PROTEOMIC STUDY OF HIV-POSITIVE IMMUNOLOGICAL NON-RESPONDER PATIENTS

Marina Flores Piñas

Final degree project

Tutor: Anna Rull Aixa

Biochemistry and molecular biology

UNIVERSITAT ROVIRA i VIRGILI

Tarragona, 2022

This Final degree project has been based on the results obtained in my external internship at the Infection and Immunity group (INIM) of the Institut d'Investigació Sanitària Pere Virgili (IISPV) under the supervision of Anna Rull Aixa.

INDEX

1. ABSTRACT

A proteomic study was designed to identify protein patterns associated with poor immune response to antiretroviral treatment (ART) administered to HIV-positive individuals.

The study of T-cells proteomic profile from seropositive patients before the start of ART revealed significant differences in both cell subtypes between control subjects, subjects initiating ART with low immune status but good immune recovery on ART and those patients with poor immune response on ART.

The significant proteins identified should be further studied as possible biomarkers for the early identification of immunological non-responder patients or as therapeutic targets for the improvement of immune recovery status to treatment.

2. INTRODUCTION

2.1. Human Immunodeficiency Virus

Three decades have passed since the first identification, in 1981, of previously healthy patients suffering from severe immune deficiency and dying from opportunistic infections – the syndrome that we now know as acquired immunodeficiency syndrome (AIDS), and which in 1983 was recognized to be the result of infection with Human Immunodeficiency Virus (HIV)[1]. Almost 30 years later, the HIV epidemic has spread throughout the world with more than 50 million people infected of which 25 million have died[2]. HIV continues to be an enormous global health challenge with immense social and economic consequences[3].

Human Immunodeficiency Virus (HIV) is a genetically related member of the *Lentivirus* genus of the *Retroviridae* family[4]. Based on genetic characteristics and differences in the viral antigens, HIV is classified into types 1 and 2 (HIV-1, HIV-2)[5]. HIV-1 has its origin from the SIV of the chimpanzee, whereas HIV-2 originated from the SIV of the sooty mangabey[6]. It is well established that blood plasma viral load is lower in HIV-2 compared with HIV-1 infection. It would therefore make sense that viral replication could largely explain the difference in pathogenicity between the two viruses[7]. HIV-1 is the cause of the global AIDS pandemic while HIV-2, although it can also cause AIDS, is considered less pathogenic and less transmissible[8].

2.2. Genome and virion structure

HIV-1, like other retroviruses, is made up of an RNA encoded genome of approximately 9.7 kilobases (kb). Both ends of the RNA genome are flanked by a long terminal repeat (LTR) promoter region[9] (*Figure 1A*).

The 5' LTR region codes for the promotor for transcription of the viral genes. In the direction 5' to 3' the reading frame of the gag gene follows, encoding the proteins of the outer core membrane (MA, p17), the capsid protein (CA, p24), the nucleocapsid (NC, p7) and a smaller, nucleic acid-stabilizing protein (TF, p6).
The gag reading frame is followed by the pol reading frame coding for the enzymes protease (PR, p12), reverse transcriptase (RT, p51) and RNase H (p15) or RT plus RNase H (together p66) and integrase (IN, p32). Adjacent to the pol gene, the env reading frame follows from which the two envelope glycoproteins gp120 (surface protein, SU) and gp41 (transmembrane protein, TM) are derived. In addition to the structural proteins, the HIV genome codes for several regulatory proteins: Tat (transactivator protein) and Rev (RNA splicing-regulator) are necessary for the initiation of HIV replication, while the other regulatory proteins Nef (negative regulating factor), Vif (viral infectivity factor), Vpr (virus protein r) and Vpu (virus protein unique) have an impact on viral replication, virus budding and pathogenesis[10,11].
HIV-2 codes for Vpx (virus protein x) instead of Vpu, which is partially responsible for the reduced pathogenicity of HIV-2[12]. The genome structure of the immunodeficiency viruses of chimpanzees (SIVcpz) and gorillas (SIVgor) is identical to that of HIV-1[13].

The mature HIV particle is round and measures approximately 100 nm in diameter, it has an outer lipid membrane as its envelope (*Figure 1B*).

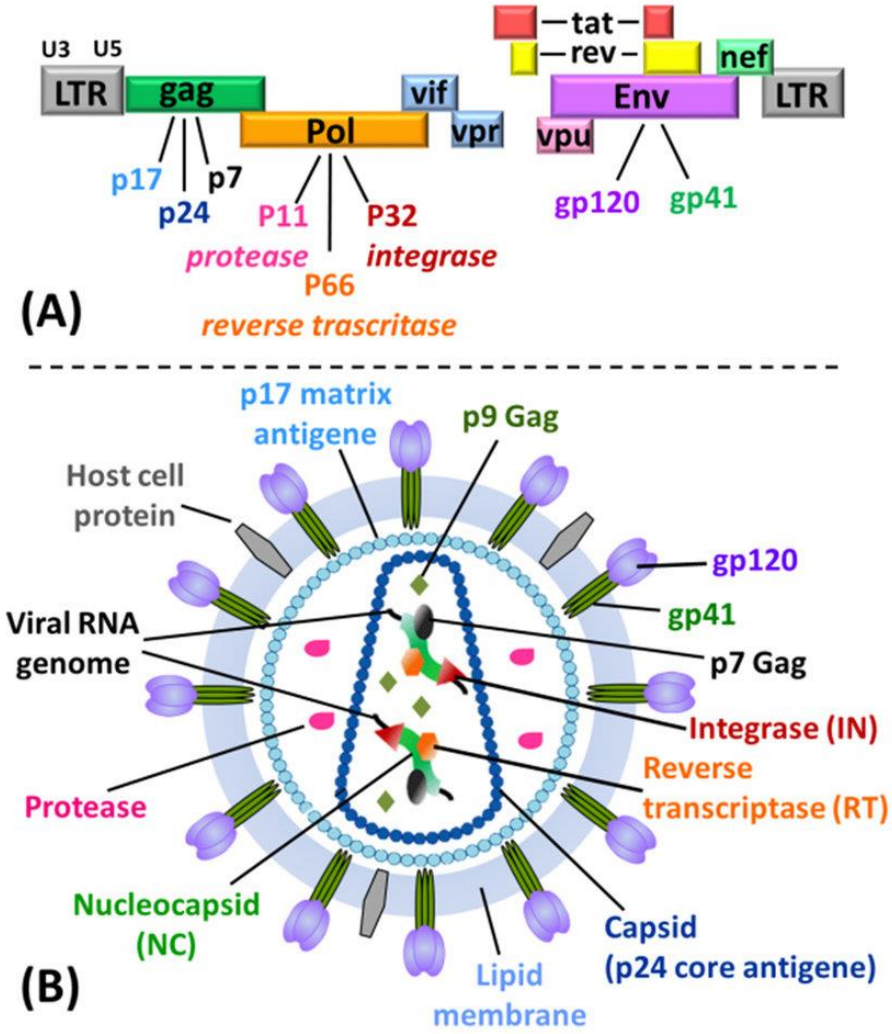

Figure 1: (A) Schematic diagram of the HIV-1 genome; (B) structure of an HIV-1 virion particle with an indication of the potential antiviral targets[14].

The envelope contains 72 knobs, composed of trimers of the Env proteins. The trimers of gp120 surface protein are anchored to the membrane by the trimers of the transmembrane protein gp41[15]. The viral envelope is composed of a lipid bilayer and, in mature virus particles, the envelope proteins gp120 and gp41. It covers the symmetrical outer capsid membrane which is formed by the matrix protein (p17).

The conical capsid is assembled from the inner capsid protein p24[16]. The tapered pole of the capsid is attached to the outer capsid membrane. Two identical molecules of viral genomic RNA are located inside the capsid and several molecules of the viral enzymes RT/RNase H and IN are bounded to the nucleic acid[17].

The VIH-1 genes with the different proteins that encode and their functions are shown in *Table 1*.

Gene	Proteins Encoded by Gene	Function of Proteins
I. Structural genes found in all retroviruses		
gag	p24, p7	Nucleocapsid
	p17	Matrix
pol	Reverse transcriptase[1]	Transcribes RNA genome into DNA
	Protease	Cleaves precursor polypeptides
	Integrase	Integrates viral DNA into host cell DNA
env	Gp120	Attachment to CD4 protein
	Gp41	Fusion with host cell
II. Regulatory genes found in human immunodeficiency virus that are required for replication		
tat	Tat	Activation of transcription of viral genes
rev	Rev	Transport of late mRNAs from nucleus to cytoplasm
III. Regulatory genes found in human immunodeficiency virus that are not required for replication (accessory genes)		
nef	Nef	Decreases CD4 proteins and class I MHC proteins on surface of infected cells; induces death of uninfected cytotoxic T cells; important for pathogenesis by SIV[2]
vif	Vif	Enhances infectivity by inhibiting the action of APOBEC3G (an enzyme that causes hypermutation in retroviral DNA)
vpr	Vpr	Transports viral core from cytoplasm into nucleus in nondividing cells
vpu	Vpu	Enhances virion release from cell

MHC = major histocompatibility complex.
[1]Reverse transcriptase also contains ribonuclease H activity, which degrades the genome RNA to allow the second strand of DNA to be made.
[2]Mutants of the *nef* gene of simian immunodeficiency virus (SIV) do not cause acquired immunodeficiency syndrome in monkeys.

Table 1: VIH-1 genes and function of the proteins encoded by the genes[17].

2.3. HIV replication cycle

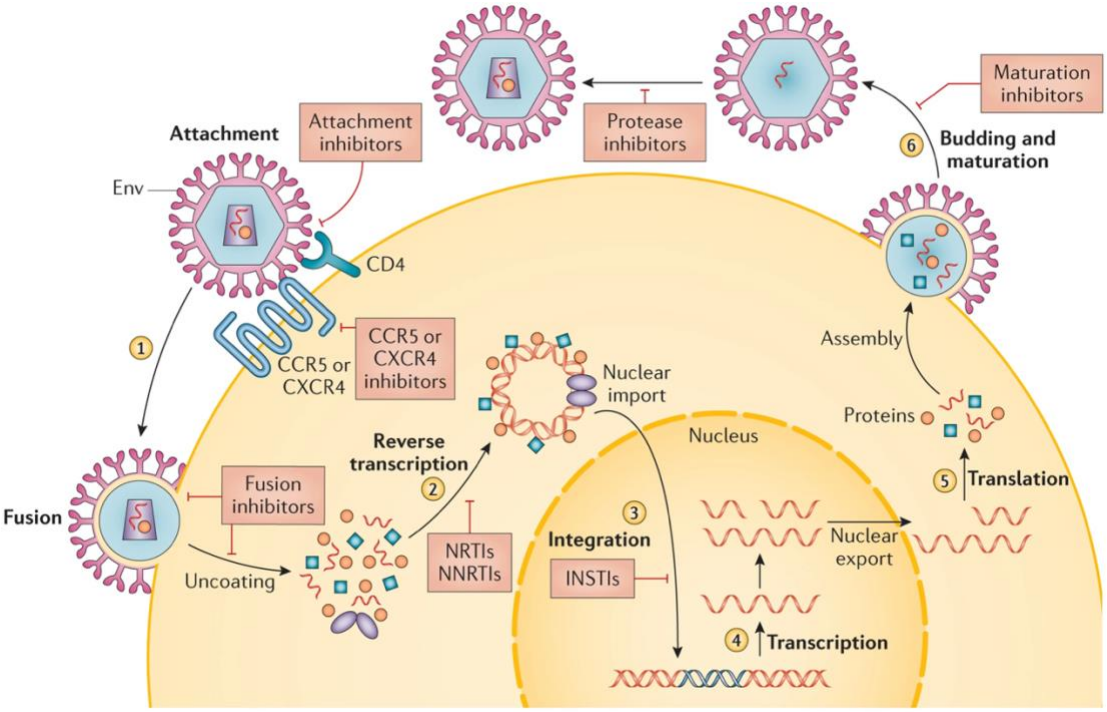

Figure 2: Representation of HIV replication cycle[19].

Viral membrane fusion is the first key step for enveloped viruses to enter host cells and establish infection[18]. The primary receptor for HIV-1 is $CD4^+$, which is expressed on the surface of T lymphocytes, monocytes, macrophages and dendritic cells. HIV also requires a co-receptor to gain entry into the host cell, typically the chemokine receptors CCR5

and CXCR4. Different HIV-1 variants typically use one or the other chemokine receptor, but some can use either; viruses that use these co-receptors for entry are called R5, X4 or R5X4 viruses, respectively. CCR5 and CXCR4 are differentially expressed on some T cell subsets, with CCR5 expressed at high levels in memory T lymphocytes but not on naive T lymphocytes, whereas CXCR4 is expressed on both. CCR5 is also expressed on macrophages and dendritic cells[19]. The receptor and the coreceptors of HIV-1 (chemokine receptors) interact with HIV-1 through interaction with envelope glycoprotein (Env), this conforms to the first step of the infection.

Fusion allows HIV to enter the $CD4^+$ T-cell. Once inside a CD4 cell, HIV releases and uses reverse transcriptase to convert its genetic material—HIV RNA—into HIV DNA. The conversion of HIV RNA to HIV DNA allows HIV to enter the $CD4^+$ T-cell nucleus and combine with the cell's genetic material—cell DNA. At this point, inside the host $CD4^+$ T-cell nucleus, HIV releases integrase, an HIV enzyme. HIV uses integrase to insert its viral DNA into the DNA of the host cell[20]. When HIV is integrated into the host $CD4^+$ T-cell DNA, mediated by host enzymes, HIV DNA is transcribed to viral mRNAs. These mRNAs are then exported to the cytoplasm where translation occurs to make viral proteins[19].

