CHARACTERIZATION OF MUTATIONAL PROFILES ASSOCIATED WITH ANTI HIV-1 AND ANTI HBV DRUGS

Candidata
Dott.ssa Maria Trignetti

Relatori
Prof. Carlo-Federico Perno
Dott.ssa Valentina Svicher

Anno accademico 2008-2009
ABSTRACT

- HIV-1
Utilizzando il modello dell’analisi di sopravvivenza è stata analizzata, durante l’interruzione terapeutica, la dinamica delle mutazioni conferenti resistenza agli inibitori nucleosidici della trascrittasi inversa (NRTI) di HIV-1.
Sono state analizzate le sequenze della trascrittasi inversa del virus estratto dal plasma di 132 pazienti per i quali erano disponibili i risultati di almeno due test di resistenza genotipica successivi: uno in corrispondenza dell’ultimo fallimento ad un regime terapeutico contenente un NRTI ed almeno uno eseguito durante l’interruzione terapeutica. Si è osservata una differente velocità di scomparsa delle singole mutazioni conferenti resistenza agli NRTI, le stesse, inoltre, scompaiono indipendentemente le una dalle altre nella maggioranza dei pazienti. La scomparsa delle mutazioni M184V/I e K65R, che si verifica nella maggior parte dei pazienti, è rapida e associata con la riemersione del ceppo wild-type, confermando l’impatto negativo di tali mutazioni sulla fitness virale. Complessivamente, la scomparsa delle mutazioni conferenti resistenza agli NRTI non sembra essere un processo ordinato e nella maggior parte dei pazienti avviene senza specifiche interazioni tra le mutazioni.

- HBV
Al fine di studiare i differenti cluster di mutazioni associate a resistenza alla lamivudina nel virus dell’epatite B di genotipo A e D, sono state analizzate le sequenze della trascrittasi inversa del virus estratto dal plasma di 89 pazienti infetti. Il prelievo è stato eseguito durante un regime terapeutico fallimentare alla lamivudina. L’associazione delle mutazioni con il genotipo virale è stata determinata con il test del Chi-quadrato e con un’analisi di regressione logistica multivariata. L’analisi di covariatione si basa su un’analisi di linkage gerarchico tramite la quale si costruisce una struttura ad albero che, a partire da un gran numero di cluster di piccole dimensioni, porta ad ottenere attraverso l’aggregazione successiva delle coppie di cluster più vicini, un unico grande cluster contenente tutte le associazioni tra mutazioni. Nel virus di genotipo A la presenza della mutazione M204V (prevalenza: 68.2%) nella trascrittasi inversa è il principale indice di fallimento alla terapia con lamivudina. L’analisi multivariata ha confermato che il genotipo A è l’unico fattore predittivo della comparsa della M204V (OR: 14.5[95% CI: 1.3-158], P=0.02). L’analisi di covariatione mostra che la M204V forma un cluster con le mutazioni della trascrittasi inversa L180M e L229V (corrispondente alla mutazione F220L nell’antigene S) e con la mutazione S207N nell’antigene S (bootstrap=0.95). Sia la mutazione F220L che la mutazione S207N sono localizzate nel quarto dominio transmembrana dell’antigene S. Al contrario, nel virus di genotipo D le mutazioni della trascrittasi M204V e M204I hanno una prevalenza simile (39.1% versus 45.3%, P=0.47) ed hanno differenti pattern di mutazioni compensatorie. La M204V forma un cluster con altre mutazioni localizzate nel dominio B della trascrittasi inversa (rtV173L, rtL180M, and rtT184A/S) (bootstrap=0.96), mentre la M204I forma un cluster con mutazioni localizzate nel dominio A della trascrittasi inversa (rtS53N, rtT54Y, and rtL80I/V) (bootstrap=0.96), senza presentare specifiche interazioni con mutazioni dell’antigene S. Il genotipo di HBV gioca, quindi, un ruolo importante nell’evoluzione della trascrittasi inversa sotto la pressione selettiva esercitata dalla terapia con lamivudina, il genotipo di HBV può’, pertanto, avere rilevanza in termini di diagnosi, risposta immunologica e progressione della malattia.
ABSTRACT

- **HIV-1**
  To date, very little information is available regarding the evolution of drug-resistance mutations during treatment interruption (TI). Using survival analysis approach, we investigated the dynamics of mutations associated with HIV-1 resistance to nucleoside-analogue reverse transcriptase inhibitors (NRTIs) during TI. Analyzing 132 patients having at least two consecutive genotypes, one at last NRTI-containing-regimen failure, and at least one during TI, we observed that the NRTI-resistance mutations disappear at different rate during TI, and are lost independently with each other in the majority of patients. The disappearance of K65R and M184I/V mutations occurred in the majority of patients, was rapid, and was associated with the re-emergence of wild-type virus, thus showing their negative impact on viral fitness. Overall, it seems that the loss of NRTI drug-resistance mutations during TI is not an ordered process, and, in the majority of patients occurs without specific interaction among mutations.

- **HBV**
  To investigate the different clusters of mutations associated with lamivudine resistance in HBV D- and A-genotypes, we analyzed HBV-RT sequences from 89 patients HBV-infected patients failing lamivudine- treatment. The association of mutations with HBV-genotypes was assessed by Chi-Squared test and multivariate logistic-regression analysis. Covariation-analysis was based on hierarchical-clustering. In A-genotype, the rtM204V (prevalence: 68.2%) is the main sign of lamivudine-failure. Multivariate-analysis confirms that A-genotype is the only predictor for rtM204V emergence (OR:14.5[95%CI:1.3-158], P=0.02). Covariation-analysis shows that rtM204V clusters with rtL180M, rtL229V (corresponding to sF220L in the HBsAg), and, interestingly, with HBsAg-mutation sS207N (bootstrap=0.95). Both sF220L and sS207N co-localize in the fourth transmembrane HBsAg-domain. Differently, in D-genotype the primary mutations rtM204V and rtM204I occur with similar prevalence (39.1% versus 45.3%, P=0.47), and show a distinct pattern of compensatory mutations. rtM204V clusters with mutations localized in the RT-B domain (rtV173L, rtL180M, and rtT184A/S) (bootstrap=0.94), while rtM204I clusters with mutations localized in the RT-A domain (rtS53N, rtT54Y, and rtL80I/V) (bootstrap=0.96) (without associations with HBsAg specific mutations). HBV-genotype plays an important role in driving RT-evolution under lamivudine-treatment, and thus can be relevant for therapeutic sequencing, immunological response and disease progression.
Table of Contents

1. General Introduction

1.1 HIV

111 Genome and morphology of the mature virion
112 Replication cycle
113 Pathogenesis
114 Epidemiology

1.2 HBV

121 Genome and morphology of the mature virion
122 HBV polymerase Similarities with HIV polymerase
123 Replication cycle and host immune response
124 Natural history of HBV infection
125 HBV Epidemiology

1.3 Antiviral Therapy

131 HIV
132 HBV
133 Antiviral Therapy for HIV-HBV Coinfected Patients

1.4 Drug Resistance

141 Mechanism of drug resistance in HIV-1 NRTIs
142 Drug resistance in HBV NRTIs

1.5 Differences and similarities between HIV-1 and HBV

1.5.1 Differences
1.5.2 Similarities

2. References
3. Results  

3.1 HIV  

3.1.1 Dynamics of NRTI Resistance Mutations During Therapy Interruption  

I) Introduction  

II) Materials and Methods  

- Study Population  
- HIV Sequencing  
- Mutations  
- Survival analysis of mutations dynamics  

III) Results  

- Baseline Patient characteristics  
- Dynamics of disappearance of individual NRTI mutations during TI  
- Revertants at RT position 215  
- Appearance of NRTI resistance mutations not present at baseline  
- The role of mutations interactions on dynamics of disappearance of NRTI resistance mutations during TI  

IV) Discussion  

V) References  

3.2 HBV  

3.2.1 The Profile of Mutational Clusters Associated with Lamivudine Resistance Can Be Constrained by HBV-genotypes  

I) Introduction  

II) Materials and Methods  

- Patients  
- HBV Sequencing  
- Mutations prevalence
III) Results

- Patients’ characteristics
- HBV RT mutations
- HBV HBsAg mutations
- Association among mutations
- Clusters of correlated mutations
- Correlation between duration of LMV-treatment and number of drug resistance mutations

IV) Discussion

V) References
1. General Introduction

1.1 HIV

Human immunodeficiency virus (HIV) is the etiological agent of the acquired immunodeficiency syndrome (AIDS) (Barre-Sinoussi et al., 1983; Coffin et al., 1986; Gallo et al., 1983; Levy et al., 1984). The basic pathology in AIDS is the loss of CD4⁺ lymphocytes and a variety of disorders in immune function, leading to the onset of opportunistic infections like Pneumocystis Carinii pneumonia and malignancies as Kaposi’s sarcoma or B-cell lymphomas (Levy, 1998). HIV was first isolated in 1983 (Barre-Sinoussi et al., 1983), and by the mid-1980s it was evident that two types of HIV, with slightly different genome structures, were circulating in human populations: the most common HIV-1, which is responsible to the world-wide AIDS epidemic and the immunologically distinct HIV-2 (Clavel et al., 1986), which is much less common and less virulent (Ariyoshi et al., 1999; Ariyoshi et al., 2000), but produces clinical findings similar to HIV-1 (Wilkins et al., 1993). Both viruses are characterized by extensive genetic diversity; HIV-1 is philogenetically divided in three groups: M, N, and O, with the M group further split into 9 subtypes and 15 circulating recombinant forms. Today, group M has a near global distribution, whereas groups N and O are restricted to individuals of West African origin. HIV-2 is also most common in individuals from West Africa and is composed of 7 subtypes. Despite its initial association with homosexual men, HIV-1 and HIV2 are now primarily transmitted by heterosexual intercourse (Walker P.R. et al., 2003) and from mother to infant.

HIV is a member of the lentivirus genus of the Retroviridae family. Like all retroviruses, HIV virions contain two identical copies of a single stranded RNA genome which are used as templates by the RNA-dependent DNA polymerase [Reverse Transcriptase (RT)] for production of DNA. Retroviruses were traditionally divided into three subfamilies, based primarily on pathogenicity rather than on genome relationship (oncoviruses which cause neoplastic disorders, spumaviruses which give cytopathic effect in tissue culture but apparently not associated with any known disease
and lentiviruses which induce slowly progressing inflammatory, neurological and immunological diseases). International committee on the taxonomy of viruses has recognized seven distinct genera in the Retroviridae family (Table 2.1) (Fields, 2001). Members belonging to the family of Retroviridae are widespread in nature and have been detected in many vertebrate species. Retroviruses show a great diversity in the interaction with their hosts. In fact, infection with some retroviruses can result in fatal diseases like AIDS, malignancies, neurological disorders and other pathological manifestations. Other retroviruses in contrast induce only benign viremia without apparent effects and some can even become established in the germ-line, where they persist as endogenous proviruses (Fields, 2001).

Table 1.1 Retroviruses Genera.

<table>
<thead>
<tr>
<th>Retrovirus Genera</th>
<th>Examples</th>
<th>Morphology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha retrovirus</td>
<td>Avian leukemia virus (ALV)</td>
<td>C-type</td>
</tr>
<tr>
<td></td>
<td>Rous sarcoma virus (RSV)</td>
<td></td>
</tr>
<tr>
<td>Beta retrovirus</td>
<td>Mouse mammary tumor virus (MMTV)</td>
<td>B-, D-type</td>
</tr>
<tr>
<td></td>
<td>Mason-Pfizer monkey virus (MPMV)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jaagsiekte sheep retrovirus (JSRV)</td>
<td></td>
</tr>
<tr>
<td>Gamma retrovirus</td>
<td>Murine leukemia viruses (MuLV)</td>
<td>C-type</td>
</tr>
<tr>
<td></td>
<td>Feline leukemia virus (FeLV)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gibbon ape leukemia virus (GaLV)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reticuloendotheliosis virus (ReT)</td>
<td></td>
</tr>
<tr>
<td>Delta retrovirus</td>
<td>Human T-lymphotropic virus (HTLV)-1, -2</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Bovine leukemia virus (BLV)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Simian T-lymphotropic virus (STLV)-1, -2, -3</td>
<td>—</td>
</tr>
<tr>
<td>Epsilon retrovirus</td>
<td>Walleye dermal sarcoma virus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Walleye epidermal hyperplasia virus 1</td>
<td></td>
</tr>
<tr>
<td>Lentivirus</td>
<td>Human immunodeficiency virus type 1 (HIV-1)</td>
<td>Rodcone core</td>
</tr>
<tr>
<td></td>
<td>HIV-2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Simian immunodeficiency virus (SIV)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Equine infectious anemia virus (EIAV)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Feline immunodeficiency virus (FIV)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Caprine arthritis encephalitis virus (CAEV)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vifamae cat virus</td>
<td></td>
</tr>
<tr>
<td>Spumavirus</td>
<td>Human foamy virus (HFV)</td>
<td>Immature</td>
</tr>
</tbody>
</table>

The retrovirus family is split up into 7 genera: the Alpharetroviruses, the Betaretroviruses, Gammaretroviruses, Deltaretroviruses, Epsilonviruses (all of which used to be classified as one genus, the oncoviruses), the Lentiviruses (which includes HIV) and the Spumaviruses.

The lentiviruses are exogenous, non-oncogenic retroviruses causing persistent infections leading to chronic diseases with long incubation periods (lenti for slow). They usually infect cells of the
immune system (dendritic cells, macrophages, T lymphocytes) and cause cytopathic effects in permissive cells, such as syncytia and cell death by apoptosis and necrosis.

The prototype members of the lentivirus family were the slow visna virus, the equine infectious anemia virus (EIAV) and the caprine arthritis-encephalitis virus (CAEV). More recent isolates include the related human and simian immunodeficiency viruses (HIV and SIV) (Table 1.1). Like the human T-cell leukemia virus (HTLV) family of primate onco-retroviruses, lentiviruses are complex retroviruses (Cullen, 1991). The significant characteristic of the complex retroviruses is the ability to regulate their own expression via virally encoded protein factors not found in other retroviruses. This property has been proposed to be essential for the long-term association of the complex retroviruses with the host and the generation of chronic active infections. The lentiviral complexity is reflected in their replication cycle, which reveals intricate regulatory pathways, unique mechanisms for viral persistence (Tang et al., 1999) and the ability to infect non-dividing cells.

1.1.1 Genome and morphology of the mature virion

Like all other retroviruses, the proviral DNA of HIV-1 (~ 9.2 kb) has three coding regions gag, pol, and env, two long terminal repeats (LTRs) with transcriptional regulatory elements, and one primary transcript (Fig. 1.1.1a and 1.1.1b) [for review see (Frankel and Young, 1998; Turner and Summers, 1999)]. The RNA genome is flanked by two short redundant (R) sequences at both termini with adjacent unique sequences, U5 and U3, found at the 5’ and 3’ ends, respectively. The gag gene encodes the large precursor polyprotein p55 that is cleaved in four proteins: the matrix p17 (MA), the "core" capsid p24 (CA), the nucleocapsid p7 (NC) and the p6 (Freed, 1998). The pol gene encodes the synthesis of three important enzymes that function at different times during the replicative cycle. The reverse transcriptase/ribonuclease H complex p51 and p66 RT acts in the early steps of the virus replication to form a double-stranded cDNA of the virus RNA. The integrase p32 (IN) mediates integration of the viral cDNA into the host chromosomal DNA. The protease p10
(PR) is responsible for the cleavage of the viral Gag and Gag-Pol polyproteins during the maturation of the viral particle. The *env* gene directs the production of an envelope precursor protein gp160, which undergoes cellular proteolytic cleavage into the outer envelope glycoprotein gp120 (SU), responsible for binding to CD4<sup>+</sup> receptors, and the transmembrane glycoprotein gp41 (TM), which catalyzes the fusion of HIV to the target cell's membrane.

In addition, HIV-1 has at least six more genes encoding viral proteins with regulatory functions (*tat* and *rev*) or accessory functions (*nef*, *vif*, *vpr* and *vpu*), for reviews see (Cullen, 1998; Emerman and Malim, 1998; Frankel and Young, 1998; Kjems and Askjaer, 2000; Piguet and Trono, 1999; Pollard and Malim, 1998; Trono, 1995).

![HIV-1 genomic organization](image)

**Fig. 1.1.1a** HIV-1 genomic organization. Like all other retroviruses, HIV-1 has three structural genes *gag*, *pol* and *env* (heavily shaded), which are flanked by the long terminal repeats (LTR’s). In addition it has six more genes, including two regulatory genes *tat* and *rev* (stippled) and four accessory genes *nef*, *vif*, *vpr* and *vpu* (white).
Fig. 1.1.1b The immature and mature forms of HIV-1. Typical lentivirus particles are spherical, about 80-110 nm in diameter, and consist of a lipid bilayer membrane surrounding a conical core. The two identical single-stranded RNA (ssRNA) molecules, of about 9.2kB each, are associated with the nucleocapsid proteins p7gag (NC). They are packed into the core along with virally encoded enzymes: reverse transcriptase (RT), integrase, and protease. P24gag comprises the inner part of the core, the capsid (CA). The p17gag protein constitutes the matrix (MA) which is located between the nucleocapsid and the virion envelope. The viral envelope is produced by the cellular plasma membrane and contains the protruding viral Env glycoproteins: gp120 surface glycoprotein (SU) and gp41 transmembrane protein (TM).
1.1.2. Replication cycle

General features of the HIV-1 replication cycle are shown in Fig. 1.1.2. The HIV-1 replication cycle begins with the recognition of the target cell by the mature virion. The major targets for HIV-1 infection are cells bearing the HLA class II receptor, CD4, on their cell surfaces. These include T-helper lymphocytes and cells of the monocyte/macrophage lineage including microglia cells in the brain. The virus-CD4 binding occurs via specific interactions between the viral outer envelope glycoprotein gp120 and the amino-terminal domain of CD4 (Dalgleish et al., 1984; Klatzmann et al., 1984). These interactions are sufficient for binding but not for infection. Subsequently the virus glycoprotein gp120 interacts with additional cell-surface proteins to promote fusion of the viral and cellular membranes. These coreceptors have been identified to be members of the chemokine receptor family and include CXCR4 and CCR5 (Alkhatib et al., 1996; Deng et al., 1996; Doranz et al., 1996; Moore, 1997). The initial binding of HIV to the CD4 receptor is mediated by conformational changes in the gp120 subunit, followed by a conformational change in the gp41 subunit, stimulated by the chemokine receptors, that allows fusion and subsequent entry of HIV.

