

1 **TITLE PAGE**

2

3 **Age-Modulated Immuno-Metabolic Proteome Profiles of Deceased Donor**

4 **Kidneys Predict 12-Month Posttransplant Outcome**

5

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

9

10 **Affiliations:**

11 <sup>1</sup>Big Data Institute, Nuffield Department of Medicine, University of Oxford; Oxford, United  
12 Kingdom.

13 <sup>2</sup>Target Discovery Institute, Centre for Medicines Discovery, Nuffield Department of Medicine,  
14 University of Oxford; Oxford, United Kingdom.

15 <sup>3</sup>Research and Development, NHS Blood and Transplant; Bristol & Oxford, United Kingdom.

16 <sup>4</sup>Nuffield Department of Surgical Sciences and Oxford Biomedical Research Centre, University  
17 of Oxford; Oxford, United Kingdom.

18 <sup>5</sup>Engineering Department, University of Oxford; United Kingdom.

19 <sup>6</sup>University Hospital Oxford; Oxford, United Kingdom.

20 <sup>7</sup>Center for Health Data Science, University of Copenhagen; Copenhagen, Denmark.

21 <sup>8</sup>Novo Nordisk Foundation Center for Protein Research, University of Copenhagen;  
22 Copenhagen, Denmark.

23 \*These authors contributed equally.

24

25 †Corresponding authors. Email: [philip.charles@ndm.ox.ac.uk](mailto:philip.charles@ndm.ox.ac.uk); [maria.kaisar@nds.ox.ac.uk](mailto:maria.kaisar@nds.ox.ac.uk)

26

27 **Keywords:**

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

30

31 **ABSTRACT**

32 Organ availability limits kidney transplantation, the best treatment for end-stage kidney disease.

33 Deceased donor acceptance criteria have been relaxed to include older donors with higher risk of

34 inferior posttransplant outcomes. More granular prediction models, based on deeper resolution

35 organ assessment and understanding of damage processes, could substantially improve donor

36 organ allocation and reduce graft dysfunction risk. Here, we profiled pre-implantation kidney

37 biopsy proteomes from 185 deceased donors by high-resolution mass spectrometry and used

38 machine learning to integrate and model these data, and donor and recipient clinical metadata to

39 predict outcome. Our analysis and orthogonal validation on an independent cohort revealed 136

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

47

## 48 **TRANSLATIONAL STATEMENT**

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

55

## 56 **INTRODUCTION**

57 Kidney transplantation is the optimal treatment for end-stage kidney disease. When  
58 compared to dialysis, transplantation profoundly increases life-expectancy, improves quality of  
59 life and is cost-effective. However, the lack of suitable donor organs for kidney transplantation  
60 limits the health benefits to patients with chronic kidney disease, and often prolongs dialysis  
61 treatment, with increased morbidity and mortality.

62           The shortage of organs, a decline in living donation in some countries and emerging  
63 ageing populations have driven an increased utilization of older donor kidneys <sup>1</sup>. More than half  
64 of all deceased donor kidneys now offered are from older donors <sup>2</sup>.

65           Ageing is associated with a time-dependent decline of organ function, which is especially  
66 evident in the kidney with the appearance of histologic lesions, such as tubular atrophy,  
67 interstitial fibrosis, glomerulosclerosis, and arteriosclerosis. Kidney function progressively  
68 declines with age, due to structural changes that cause a reduction of functioning glomeruli and  
69 renal mass, podocyte dysfunction, and malfunction of mechanisms important in cellular repair <sup>3</sup>.  
70 Glomerular diseases are in turn more common and associated with worse outcomes in older  
71 patients <sup>4</sup>. Age accelerates the transition from Acute Kidney Injury (AKI) to chronic injury <sup>5</sup> and  
72 is an independent risk factor of graft dysfunction and graft loss for deceased donor kidneys <sup>6</sup>,  
73 furthermore, older donors are more likely to suffer from additional co-morbidity risk factors such  
74 as diabetes, hypertension or cardiovascular disease.

75           Donor age has been incorporated in all clinical scoring algorithms to inform clinical  
76 decision of donor kidney allocation <sup>7,8</sup>. However, donor age alone is insufficient to consistently  
77 predict transplant outcomes. Current front-line models also include further clinical factors such  
78 as terminal serum creatinine, history of hypertension and diabetes <sup>8,9</sup> and show consistent  
79 performance across demographic cohorts but still only have a predictive accuracy of around 65%  
80 <sup>10</sup>. Factors commonly included in these models may not reflect the nuances of biological  
81 susceptibility to injury, subclinical trajectories of progression to dysfunction or capacity for renal  
82 recovery.

83           The use of pre-implantation biopsies to supplement standard clinical scoring algorithms  
84 has hitherto focused on histological scoring systems that assess evidence of chronic donor-

85 related damage, such as glomerulosclerosis or arteriolar changes <sup>11</sup>. Histology-based assessments  
86 are particularly vulnerable to biopsy reproducibility issues, even under codified protocols.  
87 Although small studies have suggested correlation with graft outcomes, there is only poor  
88 correlation with kidney function in the recipient <sup>12,13</sup>.

89 It is likely that molecular analyses could offer a clearer, higher resolution assessment of  
90 organ state; but this requires deeper understanding of molecular phenotypes specifically  
91 associated with poor outcomes rather than the immediate but potentially recoverable acute injury  
92 sustained during the donor management process and organ retrieval. For example, deceased  
93 donors are often assessed as having AKI based on serum creatinine levels <sup>14</sup>, which implies an  
94 inflammatory response, however AKI does not itself associate with longer term poor outcomes  
95 <sup>13-16</sup>, so we should not expect all inflammatory associated molecular signatures to be outcome-  
96 associated. A broad picture analysis at relevant molecular phenotype levels is necessary; an  
97 obvious first candidate is the proteome as this represents the end-state response associated with  
98 both changes in expression and post-translational regulation.

