Development of a Novel Codon-Specific Polymerase Chain Reaction for the Detection of CXCR4-Utilizing HIV Type 1 Subtype B

Sherry McLaughlin,^{1,2} Luke C. Swenson,³ Shengbo Hu,¹ Paul Hughes,¹ P. Richard Harrigan,^{3,4}
Robert W. Coombs,^{5,6} and Lisa M. Frenkel^{1,5,7,8}

Abstract

Insight concerning the switch in HIV-1 coreceptor use will lead to a better understanding of HIV-1 pathogenesis and host-virus dynamics. Predicting CXCR4 utilization by analyzing HIV-1 envelope consensus sequences is highly specific, but minority variants in the viral population are often missed resulting in low sensitivity. Commercial phenotypic assays are costly, and the development of sensitive in-house phenotypic assays to detect CXCR4-using HIV may not be feasible for some laboratories. A sensitive, inexpensive genotyping assay was developed to detect viral sequences associated with CXCR4-utilizing virus (X4). Codon-specific primer pairs were used to detect X4-associated codons at five positions in the HIV-1 envelope V3 loop (11, 13, 24, 25, and 32). Sixty plasma samples from HIV-1-infected individuals were analyzed by consensus sequencing and codon-specific PCR (CS-PCR). Forty-six of these were also phenotyped by Trofile or Enhanced Sensitivity Trofile (ESTA). CS-PCR detected X4 variants 17% more often than 11/24/25 consensus sequencing alone ($n=60$), 30% more often than Trofile ($n=27$), and in a limited data set, 16% more often than ESTA ($n=19$). CS-PCR combined with consensus sequencing had approximately 80% concordance with ESTA.

Introduction

HIV-1 PRIMARILY USES ONE of two coreceptors for entry into human cells: CC chemokine receptor 5 (CCR5) and CXC chemokine receptor 4 (CXCR4). CCR5-utilizing viral variants (R5) generally predominate early after infection and throughout most of the asymptomatic period of infection. CXCR4-utilizing variants (X4) typically emerge later in the course of disease as the immune system fails. While X4 is associated with the development of AIDS,¹ whether X4 causes or is the consequence of immune dysregulation has yet to be fully elucidated.

Detecting and monitoring coreceptor use in a laboratory setting are not trivial processes. X4 subpopulations are often at low levels—particularly early during disease progression—and even though there are multiple, well-established genotypic tools,^{2–5} population-based sequencing can generally detect only variants that are greater than 20%

of the viral population.⁶ In addition, consensus genotypes can be indeterminate when subpopulations are variable at the codons of interest, or when insertions or deletions cause overlapping reading frames; both are common in the HIV-1 envelope. Next-generation sequencing may prove to be the best method of detection, as it is not only sensitive, but can also give a fairly accurate estimation of the percent X4 in a population.^{7,8} However, next-generation sequencing is still fairly expensive and handling the data can be onerous. Phenotyping is another option for X4 detection, but many laboratories do not have the facilities or expertise to develop a sensitive culture-based assay, and commercial assays, such as Trofile, are costly. In addition, multiple amino acids in the HIV-1 envelope protein affect the efficiency of coreceptor use with progressively more changes occurring as the R5 viral population transitions to variants that can use CXCR4.⁹ How well a variant can use CXCR4 may influence how well phenotypic assays perform.

¹Seattle Children's Research Institute, Seattle, Washington.

²Department of Microbiology, University of Washington, Seattle, Washington.

³BC Centre for Excellence in HIV/AIDS, Vancouver, British Columbia, Canada.

⁴Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, Canada.

⁵Department of Laboratory Medicine, University of Washington, Seattle, Washington.

⁶Department of Medicine, University of Washington, Seattle, Washington.

⁷Department of Pediatrics, University of Washington, Seattle, Washington.

⁸Department Global Health, University of Washington, Seattle, Washington.

