REFERENCES

Additional references added as on line.

1. Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, et al. PANTHER: a library of protein families and subfamilies indexed by function. Genome Res. 2003;13(9):2129-41.
2. Mi H, Dong Q, Muruganujan A, Gaudet P, Lewis S, Thomas PD. PANTHER version 7: improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium. Nucleic Acids Res. 2010;38(Database issue):D204-10.
3. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25(1):25-9.
4. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009;10:48.
5. Hojlund K, Bowen BP, Hwang H, Flynn CR, Madireddy L, Geetha T, et al. In vivo phosphoproteome of human skeletal muscle revealed by phosphopeptide enrichment and HPLC-ESI-MS/MS. Journal of proteome research. 2009;8(11):4954-65.

Appendix as on line section

Proteomics

Proteins were extracted in radio-immuno-precipitation assay (RIPA) buffer (1% NP-40, 0.5% sodium de-oxycholate, 0.1% SDS, in PBS) containing phosphatase and protease inhibitors, and quantified using the Bradford Quick Start Protein assay following manufacturer's instructions (Bio-Rad Laboratories, USA). Aliquots of 100µg of ten samples per experiment were digested with trypsin (2.5µg trypsin per 100µg protein; 37°C, overnight), labelled with Tandem Mass Tag (TMT) ten plex reagents according to the manufacturer's protocol (Thermo Fisher Scientific, UK) and the labelled samples pooled.

For the Total proteome analysis, aliquots of 50ug of the pooled sample were evaporated to dryness and re-suspended in buffer A (20 mM ammonium hydroxide, pH 10) prior to fractionation by high pH reversed-phase (RP) chromatography using an Ultimate 3000 liquid chromatography system (Thermo Fisher Scientific). The sample was loaded onto an XBridge BEH C18 Column (130Å, 3.5 µm, 2.1 mm X 150 mm, Waters, UK) in buffer A and peptides eluted with an increasing gradient of buffer B (20 mM Ammonium Hydroxide in acetonitrile, pH 10) from 0-95% over 60 minutes. The resulting fractions were evaporated to dryness and re-suspended in 1% formic

acid prior to analysis by nano-LC MSMS using an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific).

For the Phospho-proteome analysis, the remainder of the TMT-labelled pooled sample was evaporated to dryness and subjected to titanium oxide-based phosphopeptide enrichment according to the manufacturer's instructions (Pierce). The phospho-enriched sample was evaporated to dryness and then re-suspended in 1% formic acid prior to analysis by nano-LC MSMS using an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific).

Nano-LC Mass Spectrometry

High pH RP fractions (Total proteome analysis) or the phospho-enriched fraction (Phospho-proteome analysis) were further fractionated using an Ultimate 3000 nano-HPLC system in line with an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific). In brief, peptides in 1% (vol/vol) formic acid were injected onto an Acclaim PepMap C18 nano-trap column (Thermo Scientific). After washing with 0.5% (vol/vol) acetonitrile 0.1% (vol/vol) formic acid peptides were resolved on a 250 mm × 75 μm Acclaim PepMap C18 reverse phase analytical column (Thermo Scientific) over a 150 min organic gradient, using 7 gradient segments (1-6% solvent B over 1min, 6-15% B over 58min, 15-32%B over 58min, 32-40%B over 5min, 40-90%B over 1min, held at 90%B for 6min and then reduced to 1%B over 1min) with a flow rate of 300 nl min^{-1}. Solvent A was 0.1% formic acid and Solvent B was aqueous 80% acetonitrile in 0.1% formic acid. Peptides were ionized by nano-electrospray ionization at 2.0kV using a stainless steel emitter with an internal diameter of 30 μm (Thermo Scientific) and a capillary temperature of 275°C.

All spectra were acquired using an Orbitrap Fusion Tribrid mass spectrometer controlled by Xcalibur 2.0 software (Thermo Scientific) and operated in data-dependent acquisition mode using an SPS-MS3 workflow. FTMS1 spectra were collected at a resolution of 120,000, with an

automatic gain control (AGC) target of 200,000 and a maximum injection time of 50ms. The Top N most intense ions were selected for MS/MS. Precursors were filtered according to charge state (to include charge states 2-7) and with mono-isotopic precursor selection. Previously interrogated precursors were excluded using a dynamic window (40s +/-10ppm). The MS2 precursors were isolated with a quadrupole mass filter set to a width of 1.2m/z. ITMS2 spectra were collected with an AGC target of 5,000, max injection time of 120ms and CID collision energy of 35%.

For FTMS3 analysis, the Orbitrap was operated at 60,000 resolution with an AGC target of 50,000 and a max injection time of 120ms. Precursors were fragmented by high-energy collision dissociation (HCD) at normalised collision energy of 55% to ensure maximal TMT reporter ion yield. Synchronous Precursor Selection (SPS) was enabled to include up to 5 MS2 fragment ions in the FTMS3 scan.

Data Analysis

The raw data files were processed and quantified using Proteome Discoverer software v1.4 (Thermo Scientific) and peptide sequences searched against the Uniprot Human database (134169 sequences) using the SEQUEST algorithm. Peptide precursor mass tolerance was set at 10ppm, and MS/MS tolerance was set at 0.6Da. Search criteria included oxidation of methionine (+15.9949) as a variable modification and carbamido-methylation of cysteine (+57.0214) and the addition of the TMT mass tag (+229.163) to peptide N-termini and lysine as fixed modifications. For the Phospho analysis, phosphorylation of serine, threonine and tyrosine (+79.966) were also included as variable modifications. Searches were performed with full tryptic digestion and a maximum of 1 missed cleavage was allowed. The reverse database search option was enabled and all peptide data was filtered to satisfy false discovery rate (FDR) of 5%. Only proteins that were detected in all samples were used in any analysis (accession numbers

representing cDNA were excluded). Values for each of the proteins identified are presented as a ratio to the internal standard (a pool of all samples), and represent the median of the measured peptide(s) for each protein.

High-fold differences (fold decrease/increase greater than 1.3) were determined and a log-2 transformation applied. Significance between groups was determined by an unpaired t-test (VSD compared with TOF), and –log 10 transformation applied. Only significant fold-changes ($p<0.05$) were deemed of importance.

Gene name, derived by mapping protein accession numbers within the Uniprot database from 'UniprotKB AC/ID' to 'Gene name', shall frequently be used synonymously with protein name throughout this paper (see Table 6).

Accession numbers were inputted to the Panther classification system (Database Version 11.1 (released 2016-10-24) (1, 2) to obtain functional classification information (molecular function - molecular activities; cell component - where gene products are active; biological process - pathways and larger processes made up of the activities of multiple gene products (3). Significantly altered protein accession numbers between pathologies were inputted into the 'Gene Ontology enRIchment anaLysis and visuaLizAtion' tool (GOrilla, Database update v.Feb 4 2017, (4)) to determine enriched gene ontology (GO) terms, versus a background list comprised of all proteins detected during the proteomics and phosphoproteomics analysis. A p-value threshold of 0.001 was set. Proteins were searched against those published in literature (5) and using gene ontology terms to determine any structural, or extracellular matrix proteins.