timing the ancestor of the HIV-1 pandemic strains. *Science* **288**, 1789–96.

Kothe, D. L., Decker, J. M., Li, Y., Weng, Z., Bibollet-Ruche, F., Zammit, K. P., Salazar, M. G., Chen, Y., Salazar-Gonzalez, J. F., Moldoveanu, Z., Mestecky, J., Gao, F., Haynes, B. F., Shaw, G. M., Muldoon, M., Korber, B. T. & Hahn, B. H. (2007). Antigenicity and immunogenicity of HIV-1 consensus subtype B envelope glycoproteins. *Virology* **360**, 218–34.

Kurg, R., Tekkel, H., Abroi, A. & Ustav, M. (2006). Characterization of the functional activities of the bovine papillomavirus type 1 E2 protein single-chain heterodimers. *J Virol* **80**, 11218–25.

Kutzler, M. A. & Weiner, D. B. (2008). DNA vaccines: ready for prime time? *Nat Rev Genet* **9**, 776–88.

Lai, T., Rätsep, M., Rüütel, K., Trummal, A., Kahur, K., Nielsen, S., Lauer, J., Politi, C. & Habicht, J. (2009). Modelling Estonia's concentrated HIV epidemic: Estonian National Institute of Health Development, Estonian Ministry of Social Affairs, WHO Regional Office for Europe.

Lancet (2009). A (prime) boost for HIV vaccine research? *Lancet* **374**, 1119.

Letvin, N. L. (2006). Progress and obstacles in the development of an AIDS vaccine. *Nat Rev Immunol* **6**, 930–9.

Li, L., Li, H. S., Pauza, C. D., Bukrinsky, M. & Zhao, R. Y. (2005). Roles of HIV-1 auxiliary proteins in viral pathogenesis and host-pathogen interactions. *Cell Res* **15**, 923–34.

Liitsola, K., Laukkanen, T., Denisova, A., Grishkevichius, A., Smolskaja, T., Ustina, V., Vlasov, N., Leinikki, P. & Salminen, M. O. (1996). Genetic characterization of HIV-1 strains in the Baltic countries and Russia. *Scand J Infect Dis* **28**, 537–41.

Limborska, S. A., Balanovsky, O. P., Balanovskaya, E. V., Slominsky, P. A., Schadrina, M. I., Livshits, L. A., Kravchenko, S. A., Pampuha, V. M., Khusnutdinova, E. K. & Spitsyn, V. A. (2002). Analysis of CCR5Delta32 geographic distribution and its correlation with some climatic and geographic factors. *Hum Hered* **53**, 49–54.

Liu, R., Paxton, W. A., Choe, S., Ceradini, D., Martin, S. R., Horuk, R., MacDonald, M. E., Stuhlmann, H., Koup, R. A. & Landau, N. R. (1996). Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* **86**, 367–77.

Lodish H, B. A., Matsudaira, Kaiser C A, Krieger, M, Scott, M P, Zipursky, L S, Darnell J (2004). Molecular Cell Biology, fifth edn. New York: Freeman

Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., Ingersoll, R., Sheppard, H. W. & Ray, S. C. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* **73**, 152–60.

Lu, Y., Nerurkar, V. R., Dashwood, W. M., Woodward, C. L., Ablan, S., Shikuma, C. M., Grandinetti, A., Chang, H., Nguyen, H. T., Wu, Z., Yamamura, Y., Boto, W. O., Merriwether, A., Kurata, T., Detels, R. & Yanagihara, R. (1999). Genotype and allele frequency of a 32-base pair deletion mutation in the CCR5 gene in various ethnic groups: absence of mutation among Asians and Pacific Islanders. *Int J Infect Dis* **3**, 186–91.

Lukashov, V. V., Kuiken, C. L., Boer, K. & Goudsmit, J. (1996). HIV type 1 subtypes in The Netherlands circulating among women originating from AIDS-endemic regions. *AIDS Res Hum Retroviruses* **12**, 951–3.

Lunney, J. K., Ho, C. S., Wysocki, M. & Smith, D. M. (2009). Molecular genetics of the swine major histocompatibility complex, the SLA complex. *Dev Comp Immunol* **33**, 362–74.

Lynch, R. M., Shen, T., Gnanakaran, S. & Derdeyn, C. A. (2009). Appreciating HIV type 1 diversity: subtype differences in Env. *AIDS Res Hum Retroviruses* **25**, 237–48.

Malm, M., Rollman, E., Ustav, M., Hinkula, J., Krohn, K., Wahren, B. & Blazevic, V. (2005). Cross-clade protection induced by human immunodeficiency virus-1 DNA immunogens expressing consensus sequences of multiple genes and epitopes from subtypes A, B, C, and FGH. *Viral Immunol* **18**, 678–88.

Marmor, M., Sheppard, H. W., Donnell, D., Bozeman, S., Celum, C., Buchbinder, S., Koblin, B. & Seage, G. R., 3rd (2001). Homozygous and heterozygous CCR5-Delta32 genotypes are associated with resistance to HIV infection. *J Acquir Immune Defic Syndr* **27**, 472–81.

Martinon, F., Kaldma, K., Sikut, R., Culina, S., Romain, G., Tuomela, M., Adojaan, M., Männik, A., Toots, U., Kivisild, T., Morin, J., Brochard, P., Delache, B., Tripiciano, A., Ensoli, F., Stanescu, I., Le Grand, R. & Ustav, M. (2009). Persistent Immune Responses Induced by a HIV DNA Vaccine Delivered in Association with Electroporation in the Skin of Nonhuman Primates. *Hum Gene Ther*.

Martinson, J. J., Chapman, N. H., Rees, D. C., Liu, Y. T. & Clegg, J. B. (1997). Global distribution of the CCR5 gene 32-basepair deletion. *Nat Genet* **16**, 100–3.

Mascola, J. R., Stiegler, G., VanCott, T. C., Katinger, H., Carpenter, C. B., Hanson, C. E., Beary, H., Hayes, D., Frankel, S. S., Birx, D. L. & Lewis, M. G. (2000). Protection of macaques against vaginal transmission of a pathogenic HIV-1/SIV chimeric virus by passive infusion of neutralizing antibodies. *Nat Med* **6**, 207–10.

McBurney, S. P. & Ross, T. M. (2008). Viral sequence diversity: challenges for AIDS vaccine designs. *Expert Rev Vaccines* **7**, 1405–17.

McNicholl, J. M., Smith, D. K., Qari, S. H. & Hodge, T. (1997). Host genes and HIV: the role of the chemokine receptor gene CCR5 and its allele. *Emerg Infect Dis* **3**, 261–71.

McPhillips, M. G., Ozato, K. & McBride, A. A. (2005). Interaction of bovine papillomavirus E2 protein with Brd4 stabilizes its association with chromatin. *J Virol* **79**, 8920–32.

Meyer, L., Magierowska, M., Hubert, J. B., Rouzioux, C., Deveau, C., Sanson, F., Debre, P., Delfraissy, J. F. & Theodorou, I. (1997). Early protective effect of CCR-5 delta 32 heterozygosity on HIV-1 disease progression: relationship with viral load. The SEROCO Study Group. *Aids* **11**, F73–8.

Montavon, C., Bibollet-Ruche, F., Robertson, D., Koumare, B., Mulanga, C., Esu-Williams, E., Toure, C., Mboup, S., Saman, E., Delaporte, E. & Peeters, M. (1999). The identification of a complex A/G/I/J recombinant HIV type 1 virus in various West African countries. *AIDS Res Hum Retroviruses* **15**, 1707–12.

Mooij, P. & Heeney, J. L. (2001). Rational development of prophylactic HIV vaccines based on structural and regulatory proteins. *Vaccine* **20**, 304–21.

Moore, C. B., John, M., James, I. R., Christiansen, F. T., Witt, C. S. & Mallal, S. A. (2002). Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science* **296**, 1439–43.

