Visualization of biological changes to identify significant proteins

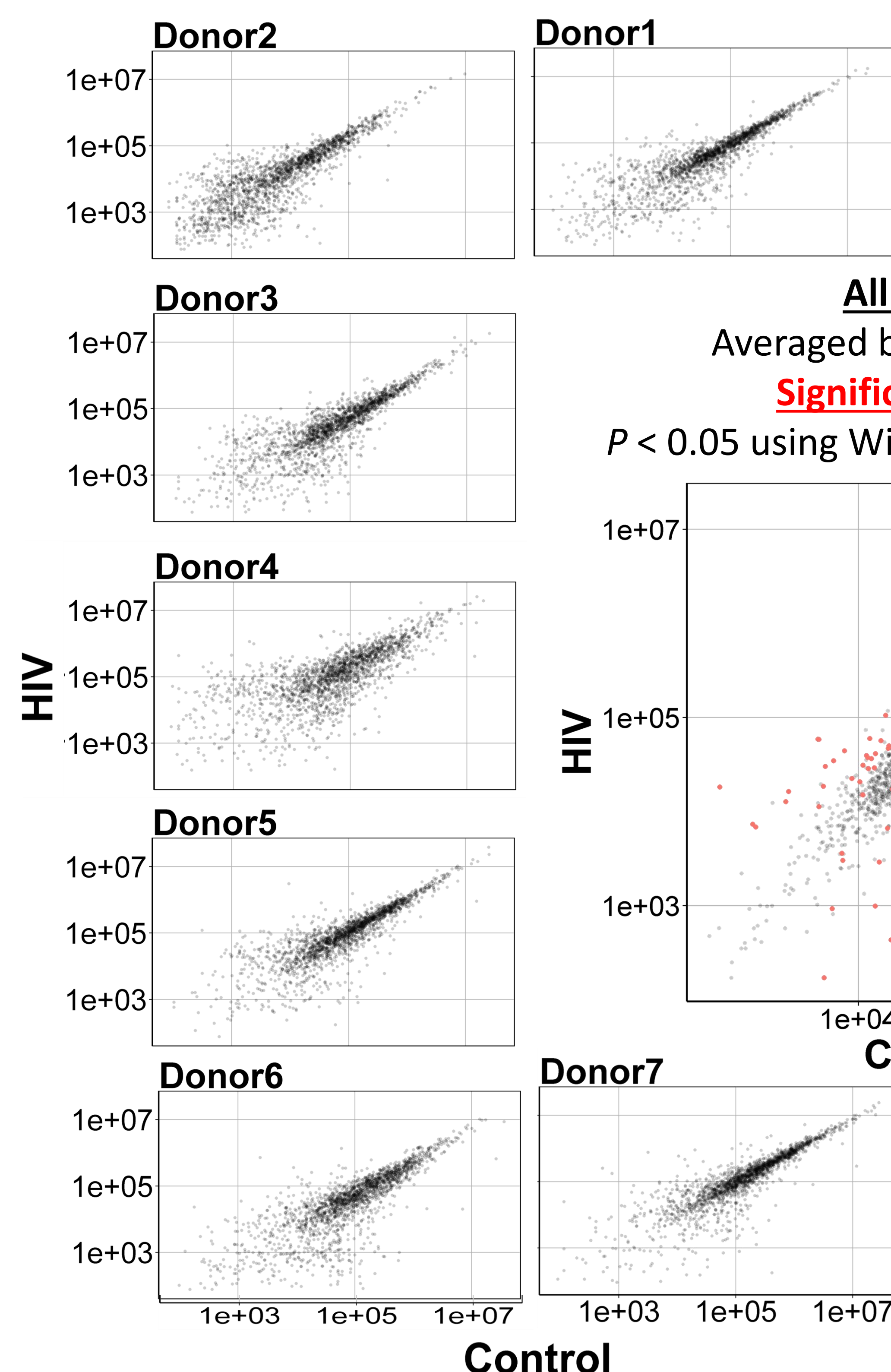Davina Faimon[1], Kristin (Reinsvold) Frederick[2], Joshua Wittenbach[1]