Various strains of HIV differ in their use of chemokine coreceptors. There are strains of HIV known as T-tropic strains, which selectively interact with the CXCR4 chemokine coreceptor of lymphocytes, while M-tropic strains of HIV interact with the CCR5 chemokine coreceptor of macrophages and dual tropic HIV strains that infect both cell types (Littman, 1998; Moore, 1997). HIV-1 infection of CD4 negative cells, such as neural cells, has also been reported (Clapham et al., 1989; Harouse et al., 1989; Kozlowski et al., 1991; Kunsch et al., 1989) but the mechanisms of HIV entry are still unclear.

Membrane fusion is followed by a poorly understood uncoating event that allows the intracellular reverse transcription.

The viral RNA is transcribed in the cytosol into double stranded DNA by the RT (Hansen et al., 1987; Muesing et al., 1985). This enzyme possesses three enzymatic activities: RNA-dependent
DNA polymerase, DNA-dependent DNA polymerase, and ribonuclease H (RNase H). The reverse transcription process takes place within a large nucleic acid-protein complex known as the preintegration complex (PIC) by the assistance of the accessory protein Vif (von Schwedler et al., 1993) and the nucleocapsid protein NC (Darlix et al., 1993). Once synthesized, the viral DNA is transported to the nucleus of the infected cell as part of the PIC that appears to include tightly condensed viral nucleic acids and the IN, MA, RT, and Vpr proteins. In contrast to other retroviruses, that require cell division and concomitant breakdown of the nuclear envelope to gain access to the nuclear compartment, the lentiviral PIC is actively imported into the nucleus during the interphase (Bukrinsky et al., 1992; Lewis and Emerman, 1994). Nuclear import of the PIC seems to be directed by the accessory protein Vpr (Fouchier et al., 1997; Heinzinger et al., 1994), the Gag matrix protein MA (Bukrinsky et al., 1993; von Schwedler et al., 1994) and the integrase IN (Gallay et al., 1997). Vpr does not contain a conventional nuclear localization signal (see later) but appears to function by connecting the PIC to the cellular nuclear import machinery (Fouchier et al., 1998; Popov et al., 1998a; Popov et al., 1998b). Recently Nitahara-Kasahara Y et al. shown that an interaction between importin alpha and the N-terminal alpha-helical domain of Vpr is indispensable, not only for the nuclear import of Vpr but also for HIV-1 replication in macrophages. The role of the MA protein in nuclear translocation of PIC remains controversial (Fouchier et al., 1997; Freed et al., 1995). The ability of lentiviruses such as HIV-1 to utilize active transport mechanisms for translocation of the PIC into the nucleus, allows these viruses to infect non-dividing cells such as differentiated macrophages, quiescent T lymphocytes and possibly neurons. In the nucleus, IN catalyzes covalent integration of the viral DNA into the host genome, where it resides permanently as a provirus. An important modification as a result of reverse transcription and integration is the duplication of the U5 and U3 sequences in the LTR, such that the provirus now is flanked by tandemly repeated sequences U3-R-U5 with important regulatory functions. The regulation of the HIV-1 transcription involves a complex interplay between cis-acting DNA and RNA elements present within the chromatin-associated proviral LTRs, cellular transcription
factors and the viral regulatory protein Tat (transcriptional transactivator). In an arrangement similar to that of several inducible cellular promoters, the HIV-1 promoter, which is located in the U3 region of the 5'LTR, contains a TATA box and binding sites for several cellular DNA-binding transcription factors, such as NF-kB, Sp1 and TBP (Jones and Peterlin, 1994). It is highly inducible and responds to the activation status of the infected cell. NF-kB is the major inducible cellular activator. It is well established that many cells in the lymphoid tissue of infected individuals are latently infected (Pantaleo et al., 1993), even though the viral replication in the body is always active. In resting T-cells, the activity of the HIV-1 promoter is minimal, leading to viral quiescence in infected primary cells. Therefore, viral activation is associated with cell activation.

The transcription of the provirus by the cellular RNA polymerase II results in a primary transcript that may serve three distinct functions: 1) it constitutes genomic RNA that is incorporated into the virion; 2) it serves as template for translation (Gag and Gag-Pol); 3) it functions as the precursor RNA for the production of diverse subgenomic mRNAs (Fig 2.3). As mentioned before, HIV-1 encodes two essential regulatory proteins Tat and Rev, which increase viral gene expression at the transcriptional and post-transcriptional levels, respectively. HIV-1 mRNA expression is biphasic and can be divided into early (Rev-independent) and late (Rev-dependent) stages (Kim et al., 1989; Pomerantz et al., 1990). First, shortly after the infection of cells, multiply spliced (~ 2kb) RNA species are formed from the primary transcript and three proteins are produced: Tat, Rev and Nef, therefore referred as early gene products (Schwartz et al., 1990). Tat [for reviews see (Cullen, 1998; Emerman and Malim, 1998; Rubartelli et al., 1998)], greatly increases transcription from the HIV promoter, by binding to a cis-acting target sequence, the trans-activator response element (TAR), which is located at the 5' end of the nascent viral RNA transcript (Berkhout et al., 1989; Dingwall et al., 1989). Tat recruits two cellular factors to this complex: cyclin T and cyclin-dependent protein kinase-9 (Cdk9). Cyclin T is proposed to bind directly Tat and to increase its affinity for the TAR RNA (Wei et al., 1998). Cdk9 phosphorylates the RNA polymerase II transcription complex and thus stimulates transcriptional elongation (Herrmann and Rice, 1995; Reines et al., 1996; Wei et al.,
1998). Rev (regulator of expression of the virion), which accumulates during the early phase of expression, initiates late gene expression by binding a unique RNA element located in the env coding region of HIV-1, the so called Rev-responsive element (RRE). This interaction promotes the stability and transport of unspliced (~ 9 kb) and partially spliced (~ 4 kb) HIV-1 mRNAs out of the nucleus. These mRNAs are responsible for the production of the viral enzymes and structural proteins (Daly et al., 1989; Felber et al., 1989; Hammarskjold et al., 1989; Malim et al., 1989). Therefore Gag, Pol, Env, Vif, Vpr, and Vpu proteins are referred to as late HIV-1 proteins.

Figure 1.1.2 Replication cycle of HIV-1.
Controversial functions have been assigned to the Nef (negative factor) protein. First, it was assumed to have an inhibitory function in the transcription of HIV-1 genes (Ahmad and Venkatesan, 1988). However, recent studies have contradicted this concept by showing that Nef is important for virus replication. Among its various functions (Piguet and Trono, 1999), it enhances viral expression in quiescent cells and mediates lymphocyte chemotaxis and activation at sites of virus replication (Kestler et al., 1991; Kirchhoff et al., 1995; Koedel et al., 1999; Miller et al., 1994; Swingler et al., 1999).

The Env precursor polyprotein (gp160) is synthesized in the endoplasmatic reticulum (ER) where it is glycosylated and appears to oligomerize to a trimeric structure post-translationally (Wyatt et al., 1998; Wyatt and Sodroski, 1998). Thereafter, it is cleaved to produce the non-covalently associated (gp41 TM - gp120 SU)3 trimeric glycoprotein complex, which is transported to the cell membrane for virus assembly. Vpu is thought to enhance this process and inhibit a premature trapping of CD4 to Env in the ER by binding CD4 molecules, which are also synthesized in the ER, and directing them to the ubiquitin-proteasome degradation pathway (Margottin et al., 1998; Schubert et al., 1998; Strebel et al., 1988; Willey et al., 1992a; Willey et al., 1992b). Similarly, the accessory protein Nef facilitates the routing of CD4 from cell-surface and Golgi apparatus to lysosomes, resulting in endosomal degradation and preventing inappropriate interaction with Env (Aiken et al., 1994). In addition, both Vpu and Nef can down-regulate expression of MHC class I molecules. The down-regulation of CD4 and MHC class I molecules on the surface of infected cells also helps infected cells to evade immune responses of the host, such as killing by cytotoxic T lymphocytes (Collins et al., 1998; Kerkau et al., 1997). During synthesis of the Gag polyprotein by ribosomes, a translational frameshift may occur, resulting in generation of smaller amount of Gag-Pol precursor polyproteins, which associate with the Gag polyprotein at the cellular membrane. The N-terminally myristoylated MA domain of the Gag/GagPol polyproteins directs insertion of the Gag precursors into the cellular membrane and interacts with the cytoplasmic tail of gp41 resulting in the anchoring of Env to the viral particle (Dorfman et al., 1994). Approximately 1200 to 2000 copies of Gag
precursor bud to form an immature particle, which encapsidates two copies of the unspliced viral RNA genome, by the ability of NC to interact with nucleic acids. Vif and Vpu proteins have been reported to play a role in packaging of the nucleoprotein core and in virion release, respectively (Hoglund et al., 1994; Lamb and Pinto, 1997). Concomitantly or immediately following the external budding, the cleavage of the Gag/Gag-Pol polyproteins by the virally encoded PR produces the structural proteins MA, CA, NC as well as the independent enzymes PR, RT and IN. This final step primes new virus particles for the next round of infection and is termed maturation.

1.1.3 Pathogenesis

AIDS

HIV infection has been associated with the acquired immunodeficiency syndrome (AIDS). A diagnosis of AIDS is made whenever a person is HIV-positive and have:

- CD4+ T cell count below 200 cells/mms
- CD4+ T cells account for fewer than 14% of all lymphocytes
- Diagnosis with one or more of the 25 AIDS defining illness, including various opportunistic infection, brain and nerve disease, certain cancers, and wasting syndrome

Approximately 10% of HIV-infected patients progress to AIDS within the first 2 to 3 years of infection, while for approximately 40% this progression is observed over a period of 10 years. 10% to 17% of HIV-infected patients may be AIDS free, some with no evidences of disease progression. These variations in responses may be due to differences in the degree of stimulation of the immune system by infection with the other pathogens as well as to viral factor, such as deletions in the nef gene or altered cell tropism (Kupfer et al., 1998).

Course of infection

Schematically, the course of infection can be divided into an acute, an asymptomatic, and symptomatic phase (Fig. 1.1.3). The acute phase accounts for the first 5-10 weeks of infection and is characterized by high virus production, and activation of lymphocytes in lymphonodes. Up to
5x10^3 infectious particles per ml of blood plasma may be found in the first days after infection. This viremia is curtailed within a few weeks and level off at the beginning of the asymptomatic phase to the so-called virological set point, that is a predictor of disease progression. During this CD4+ cells numbers decrease at a low steady rate, while virus replication remains constant at a low rate. The duration of the asymptomatic phase may last between 2 and 20 years. The end stage of disease, when the patient develops AIDS, is characterized by CD4+ cells count below 200 copies/ml and increased quantities of the virus. The number of CD8+ cytotoxic lymphocytes also decreases and lymphoid cells and tissues are damaged.

**CD4+T cell depletion**

The hypothesis that CD4+ cell depletion is caused the lysis of infected cells during viral replication has been supported by the observation of an immediate and large increase of CD4+ count after the initiation of antiretroviral therapy that blocks viral replication (Ho et al., 1995; Wei et al., 1995). This hypothesis has not withstood more detailed analyses of T cell dynamics (Roederer et al., 1998). In fact, it has been turned out that in HIV-infected patients all T cell subset are progressively destroyed, irrespective of CD4+ expression, and AIDS appear to be a disease of perturbed homeostasis. Many pathogenetic mechanisms have been proposed, including viral gene products, syncitium formation, direct virus killing of cell, apoptosis, autoimmunity, cytokine and chemokines expression, superantigens, virus directed cell mediated cytolysis and disruption of lymphoid architecture.
1.1.4 Epidemiology

Several African primates harbour lentiviruses and HIV is entered the human population in Africa by zoonotic transmission of SIV$_{cpz}$ from chimpanzee population. The first cross species transfer has been estimated to have occurred between 1915 and 1941 (Korber et al., 2000). Two types of HIV are known: the most common HIV-1, which is responsible to the world-wide AIDS epidemic and the immunologically distinct HIV-2 (Clavel et al., 1986), which is much less common and less virulent (Ariyoshi et al., 2000; Ariyoshi et al., 1999), but produces clinical findings similar to HIV-1 (Wilkins et al., 1993). The HIV-1 type itself includes three groups M, N, O, which have different geographic distributions but all produce similar clinical symptoms (Fig. 1.5). The M group is further divided into 9 pure subtypes (A, B, C, D, F, G, H, J, K), 4 sub-subtype (A1, A2, F1, F2) and 16 circulating recombinant forms on the basis of phylogenetic analysis. Almost all subtypes are present in Africa, while in Europe, North America, and Australia subtype B is more dominant, and subtype C is more common in Asia (McCutchan et al., 2000; Robertson et al., 2000).
About 40.4 million adults and children have been estimated to live with HIV/AIDS, most of them in Sub-Saharan Africa and South East Asia (Fig. 1.1.4b). Only a minority of HIV-infected individuals live in the industrialized countries and has access to the anti-HIV drugs and professional health care.

*Figure 1.1.4a. Phylogenetic relationship of primate lentiviruses.* Phylogenetic tree derived from the alignment of *pol* gene sequences of HIV-1, HIV-2 and SIV strain SIVsam2. SIV of sooty mangabeys or macaques experimentally infected with SIVsam2. Reproduced from *Human Retroviruses and AIDS* (Los Alamos, New Mexico; Theoretical Biology and Biophysics Group, Los Alamos National Laboratory) (Kuiken *et al.*, 1999).

*Figure 1.1.4b. Geographical distribution of HIV/AIDS cases.*
1.2. HBV

Hepatitis B virus (HBV) infection is a major public health problem, with approximately 350 million individuals chronically infected worldwide (Lee et al., 1997). HBV is highly endemic in sub-Saharan Africa, China and South-East Asia. It is also highly endemic in the Mediterranean basin and it is present at significant levels in most industrialized countries (Lee et al., 1997). Two forms of chronic HBV infection can be individualized according to the presence or the absence of HBe antigen, but transitional forms exist (Ganem and Prince, 2004; Lee et al., 1997). Chronic HBV carriers are exposed to a risk of complications such as chronic hepatitis, cirrhosis, and hepatocellular carcinoma, of which HBV is currently the most frequent cause (Ganem and Prince, 2004). Up to one million people die every year from the complications of HBV infection (Lee, 1997).

HBV is a member of the family *hepadnaviridae*, viruses with a double-stranded circular DNA genome that replicates through an RNA intermediate. Those most closely related to HBV have been found in woodchucks (Summers et al., 1978) and ground squirrels (Tennant et al., 1991). These viruses have about 70% sequence homology to HBV but are not known to infect humans or other primates; in contrast, HBV is infectious for the great apes. Because of similarities in DNA sequence and genome organization, the viruses infecting mammals are grouped in the genus *Orthohepadnavirus*. More distantly related viruses, with somewhat similar genome organization but almost no sequence homology, are found in ducks and geese (Mason et al., 1980; Sprengel et al., 1988). These are grouped in the genus *Avihepadnavirus*. Duck hepatitis B virus (DHBV) has been used primarily as a model system to characterize how hepadnaviruses replicate. HBV is a variable virus, due to the intrinsic properties of the HBV DNA polymerase, the enzyme that ensures viral replication. The quasispecies distribution of HBV is characterized by the coexistence of different viral populations in various proportions. Variant populations are continuously selected by the changing environment in which the virus replicates during human infection.
1.2.1. Genome and morphology of the mature virion

The HBV virus is detectable in serum in high levels and on electron microscopy is a 42 nm double-shelled particle with an outer envelop (HBsAg) and an inner nucleocapsid (HBCAg) protein. The infectious virion circulates as a 42-nm Dane particle that comprises a nucleocapsid surrounded by a lipid envelope (fig 1.2.1a). The nucleocapsid contains a partially double-stranded 3.2-kilobase genome. The compact genome encodes 4 overlapping reading frames and 4 RNA species that encode for 7 viral proteins: 3 envelope proteins—large, middle, and small that form the HBsAg, the nucleocapsid core protein (HbcAg), the secretory hepatitis B e antigen (HbeAg), the viral reverse transcriptase/polymerase, and the X protein (Dryden et. al. 2006; Ganem et al., 1987).

As shown in fig.1.2.1a the HBV virion is surrounded by a lipid envelope in which the Small-Medium and Large surface protein are inserted. The three gene products termed large (L)-, middle (M)- and small (S)-proteins are encoded by the S gene (Fig. 1.2.1b). They share the C-terminal S-domain with four transmembrane regions. The M- and L-proteins carry additional N-terminal extensions of 55 and, genotype-dependent, 107 or 118 a.a. (preS2- and preS1). The stoichiometric ratio of L/M/S proteins in virions is B1:1:4, whereas the more abundantly secreted noninfectious subviral particles contain almost exclusively S- and only traces of L-protein (Bruss et al., 2007).

During synthesis, the preS1-domain of L is myristoylated and translocated through the endoplasmic reticulum (ER). This modification and the integrity of the first 77 amino acids of preS1 are essential for infectivity (Gripon et al., 1995; LeSeyec et al., 1999). Classic HBsAg, which contains the S domain only, is also called the S-protein (24 kD). The glycosylation of the S domain gives rise to two isoforms of each protein. In addition, the M protein contains an N-linked oligosaccharide on its pre-S2-specific domain, and the L protein carries a myristic acid group in amide linkage to its amino-terminal glycine residue. While the function of M protein is still obscure, L proteins play a role in viral assembly and infectivity. L proteins carry the receptor recognition domain, which allows efficient binding to cell surface receptors. Two in-frame AUG codons are present in ORF C. Classic HBcAg (21 kD) is the product of initiation from the more internal start codon, while
initiation at the upstream AUG produces a C-related protein that is not incorporated into virions but instead is independently secreted from cells, accumulating in serum as an immunologically distinct antigen known as HBeAg (16-18 kD). The function of HBeAg is still unknown. HBcAg is the most conserved polypeptide among the mammalian hepadnaviruses with 68% homology between HBV and GSHV and 92% between GSHV and WHV. Core proteins spontaneously assemble into forms resembling core particles. The polymerase protein is a DNA-dependent DNA polymerase, a reverse transcriptase, an RNAse H, and it binds to the 5' end of HBV DNA, acting thus as a primer for reverse transcription of the pregenome, an RNA intermediate, to form negative strand DNA. Furthermore, it plays important roles in the encapsidation of the viral pregenomic RNA. The polymerase protein is quite immunogenic during both acute and chronic infection. ORF X encodes the protein X (17 kD), a transactivator for the viral core and S promoters. The X protein is the least-conserved protein among hepadnaviruses with only 33% amino acid homology between GSHV and HBV, and 71% between the two rodent viruses (WHO website 2008).