TABLE 1. PRIMERS FOR REVERSE TRANSCRIPTION, QUANTIFICATION OF TEMPLATE (qPCR), AND CODON-SPECIFIC PCR

<i>Primers to Reverse Transcribe Plasma RNA to cDNA and qPCR</i>					
	<i>F^a or R^b</i>	<i>Sequence</i>	<i>Name</i>	<i>pmol/RX^c</i>	
cDNA	R	AGT GCT TCC TGC TGC TCC CAA GAA CCC AAG	ED12		
	F	AAT TTC ACG GAC AAT GCT AAA AYC ATA ATA	envC2.F(2).pur	15	
qPCR	F	AGT ACA ATG TAC ACA TGG AAT TAR GCC AG	envC2.F(4).pur	15	
mix 1	R	TAC AGT AGA AAA ATT CCC CTC CAC AAT TA	envC3.R.pur	15	
	F	AAT TTC ACG GAC AAT GCT AAA AYC ATA ATA	envC2.F(2).pur	15	
qPCR	F	AGT ACA ATG TAC ACA TGG AAT TAR GCC AG	envC2.F(4).pur	15	
mix 2	R	GGA GGG GCA TAC ATT GCT TTT CCT AC	envC4.R3.pur	15	
<i>Codon-Specific Primers</i>					
<i>Codon</i>	<i>Reaction</i>	<i>F or R</i>	<i>Sequence</i>	<i>Name</i>	<i>pmol/RX^c</i>
11 AGG	AG+G	F	ACA AGA CCC AAC AAC AAT ACA AGA ARC AG	X4.11.AG.F	7.5
		R	AAA ATG CTC TCC CTG GTC CTA TAK KTT TC	X4.11.G.R	6.0
		R	AAA ATG CTC TCC CTG GTC CTA TAK ATT TC	X4.11.G.R(TMT) ^d	0.75
	A+GG	R	AAA ATG CTC TCC CTG GTC CTA TAC KTT TC	X4.11.G.R(MGT) ^d	0.75
		F	TAC AAG ACC CAA CAA CAA TAC AAG AAT AA	X4.11.A.F(2)	7.5
		R	AAA TGC TCT CCC TGG TCC TAT AKK TAG CC	X4.11.GG.R	7.5
11 AGA	AG+A	F	ACA AGA CCC AAC AAC AAT ACA AGA ARC AG	X4.11.AG.F	7.5
		R	AAA ATG CTC TCC CTG GTC CTA TAK KTC TT	X4.11.A.R	6.0
		R	AAA ATG CTC TCC CTG GTC CTA TAK ATC TT	X4.11.A.R(TMT)	0.75
	A+GA	R	AAA ATG CTC TCC CTG GTC CTA TAC KTC TT	X4.11.A.R(MGT)	0.75
		F	TAC AAG ACC CAA CAA CAA TAC AAG AAT AA	X4.11.A.F	7.5
		R	AAA TGC TCT CCC TGG TCC TAT AKK TAG TC	X4.11.GA.R	7.5
11 CGT	C+CGT	F	TA CAA GAC CCA ACA ACA ATA CAA GAA RTC	X4.11.C.F	7.5
		R	AA TGC TCT CCC TGG TCC YAT AKK TAT TCG	X4.11.CGT.R	7.5
25 AAA	AA+A	F	CCA GGG AGA GCA DTT TAT GCA ACA GRT (A)A	25.AA.F.LNA ^e	7.5
		R	AG TTA CAA TGT GCT TGT CTT ATA TCT CCT ATT CTT	X4.25.A.R	7.5
	A+AA	F	ACC AGG GAG AGC AKT TTA TRC AAC AGR TA	X4.25.A.F	7.5
		R	GTT ACA ATR TGC TTK YCT TAT ATC TCC TAT TAG TT	X4.25.AA.R	7.5
25 AGA	AG+A	F	CCA GGG AGA GCA DTT TAT GCA ACA GRT (A)G	25.AG.F.LNA ^e	7.5
		R	AG TTA CAA TGT GCT TGT CTT ATA TCT CCT ATT CTT	X4.25.A.R	7.5
	A+GA	F	ACC AGG GAG AGC AKT TTA TRC AAC AGR TA	X4.25.A.F	7.5
		R	GTT ACA ATG TGC TTG TCT TAT ATC TCC TAT TAG TC	X4.25.GA.R	7.5
24 AAA	A+AA	F	GGA CCA GGG AGA GCA TTT TAT RCA ACT A	24.A.F	7.5
		R	C AAT GTG CTT GTC TTA TAT CTC CTA TTA TKY ATT	X4.24.AA.R	7.5
24 AGA	A+GA	F	GGA CCA GGG AGA GCA TTT TAT RCA ACT A	24.A.F	7.5
		R	C AAT GTG CTT GTC TTA TAT CTC CTA TTA TKY ATC	X4.24.GA.R	7.5
13 CGT	C+GT	F	CCC AAC AAC AAT ACA AGA AAA RGY ATT C	13.C.F	7.5
		R	GYA TAA AAT GCT CTC CCT GGT CCY ATT C	13.GT.R	7.5
32 AAA	A+AA	F	ACA ARA GAM ATA ATA GGA RAY ATA AGR A	32.A.F(-mm)	7.5
		R	CCA TTY TGY TCT ACT AAK GTT ACA ATR TGT TT	32.AA.R(-3T)	7.5
32 CGN	C+GN	F	AAC AGG AGA MAT AAT AGG ARA TAT AAG CC	32.C.F	7.5
		R	CCA TTY TGY TCT ACT AAK GTT ACA ATR TGC AC	32.GA.R(-2a)	7.5
11 All		F	GTA CAA GAC CCA RCA ACA ATA CAA GAA RA	25.total.F	7.5
		R	TAA AAT GCT CTC CCT GGT CCY ATA KKT AT	25.total.R	7.5
25 All		F	GAC CAG GGA GAG CAT TTT ATR CAA CAG RA	11.total.F	7.5
		R	CAA TGT GCT TGT CTT ATA TCT CCT ATT AT	11.total.R	7.5

^aF, forward primer.^bR, reverse primer.^cpmol/RX, pmol of each primer per reaction.^dTMT or MGT refer to the degeneracy added to the reverse primer mix to accommodate variability at codon 13.^eLNA, locked nucleic acid, are indicated in parentheses.

CXCR4 use by HIV-1 is associated with viral genotypes that have arginine or lysine at position 11 or 25 in the variable loop 3 (V3) of the HIV-1 envelope protein.¹⁰ The identification of basic amino acids at positions 11 and 25 has >90% specificity for identifying HIV-1 subtype B X4 variants.^{2,3} Multiple three-base combinations encode the many amino acids that can be found at positions 11 and 25, including the eight that

can encode lysine or arginine. Therefore, no single base can be used to differentiate X4-associated versus R5-associated codons.

This article describes a relatively inexpensive polymerase chain reaction (PCR)-based method to detect X4 by specifically amplifying X4-associated codons found in the V3 loop of HIV-1 subtype B envelope. Codon-specific PCR

(CS-PCR) targets five codon positions that most commonly encode one or more basic amino acids in X4 variants (11, 13, 24, 25, and 32). A semiquantitative first round PCR (qPCR) was performed to ensure sufficient input of viral templates from each specimen and the resulting amplicon was used as the template for CS-PCR. The genotypic detection of X4 by CS-PCR combined with consensus sequencing was compared to consensus sequencing alone and to a commercial phenotypic assay (Trofile or Extra Sensitive Trofile Assay).

Materials and Methods

Specimens and human subjects approval

Clinical specimens, as approved by the University of Washington and the Children's Hospital of Philadelphia Institutional Review Boards, were utilized to optimize the assay. Additional test specimens with 454-pyrosequence data, graciously provided by Dr. Richard Harrigan with permission from Pfizer, were used to define cutoffs.

Isolation of RNA and generation of cDNA

Virus was pelleted from 1.0 to 1.5 ml of plasma by centrifugation at $25.2K \times g$ for 2 h at 4°C. The supernate was carefully removed leaving 50–100 μ l of plasma undisturbed above the pellet. The volume over the pellet was adjusted to 140 μ l with phosphate-buffered saline (PBS) and RNA was isolated using the Qiagen QIAamp Viral RNA Mini Kit (Valencia, CA) as per the manufacturer's protocol.

cDNA was generated using 5 μ l of the extracted RNA in a 20- μ l reaction containing 10 pmol of ED12 primer (Table 1), 500 μ M dNTP, 1 μ l 0.1 M dithiothreitol (DTT), 0.5 μ l RNasin (Promega, Madison, WI), and 1 μ l of Superscript III (Invitrogen). Primer, RNA, and dNTP were heated at 65°C for 5 min, cooled to 45°C for 5 min, at which time the enzyme, buffer, DTT mix were added, and incubation was continued at 50°C for 1 h.

Quantification of input virus by real-time PCR

The cDNA was amplified using semiquantitative real-time PCR with two primer sets (Table 1) in a BioRad iCycler (Hercules, CA), each in duplicate. Two different primer sets were used to increase the likelihood of efficient amplification of envelope sequences. Each 50 μ l reaction was prepared with 0.5 $\times$ SensiMix (Sybr Green-fluorescein formulation, Bioline, London, UK) and 15 pmol each primer. To compensate for using only half concentration real-time mix, the reaction mixture was adjusted to 1 $\times$ buffer with Bioline Biolase buffer (London, UK) and an additional 75 nmol magnesium and 5 nmol dNTPs were added. Two to 5 μ l of cDNA was added to the mix and the reactions were carried out as follows: 10 min at 95°C; followed by 50 cycles of 10 s at 92°C, 20 s at 55°C, and 20 s at 72°C concomitant with Sybr Green detection. Following the last cycle a melting curve was generated from 60° to 100°. The p2-7 plasmid (Table 2) was used to create a standard curve (10^4 , 10^3 , 10^2 , and 10 copies). The quantified input template from each of the four reactions was totaled (2 replicates $\times$ 2 primer sets). If <100 copies of input cDNA were amplified, the cDNA/qPCR was repeated to amplify a minimum of 100 copies of cDNA, or until the sample was exhausted.

