# Literature Review

## CHAPTER 1

### 1. 1 Overview of HIV/AIDS

Human Immunodeficiency Virus (HIV) is a human pathogenic virus that cause AIDS (Acquired Immunodeficiency Syndrome). HIV/AIDS has been global pandemic for over the last three decades and is depicted as the modern day plague (Quinn, 1996). The United Nations Acquired Immune Deficiency Syndrome (UNAIDS) global report 2012 estimates that by the end of 2011 approximately 34 million people were living with HIV by the end of 2011 (WHO factsheet Number 360 (<http://www.who.int/mediacentre/factsheets/fs360/en/>)) and that over 95% of them are living in low and middle in come countries (Esparza and Bhamarapravati, 2000). There is a significant variation in HIV prevalence among the countries around the globe (Figure 1.1). The UNAIDS 2012 report shows that although the global trend of new HIV infections and HIV-related deaths per year is declining, the current number of HIV infections is the highest since 1990 (Figure 1.2). The sub-Saharan region of Africa is the region most aggravated by the virus with 23.5 million people living with HIV (UNAIDS). UNAIDS estimates that approximately 1 in every 20 adults is HIV infected in this region (UNAIDS). This is 25 or more times the HIV prevalence in any other region of the world. Countries in Sub-Saharan Africa also have varying HIV prevalence with South Africa at the top followed by Nigeria (Esparza and Bhamarapravati). The next severely affected regions, besides African continent, are Asia (China, Thailand, Indonesia), Caribbean and Eastern Europe, North America, western and central Europe (Esparza and Bhamarapravati).

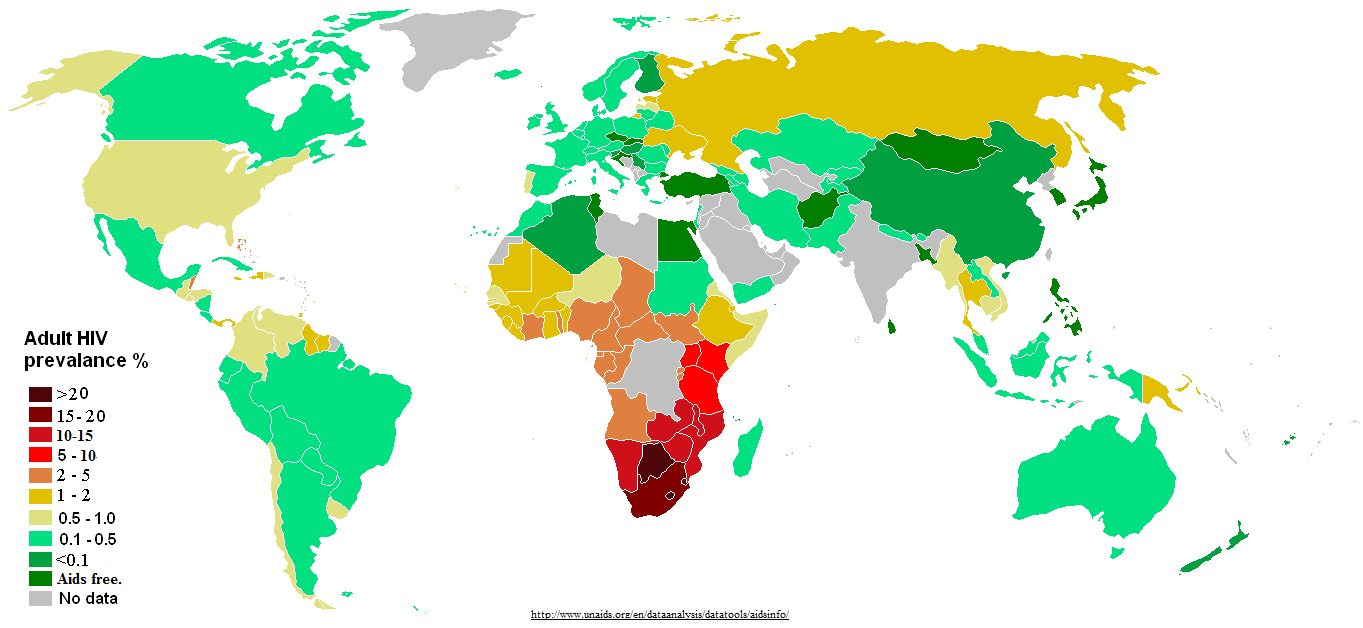


Figure 1.1: Distribution of HIV prevalence in adults around the world.

Source: http://www.unadis.org/en/dataanalysis/datatools/aidsinfo/

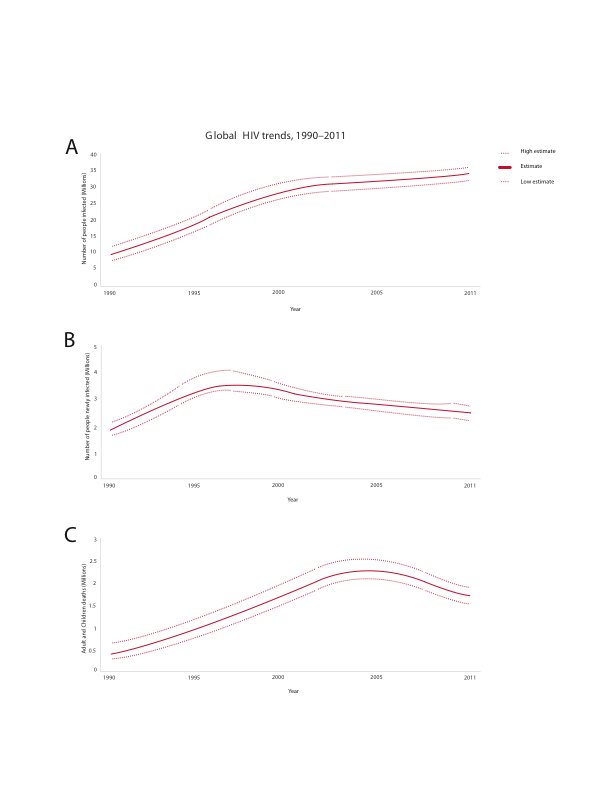


Figure 1.2: Global trend of new HIV infections from 1990 to 2011. The number of people living with HIV globally is increasing (A) while the number of people newly infected with HIV (B) and the number of Adults and child deaths due to HIV are decreasing (C) globally in the time period. This scenario can be attributed to global scale up of drugs while infected people continue transmitting the virus to uninfected people (Zaidi et al.). (Source: modified from UNAIDS 2012)

### 1.2 Discovery and characterization of HIV

As early as 1959, HIV infection cases had been documented but were unreported (Corbitt et al., 1990; Nahmias et al., 1986). Curious doctors at that time collected patient blood samples and kept frozen, which were later shown to have HIV antibodies (Zhu et al., 1998). In June 1981, a case of acute immune depletion associated secondary infection was reported in some homosexuals in the USA (Friedman-Kien, 1981; Friedman-Kien et al., 1981). Their infection was coupled with no lymphocyte proliferation (Gottlieb et al., 1981). Until 1983, the causative agent responsible for the severe immune depletion, named AIDS (Acquired Immunodeficiency Syndrome), was unknown (Francis et al., 1983; Gallo et al., 1983), when Luc Montagnier’s group at “Institut Pasteur” in Paris isolated the virus, which was initially named Human T-cell Leucamia Virus (HTLC) and later named as Human Immunodeficiency Virus (HIV) (Nahmias et al., 1986). Jay Levy’s group in San Francisco, USA also subsequently found the virus confirming the discovery in Paris (Levy et al., 1984). They found HIV was a lentivirus from Group VI retrovirus with two single strand RNA molecules (Baltimore, 1971); unique to any other previously isolated viruses and the virus can transmit from infected to healthy people (Rogers et al., 1987; Wofsy et al., 1986), mother to child through umbilical cord (Gallo et al., 1983; Ziegler et al., 1985).

Very soon, scientists around the world focused on this transmissible retrovirus. Complete sequencing of HIV genome in 1985 (Ratner et al., 1985) led scientists to know more insights of HIV including its origin, genes/proteins and life cycle (Wain-Hobson et al., 1985).

## 1.3 Origin and evolution of HIV

Exploration of the retrovirus led researchers to identify similarities between HIV and a retrovirus in African non-human primates that were then known as Simian Immunodeficiency Virus (SIV) (Gao et al., 1994). About 40 different primates, in Africa, are infected with SIV with some harboring multiple strain of SIVs (Apetrei et al., 2004). Phylogenetic analysis of SIV from African non-human primates and HIV in human provided remarkable understanding of viral transmission as zoonotic (Bailes et al., 2002) and evolution of the virus in human after transmission (Gao et al., 1999).

HIV is divided into two groups – HIV-1 and HIV-2 (Gao et al., 1999). Each group resulted from an independent cross species transmission from different African non-human primates to human (Sharp and Hahn, 2010). HIV-2 is discovered in 1986. This group is transmitted from sooty mangabey monkeys (Cercocebus atys) (Hirsch et al., 1989) and its prevalence is also high in the geographical location of these monkeys in West Africa (Santiago et al., 2005). Sooty mangabey monkeys are naturally infected by a strain of SIV (Hirsch et al., 1989). The phylogenetic analysis of HIV-2 strains shows that they closely group with the SIVsmm strain (Hirsch et al., 1989) that are non-pathogentic to its host monkeys (Gao et al., 1992). SIVsmm evolved in its host to produce multiple strains and subsequent multiple zoonotic transmissions from sooty mangabey monkeys to human (Hahn et al., 2000) gave rise to different subtypes of HIV-2. Although HIV-2 subtypes A to G are identified in human, it is assumed that more subtypes were introduced into human (Gürtler, 2004) but are lost for low adaptation fitness (Damond et al., 2004).

HIV-1 is the result of at least three cross species transmission events from chimpanzees (Pan troglodytes troglodytes (Ptt)) to human (Huet et al., 1990; Peeters et al., 1989). Phylogenetic analysis of HIV-1 sequences has shown that three independent cross transmission of the virus in to the human population, each giving rises to three sub groups: group M (Major), group O (Outlier) and group N (Non M or Non O) (Hahn et al., 2000; Keele et al., 2006). Recently a new HIV-1 strain, classified as group P, distinct from the previous three groups, has been discovered in a patient in Cameroon (Plantier et al., 2009).

Group M is the most prevalent and accounts for 98% of all infections (reviewed in (Sharp and Hahn, 2010)). Its epicenter is thought to be Kinshasha of present day Democratic Republic of Congo (Sharp and Hahn; Sharp and Hahn, 2010). Site stripping for clock detection method shows that group M and its closest simian relative branched out from their common ancestor in 17th century (Salemi et al., 2001) whereas molecular clock analysis of group M shows that the origin of it’s most recent common ancestor dates back to late 1920s (Korber et al., 2000). By 1960, long before human discovered its presence, HIV-1 group M had already diversified substantially (Worobey et al., 2008).

Group O and group N are rare and geographically confined to West African regions such as Cameroon and neighboring countries (Gao et al., 1999). It is still not understood about the non-pandemic characteristics of group O and N HIV-1 virus after the first zoonotic transmission (Ariën et al., 2005). It has been suggested that reduced replication capacity and transmission fitness are keys to their low prevalence(Ariën et al., 2005). Group O strain has at least 50% genetic identity with group M (Gürtler, 2004; VANDEN HAESEVELDE et al., 1996) and the molecular clock model of this group also showed that its origin dates back to 1920s (Lemey et al., 2004a). The date of origin of the introduction of group N into human population has been estimated to be in 1960s (Simon et al., 1998a). Phylogenetic analysis using genetic sequence under evolutionary pressure shows its close grouping with SIV from Chimpanzee (Corbet et al., 2000; Gao et al., 1999). This indicates that group N might be a recombinant strain of SIV and HIV-1 group (Simon et al., 1998a).

Group P is transmitted from gorilla as it is closely related to its SIV (Plantier et al., 2009). A study of HIV infected people in Cameroon shows its low prevalence of 0.06% (Vallari et al., 2011). Although HIV group P is discovery only in Cameroon and confined there, it can still be pandemic as it can adapt in human (Vallari et al., 2011).

### 1.4 HIV-1 Diversity

**1.4.1 HIV-1 subtypes**

HIV-1 group M is highly diversified and it is classified into nine subtypes: A, B, C, D, F, G, H, J and K (Figure 1.3) (Robertson et al., 2000b). The subtype classification is based on the phylogenetic and sequence distance analyses using gene sequence data forming major clades (Robertson et al., 2000b). “At least three epidemiologically unlinked sequences are required for defining a subtype” (Robertson et al., 2000a).

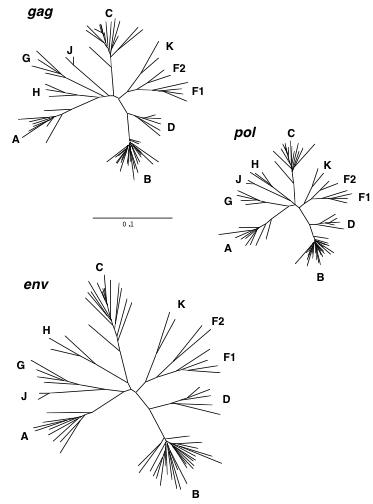


Figure 1.3: Phylogenetic tree showing HIV-1 group M diversification to subtypes A-D, F-H, J and K, inferred from nucleotide sequence alignments of gag, pol and env genes. Source: Robertson et al 2000 (Robertson et al., 2000a)

The classification of new subtype should also follow the same rule as “roughly equidistant from all previously characterized subtypes in all regions of the genome with a distinct pre-subtype branch similar to those of other subtypes” (Robertson et al., 2000b). The predominating group M subtypes are A, B, C and D (reviewed in (McCutchan, 2006)). The range of amino acid variation at gene level within a subtype and between subtypes differs from 15%– 20% and 25% - 35% respectively (Korber et al., 2001). Geographical locations of group M subtypes epidemic are show in Figure 1.4. The analysis from HIV samples collected from 70 countries in 2004 shows that “subtype C accounts for 50% of all infections worldwide” while subtypes A, B, G and D are found in decreasing order 12%, 10%, 6% and 3% respectively (Hemelaar et al., 2006). Subtypes F, H, J and K infections are rare and collectively account for only 0.94% infections (Hemelaar et al., 2006).

Subtypes can be further classified to sub-subtype based on a distinct sister clade formation (Gao et al., 2001) within a clade with the same rule of phylogenetic and distant analyses but not justifiable to call a subtype due to low genetic distance (Robertson et al., 2000b). Only subtypes A and F exhibit distinct sister clades (Gao et al., 2001). Subtype A has sub-subtypes A1 and A2 (A3 and A4 are mentioned by Taylor et al 2008(Taylor et al., 2008)); Subtype F has sub-subtypes F1 and F2 (Figure 1.3); sub-subtype F3 mentioned by Taylor et al 2008 (Taylor et al., 2008).

The lower diversity observed in Group N (Ayouba et al., 2000), O (Lemey et al., 2004b) and P (Vallari et al., 2011) correlate to low prevalence and geographical confinement to Western African countries such as Cameroon. It is still unclear if group O can be sub divided into subtypes (Roques et al., 2002). Group N as well does not show distinct sub clade within itself (Figure 1.5).

**1.4.2 HIV-1 recombination**

Initially, HIV-1 group M subtypes E (Artenstein et al., 1995; Tovanabutra et al., 2002; Wasi et al., 1995) and I (Figure 1.5) were also classified (reviewed in (McCutchan, 2006)). With availability of complete HIV genome sequence and phylogenetic analysis from it, the subtypes E and I were reclassified as circular

### 

Figure 1.4: HIV diversity around the globe, its level of prevalence in the area and number of genome sequenced. Source: McCutchan 2006

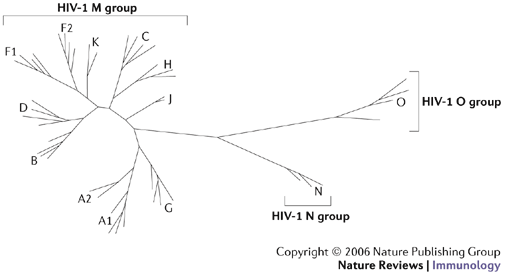


Figure 1.5: Neighboring joining phylogenetic tree showing HIV-1 group M, N and O. Group M shows distinct nine subtypes A-D, F-H, J, K while no specific subtype is observed in group N and O. Source: Letvin 2006 (Letvin, 2006)

recombinant forms CRF01\_AE (recombinant form of subtype A and E) and CRF04\_cpx (recombinant form of more than two subtypes, designated by “cpx”) respectively (reviewed in (McCutchan, 2006)). The same criterion of epidemiological unlinked isolates from three or more people applies for classification as a circular recombinant form (CRF) (Robertson et al., 2000a). A recombinant isolate discovered in single patient is termed as Unique Recombinant Form (URF) (reviewed in (McCutchan, 2006)). There are 55 CRFs listed in Los Alamos National Laboratory database for HIV sequences (<http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html>) as on July 22, 2013. The recombination breakpoints to shuffle HIV genome (Archer et al., 2008; McCutchan et al., 2002; Salminen et al., 1995; Zhang et al., 2010) from different strains of the virus are listed in <http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/breakpoints.html>, each appeared in a publication. CRFs account for at least 20% of all the HIV infections (Robertson et al., 2000a; Robertson et al., 1995; Sharp et al., 1995). CRF02\_AG is the most prevalent circulating recombinant form infecting over 9 million people on the whole (McCutchan, 2000) and geographically epidemic in the West African region (Figure 1.5). CRF01\_AE is the dominant circulating recombinant form in South-East Asia (Figure 1.5) (Taylor et al., 2008).

**1.4.3 Intra-patient HIV diversity**

HIV infection initiates mostly with a single virion (Fischer et al.; Keele et al., 2008). Evidences of multiple HIV variants transmission are also recorded (Long et al., 2000; Ping et al., 2000). Generally, HIV is genetically homogenous for a short post infection time (Delwart et al., 2002; Haase). In the long-term post infection period, virus replicates rapidly to produce genetically heterogeneous population (Long et al., 2000). This heterogeneous viral population consisting of a swarm of highly similar but genetically non-identical HIV viruses is called the HIV quasispecies (reviewed in (McCutchan, 2006)). It is observed that the diversity at a gene (*env*) in viral quasispecies can be approximately 30% (Fouchier et al., 1992). Factors that contribute to high genetic heterogeneity in viral quasispecies are high replication rate and turnover (Ho et al., 1995b), viral genome recombination (Fang et al., 2004; Gu et al., 1995; Lole et al., 1999), higher mutation rate by erroneous reverse transcriptase (Bebenek et al., 1989; Roberts et al., 1988), and host immune selection (Borrow et al., 1997; Price et al., 1997). On the whole, HIV replication (discussed in section 1.6) is the overall source of genetic heterogeneity in the viral population (reviewed in (Smyth et al.)).

Intra patientHIV genome recombination is a common event (Fang et al., 2004; Neher and Leitner). Two genomes from different viral strains from same subtype or different subtypes can be co-packed into single virion during replication (Stuhlmann and Berg, 1992). In the subsequent HIV replication, the ability of reverse transcriptase to switch between the two template genomes produces an intra subtype or inter-subtype recombined viruses at the end of the replication cycle (Ben-Artzi et al., 1996; Kostrikis et al., 2002). Genetic recombination allows rapid and efficient shuffling of advantageous genes and removing deleterious mutations, thus, increasing the viral fitness in the host (reviewed in (Smyth et al.)). Successful transmission of the recombinant forms with high viral fitness to three or more people and circulates in human population establishes Circulating Recombinant Forms (CRFs) (reviewed in (Perrin et al., 2003)), as discussed in section 1.4.2.

