SUPPLEMENTARY APPENDIX

Proteomic analysis of plasma from children with sickle cell anemia and silent cerebral infarction

Sanjay Tewari,[1] George Renney,[2] John Brewin,[1] Kate Gardner,[1] Fenella Kirkham,[3] Baba Inusa,[4] James E Barrett,[5] Stephan Menzel,[1] Swee Lay Thein,[6] Malcolm Ward[2] and David C. Rees[1]

[1]Red Cell Biology Unit, King's College Hospital, King's College London, UK; [2]Proteomics Laboratory, Institute of Psychiatry, King's College London, UK; [3]Department of Neurosciences, Institute of Child Health, University College Hospital, London, UK; [4]Evelina Children's Hospital, Guy's and St Thomas' Hospital, London, UK; [5]Division of Health & Social Care Research, King's College London, UK; [6]Sickle Cell Branch, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD, USA

 doi:10.3324/haematol.2018.187815

Received: January 10, 2018.

Accepted: March 14, 2018.

Pre-published: March 15, 2018.

Correspondence: david.rees2@nhs.net.

Supplementary Materials

Brain MRI and Neurological Examination

Many of the eligible children for this study had previously had brain MRI for routine clinical or research purposes. MRI results were considered eligible for entry into this study if performed in the 12 months preceding the date of consent. Children known to have SCIs based on previous MRIs were selectively recruited to the study. Children without an MRI in the last 12 months had an MRI scan performed specifically as part of this study within 4 weeks of blood sampling, and were assigned to either the control or SCI group. A pediatric neuroradiologist with experience of interpreting MRIs in children with sickle cell disease reported all scans. SCI was defined as a signal abnormality visible on axial and coronal planes on the T2-weighted images, which is at least 3mm in one dimension, without a history of overt stroke or a neurological deficit corresponding to the MRI lesion.

Proteomic Methods

TMT®s (Proteome Sciences plc) are a set of isobaric mass tags for labeling proteins and peptides at amine functions and subsequent mixing of up to ten different protein samples. During MS/MS, each set of tags gives rise to different reporter ions with specific molecular masses thus allowing them to share the same chemical structure but have different molecular weights dependent on the different number of heavy isotopes contained. Peptides labelled with different TMTs show identical chromatographic retention times as well as identical ionization and fragmentation behavior of tags and tagged species.

The proteins were initially digested into peptides using trypsin, and TMT labeling then performed. Off gel fractionation was then used to separate the peptides prior to LC-MS/MS. Protein discovery study was performed in a reference-controlled manner and this approach generated considerable coverageof the plasma proteome(45).

Mass spectral data were acquired using an Orbitrap Velos mass spectrometer (Thermo Scientific). Compilation and validation of the peptide and protein assignments were carried out using Proteome Discoverer v1.3 (Thermo) and Scaffold v2.6 (Proteome Software Inc) mass spectrometry software. Proteins were quantified based on the intensities of the reporter ions released from the TMT tags during MS/MS of peptides. The use of the TMT10plex reagents enabled the simultaneous analysis of five samples alongside an overall study reference comprising of a pool of all subjects. The proteomics study comprised 5 separate TMT10plex experiments, each with 10 samples and a reference channel for normalization purposes.

Data were exported for each TMT set into an Excel file, and duplicate peptides were removed. The data were filtered so that only peptides identified in all 10 sample channels were analysed further. The data were normalized and merged against the peptide sequence using R software. The merged data were further filtered such that only peptides present in all patient samples were studied. Relative amounts of each peptide were compared between the SCI and control groups, and P values calculated, for those samples with >1.3 fold differences.

Neurological biomarkers included multiplexed proximity extension assay

Plasma samples, anticoagulated with EDTA, were frozen at -80^{o}C within 6 hours of venesection, and transported as one batch to Olink Proteomics in Sweden (www.olink.com/proseek-multiplex) on dry ice for analysis using their Proseek Multiplex Neurology I panel, which includes the following markers:

ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase 1 (CD38)

Alpha-2-macroglobulin receptor-associated protein (Alpha-2-MRAP)

BDNF/NT-3 growth factors receptor (NTRK2)

Beta-nerve growth factor (Beta-NGF)

Bone morphogenetic protein 4 (BMP-4)

Brain-derived neurotrophic factor (BDNF)

Brevican core protein (BCAN)

Brorin (VWC2)

C-type lectin domain family 1 member B (CLEC1B)

C-type lectin domain family 10 member A (CLEC10A)

Cadherin-3 (CDH3)

Cadherin-6 (CDH6)

Carboxypeptidase A2 (CPA2)

Carboxypeptidase M (CPM)

Cathepsin S (CTSS)

Cell adhesion molecule 3 (CADM3)

Cell surface glycoprotein CD200 receptor 1 (CD200R1)

CMRF35-like molecule 1 (CLM-1)

CMRF35-like molecule 6 (CLM-6)

Contactin-5 (CNTN5)