The next step in the replication cycle is assembly. During assembly, new HIV RNA and HIV proteins made by the host $CD4^+$ T-cell move to the surface of the cell and assemble into immature (noninfectious) HIV. This immature particle migrates towards the cell surface. The large precursor molecules are then cleaved by the HIV-1 protease, resulting in new infectious viral particles (*Figure 3*), which bud through the host cell membrane, thus acquiring a new envelope[20]. During the budding process, the virus lipid membranes may incorporate various host cell proteins and become enriched with phospholipids and cholesterol. Differently from T-lymphocytes, where budding occurs at the cell surface and virions are released into the extracellular space, the budding process in monocytes and macrophages results in the accumulation of virions within intracellular vacuoles which are then released[21].

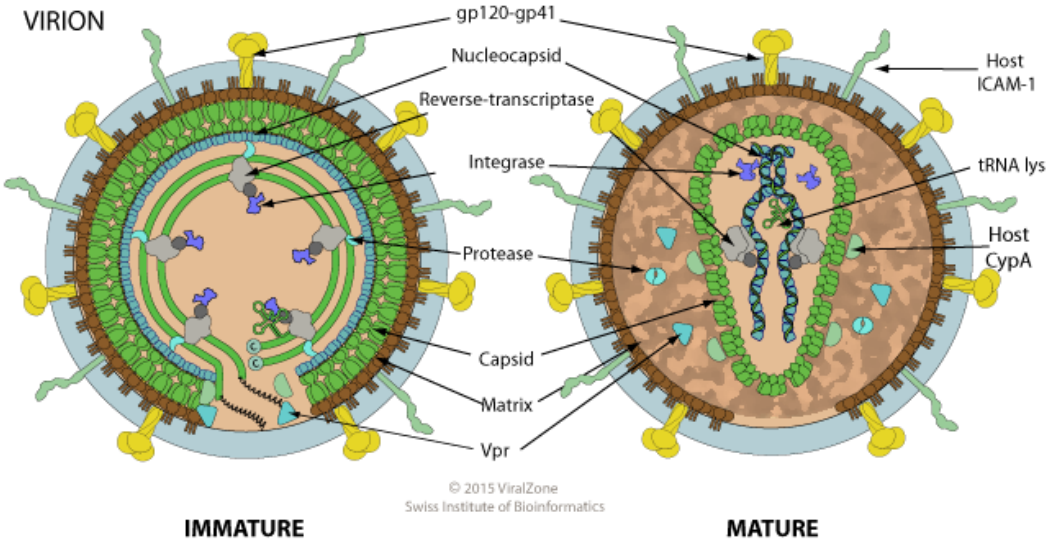

Figure 3: Immature and mature virion structure[22].

2.4. Course of infection

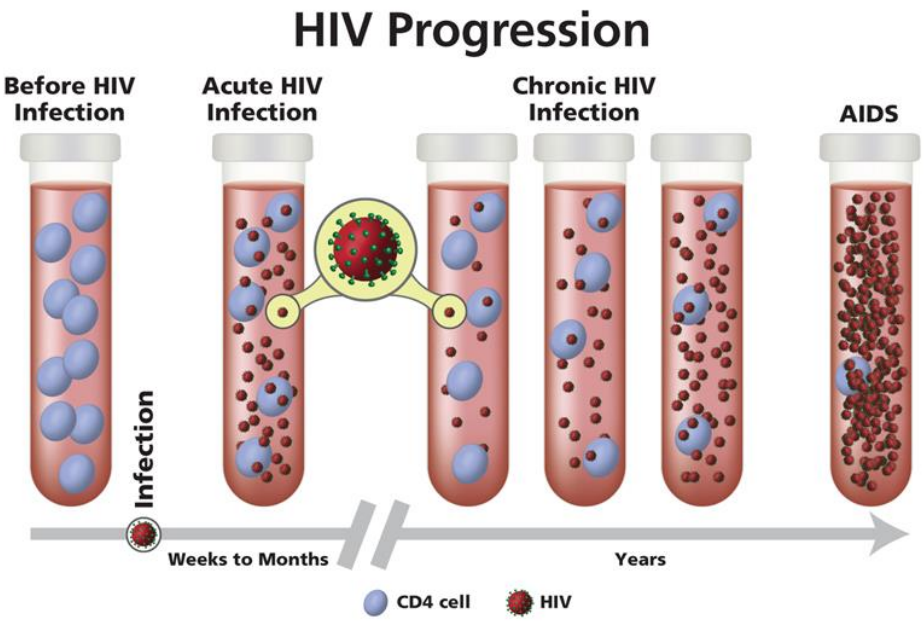

Figure 4: HIV course of infection[23].

Without treatment, HIV infection advances in stages, getting worse over time. HIV gradually destroys the immune system and eventually causes acquired immunodeficiency syndrome (AIDS). There is no cure for HIV, but treatment with HIV medicines (called antiretroviral therapy or ART) can slow or prevent HIV from advancing from one stage to the next (*Figure 4*). There are three stages of HIV infection[23]:

- **Acute HIV Infection**
  Acute HIV infection is the earliest stage of HIV infection, and it generally develops within 2 to 4 weeks after infection with HIV. During this time, some people have flu-like symptoms, such as fever, headache, and rash. In the acute stage of infection, HIV multiplies rapidly and spreads throughout the body. The virus attacks and destroys the infection-fighting $CD4^+$ T-cells ($CD4^+$ T lymphocytes) of the immune system. During the acute HIV infection stage, the level of HIV in the blood is very high, which greatly increases the risk of HIV transmission. A person may experience significant health benefits if they start ART during this stage.

- **Chronic HIV Infection**
  The second stage of HIV infection is chronic HIV infection (also called asymptomatic HIV infection or clinical latency). During this stage, HIV continues to multiply in the body but at very low levels. People with chronic HIV infection may not have any HIV-related symptoms. Without ART, chronic HIV infection usually advances to AIDS in 10 years or longer, though in some people it may advance faster. People who are taking ART may be in this stage for several decades. While it is still possible to transmit HIV to others during this stage, people who take ART exactly as prescribed and maintain an undetectable viral load have effectively no risk of transmitting HIV to an HIV-negative partner through sex.

- **AIDS**
  AIDS is the final, most severe stage of HIV infection. Because HIV has severely damaged the immune system, the body cannot fight off opportunistic infections. Opportunistic infections are infections and infection-related cancers that occur more frequently or are more severe in people with weakened immune systems than in people with healthy immune systems. People with HIV are diagnosed with AIDS if they have a $CD4^+$ count of less than 200 cells/mm3 or if they have certain opportunistic infections. Once a person is diagnosed with AIDS, they can have a high viral load and can transmit HIV to others very easily. Without treatment, people with AIDS typically survive about 3 years.

In figure 5, there is a graph of the relationship between HIV copies (viral load) and CD4 counts over the average course of untreated HIV infection.

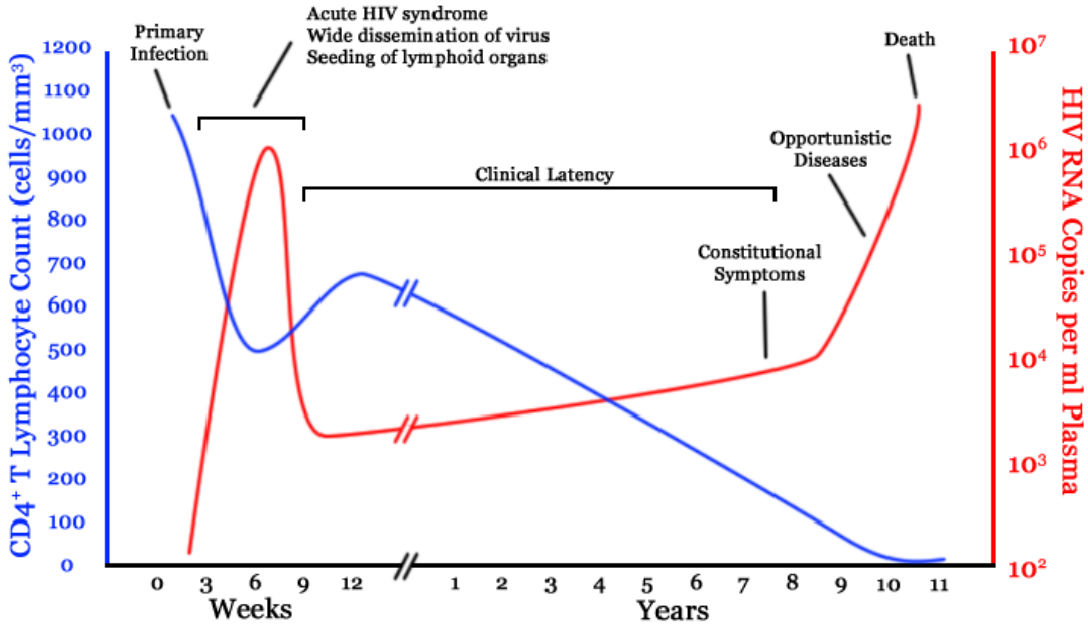

Figure 5: Viral load and $CD4^+$ changes during the course of the infection without ART[24].

2.5. Antiretroviral therapy

The treatment for HIV is called antiretroviral therapy (ART). ART involves taking a combination of HIV medicines every day[25]. Rapid initiation of antiretroviral therapy (ART) – as early as the day of HIV diagnosis – may be important in global HIV management for two main reasons. First, as means to control the HIV epidemic, in the absence of a vaccine or cure: undetectable virus means untransmissible virus[26,27]. Secondly, to optimize the health of people living with HIV (PLWH)[28].
The main goal of HIV treatment is to reduce a person's viral load to an undetectable level. An undetectable viral load means that the level of HIV in the blood is too low to be detected by a viral load test. People with HIV who maintain an undetectable viral load have effectively no risk of transmitting HIV to their HIV-negative partners through sex.

There are many HIV medicines available for HIV treatment regimens. The HIV medicines are grouped into seven drug classes according to how they fight HIV.
The choice of an HIV treatment regimen depends on a person's individual needs. When choosing an HIV treatment regimen, people with HIV and their health care providers consider many factors, including possible side effects of HIV medicines and potential drug interactions[25].

- **Nucleotide/nucleoside reverse transcriptase inhibitors (NRTI)**
  Reverse transcriptase inhibition has been a primary target in most treatment courses. These therapeutic agents are analogues of, and compete with, natural nucleotides used in DNA synthesis that lack the 3'-hydroxyl group. Once the reverse transcriptase incorporates an analogue instead of a natural nucleotide in the vDNA the process is halted because the next nucleotide cannot bind in the absence of the 3'-hydroxyl group.

- **Non-nucleotide reverse transcriptase inhibitors (NNRTI)**
  Non-nucleotide reverse transcriptase inhibitors inhibit retrotranscriptase (RT) by binding and forming a hydrophobic region close to the active site. The binding of non-nucleotide RT inhibitors changes the spatial conformation of the substrate-binding site and reduces polymerase activity. Compared to nucleotide/nucleoside reverse transcriptase inhibitors, non-nucleotide RT inhibitors do not inhibit RT of other lentiviruses such as HIV-2 and SIV (simian immunodeficiency virus)

- **Integrase inhibitors**
  Integrase inhibitors catalyze the 3' end of the viral DNA strand. All integrase inhibitors target the DNA chain transfer reaction and are also called InSTI (Integrase Strand Transfer Inhibitors). Integrase inhibitors bind to a specific complex between integrase and viral DNA and interact with two Mg2+ ion cofactors at the active site of integrase and DNA[29].

- **Protease Inhibitors (PI)**
  Protease is the enzyme responsible for the cleavage of Gag and Gag-Pol viral polyprotein precursors during virion maturation. Ten protease inhibitors are currently approved. All PIs share relatively similar chemical structures and cross-resistance is commonly observed[30].

- **Entry inhibitors**
  Entry inhibitors are classified into two distinct classes (fusion inhibitors and CCR5 antagonists) depending on the disruption or inhibition of the HIV-1 virus penetration process into the host cell.

  - **Fusion inhibitors**
    Fusion inhibitors are a relatively new class of antiretroviral drugs and include virus binding to co-receptors inhibitors - CCR5 co-receptor blockers and host cell fusion inhibitors, gp41 blockers, HIV fusion factor.

- **CCR5 antagonists**
  CCR5 antagonists bind to the hydrophobic regions within the transmembrane spirals of CCR5. This site does not overlap with the binding sites of CCR5 agonists or the HIV-1 capsule. Instead, the binding of the drug induces and stabilizes the conformation of the receptor that is not recognized by any protein. Thus, these molecules are considered allosteric inhibitors.

- **Cytochrome inhibitors**
  Represents a class of drugs whose mechanism of action is inhibition of cytochrome P450, subtype CYP3A4. They do not have an antiviral effect but increase the blood concentration of antiretroviral drugs with hepatic clearance. The class has two approved representatives for human use[29].

2.6. Diagnosis and prevention

HIV is transmitted through contact of infected body fluids with mucosal tissue, blood or broken skin[19]. Factors that increase the infectiousness of a person infected with HIV include higher levels of virus in plasma[31] or genital secretions[32], and other sexually transmitted infections[33]. Virus characteristics, such as higher envelope content, increased cell-free infectivity, increased interaction with dendritic cells and resistance to IFNα have been associated with increased infectivity[34].

