Antimicrobial Agents
and Chemotherapy

Prototypical Recombinant Multi-Protease-Inhibitor-Resistant Infectious Molecular Clones of Human Immunodeficiency Virus Type 1

Vici Varghese, Yumi Mitsuya, W. Jeffrey Fessel, Tommy F. Liu, George L. Melikian, David A. Katzenstein, Celia A. Schiffer, Susan P. Holmes and Robert W. Shafer

Antimicrob. Agents Chemother. **2013, 57(9):4290. DOI: 10.1128/AAC.00614-13.**
Published Ahead of Print 24 June 2013.

AAC
Journals.ASM.org

Prototypical Recombinant Multi-Protease-Inhibitor-Resistant Infectious Molecular Clones of Human Immunodeficiency Virus Type 1

Vici Varghese,[a] Yumi Mitsuya,[a] W. Jeffrey Fessel,[b] Tommy F. Liu,[a] George L. Melikian,[a] David A. Katzenstein,[a] Celia A. Schiffer,[c] Susan P. Holmes,[d] Robert W. Shafer[a]

Division of Infectious Diseases, Department of Medicine, Stanford University, Stanford, California, USA[a]; Kaiser-Permanente Medical Care Program—Northern California, San Francisco, California, USA[b]; Department of Biochemistry and Molecular Pharmacology, University of Massachusetts, Worcester, Massachusetts, USA[c]; Department of Statistics, Stanford University, Stanford, California, USA[d]

The many genetic manifestations of HIV-1 protease inhibitor (PI) resistance present challenges to research into the mechanisms of PI resistance and the assessment of new PIs. To address these challenges, we created a panel of recombinant multi-PI-resistant infectious molecular clones designed to represent the spectrum of clinically relevant multi-PI-resistant viruses. To assess the representativeness of this panel, we examined the sequences of the panel's viruses in the context of a correlation network of PI resistance amino acid substitutions in sequences from more than 10,000 patients. The panel of recombinant infectious molecular clones comprised 29 of 41 study-defined PI resistance amino acid substitutions and 23 of the 27 tightest amino acid substitution clusters. Based on their phenotypic properties, the clones were classified into four groups with increasing cross-resistance to the PIs most commonly used for salvage therapy: lopinavir (LPV), tipranavir (TPV), and darunavir (DRV). The panel of recombinant infectious molecular clones has been made available without restriction through the NIH AIDS Research and Reference Reagent Program. The public availability of the panel makes it possible to compare the inhibitory activities of different PIs with one another. The diversity of the panel and the high-level PI resistance of its clones suggest that investigational PIs active against the clones in this panel will retain antiviral activity against most if not all clinically relevant PI-resistant viruses.

Human immunodeficiency virus type 1 (HIV-1) protease is a highly mutable protein. More than one-half of the enzyme's 99 amino acids are subject to variation under protease inhibitor (PI) selection pressure ([1], [2]). In addition, viruses from patients treated unsuccessfully with multiple-PI-containing antiretroviral (ARV) treatment regimens often develop complex patterns of PI resistance amino acid substitutions ([3], [4]). The many genetic manifestations of PI resistance present challenges to researchers who require representative PI-resistant viruses for *in vitro* mechanistic studies and for testing of new inhibitors active against the most clinically relevant multi-PI-resistant variants.

To address these challenges, we created a panel of recombinant infectious molecular virus clones containing clinically derived protease genes with amino acid substitution patterns similar to those in the most common multiple-PI-resistant viruses. The panel was drawn from a repository of cryopreserved plasma samples from patients who received multiple PIs and harbored viruses with multiple PI resistance amino acid substitutions. To assess the validity of this panel, we performed an independent correlation network analysis of PI resistance amino acid substitutions in publicly available sequences from more than 10,000 patients.

MATERIALS AND METHODS

PI resistance amino acid substitutions and phenotypic cutoffs. Amino acid substitutions were defined as differences from the consensus B protease sequence (http://hivdb.stanford.edu/DR/asi/releaseNotes/index.html#consensusbsequences). PI resistance amino acid substitutions were defined as (i) protease substitutions that give rise to HIV-1 variants with reduced susceptibility to one or more PIs in cell culture assays ([5]); (ii) nonpolymorphic substitutions, defined here as those occurring in ≤0.5% of pooled group M viruses from PI-naive persons; and (iii) substitutions occurring in ≥0.5% of virus isolates from PI-experienced patients. Forty-one substitutions at 23 positions met two or more of these criteria: L10F, V11I, L24I, D30N, V32I, L33F, K43T, M46IL, I47AV, G48MV, I50LV, F53L, I54ALMSTV, Q58E, G73ACST, T74P, L76V, V82ACFLST, N83D, I84V, N88DS, L89V, and L90M (see Table S1 in the supplemental material). Each amino acid substitution had a prevalence of ≤0.5% in pooled group M viruses from PI-naive individuals, and with the exception of the V82C substitution, each has been shown to significantly contribute to decreased *in vitro* PI susceptibility ([5]). In addition, each amino acid substitution except for I47A (0.3%), I50L (0.4%), and V82L (0.3%) had a prevalence of ≥0.5% in pooled group M viruses from PI-experienced patients.

High-level resistance was defined according to the Monogram Biosciences (South San Francisco, CA) clinical cutoffs for the PhenoSense assay ([6]): atazanavir-ritonavir (ATV/r) at >6-fold, darunavir-ritonavir (DRV/r) at >90-fold, fosamprenavir-ritonavir (FPV/r) at >11-fold, indinavir-ritonavir (IDV/r) at >10-fold, lopinavir-ritonavir (LPV/r) at >56-fold, nelfinavir (NFV) at >4-fold, saquinavir-ritonavir (SQV/r) at >12-fold, and tipranavir-ritonavir (TPV/r) at >8-fold. The study was performed in accordance with an approved human-subjects protocol.

Creation of the prototypical multi-PI-resistant recombinant infectious molecular clones. Clinical HIV-1 isolates were obtained from cryopreserved remnant plasma samples from patients in California undergo-

Received 29 March 2013 Returned for modification 17 April 2013
Accepted 16 June 2013
Published ahead of print 24 June 2013
Address correspondence to Robert W. Shafer, rshafer@stanford.edu.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/AAC.00614-13.

doi:10.1128/AAC.00614-13

ing routine genotypic resistance testing at Stanford University Hospital. A subset of these patients also had accompanying PhenoSense resistance assays performed for clinical purposes. Available cryopreserved plasma samples were selected for the panel of recombinant infectious molecular clones based on one of the following two criteria: (i) the population-based sequence contained a pattern of PI resistance amino acid substitutions that, based on results from previously reported studies, was consistent with high-level resistance to four or more PIs (2, 3, 5, 7–9), or (ii) an accompanying phenotypic resistance test demonstrated high-level resistance to four or more PIs.