[1] Department of Mathematics, University of Nebraska at Omaha, Omaha, NE
[2] Department of Pharmacology and Experimental Neuroscience, University of Nebraska Medical Center, Omaha, NE
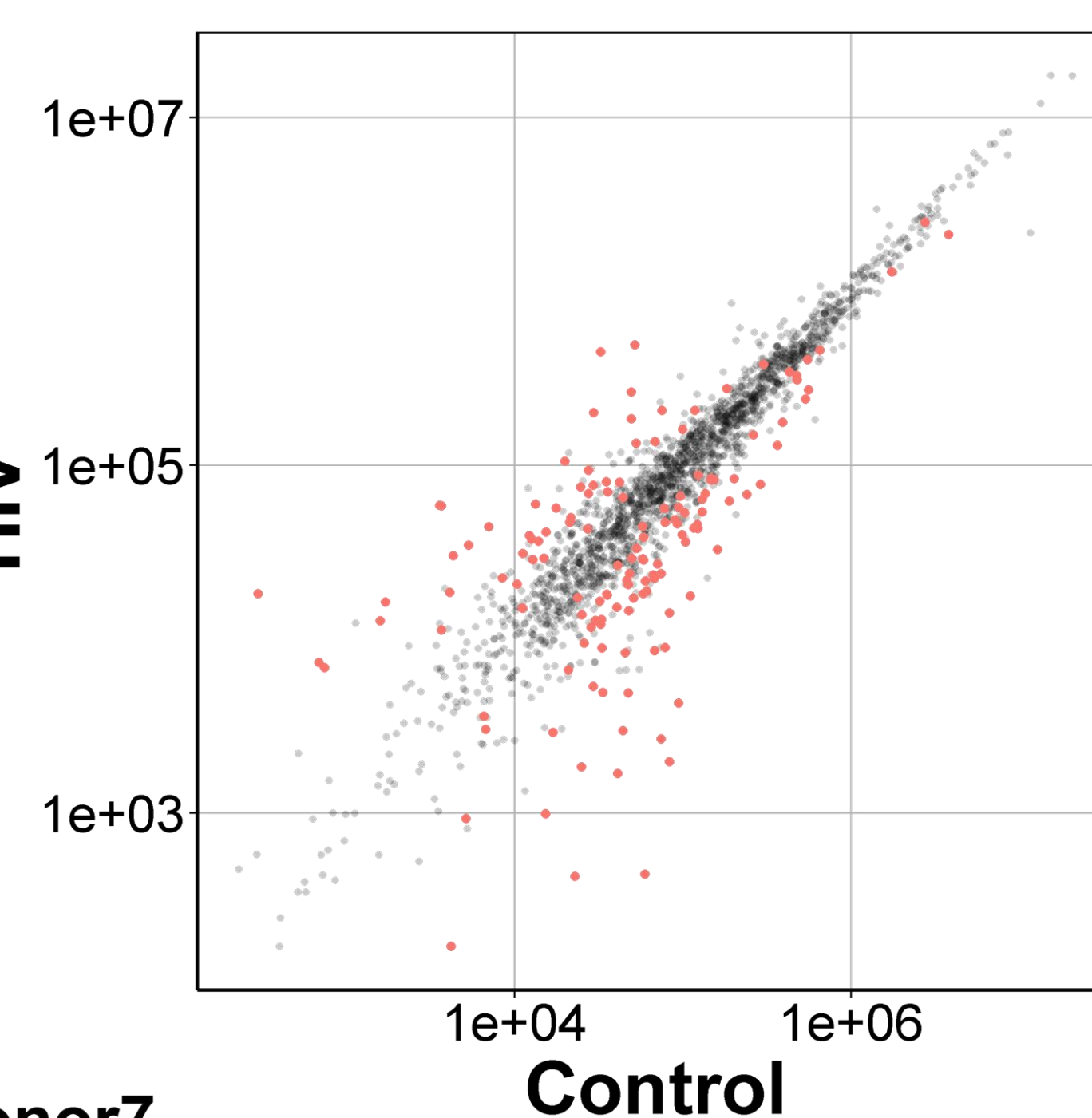
Raw data

All proteins
Raw data for each sample

Each protein from each donor was graphed to compare protein amounts in the control condition (X axis) and HIV condition (Y axis).

Proteins show linear correlation between control and HIV conditions, although it is an imperfect correlation.

The problem: Which proteins are significant?