Diagnosis
HIV testing algorithms generally have changed over time, as test accuracy has increased. Current US Centers for Disease Control and Prevention[35] and European guidelines for HIV testing[36] recommend that screening be performed with an antigen–antibody assay (these assays are considered to be the fourth generation). Positive results should be confirmed with an antibody assay that can differentiate between HIV-1 and HIV-2 infections[19].
Detection of acute infection is important to prevent onward HIV transmission. Acute infection is thought to contribute disproportionately to new infections[37,38]: reasons include higher viral levels during this period, higher infectiousness of the recently transmitted virus[39] and continued behavioural risk-taking by individuals unaware of their recently acquired HIV infection[40]. Immediate initiation of ART will reduce the symptoms of acute HIV infection, potentially reduce seeding of viral reservoirs, and maintain the health of the newly infected individual while reducing the risk of transmission to uninfected partners[19].

HIV prevention
Condom use in men has been a cornerstone of HIV prevention, as perfect use should completely prevent HIV transmission, as well as transmission of many other sexually transmitted infections[19]. However, condom effectiveness has been estimated to be approximately 80% against heterosexual transmission of HIV infection[41] and 70% against male-to-male sexual transmission[42]. Over-reporting of condom use probably contributes to these lower than expected effectiveness estimates, although improper use and condom failure also play a part[43]. Similarly, providing clean injection equipment can substantially reduce HIV transmission in injection drug users[44], although syringe

exchange programmes have not eliminated HIV transmission in drug users. Additional prevention tools are needed to augment these core strategies[19].

2.7. Immunological non-responders

The hallmark of HIV infection is the persistent destruction of $CD4^+$ T-cells, resulting in progressive immunodeficiency, opportunistic diseases, and death[45]. The increasing accessibility and use of antiretroviral therapy (ART) can suppress the HIV viral load to undetectable levels and increase the $CD4^+$ T-cell counts; therefore, the acquired immunodeficiency syndrome (AIDS)-related morbidity and mortality in HIV-1-infected individuals is sharply diminished[46–48]. However, in some patients, optimal treatment and persistent suppression of viral replication fail to restore their $CD4^+$ T-cell counts. These patients are referred to as inadequate immunological responders, immunodiscordant responders, or immunological non-responders (INRs), and an impaired immunological response is linked to an increased risk of disease progression and death for these patients[49–51].

INRs present severe immune dysfunction, and the morbidity and mortality of AIDS and non-AIDS events (such as metabolic syndrome, liver disease, nephropathy, cardiovascular disease, non-AIDS-related malignancies, and HIV-1-related neurocognitive disorder) were significantly elevated compared with those for HIV-1-infected patients who achieved complete immune reconstitution[52–57]. The duration of ART can significantly affect the magnitude of immune reconstitution in HIV-1-infected patients, thus hindering the comparison of different findings.

To date, it is commonly believed that the $CD4^+$ T-cell count is the most important predictor of immune recovery, treatment outcome, and disease progression in HIV-1 infection, but recent reports indicate the need for additional markers to supplement the $CD4^+$ T-cell count[58]. Compared with the $CD4^+$ T-cell count and viral load, the $CD4^+/CD8^+$ ratio is the potential for higher predictive and evaluative value for the recovery of immunological function, especially in patients who reached a $CD4^+$ T-cell count > 500/μl after initiation of ART[59–61].

Potential mechanisms of incomplete immune reconstitution

The underlying mechanisms of incomplete immune reconstitution are very complicated and may be multifactorial, including decreased hematopoiesis of bone marrow, insufficient thymic output, residual virus replication, aberrant immune activation, perturbations of cytokine secretion, and specific genetic or metabolic characteristics *(Figure 6)*[62–65]. However, none of these independent factors can fully explain the mechanism of incomplete immune reconstitution. At any time, the $CD4^+$ T-cell counts in HIV-1-infected individuals are associated with the production, destruction, and migration between secondary lymphoid organs and peripheral tissues[50,66]. INRs may have both reduced $CD4^+$ production and excessive destruction[58].

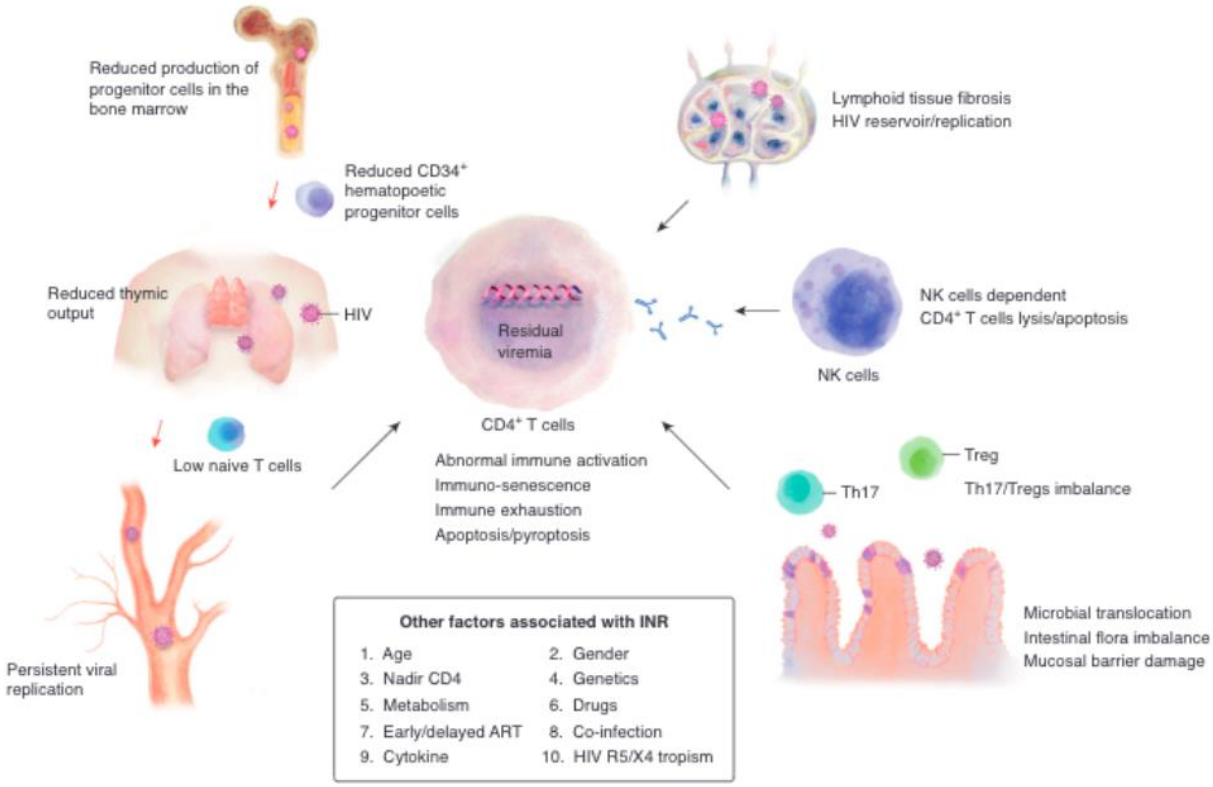

Figure 6: Factors associated with immunological non-responders[58].

2.8. Omics sciences: Proteomics

Since the process of mapping and sequencing the human genome began, new technologies have made it possible to obtain a huge number of molecular measurements within a tissue or cell. These technologies can be applied to a biological system of interest to obtain a snapshot of the underlying biology at a resolution that has never before been possible. Broadly speaking, the scientific fields associated with measuring such biological molecules in a high-throughput way are called “omics”[67]. The word omics refers to a field of study in biological sciences that ends with -omics (*Figure 7*). The ending -ome is used to address the objects of study in such fields, such as the genome, proteome, transcriptome, or metabolome, respectively[68].

Proteins are responsible for an endless number of tasks within the cell. The complete set of proteins in a cell can be referred to as its proteome and the study of protein structure and function and what every protein in the cell is doing is known as proteomics[69]. Proteomics is the science that studies those proteins as related to their biochemical properties and functional roles, and how their quantities, modifications, and structures change during growth and in response to internal and external stimuli[68].

Proteomics is crucial for early disease diagnosis, prognosis and monitoring of the disease development. Furthermore, it also has a vital role in drug development as a target molecule. So, in summary, proteomics can be defined as the characterization of the proteome, including expression, structure, functions, interactions and modifications of proteins at any stage[70].

.

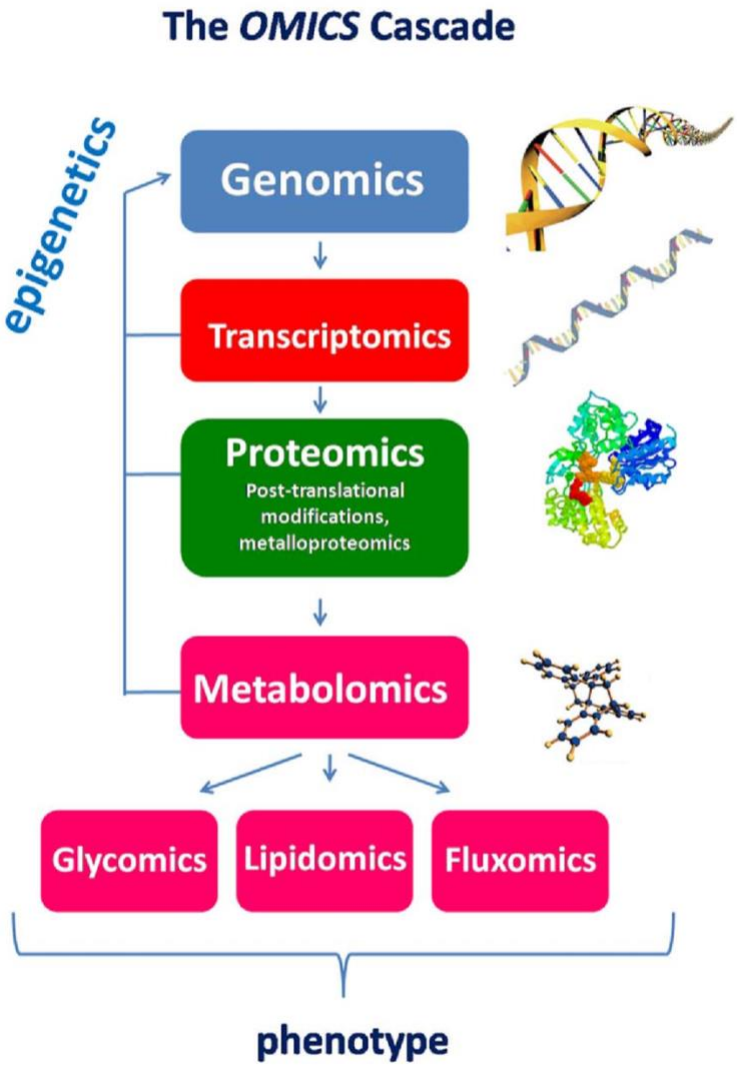

Figure 7: Schematic representation of the development of the omic fields[71].

3. HYPOTHESIS AND OBJECTIVES

Although the administration of antiretroviral therapy (ART) in HIV-positive patients has considerably reduced the morbidity and mortality associated with HIV infection, the degree of immunologic recovery among different patients varies. In most cases, the treatment meets its objectives, but there are a percentage of patients who, even when receiving ART and achieving virological suppression, do not achieve a good recovery of the number of $CD4^+$ T-cells to acceptable levels.

The main hypothesis of this work is that the immunological non-responders had a different proteomic profile before ART administration that can explain their incomplete immune reconstitution during ART. In that case, we consider that knowing the baseline $CD4^+$ and $CD8^+$ T-cells proteomic profile of HIV subjects initiating first ART with low $CD4^+$ T-cells could be of use to identify potential prognostic biomarkers for those cases in which complete recovery of the $CD4^+$ number is not going to be achieved.

The main objective is to find proteins or a profile of proteins that can be used as proteomic biomarkers in CD4+ and $CD8^+$ T-cells, to identify and differentiate patients who will be immunological ART responders and patients who will be not.

Our secondary objectives are as follows:

- Identify the metabolic pathways or biological processes that may be related to poor immune recovery and relate them to the patterns or proteins obtained.
- Identify possible therapeutic targets in non-responder patients that may allow an improvement in the immunological response to ART.

4. MATERIALS AND METHODS

4.1. Participants selection and characteristics

The study cohort included 100 HIV-patients from 5 different hospitals. These hospitals are: Hospital Universitari de Tarragona Joan XXIII, Hospital Universitari Vall d'Hebron, Hospital Universitari Germans Trias i Pujol, Hospital Clínic de Barcelona i Hospital de la Santa Creu i Sant Pau.

Patients were mainly divided into control and cases. The categorization of these groups was based on the $CD4^+$ T-cell count before starting the treatment of ART. Control patients were those with more than 200 $CD4^+$ T-cells/µl before starting treatment, while those with less than 200 $CD4^+$ T-cells/µl were classified as cases. Also, the cases were divided into Immunological Responders (IR) and Immunological Non-Responders (INR) based on the $CD4^+$ T-cell counts after 48 weeks of being on ART therapy. IRs were defined by a $CD4^+$ T-cell count greater than 250 $CD4^+$ T-cells/µl after treatment and INR patients were those that showed less or equal to 250 $CD4^+$ T-cells/µl.

Among the study patients, we found 51 control and 49 cases. Within the cases, we found 34 IR and 13 INR. There were 2 patients from cases without data on $CD4^+$ T-cells count at 48 weeks of treatment that could not be classified and should be excluded from the analyses (*Figure 8*).

