15 Xu Y & Weiss LM. The microsporidian polar tube: a highly specialized invasion organelle. *Int J Parasitol* 2005; **35**: 941–953.
16 Peuvel I, Peyret P, Metenier G, Vivares CP & Delbac F. The microsporidian polar tube: evidence for a third polar tube protein (PTP3) in *Encephalitozoon cuniculi*. *Mol Biochem Parasitol* 2002; **122**: 69–80.
17 Xu Y, Takvorian PM, Cali A, Orr G & Weiss LM. Glycosylation of the major polar tube protein of *Encephalitozoon hellem*, a microsporidian parasite that infects humans. *Infect Immun* 2004; **72**: 6341–6350.
18 van Gool T, Vetter JCM, Weinmayr B, Van Dam A, Derouin F & Dankert J. High seroprevalence of *Encephalitozoon* species in immunocompetent subjects. *J Infect Dis* 1997; **175**: 1020–1024.
19 Peek R, Delbac F, Speijer D, *et al.* Carbohydrate moieties of microsporidian polar tube proteins are targeted by immunoglobulin G in immunocompetent individuals. *Infect Immun* 2005; **73**: 7906–7913.
20 Keohane EM & Weiss LM. The structure, function, and composition of the microsporidian polar tube. In Wittner M & Weiss LM (eds): *The Microsporidia and Microsporidiosis*. Washington, DC, ASM Press, 1999: 196–224.
21 Thomarat F, Vivares CP & Gouy M. Phylogenetic analysis of the complete genome sequence of *Encephalitozoon cuniculi* supports the fungal origin of microsporidia and reveals a high frequency of fast-evolving genes. *J Mol Evol* 2004; **59**: 780–791.
22 Evering T & Weiss LM. The immunology of parasite infections in immunocompromised hosts. *Parasite Immunol* 2006; **28**: 549–565.
23 Kodjikian L, Garweg JG, Nguyen M, Schaffner T, Deplazes P & Zimmerli S. Intraocular microsporidiosis due to *Encephalitozoon cuniculi* in a patient with idiopathic $CD4^+$ T-lymphocytopenia. *Int J Med Microbiol* 2005; **294**: 529–533.
24 Sak B, Salat J, Horka H, Sakova K & Ditrich O. Antibodies enhance the protective effect of $CD4^+$ T lymphocytes in SCID mice perorally infected with *Encephalitozoon cuniculi*. *Parasite Immunol* 2006; **28**: 95–99.

CLINICAL AND VACCINE IMMUNOLOGY, Feb. 2007, p. 168–172
1556-6811/07/$08.00+0 doi:10.1128/CVI.00224-06

Vol. 14, No. 2

Detecting Immunoglobulin M Antibodies against Microsporidian *Encephalitozoon cuniculi* Polar Tubes in Sera from Healthy and Human Immunodeficiency Virus-Infected Persons in Japan[▽]

Mako Omura,[1] Koji Furuya,[1]* Shinichi Kudo,[2] Wataru Sugiura,[3] and Hiroshi Azuma[4]

Department of Parasitology, National Institute of Infectious Diseases, Tokyo, Japan[1]; Hokkaido Institute of Public Health, Sapporo, Japan[2]; AIDS Research Center, National Institute of Infectious Diseases, Tokyo, Japan[3]; and Hokkaido Red Cross Blood Center, Sapporo, Japan[4]

Received 18 June 2006/Returned for modification 1 September 2006/Accepted 6 November 2006

Encephalitozoon cuniculi*, a spore-forming obligate intracellular parasitic pathogen belonging to the phylum Microsporidia, has a unique and highly specialized organelle called the polar tube. Using an enzyme immunostaining assay in which germinated *E. cuniculi* spores were coated onto plastic surfaces, we tested healthy and human immunodeficiency virus (HIV)-infected individuals in Japan for anti-polar tube antibodies of each immunoglobulin (Ig) class. Anti-polar tube IgG was detected in just 4 of 380 healthy individuals; no anti-polar tube IgA was detected in any individuals; however, unexpectedly, anti-polar tube IgM antibodies were detected in 138 individuals (36%). When the healthy individuals were grouped by age, the highest rate of positivity to anti-polar tube IgM antibodies was seen in individuals aged 20 years old or younger. Fifty-nine percent (24/41) of the individuals aged 20 years or younger were anti-polar tube IgM antibody positive. This rate tended to decrease among individuals in older age groups. However, no anti-polar tube IgM antibodies were detected in 21 HIV-infected persons who were younger than 30 years of age and who had CD4 cell levels below 250/μl. These seroepidemiological results clearly indicate that circulating anti-polar tube IgM antibodies that are capable of strongly reacting with filaments extruded from geminated spores exist and suggest that such antibodies may play a part in protective immunity.

Encephalitozoon cuniculi is a microsporidian parasitic pathogen listed in a 1996 WHO report as an emerging infectious agent (17). The pathogen is also considered a zoonotic parasite (4). Various animals can be naturally infected by *E. cuniculi*, and its geographical distribution is worldwide (3). In Japan, *E. cuniculi* infection in rabbits and in squirrel monkeys in zoos is of current concern (1, 5). The rate of infection is considered to increase each year, and the infection has now spread throughout Japan. However, to the best of our knowledge, the only case of human microsporidiosis reported in Japan was in a 9-year-old boy in 1958 (10). Although immunological conditions of the Japanese case were not recorded, almost all other patients infected with this pathogen in other nations have been immunocompromised groups of human immunodeficiency virus (HIV)-infected patients (16). A few cases have also been found among renal transplant recipients (6, 11). *E. cuniculi* can thus be regarded as an opportunistic pathogen (2). Cases of HIV-associated infections with *E. cuniculi* are increasingly reported, although they remain less common than those due to *Encephalitozoon bieneusi* and *Encephalitozoon intestinalis* (2). Many reports on the seroprevalence of human *E. cuniculi* infection have been published (3, 8). However, the reported rates of microsporidial seropositivity vary greatly, depending on the serological technique used, probably due to the use of antigens unsuitable for measurement of specific antibodies and the use of secondary antibodies without differential specificities.

Recently, specific immunoglobulin G (IgG) antibodies against the polar tube (PT) of *E. cuniculi* were demonstrated in a healthy laboratory worker accidentally infected with *E. cuniculi* (14). The PT is a typical microsporidian spore structure with an extrusion that is essential for invasion of a host cell, as sporoplasm flows through the discharged PT and into the host cell (13).

We have recently developed an enzyme immunostaining assay (EIA) for measuring anti-*E. cuniculi* PT antibodies using 96-well microplates coated with germinated spores. This method allows us to screen human sera for anti-PT antibodies on a large scale for seroepidemiological analysis. This study reports on the screening of sera from 380 healthy persons and 78 HIV-infected persons seroepidemiologically analyzed by this particular EIA, which is capable of measuring anti-PT antibodies of each Ig class, that is, IgM, IgG, and IgA.

MATERIALS AND METHODS

Serum samples. For this study we used serum samples from 380 healthy people living in Hokkaido Prefecture, Japan; serum samples from 180 residents who underwent a serological test for parasitosis in 2000 but who showed negative results; and serum samples from 200 blood donors collected in 2005.

Serum samples from 78 HIV-infected persons, collected in 1999 from the Kanto region of Japan, were also provided for this study. These included sera from 51 persons with CD4 cell levels below 250/μl and sera from 27 persons with CD4 cell levels between 251 and 900/μl. The 51 persons in the former group were of various ages, while the 27 persons in the latter group were younger than 30 years of age. Tests for HIV infection and determination of CD4 lymphocyte counts were performed by standard laboratory protocols.

, *E. cuniculi* spores. For this study we used the *E. cuniculi* HF strain, isolated from a rabbit with encephalitozoonosis. Strain HF was then cultivated in RK-13

* Corresponding author. Mailing address: Department of Parasitology, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan. Phone: 81(3)5285-1111. Fax: 81(3)5285-1173. E-mail: kfuruya@nih.go.jp.

▽ Published ahead of print on 15 November 2006.

TABLE 1. Results of serological detection of anti-*E. cuniculi* PT antibodies in healthy residents, blood donors, and HIV-infected persons

Subject group and antibody	No. (%[a]) of individuals with antibody titer of:							
	<1/50	1/50	1/100	1/200	1/400	1/800	1/1,600	Total
Healthy residents								
IgM	114 (63.3)	26 (14.4)	16 (8.9)	16 (8.9)	5 (2.8)	3 (1.7)	0	180
IgG	177	3	0	0	0	0	0	180
IgA	180	0	0	0	0	0	0	180
Blood donors								
IgM	128 (64)	29 (14.5)	27 (13.5)	11 (5.5)	2 (1.0)	2 (1.0)	1 (0.5)	200
IgG	199	0	0	0	0	0	1	200
IgA	200	0	0	0	0	0	0	200
HIV-infected persons with CD4 cell counts below 250/μl								
IgM	47	2	1	1	0	0	0	51
IgG	49	0	1	1	0	0	0	51
IgA	51	0	0	0	0	0	0	51

[a] Rates of positivity for anti-PT IgM antibodies.

cells (ATCC CCL-37) (5). Culture supernatants of HF-infected RK-13 cells were collected, centrifuged, and used for serological tests.

Strain HF was genetically analyzed beforehand by PCR, followed by direct DNA sequencing (1). The internal transcribed spacer gene sequence revealed that strain HF was classified into genotype I, since it contained three GTTT repeats. Sequence analysis of the spore wall protein I gene revealed that the strain belonged to genotype Ia because of the amplification of a 399-bp PCR product.

Microplate enzyme immunostaining assay. Sediments containing germinated spores, nongerminated spores, and heavily infected cells detached from cell sheets, were suspended in Gibco minimal essential medium including Earle's salts and glutamine (Invitrogen Corporation, Grand Island, NY) and supplemented with 1,000 U/ml penicillin G, 1,000 μg/ml streptomycin, and 10% fetal bovine serum; this medium was also used for cultures of RK-13 and BS-C-1 (ATCC CCL-26) cells, as described below. Approximately 4×10^6 free spores (containing detached cells) were inoculated into each well of a 96-well flat bottom microplate (high-binding polystyrene; Corning Incorporated, NY) and cultured for 3 days at 35°C in an incubator with 5% CO_2. Subsequently, the wells were washed once with phosphate-buffered saline (PBS; pH 7.2) and fixed with 10% formalin in PBS for 1 h at room temperature (RT). The wells were then washed twice with PBS, treated with PBS containing 1% Tween 20 for 1 h at RT, and washed twice with PBS. The wells were further treated with blocking buffer (SuperBlock; Pierce, Rockford, IL) for 1 h at RT and were finally washed twice with PBS. The plates with wells coated with more than 100 germinated spores per well were then studied.

Twofold dilutions of each serum sample were made by using PBS containing 0.05% Tween 20 (PBS-T), starting from a 1:50 dilution; 100 μl of each of the dilutions was added to each coated well. The wells were incubated for 90 min at RT and then washed five times with 200 μl PBS-T. Subsequently, the wells were incubated with 100 μl of the secondary antibody or protein A/G, incubated for 60 min at RT, and washed five times with PBS-T. A 1:3,000 dilution of protein A/G labeled with peroxidase (PO) (Prozyme Inc., San Leardro, CA), a 1:5,000 dilution of anti-human IgM (μ-chain specific) labeled with PO (QED Bioscience Inc., San Diego, CA), and a 1:3,000 dilution of anti-human IgA (Fc specific) labeled with PO (Nordic Immunology, The Netherlands) were used to capture IgG, IgM, and IgA antibodies, respectively. Finally, the signals of PO bound to human Ig antibodies were visualized by using the substrate aminoethyl carbazole (Zymed Laboratories Inc., San Francisco, CA), according to the manufacturer's instructions. After the wells were washed with pure water, the results were observed with a light microscope. In this assay, only judgments on the serological reactions to filaments extruded from germinated spores (i.e., PTs) were made, while the reactions to the spore walls and the host cells were recorded as reference data.

The concentrations of each secondary antibody, noted above, were determined beforehand by using a dot immunoassay and the corresponding purified Ig. Sera from rabbits with encephalitozoonosis were used as positive controls (5). PO-labeled protein A/G was used for detection of rabbit IgG antibodies as the secondary antibody.

BS-C-1 cells were infected with *E. intestinalis* (ATCC 50057) and *E. hellem* (ATCC 50451) spores. The resultant germinated spores were examined by the procedures mentioned above.