Moore, J. P., Kitchen, S. G., Pugach, P. & Zack, J. A. (2004). The CCR5 and CXCR4 coreceptors--central to understanding the transmission and pathogenesis of human immunodeficiency virus type 1 infection. *AIDS Res Hum Retroviruses* **20**, 111–26.

Mulherin, S. A., O'Brien, T. R., Ioannidis, J. P., Goedert, J. J., Buchbinder, S. P., Coutinho, R. A., Jamieson, B. D., Meyer, L., Michael, N. L., Pantaleo, G., Rizzardi, G. P., Schuitemaker, H., Sheppard, H. W., Theodorou, I. D., Vlahov, D. & Rosenberg, P. S. (2003). Effects of CCR5-Delta32 and CCR2–64I alleles on HIV-1 disease progression: the protection varies with duration of infection. *Aids* **17**, 377–87.

Murphy, K., Travers, P, Walport, M (2008). Janeway's Immunobiology. New York: Garland Science.

Naganawa, S., Sato, S., Nossik, D., Takahashi, K., Hara, T., Tochikubo, O., Kitamura, K., Honda, M. & Nakasone, T. (2002). First report of CRF03_AB recombinant HIV type 1 in injecting drug users in Ukraine. *AIDS Res Hum Retroviruses* **18**, 1145–9.

Nazari, R. & Joshi, S. (2008). CCR5 as target for HIV-1 gene therapy. *Curr Gene Ther* **8**, 264–72.

Nickle, D. C., Jensen, M. A., Gottlieb, G. S., Shriner, D., Learn, G. H., Rodrigo, A. G. & Mullins, J. I. (2003). Consensus and ancestral state HIV vaccines. *Science* **299**, 1515–8; author reply 1515–8.

Nomaguchi, M., Fujita, M. & Adachi, A. (2008). Role of HIV-1 Vpu protein for virus spread and pathogenesis. *Microbes Infect* **10**, 960–7.

Novembre, J., Galvani, A. P. & Slatkin, M. (2005). The geographic spread of the CCR5 Delta32 HIV-resistance allele. *PLoS Biol* **3**, e339.

Novitsky, V., Smith, U. R., Gilbert, P., McLane, M. F., Chigwedere, P., Williamson, C., Ndung'u, T., Klein, I., Chang, S. Y., Peter, T., Thior, I., Foley, B. T., Gaolekwe, S., Rybak, N., Gaseitsiwe, S., Vannberg, F., Marlink, R., Lee, T. H. & Essex, M. (2002). Human immunodeficiency virus type 1 subtype C molecular phylogeny: consensus sequence for an AIDS vaccine design? *J Virol* **76**, 5435–51.

O'Brien, S. J. & Moore, J. P. (2000). The effect of genetic variation in chemokines and their receptors on HIV transmission and progression to AIDS. *Immunol Rev* **177**, 99–111.

Oelrichs, R. B., Workman, C., Laukkanen, T., McCutchan, F. E. & Deacon, N. J. (1998). A novel subtype A/G/J recombinant full-length HIV type 1 genome from Burkina Faso. *AIDS Res Hum Retroviruses* **14**, 1495–500.

Onafuwa-Nuga, A. & Telesnitsky, A. (2009). The remarkable frequency of human immunodeficiency virus type 1 genetic recombination. *Microbiol Mol Biol Rev* **73**, 451–80.

Op de Coul, E. L., Prins, M., Cornelissen, M., van der Schoot, A., Boufassa, F., Brettle, R. P., Hernandez-Aguado, L., Schiffer, V., McMenamin, J., Rezza, G., Robertson, R., Zangerle, R., Goudsmit, J., Coutinho, R. A. & Lukashov, V. V. (2001). Using phylogenetic analysis to trace HIV-1 migration among western European injecting drug users seroconverting from 1984 to 1997. *Aids* **15**, 257–66.

Osmanov, S., Pattou, C., Walker, N., Schwardlander, B. & Esparza, J. (2002). Estimated global distribution and regional spread of HIV-1 genetic subtypes in the year 2000. *J Acquir Immune Defic Syndr* **29**, 184–90.

Paiardini, M., Pandrea, I., Apetrei, C. & Silvestri, G. (2009). Lessons learned from the natural hosts of HIV-related viruses. *Annu Rev Med* **60**, 485–95.

Papa, A., Papadimitriou, E., Adwan, G., Clewley, J. P., Malissiovas, N., Ntoutsos, I., Alexiou, S. & Antoniadis, A. (2000). HIV-1 co-receptor CCR5 and CCR2 mutations among Greeks. *FEMS Immunol Med Microbiol* **28**, 87–9.

Paradela, A., Alvarez, I., Garcia-Peydro, M., Sesma, L., Ramos, M., Vazquez, J. & Lopez De Castro, J. A. (2000). Limited diversity of peptides related to an alloreactive T cell epitope in the HLA-B27-bound peptide repertoire results from restrictions at multiple steps along the processing-loading pathway. *J Immunol* **164**, 329–37.

Peeters, M. (2000). Recombinant HIV sequences: Their role in the global epidemic. In *HIV Sequence Compendium I*, pp. 54–71. Edited by F. B. Kuiken CL, Hahn B, Korber B, McCutchan F, Marx PA, Mellors JW, Mullins JI, Sodroski J, and Wolinksy S. Los Alamos,: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory.

Peeters, M. & Sharp, P. M. (2000). Genetic diversity of HIV-1: the moving target. *Aids* 14 Suppl 3, S129–40.

Piacentini, L., Fenizia, C., Naddeo, V. & Clerici, M. (2008). Not just sheer luck! Immune correlates of protection against HIV-1 infection. *Vaccine* **26**, 3002–7.

Plantier, J. C., Leoz, M., Dickerson, J. E., De Oliveira, F., Cordonnier, F., Lemee, V., Damond, F., Robertson, D. L. & Simon, F. (2009). A new human immunodeficiency virus derived from gorillas. *Nat Med* **15**, 871–2.

Pommier, Y., Johnson, A. A., Marchand, C. Integrase inhibitors to treat HIV/AIDS. *Nat Rev Drug Discov* **4**, 236–48.

Rambaut, A., Posada, D., Crandall, K. A. & Holmes, E. C. (2004). The causes and consequences of HIV evolution. *Nat Rev Genet* **5**, 52–61.

Ramirez, B. C., Simon-Loriere, E., Galetto, R. & Negroni, M. (2008). Implications of recombination for HIV diversity. *Virus Res* **134**, 64–73.

Renard, C., Chardon, P. & Vaiman, M. (2003). The phylogenetic history of the MHC class I gene families in pig, including a fossil gene predating mammalian radiation. *J Mol Evol* **57**, 420–34.

Renard, C., Hart, E., Sehra, H., Beasley, H., Coggill, P., Howe, K., Harrow, J., Gilbert, J., Sims, S., Rogers, J., Ando, A., Shigenari, A., Shiina, T., Inoko, H., Chardon, P. & Beck, S. (2006). The genomic sequence and analysis of the swine major histo-compatibility complex. *Genomics* **88**, 96–110.

Reynes, J., Portales, P., Segondy, M., Baillat, V., Andre, P., Reant, B., Avinens, O., Couderc, G., Benkirane, M., Clot, J., Eliaou, J. F. & Corbeau, P. (2000). CD4+ T cell surface CCR5 density as a determining factor of virus load in persons infected with human immunodeficiency virus type 1. *J Infect Dis* **181**, 927–32.

Riva, C., Romano, L., Saladini, F., Lai, A., Carr, J. K., Francisci, D., Balotta, C. & Zazzi, M. (2008). Identification of a possible ancestor of the subtype A1 HIV Type 1 variant circulating in the former Soviet Union. *AIDS Res Hum Retroviruses* **24**, 1319–25.

Robertson, D. L., Anderson, J. P., Bradac, J. A., Carr, J. K., Foley, B., Funkhouser, R. K., Gao, F., Hahn, B. H., Kalish, M. L., Kuiken, C., Learn, G. H., Leitner, T., McCutchan, F., Osmanov, S., Peeters, M., Pieniazek, D., Salminen, M., Sharp, P. M., Wolinsky, S. & Korber, B. (2000). HIV-1 nomenclature proposal. *Science* **288**, 55–6.