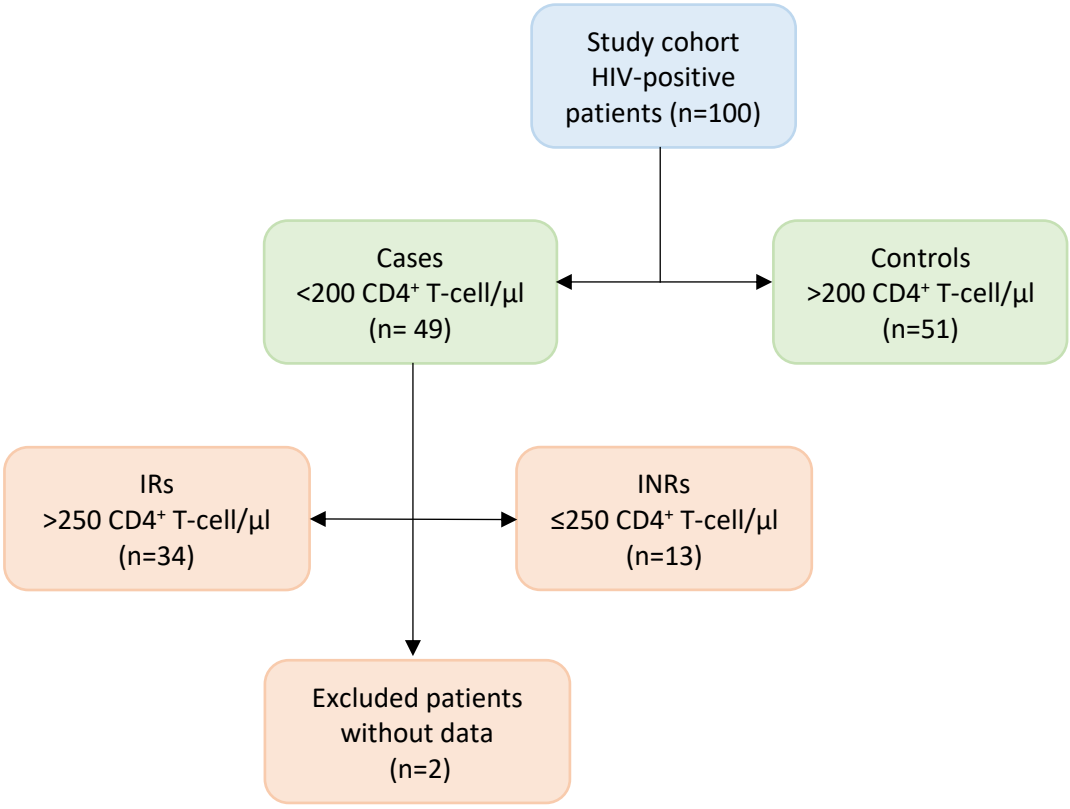

Figure 8: Flowchart of the groups of patients in the study and the number of people in each one.

4.2. $CD4^+$ and $CD8^+$ T-cells separation

From total blood as the initial sample, we processed approximately 2 EDTA 9ml Leucosep tubes from each patient for extraction of peripheral blood mononuclear cells (PBMCs) by density gradient separation (Ficoll).

Once the PBMCs were separated, the different $CD4^+$ and $CD8^+$ T-cells were isolated with Automacs using MicroBeads. The cells were stored at -80°C for subsequent proteomics analysis.

4.3. Protein extraction and identification

Protein extraction is performed according to a protocol where protease inhibitors and phosphatase inhibitors were added to prevent the proteins from degrading or being altered. The cell pellet was resuspended in the buffer with the inhibitors mentioned above, and the cells were sonicated following a 3-cycle protocol. Finally, it is centrifuged and the supernatant will be analyzed by proteomics.

There, the supernatant was used for protein digestion and 11-plex TMT peptide labelling so that it could then be analyzed by nanoLC-(Orbitrap)MS/MS. Protein identification and quantification were performed in Proteome Discovere software using multidimensional protein identification technology (MudPIT).

Protein identification and quantification were performed at Center of Omics Science (COS).

4.4. Statistical analysis

In the different statistical analyses that were performed, data from $CD4^+$ or $CD8^+$ T-cells were differentiated. The different connections between the different study groups were analyzed using IBM SPSS Statistics. We used the non-parametric Man-Whitney and Kruskal Wallis tests to indentify demographic and clinical differences among groups. We proceeded to perform the ANOVA test for the statistical analysis of the relative protein.

Then, although we managed to reduce the large number of proteins with the ANOVA analysis, we still obtained a high number of proteins and significant connections, so MetaboAnalyst Statistical Analysis was used with those significant proteins to clarify and establish clearer patterns or profiles. From this, we were also able to find useful figures to understand the results obtained and to be able to make associations more effectively. In addition, String was also used to analyze specific pathways and protein networks involving differentially expressed proteins for biological interpretation.

5. RESULTS

The baseline clinical characteristics of the 100 patients in this study are described in Table 2. The patient groups were qualitatively and quantitatively different concerning age, sex, $CD4^+$ and $CD8^+$ T-cell count and $CD4^+/CD8^+$ ratio. Comparing the different groups, the mean age of the patients is almost the same in all study groups, although the age range of the patients is more variable.

According to the patient group classification criteria, before starting ART the control group had 561 [200-1132] $CD4^+/\mu L$ T-cells and a $CD4^+/CD8^+$ ratio of 0.64 [0.20-1.84], whereas the cases (both IR and INR) had much lower $CD4^+$ T-cell counts (113 [10-200] and 75 [3-170], respectively) and a $CD4^+/CD8^+$ ratio also considerably lower (0.18 [0.02-0.49] and 0.09 [0.02-0.17], respectively).

Subject characteristics	CONTROLS	IR	INR	P value
Age (years)	41 [19-65]	42 [3-70]	42 [24-66]	0,988
Sex (male), n (%)	44 (89,79)	25 (73,53)	12 (92,31)	0,115
$CD4^+$ T-cell count (cells/µl)	561 [200-1132]	113[10-200]	75 [3-170]	>0,001
$CD8^+$ T-cell count (cells/µl)	1085 [240-2938]	922 [50-2248]	945 [98-2470]	0,446
Ratio $CD4^+/CD8^+$	0,64 [0,20-1,84]	0,18 [0,02-0,49]	0,09 [0,02-0,17]	>0,001

Table 2: **Baseline study cohort characteristics.** All qualitative variables in Table 2 were expressed by the total number of patients (as a percentage), whereas quantitative variables were expressed by the interquartile range.

5.1. Proteomic profile differences in $CD4^+$ T-cells

Initially, a total of 859 proteins were identified in $CD4^+$ T-cells from patient samples from which 103 resulted having significantly different relative concentrations among the three groups (ANOVA test).

As the main objective of the present work was to study the differential response in the INR subjects, thus, to refine the results and to be able to make concise conclusions, it was decided to only choose, for the association analyses, the significant proteins between control-INR and between IR-INR. Thus, a total of 25 proteins were highlighted (*Table 3*).

	Compounds	f.value	p.value	-log10(p)	FDR	Fisher's LSD
1	P28838	41.71	0.00	13.00	0.00	1 - 0; 2 - 0; 2 - 1
2	P07237	28.36	0.00	9.66	0.00	1 - 0; 2 - 0; 2 - 1
3	P02792	24.52	0.00	8.58	0.00	1 - 0; 2 - 0; 2 - 1
4	P05164	20.83	0.00	7.50	0.00	1 - 0; 2 - 0; 2 - 1
5	P53004	17.55	0.00	6.49	0.00	1 - 0; 2 - 0; 2 - 1
6	P61626	15.31	0.00	5.76	0.00	1 - 0; 2 - 0; 2 - 1
7	P06702	14.47	0.00	5.49	0.00	1 - 0; 2 - 0; 2 - 1
8	P07602	13.74	0.00	5.24	0.00	1 - 0; 2 - 0; 2 - 1
9	P23381	12.99	0.00	4.99	0.00	1 - 0; 2 - 0; 2 - 1
10	P08758	12.21	0.00	4.72	0.00	1 - 0; 2 - 0; 2 - 1
11	P14625	11.55	0.00	4.49	0.00	1 - 0; 2 - 0; 2 - 1
12	P02671	11.11	0.00	4.33	0.00	1 - 0; 2 - 0; 2 - 1
13	P52209	10.93	0.00	4.27	0.00	2 - 0; 2 - 1
14	P30101	10.78	0.00	4.22	0.00	1 - 0; 2 - 0; 2 - 1
15	A0A2R8Y6J3	10.62	0.00	4.16	0.00	0 - 1; 0 - 2; 1 - 2
16	U3KPS2	9.10	0.00	3.61	0.00	1 - 0; 2 - 0; 2 - 1
17	P04040	7.35	0.00	2.97	0.00	2 - 0; 2 - 1
18	P02775	7.12	0.00	2.88	0.00	2 - 0; 2 - 1
19	E5RIW3	6.99	0.00	2.83	0.00	2 - 0; 2 - 1
20	D6RB09	6.84	0.00	2.78	0.00	0 - 2; 1 - 2
21	Q96KP4	5.81	0.00	2.38	0.00	2 - 0; 2 - 1
22	Q8WYJ6	5.78	0.00	2.37	0.00	0 - 2; 1 - 2
23	P61978	5.74	0.00	2.35	0.00	0 - 2; 1 - 2
24	P09382	5.61	0.00	2.30	0.01	2 - 0; 2 - 1
25	C9J8S3	5.15	0.01	2.12	0.01	2 - 0; 2 - 1

Table 3: Important proteins of CD4⁺ T-cells identified by One-way ANOVA.

A discriminant analysis (PLS-DA) was performed with these 25 significant proteins (*Figure 7*) and a trend was shown for controls (0) and IRs (1), and as opposed to INRs (2). Controls and IRs were more homogeneous groups and were closer to each other, although the control group was clearly better defined than the IR group. INRs were more heterogeneous and more differentiated from the other study groups.

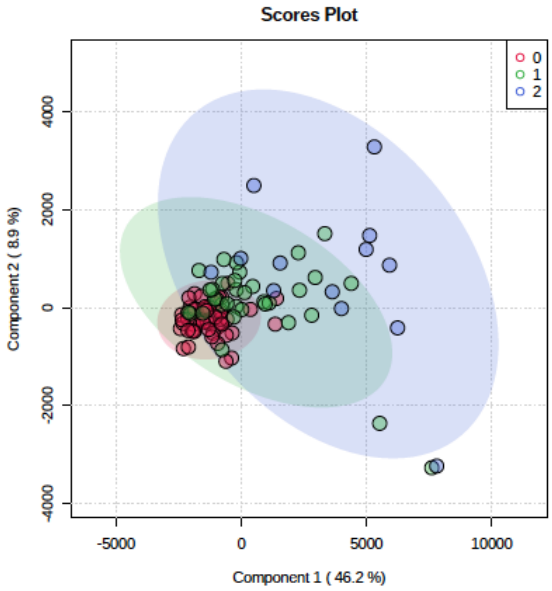

Figure 7: PLS-DA score plot of proteomic analyses in CD4⁺ T-cells distinguishes control subjects (red dots) from IR (green dots) and INR (blue dots).

The Heatmap dendrogram (*Figure 9*) confirmed how the INR pattern differed from both IRs and controls, creating a cluster including controls and IRs. This indicates that proteomic profile from controls and IRs was similar and that both control and IRs differed from INRs. Also, a specific proteomic profile could be established from each type of patient, and thus potential proteomic biomarkers for poor immune reconstitution were established.

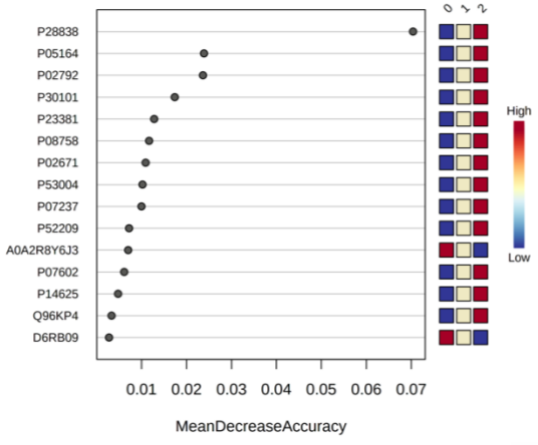

Figure 8: **Variable importance plot of the Random Forest analysis.** The variables are ordered top-to-bottom as most-to-least important in classifying between controls, IR and INR. As it can be seen, the most important protein by far is P28838, followed by protein P05164.