TABLE 2. X4 AND R5 CODON-SPECIFIC CONTROL PLASMID SEQUENCES

Codon	X4 or R5 ^a	Name	Codon tested	Amino acid	GenBank #
11 AGG	X4	3605.6-1	AGG	Arginine	HQ827788
	R5	2-7	AGT	Serine	FJ222410
11 AGA	X4	G1	AGA	Arginine	AY075875
	R5	2-7	AGT	Serine	FJ222410
11 CGT	X4	13-73	CGT	Arginine	FJ222438
	R5	2-7	AGT	Serine	FJ222410
25 AAA	X4	P2v80	AAA	Lysine	AF137843
	R5	1362B	GAA	Glutamate	DQ853436
25 AGA	X4	1365#7	AGA	Arginine	DQ853463
	R5	G2	GGA	Glycine	AY076052
24 AAA	X4	G1	AAA	Lysine	AY075875
	R5	P2v80	GAA	Glutamate	AF137843
24 AGA	X4	89.6	AGA	Arginine	U39362
	R5	1362B	GGA	Glycine	DQ853436
13 CGT	X4	1365#7	CGT	Arginine	DQ853463
	R5	1362B	CAT	Histidine	DQ853436
32 AAA	X4	13-73	AAA	Lysine	FJ222438
	R5	2-7	CAA	Glutamine	FJ222410
32 CGA	X4	C12	CGA	Arginine	AF204663
	R5	2-7	CAA	Glutamine	FJ222410

^aDesignates whether the codon is X4- or R5-associated and is accordingly used as an X4 or R5 control.

CS-PCR and generation of the consensus sequence

The qPCR from each sample was combined in a single tube and 100 μ l of the mixed amplicon was purified (QIAquick PCR Purification Kit, Qiagen). DNA was eluted with 50 μ l elution buffer, quantified using a spectrophotometer (Nanodrop, Thermo Scientific, Wilmington, DE), and normalized to 30 ng/ μ l by the addition of Tris-EDTA buffer (TE). Five microliters of the normalized template was visualized in an agarose gel to verify that the PCR products were of the desired size and of similar intensity across specimens. A consensus sequence was generated using 0.5 μ l of DNA in a 5 μ l Big-Dye (ABI, Carlsbad, CA) sequencing reaction with primer envC2.F2 and/or envC3.R1 (Table 1).

Each normalized template was diluted 1:200 in TE buffer resulting in $\sim 3 \times 10^8$ templates/ μ l (performed in strip-well tubes for efficient transfer with a multichannel pipette) and was then subjected to CS-PCR, where each codon was amplified individually. CS-PCRs were performed in 25 μ l containing 1 μ l of diluted template in 0.5 $\times$ SensiMix and 7.5 pmol of each primer (Table 1). The reaction mix was adjusted to 1 $\times$ buffer by adding Bioline Biolase buffer (without magnesium). PCR conditions were 95°C for 10 min, followed by 44 cycles of 10 s at 92°C, 1 s at either 45°C (codons 11, 13, 24AGA, and 25AGA) or 40°C (codons 24AAA, 25AAA, and 32) and 5 s at 72°C (concomitant with Sybr detection). This was followed with a melting curve from 60° to 90°C.

CS-PCR primers, controls, and set-up

Codon-specific primer pairs were designed to detect one or more X4-associated codons at positions 11, 13, 24, 25, and 32 of the V3 loop (c11, c13, c24, c25, and c32) (Table 1). Primer pairs bisect the target codons in one or two configurations (NN+N and/or N+NN, Fig. 1a and b) except for c11 CGT,

FIG. 1. Primer designs for codon 11 AGG (arginine), codon 11 CGN (arginine), and the codon 11 sample-amplification-control. Codon 11 is underlined by a bold black bar. Primer sequences are shown in bold and the engineered mismatches 2 or 3 bases from the 3' end are marked with an asterisk. (a) A + GG: the forward primer terminates at the first base of the codon and the reverse primer terminates at the second base of the codon. (b) AG + G: the forward primer terminates at the second base of the codon and the reverse primer terminates at the last base of the codon. (c) CGN: both forward and reverse primers anneal at the first base of the codon. All four base configurations (CGN) can be detected with this primer set, including CGT, the most common genotype. (d) Sample-amplification-control primers precisely flank the codon. Poor amplification of the template by these primers indicates that there are either inhibitory mismatches between the primers and the template, or that not enough template was added to the reaction.

which overlaps at the first base of the codon (Fig. 1c). To increase amplification specificity, the primers have "engineered" mismatches introduced at the second or third base from each 3' end. Each CS-PCR primer pair has matching X4-codon and consensus R5-codon control plasmids (Table 2). A 1:100 mixture of X4:R5 controls was used to monitor sensitivity. Plasmid controls were linearized with *EcoRI* and diluted to $\sim 3 \times 10^8$ copies/ μ l.

To monitor sample input concentration and gauge the effect of primer mismatches, the normalized sample templates were also amplified by primer sets that precisely flank c11 and c25 (Fig. 1d). These reactions are referred to as sample-amplification-controls (SAC).

Typically, 10 viral samples were tested with three codon-specific primer pairs and the sample-amplification-control primers, all in duplicate, in a 96-well plate (see Fig. 2). Each set of codon-specific reactions was accompanied by the respective controls: X4, R5, 1:100 mix of X4:R5, and a no DNA. Each X4 and R5 control was also amplified by the sample-amplification-control primers to monitor three parameters: whether too much sample template was submitted to CS-PCR (which was indicated by an earlier real-time cycle threshold than the R5 control), the stability of the controls, and the consistency of assay performance across experiments.

These are referred to as X4 and R5-amplification-controls (XAC and RAC).

Interpretation of consensus sequences and codon-specific PCR

Consensus sequence data were edited using Sequencher 4.2.2 (Gene Codes Corporation, Ann Arbor, MI). In chromatograms with low background, the secondary peaks $\geq 20\%$ of the upper peak were called genotypic mixtures. When background was high, usually due to subpopulations with insertions/deletions, only the highest peaks were recorded. All possible amino acid sequences were subjected to webPSSM (<http://indra.mullins.microbiol.washington.edu/webpssm/>)² and Geno2pheno (<http://coreceptor.bioinf.mpi-inf.mpg.de/index.php>)³.