HIV-1 cDNA was generated from RNA extracted from ultracentrifuged plasma samples. An 871-nucleotide amplicon encompassing 3′ *gag*, protease positions 1 to 99, and reverse transcriptase (RT) positions 1 to 24 was amplified by using the thermostable *Pfu* DNA polymerase (Promega, Madison, WI). The 3′ region of *gag* included 243 nucleotides before protease and included the 3′ part of nucleocapsid (NC), the NC/p1 cleavage site, the second spacer peptide, the p1/p6 cleavage site, and the p6 peptide. Amplicons were digested with ApaI and MscI and ligated into the vector pNLPFB (10). Following transformation into competent *Escherichia coli* cells, selected molecular clones were transfected into C8166 cells. Once taken over by syncytia, C8166 cells were cocultured with SupT1 cells. When syncytia were present in the majority of cell clusters, 20 1.0-ml aliquots of cell-free virus were harvested and stored at −70°C.

The amplicon, vector insert, and virus stock were sequenced to confirm near identity with the original population-based plasma virus sequence. Phenotypic resistance testing using the PhenoSense assay was performed for each virus clone. Virus replication was characterized by the Monogram Biosciences replication capacity (RC) assay, the virus stock p24 antigen concentration, and the virus stock 50% tissue culture infectious dose ($TCID_{50}$) during 1 week of MT2 cell culture. The size of the panel was reduced from 25 to 14 by excluding viruses that did not have high-level resistance to four or more PIs or that had a $TCID_{50}$ of <100 infectious units/ml. The DNA plasmids and virus stocks created from each recombinant infectious molecular clone were contributed to the NIH AIDS Research and Reference Reagent Program without restrictions.

Correlation network analyses of published protease sequences. We selected all protease sequences in the Stanford HIV Drug Resistance Database (HIVDB) (11). For individuals with more than one virus sequence, we included only those sequences with a nonredundant pattern of PI resistance amino acid substitutions. The first step in the network analysis was to create a list of positively correlated substitution pairs having a nonparametric Spearman correlation coefficient (rho) of >0.1 and an associated *P* value of <2e−16. To confirm the results of the Spearman correlation analysis, we also calculated the Jaccard correlation coefficient for each pair of substitutions and included only those pairs with an associated *P* value of <2e−9. The Jaccard analysis avoids exaggerating the statistical significance of pairs or rare substitutions (i.e., the double-negative category is large).

To reduce the complexity of the network analysis, we pooled amino acid substitutions at the same position that had highly correlated (defined as a rho value of ≥0.5) individual correlations with each of the amino acid substitutions at other PI resistance positions. Such substitutions included I47V/I47A (rho = 0.6), G48V/G48M (rho = 0.7), I54T/I54A/I54S (rho ≥ 0.8), I54L/I54M (rho = 0.8), G73S/G73T/G73C/G73A (rho ≥ 0.7), V82A/V82T/V82S (rho = 0.5 to 0.7), and V82F/V82L (rho = 0.5). In contrast, other amino acid substitutions at the same position, such as M46I and M46L (rho = −0.5), I50V and I50L (rho = 0), I54ML and I54V (rho = 0.1), and I54ML and I54TAS (rho ≤ −0.8), were treated as separate substitutions.

We used the R package igraph (12) to create an undirected weighted network graph from the adjacency matrix of positively correlated amino acid substitution pairs. In such a network, an edge is created between all significantly correlated amino acid substitution pairs, and the edge length is inversely correlated to the pairs' rho coefficients. The shortest path between two unconnected amino acid substitutions is the smallest sum of the edges linking each substitution. The igraph program "cliques" was used to enumerate tight clusters of amino acid substitutions (cliques) in which each substitution was significantly correlated (rho > 0.1) with each of the other substitutions in the cluster. The igraph program "maximal. cliques" was used to identify cliques that could not be extended by addition of any of the PI resistance amino acid substitutions outside the particular molecular cluster.

For each set of amino acid substitutions in a recombinant infectious molecular clone, we calculated the median of the shortest-path distances between each pair of substitutions. We then randomly sampled with replacement 100,000 sets of PI resistance amino acid substitutions containing the same number of PI resistance amino acid substitutions as the number in the matching clone (range, 3 to 11 substitutions). PI resistance amino acid substitutions were sampled according to their frequency in the HIVDB. The "cluster index" of a clone was defined as the proportion of random samples with a median shortest-path distance greater than the clone's median shortest-path distance.

RESULTS

Recombinant infectious molecular clones: associated clinical data, phenotypic resistance, and amino acid substitutions. Table 1 shows the PI resistance amino acid substitution patterns, phenotypic resistance test results, and replication capacities (RCs) of 14 multi-PI-resistant recombinant infectious molecular clones and the PI treatment history of the patients from whom the cloned samples were obtained. Among the 12 patients for whom the complete PI treatment history was available, the median number of PIs received was 4, and the median duration of treatment was 8 years.

Based on their *in vitro* susceptibility, the clones were categorized into four groups: group 1, with resistance to the earliest-approved PIs (FPV, IDV, NFV, and SQV) and ATV; group 2, with resistance to the earliest-approved PIs, ATV, and LPV; group 3, with resistance to the earliest-approved PIs, ATV, LPV, and TPV; and group 4, with resistance to all PIs. Groups 1, 2, 3, and 4 had medians of 11.5, 15, 15.5, and 24 substitutions and medians of 4.5, 5, 5, and 10 study-defined PI resistance amino acid substitutions, respectively.

Twenty-nine of the 41 study-defined PI resistance amino acid substitutions were present in one or more of the recombinant infectious molecular clones. Substitutions V32I, L33F, M46I, I54V, V82A, I84V, and L90M were found in 5 to 10 of the 14 clones. Substitutions L10F, V11I, L24I, K43T, M46L, I47V, G48V, I54M, Q58E, G73S, and I89V were found in two to four of the clones. Substitutions D30N, I47A, I50V, F53L, I54S, G73T, T74P, V82T, V82L, and N88D were each found in one clone. The accessory PI resistance amino acid substitutions L10IV, K20RIMVT, M36IL, L63P, and A71VTI each occurred in eight or more clones. Among the 14 virus stocks for which RC results were available, the median RC was 14% (range, 3% to 94%).

Each of the 14 clones had one or more previously described compensatory NC/P1 *gag* cleavage site substitutions (13): (i) A431V was present in all clones except clones 634, 1391, and 6585; (ii) K436R was present in clone 1319; and (iii) I437N was present in clones 634, 3972, 6585, and 38129. Eight of the 14 clones had one or more previously described P1/P6 cleavage site substitutions (13): (i) L449VF in clones 634, 3972, and 14311; (ii) R532S in clone 794; and (iii) P453LF in clones 1556, 4307, 18369, and 38129. P6 insertions containing a PTAP motif were present in clones 634 and 6585.