Romani, B., Engelbrecht, S. & Glashoff, R. H. (2009). Antiviral roles of APOBEC proteins against HIV-1 and suppression by Vif. *Arch Virol*.

Samson, M., Libert, F., Doranz, B. J., Rucker, J., Liesnard, C., Farber, C. M., Saragosti, S., Lapoumeroulie, C., Cognaux, J., Forceille, C., Muyldermans, G., Verhofstede, C., Burtonboy, G., Georges, M., Imai, T., Rana, S., Yi, Y., Smyth, R. J., Collman, R. G., Doms, R. W., Vassart, G. & Parmentier, M. (1996). Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* **382**, 722–5.

Santra, S., Korber, B. T., Muldoon, M., Barouch, D. H., Nabel, G. J., Gao, F., Hahn, B. H., Haynes, B. F. & Letvin, N. L. (2008). A centralized gene-based HIV-1 vaccine elicits broad cross-clade cellular immune responses in rhesus monkeys. *Proc Natl Acad Sci U S A* **105**, 10489–94.

Schinkel, J., Langendam, M. W., Coutinho, R. A., Krol, A., Brouwer, M. & Schuite-maker, H. (1999). No evidence for an effect of the CCR5 delta32/+ and CCR2b 64I/+ mutations on human immunodeficiency virus (HIV)-1 disease progression among HIV-1-infected injecting drug users. *J Infect Dis* **179**, 825–31.

Sharp, P. M. & Hahn, B. H. (2008). AIDS: prehistory of HIV-1. *Nature* **455**, 605–6.

Skar, H., Sylvan, S., Hansson, H. B., Gustavsson, O., Boman, H., Albert, J. & Leitner, T. (2008). Multiple HIV-1 introductions into the Swedish intravenous drug user population. *Infect Genet Evol* **8**, 545–52.

Stevceva, L., Yoon, V., Anastasiades, D. & Poznansky, M. C. (2007). Immune responses to HIV Gp120 that facilitate viral escape. *Curr HIV Res* **5**, 47–54.

Subbarao, S., Vanichseni, S., Hu, D. J., Kitayaporn, D., Choopanya, K., Raktham, S., Young, N. L., Wasi, C., Sutthent, R., Luo, C. C., Ramos, A. & Mastro, T. D. (2000). Genetic characterization of incident HIV type 1 subtype E and B strains from a prospective cohort of injecting drug users in Bangkok, Thailand. *AIDS Res Hum Retroviruses* **16**, 699–707.

Sullivan, A. D., Wigginton, J. & Kirschner, D. (2001). The coreceptor mutation CCR5Delta32 influences the dynamics of HIV epidemics and is selected for by HIV. *Proc Natl Acad Sci U S A* **98**, 10214–9.

Surman, S. L., Sealy, R., Jones, B. G. & Hurwitz, J. L. (2009). HIV-1 vaccine design: harnessing diverse lymphocytes to conquer a diverse pathogen. *Hum Vaccin* **5**, 268–71.

Zhuang, J., Jetzt, A. E., Sun, G., Yu, H., Klarmann, G., Ron, Y., Preston, B. D. & Dougherty, J. P. (2002). Human immunodeficiency virus type 1 recombination: rate, fidelity, and putative hot spots. *J Virol* **76**, 11273–82.

Tamura K, Dudley J, Nei M & Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24:1596–1599.

Tanaka-Matsuda, M., Ando, A., Rogel-Gaillard, C., Chardon, P. & Uenishi, H. (2009). Difference in number of loci of swine leukocyte antigen classical class I genes among haplotypes. *Genomics* **93**, 261–73.

Taylor, B. S., Sobieszczyk, M. E., McCutchan, F. E. & Hammer, S. M. (2008). The challenge of HIV-1 subtype diversity. *N Engl J Med* **358**, 1590–602.

Tebit, D. M., Nankya, I., Arts, E. J. & Gao, Y. (2007). HIV diversity, recombination and disease progression: how does fitness "fit" into the puzzle? *AIDS Rev* **9**, 75–87.

Thomson, M. M. & Najera, R. (2007). Increasing HIV-1 genetic diversity in Europe. *J Infect Dis* **196**, 1120–4.

Thomson, S. A., Jaramillo, A. B., Shoobridge, M., Dunstan, K. J., Everett, B., Rana-singhe, C., Kent, S. J., Gao, K., Medveckzy, J., Ffrench, R. A. & Ramshaw, I. A. (2005). Development of a synthetic consensus sequence scrambled antigen HIV-1 vaccine designed for global use. *Vaccine* **23**, 4647–57.

Tozzi, V. (2009). Pharmacogenetics of antiretrovirals. *Antiviral Res*.

Trono, V. P. a. D. (1999). A Structure-function analysis of the Nef Protein of Primate Lentiviruses. In *Human Retroviruses and AIDS 1999*, pp. pp. 448–459. Edited by F. B. Kuiken CL, Hahn B, Korber B, McCutchan F, Marx PA, Mellors JW, Mullins JI, Sodroski J, and Wolinksy S. Los Alamos: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory.

Uusküla, A., Kalikova, A., Zilmer, K., Tammai, L. & DeHovitz, J. (2002). The role of injection drug use in the emergence of Human Immunodeficiency Virus infection in Estonia. *Int J Infect Dis* **6**, 23–7.

Watkins, D. I., Burton, D. R., Kallas, E. G., Moore, J. P. & Koff, W. C. (2008). Non-human primate models and the failure of the Merck HIV-1 vaccine in humans. *Nat Med* **14**, 617–21.

Venkatesan, S., Petrovic, A., Van Ryk, D. I., Locati, M., Weissman, D. & Murphy, P. M. (2002). Reduced cell surface expression of CCR5 in CCR5Delta 32 heterozygotes is mediated by gene dosage, rather than by receptor sequestration. *J Biol Chem* **277**, 2287–301.

Vidal, N., Mulanga, C., Bazepeo, E. B., Lepira, F., Delaporte, E., Peeters, M. (2006). Identification and Molecular Characterization of Subsubtype A4 in Central Africa. *AIDS Res Hum Retroviruses* **22**,182–187.

Visco-Comandini, U., Hultgren, C., Brostrom, C., Birk, M., Kim, S. & Sallberg, M. (1998). Human immunodeficiency virus type 1 disease progression, CCR5 genotype, and specific immune responses. *Clin Diagn Lab Immunol* **5**, 463–6.

Voevodin, A., Samilchuk, E. & Dashti, S. (1998). A survey for 32 nucleotide deletion in the CCR-5 chemokine receptor gene (deltaccr-5) conferring resistance to human immunodeficiency virus type 1 in different ethnic groups and in chimpanzees. *J Med Virol* **55**, 147–51.

Worobey, M., Gemmel, M., Teuwen, D. E., Haselkorn, T., Kunstman, K., Bunce, M., Muyembe, J. J., Kabongo, J. M., Kalengayi, R. M., Van Marck, E., Gilbert, M. T. & Wolinsky, S. M. (2008). Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* **455**, 661–4.

Wyatt R, K. P., Hendrickson WA, Sodroski JG (1998). Structure of the core of the HIV-1 gp120 Exterior Envelope Glycoprotein. In *Human Retroviruses and AIDS 1998*, pp. pp. III-3–9 Edited by K. C. Korber B, Foley B, Hahn B, McCutchan F, Mellors JW, and Sodroski J. Los Alamos: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory.

You, J., Croyle, J. L., Nishimura, A., Ozato, K. & Howley, P. M. (2004). Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes. *Cell* **117**, 349–60.

You, J., Schweiger, M. R. & Howley, P. M. (2005). Inhibition of E2 binding to Brd4 enhances viral genome loss and phenotypic reversion of bovine papillomavirus-transformed cells. *J Virol* **79**, 14956–61.

