

Review

Retroviral Vectors for Analysis of Viral Mutagenesis and Recombination

Jonathan M.O. Rawson^{1,2} and Louis M. Mansky^{1,2,3,4,*}

¹ Institute for Molecular Virology, University of Minnesota, Moos Tower 18-242, 515 Delaware St SE, Minneapolis, MN 55455, USA; E-Mail: rawso018@umn.edu

² Molecular, Cellular, Developmental Biology & Genetics Graduate Program, University of Minnesota, 6-160 Jackson Hall, 321 Church St SE, Minneapolis, MN 55455, USA

³ Department of Diagnostic and Biological Sciences, School of Dentistry, University of Minnesota, 515 Delaware St SE, Minneapolis, MN 55455, USA

⁴ Department of Microbiology, University of Minnesota, 420 Delaware St SE, Minneapolis, MN 55455, USA

* Author to whom correspondence should be addressed; E-Mail: mansky@umn.edu; Tel.: +1-612-626-5525; Fax: +1-612-626-5515.

Received: 15 August 2014; in revised form: 15 September 2014 / Accepted: 17 September 2014 / Published: 24 September 2014

Abstract: Retrovirus population diversity within infected hosts is commonly high due in part to elevated rates of replication, mutation, and recombination. This high genetic diversity often complicates the development of effective diagnostics, vaccines, and antiviral drugs. This review highlights the diverse vectors and approaches that have been used to examine mutation and recombination in retroviruses. Retroviral vectors for these purposes can broadly be divided into two categories: those that utilize reporter genes as mutation or recombination targets and those that utilize viral genes as targets of mutation or recombination. Reporter gene vectors greatly facilitate the detection, quantification, and characterization of mutants and/or recombinants, but may not fully recapitulate the patterns of mutagenesis or recombination observed in native viral gene sequences. In contrast, the detection of mutations or recombination events directly in viral genes is more biologically relevant but also typically more challenging and inefficient. We will highlight the advantages and disadvantages of the various vectors and approaches used as well as propose ways in which they could be improved.

Keywords: retrovirus; lentivirus; reverse transcription; evolution; mutation; recombination; retroviral vector

1. Introduction

Retroviruses share a remarkable capacity to rapidly evolve, which can be attributed to several key variables: high rates of mutation, recombination, and replication, large numbers of infected cells, and strong positive selective pressures [1–3]. These factors act together to drive the expansion of a small number of viruses that are initially transmitted to a particular host into a population of diverse but interacting variants, termed the viral quasispecies. Genetic diversification enables escape from the host immune response, accelerates the emergence of drug resistance, and promotes cross-species transmission. The ability of human retroviruses like human immunodeficiency virus type-1 (HIV-1) to rapidly evolve poses a tremendous challenge to the development of effective prophylactics and therapeutics.

Retroviruses mutate at an average rate of $\sim 3 \times 10^{-5}$ mutations/base pair (bp)/cycle, corresponding to roughly one mutation per three genomes synthesized [4–8]. This mutation rate is ~ 10 –1000 fold higher than DNA viruses (10^{-6} – 10^{-8} mutations/bp) [9] and at least 10,000-fold higher than for eukaryotic DNA ($\leq 10^{-9}$ mutations/bp) [10]. Mutations could arise during several different steps of the retroviral life cycle. Reverse transcriptase (RT) converts the single-stranded viral RNA into double-stranded DNA and is thought to be one of the key drivers of viral mutagenesis, primarily due to its high error rates (typically 10^{-4} – 10^{-5} mutations/bp) [11] *in vitro*. However, RNA polymerase II (Pol II) can also generate mutations when transcribing the viral genomic RNA from the integrated proviral DNA [12], though the relative contribution of Pol II compared to RT has not been well defined. Cellular DNA polymerases can also create mutations when replicating the integrated provirus during cell division, but the high fidelity of cellular DNA replication argues that this is a relatively minor source of virus variation. Lastly, nucleic acid-editing enzymes such as the APOBEC3 family of cytosine deaminases can edit minus strand viral DNA during reverse transcription [13,14], ultimately leading to G-to-A mutations on the plus strand viral DNA. APOBEC3-mediated editing is often lethal, as multiple G-to-A mutations are usually induced in the same provirus (*i.e.*, G-to-A hypermutation), but editing has been reported to accelerate viral evolution in certain contexts [15,16]. Furthermore, while HIV-1 expresses an accessory protein called Vif to counteract APOBEC3 proteins, Vif alleles have been shown to vary widely in their abilities to counteract various APOBEC3 proteins [17–19]. In addition, G-to-A hypermutants have often been observed in patient samples [17,20–26], which provides strong evidence that APOBEC3-mediated editing occurs *in vivo*. All of these processes together contribute to the high mutation rates of retroviruses, leading to the evolution of variants that may confer drug resistance, improve transmissibility, or allow for cytotoxic T-cell or neutralizing antibody escape. Many additional variables have been shown to influence retroviral mutagenesis, including RT variants [27–31], antiviral drugs [27,28,30,32], and accessory proteins [28,33–35].

In addition to mutating at high rates, retroviruses are also able to recombine at high rates (e.g., ~ 3 –14 crossovers/genome/cycle for HIV-1) [36–42]. Retroviruses are able to recombine rapidly because (unlike other RNA viruses) they co-package two copies of the RNA genome into every viral

particle. During reverse transcription, recombination between RNA genomes often occurs but only one DNA provirus is ultimately formed, such that retroviruses are considered pseudodiploid in nature [43]. Recombination events between identical co-packaged genomes are effectively silent since new viral sequences are not generated. In contrast, recombination events between distinct genomes (*i.e.*, from heterozygous virions) can lead to novel viral variants. Recombination permits the completion of DNA synthesis even in the presence of RNA damage and also promotes the generation of novel variants through the shuffling of viral mutations. For example, recombination can link beneficial mutations together, such as low-level drug resistance mutations into a highly drug-resistant complex [44] or single-drug resistance mutations into multi-drug resistance [45]. Conversely, recombination can allow escape from deleterious or lethal mutations [46–48]. While often beneficial, recombination can also dissociate co-adapted mutations, leading to unfit viral variants [49]. The importance of recombination to the establishment of the global AIDS pandemic is clear: 65 circulating recombinant forms (CRFs) have been identified to date [50], accounting for ~20% of HIV-1 infections worldwide [51].

As a pre-requisite for recombination to occur between distinct viruses, the producer cells must become dually infected, either as a result of co-infection (near simultaneous infection) or super-infection (sequential infection). Thus, factors that alter the incidence of dual infection through either mechanism will impact the frequency of recombination. HIV-1 readily co-infects primary CD4⁺ T-cells *ex vivo* with little evidence of interference [52–55]. In fact, co-infection has been found to occur more frequently than expected from random interactions between viruses and cells [52,53], but this may be due to reactivation of silent proviruses upon co-infection [56]. In contrast to co-infection, cells infected by HIV-1 (and many other retroviruses) are resistant to re-infection, a phenomenon called super-infection resistance [57]. Super-infection of humans at the organismal level has often been documented [58–63], though initial infection may be somewhat protective of re-infection [62]. However, such individuals do not necessarily contain super-infected cells, as only a small fraction of CD4⁺ T-cells becomes infected. Nonetheless, the wide prevalence of intra-subtype, inter-subtype, and inter-group recombinants of HIV-1 demonstrates that cellular super-infection resistance is not absolute. Of note, dually infected splenocytes can readily be observed in samples from HIV-1 infected individuals [64,65] or SIV_{mac}-infected rhesus macaques [66]. However, another study found that most peripheral blood CD4⁺ T-cells harbor only a single provirus [67]. Unfortunately, it is not yet clear whether these discrepancies are due to the different compartments sampled, unique features of the infected individuals, or other factors.

In addition to the incidence of dual infection, rates of retroviral recombination are influenced by factors that alter the ability of genetically distinct genomes to co-package into the same virus particle. For example, the HIV-1 genome contains the dimerization initiation signal (DIS), a six nucleotide palindromic sequence within the first stem loop of the 5' untranslated region. The DIS sequence is the dominant factor that drives HIV-1 genomic RNA co-dimerization and co-packaging [68–71]. Thus, HIV-1 variants with matched DIS sequences ultimately recombine much more frequently than those with mismatched DIS sequences. Additionally, the rate of recombination depends on the frequency of template switching during reverse transcription. Template switching is thought to be controlled by the relative balance of polymerase and RNase H activities of RT, referred to as the dynamic copy choice model of recombination [72,73]. Factors that reduce the speed of RT-mediated DNA synthesis promote template

switching, including low dNTP pool levels [72,74–76], RNA secondary structures [77–79], and mutations in RT that impair processivity [72,75,76]. In contrast, RT mutations that reduce RNase H activity decrease the level of template switching [72,75,80]. The level of template switching also correlates fairly well with the degree of sequence homology between templates, such that closely related sequences (e.g., from the same subtype of HIV-1) recombine more frequently than divergent sequences (e.g., from different subtypes of HIV-1) [71,81].

In developed countries, the advent of highly active antiretroviral therapy (HAART) directed against HIV-1 has enabled the indefinite suppression of viral replication in the vast majority (>95%) of infected individuals, provided proper drug adherence is maintained. This raises the issue of the settings in which the high mutation and recombination rates of HIV-1 remain relevant. Prior to the initiation of treatment, rapid viral replication and diversification permit escape from CD8⁺ cytotoxic T-cell and neutralizing antibody responses, ultimately preventing effective immune system control of the virus. Defining the roles of viral mutation and recombination in these processes may inform efforts to develop an effective vaccine. Further, during HAART, a low level of ongoing viral replication may persist in areas of limited drug penetration, such as the gut-associated lymphoid tissue and central nervous system [82,83], permitting continual viral evolution. However, the presence of ongoing viral replication during HAART has remained controversial, and the predominant mechanism of viral persistence is thought to be latent infection of resting CD4⁺ T-cells [84]. HAART is also not completely suppressive in all individuals, as ~3% of infected individuals develop triple-class virological failure [85], though this may result from suboptimal drug adherence. Additionally, high rates of mutation and recombination are features often shared by other RNA viruses [9,86–88], many of which cannot be effectively countered by drug treatment. The investigation of these processes in HIV-1 will serve as a useful model system for studying the genetic diversification of other RNA viruses.

Although many advances have been made, there are still a number of fundamental questions surrounding the nature of retroviral mutagenesis and recombination that remain unanswered. For example, very few studies have examined these processes in primary cells rather than immortalized cell lines. Differences in dNTP pool levels and expression of relevant cellular factors between immortalized cell lines and primary cells could have a large impact on viral mutagenesis and recombination. Even primary cell types can differ substantially in this regard, as the most relevant targets of HIV-1 infection, activated CD4⁺ T-cells and macrophages, differ by ~130–250-fold in dNTP pool levels [89]. In addition, very few studies have examined mutagenesis or recombination in native viral genes rather than foreign reporter genes, primarily due to the increased difficulty of detecting such events. Given this, there remains a clear need for continued dissection of these processes in retroviruses, and improved vectors and methodologies to aid in these investigations.

Many different vectors and approaches have been used to examine retroviral mutagenesis and recombination. Retroviral vectors can broadly be divided into two categories: those that utilize reporter genes (e.g., antibiotic resistance genes, β -galactosidase, or fluorescent proteins) and those that utilize viral genes as targets for mutation or recombination. Reporter gene vectors greatly facilitate the detection, quantification, and characterization of mutants or recombinants, but the results may not be representative of native viral genes in regards to the frequency or types of mutation and/or recombination hotspots. In contrast, the detection of mutations or recombination events directly in viral genes is more biologically relevant, but the methods to do so are also much less efficient. Most of the

retroviral vectors utilized for studies of viral mutagenesis and recombination are not replication-competent and must be trans-complemented in order to produce infectious vector virus. The benefit of this approach is that it allows for the more accurate quantification of mutation and recombination rates, as all events must occur in a single cycle of replication. In other instances, replication-competent vectors have been used to examine retroviral mutagenesis and recombination, but they have been limited to a single round of replication by neutralizing antibodies or antiviral drugs. In contrast, multi-cycle replication assays are valuable for understanding how natural selection acts upon emerging mutants and recombinants, and some of these studies will briefly be addressed. Here, the design and application of a number of different retroviral vectors that have been employed in assays to investigate retroviral mutagenesis and recombination is reviewed. The benefits and drawbacks of each approach are highlighted and potential ways to improve current vectors and methodologies are explored.