99 Proteomic studies in donor kidneys have heretofore lacked the sample size to be  
100 representative across donor age and demographic ranges <sup>17</sup>, but advances in proteomic  
101 technology such as data independent acquisition (DIA) mass spectrometry <sup>18</sup> now allow the  
102 sensitivity and depth required for analysis of hundreds of samples to be practical. Larger  
103 analyses are central towards taking the next steps toward precision medicine, and permit higher-  
104 dimensional analysis (through the integration of proteomic profiles with clinical and  
105 demographic phenotypes) to address the heterogenous demographic factors affecting patient  
106 outcomes <sup>19</sup>. Machine learning approaches in particular offer powerful tools to extract the  
107 maximum amount of knowledge from experimental cohorts given practical and cost limitations

108 of sample acquisition and analysis, and have been applied to disease staging, prediction of  
109 disease recurrence, monitoring treatment response, and the identification of diagnostic,  
110 prognostic, and predictive biomarkers <sup>20,21</sup>.

111 Here, we benefit from the granularity provided by our MS-based proteomics profiling to  
112 report age- and immunometabolism-related proteomic signatures in pre-implantation kidney  
113 biopsies that are associated with suboptimal transplant outcomes against a clinical and  
114 demographic variability background.

115

## 116 **METHODS**

### 117 **Study Design**

118 Deceased donor pre-transplantation kidney biopsies (n=186; 1 sample was later removed  
119 during data QC) were obtained from the Quality in Organ Donor (QUOD) biobank, a national  
120 multi-center UK wide bioresource of deceased donor clinical samples procured during donor  
121 management and organ procurement.

122 Selection of biopsies of donor kidneys was based on paired 12-month post-transplant outcomes.  
123 To minimize the impact of recipient factors on outcomes we only included kidneys for which the  
124 contralateral kidney was transplanted and had similar 12-month post-transplant outcome. Donor  
125 kidneys were selected to include the continuum of transplant outcomes i.e. the full range of  
126 estimated Glomerular Filtration Rate (eGFR; specified throughout in units of ml/min/1.73 m<sup>2</sup>)  
127 from primary non-function to eGFR > 80) and, where possible, to exclude extreme demographic  
128 or clinical factors. All clinical samples were linked to corresponding donor and recipient

129 demographic and clinical metadata, provided by NHS Blood and Transplant National Registry.  
130 Detailed clinical and demographic data are summarized in Table 1.

### 131 **Study Approval and Ethics statement**

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

### 135 **Experimental Protocols and Statistical Analysis**

136 Please see Supplementary Methods.

137

## 138 **RESULTS**

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

140 Kidney biopsies were obtained from Donation after Brain Death (DBD) donors and  
141 Donation after Circulatory Death (DCD donors) at the back table immediately after kidney  
142 procurement. To minimize the impact of factors related to post-procurement and recipient  
143 characteristics, donors were selected on the basis of paired 12-month posttransplant outcomes.

144 For the purposes of exploratory analysis, we considered eGFR values in two ways.  
145 Firstly, for comparison of clinical factors, we grouped 12 months posttransplant outcomes into  
146 tertiles; Suboptimal Outcome (SO;  $eGFR \leq 39$ ), Intermediate Outcome (IO;  $40 < eGFR \leq 59$ ),  
147 and Good Outcome (GO;  $eGFR \geq 60$ ) (Figure 1). We refer to this henceforth as ‘stratified  
148 eGFR’. Secondly, all eGFR values (both recipient and donor) were rank-transformed so that,  
149 where possible, we could test/model against a continuum of outcomes while mitigating against

150 factors introducing extreme values or values recorded as 0 due to graft failure. We refer to  
151 ‘ranked eGFR’ henceforth to indicate ranked recipient eGFR at 12 months posttransplant.

152 Clinical metadata confirmed that the selected donor groups were representative of the  
153 donor population in the UK (UK Renal Registry, 2022<sup>22</sup>) (Table 1). Further, we found no  
154 significant association between donor type and ranked eGFR (t-test;  $p = 0.2028$ ). We then  
155 investigated associations between clinical variables and stratified eGFR subgroups within DBD  
156 and DCD, and between DBD and DCD within stratified eGFR subgroups (Supplementary Table  
157 1). In DBD as in DCD, donor age was significantly different between the subgroups with SO  
158 being older (Fisher’s exact test; DBD:  $p = 4.285e-7$ ; DCD:  $p = 2.080e-8$ ). Among both DBD and  
159 DCD donors there was a significant difference in UKKDRI between outcome groups (ANOVA  
160 F-test; DBD:  $p = 1.515e-5$ ; DCD:  $p = 7.006e-7$ ). Histories of hypertension were different  
161 between outcome subgroups (ANOVA F-test; DBD:  $p = 0.0010$ ; DCD:  $p = 0.0336$ ), but histories  
162 of diabetes were not significantly different (ANOVA F-test;  $p = 0.1408$ ; DCD:  $p = 0.2083$ ).

163 Terminal serum creatinine levels were similar across outcome subgroups (ANOVA F-  
164 test; DBD:  $p = 0.8601$ ; DCD:  $p = 0.1755$ ), although within the GO group it was higher in DBD  
165 than in DCD (t-test;  $p = 0.0149$ ).

166 After imputation of missing values, we examined the associations between clinical  
167 variables common to both DBD and DCD donors, calculating a composite association score (see  
168 methods) and then clustering by single linkage on association subtracted from 1 as a distance  
169 measure (Figure 2). The strongest associations with ranked eGFR were donor age (Pearson’s  $r =$   
170  $0.52$ ), and recipient age ( $r = 0.28$ ). Donor history of hypertension and cardiological disease also  
171 clustered closely using a single-linkage approach due to correlation with donor age ( $r = 0.35$  and



172  $r = 0.33$  respectively) but had a weaker direct correlation with outcome ( $r = 0.30$  and  $r = 0.23$   
173 respectively).

