Available online at www.sciencedirect.com
SCIENCE DIRECT
Antiviral Research 71 (2006) 42–52
Antiviral Research
www.elsevier.com/locate/antiviral

Increasing prevalence of HIV-1 protease inhibitor-associated mutations correlates with long-term non-suppressive protease inhibitor treatment

R.M. Kagan [a,*], P.K. Cheung [b], T.K. Huard [c], M.A. Lewinski [a]

[a] *Department of Infectious Diseases, Quest Diagnostics Nichols Institute, 33608 Ortega Hwy, San Juan Capistrano, CA 92675, United States*
[b] *BC Centre for Excellence in HIV/AIDS Vancouver, British Columbia, CanadaV6Z 1Y6*
[c] *Department of Molecular Diagnostics, Quest Diagnostics Nichols Institute, Chantilly, VA 20153, United States*

Received 5 January 2006; accepted 28 February 2006

Abstract

Treatment of human immunodeficiency virus type 1 with protease inhibitors (PIs) is associated with the emergence of resistance-associated mutations. Treatment-characterized datasets have been used to identify novel treatment-associated protease mutations. In this study, we utilized two large reference laboratory databases (>115 000 viral sequences) to identify non-established resistance-associated protease mutations. We found 20 non-established protease mutations occurring in 82% of viruses with a PI resistance score of 4–7, 62% of viruses with a resistance score of 1–3, and 35% of viruses with no predicted PI resistance. We correlated mutational prevalence to treatment duration in a treatment-characterized dataset of 2161 patients undergoing non-suppressive PI therapy. In the non-suppressed dataset, 24 mutations became more prevalent and three mutations became less prevalent after more than 48 months of non-suppressive PI-therapy. Longer durations of non-suppressive treatment correlated with higher PI resistance scores. Mutations at eight non-established positions that were more common in viruses with the longest duration of non-suppressive therapy were also more common in viruses with the highest PI resistance score. Covariation analysis of 3036 protease amino acid substitutions identified 75 positive and nine negative correlations between resistance associated positions. Our findings support the utility of reference laboratory datasets for surveillance of mutation prevalence and covariation.

Keywords: HIV-1; Protease resistance mutations; Protease inhibitors; Treatment duration; Covariation; Comutation

1. Introduction

HIV-1 protease is essential for cleavage of the viral polyprotein precursor of the gag and pol viral proteins (Kohl et al., 1988; Peng et al., 1989). Antiretroviral drugs (ARVs) targeting the active site of HIV-1 protease potently inhibit viral replication in infected individuals (Deeks et al., 1997; Palella et al., 1998). However, incomplete suppression of viral replication can select for mutations of one or more protease amino acid residues, resulting in resistance to protease inhibitors (PIs) (Condra et al., 1995; Hertogs et al., 2000; Tamalet et al., 2003; Cheung et al., 2004). The prevalence of PI resistance in tested patients was reported to be 50% or more in 1999 and subsequently declined to less than 30% from late 2003 onwards (Tamalet et al., 2003; Kagan et al., 2004a; Pillay et al., 2005). The prevalence of ARV resistance has been correlated to ARV prescription utilization (Kagan et al., 2004a). Decreased utilization of saquinavir, indinavir, nelfinavir and amprenavir and increased utilization of lopinavir/ritonavir, other ritonavir-boosted PI regimens, fos-amprenavir and atazanavir (source: prescription utilization data acquired from NDCHEALTH® http://www.ndchealth.com/products_services/products_services.htm) may have contributed to this decline. PI resistance-associated mutations have also been found in treatment-naïve patients, albeit at a much lower prevalence, likely the result of transmission of drug-resistant virus (Weinstock et al., 2004; Novak et al., 2005; Wensing et al., 2005).

Protease mutations conferring reduced susceptibility to all commercially available PIs have been identified in the HIV-1 patient population (Shafer et al., 2000; Hirsch et al., 2003; Johnson et al., 2005a, b). Established HIV-1 mutations, associated with resistance to eight FDA-approved protease inhibitors,

* Corresponding author. Tel.: +1 949 728 4018; fax: +1 949 728 4022.
E-mail address: ron.m.kagan@questdiagnostics.com (R.M. Kagan).

0166-3542/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.antiviral.2006.02.008

were defined for 21 protease positions in the International AIDS Society-USA March–April 2005 expert panel list (Johnson et al., 2005a). An additional nine positions newly defined for the most recently introduced PIs, atazanavir (3) and tipranavir (6), were added in the October–November 2005 IAS USA update (Johnson et al., 2005b). Accumulating evidence suggests that the number of positions associated with PI treatment is even greater. Wu et al. (2003) identified an additional 22 treatment-associated protease positions and confirmed the association with treatment of 23 previously identified positions. A further analysis of mutational covariation also showed 99 positive, and 16 negative correlations between treatment-associated positions (Wu et al., 2003). Analysis of an expanded treatment-characterized dataset from 3178 PI-experienced and 2689 PI-naïve patients identified 38 treatment-associated protease positions (Rhee et al., 2005). In another treatment-characterized data set of 705 PI-treated and 457 untreated patients, an additional 17 treatment-associated PR mutations were identified; 14 were positively and three were negatively associated with treatment (Svicher et al., 2005). Other findings have indicated that the HIV-1 subtype B protease amino acid sequence is conserved at 68/99 (69%) positions in treatment-naïve patients but only 45/99 (45%) positions in treated patients, suggesting that there is treatment-associated variability at more than half of the protease positions (Ceccherini-Silberstein et al., 2004).

Reference laboratories performing HIV-1 drug resistance testing typically do not obtain treatment data for tested patients. However, these laboratories often accumulate significantly larger datasets (>100 000 samples) than is feasible in clinical studies with available treatment histories. Analysis of a reference laboratory dataset comprising 40 000 protease sequences obtained from 1999 to 2002 detected positive selective pressure at 47 protease positions (Chen et al., 2004). In the present study, we assessed the utility of using large reference laboratory datasets (comprising roughly 115 000 HIV-1 subtype B protease sequences) to identify PI-associated mutation positions, using genotypic predictions of resistance to stratify our dataset and we examined the correlation between mutational prevalence and the duration of non-suppressive PI therapy in a treatment-characterized dataset.

2. Methods

2.1. Sequence determination and resistance analysis

HIV-1 RNA was extracted from clinical samples submitted for reverse transcriptase and protease genotype determination by Quest Diagnostics Nichols Institute. Reverse transcription, amplification, and dye terminator automated sequencing using ABI Prism 3700 sequencing instruments (Applied Biosystems, Foster City, CA) was performed as described previously at the Nichols Institute, San Juan Capistrano, CA laboratory (SJC; Kagan et al., 2004a), or using an ABI Prism 3100 or 3700 as previously described at the BC Centre for Excellence in HIV/AIDS (Alexander et al., 1999). At the Nichols Institute, Chantilly, VA, laboratory (CHA), the ViroSeq HIV-1 genotyping system (Celera Diagnostics, Alameda, CA) was used according to the manufacturer's instructions and automated sequencing was performed on ABI 3100 sequencing instruments (Applied Biosystems). Predicted resistance to PIs (amprenavir, atazanavir, fosamprenavir, indinavir, lopinavir, nelfinavir, ritonavir, saquinavir, tipranavir), nucleoside/nucleotide reverse transcriptase inhibitors (NRTIs; abacavir, didanosine, emtricitabine, lamivudine, stavudine, tenofovir, zalcitabine, zidovudine), and non-nucleoside reverse transcriptase inhibitors (NNRTIs; delavirdine, efavirenz, nevirapine) was determined according to the Quest-Stanford rules-based interpretive system (copyright 1997–2005, Quest Diagnostics Incorporated), an updated version of the GART study algorithm (Baxter et al., 2000; Merigan and Winters, Center for AIDS Research, Stanford University, personal communication, June 2005). The interpretive rules used in this laboratory-developed algorithm are based on mutations the laboratory considered to be associated with resistance to antiretroviral drugs based on current clinical or laboratory-based studies. Mutation and resistance data for these analyses were stored in a SQL Server 2000 relational database (Microsoft Corp. Redmond, WA). A PI resistance score of 0–7 was established according to the number of PIs to which a given virus sample showed genotypically predicted resistance. PIs with identical resistance profiles according to the Quest-Stanford algorithm (amprenavir/fosamprenavir and ritonavir/indinavir) were counted only once.