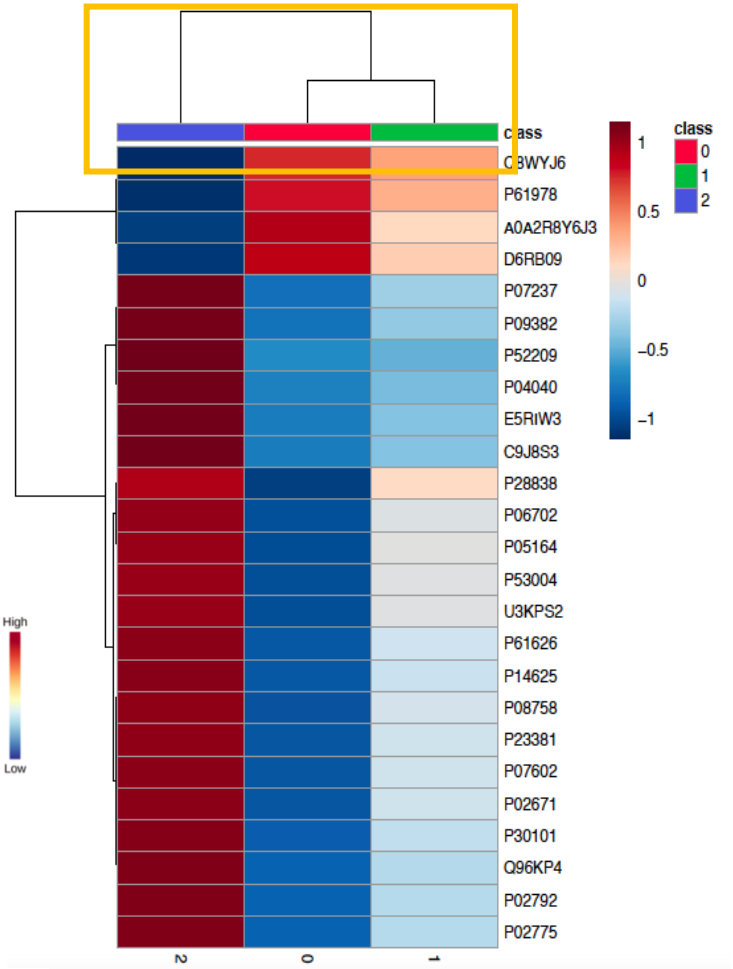

Figure 9: **Heatmap of the concentrations of measured proteins in CD4⁺ T-cells.** Class 0 refers to the control group, class 1 to the IR and class 2 to the INR. Low concentrations are shown in blue (-1) whereas higher concentrations increase to red colour (1).

Additionally, with the information provided by the Heatmap it is possible to identify 2 clusters of proteins.

1- Cluster 1: Q8WYJ6, P61978, **A0A2R8Y6J3, D6RB09**

   This cluster composed of 4 proteins has a very different pattern in both controls (1) and IR (0.5) compared to INR (-1). The most important proteins in this cluster are A0A2R8Y6J3 and D6RB09, as we can see in the Random Forest analysis (*Figure 8*) which classifies them as important variables in the differentiation of study groups.

2- Cluster 2: **P07237**, P09382, **P52209**, P04040, E5RIW3**, C9J8S3,** P28838, P06702**, P05164**, P53004, U3KPS2, P61626, P14625, P08758, P23381, P07602, P02671, P30101, Q96KP4, **P02792**, P02775

In this cluster the control and IR groups are much similar to each other while the INRs are differentiated, this would allow to directly differentiate the INRs from both subjects with good immunological status but also subjects initiating ART with low CD4+ T cell count that will have a good immunological response to ART. It can be seen how the INRs are at 1 while the controls and IR are between -0.5 and -1.

The proteins in bold are those that the Random Forest analysis identifies as the most important in differentiating the clusters, but of particular note is the P28838 protein presenting the higher mean decrease accuracy. This would be the most relevant of this cluster and surely the most influential protein in the differentiation of groups of all the results obtained so far.

Using the String database, we study the possible relationships or interactions that the different proteins selected from each cluster may have between them or with other proteins. It also allows us to obtain information about the biological process in which the proteins coincide or their subcellular localization, among others.

In the case of cluster 1 (*Table 4*), connections can be seen between all the proteins of the cluster except Septin-1, even if the connections are enlarged (*Figures 8 and 9*).

Uniprot Protein ID	Gene Symbol	Official protein name
Q8WYJ6	SEPT1	Septin-1
P61978	HNRNPK	Heterogeneous nuclear ribonucleoprotein K
A0A2R8Y6J3	RPL5	60S ribosomal protein L5
D6RB09	RPS3A	40S ribosomal protein S3a

Table 4: Different protein names of the selected proteins of the cluster 1 of $CD4^+$ T-cells.

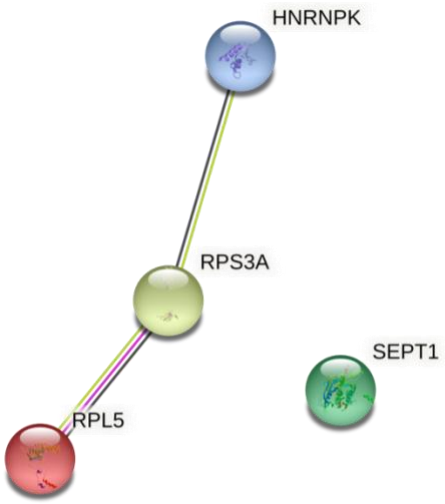

Figure 8: Protein–protein interaction analysis of the 4 proteins of cluster 1.

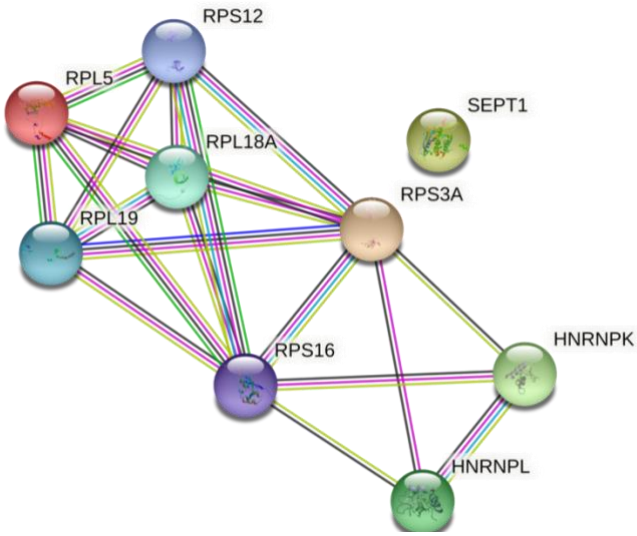

Figure 9: Analysis of the protein-protein interaction of cluster 1 extending the connections with other possible proteins.

No significant biological processes are obtained for this cluster including the three identified proteins.

On the other hand, to see what importance Sept1 might have in the immune reconstitution , we looked for potential interactions with Sept1 (*Figure 10*). It can be seen that the Sept1 protein is the main linker of the septin family with Aurora protein kinase B (AURKB), all related to cellular protein localization.

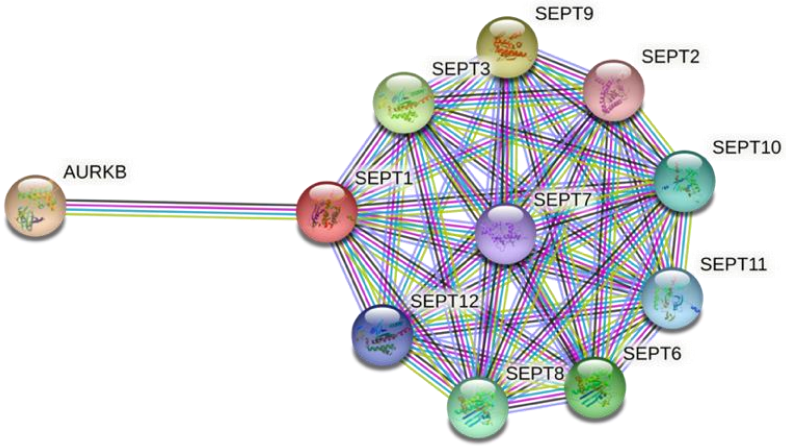

Figure 10: SEPT1 map interaction analysis. As expected, it connects with the other proteins of the septin family, but it can also be seen that it connects the AURKB protein with the septin family mentioned above.

In the case of cluster 2 (*Table 5*), most of the proteins are related to each other except TBCA and BLVRA (*Figure 10*), no matter how much we try to increase the connections.

Uniprot Protein ID	Gene Symbol	Official protein name
P07237	P4HB	Protein disulfide-isomerase
P09382	LGALS1	Galectin-1
P52209	PGD	6-phosphogluconate dehydrogenase, decarboxylating
P04040	CAT	Catalase
E5RIW3	TBCA	Tubulin-specific chaperone A
C9J8S3	RAB7A	Ras-related protein Rab-7a
P28838	LAP3	Cytosol aminopeptidase
P06702	S100A9	Protein S100-A9
P05164	MPO	Myeloperoxidase
P53004	BLVRA	Biliverdin reductase A
U3KPS2	PRTN3	Myeloblastin
P61626	LYZ	Lysozyme C
P14625	HSP90B1	Endoplasmin
P08758	ANXA5	Annexin A5
P23381	WARS1	Tryptophan--tRNA ligase, cytoplasmic
P07602	PSAP	Prosaposin
P02671	FGA	Fibrinogen alpha chain

P30101	PDIA3	Protein disulfide-isomerase A3
Q96KP4	CNDP2	Cytosolic non-specific dipeptidase
P02792	FTL	Ferritin light chain
P02775	PPBP	Platelet basic protein

Table 5: Different protein names of the selected proteins of the cluster 2 of $CD4^+$ T-cells.

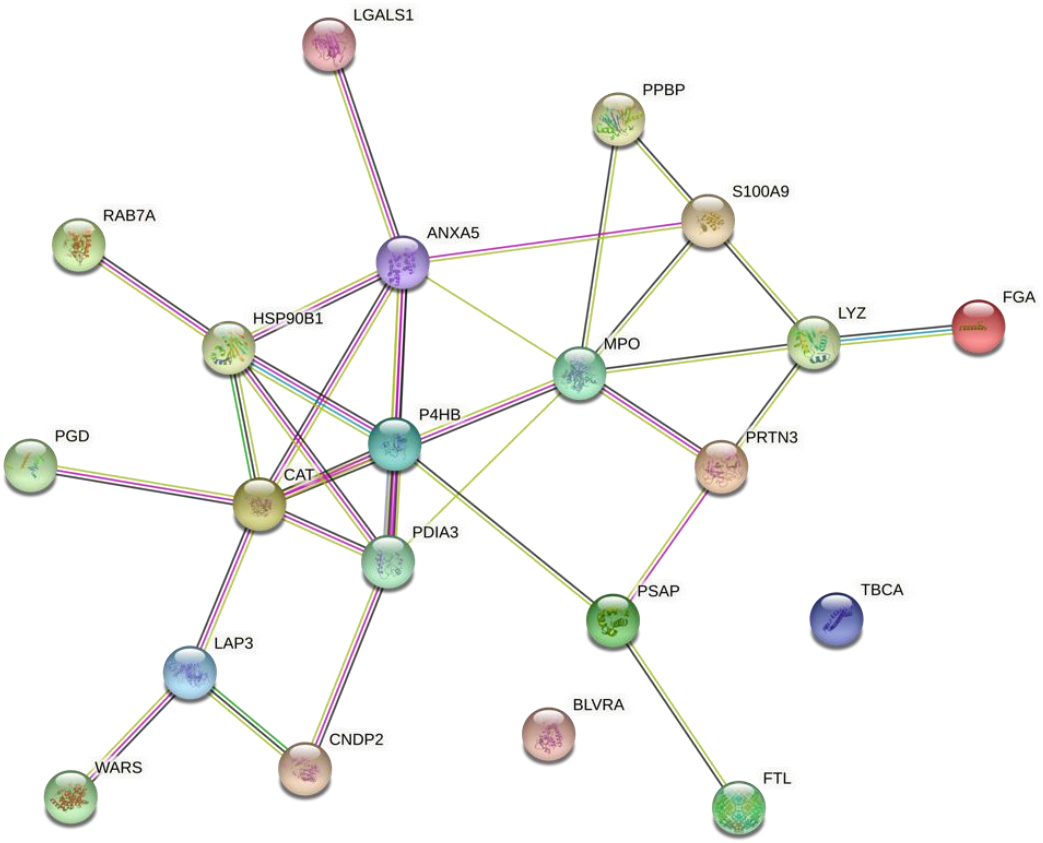

Figure 11: Protein–protein interaction analysis of the proteins of cluster 2.

We looked at the different biological processes in which the main proteins from cluster 2 are related (*Table 6*) and it was seen that, in general terms, there would be a relationship between these proteins and the immunological process.

Then, we also tried to find out if there is an important connection between TBCA and BLVRA proteins and other processes that may be of interest to our study. Although these two proteins converge in some way, the biological processes were not related to the process we are studying, so we do not follow that route.

Biological Process (Gene Ontology)			
description	*count in network*	*strength*	*false discovery rate*
Regulated exocytosis	11 of 697	1.17	4.15e-07
Leukocyte activation involved in immune r...	10 of 626	1.17	1.33e-06
Neutrophil degranulation	9 of 484	1.24	2.18e-06
Cell activation	11 of 1075	0.98	2.62e-06
Vesicle-mediated transport	12 of 1805	0.79	3.18e-05
Immune response	11 of 1588	0.81	9.91e-05
Antimicrobial humoral response	5 of 160	1.46	0.00041
Establishment of localization in cell	12 of 2375	0.67	0.00057
Immune system process	12 of 2481	0.65	0.00087
Cellular homeostasis	8 of 895	0.92	0.0012
Regulation of cell death	10 of 1696	0.74	0.0017
Defense response to bacterium	5 of 277	1.23	0.0045
Platelet degranulation	4 of 129	1.46	0.0048
Protein metabolic process	14 of 4251	0.49	0.0051
Regulation of apoptotic process	9 of 1550	0.73	0.0059
Organonitrogen compound metabolic pro...	15 of 5244	0.43	0.0095
Homeostatic process	9 of 1676	0.7	0.0102
Cell redox homeostasis	3 of 60	1.67	0.0141
Response to stress	12 of 3485	0.51	0.0206
Protein folding in endoplasmic reticulum	2 of 10	2.27	0.0237
Positive regulation of viral entry into host ...	2 of 10	2.27	0.0237
Protein folding	4 of 213	1.24	0.0237
Sequestering of metal ion	2 of 11	2.23	0.0253
Neutrophil migration	3 of 83	1.53	0.0300
Cellular oxidant detoxification	3 of 90	1.49	0.0370
Maintenance of location in cell	3 of 91	1.49	0.0373
Catabolic process	9 of 2042	0.61	0.0373
Organic substance metabolic process	17 of 7755	0.31	0.0373
Toll-like receptor signaling pathway	3 of 100	1.45	0.0451

Table 6: Biological processes in which proteins of cluster 2 are involved or related ordered by their false discovery rate (FDR), this measure describes how significant the enrichment is. In the Count In Network column, the first number indicates how many proteins in your network are annotated with a particular term and the second one indicates how many proteins in total have this term assigned. The strength describes how large the enrichment effect is.