Ethical considerations. The protocol for this study was approved by the Committee for Research on Human Subjects of the National Institute of Infectious Diseases, Tokyo, Japan. Written informed consent was obtained from the HIV-positive subjects. The use of healthy residents' sera and the use of blood donors' sera were approved by the Institutional Review Board of the Hokkaido Institute of Public Health and the Institutional Review Board of the Hokkaido Red Cross Blood Center, respectively. All serum samples included in this study were processed to protect personal information. For all serum samples, the only specific clinical information available was the sex, age, and health condition.

Statistical analysis. The sera of the healthy residents and blood donors were each divided into six groups according to age. The relationship between anti-PT IgM prevalence in each age group and the year of blood collection (2000 for healthy residents and 2005 for blood donors) was analyzed by the Mantel-Haenszel method. The relationship between increasing age and decreasing rate of positivity for anti-PT IgM antibodies in healthy subjects (healthy residents plus blood donors) was analyzed by the Cochran-Armitage test. The statistical significance of the presence of anti-PT IgM antibodies by gender was determined by the chi-square test. P values of <0.01 were considered statistically significant. Excel Statistics 2006 software for Windows (release 6.7.1; Social Survey Research Information Co. Ltd., Tokyo, Japan) was used for statistical analysis.

RESULTS

Table 1 summarizes the results of the EIA for the detection of anti-*E. cuniculi* PT antibodies in samples from healthy residents, blood donors, and HIV-infected persons. Anti-PT IgG antibodies were detected in only 3 of 180 serum samples from healthy individuals; the titers were 1:50, which was significantly lower than those for the controls (naturally infected rabbit sera), which showed titers of 1:6,400 to 1:102,400. When the 200 donor serum samples were examined by EIA, anti-*E. cuniculi* PT IgG antibodies were detected in only 1 serum sample; the titer was 1:1,600, but the positive signals were very weak even at the lower dilutions.

No anti-PT IgA antibody was detected in any of the 380 serum samples with titers below 1:50. Furthermore, 2 of the 51 HIV-infected persons with CD4 cell levels below 250/μl had anti-PT IgG antibodies, and anti-PT IgA antibodies were not detected in any of the 51 HIV-infected persons.

On the other hand, when the same serum samples described above were examined for the presence or absence of anti-PT IgM antibodies, the results were quite different from those for

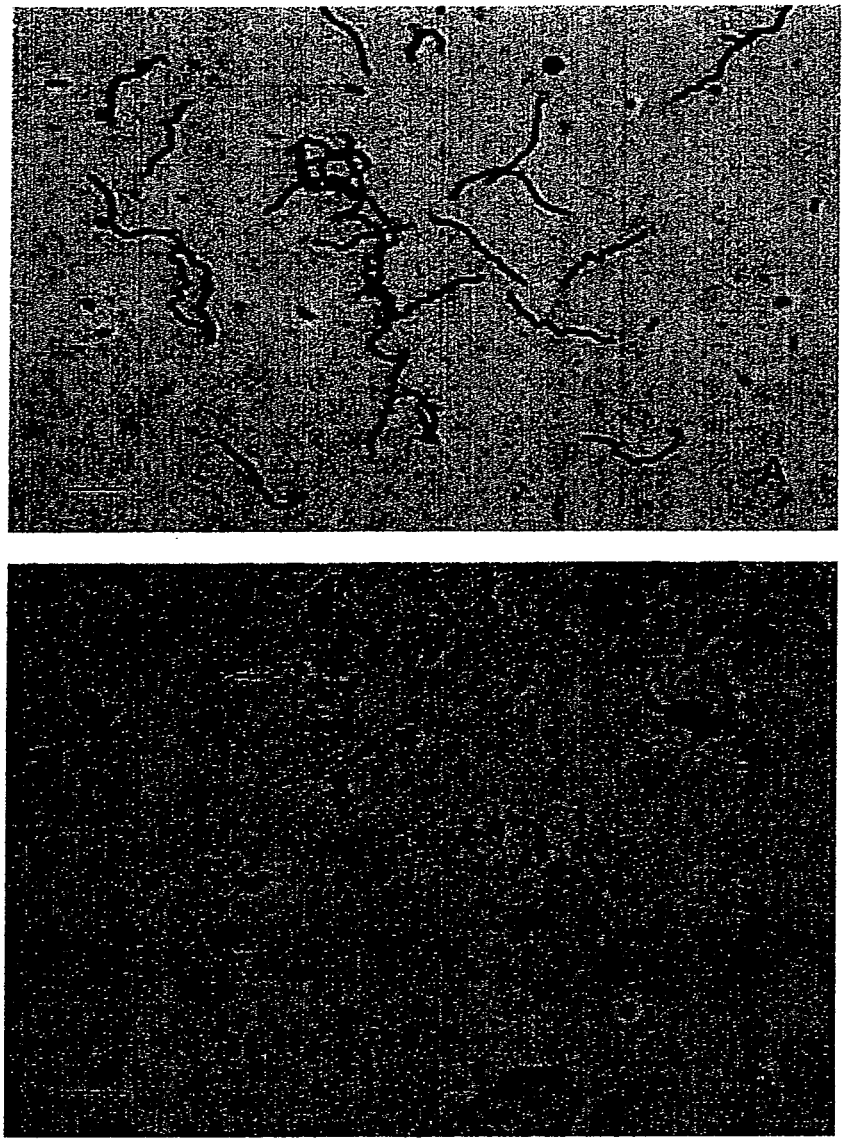

FIG. 1. Immunostaining of filaments (PTs) extruded from *E. cuniculi*-geminated spores with human IgM antibodies or rabbit IgG antibodies. (A) Positive results obtained from the serum sample from donor 197 diluted 1:200. Note the strongly positive signals on the filaments extruded from the germinated spores but the unstained spore walls. (B) Positive results obtained from a serum sample from a symptomatic rabbit with natural *E cuniculi* infection diluted 1:400. Note the positive signals on the spore wall and the filament. Bars, 10 μm.

anti-PT IgG and IgA antibodies. Anti-PT IgM antibodies were detected in 66 (36.7%) of the 180 serum samples from healthy persons, which showed titers of 1:50 to 1:800; and in 72 (36%) of the 200 serum samples from the blood donors, which showed titers of 1:50 to 1:1,600. The reactivities of the IgM antibodies with filaments were considered to be typical of IgM, because of the low titers (below 1:1,600), but the IgM antibodies had stronger reactivities than the rabbit anti-*E. cuniculi* PT IgG antibodies (Fig. 1A and B). The four serum samples from healthy persons and donors with anti-PT IgG activities also had anti-PT IgM activities.

By using almost the same procedures used for the EIA used for detection of anti-*E. cuniculi* PT IgM antibodies, sera with titers of more than 1:200 were examined for cross-reactions with PTs of *E. hellem* and *E. intestinalis*; however, no antibody activity against the filaments was detected (data not shown).

A decreasing trend in positivity rates for anti-PT IgM antibodies was observed when subjects (healthy residents plus blood donors) were grouped according to age. As seen from Table 2, the rate of positivity for anti-PT IgM antibodies was the highest among people aged 19 years or younger: 59% of healthy subjects. The seropositivity rates clearly tended to decrease among the older subjects ($P < 0.01$). The rates of positivity for anti-PT IgM antibodies in each age group were irrelevant to the year that the serum samples were collected, i.e., in 2000 (healthy residents) and 2005 (blood donors). The rate of positivity for anti-PT IgM antibodies among females was a little higher than that among males: 43.8% (39/89) for

TABLE 2. Age distribution of cases with anti-*E. cuniculi* PT IgM antibodies of healthy and HIV-infected persons

Subject group	Value for age group											
	19 yr or younger		20–29 yr		30–39 yr		40–49 yr		50–59 yr		60 yr or older	
	No. of individuals tested	No. (%) positive	No. of individuals tested	No. (%) positive	No. of individuals tested	No. (%) positive	No. of individuals tested	No. (%) positive	No. of individuals tested	No. (%) positive	No. of individuals tested	No. (%) positive
Healthy residents[a]	31	18 (58.1)	29	13 (44.8)	26	12 (46.2)	31	4 (12.9)	32	10 (31.3)	31	9 (29.0)
Blood donors[a]	10	6 (60.0)	46	19 (41.3)	54	19 (35.2)	33	7 (21.2)	44	18 (40.9)	13	3 (23.1)
Healthy residents and blood donors[b]	41	24 (58.5)	75	32 (42.7)	80	31 (38.8)	64	11 (17.2)	76	28 (36.8)	44	12 (27.3)
HIV-infected persons with CD4 cell counts below 250/μl	4	0 (0)	17	0 (0)	16	2 (12.5)	10	1 (10.0)	2	1 (50.0)	2	0 (0)

[a] Serum samples were collected in 2000 from healthy residents and in 2005 from blood donors. The relationship between anti-PT IgM prevalence in each age group and the year of blood collection was analyzed by the Mantel-Haenszel method, which did not give statistically significantly different values.
[b] An association between increasing age and a decreasing rate of positivity for anti-PT IgM was analyzed by the Cochran-Armitage test, which gave statistically significant P values of <0.01.

TABLE 3. Relationship between number of CD4 cells and rate of positivity for anti-*E. cuniculi* PT IgM antibodies among HIV-infected persons younger than 30 years of age

No. of CD4 cells/μl	No. of persons examined	No. (%) of persons with anti-*E. cuniculi* PT IgM antibodies
<250	21	0 (0)
251–399	12	3 (25)
400–900	15	9 (60)

females and 29.7% (27/91) for males. A total of 41.1% (39/95) of the female donors and 31.4% (33/105) of the male donors showed anti-PT IgM antibody titers of 1:50 or more.

Anti-*E. cuniculi* PT IgM antibodies were detected in only 4 of 51 samples from HIV-positive individuals with CD4 cell levels below 250/μl (Table 1). In particular, anti-PT IgM antibodies were not detected in persons younger than age 30 years and with <250 CD4 cells/μl (Table 2). Interestingly, a high rate of positivity for anti-PT IgM antibodies was observed among the 27 HIV-positive persons with CD4 cell counts between 251 and 900/μl; i.e., 25% (3/12) of people with CD4 cell counts between 251 and 399/μl and in 60% (9/15) of people with CD4 cell counts between 400 and 900/μl (Table 3); furthermore, all these individuals were younger than 30 years of age.

DISCUSSION

Our present results indicate that anti-*E. cuniculi* PT antibodies could be detected in 36% of the people, healthy residents and blood donors, who should be considered immunocompetent. In respect to antibodies against *Encephalitozoon* PT among immunocompetent persons, it has been reported that anti-*E. intestinalis* PT was demonstrated in 8% of Dutch blood donors and 5% of pregnant French women (15). Our anti-*E. cuniculi* PT antibodies were detected by EIA, while the anti-*E. intestinalis* PT antibodies were detected by an indirect fluorescent-antibody test (15). The sensitivities of enzyme immunoassays are generally believed to be fairly higher than those of immunofluorescence assays.

No cross-reactive relationship between human anti-*E. cuniculi* PT and human anti-*E. intestinalis* PT has been proved. We clearly showed that our human sera containing anti-PT IgM antibodies did not cross-react with filaments extruded from germinated spores of *E. intestinalis* and *E. hellem*. Our finding is in agreement with previous ones that human sera containing anti-*E. intestinalis* PT antibody activity did not immunostain the filaments extruded from *E. cuniculi* (13).

Anti-*E. cuniculi* PT antibodies, unlike the immunoglobulin class of anti-*E. intestinalis* PT antibodies, which was shown to be IgG (13, 15), were not of the IgG class but were of the IgM class. Anti-*E. cuniculi* PT IgG antibodies were detected in the sera of just four persons (one had a titer of 1:1,600, and the others showed low titers of 1:50) (Table 1). Their reactivities probably resulted from anti-*E. cuniculi* PT IgM antibodies. In our experiments, we used protein A/G instead of anti-human IgG to capture IgG antibodies. It is known that protein A strongly binds to IgG molecules but also weakly binds to some IgM molecules (9). In fact, the four serum samples all exhib-

ited elevated titers of anti-PT IgM antibodies, but their ability to stain extruded filaments was not so strong.

There was no significant difference in the rate of positivity for anti-PT IgM antibodies by gender, although the rate of positivity for anti-PT IgM antibodies was slightly higher among females than among males. However, some relationship between the prevalence of anti-PT IgM antibodies and age would be expected. The rate of positivity for anti-PT IgM antibodies was significantly higher among people <20 years of age than among people in older groups (Table 2). However, this activity for anti-PT IgM antibodies was not found among the 21 HIV-positive persons younger than 30 years of age (Table 2), all of whom had CD4 lymphocyte levels below 250/μl, indicating that they were severely immunocompromised. However, when CD4 counts were between 400 and 900 cells/μl, anti-PT IgM antibodies were detectable in 60% (9/15) of HIV-positive persons younger than 30 years of age (Table 3). Thus, we were surprised to find that increasing age and decreasing numbers of CD4 lymphocytes, factors that can induce immunosuppression, influence the production of anti-PT IgM antibodies.

