medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

TITLE PAGE

Age-Modulated Immuno-Metabolic Proteome Profiles of Deceased Donor Kidneys Predict 12-Month Posttransplant Outcome

Authors: Philip D Charles[1,2†], Sarah Fawaz[3,4*], Rebecca H Vaughan[3,4*], Simon Davis[2], Priyanka Joshi[3], Iolanda Vendrell[2], Ka Ho Tam[5], Roman Fischer[2], Benedikt M Kessler[2], Edward J Sharples[6], Alberto Santos[1,7,8], Rutger J Ploeg[3,4], Maria Kaisar[3,4†]

Affiliations:

[1]Big Data Institute, Nuffield Department of Medicine, University of Oxford; Oxford, United Kingdom.

[2]Target Discovery Institute, Centre for Medicines Discovery, Nuffield Department of Medicine, University of Oxford; Oxford, United Kingdom.

[3]Research and Development, NHS Blood and Transplant; Bristol & Oxford, United Kingdom.

[4]Nuffield Department of Surgical Sciences and Oxford Biomedical Research Centre, University of Oxford; Oxford, United Kingdom.

[5]Engineering Department, University of Oxford; United Kingdom.

[6]University Hospital Oxford; Oxford, United Kingdom.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

[7]Center for Health Data Science, University of Copenhagen; Copenhagen, Denmark.

[8]Novo Nordisk Foundation Center for Protein Research, University of Copenhagen; Copenhagen, Denmark.

*These authors contributed equally.

†Corresponding authors. Email: philip.charles@ndm.ox.ac.uk; maria.kaisar@nds.ox.ac.uk

Keywords:

Kidney transplantation, Machine Learning, Data Independent Acquisition, Proteomics, Donor age, QUOD

ABSTRACT

Organ availability limits kidney transplantation, the best treatment for end-stage kidney disease. Deceased donor acceptance criteria have been relaxed to include older donors with higher risk of inferior posttransplant outcomes. More granular prediction models, based on deeper resolution organ assessment and understanding of damage processes, could substantially improve donor organ allocation and reduce graft dysfunction risk. Here, we profiled pre-implantation kidney biopsy proteomes from 185 deceased donors by high-resolution mass spectrometry and used machine learning to integrate and model these data, and donor and recipient clinical metadata to predict outcome. Our analysis and orthogonal validation on an

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

independent cohort revealed 136 proteins predictive of outcome, 124 proteins of which showed donor-age modulated predictive effects. Observed associations with inflammatory, catabolic, lipid metabolism and apoptotic pathways may predispose donor kidneys to suboptimal posttransplant outcomes. Our work shows that integrating kidney proteome information with clinical metadata enhances the resolution of donor kidney quality stratification, and the highlighted biological mechanisms open new research directions in developing interventions during donor management or preservation to improve kidney transplantation outcome.

TRANSLATIONAL STATEMENT

We profiled the proteome of pre-implantation biopsies selected from donor kidneys on the basis of paired 12-month graft function. Our data reveal a signature of proteins which contribute to transplant outcomes, many of these show different strengths of association dependent on donor age. The biological themes of the identified candidates reinforce immuno-metabolic and catabolic mechanisms as potential contributors to donor kidney susceptibility that may reduce graft recovery after transplantation.

INTRODUCTION

Kidney transplantation is the optimal treatment for end-stage kidney disease. Compared to dialysis, transplantation increases life-expectancy, improves quality of life and is cost-effective. Limited availability of suitable donor kidneys impedes treatment of chronic kidney disease, and often prolongs dialysis, increasing morbidity and mortality. Deceased donor organ shortages, living donation decline in some countries and emerging ageing populations drive

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

increased utilization of older deceased donor kidneys, now comprising more than half of offered organs[1,2].

Ageing associates with time-dependent decline of organ function, evidenced in kidneys by histologic lesions, such as tubular atrophy, interstitial fibrosis, glomerulosclerosis, and arteriosclerosis. Older kidneys demonstrate fewer functioning glomeruli, less renal mass, podocyte dysfunction, and impaired cellular repair[3]. Glomerular diseases are more common and associated with worse outcomes in older patients[4]. Age accelerates the transition from Acute Kidney Injury (AKI) to chronic injury[5] and is an independent risk factor of graft dysfunction and loss for deceased donor kidneys[6]; furthermore, older donors are more likely to suffer from additional risk factors such as diabetes, hypertension or cardiovascular disease.

Donor age is incorporated in clinical scoring algorithms to inform kidney allocation decisions[7,8], but is insufficient to consistently predict transplant outcomes. Current front-line models incorporating further clinical factors such as terminal serum creatinine, history of hypertension and diabetes[8,9] show consistent performance across demographics but lack granular predictive accuracy[10].

Molecular analyses of biopsies plausibly offer higher resolution assessment of organ state; but require 'big picture' understanding of mechanisms associated specifically with poor outcome, rather than immediate (but potentially recoverable) acute injury. Deceased donors are frequently assessed as having sustained damage (i.e. AKI) based on serum creatinine levels[14], however this metric poorly associates with longer term outcomes[13–16].

Mass spectrometry (MS) proteomic studies can provide such a 'big picture', but have heretofore lacked cohort capacity to represent demographic diversity[17]. Advances in high-throughput techniques[18] now allow sensitivity and depth without sacrificing throughput capacity.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Developments in machine learning and nonlinear regression analyses furthermore offer tools to extract maximal knowledge from limited size experimental cohorts, with applications in disease staging, disease recurrence prediction, treatment response monitoring, and biomarker identification[19,20].

Integration of deep proteomic profiles with heterogenous clinical and demographic factors using modern statistical tools can empower the next steps toward precision medicine[21]. Here, we benefit from the granularity provided by our MS-based proteomic profiling to report age- and immunometabolism-related proteomic signatures in pre-implantation kidney biopsies associated with transplant outcomes.

METHODS

Study Design

Deceased donor pre-transplantation kidney biopsies (n=186; 1 sample excluded during data processing) were obtained from the Quality in Organ Donor (QUOD) biobank, a national multi-center UK wide bioresource of deceased donor clinical samples acquired during donor management and organ procurement. Biopsies were obtained from Donation after Brain Death (DBD) donors and Donation after Circulatory Death (DCD) donors at the back table immediately after kidney procurement.

Selection of biopsies was based on paired 12-month post-transplant outcomes. To minimize the impact of recipient factors, we only included kidneys for which the contralateral kidney was transplanted with similar outcome. Kidneys were selected to cover the outcome continuum i.e. the range of estimated Glomerular Filtration Rate (eGFR) in the recipient at 12

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

months posttransplant, from primary non-function to eGFR>80 ml/min/1.73 m^2, excluding extreme demographic or clinical factors where possible. Samples were linked to corresponding donor and recipient demographic and clinical metadata, provided by NHS Blood and Transplant National Registry.

Study Approval and Ethics statement

Informed consent from donor families was obtained prior to sample procurement. Collection of QUOD samples and research ethics approval was provided by QUOD (NW/18/0187).

Experimental Protocols and Statistical Analysis

See Supplementary Methods.

RESULTS

Donor clinical and demographic variable relevance for eGFR at 12-month posttransplant

For exploratory analysis, we considered eGFR values in two ways. Firstly, to compare clinical factors, we grouped 12 month posttransplant outcomes into tertiles; Suboptimal Outcome (SO; eGFR≤39), Intermediate Outcome (IO; 40<eGFR≤59), and Good Outcome (GO; eGFR≥60), all units ml/min/1.73 m^2 (Figure 1). We refer to this henceforth as 'stratified eGFR'. Secondly, all eGFR values (both recipient and donor) were rank-transformed so that we could model against a continuum of outcomes while mitigating extreme values, or values recorded as 0 due to graft failure. We refer to 'ranked eGFR' henceforth to indicate ranked recipient eGFR at 12 months posttransplant.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Clinical metadata illustrated that stratified eGFR donor groups gave a balanced representation of the UK donor population (Table 1), with no significant association between donor type (DBD/DCD) and ranked eGFR (t-test; p=0.2028). We investigated associations between clinical variables and stratified eGFR subgroups within each donor type, and between donor types within stratified eGFR subgroups (Supplementary Table 1). There was a significant difference in the current front-line selection score, the UK Kidney Donor Risk Index (UKKDRI)[8] between outcome groups across donor types (ANOVA F-test; DBD: p=1.298e-6; DCD: p=3.946e-7). In both donor types, the major component of UKKDRI, donor age, was significantly different between outcome subgroups with donor age in SO being older (ANOVA F-test; DBD: p=1.253e-9; DCD: p=1.196e-7). Histories of hypertension (also a component of UKKDRI) were different between subgroups in DBD (ANOVA F-test; DBD: p=0.0020; DCD: p=0.1069). Histories of diabetes (used in the US risk index, but not UKKDRI) were not significantly different (ANOVA F-test; p=0.6188; DCD: p=0.2348). Terminal serum creatinine levels were similar across outcome subgroups (ANOVA F-test; DBD: p=0.6972; DCD: p=0.6448), although within the GO group it was higher in DBD than in DCD (t-test; p=0.0443).

After imputation of missing values, we examined associations between clinical variables common to both DBD and DCD donors (Figure 2). The strongest associations with ranked eGFR were donor age (Pearson's r=-0.52), and recipient age (r=-0.28). Donor history of hypertension and cardiological disease also clustered closely due to correlation with donor age (r=0.35 and r=0.33 respectively) but had a weaker direct correlation with outcome (r=0.30 and r=0.23 respectively).

