NEXT GENERATION SEQUENCING IN REVERSE TRANSCRIPTASE AND PROTEASE OF HIV-1 HELPS TO DISCRIMINATE RECENT FROM CHRONICALLY INFECTED NEWLY DIAGNOSED PATIENTS

Jose Ángel Fernández-Caballero
• Primary HIV infection is the first period of infection, and takes from the transmission event to seroconversion.
• Despite the high number of new cases of HIV infection, primary infection is diagnosed in no more than 12%.
• Some algorithms or lab test (STARHS) have been proposed to differentiate acute from chronic infection, but they are not highly reliable and there is no gold standard.
Shortly after transmission, HIV population is highly homogeneous and progresses parallel to the viral exponential replication phase, deriving with time to different quasispecies and a greater viral diversity.

Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. Jesus F. Salazar-Gonzalez et al. JEM. 2017
• The correctly characterize primary infected patients gains relevance both for clinical assays and for the development of vaccines.
  ➔ stratification of patients in study.
  ➔ effective vaccine: early founder virus.
• Some studies have addressed the diversity of HIV population in the course of HIV infection:
  – Cornelissen et al (2007) studied how the ambiguous bases contained in pol HIV sequence were associated with the diversity of viral population.
  – Hightower et al (2012) found that this ambiguous method was similar to Shannon Entropy as a measure of diverse population in pol sequence.
Baseline routine resistance testing of the RT & protease regions may allow to estimate the timing of infection, if the diversity of the viral population that is infecting the patient may be estimated.
• To evaluate if baseline routine resistance testing of RT and Protease, produced using Next Generation Sequencing may be useful to characterize the phase of HIV infection.
Plasma samples from newly diagnosed patients sent to our laboratory for baseline resistance testing from October 2015 (when Trugene was discontinued) to January 2017 (when we moved from Roche 454 to Illumina) were analysed.

We defined primary infection as having at least one positive test, a negative Ab test in the prior 3 months with a positive one at the time of testing, and/or clinical symptoms of HIV acute infection.

We used NGS (454 GS Junior) for RT & Pro resistance testing, analysing 4 amplicons, that were obtained with AVA software and were filtered by quality values (Q>30) and length using *Usearch*. 

Blood plasma → NGS → Filtering → Phylogenetic tree → Consensus sequence → Interpretation of resistances → Epidemiological study

We used Mesquite software to build the consensus sequence of the pol gene, using a threshold value of 20% for the mixed bases count, being this an indirect measure of diversity. We chose the 20% threshold following our previous study that showed that these sequences were equivalent to Sanger sequences.

3) Muscle software was used to align the sequences and build phylogenetic trees using NJ.

4) To assess viral variability we evaluated the relative proportion of the major sequence, expressed as the mean value of the among the four amplicons analysed.
• We studied 194 patients, 159 with chronic infection and 35 recently infected.
• As it can be seen in the table, both groups were similar in terms of age, gender, origin and subtype distribution. As expected, viral load and CD4 were significantly different in both groups (p<0.001).
• Using the Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance list (Bennett et al. Plos one. 2009), TDR was found in 4 patients in the chronically infected patients and 2 in the recently infected, The details of the mutations can be observed in the figure below.
We found a greater differentiation of viral variants in patients with chronic infection. Viruses also show a higher genetic distance, reflecting viral evolution within the host.

The graphs below also show how each variant is accumulating random mutations over time of infection, resulting in greater diversity.
• Phylogenetic studies also generated a major abundance and population diversity in chronically infected patients, being viral populations more homogeneous in the primary infection group.
• As seen in the graphs below, the sequences of the patients at the early stage of infection are very similar, and with no marked genetic distance.
Here we present the relative proportion of the major sequence for both groups. As it can be seen, the major sequence represented a mean of 96.3% (IQR 95.7-98.25) of the viruses infecting patients with recent infection, and only 54.3% (28.6-71.8) of the chronically infected group (p<0.001).

This differences were also true for each of the four amplicons we generated for RT & Pro resistance studies.
As NGS may not be available in all laboratories, we evaluated the number of ambiguous bases in the 20% threshold sequences.

As expected, a higher number of ambiguous bases was observed for chronically infected patients mean value, [15 (IQR 12-21)], than in the recently infected ones [2 (IQR 1-3)], reaching statistical differences (p<0.001).

As shown in the graphs, patients with early infection never had more than 5 ambiguous bases, while chronically infected always had more than 9, peaking at 10 ambiguous bases.
Viral diversification may be identified with appropriate analysis from baseline routine testing performed with Next Generation Sequencing tests, as well as by studying the number of ambiguous bases from Sanger based sequences.

Being conservative, we propose a cutoff of 95% viral diversity and 10 ambiguous bases to discriminate chronic from recent infection in newly diagnosed patients.

These cutoffs need to be externally validated.
Acknowledgments

- Patients
- Hospital San Cecilio Research team
- Financial support
- Spanish AIDS Research Network (RIS)
- Your attention

This study has been financed by the project RD16/0025/0040, integrated into the Plan National of I+D+I and co-financed by the ISCIII-Subdirección General de Evaluación and the European regional development fund (FEDER)