#### 174 **Unsupervised analysis of pretransplant kidney proteomes**

175 The proteomic analysis quantified 2984 protein groups with 50% or less missing values  
176 (out of 7790 identified protein groups in total) over 185 samples and the 20 pool sample runs  
177 (Supplementary Figure 1A). Analysis of the pool runs showed minimal technical variance over  
178 the course of sample acquisition with a squared mean pairwise Z-corrected Pearson correlation  
179 coefficient (a multi-way R-squared) of 0.94. Furthermore, six samples were paired biopsies from  
180 the left and right kidneys of three donors. These samples showed high concordance in quantified  
181 protein intensity values between pairs, with R-squared values of 0.71, 0.92 and 0.91 for each of  
182 the three pair donors (Supplementary Figure 1B).

183 We performed an exploratory analysis of the proteomics data by Principal Component  
184 Analysis (PCA) to look for underlying data trends. Most of the variance observed in the samples  
185 was concentrated in the first two principal components (PC1: 20.01%; PC2: 13.38%; Figure 3A)  
186 and K-means clustering identified 4 distinct clusters (Figure 3A).

187 Cluster membership associated with donor type, with a preponderance of DBD samples  
188 towards Cluster 2 and a preponderance of DCD samples towards Cluster 4 (Figure 3B, upper left  
189 panel;  $p = 0.0235$ ). Sample clustering did not associate strongly with recipient eGFR at 12  
190 months post transplantation ( $p = 0.4134$ ), nor with donor eGFR ( $p = 0.1684$ ), or donor age ( $p =$   
191  $0.7907$ ) (Figure 3B, upper middle and right and lower left panels). There was a weakly  
192 significant association between cluster membership and donor BMI ( $p = 0.0350$ ) and similarly  
193 with serum creatinine ( $p = 0.0326$ ) (Figure 3B, lower middle and right panels).

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

196 To identify possible clinical variable-protein interaction effect relationships with stratified  
197 outcomes, we used a machine learning approach (Prediction Rule Ensembles<sup>23</sup>; PRE) to analyze  
198 the set of all donor type-independent clinical variables and all quantified proteins, followed by  
199 multivariate adaptive regression spline modelling<sup>24</sup> to assess individual protein relationships.

200 We first split our data into a training and test sets, excluding the six paired kidneys, and  
201 sampling equally across stratified eGFR using a 2/3:1/3 train:test split. Feature selection and  
202 predictive model training was performed on the training set only. The six paired kidneys were  
203 reserved as a second ‘biological duplicates’ test set.

204 PRE modelling finds a minimal (but not necessarily the only minimal) set of predictors  
205 for outcome in the form of decision tree, linear regression and multivariate adaptive regression  
206 spline<sup>25</sup> rules. A single application of PRE does not yield an exhaustive list of predictive  
207 candidates. We performed PRE in an iterative manner to ‘mine’ for candidate proteins, whereby  
208 proteins, but not clinical variables, identified in the final ensemble of any previous model were  
209 excluded from the dataset for future models. We performed 2000 iterations of PRE, generating  
210 3282 rules across all ensembles. An immediate observation was that the primary contributor to  
211 the rule cohort was donor age, featuring as a term in 3154 (~96.1%) rules; in comparison, protein  
212 terms (collectively) featured in 198 (~6.0%) rules, while no other non-protein term yielded a  
213 meaningful contribution – the next largest non-protein term was donor group, featuring in 5  
214 (~0.1%) rules (Figure 4A). This process generated 195 selected proteins. As iterations passed  
215 1000, the selection process became increasingly inefficient, in terms of iterations required to  
216 identify another candidate protein. However new hits were still generated up to termination at

217 iteration 2000, suggesting there was more to find. Rather than continue with further iterations,  
218 we supplemented the list of selected proteins with all other proteins that had high correlation  
219 (Pearson's  $r > 0.65$ ) with any of the 195 already selected proteins; this brought the list up to 255  
220 selected protein features.

## 221 **Regression spline modelling reveals protein associations with posttransplant outcome are** 222 **modulated by donor age**

223 Next, we tested each of the 255 protein features for individual association with outcome,  
224 beyond the effect of donor age alone. Since rank transformation is cohort specific, in order to  
225 generate results which could be generalized to other cohorts we modelled against a specific  
226 outcome binary, calibrated against a population-level threshold. Based on UK Renal Registry  
227 Reports, UK median eGFR at 12 months posttransplant was approximately  $\sim 50.25$  ml/min  
228 ( $\pm 0.24$  standard error) across the weighted average for median DBD and DCD eGFR at 12  
229 months posttransplant since 2013<sup>22</sup>; for simplicity we used a threshold of 50 (eGFR units  
230 ml/min/1.73 m<sup>2</sup>). We refer to 'sub-median outcome' henceforth to refer to recipient eGFR at 12  
231 months posttransplant less than 50. Using multivariate adaptive regression spline modelling, we  
232 generated predictive models for sub-median outcome using each protein, donor age, and any  
233 potential age:protein interactions. This was performed in a regularized framework such that only  
234 a minimal set of terms predictive of outcome were retained in each model. We then discarded  
235 any protein that either did not feature in any term in its corresponding model, or whose model  
236 gave a worse overall prediction error (Brier score<sup>26</sup>) than donor age alone.

237 After filtering we had identified 136 proteins which predicted sub-median outcome. We  
238 performed a network analysis of shared Reactome pathways (Figure 4) and, using walktrap  
239 clustering, found that these relationships were characterized by 4 major clusters; *Immune*

240 *Regulation and Complement Activation* (top 3 pathways ‘Immune System’, ‘Innate Immune  
241 System’, ‘Regulation Of Complement Cascade ’), *Protein Metabolism and Regulation* (top 3  
242 pathways: ‘Metabolism Of Proteins’, ‘Post-translational Protein Modification’ , ‘Mitochondrial  
243 Translation Elongation’), *Metabolism* (top 3 pathways: ‘Metabolism’, ‘Metabolism Of Amino  
244 Acids And Derivatives’, ‘Pyruvate Metabolism And Citric Acid (TCA) Cycle’) and *Apoptosis*  
245 (top 3 pathways: ‘Developmental Biology’, ‘Role Of GTSE1 In G2/M Progression After G2  
246 Checkpoint’, ‘Apoptosis’).