2. Vectors for Examination of Retroviral Mutagenesis

2.1. Reporter Gene Vectors for Reversion Assays

Some of the earliest vectors for quantification of retroviral mutation rates relied upon the detection of mutations that restored defective drug resistance genes (see Table 1). Many of these initial studies utilized spleen necrosis virus (SNV)-based vectors. SNV is an avian gammaretrovirus capable of infecting several types of mammalian cells and has often been employed as a model system for retrovirus biology. The first mutation rate of a retrovirus was determined using a SNV-based vector containing a dysfunctional neomycin resistance gene (*neo*) and a functional hygromycin B resistance gene (*hph*), as shown in Figure 1A [90]. The neomycin resistance gene had a single point mutation introducing a premature stop codon, and *neo* revertants could be selected for using the neomycin analog G-418. The reversion mutation frequency of SNV was determined by calculating the ratio of G-418-resistant (*neo*^r) to hygromycin-resistant (*hygro*^r) colonies and was found to be 2.2×10^{-5} . In this particular assay, the reversion mutation frequency was equivalent to the reversion mutation rate (defined as mutations/bp/cycle), as the size of the mutational target was a single base and the vector was limited to one round of replication.

Viral mutagenesis has been investigated in several other retroviruses using reversion vectors as well. The mutation rate of murine leukemia virus (MLV), another gammaretrovirus, has been determined using a *neo/hph* cassette similar to the one used for SNV [91]. More recently, the mutagenesis of HIV-1 has been investigated using a defective luciferase gene containing an insertion of eight T residues near the 5' end of the gene [30]. This vector was designed specifically to detect frameshift events, which are often triggered by homopolymeric runs [92–95]. Relative mutation frequencies were determined by comparing ratios of relative light units to the viral titers, which were determined using an integrated LTR-driven marker gene in the same cell line. In addition to luciferase, vectors containing an inactivated blasticidin resistance gene (*bsr*) and enhanced yellow fluorescent protein (EYFP) have been used to measure mutation rates of several retroviruses, including avian leukosis virus (ALV), MLV, and HIV-1 [96]. In these assays, reversion frequencies were defined as the ratio of blasticidin-resistant/EYFP⁺ cells to all EYFP⁺ cells.

Table 1. Targets for analysis of retroviral mutagenesis.

Target ¹	Assay Type ²	Virus	References
<i>neo</i>	Reporter (R)	SNV, MLV	[90,91]
<i>tk</i>	Reporter (F)	MLV, HIV-1	[8,97,98]
<i>lacZa/lacZ</i>	Reporter (F)	SNV, BLV, HTLV-1, HIV-1	[4–7,27–29,31–35,99–102]
<i>luc</i>	Reporter (R)	HIV-1	[30]
<i>bsr</i>	Reporter (R)	ALV, MLV, HIV-1	[96]
<i>hsa</i>	Reporter (F)	HIV-1	[103,104]
<i>egfp</i>	Reporter (F)	HIV-1	[105–108]
Viral	HTA	RSV	[109]
	RNase T ₁	MLV	[110]
	SSCP	HIV-1	[12]
	Sequencing (Sanger)	HIV-1	[111,112]
	Sequencing (NGS)	HIV-1	[40]

¹ Reporter gene targets include antibiotic resistance genes (*neo*, *bsr*), thymidine kinase (*tk*), β -galactosidase (*lacZa* or full-length *lacZ*), the cell surface marker heat stable antigen (*hsa*), and enhanced green fluorescent protein (*egfp*); ²Reporter gene targets are included for both reversion assays (R) as well as forward mutational assays (F). Assays that can detect mutations directly in viral genes include heteroduplex tracking (HTA), RNase T₁ fingerprinting, single-strand conformation polymorphisms (SSCP), Sanger sequencing, and next-generation sequencing (NGS) technologies.

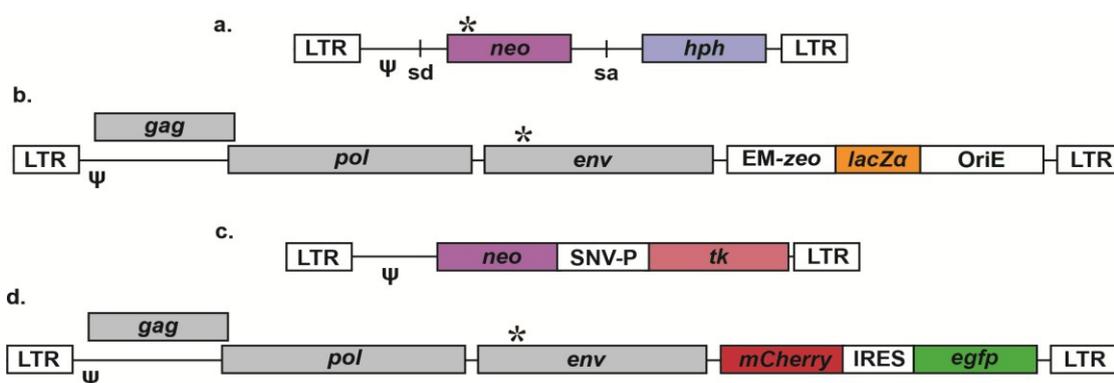


Figure 1. Vectors for the examination of retroviral mutagenesis. Most accessory genes and cis elements are not shown for simplicity; * represents inactivating mutations, while ψ represents packaging signals. **(a)** SNV vector to detect mutations that restore *neo*, conferring G-418 resistance [90]; sd = splice donor, sa = splice acceptor. **(b)** HIV-1 vector that detects mutations in *lacZa* that prevent complementation and expression of functional β -galactosidase in *E. coli*, ultimately resulting in white bacterial colonies in the presence of X-gal [101]; EM = bacterial promoter, *zeo* = zeomycin resistance gene, OriE = *E. coli* origin of replication. **(c)** MLV vector to monitor mutations that eliminate thymidine kinase (*tk*), granting resistance to the thymidine analog bromodeoxyuridine (BrdU) [8]; SNV-P = SNV promoter (U3). **(d)** HIV-1 vector to detect mutations that eliminate expression of mCherry or enhanced green fluorescent protein (EGFP) by flow cytometry [107]; IRES = internal ribosome entry site.

Although used infrequently, reversion reporter vectors do greatly facilitate the detection and quantification of viral mutants, and sequencing can determine the nature of reverting mutations. However, these vectors can only detect the few specific mutations that are able to restore a functional gene phenotype. Thus, many types of mutations cannot be detected at all, and those mutations that can be observed are limited to very narrow sequence contexts. Therefore, calculated mutation rates may not be representative of the overall mutation rate within the reporter gene or virus. In addition, the influence of factors on the mutation spectra cannot readily be addressed. For these reasons, vectors that instead rely on the elimination of functional reporter genes (*i.e.*, forward mutation assays) have been utilized much more frequently than vectors that measure reversion mutation frequencies.

2.2. Reporter Gene Vectors for Forward Mutational Assays

2.2.1. LacZ α /LacZ (β -galactosidase)

Rather than monitoring the restoration of defective reporter genes, forward mutational assays detect the elimination of functional reporter gene products. Thus, these assays can detect every mutational class, including the twelve types of substitutions, insertions, deletions, and more complex events. Furthermore, mutations can be detected in a variety of sequence contexts. Though many different reporter genes have been used in these assays (see Table 1), the lacZ α peptide gene (or, in some cases, the full-length lacZ) is one of the most widely used reporters for analyses of retroviral mutagenesis. The lacZ α gene encodes a small fragment (~150–170 bp) of β -galactosidase that can interact with another fragment (lacZ ω) to reconstitute full-length β -galactosidase. The formation of functional β -galactosidase can be detected by blue-white color screening in *E. coli*. This system has been used to determine forward mutation rates and spectra for a number of different retroviruses (see Table 1). Furthermore, this system has been used to characterize the impact of factors such as drugs, drug resistance-associated mutations, and viral accessory proteins on viral mutagenesis.

While specific details vary between laboratories, the general steps of the lacZ α assay are as follows: First, producer cells are established that contain an integrated packaging construct with a shuttle cassette. The shuttle cassette often consists of lacZ α , an antibiotic resistance gene (for selection of producer and target cells), and a bacterial origin of replication (see Figure 1B). Virus production is initiated by transient transfection with helper plasmids, and the collected virus is used to infect target cells. The virus can only replicate a single time in the target cells due to the lack of essential viral genes. Next, the lacZ α DNA is recovered, usually through a process that avoids the use of PCR, as PCR can generate background errors. For example, the lacZ α DNA can be recovered by purification with the Lac repressor protein [4] or by Hirt extraction [101], which recovers extrachromosomal DNA. For Hirt extraction, vectors have been designed with inactivating mutations in integrase to facilitate the recovery of lacZ α DNAs. The DNA can then be transformed into bacteria, selected with antibiotics, and subjected to blue-white color screening using the chromogenic substrate X-gal. Mutation frequencies are calculated by dividing the number of mutant colonies (white or light blue) by the total number of colonies. Forward mutation rates can be calculated by dividing the mutation frequencies by the size of lacZ α or by the known number of detectable sites [113,114] for mutations in lacZ α (this is possible because lacZ α has been extensively mutated). Mutation

spectra can be determined by targeted sequencing of white colonies, leading to identification of a wide array of mutations.

2.2.2. Thymidine Kinase (TK)

In addition to *lacZ α* , the herpes simplex virus thymidine kinase (*tk*) gene has been used in forward reporter vectors in analyses of retroviral mutagenesis (Table 1). Thymidine kinase is a nucleotide salvage enzyme that converts thymidine into thymidine monophosphate (TMP). Cells expressing thymidine kinase are resistant to inhibitors of *de novo* TMP synthesis (e.g., aminopterin) but sensitive to the thymidine analog bromodeoxyuridine (BrdU). These vectors contain both *tk* as well as an antibiotic resistance gene (*hygro* or *neo*) for determination of viral titers (see Figure 1C). To use these vectors, producer cells are established, virus production is initiated by transfection of helper plasmids, and TK⁻ target cells are infected, such that the viral vector is the only source of TK. Forward mutation frequencies can be calculated by dividing the number of antibiotic-resistant (*hygro*^r or *neo*^r) and BrdU-resistant (BrdU^r) colonies by all antibiotic-resistant colonies. Forward mutation rates can be calculated by dividing mutation frequencies by the size of *tk* (~1.2 kb), and spectra can be examined by PCR and sequencing. Relative to *lacZ α* , *tk* vectors offer several advantages, such as the ability to propagate and further characterize mutant proviruses that survive drug selection. As *tk* is substantially longer than *lacZ α* (~1.2 kb and ~0.2 kb, respectively), the general likelihood of observing mutations is greater. However, the increased length of *tk* also necessitates that more sequencing be performed to identify all mutations within the reporter gene. In addition, the use of PCR to recover *tk* sequences results in some level of background errors. Further, many types of cells cannot be used as target cells of infection in this assay because they express endogenous TK.

2.2.3. Fluorescent Proteins and Cell Surface Markers

HIV-1 mutagenesis has been investigated using several single-cycle reporter vectors that encode fluorescent proteins and/or cell surface markers, which can be stained by antibodies conjugated to fluorescent dyes. To construct the first vector [105], the mouse heat stable antigen (*hsa*, a cell surface marker) and enhanced green fluorescent protein (*egfp*) were inserted in place of *nef* within an HIV-1 envelope-deficient molecular clone. The reporter genes were separated by an internal ribosome entry site (IRES) element necessary for translation of EGFP. More recently, the *hsa* gene has been replaced with mCherry (Figure 1D), eliminating the need for antibody staining [107]. In addition, a similar mCherry/*egfp* vector has been constructed for mutagenesis studies in human immunodeficiency virus type-2 (HIV-2) [115]. In order to determine mutant frequencies, these vectors are co-transfected with an envelope plasmid into 293T cells. Viral stocks are collected and used to infect target cells, and flow cytometry is performed to quantify the number of infected cells (cells expressing at least one reporter) as well as the number of cells with mutant proviruses (cells expressing only one reporter). Mutant frequencies are calculated by dividing the number of mutants by the total number of infected cells. To assess mutational spectra, fluorescence-activated cell sorting (FACS) can be performed to isolate cells expressing only one reporter gene, from which genomic DNA can be purified. Lastly, the defective gene (e.g., *egfp* from mCherry⁺/EGFP⁻ cells) can be amplified by PCR, ligated into a cloning vector, and sequenced. To date, these vectors have been employed to examine the impact of small molecule

mutagens [105–108,112,115,116], drug resistance-associated mutations [117], cell type [104], and APOBEC3G [103,108] on HIV-1 mutagenesis. The primary advantage of these vectors is that they allow for rapid determination of infectivities and mutant frequencies, and thus are ideally suited for the screening of small molecules for antiviral and/or mutagenic activities. However, cell sorting is required to enrich for mutant sequences prior to further processing, as the vast majority of sequences are wild-type in the absence of sorting. Further, PCR is required to amplify, clone, and sequence defective genes. Some mutations that eliminate expression of EGFP may occur within the IRES element necessary for EGFP translation, such that the entire IRES-EGFP region (~1.3 kb) must be sequenced to identify all responsible mutations.