**Fig.1.2.1a:** The HBV virus is a double-shelled particle with an outer envelope (HBsAg) and an inner icosahedral nucleocapsid (HBcAg) protein surrounded by a lipid envelope (HbsAg).
Fig1.2.1b: Schematic representation of the HBV particle and the HBV L-protein. The partially double stranded DNA is covalently associated with the viral polymerase complex, consisting of the terminal protein (TP), the reverse transcriptase (RT) and the RNaseH. The genome is encapsulated by an icosahedral shell, comprising 120 core-protein dimers. The three HBV surface proteins (L-, M- and S-proteins) are embedded into an ER-derived lipid bilayer. The L- and M-proteins contain the complete S-domain (red), which serves as a membrane anchor.

The genome exists in two different forms (Ganem and Schneider, 2001). In the virions, the genome is a relaxed circular DNA molecule (RC-DNA) that is only partially double-stranded (Fig. 1.2.1c). One strand is complete (the L or minus strand) and even has a short terminal redundancy with a protein, the viral polymerase, covalently attached to the 5’-end. The other strand (the S or plus strand) is incomplete. The 5’-end is fixed and starts with a short oligoribonucleotide and the 3’-end is variable. The plus strand overlaps the 5’ and 3’ extremities of the minus strand, thereby assuring the circularity of RC-DNA. RC-DNA therefore contains all of the genetic information of the virus but is unsuitable as a replication template. After infection of a hepatocyte, plus strand DNA is completed, the oligoribonucleotide at the 5’-end of plus strand DNA and the terminal redundancy of minus strand DNA, along with the attached viral polymerase, are eliminated and the two strands are ligated. Importantly, these steps are apparently performed by cellular enzymes, independently of the viral polymerase. The genome is then found in the nucleus of the infected hepatocyte in the form of a covalently-closed circular DNA molecule (cccDNA) that is the real
replication matrix. The mRNAs for the viral proteins are transcribed from the cccDNA. This includes pregenomic RNA (pgRNA) which is of more-than-genome length, approximately 1.1 genomes.

**Fig 1.2.1c:** Genetic organization of the HBV genome (RC-DNA form). GRE, glucocorticoid response element; Enh, enhancer; PRE, post-transcriptional regulatory element; BCP, basal core promoter; DR, direct repeat.

### 1.2.2. HBV polymerase. Similarities with HIV polymerase

The HBV polymerase is a multifunctional enzyme. The products of the P gene are involved in multiple functions of the viral life cycle, including a priming activity to initiate minus-strand DNA synthesis, a polymerase activity, which synthesizes DNA by using either RNA or DNA templates, a nuclease activity which degrades the RNA strand of RNA-DNA hybrids, and the packaging of the RNA pregenome into nucleocapsids. Nuclear localisation signals on the polymerase mediate the transport of covalently linked viral genome through the nuclear pore.
The HBV polymerase has 4 domains: a priming region, a spacer region of unknown function, a catalytic region that functions as a RNA-dependent RNA polymerase/DNA polymerase, and a carboxy terminal region that has ribonuclease H activity (Figure 1.2.2a) (6. Seeger C, Mason WS. Hepatitis B virus biology. Microbiol Mol Biol Rev 2000;64:51–68). Although the crystal structure of HBV polymerase is unknown, much of its structure has been deduced from the human immunodeficiency virus-1 reverse transcriptase (HIV-1 RT) based on their homology. (19. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA. Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science 1992;256:1783–1790 and 21. Jacobo-Molina A) (Fig. 1.2.2b). Regardless of their amino acid sequences and differences in domain structure, all polymerases appear to have a common right-handed configuration with a thumb, a palm, and a fingers domain (Fig 1.2.2b) (Steitz TA et a., 1999; Doublie S et al., 1999). The palm domain appears to be the active site and catalyzes the phosphoryl transfer reaction; the fingers domain facilitates interactions with the incoming dNTPs as well as the template base to which it is paired; and the thumb domain may play a role in positioning the duplex DNA, processivity, and translocation (Steitz TA et a., 1999). Nucleoside analogues and dNTPs bind at a site that is located in the palm subdomain adjacent to the 3’ terminus of the primer strand (Tantillo C et al.,1994) An interesting property of the HBV polymerase seems to be its preference for nucleotides with the L-configuration in contrast to other polymerases that prefer nucleotides with the D-configuration (Davis MG et al., 1996). The catalytic region can be subdivided into 7 domains: A–G (Fig.1.2.2a). Domain A of HIV-1 RT is in close proximity to the 2 aspartic acid residues in domain C and forms part of the dNTP binding pocket (Jacobo-Molina et al., 1993). Residues in this domain are involved in the coordination of the incoming triphosphate moiety of the dNTP and the magnesium ions. Domain B for HBV RT forms an α helix with a loop region and is involved with positioning of the primer-template strand to the catalytic region (Jacobo-Molina et al.1993)? Domain C contains a sequence of 4 amino acids, tyrosine, methionine, aspartate, aspartate (YMDD), which is highly conserved among viral polymerases/reverse transcriptases that binds 2
magnesium ions and represents the active site of the enzyme (Jacobo-Molina et al., 1993; Poch O et al., 1989). Residues within domain D lie outside but may contribute to the dNTP binding site of HIV-1 RT. Mutations in this domain may indirectly affect the geometry of the dNTP binding site (Tantillo C et al., 1994). In HIV-1 RT, domain E forms part of the template-primer binding site (Kohlstaedt LA et al., 1992; Jacobo-Molina A et al., 1993). The primer strand contacts the loops between the palm and the thumb at rtM230 and rtG231.21 The methionine and glycine residues present in domain E are conserved in all HBV isolates. Domains F and G are upstream of domain A. This region may be involved in interactions with the incoming dNTP and also with the template nucleotide.

Fig. 1.2.2.a: Schematic of the HBV polymerase open reading frame illustrating the 4 functional domains and the 7 catalytic subdomains A–G. From Ghany et al., 2007.
1.2.3. Replication cycle and host immune response

The HBV virion binds to a receptor at the surface of the hepatocyte. A number of candidate receptors have been identified, including the transferrin receptor, the asialoglycoprotein receptor molecule, and human liver endonexin. The mechanism of HBsAg binding to a specific receptor to enter cells has not been established yet. Viral nucleocapsids enter the cell and reach the nucleus, where the viral genome is delivered. In the nucleus, second-strand DNA synthesis of partially double-stranded RC-DNA is completed and the gaps in both strands are repaired to yield a covalently closed circular (ccc) supercoiled DNA molecule that serves as a template for transcription of four viral RNAs that are 3.5, 2.4, 2.1, and 0.7 kb long. These transcripts are polyadenylated and transported to the cytoplasm, where they are translated into the viral nucleocapsid and precore antigen (C, pre-C), polymerase (P), envelope L (large), M (medium), S (small)), and transcriptional transactivating proteins (X).
The envelope proteins insert themselves as integral membrane proteins into the lipid membrane of the endoplasmic reticulum (ER). The 3.5 kb species, spanning the entire genome and termed pregenomic RNA (pgRNA), is packaged together with HBV polymerase and a protein kinase into core particles where it serves as a template for reverse transcription of negative-strand DNA. The RNA to DNA conversion takes place inside the particles.

The new, mature, viral nucleocapsids can then follow two different intracellular pathways, one of which leads to the formation and secretion of new virions, whereas the other leads to amplification of the viral genome inside the cell nucleus.

In the virion assembly pathway, the nucleocapsids reach the ER, where they associate with the envelope proteins and bud into the lumen of the ER, from which they are secreted via the Golgi apparatus out of the cell. In the genome amplification pathway, the nucleocapsids deliver their genome to amplify the intranuclear pool of covalently closed circular DNA (cccDNA).

The precore polypeptide is transported into the ER lumen, where its amino- and carboxy-termini are trimmed and the resultant protein is secreted as precore antigen (HBeAg).

The X protein contributes to the efficiency of HBV replication by interacting with different transcription factors, and is capable of stimulating both cell proliferation and cell death.

![HBV replication cycle](image)

Fig 1.2.3: HBV replication cycle. Figure from Guidotti et al., 2006.
HBV is non-cytopathic, and the cellular injury of hepatitis B appears immune-mediated (Chisari FV et al., 1995). Most persons exposed to HBV have a transient infection which may or may not be accompanied by symptoms and jaundice. In a proportion, however, chronic infection ensues as a result of failure of the host immune response to eliminate virus. Viral clearance is mediated by both cytopathic and noncytopathic mechanisms. Cytopathic clearance is based on direct killing of virus-harboring hepatocytes by virus-specific T cells, followed by compensatory proliferation of hepatocytes. This proliferation may also result in gradual loss of cccDNA. Uninfected hepatocytes are protected from reinfection through induction of neutralizing antibodies and innate immune mechanisms. Studies in the duck hepatitis B virus (DHBV) model support a cytopathic mechanism for viral clearance, as shown by the turn-over of large number of hepatocytes during viral clearance (Summers J et al., 2003). Studies in the HBV chimpanzee model and in transgenic animals, on the other hand, support the importance of non-cytopathic clearance as mediated by antiviral cytokines such as type I interferon and tumor necrosis factor alpha (TNFα) (Guidotti LG et al., 1999; Guidotti LG et al. 2001). In addition to the humoral response, cell-mediated immunity is crucial for the control of HBV infection (Chisari FV et al., 1995). Vigorous virus-specific CD4+ and CD8+ T cell responses which target multiple viral antigens with broad specificity are induced during transient infection, and these responses can persist for decades after viral clearance (Rehermann B et al., 1996). The continued presence of T cell responses suggest that HBV persists at low levels despite absence of detectable HBV DNA in serum and presence of antibodies to HbsAg (anti-HBs) (Liang TJ et al., 1991). The presence of low-levels of virus in the liver after apparent recovery has been demonstrated by transmission studies in chimpanzees and results of organ transplantation in humans (Liang TJ et al., 1990; Wachs ME et al., 1995). Reactivation of hepatitis B in previously recovered people during immunosuppressive therapy has been amply documented (Lok AS et al., 1991; Lau GK et al., 2007). Thus, recovery from hepatitis B may not indicate virus eradication so much as firm immunological control. (Hoofangle et al., 2007). Moreover it’s interesting to note that the immune response solicited by natural infection or vaccination can induce the emergence of
mutants in the S gene which, particularly those in the most immunogenic *a determinant*, are positively selected by their ability to evade immune defences. Since the serological diagnosis of active HBV infection is made by detecting HBsAg in different kinds of ligand assays, patients carrying immune escape mutations in HBsAg gave rise to false negative results. Recent studies performed with a series of the current commercial assays show that mutations involving residues 118–127 and 143–145 of HBsAg are of special importance, and may be implicated in some cases of occult HBV infection (Weber et al., 2006). Occult HBV is defined as persistent low level HBV replication in the absence of detectable HBsAg. Antibodies against core are usually the only detectable serologic marker, though serologically negative cases exist. Virologic characterization of occult HBV, particularly during HIV co-infection, has not been well studied.

### 1.2.4 Natural history of HBV infection

The natural history of chronic HBV infection involves a complex interplay between the virus and the host's immune system. It includes four phases of varying duration distinguished by the presence in the serum of HBeAg or its antibody (anti-HBe) and serum concentrations of HBV DNA and aminotransferases (Fig 1.2.4) (Lok AS et al., 2001; Fattovich G et al., 2003; McMahon BJ et al., 2004; and Hadziyannis SJ et al., 2006). Chronic HBV infection, which is common after neonatal or childhood infection, usually starts with an HBeAg-positive immune-tolerant phase, characterised by high viraemia, normal serum aminotransferases, and few histological changes. These patients gradually progress to the HBeAg seroconversion phase (HBeAg-positive chronic hepatitis B), which is characterised initially by positive HBeAg, high serum HBV DNA concentrations, increased aminotransferases, and active necroinflammation and/or fibrosis (Hoofnagle JH et al., 1981; Fattovich G et al., 1986). The annual probability of HBeAg seroconversion depends on the age of the patient at acute infection and HBV genotype, and is lower in Asian patients infected at birth (lower with genotype C than B) and higher in white patients infected during childhood, adolescence, or adulthood. (Fattovich G et al., 2003; Lok AS et al., 1987; Chu CJ et al., 2002). If
spontaneous HBeAg seroconversion occurs (disappearance of HBeAg and development of anti-HBe), patients progress to the HBeAg-negative phase, which is classified as either an inactive chronic HBV carrier state, characterised by low levels of viral replication, normal aminotransferases, and few histological lesions, (Lok AS et al, 2001; Fattovich G et al., 2003; McMahon BJ et al., 2004) or as HBeAg-negative chronic hepatitis B, characterised by higher viral replication, increased aminotransferases, and active liver necroinflammation and fibrosis. HBeAg-negative chronic hepatitis B may develop immediately after the HBeAg seroconversion phase or after several years of an inactive chronic carrier state (Hadziyannis SJ et al., 2006). It is associated with emergence of HBV strains that replicate but do not express HBeAg, (Hadziyannis SJ et al., 2001) and is more likely to run a fluctuating course, leading to cirrhosis and decompensation (Hadziyannis SJ et al., 2006; Brunetto et al., 1991; Lai ME et al., 1994).

Acute HBV infection is followed by the HBeAg-positive immune-tolerant phase, which is characterised by high serum HBV DNA concentrations and low alanine aminotransferase (ALT) activity. ALT activity increases and serum HBV DNA concentrations decrease during the HBeAg clearance or seroconversion phase (HBeAg-positive chronic hepatitis B), which may lead to HBeAg loss and development of its antibody, anti-HBe. Both ALT and HBV DNA concentrations remain low in the HBeAg-negative low-replicative phase (inactive chronic HBV carrier state). HBV may reactivate in a proportion of HBeAg-negative patients, resulting in high HBV replication and increased ALT activity, and thus leading to the development of HBeAg-negative chronic hepatitis B. In general, chronic HBV infection is characterised by its dynamic natural course, which includes apparent periodic activation of the host immune system against the infected hepatocytes to eradicate the virus, but this activation usually causes disease exacerbations and leads to the accumulation of fibrosis and development of cirrhosis (Fattovich G et al., 2003; McMahon BJ et al., 2004). Approximately 15–20% of patients with chronic hepatitis B develop cirrhosis within 5 years, (Liaw et al., 1988) whereas 15% of patients with compensated cirrhosis and more than 60% of those with decompensated cirrhosis die within 5 years (Weissberg JI et al., 1984). Moreover, all patients with
chronic HBV infection are at increased risk of hepatocellular carcinoma, although the risk increases substantially in patients with long-term high viraemia and cirrhosis (Chen CJ et al, 2006)

**Fig.1.2.4. Schematic representation of the phases of chronic HBV infection.** Acute HBV infection is followed by the HBeAg-positive immune-tolerant phase, which is characterised by high serum HBV DNA concentrations and low alanine aminotransferase (ALT) activity. ALT activity increases and serum HBV DNA concentrations decrease during the HBeAg clearance or seroconversion phase (HBeAg-positive chronic hepatitis B), which may lead to HBeAg loss and development of its antibody, anti-HBe. Both ALT and HBV DNA concentrations remain low in the HBeAg-negative low-replicative phase (inactive chronic HBV carrier state). HBV may reactivate in a proportion of HBeAg-negative patients, resulting in high HBV replication and increased ALT activity, and thus leading to the development of HBeAg-negative chronic hepatitis B. ULN (upper limit of normal) is 40 IU/L. Figure from Papatheodoridis et al., 2008.