5.2. Proteomic profile differences in CD8⁺ T-cells

Initially, 953 proteins were identified in CD8⁺ T-cells from patient samples. As proceeded with proteomic profile in CD4⁺ T-cells, ANOVA test was performed to refine the analysis. Finally, 35 significant proteins were obtained, with which we began to look for associations.

A discriminant analysis (PLS-DA) was performed with these 35 significant proteins. Although the distribution of proteins along the components showed an overlap between the 3 study groups (*Figure 12*), there was a certain tendency for each group, so there will be proteins that are more relevant in each case.

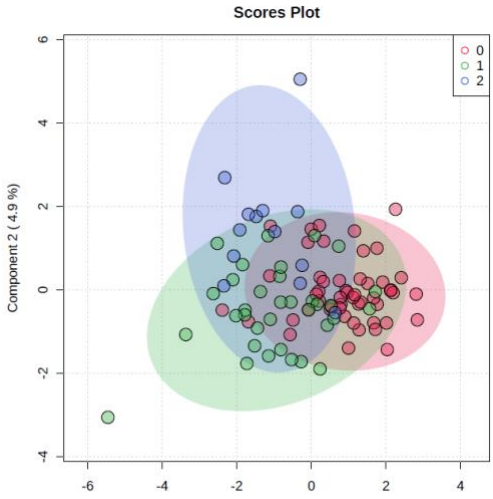

Figure 12: PLS-DA score plot of proteomic analyses of CD8+ T-cells distinguishes control subjects (red dots) from IR (green dots) and INR (blue dots).

Of the 35 significant proteins that we obtained, we decided to keep those that had significant differences between the control and IR groups with the INR group, the main target group of the study. From there, we reduced the number of proteins from 35 to 9. Table 7 shows the most significant proteins, which are the ones we will be working with from now on.

	Compounds	f.value	p.value	-log10(p)	FDR	Fisher's LSD
1	P28838	9.52	0.00	3.77	0.00	1 - 0; 2 - 0
2	P20962	9.26	0.00	3.67	0.00	1 - 0; 2 - 0
3	Q96AZ6	8.70	0.00	3.47	0.00	1 - 0; 2 - 0
4	P20591	7.87	0.00	3.16	0.01	1 - 0; 2 - 0; 2 - 1
5	P01857	7.21	0.00	2.91	0.01	1 - 0; 2 - 0
6	P46778	6.97	0.00	2.82	0.01	0 - 2; 1 - 2
7	A0A669KBI6	6.70	0.00	2.72	0.01	2 - 0; 2 - 1
8	P53634	6.51	0.00	2.65	0.01	1 - 0; 2 - 0
9	G5E9G0	5.69	0.00	2.33	0.02	0 - 1

Table 6: Important proteins of $CD8^+$ T-cells identified by One-Way ANOVA.

Figure 13 shows a Heatmap of these 9 differential proteins between groups of study. It can be seen how the control group differs from both IR and INR.
Random Forest analysis revealed which proteins are more suitable to differentiate the groups (*Figure 14*), among them, we found some of the most significant ones previously described (*Table 7*).

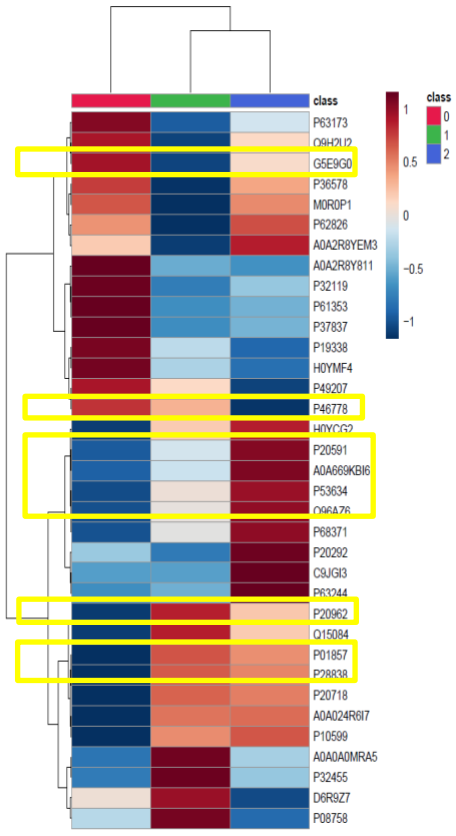

Figure 9: **Heatmap of the concentrations of measured proteins.** Low concentrations are shown in blue whereas higher concentrations increase to red colour.

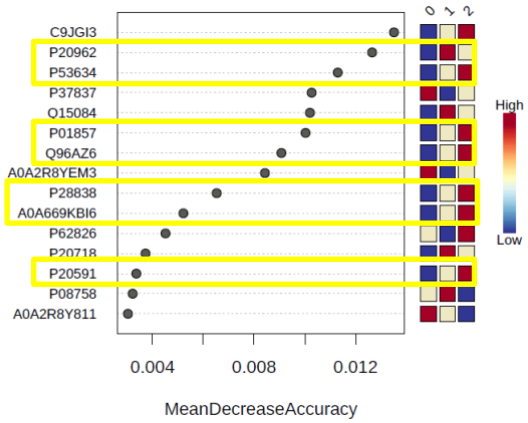

Figure 14: **Variable importance plot of the Random Forest analysis.** The variables are ordered top-to-bottom as most-to-least important in classifying between controls, IR and INR. Of the 9 important proteins previously mentioned, 7 can be seen here.

Two clusters can be identified with the proteins that have been previously selected for their significance.

- Cluster 1: P20591, AOA669KB16, **P53634**, Q96A76.
  This cluster shows a very different heatmap pattern between the control (-1) and INR (1) group. In the case of the IR, the protein pattern was closer to the control group.
  In this cluster, the P53634 protein stands out since, as seen in the Random Forest analysis, it is classified as a very important group differentiating variable.

- Cluster 2: **P20962**, P01857, P28838.
  This one, on the other hand, gives a differentiating pattern between controls and cases in general, since the INR-IR pattern is more similar in this case (both are between 0.5 and 1). The P20962 protein stands out as the main group differentiator within this cluster.

Using the String database, we study the possible relationships or interactions that the different proteins selected from each cluster may have.

Uniprot Protein ID	Gene Symbol	Official protein name
P20591	MX1	Interferon-induced GTP-binding protein Mx1
A0A669KBI6	STAT1	Signal transducer and activator of transcription 1-alpha/beta
P53634	CTSC	Dipeptidyl peptidase 1
Q96AZ6	ISG20	Interferon-stimulated gene 20 kDa protein
P46778	RPL21	60S ribosomal protein L21

Table 7: Different protein names of the selected proteins of the cluster 1 of $CD8^+$ T-cells.

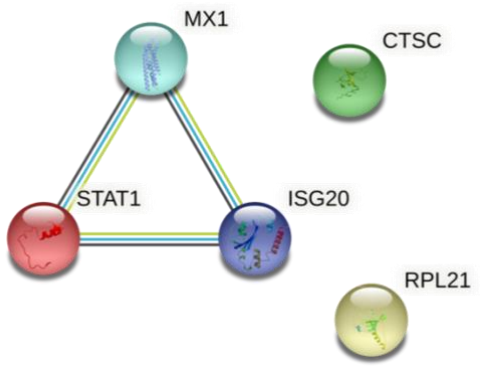

Figure 15: Protein-protein interaction map of the proteins of cluster 1.

From cluster 1 (*Table 8*), the first basic interaction scheme obtained showed how MX1-STAT1-ISFG20 interact with each other (Figure 15), which implies that they are probably in the same pathway. However, the CTSC and RPL21 proteins do not interact directly with the other ones.

We searched which pathways the 3 related proteins are involved in, and found that they are involved in the following biological processes:

Biological Process	False Discovery Rate
Adaptive immune response	2.48e-07
Positive regulation of immune system process	1.71e-05
Lymphocyte mediated immunity	5.01e-05
Regulation of macrophage activation	0.0301

Table 8: Biological processes in which proteins of cluster 1 are involved or related, with their false discovery rate.

Of these biological processes, the only one in which the 4 proteins of the cluster (the 3 we see related and CTSC) are found is in the Immune effector process.

To see if the other two proteins have any relationship, we looked for more possible interactions (*Figure 16*), the result obtained confirmed that the 3 proteins (MX1-STAT1-ISFG20) are linked to the RPL21 protein, which is precisely the one that had an inverse pattern to the proteins identified within cluster 1. But as can be seen, CTSC still has no connection with the pathways of the other proteins, even the immune effector process connexion.

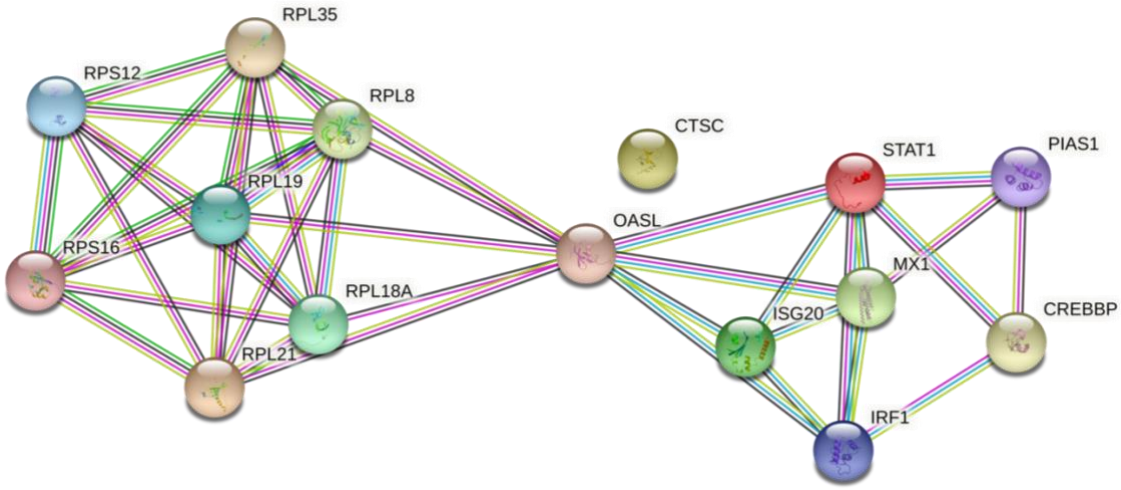

Figure 16: Analysis of the protein-protein interaction of cluster 1 extending the connections with other possible proteins.

As no relationship between CTSC with the other selected proteins has been found, an attempt is made to see the direct connections and pathways in which it is involved (*Figure 17*).

- Adaptive immune response
- Positive regulation of immune system process
- Lymphocyte mediated immunity
- Regulation of macrophage activation

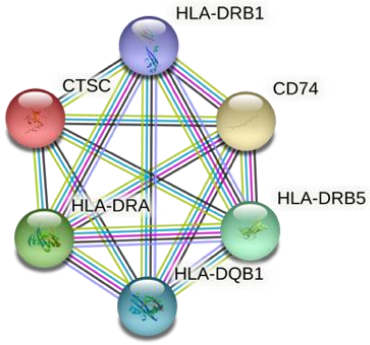

Figure 17: Interaction map of the CTSC protein.

In the case of cluster 2 (*Table 10*), the 3 proteins are analyzed but found that String does not recognize IGHG1 and the other 2 proteins are not related, so we decided to relate the proteins in cluster 2 to those in cluster 1 to see if we could find more specific relationships. As a result, we obtain the same as in Figure 16 but with the LAP3 protein entering into the network connected to STAT1, while PTMS remains unattached as it does CTSC.

Uniprot Protein ID	Gene Symbol	Official protein name
P20962	PTMS	Parathymosin
P01857	IGHG1	Immunoglobulin heavy constant gamma 1
P28838	LAP3	Cytosol aminopeptidase

Table 9: Different protein names of the selected proteins of the cluster 2 of CD8+ T-cells.

As with the CTCS protein, an attempt is made to see the direct connections and pathways in which the PTMS protein is found (*Figure 18*). We found out that PTMS may mediate immune function by blocking the effect of prothymosin alpha which confers resistance to certain opportunistic infections.

