CTSHIV: A Knowledge-Based System For the Management of HIV-infected Patients

Michael Pazzani
Ranjit Iyer

Department of Information
& Computer Science
University of California
Irvine, California 92697

Darryl See
Edison Schroeder
Jeremiah Tilles

Department of Medicine
University of California
Irvine, California 92697

Abstract

A rule-based expert system, CTSHIV (Customized Treatment Strategies for HIV) that recommends an individualized treatment strategy for HIV patients is described. Since the HIV virus mutates rapidly, a patient can develop a resistance to particular antiretroviral agents. CTSHIV addresses this problem by recommending a treatment strategy that avoids antiretroviral agents for which a resistance has developed. By monitoring the HIV virus of the patient, the treatment strategy can be switched in response to mutations of the virus. CTSHIV contains a knowledge base that encodes information from the medical literature on drug-resistant mutations. It also contains additional rules that rank and weight combinations of antiretroviral agents based upon antiviral activities, redundant mechanisms of action, and overlapping toxicities. To recommend a treatment for an HIV-infected patient, plasma is obtained from the patient and the RNA extracted. Segments of the HIV pol gene encoding the entire protease, reverse transcriptase and integrase proteins are amplified by RT-PCR (using a total of 3 primer pairs) and cloned. The patient's HIV sequencing data is entered into the expert program and the information is downloaded directly into the CTSHIV program. The program produces the five most effective 2, 3, and 4 drug regimens coupled with an explanation for their choice. Thus, the CTSHIV system couples efficient genetic sequencing with an expert program that recommends regimens based upon information in the current medical literature. Clinical trials are underway to evaluate the effectiveness of this novel approach to the management of HIV-infected patients.

Introduction

When a patient has a bacterial infection, a physician attempts to identify the particular type of bacteria (e.g., Gram Negative or Gram Positive; Aerobic or Anaerobic, etc.) and recommends a treatment that is effective against that bacteria. Currently, when a patient is infected by the HIV virus, in the most advanced treatment regimen, a combination of antiretroviral agents is administered, usually consisting of a reverse transcriptase inhibitor (AZT, ddI, ddC, d4T, 3TC, Nevirapine, and Delavirdine) and a protease inhibitor (Saquinavir, Ritonavir and Indinavir). If the patient's viral load increases, or the patient's CD4 cell count (a laboratory marker of the strength of the immune system) decreases, it is assumed that the current treatment is no longer effective due to mutation of the HIV virus and the patient is switched to another treatment. However, there is not a mechanism currently used to determine that the new treatment will be effective. We describe a knowledge-based system that is intended to manage the information needed to make treatment of HIV analogous to the treatment of bacterial infection. In particular, it contains knowledge of which strains of the HIV virus are resistant to each antiretroviral agent. This knowledge is used to determine which antiretroviral agents to administer, to switch antiretroviral agents when the virus has mutated but before the immune system is weakened, and to determine which antiretroviral agents to switch to. The strains of virus that are most prevalent in the patient are found by genetic sequencing of the virus.

Background

Many compounds have been reported to inhibit replication of the Human Immunodeficiency Virus

(HIV) both *in vitro* and *in vivo* [1,2]. There are already 9 antiretroviral agents licensed for use in the US and many more are currently undergoing preclinical and clinical evaluation. Multiple clinical trials have demonstrated that treatment with these agents can improve clinical outcomes such as mortality, occurrence of opportunistic infections and progression to AIDS [3,4,5,6,7]. Therapy with a combination of agents appears to be superior to monotherapy [8].

Unfortunately, current therapeutic regimens result in suppression but not eradication of HIV. Furthermore, antiretroviral therapy is limited by the invariable development of resistance. A variety of studies have demonstrated that resistance is associated with specific mutations in the HIV pol gene [9,10,11,12]. If it were possible to directly monitor the occurrence of such mutations in the HIV RNA sequence from a given patient, specific alternative therapies might be considered before a surrogate marker, e.g. CD4 count or viral load, could be expected to even reflect a failure of the current regimen.

The clinical application of sequence information in an individual patient is currently not feasible because of the difficulty in interpreting the sequence information and relating it to the known body of medical literature. Fortunately, knowledge-based systems technology can be applied to solve this information management problem. Rule-based expert systems [13,14,15] declaratively represent knowledge of a specialized problem and facts about a specific case and draw inferences about the case. In the program developed for the current study (Customized Treatment Strategies for HIV [CTSHIV]), the rules encode information on drug-resistant mutations of HIV and characteristics of current antiretroviral agents, the facts are the sequences of the HIV genome obtained from a specific individual, and the inference to be drawn is a set of recommended drug-regimens. The knowledge of the HIV treatment problem is represented as a set of if-then rules of the form: IF $<$ antecedent $>$ THEN $<$ consequence $>$. For example, one such rule in CTSHIV is:

IF Leucine is encoded by codon 41 of the RT portion of the pol gene, then do not use AZT [weight=1.0].