Unsupervised analysis of pretransplant kidney proteomes

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Proteomic analysis quantified 2984 protein groups with 50% or less missing values (out of 7790 identified protein groups in total) over 185 samples and 20 interspersed sample pools (Supplementary Figure 1A). Analysis of sample pools showed minimal technical variance across sample acquisition (squared mean pairwise Z-corrected Pearson's r=0.94). Six samples were paired biopsies from the left and right kidneys of three donors. These samples showed high correlation of protein intensity values between donor pairs (Pearson's r=0.71, 0.92 and 0.91; Supplementary Figure 1B).

We explored the proteomic data using Principal Component Analysis (PCA) to find underlying linear trends. Sample variance concentrated in the first two principal components (PC1: 20.01%; PC2: 13.38%; Figure 3A). K-means clustering identified 4 distinct clusters (Figure 3A) whose membership associated with donor type, with a preponderance of DBD samples towards Cluster 2 and a preponderance of DCD samples towards Cluster 4 (Figure 3B, upper left panel; p=0.0235). Clustering did not associate strongly with recipient ranked eGFR (p=0.4134), nor with donor ranked eGFR (p=0.1684), or donor age (p=0.7907) (Figure 3B, upper middle, upper right and lower left panels). There was a weakly significant association between cluster membership and donor BMI (p=0.0350) and with serum creatinine (p=0.0326) (Figure 3B, lower middle and lower right panels).

Integration of kidney proteomes with clinical metadata enhances the resolution of donor kidney quality stratification

To identify possible clinical variable-protein interaction relationships with outcome, we used machine learning (Prediction Rule Ensembles[22]; PRE) to select features from the set of quantified proteins and all donor type-independent clinical variables.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

We split our data into training and test sets, excluding the six paired kidneys, and sampling equally across stratified eGFR using a 2/3:1/3 train:test split. Test data was only used for validation (see below). The six paired kidneys were reserved as a second 'biological duplicates' test set.

A PRE model finds a minimal predictor set in the form of decision tree, linear regression and multivariate adaptive regression spline[24] rules, but does not yield an exhaustive list of predictors. We performed PRE iteratively, modelling against ranked eGFR; any proteins identified in the final ensemble model of any iteration were excluded from the dataset in future iterations, retaining only non-identified proteins plus all clinical variables. 2000 iterations of PRE generated 3282 rules across all ensembles. The most common rules involved donor age, featuring as a term in 3154 (~96.1%) rules; in comparison, protein terms (collectively) featured in 198 (~6.0%) rules, while the next largest non-protein term was donor group, featuring in 5 (~0.1%) rules (Figure 4A). Feature selection became progressively inefficient, in terms of candidate yield per iteration, however new candidates were still found up to termination at iteration 2000. This process generated 195 candidate proteins; we supplemented this list with proteins that had high correlation (Pearson's r>0.65) with any of those candidates; bringing the final list up to 255 candidates.

Regression spline modelling reveals protein associations with posttransplant outcome are modulated by donor age

Next, we tested each protein candidate for individual association with outcome. Since eGFR rank transformation is cohort specific, to generate results which generalized to other settings we modelled against an outcome binary, calibrated against a population-level threshold. Based on UK Renal Registry data, the donor type-weighted average median eGFR at 12 months

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

posttransplant since 2013[25] was approximately ~50.25 ml/min/1.73 m^2 (±0.24 standard error;); for simplicity we used a threshold of 50 ml/min/1.73 m^2. We refer to 'sub-median outcome' henceforth to refer to recipient eGFR at 12 months posttransplant less than 50 ml/min/1.73 m^2.

Using multivariate adaptive regression spline modelling[23] to assess individual protein relationships, we generated predictive models for sub-median outcome using each protein, donor age, and age:protein interaction. This was performed in a regularized framework to retain only a minimal set of predictors in each model. We discarded candidates whose model either did not feature a protein or age:protein term, or gave a worse prediction error (Brier score[26]) than donor age alone.

After filtering we had identified 136 proteins which predicted sub-median outcome (Supplementary Table 2). We performed a network analysis of shared Reactome pathways (Figure 4). Walktrap clustering revealed 4 major clusters of shared-pathway proteins (Table 2); Immune Regulation and Complement Activation, Protein Metabolism and Regulation, Metabolism and Apoptosis.

For 124 proteins, the model included an age:protein interaction term where the predictive effect of protein abundance was modulated by age, independent of the effect of age alone or protein abundance alone (Figure 5, Supplementary Table 2). To visualize these effects, we used each model to predict outcomes across increasing donor age for a high (90th percentile), median and low (10th percentile) protein abundance (Figure 5A and Supplementary Figure 2). The majority of proteins were positively associated with the chance of sub-median outcome (simplistically, more protein = worse outcome), with the effect appearing to increase up to around donor age 45-55, including representative proteins for all four major clusters including known markers Cystatin-C (CST3; nephron function) and Vitronectin (VTN; fibrosis) as well as

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

a protein known for age-modulated disease associations, Apolipoprotein E (APOE). Several proteins were negatively associated with the chance of sub-median outcome (simplistically more protein = better outcome), with the largest effects shown by MAP2K1 and SLC27A2, the latter in particular being modelled as having minimal effect in donors younger than 40. The selection and filtering steps are summarized in Figure 5B.

Orthogonal validation confirms model performance, including age-modulated immuno-metabolic impact on transplant outcomes

We adopted two orthogonal validation approaches. Firstly, we assessed the performance of each model on test data. Going from train to test data, the models showed a small increase in accuracy (Brier score; mean square error) and a small decrease in overall predictive performance as measured by the area under the curve (AUC) (Figure 6A), indicating that the models generalized well to unseen data, with most of the models (~110/136) showing almost no degradation in performance.

Secondly, we selected several cluster-representative proteins (VTN, APOE, CST3 and Prolactin Regulatory Element Binding; PREB) that had robust available antibodies (Figure 6B). We investigated the predicted pattern of associations between protein abundance and outcome (Figure 6C), by performing western blot validation of our results (Figure 6D). Selecting samples with remaining material from our cohort from the Good and Suboptimal Outcome stratified eGFR tertiles of our sample set, we compared protein abundance between younger (oldest sample 49) and older (youngest sample 58) donors. Our results were broadly consistent with the associations anticipated by our modelling:

For VTN, our model suggests a strong association between protein abundance and outcome in younger donors that strengthens towards age 40-50 and then weakens. We observed a

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

significant difference in abundance by western blot between GO and SO outcome strata in younger donors (t-test; p=2.107e-9) and a weaker but mildly significant difference in older donors (t-test; p=0.0245).

For PREB, the model suggests that the age of maximum difference is shifted towards older donors compared to VTN; we did not observe any significant difference by western (t-test; p=0.4530) in young donors but did find such a difference in older donors (t-test, p=8.800e-5).

For APOE, the model suggests that the strongest association is over the middle of the age range, where outcome changes rapidly with donor age, followed by a weaker but consistent association from donor age 49-50 onwards. We observed no difference among younger donors, although examination of the sampled age ranges indicate the area of starkest difference was under-sampled (Figure 5C; t-test, p=0.3719). In older donors, we saw a mildly significant difference in APOE abundance between outcomes (t-test, p=0.0323).

For CST3, the model again predicted a strong association in young donors which then weakens (and even reverses, such that GO samples would tend to have higher CST3 than SO samples); in younger donor western blots we observed a significant difference (t-test, p=0.0084) between GO and SO, while in older donors we saw a nonsignificant difference, but (in contrast to the prediction) still with a positive median protein abundance difference from GO to SO.

Finally, we compared the predicted outcome for each of the six paired kidneys from the second ‘biological duplicates’ test set against their actual recipient eGFR at 12 months posttransplant (Figure 6E). All three kidney pairs in this dataset had consistent outcomes across pairs; two pairs with sub-median outcome (15 and 36 ml/min/1.73 m^2; 23 and 27 ml/min/1.73 m^2) and one pair with above-median outcome (72 and 81 ml/min/1.73 m^2); all four protein models assigned the four kidneys from the two sub-median outcome donors a probability of sub-

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

median outcome greater than 0.6 (except in one case for PREB, where it was 0.48), and assigned both kidneys from the above-median outcome donor a probability of sub-median outcome less that 0.3.

DISCUSSION

Increasingly, shortages of optimal organs require utilization of kidneys from older deceased donors with increased risks of graft failure or functional decline. Here, we show that age-modulated kidney proteomic profiles improve risk stratification of donor kidney quality, revealing clinically relevant age-protein interaction effects.

Donor age remains a key contributor in these clinical decisions and is rightfully one of the most strongly weighted terms in extant scoring systems to determine kidney allocation[8,27]. In our analysis, we found no obvious difference according to age when comparing donor kidney proteomes by unbiased PCA. However, looking specifically at *outcomes*, it was the single most important factor. PCA considers only a linear combination of variables and is ill-suited to exploring nonlinear effects or interaction between variables. When we explored our data with our iterative PRE feature selection approach, a substantial number of proteins were revealed to be relevant.

The effect of donor age is not a novel finding, but integration of the age and proteomic information resulted in enhanced prediction of 12 month sub-median function without reference to other currently considered clinical factors. In particular, a factor often described as relevant to transplant outcome is donor type[28]. This is true at a clinical level in terms of donor management, and donation after circulatory death is considered an adverse factor for transplant outcome in the US[7] (although not in the UK[8,29]). Our initial PCA analysis found that non-supervised clustering of the sample proteomes did partially separate samples by type but did not extend to association

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

with outcome. Weaker association may be obscured by factors associated with the overwhelming effect of donor age, but (without disputing donor type-specific mechanisms of kidney injury[30]), our data are consistent with the idea that extent of injury (rather than the cause) is the primary contributor towards recovery potential[31].