2.2. Viral loads and clinical monitoring

2.2.1. The BC Centre for Excellence in HIV/AIDS Drug Treatment Program

In British Columbia, Canada, ARVs are provided centrally by the BC Centre for Excellence Drug Treatment Program according to guidelines established by the BC Therapeutic Guidelines Committee (Hogg et al., 2001). This program distributes more than 99% of the ARVs used in British Columbia. Individual drug utilization records are retained. Routine clinical monitoring of patients takes place at approximately 3-month intervals, at which time plasma viral load testing (Roche Amplicor Monitor Assay) and CD4 counts are performed; the data are stored in the Centre's Drug Treatment Program database. The Centre's HIV genotypic drug resistance database consists of over 25 000 samples from worldwide sources. For the purposes of this study, data were restricted to a single sample per individual patient (the latest available "on-therapy" genotype), from a period in which they had been failing a PI-based regimen (with a plasma viral load >500 copies) for the total number of months specified below. The entire interval between tests with a viral load >500 copies/ml was assumed to be spent with a detectable viral load. ARV regimen changes were permitted when the viral load was >500 copies/ml and patients were included in the study as long as they remained on any PI. Data were available from individuals tested between June 1996 and January 2006. Individuals were classified as having failed PI-based therapy for a total of 1–4 months ($n = 787$); 5–11 months ($n = 598$); 12–23 months ($n = 416$); 24–35 months ($n = 193$); 36–47 months ($n = 86$) and ≥ 48 months ($n = 81$).

2.3. Protease mutation analysis

The primary analysis utilized protease mutation data for HIV-1 subtype B samples submitted for testing from January 2002 to June 2005 at the SJC ($n = 107\,169$) or the CHA ($n = 7853$) laboratory. SJC laboratory data from an additional 39 931 samples tested from 1999 to 2001 were also available. Data were partitioned into protease sequences with no genotypically predicted resistance to any ARV (ARV-sensitive group) and sequences with genotypically predicted resistance to one or more PI. We used the χ^2-test for each position to identify mutated positions in the PI-resistant group that were present at a statistically higher frequency than in the ARV-sensitive group. To correct for multiple comparisons, we utilized the Benjamini-Hochberg correction (Benjamini and Hochberg, 1995) for 99 tests and a false discovery rate (FDR) set to 0.01. For SJC to CHA inter-laboratory comparisons we selected two time periods with available data from both laboratories: 2002–2003 and 2004–2005. Changes in mutation prevalence in the smaller BC Centre dataset of non-suppressed patients were assessed with the Fisher Exact test.

2.4. Protease mutation covariation analysis

To analyze covariation between pairs of mutations, we excluded positions with mixed amino acid assignments that included the HIV-1 protease consensus B wild type amino acid. Unmixed amino acid assignments were present in 98.6% of the SJC and 97.7% of the CHA datasets. To identify covariation, we calculated the binomial correlation coefficient ϕ for all pairs of 44 positions identified during individual protease mutation analysis (946 pairs). We utilized χ^2-tests to determine the significance of these correlations corrected for 946 multiple comparisons (FDR = 0.01). We defined a distance metric of $1 - \phi$ and constructed a pairwise distance matrix from the dataset. We utilized the Neighbor program version 3.64 in the PHYLIP phylogenetic analysis suite (Dr. J. Felsenstein, University of Washington, Dept. of Genome Sciences) to generate neighbor joining trees with three randomizations of the input distance to construct dendrograms for the identification of clusters of protease mutations. Dendrograms were rendered with TreeExplorer 2.12 (http://evolgen.biol.metro-u.ac.jp/TE/TE_man.html). To further refine the pairwise correlations, we calculated binomial correlation coefficients for 3036 pairs of 79 specific amino acid substitutions at the same 44 positions identified (3081 pairs less 45 self-pairs for the same position but different amino acid substitution, which were assigned a value of $\phi = 0$).

3. Results

3.1. Prevalence of protease mutations in reference laboratory datasets

In the SJC reference laboratory database, resistance to one or more PIs was predicted in 48.5% of the tested subtype B samples in 1999. Predicted resistance subsequently declined to 41.6% in 2000 and 35.5% in 2001. Data were available from 2002 to 2005 for both the SJC and the CHA Quest Diagnostics laboratories. The prevalence of PI resistance was 31.1% in the SJC and 34.2% in the CHA dataset in 2002, and declined to 16.1% and 21.4%, respectively, in the first half of 2005. For viruses with a PI resistance score >0, we tabulated the statistically significant (χ^2-test, FDR = 0.01 to correct for 99 comparisons) protease positions that were mutated at least 2× as frequently as in ARV-sensitive viruses and grouped them according to prevalence (Table 1). For the laboratory datasets spanning the periods 2002–2003 and 2004–2005, we identified 39–44 PI resistance-associated positions (mean = 41), 20–25 (51%–57%) of which, were not among the March–April 2005 established IAS-USA PI resistance mutations (Johnson et al., 2005a) (Table 1). Mutations at non-established positions were more commonly found (43%–75%) at lower prevalences (<4%) compared to those at established positions (0%–21%; Fisher's Exact test, $p < 0.005$) (Table 1).

The cumulative frequency of 20 non-established protease mutations that were identified in both the SJC and CHA datasets is plotted in Fig. 1. The prevalence of mutations at these positions was 34%–36% in viruses identified as PI-susceptible, but

Table 1
Prevalence of PI resistance-associated mutated positions in two reference laboratory datasets

Prevalence	Positions with mutations, total (new)[a]			
	2002–2003		2004–2005	
	SJC[b]	CHA[b]	SJC[b]	CHA[b]
<1%	2 (2)	4 (4)	1 (1)	3 (3)
1–2%	2 (2)	3 (3)	4 (4)	7 (7)
2–4%	7 (6)	7 (6)	5 (5)	9 (5)
4–8%	9 (5)	10 (6)	10 (6)	7 (5)
>8%	19 (5)	20 (6)	22 (7)	13 (0)
Total	39 (20)	44 (25)	42 (23)	39 (20)
<4% establ. (new)[c]	5% (50%)	5% (52%)	0% (43%)	21% (75%)

[a] The counts of mutated protease positions were grouped according to their prevalence in viruses with a PI resistance score of >0. New protease mutation counts (in parentheses) are positions that are not listed as having established protease mutations in the March–April 2005 International AIDS Society USA published list of resistance-associated mutations (Johnson et al., 2005a).

[b] SJC: Quest Diagnostics Nichols Institute, San Juan Capistrano, CA. CHA: Quest Diagnostics Nichols Institute, Chantilly, VA.

[c] Percent of IAS USA established mutations found at a prevalence of <4%, vs. new mutations (in parantheses) found at a prevalence of <4%.

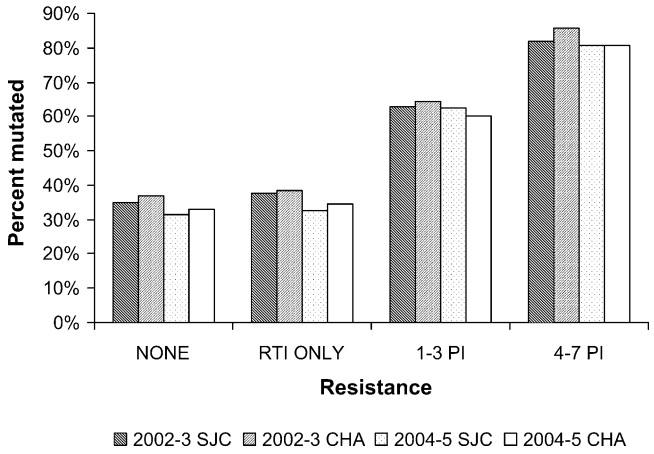

Fig. 1. Cumulative frequency of mutations at 20 non-established protease positions is greater in viruses with genotypically predicted PI resistance. The prevalence of viral sequences with an amino acid substitution at any of the 20 protease residues (4, 11, 13, 23, 34, 43, 45, 55, 58, 62, 66, 74, 75, 76, 79, 95, 89, 91, 92, 95 as described in Table 2) is shown for two reference laboratory datasets (SJC and CHA) over two time periods (2002–2003 and 2004–2005). RTI Only: Reverse transcriptase inhibitor resistance, but no Protease Inhibitor resistance was predicted for these sequences.

increased to 62% in viruses with PI resistance scores of 1–3 and 82% in viruses with PI resistance scores of 4–7 (Fig. 1). With the exceptions of positions 13 (mean and S.D.: 15.6% ± 1.1%) and 62 (17.9% ± 1.8%), mutations at these positions were uncommon in PI-susceptible viruses (Table 2, no resistance or RTI resistance only). Mutations at 14 of these 20 positions were ≥2× as frequent in viruses with a PI resistance score of 4–7 as in those with a score of 1–3 PI (Table 2, underlined positions). In contrast, little or no difference was observed between viruses with no predicted ARV resistance and predicted resistance to RTIs but not PIs (Tables 2 and 3). At 8 of the 20 positions, the most common amino acid substitution differed between PI-resistant and PI-susceptible virus (Table 2, underlined mutations). For example, most (77%) PI-resistant viruses had a threonine (K43T) at protease position 43, whereas only 6% of the PI-susceptible viruses had the K43T mutation; most (88%) PI-susceptible viruses had an arginine substitution (K43R) at this position (Table 2).