### 1.5 HIV genome and proteins – structures and functions

HIV has nine genes and produces 15 proteins (Frankel and Young, 1998). The genes are broadly grouped as accessory (vif, vpr, vpu and nef), structural (*pol*, *gag* and *env*) and regulatory (*tat* and rev). The higher number of proteins than genes is a result of post - transcriptional proteolysis of the products of structural genes (Frankel and Young, 1998).

**1.5.1 Accessory genes:**

Vif promotes the viral infectivity to the host, but has no role in viral production (Jager et al.). Vif is produced in the late stage of viral production (Sheehy et al., 2002; von Schwedler et al., 1993) to suppress the innate antiviral immunity of host (Madani and Kabat, 1998; Simon et al., 1998b). It is observed that vif is expressed only when the virus infects immune cells that express cytidine deaminase APOBEC3G (Navarro and Landau, 2004). The reason is that Vif protein prevents APOBEC3 proteins from hyper mutating HIV reverse transcripts as a mechanism of defense (Conticello et al., 2003; Mangeat et al., 2003; Mariani et al., 2003; Marin et al., 2003; Mehle et al., 2004; Schafer et al., 2004; Sheehy et al., 2003; Simon et al., 2005; Stopak et al., 2003; Wiegand et al., 2004; Zhang et al., 2003).

Vpr protein is packed in to nascent virions during budding out (Connor et al., 1995). Vpr is essential for viral core to enter and localize in the host cell nucleus after infecting the cell (Cohen et al., 1996). Vpr arrests the cell cycle during the transfer from G2 to M phase (Jowett et al., 1995; Rogel et al., 1995) by preventing the activation of the human p34cdc2/cyclin B complex (He et al., 1995). Vpr is also important for efficient viral replication in monocyte or macrophage cells, but T-cells (Connor et al., 1995).

Vpu is a protein unique to HIV-1 (Cohen et al., 1988) and is 16 kilo Dalton, 81 amino acids long (Strebel et al., 1988). Biological functions of vpu protein include degradation of CD4 in endoplasmic reticulum (Willey et al., 1992), interference of host immune cell MHC class II antigen presentation on the cell surface allowing the virus for host immune escape (Hussain et al., 2008; Nomaguchi et al., 2008) and viral maturation and release from host cell membrane (Klimkait et al., 1990).

Nef has no role in viral infectivity but plays a role during the biogenesis of viral particles (Laguette et al., 2009) and virulence (Lenassi et al.; Simmons et al., 2001). Nef down regulates the production of major histocompatibility complex type I (MHC type I) in the host cell (Blagoveshchenskaya et al., 2002; Greenberg et al., 1998; Lewis et al., 2012). This impairs the function of cytotoxic T lymphocyte cells to clear the infected cells (Adnan et al., 2006; Baur et al., 1994; Collins et al., 1998; Couillin et al., 1994; Sawai et al., 1994). Nef also down regulates CD4 on host cell surface (Garcia and Miller, 1991; Lama et al., 1999) and modulates cellular activation to evade host immune system (Baur et al., 1994; Sawai et al., 1994).

**1.5.2 Structural genes and proteins**

The Gag gene produces a precursor polyprotein (pr55gag) of ~ 500 amino acids long and weighs 55 kilodalton (Briggs et al., 2004). The Gag precursor has all the building blocks to form a fully infectious virion, even in the absence of other viral products (Wang and Barklis, 1993). A proteolytic cleavage of gag precursor yields the structural proteins – matrix, capsid, nucleocapsid and p6 (Figure 1.6) (Wiegers et al., 1998). The cleavage takes place in the nascent virus after budding out from host cell (Göttlinger et al., 1989). The matrix protein is at the N-terminal and p6 at the C-terminal of gag precursor (Borsetti et al., 1998; Wiegers et al., 1998).

All gag proteins play role at “post assembly and post processing stages in viral infectivity” (Wang and Barklis, 1993). In the HIV replication cycle, matrix domain of gag plays role in targeting gag precursor to the plasma membrane of the host cell and the viral assembly at the site (Dawson and Yu, 1998; Gheysen et al., 1989; Wang and Barklis, 1993; Zhang et al., 1998). The highly basic region in matrix mediates electrostatic association with phospholipids in plasma membrane during assembly (Chukkapalli et al.; Zhou and Resh, 1996). The capsid is a curved and closed shell consisting 250 hexamers and 12 pentamers (reviewed in (Briggs and Krausslich)). Capsid packs viral proteins, nucleocapsid and viral genome during assembly to pass on to new HIV particles (Ganser-Pornillos et al., 2004). Nucleocapsid plays role in efficient viral assembly by making pr55–pr55 inter-protein contacts (Dawson and Yu, 1998; Zhang et al., 1998) and localization of viral proteins (Larsen et al., 2008). P6 protein plays role in detaching and releasing the newly formed HIV particles (Demirov et al., 2002).

The *pol* polyprotein is produced by translational frame shift (Karacostas et al., 1993) (Figure 1.5), such that, as much as 241 nucleotides at 5’ region of the gene overlaps with 3’ region of *gag* gene (Jacks et al., 1988; Ratner et al., 1985). The proteolytic cleavage of *pol* precursor produces essential viral replication enzymes – protease (PR), reverse transcriptase (RT) and integrase (IN). The protease enzyme cleaves the *gag* and *pol* polyprotein to form the viral structure proteins and functional enzymes respectively (Darke et al., 1988b; ERICKSON-VIITANEN et al., 1989; Nutt et al., 1988). The reverse transcriptase enzyme reverse transcribes the viral RNA to produce a cDNA molecule after infecting host cell (Jacobo-Molina and Arnold, 1991; Sarafianos et al., 2009). The RNase H domain in RT degrades the viral RNA molecule following cDNA production (Davies et al., 1991). The integrase enzyme removes two bases from 3’ DNA molecule and functions strand transfer during the process of integrating the proviral DNA into the host genome (Pruss et al., 1994).

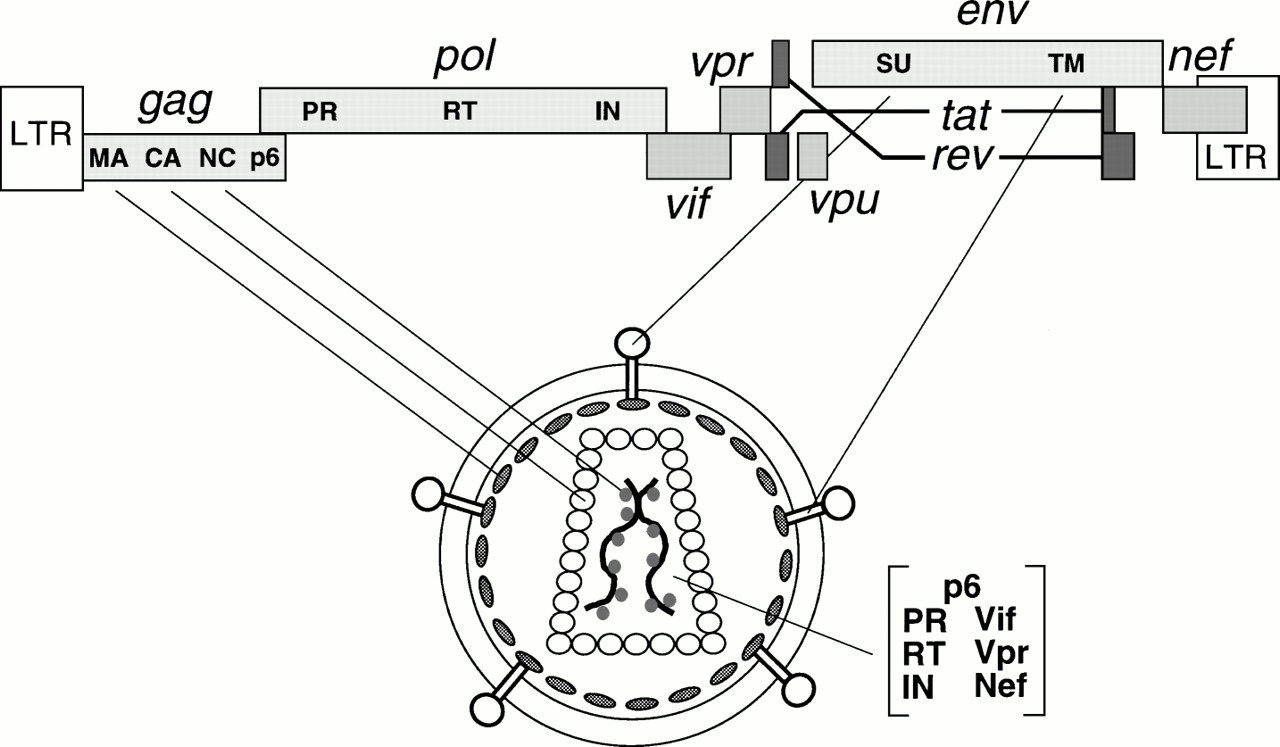


Figure 1.6: HIV genes and proteins positions in the viral genome and their viral parts. Source: Frankel and Young 1998 (Frankel and Young, 1998)

The *env* gene produces a precursor glycopolyprotein (gp160) that is processed at post-translational by human convertase enzymes - PC1 and furin to produce glycoprotein 120 (gp120, HIV-1 SU) and glycoprotein 41 (gp41, HIV-1 TM) (Decroly et al., 1994). Gp120 is a non-covalent complex of external protein and gp41 is a trans-membrane protein; both play vital role for initial steps in viral infection (Chan et al., 1997). Three gp120 molecules bound with three gp41 molecules to form envelop spikes (Pancera et al.). They are organized to form trimeric complexes on the surface of HIV and mediate HIV entry into the host cell (Liu et al., 2008). The exposed external complex gp120 binds to the CD4 receptor on the host immune cell (Rizzuto et al., 1998). This triggers a conformational shift of trimeric complex that enables a conserved gp120 region binding to a chemokine receptor, either CCR5 or CXCR4, to facilitate fusion of the viral and host membranes (Huang et al., 2007; Rizzuto et al., 1998; Wu et al., 1996; Wu et al., 1997). The gp120-CD4 complex also triggers conformational change in gp41 trans-membrane protein from native non-fusion state to fusion state (Chan et al., 1997; Kliger et al., 1997). Gp41 plays role in the viral fusion and release of viral contents in to the host cell (Furuta et al., 1998; Melikyan, 2008),. The gp41 consists of heptad repeats - HR1 and HR2 that play role in fusion process (Furuta et al., 1998; Tan et al., 1997). HR1 is a bundle of three helical motifs and HR2 is trimeric coiled coil structure (Dwyer et al., 2003). During fusion process, HR2 makes numerous contacts with HR1 to form stable six helical bundles (Melikyan et al., 2000).

**1.5.3 Regulator Genes:**

*Tat* is a trans-activating factor localized in the nucleus for HIV gene expression (Rosen and Pavlakis, 1990a; Roy et al., 1990). The HIV proviral genome integrated in to the host genome is regulated by cellular as well as the viral transcription regulatory factors (Cullen, 1991; Gaynor, 1992). Tat is the primary transcriptional regulatory factor (Marcello et al., 2001). An example of Tat action is the control of RNA polymerase II elongation during transcription, which otherwise disengages from the template DNA strand, terminating the transcription prematurely (Bourgeois et al., 2002; Chou et al.; He and Zhou). Sodroski *et al*. (1985) first explained the function of Tat (Sodroski et al., 1985).

Rev is a 19 kilo Dalton phosphoprotein (Malim et al., 1989a) trans-activating factor for HIV gene expression (Rosen and Pavlakis, 1990a). Like Tat, it is also mainly localized in the nucleus of host cell (Rosen and Pavlakis, 1990b), but cycles rapidly between the nucleus and cytoplasm as it promotes nuclear export of the transcriptional products (Fischer et al., 1995; Fischer et al., 1994; Henderson and Percipalle, 1997; Malim et al., 1989b). Rev binds at the Rev Responsive Element (RRE), which is an RNA element encoded within the *env* region of the virus (Daly et al., 1989; Malim and Cullen, 1991).

### 1.6 HIV replication

There are major 11 events in HIV replication cycle (as numbered in Figure 1.7). The initial step of viral entry (Figure 1.7 step 1) in to a host immune cell includes HIV gp120 molecule binding to CD+ receptor followed by binding to a co receptor on the surface of the host cell and fusion of the viral and host cell membranes (see section 1.5.2 for more detail).

Following the fusion, the viral core enters in to the cytoplasm of host cell. The reverse trancriptase enzyme reverse transcribes the RNA molecule to cDNA (Figure 1.7 step 2) in the intact capsid of the viral core (McDonald et al., 2002). The reverse transcriptase enzyme is not perfect at copying mRNA molecule to cDNA and has no error correct capability (Bebenek et al., 1989; Bebenek et al., 1993; Preston et al., 1988; Roberts et al., 1988). The enzyme generates mutations in its genome at the rate of 10-5 per base per replication cycle (Mansky and Temin, 1995). This is a crucial step as it contributes to generation of variations in the viral quasispecies (reviewed in (Goodenow et al., 1989; Nowak et al., 1990)).

Subsequently the capsid is dissembled, termed as uncoating (Figure 1.7 step 3) (McDonald et al., 2002; Shah et al., 2013), releasing ribonucleoprotein complex in to the cytosol (Dismuke and Aiken, 2006). The process can take an hour or less since time of post infection (Hulme et al., 2011). The capsid and nucleocapsid proteins dissociate from cDNA but the reverse transcription complex remains intact along with viral matrix, integrase, vpr and human protein high mobility group I (HMG I (Y)) forming pre-integration complex (PIC) (Bukrinsky et al., 1993; Farnet and Haseltine, 1991; Miller et al., 1997). The PIC protects cDNA from endonuclease degradation (Miller et al., 1997). In an ATP dependent process (Bukrinsky et al., 1992), PIC is transported on host microtubules towards the nuclear membrane (McDonald et al., 2002). Some studies initially suggested that formation of central DNA flap (Zennou et al., 2000), vpr, and matrix mediate nuclear import (Figure 1.7 step 4) of PIC (Popov et al., 1998), however, the no any clear evidence of these factors have been described as such for nuclear import (Riviere et al.). Integrase assists in nuclear import in association with nuclear import machinery like importin (Fassati et al., 2003) and transportin-SR2 (Christ et al., 2008). It is now established that central polypurine tract-central termination sequence (cPPT-CTS) plays role in kinetics of nuclear import (Riviere et al.).

Post-nuclear entry, integrase processes the viral DNA for integration (LaFemina et al., 1992) (Figure 1.7 step 5) into host genome as explained in section 1.5.2. Host transcriptional co-factor LEDGF/p75 and HIV integrase interact to tether to the host chromosome during the integration process (Emiliani et al., 2005).

The integrated HIV provirus hijacks the host cell transcriptional machinery for viral genes to transcribe (Figure 1.7 step 6) (Davey et al., 2011). HIV protein Tat promotes the transcription of the viral DNA (Ott et al., 2011; Razooky and Weinberger, 2011). The viral transcriptome encodes structural proteins, accessory proteins and viral enzymes necessary for a complete functional HIV (reviewed in (Karn and Stoltzfus)).

The viral Rev protein facilitates exporting (Figure 1.7 step 7) of the unprocessed viral transcriptome to cytoplasm for translation (Malim et al., 1989b). HIV has no translation system of its own; the host translational machinery is exploited for translation (Figure 1.7 step 8) of the viral transcriptome to its proteome (Cherry et al., 2005; Thompson and Sarnow, 2000).

Upon translation of all viral proteins, viral *gag* initiates virion assembly (Figure 1.7 step 9) at the cell membrane (Dong et al., 2005; Nermut et al., 1998; Saad et al., 2006). The complete assembled virion particles bud out and are released (Figure 1.7 step 10) from the plasma membrane by the host ESCRT machinery involving Tsg101 and ALIX regulatory proteins (Fujii et al., 2007; Garrus et al., 2001; Saksena et al., 2007). The maturation (Figure 1.7 step 11) of the nascent HIV virions begins concomitantly with budding out (Klimkait et al., 1990; Schubert et al., 2000). *Gag* and *Pol* polyproteins are proteolytically cleaved by protease enzyme in the maturation step (Darke et al., 1988a; Pettit et al., 2005).

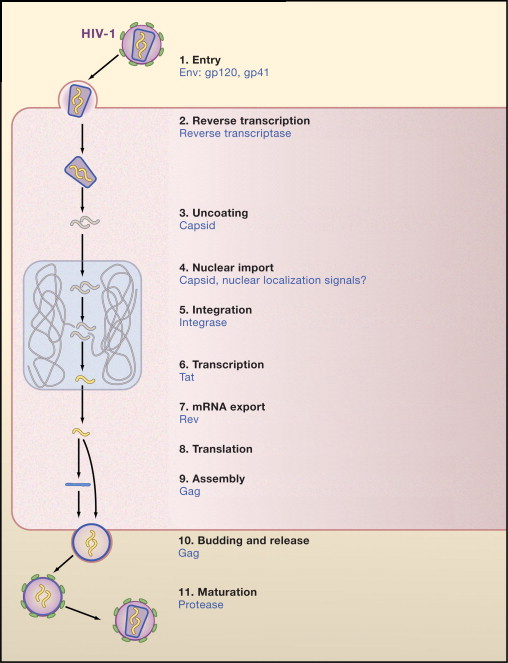


Figure 1.7: The HIV replication cycle showing major stages. Viral proteins that play role in each event are colored blue. (Source: modified from Ho and Bieniasz 2008) (Ho and Bieniasz, 2008)

Each HIV replication cycle releases new infectious virions in the order of 109 per day (Ho et al., 1995a). The number of new infecting HIV determines the replication rate of the virus (Tersmette et al., 1989). A long post infection period shows higher turnover rate associated with CD4+ cell depletion and viral population expansion (Ho et al., 1995a).

### 1.7 Anti Retro Viral Drugs

Currently different inhibitory drugs, each interfering a viral protein normal function at different stages of its replication cycle has been developed clinical practice. Depending upon the targets, drugs used in clinical HIV treatment are grouped as reverse transcriptase inhibitors, protease inhibitors and integrase inhibitors (reviewed in (De Clercq, 2002, 2005a, 2009)).

**1.7.1 Reverse Transcriptase Inhibitors**

1. **Nucleoside Reverse Transcriptase Inhibitors (NRTIs)**

NRTIs are analogs of normal nucleotides but without 3’ hydroxyl group (reviewed in (Sarafianos et al., 2004)). The drug is taken in unphosphorylated form, which cytokinases phosphorylates to form 5’ triophosphates (reviewed in (De Clercq, 2002; Ilina and Parniak, 2008)). It leads to incomplete termination of cDNA synthesis (Sluis-Cremer et al., 2000). The list of NRTI drugs shown in Table 1.1.