Correlation network analysis of PI resistance amino acid substitution patterns in the HIVDB. Of 61,989 group M HIV-1 protease-containing viruses from 59,455 individuals, 11,351 vi-

TABLE 1 Panel of prototypical multiple-PI-resistant recombinant infectious molecular clones[a]

Group and patient ID	Yr of isolation	PI(s) received	Duration of treatment (yr)	PI resistance amino acid substitutions[b]	Cluster index (%)	Fold change								RC[c] (%)
						NFV	FPV	SQV	IDV	ATV	LPV	TPV	DRV	
Resistance to earliest-approved PIs and ATV														
7932	1999	IDV	3	10I, 14R, **24I**, 35D, 36I, 37E, 41K, **46L**, **54V**, 63P, 71V, **82A**	**80**	**27**	1.8	7.6	**17**	**19**	29	0.8	0.7	26
18369	2004	IDV, NFV, SQV/r	>5	10I, 37D, **46I**, 64P, 71T, 72V, 77I, **84V**, 89M, **90M**	**84**	**45**	**16**	**21**	**28**	**7.4**	15	1.9	3.5	47
1319	2000	IDV, SQV/r	5	10I, 37D, **48V**, **53L**, 71T, 77I, **54V**, **82A**, **90M**, 93L	**12**	**29**	1.4	**>200**	**38**	**46**	34	0.2	0.4	15
1329	2003	IDV, NFV, SQV/r, LPV/r	6	**10F**, 15V, **30N**, **33F**, 35D, 36L, 37E, **43T**, 60E, 62V, 63P, **84V**, **88D**, **90M**	**94**	**>200**	**20**	**>200**	**15**	**99**	30	3.1	7.2	63
Resistance to earliest-approved PIs, ATV, and LPV														
2213	2004	IDV, NFV, IDV/r, LPV/r	7	20R, **24I**, 35D, 36I, 37E, **46I**, **54V**, 61N, 62V, 63P, 64V, 71V, 72V, 74S, **76V**, **82T**, 93L	**48**	**31**	**17**	8.2	**>200**	**7.7**	**>200**	1.4	3.6	31
634	2006	IDV, SQV/r, IDV/r, LPV/r	10	10I, 20R, **33F**, 34T, 35D, 36I, 37D, **43T**, **46I**, **48V**, **50V**, **54S**, 57K, 62V, 63P, 71V, 72V, **82A**	**19**	**130**	**112**	**>200**	**>200**	**109**	**>200**	1.6	39	63
27759	2006	PIs, LPV/r	>3	10I, 13V, **32I**, **33F**, 36I, 41K, **47A**, 63P, 70E, 71V, **82A**, **90M**, 93L	**79**	**9**	**32**	0.5	**21**	4.4	**>200**	3.5	9.2	48
Resistance to earliest-approved PIs, ATV, LPV, and TPV														
4307	2005	IDV/r, LPV/r, TPV/r	9	10V, 13V, 15V, 19I, 20T, **32I**, 36I, **46I**, **47V**, **58E**, 63P, 64V, **84V**, 92K	**83**	**29**	**26**	5.5	**17**	**33**	**78**	**16**	25	13
794	2007	IDV, SQV/r, LPV/r	12	**10F**, 16A, 19I, 20R, **33F**, 35D, 36I, 37D, **43T**, **46L**, **54V**, 55R, 60E, 62V, 63P, 71V, **82A**, **84V**, **90M**, 93L	**60**	**64**	**62**	**>200**	**63**	**59**	**127**	**24**	14	NA
3972	2006	IDV, SQV, APV, IDV/r, ATV/r, TPV/r	11	10I, 13V, 20I, 33V, 35D, 36I, 37T, **46I**, **54V**, **58E**, 62V, 63P, 64V, 71V, **74P**, **82L**, **90M**	**21**	**93**	**11**	7.4	**47**	**110**	36	**>55**	5.2	13
Resistance to earliest-approved PIs														
6585	2002	IDV, SQV/r, NFV APV/r, LPV/r	7	10V, 12V, 13V, 15V, 20M, **32I**, **33F**, **43T**, **46I**, **47V**, **54M**, 60E, 61N, 62V, 63P, 67Y, 69K, 71I, 72L, **73S**, 77I, **82A**, **89V**, **90M**	**98**	**76**	**>200**	**12**	**88**	**88**	**>200**	**12**	**112**	94
14311	2006	APV/r, ATV/r, LPV/r, FPV/LPV/r	7	**10F**, **11I**, 13V, 19Q, **32I**, 33V, 35A, 36I, **46I**, **47V**, **54M**, 57K, 62V, 63P, 64V, **73T**, 74A, **84V**, **89V**, **90M**	**99**	**50**	**>200**	**22**	**48**	**120**	**>200**	**8.0**	**>200**	4
1556	2006	IDV, NFV, SQV/r, APV/r, LPV/r, TPV/r, DRV/r	9	**10F**, **11I**, 12P, 13V, 15V, 19P, 20T, **33F**, 35G, 36I, 41K, **43T**, **46L**, **54V**, 62V, 63P, 70T, 71I, **73S**, **82A**, 79A, **84V**, **89V**, **90M**	**90**	**52**	**148**	**66**	**22**	**90**	**70**	**12**	**127**	NA
38129	2007	IDV, ATV/r, DRV/r	9	10I, **11I**, 12K, 13V, 20V, **32I**, **33F**, 35G, 36I, 37D, **46I**, **47V**, **54M**, 57K, **58E**, 63P, 64V, 66V, 71V, **73S**, **84V**, **89V**, **90M**	**99**	**47**	**>200**	**36**	**24**	**101**	**>200**	**>55**	**>200**	3

---

[a] PI resistance was defined according to the sole or upper clinical cutoff provided on the Monogram PhenoSense report: NFV at 3.6-fold, FPV/r at 11-fold, SQV/r at 12-fold, IDV/r at 10-fold, ATV/r at 5.2-fold, LPV/r at 50-fold, TPV/r at 8-fold, and DRV at 90-fold. Values greater than the PhenoSense cutoff are indicated in boldface type. NA, not applicable.

[b] Study-defined PI resistance amino acid substitutions are indicated in boldface type. The HIVDB (http://hivdb.stanford.edu/DR/asi/releaseNotes/index.html#consensusbsequences) contains the consensus subtype B sequence.

[c] RC, replication capacity determined by using the PhenoSense assay. The RC correlated with the 50% tissue culture infectious dose ($TCID_{50}$), which ranged from 300 to 36,000 infectious units after 1 week of culture with a median of 3,000 infectious units.

ruses from 10,050 individuals had one or more PI resistance amino acid substitutions. Of these 11,351 variants, 9.5% were from PI-naive individuals, 20.2% were from individuals who had received one PI, 10.0% were from individuals who had received two PIs, 7.3% were from individuals who had received three PIs, 8.9% were from individuals who had received four or more PIs, and 44.1% were from PI-treated individuals for whom the exact number of PIs received was not known. A total of 1.9% of isolates were obtained prior to 1996, 31.3% were obtained between 1996 and 2000, 58.8% were obtained between 2001 and 2005, and 9.4% were obtained between 2006 and 2012. Nearly 14% of the isolates belonged to a non-B subtype, including 3.5% of isolates in subtype F, 2.8% in subtype C, 1.7% in subtype G, 1.3% in CRF01_AE, 1.2% in subtype A, 1.1% in CRF02_AG, 1.1% in subtype D, and 1.1% belonging to miscellaneous other subtypes and circulating recombinant forms (CRFs).