All proteins
Averaged between samples
Significant proteins
$P < 0.05$ using Wilcoxon signed rank test

Statistical analysis

The mean amount for each protein from each condition was graphed to compare protein amounts in the control condition (X axis) and HIV condition (Y axis). Red points indicate proteins with significant difference as shown by Wilcoxon signed rank test.

Statistical analysis reduced the number of proteins from 2,150 to 139, which is statistically significant but not yet practical for experimental validation.

## The data: Codifying proteomic changes

- A major goal of proteomic research is to identify and quantify changes in protein levels that are critical for a thorough understanding of normal (developmental) as well as pathological processes.
- This knowledge will lead the way for development of diagnostic, prognostic, therapeutic, and preventive medical applications.
- Mass spectrometry is one of the most common research tools used for identification and quantification of proteins.

- Monocyte derived macrophages (MDMs) are integral cells of the human immune system, both innate and adaptive immunity.
- In particular, HIV infection of MDMs is a critically important component of HIV pathogenesis and progression to AIDS.

- For the present study, samples of human MDMs were collected from seven healthy donors.
- One sample from each donor was infected with HIV and one sample from each donor remained uninfected as the negative control.
- Mass spectrometry was performed to quantify amounts of each protein in the samples.
- This data was generated by Dr. Pawel Ciborowski's proteomics laboratory at UNMC.

- The goal is to identify proteins with altered levels involved in HIV-1 infection of MDMs that could serve as novel candidates in the development of novel antiretroviral treatment.

The solution: Data visualization
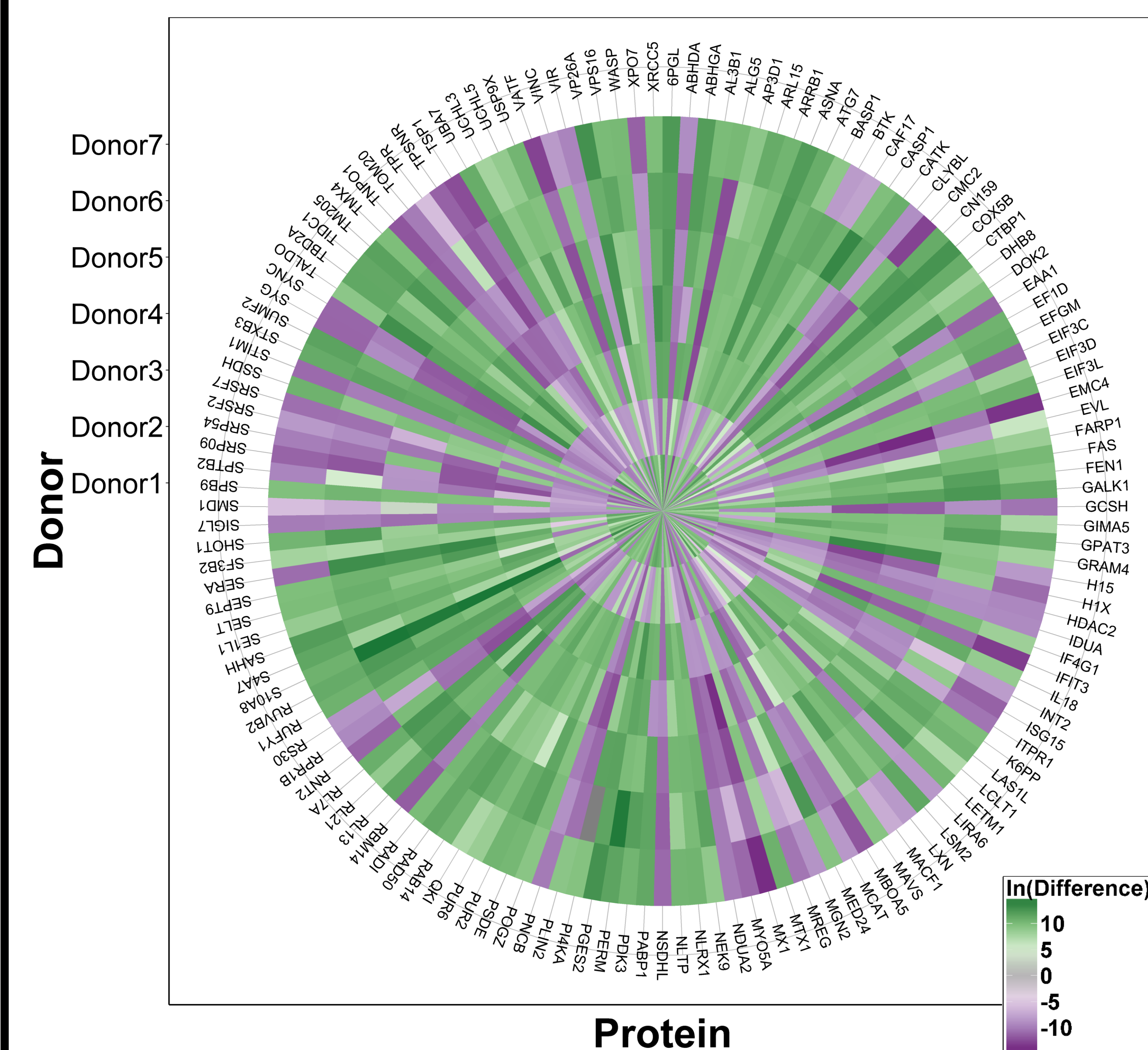
The difference in protein amounts for each statistically significant protein (φ) was graphed to compare differences between samples (ρ). Each sector is protein with each colored tile representing a donor.