3.2. *Prevalence of protease mutations during long term non-suppressive PI therapy correlates with PI resistance scores*

Incomplete suppression of viral replication by reverse transcriptase or protease inhibitors may lead to the accumulation of resistance-associated mutations over a period of months or

Table 2
Amino acid substitutions at 20 non-established protease positions are more prevalent in PI-resistant viruses[a]

Position[b]	Prevalence (mean ± S.D.)				PI-susceptible virus[c]		PI-resistant virus[c]		Study[d]
	No resistance (%)	RTI only (%)	1–3 PI (%)	4–7 PI (%)	Mutation	Distribution	Mutation	Distribution	
4	0.3 ± 0.1	0.2 ± 0.1	0.3 ± 0.2	1.9 ± 0.5	T4S/P/N	0.55/0.13/0.12	T4P/S/A	0.47/0.34/0.15	
11	0.2 ± 0.1	0.2 ± 0.1	0.5 ± 0.1	6.1 ± 1.2	V11I/A/L	0.40/0.20/0.19	V11I/L	0.80/0.17	A, D
13	14.8 ± 0.7	16.3 ± 0.9	27.2 ± 1.7	34.5 ± 3.3	I13V	0.99	I13V	0.97	A
23	0.1 ± 0.1	0.1 ± 0.1	1.0 ± 0.7	1.6 ± 0.2	L23F/I	0.70/0.21	L23I	0.89	A, D
34	0.1 ± 0.1	0.1 ± 0.1	1.2 ± 0.2	7.1 ± 1.4	E34D/G/K	0.29/0.29/0.24	E34Q/D/K	0.68/0.10	A, B, C, D
43	1.8 ± 0.3	2.0 ± 0.6	3.7 ± 0.3	12.3 ± 1.4	K43R/T	0.88/0.06	K43T/R	0.77/0.14	A, B, C, D
45	1.0 ± 0.2	1.0 ± 0.2	4.1 ± 0.8	2.3 ± 0.2	K45R	0.98	K45R/Q	0.82/0.10	A, B, C
55	0.2 ± 0.1	0.4 ± 0.1	2.9 ± 0.7	11.7 ± 2.4	K55R	0.89	K55R	0.93	A, B, C, D
58	0.3 ± 0.2	0.5 ± 0.2	3.7 ± 1.2	11.8 ± 1.4	Q58E	0.90	Q58E	1.0	A, B, C, D
62	17.6 ± 2.0	18.2 ± 1.8	35.0 ± 1.0	46.8 ± 3.5	I62V	0.99	I62V	1.0	A, C
66	0.04 ± 0.03	0.1 ± 0.1	0.8 ± 0.2	5.3 ± 0.4	I66V/F/L	0.40/0.32/0.21	I66F/V/L	0.51/0.35/0.11	A, B, D
74	0.5 ± 0.1	1.7 ± 0.2	9.2 ± 0.8	13.2 ± 0.8	T74S/A	0.55/0.29	T74S/P	0.64/0.20	A, B, C, D
75	0.2 ± 0.1	0.2 ± 0.1	1.2 ± 0.6	0.5 ± 0.3	V75I	0.99	V75I	1.0	A, D
76	0.01 ± 0.0	0.1 ± 0.1	0.9 ± 0.2	4.6 ± 0.5	L76V	0.81	L76V	0.99	A, D
79	0.4 ± 0.1	0.5 ± 0.1	0.4 ± 0.2	2.6 ± 0.8	P79S/D/A	0.42/0.18/0.15	P79A/S/D	0.37/0.23/0.15	A, D
85	0.1 ± 0.0	0.4 ± 0.1	3.1 ± 0.5	6.8 ± 1.4	I85V	0.83	I85V	0.99	A, C
89	1.2 ± 0.2	1.1 ± 0.3	3.0 ± 0.6	11.3 ± 1.5	L89M	0.85	L89V/M	0.55/0.28	A, D
91	0.2 ± 0.1	0.1 ± 0.1	0.7 ± 0.2	3.8 ± 1.0	T91V/A/I	0.48/0.20/0.12	T91S/A	0.62/0.25	
92	0.9 ± 0.1	0.9 ± 0.1	2.0 ± 0.4	5.8 ± 1.8	Q92K/E/R	0.53/0.19/0.17	Q92K/R	0.76/0.18	A, B, C, D
95	0.01 ± 0.01	0.02 ± 0.03	0.7 ± 0.1	3.7 ± 0.8	C95F/G/S	0.45/0.27/0.14	C95F	0.86	A, B, C, D

[a] Prevalences and amino acid distributions of 20 protease positions not included in the International AIDS Society-USA March-April 2005 listing (Johnson et al., 2005a) that were significantly more common in viruses with a PI resistance score >0 in all Quest Diagnostics reference laboratory datasets (SJC and CHA for 2002–3 and 2004–5 time periods). Counts were tabulated for viruses with no predicted resistance to any ARV, predicted RTI but not PI resistance, and PI resistance scores of 1–3 and 4–7.

[b] Fourteen positions that occurred at least 2× as frequently in viruses with a mutation score of 4–7 than in those with a mutation score of 1–3 (χ^2-test, $p < 0.0001$ and lower 95% confidence interval of odds ratio >2) are underlined.

[c] Right hand panel: the proportion of non-wild-type-specific amino acid substitutions at each of these mutated positions in viruses with no predicted PI resistance or with a PI resistance score >0. Only substitutions with a frequency of ≥0.1 are shown. Underlined postions: positions in PI-resistant virus at which the most frequent amino acid substitution is different from that found in PI-susceptible virus.

[d] Non-IAS protease mutations (March–April 2005 list, Johnson et al., 2005a) that have been associated with PI treatment in recent studies. (A) Wu et al., 2003; (B) Ceccherini-Silberstein et al., 2004; (C) Svicher et al., 2005; (D) Rhee et al., 2005.

Table 3
Protease mutation prevalence in patients with incomplete suppression correlates with PI resistance score[a]