Within our final list of 136 proteins associated with outcome there is a common theme of implication in immune response to kidney injury (including both chronic injury, and acute injury) particularly as a result of ischemic metabolic disruption. Our analysis of proteins associated with outcome also revealed that most (124/136) proteins showed age-moderated differences in their effect; for most proteins manifesting as a stronger negative association between abundance and outcome starting around donor age 40-50. This second-order age interaction effect, where weightings of other factors are themselves age dependent, has not (to our knowledge) been explored in transplantation, and may be key to fully understanding the effects of molecular predictors.

A prominent age-modulated example of a chronic injury associated marker in our candidate list is VTN, a primary component of the extracellular matrix involved in in cell adhesion, enhancing the activity of plasminogen activator inhibitor-1 and inhibition of the terminal complement pathway[32]. Vitronectin has been suggested as a biomarker of kidney fibrosis, although the mode of its multifaced action needs further investigation[33]. Further acute injury associated markers include components of the membrane attack complex, C5 (in the form of C5b cleavage product) and C8A, which has been associated with tissue injury resulting from ischemia/reperfusion[34,35], Complement Component 1r (C1R), part of the activation complex for the classical complement pathway[36], and Complement Factor B (CFB), a component of the alternative pathway. Another candidate associated with immune regulation is Maltase

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Glucoamylase (MGAM), characterized as an intestinal enterocyte but with expression in several tissues including kidney, and whose presence in urinary exosomes been cited as a marker of AKI in cirrhosis patients[37].

Mitogen-activated protein kinase 1 (MAP2K1, aka MEK1), a key component of the MAP kinase signal transduction pathway and closely involved in both cellular control and immune regulation (as part of TNF signaling response[38]), is notable as one of the few proteins for which higher abundance was associated with a reduced probability of sub-median outcome, indicating resilience to injury. Increased TNF is more usually associated with renal injury[39], so this result is counterintuitive. The MAPK/ERK cascade impacts many regulatory pathways so it is reasonable to assume such intuition may oversimplify the effect of increased MEK1 abundance.

Several age-moderated proteins we report are characterized as markers of protein regulation and proteasomal activity, suggesting alterations within the proteostasis network that increase susceptibility of donor grafts to subsequent injury and reduce capacity for recovery. CST3 is particularly noteworthy as, measured in serum, it is a known and effective general biomarker for kidney function and has previously been reported as having predictive power for outcomes in transplant recipients[40–42]. Our evidence indicates a further association between CST3 levels in the donor kidney tissue and outcome; moreover, that this effect is age dependent, starting around age 40. Interestingly, while serum CST3 is relatively independent of age in children and young adults[43], there is some evidence for an increase in later years[44].

We found the age-modulated candidate PREB (Prolactin Regulatory Element Binding protein) biologically interesting for three reasons. Firstly, there is a well characterized relationships between kidney dysfunction (in the form of CKD), cardiovascular disease and prolactinemia[45], with CKD patients being associated with elevated prolactinemia. Secondly, it is

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

a regulator of glucose homeostasis in the liver and therefore a plausible key node for metabolic regulation in kidneys as well[46], acknowledging the large emphasis in our pathway analysis on metabolic functions. Thirdly, it has a predicted[47] role in exit from the endoplasmic reticulum and the unfolded protein response, which has an association with CKD via NFκB -mediated inflammation[48].

Another age-moderated protein, APOE, stands out as having previously reported genetic allele age-related associations with disease and organ dysfunction including risk of Alzheimer's Disease (AD)[57] (with the strongest effect manifesting around age 65[58]), macular dysfunction, atherosclerosis and pulmonary scarring[59,60], and evidence for shared allele risk across diseases[61]. In kidneys, APOE plays an important role in lipid metabolism to regulate the growth and survival of mesangial cells and preserve organ function[49]; it is a marker for outcome in transplant recipients[50–52], and there is already evidence for *APOE* genetic allele association with kidney dysfunction risk[53–55], possibly manifested by lipidomic differences between allelic profiles[56]. We have previously observed small (not statistically significant) increases in APOE due to ischemic reperfusion injuries[62] possibly explained by a recent description of the role of APOE in mediating senescence[63]. Such evidence supports further in-depth investigation of the *APOE* genotype in outcomes across donor kidney age. There is existing evidence for similar allele dependent transplant outcome effects in another apolipoprotein (APOL1)[64], suggesting that the broader apolipoprotein allelic profile may play an important role in outcome.

Our list of outcome-associated candidates, controlling for the effect of donor age, including those for which we report a further age-moderated effect, cannot be exhaustive. Practicalities of sample acquisition limited sampling of a wide range of outcomes outside the 30-60 donor age range, especially limited good outcome events at high donor age. Organ allocation

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

algorithms impose a close link between donor and recipient age in the sample cohort, so while we interpret these age-moderated effects in terms of organ resilience in older donors, it could also represent a greater ability to repair a given level of damage in younger recipients. Further, we consider only chronological donor age, rather than a more nuanced representation of the epigenomic biological clock [65], which may to account for some variation observed with respect to both donors and recipients.

In the vast majority of proteins, the modelling suggests a plateauing effect at high donor age where the differences in outcome due to both protein and age are smaller. This effect may be an artifact of the distribution of sub-median outcomes in our UK population-representative data (>95% sub-median outcome above age 60 in our training set). The protein abundance differences between GO and SO outcomes in our western blot validation were broadly consistent with our differences expected given our prediction models. There were some differences; firstly, the differences in both VTN and PREB in older donors were larger than might be expected by examination of the prediction curves (Figure 6B, upper panel); secondly, the prediction for CST3 of a small reversal of the effect is both unexpected and biologically counterintuitive. Both examples are most readily explained as model artifacts due to lack of outcome diversity at high age ranges.

It is immediately clear from our results that the strength of the donor age factor is enormous relative to any other protein or clinical effect; this age effect is liable to dominate any prediction weighting and reduce the accuracy of estimated protein contribution. A much larger cohort could mitigate this issue. Advances in high-throughput proteomics techniques continue to increase feasible cohort sizes[66] but fundamental limitations on organ acquisition remain. Archiving at scale of clinical samples in bioresources such as the QUOD biobank to parallel

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

advancements in big data analysis and interpretation platforms is therefore necessary for future development of granular evidence-based decision making.

In this work, we profiled the proteome of pre-implantation biopsies selected from donor kidneys on the basis of paired 12-month graft function. Using machine learning and regression models, we identified 136 proteins associated with sub-median outcomes, suggesting molecular signatures which may refine models of graft dysfunction based on clinical and demographic factors alone. We also found that most of these proteins furthermore show donor-age modulated association. The biological themes of the identified proteins reinforce known immuno-metabolic mechanisms of kidney injury but raise interesting possibilities for further work, especially with regard to donor genetic background, and also suggest that the possibility of donor age-moderated weighting should be considered as a matter of course in future work.

SUPPLEMENTARY MATERIALS

Supplementary Methods

Supplementary Table 1: Clinical variable p-values for association with donor type and outcome

Supplementary Table 2: Summary of results for all candidate proteins

Supplementary Figure 1: Protein quantification quality

A: Missingness comparison: Proteins are shown ranked by the number of missing values across all samples and the twenty standard pools, excluding one run which was removed due to low signal. 2984 proteins had missing values in 50% or less runs.

B: Paired Kidney Comparison: Protein abundance values from paired kidneys (left/right) from 3 individual donors were compared, as these are effectively biological replicates. x axes: value in left kidney. y axes: value in right kidney. Inset: R-squared value

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Supplementary Figure 2: Prediction of sub-median outcome differences between high and low protein across donor age, for all shortlisted proteins with a predicted age modulation effect

Black traces: prediction at median protein abundance. Purple trace: prediction at 90th percentile of protein abundance. Orange traces: prediction at 10th percentile of protein abundance. The corresponding point on the main figure thus indicates the age at which the difference between orange and purple lines is greatest.

DATA AND MATERIALS AVAILABILITY

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [67] partner repository with the dataset identifier PXD033428.

DISCLOSURE AND FUNDING

This study was supported by NHS Blood and Transplant funding awarded to MK & RJP. SF was supported by Kidney Research UK, grant reference KS_RP_002_20210111 awarded to MK. PDC was supported by a Chinese Academy of Medical Sciences 2018-I2M-2-002 awarded to BMK. Authors declare that they have no competing interests.

Author contributions:

Conceptualization: MK

Methodology: PDC, SF, RV, SD, RF, BMK, AS, ES, RJP, MK

Investigation: PDC, SF, RV, PJ, SD, IV, KT, AS

Visualization: PDC

Funding acquisition: BMK, RJP, MK

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Project administration: MK

Supervision: RF, AS, MK

Writing – original draft: PDC, ES, MK

Writing – review & editing: All authors

REFERENCES

1. Callaghan CJ, Mumford L, Pankhurst L, Baker RJ, Bradley JA, Watson CJE. Early Outcomes of the New UK Deceased Donor Kidney Fast-Track Offering Scheme. *Transplantation*. 2017;101(12):2888-2897. doi:10.1097/TP.0000000000001860

2. Summers DM, Johnson RJ, Hudson AJ, et al. Standardized deceased donor kidney donation rates in the UK reveal marked regional variation and highlight the potential for increasing kidney donation: a prospective cohort study†. *Br J Anaesth*. 2014;113(1):83-90. doi:10.1093/bja/aet473

3. Denic A, Glassock RJ, Rule AD. The Kidney in Normal Aging: A Comparison with Chronic Kidney Disease. *Clin J Am Soc Nephrol*. 2022;17(1):137-139. doi:10.2215/CJN.10580821

4. O'Hare AM, Choi AI, Bertenthal D, et al. Age affects outcomes in chronic kidney disease. *J Am Soc Nephrol*. 2007;18(10):2758-2765. doi:10.1681/ASN.2007040422