Sectors with severe color spikes indicate potentially significant proteins to investigate.
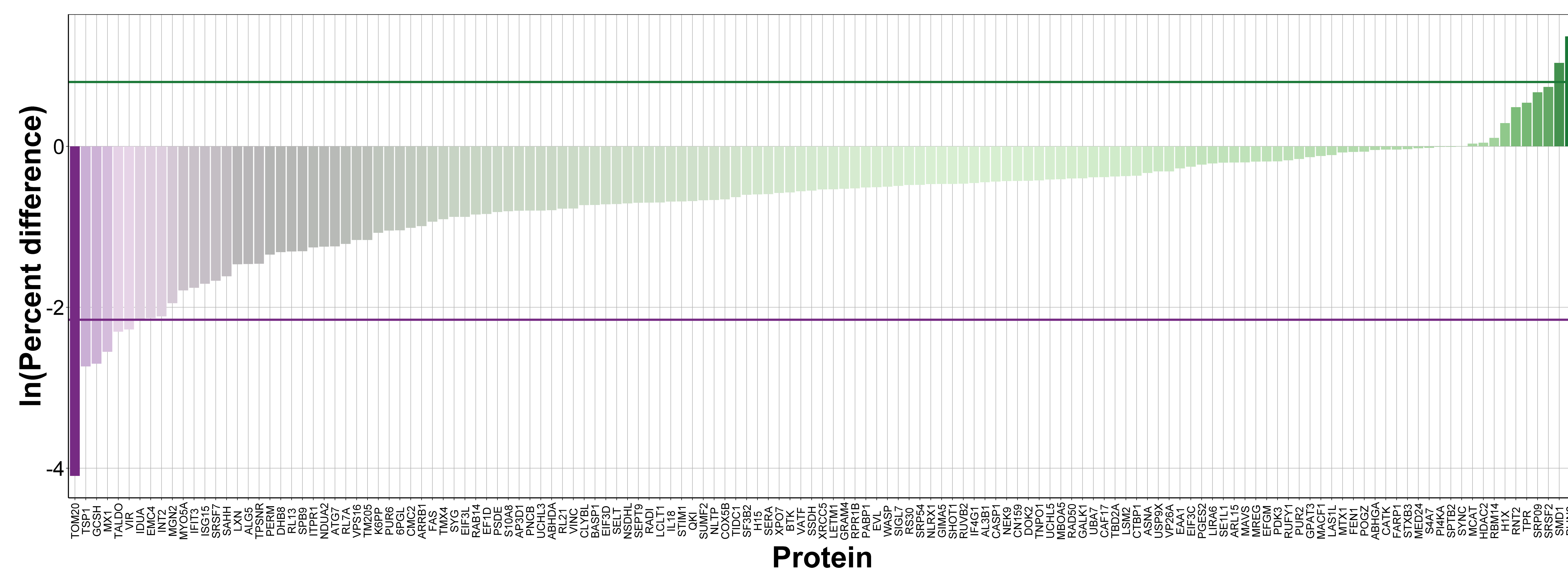
The mean percent difference (Y axis) for each statistically significant protein (X axis) was graphed to compare relative increase or decrease between control and HIV conditions. Horizontal lines indicate two standard deviations above and below the mean.