### 1.2.5 HBV Epidemiology

HBV is classified into 8 HBV genotypes, A–H, based on an 8% or more DNA sequence difference over the whole genome.39,40. Genotypes A and D are ubiquitous, although genotype D is relatively rare in Northern Europe and the Americas. Little is known about the distribution of genotype G. It has been found in Europe, USA and Japan and may be ubiquitous. Genotypes B and C are found essentially in Asia. Genotype E is found in sub-Saharan Africa and rare cases in France and Britain may be due to immigration. Genotype F is found mainly in South and Central America
Genotype H is found in Central America and in the southern part of the United States, and these latter may also be due to immigration. The significance of genotype F and H strains in Japan is not clear. However, genotype F has been described in Polynesia and at least one of the Japanese genotype F strains was isolated in Okinawa whose population is in part of Melanesian origin. Genotype F (and perhaps genotype H) may therefore be a Pacific/Latin America genotype. There is a great deal of diversity within the genotypes and this has lead to the division of some genotypes into subtypes (Table 1.2) (Devesa et al., 2004; Huy et al., 2004; Norder et al., 2004; Sugauchi et al., 2004a,b). By using subtype-specific antibodies against HBsAg, nine different serological subtypes or serotypes were defined (Couroucé-Pauty et al., 1983). One determinant is common to all subtypes (“a-determinant” of HBsAg). There are also two pairs of mutually exclusive subdeterminants (“d” or “y” and “w” or “r”). Initially, HBsAg subtypes were used for studies of geographic distribution of HBV (Perillo, 1990; Moriya et al., 2002). However, the HBsAg subtype does not reflect true genotypic variation (Moriya et al., 2002). The HBV subtype classification was based on a limited number of amino acid substitutions; sometimes the HBsAg subtype can be changed by a nucleotide point mutation of the S gene. Two amino acid residues encoded by S gene at codon positions 122 (d/y) and 160 (w/r) have been postulated to determine the different antigenic subtypes (Okamoto et al., 1988). Data suggest that HBV genotype may be related to disease outcome. In Asia, genotype C is associated with a higher risk of reactivation of hepatitis B and progression to cirrhosis than genotype B (Chu et al., 2005; Orito E et al., 2001). In Europe, genotype D is associated with more active disease than the other genotypes, but the fact that genotype D also is associated with long-standing infection (as a result of acquisition at a younger age) may constitute a bias (Grandjacques C et al., 2000; Funk ML et al., 2002) Genotype F was implicated recently in an enhanced risk of hepatocellular carcinoma in Alaskan natives (Livingston SE et al., 2007).
Peoples infected with HIV have an increased risk of developing the acquired immunodeficiency syndrome (AIDS). Thirty-three million people around the world were living with HIV in 2008 (23 in Sub-Saharan Africa and 4.8 million in need of ART) (CROI 2008) . HIV infection is a chronic infectious disease that can be treated, but not yet cured. There are effective means of preventing complications and delaying progression to AIDS. At the present time, not all persons infected with HIV have progressed to AIDS, but it is generally believed that the majority will. People with HIV infection need to receive education about the disease and treatment so that they can be active partners in decision making with their health care provider. Currently, a combination of several antiretroviral agents, termed Highly Active Anti-Retroviral Therapy (HAART), has been highly effective in reducing the number of HIV particles in the blood stream (as measured by a blood test...
called the viral load). This can improve T-cell counts. This is not a cure for HIV, and people on HAART with suppressed levels of HIV can still transmit the virus to others through sex or sharing of needles. There is good evidence that if the levels of HIV remain suppressed and the CD4 count remains greater than 200 cells/µl, then the quality and length of life can be significantly improved and prolonged. Current recommendations for treatment include at least a three drug effective regime to induce complete suppression of viral replication and reduce the incidence of new mutations. Anti-HIV drugs approved for clinical use by the U.S. Food and Drug Administration (FDA) fall into 5 classes:

1) Non Nucleoside Reverse Transcriptase Inhibitors (NNRTIs)
2) Nucleoside / Nucleotide Reverse Transcriptase Inhibitors (NRTIs, NtRTIs)
3) Protease Inhibitors (PIs)
4) Entry / Fusion Inhibitor (FI)
5) Integrase Inhibitors (INIs)
6) CCR5 Inhibitors

The NNRTIs and NRTIs block the reverse transcription of the viral RNA genome in cDNA, which is catalyzed by the reverse transcriptase. The PIs block protease mediated maturation of released virions. The entry inhibitors block HIV entry into cells. This class of drugs interferes with the binding, fusion and entry of an HIV virion to a human cell. The gp41 inhibitors binds to gp41 and interferes with its ability to approximate the two membranes. It is also referred to as a "fusion inhibitor. The CCR5 inhibitors binds to CCR5 (receptor found on the surface of CD4+ cells), preventing an interaction with gp120. It is also referred to as a "chemokine receptor antagonist".

Integrase inhibitors are a class of antiretroviral drug designed to block the action of integrase, a viral enzyme that inserts the viral genome into the DNA of the host cell.

Each of the drugs in these classes represents individual pharmacodynamics and toxicities, requiring expert knowledge to appropriately select and dose an effective combination. Despite this panel of approved antiretroviral agents and the use of combination therapy, therapeutic success is limited. A
The major cause of treatment failure is the development of drug resistance. The intra-patients virus population is a highly dynamic system, characterized by a high turnover rate and a high mutation rate. These evolutionary dynamics are the basis for a diversified population that can quickly generate drug-resistant variants in response to the therapy. Escape mutants that have a selective advantage under therapy become dominant in the population and lead to the increased virus production and eventually to therapy failure. The shifted population may be hit with a new drug combination, but finding such a potent regimen after treatment failure is challenging, because many accumulated mutations confer drug resistance not only to the administered drugs, but also to structurally and functionally similar compounds.

Therefore, resistance testing has become an important diagnostic tool in the management of HIV-infections. With the aid of HIV resistance tests, antiretroviral treatment strategies can be improved. Pharmaco-economic studies have shown that these tests are also cost-effective (Corzillius 2004).

For several years, national and international HIV treatment guidelines have recommended the use of resistance testing. With some delay, resistance tests are now covered by public health insurance in several countries. Currently, both genotypic and phenotypic tests show good intra- and inter-assay reliability. However, the interpretation of genotypic resistance profiles has become very complex and requires constant updating of the guidelines. The determination of the thresholds associated with clinically relevant phenotypic drug resistance is crucial for the effective use of phenotypic testing. Even if treatment failure requires the consideration of other causal factors, such as compliance of the patient, metabolism of drugs and drug levels, resistance testing is of great importance in antiretroviral therapy.

Finally, it needs to be emphasized that even with the benefit of well interpreted resistance tests only experienced HIV practitioners should start, stop or change antiretroviral therapy with respect to the clinical situation and the psycho-social context of the patient.
1.3.2 HBV

The optimal goal of treatment of HBV infection is to achieve HBV surface antigen (HBsAg) clearance with anti-HBV surface antibody (HBsAb) seroconversion. This, however, is only achieved in a very small number of patients. The more realistic goal is longterm suppression of HBV replication to prevent transmission and spread of HBV, halt the progression of liver disease and improve the clinical and histologic picture, and to prevent HCC. Virological responses based upon testing for levels of HBV DNA in serum are the most appropriate criteria of assessing beneficial outcome of antiviral therapy.

There are two main classes of treatment licensed in Europe and the USA for the treatment of HBV infection:

- immune modulators: aimed at helping the human immune system to mount a defence against the virus (interferon-α-2a, interferon-α-2b and Pegylated Interferon).
- antivirals: aimed at suppressing or destroying HBV by interfering with viral replication. Four nucleoside analogs have been approved so far for the treatment of chronic hepatitis B: Lamivudine, Emtricitabine, Adefovir Dipivoxil, Entecavir, Telbivudine and Tenofovir.

Interferons display a variety of properties that include antiviral, immunomodulatory, and antiproliferative effects. They enhance T-cell helper activity, cause maturation of B lymphocytes, inhibit T-cell suppressors, and enhance HLA type I expression.

During antiviral therapy with nucleoside analogs, HBV mutants harboring mutations in the viral polymerase gene that confer resistance to antiviral drugs may then be selected. Escape mutants that have a selective advantage under therapy become dominant in the population and lead to the increased virus production and eventually to therapy failure. Therefore, resistance testing has become an important diagnostic tool in the management of HBV-infections, too.
The global prevalence of chronic hepatitis B and its associated serious sequelae demand technologically advanced techniques of management. Nucleic acid testing (NAT) plays a key role in the diagnosis, surveillance, and treatment of chronic hepatitis B. NAT includes quantitative PCR-based HBV DNA assays, HBV genotyping, tests for mutations associated with resistance to antiviral medications, and assays to detect precore and core promoter mutations. Documenting the mutation that confers drug resistance has not been part of routine clinical practice. This may change with the growing armamentarium of antivirals and reported crossresistance among drugs within the same structural class (for example, cross-resistance observed between lamivudine and other L-nucleosides such as entecavir, emtricitabine, and telbivudine). While sequence determination is relatively simple to perform, interpretation of sequence data is not always straightforward. Moreover, the current anti-HBV strategy of continuous monotherapy is insufficient to completely suppress viral replication in a large number of patients. A key factor in the development of resistance is the persistence of viral replication. Several studies found a relation between ongoing viral replication and the development of resistance. Indeed, it is important to detect resistance as early as possible during treatment with nucleoside or nucleotide analogues. Genotypic testing provides information on the type of mutation which arises during treatment and if there might be decreased drug sensitivity. Knowledge of the specific mutation will be increasingly important in the future as different mutations may have a distinct influence on treatment efficacy of other compounds. Finally, HBV-genotype might influence the mutational pattern. The role of genotypes is controversial as some have reported influence of the genotype on the development of resistance, while others do not find this association (Yuen MF et al., 2001; Fung SK et al., 2006; Zollner B et al., 2001). This point has been investigated in our laboratory (see aim of work).
1.3.3 Antiviral Therapy for HIV-HBV Coinfected Patients

Co-infection with the hepatitis B virus (HBV) and the human immunodeficiency virus type-1 (HIV-1) is common and complicates treatment. Although both viruses share a highly related target for pharmaceutical intervention, there are few drugs that are approved to treat HBV and HIV-1 infection simultaneously. HBV and HIV replicate through a reverse transcription step, carried out by the virally encoded HBV DNA polymerase and the HIV-1 reverse transcriptase (RT), respectively. Lamivudine (3TC) and emtricitabine (FTC) are cytidine analogues that exert potent antiviral effects against HBV and HIV-1 targeting their DNA polymerase. Once intracellularly phosphorylated to their triphosphate forms, these drugs act as chain-terminators. However, the development of resistance can limit their clinical utility.

1.4 Drug Resistance

The development of drug resistance derived mainly by the interplay of three factors:

**Replication.** Due to the lack of proof-reading mechanism, HIV-1 and HBV reverse transcription is a highly error-prone process. It has been estimated that the mutations rate of HIV is in the range of $10^{-4}$ to $10^{-6}$ substitution per base pair per replication cycle (Mansky et al., 1998) and the mutations rate of HBV is similar. Moreover, both viruses are characterized by a short generation time (1-3 days). Thus, this high and erroneous turnover represents the driving force of viral evolution and variation within a single patient for both HIV and HBV virus.

**Diversity.** HIV and HBV exist in a single individual as a mixture of genetically different variants, described as “quasispecies”, whose distribution reflects the relative fitness of the single virus (Holland et al., 1992). Thus, it is believed that drug resistant viruses pre-exist in the population especially if the genetic distance to the dominate wild-type virus is short.

**Selection.** In presence of antiretroviral drugs, drug resistant variants may replicate better than the wild-type. Thus, if therapy failed to completely suppress viral replication, drug resistant
variants may compete with the wild-type and become dominant in the viral population (Fig. 1.4), thus leading to viral rebound. In addition to mutations that directly confer resistance, additional mutations are selected during suboptimal therapy in order to rescue losses in fitness due to the presence of resistance-conferring mutations.

Figure 1.4 Schematic representation of selection of resistant virus under drug pressure

1.4.1 Mechanism of drug resistance in HIV-1.

NRTIs

Resistance to NRTIs may be mediated by at least three different mechanisms:
Impairment of Analogue Incorporation. Several mutations or groups of mutations in reverse transcriptase can promote resistance by selectively impairing the ability of reverse transcriptase to incorporate an analogue into DNA (Fig. 1.4.1A). These mutations include M184V, K65R, L74V, Q151M, Y115F, and V75I (Fig 1.4.1). The M184V mutation causes high level resistance to lamivudine and FTC and low level resistance to ABC, ddI, and ddC (Boucher et al., 1993). Methionine 184 is located at the heart of the catalytic site of reverse transcriptase, proximal to the catalytic aspartates D185 and D186. The replacement by a valine, which has a different side chain, interferes with the proper positioning of lamivudine triphosphate within the catalytic site (Sarafianos et al., 1999).

**MUTATIONS IN THE REVERSE TRANSCRIPTASE GENE ASSOCIATED WITH RESISTANCE TO REVERSE TRANSCRIPTASE INHIBITORS**

<table>
<thead>
<tr>
<th>Multi-nRTI Resistance: 69 Insertion Complex* (affects all nRTIs currently approved by the US FDA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
</tr>
<tr>
<td>L</td>
</tr>
<tr>
<td>41</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Multi-nRTI Resistance: 151 Complex* (affects all nRTIs currently approved by the US FDA except tenofovir)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>62</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Multi-nRTI Resistance: Thymidine Analogue-associated Mutations* (TAMs; affect all nRTIs currently approved by the US FDA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
</tr>
<tr>
<td>L</td>
</tr>
<tr>
<td>41</td>
</tr>
<tr>
<td>210 215 219</td>
</tr>
</tbody>
</table>

Fig.1.4.1. NRTI Resistance Mutations. From Johnson et al., IAS 2008.

The group of mutations referred to as the Q151M complex is most often selected for in the course of the failure of regimens containing stavudine and didanosine (Iversen et al., 1996). This pathway always starts with the Q151M substitution, a residue located in the immediate vicinity of the nucleotide binding site of reverse transcriptase, and is followed by the gradual accumulation of
secondary mutations that enhance resistance and increase the activity of the enzyme (Kosalaraksa et al., 1997). The Q151M complex is relatively rare in HIV-1 (fewer than 5 percent of all HIV strains with resistance to nucleoside analogues) but can confer high-level resistance to most, but not all (lamivudine and tenofovir) NRTIs (Iversen et al., 1996). Interestingly, the Q151M complex is markedly more frequent in HIV-2 than in HIV-1. The K65R mutation is seen with increasing frequency in patients in whom therapy with nucleoside or nucleotide analogues fails, especially when the regimen includes tenofovir or abacavir. This mutation appears to confer resistance to most analogues, with the exception of zidovudine. The selection of the above mentioned mutations determine an impairment of replicative capacity directly linked to their decreased ability to incorporate the natural nucleotide substrates.

**Figure 1.4.1A-B The Two Principal Mechanisms of Resistance of HIV to Nucleoside Analogues.** In Panel A, the incorporation of a nucleoside analogue into drug-sensitive viruses results in the termination of the viral DNA chain. Mutations in drug-resistant viruses prevent the incorporation of the nucleoside analogue into the growing viral DNA chain. In Panel B, ATP in drug-sensitive viruses does not have access to a reverse transcriptase that has formed a complex with a nucleoside analogue. Mutations that cause resistance to nucleoside analogues, referred to as thymidine analogue mutations, allow ATP to bind reverse transcriptase near the 3’ end of viral DNA terminated by the incorporation of a nucleoside analogue. ATP then excises the analogue from viral DNA, allowing reverse transcription to proceed normally. Modified from Clavel and Hance, 2004.

*Removal of the Analogue from the Terminated DNA Chain.* Removal of the nucleoside analogue from the terminated DNA chain is associated with a group of mutations commonly termed “thymidine analogue mutations”. Mutations from this group are most frequently selected for after
the failure of drug combinations that include thymidine analogues, such as zidovudine and stavudine, but they can promote resistance to almost all nucleoside and nucleotide analogues, including Tenofovir (Coakley et al., 2000; Larder et al., 1989; Picard et al., 2001; Shafer et al., 1996). These mutations occur gradually, and their order of emergence can vary (Boucher et al., 1992). Thymidine analogue mutations promote resistance by fostering ATP- or pyrophosphate-mediated removal of nucleoside analogues from the 3' end of the terminated DNA strand (Fig. 1.19B) (Arion et al., 1998; Meyer et al., 1999). ATP and pyrophosphate, which are abundant in normal lymphocytes, do not participate in the DNA-polymerization reaction, but the structure of a reverse transcriptase expressing thymidine analogue mutations facilitates their entry into a site adjacent to the incorporated analogue (Boyer et al., 2001; Chamberlein et al., 2002). In this position, ATP or pyrophosphate can attack the phosphodiester bond that links the analogue to DNA, resulting in removal of the analogue (Fig. 1.4.1 B). Interestingly, the efficiency of this process, also known as “primer rescue,” can be significantly decreased by the presence of other mutations in reverse transcriptase, a phenomenon that has been best described in the case of the M184V mutation (Larder et al., 1995). As a consequence, M184V slows the selection of thymidine analogue mutations by thymidine analogues (Picard et al., 2001) and may slightly increase the residual antiviral activity of some nucleoside analogues in spite of the presence of thymidine analogue mutations.

1.4.2. Drug resistance in HBV

NRTIs

In HIV-1, both 3TC and FTC can select for the M184V mutation in the conserved YMDD motif at the active site of RT (Wainberg et al., 1995). A structurally and functionally equivalent mutation, i.e., M204V, is selected in HBV (Ling et al., 1996; Gutfreund et al. 2000; Fontaine et al., 1999). This mutation is often accompanied with other mutations, including L180M, that appear to compensate for fitness deficits that are introduced by the primary mutation (Villet et al., 2006). The incidence of
3TC-resistance is high in part because 3TC has been for a long time the only small molecule approved for the treatment of HBV. The acyclic phosphonate tenofovir is active against 3TC resistant HIV and HBV strains (Nunez et al., 2002; Van Bommel et al., 2002; Benhamou et al. 2003), while the structurally related compound adefovir has been approved for HBV treatment only, even though it has activity against HIV (Manolakopoulos et al., 2008; Yuan et al., 2007).

Entecavir (ETV) belongs to the few available drugs that retain potency against 3TC-resistant HBV (Sherman et al., 2006). This compound does not appear to select for M204V and L180M in HBV, although the pre-existence of the two mutations can decrease susceptibility to ETV (Colonno, R. et al., 2006). Additional other mutations can further amplify clinically relevant levels of resistance to this drug (Tenney, D. J. et al., 2004). In contrast, it has recently been demonstrated that ETV can select for the M184V mutation in HBV/HIV co-infected individuals (McMahon et al., 2007).

1.5 Differences and similarities between HIV-1 and HBV

1.5.1 Differences

Nevertheless HIV and HBV share similar routes of transmission, with sexual, parenteral and perinatal transmission being the most frequent modes of acquiring these infections, the exposure to these viruses is followed by an immune response which differs markedly in its ability to clear the infection. Clearance is maximal for adults exposed to HBV, negligible (or non-existent) for HIV. In contrast to HIV infection, in HBV infection the immune response is present, strong and adapted. It can take control of infection, especially when viral replication is reduced by antiviral drugs. Taking into consideration these two facts, it comes as no surprise that there is a high worldwide prevalence of coinfection with these agents. Approximately 10% of the HIV-infected population worldwide suffers from chronic hepatitis B (Figure 1.5.1). This figure may approach 20% in Southeast Asia, whereas it is 5% in North America and Western Europe. Unlike with HCV (hepatitis C virus),
infection with HBV is not eradicable and the main goal of therapy is to suppress HBV replication as much as possible and for as long as possible.