Yudin, N. S., Vinogradov, S. V., Potapova, T. A., Naykova, T. M., Sitnikova, V. V., Kulikov, I. V., Khasnulin, V. I., Konchuk, C., Vloschinskii, P. E., Ivanov, S. V., Kobzev, V. F., Romaschenko, A. G. & Voevoda, M. I. (1998). Distribution of CCR5-delta 32 gene deletion across the Russian part of Eurasia. *Hum Genet* **102**, 695–8.

Yusim, K., Kesmir, C., Gaschen, B., Addo, M. M., Altfeld, M., Brunak, S., Chigaev, A., Detours, V. & Korber, B. T. (2002). Clustering patterns of cytotoxic T-lymphocyte epitopes in human immunodeficiency virus type 1 (HIV-1) proteins reveal imprints of immune evasion on HIV-1 global variation. *J Virol* **76**, 8757–68.

SUMMARY IN ESTONIAN

HI-viiruse molekulaarne varieeruvus ja selle teadmise kasutamine vaktsiini väljatöötamisel

Käesoleva uurimistöö eesmärgiks oli Eesti populatsioonis esinevate inimese immuundefitsiitsuse (HI-) viiruse tüüpide määramine. Enne HI-viiruse epideemia algust 2000. aastal toimus nakatumine HI-viirusega peamiselt seksuaalsel teel ning levinuim viiruse tüüp Eestis, nagu mujalgi Euroopas, oli alamtüüp B. Alates 2000. aastast moodustavad nakatunutest põhiosa süstivad narkomaanid ning uurimistöös määrati nende seas levivad HI-viiruse tüübid. Viirusisolaatide nukleotiidse järjestuse analüüs näitas, et Eesti süstivate narkomaanide seas oli valdavaks viiruse tüübiks haruldane rekombinantne vorm CRF06_cpx, mis moodustas kõigist analüüsitud isolaatidest 76%. Lisaks esines märkimisväärselt suure sagedusega (17%) erinevaid unikaalseid rekombinantseid viirusvorme, mille nukleotiidne järjestus sarnanes viiruse genoomi ühes osas alamtüüp A viirustega, kuid teises osas CRF06_cpx tüüpi viirustega. Uuritud unikaalsetest rekombinantsetest viirusvormidest kolmandiku moodustasid identse mosaiikse struktuuriga isolaadid, andes tunnistust sellest, et Eesti populatsioonis ringleb uus, seni kirjeldamata rekombinantne HI-viiruse tüüp, mis hiljem määratleti uue rekombinantse vormina CRF32_06A1.

Lisaks uurisime HI-viiruse koretseptori, kemokiiniretseptori 5 (CCR5), polümorfse alleeli CCR5Δ32 esinemissagedust HI-viirusega nakatunutel Eesti populatsioonis. Inimestel, kellel esineb homosügootses olekus muteerunud alleel, mis sisaldab 32 aluspaarilist deletsiooni retseptori teist rakuvälist lingu kodeerivas järjestuses, on kogu toodetav CCR5 retseptor mittefunktionaalne ja vastavat valku raku pinnal ei ekspresseerita. Tänu sellele on need indiviidid suure tõenäosusega HI-viirusnakkusele resistentsed. Inimestel, kellel esineb heterosügootses olekus CCR5Δ32 mutatsiooniga alleel, on täheldatud väiksemat nakatumise riski ja aeglasemat serokonversiooni järgset immuunpuudulikuse (AIDS) arengut, kuigi sageli on tulemused olnud vastukäivad. Eesti populatsioon on heterosügootses olekus CCR5Δ32 ja HI-viiruse nakkuse seoste uurimiseks eriti sobiv, kuna varasemad uuringud on näidanud, et Eesti populatsioonis on CCR5Δ32 alleeli esinemissagedus üks kõrgemaid maailmas ning samaaegselt on populatsioonis väga kõrge HIV-positiivsete osakaal. Leidsime, et heterosügootses olekus CCR5Δ32 alleeli esinemise sagedus oli samasugune nii HIV positiivsetel (23,5%), kui ka HIV negatiivsetel (23,2%) patsientidel. Saadud tulemused lubavad väita, et heterosügootne CCR5Δ32/CCR5WT genotüüp ei kaitse HI-viirusesse nakatumise eest. Lisaks leidsime, et CCR5Δ32 alleeli esinemissagedus oli mõlemal viisil (seksuaalsel teel või narkootikumide veenisisesel süstimisel) nakatunutel sarnane. Ka ei erinenud CCR5Δ32 alleelisagedus unikaalsete rekombinantsete HI-viirusvormidega nakatunute seas võrreldes ülejäänud HIV-positiivsete patsientidega. Kuna unikaalsete rekombinantsete vormide esinemine on märk koinfektsioonist või superinfektsioonist,

võib väita, et CCR5Δ32/CCR5WT heterosügootne genotüüp ei taga kaitset uute nakkuste eest.

HI-viiruse alamtüüpide A, B, C, F, G ja H fülogeneetilisel analüüsil põhinevate konsensus- või eellasjärjestuse alusel konstrueeriti neli sünteetilist MultiHIV antigeeni (konsensus-A, konsensus-B, konsensus-C ja eelas-FGH), eesmärgiga kasutada neid HI-viiruse vastastes geneetilistes vaktsiinides. Kõik neli antigeeni olid sarnase ülesehituse ja suurusega, sisaldades viiruse elutsükli algfaasis ekspresseeritavaid regulatoorseid valke Rev, Nef ja Tat; struktuurseid valke Gag $_{p17}$ ja Gag $_{p24}$ ning inimese T-raku epitoopide rikkaid alasid viiruse proteaasist, pöördtranskiptaasist ning ümbrisevalkudest Env_{gp120} ja Env_{gp41}. Alamtüüp B spetsiifilist MultiHIV-B antigeeni ekspresseerivat DNA vaktsiini kasutati, et immuniseerida katseloomi – kodusigasid. Näitasime, et antigeeni vastane rakuline immuunvastus oli ulatuslik, kuna sea tsütotoksilised T-lümfotsüüdid reageerisid kõigile erinevatele viiruse komponentidele antigeenis. Katses osalenud katseloomad genotüpiseeriti MHC-I lookuste SLA-1, SLA-2 ja SLA-3 osas. Ilmnes, et kõigil katseloomadel esines unikaalne SLA-I genotüüp. Meie uurimus tõestas, et kodusiga on suurepärane katseloom T-rakulise immuunvastuse kirjeldamiseks, näiteks vaktsiiniarendusel.

ACKNOWLEDGEMENTS

This study was carried out at the company FIT Biotech Oyj Plc Eesti filiaal from 2004 to 2009.

– I would like to express my sincere gratitude to the Management and Supervisory Board of FIT Biotech for providing the opportunity to perform these studies and for supporting publishing of scientific results.
– I would like to acknowledge all co-authors and collaborators who have collected the valuable samples as well as all volunteers participating in the studies.
– I am very grateful to my supervisors prof. Richard Villems and prof. Mart Ustav for their guidance in studying this outstandingly interesting topic – HIV. It is challenging and inspiring to work under the leadership of such visionaries.
– I owe a lot to all friendly colleagues in FIT Biotech, Icosagen, and Department of Evolutionary Biology. Particularly, I would like to thank my colleagues and friends in FIT Biotech – Katrin Kaldma, Kaja Kiiver, Andres Männik, and Tarmo Mölder for good advice, critical remarks, and encouragement during the preparation of the manuscript, as well as for all good times spent together. I would like to acknowledge Anne Noop for the everyday help and for keeping collective spirit high. I am indebted to Rein Sikut for his knowing everything about immunology and cell biology, and for sharing it with those who don't.
– I thank my former colleague in the Department of Evolutionary Biology – Toomas Kivisild for always answering all my questions patiently and thoroughly, no matter if they were related to a virus, a pig or a human population.
– I am thankful to my parents, brothers, and relatives for successfully recreating the atmosphere of My Family and Other Animals – the basic knowledge and love for nature comes from the time spent with them. My passion for PCR and sequencing (amongst other good things) I have inherited from my father Jüri Parik.
– I thank Johanna, Villem, Saskia, Toomas, and Kristjan for best company one could dream of.