CS-PCR data were analyzed using the iCycler software. The analysis mode was set to "PCR Base Line Subtracted Curve Fit," baseline cycles were autocalculated, and the threshold position for each real-time PCR run was set to 50 RFU. The average cycle threshold (C_t) value for each sample was subtracted from the C_t value for its matching R5-codon control (ΔC_t). Positive ΔC_t values were recorded. Melting curves were monitored to ensure that a positive ΔC_t was not due to mispriming.

Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8	Sample 9	Sample 10	11 AG+G X4 Cont 3605	11 AG+G R5 Cont 2-7
Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8	Sample 9	Sample 10	11 AG+G X4:R5 1:100	11 AG+G no DNA
Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8	Sample 9	Sample 10	11 A+GG X4 Cont 3605	11 A+GG R5 Cont 2-7
Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8	Sample 9	Sample 10	11 A+GG X4:R5 1:100	11 A+GG no DNA
Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8	Sample 9	Sample 10	11 CGT X4 Cont 13-73	11 CGT R5 Cont 2-7
Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8	Sample 9	Sample 10	11 CGT X4:R5 1:100	11 CGT no DNA
Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8	Sample 9	Sample 10	11 Amp Control 3605	11 Amp Control 2-7
Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8	Sample 9	Sample 10	11 Amp Control 13-73	11 Amp Control 2-7

FIG. 2. Example plate format for codon-specific (CS)-PCR. This configuration tests 10 samples with three codon-specific primer pairs and the sample-amplification-control primers, all in duplicate, in a 96-well plate. Controls for each primer set are in the two right columns. These include the X4 and R5 controls (Cont), the 1:100 X4:R5 sensitivity control, and a no DNA control. In this example, the top two rows screen for c11 AG+G, the second two for c11 A+GG, the third two for c11 CGN, and the bottom two rows contain the amplification controls (Amp Control) for both the samples and the control plasmids. The specific X4 and R5 codon controls for the c11 AGG and c11 CGN codons (see Table 2) are noted in the two right columns.

The CS-PCR cutoffs were optimized by testing 78 first round *env* PCR amplicons (termed "test specimens") that had population sequences and 454-pyrosequences previously generated at the British Columbia Centre for Excellence in HIV/AIDS (Vancouver, Canada). These test specimens were from patients who entered the MOTIVATE/1029 trials for maraviroc after the 1029 study filled up. Processing of the clinical specimens and generation of the HIV envelope PCR amplicons were previously described.⁸ Prior to CS-PCR, the first-round PCR products were quantified using a double-stranded DNA quantification reagent (Quant-It PicoGreen; Invitrogen) and normalized to $\sim 3 \times 10^8$ amplicons per μ l. One microliter of the normalized template was amplified by real-time CS-PCR, in duplicate, as described above. CS-PCR results were compared to the most common X4-associated and most common R5-associated sequence obtained from each specimen by 454-pyrosequencing (using a geno2pheno cutoff of 3.5). The CS-PCR results for the test specimens, as well as corresponding % X4 by 454-pyrosequencing and geno2pheno classifications by Sanger sequencing (using a cutoff of 5.75), are presented in Supplementary Table S1 (Supplementary Data are available online at www.liebertpub.com/aid). Codon/sequence-specific effects on CS-PCR efficiency were noted and ΔC_t cutoff values were accordingly assigned (Supplementary Table S2 and Table 3). The clinical samples analyzed in this article were typed as X4 based on the ΔC_t cutoff values that were determined for each codon using the test samples. An outline of the decision-making rules is shown in Fig. 3.

Statistical comparisons

The detection of X4 viral variants as determined by CS-PCR was compared to genotyping by consensus sequencing alone, and CS-PCR combined with consensus sequencing was compared to X4 classification by the phenotypic assays Trofile and Enhanced Sensitivity Trofile Assay (ESTA) (performed at

Monogram Biosciences) using McNemar's two-tailed test. Additionally, the degree of association (kappa) was determined for the comparison with ESTA.

Results

Development of quantitative PCR

To ensure that a sufficient percentage of the viral population was assayed, a qPCR was developed to estimate the amplifiable viral RNA templates isolated from each plasma sample. Subsequently, amplicons derived from at least 100 cDNA templates were submitted to CS-PCR amplification,

TABLE 3. CODON-SPECIFIC-PCR ΔC_t CUTOFF VALUES FOR EACH CODON

Codon	ΔC_t cutoff
11 AG+G/A+GG	4 combined ^a
11 AG+A/A+GA	4 combined
11 CGT	2
13 CGT	3
24 AAA	2
24 AGA	4
25 AA+A/A+AA	4 combined
25 AG+A/A+GA	4 combined
32 AAA	2
32 CGA	2

^a ΔC_t s for the matched AR+R and A+RR reactions for codons 11 and 25 are combined to make the cutoff.

Seventy-seven samples that had previously been genotyped based on 454-pyrosequencing data were assayed by codon-specific (CS)-PCR. Codon/sequence-specific effects on CS-PCR efficiency were noted and ΔC_t cutoff values were assigned accordingly.

FIG. 3. Decision outline used to analyze the consensus sequence and CS-PCR for the X4-associated genotype.

- 1) Sequence analysis
 - a) Generate a population sequence for each sample.
 - b) Use your favorite genotyping tool to determine coreceptor use.
- 2) CS-PCR Analysis
 - a) Compare sample-amplification-controls (SAC) to R5-amplification-controls (RAC).
 - i) The average SAC Ct should not be more than 10 to 12 cycles greater than corresponding RAC Cts. If an SAC Ct is that much greater than an RAC Ct, it means that either there are multiple or critical mismatches with the primers, or that there was too little template in the reaction. Mismatches may indicate that CS-PCR will not work for that codon. If there was too little template, the amount can be increased and the sample rerun.
 - ii) The average SAC Ct should not be less than the RAC Ct. If it is, then there was too much template added to the reaction and any positive CS-PCR for that codon is suspect. The sample should be rerun using less template.
 - b) Determine the Δ Ct for each sample CS-PCR (see methods).
 - c) Designate a CS-PCR as positive if the Δ Ct is greater than the cutoff for that codon (see Table 3).
 - d) Verify that all positive CS-PCR have a melting curve that matches the X4-control melting curve.
 - e) Using the consensus sequence, verify that none of the positive CS-PCR templates have bases that are complementary to the 'engineered' mismatch in the primer.
 - f) Classify the sample as X4 if:
 - i) c11 CGT is positive
 - ii) c24 AAA or AGA is positive AND either c13 CGT or c32 AAA or CGA are positive
 - iii) c11 AG+G/A+GG or c11 AG+A/A+GA or c25 AA+A/A+AA or c25 AG+A/A+GA are positive

which ensured sufficient copies of HIV-1 to allow detection of X4 variants at a prevalence of 1–3%. The primers used to quantify the HIV-1 *env* V3 templates were designed to anneal to the relatively conserved regions flanking V3 (C2, C3, and C4). To increase the likelihood of efficient primer binding, two different qPCR primer mixtures were used in separate reactions. Each reaction contained the same two forward primers along with a unique reverse primer (Table 1). There were no mismatched bases between the primers and the control plasmid used to generate the standard curve (p2-7). Most samples amplified more efficiently with one or the other primer set indicating that mismatches between virus and the primers influenced copy number determination. However, since the standard control matched the sequence of the primers exactly, the qPCR should not have overestimated the amplifiable virus from the samples. Estimates of the number of viral RNA templates amplified from each of the 60 optimization specimens are shown in Table 4.