Moreover, HBV has a narrow spectrum of possible natural polymorphisms than HIV, due to slower viral kinetics and stronger conservatory constraints on protein sequence. Despite structural similarity of RT active site, mechanisms toward 3TC resistance are quite different in HIV and HBV. Antiviral drug resistance in chronic hepatitis B can be caused by many factors, including the viral mutation frequency, the intrinsic mutability of the antiviral target site, the selective pressure exerted by the drug, the magnitude and rate of virus replication.

1.5.2 Similarities
Involvement of reverse transcriptase activity in the replication of the HBV genome explains why this virus displays a higher mutation rate and a broader genetic variability than most DNA viruses and comparable to that of HIV.

Fig. 1.5.1: Estimated number (in millions) of subjects infected with HIV, HBV and HCV worldwide. Modified from Soriano et al., 2006.
HBV like HIV has a high rate of replication, with an estimated production of $10^{11}$ virions per day for HBV and of $10.3 \times 10^9$ virions per day for HIV (Nowak MA et al. 1996, Perelson et al. 1996). Both viral polymerases have low fidelity (Park et al. Eur J Biochem 2003). That is, it has a propensity to mispair nucleotide bases when it copies viral RNA to make viral DNA. It also lacks any proofreading activity (Park et al. Eur J Biochem 2003; Sheldon J et al. 2006), so cannot repair mistakes. This means that when a nucleotide base is misplaced it remains in the growing viral DNA as a base mutation and the new DNA has different sequence from its parent genome. The overall error rate of the HBV polymerase, that displays strong similarities with the HIV reverse transcriptase, is estimated at 1 per 1,500 nucleotides copied. This, combined with a high rate of replication, means that there is the potential for 10 billion base-pair errors per day in an infected individual. So, in any one day all possible single-base mutations could be produced (although many will yield a non-viable virus) (Locarnini et al. Antiviral Ther 2004).

In general RNA and single strand DNA (or partially double strand DNA) viruses are more likely to undergo spontaneous mutation than the genomes of eucaryotes, bacteria or double strand DNA viruses. The spontaneous mutation rate (the chance of mutation occurring in an organism or gene in each replication cycle) in RNA (HIV) and ssDNA viruses (HBV) is 1, meaning that a mutation occurs at least once every replication cycle. In contrast, the spontaneous mutation rates in eukaryotes and bacteria/DNA viruses are 0.003 and 0.01, respectively.

Moreover due to the presence of cccDNA in the nucleus of the hepatocytes, HBV infection, like HIV infection, is not curable. Indeed, when HBV infects a hepatocyte in the liver, viral replication in the hepatocyte results in the formation of cccDNA within cell. This cccDNA can serve as template for the production of new virus, but can also remain within the hepatocyte, where it is said to be archived. Archived cccDNA plays an important role in viral persistence and in the reactivation of viral replication after the cessation of antiviral therapy. By drawing comparisons with animal models of HBV infection, it is believed that drug resistance may be archived in the form of cccDNA. During the normal course of replication, cccDNA harbouring drug resistance mutations
may become archived in the hepatocytes. Once archived in this way, drug resistance may persist for long periods in the absence of drug. Archived cccDNA molecules in hepatocytes act mainly as reservoir for future viral replication and are therefore not inhibited by nucleos(t)ide analogues, which inhibit replication when they are incorporated into nucleic acid molecules during replication. Since DNA acts as reservoir for future replication, the archiving of resistant variants may lead to the persistence and expansion of this population. Similarly, latent HIV-1 persists in resting memory CD4+ T cells, even in patients receiving highly active antiretroviral therapy (HAART), precluding eradication with available antiretroviral therapy (Siliciano et al. Nat Med 2003) and HIV drug resistance mutations can accumulate, persist in the proviral genome and re-emerge, promoting further failure of rescue regimens.

Finally, anti-HBV therapy is based on the use of nucleos(t)ide analogues, several of which also display anti-HIV therapy (Lamivudine and Tenofovir).

2. References:

References:


Chu CJ, Hussain M, Lok AS. Hepatitis B virus genotype B is associated with earlier HBeAg seroconversion compared with hepatitis B virus genotype C. Gastroenterology 2002; 122: 1756–62

Chu CM, Liaw YF. 2005. Genotype C hepatitis B virus infection is associated with a higher risk of reactivation of hepatitis B and progression to cirrhosis than genotype B: a longitudinal study of hepatitis B e antigen-positive patients with normal aminotransferase levels at baseline. J Hepatol;43:411–417.


Clavel, F, D Guetard, F Brun-Vezinet, S Chamaret, MA Rey, MO Santos-Ferreira, AG Laurent, C Dauguet, C Katlama, C Rouzioux, and et al. (1986). Isolation of a new human retrovirus from West African patients with AIDS. Science. 233:343-346.


Kozlowski, MR, P Sandler, PF Lin, and A Watson (1991). Brain-derived cells contain a specific binding site for Gp120 which is not the CD4 antigen. Brain Res. 553:300-304.


Lamb, RA, and LH Pinto (1997). Do Vpu and Vpr of human immunodeficiency virus type 1 and NB of influenza B virus have ion channel activities in the viral life cycles? Virology. 229:1-11


Petersen, Nature Biotech 2008


Summers, J., J. Smolec, and R. Snyder. 1978. A virus similar to human hepatitis B virus associated


WHO website 2008


3. Results

3.1 HIV

3.1.1 Dynamics of NRTI Resistance Mutations During Therapy Interruption

Maria Trignetti\textsuperscript{a#}, Tobias Sing\textsuperscript{b#}, Valentina Svicher\textsuperscript{a}, Maria Mercedes Santoro\textsuperscript{a}, Federica Forbici\textsuperscript{c}, Roberta D’Arrigo\textsuperscript{c}, Maria Concetta Bellocchi\textsuperscript{a}, Mario Santoro\textsuperscript{a}, Patrizia Marconi\textsuperscript{c}, Mauro Zaccarelli\textsuperscript{c}, Maria Paola Trotta\textsuperscript{c}, Rita Bellagamba\textsuperscript{c}, Pasquale Narciso\textsuperscript{c}, Andrea Antinori\textsuperscript{c§}, Thomas Lengauer\textsuperscript{b}, Carlo Federico Perno\textsuperscript{c§} and Francesca Ceccherini-Silberstein\textsuperscript{c*}

\textsuperscript{a} Experimental Medicine Department, University of Rome Tor Vergata, Via Montpellier 1, 00133, Rome, Italy

\textsuperscript{b} Max Planck Institute for Informatics, Stuhlsatzenhausweg 85, 66123, Saarbrücken, Germany

\textsuperscript{c} National Institute of Infectious Diseases, “L. Spallanzani”, Via Portuense 292, 00149 Rome, Italy

Aids Research And Human Retroviruses

Volume 25, Number 1:57-64, 2009

I. INTRODUCTION

To date, more than 20 antiretroviral drugs have been approved for the therapy of HIV infection. Most of them target two viral enzymes: reverse transcriptase (RT) and protease (PR), one the integrase, while two drugs, enfuvirtide and maraviroc, target the envelope glycoproteins gp41 and gp120/CCR5 coreceptor involved in viral entry. The combined use of these drugs has substantially improved the clinical management of HIV-1 infection in terms of delaying disease progression, prolonging survival and improving quality of life. Nevertheless, when antiretroviral therapy fails to
be fully suppressive, new viral variants emerge, that make HIV-1 becoming resistant to one or more drugs by mean of accumulating mutations, either alone or in multiple and complex patterns. Regarding the nucleoside analogues inhibitors of reverse transcriptase (NRTIs), the most common backbone of antiretroviral regimens, two mechanisms are so far known to contribute to a decreased NRTI susceptibility of HIV strains. In one case, several mutations or groups of mutations in HIV-1 reverse transcriptase can promote resistance by selectively impairing the ability of the enzyme to incorporate the nucleoside analogue into DNA. These mutations include M184V/I, K65R, L74V and the Q151M complex (A62V, V75I, F77L, F116Y, and Q151M) \(^1\)-\(^3\). On the other hand, a specific set of mutations collectively termed “nucleoside analogue mutations” (NAMs; M41L, D67N, K70R, L210W, T215Y/F, and K219E/Q) can confer resistance by promoting a phosphorolysis reaction that selectively removes the nucleoside analogue from the terminated DNA chain \(^4\)-\(^6\). The NAMs accumulate gradually under the selection pressure imposed by the thymidine analogues (zidovudine and stavudine) and can promote cross resistance to almost all nucleoside and nucleotide analogues. Several studies have suggested the existence of two distinct pathways of NAM resistance, defined by different mutation patterns (NAMs-1 [M41L, L210W, and T215Y] and NAMs-2 [D67N, K70R, T215F, and K219Q/E]), whose evolution seems to be strictly influenced by viral replication \(^7\)-\(^10\).

In contrast, the dynamics of disappearance of drug resistance mutations during treatment interruption (TI) has not yet been studied systematically. From a clinical point of view, structured treatment interruptions (STI) have been expected to boost host cellular response, minimize drug toxicities, and, perhaps most importantly, allow the viral population to revert back to wild-type, thus improving the range of future drug options and/or the virological response to follow-up therapy. Unfortunately, according to a number of recent studies, STI in patients with multidrug-resistant HIV was found to have a virological benefit unable to overcome the impaired clinical and immunological outcome \(^11\)-\(^15\). Some findings indicate benefits of interrupting an individual drug in a class for which there has been an accumulation of resistance mutations, yet the existence of latent
reservoirs in which virtually every resistant variant becomes archived, presents, to our current knowledge, a fundamental limitation to extend approaches aiming at reverting resistance mutations in the viral population. Despite these limitations, the study of mutational dynamics of resistant populations in the absence of drug pressure is of crucial interest from at least two points of view. From the clinical perspective, we underline that, even if STI might not be clinically advised, "unstructured" therapy interruptions occur every day as part of life for many HIV-infected patients, for a wide variety of reasons (e.g. non-adherence, intolerance, toxicities, etc). Beside this, the study of therapy interruptions is particularly interesting also from an evolutionary point of view. In fact, the disappearance of mutations during TI is correlated with their effect upon viral fitness, i.e. it is conceivable that the shortest is the time of disappearance of a specific mutation, the greatest is its effect upon viral replicative capacity. Thus, the time to disappearance may represent a surrogate marker of the role of specific mutations in affecting viral replication.

In this study, we focused specifically on the dynamics of classical NRTI mutations ⁱ⁶, plus a number of “novel” NRTI-associated mutations ⁱ⁷-¹⁸, during TI. To our knowledge, this study is the first application of methods from survival analysis to longitudinal sequence data.

II. MATERIALS AND METHODS

Study population. The study included 323 pol sequences from 132 HIV-1 infected adult patients treated in different hospitals in Central Italy, undergoing genotypic resistance testing for routine clinical purposes between 1999 and 2006. To decrease inferences caused by viral backbones, all patients followed in this study carried a B-subtype. Patients were included in the study if, after failing an NRTI-containing regimen, they have interrupted therapy for at least one month. TI in our analysis is defined as the complete cessation of all retroviral drug administration during a certain period of time. Seven patients experienced two and one patient four cycles of TI. The decision for TI was either based on limited remaining treatment options or on patient’s choice (mainly due to side effects of the drugs). For each patient and for each cycle of TI, at least two
genotypes were available: one at failure of the NRTI-containing regimen associated to the
beginning of TI, and one or more at different time points during TI. The genotype at virological
failure was performed within 90 days (median: 31, IQR: 16-47 days) before the beginning of the TI,
and is defined as baseline genotype in all the analyses performed. Even if a total number of 746
patients (out of >6000 patients) in our DB interrupted therapy for at least one month between 1999
and 2006, we included in our analysis only a set of 132 patients that met our abovementioned study
criteria. Patients not included in our study did not have a genotype at baseline within 90 days from
TI beginning. Data for all patients were stored in a specifically designed anonymous database that
included mutational, demographic, immunologic, virologic and therapeutic parameters.

**HIV sequencing.** HIV genotype analysis was performed on plasma samples by means of a
commercially available kit (the ViroSeqTM HIV-1 Genotyping System, versions 1 and 2, Applied
Biosystems) and an automatic sequencer (ABI 377 and ABI 3100, Applied Biosystems, Foster City,
California, USA)\(^{19-20}\). Briefly, RNA was extracted, retrotranscribed by MULV RT, and amplified
with Amplitaq-Gold polymerase enzyme by using two different sequence-specific primers for 40
cycles. Pol-amplified products (containing the entire protease and the first 335 amino acids of the
reverse transcriptase open reading frame) were full-length sequenced in sense and antisense
orientations by using seven different overlapping sequence-specific primers for an automated
sequencer (ABI 3100). Mixtures of multiple residues at a single position were all considered as
present in the statistical analyses. The isolates were subtyped by comparing them to reference
sequences of known subtype [http://hivdb.stanford.edu]\(^{21}\). All sequences were confirmed to be of
subtype B.

**Mutations.** The mutations considered in the analysis were those reported to be associated
with resistance to NRTIs in the mutation lists of the International AIDS Society\(^{16}\), the Stanford
HIV Drug Resistance Database and in a comprehensive screening for “novel” mutations\(^{18}\).
Specifically, we focused on 33 NRTI resistance mutations present with a frequency ≥ 4% in our
study group of patients: 19 classical (M41L, E44D, A62V, K65R, D67N, K70R, L74V, V75I,

**Survival analysis for mutation dynamics.** Survival functions were estimated using the Kaplan-Meyer product-limit estimator. For a given mutation \( m \), all therapy interruptions with \( m \) present at baseline were considered. Observations were treated as right-censored, i.e. if mutation \( m \) was present in a sequence from time \( t_1 \), and absent in a sequence from time \( t_2 > t_1 \), mutation \( m \) was assumed to disappear exactly at time \( t_2 \). If a mutation present at baseline was still present in the last follow-up within a TI period at time \( t \), a survival time of “at least \( t \)” was assumed. Finite 95% upper confidence intervals could only be calculated for sufficiently frequent mutations.

To assess if the survival function of a mutation depends on the baseline presence of some other mutations, we proceeded as follows: we split the TI episodes with a specific \( m \) mutation present at baseline in two classes, according to the presence or absence of another \( m' \) mutation present at baseline. This procedure was performed only if the mutation was present at least in five patients in each group (with/without other mutation), this is the case for 135 of the theoretically possible 1190 pairs. We then tested the null hypothesis that the survival curves are identical using the log-rank test. The p-values calculated by log-rank test were corrected for multiple testing using the Benjamini-Hochberg method \(^{22} \).

**III. RESULTS**

**Baseline patient characteristics.** Table 1 summarizes the main baseline characteristics of our cohort of patients. All patients experienced all three major drug classes, and harboured viral strains resistant to multiple NRTIs, non-nucleoside RT inhibitors (NNRTIs) and protease inhibitors (PIs). The median time of exposure to antiretroviral treatment before TI was 7.5 years. At the time of starting TI, the patients were exposed to a median of 5 (IQR=4–5) NRTIs, 2 (IQR=1–4) PIs, and 1 (IQR=1–1) NNRTI and their median number of resistance mutations were: 4 (IQR=3–5) for
NRTIs, 3 primary (IQR=2-4) for PIs and 2 (IQR=1–2) for NNRTIs. The study population was predominantly male (97/132 patients, 73.4%), and had a median age of 40 years at baseline (IQR=36–45). At baseline, median CD4+ cell count was 489/mm³ (IQR=328–643) and median HIV-1 RNA was 3.9 log copies/ml (IQR=3.3–4.5). For 103 out of 132 patients, a single HIV-1 genotype during TI was available, whereas two genotypes during TI were available for 24 patients, three and four genotypes were given for 4 and 1 patients during TI, respectively. Among the 132 patients, 7 patients did more than one TI cycle with available genotypes at baseline and during TI. The median time of TIs was 4.3 months, with an IQR from 3.1 to 6.8 months, from a minimum of 1 month to a maximum of 28.8 months.