247 For 124 proteins, the prediction included an age:protein interaction term where the  
248 predictive effect of protein abundance was modulated by age, independent of the effect of age  
249 alone or protein abundance alone (Figure 5). To visualize these effects, we used each model to  
250 simulate the effect of increasing donor age with a high (90<sup>th</sup> percentile), median and low (10<sup>th</sup>  
251 percentile) abundance value for the protein (Figure 5A and Supplementary Figure 2). In all cases,  
252 the chance of sub-median outcome was (unsurprisingly) driven primarily by donor age, with a  
253 plateau at high donor age as a result of the low frequency of high-donor age, above-median  
254 outcomes in both our data (<5% above age 60 in our training set) but also in the general  
255 population. The difference in the modelled effect of changing protein abundance was striking;  
256 the majority of proteins were positively associated with the chance of sub-median outcome  
257 (simplistically, more protein = worse outcome), with the effect appearing to increase up to  
258 around donor age 45-55. These included representative proteins for all four major clusters  
259 including the known markers Cystatin-C (CST3; nephron function) and Vitronectin (VTN;  
260 fibrosis) as well as a protein notorious for age-modulated associations with other diseases,  
261 Apolipoprotein E (APOE). In contrast, several proteins were negatively associated with the  
262 chance of sub-median outcome (simplistically more protein = better outcome), with the largest

263 effects shown by MAP2K1 and SLC27A2, with the latter in particular being modelled as having  
264 minimal effect in donors younger than 40.

265 The full list of all 136 proteins, including a summary of age modulation effect, is given in  
266 Supplementary Table 2. The selection and filtering steps are summarized in Figure 5B.

### 267 **Orthogonal validation confirms model performance, including age-modulated immuno-** 268 **metabolic impact on transplant outcomes**

269 We adopted two orthogonal approaches to validate our results. Firstly, we assessed the  
270 performance of each model on test data. Going from train to test data, the models showed a small  
271 increase in accuracy (Brier score; mean square error) and a small decrease in overall predictive  
272 performance as measured by the area under the curve (AUC) (Figure 6A), indicating that the  
273 models are generalizable. Most of the models (~110/136) showed high levels of accuracy with  
274 negligible differences between test and train model performance.

275 Secondly, we selected several cluster-representative proteins (VTN, APOE, CST3 and  
276 Prolactin Regulatory Element Binding; PREB) with well characterized available antibodies that  
277 had good (AUC > 0.75) predicted outcome performance on our test data set (Figure 6B). We  
278 investigated the predicted pattern of associations between protein abundance and outcome  
279 (Figure 6C), by performing western blot validation of our results (Figure 6D). We selected  
280 samples with remaining material from our cohort from the Good and Suboptimal Outcome  
281 tertiles of our sample set, to compare protein abundance between younger (oldest sample 49) and  
282 older (youngest sample 58) donors. Our western blot results were broadly consistent with the  
283 associations anticipated by our modelling. For VTN, our model suggests a strong (and  
284 strengthening with age) association between protein abundance and outcome in younger donors,  
285 and we observed a significant difference in abundance between GO and SO outcome strata (t-

286 test;  $p=2.107e-9$ ). In older donors, the model suggests a weaker (and weakening with age)  
287 association, and we observed a mildly significant difference between abundance and outcome (t-  
288 test;  $p=0.2450$ ). For PREB, the model suggests a weaker association in younger donors than in  
289 VTN since the age of maximum difference is shifted towards older donors. We did not observe  
290 any significant difference by western (t-test;  $p=0.4530$ ) in young donors. At older ages, the  
291 model suggests a larger association in the older age range, weakening at an older age than VTN;  
292 we observe a corresponding significant difference in PREB abundance by western blot (t-test,  
293  $p=8.800e-5$ ). For APOE, the model suggests that the strongest association is over the middle of  
294 the age range, where outcome changes rapidly with donor age. Younger donor samples did not  
295 provide good coverage of the age range region of expected difference, and none was observed (t-  
296 test,  $p=0.3719$ ). In older donors, the effect of APOE difference weakens but remains present, and  
297 consistent with that we saw a mildly significant difference in APOE abundance between  
298 outcomes (t-test,  $p=0.0323$ ). For CST3, the model again predicted a strong association in young  
299 donors, which was consistent with the significant difference observed by western blot (t-test,  
300  $p=0.0084$ ). In older donors the model actually predicts the association to weaken and even  
301 reverse, such that GO samples would tend to have higher CST3 than SO samples; by western  
302 blot, we saw a nonsignificant difference, but (in contrast to the prediction) still with a positive  
303 median protein abundance difference from GO to SO. As a final check, we compared the  
304 predicted outcome for each of the six paired kidneys from the second 'biological duplicates' test  
305 set against their actual recipient eGFR at 12 months posttransplant (Figure 6E). All three kidney  
306 pairs in this dataset had very consistent outcomes across pairs; two pairs with sub -median  
307 outcome (15 and 36 ml/min; 23 and 27 ml/min) and one pair with much better outcome (72 and  
308 81 ml/min). All four models assigned the four kidneys from two of the donors that went on to

309 have a sub-median outcome a probability of sub-median outcome greater than 0.5 (save in one  
310 case for PREB, where the probability was ~0.48), and likewise assigned both kidneys from the  
311 pair which both went on to have a good outcome a low probability of sub-median outcome.

## 312 **DISCUSSION**

313         Increasingly, kidneys from older donors are utilized for transplantation to expand the  
314 deceased donor organ supply. The shortage of optimal organs, combined with increased  
315 utilization of organs from older donors, increases risks of graft failure or functional decline. For  
316 any individual donor organ, post-transplant outcomes remain hard to predict at the point of organ  
317 acceptance. Here, we show that age-modulated kidney proteomic profiles improve risk  
318 stratification of donor kidney quality, revealing clinically relevant age-protein interaction effects.