The rule is derived from the literature which reports that a Methionine to Leucine mutation at codon 41 confers high-level resistance to AZT (>100-fold resistance [9]). The weight associated with a rule is not a confidence as in many expert systems. Rather, it should be viewed as part of the conclusion. The conclusions are weighted from 0.1 (low) to 1.0 (high) based upon expert advise and the reported level of resistance in the literature. The weight reflects the level of resistance to using a particular drug. The weights are additive. If there are several reasons for not using a drug, the total amount of resistance of the drug is the sum of the individual resistances. This reflects the underlying causal mechanism where a mutation changes the HIV virus in a particular location that is critical to the activity of the drug.

The level of resistance to each individual drug is determined by using a set of approximately 40 rules from the medical literature. Many rules exclude more than one drug from consideration. Base substitutions coding for either *in vitro* and/or *in vivo* resistance to specific drugs [16,17,18,19,20] were identified in the literature and included as rules.

CTSHIV also contains a second set of rules that rate combinations of antiretroviral agents rather than individual ones. Here, we exclude antiretroviral combinations with overlapping toxicities or redundant mechanisms of action and unproved efficacy and/or safety. There are fewer than 10 of these rules in CTSHIV. For example, one such rule in CTSHIV is:

If the combination contains Indinavir, Ritonavir or Saquinavir with either Nevirapine or Delavirdine
Then do not use the combination [weight=1.0].

With this information, CTSHIV uses its knowledge-base to find the level of resistance of each drug, and can find the weighting of each combination of drugs. No more than 4 antiretroviral agents are administered in practice, so this limits the numbers of combinations considered. For efficiency, a branch-and-bound search is used to find the five best combinations of 2 drugs, the five best combinations of 3 drugs and the five best combinations of 4 drugs. CTSHIV associates citations from the medical literature with each rule and maintains a trace of rule execution. Therefore, it can explain its conclusions and point the physician to the relevant medical literature.

CTSHIV goes through the following stages:

1. Blood is obtained from HIV-infected individuals after informed consent was obtained using a form approved by the Institutional Review Board at UC Irvine. All patients in the study have weakened immune systems (CD4 counts < 500 cells/mm^3). The protease, reverse transcriptase, and integrase segments of the HIV pol gene are cloned. Five clones from each patient are sequenced so that different strains of the virus within the same patient can be detected.

2. The level of resistance of each antiretroviral agent is found. The weight of the conclusion of each rule is multiplied by the fraction of the virus of each patient that has the specified mutation.

3. A branch-and-bound search algorithm finds the combinations of antiretroviral agents with the minimum weight. The weight of a combination is the sum of the weights of the individual drugs plus the weight of the combination. At most 10,000 combinations are considered, but the branch-and-bound search algorithm often reduces this by an order of magnitude.

4. A ranked ordering of drug regimens and the explanation for excluding any drug is printed. In the case that two drugs or combinations have equal weights, a default preference ordering of the drugs is used to favor one therapy over another.

CTSHIV is implemented in JAVA. Sample output in Figure 1 shows how CTSHIV displays recommendations (top), and how the recommendations are explained (lower). CTSHIV displays all mutations of the HIV virus and highlights the ones that result in drug resistence.

Clinical trials using CTSHIV were started in January 1997. Subjects in the study are re-evaluated by CTSHIV every three months to determine whether a change in therapy is warranted. The initial results of the clinical trials have been very promising.

Future Work

The CTSHIV system is limited by an incomplete understanding of the relationship between genomic mutations conferring drug resistance and clinical and surrogate marker outcomes. Furthermore, the expert program can only incorporate rules from the current literature. Like all knowledge-based systems, CTSHIV is designed to allow an expert to easily add rules to the system and to understand the effects of adding rules.

Our future research in CTSHIV will involve two directions. First, the weights in the conclusions of CTSHIV might be put on a more principled basis. The weighting of individual drugs and combinations follows the procedure that was previously done manually. Automating this procedure in CTSHIV saves a large amount of tedious and error-prone work. Nonetheless, the current implementation of weights is somewhat ad hoc. Fortunately, for the subjects in the current trial, CTSHIV has been able to find combination therapies with weights of 0, reducing the importance of the particular weighting scheme. Furthermore, the current procedure used in practice for the treatment of HIV-infected patients is considerably more ad hoc and less informed than the procedure implemented in CTSHIV.