Anti-PT IgG antibodies were not detectable in most of our subjects, as noted above. The only human case demonstrating elevated anti-PT IgG antibodies involved an accidental *E. cuniculi* infection (14). In respect to the specific immune responses to the *E. cuniculi* infection, anti-spore wall IgG was observed to precede anti-PT IgG (14). Most of our subjects were negative for anti-spore wall IgM and IgG (data not shown). Additionally, the rate of positivity for anti-PT IgM antibodies for the serum samples collected from blood donors in 2005 was almost the same as that for the serum samples collected from healthy residents in 2000 (Table 1). These findings suggest that most of our anti-PT IgM antibodies do not belong to the class of early IgM antibodies found after infection by *E. cuniculi*. Although anti-PT IgM reacted only with the *E. cuniculi* PTs of the *Encephalitozoon* sp. tested, as described above, further research concerning the specificities and immunoreactivities of human anti-*E. cuniculi* PT IgM antibodies needs to be undertaken.

E. cuniculi infection in immunocompromised patients results in disseminated disease that is clinically manifested in symptoms such as keratoconjunctivitis, hepatitis, and peritonitis (16). However, no symptomatic cases of infection with *E. cuniculi* among immunocompetent persons have been described (16), apart from the accidentally infected French individual, who had severe keratoconjunctivitis (14). A Japanese child with encephalitozoonosis due to *E. cuniculi* infection in 1958 is considered the only case due to natural infection, but unfortunately, the patient's immune status was not recorded (10). A few cases of *E. cuniculi* infection have also been reported in transplant patients (6, 11). Thus, apart from some extremely rare situations, it is most unlikely that *E. cuniculi* causes microsporidiosis in immunocompetent persons (12, 16). Considering that almost all human encephalitozoonosis cases occurred in immunocompromised patients infected with HIV (16), one can speculate that protective immunity plays a very important role against *E. cuniculi* infection. In experimental models, the protective immune response against *E. cuniculi* is noted to be mediated by cytotoxic $CD8^+$ T cells (7). Also, the in vitro infectivity of microsporidia has been observed to be reduced by treatment with monoclonal and polyclonal antibodies to the polar tube protein (7), suggesting that anti-PT antibody may constitute a first line of defense against infection by *E. cuniculi*. Our study clearly indicates that there are circulating IgM antibodies that are capable of strongly reacting with the filaments that extrude from germinated *E. cuniculi* spores. We believe that this is the first study to provide seroepidemiological data on human anti-PT IgM antibodies. Further studies focused on human anti-*E. cuniculi* PT IgM antibodies need to be performed from the parallel perspectives of protective immunity and preventive medicine.

ACKNOWLEDGMENTS

We appreciate the assistance with serological testing and cell culture given by Tokiko Asakura.

This work was supported by grant-in-aid no. 16390177 from the Ministry of Education, Science, Sports and Culture of Japan.

REFERENCES

1. **Asakura, T., S. Nakamura, M. Ohta, Y. Une, and K. Furuya.** 2006. Genetically unique microsporidian *Encephalitozoon cuniculi* strain type III isolated from squirrel monkeys. Parasitol. Int. 55:159–162.
2. **Bryan, R. T.** 1995. Microsporidiosis as an AIDS-related opportunistic infection. Clin. Infect. Dis. 21(Suppl. 1):S62–S65.
3. **Canning, E. U., and W. S. Hollister.** 1987. Microsporidia of mammals—widespread pathogens or opportunistic curiosities? Parasitol. Today 3:267–273.
4. **Deplazes, P., A. Mathis, R. Baumgartner, I. Tanner, and R. Weber.** 1996. Immunologic and molecular characteristics of *Encephalitozoon*-like microsporidia isolated from humans and rabbits indicate that *Encephalitozoon cuniculi* is a zoonotic parasite. Clin. Infect. Dis. 22:557–559.
5. **Furuya, K., D. Fukui, M. Yamaguchi, Y. Nakaoka, G. Bando, and M. Kosuge.** 2001. Isolation of *Encephalitozoon cuniculi* using primary tissue culture techniques from a rabbit in a colony showing encephalitozoonosis. J. Vet. Med. Sci. 63:203–206.
6. **Gamboa-Dominguez, A., J. De Anda, J. Donis, F. Ruiz-Maza, G. S. Visvesvara, and H. Diliz.** 2003. Disseminated *Encephalitozoon cuniculi* infection in a Mexican kidney transplant recipient. Transplantation 75:1898–1900.
7. **Khan, I. A., M. Moretto, and L. M. Weiss.** 2001. Immune response to *Encephalitozoon cuniculi* infection. Microbes Infect. 3:401–405.
8. **Kucerova-Pospisilova, Z., and O. Ditrich.** 1998. The serological surveillance of several groups of patients using antigens of *Encephalitozoon hellem* and *E. cuniculi* antibodies to microsporidia in patients. Folia Parasitol. (Praha) 45:108–112.
9. **Langone, J. J.** 1982. Protein A of *Staphylococcus aureus* and related immunoglobulin receptors produced by streptococci and pneumonococci. Adv. Immunol. 32:157–252.
10. **Matsubayashi, H., T. Koike, I. Mikata, H. Takei, and S. Hagiwara.** 1959. A case of *Encephalitozoon*-like body infection in man. Arch. Pathol. 67:181–187.
11. **Mohindra, A. R., M. W. Lee, G. Visvesvara, H. Moura, R. Parasuraman, G. J. Leitch, L. Xiao, J. Yee, and R. del Busto.** 2002. Disseminated microsporidiosis in a renal transplant recipient. Transplant. Infect. Dis. 4:102–107.
12. **Orenstein, J. M., H. P. Gaetz, A. T. Yachnis, S. S. Frankel, R. B. Mertens, and E. S. Didier.** 1997. Disseminated microsporidiosis in AIDS: are any organs spared? AIDS 11:385–386.
13. **Peek, R., F. Delbac, D. Speijer, V. Polonais, S. Greve, E. Wentink-Bonnema, J. Ringrose, and T. van Gool.** 2005. Carbohydrate moieties of microsporidian polar tube proteins are targeted by immunoglobulin G in immunocompetent individuals. Infect. Immun. 73:7906–7913.
14. **van Gool, T., C. Biderre, F. Delbac, E. Wentink-Bonnema, R. Peek, and C. P. Vivares.** 2004. Serodiagnostic studies in an immunocompetent individual infected with *Encephalitozoon cuniculi*. J. Infect. Dis. 189:2243–2249.
15. **van Gool, T., J. C. M. Vetter, B. Weinmayr, A. Van Dam, F. Derouin, and J. Dankert.** 1997. High seroprevalence of *Encephalitozoon* species in immunocompetent subjects. J. Infect. Dis. 175:1020–1024.
16. **Weber, R., D. A. Schwartz, and P. Deplazes.** 1999. Laboratory diagnosis of microsporidiosis, p. 315–362. *In* M. Wittner and L. M. Weiss (ed.), The microsporidia and microsporidiosis. ASM Press, Washington, DC.
17. **World Health Organization.** 1996. Fighting disease fostering development, p. 112. *In* The world health report. World Health Organization, Geneva, Switzerland.

AIDS RESEARCH AND HUMAN RETROVIRUSES
Volume 23, Number 4, 2007, pp. 564–568

DOI: 10.1089/aid.2006.0270

Sequence Note

Molecular Epidemiology of HIV Type 1 in Treatment-Naive Patients in North Ethiopia

AFEWORK KASSU,[1,2,*] MASAYUKI FUJINO,[3,*] MASAKAZU MATSUDA,[3] MASAKO NISHIZAWA,[3] FUSAO OTA,[1] and WATARU SUGIURA[3]

ABSTRACT

To understand the predominant HIV subtype and drug-resistant viruses in northwest Ethiopia, isolates from 92 antiretroviral drug-naive HIV-1-infected tuberculosis patients were analyzed. Of these patients, 90 (97.8%) were found to be infected with viral subtype C. Other isolates had subtype A (1.1%) and subtype D (1.1%). No primary mutations were associated with protease inhibitor drug resistance. One case (1.1%) had the reverse-transcriptase mutation, V75I. Two patients (2.2%) had the G190A mutation, which confers resistance to the nonnucleoside reverse transcriptase inhibitor, nevirapine. Our study demonstrates that subtype C is the major HIV-1 subtype in northwest Ethiopia. Our results also reveal that the population in the study area had been exposed to antiretrovirals and that treatment-naive patients had drug resistance mutations. Thus, our results emphasize the need for routine drug resistance monitoring in northwest Ethiopia.

ETHIOPIA, A COUNTRY IN THE HORN OF AFRICA and neighboring Sudan, Kenya, Djibouti, and Somalia, has been facing an epidemic of HIV-1 subtype C with low frequency of subtypes A and D.[1–3] Since the first two reported AIDS cases in 1986, the epidemic has been spreading rapidly throughout the country. At the end of 2003, 1.5 million people in Ethiopia were living with HIV/AIDS. The adult prevalence of HIV in 2003 was estimated at 4.4%, and the number of people infected with the virus is projected to exceed 4 million by the year 2014. About 245,000 people living with HIV/AIDS in 2003 were estimated to need antiretroviral therapy. To control the epidemic, the government of Ethiopia has developed a national policy to guide the implementation of successful programs to handle HIV/AIDS, including the provision of highly active antiretroviral therapy (HAART). As a result, antiretroviral (ARV) therapy has officially been offered since 2003 in the country. However, the use of these drugs has created a major public health concern about the possible emergence of drug-resistant mutants that would lead to treatment failure. Given the limited availability of treatment options, optimizing initial ARV therapy depends on characterizing baseline polymorphisms in the antiretroviral target genomes and their impact on the development of drug resistance.

To clarify the distribution of HIV-1 subtypes and prevalence of drug resistance-related mutations in antiretroviral drug-naive HIV-1-infected patients in Gondar, northwest Ethiopia, we collected and analyzed serum from 92 ARV drug-naive, HIV-1-positive tuberculosis patients diagnosed from January to August 2003 at the University of Gondar Hospital. Selection of patients and sample collection have been reported elsewhere.[4] Nucleotide sequences of HIV-1 *gag* p17, protease (PR), reverse transcriptase (RT), and envelope C2V3 regions were determined as previously described.[5] In brief, viral RNA was extracted from 200 μl of serum using a commercially available kit (Boehringer Mannheim GmbH, Mannheim, Germany). Four regions were separately reverse transcribed and amplified using AMV-RT and Taq polymerase (TaKaRa, Osaka, Japan) and region-specific primer pairs. Primary PCR products were fur-

[1]Department of Preventive Environment and Nutrition, Institute of Health Biosciences, University of Tokushima, Tokushima 770-8503, Japan.
[2]Department of Microbiology and Parasitology, Gondar College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia.
[3]AIDS Research Center, National Institute of Infectious Diseases, Tokyo 2080011, Japan.
*These authors contributed equally to this study.

TABLE 1. SOCIODEMOGRAPHIC CHARACTERISTICS OF PATIENTS

	Frequency	*%*
Sex		
Male	36	39.1
Female	56	60.9
Age in years		
<20	1	1.1
20–29	38	41.3
30–39	38	41.3
40–49	11	12
>50	4	4.3
Religion		
Christian	86	93.5
Muslim	5	5.4
Other	1	1.1
Residence		
Gondar city	75	81.5
Surrounding Gondar	17	18.5
Marital status		
Single	36	39.1
Married	24	26.1
Divorced	21	22.8
Widowed	11	12

ther amplified with a high-fidelity DNA polymerase (KOD DNA polymerase, TOYOBO, Osaka, Japan). Sequence reactions were performed using the BigDye terminator cycle sequencing kit (Applied Biosystems, Tokyo, Japan), and nucleotide sequences were determined by direct sequencing using the 3730 DNA ANALYSER (Applied Biosystems). The results were analyzed by Sequence Navigator version 1.0.1 software (Applied Biosystems). The virus subtype of each sample was determined by phylogenetic analysis of *env* C2V3, *gag* p17, PR, and RT gene sequences. The sequences were aligned with a set of reference sequences obtained from the Los Alamos sequence database by Clustal W of the Philip program. The genetic distances were calculated by Kimura's two-parameter analyses, and phylogenetic trees were depicted by the neighbor-joining method using the MEGA version 3.0 program. To determine the prevalence and pattern of drug resistance mutations, sequences of PR and RT were compared with the HXB2 reference sequence and mutation points were determined according to the IAS-USA drug-resistance chart.[6]

Sequences for PR, RT, and gag p17 were successfully obtained from all 92 specimens. Envelop C2V3 sequences were obtained only from 90 samples. The baseline characteristics of the 92 subjects are shown in Table 1. The majority of patients (82.6%) were in the age group 20–39 years. Women, residents of Gondar city, and Christians represented 60.9, 81.5, and 93.5% of the patients, respectively. Phylogenetic analyses of sequences clarified that 90 (97.8%) of the 92 subjects were infected with HIV-1 subtype C. Two nonsubtype C cases were identified: one subtype A (1.1%) and one subtype D (1.1%). All four segments amplified by RT-PCR demonstrated identical subtypes and no recombinant forms were identified. Thus, our results show that subtype C is the predominant subtype among patients living in the study area. Our result is consistent with previous reports that the Ethiopian HIV/AIDS epidemic is dominated by subtype C viruses.[1–3] These studies also reported very low frequencies of subtypes A and D.[2,3] The dominance of subtype C in the Ethiopian epidemic might be due to the rapid saturation of this subtype among commercial sex workers and their network.[1,3] Indeed, heterosexual contact was reported as the major risk behavior in the Ethiopian HIV epidemic. Our finding of the rarely reported subtypes A and D in patients from northwest Ethiopia, in addition to the dominant subtype C, suggests the introduction of such clades from neighboring countries such as Sudan[7] and Kenya,[8] where subtype D and A viruses predominate.

