# Determination of HLA-A, B and C by next-generation sequencing and affinity prediction with HIV-1 epitopes of patients from different regions of Brazil

**ISABEL M. PRELLWITZ[1], BRUNNA L. M. ALVES[1], MARIANNE MONTEIRO[2], SAYONARA ROCHA[2], ELIZABETH S. MACHADO[3], VANUSA P. DA HORA[4], JULIANA D. SIQUEIRA[1], HECTOR N. SEUÁNEZ[5], MARCELO A. SOARES[1], ESMERALDA A. SOARES[1]**

[1]Programa de Oncovirologia, Instituto Nacional de Câncer, Rio de Janeiro, Brazil; [2]Hospital Federal de Ipanema, Rio de Janeiro, Brazil; [3]Hospital Universitário Clementino Fraga Filho, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil; [4]Hospital Universitário Doutor Miguel Riet Corrêa Junior, Rio Grande, RS, Brazil; [5]Programa de Genética, Instituto Nacional de Câncer, Rio de Janeiro, Brazil.

## INTRODUCTION

HIV eradication is still a major challenge and a large number of people infected with HIV are under antiretroviral therapy (ART). Even under ART, HIV patients have higher risk and worse prognosis for cancer. In light of this, new strategies are emerging that are based on the therapeutic vaccination of HIV-infected individuals, eliciting immune responses that may help controlling HIV replication after discontinuation of ART.

The ability to control HIV infection has been correlated with certain human leukocyte antigen (HLA) alleles. HLA plays an important role in the cytotoxic T-cell-mediated immunity responsible for the control of HIV-1 viremia. Cellular responses are dependent on the individual's immunogenic constitution, and especially on their HLA class I allele types. The combination of the composition of the HLA alleles and HIV epitopes that are restricted by these alleles will allow proof of principle for the design of customized therapeutic vaccines against HIV.

In the present proposal, we will be able to determine the composition of the HLA class I alleles (A, B and C loci) under a new approach. Next-generation sequencing (NGS) technology applied to the determination of the HLA sequence will generate allele sequences with ultra-deep resolution, enabling the description of novel HLA alleles present in the Brazilian population.

This study is part of an international multicentric study conducted in collaboration with several researchers in Brazil, Canada and France, whose main purpose is to correlate the composition of HLA alleles to HIV epitopes that are restricted by those alleles. Those epitopes would be able to elicit effective cytotoxic responses that may help control infection after discontinuation of antiretroviral treatment.

## OBJECTIVES

The present study aims to determine the allele composition of the HLA class I A, B and C loci under a new approach by NGS that will generate allele sequences with ultra-deep resolution and predict the affinity of the most frequent haplotypes found to HIV-1 conserved epitopes. .

## METHODOLOGY

Forty four HIV-1+ adults were selected from the Hospital Federal de Ipanema and the Hospital Universitário Clementino Fraga Filho in Rio de Janeiro (RJ) and 40 individuals from the Hospital Universitário Dr. Miguel Riet Correa Jr. from Rio Grande (RG). The inclusion criteria were age equal or greater than 18 years, being under first-line HAART and being under therapeutic success (undetectable HIV viral load) for at least 12 months.

A peripheral whole blood sample was collected during follow-up in routine outpatient follow-up along with clinical and epidemiological data and the genomic DNA from the samples was extracted. For the analysis of the HLA alleles, separate amplifications of the HLA-A, B and C loci were performed which were subsequently pooled for the construction of the genomic libraries using the Nextera XT DNA Sample Preparation kit. These were sequenced in an Illumina MiSeq platform.

The HLA alleles were typed using the comemercial softwares Assign 2.0 TruSight HLA Analysis and HLA-Twin (Omixon). HLA alleles indicative of being novel are being analyzed manually using an in-house developed pipeline. The affinity between the HLA groove and the HIV-1 peptides variants was predicted using the immunogenic epitope database considering only those with IC50< 50nM, and analyzed manually using an in-house developed pipeline.