PUBLICATIONS

CURRICULUM VITAE

Maarja Adojaan

Date and place of birth:	2.05.1974, Tartu
Citizenship:	Estonian
Family status:	married, four children
E-mail:	maarja.adojaan@fitbiotech.com
Current position:	FIT Biotech Oy Eesti AS, scientist

Education:

1981–1992	Miina Härma Gymnasium
1992–1996	University of Tartu, Faculty of Biology and Geography, BSc (biology)
1996–1999	University of Tartu, Institute of Molecular and Cell Biology, Department of Evolutionary biology, MSc (molecular and cell biology) "Diversity of Y chromosome haplogroups in populations of Eastern Europe, Caucasus region and Anatolia"
1999–	University of Tartu, Institute of Molecular and Cell Biology, Department of Evolutionary biology, doctorate studies

Professional employment:

1999–2001	Estonian Biocentre, technician
2001–	FIT Biotech Oy Eesti AS, scientist

List of publications:

Adojaan Maarja, Kivisild Toomas, Männik Andres, Krispin Tõnu, Ustina Valentina, Zilmer Kai, Liebert Elo, Jaroslavtsev Nikolai, Priimägi Ludmilla, Tefanova Valentina, Schmidt Jelena, Krohn Kai, Villems Richard, Salminen Mika, Ustav Mart. Predominance of a Rare Type of HIV-1 in Estonia. (2005) *JAIDS Journal of Acquired Immune Deficiency Syndromes* 39: 598–605.

Adojaan Maarja, Mölder Tarmo, Männik Andres, Kivisild Toomas, Villems Richard, Krispin Tõnu, Ustav Mart. High Prevalence of the CCR5Δ32 HIV-Resistance Mutation among Estonian HIV Type 1-Infected Individuals. (2007) *AIDS Research and Human Retroviruses* 23:193–197.

Mölder Tarmo, **Adojaan Maarja**, Kaldma Katrin, Ustav Mart, Sikut Rein. Elicitation of broad CTL-response against HIV-1 by the DNA vaccine encoding artificial multi-component fusion protein MultiHIV – Study in domestic pigs. (2009) *Vaccine* (E-published ahead of print).

Martinon Frederic, Kaldma Katrin, Sikut Rein, Culina Slobodan, Romain Gabrielle, Tuomela Mari, **Adojaan Maarja**, Männik Andres, Toots Urve, Kivisild Toomas, Morin Julie, Brochard Patricia, Delache Benoit, Tripiciano Antonella, Ensoli Fabrizio, Stanescu Ioana, Le Grand Roger, Ustav Mart. (2009). Persistent Immune Responses Induced by a HIV DNA Vaccine Delivered in Association with Electroporation in the Skin of Nonhuman primates. (2009) *Human Gene Therapy*, 20:1–17.

CURRICULUM VITAE

Maarja Adojaan

Sünniaeg ja koht:	2.05.1974
Perekonnaseis:	abielus, 4 last
E-post:	maarja.adojaan@fitbiotech.com
Praegune töökoht, amet:	FIT Biotech Oy Eesti AS, teadur

Haridus:

1981–1992	Miina Härma Gümnaasium
1992–1996	Tartu Ülikool, Bioloogia- ja geograafiateaduskond, BSc (bioloogia erialal)
1996–1999	Tartu Ülikool, Molekulaar-ja Rakubioloogia Instituut, Evolutsioonilse bioloogia õppetool, MSc (molekulaar- ja rakubioloogia) "Diversity of Y chromosome haplogroups in populations of Eastern Europe, Caucasus region and Anatolia"
1999–	Tartu Ülikool, Molekulaar- ja Rakubioloogia Instituut, Evolutsioonilise bioloogia õppetool, doktoriõpe

Töökogemus:

1999–2001	Eesti Biokeskus, laborant
2001–	FIT Biotech Oy Eesti AS, teadur

Publikatsioonide loetelu

Adojaan Maarja, Kivisild Toomas, Männik Andres, Krispin Tõnu, Ustina Valentina, Zilmer Kai, Liebert Elo, Jaroslavtsev Nikolai, Priimägi Ludmilla, Tefanova Valentina, Schmidt Jelena, Krohn Kai, Villems Richard, Salminen Mika, Ustav Mart. Predominance of a Rare Type of HIV-1 in Estonia. (2005) *JAIDS Journal of Acquired Immune Deficiency Syndromes* 39: 598–605.

Adojaan Maarja, Mölder Tarmo, Männik Andres, Kivisild Toomas, Villems Richard, Krispin Tõnu, Ustav Mart. High Prevalence of the CCR5Δ32 HIV-Resistance Mutation among Estonian HIV Type 1-Infected Individuals. (2007) *AIDS Research and Human Retroviruses* 23:193–197.

Mölder Tarmo, **Adojaan Maarja**, Kaldma Katrin, Ustav Mart, Sikut Rein. Elicitation of broad CTL-response against HIV-1 by the DNA vaccine encoding artificial multi-component fusion protein MultiHIV – Study in domestic pigs. (2009) *Vaccine* (E-published ahead of print).

Martinon Frederic, Kaldma Katrin, Sikut Rein, Culina Slobodan, Romain Gabrielle, Tuomela Mari, **Adojaan Maarja**, Männik Andres, Toots Urve, Kivisild Toomas, Morin Julie, Brochard Patricia, Delache Benoit, Tripiciano Antonella, Ensoli Fabrizio, Stanescu Ioana, Le Grand Roger, Ustav Mart. (2009). Persistent Immune Responses Induced by a HIV DNA Vaccine Delivered in Association with Electroporation in the Skin of Nonhuman primates. (2009) *Human Gene Therapy*, 20:1–17.

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS

1. **Toivo Maimets**. Studies of human oncoprotein p53. Tartu, 1991, 96 p.
2. **Enn K. Seppet**. Thyroid state control over energy metabolism, ion transport and contractile functions in rat heart. Tartu, 1991, 135 p.
3. **Kristjan Zobel**. Epifüütsete makrosamblike väärtus õhu saastuse indikaatoritena Hamar-Dobani boreaalsetes mägimetsades. Tartu, 1992, 131 lk.
4. **Andres Mäe**. Conjugal mobilization of catabolic plasmids by transposable elements in helper plasmids. Tartu, 1992, 91 p.
5. **Maia Kivisaar**. Studies on phenol degradation genes of *Pseudomonas* sp. strain EST 1001. Tartu, 1992, 61 p.
6. **Allan Nurk**. Nucleotide sequences of phenol degradative genes from *Pseudomonas sp.* strain EST 1001 and their transcriptional activation in *Pseudomonas putida.* Tartu, 1992, 72 p.
7. **Ülo Tamm**. The genus *Populus* L. in Estonia: variation of the species biology and introduction. Tartu, 1993, 91 p.
8. **Jaanus Remme**. Studies on the peptidyltransferase centre of the *E.coli* ribosome. Tartu, 1993, 68 p.
9. **Ülo Langel**. Galanin and galanin antagonists. Tartu, 1993, 97 p.
10. **Arvo Käärd**. The development of an automatic online dynamic fluorescense-based pH-dependent fiber optic penicillin flowthrought biosensor for the control of the benzylpenicillin hydrolysis. Tartu, 1993, 117 p.
11. **Lilian Järvekülg**. Antigenic analysis and development of sensitive immunoassay for potato viruses. Tartu, 1993, 147 p.
12. **Jaak Palumets**. Analysis of phytomass partition in Norway spruce. Tartu, 1993, 47 p.
13. **Arne Sellin**. Variation in hydraulic architecture of *Picea abies* (L.) Karst. trees grown under different enviromental conditions. Tartu, 1994, 119 p.
13. **Mati Reeben**. Regulation of light neurofilament gene expression. Tartu, 1994, 108 p.
14. **Urmas Tartes**. Respiration rhytms in insects. Tartu, 1995, 109 p.
15. **Ülo Puurand**. The complete nucleotide sequence and infections *in vitro* transcripts from cloned cDNA of a potato A potyvirus. Tartu, 1995, 96 p.
16. **Peeter Hõrak**. Pathways of selection in avian reproduction: a functional framework and its application in the population study of the great tit (*Parus major*). Tartu, 1995, 118 p.
17. **Erkki Truve**. Studies on specific and broad spectrum virus resistance in transgenic plants. Tartu, 1996, 158 p.
18. **Illar Pata**. Cloning and characterization of human and mouse ribosomal protein S6-encoding genes. Tartu, 1996, 60 p.
19. **Ülo Niinemets**. Importance of structural features of leaves and canopy in determining species shade-tolerance in temperature deciduous woody taxa. Tartu, 1996, 150 p.