1. **Non-nucleoside reverse transcriptase inhibitors (NNRTIs)**

NNRTIs specifically bind at an allosteric site 10 [Å](http://en.wikipedia.org/wiki/%C3%85) from polymerase active site of the HIV-1 reverse transcriptase (Himmel et al., 2006; Sarafianos et al., 2009), close to substrate-binding site. The binding induces conformational

| Drugs | Mechanism of Action | Mechanisms of Resistance |
| --- | --- | --- |
| Fusion and entry inhibitors | | |
| Enfuvirtide (T-20) | 36 amino acid peptide derived from the HR2 domain of glycoprotein 41  Interferes with glycoprotein 41 dependent membrane fusion | Mutations affect HR1, a domain of glycoprotein 41 whose interaction with HR2 promotes membrane fusion |
| Maraviroc | Binds to CCR5 co receptor | Development of CXCR4 using HIV HIV-1 exploits CCR5 conformational heterogeneity |
| Nucleoside reverse transcriptase inhibitors (NRTls) | | |
| Zidovudine (AZT) | Analogues of normal nucleosides  Active as triphosphate derivatives  Incorporated into nascent viral DNA  Prematurely terminate HIV DNA synthesis | Thymidine analogue mutations promote ATP-mediated and pyrophosphate-mediated excision of the incorporated terminator |
| Didanosine (ddl) |
| Zalcitabine (ddC) |
| Stavudine (d4T) |
| Lamivudine (3TC) |
| Abacavir (ABC) |
| Tenofovir disoproxil (TVD) |
| Emtricitabine (FTC) |
| Nucleotide reverse transcriptase inhibitors (NtRTls) | | |
| Tenofovir | Same as nucleoside analogues | K65R impairs incorporation of tenofovir into DNA  Thymidine analogue mutations often associated with cross-resistance to tenofovir |
| Non-nucleoside reverse transcriptase inhibitors (NNRTls) | | |
| Nevirapine (NVP) | Bind a hydrophobic pocket of HIV type 1 reverse transcriptase  Block polymerization of viral DNA  Inactive against HIV type 2 | Mutations reduce affinity of the inhibitors for the enzyme  Single mutations generally sufficient to induce high level of resistance |
| Delavirdine (DLV) |
| Efavirenz (EFV) |
| Etravirine (ETR) |
| Protease Inhibitors | | |
| Saquinavir (SQV) | Structure derived from natural peptidic substrates of the HIV type 1 protease  Bind the active site of the protease | Mutations reduce affinity of the inhibitors for the enzyme  High-level resistance requires a accumulation of mutations |
| Ritonavir (RTV) |
| Indinavir (IDV) |
| Nelflnavir (NFV) |
| Amprenavir (APV) |
| Lopinavir + Ritonavir (LPV/r) |
| Fosamprenavir (FPV) |
| Tipranavir (TPV) |
| Darunavir (DRV) |
| Atazanavir (ATV) |
|  |  |  |
| Integrase Inhibitors | | |
| Raltegravir | Binds selectively to the enzyme complexes that results in inhibiting strand transfer of viral and host DNA | Mutations at conserved carboxylate residues (Asp64 and Asp116) |

Table 1.1: Antiretroviral drugs used in the treatment of HIV infection, their mechanisms of action and mechanisms of resistance.

Source: Adapted from François Clavel and Allan J. Hance 2004 (Clavel and Hance, 2004) and (Ammaranond and Sanguansittianan, 2012; Colin et al., 2013)

change in the enzyme, which distorts the catalytic aspartate triad of active site and inhibits the function of the enzyme (Balzarini, 2004; Esnouf et al., 1995). NNRTIs drug are administered along with NRTI drugs in ART at non-toxic concentration to human cells (De Clercq, 1996). NNRTIs drugs from first generation have inherent disadvantageous characteristics like cross-resistance and low genetic barrier i.e. require only one mutation to select (Adams et al., 2010). The list of NNRTI drugs is shown in Table 1.1.

**1.7.2 Protease Inhibitors (PI)**

Generally, protease inhibitors are added to second line ART therapy following failure of first line therapy (Pujades-Rodriguez et al., 2008). Protease Inhibitors interfere with cleavage of *gag*-*pol* polypeptide (Seelmeier et al., 1988) as competitive peptidomimetic inhibitors. The hydroxyethylene core in the inhibitors prohibits the cleavage action of the HIV protease enzyme (Vacca et al., 1994; Vacca et al., 1991). However, as an adverse side effect, patients consuming the inhibitors have developed lipodystrophy and hyperlipidemia (Carr et al., 2000; Carr et al., 1998a; Carr et al., 1998b, c; Liang et al., 2001; Miller et al., 2000; Tsiodras et al., 2000). The list of protease inhibitors is show in Table 1.1.

**1.7.3 Integrase Inhibitors**

Although the anti retroviral therapy could suppress the viral replication, they are expensive, can lead to multiple drug resistance and, thus, require high levels of adherence (Cohen, 2002). These necessitate alternative drugs that target the third enzyme, the integrase. The feasibility and efficacy of integrase inhibitors have been tested in Rhesus Macaques (Hazuda et al., 2004). Most of the integrase inhibitors target strand transfer function of the enzyme (Bera et al., 2011; Espeseth et al., 2000; Hazuda et al., 2000; McColl and Chen, 2010; Pannecouque et al., 2002). X-ray structure of the integrase enzyme has revealed the active site model of the enzyme complexes with the DNA (Chen et al., 2008). The only integrase inhibitor showing promising anti retroviral effect through animal model is Raltegravir and is currently in clinical trials (reviewed in (Ammaranond and Sanguansittianan, 2012)).

**1.7.4 Cell entry inhibitors**

The cell entry inhibitors interfere viral binding or fusion of HIV to a host cell. These drugs are added to salvage drug regimen for patients showing multiple high anti viral drug resistance. Following are the two classes of cell entry inhibitors.

1. **CCR5 co receptor antagonist:**

Maraviroc is the only CCR5 antagonist in clinical use (De Clercq, 2005a, b; Fätkenheuer et al., 2005; Rosario et al., 2005; Rosario et al., 2006; Wheeler et al., 2007). It is the only anti retroviral drug that does not target any viral enzyme or protein molecule, but host cell receptor CCR5 (Westby and van der Ryst, 2005). Binding of CCR5 antagonist renders unavailability of the co receptor for HIV gp120 binding, thereby preventing the viral entry in to the cell (Fätkenheuer et al., 2005). However, it is important for HIV tropism test before administrating this drug, as the drug is ineffective against CXCR4 co receptor using viruses (Raymond et al., 2010).

1. **Fusion Inhibitors**

Fusion inhibitor design is based on targeting the heptad regions HR1 or HR2 of gp41 disabling the virus to make fusion pore (reviewed in (Baldwin et al., 2003)). Enfuvirtide (Duffalo and James, 2003; Poveda et al., 2005) is a synthetic peptide, approved for clinical use in 2003 (Robertson, 2003), which can bind to gp41 HR1 region (Wild et al., 1993). Emergence of Enfuvirtide resistant viral strains lead to discontinue its use for clinical use in 2004 (Briz et al., 2006; Melby et al., 2007). Sifuvirtide is another HIV fusion inhibitor peptide under research (Wang et al., 2009).

### 1.8 HIV Treatment

The World Health Organization (WHO) has produced year 2013 clinical guideline (<http://www.who.int/hiv/pub/guidelines/arv2013/art/en/index.html>) for HIV treatment. From the treatment point of view, the guideline has grouped HIV infected people as Adult, pregnant and breast feeding women, children less than 3 years old, children 3 or more years old and HIV people co-infected with other diseases. Regardless of the grouping, the guideline recommends treatment initiation to all confirmed HIV infected people with CD+ cell count less or equal to 350 per mm3. However, in resource poor settings, where laboratory tests for CD+ count may not be feasible, WHO clinical stages should be used as a guide for treatment initiation (Weinberg and Kovarik, 2010).

According to the WHO consolidated anti retroviral guideline recommendation for HIV treatment, a combination of three reverse transcriptase inhibitors (2 NRTIs + 1 NNRTI) are used as first line therapy. Addition of protease inhibitor is recommended for children below 3 years old. On virologic failure to first line, second line drug regimen containing 2 NRTIs and a ritonavir boosted protease inhibitor is recommended. Following failure to second line therapy, a new drug class – integrase inhibitor is introduced in third line therapy along with a reverse transcriptase and a protease inhibitor.

Patients under ART therapy are monitored for effectiveness of the treatment at a defined interval. Laboratory test for viral load should be done for monitoring the treatment response. The viral load of greater than 1000 viral RNA copies/ml blood sample indicates virologic failure to the treatment and the patient is recommended to switch to new drug regimen (Figure 1.8). If viral load test is not feasible routinely, CD4+ count and clinical monitoring need to be used.

### 1.8 HIV Drug Resistance

Anti retroviral drug therapy has been successful at controlling HIV replication to low viral load within a host (Rosenberg et al., 2000). The therapy has been compromised and threatened for long term sustainability by the emergence of drug resistant HIV viral variants (reviewed in (Clavel and Hance, 2004)). HIV mutates spontaneously in every replication cycle (di Marzo Veronese et al., 1993; Dumonceaux et al., 1998); the error prone reverse transcriptase introduces mutations randomly at the rate of 10-5 per base per replication cycle (Mansky, 1996; Mansky and Temin, 1995). The accumulation of mutations and viral recombination contribute the most for the emergence of drug resistant variants (reviewed in (Menéndez-Arias, 2002; Smyth et al., 2012; Tebit et al., 2007)). In the presence of anti retro viral drugs the HIV variants in the quasispecies get selected (Leslie et al., 2004; Nijhuis et al., 1999). The selected variants proliferate to high number in subsequent replication cycles (Hosseinipour et al., 2013). Transmission of resistant virus also confers drug resistance in drug naïve patients (Boden and Markowitz, 1998; Wensing et al., 2005; Yerly et al., 1999). Although drug resistant mutations escape drug effect, they confer decrease in viral fitness and replication rate at certain level (Lucas, 2005; Mammano et al., 2000). Other mutations counteract the effect of drug resistant mutations to boost up viral fitness and replication rate (Nijhuis et al., 1999).

ART drugs are designed based on three dimension structures of the viral proteins to bind at a specific active region of the proteins and interfere with its normal function (reviewed in (Tantillo et al., 1994)). In general, mutations at the active site amino acid

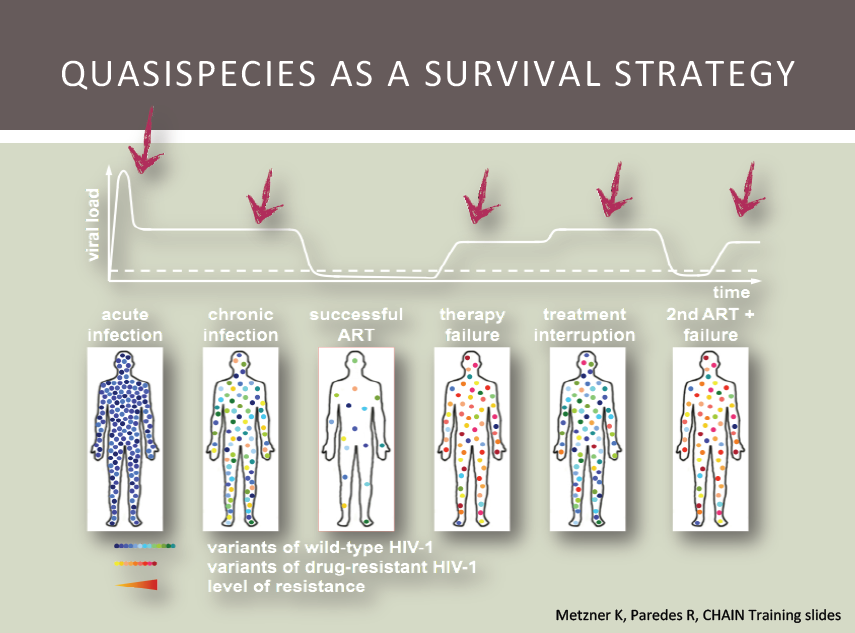


Figure 1.8: Dynamics of HIV viral load within an infected patient. Viral load peaks at acute infection followed by significant drop due to less CD4+ count. Introduction of highly active cocktail of anti retroviral drugs arrest viral replication cycle decreases the viral load to safe level. Spontaneous mutations give rise to resistant viruses that replicate with high turnover increasing the viral load and the therapy fails. Interruption of drugs contributes to higher viral load. Introduction of second line anti retroviral drugs suppress the viral load back to safe level again and the cycle of viral load repeats. Source: Roger Paredes and personal communication

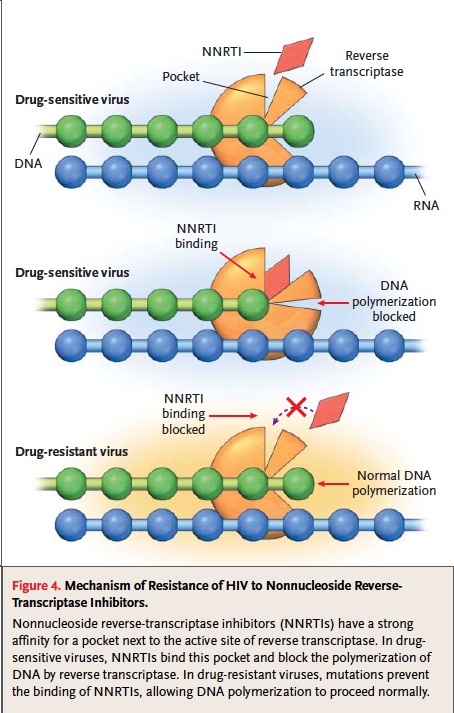


Figure 1.9: Development of HIV drug resistance again NNRTI. Accumulation of Drug resistant mutations changes three-dimensional confirmation of a viral protein, disabling drug binding and carrying out its normal function. Source: (Clavel and Hance, 2004)

residues of HIV proteins change their three-dimensional shape, inactivating the drug functions (Figure 1.9) (Fikkert et al., 2003; Hsiou et al., 2001; Miller and Larder, 2000; Tantillo et al., 1994). For more details on mechanisms of HIV resistance, see Table 1.1.

The mutations attributing resistance to drugs have been well documented with associated drug and resistance level (Shafer, 2006). The viral quasispecies sequence information can be used to align against the reference sequence and infer the presence of any drug resistant mutation (Kuiken et al., 2003).

### 1.9 HIV genotyping for Resistance Test

### Conventional Population Based HIV Drug Resistance Genotyping

The pattern of drug resistant mutations in HIV has strong relation with virologic failure (Condra et al., 1995; Larder et al., 1989; Larder and Kemp, 1989; Lorenzi et al., 1999; Molla et al., 1996; van Leeuwen et al., 1995; Zolopa et al., 1999). Given the database management of all the drug resistant mutations, the application of HIV genotype information on therapy optimization has shown virologic benefits in clinical settings (Baxter et al., 2000; Durant et al., 1999; Van Vaerenbergh, 2001). Conventional HIV genotyping involves Sanger dideoxy termination sequencing method (Ewing et al., 1998; Metzker, 2005; Sanger et al., 1977). The method produces a chromatogram that represents an average genotype sequence of the viral population in the host (Struck et al., 2012).

Conventional population based sequencing method has limited sensitivity, as it has been able to detect only the viral variants of frequency greater or equal to 20%; low prevalent variants go undetected (Ji et al., 2011; Johnson and Geretti). Thus, sequencing technology underestimates the total variants in the viral population (Palmer et al., 2005). Undetected low frequency HIV variants are clinically significant for designing salvage therapy, as they are usually drug resistant (Paredes et al., 2010; Rowley et al.; Simen et al., 2007; Simen et al., 2009). Those minor variants rebound in the presence of drugs, leading to ultimate virologic failure (Paredes et al., 2010; Rowley et al.). This necessitates improved and highly sensitive sequencing technologies able to detect minor HIV variants in the viral quasispecies (reviewed in (Metzker)) .

### 1.10 Next Generation Sequencing technologies

The newer sequencing technologies developed as an alternative to automated Sanger sequencing is referred as Next Generation Sequencing (NGS) technologies. Commercially available NGS technologies in the market are Roche 454, Illumina, Applied Biosystems SOLiD technology and Ion torrent (reviewed in (Metzker)). All the NGS systems has PCR step for DNA amplification but differ in other features (Shokralla et al.); each one outperforms the others at a particular application (reviewed in (Metzker)).

Roche 454 and Illumina implement ‘Sequencing by synthesis’ (SBS) technique for DNA sequencing (Margulies et al., 2005b). Both systems detect fluorescent light emitted from each nucleotide incorporation event during DNA synthesis (Margulies et al., 2005b) (reviewed in (Shendure and Ji, 2008)). The intensity of fluorescent light emitted per base addition is too small for detection; this requires amplification of the template DNA to million copies. Simultaneous addition of same nucleotide in million copies produces higher intensity of detectable fluorescent light (reviewed in (Metzker; Shendure and Ji, 2008)). A defined order of free nucleotide molecules are flown in the reaction plate, allow nucleotide to incorporate, detect fluorescent light and wash off any unincorporated nucleotides for next cycle. Roche 454 and Illumina differ only at the sequencing step. In Roche 454, polymerase continues nucleotide addition reactions until the base flowing in the reaction plate is complementary to the template sequence. The intensity of fluorescent light emission is detected and the number of bases subsequently added, as a homopolymer run, is calculated based on the light intensity in a particular reaction cycle (Margulies et al., 2005b). Illumina, on the other hand, stops the reaction after single nucleotide addition (Bentley et al., 2008). The light emission is detected that determines the base added in that particular cycle of base flow ((Bentley et al., 2008), reviewed in (Shendure and Ji, 2008)). These different sequencing techniques have high impact on sequence read length. Roche 454 produces low number of sequences but longer read length (Table 1.1). Comparatively, Illumina produces shorter reads in high quantity (Table 1.1).

Applied Biosystems SOLiD implements ‘Sequencing by ligation’ technique for DNA sequencing; thus, bypassing any DNA polymerase related sequencing error (Pandey et al., 2008). The template DNA is amplified in similar way to previous NGS technologies. During sequencing, a universal primer and a library of pre-designed 1,2-probes of 8 nucleotides (or dibase probe) along with DNA ligase enzyme is added. The probes hybridize to the complementary template sequence and the fluorescence of the probe is read. The probe hybridization is repeated for seven cycles extending read length to 35 bases. In the next cycle, a new universal primer is hybridized at offset position of one base (n-1) to the previous primer position followed by ligation sequencing process. The primer resetting cycle is repeated five times providing dual measurement of each base and the final sequence is decoded from color code information using 4 by 4 color code (reviewed in (Mardis, 2008)).

| Platform | Read length (bp) | Reads/run | Sequencing output/run | Run time | Advantages | Primary applications |
| --- | --- | --- | --- | --- | --- | --- |
| Roche 454 GS FLX | 400–500 | 1 × 106 | ≤500 Mb | 10 h | Longest read lengths among 2nd generation, high throughput compared to 1st generation sequencing | 1\*, 2, 3\*, 4, 7, 8\* |
| Roche 454 GS FLX+ | 600–800 | 1 × 106 | ≤700 Mb | 23 h |
| Roche 454 GS Junior | 400–450 | 1 × 105 | ∼35 Mb | 10 h |
| Illumina HiSeq 2000 | 100–200 | 6 × 109 | ≤540–600 Gb | 11 d | Very high throughput | 1\*, 2, 3\*, 4, 5, 6, 7, 8 |
| Illumina HiSeq 1000 | 100–200 | 3 × 109 | ≤270–300 Gb | 8.5 d |
| Illumina GAIIx | 50–75 | 6.4 × 108 | ≤95 Gb | 7.5–14.5 d |
| Illumina MiSeq | 100–150 | 7 × 106 | ≤1–2 Gb | 19–27 h |
| AB SOLiD 5500 system | 35–75 | 2.4 × 109 | ∼100 Gb | 4 d | Very high throughput; lowest reagent cost needed to reassemble a human genome among the widely accepted 2nd generation platforms, lower error rate | 3\*, 5, 6, 8 |
| AB SOLiD 5500 xl system | 35–75 | 6 × 109 | ∼250 Gb | 7–8 d |
| Ion Torrent -314 chip | 100–200 | 1 × 106 | ≥10 Mb | 3.5 h | Direct measurement of nucleotide incorporation events; DNA synthesis reaction operates under natural conditions (no need for modified DNA bases) | 1, 2, 3, 4, 8 |
| Ion Torrent -316 chip | 100–200 | 6 × 106 | ≥100 Mb | 4.7 h |
| Ion Torrent -318 chip | 100–200 | 11 × 106 | ≥1 Gb | 5.5 h |

Bold indicates applications that are most often used, economical or growing

1 = de novo BACs, plastids, microbial genomes. 2 = transcriptome characterization.