Twenty-five percent of sequences had one study-defined PI resistance amino acid substitution, 17% had two PI resistance amino acid substitutions, 16% had three PI resistance amino acid substitutions, 14% had four PI resistance amino acid substitutions, 11% had five PI resistance amino acid substitutions, and 17% had six or more PI resistance amino acid substitutions. The 11,351 isolates had 3,139 unique patterns of PI resistance amino acid substitutions. L90M was the most common PI resistance amino acid substitution, occurring in 51% of isolates with one or more PI resistance amino acid substitutions; substitutions I54V, V82A, M46I, I84V, L33F, G73S, M46L, and L10F occurred in 10% to 34% of isolates; substitutions D30N, L24I, N88D, V32I, F53L, Q58E, I47V, G48V, L89V, V82T, L76V, I54L, I54M, G73T, I50V, V82F, T74P, and N88S occurred in 2% to 9% of isolates; and substitutions G73C, I54A, V82S, I54T, I54S, V82C, I50L, G48M, G73A, I47A, and V82L occurred in <2% of isolates. Although most PI resistance amino acid substitutions increased in frequency with each calendar year, the L33F, V11I, I47V, Q58E, L10F, I84V, K43T, and I54ML substitutions demonstrated the greatest increase in prevalence per year since 2000.

Table 2 shows those amino acid substitution pairs with the highest levels of pairwise correlation using the Spearman correlation coefficient (rho). The two most highly correlated substitutions, D30N and N88D, displayed no significant positive correlations with other PI resistance amino acid substitutions. In contrast, most of the other PI resistance amino acid substitutions were correlated ($\text{rho} > 0.1$; $P < 1\text{e}{-16}$) with two or more other PI resistance amino acid substitutions. The V32I, L33F, I47V, V82A, and I84V substitutions were significantly correlated with 10 to 11 other PI resistance amino acid substitutions; M46I, I54V, I54M, L89V, and L90M were correlated with 7 to 9 other PI resistance amino acid substitutions; and L10F, K43T, F53L, I54L, G73S, and L90M were correlated with 5 to 6 other PI resistance amino acid substitutions. The I50L, Q58E, V82FL, N83D, and N88S substitutions were not correlated with any other PI resistance amino acid substitution.

Figure 1 shows the protease inhibitor amino acid substitution patterns of group 1 to 4 recombinant infectious molecular clones superimposed on the correlation network created from the adjacency matrix of amino acid substitution correlations in more than 10,000 published protease sequences. As indicated in Materials and Methods, the following amino acid substitutions at the same position were represented by a single amino acid: I47VA, G48VM, I54TAS, I54ML, G73STCA, and V82ATS. The graphs do not show the PI resistance amino acid substitutions I50L, Q58E, V82FL, N83D, and N88S, which were not correlated with any other PI resistance amino acid substitutions. Substitutions D30N and N88D, which correlated only with one another, were not shown either.

Three of the four group 1 clones and the pan-PI-resistant group 4 clones had cluster indexes of between 80 and 99%, indicating that the medians of their pairwise shortest-path distances were lower than those of matched random substitution patterns (Table 1). In contrast, the group 2 and group 3 clones had clustering indexes that did not differ from matched random substitution patterns.

Table 3 lists the 27 maximal cliques generated from the matrix of significantly correlated pairs of PI resistance amino acid substitutions. Two maximal cliques contained 2 amino acid substitutions, 7 contained 3 amino acid substitutions, and 18 contained 4 or 5 amino acid substitutions. Nine of the 27 substitution patterns comprising a maximal clique were found in a clinical virus sample without any other study-defined PI resistance amino acid substitution. The remaining 18 patterns occurred only in the presence of one or more additional study-defined PI resistance substitutions. The substitutions comprising all but four of the cliques were present in one or more of the recombinant infectious molecular clones.

DISCUSSION

We created a panel of multi-PI-resistant clones and assessed its validity by performing an independent correlation network analysis of PI resistance amino acid substitutions in publicly available sequences from more than 10,000 patients. The clinical samples used to create the panel were selected based upon sample availability and findings from previously reported studies. The correlation network analysis was performed to assess the concordance of the amino acid substitution patterns of the panel's clones with the amino acid substitution patterns in publicly available protease sequences.

In the 1990s, three research groups contributed infectious PI-resistant clones to the NIH AIDS Research and Reference Reagent Program. These clones contained the following PI resistance amino acid substitutions: (i) G48V plus L90M (14), (ii) M46I plus L63P plus V82T plus I84V with or without L10R (15); and (iii) V82A, V82F, I84V, and V82F plus I84V (16). The panel that we have created and submitted to the NIH AIDS Research and Reference Reagent Program increases the number and variety of clinically relevant PI-resistant clones available to researchers.

The large number of PI resistance amino acid substitutions makes it challenging to identify representative multi-PI-resistant viruses. Indeed, the publicly available sequence data set contained more than 3,000 permutations of the 41 PI resistance amino acid substitutions used in our analysis. We performed a correlation network analysis because of the suitability of this approach for representing the overlapping nature of the patterns of protease amino acid substitutions in multi-PI-resistant viruses. In contrast, other more commonly used clustering approaches, such as hierarchical and k-medioid clustering, fail to adequately represent overlapping patterns.

The correlation network analysis revealed an underlying structure in the patterns of PI resistance amino acid substitutions. It showed two main clusters of amino acid substitutions: one large

TABLE 2 Positively correlated PI resistance amino acid substitutions ranked by Spearman's correlation coefficient