Position[b]	BC Centre for Excellence in HIV/AIDS						Quest Diagnostics Nichols Institute			
	Months on PI therapy with a measurable viral load (n) (%)						Prevalence (mean ± S.D.) by PI resistance score (%)			
	1–4 (787)	5–11 (598)	12–23 (416)	24–35 (193)	36–47 (86)	≥48 (81)	None	RTI only	1–3 PI	4–7 PI
4	0.6	0.3	0.2	1.6	1.2	3.7	0.3 ± 0.1	0.2 ± 0.1	0.3 ± 0.2	1.9 ± 0.5
10	13.9	15.7	25.2	39.4	57.0	69.1	8.4 ± 0.1	10.8 ± 0.8	36.5 ± 1.7	86.9 ± 1.1
11	0.8	0.0	0.5	0.5	1.2	4.9	0.2 ± 0.1	0.2 ± 0.1	0.5 ± 0.1	6.1 ± 1.2
13	13.6	13.4	11.5	15.0	22.1	24.7	14.8 ± 0.7	16.3 ± 0.9	27.2 ± 1.7	34.5 ± 3.3
16	4.5	3.5	3.6	4.7	2.3	9.9	4.2 ± 0.3	3.6 ± 0.4	3.7 ± 0.5	6.1 ± 1.1
18	2.2	3.2	3.4	6.7	3.5	0.0	1.8 ± 0.2	2.0 ± 0.3	2.8 ± 0.6	3.7 ± 0.5
20	5.7	5.4	8.7	19.7	25.6	43.2	2.3 ± 0.4	3.7 ± 0.5	22.0 ± 0.8	42.3 ± 3.6
23	0.0	0.0	0.0	1.0	0.0	0.0	0.1 ± 0.1	0.1 ± 0.1	1.0 ± 0.7	1.6 ± 0.2
24	0.1	0.8	1.9	3.6	3.5	8.6	0.03 ± 0.03	0.2 ± 0.1	2.4 ± 0.5	12.7 ± 0.9
30	1.1	5.4	4.6	5.2	7.0	0.0	0.02 ± 0.02	0.02 ± 0.02	25.6 ± 1.6	1.8 ± 0.6
32	0.4	0.2	2.2	3.1	3.5	7.4	0.03 ± 0.03	0.20 ± 0.10	2.4 ± 0.8	13.7 ± 0.7
33	3.6	6.7	2.9	6.2	8.1	19.8	3.6 ± 0.2	3.9 ± 0.4	8.6 ± 0.8	34.6 ± 2.5
34	0.8	0.3	0.2	1.0	2.3	6.2	0.1 ± 0.1	0.1 ± 0.1	1.2 ± 0.2	7.1 ± 1.4
35	33.4	35.0	39.4	37.8	32.6	39.5	21.2 ± 2.7	20.3 ± 1.4	30.5 ± 3.0	34.3 ± 1.8
36	26.4	26.4	30.8	38.9	37.2	58.0	14.0 ± 0.7	16.1 ± 1.3	32.3 ± 1.0	44.7 ± 4.4
43	3.6	3.7	5.5	4.7	3.5	13.6	1.8 ± 0.3	2.0 ± 0.6	3.7 ± 0.3	12.3 ± 1.4
45	1.9	2.0	2.6	3.1	1.2	1.2	1.0 ± 0.2	1.0 ± 0.2	4.1 ± 0.8	2.3 ± 0.2
46	2.5	5.9	12.3	21.2	34.9	48.2	0.2 ± 0.1	1.5 ± 0.5	20.8 ± 0.8	57.4 ± 3.7
47	0.3	0.0	1.4	1.6	3.5	3.7	0.03 ± 0.04	0.03 ± 0.03	1.7 ± 0.5	10.6 ± 1.4
48	0.4	0.3	2.9	7.3	16.3	21.0	0.02 ± 0.03	0.02 ± 0.03	1.3 ± 0.3	9.5 ± 1.9
50	0.0	0.2	0.0	1.0	1.2	3.7	0.03 ± 0.03	0.03 ± 0.04	2.5 ± 1.2	3.9 ± 0.8
53	0.0	0.3	2.2	2.6	8.1	14.8	0.01 ± 0.02	0.03 ± 0.03	1.1 ± 0.2	10.2 ± 1.5
54	0.8	3.2	6.5	18.1	30.2	55.6	0.04 ± 0.05	0.38 ± 0.18	12.9 ± 1.5	69.6 ± 1.2
55	0.4	0.2	1.4	4.2	3.5	9.9	0.2 ± 0.1	0.4 ± 0.1	2.9 ± 0.7	11.7 ± 2.4
58	0.5	1.3	1.9	3.1	5.8	7.4	0.3 ± 0.2	0.5 ± 0.2	3.7 ± 1.2	11.8 ± 1.4
60	7.0	7.9	6.7	10.4	17.4	9.9	8.2 ± 0.8	8.2 ± 0.5	8.9 ± 1.1	14.8 ± 1.2
61	4.3	4.2	5.3	5.7	8.1	9.9	4.0 ± 0.4	3.9 ± 0.2	4.1 ± 0.6	6.7 ± 1.5
62	25.9	27.6	34.4	40.9	37.2	46.9	17.6 ± 2.0	18.2 ± 1.8	35.0 ± 1.0	46.8 ± 3.5
66	0.4	0.7	0.5	1.6	3.5	3.7	0.04 ± 0.03	0.09 ± 0.07	0.8 ± 0.2	5.3 ± 0.4
71	14.4	18.2	28.6	37.3	53.5	51.9	7.9 ± 1.2	9.3 ± 0.6	37.3 ± 3.9	77.0 ± 2.4
72	13.6	17.2	16.8	14.0	22.1	18.5	12.5 ± 1.0	13.0 ± 1.2	17.1 ± 0.6	24.2 ± 3.3
73	1.1	2.7	6.3	12.4	16.3	24.7	0.02 ± 0.02	0.1 ± 0.1	7.8 ± 0.4	23.2 ± 3.7
74	1.0	1.0	3.9	3.6	8.1	11.1	0.5 ± 0.1	1.7 ± 0.2	9.2 ± 0.8	13.2 ± 0.8
75	0.3	0.5	0.2	2.1	2.3	1.2	0.2 ± 0.1	0.2 ± 0.1	1.2 ± 0.6	0.5 ± 0.3
76	0.3	0.2	0.5	1.0	0.0	0.0	0.01 ± 0.01	0.1 ± 0.09	0.9 ± 0.2	4.6 ± 0.5
79	0.5	1.0	0.7	1.6	2.3	2.5	0.4 ± 0.1	0.5 ± 0.1	0.4 ± 0.2	2.6 ± 0.8
82	3.3	5.4	12.3	19.7	37.2	46.9	1.5 ± 0.3	1.8 ± 0.3	17.4 ± 2.2	59.9 ± 2.0
84	0.5	2.3	4.6	9.3	15.1	19.8	0.01 ± 0.01	0.05 ± 0.06	2.1 ± 0.4	38.5 ± 2.3
85	0.3	0.8	2.2	2.6	8.1	14.8	0.1 ± 0.0	0.4 ± 0.1	3.1 ± 0.5	6.8 ± 1.4
88	0.6	2.2	5.3	7.3	8.1	2.5	0.05 ± 0.03	1.1 ± 0.3	22.7 ± 1.1	5.0 ± 0.6
89	5.0	2.5	3.6	0.5	2.3	7.4	1.2 ± 0.2	1.1 ± 0.3	3.0 ± 0.6	11.3 ± 1.5
90	2.7	5.4	13.7	30.6	43.0	43.2	0.01 ± 0.01	0.01 ± 0.02	41.4 ± 3.1	66.7 ± 2.2
91	0.4	0.3	0.5	0.0	3.5	7.4	0.2 ± 0.1	0.1 ± 0.1	0.7 ± 0.2	3.8 ± 1.0
92	2.9	2.7	4.1	3.6	3.5	6.2	0.9 ± 0.1	0.9 ± 0.1	2.0 ± 0.4	5.8 ± 1.8
93	39.5	43.7	45.9	48.2	48.8	53.1	23.2 ± 2.0	22.2 ± 1.9	40.9 ± 1.2	45.3 ± 1.2
95	0.1	0.0	0.2	2.1	3.5	3.7	0.01 ± 0.01	0.02 ± 0.03	0.7 ± 0.1	3.7 ± 0.8

[a] The prevalence of mutations at 46 protease positions found in patients with non-suppressive PI therapy (*n*: number of patients in each group) for 1–4 to ≥48 months compared to the frequency of mutations in two Quest Diagnostics Nichols Institute databases (San Juan Capistrano, CA, and Chantilly, VA) during two time periods (2002–2003 and 2004–2005).

[b] Twenty four positions that were mutated more frequently at ≥48 compared to 1–4 months of non-suppressive therapy (Fisher Exact test, and corrected for 46 multiple comparisons with a Benjamini-Hochberg false discovery rate of 0.1) are underlined.

years (Barbour et al., 2002; Kantor et al., 2004). The prevalence of protease mutations for a cohort of patients (BC Centre for Excellence in HIV/AIDS, as described in Section 2) undergoing PI therapy with incomplete suppression of viral replication for 1–≥48 months was tabulated for 46 protease positions, including 21 IAS-USA resistance-associated positions (March/April 2005 listing; Johnson et al., 2005a) and 25 non-established positions (Table 3). The mutation prevalence at 24 positions was greater after ≥48 months compared to 1–4 months of non-suppressive therapy (Table 3; Fisher Exact test and corrected for 46 multiple comparisons with a Benjamini-Hochberg false discovery rate of 0.1) including eight non-established positions (34, 43, 55, 58, 62,

74, 85 and 91; 13, 93 and 95 were also significant before correction for multiple comparisons). To examine whether mutations observed after long term non-suppressive PI therapy existed at baseline or were likely selected by PI therapy, we tabulated the amino acid substitution frequencies at several polymorphic positions identified in the reference laboratory dataset (Table 2) in the non-suppressed ≥48 month treatment group and in the 1–4 month group. For position 34, E34Q comprised 80% of the non-synonymous substitutions in the ≥48 month group but only 17% in the 1–4 month group, where other substitutions (A/D/G/K/L) predominated. Likewise, K43T was seen in 64% of the patients in the ≥48 month group but in only 7% of the 1–4 month group, where K43R comprised 79% of the non-synononymous substitutions. At other non-polymorphic positions (Table 2) the dominant amino acid substitution was the same as that shown for the larger reference laboratory dataset (not shown).

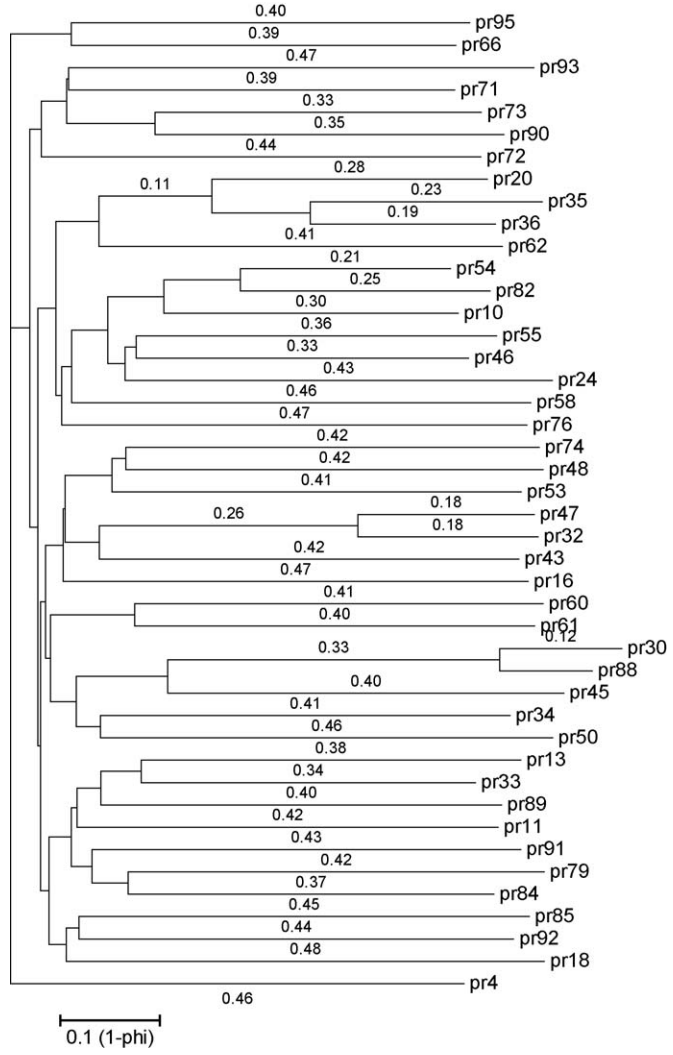

Fig. 2. Neighbor joining analysis of 44 protease positions in PI-resistant viruses. The binomial correlation coefficient ϕ was calculated for all 946 possible pairwise combinations and a distance of $1-\phi$ was used to construct a distance matrix for the Neighbor program in PHYLIP 3.64. The resulting unrooted neighbor joining tree shows the clustering and covariation of mutated protease positions in PI-resistant virus.