2.3. Detection of Mutations in Viral Genes

2.3.1. Biochemical and Sanger Sequencing-Based Approaches

Rather than detecting mutations in reporter genes, some studies have assessed viral mutagenesis by directly detecting mutations in viral genes (Table 1). As reporter genes are unnecessary for this approach, a wide variety of retroviral vectors can be used, provided that they can be limited to a single round of replication. In addition to sequencing, mutations can be identified in viral genes by biochemical assays, including heteroduplex tracking (HTA), RNase T1 fingerprinting, and detection of single-strand conformation polymorphisms (SSCP). In HTA [109], viral RNA is annealed to a radiolabeled probe and subjected to denaturing-gradient gel electrophoresis. Mismatches caused by viral mutations, particularly in low-melting regions, can alter the thermodynamic stability of the heteroduplex and lead to altered migration during electrophoresis. For RNase T1 fingerprinting [110], radiolabeled viral RNA is digested with RNase T1, which cleaves at G residues in single-stranded RNA. The digested viral RNA is subjected to two-dimensional electrophoresis-homochromatography, generating a map, or fingerprint, of RNase T1-resistant oligonucleotides. Mutations within the larger RNase T1-resistant oligonucleotides often alter migration, allowing for their detection. Lastly, in SSCP [12], PCR of viral genes is performed from infected cell clones. The PCR products are radiolabeled, denatured, and analyzed by non-denaturing gel electrophoresis. Again, mutations are detected by altered mobility during electrophoresis. The primary drawback of these biochemical approaches is that they are labor-intensive, and are thus not conducive to identifying and characterizing large numbers of viral mutants. Additionally, they cannot detect all mutations (RNase T1 fingerprinting and SSCP have been estimated to detect ~80% of changes) [12,110], and thus calculated mutation rates underestimate true mutation rates.

In place of biochemical assays, other groups have simply amplified viral genes by PCR and identified mutations by Sanger sequencing [111,112]. In this instance, nearly all mutations can be identified, provided that they do not prevent amplification. Unfortunately, like biochemical assays, this approach is extremely low-throughput. For example, assuming a mutation rate of 3×10^{-5} mutations/bp/cycle (the retroviral average) and a target size of ~800 bp (the length of a typical Sanger read), ~0.024 mutations will be identified per read. Thus, in order to obtain 50 mutations, ~2000 Sanger reads would be required. Accordingly, studies following this approach have reported very low numbers of mutations, which hinders a robust determination of mutation rates and spectra. Additionally, without

enriching for mutated sequences, the presence of background errors due to PCR and sequencing becomes a much greater concern, and the necessary controls to account for this must be included. To circumvent the inefficiency of Sanger sequencing, some groups have sequenced virus passaged multiple (~5–10) times [118,119], but mutation rates cannot easily be calculated in these experiments due to purifying selection and uncertainty regarding the number of rounds of replication.

2.3.2. Next-Generation Sequencing-Based Approaches

The advent of next-generation sequencing (NGS) technologies has revolutionized biology, and retrovirology has been no exception. NGS has been used to infer co-receptor usage [120,121], examine immune escape variants [122], and identify minority variants [123–126] that may contribute to drug resistance in samples from HIV-1-infected individuals. Illumina platforms currently dominate the market due to their high outputs and low costs. Illumina can presently generate ~45–50 million 2×300 bp paired-end reads (MiSeq) or ~600 million 2×150 bp paired-end reads (HiSeq 2500 in rapid mode). While NGS can identify thousands of variants, NGS technologies are currently hindered by high error rates, due to both PCR and sequencing. Although sequencing libraries can be created in the absence of amplification, PCR is often employed to generate more material for sequencing and to enrich for adapter-ligated templates. The background error rates from NGS have been estimated to range from 10^{-2} mutations/bp for unprocessed data to 10^{-3} or 10^{-4} mutations/bp for processed data [127–131]. Given this, the average retroviral mutation rate is minimally ~10 fold lower than the background error rate, which prevents distinction of biological mutations from background errors. This problem is illustrated by a recent study [40] in which HIV-1 was allowed to replicate for a single cycle in primary CD4⁺ T-cells, amplified by PCR, and subjected to 454 sequencing. The background error rate was determined by amplifying and sequencing plasmid DNA and found to be 7.4×10^{-5} mutations/bp, while the error rate in the biological sample was found to be 1.2×10^{-4} mutations/bp. Thus, assuming equivalent background levels, ~60% of the mutations were artifactual, and these mutations could not be distinguished from those induced during the viral life cycle. Nonetheless, the mutation rate of HIV-1 was estimated by subtracting the background error rate from the error rate of the biological sample, leading to an estimate of $\sim 4.6 \times 10^{-5}$ mutations/bp/cycle. Recently, two key strategies have been developed to help circumvent the high error rates of NGS. In the first strategy, which has been referred to as low-error amplicon sequencing (LEA-Seq) [132], unique identifier tags (UIDs), such as twelve random bases, are assigned to each template molecule, either through ligation, reverse transcription, or linear PCR [125,132–135]. Next, exponential PCR of the tagged molecules is performed, followed by redundant sequencing. Consensus families can be built from sequences with identical tags, allowing the identification and exclusion of most PCR and sequencing errors. LEA-Seq has been used for 16S rRNA sequencing in microbial metagenomics [132,134], identification of mutations in human mitochondrial and nuclear DNA [133,135], determination of DNA polymerase fidelity during PCR [135], and tracking minority variants in the HIV-1 protease during drug exposure [125]. In the second strategy, called circular resequencing (CirSeq), rolling circle amplification of template molecules is used to generate three copies from the same starting template molecule that are linked in tandem [136–138]. After the linked copies are sequenced, consensus families are again built, permitting the removal of most background errors. CirSeq was recently used to

determine the mutation rate and spectrum of poliovirus, as well as to track the fitness of poliovirus variants across serially passaged populations [138]. While neither method has yet been used to determine mutation rates or spectra of retroviruses, both methods are likely to advance our understanding of how mutations arise in viral genes during retroviral replication.

3. Vectors for Examination of Retroviral Recombination

3.1. Reporter Gene Vectors

3.1.1. Antibiotic Resistance Genes

Similar to the initial vectors used to investigate retroviral mutation rates, the first vectors used to quantify a retroviral recombination rate relied upon a cassette containing two antibiotic resistance genes (*neo* and *hph*) in an SNV-derived vector (see Table 2 and Figure 2A) [43]. Two similar vectors were engineered: one with a dysfunctional *neo* gene, and the other with a dysfunctional *hph* gene. In both cases, the respective antibiotic resistance genes were disrupted by 4-bp insertions (~1 kb apart), which were found to spontaneously revert at low rates. To examine recombination, dually-infected producer cell clones were created by sequential infection with the parental viruses at low multiplicities of infection (MOI). PCR and enzyme digestion were performed to ensure that producer cell clones contained a single copy of each parental virus and that recombination did not occur during the generation of producer cell clones. While more laborious than simply co-transfecting viral vectors, this procedure eliminates transfection-induced recombination and more closely mimics natural viral replication. Viruses were collected from producer cell clones and used to infect target cells at low MOIs. The recombination rate was determined by dividing the titer of double drug-resistant (*neo^r/hygro^r*) colonies by the lower of the two single drug-resistant titers. Using this method, the recombination rate of SNV was found to be ~2%/kilobase/cycle [43]. Similar MLV-based vectors were later devised and used to show that recombination rates increase with the distance between genetic markers [139]. Additionally, MLV-based vectors have been used to characterize the effects of moving the dimerization linkage structure on recombination rates [140]. More recently, HIV-1-based *neo/hph* vectors were used to demonstrate that the HIV-1 recombination rate is ~10-fold higher than in gammaretroviruses [41]. Notably, experiments with these vectors must be performed at low MOIs, as co-infection can produce the same phenotype as a recombinant. Often, Southern blotting of recombinants is required to definitively demonstrate that double-resistant cell clones are not due to co-infection. In addition, this system is not easily amenable to other cell lines and primary cell types, as antibiotic sensitivities can vary widely between cell types.

Table 2. Targets for analysis of retroviral recombination.

Target ¹	Assay Type ²	Virus	References
<i>neo/hph</i>	Reporter Gene	SNV, MLV, HIV-1	[41,43,139–143]
<i>tk</i>	Reporter Gene	MLV	[144]
<i>lacZ</i>	Reporter Gene	MLV, HIV-1	[42]
<i>gfp</i>	Reporter Gene	MLV, HIV-1, HIV-2, SIV _{agm}	[38,49,68–71,145–147]

Table 2. Cont.

Target ¹	Assay Type ²	Virus	References
<i>ecfp/eyfp</i>	Reporter Gene	HIV-1	[55]
viral	Gag Reconstitution	HIV-1	[71]
	Modified LacZ	HIV-1	[77,148,149]
	PCR/qPCR	HIV-1	[148,150–152]
	HTA	HIV-1	[36,37,152]
	Restriction Mapping	HIV-1	[37,45]
	Sequencing (Sanger)	HIV-1	[37,39,45,151]
	Sequencing (NGS)	HIV-1	[40,153]

¹ Reporter gene targets include antibiotic resistance genes (*neo*, *hph*), thymidine kinase (*tk*), β -galactosidase (*lacZ*), and fluorescent proteins (*gfp*, *ecfp*, *eyfp*); ²Reporter gene assays typically detect recombination events that rescue defective gene products. Other assays have been used to detect recombination events in viral genes, such as a Gag reconstitution assay (in which Gag itself is treated as a reporter gene), a modified LacZ assay, PCR or quantitative PCR, heteroduplex tracking (HTA), restriction digest mapping, Sanger sequencing, and next-generation sequencing (NGS) technologies.

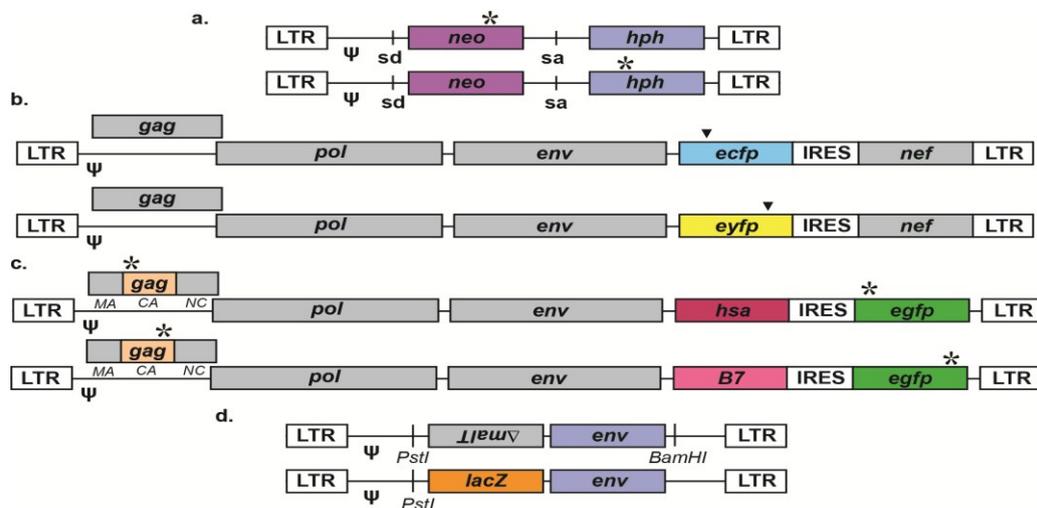


Figure 2. Vector pairs for the investigation of retroviral recombination. Most accessory genes and cis elements are not shown for simplicity; * represents inactivating mutations, while ψ represents packaging signals. **(a)** SNV vector pair that detects recombination events that restore *neo* and *hph*, resulting in viruses that confer G-418 and hygromycin B double drug-resistance to infected cells [43]; sd = splice donor, sa = splice acceptor. **(b)** HIV-1 vector pair that detects recombination events between *ecfp* and *eyfp*, which leads to expression of a modified GFP [55]; ∇ = essential mutations in *ecfp* and *eyfp*. **(c)** HIV-1 vector pair that can simultaneously detect recombination events that restore the capsid (CA) portion of the Gag polyprotein (by staining with an appropriate antibody) and enhanced green fluorescent protein (EGFP) [71]. Viral infection is monitored with cell surface markers like the mouse heat stable antigen (HSA) and B7; IRES = internal ribosome entry site. **(d)** HIV-1 vector pair that detects recombination events that occur within a region of homology (such as the viral env) inserted between β -galactosidase (*lacZ*) and a BamHI restriction site [77].