## RESULTS

Untill now, all 84 patients were collected and had the HLA-A, B and C alleles amplified. Eighty-three patients were sequenced and analyzed. Of these, 79 had at least one HLA allele successfully typed, totaling 442 typed alleles. The clinical and epidemiological data of these patients were compiled and can be seen in Table 1.

Thirteen patients presented homozygosity at least in one locus. The most frequent alleles from Rio de Janeiro were A*02:01:01 (24%), B*07:02:01 (14%), C*4:01:01 (13%) and from Rio Grande do Sul were A*01:01:01:01 (20%), B*08:01:01:01 (9%) and C*03:03:01:01 (8%) (Figure 1). Three patients from RJ had the B*27:05:02 allele that is associated with the control of HIV infection and four patients from RJ and five patients from RS have the B*35:01:01 or B*35:03:01 alleles that are associated with faster progression to disease. Nine (2%) alleles presented evidence of being novel and need to be confirmed.

Two epitopes, KARVLAEAM (nts 1410-1436, HXB2 reference sequence) and EMMTACQGV (nts 1368-1395, HXB2) of the Gag protein were present in the 19 HIV sequences analyzed from Rio de Janeiro and had high affinity for the HLA-A, B and C alleles of these patients (Figure 2). These epitopes were analyzed for intrapatient variation and only five patients presented variant sequences with frequency >1% for the epitope KARVLAEAM and four patients for the epitope EMMTACQGV. Interestingly, these variants also had high affinity for the respective HLA-A, B and C alleles (Figure 3).

## PERSPECTIVES

- The possible novel alleles will be confirmed by other methodologies and submitted to the IMGT-HLA database.
- Identification of epitopes in other regions of HIV: Pol, Env, Vif and Nef.
- The combination of the composition of the HLA alleles and the affinity to HIV-1 epitopes that are restricted intrapatient will allow proof of principle for the design of custom therapeutic vaccines against HIV.

**Table 1.** Epidemiological profile of patients.

| | Hospital Federal de Ipanema | Hospital Universitário Clementino Fraga Filho – UFRJ | Hospital Universitário Dr. Miguel Riet Corrêa Jr. - FURG |
|---|---|---|---|
| Number of patients (n) | 32 | 12 | 40 |
| Median age (years) | 38 | 43,5 | 43 |
| Male | 24 (75%) | 8 (67%) | 15 (37%) |
| Median time of diagnosis at the visit (months) | 56.5 | 159,5 | 56 |
| Median CD4 count at collection (cel/mm3) | 712.5 | 813,5 | 780,5 |
| Median CD8 count at collection (cel/mm3) | 657.5 | 930 | 862 |
| Median time to initiation of treatment (months) | 14 | 20,5 | 10,5 |
| Median time of treatment (months) | 37.5 | 112 | 43 |
| Sexual transmission Men who have sex with men (MSM) | 7 (22%) | 1 (8%) | 2 (5%) |
| Sexual transmission Non-MSM | 4 (13%) | 7 (59%) | 2 (5%) |
| Drug user | 0 | 0 | 2 (5%) |