3 = targeted re-sequencing. 4 = de novo plant and animal genomes.

5 = re-sequencing and transcript counting. 6 = mutation detection.

7 = metagenomics. 8 = other

\*Pooling multiple samples with sequence tags (i.e. MIDs or indexes) is required for efficient use of this application

Table 1.1: Comparative analysis of different NGS systems. Source: adapted from Shokralla et al 2012 (Shokralla et al.), Niedringhaus et al 2011 (Niedringhaus et al., 2011) and Glenn 2011 (Glenn, 2011)

Ion torrent technology implements SBS and electronic sensors connected to CMOS integrated circuit with a microprocessor for signal processing (Jakobson et al., 2002; Milgrew et al., 2004). The sequencing step is similar to Roche 454 homopolymer sequencing but the base detection is completely electronic. During DNA sequencing, a base incorporation releases a hydroxyl ion (H+) that shifts pH of the surrounding solution and this correlates directly to the number of nucleotides incorporated in that particular base flow cycle (reviewed in (Niedringhaus et al., 2011).

This change in pH is detected by sensor at the bottom of each well, converted to a voltage and digitalized by semi conductor CMOS integrated circuits (Pennisi, 2010). Signal processing software is used to convert the data for measurement of incorporation in that flow using a physical model (Rothberg et al., 2011).

### 1.11 HIV-1 Drug resistance Genotyping in the era of high throughput sequencing

The next generation sequencing technologies are highly demanded in the current genomic research in comparison to Sanger sequencing method for two reasons: low cost and ultra deep high throughput. The application of NGS has contributed to many research findings in organisms like HIV that exhibit high heterogeneity at genetic sequence. HIV mutations has strong link to viral fitness and survival in the presence of antiretroviral drugs and every mutation known to correlate drug resistance is maintained in a proper database. Therefore, it is highly essential to accurately detect HIV drug resistant mutations while assessing genetically diverse viral quasispecies.

Genotyping of HIV quasispecies are essential to identify the mutations responsible for drug resistance. HIV genotyping with NGS include amplicon-based approach of resequencing HIV genomic section that covers drug resistant mutations. Longer sequencing reads from NGS that cover the amplicon region is more advantageous to short reads for assessing drug resistant (reviewed in (Droege and Hill, 2008)). Roche 454 system with longest read length (Table 1.1) NGS system is, thus, increasingly used for genotyping HIV quasispecies.

Although Roche 454 produces large amount of sequence data with long reads, it is error prone and that increases the artificial diversity in the sample confounding the real variation. The PCR amplification step has embedded errors like DNA recombination (Kanagawa, 2003; Meyerhans et al., 1990; Yang et al., 1996), DNA synthesis error (Hughes and Totten, 2003; Mansky and Temin, 1995) and DNA re-sampling error (Liu et al., 1996). In addition to this, ultra deep sequencing underlying algorithm has inherent base miscall error (Quinlan et al., 2008). The frequency of base call error is six times in homopolymeric region than non-homopolymeric region (Margulies et al., 2005a; Wang et al., 2007). Enhancement in sequencing chemistry and base calling algorithm in Roche 454 Genome Sequencer FLX system improved accuracy of data to >99.5% including errors at homopolymer region (reviewed in (Droege and Hill, 2008)). At the coverage of 10x, the FLX system is able to identify all sequence variations that is known from Sanger sequencing (Bordoni et al., 2008). An accuracy of 99.28% is reported, all errors at homopolymer region but given the coverage ≥ 30x, FLX system correctly identified all sequencing error (Bordoni et al., 2008).

Ultra deep high throughput sequencers associates quality score to every base. The score denotes probability that the corresponding base call is correct (Cock et al., 2010). It is essential to trim out the poor quality scored bases ensure high quality data in the downstream analysis.

Errors generated in early cycles of PCR get amplified in subsequent cycles and that they might be associated with high quality score during ultra deep sequencing. These erroneous bases can escape at quality trimming given their high quality score. A novel technique of tagging every template polynucleotide with a primer ID is introduced to track every ultra deep sequence read that originates from same source template sequence (Jabara et al., 2011). Consensus sequence representing sequences with same Primer ID denotes the original source template and that corrects every known sequencing error (Jabara et al., 2011).

Ultra deep high throughput sequencers produce large volume of data that often is difficult to analyze at human scale. Analysis of data often requires multiple processing steps of specific task like grouping the data with similar information, quality control, mapping the sequence data to a reference sequence to infer any mutations. Reference mapped HIV quasispecies sequence data can be used to query a standard HIV drug resistance database e.g. Stanford database. Bioinformatics computational tools are used for each specific task but still they needs to be run manually. An automated computational pipeline where all the tasks can be run in step-wise fashion is a current necessity.

### 1.12 Thesis Outline

**Chapter 2**: The chapter introduces and describes, in detail, a novel algorithm for quality trimming called QTrim. The tool is designed currently to quality control Roche 454 ultra deep sequence data and optimized for both poor and high quality data. QTrim is compared to other widely used quality trimming tools and the comparative analysis result is also presented. The entire chapter is produced as a paper for publication entitled “**QTrim: A novel tool for the quality trimming of sequence reads generated using the Roche/454 sequencing platform” (Shrestha, RK and Travers, Simon; being reviewed)**.

**Chapter 3**: The chapter introduces Primer ID technique recently developed to correct ultra deep sequencing errors. The chapter also discusses a comparative analysis of sequence data with and without Primer ID approach and the effect on downstream analysis. The Primer ID approach sequence data is obtained from a study conducted to observe HIV virus response on the vaccine. The study was conducted in the University of Cape Town.

**Chapter 4**: The chapter introduces and describes the computational pipeline called Seq2Res that facilitates low cost HIV drug resistance test. The chapter describes a workflow of the pipeline, the requirements of the pipeline, HIV drug resistance output files and plots that summaries overall analysis. The test data is obtained from a study called CIPRA-SA, which is a trial of anti retroviral drug monitoring strategy in resource poor setting. The chapter is included in paper publication.

**Chapter 5**: This chapter discusses on comparison of clinical data and sequence data analyzed using Seq2Res pipeline. Excitingly, the sequence data analyzed using Seq2Res agrees to clinical data.

**Bibliography**

Adams, J, Patel, N, Mankaryous, N, Tadros, M, Miller, CD (2010) Nonnucleoside Reverse Transcriptase Inhibitor Resistance and the Role of the Second-Generation Agents. *The Annals of Pharmacotherapy* **44**: 157-165.

Adnan, S, Balamurugan, A, Trocha, A, Bennett, MS, Ng, HL, Ali, A, Brander, C, Yang, OO (2006) Nef interference with HIV-1–specific CTL antiviral activity is epitope specific. *Blood* **108**: 3414-3419.

Ammaranond, P, Sanguansittianan, S (2012) Mechanism of HIV antiretroviral drugs progress toward drug resistance. *Fundamental & Clinical Pharmacology* **26**: 146–161.

Apetrei, C, Robertson, DL, Marx, PA (2004) The history of SIVS and AIDS: epidemiology, phylogeny and biology of isolates from naturally SIV infected non-human primates (NHP) in Africa. *Frontiers in bioscience: a journal and virtual library* **9**: 225-254.

Archer, J, Pinney, JW, Fan, J, Simon-Loriere, E, Arts, EJ, Negroni, M, Robertson, DL (2008) Identifying the important HIV-1 recombination breakpoints. *PLoS computational biology* **4**: e1000178.

Ariën, KK, Abraha, A, Quiñones-Mateu, ME, Kestens, L, Vanham, G, Arts, EJ (2005) The Replicative Fitness of Primary Human Immunodeficiency Virus Type 1 (HIV-1) Group M, HIV-1 Group O, and HIV-2 Isolates. *Journal of Virology* **79**: 8979-8990.

Artenstein, AW, Coppola, J, Brown, AE, Carr, JK, Sanders-Buell, E, Galbarini, E, Mascola, JR, VanCott, TC, Schonbrood, P, McCutchan, FE, et al. (1995) Multiple introductions of HIV-1 subtype E into the western hemisphere. *Lancet* **346**: 1197-1198.

Ayouba, A, Souquieres, S, Njinku, B, Martin, PM, Muller-Trutwin, MC, Roques, P, Barre-Sinoussi, F, Mauclere, P, Simon, F, Nerrienet, E (2000) HIV-1 group N among HIV-1-seropositive individuals in Cameroon. *AIDS* **14**: 2623-2625.

Bailes, E, Chaudhuri, RR, Santiago, ML, Bibollet-Ruche, F, Hahn, BH, Sharp, PM (2002) The Evolution of Primate Lentiviruses and the Origins of AIDS. In: The Molecular Epidemiology of Human Viruses, Springer US, pp. 65-96.

Baldwin, CE, Sanders, RW, Berkhout, B (2003) Inhibiting HIV-1 entry with fusion inhibitors. *Curr Med Chem* **10**: 1633-1642.

Baltimore, D (1971) Expression of animal virus genomes. *Bacteriological Reviews* **35**: 235.

Balzarini, J (2004) Current Status of the Non-nucleoside Reverse Transcriptase Inhibitors of Human Immunodeficiency Virus Type 1. *Current Topics in Medicinal Chemistry* **4**: 921-944.

Baur, AS, Sawai, ET, Dazin, P, Fantl, WJ, Cheng-Mayer, C, Peterlin, BM (1994) HIV-1 Nef leads to inhibition or activation of T cells depending on its intracellular localization. *Immunity* **1**: 373-384.

Baxter, JD, Mayers, DL, Wentworth, DN, Neaton, JD, Hoover, ML, Winters, MA, Mannheimer, SB, Thompson, MA, Abrams, DI, Brizz, BJ, Ioannidis, JP, Merigan, TC (2000) A randomized study of antiretroviral management based on plasma genotypic antiretroviral resistance testing in patients failing therapy. CPCRA 046 Study Team for the Terry Beirn Community Programs for Clinical Research on AIDS. *AIDS* **14**: F83-93.

Bebenek, K, Abbotts, J, Roberts, JD, Wilson, SH, Kunkel, TA (1989) Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase. *J Biol Chem* **264**: 16948-16956.

Bebenek, K, Abbotts, J, Wilson, SH, Kunkel, TA (1993) Error-prone polymerization by HIV-1 reverse transcriptase. Contribution of template-primer misalignment, miscoding, and termination probability to mutational hot spots. *J Biol Chem* **268**: 10324-10334.

Ben-Artzi, H, Shemesh, J, Zeelon, E, Amit, B, Kleiman, L, Gorecki, M, Panet, A (1996) Molecular analysis of the second template switch during reverse transcription of the HIV RNA template. *Biochemistry* **35**: 10549-10557.

Bentley, DR, Balasubramanian, S, Swerdlow, HP, Smith, GP, Milton, J, Brown, CG, Hall, KP, Evers, DJ, Barnes, CL, Bignell, HR, Boutell, JM, Bryant, J, Carter, RJ, Keira Cheetham, R, Cox, AJ, Ellis, DJ, Flatbush, MR, Gormley, NA, Humphray, SJ, Irving, LJ, Karbelashvili, MS, Kirk, SM, Li, H, Liu, X, Maisinger, KS, Murray, LJ, Obradovic, B, Ost, T, Parkinson, ML, Pratt, MR, Rasolonjatovo, IM, Reed, MT, Rigatti, R, Rodighiero, C, Ross, MT, Sabot, A, Sankar, SV, Scally, A, Schroth, GP, Smith, ME, Smith, VP, Spiridou, A, Torrance, PE, Tzonev, SS, Vermaas, EH, Walter, K, Wu, X, Zhang, L, Alam, MD, Anastasi, C, Aniebo, IC, Bailey, DM, Bancarz, IR, Banerjee, S, Barbour, SG, Baybayan, PA, Benoit, VA, Benson, KF, Bevis, C, Black, PJ, Boodhun, A, Brennan, JS, Bridgham, JA, Brown, RC, Brown, AA, Buermann, DH, Bundu, AA, Burrows, JC, Carter, NP, Castillo, N, Chiara, ECM, Chang, S, Neil Cooley, R, Crake, NR, Dada, OO, Diakoumakos, KD, Dominguez-Fernandez, B, Earnshaw, DJ, Egbujor, UC, Elmore, DW, Etchin, SS, Ewan, MR, Fedurco, M, Fraser, LJ, Fuentes Fajardo, KV, Scott Furey, W, George, D, Gietzen, KJ, Goddard, CP, Golda, GS, Granieri, PA, Green, DE, Gustafson, DL, Hansen, NF, Harnish, K, Haudenschild, CD, Heyer, NI, Hims, MM, Ho, JT, Horgan, AM, Hoschler, K, Hurwitz, S, Ivanov, DV, Johnson, MQ, James, T, Huw Jones, TA, Kang, GD, Kerelska, TH, Kersey, AD, Khrebtukova, I, Kindwall, AP, Kingsbury, Z, Kokko-Gonzales, PI, Kumar, A, Laurent, MA, Lawley, CT, Lee, SE, Lee, X, Liao, AK, Loch, JA, Lok, M, Luo, S, Mammen, RM, Martin, JW, McCauley, PG, McNitt, P, Mehta, P, Moon, KW, Mullens, JW, Newington, T, Ning, Z, Ling Ng, B, Novo, SM, O'Neill, MJ, Osborne, MA, Osnowski, A, Ostadan, O, Paraschos, LL, Pickering, L, Pike, AC, Chris Pinkard, D, Pliskin, DP, Podhasky, J, Quijano, VJ, Raczy, C, Rae, VH, Rawlings, SR, Chiva Rodriguez, A, Roe, PM, Rogers, J, Rogert Bacigalupo, MC, Romanov, N, Romieu, A, Roth, RK, Rourke, NJ, Ruediger, ST, Rusman, E, Sanches-Kuiper, RM, Schenker, MR, Seoane, JM, Shaw, RJ, Shiver, MK, Short, SW, Sizto, NL, Sluis, JP, Smith, MA, Ernest Sohna Sohna, J, Spence, EJ, Stevens, K, Sutton, N, Szajkowski, L, Tregidgo, CL, Turcatti, G, Vandevondele, S, Verhovsky, Y, Virk, SM, Wakelin, S, Walcott, GC, Wang, J, Worsley, GJ, Yan, J, Yau, L, Zuerlein, M, Mullikin, JC, Hurles, ME, McCooke, NJ, West, JS, Oaks, FL, Lundberg, PL, Klenerman, D, Durbin, R, Smith, AJ (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**: 53-59.

Bera, S, Pandey, KK, Vora, AC, Grandgenett, DP (2011) HIV-1 Integrase Strand Transfer Inhibitors Stabilize an Integrase–Single Blunt-Ended DNA Complex. *Journal of Molecular Biology* **410**: 831-846.

Blagoveshchenskaya, AD, Thomas, L, Feliciangeli, SF, Hung, CH, Thomas, G (2002) HIV-1 Nef downregulates MHC-I by a PACS-1- and PI3K-regulated ARF6 endocytic pathway. *Cell* **111**: 853-866.

Boden, D, Markowitz, M (1998) Resistance to Human Immunodeficiency Virus Type 1 Protease Inhibitors. *Antimicrobial Agents and Chemotherapy* **42**: 2775-2783.

Bordoni, R, Bonnal, R, Rizzi, E, Carrera, P, Benedetti, S, Cremonesi, L, Stenirri, S, Colombo, A, Montrasio, C, Bonalumi, S, Albertini, A, Bernardi, LR, Ferrari, M, De Bellis, G (2008) Evaluation of human gene variant detection in amplicon pools by the GS-FLX parallel Pyrosequencer. *BMC Genomics* **9**: 464.

Borrow, P, Lewicki, H, Wei, X, Horwitz, MS, Peffer, N, Meyers, H, Nelson, JA, Gairin, JE, Hahn, BH, Oldstone, MB, Shaw, GM (1997) Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. *Nat Med* **3**: 205-211.

Borsetti, A, Ohagen, A, Gottlinger, HG (1998) The C-terminal half of the human immunodeficiency virus type 1 Gag precursor is sufficient for efficient particle assembly. *J Virol* **72**: 9313-9317.

Bourgeois, CF, Kim, YK, Churcher, MJ, West, MJ, Karn, J (2002) Spt5 cooperates with human immunodeficiency virus type 1 Tat by preventing premature RNA release at terminator sequences. *Mol Cell Biol* **22**: 1079-1093.

Briggs, JA, Krausslich, HG The molecular architecture of HIV. *J Mol Biol* **410**: 491-500.

Briggs, JA, Simon, MN, Gross, I, Krausslich, HG, Fuller, SD, Vogt, VM, Johnson, MC (2004) The stoichiometry of Gag protein in HIV-1. *Nat Struct Mol Biol* **11**: 672-675.

Briz, V, Poveda, E, Soriano, V (2006) HIV entry inhibitors: mechanisms of action and resistance pathways. *Journal of Antimicrobial Chemotherapy* **57**: 619-627.

Bukrinsky, MI, Sharova, N, Dempsey, MP, Stanwick, TL, Bukrinskaya, AG, Haggerty, S, Stevenson, M (1992) Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. *Proceedings of the National Academy of Sciences* **89**: 6580-6584.

Bukrinsky, MI, Sharova, N, McDonald, TL, Pushkarskaya, T, Tarpley, WG, Stevenson, M (1993) Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc Natl Acad Sci U S A* **90**: 6125-6129.

Carr, A, Miller, J, Law, M, Cooper, DA (2000) A syndrome of lipoatrophy, lactic acidaemia and liver dysfunction associated with HIV nucleoside analogue therapy: contribution to protease inhibitor-related lipodystrophy syndrome. *AIDS* **14**: F25-32.

Carr, A, Samaras, K, Burton, S, Law, M, Freund, J, Chisholm, DJ, Cooper, DA (1998a) A syndrome of peripheral lipodystrophy, hyperlipidaemia and insulin resistance in patients receiving HIV protease inhibitors. *AIDS* **12**: F51-58.