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

327         The effect of donor age is not a novel finding, but it is particularly interesting to note that  
328 pretransplant kidney proteomes were the second most powerful contributor to outcomes; and that  
329 integration of the age and proteomic information resulted in enhanced prediction of 12 month  
330 sub-median function. One factor often described as relevant to transplant outcome is donor type  
331 <sup>28</sup>. This is certainly true at a clinical level in terms of donor management, and donation after

332 circulatory death is considered an adverse factor for transplant outcome in the US <sup>7</sup> (although not  
333 in the UK <sup>8,29</sup>). Our initial PCA analysis found that non-supervised clustering of the sample  
334 proteomes did partially separate samples by type. However, these differences did not extend to  
335 association with outcome in our cohort. Not only did we fail to observe any direct association  
336 between donor type and outcome, but donor type did not feature as an interaction term with any  
337 protein identified by our iterative modelling. It is possible that such association may be obscured  
338 by other sample characteristics associated with the overwhelming effect of donor age, especially  
339 if the effect is relatively small. However, without disputing different donor type-specific  
340 mechanisms for how kidney injury may be sustained <sup>30</sup>, our data are consistent with previous  
341 reports that the level of injury (rather than the cause) is the primary contributor towards the  
342 potential for recovery <sup>31</sup>.

343         Within our final list of 136 proteins associated with outcome there is a common theme of  
344 implication in immune response to kidney injury (including both chronic injury, and acute  
345 injury) particularly as a result of ischemic metabolic disruption. Our analysis of proteins  
346 associated with outcome revealed that most (124/136 candidates) showed evidence of age-  
347 moderated differences in their effect; in most cases, that as donor age increases, higher levels of  
348 the protein become more negatively associated with outcome. The effect of second-order age  
349 interactions, where the weightings of other terms are themselves age dependent, has not (to our  
350 knowledge) been explored in terms of transplant outcome. Many of the proteins we find to be  
351 age-moderated have known relevance, which is reassuring, however several have not previously  
352 been reported in the context of transplant outcome or kidney dysfunction; it is possible the effect  
353 of age modulation is key to fully understanding the links between molecular predictors and  
354 transplant outcomes.



355           A prominent age-modulated example of a chronic injury associated marker in our  
356 candidate list is VTN, a primary component of the extracellular matrix involved in cell  
357 adhesion, enhancing the activity of plasminogen activator inhibitor-1 (PAI-1), and inhibition of  
358 the terminal complement pathway<sup>32</sup>. Vitronectin has been suggested as a biomarker of kidney  
359 fibrosis although the mode of its multifaced action needs further investigation<sup>33</sup>. In addition,  
360 examples for known acute injury associated markers include several complement components,  
361 including two components of the membrane attack complex, C5 (in the form of C5b cleavage  
362 product) and C8A, which has been associated with tissue injury resulting from  
363 ischemia/reperfusion<sup>34,35</sup>, Complement Component 1r (C1R), part of the activation complex for  
364 the classical complement pathway<sup>36</sup>, and Complement Factor B (CFB), a component of the  
365 alternative pathway. Another candidate associated with immune regulation is Maltase  
366 Glucoamylase (MGAM), characterized as an intestinal enterocyte but with expression in several  
367 tissues including kidney, and whose presence in urinary exosomes been cited as a marker of AKI  
368 in cirrhosis patients<sup>37</sup>. Finally, Mitogen-activated protein kinase 1 (MAP2K1, aka MEK1), as a  
369 key component of the MAP kinase signal transduction pathway and therefore closely involved in  
370 both cellular control and immune regulation (as part of TNF $\alpha$  signaling response<sup>38</sup>), is notable as  
371 one of the few proteins for which higher abundance was associated with a reduced probability of  
372 sub-median outcome, indicating resilience to injury. Increased TNF $\alpha$  is more usually associated  
373 with renal injury<sup>39</sup>, so this result is counterintuitive. The MAPK/ERK cascade impacts many  
374 regulatory pathways so it is reasonable to assume such intuition may oversimplify the effect of  
375 increased MEK1 abundance.

376           Several proteins in the age-moderated group are characterized as markers of protein  
377 regulation and proteasomal activity. These may indicate alterations within the proteostasis

378 network that increase susceptibility of donor grafts to subsequent injury and reduce capacity for  
379 recovery. CST3 is particularly noteworthy as, measured in serum, it is a known and effective  
380 general biomarker for kidney function and has previously been reported as having predictive  
381 power for outcomes in transplant recipients<sup>40-42</sup>. Our evidence indicates a further association  
382 between CST3 levels in the donor kidney tissue and outcome; moreover, that this effect is age  
383 dependent, starting around age 40. Interestingly, serum CST3 is relatively independent of age in  
384 children and young adults<sup>43</sup>, but numerous studies have demonstrated an increase in later years,  
385 as summarized by Edinga-Melenge *et al.*<sup>44</sup>. These two age related effects may be  
386 interconnected.

387 We found the age-modulated candidate PREB (Prolactin Regulatory Element Binding  
388 protein) biologically interesting for three reasons. Firstly, there is a well characterized  
389 relationships between kidney dysfunction (in the form of CKD), cardiovascular disease and  
390 prolactinemia<sup>45</sup>, with CKD patients being associated with elevated prolactinemia. Secondly, it is  
391 a regulator of glucose homeostasis in the liver and therefore a plausible key node for metabolic  
392 regulation in kidneys as well<sup>46</sup>, acknowledging the large emphasis in our pathway analysis on  
393 metabolic functions. Thirdly, it has a predicted<sup>47</sup> role in exit from the endoplasmic reticulum and  
394 the unfolded protein response, which has an association with CKD via NFκB -mediated  
395 inflammation<sup>48</sup>.