Development of CS-PCR for detection of X4-associated codons c11 (AGG/AGA/CGN) and c25 (AAA/AGA)

The PCR was designed to target whole codons with "codon-specific" primer pairs. The forward and reverse primers bisect the codon such that all three bases of the targeted codon must be present for efficient amplification to occur (Fig. 4). The codon was split A + RR or AR + R, utilizing two sets of primers. One configuration was typically more efficient and/or specific. The c11 CGN arginine primers were designed differently to accommodate all four CGN codons in a single reaction (Fig. 1c). Not all possible codons for arginine and lysine were included in the assay because three codons represent 99% of the basic residues at c11 and two codons represent 95% of X4 variants at c25 (tabulated from an

alignment of 6805 HIV-1 subtype B sequences obtained from the Los Alamos National Laboratory HIV-1 Sequence Database).

c24 (AAA/AGA), c32 (AAA/CGA), and c13 (CGT)

Based on multiple observations that basic amino acids at c24, c13, or c32 are associated with X4 variants,^{1,11–15} these codons were included in the assay, but any one alone was not sufficient to classify the sample as X4. Codons 24 (AAA and AGA), 13 (CGA), and 32 (AAA and CGA) are detected with single reactions in the N+NN configuration.

Discrimination between X4 and R5-associated codons

During development of CS-PCR, multiple primer pairs and reaction conditions were tested in an attempt to balance primer binding at low stringency—which would better accommodate the high degree of HIV-1 envelope diversity—with the tendency of Taq polymerase to indiscriminately extend from the 3' end of a primer that was not complementary to the template. Mismatches were engineered at the –2 or –3 position of each primer to increase the specificity of the reaction. This created two mismatches with nontargeted codons at or within one base of the primer's 3' end (Fig. 4), which greatly decreased nonspecific amplification. The base chosen to introduce a mismatch in the primer was rarely found at that location in sequences obtained from the Los Alamos National Laboratory HIV-1 Sequence Database.

In addition to the engineered mismatches, a low magnesium concentration was required to inhibit nonspecific extension by Taq. The 1 × SensiMix real-time mix has 3 mM Mg²⁺. Half concentration of the real-time mix (1.5 mM Mg²⁺) increased the specificity thereby reducing false-positive reactions. In addition, the lower Mg²⁺ concentration decreased the incidence of primer-

TABLE 4. COMPARISON OF INDIVIDUAL CODON CODON-SPECIFIC-PCR RESULTS TO CONSENSUS SEQUENCING AND COMMERCIAL PHENOTYPING

ID	VL per ml	Copies screened	c11		c11		c25		c25		c24		c24		c13		c32		11/24/25 con seq ^a	CS-PCR ^b	G2P ^c	PSSM ^d	Tprofile or ESTA ^e
			AG+ G	A+ GG	AG+ A	A+ GA	c11	AA+ A	A+ AA	AG+ A	A+ GA	c24	A+ AA	A+ GA	C+ GT	C+ GN	A+ AA	c32					
1	33,754	147							3.0	6.5							6.0		R5	X4	R5	R5	R5
2	45,670	123															8.7		R5	R5	R5	R5	R5
3	71,801	123															8.5	2.0	R5	R5	R5	R5	R5
4	44,127	142																	R5	R5	R5	R5	Non-rep
5	1,544	100	2.6	seq										2.1			3.2	X4	R5	R5	FrSh	FrSh	D/M
6	29,824	111				6.3			1.3	1.8						13.9	5.1	seq	R5	X4	R5	R5	R5
7	4,677	244														5.1	5.5		R5	R5	R5	R5	R5
8	13,266	149															8.3		R5	R5	R5	R5	R5
9	14,280	573	18.7	13.8						2.6						9.5	seq		X4	X4	X4	X4	R5
10	100,000	917		4.0		4.4	14.6										13.3		X4	X4	X4	X4	D/M
11	49,101	125															3.9		R5	R5	R5	R5	R5
12	72,653	80							4.2	5.1						2.4	5.1		R5	X4	R5	R5	R5
13	6,603	28															4.1		R5	R5	R5	R5	R5
14	8,344	101															6.7		R5	R5	R5	R5	R5
15	30,440	322											4.7			6.1			R5	X4	X4	R5	D/M
16	26,348	103							seq	8.2						9.5	7.7	seq	X4	X4	R5	X4	R5
17	14,183	117								7.9						5.0			R5	X4	R5	R5	R5
18	51,871	121																5.8	R5	X4	R5	R5	R5
19	5,413	111		5.9		4.7										7.9			R5	R5	R5	R5	R5
20	26,474	128																	R5	R5	R5	R5	R5
21	25,161	126														9.9	3.0		R5	R5	R5	R5	R5
22	13,181	119		2.4													2.7		R5	R5	R5	R5	R5
23	5,231	149																	R5	R5	R5	R5	R5
24	<50	400																	X4	R5	X4	X4	Non-rep
25	3,938	125																	R5	R5	R5	R5	Non-rep
26	4,457	1,011							7.7	8.1						11.4	4.7		X4	X4	R5	R5	not done
27	47,306	276																	R5	R5	R5	R5	R5
28	18,090	107							4.8	5.0				6.4		3.2	5.8		X4	X4	R5	X4	R5
29	41,600	148																	R5	R5	R5	R5	R5
30	1,314	154																	R5	R5	R5	R5	not done
31	45,714	331														3.5			R5	R5	R5	R5	R5
32	47,046	586																	R5	R5	R5	R5	R5
33	101,600	581															4.8		R5	R5	R5	R5	R5
34	88	334							1.4	2.3									R5	R5	X4	X4	R5
35	UK	315															4.6		R5	R5	R5	R5	not done
36	225	76														10.5	2.8	3.0	R5	R5	R5	R5	R5
37	4,750	210							8.0	3.7						14.0			X4	X4	X4	R5	R5
38	3,190	16,350							3.3	10.8	3.1	5.3							X4	X4	X4	R5	not done
39	50,000	138																3.0	R5	R5	R5	R5	not done
40	UK	1,658								6.5						6.2	11.6		R5	X4	R5	R5	not done

(continued)

TABLE 4. (CONTINUED)

ID	VL per ml	Copies screened	c11			c11			c25			c25			c24			c13			c32			11/24/25 con seq ^a			PSSM ^d	G2P ^c	CS-PCR ^b	Tprofile or ESTA ^e			
			AG+	A+	GG	AG+	A	GA	AA+	A+	AA	AG+	A	GA	A+	AA	GA	C+	C+	GT	C+	AA	AA	seq	AA	seq					AA		
41	58,000	500																															
42	2,990	110																															
43	UK	548																															
44	553,000	158																															
45	71,500	273																															
46	11,800	307																															
47	2,430	30																															
48	107,000	2,278																															
49	75,300	450																															
50	4,080	1,765																															
51	46,200	162																															
52	36,000	1,768																															
53	8,030	101																															
54	21,460	29																															
55	112,300	200																															
56	59,430	155																															
57	47,700	160																															
58	2,502	190																															
59	6,048	185																															
60	117,400	194																															

^aThe genotype is assigned by the occurrence of arginine or lysine at positions 11 or 25, or at 24 plus 13 or 32 in the consensus sequence.

