SUPPLEMENTARY INFORMATION

QUANTITATIVE HDL PROTEOMICS IDENTIFIES PEROXIREDOXIN-6 AS A BIOMARKER OF HUMAN ABDOMINAL AORTIC ANEURYSM

*Elena Burillo1# , Inmaculada Jorge2# , Diego Martínez-López1 , Emilio Camafeita2 , Luis Miguel Blanco-Colio¹ , Marco Trevisan-Herraz2 , Iakes Ezkurdia² , Jesús Egido³ , Jean-Baptiste Michel4 , Olivier Meilhac5 , Jesús Vázquez2# * and Jose Luis Martin-Ventura1# **

In alphabetical order; these authors contributed equally to this work

Supplementary Methods

HDL isolation

Lipoproteins were isolated from individual EDTA plasma samples by ultracentrifugation. Plasma density was adjusted to $p=1.063$ g/mL with KBr and the sample was overlaid with KBr saline solution ($p=1.063$ g/mL). Samples were ultracentrifuged at 53000 rpm for 18 h at 4°C in a TLA-100 rotor (Beckman). After centrifugation, the upper, apolipoprotein-B-containing fraction was recovered and stored. The bottom fraction, containing HDLs, was adjusted to 1.25 g/mL with KBr and overlaid with KBr saline solution (ρ=1.22 g/mL). The HDL-containing fraction was then ultracentrifuged at 63000 rpm for 24 h at 4°C. After this step, HDLs, in the top layer, were recovered as a single band and were extensively rinsed with saline and concentrated using a centrifugal concentrating device (cutoff 10 kDa). All HDL samples were desalted by centrifugation and 3 washes with saline.

Proteomic study

Experimental design

In each experiment, samples from 2 AAA patients and 1 control subject were compared with an internal control, prepared by pooling protein extracts from all subjects of the study. All the comparatives included independent biological preparations, making a total 14 comparatives between AAA samples vs internal control sample and 7 comparatives between control samples vs internal control sample. The use of internal control allowed comparing data from different individuals in different experiments, as described in the Methods section (FIGURE 1). Only 31 outlier proteins showed significant abundance differences at a 1% FDR threshold in all comparisons of the 7 integration experiments. Proteins identified in at least three independent comparatives at 1% false discovery rate (FDR) were considered for the analysis. Quantitative data were analysed using the weighted spectrum, peptide, and protein (WSPP) model 1 , which allowed decomposition of the total variance into the spectral, peptide, and protein variance components. The distribution of quantitative data at the 3 levels, spectrum, peptide and protein, were in good agreement with the predictions of the statistical model, as depicted from the analysis of sigmoidal distributions obtained in all iTRAQ comparisons (Supplementary Figure S3). The variances at the spectrum and peptide levels were similar in all the pairwise comparisons (Supplementary Table S3), and were within the values expected according to results previously obtained in our laboratory ¹⁻⁵. A FDR threshold of 1% was set to detect significant protein-abundance changes from the *Z* values. The high inter-individual differences in protein variance observed in human HDL preparations ⁴ prompted us to use a shared all-experiment internal control as a reference for quantifications; using this approach, protein variance was consistently below 0.4, a value similar to that we obtained working with pooled human HDL samples ⁴ or human HDL samples from AAA patients ⁶ (Supplementary Figure S4).

iTRAQ labelling of peptides

The dried peptides were taken up in 15 µL of iTRAQ solution buffer provided with the iTRAQ kit (ABSciex) and labeled by adding 35 µL of the corresponding iTRAQ reagent in ethanol, followed by 1h incubation at room temperature in 70% ethanol in 180 mM triethylammonium bicarbonate (TEAB), pH 8.53. HDL samples from control subjects were labeled with iTRAQ tags 114, and HDL

samples from AAA patients were labeled with tags 115 and 116. In addition, a pool of all 21 samples was used as an internal control in every iTRAQ experiment and was labeled with tag 117. After quenching with 50 µL 0.5% TFA (v/v) for 30 min, samples were fully dried to completely stop the labeling reaction. The four labeled samples were resuspended in 100 µL 0.1% TFA (v/v), combined in one tube, and cleaned up with C18 Oasis cartridges using 50% ACN (v/v) in 0.1% TFA (v/v) as elution solvent. Samples were dried down prior to LC-MS/MS analysis.

Statistics

The $log₂$ ratio of concentrations in the two samples being compared, A and B, determined by spectrum *s* of peptide *p* derived from protein *q*, is expressed as X_{eqps} =log₂(A/B). The log₂-ratio value associated with each peptide, X_{eqp} , is then calculated as a weighted average of the spectra used to quantify the peptide, and the value associated with each protein, X_{eq} , is similarly the weighted average of its peptides. In addition, a grand mean, *Xe*, is calculated in each experiment as a weighted average of the protein values. The global distribution of values at each level is described using a standardized variable, *Z*, that expresses the quantitative values in units of standard deviation and that in the null hypothesis is expected to follow a N (0,1) normal distribution. Our statistical model was used to average protein quantifications from each comparison between control samples and the internal control and between AAA patient samples and the internal control 1.5 . These averaged protein quantifications yielded integrated data for comparison of AAA patient samples with samples from control participants (FIGURE 1, green arrows), thereby increasing the statistical power to detect protein alterations 1 . Again, the distribution of the standardized variable describing the variability between different experiments within the same protein was very close to the null hypothesis (Supplementary Figure S3), demonstrating that the quantitative results from all experiments were generally reproducible and therefore can be integrated. The final distribution of averaged protein abundance ratios between AAA patients and controls was also close to that of the null hypothesis, allowing the detection of significant protein abundance changes (Supplementary Figure S3).