Carr, A, Samaras, K, Chisholm, DJ, Cooper, DA (1998b) Abnormal fat distribution and use of protease inhibitors. *Lancet* **351**: 1736.

Carr, A, Samaras, K, Chisholm, DJ, Cooper, DA (1998c) Pathogenesis of HIV-1-protease inhibitor-associated peripheral lipodystrophy, hyperlipidaemia, and insulin resistance. *Lancet* **351**: 1881-1883.

Chan, DC, Fass, D, Berger, JM, Kim, PS (1997) Core Structure of gp41 from the HIV Envelope Glycoprotein. *Cell* **89**: 263-273.

Chen, X, Tsiang, M, Yu, F, Hung, M, Jones, GS, Zeynalzadegan, A, Qi, X, Jin, H, Kim, CU, Swaminathan, S, Chen, JM (2008) Modeling, Analysis, and Validation of a Novel HIV Integrase Structure Provide Insights into the Binding Modes of Potent Integrase Inhibitors. *Journal of Molecular Biology* **380**: 504-519.

Cherry, S, Doukas, T, Armknecht, S, Whelan, S, Wang, H, Sarnow, P, Perrimon, N (2005) Genome-wide RNAi screen reveals a specific sensitivity of IRES-containing RNA viruses to host translation inhibition. *Genes Dev* **19**: 445-452.

Chou, S, Upton, H, Bao, K, Schulze-Gahmen, U, Samelson, AJ, He, N, Nowak, A, Lu, H, Krogan, NJ, Zhou, Q, Alber, T HIV-1 Tat recruits transcription elongation factors dispersed along a flexible AFF4 scaffold. *Proc Natl Acad Sci U S A* **110**: E123-131.

Christ, F, Thys, W, De Rijck, J, Gijsbers, R, Albanese, A, Arosio, D, Emiliani, S, Rain, JC, Benarous, R, Cereseto, A, Debyser, Z (2008) Transportin-SR2 imports HIV into the nucleus. *Curr Biol* **18**: 1192-1202.

Chukkapalli, V, Oh, SJ, Ono, A Opposing mechanisms involving RNA and lipids regulate HIV-1 Gag membrane binding through the highly basic region of the matrix domain. *Proc Natl Acad Sci U S A* **107**: 1600-1605.

Clavel, F, Hance, AJ (2004) HIV Drug Resistance. *New England Journal of Medicine* **350**: 1023-1035.

Cock, PJA, Fields, CJ, Goto, N, Heuer, ML, Rice, PM (2010) The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research* **38**: 1767-1771.

Cohen, EA, Subbramanian, RA, Gottlinger, HG (1996) Role of auxiliary proteins in retroviral morphogenesis. *Curr Top Microbiol Immunol* **214**: 219-235.

Cohen, EA, Terwilliger, EF, Sodroski, JG, Haseltine, WA (1988) Identification of a protein encoded by the vpu gene of HIV-1. *Nature* **334**: 532-534.

Cohen, J (2002) Confronting the Limits of Success. *Science* **296**: 2320-2324.

Colin, P, Benureau, Y, Staropoli, I, Wang, Y, Gonzalez, N, Alcami, J, Hartley, O, Brelot, A, Arenzana-Seisdedos, F, Lagane, B (2013) HIV-1 exploits CCR5 conformational heterogeneity to escape inhibition by chemokines. *Proc Natl Acad Sci U S A* **110**: 9475-9480.

Collins, KL, Chen, BK, Kalams, SA, Walker, BD, Baltimore, D (1998) HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. *Nature* **391**: 397-401.

Condra, JH, Schleif, WA, Blahy, OM, Gabryelski, LJ, Graham, DJ, Quintero, JC, Rhodes, A, Robbins, HL, Roth, E, Shivaprakash, M, et al. (1995) In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors. *Nature* **374**: 569-571.

Connor, RI, Chen, BK, Choe, S, Landau, NR (1995) Vpr Is Required for Efficient Replication of Human Immunodeficiency Virus Type-1 in Mononuclear Phagocytes. *Virology* **206**: 935-944.

Conticello, SG, Harris, RS, Neuberger, MS (2003) The Vif Protein of HIV Triggers Degradation of the Human Antiretroviral DNA Deaminase APOBEC3G. *Current Biology* **13**: 2009-2013.

Corbet, S, Müller-Trutwin, MC, Versmisse, P, Delarue, S, Ayouba, A, Lewis, J, Brunak, S, Martin, P, Brun-Vezinet, Ft, Simon, Ft (2000) env sequences of simian immunodeficiency viruses from chimpanzees in Cameroon are strongly related to those of human immunodeficiency virus group N from the same geographic area. *Journal of virology* **74**: 529–534.

Corbitt, G, Bailey, A, Williams, G (1990) HIV infection in Manchester, 1959. *The Lancet* **336**: 51.

Couillin, I, Culmann-Penciolelli, B, Gomard, E, Choppin, J, Levy, JP, Guillet, JG, Saragosti, S (1994) Impaired cytotoxic T lymphocyte recognition due to genetic variations in the main immunogenic region of the human immunodeficiency virus 1 NEF protein. *The Journal of Experimental Medicine* **180**: 1129-1134.

Cullen, BR (1991) Regulation of HIV-1 gene expression. *FASEB J* **5**: 2361-2368.

Daly, TJ, Cook, KS, Gray, GS, Maione, TE, Rusche, JR (1989) Specific binding of HIV-1 recombinant Rev protein to the Rev-responsive element in vitro. *Nature* **342**: 816-819.

Damond, F, Worobey, M, Campa, P, Farfara, I, Colin, G, Matheron, S, Brun-Vézinet, Ft, Robertson, DL, Simon, Ft (2004) Identification of a highly divergent HIV type 2 and proposal for a change in HIV type 2 classification. *AIDS research and human retroviruses* **20**: 666–672.

Darke, PL, Nutt, RF, Brady, SF, Garsky, VM, Ciccarone, TM, Leu, C-T, Lumma, PK, Freidinger, RM, Veber, DF, Sigal, IS (1988a) HIV-1 protease specificity of peptide cleavage is sufficient for processing of gag and pol polyproteins. *Biochemical and Biophysical Research Communications* **156**: 297-303.

Darke, PL, Nutt, RF, Brady, SF, Garsky, VM, Ciccarone, TM, Leu, CT, Lumma, PK, Freidinger, RM, Veber, DF, Sigal, IS (1988b) HIV-1 protease specificity of peptide cleavage is sufficient for processing of gag and pol polyproteins. *Biochem Biophys Res Commun* **156**: 297-303.

Davey, NE, Travé, G, Gibson, TJ (2011) How viruses hijack cell regulation. *Trends in Biochemical Sciences* **36**: 159-169.

Davies, JF, Hostomska, Z, Hostomsky, Z, Jordan, Matthews, DA (1991) Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. *Science* **252**: 88-95.

Dawson, L, Yu, X-F (1998) The Role of Nucleocapsid of HIV-1 in Virus Assembly. *Virology* **251**: 141-157.

De Clercq, E (1996) Non-nucleoside reverse transcriptase inhibitors (NNRTIs) for the treatment of human immunodeficiency virus type 1 (HIV-1) infections: Strategies to overcome drug resistance development. *Medicinal Research Reviews* **16**: 125–157.

De Clercq, E (2002) Strategies in the design of antiviral drugs. *Nature Reviews Drug Discovery* **1**: 13-25.

De Clercq, E (2005a) Emerging anti-HIV drugs.

De Clercq, E (2005b) New Approaches toward Anti-HIV Chemotherapy‡. *Journal of Medicinal Chemistry* **48**: 1297-1313.

De Clercq, E (2009) Anti-HIV drugs: 25 compounds approved within 25 years after the discovery of HIV. *International journal of antimicrobial agents* **33**: 307–320.

Decroly, E, Vandenbranden, M, Ruysschaert, JM, Cogniaux, J, Jacob, GS, Howard, SC, Marshall, G, Kompelli, A, Basak, A, Jean, F (1994) The convertases furin and PC1 can both cleave the human immunodeficiency virus (HIV)-1 envelope glycoprotein gp160 into gp120 (HIV-1 SU) and gp41 (HIV-I TM). *Journal of Biological Chemistry* **269**: 12240-12247.

Delwart, E, Magierowska, M, Royz, M, Foley, B, Peddada, L, Smith, R, Heldebrant, C, Conrad, A, Busch, M (2002) Homogeneous quasispecies in 16 out of 17 individuals during very early HIV-1 primary infection. *AIDS* **16**: 189-195.

Demirov, DG, Orenstein, JM, Freed, EO (2002) The late domain of human immunodeficiency virus type 1 p6 promotes virus release in a cell type-dependent manner. *J Virol* **76**: 105-117.

di Marzo Veronese, F, Reitz, MS, Jr., Gupta, G, Robert-Guroff, M, Boyer-Thompson, C, Louie, A, Gallo, RC, Lusso, P (1993) Loss of a neutralizing epitope by a spontaneous point mutation in the V3 loop of HIV-1 isolated from an infected laboratory worker. *J Biol Chem* **268**: 25894-25901.

Dismuke, DJ, Aiken, C (2006) Evidence for a functional link between uncoating of the human immunodeficiency virus type 1 core and nuclear import of the viral preintegration complex. *J Virol* **80**: 3712-3720.

Dong, X, Li, H, Derdowski, A, Ding, L, Burnett, A, Chen, X, Peters, TR, Dermody, TS, Woodruff, E, Wang, JJ, Spearman, P (2005) AP-3 directs the intracellular trafficking of HIV-1 Gag and plays a key role in particle assembly. *Cell* **120**: 663-674.

Droege, M, Hill, B (2008) The Genome Sequencer FLX System--longer reads, more applications, straight forward bioinformatics and more complete data sets. *J Biotechnol* **136**: 3-10.

Duffalo, ML, James, CW (2003) Enfuvirtide: A Novel Agent for the Treatment of HIV-1 Infection. *The Annals of Pharmacotherapy* **37**: 1448-1456.

Dumonceaux, J, Nisole, S, Chanel, C, Quivet, L, Amara, A, Baleux, F, Briand, P, Hazan, U (1998) Spontaneous mutations in the env gene of the human immunodeficiency virus type 1 NDK isolate are associated with a CD4-independent entry phenotype. *J Virol* **72**: 512-519.

Durant, J, Clevenbergh, P, Halfon, P, Delgiudice, P, Porsin, S, Simonet, P, Montagne, N, Boucher, CA, Schapiro, JM, Dellamonica, P (1999) Drug-resistance genotyping in HIV-1 therapy: the VIRADAPT randomised controlled trial. *Lancet* **353**: 2195-2199.

Dwyer, JJ, Hasan, A, Wilson, KL, White, JM, Matthews, TJ, Delmedico, MK (2003) The hydrophobic pocket contributes to the structural stability of the N-terminal coiled coil of HIV gp41 but is not required for six-helix bundle formation. *Biochemistry* **42**: 4945-4953.

Emiliani, S, Mousnier, A, Busschots, K, Maroun, M, Maele, BV, Tempé, D, Vandekerckhove, L, Moisant, F, Ben-Slama, L, Witvrouw, M, Christ, F, Rain, J-C, Dargemont, C, Debyser, Z, Benarous, R (2005) Integrase Mutants Defective for Interaction with LEDGF/p75 Are Impaired in Chromosome Tethering and HIV-1 Replication. *Journal of Biological Chemistry* **280**: 25517-25523.

ERICKSON-VIITANEN, S, MANFREDI, J, VIITANEN, P, TRIBE, DE, TRITCH, R, HUTCHISON III, CA, LOEB, DD, SWANSTROM, R (1989) Cleavage of HIV-1 gag polyprotein synthesized in vitro: sequential cleavage by the viral protease. *AIDS research and human retroviruses* **5**: 577–591.

Esnouf, R, Ren, J, Ross, C, Jones, Y, Stammers, D, Stuart, D (1995) Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors. *Nature Structural & Molecular Biology* **2**: 303–308.

Esparza, J, Bhamarapravati, N (2000) Accelerating the development and future availability of HIV-1 vaccines: why, when, where, and how? *Lancet* **355**: 2061-2066.

Espeseth, AS, Felock, P, Wolfe, A, Witmer, M, Grobler, J, Anthony, N, Egbertson, M, Melamed, JY, Young, S, Hamill, T, Cole, JL, Hazuda, DJ (2000) HIV-1 integrase inhibitors that compete with the target DNA substrate define a unique strand transfer conformation for integrase. *Proc Natl Acad Sci U S A* **97**: 11244-11249.

Ewing, B, Hillier, L, Wendl, MC, Green, P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* **8**: 175-185.

Fang, G, Weiser, B, Kuiken, C, Philpott, SM, Rowland-Jones, S, Plummer, F, Kimani, J, Shi, B, Kaul, R, Bwayo, J, Anzala, O, Burger, H (2004) Recombination following superinfection by HIV-1. *AIDS* **18**: 153-159.

Farnet, CM, Haseltine, WA (1991) Determination of viral proteins present in the human immunodeficiency virus type 1 preintegration complex. *J Virol* **65**: 1910-1915.

Fassati, A, Gorlich, D, Harrison, I, Zaytseva, L, Mingot, JM (2003) Nuclear import of HIV-1 intracellular reverse transcription complexes is mediated by importin 7. *EMBO J* **22**: 3675-3685.

Fätkenheuer, G, Pozniak, AL, Johnson, MA, Plettenberg, A, Staszewski, S, Hoepelman, AIM, Saag, MS, Goebel, FD, Rockstroh, JK, Dezube, BJ, Jenkins, TM, Medhurst, C, Sullivan, JF, Ridgway, C, Abel, S, James, IT, Youle, M, van der Ryst, E (2005) Efficacy of short-term monotherapy with maraviroc, a new CCR5 antagonist, in patients infected with HIV-1. *Nature Medicine* **11**: 1170-1172.

Fikkert, V, Maele, BV, Vercammen, J, Hantson, A, Remoortel, BV, Michiels, M, Gurnari, C, Pannecouque, C, Maeyer, MD, Engelborghs, Y, Clercq, ED, Debyser, Z, Witvrouw, M (2003) Development of Resistance against Diketo Derivatives of Human Immunodeficiency Virus Type 1 by Progressive Accumulation of Integrase Mutations. *Journal of Virology* **77**: 11459-11470.

Fischer, U, Huber, J, Boelens, WC, Mattaj, IW, Luhrmann, R (1995) The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* **82**: 475-483.

Fischer, U, Meyer, S, Teufel, M, Heckel, C, Luhrmann, R, Rautmann, G (1994) Evidence that HIV-1 Rev directly promotes the nuclear export of unspliced RNA. *EMBO J* **13**: 4105-4112.

Fischer, W, Ganusov, VV, Giorgi, EE, Hraber, PT, Keele, BF, Leitner, T, Han, CS, Gleasner, CD, Green, L, Lo, CC, Nag, A, Wallstrom, TC, Wang, S, McMichael, AJ, Haynes, BF, Hahn, BH, Perelson, AS, Borrow, P, Shaw, GM, Bhattacharya, T, Korber, BT (2010) Transmission of single HIV-1 genomes and dynamics of early immune escape revealed by ultra-deep sequencing. *PLoS One* **5**: e12303.

Fouchier, RA, Groenink, M, Kootstra, NA, Tersmette, M, Huisman, HG, Miedema, F, Schuitemaker, H (1992) Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. *J Virol* **66**: 3183-3187.

Francis, DP, Curran, JW, Essex, M (1983) Epidemic acquired immune deficiency syndrome: epidemiologic evidence for a transmissible agent. *Journal of the National Cancer Institute* **71**: 5–9.

Frankel, AD, Young, JAT (1998) HIV-1: Fifteen Proteins and an RNA. *Annual Review of Biochemistry* **67**: 1-25.

Friedman-Kien, AE (1981) Disseminated Kaposi's sarcoma syndrome in young homosexual men. *Journal of the American Academy of Dermatology* **5**: 468–471.

Friedman-Kien, AE, Laubenstein, L, Marmor, M, Hymes, K, Green, J, Ragaz, A, Gottleib, J, Muggia, F, Demopoulos, R, Weintraub, M (1981) Kaposi’s sarcoma and Pneumocystis pneumonia among homosexual men—New York City and California. *MMWR* **30**: 305–308.

Fujii, K, Hurley, JH, Freed, EO (2007) Beyond Tsg101: the role of Alix in'ESCRTing'HIV-1. *Nature Reviews Microbiology* **5**: 912–916.

Furuta, RA, Wild, CT, Weng, Y, Weiss, CD (1998) Capture of an early fusion-active conformation of HIV-1 gp41. *Nature Structural & Molecular Biology* **5**: 276-279.

Gallo, RC, Sarin, PS, Gelmann, EP, Robert-Guroff, M, Richardson, E, Kalyanaraman, VS, Mann, D, Sidhu, GD, Stahl, RE, Zolla-Pazner, S, Leibowitch, J, Popovic, M (1983) Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science (New York, NY)* **220**: 865-867.

Ganser-Pornillos, BK, von Schwedler, UK, Stray, KM, Aiken, C, Sundquist, WI (2004) Assembly properties of the human immunodeficiency virus type 1 CA protein. *J Virol* **78**: 2545-2552.

Gao, F, Bailes, E, Robertson, DL, Chen, Y, Rodenburg, CM, Michael, SF, Cummins, LB, Arthur, LO, Peeters, M, Shaw, GM (1999) Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature* **397**: 436-441.

Gao, F, Vidal, N, Li, Y, Trask, SA, Chen, Y, Kostrikis, LG, Ho, DD, Kim, J, Oh, M-D, Choe, K, Salminen, M, Robertson, DL, Shaw, GM, Hahn, BH, Peeters, M (2001) Evidence of Two Distinct Subsubtypes within the HIV-1 Subtype A Radiation. *AIDS Research and Human Retroviruses* **17**: 675-688.

Gao, F, Yue, L, Robertson, DL, Hill, SC, Hui, H, Biggar, RJ, Neequaye, AE, Whelan, TM, Ho, DD, Shaw, GM (1994) Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology. *Journal of virology* **68**: 7433–7447.

Gao, F, Yue, L, White, AT, Pappas, PG, Barchue, J, Hanson, AP, Greene, BM, Sharp, PM, Shaw, GM, Hahn, BH (1992) Human infection by genetically diverse SIVSM-related HIV-2 in West Africa. *Nature* **358**: 495-499.

Garcia, JV, Miller, AD (1991) Serine phosphorylation-independent downregulation of cell-surface CD4 by nef. *Nature* **350**: 508–511.

Garrus, JE, von Schwedler, UK, Pornillos, OW, Morham, SG, Zavitz, KH, Wang, HE, Wettstein, DA, Stray, KM, Cote, M, Rich, RL, Myszka, DG, Sundquist, WI (2001) Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* **107**: 55-65.

Gaynor, R (1992) Cellular transcription factors involved in the regulation of HIV-1 gene expression. *AIDS* **6**: 347-363.

Gheysen, D, Jacobs, E, de Foresta, F, Thiriart, C, Francotte, M, Thines, D, De Wilde, M (1989) Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. *Cell* **59**: 103-112.

Glenn, TC (2011) Field guide to next-generation DNA sequencers. *Mol Ecol Resour* **11**: 759-769.

Goodenow, M, Huet, T, Saurin, W, Kwok, S, Sninsky, J, Wain-Hobson, S (1989) HIV-1 isolates are rapidly evolving quasispecies: evidence for viral mixtures and preferred nucleotide substitutions. *J Acquir Immune Defic Syndr* **2**: 344-352.