^bThe genotype is assigned by the detection of arginine or lysine at positions 11 or 25, or at 24 plus 13 or 32 using CS-PCR.

^cThe genotype is assigned by G2P³ using the consensus sequence. The false-positive rate was set to the "Optimized cutoffs" based on analysis of clinical data from MOTIVATE (2% and 5.75% FPR). In some cases a secondary sequence with an indel could be discerned with some confidence and was also evaluated by G2P.

^dThe genotype is assigned by PSSM² using the consensus sequence. In some cases a secondary sequence with an indel could be discerned with some confidence and was also evaluated by PSSM.

^ePhenotype determined at Monogram Biosciences. The first 32 samples were assayed by the original Tprofile and the remainder by the Enhanced Sensitivity Tprofile Assay (ESTA). Specimens from 60 HIV-1-infected individuals (rows 1–60) were examined for X4 virus. The plasma viral loads (VL) obtained clinically are shown in the second column, and the minimum copies of HIV-1 RNA assayed by codon-specific (CS)-PCR and consensus sequencing are in the third column. Values shown in columns 4–17 are the ΔC_s between a positive CS-PCR and its corresponding R5 control (only those ≥ 2 cycles are shown unless they add support to an X4 classification). Those ΔC_s that are highlighted in darker gray indicate that the codon was also visible in the sequence chromatogram.

seq, an X4-associated codon could be visualized in the sequencing chromatogram, but was not detected by CS-PCR; *F/S/I*, the consensus sequence had a combination of insertions and deletions that made it impossible to analyze by PSSM; *Nonrep*, nonreportable designation was assigned by Monogram Biosciences; *D/M*, dual-mixed. Indicates that both coreceptors were utilized by the viruses in the sample. *D/M* is equivalent to a genotype of X4.

FIG. 4. Codon-specific primers by design do not amplify R5-associated codons. The primers designed to amplify X4 codons are shown in bold with the engineered mismatches marked with an asterisk. Representative double-stranded R5-associated sequences are shown in regular type with the target codon underlined. Mismatches between the sequence and the primers are indicated by vertical lines drawn through the mismatched bases. The primers are designed so that extension is inhibited in at least one direction, hindering amplification. An "X" indicates where extension is impeded and an arrow shows where extension can occur. (a) When c11 is AGT (serine), the A+GG primer set will extend in the forward direction, but will have mismatches at the second and third (−2, −3) base from the 3' end of the reverse primer, which interferes with extension. (b) When c11 is GGG (glycine), the A+GG primer set will extend in the reverse direction, but will have mismatches at −1 and −3 in the forward primer, which will impede extension. (c, d) A similar situation is shown for the AG+G primer set.

dimer at low annealing temperatures. Although use of a diluted mix detracts from the simplicity offered by a commercially prepared mix, it halves the amount of an expensive reagent.

Reaction conditions that were sufficiently stringent to completely stop amplification of the R5 controls generally severely reduced the amplification efficiency of the X4 controls. Therefore, optimal reaction conditions were empirically established that resulted in at least a 10 cycle ΔC_t between the X4 and R5-codon controls. ΔC_t s were calculated for each sample CS-PCR that crossed the threshold before the R5-codon control. ΔC_t cutoff values for each codon were established on a "training set" of test specimens that had accompanying 454-pyrosequencing data (Table 3 and Supplementary Tables 1 and 2).¹¹ ΔC_t values for CS-PCR that were greater than the cutoff were considered positive for that codon and X4 determinations were assigned accordingly.

Initially, differentiation of c25 GAA (R5) from c25 AAA (X4) was problematic, with only a few cycles difference between the X4 and R5 control C_t values. To increase the specificity of the c25 AA+A primer set, a "locked" nucleic acid (a modified nucleotide with the conformation fixed to increase hybridization specificity^{16,17}) was incorporated at the −2 position of the forward primer, which corresponds to the first G of the R5-associated GAN codon. This modification increased the specificity such that the difference in C_t values between X4 and R5 controls was greater than 10 cycles.

Optimization of CS-PCR using clinical specimens that were also analyzed by consensus sequencing and Tprofile

Overview of CS-PCR and sequencing results. Sixty plasma specimens from HIV-1-infected individuals were analyzed by CS-PCR and consensus sequencing during development of the assay (Table 4). These samples are separate from the test specimens used to establish the ΔC_t cutoffs. The V3 consensus sequence from 34 of the 60 samples (55%) had minority multivariant or indel subpopulations that could not be resolved by the sequence analysis software. Inspection of these chromatograms revealed indel variants in 13 of the 34, including four (#9, 17, 24, and 28) with peaks that had heights greater than 20% of the primary variant.

Discordant results between the two primer sets for c11 or c25. There were discordant results between the two primer sets targeting the same codon (A+RR and AR+R) in 31 (66%) of the 47 reactions that detected X4-associated codons at c11 or c25. In a minority of cases ($n=3$) mismatches between the viral target and one of the primers was evident in the sequencing chromatogram. For example, c11 in sample #52 was 100% AGG in the consensus sequence. While the c11 A+GG reaction had a high ΔC_t , the c11 AG+G reaction was negative due to GTA at c12 rather than the consensus ATA. This

substitution created a double mismatch at the second and third bases of the reverse primer in the AG+G reaction, but only a single mismatch at the third base in the A+GG reaction. As another example, testing of sample #49 revealed that c25 GCA could misprime with the c25 AA+A primer set, leading to a false-positive discordant result.

When the paired CS-PCR were discordant, the sample was not typed as X4 unless the combined ΔC_t for the paired set was at least four cycles. This was done to decrease the number of false-positive reactions that might be due to alternate or less stringent primer binding with the engineered mismatch.

Comparison of CS-PCR, position-specific scored matrices (PSSM), Trofile, and ESTA

CS-PCR alone detected X4-associated sequences in 25 of the 60 samples (Table 4). Three additional samples were classified as X4 based on a genotype revealed by consensus sequencing. These were not detected by CS-PCR due to multiple mismatches or deletions within the primer binding sites. In all, 28 samples were classified as containing X4-associated sequences.