Second, when enough patient data has been collected, a learning algorithm such as FOCL [14] might be applied to the program to suggest rule modifications based upon outcome measures from patients such as CD4 count, viral load, and progression to AIDS. New rules can be proposed by the learning system for review and confirmation. We hope to focus the learning of new rules by providing 3-dimensional structural knowledge of the HIV virus and the antiretroviral agents. By providing knowledge of the causal mechanisms of drug resistance, it is anticipated that a learning algorithm will be able to help identify possible mutation rules from reasonably-sized databases containing patients currently enrolled in HIV studies.

Acknowledgments

The technical assistance of Tonya Clark is appreciated. Many thanks to Douglas Richman MD for reviewing and revising the expert rules. The sequencing was performed by Erik Avajani and Clifford Brunk. This work was supported by the California University-wide AIDS Research Program through the California Collaborative Treatment Group (CCTG).

References

1. DeClercq E. Antiviral therapy for human immunodeficiency virus infections. Clin Microbiol Rev 1995;8:200-239.

2. Frost S, McLean A. Quasispecies dynamics and the emergence of drug resistance during zidovudine therapy of HIV infection. AIDS 1994;8:323-332.

3. Ho D, Moudgil D, Alan M. Quantitation of human immunodeficiency virus type 1 in the blood of infected persons. 1989;321:1621-1625.

4. Holodniy M, Katzenstein D, Winters M, Montoya J, Shafer R, Kozal M, Ragni M, Merigan T. Measurement of HIV virus load and genotypic resistance by gene amplification in asymptomatic subjects treated with combination therapy. J Acquired Immune Defic Syndr 1993;6;366-369.

5. Kahn J, Lagakos S, Richman D, Cross A, Petinelli C, Liou S, Brown M, Volberding P, Crumpacker C, Beall G, Sqacks S, Merigan T, Beltangady M, Smaldone L, Dolin R. A controlled trial comparing zidovudine with didanosine in human immunodeficiency virus infection. The NIAID AIDS Clinical Trials Group. N Engl J Med 1992;327;581-587.

6. Schooley RT, Ramirez-Ronda C, Lange JM, Cooper DA, Lavelle J, Lefkowitz L, Moore M, Larder BA, St. Clair M, Mulder JW. Virologic and immunologic benefits of initial combination therapy with zidovudine and zalcitabine or didanosine compared with zidovudine monotherapy. Wellcome Resistance Study Collaborative Group. J Infect Dis 1996;173;1354-66.

7. Zhu QY, Scarborough A, Polsky B, Chou TC. Drug combinations and effect parameters of zidovudine, stavudine, and nevirapine in standardized drug-sensitive and resistant HIV type 1 strains. AIDS Research and Human Retroviruses 1996;12;507-17.

8. Lange JM. Triple combinations: present and future. Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology 1995;10 Suppl 1;S77-82.

9. Larder B, Darby G, Richman D. HIV with reduced sensitivity to zidovudine (AZT) isolated during prolonged therapy. Science 1991;243;1731-1734.

10. Larder B, Kellam P, Kemp S. Zidovudine resistance predicted by direct detection of mutations in DNA from HIV-infected lymphocytes. AIDS 1991;5:137-144.

11. Larder B, Kemp S. Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT). Science 1989;246:1155-1158.

12. Frenkel L, Wagner M, Atwood S, Cummins T, Dewhurst S. Specific, sensitive, and rapid assay for HIV pol mutations associated with resistance to zidovudine and didanosine. 1995;33:342-347.

13. Jackson P. Introduction to Expert Systems. Addison-Wesley, Reading MA, 1990.

14. Pazzani M, Kibler D. The utility of knowledge in Inductive learning. Machine Learning 1992;9:57-94.

15. Ohno-Machado L, Parra E, Henry S, Tu S, Musen M. AIDS 2: a decision-support tool for decreasing physician's uncertainty regarding patient eligibility for HIV treatment protocols. Proceedings of the 17th Annual Symposium on Computer Applications in Medical Care, Washington DC, 1993, 429-433.

16. St Clair M, Martin J, Tudor W. Resistance to ddI and sensitivity to AZT induced by a mutation in HIV-1 reverse transcriptase. Science 1991;253:1557-1559.

17. Gu Z, Gao Q, Li X, Parniak M, Wainberg M. Novel mutation in the human immunodeficiency virus type 1 reverse transcriptase gene that encodes cross resistance to 2',3'-dideoxyinosine and 2',3'-dideoxycytidine. J Virol 1992;66;7128-7135.