Gottlieb, MS, Schroff, R, Schanker, HM, Weisman, JD, Fan, PT, Wolf, RA, Saxon, A (1981) \textitPneumocystis carinii Pneumonia and Mucosal Candidiasis in Previously Healthy Homosexual Men. *New England Journal of Medicine* **305**: 1425-1431.

Göttlinger, HG, Sodroski, JG, Haseltine, WA (1989) Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. *Proceedings of the National Academy of Sciences* **86**: 5781-5785.

Greenberg, ME, Iafrate, AJ, Skowronski, J (1998) The SH3 domain-binding surface and an acidic motif in HIV-1 Nef regulate trafficking of class I MHC complexes. *EMBO J* **17**: 2777-2789.

Gu, Z, Gao, Q, Faust, EA, Wainberg, MA (1995) Possible involvement of cell fusion and viral recombination in generation of human immunodeficiency virus variants that display dual resistance to AZT and 3TC. *J Gen Virol* **76 ( Pt 10)**: 2601-2605.

Gürtler, L (2004) [Zoonotic infections stimulation]. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz* **47**: 609-610.

Haase, AT Targeting early infection to prevent HIV-1 mucosal transmission. *Nature* **464**: 217-223.

Hahn, BH, Shaw, GM, De, KM, Sharp, PM (2000) AIDS as a zoonosis: scientific and public health implications. *Science* **287**: 607–614.

Hazuda, DJ, Anthony, NJ, Gomez, RP, Jolly, SM, Wai, JS, Zhuang, L, Fisher, TE, Embrey, M, Guare, JP, Egbertson, MS, Vacca, JP, Huff, JR, Felock, PJ, Witmer, MV, Stillmock, KA, Danovich, R, Grobler, J, Miller, MD, Espeseth, AS, Jin, L, Chen, I-W, Lin, JH, Kassahun, K, Ellis, JD, Wong, BK, Xu, W, Pearson, PG, Schleif, WA, Cortese, R, Emini, E, Summa, V, Holloway, MK, Young, SD (2004) A naphthyridine carboxamide provides evidence for discordant resistance between mechanistically identical inhibitors of HIV-1 integrase. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 11233-11238.

Hazuda, DJ, Felock, P, Witmer, M, Wolfe, A, Stillmock, K, Grobler, JA, Espeseth, A, Gabryelski, L, Schleif, W, Blau, C, Miller, MD (2000) Inhibitors of Strand Transfer That Prevent Integration and Inhibit HIV-1 Replication in Cells. *Science* **287**: 646-650.

He, J, Choe, S, Walker, R, Marzio, PD, Morgan, DO, Landau, NR (1995) Human immunodeficiency virus type 1 viral protein R (Vpr) arrests cells in the G2 phase of the cell cycle by inhibiting p34cdc2 activity. *Journal of Virology* **69**: 6705-6711.

He, N, Zhou, Q New insights into the control of HIV-1 transcription: when Tat meets the 7SK snRNP and super elongation complex (SEC). *J Neuroimmune Pharmacol* **6**: 260-268.

Hemelaar, J, Gouws, E, Ghys, PD, Osmanov, S (2006) Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* **20**: W13-W23.

Henderson, BR, Percipalle, P (1997) Interactions between HIV Rev and nuclear import and export factors: the Rev nuclear localisation signal mediates specific binding to human importin-beta. *J Mol Biol* **274**: 693-707.

Himmel, DM, Sarafianos, SG, Dharmasena, S, Hossain, MM, McCoy-Simandle, K, Ilina, T, Clark, AD, Jr., Knight, JL, Julias, JG, Clark, PK, Krogh-Jespersen, K, Levy, RM, Hughes, SH, Parniak, MA, Arnold, E (2006) HIV-1 reverse transcriptase structure with RNase H inhibitor dihydroxy benzoyl naphthyl hydrazone bound at a novel site. *ACS Chem Biol* **1**: 702-712.

Hirsch, VM, Olmsted, RA, Murphey-Corb, M, Purcell, RH, Johnson, PR (1989) An African primate lentivirus (SIVsmclosely related to HIV-2.

Ho, DD, Bieniasz, PD (2008) HIV-1 at 25. *Cell* **133**: 561-565.

Ho, DD, Neumann, AU, Perelson, AS, Chen, W, Leonard, JM, Markowitz, M (1995a) Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**: 123–126.

Ho, DD, Neumann, AU, Perelson, AS, Chen, W, Leonard, JM, Markowitz, M (1995b) Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**: 123-126.

Hosseinipour, MC, Gupta, RK, Van Zyl, G, Eron, JJ, Nachega, JB (2013) Emergence of HIV drug resistance during first- and second-line antiretroviral therapy in resource-limited settings. *J Infect Dis* **207 Suppl 2**: S49-56.

Hsiou, Y, Ding, J, Das, K, Clark Jr, AD, Boyer, PL, Lewi, P, Janssen, PA, Kleim, J-P, Rösner, M, Hughes, SH (2001) The Lys103Asn mutation of HIV-1 RT: a novel mechanism of drug resistance. *Journal of molecular biology* **309**: 437-445.

Huang, C-c, Lam, SN, Acharya, P, Tang, M, Xiang, S-H, Hussan, SS-u, Stanfield, RL, Robinson, J, Sodroski, J, Wilson, IA, Wyatt, R, Bewley, CA, Kwong, PD (2007) Structures of the CCR5 N Terminus and of a Tyrosine-Sulfated Antibody with HIV-1 gp120 and CD4. *Science* **317**: 1930-1934.

Huet, T, Cheynier, R, Meyerhans, A, Roelants, G, Wain-Hobson, S (1990) Genetic organization of a chimpanzee lentivirus related to HIV-1.

Hughes, JP, Totten, P (2003) Estimating the accuracy of polymerase chain reaction-based tests using endpoint dilution. *Biometrics* **59**: 505-511.

Hulme, AE, Perez, O, Hope, TJ (2011) Complementary assays reveal a relationship between HIV-1 uncoating and reverse transcription. *Proc Natl Acad Sci U S A* **108**: 9975-9980.

Hussain, A, Wesley, C, Khalid, M, Chaudhry, A, Jameel, S (2008) Human immunodeficiency virus type 1 Vpu protein interacts with CD74 and modulates major histocompatibility complex class II presentation. *Journal of virology* **82**: 893–902.

Ilina, T, Parniak, MA (2008) Inhibitors of HIV‐1 Reverse Transcriptase. In: Advances in Pharmacology, Academic Press, pp. 121-167.

Jabara, CB, Jones, CD, Roach, J, Anderson, JA, Swanstrom, R (2011) Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID. *Proc Natl Acad Sci U S A* **108**: 20166-20171.

Jacks, T, Power, MD, Masiarz, FR, Luciw, PA, Barr, PJ, Varmus, HE (1988) Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* **331**: 280-283.

Jacobo-Molina, A, Arnold, E (1991) HIV reverse transcriptase structure-function relationships. *Biochemistry* **30**: 6351–6361.

Jager, S, Kim, DY, Hultquist, JF, Shindo, K, LaRue, RS, Kwon, E, Li, M, Anderson, BD, Yen, L, Stanley, D, Mahon, C, Kane, J, Franks-Skiba, K, Cimermancic, P, Burlingame, A, Sali, A, Craik, CS, Harris, RS, Gross, JD, Krogan, NJ Vif hijacks CBF-beta to degrade APOBEC3G and promote HIV-1 infection. *Nature* **481**: 371-375.

Jakobson, CG, Dinnar, U, Feinsod, M, Nemirovsky, Y (2002) Ion-sensitive field-effect transistors in standard CMOS fabricated by post processing. *IEEE Sensors Journal* **2**: 279-287.

Ji, H, Li, Y, Graham, M, Liang, BB, Pilon, R, Tyson, S, Peters, G, Tyler, S, Merks, H, Bertagnolio, S, Soto-Ramirez, L, Sandstrom, P, Brooks, J (2011) Next-generation sequencing of dried blood spot specimens: a novel approach to HIV drug-resistance surveillance. *Antivir Ther* **16**: 871-878.

Johnson, JA, Geretti, AM Low-frequency HIV-1 drug resistance mutations can be clinically significant but must be interpreted with caution. *J Antimicrob Chemother* **65**: 1322-1326.

Jowett, JB, Planelles, V, Poon, B, Shah, NP, Chen, M-L, Chen, IS (1995) The human immunodeficiency virus type 1 vpr gene arrests infected T cells in the G2+ M phase of the cell cycle. *Journal of virology* **69**: 6304–6313.

Kanagawa, T (2003) Bias and artifacts in multitemplate polymerase chain reactions (PCR). *J Biosci Bioeng* **96**: 317-323.

Karacostas, V, Wolffe, EJ, Nagashima, K, Gonda, MA, Moss, B (1993) Overexpression of the HIV-1 gag-pol polyprotein results in intracellular activation of HIV-1 protease and inhibition of assembly and budding of virus-like particles. *Virology* **193**: 661–671.

Karn, J, Stoltzfus, CM Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harbor Perspectives in Medicine* **2**.

Keele, BF, Giorgi, EE, Salazar-Gonzalez, JF, Decker, JM, Pham, KT, Salazar, MG, Sun, C, Grayson, T, Wang, S, Li, H, Wei, X, Jiang, C, Kirchherr, JL, Gao, F, Anderson, JA, Ping, LH, Swanstrom, R, Tomaras, GD, Blattner, WA, Goepfert, PA, Kilby, JM, Saag, MS, Delwart, EL, Busch, MP, Cohen, MS, Montefiori, DC, Haynes, BF, Gaschen, B, Athreya, GS, Lee, HY, Wood, N, Seoighe, C, Perelson, AS, Bhattacharya, T, Korber, BT, Hahn, BH, Shaw, GM (2008) Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc Natl Acad Sci U S A* **105**: 7552-7557.

Keele, BF, Heuverswyn, FV, Li, Y, Bailes, E, Takehisa, J, Santiago, ML, Bibollet-Ruche, F, Chen, Y, Wain, LV, Liegeois, F, Loul, S, Ngole, EM, Bienvenue, Y, Delaporte, E, Brookfield, JFY, Sharp, PM, Shaw, GM, Peeters, M, Hahn, BH (2006) Chimpanzee Reservoirs of Pandemic and Nonpandemic HIV-1. *Science* **313**: 523-526.

Kliger, Y, Aharoni, A, Rapaport, D, Jones, P, Blumenthal, R, Shai, Y (1997) Fusion peptides derived from the HIV type 1 glycoprotein 41 associate within phospholipid membranes and inhibit cell-cell Fusion. Structure-function study. *J Biol Chem* **272**: 13496-13505.

Klimkait, T, Strebel, K, Hoggan, MD, Martin, MA, Orenstein, JM (1990) The human immunodeficiency virus type 1-specific protein vpu is required for efficient virus maturation and release. *Journal of Virology* **64**: 621-629.

Korber, B, Gaschen, B, Yusim, K, Thakallapally, R, Kesmir, C, Detours, V (2001) Evolutionary and immunological implications of contemporary HIV-1 variation. *British Medical Bulletin* **58**: 19-42.

Korber, B, Muldoon, M, Theiler, J, Gao, F, Gupta, R, Lapedes, A, Hahn, BH, Wolinsky, S, Bhattacharya, T (2000) Timing the ancestor of the HIV-1 pandemic strains. *Science* **288**: 1789-1796.

Kostrikis, LG, Touloumi, G, Karanicolas, R, Pantazis, N, Anastassopoulou, C, Karafoulidou, A, Goedert, JJ, Hatzakis, A (2002) Quantitation of human immunodeficiency virus type 1 DNA forms with the second template switch in peripheral blood cells predicts disease progression independently of plasma RNA load. *J Virol* **76**: 10099-10108.

Kuiken, C, Korber, B, Shafer, RW (2003) HIV sequence databases. *AIDS reviews* **5**: 52.

LaFemina, RL, Schneider, CL, Robbins, HL, Callahan, PL, LeGrow, K, Roth, E, Schleif, WA, Emini, EA (1992) Requirement of active human immunodeficiency virus type 1 integrase enzyme for productive infection of human T-lymphoid cells. *Journal of Virology* **66**: 7414-7419.

Laguette, N, Benichou, S, Basmaciogullari, S (2009) Human Immunodeficiency Virus Type 1 Nef Incorporation into Virions Does Not Increase Infectivity. *Journal of Virology* **83**: 1093-1104.

Lama, J, Mangasarian, A, Trono, D (1999) Cell-surface expression of CD4 reduces HIV-1 infectivity by blocking Env incorporation in a Nef- and Vpu-inhibitable manner. *Current biology: CB* **9**: 622-631.

Larder, BA, Darby, G, Richman, DD (1989) HIV with reduced sensitivity to zidovudine (AZT) isolated during prolonged therapy. *Science* **243**: 1731-1734.

Larder, BA, Kemp, SD (1989) Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT). *Science* **246**: 1155-1158.

Larsen, LS, Beliakova-Bethell, N, Bilanchone, V, Zhang, M, Lamsa, A, Dasilva, R, Hatfield, GW, Nagashima, K, Sandmeyer, S (2008) Ty3 nucleocapsid controls localization of particle assembly. *J Virol* **82**: 2501-2514.

Lemey, P, Pybus, OG, Rambaut, A, Drummond, AJ, Robertson, DL, Roques, P, Worobey, M, Vandamme, A-M (2004a) The Molecular Population Genetics of HIV-1 Group O. *Genetics* **167**: 1059-1068.

Lemey, P, Pybus, OG, Rambaut, A, Drummond, AJ, Robertson, DL, Roques, P, Worobey, M, Vandamme, AM (2004b) The molecular population genetics of HIV-1 group O. *Genetics* **167**: 1059-1068.

Lenassi, M, Cagney, G, Liao, M, Vaupotic, T, Bartholomeeusen, K, Cheng, Y, Krogan, NJ, Plemenitas, A, Peterlin, BM HIV Nef is secreted in exosomes and triggers apoptosis in bystander CD4+ T cells. *Traffic* **11**: 110-122.

Leslie, AJ, Pfafferott, KJ, Chetty, P, Draenert, R, Addo, MM, Feeney, M, Tang, Y, Holmes, EC, Allen, T, Prado, JG, Altfeld, M, Brander, C, Dixon, C, Ramduth, D, Jeena, P, Thomas, SA, John, AS, Roach, TA, Kupfer, B, Luzzi, G, Edwards, A, Taylor, G, Lyall, H, Tudor-Williams, G, Novelli, V, Martinez-Picado, J, Kiepiela, P, Walker, BD, Goulder, PJR (2004) HIV evolution: CTL escape mutation and reversion after transmission. *Nature Medicine* **10**: 282-289.

Letvin, NL (2006) Progress and obstacles in the development of an AIDS vaccine. *Nature Reviews Immunology* **6**: 930-939.

Levy, JA, Hoffman, AD, Kramer, SM, Landis, JA, Shimabukuro, JM, Oshiro, LS (1984) Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science* **225**: 840-842.

Lewis, MJ, Lee, P, Ng, HL, Yang, OO (2012) Immune Selection In Vitro Reveals Human Immunodeficiency Virus Type 1 Nef Sequence Motifs Important for Its Immune Evasion Function In Vivo. *Journal of Virology* **86**: 7126-7135.

Liang, J-S, Distler, O, Cooper, DA, Jamil, H, Deckelbaum, RJ, Ginsberg, HN, Sturley, SL (2001) HIV protease inhibitors protect apolipoprotein B from degradation by the proteasome: A potential mechanism for protease inhibitor-induced hyperlipidemia. *Nature Medicine* **7**: 1327-1331.

Liu, J, Bartesaghi, A, Borgnia, MJ, Sapiro, G, Subramaniam, S (2008) Molecular architecture of native HIV-1 gp120 trimers. *Nature* **455**: 109-113.

Liu, SL, Rodrigo, AG, Shankarappa, R, Learn, GH, Hsu, L, Davidov, O, Zhao, LP, Mullins, JI (1996) HIV quasispecies and resampling. *Science* **273**: 415-416.

Lole, KS, Bollinger, RC, Paranjape, RS, Gadkari, D, Kulkarni, SS, Novak, NG, Ingersoll, R, Sheppard, HW, Ray, SC (1999) Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* **73**: 152-160.

Long, EM, Martin, HL, Jr., Kreiss, JK, Rainwater, SM, Lavreys, L, Jackson, DJ, Rakwar, J, Mandaliya, K, Overbaugh, J (2000) Gender differences in HIV-1 diversity at time of infection. *Nat Med* **6**: 71-75.

Lorenzi, P, Opravil, M, Hirschel, B, Chave, JP, Furrer, HJ, Sax, H, Perneger, TV, Perrin, L, Kaiser, L, Yerly, S (1999) Impact of drug resistance mutations on virologic response to salvage therapy. Swiss HIV Cohort Study. *AIDS* **13**: F17-21.

Lucas, GM (2005) Antiretroviral adherence, drug resistance, viral fitness and HIV disease progression: a tangled web is woven. *Journal of Antimicrobial Chemotherapy* **55**: 413-416.

Madani, N, Kabat, D (1998) An endogenous inhibitor of human immunodeficiency virus in human lymphocytes is overcome by the viral Vif protein. *Journal of virology* **72**: 10251-10255.

Malim, MH, Bohnlein, S, Hauber, J, Cullen, BR (1989a) Functional dissection of the HIV-1 Rev trans-activator--derivation of a trans-dominant repressor of Rev function. *Cell* **58**: 205-214.

Malim, MH, Cullen, BR (1991) HIV-1 structural gene expression requires the binding of multiple Rev monomers to the viral RRE: implications for HIV-1 latency. *Cell* **65**: 241-248.

Malim, MH, Hauber, J, Le, S-Y, Maizel, JV, Cullen, BR (1989b) The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* **338**: 254–257.

Mammano, F, Trouplin, V, Zennou, V, Clavel, F (2000) Retracing the evolutionary pathways of human immunodeficiency virus type 1 resistance to protease inhibitors: virus fitness in the absence and in the presence of drug. *J Virol* **74**: 8524-8531.

Mangeat, B, Turelli, P, Caron, G, Friedli, M, Perrin, L, Trono, D (2003) Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* **424**: 99–103.

Mansky, LM (1996) Forward mutation rate of human immunodeficiency virus type 1 in a T lymphoid cell line. *AIDS Res Hum Retroviruses* **12**: 307-314.

Mansky, LM, Temin, HM (1995) Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *Journal of Virology* **69**: 5087-5094.

Marcello, A, Zoppe, M, Giacca, M (2001) Multiple modes of transcriptional regulation by the HIV-1 Tat transactivator. *IUBMB Life* **51**: 175-181.

Mardis, ER (2008) The impact of next-generation sequencing technology on genetics. *Trends Genet* **24**: 133-141.