Table 1. Baseline characteristics of our cohort of patients

<table>
<thead>
<tr>
<th>Patients, N</th>
<th>132</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with more than one TI cycle, N</td>
<td>7</td>
</tr>
<tr>
<td>Male, N (%)</td>
<td>97 (73.4)</td>
</tr>
<tr>
<td>Median age, (years) (IQR)</td>
<td>39.5 (35.7-45.0)</td>
</tr>
<tr>
<td>Risk factor, N (%)</td>
<td></td>
</tr>
<tr>
<td>Sexual</td>
<td>54 (40.9)</td>
</tr>
<tr>
<td>Drug addict</td>
<td>19 (14.3)</td>
</tr>
<tr>
<td>Perinatal</td>
<td>4 (3.0)</td>
</tr>
<tr>
<td>Iatrogen</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td>Not known</td>
<td>54 (40.9)</td>
</tr>
<tr>
<td>Patients at CDC stage C, N (%)</td>
<td>37 (28.0)</td>
</tr>
<tr>
<td>Median (IQR) viremia (log HIV RNA copies/ml)</td>
<td>3.9 (3.3-4.5)</td>
</tr>
<tr>
<td>Median (IQR) CD4 cell count (cells/µl)</td>
<td>489 (328-643)</td>
</tr>
<tr>
<td>Median (IQR) number of years since diagnosis</td>
<td>10.5 (9.8-11.1)</td>
</tr>
<tr>
<td>Median (IQR) number of years under treatment</td>
<td>7.5 (5.6-9.1)</td>
</tr>
<tr>
<td>Median (IQR) number of months remaining under TI</td>
<td>4.3 (3.1-6.8)</td>
</tr>
<tr>
<td>Median number (IQR) of previously received drugs</td>
<td></td>
</tr>
<tr>
<td>NRTI</td>
<td>5 (4-5)</td>
</tr>
<tr>
<td>NNRTI</td>
<td>1 (1-1)</td>
</tr>
<tr>
<td>PI</td>
<td>2 (1-4)</td>
</tr>
<tr>
<td>Median number (IQR) of resistance associated mutations a</td>
<td></td>
</tr>
<tr>
<td>NRTI</td>
<td>4 (3-5)</td>
</tr>
<tr>
<td>NNRTI</td>
<td>2 (1-2)</td>
</tr>
<tr>
<td>PI</td>
<td>3 (2-4)b</td>
</tr>
</tbody>
</table>

aThe drug resistance mutations considered are those listed by the International AIDS Society (IAS 2007)
bIQR: Interquartile range cWe considered only major PI resistance mutations.
Dynamics of disappearance of individual NRTI mutations during TI. By using a survival analysis approach, we found that the NRTI resistance mutations showed different dynamics of disappearance during TI (Table 2). In particular, the K65R, M184I and the novel mutation K43E were the fastest mutations to disappear with a median time of disappearance of less than 4 months. Among them, the K65R mutation disappeared in all patients already within one month of TI. Differently, the majority of the classical and recently associated NRTI resistance mutations showed a median time of disappearance between 4 and 6 months. Among them, the M184V and Y115F disappeared in 91.5% and 100% patients, and with a median time of disappearance of 4.3 and 5.2 months, respectively. Of note, the loss of the M184V mutation mainly coincided with the re-emergence of the wild-type strains. We observed a median time of disappearance longer than 6 months for the classical NRTI resistance mutations L210W and F116Y, and for the NRTI resistance-associated polymorphisms K122E, G196E, F214L. In sharp contrast, the polymorphisms I50V and R83K, which have been negatively associated with NRTI failure\textsuperscript{18}, did not disappear during TI.
Table 2: Statistics for the estimated survival functions of NRTI resistance mutations

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Events of disappearance</th>
<th>Median Time of disappearance (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N/Tot (%)</td>
<td></td>
</tr>
<tr>
<td><strong>Disappearance within 4 months</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K65R</td>
<td>8/8 (100)</td>
<td>n.d.(^b)</td>
</tr>
<tr>
<td>M184I</td>
<td>5/5 (100)</td>
<td>3.10 (2.70 –∞)</td>
</tr>
<tr>
<td>K43E</td>
<td>5/9 (55.5)</td>
<td>3.60 (3.30 –∞)</td>
</tr>
<tr>
<td><strong>Disappearance between 4 and 6 months</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T215F</td>
<td>28/36 (78)</td>
<td>4.20 (3.50 – 5.60)</td>
</tr>
<tr>
<td>M184V</td>
<td>87/95 (91.5)</td>
<td>4.30 (3.90 – 5.00)</td>
</tr>
<tr>
<td>L74V</td>
<td>27/32 (84.3)</td>
<td>4.40 (3.90 – 7.40)</td>
</tr>
<tr>
<td>D67N</td>
<td>63/84 (75)</td>
<td>4.40 (3.80 – 5.60)</td>
</tr>
<tr>
<td>D218E</td>
<td>14/19 (73.6)</td>
<td>4.50 (3.50 – 7.40)</td>
</tr>
<tr>
<td>T39A</td>
<td>9/17 (53)</td>
<td>4.60 (3.30 –∞)</td>
</tr>
<tr>
<td>M41I</td>
<td>46/68 (67.4)</td>
<td>4.70 (3.90 – 6.30)</td>
</tr>
<tr>
<td>E44D</td>
<td>15/25 (60)</td>
<td>4.80 (3.60 –∞)</td>
</tr>
<tr>
<td>V75I</td>
<td>7/8 (87.5)</td>
<td>4.90 (4.20 – ∞)</td>
</tr>
<tr>
<td>V118I</td>
<td>22/35 (63)</td>
<td>4.90 (4.20 – 8.00)</td>
</tr>
<tr>
<td>E203K</td>
<td>9/14 (64.2)</td>
<td>5.10 (3.10 – ∞)</td>
</tr>
<tr>
<td>Y115F</td>
<td>8/8 (100)</td>
<td>5.20 (4.00 – ∞)</td>
</tr>
<tr>
<td>K70R</td>
<td>43/59 (72.8)</td>
<td>5.20 (4.40 – 6.80)</td>
</tr>
<tr>
<td>T215Y</td>
<td>44/61 (72)</td>
<td>5.20 (4.10 – 7.30)</td>
</tr>
<tr>
<td>K219E</td>
<td>10/14 (71.4)</td>
<td>5.20 (3.90 – ∞)</td>
</tr>
<tr>
<td>K219Q</td>
<td>29/41 (70.7)</td>
<td>5.20 (3.40 – 6.50)</td>
</tr>
<tr>
<td>Q151M</td>
<td>8/9 (89.0)</td>
<td>5.60 (4.20 – ∞)</td>
</tr>
<tr>
<td>A62V</td>
<td>6/10 (60)</td>
<td>5.60 (5.20 – ∞)</td>
</tr>
<tr>
<td>K20R</td>
<td>13/26 (50)</td>
<td>5.60 (3.80 – ∞)</td>
</tr>
<tr>
<td>H208Y</td>
<td>7/13 (53.8)</td>
<td>5.70 (4.40 – ∞)</td>
</tr>
<tr>
<td><strong>Disappearance after 6 months</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L210W</td>
<td>27/45 (60)</td>
<td>6.30 (4.10 – 8.00)</td>
</tr>
<tr>
<td>V35M</td>
<td>6/15 (40)</td>
<td>6.30 (5.20 – ∞)</td>
</tr>
<tr>
<td>F116Y</td>
<td>7/7 (100)</td>
<td>6.80 (5.50 – ∞)</td>
</tr>
<tr>
<td>K43Q</td>
<td>3/5 (60)</td>
<td>8.00 (3.90 – ∞)</td>
</tr>
<tr>
<td>F214L</td>
<td>11/30 (36.6)</td>
<td>8.00 (5.20 – ∞)</td>
</tr>
<tr>
<td>K122E</td>
<td>19/60 (31.6)</td>
<td>8.30 (6.50 – ∞)</td>
</tr>
<tr>
<td>K43N</td>
<td>2/5 (40)</td>
<td>10.80 (3.10 – ∞)</td>
</tr>
<tr>
<td>G196E</td>
<td>9/50 (18)</td>
<td>11.60 (9.40 – ∞)</td>
</tr>
<tr>
<td><strong>No disappearance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I50V</td>
<td>0/5 (0)</td>
<td>∞ (∞ –∞)</td>
</tr>
<tr>
<td>R83K</td>
<td>1/28 (3.5)</td>
<td>∞ (∞ –∞)</td>
</tr>
</tbody>
</table>

\(^a\) NRTI resistance mutations considered are those present at the baseline of treatment interruption with a frequency ≥ 4%. We reported in bold the classical NRTI resistance mutations.

\(^b\) The median time of disappearance of the K65R mutation was not calculated since this mutation disappeared already within the first month of interruption.

**Revertants at RT position 215.** In our study, we observed that the complete disappearance of the T215Y/F mutations during TI was associated in the majority of patients with the appearance
of the wild-type amino acid (T) at position 215 alone and very rarely with the concomitant appearance of other mutations at that position, the T215 revertants, (Table 3). In contrast, the maintenance of the T215Y/F mutations during TI was different according to which mutation was present at baseline: the maintenance of T215Y mutation was frequently observed alone (26%) or together with the T215 revertants (13%), while the maintenance of T215F mutation was observed less frequently alone (7.4%) and more frequently together with some 215 revertants (25.9%). In particular, among the 54 patients with T215Y mutation at baseline, the appearance of the wild-type amino acid at position 215 occurred in 31 (57.4%) patients, while the appearance of T215 variants occurred in 9 (16.6%) patients (Table 3). Among the 27 patients with T215F and no revertants at baseline, the appearance of the wild-type amino acid at position 215 occurred in 18 (66.6%) patients, and the appearance of the T215 variants in 8 (29.6%) patients. Some differences, yet not statistically significant, can be noted in term of patterns of reversion from 215Y (consistent presence of S/N/C) and those from 215F (more variable) suggesting different replicative capacity for all these variants (Table 3). In addition, we found that the presence of the revertants is associated with a median length of TI that lies between median time of maintenance of T215Y/F and median time of TI for complete disappearance of such mutations. This supports the role of the revertants as intermediate forms arising from the T215Y or T215F mutation through a genetic drift process.
Table 3: Revertants at RT position 215

| Mutation at baseline | N (%) | Mutation at last TI episode | N (%) | Median TI length
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>months (IQR)</td>
</tr>
<tr>
<td><strong>T215Y</strong></td>
<td>54 (41)</td>
<td>T215Y</td>
<td>14 (26)</td>
<td>3.1 (2.3-3.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T215swt</td>
<td>31 (57.4)</td>
<td>4.8 (3.3-7.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T215S/N/C/Twt</td>
<td>2 (3.7)</td>
<td>5.5 (4.0-7.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T215S/N/C/Y</td>
<td>2 (3.7)</td>
<td>3.6 (3.4-3.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T215S/N/C/Y/Twt</td>
<td>5 (9.2)</td>
<td>3.5 (3.3-4)</td>
</tr>
<tr>
<td><strong>T215F</strong></td>
<td>27* (20.4)</td>
<td>T215F</td>
<td>2* (7.4)</td>
<td>2.5 (2.4-2.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T215swt</td>
<td>18 (66.6)</td>
<td>4 (2.8-5.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T215F/I/S/Twt</td>
<td>4 (14.8)</td>
<td>3.3 (3.1-3.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T215F/S</td>
<td>1* (3.7)</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T215F/I/P/L/S/Twt</td>
<td>1 (3.7)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T215S/Y</td>
<td>1 (3.7)</td>
<td>5.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T215F/Y</td>
<td>1 (3.7)</td>
<td>3.4</td>
</tr>
<tr>
<td><strong>T215S/Y</strong></td>
<td>1 (7.57)</td>
<td>T215S/N/C/Y/Twt</td>
<td>1 (100)</td>
<td>2.3</td>
</tr>
</tbody>
</table>

*Frequency of resistance mutations at position 215 in 132 patients before starting TI. The percentages were calculated for patients containing T215Y or T215F without revertants at baseline.

**Appearance of NRTI resistance mutations not present at baseline.** Of interest, in addition to the appearance of the 215 revertants, we also observed during TI a rare appearance of some NRTI resistance mutations. In particular, the classical L210W, M184V, and E44D and the recently associated NRTI resistance mutations K20R, V35M, T39A, K43E appeared each in 1 (0.75%) patient, the T215Y and V118I mutations in 2 (1.5%) patients, and the K70R mutation in 3 (2.3%) patients. Interestingly, the I50V, F214L and R83K mutations, that have been negatively associated with NRTI failure and/or with NRTI resistance mutations, appeared in 4 (3%), 6 (4.5%) and 13 (9.8%) patients, respectively, not carrying them at baseline.

**The role of mutation interactions on dynamics of disappearance of NRTI resistance mutations during TI.** To assess if the survival probability of maintenance/disappearance of a specific mutation depends on baseline presence of some other mutations, separate survival probability curves were estimated for several pairs of mutations. Among 135 pairs of mutations
analysed, we found only 7 mutation pairs with significant (p<0.05, log-rank test) interactions (K70R with D67N, K122E with K219Q, K122E with D67N, K122E with K70R, G196E with E44D, F214L with K219Q). For example, the presence of NAM2 D67N was associated with a faster disappearance of NAM2 K70R, indeed the median time of disappearance of K70R alone was longer than that of K70R with D67N (6.8 months \textit{versus} 4.6 months, p=0.046), or the median time of disappearance of K122E alone was longer than that of K122E with the presence of K219Q (12.6 months \textit{versus} 6.5 months, p=0.002). Nevertheless after correction for multiple comparisons, all of the potential interactions became insignificant. This suggests that, at least with this model, the disappearance of NRTI resistance mutations may proceed without significant strong interactions among such mutations.

IV. DISCUSSION

Our study shows that the NRTI resistance mutations disappear during therapy interruption with different dynamics and largely independently from each other, without specific agonistic or antagonistic behaviour. Indeed, despite extensive screening for pair-wise or higher-order interactions among mutations during the disappearance process, no significant interactions among mutations (after correction for multiple testing) were found. This behaviour seems to be not consistent with the ordered and stepwise accumulation of resistance mutations that characterizes the viral evolution under drug pressure, and that leads to the emergence of well-defined clusters of mutations\textsuperscript{7,18,24-27}.

Although different patterns in the evolution of drug resistant mutations can be observed during TI\textsuperscript{28-29}, most of the literature data show the total or partial time-dependent re-emergence of wild-type variants after drug-selective pressure is withdrawn\textsuperscript{30-37}. Similarly, our results show that therapy interruption induces a selection process that mainly consists in the loss of NRTI-resistance mutations over time. In particular, the estimated median survival time of drug resistance mutations analyzed in our study ranges between 3 and 12 months of TI (table 2).
Among the NRTI resistance mutations, the K65R, M184I and M184V mutations disappear rapidly (median survival time <4.3 months) and in almost all patients (100%, 100%, and 91.5%, respectively). Interestingly, the NRTI resistance mutations K65R and M184V share peculiar characteristics. Indeed, both mutations confer resistance to NRTIs by reducing the ability of reverse transcriptase to recognize and incorporate the nucleotide analogue. At the same time they confer zidovudine hypersusceptibility, suppress zidovudine resistance, and impair the rescue of chain-terminated DNA synthesis. Moreover, these mutations are localized in domains that are critical for the activity of the reverse transcriptase enzyme, and thus their appearance is associated with a marked decrease in viral replication capacity. Despite this replicative defect, the emergence of K65R or M184V is not associated with the accumulation of compensatory mutations able to restore viral replication capacity. The rapid loss of the K65R and M184V mutations can be just related to their ability to drastically impair viral fitness, combined to the inability of the virus to build a network of compensatory mutations. The rapid dynamics of appearance and disappearance of M184V/I mutations was already observed after 3TC failure and interruption, suggesting easy development of such a mutation, but also suggesting a remarkable growth disadvantage for HIV harbouring the mutation. Of note, we observed a trend in which the loss of the M184V mutation mainly coincided with the re-emergence of the wild-type strains. These results suggests that the detrimental presence of the M184V mutation may drive the rapid loss of other drug resistance mutations, and, presumably, a shift from drug-resistant strains to wild-type strains takes place. Thus, it is conceivable to hypothesize that the complete interruption of the antiretroviral regimens triggers a competition between the viral strain(s) with M184V (+/- other resistant mutations) and the wild-type strain(s) (present as minority quasispecie under the drug pressure). This competition leads to the outbreak of the wild-type strain that in the absence of drug pressure has a replication capacity higher than the drug-resistant strain. The wild-type strain rapidly overcomes the drug-resistant strains and becomes the predominant viral specie. The re-emergence of the wild-type virus depends by its presence in cellular/compartment reservoirs. We know that the nature of the archived copies
of viral DNA, may depend on how long the patient was infected before starting therapy, how long the patient was on therapy, and how long was on a failing therapy. Therefore, all these factors are going to make major contributions to the ability of an archival wild-type virus to re-emerge.

Similarly to K65R and M184V/I mutations, we found that the novel NRTI mutation K43E is one of the fastest mutation to disappear (median survival time 3.6 months). The involvement of K43E in NRTI resistance has been recently investigated in a recent work by our group\(^\text{18}\). In this study, we observed that the K43E was completely absent in drug-naïve patients, significantly increased in NRTI-treated patients (around 8.4%) and was significantly correlated with the TAM1 pathway. In addition, we observed that the co-presence of K43E with TAMs1 is associated with an increased zidovudine resistance and high level of viremia at virological failure. Consistent with this finding, another recent work, showed a high correlation of K43E with TAMs1 and that the K43E change increased viral replication when combined with another novel mutation (E40F)\(^\text{38}\). Therefore, the fact that K43E novel mutation disappears during TI with a dynamics of disappearance similar to other classical NRTI mutations, supports the hypothesis that this novel mutation is involved in the complex mechanisms underlying NRTI resistance.

Regarding the complexity and different patterns of mutations regulating resistance and viral fitness, as expected, in a subset of 50 patients that interrupted therapy containing also a NNRTI, we observed that NNRTI mutations persist relatively longer compared to the NRTI substitutions. In particular, among the major and most frequent NNRTI mutations, the G190A and Y181C mutations disappeared after 6 and 8 months of TI, respectively, while the K103N mutation showed a slower dynamics of disappearance, being detected after 9 months of TI in 20% of patients (data not shown).

The competition between drug resistant and wild-type viruses is not the only mechanism that can explain the loss of drug resistance mutations during TI. Indeed, in our study we showed that the disappearance of the T215F/Y mutations can be associated with the appearance of the T215 revertants (T215C/I/N/S). It is known that the emergence of the T215Y or T215F mutations
requires two nucleotide mutations from ACT/C to TAT/C, and from ACT/C to TTT/C, respectively. It is conceivable that during TI, a viral strain with the T215Y/F mutations can undergo a process of genetic drift that implies the acquisition of single nucleotide mutations at position 215 and consequently the appearance of the T215 revertants. We observed that these revertants can persist in the absence of drug pressure within a median length of TI that lies between median time of maintenance of T215Y/F and median time of TI for complete disappearance of such mutations. This persistence can be due to the fact that these revertant forms should not confer a high disadvantage in terms of viral fitness. Indeed, a previous study demonstrated that in presence of the T215 revertants, viral replication capacity is similar to that observed for the wild-type virus. The persistence of the T215 revertants during TI represents an important clinical issue since in vitro studies have demonstrated that the presence of the T215 revertants can speed up the appearance of the T215Y mutation, thus decreasing the genetic barrier to drug resistance. In addition, other studies demonstrated that the presence of T215 revertants confer increased risk of virologic failure of zidovudine or stavudine in antiretroviral-naive patients. Thus, further studies are necessary to investigate whether the presence of T215 revertants at the re-start of an antiretroviral regimen after therapy interruption is associated with a poorer virological response and virological failure.

During TI, viral evolution is not constrained by drug pressure, and thus the different viral quasispecies can evolve independently of each other by different mechanisms. The independent evolution of the viral quasispecies can explain why we did not observe any significant interaction among mutations on their disappearance during therapy interruption. Though, to support this hypothesis for all mutations, larger studies with a greater number of patients and samples are warranted.