396 Another age-moderated protein, APOE stands out as having previously reported age-  
397 related associations with disease and organ dysfunction, in both cases including Alzheimer's  
398 Disease (AD). APOE plays an important role in lipid metabolism to regulate the growth and  
399 survival of mesangial cells and preserve kidney function<sup>49</sup>; not only is APOE protein a marker  
400 for outcome in transplant recipients<sup>50-52</sup>, there is also evidence for particular *APOE* genetic

401 alleles being associated with dysfunction risk <sup>53-55</sup>, possibly manifested directly by lipidomic  
402 differences between allelic profiles <sup>56</sup>. *APOE* alleles are further implicated in a host of disorders  
403 including age-moderated association with AD risk <sup>57</sup> (with the strongest effect manifesting  
404 around age 65 <sup>58</sup>), macular dysfunction, atherosclerosis and pulmonary scarring <sup>59,60</sup>. There is  
405 also evidence of allele risk association between disparate pathologies <sup>61</sup>. Interestingly, we have  
406 previously observed small (not statistically significant) increases in APOE due to ischemic  
407 reperfusion injuries <sup>62</sup> which may be explained by a recently description of the role of APOE in  
408 mediating senescence<sup>63</sup>. Such evidence suggests cause for further investigation of the *APOE*  
409 genotype with respect to donor kidneys of varying age. Indeed, the lipoprotein genetic  
410 background in general, with the close links between lipidome status and kidney function, may be  
411 associated with outcome. There is evidence of at least one similar allele dependent effects in  
412 another apolipoprotein (APOL1) associated with transplant outcome <sup>64</sup>.

413         Organ allocation algorithms impose a close link between donor and recipient age in the  
414 sample cohort, so these two variables cannot be conclusively separated. It is difficult to assess  
415 whether the age-moderated effects we observe represent a tendency for greater damage in older  
416 donors, or alternatively a greater ability to repair a given level of damage in younger recipients.  
417 Our analysis considers only chronological donor age, rather than a more nuanced representation  
418 of the epigenomic biological clock <sup>65</sup>, which is likely to account for some variation observed  
419 with respect to both donors and recipients.

420         Our list of outcome-associated candidates, controlling for the effect of donor age,  
421 including those for which we further report an additional age-moderated effect, is almost  
422 certainly not exhaustive. Practicalities of sample acquisition limited sampling of a range of  
423 outcomes for donors outside the 30-60 age range. In our dataset, as in the general population,

424 there is a trend towards lower eGFR (fewer good outcome events) at high donor age (Table 1). In  
425 the vast majority of proteins, the modelling suggests a plateauing effect at high donor age where  
426 the differences in outcome due to both protein and age are smaller. This effect may be an artifact  
427 of the distribution of sub-median outcomes. It is also immediately clear from our results that the  
428 strength of the donor age factor is enormous relative to any one protein effect; this must be borne  
429 in mind when interpreting model fits; at high donor age, the age effect is liable to dominate any  
430 prediction weighting and reduce the accuracy of estimated protein contribution. Both of these  
431 considerations should ideally be explored in a much larger cohort. High-throughput proteomics  
432 techniques continue to advance rapidly and larger cohort sizes are ever more feasible<sup>66</sup> but  
433 fundamental limitations on organ acquisition remain. Archiving at scale of clinical samples in  
434 bioresources such as the Quality in Organ Donation (QUOD) biobank to parallel advancements  
435 in big data analysis and interpretation platforms will improve the granularity in evidence-based  
436 decision making and accelerate the translation and application to clinical practice.

437 In our orthogonal validation of selected proteins, the expected protein abundance  
438 difference between extremes of the eGFR continuum across donor age distribution tails, and the  
439 plateauing/narrowing of outcome differences at high donor age were broadly consistent with the  
440 western blot results. In younger donors the predicted strengths of association were very  
441 consistent with the observed differences in protein abundance in selected samples by western  
442 blot. In older donors, the western blot results were still broadly consistent (relative to the effect at  
443 in younger donors) in terms of a much weaker association, but the differences in both VTN and  
444 PREB were larger than might be expected by examination of the prediction curves (Figure 6B,  
445 upper panel) and the prediction for CST3 of a small reversal of the effect is both unexpected and  
446 biologically counterintuitive. A better explanation is that the model fit is impacted by the

447 (unavoidable) lack of outcome diversity at high age ranges, with the CST3 ‘flip’ in terms of  
448 protein effect likely being artefactual.

449 In this work, we profiled the proteome of pre-implantation biopsies selected from donor  
450 kidneys on the basis of paired 12-month graft function, limiting potential surgical and post-  
451 transplant biases on transplant outcomes. Using machine learning and multivariate adaptive  
452 spline regression models, we identified 136 candidate proteins associated with sub-median  
453 outcomes, suggesting molecular signatures which may refine models of graft dysfunction based  
454 on clinical and demographic factors alone. We also found that most (124) of these candidates  
455 furthermore show different strengths of association dependent on the age of the donor. The  
456 biological themes of the identified candidates reinforce known immuno-metabolic mechanisms  
457 of kidney injury but raise interesting possibilities for further work, especially with regard to  
458 donor genetic background. Furthermore, our results strongly suggest that for any studies of  
459 subclinical molecular indicators with regard to kidney transplant outcomes, the possibility of  
460 donor age-moderated weighting should be considered as a matter of course.

461

## 462 **SUPPLEMENTARY MATERIALS**

463 Supplementary Methods

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

465 Supplementary Table 2: Summary of results for all candidate proteins

466 Supplementary Figure 1: Protein Quantification Missingness

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

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

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

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

479

## 480 **DATA AND MATERIALS AVAILABILITY**

481 The mass spectrometry proteomics data have been deposited to the ProteomeXchange  
482 Consortium via the PRIDE <sup>67</sup> partner repository with the dataset identifier PXD033428.

483

## 484 **DISCLOSURE AND FUNDING**

485 This study was funded by NHS Blood and Transplant WP15-07 awarded to MK & RJP. SF was  
486 supported by Kidney Research UK, grant reference KS\_RP\_002\_20210111 awarded to MK.  
487 PDC was supported by a Chinese Academy of Medical Sciences 2018-I2M-2-002 awarded to  
488 BMK. Authors declare that they have no competing interests.