20. **Ants Kurg**. Bovine leukemia virus: molecular studies on the packaging region and DNA diagnostics in cattle. Tartu, 1996, 104 p.
21. **Ene Ustav**. E2 as the modulator of the BPV1 DNA replication. Tartu, 1996, 100 p.
22. **Aksel Soosaar**. Role of helix-loop-helix and nuclear hormone receptor transcription factors in neurogenesis. Tartu, 1996, 109 p.
23. **Maido Remm**. Human papillomavirus type 18: replication, transformation and gene expression. Tartu, 1997, 117 p.
24. **Tiiu Kull**. Population dynamics in *Cypripedium calceolus* L. Tartu, 1997, 124 p.
25. **Kalle Olli**. Evolutionary life-strategies of autotrophic planktonic microorganisms in the Baltic Sea. Tartu, 1997, 180 p.
26. **Meelis Pärtel**. Species diversity and community dynamics in calcareous grassland communities in Western Estonia. Tartu, 1997, 124 p.
27. **Malle Leht**. The Genus *Potentilla* L. in Estonia, Latvia and Lithuania: distribution, morphology and taxonomy. Tartu, 1997, 186 p.
28. **Tanel Tenson**. Ribosomes, peptides and antibiotic resistance. Tartu, 1997, 80 p.
29. **Arvo Tuvikene**. Assessment of inland water pollution using biomarker responses in fish *in vivo* and *in vitro*. Tartu, 1997, 160 p.
30. **Urmas Saarma**. Tuning ribosomal elongation cycle by mutagenesis of 23S rRNA. Tartu, 1997, 134 p.
31. **Henn Ojaveer**. Composition and dynamics of fish stocks in the gulf of Riga ecosystem. Tartu, 1997, 138 p.
32. **Lembi Lõugas**. Post-glacial development of vertebrate fauna in Estonian water bodies. Tartu, 1997, 138 p.
33. **Margus Pooga**. Cell penetrating peptide, transportan, and its predecessors, galanin-based chimeric peptides. Tartu, 1998, 110 p.
34. **Andres Saag**. Evolutionary relationships in some cetrarioid genera (Lichenized Ascomycota). Tartu, 1998, 196 p.
35. **Aivar Liiv**. Ribosomal large subunit assembly *in vivo*. Tartu, 1998, 158 p.
36. **Tatjana Oja**. Isoenzyme diversity and phylogenetic affinities among the eurasian annual bromes (*Bromus* L., Poaceae). Tartu, 1998, 92 p.
37. **Mari Moora**. The influence of arbuscular mycorrhizal (AM) symbiosis on the competition and coexistence of calcareous crassland plant species. Tartu, 1998, 78 p.
38. **Olavi Kurina**. Fungus gnats in Estonia (*Diptera: Bolitophilidae, Keroplatidae, Macroceridae, Ditomyiidae, Diadocidiidae, Mycetophilidae*). Tartu, 1998, 200 p.
39. **Andrus Tasa**. Biological leaching of shales: black shale and oil shale. Tartu, 1998, 98 p.
40. **Arnold Kristjuhan.** Studies on transcriptional activator properties of tumor suppressor protein p53. Tartu, 1998, 86 p.

41. **Sulev Ingerpuu.** Characterization of some human myeloid cell surface and nuclear differentiation antigens. Tartu, 1998, 163 p.
42. **Veljo Kisand.** Responses of planktonic bacteria to the abiotic and biotic factors in the shallow lake Võrtsjärv. Tartu, 1998, 118 p.
43. **Kadri Põldmaa.** Studies in the systematics of hypomyces and allied genera (Hypocreales, Ascomycota). Tartu, 1998, 178 p.
44. **Markus Vetemaa.** Reproduction parameters of fish as indicators in environmental monitoring. Tartu, 1998, 117 p.
45. **Heli Talvik.** Prepatent periods and species composition of different *Oesophagostomum* spp. populations in Estonia and Denmark. Tartu, 1998, 104 p.
46. **Katrin Heinsoo.** Cuticular and stomatal antechamber conductance to water vapour diffusion in *Picea abies* (L.) karst. Tartu, 1999, 133 p.
47. **Tarmo Annilo.** Studies on mammalian ribosomal protein S7. Tartu, 1998, 77 p.
48. **Indrek Ots.** Health state indicies of reproducing great tits (*Parus major*): sources of variation and connections with life-history traits. Tartu, 1999, 117 p.
49. **Juan Jose Cantero.** Plant community diversity and habitat relationships in central Argentina grasslands. Tartu, 1999, 161 p.
50. **Rein Kalamees.** Seed bank, seed rain and community regeneration in Estonian calcareous grasslands. Tartu, 1999, 107 p.
51. **Sulev Kõks.** Cholecystokinin (CCK) — induced anxiety in rats: influence of environmental stimuli and involvement of endopioid mechanisms and erotonin. Tartu, 1999, 123 p.
52. **Ebe Sild.** Impact of increasing concentrations of O_3 and CO_2 on wheat, clover and pasture. Tartu, 1999, 123 p.
53. **Ljudmilla Timofejeva.** Electron microscopical analysis of the synaptonemal complex formation in cereals. Tartu, 1999, 99 p.
54. **Andres Valkna.** Interactions of galanin receptor with ligands and G-proteins: studies with synthetic peptides. Tartu, 1999, 103 p.
55. **Taavi Virro.** Life cycles of planktonic rotifers in lake Peipsi. Tartu, 1999, 101 p.
56. **Ana Rebane.** Mammalian ribosomal protein S3a genes and intron-encoded small nucleolar RNAs U73 and U82. Tartu, 1999, 85 p.
57. **Tiina Tamm.** Cocksfoot mottle virus: the genome organisation and translational strategies. Tartu, 2000, 101 p.
58. **Reet Kurg.** Structure-function relationship of the bovine papilloma virus E2 protein. Tartu, 2000, 89 p.
59. **Toomas Kivisild.** The origins of Southern and Western Eurasian populations: an mtDNA study. Tartu, 2000, 121 p.
60. **Niilo Kaldalu.** Studies of the TOL plasmid transcription factor XylS. Tartu 2000. 88 p.