Nevertheless, our results underline that viral fitness plays a key role in modulating viral evolution during TI. In general, mutations detrimental for viral fitness are rapidly lost (as in the case of K65R, and M184V), while mutations that do not drastically impair viral fitness or that confer a replicative advantage are lost more slowly or are maintained. This is the case of the mutations...
K122E, G196E, and F214L. These mutations have been recently associated with NRTI treatment, are common polymorphisms in drug-naïve patients, and have a notably long time of disappearance during TI (8-11.6 months), thus suggesting that they contribute positively to the viral replication both in presence and in absence of drug pressure. In particular, the F214L mutation is a natural polymorphism detected in around 18% of drug-naive and NRTI-treated patients, that has been negatively associated with type-1 NAMs, and positively associated with type-2 NAMs. As a confirmation of this finding, we found that the disappearance of the F214L mutation during TI is associated with the disappearance of the NAM2 K219Q, while, in other patients, the appearance of F214L mutation during TI is associated with the disappearance of NAMs1 in 2 out of 6 patients. These results support that this mutational shift may represent a determinant for the choice of a specific NAM pathway under thymidine analogues pressure.

Finally, another mutation worth to be discussed is the R83K, a common polymorphism in drug-naïve patients, that has been negatively associated with NRTI failure and with the presence of NAMs. During TI, the R83K mutation does not disappear at all; if something, it appears in 9.8% of patients undergoing TI, and its appearance was always associated with the disappearance of both NAM1 and NAM2. These findings suggest that the R83K mutation is selectively neutral in wild-type strains but potentially deleterious for viral replication in the presence of NAMs. The antagonism between R83K and the NAMs strongly supports the R83K as a “protective” polymorphism able to interfere with the accumulation of NAMs, and consequently to prolong the benefits of antiretroviral regimens. Consistent with this hypothesis, a recent study demonstrated that the presence of R83K at baseline is associated with a favourable virological response to thymidine analogues-containing regimens, and with very limited emergence of NAMs development.

In conclusion, even if this study includes a limited number of patients and of follow-up sequences, we showed that the loss of NRTI drug resistance mutations during TI is not a stepwise or ordered process, and seems to proceed largely without specific interaction among mutations, and
that both genetic drift and shift may drive the viral evolution towards the emergence of the most fitted virus.

**ACKNOWLEDGMENTS**

This work was financially supported by grants from the Italian National Institute of Health, the Ministry of University and Scientific Research, Current and Finalized Research of the Italian Ministry of Health and the European Community (QLK2-CT-2000-00291, and the Descartes Prize HPAW-90001). The research at Max-Planck-Institute for Informatics was performed in the context of the EU network of Excellence BioSapiens (EU grant No. LSHG-CT-2003-503265). We thank Caterina Gori, Ada Bertoli, Fabio Continenza, Daniele Pizzi, Andrea Biddittu, Sara Bono, Lavinia Fabeni, Francesca Stazi and Sandro Bonfigli for sequencing and data management, all other participants and members of INMI-Collaborative Group for Clinical Use of HIV Genotype Resistance Test.

Members of the Collaborative Group for Clinical Use of HIV-1 Genotype Resistance Test (GRT) at National Institute for Infectious Diseases ‘Lazzaro Spallanzani’ in Rome are: Andrea Antinori (Co-Chair), Adriana Ammassari, Gianfranco Anzidei, Francesco Baldini, Rita Bellagamba, Maria Concetta Bellocchi, Ada Bertoli, Sandro Bonfigli, Evangelo Boumis, Francesca Ceccherini-Silberstein, Bruno Christian Ciancio, Fabio Continenza, Roberta D’Arrigo, Patrizio De Longis, Gianpiero D’Offizi, Federica Forbici, Sara Giannella, Enrico Girardi, Caterina Gori, Giuseppina Liuzzi, Patrizia Lorenzini, Patrizia Marconi, Pasquale Narciso, Emanuele Nicasiri, Pasquale Noto, Carlo Federico Perno (Co-Chair), Pietro Sette, Fabio Soldani, Maria Paola Trotta, Valerio Tozzi, Ilaria Uccella, Ubaldo Visco-Comandini, Mauro Zaccarelli, Daniela Zinzi.
V. References


3.2 HBV

3.2.1 The Profile of Mutational Clusters Associated with Lamivudine Resistance

Can Be Constrained by HBV-genotypes

Running Title: Characterization of LMV Resistance Profiles

Valentina Svicher⁴⁺, Caterina Gori²⁺, Maria Trignetti¹, Michela Visca³, Valeria Micheli⁴,
Martina Bernassola³, Romina Salpini¹, Guido Gubertini⁴, Roberta Longo³, Fosca Niero⁴,
Francesca Ceccherini-Silberstein¹⁺², Giuseppe Maria De Sanctis³, Alberto Spanò³,
Giuseppina Cappiello³, Carlo Federico Perno¹⁺²⁺

Department of Experimental Medicine, University of Rome “Tor Vergata”, Rome,¹
National Institute for Infectious Diseases “L. Spallanzani”, Rome,² “Sandro Pertini” Hospital, Rome, Italy,³
“Luigi Sacco” Hospital, Milan, Italy,⁴ University of “La Sapienza”, Rome, Italy⁵


I. Introduction

To date, interferon, and four nucleos(t)ide reverse transcriptase (RT) inhibitors (NsRTIs: LMV-lamivudine, ETV-entecavir, TBV-telbivudine; NtRTIs: ADV-adefovir) have been approved for the treatment of chronic HBV-infection. Ns(t)RTIs can efficiently suppress HBV-replication in most patients and delay disease progression [1-3]. Unfortunately, when antiviral therapy fails to be fully suppressive, new viral variants emerge allowing HBV to become resistant to one or more Ns(t)RTIs by accumulating mutations either alone or in clusters.

LMV has the highest rate of drug-resistance emergence. LMV-resistance increases at a rate of 14-24% per year, reaching ~70% after 4-5 years of treatment [4]. In contrast to HIV, HBV can develop different patterns of LMV-resistance, and can add compensatory mutations able to restore viral fitness and/or increase the level of (cross)-resistance [5-9]. Although a relatively high number of mutations has been associated with LMV-resistance (together with the classical M204I/V), the patterns underlying LMV-resistance have not yet been completely characterized. This is mainly
related to the fact that the methodology of mutation-assessment is not standardized, since the majority of the studies evaluate only a short region encompassing the RT YMDD region. In addition, few information is still available regarding the factors that can affect drug-resistance development. Preliminary findings indicate that the HBV-genotype can drive drug-resistance evolution under drug-pressure. For instance, a higher occurrence of M204V was observed in A- and B-genotypes than in D- and C-genotype [10,11]. Nevertheless, no information is so far available regarding the impact of these genotypes on the other LMV-resistance mutations.

In addition, since the HBV genome is organized in such a way that the RT-gene is overlapping with the HBV surface antigen (HBsAg) gene, some LMV-resistance RT mutations produce mutations in the HBsAg that can result in a reduced HBsAg antigenicity [12-15]. It is also conceivable that some HBsAg mutations can act as compensatory mutations to preserve the correct HBsAg structure/function impaired by the LMV-resistance mutations. This point has not been so far investigated.

Based on this scenario, this study is aimed to i) provide, by using statistical and computational methods, a characterization of LMV-resistance profile in a well-defined cohort of HBV-infected patients, ii) identify factors that can influence their development, and iii) correlate LMV-resistance mutations with those observed in the HBsAg.

II. Patients and Methods

Patients. This study includes 89 HBV chronically-infected patients (61 HBV mono-infected patients and 28 HBV+HIV co-infected patients) followed in clinical centres in North and Central Italy between 2000-2007; all patients were failing their first line LMV-treatment at the time of genotypic-resistance testing. Virological-breakthrough was defined by a rebound of serum HBV-DNA of >1 log IU/ml from the nadir value confirmed by 2 consecutive determinations. Serum HBV-DNA was quantified using the COBAS AmpliPrep-Cobas TaqMan HBV test with lower limit of quantification 12IU/ml (CAP-CTM, Roche Molecular System, Branchburg, NJ, USA).
**HBV sequencing.** The sequencing of the HBV-RT (A, B, C, D, E domains) was performed on plasma samples as follows. HBV-DNA was extracted using a commercially available kit (QIAmp DNA blood mini-kit, QIagen Inc, USA), and then amplified with Amplitaq-Gold polymerase enzyme using the following primers’ pairs: 5’GGTCACCATATTCTTGGGAA and 5’GTGGGGGTGTCCGTGAGCAAA. PCR conditions were: one cycle at 93°C for 12min, 40 cycles (94°C 50s, 53°C 50s, 72°C 1min and 30sec), and a final cycle at 72° for 10min. PCR-products were sequenced by using 8 different overlapping sequence-specific primers, a BigDye terminator v. 3.1 cycle sequencing kit (Applied-Biosystems) and an automated sequencer (ABI-3100). The sequences were analyzed using SeqScape-v.2.0 software. The quality endpoint for each individual, was ensured by a coverage of the RT-sequence by at least two sequence segments. Sequences having a mixture of wild-type and mutant residues at single positions were considered to have the mutant(s) at that position. HBV-RT sequences were obtained at the failure of LAM-treatment. Being all these lamivudine-treated patients undergoing their first therapeutic regimen, no mutations pre-existing therapy are foreseen; thus pre-treatment samples were not sequenced.

To define HBV-genotype, sequences were compared to at least 30 Genbank sequences representative of all known HBV-genotypes. HBV-genotype was confirmed by the construction of phylogenetic-trees generated with the Kimura’s two-parameter model. The statistical robustness within each phylogenetic tree was confirmed with a bootstrap-analysis using 1,000 replicates. Calculations were performed with Phylip-v.3.65 software.

**Statistical analysis.**

**Mutation prevalence.** The frequency of mutations was calculated by stratifying our cohort of patients according with HIV co-infection, and with HBV-genotypes. Statistically significant differences were assessed by Chi-Squared tests, and corrected for multiple-hypothesis testing with the Benjamini-Hochberg method at a false discovery rate (FDR) of 0.05 [16]. The attention was focused on the LMV-associated mutations reported in the references #5-9, and on the immune- or diagnostic-escape mutations reported in the references #12-15.
The association of mutations with specific predictive variables was assessed by multivariate logistic-regression analysis (predictor variables considered are: subject-demographic, HBV-genotype, HIV co-infection, ADV co-administration, therapy duration).

The correlation between drug-resistance mutations number and LMV-treatment duration of was assessed by Mann-Whitney test.

Mutation covariation. In our cohort of HBV-infected patients, the covariation among mutations was investigated by calculating the binomial-correlation coefficient (phi) for the simultaneous presence of mutations at two positions in the same isolate. Statistically significant pairs of mutations were assessed by Fisher’s exact test and corrected for multiple-testing with the Benjamini-Hochberg method (FDR=0.05). Samples having a mixture of two or more mutations at a given pair of positions were ignored in calculating the covariation, due to the impossibility of identifying whether these mutations are located in the same viral genome.

Cluster-analysis. Mutational clusters were defined as clusters of three or more mutated positions in which each position was significantly correlated with each other, and were identified using an average linkage hierarchical agglomerative clustering, as described elsewhere [17-20]. Briefly, the phi coefficients were transformed into dissimilarity values by mapping phi=1 (maximal positive association) to dissimilarity 0, and phi=-1 (maximal negative association) to dissimilarity 1, with linear interpolation in between. The dissimilarity of mutations at the same position was left undefined, as such pairs never co-occur in a single sequence (except from mixtures) and would lead to artefacts in the resulting dendrogram. The resulting partial dissimilarity matrix was then used as input for average linkage hierarchical agglomerative clustering, and undefined dissimilarity values were ignored in computing average dissimilarities between clusters. The statistical robustness of the dendrogram was confirmed with a bootstrap-analysis using 500 replicates. A bootstrap value of 1 for an edge in the dendrogram means that the set of mutations in the induced subtree occurs as a subtree in all dendrograms from the different bootstrap replicates. Thus, higher bootstrap values indicate that the association of mutations into a group is not due to sampling bias.
III. RESULTS

**Patients’ characteristics.** Our cohort of 89 patients included 61 HBV mono-infected patients, and 28 HBV+HIV co-infected patients. These two groups of patients were homogenous, with significant difference only for the age, that was higher in HBV mono-infected than in HBV+HIV co-infected patients, and for the proportion of negative serum HBeAg patients, that was higher in HBV mono-infected than in HBV+HIV co-infected ones (Table 1). HBV mono-infected patients were failing their first line LMV-treatment after a median time of 44 months (IQR:24-61) and with a median viremia of 3.9 logIU/ml (IQR:2.7-5.0) (Table 1). LMV was given as monotherapy (100mg/daily) in all patients but 10 (treated also with ADV). HBV+HIV co-infected patients were failing their first line LMV-treatment (administered at 300mg/daily) after a median time of 32 months (IQR:17-57) and with a median viremia of 4.8 logIU/ml (IQR:2.8-6.9) (Table 1). No patients were treated with drugs potentially effective against HBV other than LMV.

Since previous European studies [21,22] highlighted a different distribution of HBV-genotypes in HBV mono- and HBV+HIV co-infected patients, we performed a genetic analysis to verify this result also in our cohort from Italy. In agreement with them, we found that D-genotype is the predominant in HBV mono-infected patients (prevalence, 90.1%), while A-genotype is the most prevailing one in HBV+HIV co-infected patients (prevalence, 57.1%) (Table 1). We also detected the E, F, G genotypes in one HBV+HIV co-infected patient, respectively.

**HBV RT mutations.** On these bases, we assessed the prevalence and type of HBV RT mutations. LMV-associated mutations were detected in 85 out 89 LMV-failing patients (median number=3 [IQR:2-3]). The primary mutations rtM204V, rtM204I, and rtM204I/V were observed in 43 (48.3%), 32 (35.9%), and 7 (7.8%) patients, respectively, and rtL180M was the most common LMV-resistance mutation observed (prevalence, 65 [73%]). rtS135Y was detected in the 3 patients failing LMV-treatment without rtM204V or rtM204I.

The prevalence of LMV-resistance mutations was not significantly different in HBV mono- and HBV+HIV co-infected patients, with the exception of rtM204I and rtQ215S, less prevalent in
HBV+HIV co-infected than in HBV mono-infected patients (17.8% versus 44.2%, P=0.01; 0 versus 13.1%, P=0.04, respectively). rtQ215S (known to confer also ADV-resistance [5-8]) was detected in 7 patients failing LMV-monotherapy, and in 1 patient failing LMV+ADV.

The differences in the prevalence of LMV-resistance mutations were more accentuated when we stratified our cohort of patients according to HBV-genotypes (Table 2).

Table 1. Main Characteristics of the Study Population

<table>
<thead>
<tr>
<th></th>
<th>Patients infected with</th>
<th>P values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HBV N (%/61)</td>
<td>HBV+HIV N (%/28)</td>
</tr>
<tr>
<td>Country of origin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>56 (91.8)</td>
<td>23 (82.4)</td>
</tr>
<tr>
<td>Other countries</td>
<td>5 (8.2)</td>
<td>5 (17.8)</td>
</tr>
<tr>
<td>Median (IQR) age, years</td>
<td>55 (45-63)</td>
<td>43 (41-49)</td>
</tr>
<tr>
<td>Negative serum HBeAg, N (%)</td>
<td>39 (81.2)</td>
<td>9 (52.9)</td>
</tr>
<tr>
<td>Median (IQR) treatment duration, months</td>
<td>44 (24-61)</td>
<td>32 (17-57)</td>
</tr>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LMV</td>
<td>51 (83.6)</td>
<td>28 (100)</td>
</tr>
<tr>
<td>LMV+ADV</td>
<td>10 (16.3)</td>
<td>0</td>
</tr>
<tr>
<td>Median viremia (IQR) at failure, log IU/ml</td>
<td>3.9 (2.7-5.0)</td>
<td>4.8 (2.8-6.9 )</td>
</tr>
<tr>
<td>Median ALT (IQR) at failure, IU/L</td>
<td>44.5 (30-77)</td>
<td>54 (32-109)</td>
</tr>
<tr>
<td>Median AST (IQR) at failure, IU/L</td>
<td>37 (23-53)</td>
<td>40 (26.7-76.2)</td>
</tr>
<tr>
<td>HBV genotype</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>55 (90.1)</td>
<td>9 (32.1)</td>
</tr>
<tr>
<td>A</td>
<td>6 (9.8)</td>
<td>16 (57.1)</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>1 (3.6)</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>1 (3.6)</td>
</tr>
<tr>
<td>G</td>
<td>0</td>
<td>1 (3.6)</td>
</tr>
</tbody>
</table>

a The percentages were calculated upon 61 and 28 patients for HBV and HBV+HIV, respectively, except for HBeAg status whose data were available for 48 and 17 patients, respectively.
b P values were assessed by Mann Whitney test.
c P values were assessed by Chi Squared test. Abbreviation: LMV, lamivudine; ADF, adefovir; HBeAg, HBV e antigen; ALT, alanine aminotransferase; AST, aspartate amino transferase; IQR, interquartile range; n.s., not significant
In particular, we found that A-genotype was more prone to develop rtM204V than rtM204I (68.2% versus 13.6%, P=0.001). This result was also supported by multivariate-analysis showing that A-genotype is the only predictor variable for the emergence of rtM204V (OR:14.5[95%CI:1.3-
Thus, this result supports that rtM204V is the main sign of LMV-failure in A-genotype. This situation was not observed in D-genotype, characterized by a similar occurrence of rtM204V and rtM204I (39.1% versus 45.3%, P=0.47).

Table 3. Odds ratio for the presence of the M204V mutation from fitting a multivariable logistic regression model in patients failing lamivudine treatment.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Univariable</th>
<th>Multivariable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>95% CI (P value)</td>
</tr>
<tr>
<td>Genotype</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>1.00</td>
<td>3.85</td>
</tr>
<tr>
<td>A</td>
<td>1.00</td>
<td>1.27-11.71 (P=0.01)</td>
</tr>
<tr>
<td>Co-infection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>without HIV</td>
<td>1.00</td>
<td>1.08</td>
</tr>
<tr>
<td>with HIV</td>
<td>1.00</td>
<td>0.42-2.75 (P=0.87)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>per 10 years older</td>
<td>1.03</td>
<td>0.92-1.15 (P=0.38)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>1.00</td>
<td>2.62</td>
</tr>
<tr>
<td>M</td>
<td>1.00</td>
<td>0.87-7.91 (P=0.08)</td>
</tr>
<tr>
<td>Length of treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>per month longer</td>
<td>1.00</td>
<td>0.99-1.00 (P=0.36)</td>
</tr>
<tr>
<td>ADV treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>3.22</td>
</tr>
<tr>
<td>Yes</td>
<td>1.00</td>
<td>3.22</td>
</tr>
</tbody>
</table>

In addition, we observed that the prevalence of rtL229V under LMV-failure was significantly greater in A- than D-genotype (27.2% versus 7.8%, P=0.02), while the prevalence of rtL80I and rtS135Y was significantly greater in D- than A-genotype (23.4% versus 4.5%, P=0.04; 17.2% versus 0, P=0.003, respectively). Differently from rtL80I, the prevalence of rtL80V was similar between the two HBV-genotypes (17.2% versus 13.6%, P=0.69).