To understand subtypes in more detail, phylogenetic analyses were performed. Figure 1 shows an unrooted phylogenetic tree of the sequences for *env* C2V3, *gag* p17, PR, and RT. Interestingly, two distinct clusters were observed in *env* C2V3 sequences (Fig. 1a). Of the 88 *env* C2V3 sequences analyzed, 39 clustered with an Ethiopian reference HIV-1 subtype C sequence (Eth2220) (subcluster C), and the remaining 49 sequences clustered separately (subcluster C′). Similar clustering was also observed in *gag* p17 sequences (Fig. 1b), where 43 sequences clustered as C and 35 as C′, whereas 8 sequences were not separable. Our findings strengthen previous studies[1,2] that demonstrated that the Ethiopian subtype C had a genetic subcluster designated C′ with a significant bootstrap value in the *gag* and *env* regions.

To clarify the significance of the two subclusters, we analyzed the linkage between subclustering of the *env* and *gag* regions. As shown in Table 2, the subclusters of these two regions matched each other significantly ($\chi^2 = 30.82$, df = 2, $p < 0.001$), as subclusters C and C′ had concordance rates of 31.8% and 36.4%, respectively.

In contrast to the C2V3 and p17 regions, sequences for PR (Fig. 1c) and RT (Fig. 1d) did not show obvious subclustering. Furthermore, subclustering was not associated with the patients' sociodemographic features such as age, sex, religion, residence, and ethnic group or clinical signs and symptoms of the disease. The C and C′ strains have been cocirculating in the population independent of geography, time of sample collection, and risk group.[2]

Drug resistance mutations in the 90 HIV-1 subtype C sequences were analyzed by examining amino acid mutations at PR and RT sites associated with ARV drug resistance.[6] As shown in Table 3, no subject displayed major resistance mutations in PR. Minor mutations that might contribute to protease inhibitor (PI) resistance were found at L10I (2.2%), L10V (2.2%), K20R (9.8%), M36I (97.8%), L63P (22.8%), A71T (2.2%), V77I (1.1%), and V82I (8.7%). Analysis of sites in the

TABLE 2. SUBCLUSTERS IN ENVELOPE AND GAG SEQUENCES DEMONSTRATE HIGH CONCORDANCE RATE

	gagp17			
	C	*C′*	*NS*	*Total*
*env*C2V3				
C	28 (31.8%)	3 (3.4%)	8 (9.1%)	39 (44.3%)
C′	15 (17.0%)	32 (36.4%)	2 (2.3%)	49 (55.7%)

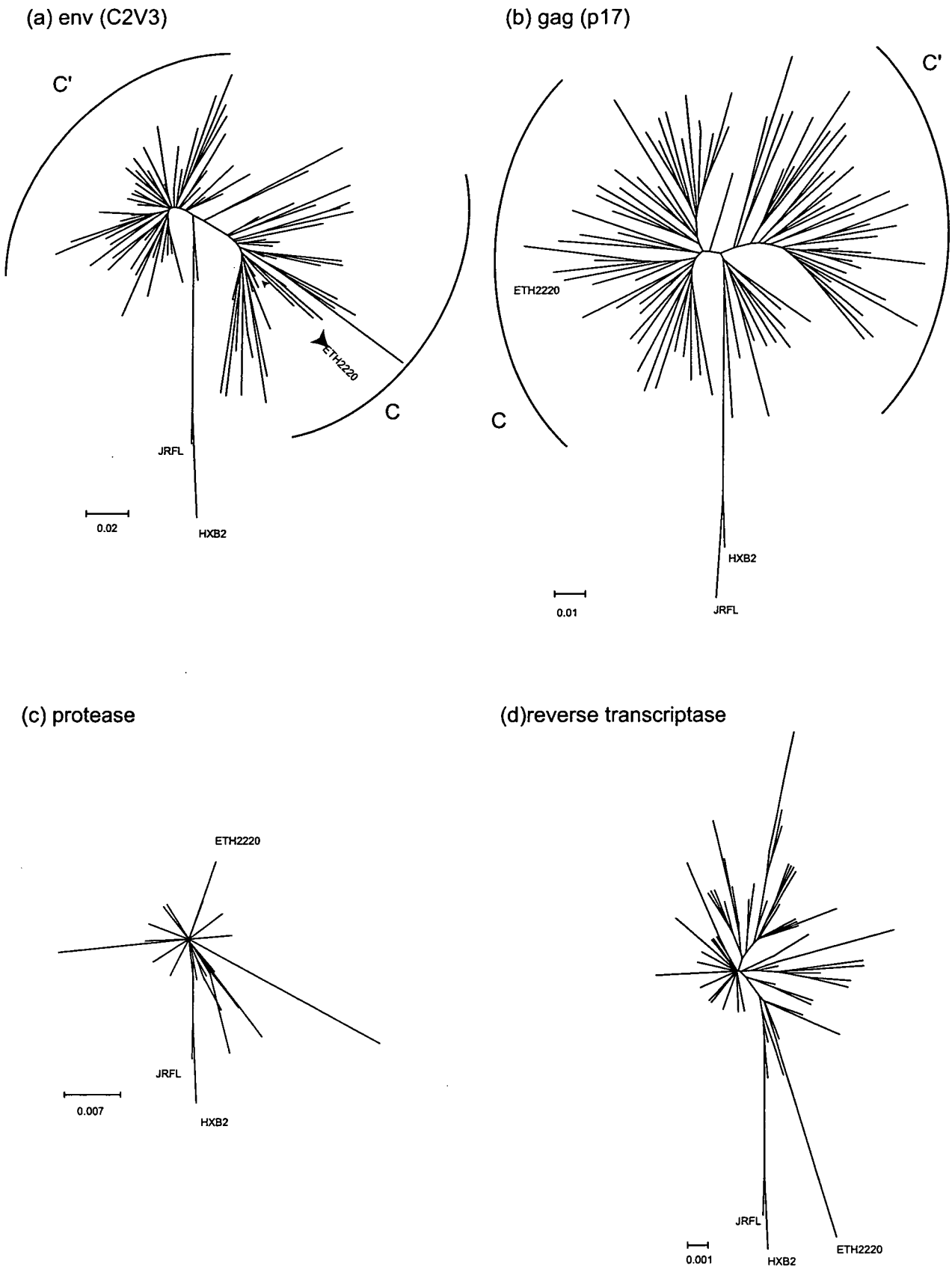

FIG. 1. Unrooted phylogenetic trees depicted by nucleotide sequences of 92 Ethiopian HIV-1-infected cases. (**a**) *env* C2V3, (**b**) *gag* p17, (**c**) protease, and (**d**) reverse transcriptase regions.

TABLE 3. PREVALENCE OF PROTEASE AND REVERSE TRANSCRIPTASE MUTATIONS IN HIV-1 SUBTYPE C ISOLATES[a]

	Ethiopia (present study) 2003											
Mutations	*All subtype C* (n = 90)	*C* (n = 39)	*C'* (n = 49)	p *value*	*Djibouti*[9] *2002* (n = 47)	*Malawi*[10] *1996–2001* (n = 21)	*Zambia*[11] *2000* (n = 25)	*UK*[12] *2000* (n = 25)	*Israel 1999–2000* (n = 20)	*Zimbabwe*[14] *1995* (n = 129)	*Worldwide 1986–1998* (n = 12)	*South Africa*[16] *2000* (n = 37)
Protease												
L10I/V	2, 2	3, 5	2, 0	1, 0.2	4, 6	10	0	0	0	0	8	ND[b]
I13V	7	8	6	1	0	0	4	0	0	0	0	ND
G16E	16	21	10	0.4	0	0	17	50	0	0	8	ND
K20R	12	21	6	0.06	6	29	4	36	25	8	1	ND
L33F	0	0	0	ND	0	0	0	0	7	0	0	ND
E35D	6	5	6	1	2	0	33	0	0	8	8	ND
M36I	97	100	94	1	43	91	84	100	95	100	100	ND
N37D/S	3, 1	5, 0	2, 2	0.6, 1	0	0	54	0	0	0	17	ND
M46I	0	0	0	ND	0	0	0	0	10	0	0	ND
D60E	8	8	8	1	0	0	4	8	0	33	0	ND
L63A/P/V	10, 32, 10	13, 33, 13	6, 33, 8	0.5, 1, 0.5	0, 2, 0	38	8, 25, 4	44	20	0, 8, 17	0, 17, 0	
I/L64V	1	3	0	0.4	0	0	4	0	0	0	0	ND
A71T	2	3	2	0.4	0	0	0	4	0	0	0	ND
T74S	11	13	10	0.7	0	0	13	24	0	0	0	ND
V77I	1	3	0	0.4	0	14	0	12	0	0	0	ND
V82I	9	10	8	1	0	0	0	0	0	0	8	ND
L89M	93	92	94	0.3	0	0	80	0	93	83	92	ND
I93L	98	97	98	0.4	0	100	96	100	80	100	ND	ND
Reverse												
E44	1	0	1	1	0	0	0	0	0	0	ND	0
S48T	45	21	65	<0.001	0	0	63	ND	0	100	ND	100
V60I	14	15	12	0.8	0	0	4	ND	0	33	ND	0
S68G	0	0	0	ND	0	0	4	ND	0	0	ND	0
A98G/S	0, 31	0, 49	0, 14	0, 0.002	0	0	0	ND	20	0	ND	0, 3
K102Q	3	0	6	0.3	0	0	4	ND	0	0	ND	0
I135T/V	19, 8	15, 8	22, 8	0.4, 1	0	0	8, 4	ND	0	8	ND	5, 0
K166R	20	13	27	0.2	0	0	29	ND	0	17	ND	100
I178M	8	10	6	0.7	0	0	8	ND	0	8	ND	0
G190A	2	2	0	0.2	0	0	0	ND	0	0	ND	0
I202V	17	8	24	0.05	0	0	4	ND	0	0	ND	0
R211K	63	64	61	0.8	0	76	71	ND	0	33	ND	0

[a]Numbers demonstrate prevalence of mutations (%).
[b]ND, no data.

RT gene associated with resistance to RT inhibitors showed mutations V75I and G190A in one case (1.1%) and two cases (2.2%), respectively. The V75I mutation was observed in a subtype D isolate while the G190A mutation was seen in subtype C viruses. The population in the area sampled appears not to have been highly exposed to antiretrovirals when our samples were collected, consistent with the history of ARV therapy in Ethiopia, which began in 2003 at selected health institutions, including the University of Gondar Hospital. However, single-dose nevirapine therapy to prevent mother-to-child transmission (MTCT) of HIV was implemented in the country before our study. One of the two G190A cases was female, but previous exposure to nevirapine by the MTCT program was not confirmed. For the other case (a male), transmission of G190A HIV from nevirapine-exposed patient was suspected. Thus, our observation that two patients (a male and a female) had G190A mutations indicates that nevirapine-resistant HIV may be circulating in the study area.

Although no significant clustering was observed in phylogenetic analyses of PR and RT regions, significantly higher prevalences of S48T ($p < 0.001$) and V202I ($p = 0.05$) were found in C′, whereas A98S was significantly more prevalent in C ($p = 0.002$). The mutation patterns of our sample are compared in Table 3 to those of HIV-1 isolates from ARV therapy-naive patients of Djibouti,[9] Malawi,[10] Zambia,[11] UK,[12], Zimbabwe,[14] worldwide,[15] and South Africa.[16] Among the reported subtype C cases, it is interesting to see diverse polymorphism. The mutation pattern of our study population appears to resemble patterns from Zambia[11] and Zimbabwe.[14] The low prevalence of drug resistance mutations in these studies might reflect the lack of widespread usage of ARV drugs in those countries and their surroundings.