3.1.2. Fluorescent Proteins

Retroviral recombination rates have often been measured with fluorescent proteins (Table 2), eliminating the need for antibiotic selection of target cells. Importantly, these vectors facilitate the analysis of recombination in a variety of cell types (including primary cells), as the antibiotic sensitivities of different cell lines do not need to be considered. In one vector system that has often been used for HIV-1, recombination rates are measured using combinations of vectors with distinct inactivating mutations in *gfp* [38]. Specifically, these vectors contain cassettes encoding a cell surface marker (HSA or Thy-1) used to track infectivity, an IRES element, and an inactivated GFP as the recombination target. The expression of GFP can be eliminated by introducing mutations at various positions allowing for measurement of recombination at distances ranging from ~100 to 600 bp. To determine recombination rates, producer cell clones were first created by sequential infection at low MOIs (0.05–0.1). Cell sorting of infected cells was used to purify dually infected producer cells, which express both cell surface markers (HSA and Thy-1). Virus production from producer cells was initiated by the transfection of helper plasmids, and collected viruses were used to infect target cells. Recombination rates in target cells were calculated by comparing the frequency of GFP⁺ cells to all infected cells (HSA⁺ and/or Thy-1⁺). As the experiments were performed at relatively high MOIs (0.4–0.5), some cells were very likely co-infected. In contrast to the dual-antibiotic resistance vectors described in Section 3.1.1., co-infection with the inactivated GFP vectors does not result in a recombinant phenotype. However, co-infection does influence the calculation of the total number of infected cells, which in turn could alter calculated recombination rates. Therefore, the numbers of infected cells were converted to MOIs based upon a Poisson distribution of co-infection in order to more accurately estimate recombination rates [38]. In some studies, further analyses were done to confirm recombination events in *gfp* and map recombination breakpoints in viral genes [70,147]. In order to accomplish this, single GFP⁺ cell clones were isolated by cell sorting and propagated independently in order to prevent PCR-mediated recombination between proviruses. PCR and restriction mapping were used to verify reconstitution of functional *gfp* genes, as the inactivating mutations in *gfp* were designed to introduce specific restriction sites. Sanger sequencing of PCR products was used to further confirm recombination in *gfp*. Additional PCR and Sanger sequencing of GFP⁺ cell clones has often been performed to identify crossovers in viral sequences as well. Of note, rather than isolating single cell clones, a modified type of PCR called single-genome sequencing (SGS) has also been used to prevent PCR-mediated recombination [49]. In SGS, templates (e.g., viral RNA or proviral DNA) are diluted such that <30% of reactions result in products. Under these conditions, the vast majority of reactions contain only a single template, thus preventing most PCR-mediated recombination between templates.

Initially, vectors based on inactivated GFP were used to examine the impact of distance, cell type (primary vs. immortalized T-cells), and accessory genes on the recombination rate of HIV-1 [38]. Similar vectors were later used to measure recombination rates in non-B subtypes of HIV-1 as well as rates of intersubtype or intergroup recombination [49,68–71]. Further, these vectors have been used to study the impact of cell type (T-cells, monocytes, and macrophages) [145] on recombination. In addition to HIV-1, analogous vectors have been engineered for HIV-2 and simian immunodeficiency virus from African green monkeys (SIV_{agm}), both of which were found to recombine at similar rates to

HIV-1 [146]. These vectors have also been used to demonstrate that HIV-1 and HIV-2 are able to co-package and recombine [147], albeit at very low rates (recombination frequency of 0.2%, ~35-fold lower than for intrasubtype recombination in HIV-1).

In addition to GFP, several other fluorescent proteins have been used in recombination vectors such as enhanced yellow and cyan fluorescent proteins (EYFP and ECFP). These vectors take advantage of the high level of homology between different fluorescent proteins. Specifically, infectious molecular clones of HIV-1 have been developed that encode EYFP or ECFP in place of the accessory protein Nef, with Nef expressed from a downstream IRES element (see Figure 2B) [55]. As the critical residues in *yfp* and *cfp* are ~400 bp apart, recombination between these residues gives rise to a modified *gfp* gene (called *gfp**) that can be spectrally separated from *yfp* with the appropriate filters. Thus, recombination frequencies were calculated by determining the ratio of GFP⁺ to EYFP or ECFP⁺ cells during flow cytometry. To map recombination breakpoints in *gfp**, single GFP⁺ cell clones were isolated by FACS, amplified by PCR, and sequenced. Sequencing was also performed to map crossovers in viral genes, which could easily be identified as the vectors were engineered in two distinct isolates of HIV-1. Although these vectors express all viral genes and are replication-competent, replication can still be limited to a single round of replication through use of an antiviral drug.

3.1.3. LacZ (β -galactosidase)

Recombination rates have occasionally been determined using other reporter genes such as the full-length *lacZ* gene encoding β -galactosidase [42]. For example, HIV-1 and MLV vector pairs have been designed in which one encodes a functional *lacZ* gene but no antibiotic resistance gene, while the other encodes a defective *lacZ* gene and a puromycin resistance gene (*puro*). Recombination rates were calculated by determining the ratio of *lacZ*⁺/puromycin-resistant (*puro*^r) colonies (*i.e.*, *puro*^r colonies that stain blue in the presence of X-gal) to all *puro*^r colonies. These vectors have been used to compare frequencies of recombination between MLV and HIV-1, as well as to investigate factors responsible for observed differences between these viruses.

3.1.4. Direct Repeat Reporter Vectors

Directly repeated sequences in retroviral genomes and vectors have long been known to delete at high frequencies during reverse transcription [154–157]. Deletions occur at high rates regardless of whether the repeats are in tandem or separated by additional sequences. In one of the earliest observations of this phenomenon, the *src* oncogene was found to be frequently lost during RSV replication due to flanking direct repeats [155,156]. Direct repeat deletions also contribute to the variable copy number of repeats that have been observed in the enhancer sequences of retroviral LTRs [158,159]. These deletions have been shown to occur primarily through intramolecular template switching events, rather than through homologous recombination between separate viral genomes [141]. Retroviral vectors have been developed that take advantage of the high efficiency of direct repeat deletion to assay the impact of various factors on template switching. These vectors have large (>100 bp) inactivating duplications in reporter genes, such as *tk*, *lacZ*, or *gfp*. Direct repeat deletion eliminates the duplicated sequence, thus restoring the reporter gene phenotype. In one early study, an MLV vector was engineered with a *tk/neo* cassette in which *tk* contained an internal 700 bp

duplication [144]. Deletion frequencies were first assessed by comparing the ratio of *tk*⁺ (*i.e.*, aminopterin-resistant)/*neo*^r colonies to all *neo*^r colonies. Southern blotting and densitometry were performed to determine deletion frequencies in a more quantitative fashion. In several other studies, MLV and HIV-1-based vectors were engineered in which *lacZ* reporter genes were inactivated by internal duplications [42,74,80,160,161]. Deletion frequencies were calculated by comparing titers of *lacZ*⁺/*puro*^r cell colonies to all *puro*^r colonies. Lastly, MLV and HIV-1 vectors with internal duplications (of 200–250 bp) in *gfp* have been developed [72,73,75]. With these vectors, direct repeat deletion leads to reconstitution of functional GFP, and deletion frequencies can be determined as the ratio of GFP⁺/antibiotic-resistant cell colonies to all antibiotic-resistant cell colonies. More recently, a similar HIV-1 vector was developed in which the antibiotic resistance gene was replaced with the cell surface marker *hsa*, facilitating use in primary cells [162]. As direct repeat deletion is very efficient during reverse transcription, the ability of various factors to impact deletion frequencies can readily be determined, even when such effects are relatively small. However, as discussed, direct repeat deletions primarily occur through intramolecular template switching events rather than intermolecular template switching events [141]. Thus, the outcomes observed in these studies may not necessarily be representative of retroviral recombination, which requires intermolecular template switching events between distinct viruses.

3.2. Detection of Recombinants in Viral Genes

3.2.1. Gag Reconstitution Assay

Recently, a new HIV-1 vector system has been engineered that is based on the reconstitution of functional Gag protein (Figure 2C) [71]. This vector system combines many of the best features of reporter assays with the increased biological relevance of viral genes. The concept behind the vector system was straightforward: Gag is expressed at high levels on the plasma membranes of infected target cells, such that Gag itself can be treated as a reporter gene with a reliable anti-Gag antibody. More specifically, multiple pairs of vectors were engineered in different subtypes and groups of HIV-1 with distinct inactivating mutations separated by ~300 bp in the capsid portion of the Gag polyprotein. In addition, the *gfp* genes were inactivated in these vectors, allowing for simultaneous detection of recombination in identical (*gfp*) and non-identical (*gag*) target sequences. Cell-surface markers such as HSA and B7 (*i.e.*, human CD80) permitted tracking of viral infectivity. These vectors were sequentially introduced into producer cells, complemented with Gag-Pol in trans to produce infectious virus, and used to infect target cells. Target cells were stained with a PE-conjugated anti-p24 (capsid) antibody to detect recombination events that led to expression of full-length Gag. The authors showed that the particular antibody they used could reliably detect a variety of parental Gag proteins from different subtypes and groups of HIV-1, as well as recombinants between subtypes and groups. Recombination rates were determined by comparing the frequency of Gag⁺ cells to all infected cells (as determined by the cell surface markers). Furthermore, since these vectors also contain defective forms of *gfp*, recombination rates could simultaneously be determined in identical (albeit foreign) gene sequences across viruses. While representing a tremendous improvement over standard reporter gene vectors, these vectors are limited by a small dynamic range of the assay due to the fixed distance

(~300 bp of capsid) of the mutations. However, suitable antibodies that recognize other regions of the unprocessed Gag polyprotein may be identified in the future, allowing for analysis of a larger region of Gag. Generally, this approach would likely not be amenable for identifying recombinants in other viral genes and non-coding regions of the genome. Additionally, some Gag recombinants may not be identified if, for example, they prevent proper folding or recognition by the antibody.

3.2.2. Use of LacZ (β -galactosidase) to Detect Recombinants in Viral Genes

Additional retroviral vectors have been developed that take advantage of reporter gene phenotypes but actually detect recombination events within viral targets. Specifically, vector pairs for different groups and subtypes of HIV-1 have been designed that contain either a functional *lacZ* or a completely unrelated sequence in its place (such as a defective *malt* gene in reverse orientation), as demonstrated in Figure 2D [77,148,149]. After the *lacZ* gene, *env* genes from various viral isolates were inserted, followed by a BamHI restriction site (in the *lacZ*⁻ vector) or no restriction site (in the *lacZ*⁺ vector). After PCR amplification, proviral DNA was digested with the BamHI restriction enzyme (in conjunction with PstI), ligated into a cloning vector, and subjected to blue-white color screening. Recombination in *env* results in *lacZ*⁺ proviruses with the BamHI restriction site, which can then efficiently ligate into the cloning vector. The recombination frequency is then computed by determining the ratio of *lacZ*⁺ colonies to all colonies. In this approach, all observed recombination events must have occurred within *env* rather than *lacZ*, as one vector completely lacks the *lacZ* gene. When *env* genes from different viral isolates are used, sequencing can be performed to identify the exact breakpoints. While this approach is able to identify recombinants in viral genes, it is somewhat laborious and involves the placement of viral genes outside of their native sequence contexts.

3.2.3. Heteroduplex Tracking, Quantitative PCR, and Restriction Mapping

Numerous biochemistry and molecular biology-based techniques have been used to directly identify and quantify retroviral recombinants occurring within viral genes, without any assistance from reporter genes. Most often, these assays rely upon the detection of sequence differences between distinct viral isolates, either from the same subtype (intrasubtype) or different subtypes (intersubtype), groups, or types of HIV. Thus, recombination frequencies from these assays may somewhat underestimate recombination rates between highly homologous viruses. Even when these assays are able to detect recombination between nearly identical viruses (by, for example, restriction mapping), the recombination breakpoints cannot be better defined unless silent point mutations have been introduced to create intervening intervals. These types of approaches have been used to identify recombinants in viral genes during both single cycle and multi-cycle infections. In spreading infections, viral recombinants are subject to strong purifying selection, which can restrict observable recombination events. In some cases, parental retroviral vectors have been used that contain different defects in essential viral genes (such as deletions in *pol* and *env*), such that only recombinants are able to spread after the initial round of replication [49,163,164].

In several reports, heteroduplex tracking (HTA) has been performed to identify recombinants in viral genes [36,37,152], which can detect recombinants (as well as mutants) when there are sufficient sequence differences between the parental viruses. Recombination can alter the affinity of a viral

sequence for a radioactive probe (relative to the two parental DNAs), which can be detected as a shift in band pattern during gel electrophoresis. HTA has been estimated to reliably detect sequence divergence when the mismatches exceed ~1% [165,166], and thus this technique can only identify crossovers introducing >1% mismatch relative to the parental viruses. Therefore, some recombinants cannot be detected by this method, particularly when the parental isolates are quite similar (e.g., from the same HIV-1 subtype). This method also requires sequencing of fragments to confirm that gel shifts were not caused by mutations, insertions, or deletions. Restriction mapping, Southern blotting (with short oligonucleotides specific to one of the parental isolates), and/or Sanger sequencing can be used to further refine recombination breakpoints between the viruses.

As an alternative to HTA, several studies have used PCR or restriction mapping to detect recombination [37,45,148,150–152]. PCR has been performed using primers that are specific to one of the two parental viruses. By combining one primer from each virus, recombinants can be detected. Additionally, different forms of quantitative PCR have been used to determine starting amounts of recombinant and parental DNAs, enabling calculation of recombination frequencies. As for HTA, Southern blotting or Sanger sequencing can be used to more precisely identify crossover points between distinct viral isolates. These approaches have been used to track HIV-1 intersubtype viral recombinants, and have led to identification of recombination hotspots as well as characterization of the potential mutagenicity of recombination.