Substitution 1	Substitution 2	No. of isolates with substitution 1 alone	No. of isolates with substitution 2 alone	No. of isolates with substitution 1 + substitution 2	No. of isolates with neither substitution	Spearman's rho	*P* value
D30N	N88D	385	176	710	10,112	0.70	<2e−16
V32I	I47V	384	259	467	10,273	0.56	<2e−16
I54V	V82A	1,567	1,416	2,199	6,201	0.40	<2e−16
I47V	I54M	508	187	218	10,470	0.37	<2e−16
G48V	I54T	540	39	120	10,684	0.36	<2e−16
G73S	L90M	136	4,475	1,123	5,649	0.28	<2e−16
M46L	V82A	579	2,607	1,008	7,189	0.28	<2e−16
V32I	I54M	670	224	181	10,308	0.27	<2e−16
G48V	I54S	586	33	74	10,690	0.26	<2e−16
L24I	I54V	266	3,113	653	7,351	0.24	<2e−16
M46I	I47V	3,073	188	538	7,584	0.24	<2e−16
L33F	I54L	1,630	185	273	9,295	0.24	<2e−16
G48V	V82A	164	3,119	496	7,604	0.23	<2e−16
L24I	V82A	301	2,997	618	7,467	0.23	<2e−16
M46I	I84V	2,412	1,140	1,199	6,632	0.21	<2e−16
L10F	I84V	750	1,759	580	8,294	0.21	<2e−16
I84V	L90M	714	3,973	1,625	5,071	0.21	<2e−16
G73T	I84V	139	2,087	252	8,905	0.21	<2e−16
V32I	M46I	300	3,060	551	7,472	0.20	<2e−16
I54M	L89V	290	476	115	10,502	0.20	<2e−16
M46I	L76V	3,254	107	357	7,665	0.20	<2e−16
L24I	M46L	584	1,252	335	9,212	0.19	<2e−16
I47V	L89V	575	440	151	10,217	0.18	<2e−16
V32I	I54L	710	317	141	10,215	0.18	<2e−16
K43T	V82A	297	3,120	495	7,471	0.18	<2e−16
I47V	I54L	599	331	127	10,326	0.18	<2e−16
V32I	L89V	692	432	159	10,100	0.17	<2e−16
M46L	I54V	742	2,921	845	6,875	0.17	<2e−16
G73T	L90M	22	5,229	369	5,763	0.17	<2e−16
L10F	M46I	640	2,921	690	7,132	0.16	<2e−16
L33F	I54M	1,713	215	190	9,265	0.16	<2e−16
I54L	I84V	224	2,105	234	8,820	0.16	<2e−16
L33F	I47V	1,622	445	281	9,035	0.15	<2e−16
V32I	L33F	538	1,590	313	8,942	0.15	<2e−16
L33F	T74P	1,252	1,688	651	7,792	0.15	<2e−16
L33F	I84V	1,771	108	132	9,372	0.15	<2e−16
L33F	K43T	1,609	498	294	8,982	0.15	<2e−16
G73S	I84V	786	1,866	473	8,258	0.15	<2e−16
F53L	I54V	322	3,315	451	7,295	0.15	<2e−16
G48V	V82S	604	112	56	10,611	0.14	<2e−16
V32I	K43T	685	626	166	9,906	0.14	<2e−16
M46I	G73T	3,352	132	259	7,640	0.14	<2e−16
F53L	V82A	341	3,183	432	7,427	0.14	<2e−16
L33F	V82A	1,030	2,742	873	6,738	0.14	<2e−16
M46I	L90M	1,480	3,467	2,131	4,305	0.13	<2e−16
G73S	L89V	1,088	420	171	9,704	0.13	<2e−16
I47V	F53L	586	633	140	10,024	0.13	<2e−16
I84V	L89V	2,087	339	252	8,705	0.13	<2e−16
L10F	L89V	1,164	425	166	9,628	0.12	<2e−16
M46I	I54M	3,365	159	246	7,613	0.12	<2e−16
I54V	V82S	3,635	37	131	7,580	0.12	<2e−16
I54T	V82A	36	3,492	123	7,732	0.12	<2e−16
T74P	I84V	114	2,213	126	8,930	0.12	<2e−16
L24I	V82S	863	112	56	10,352	0.11	<2e−16
V11I	L89V	486	499	92	10,306	0.11	<2e−16
G48V	I50V	597	222	63	10,501	0.11	<2e−16
K43T	I47V	663	597	129	9,994	0.11	<2e−16
I54M	I84V	227	2,161	178	8,817	0.11	<2e−16
L89V	L90M	161	5,168	430	5,624	0.11	<2e−16

(Continued on following page)

TABLE 2 (Continued)

Substitution 1	Substitution 2	No. of isolates with substitution 1 alone	No. of isolates with substitution 2 alone	No. of isolates with substitution 1 + substitution 2	No. of isolates with neither substitution	Spearman's rho	*P* value
L10F	G73T	1,212	273	118	9,780	0.11	<2e−16
L10F	I47V	1,149	545	181	9,508	0.11	<2e−16
V32I	F53L	712	634	139	9,898	0.11	<2e−16
V32I	V82A	430	3,194	421	7,338	0.11	<2e−16
I54S	V82A	18	3,526	89	7,750	0.11	<2e−16
I54L	L90M	113	5,253	345	5,672	0.11	<2e−16
L10F	L33F	969	1,542	361	8,511	0.10	<2e−16
L33F	F53L	1,665	535	238	8,945	0.10	<2e−16
K43T	I54V	393	3,367	399	7,224	0.10	<2e−16
I54M	L90M	99	5,292	306	5,686	0.10	<2e−16
V11I	G73S	436	1,117	142	9,688	0.10	<2e−16
I47V	L90M	230	5,102	496	5,555	0.10	<2e−16
V11I	I84V	359	2,120	219	8,685	0.10	<2e−16
L33F	M46L	1,494	1,178	409	8,302	0.10	<2e−16
I54V	V82T	3,490	231	276	7,386	0.10	<2e−16
M46I	G73S	3,054	702	557	7,070	0.10	<2e−16

cluster (cluster 1) consisted of substitutions surrounding V82ATS that included, in order of decreasing frequency, I54V, M46L, L24I, G48VM, and I54TAS. The second large cluster (cluster 2) included, in order of decreasing frequency, the L90M, M46I, G73STCA, I84V, I54ML, V32I, I47VA, L10F, L89V, and V11I substitutions. V82A and V32I were the only positively correlated PI resistance amino acid substitutions that were present in different clusters (V82A in cluster 1 and V32I in cluster 2).

A third set of PI resistance amino acid substitutions, L33F, K43T, and F53L, consisted of indirectly linking substitutions because each of these substitutions correlated with two or more cluster 1 and two or more cluster 2 substitutions. Although the I50V and L76V substitutions are among the most potent DRV resistance amino acid substitutions, they rarely occur in pan-PI-resistant viruses, possibly because they increase susceptibility to one or more PIs and reduce protease function (17–20). The occurrence of the L76V substitution in combination M46I was described previously (21). However, the occurrence of the I50V substitution in combination with the G48V and/or I54TAS substitution has not been previously described.

Biophysical interactions are likely to explain some of the patterns of observed PI resistance amino acid substitutions. In cluster 1, only G48 directly contacts substrates (22); thus, most of the cluster 1 substitutions are consistent with the substrate envelope hypothesis in that they preserve substrate recognition while decreasing inhibitor binding (23). In cluster 2, several residues, including V32, I47, and I84, are involved in substrate recognition (24). However, only the substitution V32I results in an increase in residue volume, potentially impinging on substrate recognition. Most of the remaining PI resistance amino acid substitutions alter protease flap dynamics or modulate the hydrophobic sliding of the enzyme core (25).

The panel of recombinant infectious molecular clones comprised 29 of the 41 study-defined PI resistance amino acid substitutions, 27 of the 33 substitutions that were significantly correlated with another PI resistance substitution, and 23 of the 27 amino acid substitution cliques. Based on their phenotypic properties, the clones were classified into four groups with increasing cross-resistance to the PIs most commonly used for salvage therapy: LPV, TPV, and DRV (26). However, TPV-resistant viruses that retain susceptibility to LPV and DRV-resistant viruses that retain susceptibility to TPV do occur and are not present in our panel (26). In addition, two of the group 1 viruses were susceptible to FPV, and one of the group 3 viruses was susceptible to SQV.