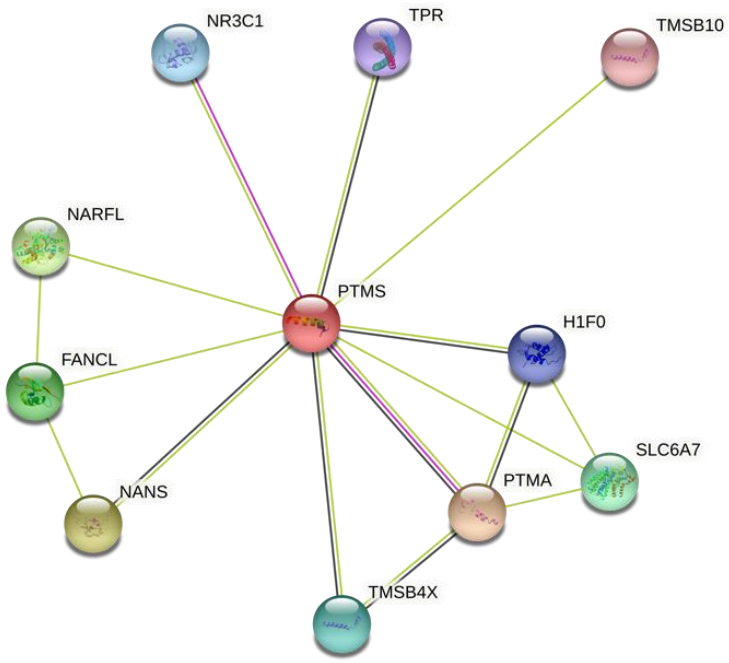

Figure 18: Interacion map of PTMS protein.

5.3. Significant common proteins between CD4⁺ and CD8⁺ T-cells

The only protein that can be found significantly altered between groups in both CD4⁺ and CD8⁺ T-cells is P28838, also known as Cytosol aminopeptidase.

It has already been mentioned that this protein in CD4⁺ T-cells could be the most important in terms of differentiating the different groups of patients since in the heatmap (*Figure 9*) the controls are at -1, the IR at 0 and the INR at 1 and the RandomForest analysis confirms this.

On the other hand, in $CD8^+$ T-cells, this protein is also relevant but not as much as in the case of $CD4^+$, in this case in the heatmap (*Figure 13*) we see that this protein allows us to differentiate from control (-1) to cases (1), but we could not differentiate the different cases only by looking at the relative concentrations of this protein in $CD8^+$ T-cells.

Figures 19 and 20 show the difference in the relative concentrations of this protein in the different T-cells and the different groups of patients, being 0 control group, 1 IR group and 2 INR group, in the two graphs considered. The pattern obtained in the respective Heatmaps is now confirmed.

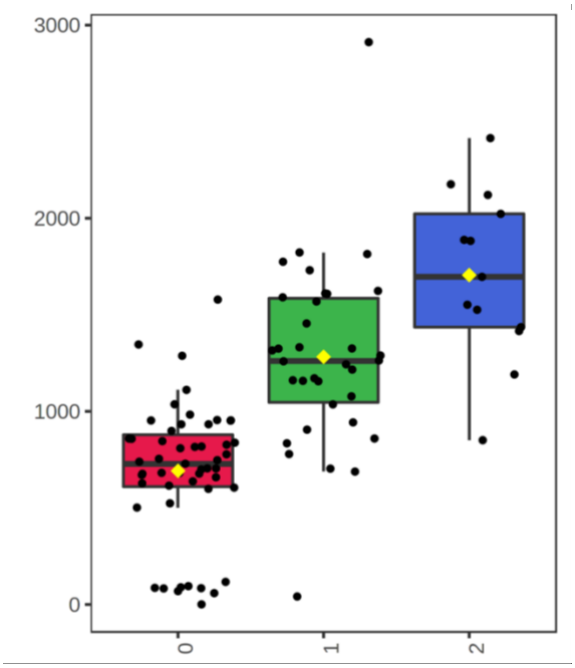

Figure 19: $CD4^+$ T-cell 28838 concentration.

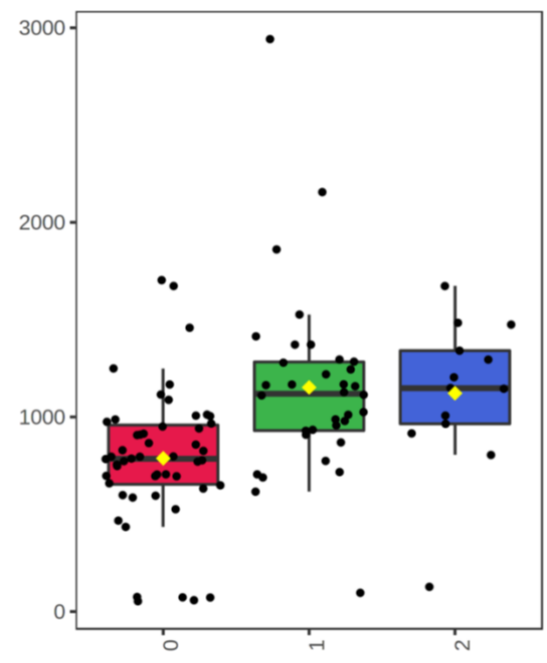

Figure 20: $CD8^+$ T-cell 28838 concentration.

6. DISCUSSION

Multiple complex mechanisms are involved in the failure of immune recovery in HIV-positive individuals who achieve successful virologic suppression. Although some factors that have been commonly associated with poor immune recovery continue to be investigated, the complete mechanism preceding poor immune recovery associated with ART remains unclear.

In this study, it was initially postulated that there may be proteomic differences related to differential immune response among HIV-positives that may be related to poor immune progression. It was hoped that proteins may play a key role in identifying different classes of patients based on patterns that may be repetitive among HIV-positive individuals. Indeed, this study has revealed that there are proteomic differences between HIV-positives showing different immune profiles prior to ART administration that may be directly related to treatment response, Thus, there are specific proteins that could allow to identify those patients who will not respond well to treatment and, therefore, look for another approach to treat the infection.

Different patterns have been obtained between the 2 types of T-cells, so we proceed to comment on the most relevant or most useful aspects and findings that may be present in each subtype.

In $CD4^+$ T-cells, one of the proteins that stand out in cluster 1 with good discriminatory power using Random Forest to discriminate INR patients compared to controls and IR patients is A0A2R8Y6J3 (RPL5). It has been previously demonstrated that this protein, which gene expression seems to be altered with the administration of ART, is mainly related to the biological process of negative regulation of gene expression [72]. Therefore, this protein, in its basal state, could be affecting certain factors related to poor immune recovery, preventing it, and when the treatment is administered, by suffering modifications, it could stop decreasing this range of gene expression and could allow the process of poor immune recovery to continue.

In cluster 2 of $CD4^+$ T-cells, the significance of the P05164 protein (MPO) stands out, which is found in high concentrations in INR patients compared to the other study groups. Myeloperoxidase (MPO) is a heme-containing peroxidase expressed mainly in neutrophils. In the presence of hydrogen peroxide and halides, MPO catalyzes the formation of reactive oxygen intermediates, including hypochlorous acid (HOCl)[73]. HIV-infected people often experience a decrease in peripheral blood neutrophil count compared to uninfected people. Among the major immunological factors contributing to this situation, it has been proposed that HIV-induced cytotoxicity contributes to neutropenia[74]. Because neutrophils are responsible for releasing enzymes such as MPO, this situation of neutrophil apoptosis responds to the high levels of MPO found in $CD4^+$ T-cells. The MPO protein features, among other things, as a tumour differentiation antigen of the disease known as leukaemia, a disease based on the malignant neoplasia of mature $CD4^+$ lymphocytes[75]. This relationship previously described between $CD4^+$ T-cells and MPO could explain the MPO levels initially found in INR patients since this

enzyme is found to be acting in diseases related to problems with $CD4^+$ T-cells, the same cells that INRs fail to recover.

In $CD8^+$ T-cells, one of the proteins that stands out in cluster 1 is P53634 (CTSC). This protein is found in higher concentrations in the INR group compared to the other study groups. This protein, known as Dipeptidyl peptidase 1 (CTSC) is a thiol protease and has dipeptidylpeptidase activity. Because when looking at connections with other proteins in its cluster it had virtually no links previously described, we looked at the main biological processes in which it is involved and it was strongly related to the immune process. This fact fits with the results observed, which is that the activation of the immune response within the infection could enhance the relative concentration of this protein in this cell subtype. However, on the other hand, the administration of ART in these patients may negative influence the function of this protein preventing efficient immune reconstruction.

What we could ask ourselves is why this protein is found in higher relative concentrations in INR patients with respect to those in which they are going to achieve normal immunity. INR, in whom the treatment will not achieve immune reconstruction, may have immunological differences that the immune system itself recognizes and potentiate the overexpression of CTSC protein to try to enhance the recovery of $CD4^+$ T-cells, although it is a complicated field of study since it is not yet known the process by which patients fail to respond to treatment.

In cluster 2 of $CD8^+$ T-cells, the most significant protein is P20962 (PTMS). This protein gives a differentiating pattern between Controls and Cases in general, since the INR-IR pattern is more similar in this case, being the relative concentrations of this protein higher in the cases than in the control. Parathymosin (PTMS) may mediate immune function by blocking the effect of prothymosin alpha which confers resistance to certain opportunistic infections. This protein function fits with the elevated levels in those patients where $CD4^+$ T-cell levels are below the optimal threshold before ART, i.e., this protein was overexpressed in patients classified as cases who is expected to achieve immunologic recovery once ART is initiated. This protein, however, does not allow us to differentiate patients in IR and INR.

Finally, we are going to focus on the P28838 protein, also known as Cytosol aminopeptidase (LAP3), the only protein found in both cell subtypes and with a high significance in relation to INR participants. This protein is a cytosolic metallopeptidase and its main function is to catalyze the removal of unsubstituted N-terminal hydrophobic amino acids from various peptides. In the case of $CD4^+$, the P28838 protein has more differentiated concentrations between the groups and, therefore, is more easily associated with the identification of possible patients ($p < 0.001$), while in the case of $CD8^+$ the concentrations of the groups considered as cases (IR and INR) are quite similar between them and differ from the control group ($p<0.001$). It is known that in infected cells the main metabolic pathways such as glycolysis and the tricarboxylic acid cycle are altered[76,77], this could be one of the factors related to poor immunological recovery. The biological processes that embrace P28838 are closely related to the TAC

cycle and glycolysis, so in those individuals who have a predisposition to poor immunological recovery, an overproduction of this protein may be a response of the organism to try to rebuild the basal functioning of these metabolic pathways and to reconstitute the levels of $CD4^+$ T cells.

To conclude the discussion, it should be noted that our study had some notable limitations. The number of patients per group was relatively small, which made it difficult to extrapolate data on a large scale and affirm the patterns obtained. In addition, the intrinsic variability that human patients implies must be taken into account, since a large number of factors affect it. There is no doubt that more studies with a larger initial cohort of seropositive individuals are needed to confirm or complete these relationships obtained between proteomic profiles and different types of patients.

Also, no standard definition for the immunological response is available and therefore the threshold of 250 for $CD4^+$ T cells/ul could seem somewhat arbitrarily set for the present study. In this regard, we have previously validated that patients receiving ART with $CD4^+$ T-cell counts persistently below 250 cells/µl are poor immunological responders and are associated with worse clinical outcomes[78].

7. CONCLUSION

In conclusion, it has been shown that there is a specific $CD4^+$ and $CD8^+$ proteomic pattern in HIV-positives before undergoing ART that is distinctive among subjects with different immunological response to ART. So in the future, it would be possible to differentiate patients with poor immune recovery with a simple analysis focused on a proteomic pattern.

The proteins identified in this study demonstrated significant differences among subjects initiating ART with good immune status (control), subjects initiating ART with low immune status but good immune recovery on ART (immunological responders, IR) and subjects initiating ART with low immune status that maintain poor immune recovery on ART (immunological non- responders, INR). The LAP3 protein, a cytosolic metallopeptidase, could be highlighted, which relative concentration in $CD4^+$ T-cells could easily allow the identification of HIV-positive with a potential immune failure to ART and which relative concentration in $CD8^+$ T-cells may identify HIV-positives without ART with low $CD4^+$ T-cell counts.

This study has allowed the first screening of proteins in $CD4^+$ and $CD8^+$ T-cells in HIV-positives with different immunological response prior to ART administration, but more in-depth further studies are needed to identify whether this proteins may serve as biomarkers of poor immune recovery and whether this could have any chance of being used as a therapeutic target to achieve treatment success.

8. REFERENCES

1. Barré-Sinoussi, F. *et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**, 868–871 (1983).
2. UNAIDS Global report 2010 | UNAIDS Report on the global AIDS epidemic 2010. https://www.unaids.org/globalreport/Global_report.htm.
3. Medical, D. & Med, D. DOCTOR OF MEDICAL SCIENCE DANISH MEDICAL JOURNAL. *J (Basel)* **60**, 4622 (2013).
4. Fanales-Belasio, E., Raimondo, M., Suligoi, B. & Buttò, S. HIV virology and pathogenetic mechanisms of infection: a brief overview. *Annali dell'Istituto Superiore di Sanità* **46**, 5–14 (2010).
5. Blut), G. A. C. B. (Arbeitskreis & Blood', S. 'Assessment of P. T. by. Human Immunodeficiency Virus (HIV). *Transfusion Medicine and Hemotherapy* **43**, 203 (2016).
6. Gao, F. *et al.* Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature 1999 397:6718* **397**, 436–441 (1999).
7. Esbjörnsson, J. *et al.* HIV-2 as a model to identify a functional HIV cure. doi:10.1186/s12981-019-0239-x.
8. Delgado, R. Características virológicas del VIH. *Enfermedades Infecciosas y Microbiología Clínica* **29**, 58–65 (2011).
9. Nkeze, J., Li, L., Benko, Z., Li, G. & Zhao, R. Y. Molecular characterization of HIV-1 genome in fission yeast Schizosaccharomyces pombe. *Cell and Bioscience* **5**, 1–13 (2015).
10. Levy, J. A. HIV and the pathogenesis of AIDS. *Published in **2007** in Washington DC by ASM Press* (2007).
11. Sauter, D. *et al.* Human tetherin exerts strong selection pressure on the HIV-1 group N Vpu protein. *PLoS Pathog* **8**, (2012).
12. Vicenzi, E. & Poli, G. Novel factors interfering with human immunodeficiency virus-type 1 replication in vivo and in vitro. *Tissue Antigens* **81**, 61–71 (2013).
13. Foley, B. T. *et al.* HIV Sequence Compendium 2018. (2018) doi:10.2172/1458915.
14. Musumeci, D., Riccardi, C. & Montesarchio, D. G-Quadruplex forming oligonucleotides as Anti-HIV agents. *Molecules* **20**, 17511–17532 (2015).
15. The Primate Lentivirus-Encoded Nef Protein Can Regulate Several Steps of the Viral Replication Cycle | Elsevier Enhanced Reader. https://reader.elsevier.com/reader/sd/pii/S0042682299900533?token=541F7088F338A7C37A01CE82783E5F07E9304E2DA1A4477C842469ECF8D9166AE2F11629A1835D90676F372F8983A9DD&originRegion=eu-west-1&originCreation=20220204103013.
16. Niedrig, M. *et al.* Inhibition of infectious human immunodeficiency virus type 1 particle formation by Gag protein-derived peptides. *J Gen Virol* **75 (Pt 6)**, 1469–1474 (1994).
17. Humans, I. W. G. on the E. of C. R. to. Human Immunodeficiency Viruses. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans* **67**, 31–259 (1996).
18. Chen, B. Molecular Mechanism of HIV-1 Entry. *Trends Microbiol* **27**, 878 (2019).