The eight proteins outside two standard deviations are potentially the most biologically sensitive to HIV infection and warrant further investigation.
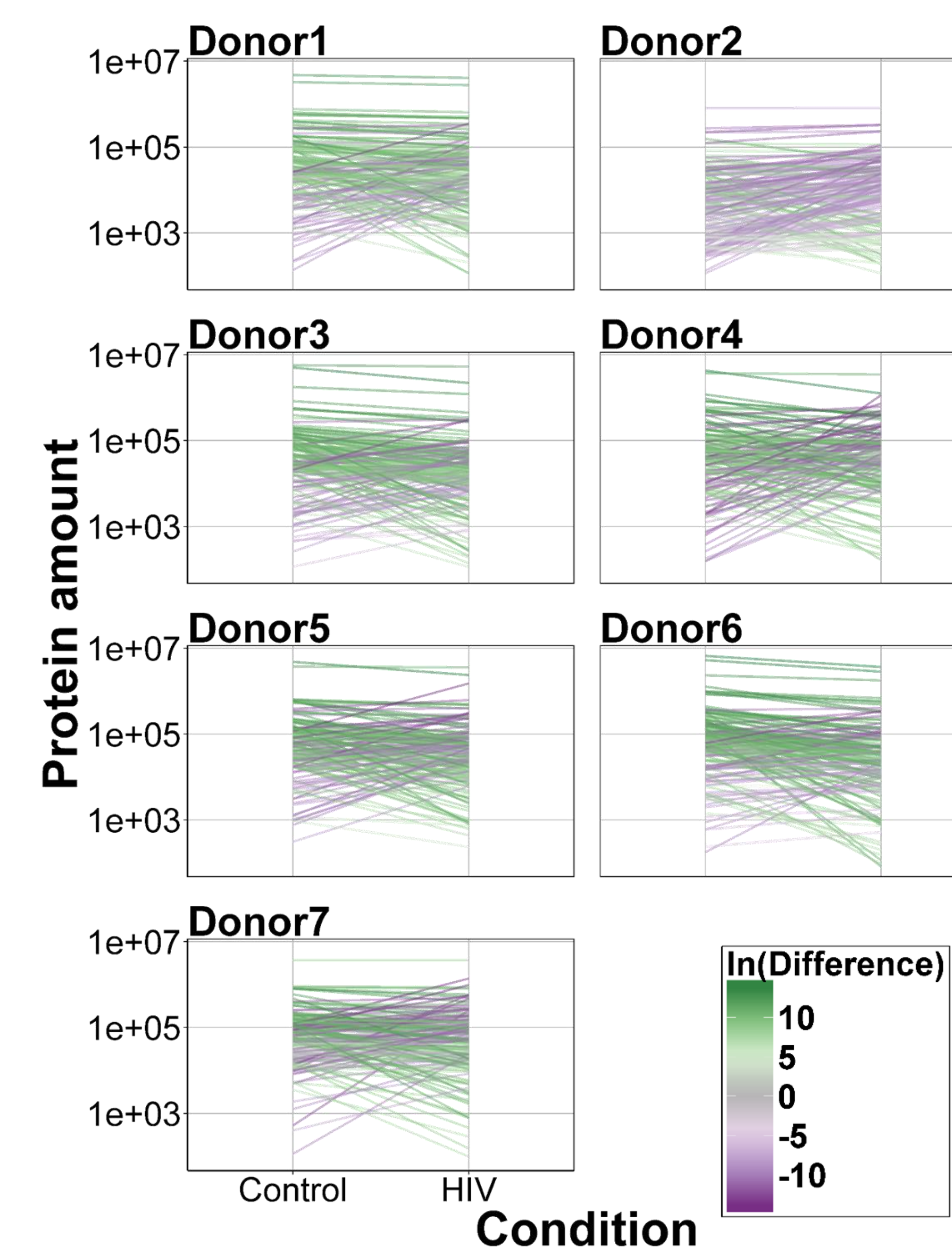
The amount for each statistically significant protein from sample was graphed to compare the differences in protein amounts (Y axis) for each condition (X axis).

Lines with greater slope indicate proteins with a greater amount of change between conditions. Green indicates the protein amount decreased after HIV infection while purple indicates an increase in protein amount.