61. **Dina Lepik.** Modulation of viral DNA replication by tumor suppressor protein p53. Tartu 2000. 106 p.
62. **Kai Vellak.** Influence of different factors on the diversity of the bryophyte vegetation in forest and wooded meadow communities. Tartu 2000. 122 p.
63. **Jonne Kotta.** Impact of eutrophication and biological invasionas on the structure and functions of benthic macrofauna. Tartu 2000. 160 p.
64. **Georg Martin.** Phytobenthic communities of the Gulf of Riga and the inner sea the West-Estonian archipelago. Tartu, 2000. 139 p.
65. **Silvia Sepp.** Morphological and genetical variation of *Alchemilla L.* in Estonia. Tartu, 2000. 124 p.
66. **Jaan Liira.** On the determinants of structure and diversity in herbaceous plant communities. Tartu, 2000. 96 p.
67. **Priit Zingel.** The role of planktonic ciliates in lake ecosystems. Tartu 2001. 111 p.
68. **Tiit Teder.** Direct and indirect effects in Host-parasitoid interactions: ecological and evolutionary consequences. Tartu 2001. 122 p.
69. **Hannes Kollist.** Leaf apoplastic ascorbate as ozone scavenger and its transport across the plasma membrane. Tartu 2001. 80 p.
70. **Reet Marits.** Role of two-component regulator system PehR-PehS and extracellular protease PrtW in virulence of *Erwinia Carotovora* subsp. *Carotovora*. Tartu 2001. 112 p.
71. **Vallo Tilgar.** Effect of calcium supplementation on reproductive performance of the pied flycatcher *Ficedula hypoleuca* and the great tit *Parus major,* breeding in Nothern temperate forests. Tartu, 2002. 126 p.
72. **Rita Hõrak.** Regulation of transposition of transposon Tn*4652* in *Pseudomonas putida*. Tartu, 2002. 108 p.
73. **Liina Eek-Piirsoo.** The effect of fertilization, mowing and additional illumination on the structure of a species-rich grassland community. Tartu, 2002. 74 p.
74. **Krõõt Aasamaa.** Shoot hydraulic conductance and stomatal conductance of six temperate deciduous tree species. Tartu, 2002. 110 p.
75. **Nele Ingerpuu.** Bryophyte diversity and vascular plants. Tartu, 2002. 112 p.
76. **Neeme Tõnisson.** Mutation detection by primer extension on oligonucleotide microarrays. Tartu, 2002. 124 p.
77. **Margus Pensa.** Variation in needle retention of Scots pine in relation to leaf morphology, nitrogen conservation and tree age. Tartu, 2003. 110 p.
78. **Asko Lõhmus.** Habitat preferences and quality for birds of prey: from principles to applications. Tartu, 2003. 168 p.
79. **Viljar Jaks.** p53 — a switch in cellular circuit. Tartu, 2003. 160 p.
80. **Jaana Männik.** Characterization and genetic studies of four ATP-binding cassette (ABC) transporters. Tartu, 2003. 140 p.
81. **Marek Sammul.** Competition and coexistence of clonal plants in relation to productivity. Tartu, 2003. 159 p

82. **Ivar Ilves.** Virus-cell interactions in the replication cycle of bovine papillomavirus type 1. Tartu, 2003. 89 p.
83. **Andres Männik.** Design and characterization of a novel vector system based on the stable replicator of bovine papillomavirus type 1. Tartu, 2003. 109 p.
84. **Ivika Ostonen.** Fine root structure, dynamics and proportion in net primary production of Norway spruce forest ecosystem in relation to site conditions. Tartu, 2003. 158 p.
85. **Gudrun Veldre.** Somatic status of 12–15-year-old Tartu schoolchildren. Tartu, 2003. 199 p.
86. **Ülo Väli.** The greater spotted eagle *Aquila clanga* and the lesser spotted eagle *A. pomarina*: taxonomy, phylogeography and ecology. Tartu, 2004. 159 p.
87. **Aare Abroi.** The determinants for the native activities of the bovine papillomavirus type 1 E2 protein are separable. Tartu, 2004. 135 p.
88. **Tiina Kahre.** Cystic fibrosis in Estonia. Tartu, 2004. 116 p.
89. **Helen Orav-Kotta.** Habitat choice and feeding activity of benthic suspension feeders and mesograzers in the northern Baltic Sea. Tartu, 2004. 117 p.
90. **Maarja Öpik.** Diversity of arbuscular mycorrhizal fungi in the roots of perennial plants and their effect on plant performance. Tartu, 2004. 175 p.
91. **Kadri Tali.** Species structure of *Neotinea ustulata*. Tartu, 2004. 109 p.
92. **Kristiina Tambets.** Towards the understanding of post-glacial spread of human mitochondrial DNA haplogroups in Europe and beyond: a phylogeographic approach. Tartu, 2004. 163 p.
93. **Arvi Jõers.** Regulation of p53-dependent transcription. Tartu, 2004. 103 p.
94. **Lilian Kadaja.** Studies on modulation of the activity of tumor suppressor protein p53. Tartu, 2004. 103 p.
95. **Jaak Truu.** Oil shale industry wastewater: impact on river microbial community and possibilities for bioremediation. Tartu, 2004. 128 p.
96. **Maire Peters.** Natural horizontal transfer of the *pheBA* operon. Tartu, 2004. 105 p.
97. **Ülo Maiväli.** Studies on the structure-function relationship of the bacterial ribosome. Tartu, 2004. 130 p.
98. **Merit Otsus.** Plant community regeneration and species diversity in dry calcareous grasslands. Tartu, 2004. 103 p.
99. **Mikk Heidemaa.** Systematic studies on sawflies of the genera *Dolerus, Empria,* and *Caliroa* (Hymenoptera: Tenthredinidae). Tartu, 2004. 167 p.
100. **Ilmar Tõnno.** The impact of nitrogen and phosphorus concentration and N/P ratio on cyanobacterial dominance and N_2 fixation in some Estonian lakes. Tartu, 2004. 111 p.
101. **Lauri Saks.** Immune function, parasites, and carotenoid-based ornaments in greenfinches. Tartu, 2004. 144 p.

102. **Siiri Rootsi.** Human Y-chromosomal variation in European populations. Tartu, 2004. 142 p.
103. **Eve Vedler.** Structure of the 2,4-dichloro-phenoxyacetic acid-degradative plasmid pEST4011. Tartu, 2005. 106 p.
104. **Andres Tover.** Regulation of transcription of the phenol degradation *pheBA* operon in *Pseudomonas putida*. Tartu, 2005. 126 p.
105. **Helen Udras.** Hexose kinases and glucose transport in the yeast *Hansenula polymorpha*. Tartu, 2005. 100 p.
106. **Ave Suija.** Lichens and lichenicolous fungi in Estonia: diversity, distribution patterns, taxonomy. Tartu, 2005. 162 p.
107. **Piret Lõhmus.** Forest lichens and their substrata in Estonia. Tartu, 2005. 162 p.
108. **Inga Lips.** Abiotic factors controlling the cyanobacterial bloom occurrence in the Gulf of Finland. Tartu, 2005. 156 p.
109. **Kaasik, Krista.** Circadian clock genes in mammalian clockwork, metabolism and behaviour. Tartu, 2005. 121 p.
110. **Juhan Javoiš.** The effects of experience on host acceptance in ovipositing moths. Tartu, 2005. 112 p.
111. **Tiina Sedman.** Characterization of the yeast *Saccharomyces cerevisiae* mitochondrial DNA helicase Hmi1. Tartu, 2005. 103 p.
112. **Ruth Aguraiuja.** Hawaiian endemic fern lineage *Diellia* (Aspleniaceae): distribution, population structure and ecology. Tartu, 2005. 112 p.
113. **Riho Teras.** Regulation of transcription from the fusion promoters generated by transposition of Tn*4652* into the upstream region of *pheBA* operon in *Pseudomonas putida*. Tartu, 2005. 106 p.
114. **Mait Metspalu.** Through the course of prehistory in india: tracing the mtDNA trail. Tartu, 2005. 138 p.
115. **Elin Lõhmussaar.** The comparative patterns of linkage disequilibrium in European populations and its implication for genetic association studies. Tartu, 2006. 124 p.
116. **Priit Kupper.** Hydraulic and environmental limitations to leaf water relations in trees with respect to canopy position. Tartu, 2006. 126 p.
117. **Heili Ilves.** Stress-induced transposition of Tn*4652* in *Pseudomonas Putida.* Tartu, 2006. 120 p.
118. **Silja Kuusk.** Biochemical properties of Hmi1p, a DNA helicase from *Saccharomyces cerevisiae* mitochondria. Tartu, 2006. 126 p.
119. **Kersti Püssa.** Forest edges on medium resolution landsat thematic mapper satellite images. Tartu, 2006. 90 p.
120. **Lea Tummeleht.** Physiological condition and immune function in great tits (*Parus major* l.): Sources of variation and trade-offs in relation to growth. Tartu, 2006. 94 p.
121. **Toomas Esperk.** Larval instar as a key element of insect growth schedules. Tartu, 2006. 186 p.