Mutations at position 30, associated with nelfinavir resistance (Patick et al., 1998), were found in 4.6%–7.0% of patients with incomplete suppression after 5–47 months, but in none of the patients with ≥48 months of non-suppressive PI therapy (Table 3). However, mutations at position 90 (L90M), seen more often in heavily PI-experienced patients, was found in 43.2% of patients after ≥48 months of non-suppressive therapy (Table 3). In the reference laboratory dataset, the prevalence of the position 30 mutation was highest in patients with a PI resistance score of 1–3 (Table 3: 25.6%), but this mutation was much less frequent in those with a PI resistance score of 4–7 (Table 3: 1.8%) and position 90 mutations were seen in 66.7% of patients with a score of 4–7 (Table 3). Mutations at positions 88 and 45, which are often seen together with the D30N mutation (Fig. 2 and Svicher et al., 2005), were also less common in patients with ≥48 months of non-suppressive PI therapy (Table 3).

Mutations at position 76 were very uncommon in the non-suppressed patients but appeared in 4.6% of the reference laboratory samples with a PI resistance score of 4–7 (Table 3). The L76V mutation has become more prevalent in PI-resistant viruses in the Quest Diagnostics (SJC) database since 2001 (1999–2000: 0.48%; 2001: 1.7%; 2005: 2.8%) and is thought to be selected by lopinavir during therapy failure of salvage therapy (Mueller et al., 2004).

The mutation prevalences of both established and non-established protease mutations correlated with the PI resistance scores seen in the Quest Diagnostics reference laboratory datasets (Table 4). Mutation prevalences in virus from patients after 1–4 months of non-suppressive therapy correlated best with those of PI-susceptible virus (Table 4). Conversely, mutation prevalences in samples from patients with 36–47 and ≥48 months of non-suppressive treatment correlated best with those of viruses with PI resistance scores of 4–7. For intermediate durations of non-suppressive treatment (5–35 months), mutation prevalence correlated best with prevalences in viruses with PI resistance scores of 1–3 (Table 4).

Table 4
Correlation of non-suppressive treatment duration with PI resistance score[a,b]

Non-suppressive PI therapy (months)	PI resistance score		
	RTI only	1–3 PI	4–7 PI
1–4	**0.83**	0.75	0.59
5–11	0.78	**0.85**	0.65
12–23	0.68	**0.89**	0.78
24–35	0.61	**0.86**	0.77
36–47	0.48	0.81	**0.84**
≥48	0.50	0.67	**0.92**

[a] Spearman correlation coefficients were calculated for mutation prevalences at 46 protease positions in the non-suppressed dataset vs. the mean mutation prevalence at these positions in the Quest Diagnostics reference laboratory datasets shown in Table 3, for viruses with RTI resistance only and for resistance scores of 1–3 and 4–7 PIs.

[b] Bold, boxed coefficients represent the best correlations for each resistance score group.

3.3. Covariation of protease mutations and neighbor joining clusters

Covariation of protease mutations has been previously reported (Hoffman et al., 2003; Wu et al., 2003; Kagan et al., 2004b; Svicher et al., 2005). We performed a clustering analysis of 44 PI resistance-associated mutations in the reference laboratory datasets that were also identified in the BC Centre for Excellence non-suppressive therapy dataset (Table 3) by calculating the binomial correlation coefficient ϕ for all 946 possible pairs of positions and then constructing unrooted neighbor joining trees based on the pairwise distances ($d = 1 - \phi$), where branch lengths are proportional to the degree of co-mutation (Fig. 2; data shown for SJC 2005 dataset). A number of non-established mutations could be placed into clusters with known PI resistance-associated mutations. The tightest association was between nelfinavir resistance-associated positions 30 + 88 (SJC: $d = 0.21$; CHA: $d = 0.23$) in a cluster that included residue 45, followed by the amprenavir resistance-associated positions 32 + 47 (SJC: $d = 0.36$; CHA: $d = 0.30$). A large cluster of resistance-related protease mutations included residues ((54 + 82) + 10) in both the SJC and CHA datasets has also been described previously (Wu et al., 2003). Mutations at position 55 were associated with position 46 in the cluster ((55 + 46) + 24) in both Quest Diagnostics SJC and CHA datasets. An association between K55R and M46I/L has been observed previously (Morgan et al., 2003). Positions 35 and 36 were the third most strongly associated pair (SJC: $d = 0.42$; CHA: $d = 0.43$) and clustered with residues 20 and 62 in both datasets. To further define the associations between protease positions in resistant virus, we calculated the correlations between 79 specific amino acid substitutions and 3036 mutation pairs. Common substitutions at polymorphic positions (e.g., M46I and M46L) were analyzed separately (Fig. 3). Substitution-specific associations that were not observed when the amino acid information was omitted (Fig. 2) included I54M associated with I47V + V32I, an amprenavir resistance-related mutation cluster (Maguire et al., 2002). I54V, in contrast, clustered more closely with V82A ($d = 0.57$) in a cluster that included L24I and M46L (Fig. 3). M46I clustered with K55R ($d = 0.76$), and L76V appeared to cluster with both M46I and K55R (Fig. 3). Table 5 shows 75/3036 statistically significant positive mutation associations (FDR = 0.01, $\phi > 0.15$) and 9/3036 negative associations (FDR = 0.01, $\phi < -0.15$; Table 5). Amino acid substitution-specific associations included V11**I** + (one or more of K20I, A71V, G73S, I84V, L89V) versus V11**L** + T91S, L33F + (one or more of E34Q, K43T, M46L, F53L, I54M, K55R, T74P, V82A, I84V) versus L33**I** + M36I, I54**M** + I47V + V32I versus I54**V** + V82A and I66**F** + L89V versus I66**V** + C95F (Fig. 3 and Table 5). Negative associations between specific amino acid substitutions (bottom of Table 5) primarily involved the nelfinavir-related mutations D30N + N88D and L90M, as reported previously (Sugiura et al., 2002), and were also noted between D30N and/or N88D and L10I, I54V, V82A, and I84V. The L24I mutation was also negatively associated with L90M, in agreement with a previous report (Wu et al., 2003).