5. Ishani A, Xue JL, Himmelfarb J, et al. Acute Kidney Injury Increases Risk of ESRD among Elderly. *J Am Soc Nephrol*. 2009;20(1):223-228. doi:10.1681/ASN.2007080837

6. Summers DM, Johnson RJ, Hudson A, Collett D, Watson CJ, Bradley JA. Effect of donor age and cold storage time on outcome in recipients of kidneys donated after circulatory death in the UK: a cohort study. *Lancet*. 2013;381(9868):727-734. doi:10.1016/S0140-6736(12)61685-7

7. Rao PS, Schaubel DE, Guidinger MK, et al. A comprehensive risk quantification score for deceased donor kidneys: the kidney donor risk index. *Transplantation*. 2009;88(2):231-236. doi:10.1097/TP.0b013e3181ac620b

8. Watson CJE, Johnson RJ, Birch R, Collett D, Bradley JA. A Simplified Donor Risk Index for Predicting Outcome After Deceased Donor Kidney Transplantation. *Transplantation*. 2012;93(3):314-318. doi:10.1097/TP.0b013e31823f14d4

9. Neuberger J, Callaghan C. Organ utilization - the next hurdle in transplantation? *Transpl Int*. 2020;33(12):1597-1609. doi:10.1111/tri.13744

10. Clayton PA, Dansie K, Sypek MP, et al. External validation of the US and UK kidney donor risk indices for deceased donor kidney transplant survival in the Australian and New Zealand population. *Nephrol Dial Transplant*. 2019;34(12):2127-2131. doi:10.1093/NDT/GFZ090

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

11. Remuzzi G, Cravedi P, Perna A, et al. Long-Term Outcome of Renal Transplantation from Older Donors. *N Engl J Med*. 2006;354(4):343-352. doi:10.1056/NEJMoa052891

12. Husain SA, Chiles MC, Lee S, et al. Characteristics and Performance of Unilateral Kidney Transplants from Deceased Donors. *Clin J Am Soc Nephrol*. 2018;13(1):118-127. doi:10.2215/CJN.06550617

13. Liu C, Hall IE, Mansour S, Thiessen Philbrook HR, Jia Y, Parikh CR. Association of Deceased Donor Acute Kidney Injury With Recipient Graft Survival. *JAMA Network Open*. 2020;3(1):e1918634. doi:10.1001/jamanetworkopen.2019.18634

14. Yu K, Husain SA, King K, Stevens JS, Parikh CR, Mohan S. Kidney nonprocurement in deceased donors with acute kidney injury. *Clin Transplant*. Published online August 4, 2022:e14788. doi:10.1111/ctr.14788

15. Hall IE, Akalin E, Bromberg JS, et al. Deceased-donor acute kidney injury is not associated with kidney allograft failure. *Kidney Int*. 2019;95(1):199-209. doi:10.1016/j.kint.2018.08.047

16. Mansour SG, Khoury N, Kodali R, et al. Clinically adjudicated deceased donor acute kidney injury and graft outcomes. *PLoS One*. 2022;17(3):e0264329. doi:10.1371/journal.pone.0264329

17. von Moos S, Akalin E, Mas V, Mueller TF. Assessment of Organ Quality in Kidney Transplantation by Molecular Analysis and Why It May Not Have Been Achieved, Yet. *Front Immunol*. 2020;11:833. doi:10.3389/fimmu.2020.00833

18. Gillet LC, Navarro P, Tate S, et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics*. 2012;11(6):O111.016717. doi:10.1074/mcp.O111.016717

19. Geyer PE, Holdt LM, Teupser D, Mann M. Revisiting biomarker discovery by plasma proteomics. *Mol Syst Biol*. 2017;13(9):942. doi:10.15252/msb.20156297

20. Connor KL, O'Sullivan ED, Marson LP, Wigmore SJ, Harrison EM. The Future Role of Machine Learning in Clinical Transplantation. *Transplantation*. 2021;105(4):723-735. doi:10.1097/TP.0000000000003424

21. Raynaud M, Aubert O, Divard G, et al. Dynamic prediction of renal survival among deeply phenotyped kidney transplant recipients using artificial intelligence: an observational, international, multicohort study. *The Lancet Digital Health*. 2021;3(12):e795-e805. doi:10.1016/S2589-7500(21)00209-0

22. Fokkema M. Fitting prediction rule ensembles with R package pre. *J Stat Softw*. 2020;92(1):1-30. doi:10.18637/jss.v092.i12

23. Friedman JH. Multivariate Adaptive Regression Splines. *aos*. 1991;19(1):1-67. doi:10.1214/aos/1176347963

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

24. Milborrow. Derived from mda:mars by T. Hastie and R. Tibshirani. S. *Earth: Multivariate Adaptive Regression Splines*.; 2011. http://CRAN.R-project.org/package=earth

25. UK Renal Registry. *UK Renal Registry 24th Annual Report - Data to 31/12/2020*.; 2022. https://renal.org/audit-research/annual-report

26. Brier GW. Verification Of Forecasts Expressed In Terms Of Probability. *Mon Weather Rev*. 1950;78(1):1-3. doi:10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2

27. Li B, Cairns JA, Robb ML, et al. Predicting patient survival after deceased donor kidney transplantation using flexible parametric modelling. *BMC Nephrol*. 2016;17(1):51. doi:10.1186/s12882-016-0264-0

28. Gill J, Rose C, Lesage J, Joffres Y, Gill J, O'Connor K. Use and Outcomes of Kidneys from Donation after Circulatory Death Donors in the United States. *J Am Soc Nephrol*. 2017;28(12):3647-3657. doi:10.1681/asn.2017030238

29. Robinson C. *NHSBT Policy POL186/10. Kidney Transplantation: Deceased Donor Allocation*.; 2019. https://nhsbtdbe.blob.core.windows.net/umbraco-assets-corp/16915/kidney-allocation-policy-pol186.pdf

30. Vaughan RH, Kresse JC, Farmer LK, et al. Cytoskeletal protein degradation in brain death donor kidneys associates with adverse post-transplant outcomes. *Am J Transplant*. Published online December 8, 2021. doi:10.1111/ajt.16912

31. Kosmoliaptsis V, Salji M, Bardsley V, et al. Baseline donor chronic renal injury confers the same transplant survival disadvantage for DCD and DBD kidneys. *Am J Transplant*. 2015;15(3):754-763. doi:10.1111/ajt.13009

32. De Lorenzi V, Sarra Ferraris GM, Madsen JB, Lupia M, Andreasen PA, Sidenius N. Urokinase links plasminogen activation and cell adhesion by cleavage of the RGD motif in vitronectin. *EMBO Rep*. 2016;17(7):982-998. doi:10.15252/embr.201541681

33. Carreras-Planella L, Cucchiari D, Cañas L, et al. Urinary vitronectin identifies patients with high levels of fibrosis in kidney grafts. *J Nephrol*. 2021;34(3):861-874. doi:10.1007/s40620-020-00886-y

34. Diepenhorst GMP, van Gulik TM, Hack CE. Complement-mediated ischemia-reperfusion injury: lessons learned from animal and clinical studies. *Ann Surg*. 2009;249(6):889-899. doi:10.1097/SLA.0b013e3181a38f45

35. Arias-Cabrales CE, Riera M, Pérez-Sáez MJ, et al. Activation of final complement components after kidney transplantation as a marker of delayed graft function severity. *Clin Kidney J*. 2021;14(4):1190-1196. doi:10.1093/ckj/sfaa147

36. Almitairi JOM, Venkatraman Girija U, Furze CM, et al. Structure of the C1r-C1s interaction of the C1 complex of complement activation. *Proc Natl Acad Sci U S A*. 2018;115(4):768-773. doi:10.1073/pnas.1718709115

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

37. Awdishu L, Tsunoda S, Pearlman M, et al. Identification of Maltase Glucoamylase as a Biomarker of Acute Kidney Injury in Patients with Cirrhosis. *Crit Care Res Pract*. 2019;2019:5912804. doi:10.1155/2019/5912804

38. Sabio G, Davis RJ. TNF and MAP kinase signalling pathways. *Semin Immunol*. 2014;26(3):237-245. doi:10.1016/j.smim.2014.02.009

39. Ernandez T, Mayadas TN. Immunoregulatory role of TNFalpha in inflammatory kidney diseases. *Kidney Int*. 2009;76(3):262-276. doi:10.1038/ki.2009.142

40. Coll E, Botey A, Alvarez L, et al. Serum cystatin C as a new marker for noninvasive estimation of glomerular filtration rate and as a marker for early renal impairment. *Am J Kidney Dis*. 2000;36(1):29-34. doi:10.1053/ajkd.2000.8237

41. Christensson A, Ekberg J, Grubb A, Ekberg H, Lindström V, Lilja H. Serum cystatin C is a more sensitive and more accurate marker of glomerular filtration rate than enzymatic measurements of creatinine in renal transplantation. *Nephron Physiol*. 2003;94(2):19-27. doi:10.1159/000071287

42. Rodrigo E, Ruiz JC, Fernández-Fresnedo G, et al. Cystatin C and albuminuria as predictors of long-term allograft outcomes in kidney transplant recipients. *Clin Transplant*. 2013;27(2):E177-83. doi:10.1111/ctr.12082

43. Bökenkamp A, Domanetzki M, Zinck R, Schumann G, Byrd D, Brodehl J. Cystatin C--a new marker of glomerular filtration rate in children independent of age and height. *Pediatrics*. 1998;101(5):875-881. doi:10.1542/peds.101.5.875

44. Edinga-Melenge BE, Yakam AT, Nansseu JR, et al. Reference intervals for serum cystatin C and serum creatinine in an adult sub-Saharan African population. *BMC Clin Pathol*. 2019;19:4. doi:10.1186/s12907-019-0086-7