Beyond the LMV-associated mutations, we detected the amino acid variants rtS53N, rtT54H, rtT54Y, and rtQ130P specifically associated with D-genotype (72.3% versus 0, P=0.007; 43.6% versus 4.5%, P=0.001; 43.6% versus 0, P=<0.001; 62.5% versus 0, P<0.001).

Overall results suggest that, beyond HIV-1 co-infection, HBV-genotype plays a role in driving the evolution of HBV drug-resistance under drug-pressure.
**HBV HBsAg mutations.**

Due to the overlapping between the RT and the HBsAg genes, some LMV-resistance mutations produce mutations also in the HBsAg [12-15]. Among the immune-escape mutations associated with LMV-treatment, the pairs sI195M+sE164D (corresponding to rtM204V+rtV173L), known to drastically reduce HBsAg antigenicity [13], was observed in 4.7% and 9.1% of patients infected with D- and A-genotype, respectively.

Beyond the known HBsAg mutations, we also detected sE164V and sS207N associated with A-genotype (18.2% *versus* 0, P=0.003; 36.4% *versus* 12.5%, P=0.013). Both mutations do not correspond to amino acid changes in RT. sE164V, localized in the “a-determinant”, is a novel variant at a position already associated with immune-escape, while sS207N is localized in the fourth transmembrane HBsAg domain.

As an additional data (not associated with LMV-treatment) we found that the threonine (T) is the wild-type amino acid at the HBsAg position 131 of D genotype: the mutation sT131N was observed only in 3.6% of these patients. In contrast, the asparagine (N) is the wild-type amino acid at position 131 in the consensus sequence of HBsAg A genotype, and was found in 95.4% of patients carrying this genotype. Since the N at position 131 has been associated with HBV immune- and diagnostic-escape [15], this suggests that A genotype can have potentially an intrinsic propensity to an easier escape from the immune pressure.

**Associations among mutations.** To identify significant patterns of pair-wise correlations between mutations observed in 64 D-genotype infected- and 22 A-genotype infected-patients, we calculated the phi-coefficient and its statistical significance for each pair of mutations (Table 4). A positive and statistically significant correlation between mutations at two specific positions (0<phi<1, P<0.05) indicates that under drug-pressure these two positions mutate in a correlated manner in order to confer an advantage in term of viral-fitness or drug-resistance, thus it indicates that the co-occurrence of mutations is not due by chance.
We found that in both genotypes rtM204V always (100%) occurred with rtL180M (D-genotype: phi=0.60; A-genotype: phi=0.87). In D-genotype, rtM204V was also correlated with rtT184A/S (phi=0.34), and rtV173L (phi=0.28): rtT184A/S and rtV173L were never found (0%) without rtM204V. In A-genotype, rtM204V was correlated with rtL229V (phi=0.42): rtL229V was never found (0%) without rtM204V. These results indicate that the HBV-genotype can influence the type of compensatory mutations observed with rtM204V.

Of note, rtL229V in RT corresponds to sF220L in the HBsAg. sF220L was correlated with the HBsAg variant sS207N (phi=0.59). Independently from the co-infection status, this pair of mutations was observed only in patients infected with A-genotype. Both sS207N and sF220L localized in the fourth transmembrane HBsAg domain.

rtM204I showed a behaviour different from that observed for rtM204V. First of all, this mutation occurred also alone without any other LMV-resistance mutation (in contrast with rtM204V) in 9 (14.1%) and in 1 (4.5%) of patients infected of D and A-genotype, respectively. In addition, it was positively correlated with the rtL80I/V (D: phi=0.33; A: phi=0.86). This correlation

### TABLE 4. Statistically significant correlated pairs of HBV reverse transcriptase and HBsAg mutations

<table>
<thead>
<tr>
<th>RT P-value&lt;sup&gt;a&lt;/sup&gt; mutations</th>
<th>Frequency</th>
<th>Correlated</th>
<th>Frequency</th>
<th>Covariation</th>
<th>Phi&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N&lt;sub&gt;0&lt;/sub&gt; (%)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Mutation&lt;sup&gt;a&lt;/sup&gt;</td>
<td>N&lt;sub&gt;0&lt;/sub&gt; (%)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Frequency (%)&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td><strong>D genotype</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rtM204I</td>
<td>29 (45.3)</td>
<td>rtL80I/V</td>
<td>24 (37.5)</td>
<td>18 (62.1)</td>
<td>0.33</td>
</tr>
<tr>
<td>rtM204V</td>
<td>25 (39.1)</td>
<td>rtL180I/M</td>
<td>45 (70.3)</td>
<td>25 (100)</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rtT184A/S</td>
<td>6 (9.4)</td>
<td>6 (24.0)</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rtV173L</td>
<td>3 (4.7)</td>
<td>3 (12.0)</td>
<td>0.28</td>
</tr>
<tr>
<td>rtL80I/V</td>
<td>16 (36.4)</td>
<td>rtT54Y</td>
<td>17 (46.3)</td>
<td>10 (62.5)</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rtS53N</td>
<td>17 (72.3)</td>
<td>16 (100)</td>
<td>0.46</td>
</tr>
<tr>
<td><strong>A genotype</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rtM204I</td>
<td>3 (13.6)</td>
<td>rtL80I/V</td>
<td>4 (18.2)</td>
<td>2 (66.7)</td>
<td>0.86</td>
</tr>
<tr>
<td>rtM204V</td>
<td>15 (68.2)</td>
<td>rtL180I/M</td>
<td>18 (81.8)</td>
<td>15 (100)</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rtL229V</td>
<td>6 (27.2)</td>
<td>6 (40)</td>
<td>0.42</td>
</tr>
<tr>
<td>rtL229V</td>
<td>6 (27.2)</td>
<td>sS207N</td>
<td>8 (36.4)</td>
<td>5 (83.3)</td>
<td>0.59</td>
</tr>
</tbody>
</table>

<sup>a</sup>The frequency was determined in 64 and 22 RT sequences from HBV D and A genotype infected patients with virological failure to LMV, respectively. Only, the correlation of L80I/V with S53N and T54Y was determined in a subset of 44 RT sequences encompassing the amino acids 50-345 from D genotype infected patients.

<sup>b</sup>Percentages were calculated for patients containing each specific mutation reported in the first column.

<sup>c</sup>Positive and negative correlations with phi > 0.20 and phi < -0.20, respectively, are shown.

<sup>d</sup>P values in boldface remained significant also after correction for multiple hypothesis testing (Benjamini & Hochberg, 1995).

<sup>e</sup>S207N is a HBsAg mutation
Clusters of correlated mutations. The topology of the dendrograms suggests different patterns of LMV-resistance mutations in patients infected with HBV D and A-genotype, respectively (Fig. 1). In particular, in D-genotype, LMV-resistance is mainly mediated by two different clusters of mutations. One cluster is formed by rtM204V+rtL180M (bootstrap=1) (prevalence, 45.3%), that was linked to rtV173L and rtT184A/S (bootstrap=0.94). The other cluster is formed by rtM204I+rtL80I/V (bootstrap=0.56) (prevalence, 28.1%), linked to rtS53N and rtT54Y (bootstrap=0.96). It is interesting to note that in D-genotype, the primary rtM204V mutation clusters with mutations (rtV173L, rtL180M, rtT184A/S) localized in the RT B domain, while M204I with mutations (rtS53N, rtT54Y, rtL80I/V) localized in the RT A domain, suggesting that rtM204V and rtM204I require different patterns of mutations to compensate for viral fitness.

In A-genotype, the cluster formed by rtM204V+rtL180M, with a prevalence of 68.2%, is the main sign of LMV-resistance and failure (bootstrap=0.99). The rtM204V and rtL180M were also linked to rtL229V and sS207N (bootstrap=0.95).

The other less prevalent cluster of mutations involved rtM204I and rtL80I/V (bootstrap=0.99) (prevalence, 9.1%). These two mutations were not correlated with any other RT or HBsAg mutations.
**D genotype**

![Dendrogram for D genotype](image)

Prevalence 28.1%  
Prevalence 45.3%

**A genotype**

![Dendrogram for A genotype](image)

Prevalence 9.1%  
Prevalence 68.2%

**Legends to figures**

**Figure 1** Dendrogram obtained from average linkage hierarchical agglomerative clustering, showing clusters of LMV-resistance mutations in HBV D and A-genotype infected patients. The length of branches reflects distances between mutations in the original distance matrix. Bootstrap values, indicating the significance of clusters, are reported in the boxes. A bootstrap value of 1 for an edge in the dendrogram means that the set of mutations in the induced subtree occurs as a subtree in all dendrograms from the different bootstrap replicates. Thus, a bootstrap values >0.75 indicate that the association of mutations into a group is not due to sampling bias.
Correlation between duration of LMV-treatment and number of drug-resistance mutations.

The correlation among the number of LMV-resistance mutations and the duration of LMV-treatment showed that the presence of rtM204V+rtL180M+rtT184A/S was significantly associated with a LMV-treatment longer than that observed in presence of rtM204V+rtL180M alone (2,063 days versus 944 days, P=0.039) (Table 5), thus indicating that the maintenance of a failing regimen determines the accumulation of LMV-resistance mutations.

Table 5. Association between the number of reverse transcriptase mutations and the duration of lamivudine treatment

<table>
<thead>
<tr>
<th>Pattern of RT mutations</th>
<th>N° of patients N (%)</th>
<th>Duration of lamivudine treatment Median days (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M204I</td>
<td>9 (14.1)</td>
<td>673 (381-1,439)</td>
</tr>
<tr>
<td>M204I + L80I/V or L180M</td>
<td>12 (18.7)</td>
<td>1,253 (815-1,524)</td>
</tr>
<tr>
<td>M204I + L80I/V + L180M</td>
<td>12 (18.7)</td>
<td>1,804 (1,048-2,035)</td>
</tr>
<tr>
<td>M204V</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>M204V + L180M</td>
<td>25 (39.1)</td>
<td>944 (1,732-6,162)</td>
</tr>
<tr>
<td>M204V + L180M + T184A/S</td>
<td>6 (9.4)</td>
<td>2,063 (1,962-2,127)</td>
</tr>
</tbody>
</table>

* The analysis was performed in 64 HBV D genotype infected patients.

\* The difference in the duration of lamivudine treatment between M204V and M204V+L180M+T184A/S is statistically significant (P=0.039, Mann Whitney test).

### IV. DISCUSSION

To our knowledge, this is the first study that characterizes the mutational clusters induced by LMV-pressure in two different HBV-genotypes.

Starting from previously published data [10], we confirmed that the A-genotype is more prone to develop rtM204V than rtM204I at LMV-failure. Another study showed that the occurrence of rtM204V is higher in patients infected with B- than in C-genotype [11]. The role of HBV-genotype as a factor potentially affecting virological-response to LMV is still matter of controversy [23-25],

98
yet overall findings suggest a potential effect of HBV-genotype in modulating resistance development under LMV-pressure.

We found that in D-genotype LMV-resistance is mainly mediated by two distinct clusters of mutations. One cluster is formed by the primary mutation rtM204V plus mutations localized in the HBV RT-B domain (rtL180M, rtV173L, and rtT184A/S). This cluster is consistent with previous studies showing that rtL180M can compensate the loss of replicative capacity due to rtM204V and also increase the level of LMV-resistance [5,26,27], while rtV173L can only act as compensatory mutation by positioning either the template strand of HBV nucleic acid or other residues critical for the polymerization reaction [28]. The other mutational cluster observed in D-genotype is formed by the primary mutation M204I plus mutations localized in the HBV RT-A domain (rtS53N, rtT54Y, rtL80I/V). The association between rtM204I and rtL80I/V is consistent with a previous study showing the ability rtL80I/V to enhance HBV replication impaired by rtM204I [29].

Differently, in A-genotype the main sign of LMV-resistance involves the rtM204V and rtL180M that is linked to the pairs rtL229V (corresponding to sF220L in the HBsAg) and sS207N. Based on the HBsAg-structure proposed by Stirk [30], both sF220L and sS207N are localized near with each other in the fourth transmembrane HBsAg domain. It is conceivable that the structural modification induced by sF220L in the fourth transmembrane HBsAg domain may be compensated by a mutational change at position 207. This supports that HBsAg mutations can play an indirect role in drug-resistance presumably by preserving the correct structure/function of the HBsAg, and thus sufficient levels of replication capacity.

A potential limitation of this study is the unavailability of pre-treatment samples, thus we cannot totally exclude that the novel amino-acid variants analyzed can be already present before the beginning of treatment. To add information on this point, we selected from our database a set of 35 and 15 drug-naïve patients infected with D- and A-genotype respectively, and we compared the prevalence of these variants in drug-naïve and in LMV-failed patients. The A-genotype associated rtL229V and sS207N occurred in 0 and in 1 (6.7%) drug-naive patient, and their prevalence
significantly increased to 27.2% and 36.4% in LMV-failed patients, respectively (P=0.027-0.039), thus suggesting their association with NRTI-treatment. This is consistent with a recent study showing the rtL229V emergence under ETV-treatment [31]. The D-genotype associated rtS135Y is present in 8.5% of drug-naïve patients, in 17.2% of LMV-failed patients, and it has been observed at baseline in some patients with a sub-optimal tenofovir-response [32]. Similarly rtQ215S, known to confer LMV- and ADV-resistance [5-8], occurred in 5.7% of drug-naïve and in 12.5% of LMV-failed patients infected with D-genotype. While further data remain necessary to define this point, the limited presence/absence of the above-mentioned mutations in drug-naïve patients suggests that they are mostly associated with LMV-failure.

Differently, rtS53N and rtT54Y are two common variants in both drug-naïve and LMV-failing patients infected with D-genotype (rtS53N: 62.8% and 72.3, rtT54N: 37.1% and 43.6% in drug-naïve and in LMV-failed patients, respectively). In this case, their association with the rtL80I/V mutations suggests that they could act as genetic determinants able to favor the development of rtL80I/V, thus stabilizing the secondary structure of the RT-A domain. Although further studies are necessary, this could potentially explain the low genetic barrier of D-genotype for L80I/V development observed by our study and also by the study of Warner and co-workers [29].

The different patterns of LMV-resistance mutations between D and A-genotype arise some important considerations in term of cross-resistance to the new anti-HBV drugs ETV and TBV. It has been demonstrated that the baseline presence of rtM204V+rtL180M, but not that of the rtM204I, favors the appearance of mutations associated with ETV-resistance at position 184, 202, and 250 [33]. The co-presence of these mutations with rtM204V+rtL180M has been shown to decrease the access to the ETV-binding pocket, thus conferring higher level ETV-resistance [34]. Among them, rtT184A (when present with rtM204V+rtL180M) was shown to be associated with high level ETV-resistance, and with increased viral replication in in vitro experiments [35], and with clinical virological breakthrough in LMV-refractory patients receiving ETV-treatment [33,35].
Of note, in our cohort of ETV-naïve patients, the cluster rtM204V+rtL180M+rtT184A was detected only in D-genotype with a prevalence of 4.7%, and was significantly associated with a longer duration of LMV-treatment. These findings support the potential use of genotypic resistance testing at baseline in order to set-up a rationale anti-HBV therapy, and highlight the importance to detect drug-resistance early to avoid the accumulation of drug-resistance mutations, and to decrease the probability of cross-resistance to the other anti-HBV drugs.

Recently, the Globe-Trial demonstrated that \textit{in vitro} TBV-resistance is conferred by rtM204I, and by rtM204V+rtL180M, but not by the rtM204V single mutation, while \textit{in vivo} TBV-resistance is almost exclusively due to the presence of M204I [36]. The activity of TBV against M204V opens important perspectives regarding the correct TBV positioning in patients failing LMV. It is conceivable that the switch from LMV to TBV may be warranted at the early virological-failure in presence of the rtM204V alone (and perhaps of rtM204V+other mutations different than rtL180M). We never found rtM204V alone in either D- and A-genotype, however, we can not exclude that this can occur at earlier stages of virological failure (i.e. immediately after rebound of viral load), before virus had time and chance to build mutations (such as L180M) compensatory to M20V-induced loss of fitness. This strategy could be particularly applied in patients infected with A-genotype where the pattern rtM204V+rtL180M is the main sign of LMV-failure.

In conclusion, our study shows that HBV-genotype plays a role in driving HBV-evolution under drug-pressure. The characterization of LMV-resistance profile in the different HBV-genotypes is important for the set up of a correct sequencing of the anti-HBV drugs.

**ACKNOWLEDGMENTS**

We thank Sara Romano, Valentina Gallinaro, and Amalia Mastrofrancesco for data-management.
V. REFERENCES


**Acknowledgment**

I would like to express my deepest gratitude to everyone who has contributed to this thesis. In particular I would like to thank:

- My supervisors Dr Valentina Svicher and Dr Francesca Ceccherini-Silberstein for their scientific guidance.
- Professor Perno for giving me the opportunity to work in stimulating research group constituted by persons with high human and scientific profile.
- Prof Stefano Aquaro
- All people working in Prof Perno’s group: Maria Santoro, Michela Pollicita, Ada Bertoli, Federica Forbici, Caterina Gori, Roberta D’Arrigo, Andrea Bidittu, Claudia Alteri, Danilo Armenia, Romina Salpini, Lavinia Fabeni, Francesca Stazi and Salvatore Dimonte for their excellent skill in HIV and HBV sequencing and data management. They have provided all the sequences and the laboratory data used to perform the studies here presented.
- Tobias Sing for introducing me to computational biology.