3.2.4. Direct Sequencing of Viral Genes-Sanger

As discussed above, Sanger sequencing of viral genes has often been employed to further characterize recombinants identified by techniques such as HTA, PCR, or restriction mapping. Sanger sequencing has often also been used to find crossovers in viral genes from proviruses that are recombinants in reporter genes. In addition, recombinants have been identified simply by direct amplification and sequencing of viral genes from infected cells [37,39,151]. The low-throughput nature of direct Sanger sequencing is somewhat less problematic for studies of retroviral recombination than for mutagenesis. This is because some retroviruses, such as lentiviruses, recombine ~10 times more frequently than they mutate. While sequencing is often used to map recombination breakpoints between heterogeneous viruses, HIV-1 vectors have recently been engineered that allow detection of recombination between highly homologous viruses, thus more accurately representing the viral population found within infected individuals [39]. Specifically, silent mutations that could be detected by PCR and Sanger sequencing were introduced throughout the *gag* gene of an HIV-1 infectious molecular clone. These silent mutations served as markers for recombination and did not alter the protein sequences or any known cis elements. Further, they were shown to have no effect on viral infectivity. Most markers actually consisted of two mutations, typically spaced three bases apart, to facilitate differentiation of markers from mutations or background errors. Notably, the vectors they designed encode functional open reading frames for all viral proteins. In order to investigate HIV-1 recombination, viral stocks were first produced by co-transfecting the wild-type and marker plasmids into 293T cells. This is in contrast to many earlier studies in which producer cell clones were generated by sequential infection or, in some cases, co-infection with parental viruses [38,41,43,68–70,139–143,145–147]. While sequential infection is more time-consuming, it eliminates the possibility of transfection-induced

recombination (TIR) and more closely resembles native viral replication. However, the investigators of this study measured the frequency of TIR by direct sequencing of plasmid DNA from co-transfected cells and found that the frequency of TIR was much lower than the recombination rates measured for HIV-1. Next, the authors collected virus from co-transfected cells, infected primary T-cells, performed PCR, and identified recombinants by Sanger sequencing. Virus replication was limited to a single round using the T-20 fusion inhibitor. All PCR reactions were terminated in the log-linear phase of amplification, which reduces the frequency of PCR-mediated recombination [167]. Additional controls were performed to show that PCR-induced and intervirion recombination did not substantially contribute to the observed data. Further, the authors estimated the fraction of heterozygous virions directly from the sequencing data and developed a mathematical model that accounted for the possibility of multiple template switches, based on distances between genetic markers and observed recombination rates. Using this vector system, the HIV-1 recombination rate was observed to be 0.81×10^{-3} recombination events per nucleotide (REPN), corresponding to a true recombination rate of 1.35×10^{-3} REPN (or 12.5 crossovers/genome) after considering the possibility of multiple template switches.

3.2.5. Direct Sequencing of Viral Genes-NGS

In several subsequent studies, similar HIV-1 vectors with marker mutations throughout the *gag* and *pol* genes have been applied to 454 sequencing [40,153]. As described above, wild-type and marker vectors were co-transfected into 293T cells, and viruses were collected and used to infect primary blood mononuclear cells. Although replication-competent, the virus was limited to a single round of replication with fusion inhibitor. Proviral DNA was amplified by PCR, sequencing libraries were prepared, and 454 sequencing was performed, generating ~1 million reads per run. Experimental sources of recombination, such as transfection-induced recombination, intervirion recombination, and PCR, were carefully measured in parallel. These background sources were found to result in recombination rates less than 10% of the recombination rate measured in the biological samples. Using these vectors and 454 sequencing, the overall rates of recombination in HIV-1 *gag* and *pol* were found to be $1.5\text{--}2 \times 10^{-3}$ REPN, or ~14–19 crossovers per genome. In addition, using the large amounts of data generated by 454, recombination hot and cold spots could be identified in viral genes [153]. Further, ~15%–20% of mutations were found to be associated with recombination, though the direction of causality could not be determined [40]. As the association of recombination and mutation is itself a topic of great research interest, the high background error rates of NGS technologies are also somewhat problematic for studies of viral recombination. Thus, recent improvements aimed at lowering NGS error rates (see Section 2.3.2.) will surely benefit recombination studies as well. Additionally, these improvements often permit the identification and exclusion of PCR-induced crossovers as well as errors, thus limiting the effects of PCR-mediated recombination on the observed data.

4. Conclusions

Given the wide variety of retroviral vectors available, choosing a particular vector to assess retroviral mutagenesis and/or recombination can be daunting. Retroviral vectors based on reporter genes offer a number of distinct benefits that have led to their widespread use in mutation and recombination studies. First, mutants or recombinants can be efficiently detected based on phenotypes that are easily

observed and quantified, while mutants or recombinants in viral targets must be detected through much more laborious means. Targeted sequencing of mutant viruses can determine the nature of the underlying mutations, while targeted sequencing of recombinants can confirm that recombination took place (though the breakpoint locations cannot be precisely defined). In addition, reporter gene vectors are generally somewhat less dependent on PCR for identification of mutants or recombinants, and thus relatively less impeded by PCR-induced errors or crossovers. Nonetheless, reporter gene-based approaches are limited in that they may not necessarily be reflective of viral gene sequences. Large amounts of evidence have shown that patterns of mutation and recombination are affected by factors like homopolymeric runs and secondary structures [77–79,92–95,152], which may vary between template sequences. Given this, the degree to which reporter genes accurately represent viral genes cannot be easily ascertained at present, as so few studies have examined large numbers of mutations or crossovers in viral genes. Additionally, reporter gene vectors typically do not express all viral genes, though some exceptions do exist. In contrast, native viral genes are viewed as more desirable targets of mutation and recombination, but detection of mutation or recombination is often very laborious and time-consuming. Next-generation sequencing (NGS) technologies offer an obvious solution to these barriers, but standard NGS methods result in high background error rates, hindering identification of true mutations. Additionally, NGS is heavily reliant on PCR for sample preparation, which, under standard conditions, generates high levels of artifactual mutations and recombination events. However, several strategies for NGS sample preparation have recently been established that allow for identification and exclusion of the vast majority of background errors and crossovers due to PCR or sequencing [125,133,135,136,138]. While these methods have not yet been used to quantify retroviral mutation or recombination rates, they will be key to advancing our understanding of retroviral mutagenesis and recombination in viral genes. In the near future, standardized retroviral protocols, vectors, and assays will hopefully be developed that allow for application of these improvements to a wide variety of questions in retrovirology.

Acknowledgments

This work was supported by NIH grant R01 GM105876. J.R. was supported by the Institute for Molecular Virology Training Program (NIH grant T32 AI083196) and a University of Minnesota Doctoral Dissertation Fellowship. The authors would like to thank Christine Clouser for critical comments on the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References and Notes

1. Delviks-Frankenberry, K.; Galli, A.; Nikolaitchik, O.; Mens, H.; Pathak, V.K.; Hu, W.S. Mechanisms and factors that influence high frequency retroviral recombination. *Viruses* **2011**, *3*, 1650–1680.

2. Smyth, R.P.; Davenport, M.P.; Mak, J. The origin of genetic diversity in hiv-1. *Virus Res.* **2012**, *169*, 415–429.
3. Hu, W.S.; Hughes, S.H. Hiv-1 reverse transcription. *Cold Spring Harb. Perspect. Med.* **2012**, *2*, doi:10.1101/cshperspect.a006882.
4. Mansky, L.M.; Temin, H.M. Lower *in vivo* mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J. Virol.* **1995**, *69*, 5087–5094.
5. Pathak, V.K.; Temin, H.M. Broad spectrum of *in vivo* forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: Substitutions, frameshifts, and hypermutations. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 6019–6023.
6. Mansky, L.M. *In vivo* analysis of human t-cell leukemia virus type 1 reverse transcription accuracy. *J. Virol.* **2000**, *74*, 9525–9531.
7. Mansky, L.M.; Temin, H.M. Lower mutation rate of bovine leukemia virus relative to that of spleen necrosis virus. *J. Virol.* **1994**, *68*, 494–499.
8. Varela-Echavarria, A.; Prorock, C.M.; Ron, Y.; Dougherty, J.P. High rate of genetic rearrangement during replication of a moloney murine leukemia virus-based vector. *J. Virol.* **1993**, *67*, 6357–6364.
9. Sanjuan, R.; Nebot, M.R.; Chirico, N.; Mansky, L.M.; Belshaw, R. Viral mutation rates. *J. Virol.* **2010**, *84*, 9733–9748.
10. McCulloch, S.D.; Kunkel, T.A. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res.* **2008**, *18*, 148–161.
11. Menendez-Arias, L. Mutation rates and intrinsic fidelity of retroviral reverse transcriptases. *Viruses* **2009**, *1*, 1137–1165.
12. O'Neil, P.K.; Sun, G.; Yu, H.; Ron, Y.; Dougherty, J.P.; Preston, B.D. Mutational analysis of hiv-1 long terminal repeats to explore the relative contribution of reverse transcriptase and rna polymerase ii to viral mutagenesis. *J. Biol. Chem.* **2002**, *277*, 38053–38061.
13. Refsland, E.W.; Harris, R.S. The apobec3 family of retroelement restriction factors. *Curr. Top. Microbiol. Immunol.* **2013**, *371*, 1–27.
14. Desimmie, B.A.; Delviks-Frankenberry, K.A.; Burdick, R.C.; Qi, D.; Izumi, T.; Pathak, V.K. Multiple apobec3 restriction factors for hiv-1 and one vif to rule them all. *J. Mol. Biol.* **2014**, *426*, 1220–1245.
15. Mulder, L.C.; Harari, A.; Simon, V. Cytidine deamination induced hiv-1 drug resistance. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5501–5506.
16. Kim, E.Y.; Bhattacharya, T.; Kunstman, K.; Swantek, P.; Koning, F.A.; Malim, M.H.; Wolinsky, S.M. Human apobec3g-mediated editing can promote hiv-1 sequence diversification and accelerate adaptation to selective pressure. *J. Virol.* **2010**, *84*, 10402–10405.
17. Simon, V.; Zennou, V.; Murray, D.; Huang, Y.; Ho, D.D.; Bieniasz, P.D. Natural variation in vif: Differential impact on apobec3g/3f and a potential role in hiv-1 diversification. *PLoS Pathog.* **2005**, *1*, e6.
18. Iwabu, Y.; Kinomoto, M.; Tatsumi, M.; Fujita, H.; Shimura, M.; Tanaka, Y.; Ishizaka, Y.; Nolan, D.; Mallal, S.; Sata, T.; *et al.* Differential anti-apobec3g activity of hiv-1 vif proteins derived from different subtypes. *J. Biol. Chem.* **2010**, *285*, 35350–35358.