19. Deeks, S. G., Overbaugh, J., Phillips, A. & Buchbinder, S. HIV infection. *Nat Rev Dis Primers* **1**, (2015).
20. - HIVinfo, N. & Info, C. Offering information on HIV/AIDS treatment, prevention, and research. (2021).
21. Fanales-Belasio, E., Raimondo, M., SuligoI., B. & Buttò, S. HIV virology and pathogenetic mechanisms of infection: a brief overview. *Annali dell'Istituto superiore di sanita* **46**, 5–14 (2010).
22. HIV Virion ~ ViralZone. https://viralzone.expasy.org/5182.
23. HIVinfo | Information on HIV/AIDS Treatment, Prevention and Research | NIH. https://hivinfo.nih.gov/home-page.
24. (PDF) Oscillations in a Model for HIV Infection with Three Intracellular Delays and RTI: Delays Can Induce Viral Blips. https://www.researchgate.net/publication/242611621_Oscillations_in_a_Model_for_HIV_Infection_with_Three_Intracellular_Delays_and_RTI_Delays_Can_Induce_Viral_Blips.
25. HIV Treatment: The Basics | NIH. https://hivinfo.nih.gov/understanding-hiv/fact-sheets/hiv-treatment-basics.
26. Bavinton, B. R. *et al.* Viral suppression and HIV transmission in serodiscordant male couples: an international, prospective, observational, cohort study. *Lancet HIV* **5**, e438–e447 (2018).
27. The Lancet HIV. U=U taking off in 2017. *Lancet HIV* **4**, e475 (2017).
28. Ford, N. *et al.* Benefits and risks of rapid initiation of antiretroviral therapy. *AIDS* **32**, 17–23 (2018).
29. ANTIRETROVIRAL THERAPY. MECHANISMS OF ACTION AND - ProQuest. https://www.proquest.com/docview/2544548452.
30. Arts, E. J. & Hazuda, D. J. HIV-1 Antiretroviral Drug Therapy. *Cold Spring Harbor Perspectives in Medicine* **2**, (2012).
31. Quinn, T. C. *et al.* Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. *N Engl J Med* **342**, 921–929 (2000).
32. Baeten, J. M. *et al.* Genital HIV-1 RNA predicts risk of heterosexual HIV-1 transmission. *Sci Transl Med* **3**, (2011).
33. Røttingen, J. A., Cameron, W. D. & Garnett, G. P. A systematic review of the epidemiologic interactions between classic sexually transmitted diseases and HIV: how much really is known? *Sex Transm Dis* **28**, 579–597 (2001).
34. Parrish, N. F. *et al.* Phenotypic properties of transmitted founder HIV-1. *Proc Natl Acad Sci U S A* **110**, 6626–6633 (2013).
35. Branson, B. M. *et al.* Laboratory testing for the diagnosis of HIV infection : updated recommendations. (2014) doi:10.15620/CDC.23447.
36. Gökengin, D. *et al.* 2014 European Guideline on HIV testing. *Int J STD AIDS* **25**, 695–704 (2014).
37. Bellan, S. E., Dushoff, J., Galvani, A. P. & Meyers, L. A. Reassessment of HIV-1 Acute Phase Infectivity: Accounting for Heterogeneity and Study Design with Simulated Cohorts. *PLOS Medicine* **12**, e1001801 (2015).
38. Hollingsworth, T. D., Anderson, R. M. & Fraser, C. HIV-1 transmission, by stage of infection. *J Infect Dis* **198**, 687–693 (2008).
39. Cassell, M. M., Wilcher, R., Ramautarsing, R. A., Phanuphak, N. & Mastro, T. D. Go Where the Virus Is: An HIV Micro-epidemic Control Approach to Stop HIV Transmission. *Global Health: Science and Practice* **8**, 614 (2020).

40. Marks, G., Crepaz, N. & Janssen, R. S. Estimating sexual transmission of HIV from persons aware and unaware that they are infected with the virus in the USA. *AIDS* **20**, 1447–1450 (2006).
41. Weller, S. C. & Davis-Beaty, K. Condom effectiveness in reducing heterosexual HIV transmission. *Cochrane Database of Systematic Reviews* **2012**, (2002).
42. Smith, D. K., Herbst, J. H., Zhang, X. & Rose, C. E. Condom effectiveness for HIV prevention by consistency of use among men who have sex with men in the United States. *J Acquir Immune Defic Syndr* **68**, 337–344 (2015).
43. Remis, R. S., Alary, M., Liu, J., Kaul, R. & Palmer, R. W. H. HIV Transmission among Men Who Have Sex with Men due to Condom Failure. *PLOS ONE* **9**, e107540 (2014).
44. Aspinall, E. J. *et al.* Are needle and syringe programmes associated with a reduction in HIV transmission among people who inject drugs: a systematic review and meta-analysis. *Int J Epidemiol* **43**, 235–248 (2014).
45. Lucas, S. & Nelson, A. M. HIV and the spectrum of human disease. *J Pathol* **235**, 229–241 (2015).
46. Prabhu, S., Harwell, J. I. & Kumarasamy, N. Advanced HIV: diagnosis, treatment, and prevention. *Lancet HIV* **6**, e540–e551 (2019).
47. Saag, M. S. *et al.* Antiretroviral Drugs for Treatment and Prevention of HIV Infection in Adults: 2018 Recommendations of the International Antiviral Society-USA Panel. *JAMA* **320**, 379–396 (2018).
48. Ghosn, J., Taiwo, B., Seedat, S., Autran, B. & Katlama, C. HIV. *Lancet* **392**, 685–697 (2018).
49. Gazzola, L., Tincati, C., Bellistrì, G. M., Monforte, A. D. A. & Marchetti, G. The absence of CD4+ T cell count recovery despite receipt of virologically suppressive highly active antiretroviral therapy: clinical risk, immunological gaps, and therapeutic options. *Clin Infect Dis* **48**, 328–337 (2009).
50. Corbeau, P. & Reynes, J. Immune reconstitution under antiretroviral therapy: the new challenge in HIV-1 infection. *Blood* **117**, 5582–5590 (2011).
51. Battegay, M., Nüesch, R., Hirschel, B. & Kaufmann, G. R. Immunological recovery and antiretroviral therapy in HIV-1 infection. *Lancet Infect Dis* **6**, 280–287 (2006).
52. van Lelyveld, S. F. L. *et al.* Long-term complications in patients with poor immunological recovery despite virological successful HAART in Dutch ATHENA cohort. *AIDS* **26**, 465–474 (2012).
53. Takuva, S. *et al.* Poor CD4 recovery and risk of subsequent progression to AIDS or death despite viral suppression in a South African cohort. *J Int AIDS Soc* **17**, (2014).
54. Engsig, F. N. *et al.* Long-term mortality in HIV-positive individuals virally suppressed for >3 years with incomplete CD4 recovery. *Clin Infect Dis* **58**, 1312–1321 (2014).
55. Pacheco, Y. M. *et al.* Increased risk of non-AIDS-related events in HIV subjects with persistent low CD4 counts despite cART in the CoRIS cohort. *Antiviral Res* **117**, 69–74 (2015).
56. Baker, J. v. *et al.* Poor initial CD4+ recovery with antiretroviral therapy prolongs immune depletion and increases risk for AIDS and non-AIDS diseases. *J Acquir Immune Defic Syndr* **48**, 541–546 (2008).

57. Young, J. *et al.* CD4 cell count and the risk of AIDS or death in HIV-Infected adults on combination antiretroviral therapy with a suppressed viral load: a longitudinal cohort study from COHERE. *PLoS Med* **9**, (2012).
58. Yang, X. *et al.* Incomplete immune reconstitution in HIV/AIDS patients on antiretroviral therapy: Challenges of immunological non-responders. *J Leukoc Biol* **107**, 597–612 (2020).
59. Buggert, M. *et al.* Multiparametric bioinformatics distinguish the CD4/CD8 ratio as a suitable laboratory predictor of combined T cell pathogenesis in HIV infection. *J Immunol* **192**, 2099–2108 (2014).
60. Lu, W. *et al.* CD4:CD8 ratio as a frontier marker for clinical outcome, immune dysfunction and viral reservoir size in virologically suppressed HIV-positive patients. *J Int AIDS Soc* **18**, (2015).
61. Serrano-Villar, S. *et al.* HIV-infected individuals with low CD4/CD8 ratio despite effective antiretroviral therapy exhibit altered T cell subsets, heightened CD8+ T cell activation, and increased risk of non-AIDS morbidity and mortality. *PLoS Pathog* **10**, (2014).
62. Gunda, D. W., Kilonzo, S. B., Kamugisha, E., Rauya, E. Z. & Mpondo, B. C. Prevalence and risk factors of poor immune recovery among adult HIV patients attending care and treatment centre in northwestern Tanzania following the use of highly active antiretroviral therapy: a retrospective study. *BMC Res Notes* **10**, (2017).
63. Utay, N. S. & Hunt, P. W. Role of immune activation in progression to AIDS. *Curr Opin HIV AIDS* **11**, 131–137 (2016).
64. Rodríguez-Gallego, E. *et al.* A baseline metabolomic signature is associated with immunological CD4+ T-cell recovery after 36 months of antiretroviral therapy in HIV-infected patients. *AIDS* **32**, 565–573 (2018).
65. Massanella, M., Negredo, E., Clotet, B. & Blanco, J. Immunodiscordant responses to HAART--mechanisms and consequences. *Expert Rev Clin Immunol* **9**, 1135–1149 (2013).
66. Gaardbo, J. C., Hartling, H. J., Gerstoft, J. & Nielsen, S. D. Incomplete immune recovery in HIV infection: mechanisms, relevance for clinical care, and possible solutions. *Clin Dev Immunol* **2012**, (2012).
67. CM, M., SJ, N. & GS, O. Evolution of Translational Omics: Lessons Learned and the Path Forward. *Evolution of Translational Omics* (2012) doi:10.17226/13297.
68. Vailati-Riboni, M., Palombo, V. & Loor, J. J. What Are Omics Sciences? *Periparturient Diseases of Dairy Cows: A Systems Biology Approach* 1–7 (2017) doi:10.1007/978-3-319-43033-1_1.
69. "Omics" Sciences: Genomics, Proteomics, and Metabolomics | ISAAA.org. https://www.isaaa.org/resources/publications/pocketk/15/default.asp.
70. Domon, B. & Aebersold, R. Mass spectrometry and protein analysis. *Science* **312**, 212–217 (2006).
71. [PDF] Proteomics, Metabolomics and Lipidomics in Reproductive Biotechnologies: The MS Solutions | Semantic Scholar. https://www.semanticscholar.org/paper/Proteomics%2C-Metabolomics-and-Lipidomics-in-The-MS-Ferreira-Guimar%C3%A3es/68d39cfd4273d1b910b9519b860fc8d672fb226a.
72. Massanella, M. *et al.* Differential gene expression in HIV-infected individuals following ART. *Antiviral Res* **100**, 420–428 (2013).

73. Aratani, Y. Myeloperoxidase: Its role for host defense, inflammation, and neutrophil function. *Archives of Biochemistry and Biophysics* **640**, 47–52 (2018).
74. Hensley-McBain, T. & Klatt, N. R. The Dual Role of Neutrophils in HIV Infection. *Curr HIV/AIDS Rep* **15**, 1 (2018).
75. Detección de anticuerpos contra los antígenos de diferenciación tumoral proteinasa 3 (PR3) y mieloperoxidasa (MPO) en la leucemia promielocítica. http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-02892010000200004.
76. Clerc, I. *et al.* Entry of glucose- and glutamine-derived carbons into the citric acid cycle supports early steps of HIV-1 infection in CD4 T cells. *Nat Metab* **1**, 717–730 (2019).
77. Sáez-Cirión, A. & Sereti, I. Immunometabolism and HIV-1 pathogenesis: food for thought. *Nat Rev Immunol* **21**, 5–19 (2021).
78. Rosado-Sánchez, I. *et al.* Glutaminolysis and lipoproteins are key factors in late immune recovery in successfully treated HIV-infected patients. *Clin Sci (Lond)* **133**, 997–1010 (2019).