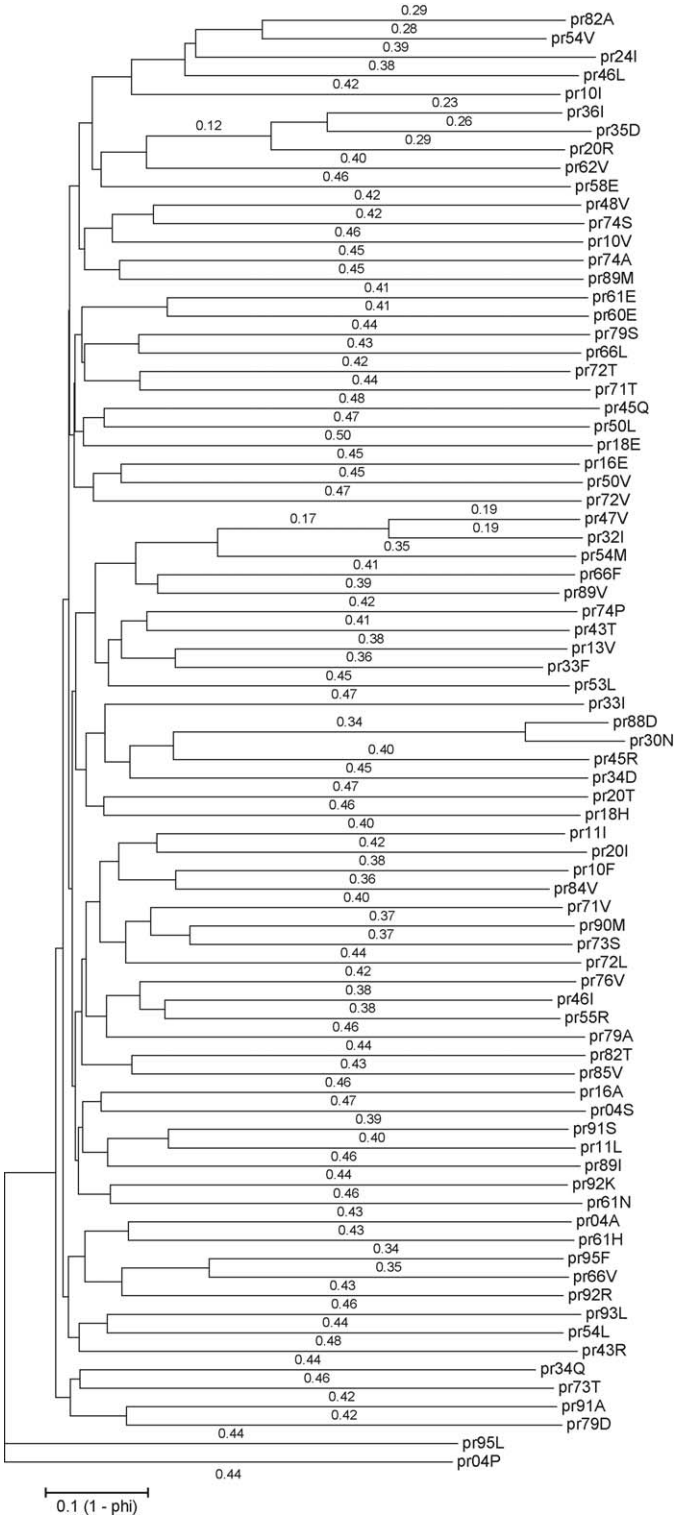

Fig. 3. Neighbor joining analysis of 79 protease amino acid substitutions at 44 positions in PI-resistant viruses. Covariation was analyzed as described for Fig. 2, with the addition of the specific amino acid substitutions occurring at each of the 44 variant positions in PI-resistant virus.

Table 5
Pairwise correlations between protease amino acid substitutions in PI-resistant virus[a]

Pos 1	Pos 2	ϕ	OR[b]	Pos 1	Pos 2	ϕ	OR[b]
L10F	M46I	0.17	1.6	E35D	M36I	0.51	2.0
L10F	I84V	0.26	2.1	E35D	I62V	0.18	1.3
L10I	L24I	0.18	1.9	M36I	I62V	0.23	1.3
L10I	I54V	0.20	1.4	K43T	T74P	0.17	4.2
L10I	V82A	0.21	1.4	K43T	V82A	0.24	2.3
V11I	K20I	0.19	4.0	K45R	N88D	0.18	3.8
V11I	A71V	0.15	2.0	M46I	I47V	0.20	2.1
V11I	G73S	0.17	3.4	M46I	K55R	0.24	2.1
V11I	I84V	0.17	2.7	M46I	L76V	0.21	2.9
V11I	L89V	0.18	5.1	M46I	I84V	0.19	1.5
V11L	T91S	0.21	18	M46L	I54V	0.20	1.8
I13V	K20T	0.18	2.0	M46L	V82A	0.31	2.3
I13V	L33F	0.26	1.7	I47V	I54M	0.33	6.7
I13V	N88D	0.18	1.7	G48V	T74S	0.16	3.9
I13V	L89V	0.21	2.3	G48V	V82A	0.21	2.7
K20I	M36I	0.18	1.8	F53L	V82A	0.17	2.0
K20I	L89V	0.16	3.2	I54M	L89V	0.21	5.2
K20I	L90M	0.16	1.5	I54V	K55R	0.15	1.7
K20R	E35D	0.40	2.3	I54V	A71V	0.17	1.3
K20R	M36I	0.44	2.4	I54V	V82A	0.43	2.0
K20R	I62V	0.16	1.4	I54V	V82T	0.16	2.2
K20R	V82A	0.18	1.7	Q58E	V82A	0.15	1.7
L24I	M46L	0.22	3.4	D60E	Q61E	0.18	3.6
L24I	I54V	0.24	2.4	I66F	L89V	0.20	7.8
L24I	K55R	0.15	2.9	I66V	C95F	0.31	19
L24I	V82A	0.29	2.7	A71V	G73S	0.25	1.8
D30N	N88D	0.82	7.1	A71V	V82A	0.20	1.4
V32I	L33F	0.17	2.0	I72L	G73S	0.19	3.8
V32I	M46I	0.22	2.0	G73S	L90M	0.25	1.6
V32I	I47V	0.63	8.5	P79D	T91A	0.16	32
V32I	I54M	0.26	4.6	I84V	T91S	0.18	3.2
V32I	I66F	0.18	5.3	L90M	I93L	0.17	1.2
V32I	L89V	0.16	3.1				
L33F	E34Q	0.21	3.1				
L33F	K43T	0.20	2.4	L10I	D30N	−0.18	0.34
L33F	M46L	0.17	1.9	L10I	N88D	−0.16	0.43
L33F	F53L	0.16	2.2	L24I	L90M	−0.29	<0.012[c]
L33F	I54M	0.22	2.9	D30N	I54V	−0.17	0.29
L33F	K55R	0.18	2.1	D30N	V82A	−0.19	0.19
L33F	T74P	0.19	2.9	D30N	I84V	−0.17	0.13
L33F	V82A	0.29	1.9	D30N	L90M	−0.31	0.24
L33F	I84V	0.18	1.6	V82A	N88D	−0.17	0.30
L33I	M36I	0.18	2.4	N88D	L90M	−0.23	0.47

[a] A total of 3036 specific amino acid substitutions at 44 protease positions in 2436 PI-resistant viral sequences from the Quest Diagnostics Nichols Institute, San Juan Capistrano, reference laboratory in 2005 were analyzed for pairwise associations by calculating the binomial correlation coefficient, ϕ. The top 75 significant positions (χ^2 analysis with an FDR = 0.01 to correct for 3036 comparisons) and $-0.15 < \phi < 0.15$ are shown. Negative associations between mutations are shown in the lower right-hand part of the table.

[b] OR: odds ratio between the frequency of occurrence of the respective mutation in viruses with a mutation score >0 and antiretroviral-sensitive viruses.

[c] No cases of L24I occurred together with L90M in the 2005 dataset; therefore the OR is estimated for the detection of a single case of L24I + L90M.

4. Discussion

In this work, we analyzed HIV-1 subtype B protease sequences in two independent reference laboratory datasets over two time periods, 2002–2003 and 2004–2005. The overall number of protease positions associated with PI resistance was found to be between 39 and 44 and was not dependent on the time period or the laboratory. These values fall within a proposed upper limit of protease variation in treated patients of approximately 55% of the protein (Ceccerini-Silberstein et al., 2004). The prevalence of mutations at 20 protease positions not used for genotypic predictions of PI resistance was generally low when compared to established mutations, but increased for higher PI resistance scores and for eight positions, a different amino acid substitution bias was seen in viruses with predicted PI resistance compared to PI-sensitive viruses. These data suggest that these mutations are selected in PI-resistant virus and may contribute to resistance or play a compensatory role.

Recent studies of several hundred to a few thousand PI-treated and PI-naïve patients have identified mutations at 17–22 positions not considered as established resistance-associated positions at the time (Wu et al., 2003; Ceccherini-Silberstein et al.,

2004; Rhee et al., 2005; Svicher et al., 2005). In this study we applied rules-based genotypic predictions to stratify large reference laboratory datasets according to the presence or absence of predicted PI resistance. Eighteen of 20 non-established protease positions identified by our approach were also previously identified in treatment-association studies cited above. The cumulative prevalence of mutations at these positions was highest in viruses with a PI resistance score of 4–7, suggesting that they are more prevalent in heavily PI-experienced patients.

The mutation prevalence at both established and non-established positions also increased with the duration of non-suppressive PI therapy. Emerging mutations that become more prevalent with increased duration of non-suppressive PI therapy may have a compensatory role that, over time, improves the fitness of the viral protease in the face of continued selective pressure from protease inhibitors (Resch et al., 2005). In some cases, novel mutations may take on new roles and contribute to the loss of susceptibility to newly introduced PIs. For example, I13V + L33F has been associated with tipranavir resistance (Doyon et al., 2005) and may be expected to increase in frequency once the use of this PI becomes more widespread. Indeed, the latest IAS-USA chart (October–November, 2005; Johnson et al., 2005b) has categorized several of the non-established mutations identified in the reference laboratory databases or the non-suppressed cohort as associated with resistance to the new PIs atazanavir (85 and 93) or tipranavir (13, 16, 35, 43, 58 and 74).

The L76V mutation, which has become more prevalent in the reference laboratory datasets in patients with a PI resistance index of 4–7, was rarely seen in the non-suppressed dataset. This mutation has been reported to be selected by lopinavir (LPV/r) during therapy failure of salvage therapy and was very uncommon in clinical datasets prior to 2001 when lopinavir was introduced into clinical use (Mueller et al., 2004). Therapy data was not available for the reference laboratory data, thus we could not assess whether the lower prevalence of L76V in the non-supressed dataset may reflect differing patterns of lopinavir usage. In general, however, the mutation prevalence in patients with the longest duration of non-suppressive therapy ($\geq$48 months) correlated with the prevalence in patients with the highest PI resistance scores (4–7) in the reference laboratory datasets. Continuation of PI regimens without viral suppression, therefore, is likely to lead to the continued accumulation of protease mutations.