Margulies, M, Egholm, M, Altman, WE, Attiya, S, Bader, JS, Bemben, LA, Berka, J, Braverman, MS, Chen, Y-J, Chen, Z, Dewell, SB, Du, L, Fierro, JM, Gomes, XV, Godwin, BC, He, W, Helgesen, S, Ho, CH, Irzyk, GP, Jando, SC, Alenquer, MLI, Jarvie, TP, Jirage, KB, Kim, J-B, Knight, JR, Lanza, JR, Leamon, JH, Lefkowitz, SM, Lei, M, Li, J, Lohman, KL, Lu, H, Makhijani, VB, McDade, KE, McKenna, MP, Myers, EW, Nickerson, E, Nobile, JR, Plant, R, Puc, BP, Ronan, MT, Roth, GT, Sarkis, GJ, Simons, JF, Simpson, JW, Srinivasan, M, Tartaro, KR, Tomasz, A, Vogt, KA, Volkmer, GA, Wang, SH, Wang, Y, Weiner, MP, Yu, P, Begley, RF, Rothberg, JM (2005a) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**: 376-380.

Margulies, M, Egholm, M, Altman, WE, Attiya, S, Bader, JS, Bemben, LA, Berka, J, Braverman, MS, Chen, YJ, Chen, Z, Dewell, SB, Du, L, Fierro, JM, Gomes, XV, Godwin, BC, He, W, Helgesen, S, Ho, CH, Irzyk, GP, Jando, SC, Alenquer, ML, Jarvie, TP, Jirage, KB, Kim, JB, Knight, JR, Lanza, JR, Leamon, JH, Lefkowitz, SM, Lei, M, Li, J, Lohman, KL, Lu, H, Makhijani, VB, McDade, KE, McKenna, MP, Myers, EW, Nickerson, E, Nobile, JR, Plant, R, Puc, BP, Ronan, MT, Roth, GT, Sarkis, GJ, Simons, JF, Simpson, JW, Srinivasan, M, Tartaro, KR, Tomasz, A, Vogt, KA, Volkmer, GA, Wang, SH, Wang, Y, Weiner, MP, Yu, P, Begley, RF, Rothberg, JM (2005b) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**: 376-380.

Mariani, R, Chen, D, Schröfelbauer, B, Navarro, F, König, R, Bollman, B, Münk, C, Nymark-McMahon, H, Landau, NR (2003) Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell* **114**: 21–31.

Marin, M, Rose, KM, Kozak, SL, Kabat, D (2003) HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nature medicine* **9**: 1398–1403.

McColl, DJ, Chen, X (2010) Strand transfer inhibitors of HIV-1 integrase: Bringing IN a new era of antiretroviral therapy. *Antiviral Research* **85**: 101-118.

McCutchan, FE (2000) Understanding the genetic diversity of HIV-1. *AIDS* **14 Suppl 3**: S31-44.

McCutchan, FE (2006) Global epidemiology of HIV. *Journal of Medical Virology* **78**: S7–S12.

McCutchan, FE, Carr, JK, Murphy, D, Piyasirisilp, S, Gao, F, Hahn, B, Yu, X-F, Beyrer, C, Birx, DL (2002) Precise mapping of recombination breakpoints suggests a common parent of two BC recombinant HIV type 1 strains circulating in China. *AIDS research and human retroviruses* **18**: 1135–1140.

McDonald, D, Vodicka, MA, Lucero, G, Svitkina, TM, Borisy, GG, Emerman, M, Hope, TJ (2002) Visualization of the intracellular behavior of HIV in living cells. *The Journal of Cell Biology* **159**: 441-452.

Mehle, A, Strack, B, Ancuta, P, Zhang, C, McPike, M, Gabuzda, D (2004) Vif overcomes the innate antiviral activity of APOBEC3G by promoting its degradation in the ubiquitin-proteasome pathway. *Journal of Biological Chemistry* **279**: 7792–7798.

Melby, T, Demasi, R, Cammack, N, Miralles, GD, Greenberg, ML (2007) Evolution of Genotypic and Phenotypic Resistance during Chronic Treatment with the Fusion Inhibitor T-1249. *AIDS Research and Human Retroviruses* **23**: 1366-1373.

Melikyan, GB (2008) Common principles and intermediates of viral protein-mediated fusion: the HIV-1 paradigm. *Retrovirology* **5**: 111.

Melikyan, GB, Markosyan, RM, Hemmati, H, Delmedico, MK, Lambert, DM, Cohen, FS (2000) Evidence That the Transition of HIV-1 Gp41 into a Six-Helix Bundle, Not the Bundle Configuration, Induces Membrane Fusion. *The Journal of Cell Biology* **151**: 413-424.

Menéndez-Arias, L (2002) Targeting HIV: antiretroviral therapy and development of drug resistance. *Trends in Pharmacological Sciences* **23**: 381-388.

Metzker, ML Sequencing technologies - the next generation. *Nat Rev Genet* **11**: 31-46.

Metzker, ML (2005) Emerging technologies in DNA sequencing. *Genome Res* **15**: 1767-1776.

Meyerhans, A, Vartanian, JP, Wain-Hobson, S (1990) DNA recombination during PCR. *Nucleic Acids Res* **18**: 1687-1691.

Milgrew, MJ, Hammond, PA, Cumming, DRS (2004) The development of scalable sensor arrays using standard CMOS technology. *Sensors and Actuators B: Chemical* **103**: 37-42.

Miller, J, Carr, A, Smith, D, Emery, S, Law, MG, Grey, P, Cooper, DA (2000) Lipodystrophy following antiretroviral therapy of primary HIV infection. *AIDS* **14**: 2406-2407.

Miller, MD, Farnet, CM, Bushman, FD (1997) Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. *Journal of Virology* **71**: 5382-5390.

Miller, V, Larder, BA (2000) Mutational patterns in the HIV genome and cross-resistance following nucleoside and nucleotide analogue drug exposure. *Antiviral therapy* **6**: 25-44.

Molla, A, Korneyeva, M, Gao, Q, Vasavanonda, S, Schipper, PJ, Mo, HM, Markowitz, M, Chernyavskiy, T, Niu, P, Lyons, N, Hsu, A, Granneman, GR, Ho, DD, Boucher, CA, Leonard, JM, Norbeck, DW, Kempf, DJ (1996) Ordered accumulation of mutations in HIV protease confers resistance to ritonavir. *Nat Med* **2**: 760-766.

Nahmias, A, Weiss, J, Yao, X, Lee, F, Kodsi, R, Schanfield, M, Matthews, T, Bolognesi, D, Durack, D, Motulsky, A (1986) Evidence for human infection with an HTLV III/LAV-like virus in Central Africa, 1959. *The Lancet* **327**: 1279-1280.

Navarro, F, Landau, NR (2004) Recent insights into HIV-1 Vif. *Current Opinion in Immunology* **16**: 477-482.

Neher, RA, Leitner, T Recombination rate and selection strength in HIV intra-patient evolution. *PLoS Comput Biol* **6**: e1000660.

Nermut, MV, Hockley, DJ, Bron, P, Thomas, D, Zhang, WH, Jones, IM (1998) Further evidence for hexagonal organization of HIV gag protein in prebudding assemblies and immature virus-like particles. *J Struct Biol* **123**: 143-149.

Niedringhaus, TP, Milanova, D, Kerby, MB, Snyder, MP, Barron, AE (2011) Landscape of next-generation sequencing technologies. *Anal Chem* **83**: 4327-4341.

Nijhuis, M, Schuurman, R, de Jong, D, Erickson, J, Gustchina, E, Albert, J, Schipper, P, Gulnik, S, Boucher, CA (1999) Increased fitness of drug resistant HIV-1 protease as a result of acquisition of compensatory mutations during suboptimal therapy. *Aids* **13**: 2349–2359.

Nomaguchi, M, Fujita, M, Adachi, A (2008) Role of HIV-1 Vpu protein for virus spread and pathogenesis. *Microbes and Infection* **10**: 960–967.

Nowak, MA, May, RM, Anderson, RM (1990) The evolutionary dynamics of HIV-1 quasispecies and the development of immunodeficiency disease. *AIDS* **4**: 1095-1103.

Nutt, RF, Brady, SF, Darke, PL, Ciccarone, TM, Colton, CD, Nutt, EM, Rodkey, JA, Bennett, CD, Waxman, LH, Sigal, IS, et al. (1988) Chemical synthesis and enzymatic activity of a 99-residue peptide with a sequence proposed for the human immunodeficiency virus protease. *Proc Natl Acad Sci U S A* **85**: 7129-7133.

Ott, M, Geyer, M, Zhou, Q (2011) The Control of HIV Transcription: Keeping RNA Polymerase II on Track. *Cell Host & Microbe* **10**: 426-435.

Palmer, S, Kearney, M, Maldarelli, F, Halvas, EK, Bixby, CJ, Bazmi, H, Rock, D, Falloon, J, Davey, RT, Jr., Dewar, RL, Metcalf, JA, Hammer, S, Mellors, JW, Coffin, JM (2005) Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. *J Clin Microbiol* **43**: 406-413.

Pancera, M, Majeed, S, Ban, YE, Chen, L, Huang, CC, Kong, L, Kwon, YD, Stuckey, J, Zhou, T, Robinson, JE, Schief, WR, Sodroski, J, Wyatt, R, Kwong, PD Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility. *Proc Natl Acad Sci U S A* **107**: 1166-1171.

Pandey, V, Nutter, RC, Prediger, E (2008) Applied Biosystems SOLiD™ System: Ligation‐Based Sequencing. *Next Generation Genome Sequencing: Towards Personalized Medicine*: 29-42.

Pannecouque, C, Pluymers, W, Van Maele, B, Tetz, V, Cherepanov, P, De Clercq, E, Witvrouw, M, Debyser, Z (2002) New Class of HIV Integrase Inhibitors that Block Viral Replication in Cell Culture. *Current Biology* **12**: 1169-1177.

Paredes, R, Lalama, CM, Ribaudo, HJ, Schackman, BR, Shikuma, C, Giguel, F, Meyer, WA, 3rd, Johnson, VA, Fiscus, SA, D'Aquila, RT, Gulick, RM, Kuritzkes, DR (2010) Pre-existing minority drug-resistant HIV-1 variants, adherence, and risk of antiretroviral treatment failure. *J Infect Dis* **201**: 662-671.

Peeters, M, Honoré, C, Huet, T, Bedjabaga, L, Ossari, S, Bussi, P, Cooper, RW, Delaporte, E (1989) Isolation and partial characterization of an HIV-related virus occurring naturally in chimpanzees in Gabon. *Aids* **3**: 625–630.

Pennisi, E (2010) Genomics. Semiconductors inspire new sequencing technologies. *Science* **327**: 1190.

Perrin, L, Kaiser, L, Yerly, S (2003) Travel and the spread of HIV-1 genetic variants. *Lancet Infect Dis* **3**: 22-27.

Pettit, SC, Lindquist, JN, Kaplan, AH, Swanstrom, R (2005) Processing sites in the human immunodeficiency virus type 1 (HIV-1) Gag-Pro-Pol precursor are cleaved by the viral protease at different rates. *Retrovirology* **2**: 66.

Ping, LH, Cohen, MS, Hoffman, I, Vernazza, P, Seillier-Moiseiwitsch, F, Chakraborty, H, Kazembe, P, Zimba, D, Maida, M, Fiscus, SA, Eron, JJ, Swanstrom, R, Nelson, JA (2000) Effects of genital tract inflammation on human immunodeficiency virus type 1 V3 populations in blood and semen. *J Virol* **74**: 8946-8952.

Plantier, JC, Leoz, M, Dickerson, JE, De Oliveira, F, Cordonnier, F, Lemee, V, Damond, F, Robertson, DL, Simon, F (2009) A new human immunodeficiency virus derived from gorillas. *Nat Med* **15**: 871-872.

Popov, S, Rexach, M, Ratner, L, Blobel, G, Bukrinsky, M (1998) Viral Protein R Regulates Docking of the HIV-1 Preintegration Complex to the Nuclear Pore Complex. *Journal of Biological Chemistry* **273**: 13347-13352.

Poveda, E, Briz, V, Soriano, V (2005) Enfuvirtide, the first fusion inhibitor to treat HIV infection. *Aids Rev* **7**: 139–147.

Preston, BD, Poiesz, BJ, Loeb, LA (1988) Fidelity of HIV-1 reverse transcriptase. *Science* **242**: 1168-1171.

Price, DA, Goulder, PJ, Klenerman, P, Sewell, AK, Easterbrook, PJ, Troop, M, Bangham, CR, Phillips, RE (1997) Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. *Proc Natl Acad Sci U S A* **94**: 1890-1895.

Pruss, D, Reeves, R, Bushman, FD, Wolffe, AP (1994) The influence of DNA and nucleosome structure on integration events directed by HIV integrase. *Journal of Biological Chemistry* **269**: 25031-25041.

Pujades-Rodriguez, M, O'Brien, D, Humblet, P, Calmy, A (2008) Second-line antiretroviral therapy in resource-limited settings: the experience of Medecins Sans Frontieres. *AIDS* **22**: 1305-1312.

Quinlan, AR, Stewart, DA, Stromberg, MP, Marth, GT (2008) Pyrobayes: an improved base caller for SNP discovery in pyrosequences. *Nat Methods* **5**: 179-181.

Quinn, TC (1996) Global burden of the HIV pandemic. *Lancet* **348**: 99-106.

Ratner, L, Haseltine, W, Patarca, R, Livak, KJ, Starcich, B, Josephs, SF, Doran, ER, Rafalski, JA, Whitehorn, EA, Baumeister, K (1985) Complete nucleotide sequence of the AIDS virus, HTLV-III.

Raymond, S, Delobel, P, Mavigner, M, Cazabat, M, Encinas, S, Souyris, C, Bruel, P, Sandres-Saune, K, Marchou, B, Massip, P, Izopet, J (2010) CXCR4-using viruses in plasma and peripheral blood mononuclear cells during primary HIV-1 infection and impact on disease progression. *AIDS* **24**: 2305-2312.

Razooky, BS, Weinberger, LS (2011) Mapping the architecture of the HIV-1 Tat circuit: A decision-making circuit that lacks bistability and exploits stochastic noise. *Methods* **53**: 68-77.

Riviere, L, Darlix, JL, Cimarelli, A Analysis of the viral elements required in the nuclear import of HIV-1 DNA. *J Virol* **84**: 729-739.

Rizzuto, CD, Wyatt, R, Hernandez-Ramos, N, Sun, Y, Kwong, PD, Hendrickson, WA, Sodroski, J (1998) A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding. *Science* **280**: 1949-1953.

Roberts, JD, Bebenek, K, Kunkel, TA (1988) The accuracy of reverse transcriptase from HIV-1. *Science* **242**: 1171-1173.

Robertson, D (2003) US FDA approves new class of HIV therapeutics. *Nature Biotechnology* **21**: 470-471.

Robertson, DL, Anderson, JP, Bradac, JA, Carr, JK, Foley, B, Funkhouser, RK, Gao, F, Hahn, BH, Kalish, ML, Kuiken, C (2000a) HIV-1 nomenclature proposal. *Science* **288**: 55–55.

Robertson, DL, Anderson, JP, Bradac, JA, Carr, JK, Foley, B, Funkhouser, RK, Gao, F, Hahn, BH, Kalish, ML, Kuiken, C, Learn, GH, Leitner, T, McCutchan, F, Osmanov, S, Peeters, M, Pieniazek, D, Salminen, M, Sharp, PM, Wolinsky, S, Korber, B (2000b) HIV-1 nomenclature proposal. *Science* **288**: 55-56.

Robertson, DL, Sharp, PM, McCutchan, FE, Hahn, BH (1995) Recombination in HIV-1. *Nature* **374**: 124-126.

Rogel, ME, Wu, LI, Emerman, M (1995) The human immunodeficiency virus type 1 vpr gene prevents cell proliferation during chronic infection. *Journal of virology* **69**: 882–888.

Rogers, MF, Thomas, PA, Starcher, ET, Noa, MC, Bush, TJ, Jaffe, HW (1987) Acquired Immunodeficiency Syndrome in Children: Report of the Centers for Disease Control National Surveillance, 1982 to 1985. *Pediatrics* **79**: 1008-1014.

Roques, P, Robertson, DL, Souquière, S, Damond, F, Ayouba, A, Farfara, I, Depienne, C, Nerrienet, E, Dormont, D, Brun-Vézinet, F, Simon, F, Mauclère, P (2002) Phylogenetic Analysis of 49 Newly Derived HIV-1 Group O Strains: High Viral Diversity but No Group M-like Subtype Structure. *Virology* **302**: 259-273.

Rosario, MC, Jacqmin, P, Dorr, P, van der Ryst, E, Hitchcock, C (2005) A pharmacokinetic-pharmacodynamic disease model to predict in vivo antiviral activity of maraviroc. *Clinical Pharmacology & Therapeutics* **78**: 508-519.

Rosario, MC, Poland, B, Sullivan, J, Westby, M, van der Ryst, E (2006) A pharmacokinetic-pharmacodynamic model to optimize the phase IIa development program of maraviroc. *JAIDS Journal of Acquired Immune Deficiency Syndromes* **42**: 183–191.

Rosen, CA, Pavlakis, GN (1990a) Tat and Rev: positive regulators of HIV gene expression. *AIDS* **4**: 499-509.

Rosen, CA, Pavlakis, GN (1990b) Tat and Rev: positive regulators of HIV gene expression. *AIDS* **4**: A51.

Rosenberg, ES, Altfeld, M, Poon, SH, Phillips, MN, Wilkes, BM, Eldridge, RL, Robbins, GK, D'Aquila, RT, Goulder, PJ, Walker, BD (2000) Immune control of HIV-1 after early treatment of acute infection. *Nature* **407**: 523-526.

Rothberg, JM, Hinz, W, Rearick, TM, Schultz, J, Mileski, W, Davey, M, Leamon, JH, Johnson, K, Milgrew, MJ, Edwards, M, Hoon, J, Simons, JF, Marran, D, Myers, JW, Davidson, JF, Branting, A, Nobile, JR, Puc, BP, Light, D, Clark, TA, Huber, M, Branciforte, JT, Stoner, IB, Cawley, SE, Lyons, M, Fu, Y, Homer, N, Sedova, M, Miao, X, Reed, B, Sabina, J, Feierstein, E, Schorn, M, Alanjary, M, Dimalanta, E, Dressman, D, Kasinskas, R, Sokolsky, T, Fidanza, JA, Namsaraev, E, McKernan, KJ, Williams, A, Roth, GT, Bustillo, J (2011) An integrated semiconductor device enabling non-optical genome sequencing. *Nature* **475**: 348-352.

Rowley, CF, Boutwell, CL, Lee, EJ, MacLeod, IJ, Ribaudo, HJ, Essex, M, Lockman, S Ultrasensitive detection of minor drug-resistant variants for HIV after nevirapine exposure using allele-specific PCR: clinical significance. *AIDS Res Hum Retroviruses* **26**: 293-300.

Roy, S, Delling, U, Chen, CH, Rosen, CA, Sonenberg, N (1990) A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated trans-activation. *Genes Dev* **4**: 1365-1373.

Saad, JS, Miller, J, Tai, J, Kim, A, Ghanam, RH, Summers, MF (2006) Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proc Natl Acad Sci U S A* **103**: 11364-11369.

Saksena, S, Sun, J, Chu, T, Emr, SD (2007) ESCRTing proteins in the endocytic pathway. *Trends Biochem Sci* **32**: 561-573.

Salemi, M, Strimmer, K, Hall, WW, Duffy, M, Delaporte, E, Mboup, S, Peeters, M, Vandamme, AM (2001) Dating the common ancestor of SIVcpz and HIV-1 group M and the origin of HIV-1 subtypes using a new method to uncover clock-like molecular evolution. *FASEB J* **15**: 276-278.