45. Dourado M, Cavalcanti F, Vilar L, Cantilino A. Relationship between Prolactin, Chronic Kidney Disease, and Cardiovascular Risk. *Int J Endocrinol*. 2020;2020:9524839. doi:10.1155/2020/9524839

46. Park JM, Kim MY, Kim TH, Min DK, Yang GE, Ahn YH. Prolactin regulatory element-binding (PREB) protein regulates hepatic glucose homeostasis. *Biochim Biophys Acta Mol Basis Dis*. 2018;1864(6 Pt A):2097-2107. doi:10.1016/j.bbadis.2018.03.024

47. Gaudet P, Livstone MS, Lewis SE, Thomas PD. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. *Brief Bioinform*. 2011;12(5):449-462. doi:10.1093/bib/bbr042

48. Mohammed-Ali Z, Cruz GL, Dickhout JG. Crosstalk between the unfolded protein response and NF-κB-mediated inflammation in the progression of chronic kidney disease. *J Immunol Res*. 2015;2015:428508. doi:10.1155/2015/428508

49. Chen G, Paka L, Kako Y, Singhal P, Duan W, Pillarisetti S. A Protective Role for Kidney Apolipoprotein E: REGULATION OF MESANGIAL CELL PROLIFERATION AND

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

MATRIX EXPANSION. *J Biol Chem*. 2001;276(52):49142-49147. doi:10.1074/JBC.M104879200

50. Kahraman S, Kiykim AA, Altun B, et al. Apolipoprotein E gene polymorphism in renal transplant recipients: effects on lipid metabolism, atherosclerosis and allograft function. *Clin Transplant*. 2004;18(3):288-294. doi:10.1111/j.1399-0012.2004.00162.x

51. Hernández D, Salido E, Linares J, et al. Role of apolipoprotein E epsilon 4 allele on chronic allograft nephropathy after renal transplantation. *Transplant Proc*. 2004;36(10):2982-2984. doi:10.1016/j.transproceed.2004.10.038

52. Cofán F, Cofan M, Rosich E, et al. Effect of apolipoprotein E polymorphism on renal transplantation. *Transplant Proc*. 2007;39(7):2217-2218. doi:10.1016/j.transproceed.2007.06.011

53. Czaplińska M, Ćwiklińska A, Sakowicz-Burkiewicz M, et al. Apolipoprotein e gene polymorphism and renal function are associated with apolipoprotein e concentration in patients with chronic kidney disease. *Lipids Health Dis*. 2019;18(1):1-9. doi:10.1186/s12944-019-1003-x

54. Hsu CC, Kao WHL, Coresh J, et al. Apolipoprotein E and progression of chronic kidney disease. *J Am Med Assoc*. 2005;293(23):2892-2899. doi:10.1001/jama.293.23.2892

55. Saito T, Ishigaki Y, Oikawa S, Yamamoto TT. Etiological significance of apolipoprotein E mutations in lipoprotein glomerulopathy. *Trends Cardiovasc Med*. 2002;12(2):67-70. doi:10.1016/S1050-1738(01)00148-7

56. Sienski G, Narayan P, Bonner JM, et al. APOE4 disrupts intracellular lipid homeostasis in human iPSC-derived glia. *Sci Transl Med*. 2021;13(583):eaaz4564. doi:10.1126/scitranslmed.aaz4564

57. Liu CC, Kanekiyo T, Xu H, Bu G. Apolipoprotein E and Alzheimer disease: risk, mechanisms, and therapy. *Nat Rev Neurol*. 2013;9(2):106-118. doi:10.1038/nrneurol.2012.263

58. Saddiki H, Fayosse A, Cognat E, et al. Age and the association between apolipoprotein E genotype and Alzheimer disease: A cerebrospinal fluid biomarker-based case-control study. *PLoS Med*. 2020;17(8):e1003289. doi:10.1371/journal.pmed.1003289

59. Mahley RW, Rall SC. Apolipoprotein E: far more than a lipid transport protein. *Annu Rev Genomics Hum Genet*. 2000;1:507-537. doi:10.1146/annurev.genom.1.1.507

60. Yao X, Gordon EM, Figueroa DM, Barochia AV, Levine SJ. Emerging Roles of Apolipoprotein E and Apolipoprotein A-I in the Pathogenesis and Treatment of Lung Disease. *Am J Respir Cell Mol Biol*. 2016;55(2):159-169. doi:10.1165/rcmb.2016-0060TR

61. Miwa K, Tanaka M, Okazaki S, et al. Chronic kidney disease is associated with dementia independent of cerebral small-vessel disease. *Neurology*. 2014;82(12):1051-1057. doi:10.1212/WNL.0000000000000251

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

62. Huang H, van Dullemen LFA, Akhtar MZ, et al. Proteo-metabolomics reveals compensation between ischemic and non-injured contralateral kidneys after reperfusion. *Sci Rep*. 2018;8(1):8539. doi:10.1038/s41598-018-26804-8

63. Zhao H, Ji Q, Wu Z, et al. Destabilizing heterochromatin by APOE mediates senescence. *Nature Aging*. 2022;2(4):303-316. doi:10.1038/s43587-022-00186-z

64. Julian BA, Gaston RS, Brown WM, et al. Effect of Replacing Race with Apolipoprotein L1 Genotype in Calculation of Kidney Donor Risk Index. *Am J Transplant*. 2017;17(6):1540. doi:10.1111/AJT.14113

65. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat Rev Genet*. 2018;19(6):371-384. doi:10.1038/s41576-018-0004-3

66. Meier F, Brunner AD, Frank M, et al. diaPASEF: parallel accumulation–serial fragmentation combined with data-independent acquisition. *Nat Methods*. 2020;17(12):1229-1236. doi:10.1038/s41592-020-00998-0

67. Perez-Riverol Y, Csordas A, Bai J, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res*. 2019;47(D1):D442-D450. doi:10.1093/nar/gky1106

ACKNOWLEDGMENTS

We thank the UK QUOD Consortium and NHS Blood and Transplant UK Registry for providing the samples and the associated clinical and demographic metadata. ; in particular we thank Sheba Ziyenge, Lewis Simmonds and Dr Sarah Cross, Dr Sergei Maslau and Mr Tomas Surik for their support on the QUOD sample selection.

We thank members of the Discovery Proteomics Facility within the TDI Mass Spectrometry Laboratory for expert help with mass spectrometry analysis, and members of the Lindgren group at the BDI for informative discussions regarding statistical modelling.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