To further elucidate the possible roles of the non-established mutations reported in the reference laboratory datasets and the BC Centre clinical dataset, we analyzed the co-mutation rates and clustering between 44 protease positions identified in the non-suppressed patients without or with the consideration of specific amino acid substitutions. Other approaches, including mutual information (Hoffman et al., 2003) and binomial correlations, have been used to detect mutation covariation (Wu et al., 2003; Rhee et al., 2005; Svicher et al., 2005). We expanded on this approach by adding neighbor joining analysis to better group the mutations into clusters. Neighbor joining cluster analysis typically employs a distance table derived from a nucleotide or protein sequence alignment, in which each sequence is a taxonomic unit (Saitou and Nei, 1987). A neighbor joining analysis of the 2436 PI-resistant sequences in the SJC 2005 dataset would require 2 965 830 calculated distances between sequence pairs; such an analysis would be computationally intensive and the resulting dendrogram (with 2436 nodes) would be difficult to visualize. Rather than consider each sequence individually, we employed a variation of this approach by treating the protease positions of interest as taxonomic units, and utilized binomial correlations between the positions to measure distances. Thus, for 44 positions, only 946 distance calculations were necessary and the cluster relationships between these positions could be visualized in a dendrogram with 44 nodes (or 79 nodes when specific amino acid substitutions were taken into consideration). In earlier work (Kagan et al., 2004b) we showed that another distance-based method, the Fitch-Margoliash tree-building algorithm (Fitch and Margoliash, 1967), could be applied to these data to yield similar results; however, this method is much more computationally intensive. The co-mutation associations and clusters described in our reference laboratory datasets closely matched those uncovered in treatment association studies. The strongest associations were between D30N + N88D (nelfinavir resistance) and I47V + V32I (amprenavir resistance). A number of non-established mutations also showed evidence of co-mutation, such as I66V + C95F and I66F + L89V. We also confirmed a number of negative associations between specific pairs of mutations including D30N + L90M and L24I + L90M. In patients failing nelfinavir therapy, the D30N and L90M mutations rarely appear together on the same viral clone and the D30N + L90M double mutant is severely impaired (Sugiura et al., 2002). In clinical datasets, the low prevalence of this double mutant likely reflects that individuals tend to switch from failing regimens based on nelfinavir alone. In the BC Centre dataset described here, only one patient remained on nelfinavir after 48 months of non-suppressive therapy. The lower replicative capacity of the D30N mutant (Martinez-Picado et al., 1999; Devereux et al., 2001) would likely promote the disappearance of D30N in this highly treatment-experienced group. The structural basis for the apparent antagonism between L24I and L90M is not known. However, in the crystal structure of HIV-1 protease bound to saquinavir, there are van der Waals interactions between the side chain of L90 and L24 and D25, and the L90M mutation appears to cause a repositioning of residues 24–29 (Hong et al., 2000). The L24I mutation may introduce further structural perturbations within this context and further investigation will be required to better understand these interactions.

The underlying structural basis for co-variation of some protease mutations has been described in detail (Wu et al., 2003; Rhee et al., 2005; Svicher et al., 2005). For example, positively associated positions 46 and 55 are both flap residues positioned only 3.8 Å apart, as are negatively associated positions 24 and 90 (Wu et al., 2003). Likewise, the N88D and K45R mutations may create favorable charge interactions in the D30N nelfinavir-resistant mutant, thus playing a compensatory role (Svicher et al., 2005). Further biochemical and structural analysis will be needed to understand many more of these interactions.

5. Conclusions

In summary, we have utilized large reference laboratory datasets to identify non-established PI resistance-associated mutations and correlated increases in mutational prevalence and PI resistance scores with longer durations of non-suppressive PI therapy. Large reference laboratory datasets, although lacking specific PI treatment histories, can corroborate results derived from studies with available treatment histories. Genotypically-based PI resistance scores may also serve as a useful marker for non-suppressive PI treatment duration. We have also presented a novel adaptation of neighbor joining trees to represent mutational covariation and clustering in large datasets. These techniques contribute to the continued surveillance of resistance-associated mutation patterns and refinement of resistance prediction algorithms as new data on the contributions of these mutations to resistance become available.

Acknowledgements

The authors thank Richard Harrigan (BC Centre for Excellence in HIV/AIDS) for contributing the BC Centre non-suppressed cohort data to this study and for helpful discussions. The authors also wish to thank Jeff Radcliff (Quest Diagnostics Nichols Institute SJC) for carefully reviewing and editing the manuscript.

References

Alexander, C.S., Dong, W., Schechter, M.T., O'Shaughnessy, M.V., Strathdee, S.A., Mo, T., Montaner, J.S., Harrigan, P.R., 1999. Prevalence of primary HIV drug resistance among seroconverters during an explosive outbreak of HIV infection among injecting drug users. AIDS 13, 981–985.

Barbour, J.D., Wrin, T., Grant, R.M., Martin, J.N., Segal, M.R., Petropoulos, C.J., Deeks, S.G., 2002. Evolution of phenotypic drug susceptibility and viral replication capacity during long-term virologic failure of protease inhibitor therapy in human immunodeficiency virus-infected adults. J. Virol. 76, 11104–11112.

Baxter, J.D., Mayers, D.L., Wentworth, D.N., Neaton, J.D., Hoover, M.L., Winters, M.A., Mannheimer, S.B., Thompson, M.A., Abrams, D.I., Brizz, B.J., Ioannidis, J.P., Merigan, T.C., 2000. A randomized study of antiretroviral management based on plasma genotypic antiretroviral resistance testing in patients failing therapy. CPCRA 046 Study Team for the Terry Beirn Community Programs for Clinical Research on AIDS. AIDS 14, F83–F93.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. JR. Stat. Soc. Ser. B. Stat. Methodol. 57, 289–300.

Ceccherini-Silberstein, F., Erba, F., Gago, F., Bertoli, A., Forbici, F., Bellocchi, M.C., Gori, C., D'Arrigo, R., Marcon, L., Balotta, C., Antinori, A., Monforte, A.D., Perno, C.F., 2004. Identification of the minimal conserved structure of HIV-1 protease in the presence and absence of drug pressure. AIDS 18, F11–F19.

Chen, L., Perlina, A., Lee, C.J., 2004. Positive selection detection in 40,000 human immunodeficiency virus (HIV) type 1 sequences automatically identifies drug resistance and positive fitness mutations in HIV protease and reverse transcriptase. J. Virol. 78, 3722–3732.

Cheung, P.K., Wynhoven, B., Harrigan, P.R., 2004. Which HIV-1 drug resistance mutations are common in clinical practice? AIDS Rev. 6, 107–116.

Condra, J.H., Schleif, W.A., Blahy, O.M., Gabryelski, L.J., Graham, D.J., Quintero, J.C., Rhodes, A., Robbins, H.L., Roth, E., Shivaprakash, M., Titus, D., Yang, T., Tepplert, H., Squires, K.E., Deutsch, P.J., Emini, E.A., 1995. In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors. Nature 374, 569–571.

Deeks, S.G., Smith, M., Holodniy, M., Kahn, J.O., 1997. HIV-1 protease inhibitors. A review for clinicians. JAMA 277, 145–153.

Devereux, H.L., Emery, V.C., Johnson, M.A., Loveday, C., 2001. Replicative fitness in vivo of HIV-1 variants with multiple drug resistance-associated mutations. J. Med. Virol. 65, 218–224.

Doyon, L., Tremblay, S., Bourgon, L., Wardrop, E., Cordingley, M.G., 2005. Selection and characterization of HIV-1 showing reduced susceptibility to the non-peptidic protease inhibitor tipranavir. Antiviral Res. 68, 27–35.

Fitch, W.M., Margoliash, E., 1967. Construction of phylogenetic trees. Science 155, 279–284.

Hertogs, K., Bloor, S., Kemp, S.D., Van den Eynde, C., Alcorn, T.M., Pauwels, R., Van Houtte, M., Staszewski, S., Miller, V., Larder, B.A., 2000. Phenotypic and genotypic analysis of clinical HIV-1 isolates reveals extensive protease inhibitor cross-resistance: a survey of over 6000 samples. AIDS 14, 1203–1210.

Hirsch, M.S., Brun-Vezinet, F., Clotet, B., Conway, B., Kuritzkes, D.R., D'Aquila, R.T., Demeter, L.M., Hammer, S.M., Johnson, V.A., Loveday, C., Mellors, J.W., Jacobsen, D.M., Richman, D.D., 2003. Antiretroviral drug resistance testing in adults infected with human immunodeficiency virus type 1: 2003 recommendations of an International AIDS Society-USA Panel. Clin. Infect. Dis. 37, 113–128.

Hoffman, N.G., Schiffer, C.A., Swanstrom, R., 2003. Covariation of amino acid positions in HIV-1 protease. Virology 314, 536–548, Erratum in: Virology 2005. 331, 206–207.