19. Binka, M.; Ooms, M.; Steward, M.; Simon, V. The activity spectrum of vif from multiple hiv-1 subtypes against apobec3g, apobec3f, and apobec3h. *J. Virol.* **2012**, *86*, 49–59.
20. Fourati, S.; Lambert-Niclot, S.; Soulie, C.; Wirden, M.; Malet, I.; Valantin, M.A.; Tubiana, R.; Simon, A.; Katlama, C.; Carcelain, G.; *et al.* Differential impact of apobec3-driven mutagenesis on hiv evolution in diverse anatomical compartments. *AIDS* **2014**, *28*, 487–491.
21. Janini, M.; Rogers, M.; Birx, D.R.; McCutchan, F.E. Human immunodeficiency virus type 1 DNA sequences genetically damaged by hypermutation are often abundant in patient peripheral blood mononuclear cells and may be generated during near-simultaneous infection and activation of cd4(+) t cells. *J. Virol.* **2001**, *75*, 7973–7986.
22. Kieffer, T.L.; Kwon, P.; Nettles, R.E.; Han, Y.; Ray, S.C.; Siliciano, R.F. G-->a hypermutation in protease and reverse transcriptase regions of human immunodeficiency virus type 1 residing in resting cd4+ t cells *in vivo*. *J. Virol.* **2005**, *79*, 1975–1980.
23. Kijak, G.H.; Janini, L.M.; Tovanabuttra, S.; Sanders-Buell, E.; Arroyo, M.A.; Robb, M.L.; Michael, N.L.; Birx, D.L.; McCutchan, F.E. Variable contexts and levels of hypermutation in hiv-1 proviral genomes recovered from primary peripheral blood mononuclear cells. *Virology* **2008**, *376*, 101–111.
24. Pace, C.; Keller, J.; Nolan, D.; James, I.; Gaudieri, S.; Moore, C.; Mallal, S. Population level analysis of human immunodeficiency virus type 1 hypermutation and its relationship with apobec3g and vif genetic variation. *J. Virol.* **2006**, *80*, 9259–9269.
25. Wei, M.; Xing, H.; Hong, K.; Huang, H.; Tang, H.; Qin, G.; Shao, Y. Biased g-to-a hypermutation in hiv-1 proviral DNA from a long-term non-progressor. *AIDS* **2004**, *18*, 1863–1865.
26. Wood, N.; Bhattacharya, T.; Keele, B.F.; Giorgi, E.; Liu, M.; Gaschen, B.; Daniels, M.; Ferrari, G.; Haynes, B.F.; McMichael, A.; *et al.* Hiv evolution in early infection: Selection pressures, patterns of insertion and deletion, and the impact of apobec. *PLoS Pathog.* **2009**, *5*, e1000414.
27. Mansky, L.M.; Bernard, L.C. 3'-azido-3'-deoxythymidine (azt) and azt-resistant reverse transcriptase can increase the *in vivo* mutation rate of human immunodeficiency virus type 1. *J. Virol.* **2000**, *74*, 9532–9539.
28. Mansky, L.M.; le Rouzic, E.; Benichou, S.; Gajary, L.C. Influence of reverse transcriptase variants, drugs, and vpr on human immunodeficiency virus type 1 mutant frequencies. *J. Virol.* **2003**, *77*, 2071–2080.
29. Weiss, K.K.; Chen, R.; Skasko, M.; Reynolds, H.M.; Lee, K.; Bambara, R.A.; Mansky, L.M.; Kim, B. A role for dntp binding of human immunodeficiency virus type 1 reverse transcriptase in viral mutagenesis. *Biochemistry* **2004**, *43*, 4490–4500.
30. Chen, R.; Yokoyama, M.; Sato, H.; Reilly, C.; Mansky, L.M. Human immunodeficiency virus mutagenesis during antiviral therapy: Impact of drug-resistant reverse transcriptase and nucleoside and nonnucleoside reverse transcriptase inhibitors on human immunodeficiency virus type 1 mutation frequencies. *J. Virol.* **2005**, *79*, 12045–12057.
31. Abram, M.E.; Ferris, A.L.; Das, K.; Quinones, O.; Shao, W.; Tuske, S.; Alvord, W.G.; Arnold, E.; Hughes, S.H. Mutations in hiv-1 reverse transcriptase affect the errors made in a single cycle of viral replication. *J. Virol.* **2014**, *88*, 7589–7601.

32. Mansky, L.M. Mutagenic outcome of combined antiviral drug treatment during human immunodeficiency virus type 1 replication. *Virology* **2003**, *307*, 116–121.
33. Mansky, L.M. The mutation rate of human immunodeficiency virus type 1 is influenced by the vpr gene. *Virology* **1996**, *222*, 391–400.
34. Mansky, L.M.; Preveral, S.; Selig, L.; Benarous, R.; Benichou, S. The interaction of vpr with uracil DNA glycosylase modulates the human immunodeficiency virus type 1 *in vivo* mutation rate. *J. Virol.* **2000**, *74*, 7039–7047.
35. Chen, R.; le Rouzic, E.; Kearney, J.A.; Mansky, L.M.; Benichou, S. Vpr-mediated incorporation of ung2 into hiv-1 particles is required to modulate the virus mutation rate and for replication in macrophages. *J. Biol. Chem.* **2004**, *279*, 28419–28425.
36. Jetzt, A.E.; Yu, H.; Klarmann, G.J.; Ron, Y.; Preston, B.D.; Dougherty, J.P. High rate of recombination throughout the human immunodeficiency virus type 1 genome. *J. Virol.* **2000**, *74*, 1234–1240.
37. Zhuang, J.; Jetzt, A.E.; Sun, G.; Yu, H.; Klarmann, G.; Ron, Y.; Preston, B.D.; Dougherty, J.P. Human immunodeficiency virus type 1 recombination: Rate, fidelity, and putative hot spots. *J. Virol.* **2002**, *76*, 11273–11282.
38. Rhodes, T.D.; Nikolaitchik, O.; Chen, J.; Powell, D.; Hu, W.S. Genetic recombination of human immunodeficiency virus type 1 in one round of viral replication: Effects of genetic distance, target cells, accessory genes, and lack of high negative interference in crossover events. *J. Virol.* **2005**, *79*, 1666–1677.
39. Schlub, T.E.; Smyth, R.P.; Grimm, A.J.; Mak, J.; Davenport, M.P. Accurately measuring recombination between closely related hiv-1 genomes. *PLoS Comput. Biol.* **2010**, *6*, e1000766.
40. Schlub, T.E.; Grimm, A.J.; Smyth, R.P.; Cromer, D.; Chopra, A.; Mallal, S.; Venturi, V.; Waugh, C.; Mak, J.; Davenport, M.P. Fifteen to twenty percent of hiv substitution mutations are associated with recombination. *J. Virol.* **2014**, *88*, 3837–3849.
41. Rhodes, T.; Wargo, H.; Hu, W.S. High rates of human immunodeficiency virus type 1 recombination: Near-random segregation of markers one kilobase apart in one round of viral replication. *J. Virol.* **2003**, *77*, 11193–11200.
42. Onafuwa, A.; An, W.; Robson, N.D.; Telesnitsky, A. Human immunodeficiency virus type 1 genetic recombination is more frequent than that of moloney murine leukemia virus despite similar template switching rates. *J. Virol.* **2003**, *77*, 4577–4587.
43. Hu, W.S.; Temin, H.M. Genetic consequences of packaging two rna genomes in one retroviral particle: Pseudodiploidy and high rate of genetic recombination. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 1556–1560.
44. Kellam, P.; Larder, B.A. Retroviral recombination can lead to linkage of reverse transcriptase mutations that confer increased zidovudine resistance. *J. Virol.* **1995**, *69*, 669–674.
45. Moutouh, L.; Corbeil, J.; Richman, D.D. Recombination leads to the rapid emergence of hiv-1 dually resistant mutants under selective drug pressure. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 6106–6111.

46. Kim, E.Y.; Busch, M.; Abel, K.; Fritts, L.; Bustamante, P.; Stanton, J.; Lu, D.; Wu, S.; Glowczwskie, J.; Rourke, T.; *et al.* Retroviral recombination *in vivo*: Viral replication patterns and genetic structure of simian immunodeficiency virus (siv) populations in rhesus macaques after simultaneous or sequential intravaginal inoculation with sivmac239deltavpx/deltavpr and sivmac239deltanef. *J. Virol.* **2005**, *79*, 4886–4895.
47. Quan, Y.; Liang, C.; Brenner, B.G.; Wainberg, M.A. Multidrug-resistant variants of hiv type 1 (hiv-1) can exist in cells as defective quasispecies and be rescued by superinfection with other defective hiv-1 variants. *J. Infect. Dis.* **2009**, *200*, 1479–1483.
48. Quan, Y.; Xu, H.; Wainberg, M.A. Defective hiv-1 quasispecies in the form of multiply drug-resistant proviral DNA within cells can be rescued by superinfection with different subtype variants of hiv-1 and by hiv-2 and siv. *J. Antimicrob. Chemother.* **2014**, *69*, 21–27.
49. Galli, A.; Kearney, M.; Nikolaitchik, O.A.; Yu, S.; Chin, M.P.; Maldarelli, F.; Coffin, J.M.; Pathak, V.K.; Hu, W.S. Patterns of human immunodeficiency virus type 1 recombination *ex vivo* provide evidence for coadaptation of distant sites, resulting in purifying selection for intersubtype recombinants during replication. *J. Virol.* **2010**, *84*, 7651–7661.
50. Los alamos national laboratory hiv databases. Available online: <http://www.hiv.lanl.gov/> (accessed on 14 August 2014).
51. Hemelaar, J.; Gouws, E.; Ghys, P.D.; Osmanov, S. Global trends in molecular epidemiology of hiv-1 during 2000–2007. *AIDS* **2011**, *25*, 679–689.
52. Chen, J.; Dang, Q.; Unutmaz, D.; Pathak, V.K.; Maldarelli, F.; Powell, D.; Hu, W.S. Mechanisms of nonrandom human immunodeficiency virus type 1 infection and double infection: Preference in virus entry is important but is not the sole factor. *J. Virol.* **2005**, *79*, 4140–4149.
53. Dang, Q.; Chen, J.; Unutmaz, D.; Coffin, J.M.; Pathak, V.K.; Powell, D.; KewalRamani, V.N.; Maldarelli, F.; Hu, W.S. Nonrandom hiv-1 infection and double infection via direct and cell-mediated pathways. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 632–637.
54. Del Portillo, A.; Tripodi, J.; Najfeld, V.; Wodarz, D.; Levy, D.N.; Chen, B.K. Multiploid inheritance of hiv-1 during cell-to-cell infection. *J. Virol.* **2011**, *85*, 7169–7176.
55. Levy, D.N.; Aldrovandi, G.M.; Kutsch, O.; Shaw, G.M. Dynamics of hiv-1 recombination in its natural target cells. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4204–4209.
56. Bregnard, C.; Pacini, G.; Danos, O.; Basmaciogullari, S. Suboptimal provirus expression explains apparent nonrandom cell coinfection with hiv-1. *J. Virol.* **2012**, *86*, 8810–8820.
57. Nethe, M.; Berkhout, B.; van der Kuyl, A.C. Retroviral superinfection resistance. *Retrovirology* **2005**, *2*, 52.
58. Chohan, B.; Lavreys, L.; Rainwater, S.M.; Overbaugh, J. Evidence for frequent reinfection with human immunodeficiency virus type 1 of a different subtype. *J. Virol.* **2005**, *79*, 10701–10708.
59. Piantadosi, A.; Chohan, B.; Chohan, V.; McClelland, R.S.; Overbaugh, J. Chronic hiv-1 infection frequently fails to protect against superinfection. *PLoS Pathog.* **2007**, *3*, e177.
60. Piantadosi, A.; Ngayo, M.O.; Chohan, B.; Overbaugh, J. Examination of a second region of the hiv type 1 genome reveals additional cases of superinfection. *AIDS Res. Hum. Retrovir.* **2008**, *24*, 1221.

61. Redd, A.D.; Mullis, C.E.; Serwadda, D.; Kong, X.; Martens, C.; Ricklefs, S.M.; Tobian, A.A.; Xiao, C.; Grabowski, M.K.; Nalugoda, F.; *et al.* The rates of hiv superinfection and primary hiv incidence in a general population in rakai, uganda. *J. Infect. Dis.* **2012**, *206*, 267–274.
62. Ronen, K.; McCoy, C.O.; Matsen, F.A.; Boyd, D.F.; Emery, S.; Odem-Davis, K.; Jaoko, W.; Mandaliya, K.; McClelland, R.S.; Richardson, B.A.; *et al.* Hiv-1 superinfection occurs less frequently than initial infection in a cohort of high-risk kenyan women. *PLoS Pathog.* **2013**, *9*, e1003593.
63. Smith, D.M.; Wong, J.K.; Hightower, G.K.; Ignacio, C.C.; Koelsch, K.K.; Daar, E.S.; Richman, D.D.; Little, S.J. Incidence of hiv superinfection following primary infection. *JAMA: J. Am. Med. Assoc.* **2004**, *292*, 1177–1178.
64. Gratton, S.; Cheynier, R.; Dumaurier, M.J.; Oksenhendler, E.; Wain-Hobson, S. Highly restricted spread of hiv-1 and multiply infected cells within splenic germinal centers. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 14566–14571.
65. Jung, A.; Maier, R.; Vartanian, J.P.; Bocharov, G.; Jung, V.; Fischer, U.; Meese, E.; Wain-Hobson, S.; Meyerhans, A. Recombination: Multiply infected spleen cells in hiv patients. *Nature* **2002**, *418*, 144.
66. Schultz, A.; Sopper, S.; Sauermann, U.; Meyerhans, A.; Suspene, R. Stable multi-infection of splenocytes during siv infection--the basis for continuous recombination. *Retrovirology* **2012**, *9*, 31.
67. Josefsson, L.; King, M.S.; Makitalo, B.; Brannstrom, J.; Shao, W.; Maldarelli, F.; Kearney, M.F.; Hu, W.S.; Chen, J.; Gaines, H.; *et al.* Majority of cd4+ t cells from peripheral blood of hiv-1-infected individuals contain only one hiv DNA molecule. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11199–11204.
68. Chin, M.P.; Rhodes, T.D.; Chen, J.; Fu, W.; Hu, W.S. Identification of a major restriction in hiv-1 intersubtype recombination. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 9002–9007.
69. Chin, M.P.; Chen, J.; Nikolaitchik, O.A.; Hu, W.S. Molecular determinants of hiv-1 intersubtype recombination potential. *Virology* **2007**, *363*, 437–446.
70. Chin, M.P.; Lee, S.K.; Chen, J.; Nikolaitchik, O.A.; Powell, D.A.; Fivash, M.J., Jr.; Hu, W.S. Long-range recombination gradient between hiv-1 subtypes b and c variants caused by sequence differences in the dimerization initiation signal region. *J. Mol. Biol.* **2008**, *377*, 1324–1333.
71. Nikolaitchik, O.A.; Galli, A.; Moore, M.D.; Pathak, V.K.; Hu, W.S. Multiple barriers to recombination between divergent hiv-1 variants revealed by a dual-marker recombination assay. *J. Mol. Biol.* **2011**, *407*, 521–531.
72. Svarovskaia, E.S.; Delviks, K.A.; Hwang, C.K.; Pathak, V.K. Structural determinants of murine leukemia virus reverse transcriptase that affect the frequency of template switching. *J. Virol.* **2000**, *74*, 7171–7178.
73. Hwang, C.K.; Svarovskaia, E.S.; Pathak, V.K. Dynamic copy choice: Steady state between murine leukemia virus polymerase and polymerase-dependent rna h activity determines frequency of *in vivo* template switching. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12209–12214.
74. Pfeiffer, J.K.; Topping, R.S.; Shin, N.H.; Telesnitsky, A. Altering the intracellular environment increases the frequency of tandem repeat deletion during moloney murine leukemia virus reverse transcription. *J. Virol.* **1999**, *73*, 8441–8447.