18. Gu Z, Gao Q, Fang H, Parniak M, Brenner B, Wainberg M. Identification of novel mutations that confer drug resistance in the human immunodeficiency virus polymerase gene. Leukemia 1994;8S1:5166-5169.

19. Richman D, Shih C, Lowy I, Rose J, Prodanovich P, Goff S, Griffin J.

20. Human immunodeficiency virus type 1 mutants resistant to nonnucleoside inhibitors of reverse transcriptase arise in tissue culture. Proc Natl Acad Sci Usa. 1991;88;11241-11245.

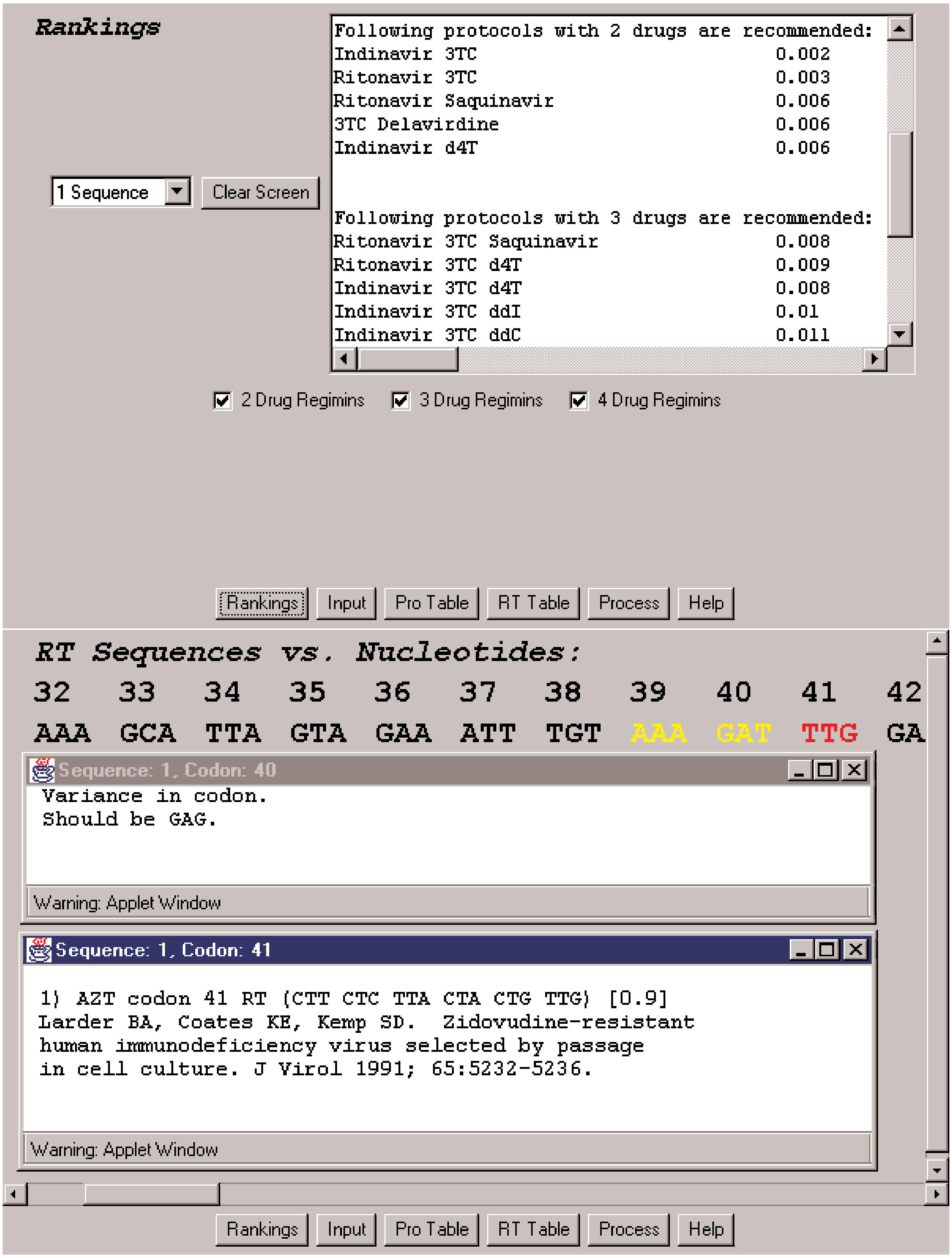

Figure 1: The output of CTSHIV rank orders combination therapy and explains the rankings.