Hogg, R.S., Yip, B., Chan, K.J., Wood, E., Craib, K.J., O'Shaughnessy, M.V., Montaner, J.S., 2001. Rates of disease progression by baseline CD4 cell count and viral load after initiating triple-drug therapy. JAMA 286, 2568–2577.

Hong, L., Zhang, X.C., Hartsuck, J.A., 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.

Johnson, V.A., Brun-Vezinet, F., Clotet, B., Conway, B., Kuritzkes, D.R., Pillay, D., Schapiro, J., Telenti, A., Richman, D., 2005a. Update of the drug resistance mutations in HIV-1: 2005. Top. HIV Med. 13, 51–57.

Johnson, V.A., Brun-Vezinet, F., Clotet, B., Conway, B., Kuritzkes, D.R., Pillay, D., Schapiro, J.M., Telenti, A., Richman, D.D., 2005b. Update of the drug resistance mutations in HIV-1: Fall 2005b. Top. HIV Med. 13, 125–131.

Kagan, R., Winters, M., Merigan, T., Heseltine, P., 2004a. HIV type 1 genotypic resistance in a clinical database correlates with antiretroviral utilization. AIDS Res. Hum. Retrov. 20, 1–9.

Kagan, R., Winters, M., Merigan, T., Heseltine, P., 2004b. Large scale cluster analysis of HIV-1 protease mutations in a clinical database. Antiviral Ther. 9, S92.

Kantor, R., Shafer, R.W., Follansbee, S., Taylor, J., Shilane, D., Hurley, L., Nguyen, D.P., Katzenstein, D., Fessel, W.J., 2004. Evolution of resistance to drugs in HIV-1-infected patients failing antiretroviral therapy. AIDS 18, 1503–1511.

Kohl, N.E., Emini, E.A., Schleif, W.A., Davis, L.J., Heimbach, J.C., Dixon, R.A., Scolnick, E.M., Sigal, I.S., 1988. Active human immunodeficiency virus protease is required for viral infectivity. Proc. Natl. Acad. Sci. USA 85, 4686–4690.

Maguire, M., Shortino, D., Klein, A., Harris, W., Manohitharajah, V., Tisdale, M., Elston, R., Yeo, J., Randall, S., Xu, F., Parker, H., May, J., Snowden, W., 2002. Emergence of resistance to protease inhibitor amprenavir in human immunodeficiency virus type 1-infected patients: selection of four alternative viral protease genotypes and influence of viral susceptibility to coadministered reverse transcriptase nucleoside inhibitors. Antimicrob. Agents Chemother. 46, 731–738.

Martinez-Picado, J., Savara, A.V., Sutton, L., D'Aquila, R.T., 1999. Replicative fitness of protease inhibitor-resistant mutants of human immunodeficiency virus type 1. J. Virol. 73, 3744–3752.

Morgan, E., Tisdale, M., Maguire, M., Macmanus, S., Yates, P., Griffin, P., Xu, F., Elston, R., Kleim, J.P., Pillay, D., Cane, P., 2003. The HIV-1 protease mutation K55R is associated with the presence of the M46I/L mutation. Antiviral Ther. 8, S57.

Mueller, S.M., Daeumer, M., Kaiser, R., Walter, H., Colonno, R., Korn, K., 2004. Susceptibility to saquinavir and atazanavir in highly protease inhibitor (PI) resistant HIV-1 is caused by lopinavir-induced drug resistance mutation L76V. Antiviral Ther. 9, S44.

Novak, R.M., Chen, L., MacArthur, R.D., Baxter, J.D., Huppler Hullsiek, K., Peng, G., Xiang, Y., Henely, C., Schmetter, B., Uy, J., van den Berg-Wolf, M., Kozal, M., Terry Beirn Community Programs for Clinical Research on AIDS 058 Study Team, 2005. Prevalence of antiretroviral drug resistance mutations in chronically HIV-infected, treatment-naive patients: implications for routine resistance screening before initiation of antiretroviral therapy. Clin. Infect. Dis. 40, 468–474.

Patick, A.K., Duran, M., Cao, Y., Shugarts, D., Keller, M.R., Mazabel, E., Knowles, M., Chapman, S., Kuritzkes, D.R., Markowitz, M., 1998. Genotypic and phenotypic characterization of human immunodeficiency virus type 1 variants isolated from patients treated with the protease inhibitor nelfinavir. Antimicrob. Agents Chemother. 42, 2637–2644.

Palella Jr., F.J., Delaney, K.M., Moorman, A.C., Loveless, M.O., Fuhrer, J., Satten, G.A., Aschman, D.J., Holmberg, S.D., 1998. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. N. Engl. J. Med. 338, 853–860.

Peng, C., Ho, B.K., Chang, T.W., Chang, N.T., 1989. Role of human immunodeficiency virus type 1-specific protease in core protein maturation and viral infectivity. J. Virol. 63, 2550–2556.

Pillay, D., Green, H., Matthias, R., Dunn, D., Phillips, A., Sabin, C., Evans, B., U.K. Collaborative Group on HIV Drug Resistance, 2005. Estimating HIV-1 drug resistance in antiretroviral-treated individuals in the United Kingdom. J. Infect. Dis. 192, 967–973.

Resch, W., Parkin, N., Watkins, T., Harris, J., Swanstrom, R., 2005. Evolution of human immunodeficiency virus type 1 protease genotypes and phenotypes in vivo under selective pressure of the protease inhibitor ritonavir. J. Virol. 79, 10638–10649.

Rhee, S.Y., Fessel, W.J., Zolopa, A.R., Hurley, L., Liu, T., Taylor, J., Nguyen, D.P., Slome, S., Klein, D., Horberg, M., Flamm, J., Follansbee, S., Schapiro, J.M., Shafer, R.W., 2005. HIV-1 Protease and reverse-transcriptase mutations: correlations with antiretroviral therapy in subtype B isolates and implications for drug-resistance surveillance. J. Infect. Dis. 192, 456–465.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Shafer, R.W., Kantor, R., Gonzales, J., 2000. The genetic basis of HIV-1 reverse transcriptase and protease inhibitors. AIDS Rev. 2, 211–228.

Sugiura, W., Matsuda, Z., Yokomaku, Y., Hertogs, K., Larder, B., Oishi, T., Okano, A., Shiino, T., Tatsumi, M., Matsuda, M., Abumi, H., Takata, N., Shirahata, S., Yamada, K., Yoshikura, H., Nagai, Y., 2002. Interference between D30N and L90M in selection and development of protease inhibitor-resistant human immunodeficiency virus type 1. Antimicrob. Agents Chemother. 46, 708–715.

Svicher, V., Ceccherini-Silberstein, F., Erba, F., Santoro, M., Gori, C., Bellocchi, M.C., Giannella, S., Trotta, M.P., Monforte, A., Antinori, A., Perno, C.F., 2005. Novel human immunodeficiency virus type 1 protease mutations potentially involved in resistance to protease inhibitors. Antimicrob. Agents Chemother. 49, 2015–2025.

Tamalet, C., Fantini, J., Tourres, C., Yahi, N., 2003. Resistance of HIV-1 to multiple antiretroviral drugs in France: a 6-year survey (1997–2002) based on an analysis of over 7000 genotypes. AIDS 17, 2383–2388.

Weinstock, H.S., Zaidi, I., Heneine, W., Bennett, D., Garcia-Lerma, J.G., Douglas Jr., J.M., LaLota, M., Dickinson, G., Schwarcz, S., Torian, L., Wendell, D., Paul, S., Goza, G.A., Ruiz, J., Boyett, B., Kaplan, J.E., 2004. The epidemiology of antiretroviral drug resistance among drug-naïve HIV-1-infected persons in 10 US cities. J. Infect. Dis. 189, 2174–2180.

Wensing, A.M., van de Vijver, D.A., Angarano, G., Asjo, B., Balotta, C., Boeri, E., Camacho, R., Chaix, M.L., Costagliola, D., De Luca, A., Derdelinckx, I., Grossman, Z., Hamouda, O., Hatzakis, A., Hemmer, R., Hoepelman, A., Horban, A., Korn, K., Kucherer, C., Leitner, T., Loveay, C., MacRae, E., Maljkovic, I., de Mendoza, C., Meyer, L., Nielsen, C., Op de Coul, E.L., Ormaasen, V., Paraskevis, D., Perrin, L., Puchhammer-Stockl, E., Ruiz, L., Salminen, M., Schmit, J.C., Schneider, F., Schuurman, R., Soriano, V., Stanczak, G., Stanojevic, M., Vandamme, A.M., Van Laethem, K., Violin, M., Wilbe, K., Yerly, S., Zazzi, M., Boucher, C.A., SPREAD Programme, 2005. Prevalence of drug-resistant HIV-1 variants in untreated individuals in Europe: implications for clinical management. J. Infect. Dis. 192, 958–966.

Wu, T.D., Schiffer, C.A., Gonzales, M.J., Taylor, J., Kantor, R., Chou, S., Israelski, D., Zolopa, A.R., Fessel, W.J., Shafer, R.W., 2003. Mutation patterns and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments. J. Virol. 77, 4836–4847.