FIGURES AND TABLES

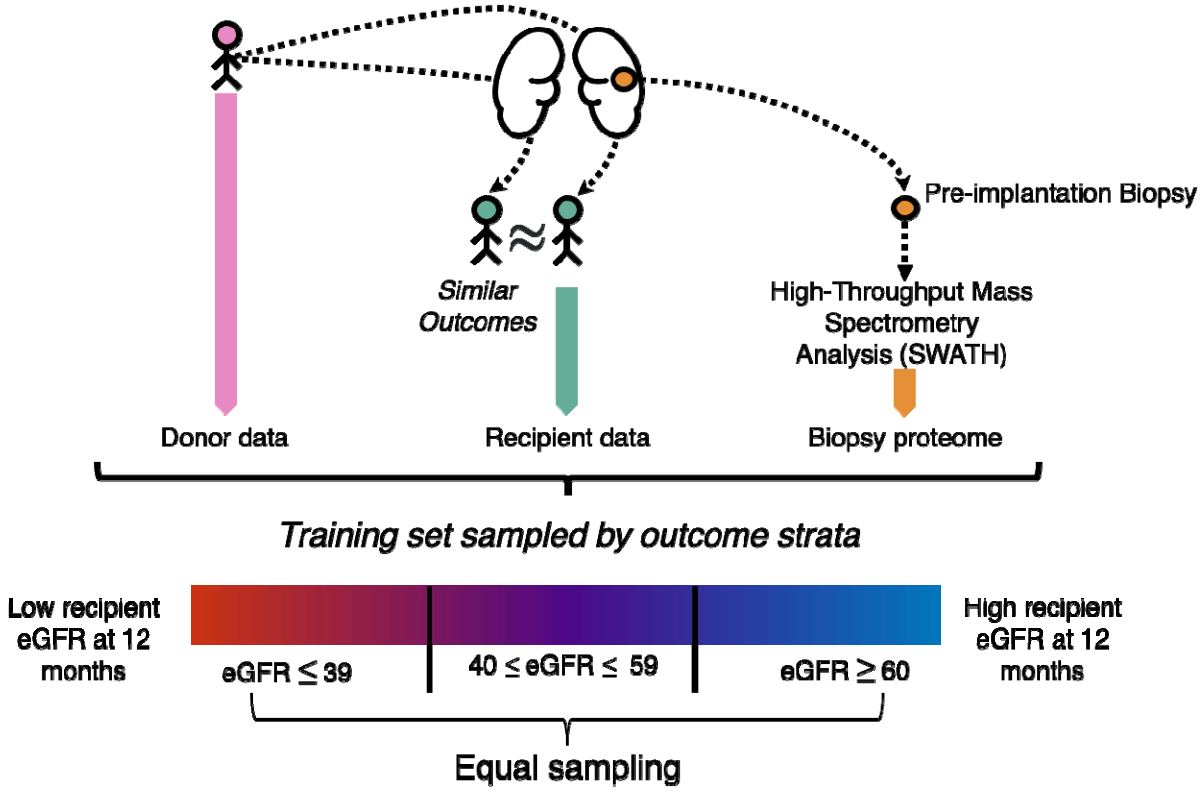


Figure 1: Experimental design to discover donor kidney proteome associations with transplant outcome One kidney from each donor pair was biopsied at the back table. Donor kidney samples were selected randomly from pairs where both recipients had similar outcomes. The biopsy samples were subjected to proteomic analysis to yield a snapshot of the organ proteome before transplantation. We analyzed donor characteristics and clinical variables, recipient characteristics and protein abundances in a combined model against outcome. eGFR units for stratification given in ml/min/1.73 m^2.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Donor Type	DBD			DCD		
Outcome Tertile *(eGFR in ml/min/1.73 m²)*	1st: Suboptimal *(eGFR≤39)*	2nd: Intermediate *(40≤eGFR≤59)*	3rd: Good *(eGFR≥60)*	1st: Suboptimal *(eGFR≤39)*	2nd: Intermediate *(40≤eGFR≤59)*	3rd: Good *(eGFR≥60)*
n	31	31	38	31	28	26
Donor Age, y	56.84 ± 12.29	51.32 ± 12.24	39.05 ± 14.12	55.48 ± 9.34	53.57 ± 9.75	38.31 ± 12.28
Donor Sex						
Male	15 (48.4%)	16 (51.6%)	19 (50.0%)	22 (71.0%)	16 (57.1%)	16 (61.5%)
Female	16 (51.6%)	15 (48.4%)	19 (50.0%)	9 (29.0%)	12 (42.9%)	10 (38.5%)
Donor Ethnicity						
White	30 (96.8%)	30 (96.8%)	36 (94.7%)	30 (96.8%)	28 (100.0%)	25 (96.2%)
Other	1 (3.2%)	1 (3.2%)	2 (5.3%)	1 (3.2%)	0 (0.0%)	1 (3.8%)
Donor Weight, kg	82.53 ± 18.20	76.61 ± 18.07	81.38 ± 17.72	80.58 ± 14.95	82.43 ± 17.20	78.67 ± 13.96
Donor Height, cm	168.42 ± 9.37	169.52 ± 7.67	174.82 ± 11.16	169.97 ± 7.98	171.64 ± 9.73	174.65 ± 8.98
Donor S-Cr terminal, µmol/l	86.54 ± 40.81	82.57 ± 49.65	90.19 ± 67.36	73.37 ± 19.03	70.31 ± 39.02	59.60 ± 22.39
Donor CIT, h	15.80 ± 3.88	14.20 ± 4.60	13.42 ± 4.67	13.65 ± 5.20	11.72 ± 3.55	12.80 ± 4.49
Donor COD						
Trauma	1 (3.2%)	3 (9.7%)	3 (7.9%)	4 (12.9%)	3 (10.7%)	4 (15.4%)
Other	30 (96.8%)	28 (90.3%)	35 (92.1%)	27 (87.1%)	25 (89.3%)	22 (84.6%)
Donor UKKDRI	1.41 ± 0.52	1.10 ± 0.36	0.85 ± 0.36	1.31 ± 0.37	1.21 ± 0.40	0.73 ± 0.35
Recipient Age, y	53.03 ± 12.21	52.10 ± 14.61	39.71 ± 16.03	51.90 ± 9.85	50.93 ± 11.04	44.92 ± 12.87
Recipient Sex						
Female	15 (48.4%)	8 (25.8%)	12 (31.6%)	11 (35.5%)	9 (32.1%)	5 (19.2%)
Male	16 (51.6%)	23 (74.2%)	26 (68.4%)	20 (64.5%)	19 (67.9%)	21 (80.8%)
Recipient Ethnicity						
White	24 (77.4%)	21 (67.7%)	29 (76.3%)	22 (71.0%)	23 (82.1%)	20 (76.9%)
Other	7 (22.6%)	10 (32.3%)	9 (23.7%)	9 (29.0%)	5 (17.9%)	6 (23.1%)
Recipient Posttransplant Kidney Function (mean eGFR, ml/min/1.73 m²)						
3 months	29.71 ± 12.06	50.32 ± 17.28	78.54 ± 25.97	31.72 ± 12.24	46.50 ± 10.34	77.88 ± 18.57
12 months	26.58 ± 11.98	49.58 ± 6.10	85.58 ± 35.84	25.10 ± 12.01	48.24 ± 6.29	80.26 ± 15.91

Table 1: Donor and recipient clinical and demographic variables
Donor kidney associated metadata. Samples are subdivided by donor type and by final assigned outcome tertile. Numerical variables are given ± standard deviation. Categorial variables are given alongside percentage of total cohort

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

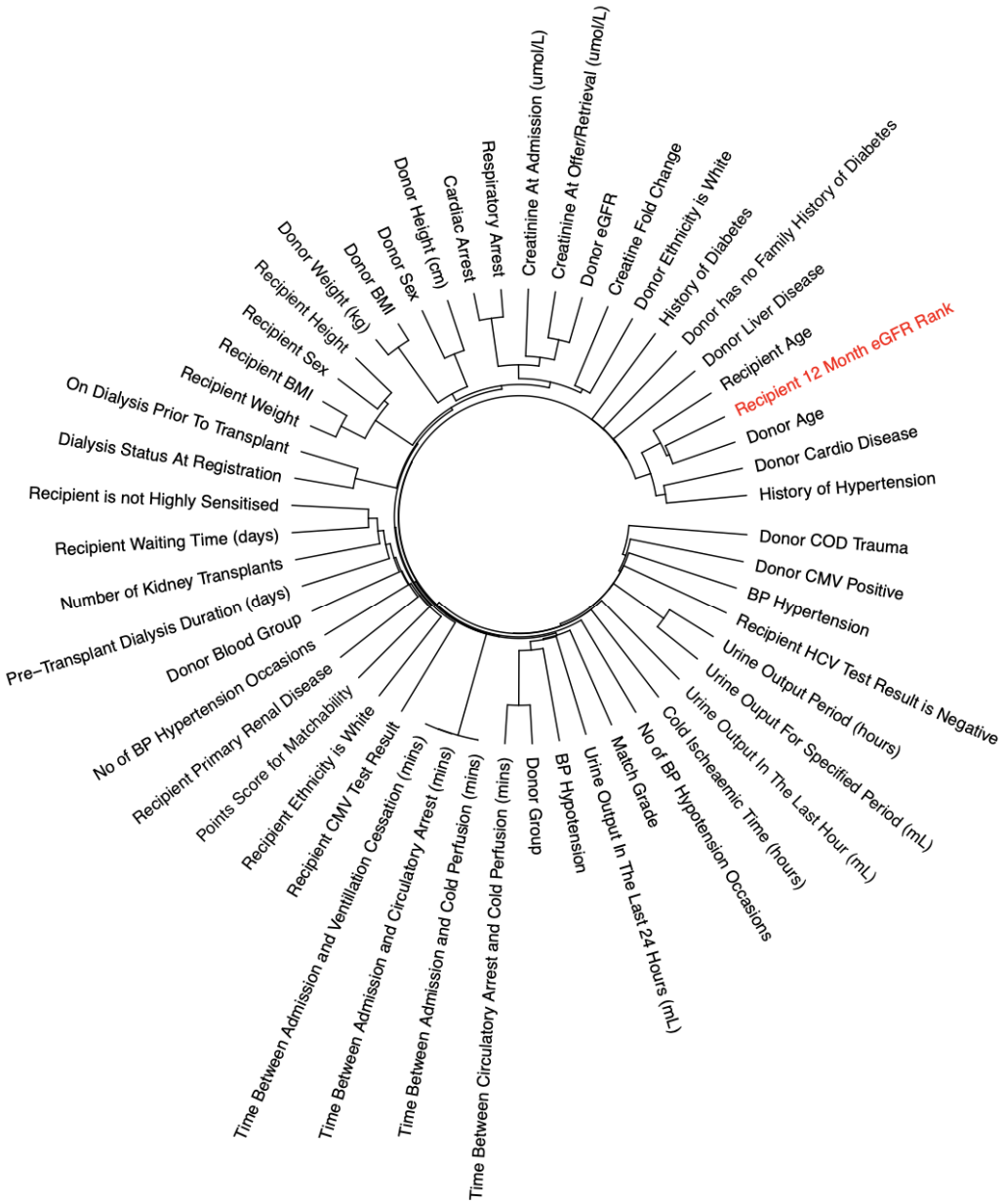


Figure 2: Donor and recipient clinical and demographic data association with recipient 12 month eGFR rank
Single-linkage hierarchical clustering of curated, imputed clinical variables by relative association strength (taking distance as 1-association). The outcome variable (ranked recipient eGFR at 12 months post-transplantation) is highlighted in red.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