## 489 **Author contributions:**

490 Conceptualization: MK

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

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

493 Visualization: PDC

494 Funding acquisition: BMK, RJP, MK

495 Project administration: MK

496 Supervision: RF, AS, MK

497 Writing – original draft: PDC, ES, MK

498 Writing – review & editing: All authors

499

## 500 REFERENCES

- 501 1. Callaghan CJ, Mumford L, Pankhurst L, Baker RJ, Bradley JA, Watson CJE. Early  
502 Outcomes of the New UK Deceased Donor Kidney Fast-Track Offering Scheme.  
503 *Transplantation*. 2017;101(12):2888-2897. doi:10.1097/TP.0000000000001860
- 504 2. Summers DM, Johnson RJ, Hudson AJ, et al. Standardized deceased donor kidney donation  
505 rates in the UK reveal marked regional variation and highlight the potential for increasing  
506 kidney donation: a prospective cohort study†. *Br J Anaesth*. 2014;113(1):83-90.  
507 doi:10.1093/bja/aet473
- 508 3. Denic A, Glasscock RJ, Rule AD. The Kidney in Normal Aging: A Comparison with Chronic  
509 Kidney Disease. *Clin J Am Soc Nephrol*. 2022;17(1):137-139. doi:10.2215/CJN.10580821
- 510 4. O’Hare AM, Choi AI, Bertenthal D, et al. Age affects outcomes in chronic kidney disease. *J*  
511 *Am Soc Nephrol*. 2007;18(10):2758-2765. doi:10.1681/ASN.2007040422
- 512 5. Ishani A, Xue JL, Himmelfarb J, et al. Acute Kidney Injury Increases Risk of ESRD among  
513 Elderly. *J Am Soc Nephrol*. 2009;20(1):223-228. doi:10.1681/ASN.2007080837
- 514 6. Summers DM, Johnson RJ, Hudson A, Collett D, Watson CJ, Bradley JA. Effect of donor  
515 age and cold storage time on outcome in recipients of kidneys donated after circulatory death  
516 in the UK: a cohort study. *Lancet*. 2013;381(9868):727-734. doi:10.1016/S0140-  
517 6736(12)61685-7
- 518 7. Rao PS, Schaubel DE, Guidinger MK, et al. A comprehensive risk quantification score for  
519 deceased donor kidneys: the kidney donor risk index. *Transplantation*. 2009;88(2):231-236.  
520 doi:10.1097/TP.0b013e3181ac620b
- 521 8. Watson CJE, Johnson RJ, Birch R, Collett D, Bradley JA. A Simplified Donor Risk Index  
522 for Predicting Outcome After Deceased Donor Kidney Transplantation. *Transplantation*.  
523 2012;93(3):314-318. doi:10.1097/TP.0b013e31823f14d4

- 524 9. Neuberger J, Callaghan C. Organ utilization - the next hurdle in transplantation? *Transpl Int*.  
525 2020;33(12):1597-1609. doi:10.1111/tri.13744
- 526 10. Clayton PA, Dansie K, Sypek MP, et al. External validation of the US and UK kidney donor  
527 risk indices for deceased donor kidney transplant survival in the Australian and New Zealand  
528 population. *Nephrol Dial Transplant*. 2019;34(12):2127-2131. doi:10.1093/NDT/GFZ090
- 529 11. Remuzzi G, Cravedi P, Perna A, et al. Long-Term Outcome of Renal Transplantation from  
530 Older Donors. *N Engl J Med*. 2006;354(4):343-352. doi:10.1056/NEJMoa052891
- 531 12. Husain SA, Chiles MC, Lee S, et al. Characteristics and Performance of Unilateral Kidney  
532 Transplants from Deceased Donors. *Clin J Am Soc Nephrol*. 2018;13(1):118-127.  
533 doi:10.2215/CJN.06550617
- 534 13. Liu C, Hall IE, Mansour S, Thiessen Philbrook HR, Jia Y, Parikh CR. Association of  
535 Deceased Donor Acute Kidney Injury With Recipient Graft Survival. *JAMA Network Open*.  
536 2020;3(1):e1918634. doi:10.1001/jamanetworkopen.2019.18634
- 537 14. Yu K, Husain SA, King K, Stevens JS, Parikh CR, Mohan S. Kidney nonprocurement in  
538 deceased donors with acute kidney injury. *Clin Transplant*. Published online August 4,  
539 2022:e14788. doi:10.1111/ctr.14788
- 540 15. Hall IE, Akalin E, Bromberg JS, et al. Deceased-donor acute kidney injury is not associated  
541 with kidney allograft failure. *Kidney Int*. 2019;95(1):199-209.  
542 doi:10.1016/j.kint.2018.08.047
- 543 16. Mansour SG, Khoury N, Kodali R, et al. Clinically adjudicated deceased donor acute kidney  
544 injury and graft outcomes. *PLoS One*. 2022;17(3):e0264329.  
545 doi:10.1371/journal.pone.0264329
- 546 17. von Moos S, Akalin E, Mas V, Mueller TF. Assessment of Organ Quality in Kidney  
547 Transplantation by Molecular Analysis and Why It May Not Have Been Achieved, Yet.  
548 *Front Immunol*. 2020;11:833. doi:10.3389/fimmu.2020.00833
- 549 18. Gillet LC, Navarro P, Tate S, et al. Targeted data extraction of the MS/MS spectra generated  
550 by data-independent acquisition: a new concept for consistent and accurate proteome  
551 analysis. *Mol Cell Proteomics*. 2012;11(6):O111.016717. doi:10.1074/mcp.O111.016717
- 552 19. Raynaud M, Aubert O, Divard G, et al. Dynamic prediction of renal survival among deeply  
553 phenotyped kidney transplant recipients using artificial intelligence: an observational,  
554 international, multicohort study. *The Lancet Digital Health*. 2021;3(12):e795-e805.  
555 doi:10.1016/S2589-7500(21)00209-0
- 556 20. Geyer PE, Holdt LM, Teupser D, Mann M. Revisiting biomarker discovery by plasma  
557 proteomics. *Mol Syst Biol*. 2017;13(9):942. doi:10.15252/msb.20156297
- 558 21. Connor KL, O'Sullivan ED, Marson LP, Wigmore SJ, Harrison EM. The Future Role of  
559 Machine Learning in Clinical Transplantation. *Transplantation*. 2021;105(4):723-735.  
560 doi:10.1097/TP.0000000000003424