In conclusion, our study demonstrates the predominance of HIV-1 subtype C and low prevalence of drug resistance cases in Gondar, Ethiopia. However, two cases of nevirapine-resistant cases found in the study population presage the eventual outspread of drug-resistant HIV in the area. Thus, continuous monitoring of treatment-naive HIV-1-infected populations should be considered.

ACKNOWLEDGMENTS

The authors thank Ms. Claire Baldwin for her help in preparing the manuscript. This study was supported by a Grant-in-Aid for AIDS research from the Ministry of Health, Labor, and Welfare of Japan (H15-AIDS-001) and in part by a grant from the National Institute of Biomedical Innovation (01–04).

REFERENCES

1. Abebe A, Kuiken CL, Goudsmit J, *et al.*: HIV type 1 subtype C in Addis Ababa, Ethiopia. AIDS Res Hum Retroviruses 1997;13: 1071–1075.
2. Abebe A, Pollakis G, Fontanet AL, *et al.*: Identification of a genetic subcluster of HIV type 1 subtype C (C′) widespread in Ethiopia. AIDS Res Hum Retroviruses 2000;16:1909–1914.
3. Hussein M, Abebe A, Pollakis G, *et al.*: HIV-1 subtype C in commerical sex workers in Addis Ababa, Ethiopia. J Acquir Immune Defic Syndr 2000;23:120–127.
4. Kassu A, Mohammad A, Fujimaki Y, *et al.*: Serum IgE levels of tuberculosis patients in a tropical setup with high prevalence of HIV and intestinal parasitoses. Clin Exp Immunol 2004;138:122–127.
5. Sugiura W, Matsuda M, Abumi H, *et al.*: Prevalence of drug resistance-related mutations among HIV-1s in Japan. Jpn J Infect Dis 1999;52:21–22.
6. Johnson VA, Brun-Vézinet F, Clotet B, *et al.*: Update of the drug resistance mutations in HIV-1. Topics HIV Med 2006;14:125–130.
7. Hierholzer M, Graham RR, El Khidir I, *et al.*: HIV type 1 strains from East and West Africa are intermixed in Sudan. AIDS Res Hum Retroviruses 2002;18:1163–1166.
8. Neilson JR, John GC, Carr JK, *et al.*: Subtypes of human immunodeficiency virus type 1 and disease stage among women in Nairobi, Kenya. J Virol 1999;73:4393–4403.
9. Maslin J, Rogier C, Caron M, *et al.*: Antiretroviral drug resistance among drug-naive HIV-1-infected individuals in Djibouti (Horn of Africa). Antiviral Ther 2005;10:855–859.
10. Petch LA, Hoffman IF, Jere CS, *et al.*: Genotypic analysis of the protease and reverse transcriptase of HIV type 1 subtype C isolates from antiretroviral drug-naive adults in Malawi. AIDS Res Hum Retroviruses 2005;21:799–805.
11. Handema R, Terunuma H, Kasolo F, *et al.*: Prevalence of drug-resistance-associated mutations in antiretroviral drug-naive Zambians infected with subtype C HIV-1. AIDS Res Hum Retroviruses 2003;19:151–160.
12. Cane PA, de Ruiter A, Rice P, *et al.*: Resistance-associated mutations in the human immunodeficiency virus type 1 subtype C protease gene from treated and untreated patients in the United Kingdom. J Clin Microbiol 2001;39:2652–2654.
13. Grossman Z, Vardinon N, Chemtob D, *et al.*: Genotypic variation of HIV-1 reverse transcriptase and protease: Comparative analysis of clade C and clade B. AIDS 2001;15:1453–1460.
14. Shafer RW, Chuang TK, Hsu P, White CB, and Katzenstein DA: Sequence and drug susceptibility of subtype C protease from human immunodeficiency virus type 1 seroconverters in Zimbabwe. AIDS Res Hum Retroviruses 1999;15:65–69.
15. Pillay C, Bredell H, McIntyre J, Gray G, and Morris L: HIV-1 subtype C reverse transcriptase sequences from drug-naive pregnant women in South Africa. AIDS Res Hum Retroviruses 2002;18: 605–610.
16. Pieniazek D, Rayfield M, Hu DJ, *et al.*: Protease sequences from HIV-1 group M subtypes A-H reveal distinct amino acid mutation patterns associated with protease resistance in protease inhibitor-naive individuals worldwide. HIV Variant Working Group. AIDS 2000;14:1489–1495.

Address reprint requests to:
Wataru Sugiura
AIDS Research Center
National Institute of Infectious Diseases
Tokyo 2080011, Japan

E-mail: wsugiura@nih.go.jp

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution xxx (2007) xxx–xxx
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Bayesian network analysis of resistance pathways against HIV-1 protease inhibitors

K. Deforche [a,*], R. Camacho [b], Z. Grossman [c], T. Silander [d], M.A. Soares [e], Y. Moreau [f], R.W. Shafer [g], K. Van Laethem [a,h], A.P. Carvalho [b], B. Wynhoven [i], P. Cane [j], J. Snoeck [a], J. Clarke [k], S. Sirivichayakul [l], K. Ariyoshi [m], A. Holguin [n], H. Rudich [c], R. Rodrigues [o], M.B. Bouzas [p], P. Cahn [p], L.F. Brigido [o], V. Soriano [n], W. Sugiura [m], P. Phanuphak [l], L. Morris [q], J. Weber [k], D. Pillay [r], A. Tanuri [e], P.R. Harrigan [i], J.M. Shapiro [s], D.A. Katzenstein [g], R. Kantor [g], A.-M. Vandamme [a]

[a] *Rega Institute for Medical Research, Katholieke Universiteit Leuven, Leuven, Belgium*
[b] *Virology Laboratory, Hospital Egas Moniz, Lisbon, Portugal*
[c] *Chaim Sheba Medical Center, Ministry of Health, Tel-Aviv, Israel*
[d] *Helsinki Institute for Information Technology, Helsinki, Finland*
[e] *Departamento de Genética, Universidade Federal do Rio de Janeiro, Brazil*
[f] *ESAT, Katholieke Universiteit Leuven, Leuven, Belgium*
[g] *Division of Infectious Diseases, Brown University, Providence RI, USA*
[h] *AIDS Reference Laboratory, University Hospitals, Leuven, Belgium*
[i] *BC Centre for Excellence in HIV/AIDS, Vancouver, Canada*
[j] *Reference and Microbiology Division, Health Protection Agency, Porton Down, UK*
[k] *Department of GUM & Communicable Diseases, Wright Fleming Institute, London, UK*
[l] *Department of Medicine, Chulalongkorn University, Bangkok, Thailand*
[m] *Department of Pathology, National Institute of Infectious Diseases, Tokyo, Japan*
[n] *Department of Infectious Diseases, Hospital Carlos III, Madrid, Spain*
[o] *Laboratorio de Retrovirologia, Instituto Adolfo Lutz, São Paulo, Brazil*
[p] *Fundación Huesped, Buenos Aires, Argentina*
[q] *AIDS Unit, National Institute for Communicable Diseases, Johannesburg, South Africa*
[r] *Antiviral Susceptibility Reference Unit, Health Protection Agency, Birmingham, UK*
[s] *National Hemophilia Center, Sheba Medical Center, Tel Aviv, Israel*

Received 17 February 2006; received in revised form 8 September 2006; accepted 11 September 2006

Abstract

Interpretation of Human Immunodeficiency Virus 1 (HIV-1) genotypic drug resistance is still a major challenge in the follow-up of antiviral therapy in infected patients. Because of the high degree of HIV-1 natural variation, complex interactions and stochastic behaviour of evolution, the role of resistance mutations is in many cases not well understood. Using Bayesian network learning of HIV-1 sequence data from diverse subtypes (A, B, C, F and G), we could determine the specific role of many resistance mutations against the protease inhibitors (PIs) nelfinavir (NFV), indinavir (IDV), and saquinavir (SQV). Such networks visualize relationships between treatment, selection of resistance mutations and presence of polymorphisms in a graphical way. The analysis identified 30N, 88S, and 90M for nelfinavir, 90M for saquinavir, and 82A/T and 46I/L for indinavir as most probable major resistance mutations. Moreover we found striking similarities for the role of many mutations against all of these drugs. For example, for all three inhibitors, we found that the novel mutation 89I was minor and associated with mutations at positions 90 and 71. Bayesian network learning provides an autonomous method to gain insight in the role of resistance mutations and the influence of HIV-1 natural variation.

* Corresponding author. Tel.: +32 16 332160; fax: +32 16 332131.
E-mail address: koen.deforche@gmail.com (K. Deforche).

1567-1348/$ – see front matter © 2007 Published by Elsevier B.V.
doi:10.1016/j.meegid.2006.09.004

Please cite this article in press as: Deforche, K. et al., Bayesian network analysis of resistance pathways against HIV-1 protease inhibitors, Infect. Genet. Evol. (2007), doi:10.1016/j.meegid.2006.09.004

We successfully applied the method to three protease inhibitors. The analysis shows differences with current knowledge especially concerning resistance development in several non-B subtypes.

Keywords: HIV; Protease; Nelfinavir; Indinavir; Saquinavir

1. Introduction

Human Immunodeficiency Virus (HIV) escapes the inhibitory effect of antiretroviral drugs by selection of mutations that increase resistance against those drugs. To obtain an effective therapy, it is thus necessary to use antiretroviral drugs for which the virus remains susceptible. Genotypic interpretation systems predict the susceptibility or therapy response for various drugs (Shafer, 2002; Van Laethem et al., 2002), based on the presence of mutations at positions associated with drug resistance. Unfortunately, the role of many resistance mutations remains unsufficiently known, as well as the role of HIV-1 natural variation. This variation within the HIV main group is reflected in a subtype system with 9 identified subtypes and 16 Circulating Recombinant Forms (CRFs). In addition, unclassified strains and new recombinants are increasingly reported. Different prevalences of known resistance-associated mutations and new mutations are seen in different subtypes (Frater et al., 2001; Grossman et al., 2001; Brindeiro et al., 2002; Ariyoshi et al., 2003; Parkin and Schapiro, 2004). With a few exceptions, these differences in prevalence could not be explained by different genetic barriers because of different codon usage (Turner et al., 2004). In previous work, we used Bayesian network (BN) learning to demonstrate how polymorphisms may influence how drug-associated mutations get selected. These explained some notable subtype differences that have been observed for resistance development against nelfinavir (Deforche et al., 2006).

In this work we present the application of Bayesian network learning to study development of resistance against three protease inhibitors (PIs): nelfinavir (NFV), indinavir (IDV), and saquinavir (SQV). Results were compared in the context of cross-resistance within the class of PIs.

A Bayesian network (BN) is a probabilistic model that describes statistical independencies between multiple variables. In this work, we learn Bayesian networks from observations of the variables. In this way, the best Bayesian network is searched that explains a maximum of the observed correlations in the data using a minimum number of *direct influences*. Dependencies are visualized in a directed acyclic graph and form the qualitative component of the BN. In this graph, each node corresponds to a variable, and a directed arc (arrow) between nodes represents a direct influence. Mathematically, a Bayesian network provides a refactoring of the Joint Probability Distribution (JPD) of the data, using Bayes' rule. As a BN simplifies the JPD, it provides an effective model that summarizes statistical properties of the data.

Within the study of drug resistance, one often refers to a mutation that is selected as a first mutation as a *major mutation* (Shafer, 2002; Johnson et al., 2004). Similarly, a *minor mutation* further increases resistance only in presence of other mutations, or compensates for a possible fitness impact of other mutations, and is therefore selected only in presence of these other mutations. Although these concepts are not rigorously defined, conditional independencies in the networks allow us to identify major and minor mutations, in agreement with these definitions.

2. Materials and methods

Data was derived from five clinical databases: Portugal, Belgium, Israel, Brazil and an international database containing sequences from subtypes other than subtype B. In total we had access to 4911 sequences. Protease (PRO) and partial reverse transcriptase (RT) HIV-1 sequences from protease inhibitor (PI) naive patients and from patients treated with only experience to NFV, IDV, or SQV as only PI, either unboosted or boosted with ritonavir, were trimmed to the first 350 amino-acids. At most one treated sequence and one naive sequence per patient were included and identical sequences were removed. RT inhibitor experienced patients were included in the PI naive patient population, since no resistance to RT inhibitors is expected in the protease gene.