122. **Harri Valdmann.** Lynx (*Lynx lynx*) and wolf (*Canis lupus*) in the Baltic region: Diets, helminth parasites and genetic variation. Tartu, 2006. 102 p.
123. **Priit Jõers.** Studies of the mitochondrial helicase Hmi1p in *Candida albicans* and *Saccharomyces cerevisia*. Tartu, 2006. 113 p.
124. **Kersti Lilleväli.** Gata3 and Gata2 in inner ear development. Tartu, 2007. 123 p.
125. **Kai Rünk.** Comparative ecology of three fern species: *Dryopteris carthusiana* (Vill.) H.P. Fuchs, *D. expansa* (C. Presl) Fraser-Jenkins & Jermy and *D. dilatata* (Hoffm.) A. Gray (Dryopteridaceae). Tartu, 2007. 143 p.
126. **Aveliina Helm.** Formation and persistence of dry grassland diversity: role of human history and landscape structure. Tartu, 2007. 89 p.
127. **Leho Tedersoo.** Ectomycorrhizal fungi: diversity and community structure in Estonia, Seychelles and Australia. Tartu, 2007. 233 p.
128. **Marko Mägi.** The habitat-related variation of reproductive performance of great tits in a deciduous-coniferous forest mosaic: looking for causes and consequences. Tartu, 2007. 135 p.
129. **Valeria Lulla.** Replication strategies and applications of Semliki Forest virus. Tartu, 2007. 109 p.
130. **Ülle Reier**. Estonian threatened vascular plant species: causes of rarity and conservation. Tartu, 2007. 79 p.
131. **Inga Jüriado**. Diversity of lichen species in Estonia: influence of regional and local factors. Tartu, 2007. 171 p.
132. **Tatjana Krama.** Mobbing behaviour in birds: costs and reciprocity based cooperation. Tartu, 2007.
133. **Signe Saumaa.** The role of DNA mismatch repair and oxidative DNA damage defense systems in avoidance of stationary phase mutations in *Pseudomonas putida.* Tartu, 2007. 172 p.
134. **Reedik Mägi**. The linkage disequilibrium and the selection of genetic markers for association studies in european populations. Tartu, 2007. 96 p.
135. **Priit Kilgas.** Blood parameters as indicators of physiological condition and skeletal development in great tits (*Parus major*): natural variation and application in the reproductive ecology of birds. Tartu, 2007. 129 p.
136. **Anu Albert**. The role of water salinity in structuring eastern Baltic coastal fish communities. Tartu, 2007. 95 p.
137. **Kärt Padari.** Protein transduction mechanisms of transportans. Tartu, 2008. 128 p.
138. **Siiri-Lii Sandre.** Selective forces on larval colouration in a moth. Tartu, 2008. 125 p.
139. **Ülle Jõgar.** Conservation and restoration of semi-natural floodplain meadows and their rare plant species. Tartu, 2008. 99 p.
140. **Lauri Laanisto.** Macroecological approach in vegetation science: generality of ecological relationships at the global scale. Tartu, 2008. 133 p.
141. **Reidar Andreson**. Methods and software for predicting PCR failure rate in large genomes. Tartu, 2008. 105 p.

142. **Birgot Paavel.** Bio-optical properties of turbid lakes. Tartu, 2008. 175 p.
143. **Kaire Torn.** Distribution and ecology of charophytes in the Baltic Sea. Tartu, 2008, 98 p.
144. **Vladimir Vimberg.** Peptide mediated macrolide resistance. Tartu, 2008, 190 p.
145. **Daima Örd.** Studies on the stress-inducible pseudokinase TRB3, a novel inhibitor of transcription factor ATF4. Tartu, 2008, 108 p.
146. **Lauri Saag.** Taxonomic and ecologic problems in the genus *Lepraria* (*Stereocaulaceae*, lichenised *Ascomycota*). Tartu, 2008, 175 p.
147. **Ulvi Karu.** Antioxidant protection, carotenoids and coccidians in greenfinches – assessment of the costs of immune activation and mechanisms of parasite resistance in a passerine with carotenoid-based ornaments. Tartu, 2008, 124 p.
148. **Jaanus Remm.** Tree-cavities in forests: density, characteristics and occupancy by animals. Tartu, 2008, 128 p.
149. **Epp Moks.** Tapeworm parasites *Echinococcus multilocularis* and *E. granulosus* in Estonia: phylogenetic relationships and occurrence in wild carnivores and ungulates. Tartu, 2008, 82 p.
150. **Eve Eensalu.** Acclimation of stomatal structure and function in tree canopy: effect of light and CO_2 concentration. Tartu, 2008, 108 p.
151. **Janne Pullat**. Design, functionlization and application of an *in situ* synthesized oligonucleotide microarray. Tartu, 2008, 108 p.
152. **Marta Putrinš.** Responses of *Pseudomonas putida* to phenol-induced metabolic and stress signals. Tartu, 2008, 142 p.
153. **Marina Semtšenko.** Plant root behaviour: responses to neighbours and physical obstructions. Tartu, 2008, 106 p.
154. **Marge Starast.** Influence of cultivation techniques on productivity and fruit quality of some *Vaccinium* and *Rubus* taxa. Tartu, 2008, 154 p.
155. **Age Tats.** Sequence motifs influencing the efficiency of translation. Tartu, 2009, 104 p.
156. **Radi Tegova.** The role of specialized DNA polymerases in mutagenesis in *Pseudomonas putida.* Tartu, 2009, 124 p.
157. **Tsipe Aavik.** Plant species richness, composition and functional trait pattern in agricultural landscapes – the role of land use intensity and landscape structure. Tartu, 2008, 112 p.
158. **Kaja Kiiver.** Semliki forest virus based vectors and cell lines for studying the replication and interactions of alphaviruses and hepaciviruses. Tartu, 2009, 104 p.
159. **Meelis Kadaja.** Papillomavirus Replication Machinery Induces Genomic Instability in its Host Cell. Tartu, 2009, 126 p.
160. **Pille Hallast.** Human and chimpanzee Luteinizing hormone/Chorionic Gonadotropin beta (*LHB/CGB*) gene clusters: diversity and divergence of young duplicated genes. Tartu, 2009, 168 p.

161. **Ain Vellak.** Spatial and temporal aspects of plant species conservation. Tartu, 2009, 86 p.
162. **Triinu Remmel.** Body size evolution in insects with different colouration strategies: the role of predation risk. Tartu, 2009, 168 p.
163. **Jaana Salujõe.** Zooplankton as the indicator of ecological quality and fish predation in lake ecosystems. Tartu, 2009, 129 p.
164. **Ele Vahtmäe.** Mapping benthic habitat with remote sensing in optically complex coastal environments. Tartu, 2009, 109 p.
165. **Liisa Metsamaa.** Model-based assessment to improve the use of remote sensing in recognition and quantitative mapping of cyanobacteria. Tartu, 2009, 114 p.
166. **Pille Säälik.** The role of endocytosis in the protein transduction by cell-penetrating peptides. Tartu, 2009, 155 p.
167. **Lauri Peil.** Ribosome assembly factors in *Escherichia coli.* Tartu, 2009, 147 p.
168. **Lea Hallik.** Generality and specificity in light harvesting, carbon gain capacity and shade tolerance among plant functional groups. Tartu, 2009, 99 p.
169. **Mariliis Tark.** Mutagenic potential of DNA damage repair and tolerance mechanisms under starvation stress. Tartu, 2009, 191 p.
170. **Riinu Rannap.** Impacts of habitat loss and restoration on amphibian populations. Tartu, 2009, 117 p.