Genotypic classifications of viral samples using the 11/24/25 rule, geno2pheno, and PSSM are shown in Table 4. The presence of a basic amino acid at codons 11, 24, and/or 25 (i.e., the 11/24/25 rule)¹⁰ in the consensus sequences resulted in classification of 15 samples as X4. Twelve of the 15 codons were detected by CS-PCR. Analysis of the consensus sequences by PSSM, which does not classify sequences based strictly on the sequence at 11/24/25, classified 12 samples as X4. All of these had X4-associated codons by CS-PCR except for sample #24.

Twenty-seven of the 60 samples had phenotypic testing by Trofile (which is reported to have 85% sensitivity when X4 comprises 5% of the population using mixtures of clonal specimens¹⁸). Of the 27, only three were found to harbor X4 variants (Tables 4 and 5). Combined CS-PCR and 11/24/25 consensus sequencing typed these three and 10 of the remaining 24 as X4. Specimens from 19 subjects had phenotypic testing by the more sensitive ESTA; seven of these were found to have X4 variants. All seven plus four additional specimens

were classified as X4 by combined CS-PCR and 11/24/25 consensus sequencing (Tables 4 and 5).

CS-PCR detected X4-virus in a significantly greater number of specimens compared to consensus sequencing alone ($p=0.024$) and CS-PCR combined with consensus sequencing detected X4-virus in a significantly greater number of specimens compared to phenotypic analysis by Trofile ($n=27$; $p<0.004$). While only a small number of specimens ($n=19$) were phenotyped by ESTA, the difference in detection between CS-PCR combined with consensus sequencing versus ESTA was not significantly different ($p=0.134$) and the strength of agreement was considered moderate (kappa = 0.596; 95% confidence interval 0.274, 0.918).

With ESTA designated as the gold standard, CS-PCR combined with consensus sequencing had a concordance of 79% (15/19), a sensitivity of 100% (7/7), and a specificity of 67% (8/12).

Discussion

An economical and sensitive PCR-based assay was developed to determine whether research or clinical specimens harbor HIV-1 variants that can use the CXCR4 coreceptor. Semiquantification of the input viral target ensures that the assay is sensitive to detect minor viral variants (~1–3%). X4 genotypes are identified by combining data from a consensus sequence and CS-PCR. CS-PCR is relatively inexpensive compared to phenotypic assays, and is more sensitive compared to consensus sequencing alone.

During and following the development of CS-PCR, 60 plasma specimens from HIV-1 subtype B-infected individuals were tested. CS-PCR detected X4 variants 17% more often compared to 11/24/25 consensus sequencing alone, and in a limited comparison, 16% more often than phenotypic analysis by ESTA. The increased detection of X4 by CS-PCR compared to ESTA may be due to quantification of the input viral templates prior to CS-PCR. Input viral RNA templates are not quantified prior to testing by ESTA, and so the assay sensitivity is limited by the number of *env* cDNA that actually get screened for coreceptor usage. Recent massive parallel sequencing also detected minor X4 populations at a higher frequency compared to ESTA.^{19–21} Additionally, the presence of basic amino acids at codons 11, 24, or 25 does not guarantee that the variant will use CXCR4 efficiently enough to be detected using a culture-based assay. Given that HIV-1 utilization of CXCR4 evolves over time,^{1,9} genotypic, but not phenotypic, assays can potentially detect viruses that have begun the evolutionary process toward use of CXCR4. CS-PCR permits the detection of these variants before they become fit, allowing one to track the evolution of low-level variants while monitoring other host variables (i.e., immune system markers).

Multiple caveats inherent to PCR constrain the use of CS-PCR. Most relevant is the high degree of variability in HIV-1 *env* that can lead to false-negative and false-positive classifications. Minor viral populations with mismatches or insertions/deletions near the target codons will go undetected. X4 sequences can vary extensively from the consensus and in some cases are associated with insertions close to c11 and deletions close to c25 (personal observations). A prime example in this study was sample #5, which was classified based solely on the sequence of codon 11. An insertion beyond c14 and a

TABLE 5. CONTINGENCY TABLES COMPARING X4 AND R5 CLASSIFICATIONS BY (a) CODON-SPECIFIC-PCR AND CONSENSUS SEQUENCING, OR (b) CODON-SPECIFIC-PCR COMBINED WITH CONSENSUS SEQUENCING AND ESTA

		<i>Consensus sequencing</i>	
	(a)	R5	X4
CS-PCR	R5	32	3
	X4	45	15
		<i>ESTA</i>	
	(b)	R5	X4
CS-PCR + consensus sequence	R5	8	0
	X4	4	7

probable deletion at c25 made the consensus sequence indecipherable. Conversely, use of a low annealing temperature to amplify divergent sequences may allow nonspecific amplification, especially if the engineered primer mismatches happen to be complementary to the target HIV-1 sequence. A computer algorithm to screen the consensus sequence for bases found to cause false-positive CS-PCR can be written to simplify analysis.

At this juncture, CS-PCR is designed to detect the most common codons associated with CXCR4-utilizing HIV-1. Rare codons or changes at other associated sites, including those located outside the V3 loop, will not be detected. Adding reactions to screen for additional codons would increase the cost and complexity of the assay. However, it would be relatively easy to design CS-PCR for any target codon of interest—for example, monitoring the loss of N-linked glycosylation sites associated with coreceptor utilization.^{22,23} An additional constraint is that CS-PCR was specifically designed for subtype B. Additional primers would be needed to test other subtypes.

CS-PCR is also limited in that it cannot quantify X4 variants. Detection of an apparent low frequency of X4-associated sequence could be because the target codon is truly at a low concentration, or because of multiple mismatches between the viral sequence and the primers. Recent studies indicate that the absolute number²⁴ and/or the concentration⁷ of X4 variants appear to have clinical relevance. These reports note that individuals with low levels of X4 (<2% or <2.8 log copies X4/ml plasma) estimated by 454-pyrosequencing benefited from the addition of maraviroc, a CCR5 antagonist, to combination antiretroviral therapy. CS-PCR has not been evaluated to determine if it can predict clinical outcomes of treatment with CCR5 antagonists, although these studies are underway.

In summary, a PCR-based protocol to detect HIV-1 X4 variants was developed, and in preliminary analysis is more sensitive than consensus sequencing alone and is comparable to the phenotype assay ESTA. Importantly, the cost of performing CS-PCR is relatively modest by comparison to commercial phenotypic assays. Several potential weaknesses of CS-PCR are noted; most prominent is that sequence variation in the V3 loop of HIV-1 *env* will likely result in occasional false-negative and false-positive classifications.