Systems biology analysis

Systems biology analysis were performed using the SBT model ⁵. Proteins were classified as belonging to a functional *class* or *category*, and statistical analyses were performed to determine the relative abundance of each category by integrating the quantitative values of its protein components, after eliminating protein-category outliers. Proteins were functionally annotated using the Ingenuity Knowledge Database ^{7,8}, CORUM ⁹, and DAVID ^{10,11}. Network analysis were obtained using STRINGS 12 .

References

- 1 Navarro, P. *et al.* General statistical framework for quantitative proteomics by stable isotope labeling. *J Proteome Res* 13, 1234-1247, doi:10.1021/pr4006958 (2014).
- 2 Bonzon-Kulichenko, E. et al. A robust method for quantitative high-throughput analysis of proteomes by 18O labeling. *Mol Cell Proteomics* 10, M110 003335, doi:M110.003335 [pii] 10.1074/mcp.M110.003335 (2011).
- 3 Jorge," I.*! et! al.*" Statistical" model" to" analyze" quantitative" proteomics" data" obtained by 18O/16O labeling and linear ion trap mass spectrometry: application to the study of vascular endothelial growth factor-induced angiogenesis in endothelial cells. Mol Cell Proteomics 8, 1130-1149, doi:M800260-MCP200 [pii] 10.1074/mcp.M800260-MCP200 (2009).
- 4 Jorge, I. *et al.* The human HDL proteome displays high inter-individual variability and is altered dynamically in response to angioplasty-induced

atheroma plaque rupture. J Proteomics 106, 61-73, doi:10.1016/j.jprot.2014.04.010 (2014).

- 5 García-Marqués, F. *et al.* A novel systems-biology algorithm for the analysis of coordinated" protein" responses" using" quantitative" proteomics." M*ol! Cell! Proteomics,*"doi:10.1074/mcp.M115.055905"(2016).
- 6 Burillo, E. et al. ApoA-I/HDL-C levels are inversely associated with abdominal aortic aneurysm progression. Thromb Haemost 113, doi:10.1160/TH14-10-0874 (2015).
- 7 Ficenec, D. et al. Computational knowledge integration in biopharmaceutical research. Brief Bioinform 4, 260-278 (2003).
- 8 Calvano, S. E. *et al.* A network-based analysis of systemic inflammation in humans. Nature 437, 1032-1037, doi:10.1038/nature03985 (2005).
- 9 Ruepp, A. *et al.* CORUM: the comprehensive resource of mammalian protein complexes."N*ucleic!Acids!Res!*3**6,**"D646@650,"doi:10.1093/nar/gkm936"(2008).
- 10 Huang, d. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4, 44-57, doi:10.1038/nprot.2008.211 (2009).
- 11 Huang, d. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic *Acids!Res!*3**7,**"1@13,"doi:10.1093/nar/gkn923"(2009).
- 12 Szklarczyk, D. et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43, D447-452, doi:10.1093/nar/gku1003 (2015).

Supplementary Table S1.- Proteins identification in at least three iTRAQ experiments. Number of spectra and peptide for each protein identified at 1% FDR and the percent

coverage in either are indicated. Accession number from Uniprot database.

Supplementary Table S2.- Protein abundance changes in HDL proteome from AAA vs control patients.

 $\begin{array}{|c|c|c|c|c|}\hline -2.5 & -1.5 & 0 & 1.5 & 3 \ \hline \end{array}$

decrease in AAA

incease in AAA

Accession number from the Uniprot database (http://www.uniprot.org/).

Supplementary Table S3.- Statistical parameters estimated for all the AAA/IC and C/IC comparatives from all the iTRAQ experiments

Supplementary Figure S2. Western-blot of validated proteins (full-image). Representative western-blot of PON1, PRDX6, HLA-I and APOA1, and Ponceau staining of HDLs from controls and patients. Red square indicates the band of ApoA1 used for densitometric analysis (based on their molecular weight) and used as loading control (since ApoA1 is the majority protein in HDL as observed in the ponceau staining).

Supplementary Figure S3. Integrative analysis of the HDL protein abundance changes in AAA patients with respect to control subjects. Data from quantitative iTRAQ experiments comparing control subjects with the internal control (C1, C2, C3, C4, C5, C7, and C9 vs IC) were integrated to obtain average protein values for control (C) vs IC. Data comparing AAA patients with the internal control (AAA1, AAA2, AAA3, AAA4, AAA5, AAA6, AAA7, AAA8, AAA9, AAA10, AAA11, AAA12, AAA13, and AAA14 vs IC) were integrated to obtain average protein values for AAA vs IC. Finally, quantitative data for comparison of AAA vs C were obtained after integration of the comparisons of C vs IC and AAA vs IC. Plots represent the cumulative distribution of the standardized variable in logarithmic scale (log₂-ratios) at the different level of the integration data.

Supplementary Figure S4. Biological variability of individual and pooled samples in different experiments. Individual human plasma samples from AAA patients and control subjects were compared with a pooled internal control in different iTRAQ experiments (Burillo E, Lindholt JS, Molina-Sánchez P,et al. Thromb Haemost. 2015 May 26;113(6):1335-46) (green). Pooled HDL samples from CAD patients before and after angioplasty were compared in O16/O18 experiments (red). Individual HDL samples from two CAD patients before and after angioplasty were compared in different iTRAQ experiments (blue) (Jorge I, Burillo E, J. Proteomics, 2014). In the current iTRAQ study, individual HDL samples from AAA patients and control subjects were compared with the pooled internal control (yellow).