Salminen, MO, Carr, JK, Burke, DS, McCutchan, FE (1995) Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *AIDS Res Hum Retroviruses* **11**: 1423-1425.

Sanger, F, Nicklen, S, Coulson, AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**: 5463-5467.

Santiago, ML, Range, F, Keele, BF, Li, Y, Bailes, E, Bibollet-Ruche, F, Fruteau, C, Noë, R, Peeters, M, Brookfield, JF (2005) Simian immunodeficiency virus infection in free-ranging sooty mangabeys (Cercocebus atys atys) from the Tai Forest, Cote d'Ivoire: implications for the origin of epidemic human immunodeficiency virus type 2. *Journal of virology* **79**: 12515–12527.

Sarafianos, SG, Hughes, SH, Arnold, E (2004) Designing anti-AIDS drugs targeting the major mechanism of HIV-1 RT resistance to nucleoside analog drugs. *The International Journal of Biochemistry & Cell Biology* **36**: 1706-1715.

Sarafianos, SG, Marchand, B, Das, K, Himmel, DM, Parniak, MA, Hughes, SH, Arnold, E (2009) Structure and Function of HIV-1 Reverse Transcriptase: Molecular Mechanisms of Polymerization and Inhibition. *Journal of Molecular Biology* **385**: 693-713.

Sawai, ET, Baur, A, Struble, H, Peterlin, BM, Levy, JA, Cheng-Mayer, C (1994) Human immunodeficiency virus type 1 Nef associates with a cellular serine kinase in T lymphocytes. *Proceedings of the National Academy of Sciences* **91**: 1539-1543.

Schafer, A, Bogerd, HP, Cullen, BR (2004) Specific packaging of APOBEC3G into HIV-1 virions is mediated by the nucleocapsid domain of the gag polyprotein precursor. *Virology* **328**: 163-168.

Schubert, U, Ott, DE, Chertova, EN, Welker, R, Tessmer, U, Princiotta, MF, Bennink, JR, Krausslich, HG, Yewdell, JW (2000) Proteasome inhibition interferes with gag polyprotein processing, release, and maturation of HIV-1 and HIV-2. *Proc Natl Acad Sci U S A* **97**: 13057-13062.

Seelmeier, S, Schmidt, H, Turk, V, von der Helm, K (1988) Human immunodeficiency virus has an aspartic-type protease that can be inhibited by pepstatin A. *Proc Natl Acad Sci U S A* **85**: 6612-6616.

Shafer, Robert W (2006) Rationale and Uses of a Public HIV Drug‐Resistance Database. *The Journal of Infectious Diseases* **194**: S51-S58.

Shah, VB, Shi, J, Hout, DR, Oztop, I, Krishnan, L, Ahn, J, Shotwell, MS, Engelman, A, Aiken, C (2013) The host proteins transportin SR2/TNPO3 and cyclophilin A exert opposing effects on HIV-1 uncoating. *J Virol* **87**: 422-432.

Sharp, PM, Hahn, BH The evolution of HIV-1 and the origin of AIDS. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**: 2487-2494.

Sharp, PM, Hahn, BH (2010) The evolution of HIV-1 and the origin of AIDS. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**: 2487-2494.

Sharp, PM, Robertson, DL, Hahn, BH (1995) Cross-Species Transmission and Recombination of 'AIDS' Viruses. *Philosophical Transactions: Biological Sciences* **349**: 41-47.

Sheehy, AM, Gaddis, NC, Choi, JD, Malim, MH (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418**: 646-650.

Sheehy, AM, Gaddis, NC, Malim, MH (2003) The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nature medicine* **9**: 1404–1407.

Shendure, J, Ji, H (2008) Next-generation DNA sequencing. *Nat Biotechnol* **26**: 1135-1145.

Shokralla, S, Spall, JL, Gibson, JF, Hajibabaei, M Next-generation sequencing technologies for environmental DNA research. *Mol Ecol* **21**: 1794-1805.

Simen, BB, Huppler Hullsiek, K, Novak, RM, MacArthur, RD, Baxter, JD, Huang, C, Lubeski, C, Turenchalk, GS, Braverman, MS, Desany, B (2007) Prevalence of low abundant drug-resistant variants by ultra-deep sequencing in chronically HIV-infected antiretroviral (ARV)-naıve patients and the impact on virological outcomes. In: 16th International HIV Drug Resistance Workshop Barbados.

Simen, BB, Simons, JF, Hullsiek, KH, Novak, RM, MacArthur, RD, Baxter, JD, Huang, C, Lubeski, C, Turenchalk, GS, Braverman, MS, Desany, B, Rothberg, JM, Egholm, M (2009) Low-Abundance Drug-Resistant Viral Variants in Chronically HIV-Infected, Antiretroviral Treatment–Naive Patients Significantly Impact Treatment Outcomes. *Journal of Infectious Diseases* **199**: 693-701.

Simmons, A, Aluvihare, V, McMichael, A (2001) Nef triggers a transcriptional program in T cells imitating single-signal T cell activation and inducing HIV virulence mediators. *Immunity* **14**: 763-777.

Simon, F, Mauclère, P, Roques, P, Loussert-Ajaka, I, Müller-Trutwin, MC, Saragosti, S, Georges-Courbot, MC, Barré-Sinoussi, F, Brun-Vézinet, F (1998a) Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nature Medicine* **4**: 1032-1037.

Simon, JHM, Gaddis, NC, Fouchier, RAM, Malim, MH (1998b) Evidence for a newly discovered cellular anti-HIV-1 phenotype. *Nature Medicine* **4**: 1397-1400.

Simon, V, Zennou, V, Murray, D, Huang, Y, Ho, DD, Bieniasz, PD (2005) Natural variation in Vif: differential impact on APOBEC3G/3F and a potential role in HIV-1 diversification. *PLoS pathogens* **1**: e6.

Sluis-Cremer, N, Arion, D, Parniak\*, MA (2000) Molecular mechanisms of HIV-1 resistance to nucleoside reverse transcriptase inhibitors (NRTIs). *Cellular and Molecular Life Sciences CMLS* **57**: 1408-1422.

Smyth, RP, Davenport, MP, Mak, J (2012) The origin of genetic diversity in HIV-1. *Virus Res* **169**: 415-429.

Sodroski, J, Rosen, C, Wong-Staal, F, Salahuddin, SZ, Popovic, M, Arya, S, Gallo, RC, Haseltine, WA (1985) Trans-acting transcriptional regulation of human T-cell leukemia virus type III long terminal repeat. *Science* **227**: 171-173.

Stopak, K, de Noronha, C, Yonemoto, W, Greene, WC (2003) HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. *Molecular cell* **12**: 591–601.

Strebel, K, Klimkait, T, Martin, MA (1988) A novel gene of HIV-1, vpu, and its 16-kilodalton product. *Science* **241**: 1221-1223.

Struck, D, Wallis, CL, Denisov, G, Lambert, C, Servais, JY, Viana, RV, Letsoalo, E, Bronze, M, Aitken, SC, Schuurman, R, Stevens, W, Schmit, JC, Rinke de Wit, T, Perez Bercoff, D (2012) Automated sequence analysis and editing software for HIV drug resistance testing. *J Clin Virol* **54**: 30-35.

Stuhlmann, H, Berg, P (1992) Homologous recombination of copackaged retrovirus RNAs during reverse transcription. *J Virol* **66**: 2378-2388.

Tan, K, Liu, J-h, Wang, J-h, Shen, S, Lu, M (1997) Atomic structure of a thermostable subdomain of HIV-1 gp41. *Proceedings of the National Academy of Sciences* **94**: 12303-12308.

Tantillo, C, Ding, J, Jacobo-Molina, A, Nanni, RG, Boyer, PL, Hughes, SH, Pauwels, R, Andries, K, Janssen, PA, Arnold, E (1994) Locations of anti-AIDS drug binding sites and resistance mutations in the three-dimensional structure of HIV-1 reverse transcriptase. Implications for mechanisms of drug inhibition and resistance. *J Mol Biol* **243**: 369-387.

Taylor, BS, Sobieszczyk, ME, McCutchan, FE, Hammer, SM (2008) The Challenge of HIV-1 Subtype Diversity. *New England Journal of Medicine* **358**: 1590-1602.

Tebit, DM, Nankya, I, Arts, EJ, Gao, Y (2007) HIV diversity, recombination and disease progression: how does fitness “fit” into the puzzle. *AIDS Rev* **9**: 75–87.

Tersmette, M, Gruters, RA, Wolf, Fd, Goede, REd, Lange, JM, Schellekens, PT, Goudsmit, J, Huisman, HG, Miedema, F (1989) Evidence for a role of virulent human immunodeficiency virus (HIV) variants in the pathogenesis of acquired immunodeficiency syndrome: studies on sequential HIV isolates. *Journal of Virology* **63**: 2118-2125.

Thompson, SR, Sarnow, P (2000) Regulation of host cell translation by viruses and effects on cell function. *Curr Opin Microbiol* **3**: 366-370.

Tovanabutra, S, Robison, V, Wongtrakul, J, Sennum, S, Suriyanon, V, Kingkeow, D, Kawichai, S, Tanan, P, Duerr, A, Nelson, KE (2002) Male viral load and heterosexual transmission of HIV-1 subtype E in northern Thailand. *J Acquir Immune Defic Syndr* **29**: 275-283.

Tsiodras, S, Mantzoros, C, Hammer, S, Samore, M (2000) Effects of protease inhibitors on hyperglycemia, hyperlipidemia, and lipodystrophy: a 5-year cohort study. *Arch Intern Med* **160**: 2050-2056.

UNAIDS (2012) Global Report 2012: UNAIDS Report on the Global AIDS Epidemic. ebookpartnership. com.

Vacca, JP, Dorsey, BD, Schleif, WA, Levin, RB, McDaniel, SL, Darke, PL, Zugay, J, Quintero, JC, Blahy, OM, Roth, E (1994) L-735,524: an orally bioavailable human immunodeficiency virus type 1 protease inhibitor. *Proceedings of the National Academy of Sciences* **91**: 4096-4100.

Vacca, JP, Guare, JP, deSolms, SJ, Sanders, WM, Giuliani, EA, Young, SD, Darke, PL, Zugay, J, Sigal, IS, Schleif, WA, et al. (1991) L-687,908, a potent hydroxyethylene-containing HIV protease inhibitor. *J Med Chem* **34**: 1225-1228.

Vallari, A, Holzmayer, V, Harris, B, Yamaguchi, J, Ngansop, C, Makamche, F, Mbanya, D, Kaptué, L, Ndembi, N, Gürtler, L, Devare, S, Brennan, CA (2011) Confirmation of Putative HIV-1 Group P in Cameroon. *Journal of Virology* **85**: 1403-1407.

van Leeuwen, R, Katlama, C, Kitchen, V, Boucher, CA, Tubiana, R, McBride, M, Ingrand, D, Weber, J, Hill, A, McDade, H, et al. (1995) Evaluation of safety and efficacy of 3TC (lamivudine) in patients with asymptomatic or mildly symptomatic human immunodeficiency virus infection: a phase I/II study. *J Infect Dis* **171**: 1166-1171.

Van Vaerenbergh, K (2001) Study of the impact of HIV genotypic drug resistance testing on therapy efficacy. *Verhandelingen - Koninklijke Academie voor Geneeskunde van België* **63**: 447-473.

VANDEN HAESEVELDE, MM, Peeters, M, JANNES, G, JANSSENS, W, VAN DER GROEN, G, SHARP, PM, SAMAN, E (1996) Sequence analysis of a highly divergent HIV-1-related lentivirus isolated from a wild captured chimpanzee. *Virology* **221**: 346–350.

von Schwedler, U, Song, J, Aiken, C, Trono, D (1993) Vif is crucial for human immunodeficiency virus type 1 proviral DNA synthesis in infected cells. *Journal of virology* **67**: 4945-4955.

Wain-Hobson, S, Sonigo, P, Danos, O, Cole, S, Alizon, M (1985) Nucleotide sequence of the AIDS virus, LAV. *Cell* **40**: 9–17.

Wang, C, Mitsuya, Y, Gharizadeh, B, Ronaghi, M, Shafer, RW (2007) Characterization of mutation spectra with ultra-deep pyrosequencing: application to HIV-1 drug resistance. *Genome Res* **17**: 1195-1201.

Wang, C-T, Barklis, E (1993) Assembly, processing, and infectivity of human immunodeficiency virus type 1 gag mutants. *Journal of virology* **67**: 4264–4273.

Wang, R-R, Yang, L-M, Wang, Y-H, Pang, W, Tam, S-C, Tien, P, Zheng, Y-T (2009) Sifuvirtide, a potent HIV fusion inhibitor peptide. *Biochemical and Biophysical Research Communications* **382**: 540-544.

Wasi, C, Herring, B, Raktham, S, Vanichseni, S, Mastro, TD, Young, NL, Rubsamen-Waigmann, H, von Briesen, H, Kalish, ML, Luo, CC, et al. (1995) Determination of HIV-1 subtypes in injecting drug users in Bangkok, Thailand, using peptide-binding enzyme immunoassay and heteroduplex mobility assay: evidence of increasing infection with HIV-1 subtype E. *AIDS* **9**: 843-849.

Weinberg, JL, Kovarik, CL (2010) The WHO Clinical Staging System for HIV/AIDS. *Virtual Mentor* **12**: 202-206.

Wensing, AMJ, Vijver, DAvd, Angarano, G, Åsjö, B, Balotta, C, Boeri, E, Camacho, R, Chaix, M-L, Costagliola, D, Luca, AD, Derdelinckx, I, Grossman, Z, Hamouda, O, Hatzakis, A, Hemmer, R, Hoepelman, A, Horban, A, Korn, K, Kücherer, C, Leitner, T, Loveday, C, MacRae, E, Maljkovic, I, Mendoza, Cd, Meyer, L, Nielsen, C, Coul, ELOd, Ormaasen, V, Paraskevis, D, Perrin, L, Puchhammer-Stöckl, E, Ruiz, L, Salminen, M, Schmit, J-C, Schneider, F, Schuurman, R, Soriano, V, Stanczak, G, Stanojevic, M, Vandamme, A-M, Laethem, KV, Violin, M, Wilbe, K, Yerly, S, Zazzi, M, Boucher, CA (2005) Prevalence of Drug-Resistant HIV-1 Variants in Untreated Individuals in Europe: Implications for Clinical Management. *Journal of Infectious Diseases* **192**: 958-966.

Westby, M, van der Ryst, E (2005) CCR5 antagonists: host-targeted antivirals for the treatment of HIV infection. *Antivir Chem Chemother* **16**: 339-354.

Wheeler, J, McHale, M, Jackson, V, Penny, M (2007) Assessing theoretical risk and benefit suggested by genetic association studies of CCR5: experience in a drug development programme for maraviroc. *Antiviral therapy* **12**: 233.

Wiegand, HL, Doehle, BP, Bogerd, HP, Cullen, BR (2004) A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins. *The EMBO Journal* **23**: 2451-2458.

Wiegers, K, Rutter, G, Kottler, H, Tessmer, U, Hohenberg, H, Krausslich, HG (1998) Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. *J Virol* **72**: 2846-2854.

Wild, C, Greenwell, T, Matthews, T (1993) A synthetic peptide from HIV-1 gp41 is a potent inhibitor of virus-mediated cell-cell fusion. *AIDS research and human retroviruses* **9**: 1051-1053.

Willey, RL, Maldarelli, F, Martin, MA, Strebel, K (1992) Human immunodeficiency virus type 1 Vpu protein induces rapid degradation of CD4. *Journal of Virology* **66**: 7193-7200.

Wofsy, C, Hauer, L, Michaelis, B, Cohen, J, Padian, N, Evans, L, Levy, J (1986) Isolation of AIDS-associated retrovirus from genital secretions of women with antibodies to the virus. *The Lancet* **327**: 527–529.

Worobey, M, Gemmel, M, Teuwen, DE, Haselkorn, T, Kunstman, K, Bunce, M, Muyembe, JJ, Kabongo, JM, Kalengayi, RM, Van Marck, E, Gilbert, MT, Wolinsky, SM (2008) Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* **455**: 661-664.

Wu, L, Gerard, NP, Wyatt, R, Choe, H, Parolin, C, Ruffing, N, Borsetti, A, Cardoso, AA, Desjardin, E, Newman, W, Gerard, C, Sodroski, J (1996) CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. *Nature* **384**: 179-183.

Wu, L, LaRosa, G, Kassam, N, Gordon, CJ, Heath, H, Ruffing, N, Chen, H, Humblias, J, Samson, M, Parmentier, M, Moore, JP, Mackay, CR (1997) Interaction of Chemokine Receptor CCR5 with its Ligands: Multiple Domains for HIV-1 gp120 Binding and a Single Domain for Chemokine Binding. *The Journal of Experimental Medicine* **186**: 1373-1381.

Yang, YL, Wang, G, Dorman, K, Kaplan, AH (1996) Long polymerase chain reaction amplification of heterogeneous HIV type 1 templates produces recombination at a relatively high frequency. *AIDS Res Hum Retroviruses* **12**: 303-306.

Yerly, S, Kaiser, L, Race, E, Bru, J-P, Clavel, F, Perrin, L (1999) Transmission of antiretroviral-drug-resistant HIV-1 variants. *The Lancet* **354**: 729.

Zaidi, J, Grapsa, E, Tanser, F, Newell, M-L, Bärnighausen, T Dramatic increases in HIV prevalence after scale-up of antiretroviral treatment: a longitudinal population-based HIV surveillance study in rural kwazulu-natal. *AIDS* **27**: 000-000.

Zennou, V, Petit, C, Guetard, D, Nerhbass, U, Montagnier, L, Charneau, P (2000) HIV-1 Genome Nuclear Import Is Mediated by a Central DNA Flap. *Cell* **101**: 173-185.

Zhang, H, Yang, B, Pomerantz, RJ, Zhang, C, Arunachalam, SC, Gao, L (2003) The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* **424**: 94-98.

Zhang, M, Foley, B, Schultz, AK, Macke, JP, Bulla, I, Stanke, M, Morgenstern, B, Korber, B, Leitner, T (2010) The role of recombination in the emergence of a complex and dynamic HIV epidemic. *Retrovirology* **7**: 25.

Zhang, Y, Qian, H, Love, Z, Barklis, E (1998) Analysis of the Assembly Function of the Human Immunodeficiency Virus Type 1 Gag Protein Nucleocapsid Domain. *Journal of Virology* **72**: 1782-1789.

Zhou, W, Resh, MD (1996) Differential membrane binding of the human immunodeficiency virus type 1 matrix protein. *J Virol* **70**: 8540-8548.

Zhu, T, Korber, BT, Nahmias, AJ, Hooper, E, Sharp, PM, Ho, DD (1998) An African HIV-1 sequence from 1959 and implications for the origin of the epidemic. *Nature* **391**: 594-597.

Ziegler, J, Johnson, R, Cooper, D, Gold, J (1985) Postnatal transmission of AIDS-associated retrovirus from mother to infant. *The Lancet* **325**: 896–898.

Zolopa, AR, Shafer, RW, Warford, A, Montoya, JG, Hsu, P, Katzenstein, D, Merigan, TC, Efron, B (1999) HIV-1 genotypic resistance patterns predict response to saquinavir-ritonavir therapy in patients in whom previous protease inhibitor therapy had failed. *Ann Intern Med* **131**: 813-821.