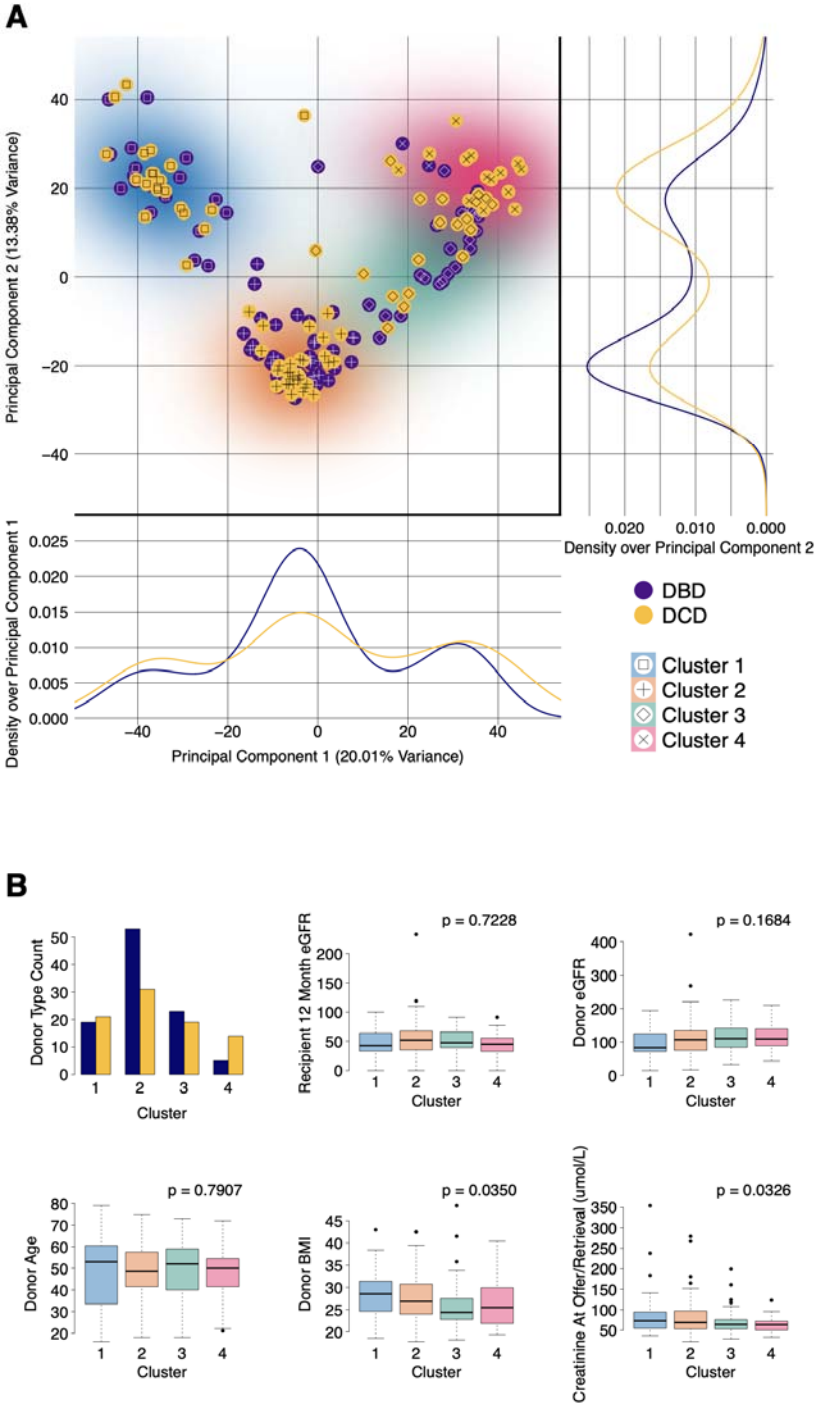


Figure 3: Unbiased analysis of pretransplant kidney proteomes and cluster associations

A: Unbiased analysis of proteomic data by k-means clustering. Sample separation by Principal Component Analysis. Top Left: Samples were assigned to four clusters by k-means. Bottom & Right: There was a difference in the distribution of DBD and DCD donors across clusters, with the DBD donors being more heavily concentrated in Cluster 2 ('+' symbol; orange shading), and DCD in Cluster 4 ('x' symbol; pink shading)

B: There were no associations between proteome clusters and most donor and recipient factors, except for mildly significant differences in donor BMI and creatinine (selected comparisons shown; left-right, top-bottom: donor type, recipient 12-month posttransplant eGFR (outcome), donor eGFR, donor age, donor BMI, donor creatinine at retrieval).

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

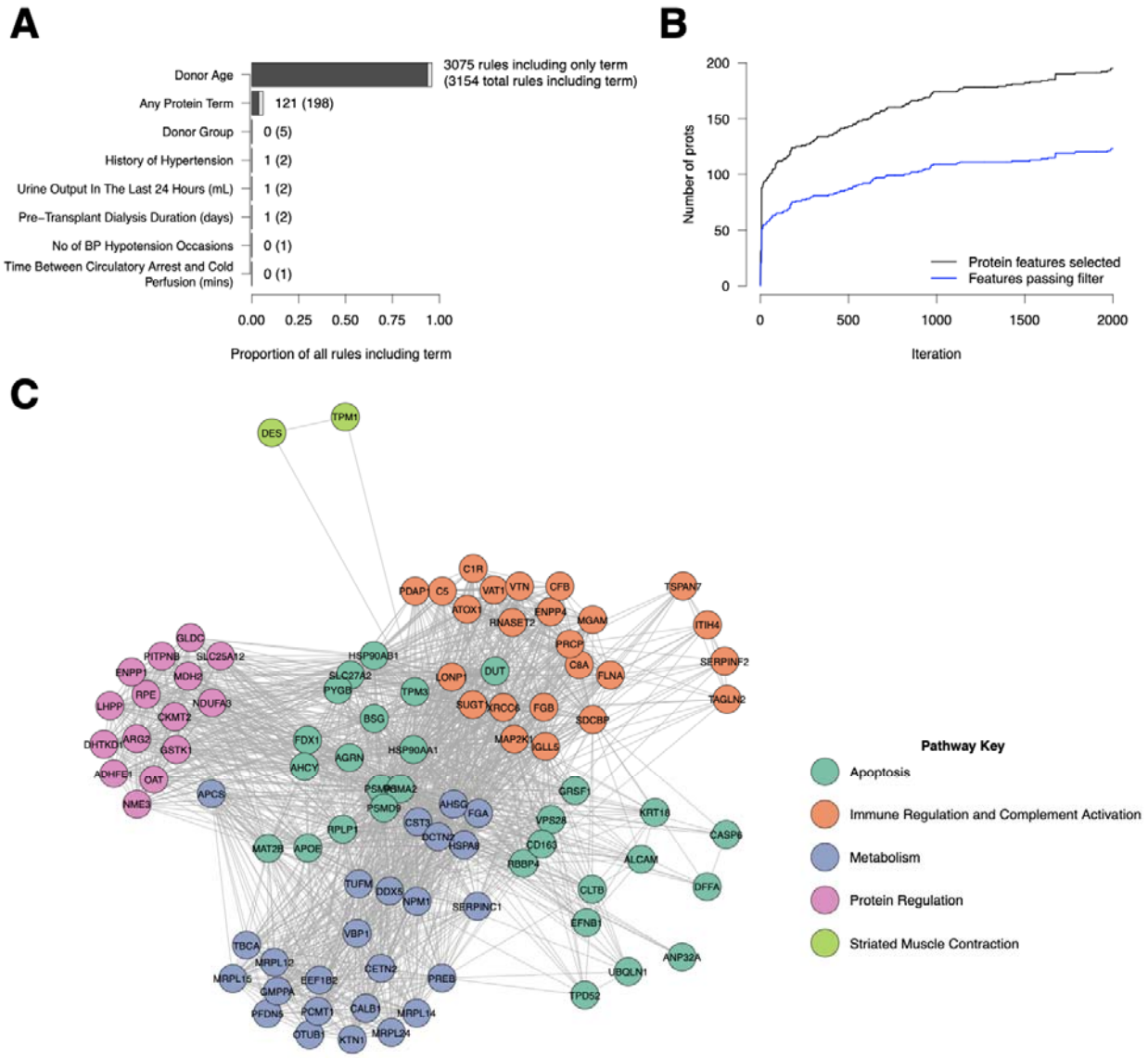


Figure 4: Age and combined age:protein related associations link to construction of age-modulated immune metabolic biological networks

A: Prediction Rule Ensemble (PRE) modelling was performed in an iterative manner to select protein and clinical variable associations with ranked eGFR. At each iteration, only proteins not previously featured in a model were considered. The rules found across all iterations were dominated by donor age terms.

B: Cumulative protein features identified at each iteration. Black line: all features identified by feature selection approach. Blue line: features passing the secondary filter for predictive power and accuracy.

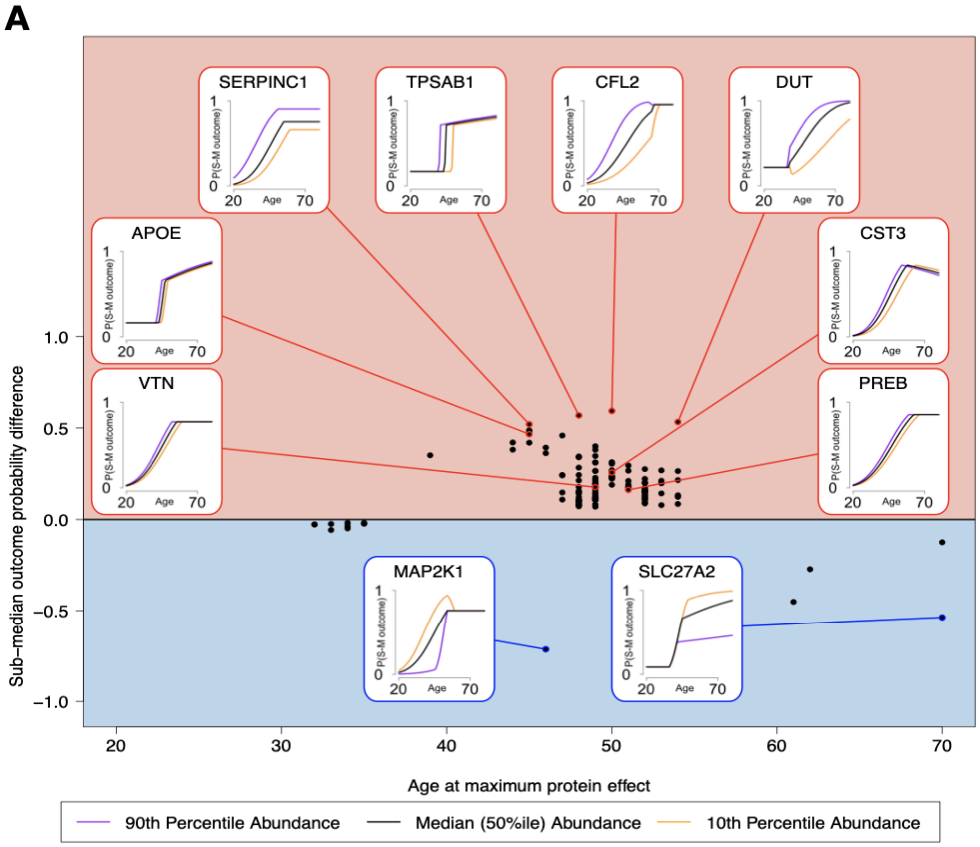
C: Shared Reactome pathway membership network analysis of filtered features. Nodes are colored by assigned cluster, and the clusters are annotated according to the top three most enriched pathways within each cluster.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Cluster Label	Top 3 Shared Pathways
Immune Regulation and Complement Activation	Immune System Innate Immune System Regulation Of Complement Cascade
Protein Metabolism and Regulation	Metabolism Of Proteins Post-translational Protein Modification Mitochondrial Translation Elongation
Metabolism	Metabolism Metabolism Of Amino Acids And Derivatives Pyruvate Metabolism And Citric Acid (TCA) Cycle
Apoptosis	Developmental Biology Role Of GTSE1 In G2/M Progression After G2 Checkpoint Apoptosis
Striated Muscle Contraction	Striated Muscle Contraction