75. Nikolenko, G.N.; Svarovskaia, E.S.; Delviks, K.A.; Pathak, V.K. Antiretroviral drug resistance mutations in human immunodeficiency virus type 1 reverse transcriptase increase template-switching frequency. *J. Virol.* **2004**, *78*, 8761–8770.
76. Operario, D.J.; Balakrishnan, M.; Bambara, R.A.; Kim, B. Reduced dntp interaction of human immunodeficiency virus type 1 reverse transcriptase promotes strand transfer. *J. Biol. Chem.* **2006**, *281*, 32113–32121.
77. Galetto, R.; Moumen, A.; Giacomoni, V.; Veron, M.; Charneau, P.; Negroni, M. The structure of hiv-1 genomic rna in the gp120 gene determines a recombination hot spot *in vivo*. *J. Biol. Chem.* **2004**, *279*, 36625–36632.
78. DeStefano, J.J.; Mallaber, L.M.; Rodriguez-Rodriguez, L.; Fay, P.J.; Bambara, R.A. Requirements for strand transfer between internal regions of heteropolymer templates by human immunodeficiency virus reverse transcriptase. *J. Virol.* **1992**, *66*, 6370–6378.
79. Wu, W.; Blumberg, B.M.; Fay, P.J.; Bambara, R.A. Strand transfer mediated by human immunodeficiency virus reverse transcriptase *in vitro* is promoted by pausing and results in misincorporation. *J. Biol. Chem.* **1995**, *270*, 325–332.
80. Brincat, J.L.; Pfeiffer, J.K.; Telesnitsky, A. Rnase h activity is required for high-frequency repeat deletion during moloney murine leukemia virus replication. *J. Virol.* **2002**, *76*, 88–95.
81. An, W.; Telesnitsky, A. Effects of varying sequence similarity on the frequency of repeat deletion during reverse transcription of a human immunodeficiency virus type 1 vector. *J. Virol.* **2002**, *76*, 7897–7902.
82. Chun, T.W.; Nickle, D.C.; Justement, J.S.; Meyers, J.H.; Roby, G.; Hallahan, C.W.; Kottlilil, S.; Moir, S.; Mican, J.M.; Mullins, J.I.; *et al.* Persistence of hiv in gut-associated lymphoid tissue despite long-term antiretroviral therapy. *J. Infect. Dis.* **2008**, *197*, 714–720.
83. Gras, G.; Kaul, M. Molecular mechanisms of neuroinvasion by monocytes-macrophages in hiv-1 infection. *Retrovirology* **2010**, *7*, 30.
84. Ruelas, D.S.; Greene, W.C. An integrated overview of hiv-1 latency. *Cell* **2013**, *155*, 519–529.
85. Pursuing Later Treatment Options II (PLATO II) project team; Collaboration of Observational HIV Epidemiological Research Europe (COHERE) Group; Nakagawa, F.; Lodwick, R.; Costagliola, D.; van Sighem, A.; Torti, C.; Podzamczar, D.; Mocroft, A.; Ledergerber, B.; *et al.* Calendar time trends in the incidence and prevalence of triple-class virologic failure in antiretroviral drug-experienced people with hiv in europe. *J. Acquir. Immune Defic. Syndr.* **2012**, *59*, 294–299.
86. Galli, A.; Bukh, J. Comparative analysis of the molecular mechanisms of recombination in hepatitis c virus. *Trends Microbiol.* **2014**, *22*, 354–364.
87. Lin, C.C.; Yang, Z.W.; Iang, S.B.; Chao, M. Reduced genetic distance and high replication levels increase the rna recombination rate of hepatitis delta virus. *Virus Res.* **2014**, doi:10.1016/j.virusres.2014.08.011.
88. Lowry, K.; Woodman, A.; Cook, J.; Evans, D.J. Recombination in enteroviruses is a biphasic replicative process involving the generation of greater-than genome length 'imprecise' intermediates. *PLoS pathogens* **2014**, *10*, e1004191.

89. Diamond, T.L.; Roshal, M.; Jamburuthugoda, V.K.; Reynolds, H.M.; Merriam, A.R.; Lee, K.Y.; Balakrishnan, M.; Bambara, R.A.; Planelles, V.; Dewhurst, S.; *et al.* Macrophage tropism of hiv-1 depends on efficient cellular dntp utilization by reverse transcriptase. *J. Biol. Chem.* **2004**, *279*, 51545–51553.
90. Dougherty, J.P.; Temin, H.M. Determination of the rate of base-pair substitution and insertion mutations in retrovirus replication. *J. Virol.* **1988**, *62*, 2817–2822.
91. Varela-Echavarria, A.; Garvey, N.; Preston, B.D.; Dougherty, J.P. Comparison of moloney murine leukemia virus mutation rate with the fidelity of its reverse transcriptase *in vitro*. *J. Biol. Chem.* **1992**, *267*, 24681–24688.
92. Bebenek, K.; Abbotts, J.; Roberts, J.D.; Wilson, S.H.; Kunkel, T.A. Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase. *J. Biol. Chem.* **1989**, *264*, 16948–16956.
93. Boyer, J.C.; Bebenek, K.; Kunkel, T.A. Unequal human immunodeficiency virus type 1 reverse transcriptase error rates with rna and DNA templates. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 6919–6923.
94. Eckert, K.A.; Kunkel, T.A. Fidelity of DNA synthesis catalyzed by human DNA polymerase alpha and hiv-1 reverse transcriptase: Effect of reaction ph. *Nucleic Acids Res.* **1993**, *21*, 5212–5220.
95. Ji, J.; Loeb, L.A. Fidelity of hiv-1 reverse transcriptase copying a hypervariable region of the hiv-1 env gene. *Virology* **1994**, *199*, 323–330.
96. Laakso, M.M.; Sutton, R.E. Replicative fidelity of lentiviral vectors produced by transient transfection. *Virology* **2006**, *348*, 406–417.
97. Parthasarathi, S.; Varela-Echavarria, A.; Ron, Y.; Preston, B.D.; Dougherty, J.P. Genetic rearrangements occurring during a single cycle of murine leukemia virus vector replication: Characterization and implications. *J. Virol.* **1995**, *69*, 7991–8000.
98. Huang, K.J.; Wooley, D.P. A new cell-based assay for measuring the forward mutation rate of hiv-1. *J. Virol. Methods* **2005**, *124*, 95–104.
99. Kim, T.; Mudry, R.A., Jr.; Rexrode, C.A., 2nd; Pathak, V.K. Retroviral mutation rates and a-to-g hypermutations during different stages of retroviral replication. *J. Virol.* **1996**, *70*, 7594–7602.
100. Mansky, L.M. Forward mutation rate of human immunodeficiency virus type 1 in a t lymphoid cell line. *AIDS Res. Hum. Retrovir.* **1996**, *12*, 307–314.
101. Abram, M.E.; Ferris, A.L.; Shao, W.; Alvord, W.G.; Hughes, S.H. Nature, position, and frequency of mutations made in a single cycle of hiv-1 replication. *J. Virol.* **2010**, *84*, 9864–9878.
102. Pathak, V.K.; Temin, H.M. Broad spectrum of *in vivo* forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: Deletions and deletions with insertions. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 6024–6028.
103. Sadler, H.A.; Stenglein, M.D.; Harris, R.S.; Mansky, L.M. Apobec3g contributes to hiv-1 variation through sublethal mutagenesis. *J. Virol.* **2010**, *84*, 7396–7404.
104. Holtz, C.M.; Mansky, L.M. Variation of hiv-1 mutation spectra among cell types. *J. Virol.* **2013**, *87*, 5296–5299.
105. Dapp, M.J.; Clouser, C.L.; Patterson, S.; Mansky, L.M. 5-azacytidine can induce lethal mutagenesis in human immunodeficiency virus type 1. *J. Virol.* **2009**, *83*, 11950–11958.

106. Clouser, C.L.; Patterson, S.E.; Mansky, L.M. Exploiting drug repositioning for discovery of a novel hiv combination therapy. *J. Virol.* **2010**, *84*, 9301–9309.
107. Rawson, J.M.; Heineman, R.H.; Beach, L.B.; Martin, J.L.; Schnettler, E.K.; Dapp, M.J.; Patterson, S.E.; Mansky, L.M. 5,6-dihydro-5-aza-2'-deoxycytidine potentiates the anti-hiv-1 activity of ribonucleotide reductase inhibitors. *Bioorganic Med. Chem.* **2013**, *21*, 7222–7228.
108. Dapp, M.J.; Holtz, C.M.; Mansky, L.M. Concomitant lethal mutagenesis of human immunodeficiency virus type 1. *J. Mol. Biol.* **2012**, *419*, 158–170.
109. Leider, J.M.; Palese, P.; Smith, F.I. Determination of the mutation rate of a retrovirus. *J. Virol.* **1988**, *62*, 3084–3091.
110. Monk, R.J.; Malik, F.G.; Stokesberry, D.; Evans, L.H. Direct determination of the point mutation rate of a murine retrovirus. *J. Virol.* **1992**, *66*, 3683–3689.
111. Gao, F.; Chen, Y.; Levy, D.N.; Conway, J.A.; Kepler, T.B.; Hui, H. Unselected mutations in the human immunodeficiency virus type 1 genome are mostly nonsynonymous and often deleterious. *J. Virol.* **2004**, *78*, 2426–2433.
112. Dapp, M.J.; Bonnac, L.; Patterson, S.E.; Mansky, L.M. Discovery of novel ribonucleoside analogs with activity against human immunodeficiency virus type 1. *J. Virol.* **2014**, *88*, 354–363.
113. Bebenek, K.; Kunkel, T.A. Analyzing fidelity of DNA polymerases. *Methods Enzymol.* **1995**, *262*, 217–232.
114. Boyer, J.C.; Bebenek, K.; Kunkel, T.A. Analyzing the fidelity of reverse transcription and transcription. *Methods Enzymol.* **1996**, *275*, 523–537.
115. Beach, L.B.; Rawson, J.M.; Kim, B.; Patterson, S.E.; Mansky, L.M. Novel inhibitors of human immunodeficiency virus type 2 infectivity. *J. Gen. Virol.* **2014**, doi:10.1099/vir.0.069864-0.
116. Clouser, C.L.; Chauhan, J.; Bess, M.A.; van Oploo, J.L.; Zhou, D.; Dimick-Gray, S.; Mansky, L.M.; Patterson, S.E. Anti-hiv-1 activity of resveratrol derivatives and synergistic inhibition of hiv-1 by the combination of resveratrol and decitabine. *Bioorganic Med. Chem. Lett.* **2012**, *22*, 6642–6646.
117. Dapp, M.J.; Heineman, R.H.; Mansky, L.M. Interrelationship between hiv-1 fitness and mutation rate. *J. Mol. Biol.* **2013**, *425*, 41–53.
118. Loeb, L.A.; Essigmann, J.M.; Kazazi, F.; Zhang, J.; Rose, K.D.; Mullins, J.I. Lethal mutagenesis of hiv with mutagenic nucleoside analogs. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 1492–1497.
119. Harris, K.S.; Brabant, W.; Styrchak, S.; Gall, A.; Daifuku, R. Kp-1212/1461, a nucleoside designed for the treatment of hiv by viral mutagenesis. *Antivir. Res.* **2005**, *67*, 1–9.
120. Swenson, L.C.; Mo, T.; Dong, W.W.; Zhong, X.; Woods, C.K.; Jensen, M.A.; Thielen, A.; Chapman, D.; Lewis, M.; James, I.; *et al.* Deep sequencing to infer hiv-1 co-receptor usage: Application to three clinical trials of maraviroc in treatment-experienced patients. *J. Infect. Dis.* **2011**, *203*, 237–245.
121. Vandenbroucke, I.; van Marck, H.; Mostmans, W.; van Eygen, V.; Rondelez, E.; Thys, K.; van Baelen, K.; Fransen, K.; Vaira, D.; Kabeya, K.; *et al.* Hiv-1 v3 envelope deep sequencing for clinical plasma specimens failing in phenotypic tropism assays. *AIDS Res. Ther.* **2010**, *7*, 4.
122. Fischer, W.; Ganusov, V.V.; Giorgi, E.E.; Hraber, P.T.; Keele, B.F.; Leitner, T.; Han, C.S.; Gleasner, C.D.; Green, L.; Lo, C.C.; *et al.* Transmission of single hiv-1 genomes and dynamics of early immune escape revealed by ultra-deep sequencing. *PLoS One* **2010**, *5*, e12303.