The analysis followed closely the method described in Deforche et al. (2006). Subtyping was done using a phylogenetic analysis (de Oliveira et al., 2005). We identified wild type polymorphisms based on a prevalence greater than 10% in untreated patients and determined treatment associated mutations by testing for independence from treatment using a Cochran–Mantel–Haenszel χ^2 test, stratifying in each combination of subtype and database. The statistical analysis was corrected for multiple comparisons using Benjamini & Hochberg with a False Discovery Rate of 0.05. The data sets for Bayesian network were also stratified for an equal ratio of treated and untreated sequences within each combination of subtype and database, and included next to treatment experience, Boolean variables indicating presence of each treatment associated mutation and presence of polymorphic amino acids. Bayesian network learning was done by searching using a simulated annealing heuristic for the most probable network structure using a Bayesian scoring metric. A non-parametric bootstrap was performed by resampling from the sequences, to assess the robustness of network features.

In the final networks, we do not show the obvious strong antagonistic direct influences between different amino acids at single residue. Only network features (presence or absence of arcs) with a bootstrap higher than 65% were considered robust, and only robust arcs are shown. To reduce the

Please cite this article in press as: Deforche, K. et al., Bayesian network analysis of resistance pathways against HIV-1 protease inhibitors, Infect. Genet. Evol. (2007), doi:10.1016/j.meegid.2006.09.004

complexity of the graphs, polymorphic positions that did not directly influence any treatment-associated mutations were omitted. Arcs were colored according to their function to improve reading the graph, but this coloring is only indicative. For each drug, known resistance mutations are those that are defined for that drug in either the International AIDS Society list of resistance mutations of 2005 (Johnson et al., 2005) or that are included in the resistance score in at least one of the latest versions of public resistance interpretation systems ANRS 2004.09, REGA 6.2 or HIVDB 2004.12 (Kantor et al., 2001).

To interpret the Bayesian network in the context of antiretroviral resistance, we considered the meaning of an arc between two mutations that was derived in Deforche et al. (2006). A major mutation is unconditionally dependent on

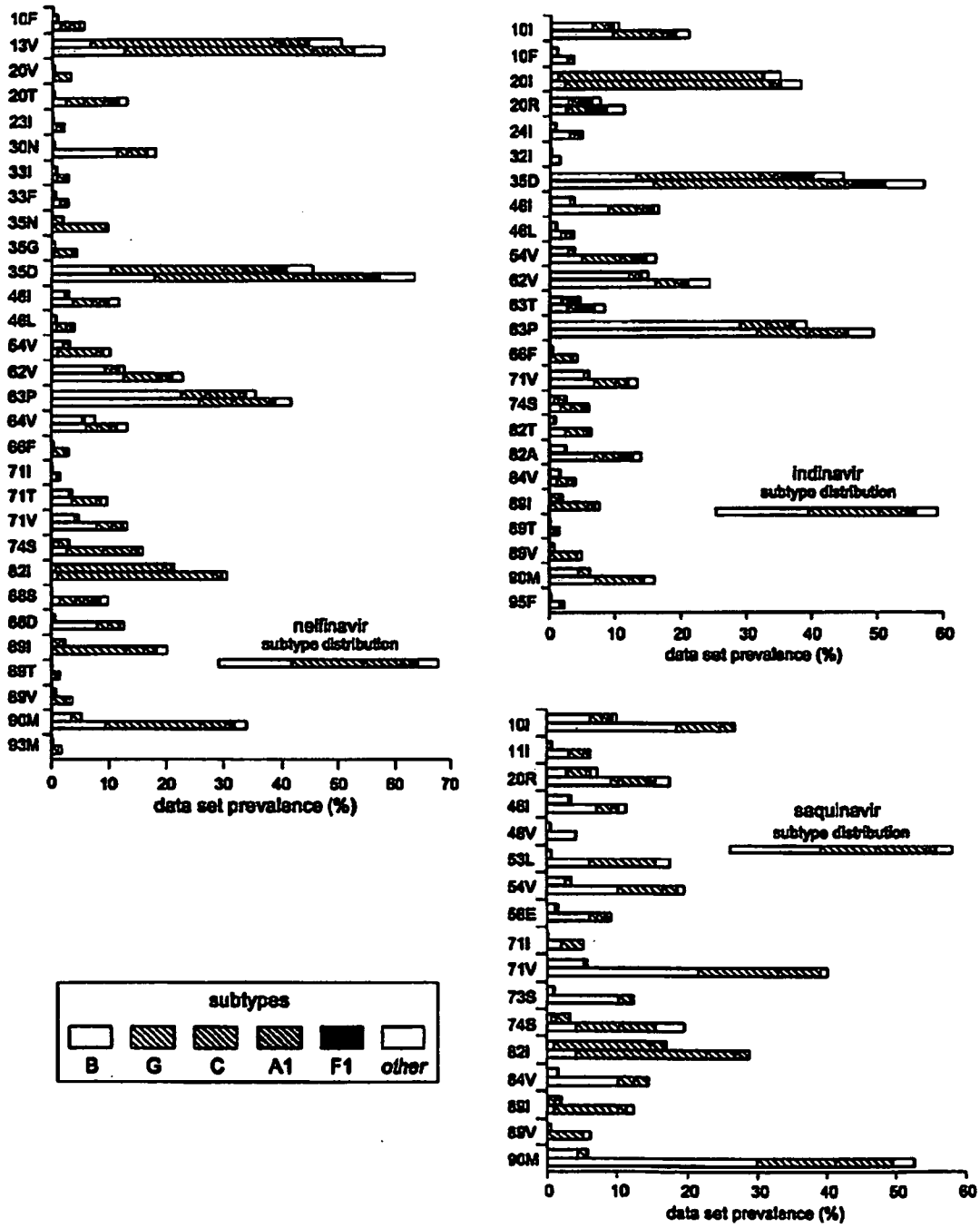

Fig. 1. Dataset prevalence (%) of NFV, IDV, and SQV treatment-associated mutations in sequences from untreated (top bar) and treated (bottom bar) patients. For each drug, the data was stratified for the overall subtype distribution of the sequences to be identical for treated and untreated patients.

Please cite this article in press as: Deforche, K. et al., Bayesian network analysis of resistance pathways against HIV-1 protease inhibitors, Infect. Genet. Evol. (2007), doi:10.1016/j.meegid.2006.09.004

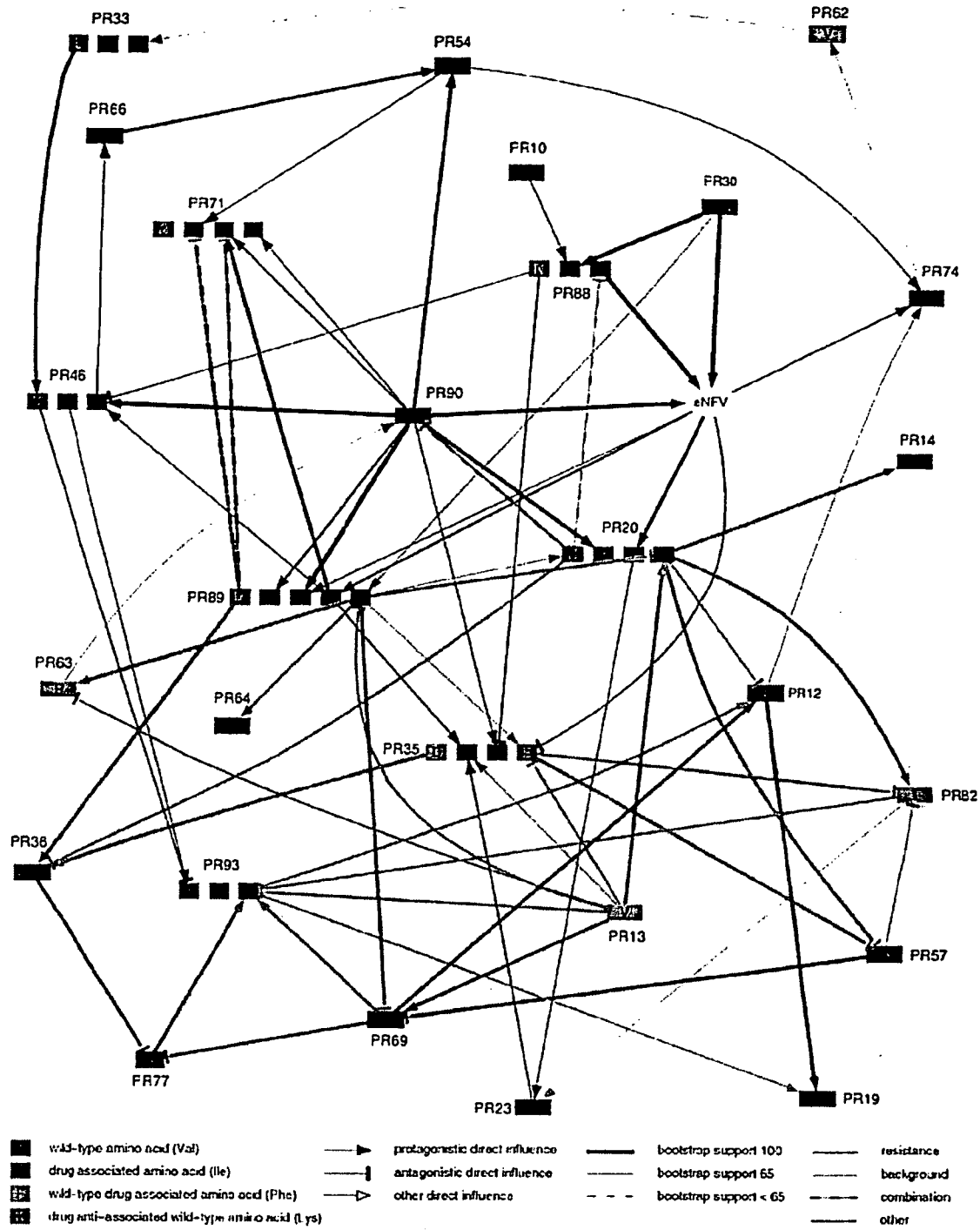

Fig. 2. Annotated NFV experience Bayesian network showing direct influences between NFV-associated mutations, polymorphisms and NFV treatment (eNFV). An arc represents a direct dependency between the corresponding variables and thickness is proportional to bootstrap support. Arc color indicates whether it is a direct influence between NFV-associated mutations (black), an influence from background polymorphisms on NFV-associated mutations (blue), or a combination of these (blue–black dashed) or merely an association between background polymorphisms (green). An antagonistic arc with a wild type was treated the same as a synergistic arc with mutations at this position. Arc direction has no causal meaning, but may indicate a non-additive multivariate effect. Unconditional dependencies with treatment with bootstrap support between 35% and 65% are shown dashed.

Please cite this article in press as: Deforche, K. et al., Bayesian network analysis of resistance pathways against HIV-1 protease inhibitors, Infect. Genet. Evol. (2007), doi:10.1016/j.meegid.2006.09.004

treatment, which is indicated in the networks as the robust presence of an arc between the mutation and treatment. In this respect, an antagonistic arc with a wild type was treated the same as a synergistic arc with mutations at this position. Similarly, a minor mutation is expected to be conditionally independent from treatment but dependent on other resistance mutations, and thus indicated by the robust absence of an arc between the mutation and treatment. As was discussed in Deforche et al. (2006), a minor mutation may still be connected to treatment, when the cost is lower to connect to treatment instead of all the resistance mutations it is associated with. Where appropriate, we used the multivariate effect implied by arc directions as well to narrow down the list of major mutations.

3. Results

3.1. Subtyping

The subtype could be determined for 85% of the sequences. The overall subtype distribution of the sequences was subtype G (31%), B (27%), C (24%), A1 (12%), D (3%), F1 (3%) and other subtypes (<1%). The subtype distribution was different for the untreated and each of the NFV, IDV and SQV treated populations. As a result the subtype distributions for each analysis were slightly different (Fig. 1).

3.2. Treatment-associated mutations

The data used to determine mutations associated with treatment with each of the drugs, included 479 (NFV), 539 (IDV), and 97 (SQV) sequences from patients with experience with that drug as sole PI.

Fig. 1 shows the prevalence of treatment associated mutations in naive and treated patients, that were identified for each of the three drugs, using a χ^2 statistical analysis.

The most notable discordances with known resistance mutations were the novel mutations 20V, 33I/F, 35D/G/N, 62V, 64V, 66F, 74S, 89I/T/V and 93M for NFV; 35D, 62V, 63T, 66F, 74S, 89I/T/V and 95F for IDV; and 11I, 58E, 74S, 82I, and 89I/V for SQV. These mutations, except for 20V, 35N, and 89T have been previously described to be associated with PI experience in different studies (Wu et al., 2003; Svicher et al., 2005), but not with specific inhibitors. Some of these novel mutations were associated with treatment by all three drugs (74S and 89I/V) or by two drugs (35D, 58E, 66F, 89T, and 95F). The selection of some of these mutations was more pronounced than selection of mutations that have been widely accepted as resistance mutations.