Although the median RC of the panel was 20% (range, 3% to 94%), each of the 14 virus stocks had 50% tissue culture infectious doses (TCID$_{50}$) high enough for *in vitro* susceptibility testing (range, 300 to 36,000 infectious units) of novel compounds. Each of the clones also had previously described compensatory *gag* cleavage site amino acid substitutions, changes that are often required for the replication of drug-resistant mutant viruses.

The four group 1 recombinant viruses (resistance to the earliest-approved PIs and ATV) had amino acid substitutions that were localized within cluster 1 or cluster 2 but not both clusters (Fig. 1A). Five of the six group 2 and 3 recombinant viruses (resistance to the earliest-approved PIs, ATV, LPV, and/or TPV) had amino acid substitutions that included both cluster 1 and cluster 2 amino acid substitutions (Fig. 1B and C). Two of the four group 4 recombinant viruses (resistance to all PIs) contained many cluster 2 amino acid substitutions, and two contained many cluster 2 amino acid substitutions in combination with the cluster 1 amino acid substitutions V82A with or without M46L and I54V (Fig. 1D). The progression to pan-PI resistance may therefore be associated with fewer cluster 1 and more cluster 2 amino acid substitutions.

In conclusion, we have created a panel of 14 recombinant infectious molecular clones comprising most PI resistance amino acid substitutions and most clusters of PI resistance amino acid substitutions. It also contains viruses with the most common phenotypic patterns of high-level multiple-PI resistance. The public availability of the panel makes it possible to compare the inhibitory activities of different PIs with one another. The diversity of the panel and the high level of PI resistance of its clones make it likely that investigational PIs active against the clones in this panel

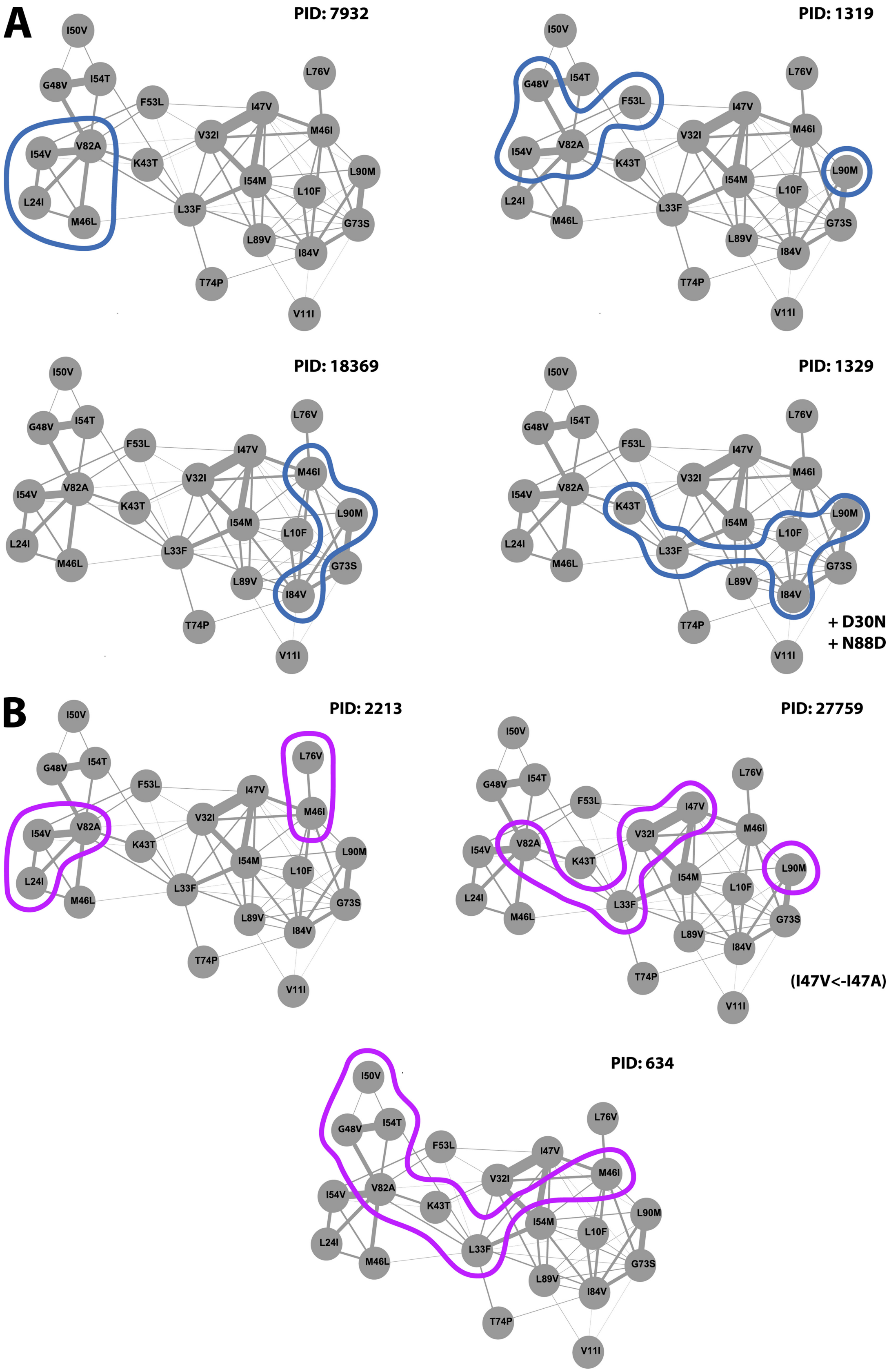

FIG 1 Protease inhibitor (PI) amino acid substitution patterns in the panel of 14 multiple-PI-resistant recombinant infectious clones. (A) Group 1 clones; (B) group 2 clones; (C) group 3 clones; (D) group 4 clones. The amino acid substitutions within a clone are circled and shown against the backdrop of the correlation network created from the adjacency matrix of correlations between PI resistance amino acid substitutions in more than 10,000 independent publicly available protease sequences. Each edge in the network represents a strongly significant correlation (Spearman's rho of >0.1; $P < 2e-16$) between a pair of substitutions, and the thickness of the edge is proportional to the strength of the correlation (rho). PI resistance amino acid substitutions, indicated in plain text, are in the clone's sequence but were not correlated with other PI resistance amino acid substitutions in theTable 1.> network. PID, patient identification.