123. Simen, B.B.; Simons, J.F.; Hullsiek, K.H.; Novak, R.M.; Macarthur, R.D.; Baxter, J.D.; Huang, C.; Lubeski, C.; Turenchalk, G.S.; Braverman, M.S.; *et al.* Low-abundance drug-resistant viral variants in chronically hiv-infected, antiretroviral treatment-naive patients significantly impact treatment outcomes. *J. Infect. Dis.* **2009**, *199*, 693–701.
124. Hedskog, C.; Mild, M.; Jernberg, J.; Sherwood, E.; Bratt, G.; Leitner, T.; Lundeberg, J.; Andersson, B.; Albert, J. Dynamics of hiv-1 quasispecies during antiviral treatment dissected using ultra-deep pyrosequencing. *PLoS One* **2010**, *5*, e11345.
125. Jabara, C.B.; Jones, C.D.; Roach, J.; Anderson, J.A.; Swanstrom, R. Accurate sampling and deep sequencing of the hiv-1 protease gene using a primer id. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20166–20171.
126. Li, J.Z.; Chapman, B.; Charlebois, P.; Hofmann, O.; Weiner, B.; Porter, A.J.; Samuel, R.; Vardhanabhuti, S.; Zheng, L.; Eron, J.; *et al.* Comparison of illumina and 454 deep sequencing in participants failing raltegravir-based antiretroviral therapy. *PLoS One* **2014**, *9*, e90485.
127. Nakamura, K.; Oshima, T.; Morimoto, T.; Ikeda, S.; Yoshikawa, H.; Shiwa, Y.; Ishikawa, S.; Linak, M.C.; Hirai, A.; Takahashi, H.; *et al.* Sequence-specific error profile of illumina sequencers. *Nucleic Acids Res.* **2011**, *39*, e90.
128. Gilles, A.; Meglec, E.; Pech, N.; Ferreira, S.; Malausa, T.; Martin, J.F. Accuracy and quality assessment of 454 gs-flx titanium pyrosequencing. *BMC Genomics* **2011**, *12*, 245.
129. Vandenbroucke, I.; van Marck, H.; Verhasselt, P.; Thys, K.; Mostmans, W.; Dumont, S.; van Eygen, V.; Coen, K.; Tuefferd, M.; Aerssens, J. Minor variant detection in amplicons using 454 massive parallel pyrosequencing: Experiences and considerations for successful applications. *BioTechniques* **2011**, *51*, 167–177.
130. Minoche, A.E.; Dohm, J.C.; Himmelbauer, H. Evaluation of genomic high-throughput sequencing data generated on illumina hiseq and genome analyzer systems. *Genome Biol.* **2011**, *12*, R112.
131. Dohm, J.C.; Lottaz, C.; Borodina, T.; Himmelbauer, H. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Res.* **2008**, *36*, e105.
132. Faith, J.J.; Guruge, J.L.; Charbonneau, M.; Subramanian, S.; Seedorf, H.; Goodman, A.L.; Clemente, J.C.; Knight, R.; Heath, A.C.; Leibel, R.L.; *et al.* The long-term stability of the human gut microbiota. *Science* **2013**, *341*, 1237439.
133. Schmitt, M.W.; Kennedy, S.R.; Salk, J.J.; Fox, E.J.; Hiatt, J.B.; Loeb, L.A. Detection of ultra-rare mutations by next-generation sequencing. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 14508–14513.
134. Lundberg, D.S.; Yourstone, S.; Mieczkowski, P.; Jones, C.D.; Dangl, J.L. Practical innovations for high-throughput amplicon sequencing. *Nat. Methods* **2013**, *10*, 999–1002.
135. Kinde, I.; Wu, J.; Papadopoulos, N.; Kinzler, K.W.; Vogelstein, B. Detection and quantification of rare mutations with massively parallel sequencing. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 9530–9535.
136. Lou, D.I.; Hussmann, J.A.; McBee, R.M.; Acevedo, A.; Andino, R.; Press, W.H.; Sawyer, S.L. High-throughput DNA sequencing errors are reduced by orders of magnitude using circle sequencing. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 19872–19877.

137. Acevedo, A.; Andino, R. Library preparation for highly accurate population sequencing of rna viruses. *Nat. Protoc.* **2014**, *9*, 1760–1769.
138. Acevedo, A.; Brodsky, L.; Andino, R. Mutational and fitness landscapes of an rna virus revealed through population sequencing. *Nature* **2014**, *505*, 686–690.
139. Anderson, J.A.; Bowman, E.H.; Hu, W.S. Retroviral recombination rates do not increase linearly with marker distance and are limited by the size of the recombining subpopulation. *J. Virol.* **1998**, *72*, 1195–1202.
140. Anderson, J.A.; Pathak, V.K.; Hu, W.S. Effect of the murine leukemia virus extended packaging signal on the rates and locations of retroviral recombination. *J. Virol.* **2000**, *74*, 6953–6963.
141. Hu, W.S.; Bowman, E.H.; Delviks, K.A.; Pathak, V.K. Homologous recombination occurs in a distinct retroviral subpopulation and exhibits high negative interference. *J. Virol.* **1997**, *71*, 6028–6036.
142. Bircher, L.A.; Rigano, J.C.; Ponferrada, V.G.; Wooley, D.P. High fidelity of homologous retroviral recombination in cell culture. *Arch. Virol.* **2002**, *147*, 1665–1683.
143. Anderson, J.A.; Teufel, R.J., 2nd; Yin, P.D.; Hu, W.S. Correlated template-switching events during minus-strand DNA synthesis: A mechanism for high negative interference during retroviral recombination. *J. Virol.* **1998**, *72*, 1186–1194.
144. Delviks, K.A.; Pathak, V.K. Effect of distance between homologous sequences and 3' homology on the frequency of retroviral reverse transcriptase template switching. *J. Virol.* **1999**, *73*, 7923–7932.
145. Chen, J.; Rhodes, T.D.; Hu, W.S. Comparison of the genetic recombination rates of human immunodeficiency virus type 1 in macrophages and t cells. *J. Virol.* **2005**, *79*, 9337–9340.
146. Chen, J.; Powell, D.; Hu, W.S. High frequency of genetic recombination is a common feature of primate lentivirus replication. *J. Virol.* **2006**, *80*, 9651–9658.
147. Motomura, K.; Chen, J.; Hu, W.S. Genetic recombination between human immunodeficiency virus type 1 (hiv-1) and hiv-2, two distinct human lentiviruses. *J. Virol.* **2008**, *82*, 1923–1933.
148. Baird, H.A.; Gao, Y.; Galetto, R.; Lalonde, M.; Anthony, R.M.; Giacomoni, V.; Abreha, M.; Destefano, J.J.; Negroni, M.; Arts, E.J. Influence of sequence identity and unique breakpoints on the frequency of intersubtype hiv-1 recombination. *Retrovirology* **2006**, *3*, 91.
149. Simon-Loriere, E.; Galetto, R.; Hamoudi, M.; Archer, J.; Lefeuvre, P.; Martin, D.P.; Robertson, D.L.; Negroni, M. Molecular mechanisms of recombination restriction in the envelope gene of the human immunodeficiency virus. *PLoS Pathog.* **2009**, *5*, e1000418.
150. Kuwata, T.; Miyazaki, Y.; Igarashi, T.; Takehisa, J.; Hayami, M. The rapid spread of recombinants during a natural *in vitro* infection with two human immunodeficiency virus type 1 strains. *J. Virol.* **1997**, *71*, 7088–7091.
151. Iglesias-Sanchez, M.J.; Lopez-Galindez, C. Analysis, quantification, and evolutionary consequences of hiv-1 *in vitro* recombination. *Virology* **2002**, *304*, 392–402.
152. Quinones-Mateu, M.E.; Gao, Y.; Ball, S.C.; Marozsan, A.J.; Abreha, A.; Arts, E.J. *In vitro* intersubtype recombinants of human immunodeficiency virus type 1: Comparison to recent and circulating *in vivo* recombinant forms. *J. Virol.* **2002**, *76*, 9600–9613.

153. Smyth, R.P.; Schlub, T.E.; Grimm, A.J.; Waugh, C.; Ellenberg, P.; Chopra, A.; Mallal, S.; Cromer, D.; Mak, J.; Davenport, M.P. Identifying recombination hot spots in the hiv-1 genome. *J. Virol.* **2014**, *88*, 2891–2902.
154. Czernilofsky, A.P.; Levinson, A.D.; Varmus, H.E.; Bishop, J.M.; Tischer, E.; Goodman, H.M. Nucleotide sequence of an avian sarcoma virus oncogene (src) and proposed amino acid sequence for gene product. *Nature* **1980**, *287*, 198–203.
155. Hughes, S.; Kosik, E. Mutagenesis of the region between env and src of the sr-a strain of rous sarcoma virus for the purpose of constructing helper-independent vectors. *Virology* **1984**, *136*, 89–99.
156. Omer, C.A.; Pogue-Geile, K.; Guntaka, R.; Staskus, K.A.; Faras, A.J. Involvement of directly repeated sequences in the generation of deletions of the avian sarcoma virus src gene. *J. Virol.* **1983**, *47*, 380–382.
157. Rhode, B.W.; Emerman, M.; Temin, H.M. Instability of large direct repeats in retrovirus vectors. *J. Virol.* **1987**, *61*, 925–927.
158. Lovmand, S.; Kjeldgaard, N.O.; Jorgensen, P.; Pedersen, F.S. Enhancer functions in u3 of akv virus: A role for cooperativity of a tandem repeat unit and its flanking DNA sequences. *J. Virol.* **1990**, *64*, 3185–3191.
159. Trainor, C.D.; Scott, M.L.; Josephs, S.F.; Fry, K.E.; Reitz, M.S., Jr. Nucleotide sequence of the large terminal repeat of two different strains of gibbon ape leukemia virus. *Virology* **1984**, *137*, 201–205.
160. Pfeiffer, J.K.; Georgiadis, M.M.; Telesnitsky, A. Structure-based moloney murine leukemia virus reverse transcriptase mutants with altered intracellular direct-repeat deletion frequencies. *J. Virol.* **2000**, *74*, 9629–9636.
161. An, W.; Telesnitsky, A. Frequency of direct repeat deletion in a human immunodeficiency virus type 1 vector during reverse transcription in human cells. *Virology* **2001**, *286*, 475–482.
162. Nguyen, L.A.; Kim, D.H.; Daly, M.B.; Allan, K.C.; Kim, B. Host samhd1 protein promotes hiv-1 recombination in macrophages. *J. Biol. Chem.* **2014**, *289*, 16642.
163. Clavel, F.; Hoggan, M.D.; Willey, R.L.; Strebel, K.; Martin, M.A.; Repaske, R. Genetic recombination of human immunodeficiency virus. *J. Virol.* **1989**, *63*, 1455–1459.
164. St Louis, D.C.; Gotte, D.; Sanders-Buell, E.; Ritchey, D.W.; Salminen, M.O.; Carr, J.K.; McCutchan, F.E. Infectious molecular clones with the nonhomologous dimer initiation sequences found in different subtypes of human immunodeficiency virus type 1 can recombine and initiate a spreading infection *in vitro*. *J. Virol.* **1998**, *72*, 3991–3998.
165. Delwart, E.L.; Shpaer, E.G.; Louwagie, J.; McCutchan, F.E.; Grez, M.; Rubsamen-Waigmann, H.; Mullins, J.I. Genetic relationships determined by a DNA heteroduplex mobility assay: Analysis of hiv-1 env genes. *Science* **1993**, *262*, 1257–1261.
166. Zhu, T.; Wang, N.; Carr, A.; Nam, D.S.; Moor-Jankowski, R.; Cooper, D.A.; Ho, D.D. Genetic characterization of human immunodeficiency virus type 1 in blood and genital secretions: Evidence for viral compartmentalization and selection during sexual transmission. *J. Virol.* **1996**, *70*, 3098–3107.

167. Smyth, R.P.; Schlub, T.E.; Grimm, A.; Venturi, V.; Chopra, A.; Mallal, S.; Davenport, M.P.; Mak, J. Reducing chimera formation during pcr amplification to ensure accurate genotyping. *Gene* **2010**, *469*, 45–51.

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).