At the same time we did not find selection of mutations 82A/F/T/S or 84V by NFV, even though they are considered important for NFV resistance by all algorithms.

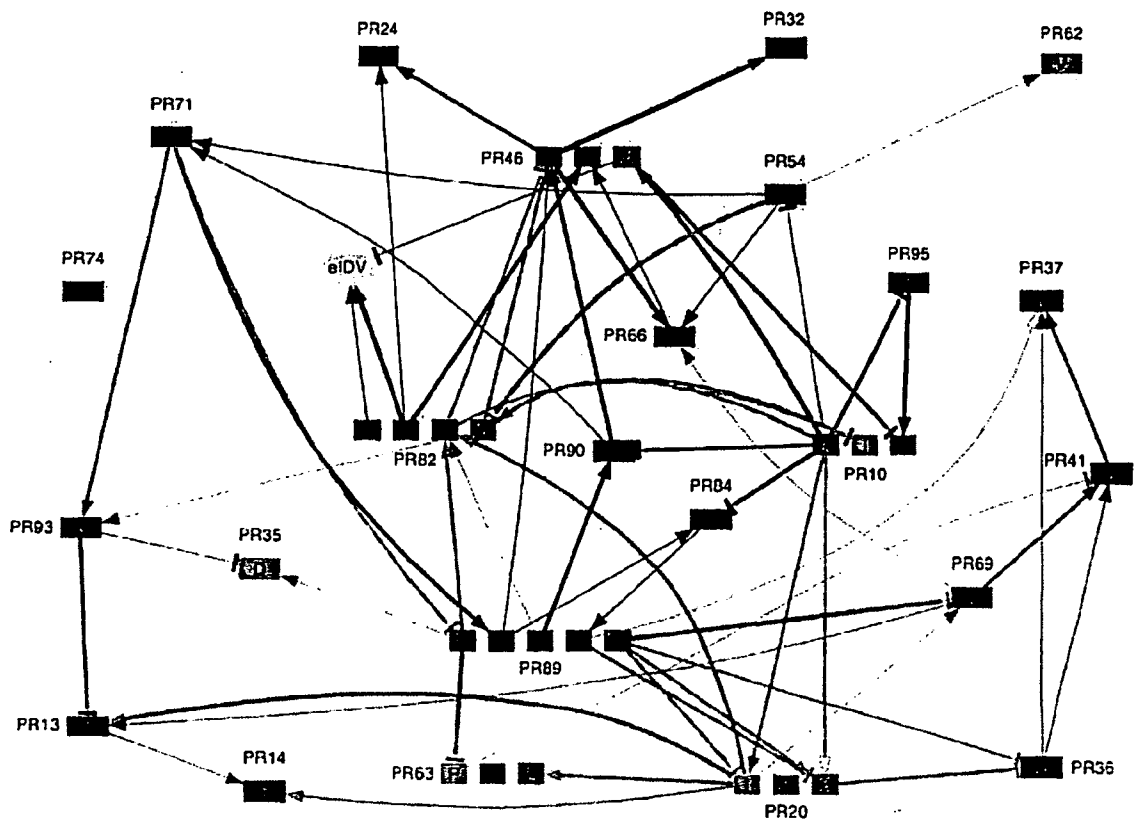

Fig. 3. Annotated IDV experience Bayesian network showing direct influences between IDV-associated mutations, polymorphisms and IDV treatment (eIDV). Legend as in Fig. 2.

Please cite this article in press as: Deforche, K. et al., Bayesian network analysis of resistance pathways against HIV-1 protease inhibitors, Infect. Genet. Evol. (2007), doi:10.1016/j.meegid.2006.09.004

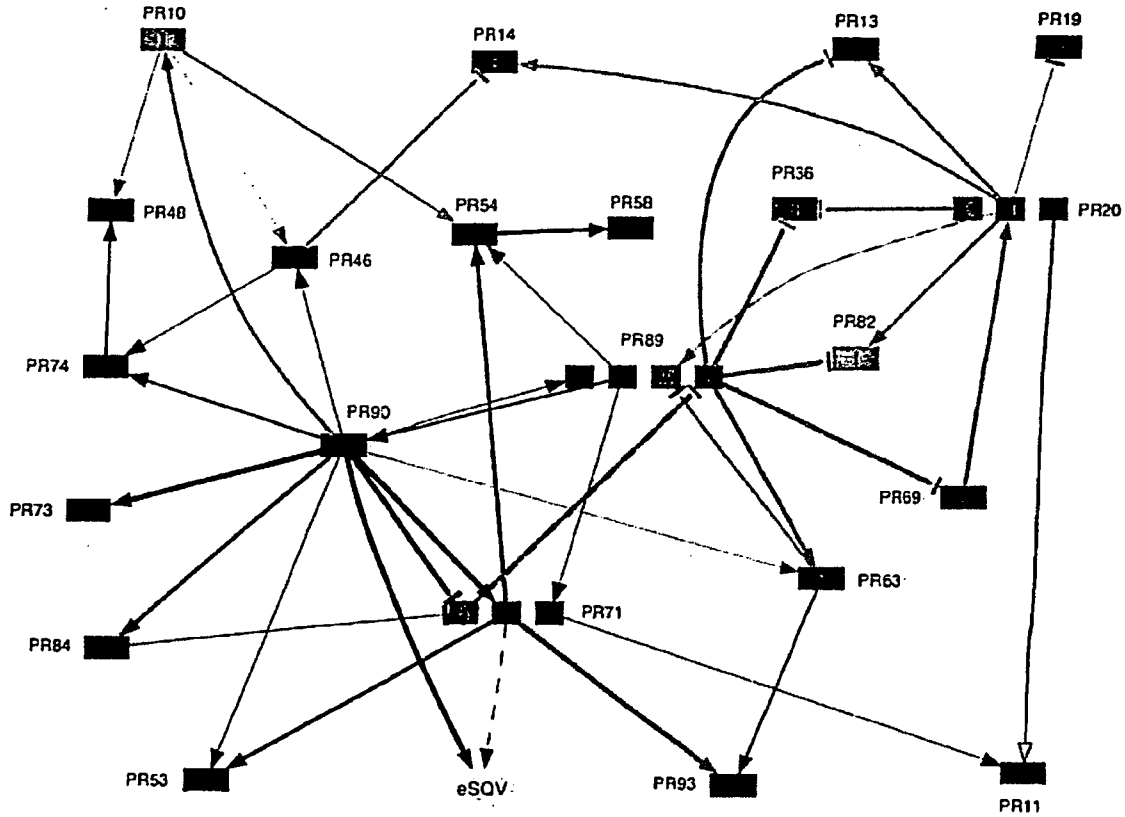

Fig. 4. Annotated SQV experience Bayesian network showing direct influences between SQV-associated mutations, polymorphisms and SQV treatment (eSQV). Legend as in Fig. 2.

3.3. Bayesian network learning

The data sets for Bayesian network learning included 340 (NFV), 288 (IDV), and 31 (SQV) sequences from patients on treatment, and respectively 967, 925, and 716 sequences from PI naive patients. Because of the stratification, the overall ratios of treated to untreated sequences were the same within every combination of data set and subtype.

The Bayesian network learning discovered many robust interactions between the variables in our data sets: the fraction of arcs with bootstrap support over 65% increased with available data and ranged from 44% (for the SQV network) to 68% (for the IDV network). The networks are shown in Figs. 2–4. For resistance against NFV, the network indicated 30N, 88S and 90M as major mutations, since they show a robust unconditional dependence on treatment, and analysis of the non-additive multivariable effect implied by arc directions at the treatment node, indicated that these three mutations occur mostly independently. The amount of selection of one of these major mutations is different for different subtypes (see Fig. 1). In subtype B, 30N is selected most. In subtypes C and G, 90M is selected most. Finally, in subtypes A1 and F1, 88S is selected most. For resistance against IDV, the network indicated 82A/T and 46I/L as major mutations. Mutation 84V is also selected by IDV in our data set, but we find it to be selected only after accumulation of mutations 82A/T or 46I/L, and 10I/F. The results for SQV were less conclusive, as more network features (presence or absence of arcs) were not robust because of the low amount of data. The network indicated 90M as major mutations but could not exclude 71V as additional major mutation (which was also unconditionally dependent on treatment in the most probable network, but with bootstrap support 47%), and 46I, 48V, 53L, 58E, 73S, 74S, or 89V as alternative major mutations, ordered by likelihood. These mutations were indicated by the most probable network as conditionally independent from treatment but this independence was not robust. Mutation 48V only occurred in subtype B, and the most probable network indicated that it appeared only after accumulation of mutations 90M and 10I or 74S in our data set.

For most minor mutations, that are conditionally independent from treatment, the networks suggest their role in more detail by indicating robust interactions with other resistance mutations in whose presence they are selected, and thus contribute to a selective advantage of the virus. The network for NFV shows that minor mutations 20T/V, 35N, 46I/L, 54V, 71I/T/V, and 89I/T/V, and the polymorphisms 63P and 89L directly influence a major mutation, while minor mutations 10F, 23I, 33I/F, 66F, or 93M are further away in the resistance pathway. Similarly, for IDV, mutations 10I/F, 24I, 32I, 54V, 66F, or 90M directly influence a major mutation, and mutations 20R, 71V, 74S, 84V, 89I/T/V, and 95F are further away in the resistance pathway.

Please cite this article in press as: Deforche, K. et al., Bayesian network analysis of resistance pathways against HIV-1 protease inhibitors, Infect. Genet. Evol. (2007), doi:10.1016/j.meegid.2006.09.004

The networks indicate robust interactions between polymorphisms and resistance mutations, which may explain subtype differences. For example, the NFV network indicated a protagonistic interaction of the 89L polymorphism on development of the 30N mutation, explaining the higher prevalence of 30N in subtype B (Grossman et al., 2004; Abecasis et al., 2005). This effect was recently confirmed in *in vitro* experiments (Calazans et al., 2005). Similarly, both the IDV and SQV networks suggested a protagonistic interaction of the 93L polymorphism on development of the 71V mutation.

There are striking similarities when comparing the networks for different drugs, especially when considering arcs with bootstrap below 65% as well (which are not shown). All three networks indicated interactions between resistance mutations 90M and 89I, 90M and 71V, and 90M and 54V, and between the polymorphism 89L/M on mutations at position 71. In addition, interactions between polymorphisms and resistance mutations found in two networks were interactions of polymorphism T12S on mutation 74S, L63P on mutation 90M, I93L on mutation 90M, I93L on mutation 71V, and L10I on mutations 54V and 90M.

4. Discussion

Based on higher prevalence in sequences from treated versus untreated patients, we confirmed the selection of many known mutations by antiviral drugs, but also identified selection of novel mutations. As can be seen from Fig. 1, selection of these novel mutations was often more pronounced or sometimes exclusively in non-B subtypes.

The low level of selection of mutations at position 82 and 84 during NFV treatment is confirmed in other data sets (Kantor et al., 2001). In Shafer (2002) it is argued that mutations at position 82 have no phenotypic effect on their own for resistance against NFV, but contribute to resistance together with other mutations. A possible explanation for this discrepancy may be that selection of these mutations depends on the presence of other mutations that are not commonly present in untreated patients, or that are not selected by NFV, but are more common in patients exposed to other PI treatment.

The learned Bayesian networks indicated major mutations, largely in agreement with current knowledge (Johnson et al., 2004), with some exceptions. For NFV we found that 88S has a different role as 88D, and should be considered a major mutation, and may be more important than 30N or 90M in subtypes A1 and F1. Published phenotypic data supports this finding by indicating a phenotypic fold change in EC_{50} of 8.9 for 88S alone (Kantor et al., 2001). The IDV network indicated that 84V is not a major mutation, while it is widely considered so. As it is documented that it rarely develops as a first mutation, but only appears in isolates that already have a 90M (Shafer, 2002), this discrepancy is explained by a discordant definition of a major mutation. Similarly, according to our semantics, the SQV network indicates that 48V is not a major resistance mutation for SQV, since it virtually never occurred without mutation 90M. This is not due to its low prevalence in SQV failing sequences, which was comparable to the prevalence of the major IDV mutation 82T. However, the SQV dataset was rather small to make final conclusions.

The power of Bayesian network learning to find robust (in)dependencies in the data, depends on the sample size, the number of variables, and the actual number of independencies in the data set. It has previously been observed that resistance against IDV is less structured than resistance development against SQV (Beerenwinkel et al., 2004), which may explain why a similar amount of robust dependencies were observed for resistance against IDV as for SQV, despite the fact that the IDV data set was several times larger.