Table 2: Shared Pathway Network Clusters
Proteins in Figure 4C were clustered by pathway membership, forming 4 major clusters and one minor cluster (Striated Muscle Contraction). We assigned summary labels to each cluster based on the top 3 pathways with shared membership in each cluster.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .



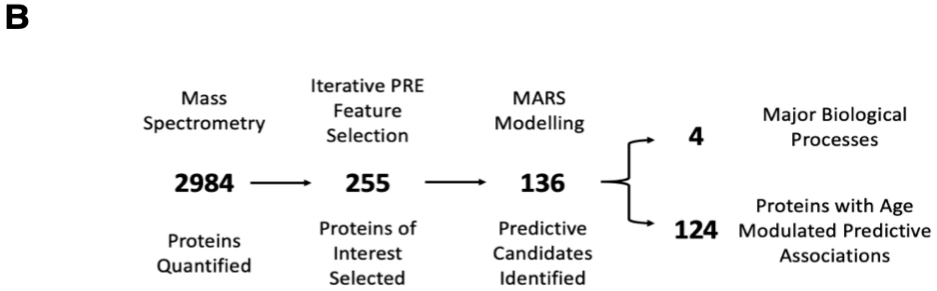


Figure 5: Modelled associations between proteins and kidney transplant outcome change with donor age

A: Ages at which the predicted probability of sub-median outcome is most different between the 10th percentile and 90th percentile of protein abundance. *x* axis: age at which difference is greatest (i.e. when protein has greatest effect). *y* axis: greatest difference.

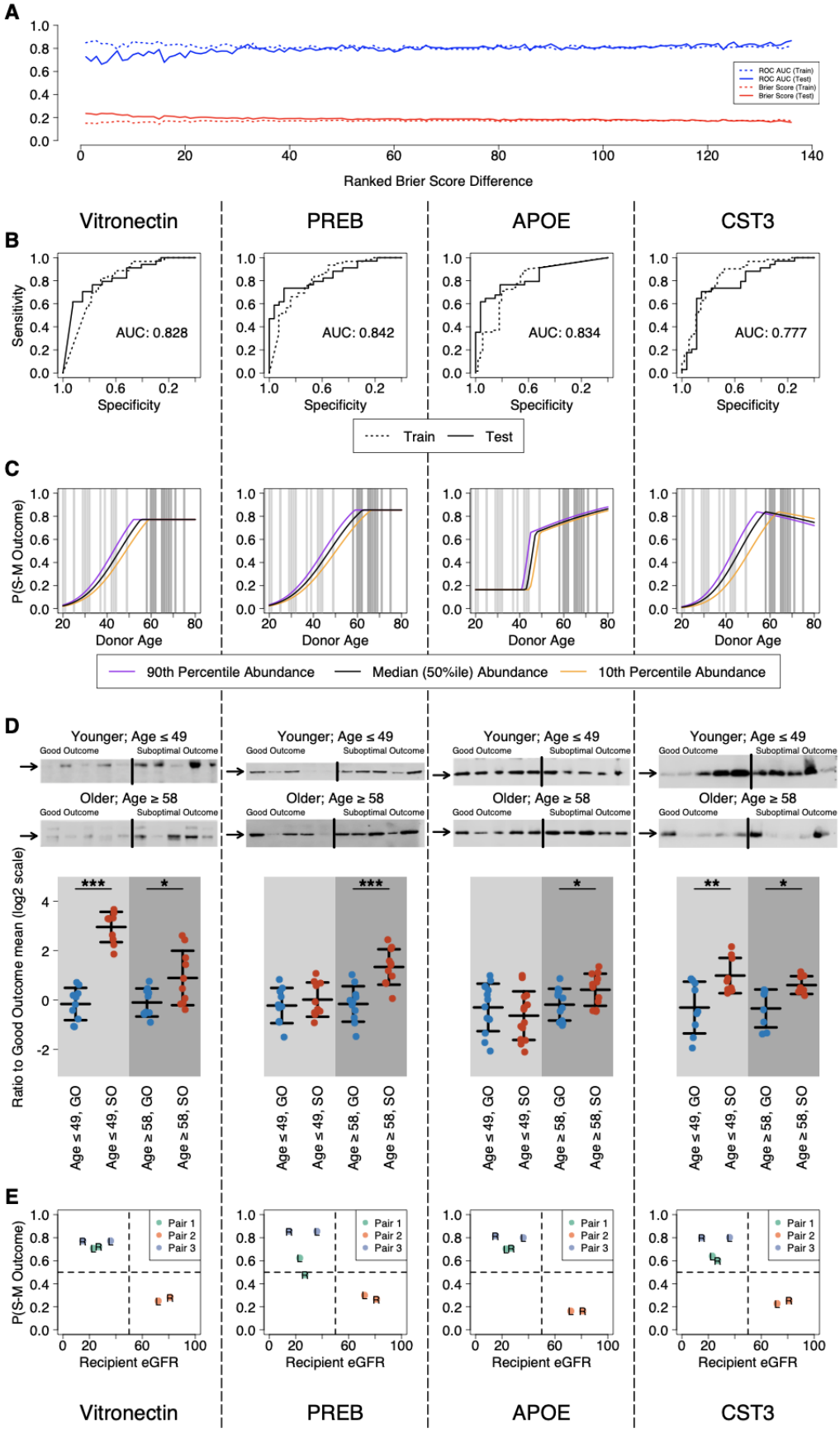
Proteins above *x*=0 are modelled as having a more negative association with outcome when the protein abundance is high, at that donor age. Proteins below *x*=0 are modelled to have a more positive association with outcome when protein abundance is high, at that donor age.

Proteins with absolute net difference >0.5 are labelled, as well as the selected proteins VTN, PREB, APOE and CST3.

The inset graphs indicate how the prediction of sub-median outcome ("P(S-M outcome)"; *y* axes) changes with donor age (*x* axes) for labelled proteins. Black trace: prediction at median protein abundance. Purple trace: prediction at 90th percentile of protein abundance. Orange trace: prediction at 10th percentile of protein abundance. The corresponding point on the main figure thus indicates the age at which the difference between orange and purple lines is greatest.

B: Summary of feature selection and modelling analysis

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .



682

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23288011; this version posted April 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Figure 6: Orthogonal validation confirms age-modulated immuno-metabolic proteins predict 12-month transplant outcomes

A: Validation of models in test dataset. Models are plotted in order of decreasing Brier score (mean squared prediction error) *difference* between test and train data along the x axis. The lower two traces indicate the Brier score in train (purple) and test (green) data. The upper two traces indicate the AUC from the corresponding ROC analyses in train (orange) and test (blue) data.

B-E: Validation of four selected proteins. Left-Right: VTN, PREB, APOE, CST3.

B: Final ROC curves and AUC values for models trained on each protein (and donor age) against test data. The dotted line indicates the original performance against training data.

C: Change in the prediction of sub-median outcome ("P(S-M Outcome)"; *y* axes) with donor age (*x* axes) for each protein. Black trace: prediction at median protein abundance. Purple trace: prediction at 90th percentile of protein abundance. Orange trace: prediction at 10th percentile of protein abundance. (These are the same as the inset graphs in Figure 5). The light grey and dark grey vertical lines, respectively, indicate the corresponding 'younger' and 'older' sampled ages for the western blots below.

D: Western blots comparing younger (age $\leq$ 49) and older (age $\geq$ 58) donors between Good Outcome (GO; eGFR $\geq$ 60) and Suboptimal Outcome (SO; eGFR $\leq$ 40) outcome tertiles. Top row: representative western blots (n=5 per group) from comparison of younger donors. Middle row: representative western blots (n=5 per group) from comparison of older donors. Bottom row: result values for all quantified samples relative to the GO mean. Error bars indicate ±1 standard deviation; the central wider bar indicates mean. Significance stars indicate t-test comparison p-values (***: < 0.001, *: < 0.05).

E: Predicted outcome for six paired Left (L) and Right (R) kidneys from three donors. *x* axes: recipient eGFR at 12 months (i.e. actual outcome). *y* axes: predicted probability of sub-median outcome ("P(S-M Outcome)") using models trained on each protein with donor age. Vertical dotted line indicated median outcome (eGFR = 50). Horizontal dotted line indicates P(Sub-Median Outcome) = 0.5