Acknowledgments

The authors acknowledge and greatly appreciate Dr. Steven Douglas, Nancy Tustin, and their IMPAACT Immunology Specialty Laboratory (NIH IMPAACT Grant AI068632) for facilitating the transfer of samples and data from Children's Hospital of Philadelphia. Support for sample collection was provided by Grant P01 MH076388; Neurokinin-1R (SP Receptor) Antagonists for HIV-1 Therapy (Douglas, Steven D.); Project 4 (Tebas, Pablo); and Core B (Manak, Mark). We also want to acknowledge Joan Dragavon and Erin Goecker for facilitating the transfer of samples and data from the University of Washington CFAR Clinical Retrovirology Laboratory Core (AI-27757). We acknowledge and are exceedingly thankful to Dr. Richard Harrigan, Winnie Dong, Chanson Brumme, and Luke Swenson at the British Columbia Centre for Excellence in HIV/AIDS (Vancouver, Canada) and the Pfizer MOTIVATE team for providing the PCR products and 454-pyrosequencing data to optimize the ΔC_t cutoffs. This work was supported by a Virology Developmental award

from the International Maternal Pediatric Adolescent AIDS Clinical Trials Group (IMPAACT) (1 U01 AI068632 to LMF). The overall support for the IMPAACT was provided by the National Institute of Allergy and Infectious Diseases (NIAID) (U01 AI068632), the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), and the National Institute of Mental Health (NIMH) (AI068632). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Author Disclosure Statement

No competing financial interests.

References

- Shankarappa R, Margolick JB, Gange SJ, *et al.*: Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J Virol* 1999;73:10489–10502.
- Jensen MA, Li FS, van 't Wout AB, *et al.*: Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 *env* V3 loop sequences. *J Virol* 2003;77:13376–13388.
- Sing T, Low AJ, Beerenwinkel N, *et al.*: Predicting HIV coreceptor usage on the basis of genetic and clinical covariates. *Antivir Ther* 2007;12:1097–1106.
- Poveda E, Seclen E, Gonzalez Mdel M, *et al.*: Design and validation of new genotypic tools for easy and reliable estimation of HIV tropism before using CCR5 antagonists. *J Antimicrob Chemother* 2009;63:1006–1010.
- Resch W, Hoffman N, and Swanstrom R: Improved success of phenotype prediction of the human immunodeficiency virus type 1 from envelope variable loop 3 sequence using neural networks. *Virology* 2001;288:51–62.
- Schuurman R, Brambilla D, de Groot T, *et al.*: Underestimation of HIV type 1 drug resistance mutations: Results from the ENVA-2 genotyping proficiency program. *AIDS Res Hum Retroviruses* 2002;18:243–248.
- Swenson LC, Mo T, Dong WWY, *et al.*: Deep sequencing to infer HIV-1 co-receptor usage: Application to three clinical trials of maraviroc in treatment-experienced patients. *J Infect Dis* 2011;203:237–245.
- Swenson LC, Mo T, Dong WW, *et al.*: Deep V3 sequencing for HIV type 1 tropism in treatment-naïve patients: A re-analysis of the MERIT trial of maraviroc. *Clin Infect Dis* 2011;53:732–742.
- Coetzer M, Nedellec R, Salkowitz J, *et al.*: Evolution of CCR5 use before and during coreceptor switching. *J Virol* 2008;82:11758–11766.
- De Jong JJ, De Ronde A, Keulen W, Tersmette M, and Goudsmit J: Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: Analysis by single amino acid substitution. *J Virol* 1992;66:6777–6780.
- McLaughlin S, Hughes P, Hu S, *et al.*: Optimization of a Novel Codon-Specific PCR-Based Assay to Detect CXCR4-Utilizing HIV-1 Subtype B Variants. *CROI 2012: Abstract #712*.
- Milich L, Margolin B, and Swanstrom R: V3 loop of the human immunodeficiency virus type 1 *Env* protein: Interpreting sequence variability. *J Virol* 1993;67:5623–5634.
- Milich L, Margolin BH, and Swanstrom R: Patterns of amino acid variability in NSI-like and SI-like V3 sequences and a

- linked change in the CD4-binding domain of the HIV-1 Env protein. *Virology* 1997;239:108–118.
14. Cardozo T, Kimura T, Philpott S, *et al.*: Structural basis for coreceptor selectivity by the HIV type 1 V3 loop. *AIDS Res Hum Retroviruses* 2007;23:415–426.
 15. De Jong JJ, Goudsmit J, Keulen W, *et al.*: Human immunodeficiency virus type 1 clones chimeric for the envelope V3 domain differ in syncytium formation and replication capacity. *J Virol* 1992;66:757–765.
 16. Arora A, Kaur H, Wengel J, and Maiti S: Effect of locked nucleic acid (LNA) modification on hybridization kinetics of DNA duplex. *Nucleic Acids Symp Ser (Oxf)* 2008;52:417–418.
 17. You Y, Moreira BG, Behlke MA, and Owczarzy R: Design of LNA probes that improve mismatch discrimination. *Nucleic Acids Res* 2006;34:e60.
 18. Whitcomb JM, Huang W, Fransen S, *et al.*: Development and characterization of a novel single-cycle recombinant-virus assay to determine human immunodeficiency virus type 1 coreceptor tropism. *Antimicrob Agents Chemother* 2007;51:566–575.
 19. Daumer M, Kaiser R, Klein R, *et al.*: Inferring viral tropism from genotype with massively parallel sequencing: Qualitative and quantitative analysis. *Antiviral Ther* 2008;13:A101.
 20. Swenson LC, Moores A, Low AJ, *et al.*: Improved detection of CXCR4-using HIV by V3 genotyping: Application of population-based and “deep” sequencing to plasma RNA and proviral DNA. *J Acquir Immune Defic Syndr* 2010;54:506–510.
 21. Archer J, Braverman MS, Taillon BE, *et al.*: Detection of low-frequency pretherapy chemokine (CXC motif) receptor 4 (CXCR4)-using HIV-1 with ultra-deep pyrosequencing. *AIDS* 2009;23:1209–1218.
 22. Polzer S, Dittmar MT, Schmitz H, and Schreiber M: The N-linked glycan g15 within the V3 loop of the HIV-1 external glycoprotein gp120 affects coreceptor usage, cellular tropism, and neutralization. *Virology* 2002;304:70–80.
 23. Clevestig P, Pramanik L, Leitner T, and Ehrnst A: CCR5 use by human immunodeficiency virus type 1 is associated closely with the gp120 V3 loop N-linked glycosylation site. *J Gen Virol* 2006;87:607–612.
 24. Valdez H, Chapman D, Biswai P, *et al.*: Influence of amount and percentage of CXCR4-using virus in predicting week 48 responses to maraviroc in treatment-naive patients. *J Int AIDS Soc* 2010;13:O11.

Address correspondence to:

Sherry McLaughlin
Seattle Children's Research Institute
1900 9th Avenue
Seattle, Washington 98101

E-mail: shegato@u.washington.edu