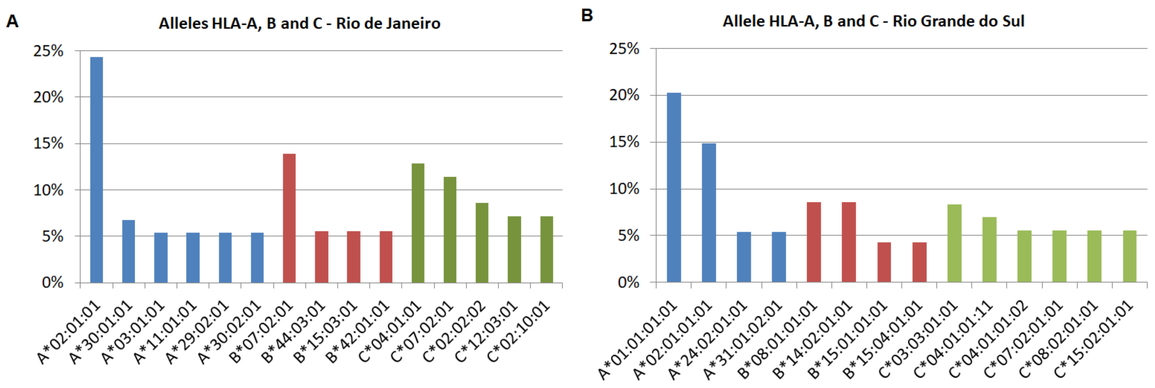


**Figure 1.** Most common HLA-A, B and C alleles found. A: patients from Rio de Janeiro. HLA-A n=74, HLA-B n=72, HLA-C n=70. B: patients from Rio Grande do Sul. HLA-A n=74, HLA-B n=70, HLA-C n=72

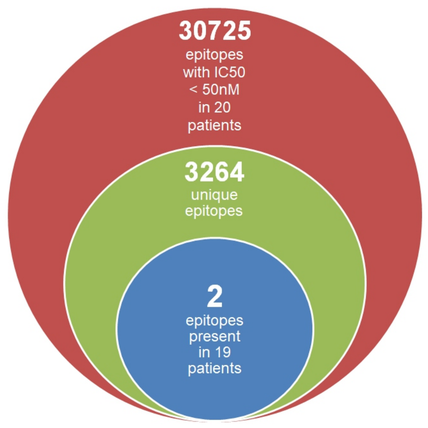


**Figure 2.** Flowchart of the epitope identification analysis from the HIV Gag protein with high affinity (IC50 < 50nM) to the two most common HLA-A, B and C alleles among patients from Rio de Janeiro (n = 20).

**A1**

| variant | HLA-A | | HLA-B | | HLA-C | |
|---|---|---|---|---|---|---|
| HIV11 | KIITACQGV | A*02:01 | A*24:03 | B*41:02 | B*44:03 | C*07:06 | C*17:03 |
| HIV19 | KMMTACQGV | A*24:03 | A*30:02 | x | x | x | x |
| HIV23 | EMMTACQKV | A*30:02 | A*31:01 | B*15:03 | B*15:17 | C*02:10 | C*05:01 |
| HIV32 | EMMTACRGV | A*25:01 | A*29:02 | B*18:01 | B*56:42 | x | x |

**A2**


**B1**

| variant | HLA-A | | HLA-B | | HLA-C | |
|---|---|---|---|---|---|---|
| HIV11 | KARVLAKAM | A*02:01 | A*24:03 | B*41:02 | B*44:03 | C*07:06 | C*17:03 |
| HIV12 | EARVLAEAM | A*68:02 | A*01:01 | B*39:13 | B*44:03 | C*03:04 | C*06:02 |
| HIV13 | KARVLAEAV | A*68:01 | A*01:01 | B*81:01 | B*81:01 | C*07:01 | C*08:04 |
| HIV23 | KAKVLAEAM | A*30:02 | A*31:01 | B*15:03 | B*15:17 | C*02:10 | C*05:01 |
| HIV28 | KVRVLAEAM | A*24:05 | A*01:01 | B*15:17 | B*39:01 | C*07:01 | C*12:03 |

**B2**


**Figure 3.** Epitope variants with frequency >1% in the viral population. A1 and B1: Table with variants for epitopes EMMTACQGV and KARVLAEAM respectively, and the HLA-A, B and C alleles for each patient. In green, the alleles to which the variant showed high affinity (IC50 < 50nM) and in red, the alleles to which the variant did not show high affinity (IC50 < 50nM). A2 and B2: Representation (patient HIV23) of amino acid conservation in the epitopes EMMTACQGV and KARVLAEAM, respectively.

SUPPORT

FAPERJ · inct · CNPq · HUCFF · Rede Hospitalar Federal no Rio de Janeiro · Université Bordeaux 1 Sciences Technologies

Projeto Gráfico: Área de Edição e Produção de Materiais Técnico-Científicos / INCA

INCA · SUS · MINISTÉRIO DA SAÚDE