Cytotoxic and regulatory T-cell molecule (CRTAM)

Dickkopf-related protein 4 (Dkk-4)

Dipeptidyl peptidase 1 (CTSC)

Disintegrin and metalloproteinase domain-containing protein 22 (ADAM 22)

Disintegrin and metalloproteinase domain-containing protein 23 (ADAM 23)

Draxin (DRAXIN)

Ephrin type-B receptor 6 (EPHB6)

Ephrin-A4 (EFNA4)

Epithelial discoidin domain-containing receptor 1 (DDR1)

Ezrin (EZR)

Fc receptor-like protein 2 (FcRL2)

Galectin-8 (gal-8)

GDNF family receptor alpha-1 (GFR-alpha-1)

GDNF family receptor alpha-3 (GDNFR-alpha-3)

Glial cell line-derived neurotrophic factor (GDNF)

Glypican-5 (GCP5)

Granulocyte Colony-Stimulating Factor (G-CSF)

Granulocyte-macrophage colony-stimulating factor receptor subunit alpha (GM-CSF-R-alpha)

Granzyme A (GZMA)

Growth/differentiation factor 8 (GDF-8)

Hydroxyacylglutathione hydrolase, mitochondrial (HAGH)

Interleukin-12 subunit beta, Interleukin-12 subunit alpha (IL-12B, IL-12A)

Interleukin-5 receptor subunit alpha (IL-5R-alpha)

Junctional adhesion molecule B (JAM-B)

Kynureninase (KYNU)

Latexin (LXN)

Layilin (LAYN)

Leucine-rich repeat transmembrane protein FLRT2 (FLRT2)

Leukocyte-associated immunoglobulin-like receptor 2 (LAIR-2)

Linker for activation of T-cells family member 1 (LAT)

Lysosome membrane protein 2 (SCARB2)

Macrophage scavenger receptor types I and II (MSR1)

MAM domain-containing glycosylphosphatidylinositol anchor protein 1 (MDGA1)

Matrilin-3 (MATN3)

Mesencephalic astrocyte-derived neurotrophic factor (MANF)

N-acylethanolamine-hydrolyzing acid amidase (NAAA)

Neprilysin (NEP)

Netrin receptor UNC5C (UNC5C)

Neuroblastoma suppressor of tumorigenicity 1 (NBL1)

Neurocan core protein (NCAN)

Neuronal cell adhesion molecule (Nr-CAM)

Neuropilin-2 (NRP2)

Neutral ceramidase (N-CDase)

Nicotinamide/nicotinic acid mononucleotide adenylyltransferase 1 (NMNAT1)

NKG2D ligand 2 (N2DL-2)

NT-3 growth factor receptor (NTRK3)

OX-2 membrane glycoprotein (CD200)

Platelet-derived growth factor receptor alpha (PDGF-R-alpha)

Plexin-B1 (PLXNB1)

Plexin-B3 (PLXNB3)

Poliovirus receptor (PVR)

Protogenin (PRTG)

R-spondin-1 (RSPO1)

Repulsive guidance molecule A (RGMA)

RGM domain family member B (RGMB)

Roundabout homolog 2 (ROBO2)

Scavenger receptor class A member 5 (SCARA5)

Scavenger receptor class F member 2 (SCARF2)

Secreted frizzled-related protein 3 (sFRP-3)

Serine/threonine-protein kinase receptor R3 (SKR3)

Sialic acid-binding Ig-like lectin 9 (Siglec-9)

Sialoadhesin (SIGLEC1)

SPARC-related modular calcium-binding protein 2 (SMOC2)

Sphingomyelin phosphodiesterase (SMPD1)

Tenascin-R (TN-R)

Testican-1 (SPOCK1)

Thy-1 membrane glycoprotein (THY 1)

Transmembrane protease serine 5 (TMPRSS5)

Tumor necrosis factor receptor superfamily member 12A (TNFRSF12A)

Tumor necrosis factor receptor superfamily member 21 (TNFRSF21)

Tumor necrosis factor receptor superfamily member 27 (EDA2R)

WAP, Kazal, immunoglobulin, Kunitz and NTR domain-containing protein 1 (WFIKKN1)

Genetic Association study

The adult cohort used in this part of the study had previously had genome-wide genotyping array performed using Illumina's Infinium MEGA chip (Multi Ethnic Genotyping Array). Genotyping data were quality controlled and imputed to 1000 Genomes phase 3 reference set using the online Michigan server. 359 individuals with SCD had genotyping as well as MRI data available, and were classified as having SCI or not; patients with large vessel vasculopathy, overt stroke or no brain MRI were excluded. We performed a mixed modeling analysis incorporating a genetic relatedness matrix to take account of relatedness (including population stratification), sex, age at MRI, and sickle genotype as fixed covariates. Each gene analysed contained multiple genetic variants; we corrected for multiple testing after quantifying the linkage disequilibrium within each region and used this to calculate appropriate significance levels for that region