The biological role of minor mutations is to further increase resistance, and/or to compensate for a loss in replication capacity caused by the major mutation. Minor mutations that only improve replication capacity that was compromised by a resistance mutation in the virus should develop in the context of the same resistance mutations regardless of the inhibitor used. Indeed, these mutations may even develop in absence of the inhibitor, to improve replication capacity compromised by other resistance mutations (van Maarseveen et al., 2006). The similarities observed in the networks for different drugs, could thus indicate that mutations 10I, 12S, 54V, 63P, 71V, 89I, and 93L improve replication capacity compromised by other mutations, although their role in increasing resistance cannot be excluded.

The method of identifying possible resistance mutations by considering mutations associated with treatment in a cross-sectional data set can be confounded by drift. Drift may be the reason for a higher prevalence of a mutation in the treated population, and this is more likely for polymorphisms that occur frequently in the untreated population. Even after stratifying in combinations of database and subtype, we cannot exclude this effect of drift. At the same time, the Bayesian networks could not be used in most cases to reliably determine the role of these polymorphic resistance mutations, since they mostly ignored the relative low amount of variation associated with treatment while explaining the larger amount of variation at these polymorphisms in association with other polymorphisms. Mutations 10I and 63P however show similar linkage in the networks for different PIs, indicating their role. The interactions we found between 63P and the major mutation 90M are in agreement with earlier reports on the role of L63P (Martinez-Picado et al., 1999; Sune et al., 2004). The Bayesian networks could not clarify in a consistent way the role of other polymorphisms that we found to be associated with treatment (13V, 20I, 35D, 36I, 62V, 82I), despite the possibly important clinical implications.

For the analysis of SQV and IDV, data from both boosted and unboosted regimens were combined. The effect of boosting is suggested to primarily increase the genetic barrier to develop a clinically relevant level of resistance, by increasing the intracellular concentration of the drug. Whether this changes the patterns of drug resistance mutations has not been investigated yet. If using boosted regimen changes resistance pathways, then this would have blurred the analysis only yielding lower bootstrap confidence per pathway.

Please cite this article in press as: Deforche, K. et al., Bayesian network analysis of resistance pathways against HIV-1 protease inhibitors, Infect. Genet. Evol. (2007), doi:10.1016/j.meegid.2006.09.004

A difficult question in resistance interpretation is how to score presence at baseline of minor mutations. These generally do not have an effect on resistance on their own, and may even represent a fitness penalty with respect to the wild type. Thus, without considering further evolution, the virus remains fully susceptible to the drug. However, some of these mutations improve the fitness or resistance impact of the corresponding major mutation. Therefore, the presence of these mutations will speed up the selection and increase the impact of these major mutations. Other minor mutations do not directly influence the major mutation, and thus do not have the same clinical significance when present at baseline. Therefore, we predict that for NFV, in absence of major mutations 30N, 88S, or 90M, presence of mutations 20T/V, 35N, 46I/L, 54V, 71I/T/V, or 89I/T/V, or of polymorphisms 63P or 89L, should impact clinical outcome to a greater extent than mutations 10F, 23I, 33I/F, 66F, or 93M. Similarly, for IDV, in absence of major mutations 82A/T and 46I/V, presence of mutations 10I/F, 24I, 32I, 54V, 66F, or 90M should have a higher impact on clinical outcome than mutations 20R, 71V, 74S, 84V, 89I/T/V, and 95F.

The power of Bayesian network learning lies in its sound mathematical foundation to distill likely direct interactions (which in many cases could be causalities) from the many observed associations between different residues. Bayesian network learning has previously successfully been applied to amino acid sequence data, in the context of secondary structure prediction (Klingler and Brutlag, 1994), but also in the field of HIV Drug Resistance (Beerenwinkel et al., 2004). In the latter analysis, Bayesian network models were constrained to trees with a special structure of the Conditional Probability Distributions (CPDs). In this way, the models described ordered accumulation of mutations. In contrast, we use unconstrained Bayesian networks, and added information on background polymorphisms to the analysis. As a consequence, both antagonistic and synergistic interactions between treatment associated mutations and polymorphisms were learned, without the prior assumption of a strict ordered accumulation of mutations.

5. Conclusions

We applied Bayesian network learning to HIV-1 protease sequence data and exposure to protease inhibitors to learn many aspects of resistance development against three protease inhibitors. We used the structure of the network to infer hypotheses about the role of resistance mutations. Our analysis confirmed current knowledge, especially for resistance development in subtype B viruses. Our analysis shows an important impact of polymorphisms on resistance development that could explain subtype differences in resistance development. Our results may suggest new *in vitro* experiments, to confirm the hypothesised role of novel resistance mutations, or be used to update genotypic resistance interpretation systems.

Acknowledgments

The authors wish to thank Ana Abecasis and Esmeralda AJM Soares for data collection. Koen Deforche was funded by a Ph.D. grant of the Institute for the Promotion of Innovation through Sciences and Technology in Flanders (IWT). Yves Moreau is a post-doctoral researcher with the FWO-Vlaanderen; his research is supported by KULeuven GOA-Mefisto-666 and GOA-Ambiorics, Belspo IUAP V-22, and EU FP6 NoE Biopattern. This work was supported by the AIDS Reference Laboratory of Leuven that receives support from the Belgian Ministry of Social Affairs through a fund within the Health Insurance System, and funded through FWO-Vlaanderen grant G.0266.04, and by the Katholieke Universiteit Leuven through grant OT/04/43.

References

Abecasis, A.B., Deforche, K., Snoeck, J., Bacheler, L.T., Kenna, P.M., Carvalho, A.P., Gomes, P., Camacho, R.J., Vandamme, A.-M., 2005. Protease mutation M89I/V is linked to therapy failure in patients infected with the HIV-1 non-B subtypes C, F or G. AIDS 19 (16), 1799–1806.

Ariyoshi, K., Matsuda, M., Miura, H., Tateishi, S., Yamada, K., Sugiura, W., 2003. Patterns of point mutations associated with antiretroviral drug treatment failure in CRF01_AE (subtype E) infection differ from subtype B infection. J. Acquir. Immune Defic. Syndr. 33 (3), 336–342.

Beerenwinkel, N., Rahnenführer, J., Daumer, M., Hoffmann, D., Kaiser, R., Selbig, J., Lengauer, T., 2004. Learning multiple evolutionary pathways from cross-sectional data. In: RECOMB. ACM Press, pp. 36–44.

Brindeiro, P.A., Brindeiro, R.M., Mortensen, C., Heringa, K., De Vroey, V., Rubini, N.P.M., Sion, F.S., De Sa, C.A.M., Machado, D.M., Succi, R.C.M., Tanuri, A., 2002. Testing genotypic and phenotypic resistance in human immunodeficiency virus type 1 isolates of clade B and other clades from children failing antiretroviral therapy. J. Clin. Microbiol. 40 (12), 4512–4519.

Calazans, A., Brindeiro, R., Brindeiro, P., Verli, H., Arruda, M.B., Gonzalez, L.M.F., Guimaraes, J.A., Diaz, R.S., Antunes, O.A.C., Tanuri, A., 2005. Low accumulation of L90M in protease from subtype F HIV-1 with resistance to protease inhibitors is caused by the L89M polymorphism. J. Infect. Dis. 191 (11), 1961–1970.

de Oliveira, T., Deforche, K., Cassol, S., Salminen, M., Paraskevis, D., Seebregts, C., Snoeck, J., van Rensburg, E.J., Wensing, A.M.J., van de Vijver, D.A., Boucher, C.A., Camacho, R., Vandamme, A.-M., 2005. An automated genotyping system for analysis of HIV-1 and other microbial sequences. Bioinformatics 21 (19), 3797–3800.

Deforche, K., Silander, T., Camacho, R., Grossman, Z., Soares, M.A., Van Laethem, K., Kantor, R., Moreau, Y., Vandamme, A.-M., 2006. On behalf of the Non-B Workgroup. Analysis of HIV-1 pol sequences using Bayesian Networks: implications for drug resistance. Bioinformatics, in press, http://bioinformatics.oxfordjournals.org/cgi/reprint/btl508v1.

Frater, A.J., Beardall, A., Ariyoshi, K., Churchill, D., Galpin, S., Clarke, J.R., Weber, J.N., McClure, M.O., 2001. Impact of baseline polymorphisms in RT and protease on outcome of highly active antiretroviral therapy in HIV-1-infected African patients. AIDS 15 (12), 1493–1502.

Grossman, Z., Paxinos, E.E., Averbuch, D., Maayan, S., Parkin, N.T., Engelhard, D., Lorber, M., Istomin, V., Shaked, Y., Mendelson, E., Ram, D., Petropoulos, C.J., Schapiro, J.M., 2004. Mutation D30N is not preferentially selected by human immunodeficiency virus type 1 subtype C in the development of resistance to nelfinavir. Antimicrob. Agents Chemother. 48 (6), 2159–2165.

Grossman, Z., Vardinon, N., Chemtob, D., Alkan, M.L., Bentwich, Z., Burke, M., Gottesman, G., Istomin, V., Levi, I., Maayan, S., Shahar, E., Schapiro, J.M., 2001. Genotypic variation of HIV-1 reverse transcriptase and protease: comparative analysis of clade C and clade B. AIDS 15 (12), 1453–1460.

Johnson, V.A., Brun-Vézinet, F., Onaventura, C., Conway, B., D'Aquila, R.T., Demeter, L.M., Kuritzkes, D.R., Pillay, D., Schapiro, J.M., Telenti, A., Richman, D.D., 2004. Update of the drug resistance mutations in HIV-1: 2004. Top HIV Med. 12 (4), 119–123.

Johnson, V.A., Brun-Vézinet, F., Onaventura, C., Conway, B., Kuritzkes, D.R., Pillay, D., Schapiro, J.M., Telenti, A., Richman, D.D., 2005. Update

of the Drug Resistance Mutations in HIV-1: Fall 2005. Top HIV Med. 13 (4), 125–131.

Kantor, R., Machekano, R., Gonzales, M.J., Dupnik, K., Schapiro, J.M., Shafer, R.W., 2001. Human immunodeficiency virus reverse transcriptase and protease sequence database: an expanded data model integrating natural language and sequence analysis programs. Nucleic Acids Res. 29 (1), 296–299.

Klingler, T.M., Brutlag, D.L., 1994. Discovering structural correlations in α-helices. Protein Sci. 3 (10), 1847–1857.

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 (5), 3744–3752.

Parkin, N.T., Schapiro, J.M., 2004. Antiretroviral drug resistance in non-subtype B HIV-1, HIV-2 and SIV. Antivir. Ther. 9 (1), 3–12.

Shafer, R.W., 2002. Genotypic testing for human immunodeficiency virus type 1 drug resistance. Clin. Microbiol. Rev. 15 (2), 247–277.

Sune, C., Brennan, L., Stover, D.R., Klimkait, T., 2004. Effect of polymorphisms on the replicative capacity of protease inhibitor-resistant HIV-1 variants under drug pressure. Clin. Microbiol. Infect. 10 (2), 119–1119.

Svicher, V., Ceccherini-Silberstein, F., Erba, F.u., Santoro, M., Gori, C., Bellocchi, M.C., Giannella, S., Trotta, M.P., Monforte, A.d., 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 (5), 2015–2025.

Turner, D., Brenner, B., Moisi, D., Detorio, M., Cesaire, R., Kurimura, T., Mori, H., Essex, M., Maayan, S., Wainberg, M.A., 2004. Nucleotide and amino acid polymorphisms at drug resistance sites in non-B-subtype variants of human immunodeficiency virus type 1. Antimicrob. Agents Chemother. 48 (8), 2993–2998.

Van Laethem, K., De Luca, A., Antinori, A., Cingolani, A., Perno, C.F., Vandamme, A.-M., 2002. A genotypic drug resistance interpretation algorithm that significantly predicts therapy response in HIV-1 infected patients. Antivir. Ther. 7 (2), 1359–6535.

van Maarseveen, N.M., de Jong, D., Boucher, C.A.B., Nijhuis, M., 2006. An increase in viral replicative capacity drives the evolution of protease inhibitor-resistant human immunodeficiency virus type 1 in the absence of drugs. J. Acquir. Immune Defic. Syndr. 42 (2), 162–168.

Wu, T.D., Shiffer, 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 (8), 4836–4847.

Please cite this article in press as: Deforche, K. et al., Bayesian network analysis of resistance pathways against HIV-1 protease inhibitors, Infect. Genet. Evol. (2007), doi:10.1016/j.meegid.2006.09.004