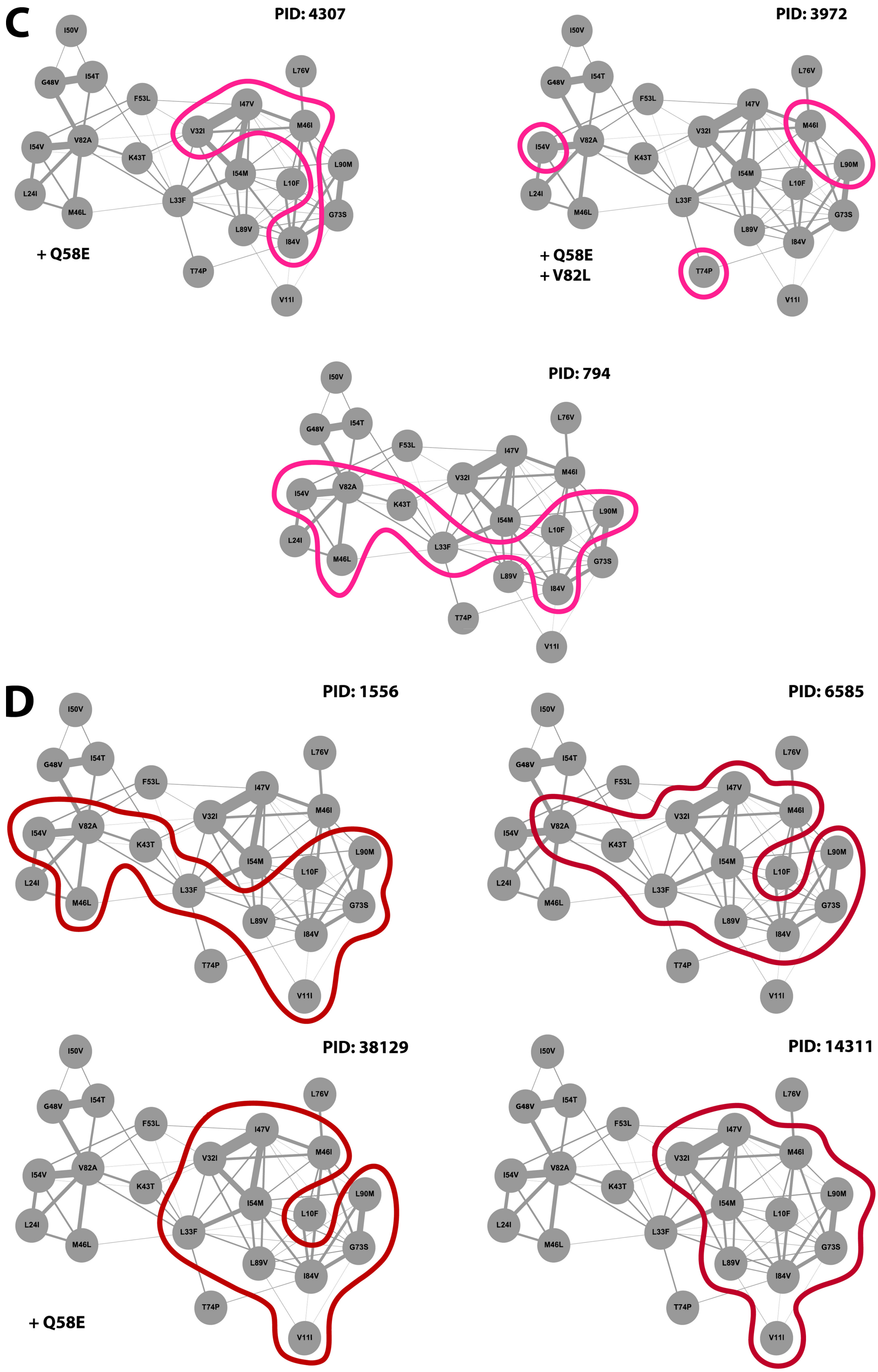

FIG 1 continued

TABLE 3 Presence of maximal cliques of PI resistance amino acid substitutions created from strongly correlated pairs of PI resistance amino acid substitutions in the panel of recombinant infectious molecular clones[a]

Substitution pattern	No. of sequences in the HIVDB without additional DRSs[b]	No. of sequences in the HIVDB with or without additional DRSs[c]	Group in PI resistance panel
30N, 88D	401	525	1
46I, 76V	27	234	2
43T, 54V, 82A	6	186	3
53L, 54V, 82A	11	174	1
33F, 46L, 82A	0	169	3
48V, 54T, 82A	51	131	2
33F, 54T, 82A	3	78	None
33F, 74P, 84V	0	78	None
48V, 50V, 54T	0	26	3
24I, 46L, 54V, 82A	89	150	1
32I, 46I, 47V, 54M	0	134	4
32I, 47V, 54M, 89V	0	71	4
32I, 33F, 47V, 53L	0	42	None
32I, 33F, 53L, 82A	0	38	None
32I, 33F, 43T, 82A	0	31	4
11I, 73S, 84V, 89V	0	30	4
46I, 47V, 54M, 73S, 90M	0	70	4
46I, 54M, 73S, 84V, 90M	2	62	4
54M, 73S, 84V, 89V, 90M	0	45	4
47V, 54M, 73S, 89V, 90M	0	45	4
32I, 33F, 43T, 47V, 54M	0	30	4
10F, 46I, 54M, 73S, 84V	1	26	4
10F, 33F, 54M, 73S, 84V	0	23	4
10F, 54M, 73S, 84V, 89V	0	21	4
10F, 46I, 47V, 54M, 73S	0	17	4
10F, 47V, 54M, 73S, 89V	0	14	4
10F, 33F, 47V, 54M, 73S	0	11	4

[a] This analysis was done after pooling amino acid substitutions at the same position displaying highly similar patterns of correlation: I47VA, G48VM, I54TAS, I54ML, G73STCA, and V82AST.

[b] Number of persons having a virus with the indicated amino acid substitution pattern but none of the other study-defined amino acid substitutions. DRSs, drug resistance substitutions.

[c] Number of persons having a virus with the indicated amino acid substitution pattern with or without other study-defined amino acid substitutions.

will retain antiviral activity against most, if not all, clinically relevant PI-resistant viruses.

ACKNOWLEDGMENT

This research was supported by funding from AI068581 and AI46148.

REFERENCES

1. **Ceccherini-Silberstein F, Erba F, Gago F, Bertoli A, Forbici F, Bellocchi MC, Gori C, D'Arrigo R, Marcon L, Balotta C, Antinori A, Monforte AD, Perno CF.** 2004. Identification of the minimal conserved structure of HIV-1 protease in the presence and absence of drug pressure. AIDS **18:** 11–19.
2. **Wu TD, Schiffer CA, Gonzales MJ, Taylor J, Kantor R, Chou S, Israelski D, Zolopa AR, Fessel WJ, Shafer RW.** 2003. Mutation patterns and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments. J. Virol. **77:**4836–4847.
3. **Babrzadeh F, Varghese V, Pacold M, Liu TF, Nyren P, Schiffer C, Fessel WJ, Shafer RW.** 2013. Collinearity of protease mutations in HIV-1 samples with high-level protease inhibitor class resistance. J. Antimicrob. Chemother. **68:**414–418.
4. **Baxter JD, Schapiro JM, Boucher CA, Kohlbrenner VM, Hall DB, Scherer JR, Mayers DL.** 2006. Genotypic changes in human immunodeficiency virus type 1 protease associated with reduced susceptibility and virologic response to the protease inhibitor tipranavir. J. Virol. **80:**10794–10801.
5. **Rhee SY, Taylor J, Fessel WJ, Kaufman D, Towner W, Troia P, Ruane P, Hellinger J, Shirvani V, Zolopa A, Shafer RW.** 2010. HIV-1 protease mutations and protease inhibitor cross-resistance. Antimicrob. Agents Chemother. **54:**4253–4261.
6. **Petropoulos CJ, Parkin NT, Limoli KL, Lie YS, Wrin T, Huang W, Tian H, Smith D, Winslow GA, Capon DJ, Whitcomb JM.** 2000. A novel phenotypic drug susceptibility assay for human immunodeficiency virus type 1. Antimicrob. Agents Chemother. **44:**920–928.
7. **Mitsuya Y, Winters MA, Fessel WJ, Rhee SY, Hurley L, Horberg M, Schiffer CA, Zolopa AR, Shafer RW.** 2006. N88D facilitates the co-occurrence of D30N and L90M and the development of multidrug resistance in HIV type 1 protease following nelfinavir treatment failure. AIDS Res. Hum. Retroviruses **22:**1300–1305.
8. **Agniswamy J, Shen CH, Aniana A, Sayer JM, Louis JM, Weber IT.** 2012. HIV-1 protease with 20 mutations exhibits extreme resistance to clinical inhibitors through coordinated structural rearrangements. Biochemistry **51:**2819–2828.
9. **Saskova KG, Kozisek M, Rezacova P, Brynda J, Yashina T, Kagan RM, Konvalinka J.** 2009. Molecular characterization of clinical isolates of human immunodeficiency virus resistant to the protease inhibitor darunavir. J. Virol. **83:**8810–8818.
10. **Imamichi T, Berg SC, Imamichi H, Lopez JC, Metcalf JA, Falloon J, Lane HC.** 2000. Relative replication fitness of a high-level 3′-azido-3′-deoxythymidine-resistant variant of human immunodeficiency virus type 1 possessing an amino acid deletion at codon 67 and a novel substitution (Thr→Gly) at codon 69. J. Virol. **74:**10958–10964.
11. **Rhee SY, Gonzales MJ, Kantor R, Betts BJ, Ravela J, Shafer RW.** 2003. Human immunodeficiency virus reverse transcriptase and protease sequence database. Nucleic Acids Res. **31:**298–303.
12. **Csardi G, Nepusz T.** 2006. The igraph software package for complex network research. InterJournal Complex Systems 1695. http://igraph.sf.net/.
13. **Fun A, Wensing AM, Verheyen J, Nijhuis M.** 2012. Human immunodeficiency virus Gag and protease: partners in resistance. Retrovirology **9:**63. doi:10.1186/1742-4690-9-63.
14. **Jacobsen H, Yasargil K, Winslow DL, Craig JC, Krohn A, Duncan IB, Mous J.** 1995. Characterization of human immunodeficiency virus type 1 mutants with decreased sensitivity to proteinase inhibitor Ro 31-8959. Virology **206:**527–534.
15. **Condra JH, Schleif WA, Blahy OM, Gabryelski LJ, Graham DJ, Quintero JC, Rhodes A, Robbins HL, Roth E, Shivaprakash M, Titus D, Yang T, Teppler H, Squires KE, Deutsch PJ, Emini EA.** 1995. In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors. Nature **374:**569–571.
16. **Otto MJ, Garber S, Winslow DL, Reid CD, Aldrich P, Jadhav PK, Patterson CE, Hodge CN, Cheng YS.** 1993. In vitro isolation and identification of human immunodeficiency virus (HIV) variants with reduced sensitivity to C-2 symmetrical inhibitors of HIV type 1 protease. Proc. Natl. Acad. Sci. U. S. A. **90:**7543–7547.
17. **Louis JM, Zhang Y, Sayer JM, Wang YF, Harrison RW, Weber IT.** 2011. The L76V drug resistance mutation decreases the dimer stability and rate of autoprocessing of HIV-1 protease by reducing internal hydrophobic contacts. Biochemistry **50:**4786–4795.
18. **Wiesmann F, Vachta J, Ehret R, Walter H, Kaiser R, Sturmer M, Tappe A, Daumer M, Berg T, Naeth G, Braun P, Knechten H.** 2011. The L76V mutation in HIV-1 protease is potentially associated with hypersusceptibility to protease inhibitors atazanavir and saquinavir: is there a clinical advantage? AIDS Res. Ther. **8:**7. doi:10.1186/1742-6405-8-7.
19. **Bethell R, Scherer J, Witvrouw M, Paquet A, Coakley E, Hall D.** 2012. Short communication: phenotypic protease inhibitor resistance and cross-resistance in the clinic from 2006 to 2008 and mutational prevalences in HIV from patients with discordant tipranavir and darunavir susceptibility phenotypes. AIDS Res. Hum. Retroviruses **28:** 1019–1024.
20. **Young TP, Parkin NT, Stawiski E, Pilot-Matias T, Trinh R, Kempf DJ, Norton M.** 2010. Prevalence, mutation patterns, and effects on protease inhibitor susceptibility of the L76V mutation in HIV-1 protease. Antimicrob. Agents Chemother. **54:**4903–4906.
21. **Nijhuis M, Wensing AM, Bierman WF, de Jong D, Kagan R, Fun A, Jaspers CA, Schurink KA, van Agtmael MA, Boucher CA.** 2009. Failure of treatment with first-line lopinavir boosted with ritonavir can be explained by novel resistance pathways with protease mutation 76V. J. Infect. Dis. **200:**698–709.
22. **Hong L, Zhang XC, Hartsuck JA, Tang J.** 2000. Crystal structure of an in vivo HIV-1 protease mutant in complex with saquinavir: insights into the mechanisms of drug resistance. Protein Sci. **9:**1898–1904.

23. **King NM, Prabu-Jeyabalan M, Nalivaika EA, Schiffer CA.** 2004. Combating susceptibility to drug resistance: lessons from HIV-1 protease. Chem. Biol. **11:**1333–1338.
24. **Ozen A, Haliloglu T, Schiffer CA.** 2011. Dynamics of preferential substrate recognition in HIV-1 protease: redefining the substrate envelope. J. Mol. Biol. **410:**726–744.
25. **Ali A, Bandaranayake RM, Cai Y, King NM, Kolli M, Mittal S, Murzycki JF, Nalam MN, Nalivaika EA, Ozen A, Prabu-Jeyabalan MM, Thayer K, Schiffer CA.** 2010. Molecular basis for drug resistance in HIV-1 protease. Viruses **2:**2509–2535.
26. **Doherty KM, Nakka P, King BM, Rhee SY, Holmes SP, Shafer RW, Radhakrishnan ML.** 2011. A multifaceted analysis of HIV-1 protease multidrug resistance phenotypes. BMC Bioinformatics **12:**477. doi:10.1186/1471-2105-12-477.