- 561 22. UK Renal Registry. *UK Renal Registry 24th Annual Report - Data to 31/12/2020.*; 2022.  
562 <https://renal.org/audit-research/annual-report>
- 563 23. Fokkema M. Fitting prediction rule ensembles with R package pre. *J Stat Softw.*  
564 2020;92(1):1-30. doi:10.18637/jss.v092.i12
- 565 24. Friedman JH. Multivariate Adaptive Regression Splines. *aos.* 1991;19(1):1-67.  
566 doi:10.1214/aos/1176347963
- 567 25. Milborrow. Derived from mda:mars by T. Hastie and R. Tibshirani. S. *Earth: Multivariate*  
568 *Adaptive Regression Splines.*; 2011. <http://CRAN.R-project.org/package=earth>
- 569 26. Brier GW. Verification Of Forecasts Expressed In Terms Of Probability. *Mon Weather Rev.*  
570 1950;78(1):1-3. doi:10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2
- 571 27. Li B, Cairns JA, Robb ML, et al. Predicting patient survival after deceased donor kidney  
572 transplantation using flexible parametric modelling. *BMC Nephrol.* 2016;17(1):51.  
573 doi:10.1186/s12882-016-0264-0
- 574 28. Gill J, Rose C, Lesage J, Joffres Y, Gill J, O'Connor K. Use and Outcomes of Kidneys from  
575 Donation after Circulatory Death Donors in the United States. *J Am Soc Nephrol.*  
576 2017;28(12):3647-3657. doi:10.1681/asn.2017030238
- 577 29. Robinson C. *NHSBT Policy POL186/10. Kidney Transplantation: Deceased Donor*  
578 *Allocation.*; 2019. [https://nhsbt.dbe.blob.core.windows.net/umbraco-assets-](https://nhsbt.dbe.blob.core.windows.net/umbraco-assets-corp/16915/kidney-allocation-policy-pol186.pdf)  
579 [corp/16915/kidney-allocation-policy-pol186.pdf](https://nhsbt.dbe.blob.core.windows.net/umbraco-assets-corp/16915/kidney-allocation-policy-pol186.pdf)
- 580 30. Vaughan RH, Kresse JC, Farmer LK, et al. Cytoskeletal protein degradation in brain death  
581 donor kidneys associates with adverse post-transplant outcomes. *Am J Transplant.* Published  
582 online December 8, 2021. doi:10.1111/ajt.16912
- 583 31. Kosmoliaptsis V, Salji M, Bardsley V, et al. Baseline donor chronic renal injury confers the  
584 same transplant survival disadvantage for DCD and DBD kidneys. *Am J Transplant.*  
585 2015;15(3):754-763. doi:10.1111/ajt.13009
- 586 32. De Lorenzi V, Sarra Ferraris GM, Madsen JB, Lupia M, Andreasen PA, Sidenius N.  
587 Urokinase links plasminogen activation and cell adhesion by cleavage of the RGD motif in  
588 vitronectin. *EMBO Rep.* 2016;17(7):982-998. doi:10.15252/embr.201541681
- 589 33. Carreras-Planella L, Cucchiari D, Cañas L, et al. Urinary vitronectin identifies patients with  
590 high levels of fibrosis in kidney grafts. *J Nephrol.* 2021;34(3):861-874. doi:10.1007/s40620-  
591 020-00886-y
- 592 34. Diepenhorst GMP, van Gulik TM, Hack CE. Complement-mediated ischemia-reperfusion  
593 injury: lessons learned from animal and clinical studies. *Ann Surg.* 2009;249(6):889-899.  
594 doi:10.1097/SLA.0b013e3181a38f45
- 595 35. Arias-Cabrales CE, Riera M, Pérez-Sáez MJ, et al. Activation of final complement  
596 components after kidney transplantation as a marker of delayed graft function severity. *Clin*  
597 *Kidney J.* 2021;14(4):1190-1196. doi:10.1093/ckj/sfaa147

- 598 36. Almitairi JOM, Venkatraman Girija U, Furze CM, et al. Structure of the C1r-C1s interaction  
599 of the C1 complex of complement activation. *Proc Natl Acad Sci U S A*. 2018;115(4):768-  
600 773. doi:10.1073/pnas.1718709115
- 601 37. Awdishu L, Tsunoda S, Pearlman M, et al. Identification of Maltase Glucoamylase as a  
602 Biomarker of Acute Kidney Injury in Patients with Cirrhosis. *Crit Care Res Pract*.  
603 2019;2019:5912804. doi:10.1155/2019/5912804
- 604 38. Sabio G, Davis RJ. TNF and MAP kinase signalling pathways. *Semin Immunol*.  
605 2014;26(3):237-245. doi:10.1016/j.smim.2014.02.009
- 606 39. Hernandez T, Mayadas TN. Immunoregulatory role of TNFalpha in inflammatory kidney  
607 diseases. *Kidney Int*. 2009;76(3):262-276. doi:10.1038/ki.2009.142
- 608 40. Coll E, Botey A, Alvarez L, et al. Serum cystatin C as a new marker for noninvasive  
609 estimation of glomerular filtration rate and as a marker for early renal impairment. *Am J*  
610 *Kidney Dis*. 2000;36(1):29-34. doi:10.1053/ajkd.2000.8237
- 611 41. Christensson A, Ekberg J, Grubb A, Ekberg H, Lindström V, Lilja H. Serum cystatin C is a  
612 more sensitive and more accurate marker of glomerular filtration rate than enzymatic  
613 measurements of creatinine in renal transplantation. *Nephron Physiol*. 2003;94(2):19-27.  
614 doi:10.1159/000071287
- 615 42. Rodrigo E, Ruiz JC, Fernández-Fresnedo G, et al. Cystatin C and albuminuria as predictors  
616 of long-term allograft outcomes in kidney transplant recipients. *Clin Transplant*.  
617 2013;27(2):E177-83. doi:10.1111/ctr.12082
- 618 43. Bökenkamp A, Domanetzki M, Zinck R, Schumann G, Byrd D, Brodehl J. Cystatin C--a  
619 new marker of glomerular filtration rate in children independent of age and height.  
620 *Pediatrics*. 1998;101(5):875-881. doi:10.1542/peds.101.5.875
- 621 44. Edinga-Melenge BE, Yakam AT, Nansseu JR, et al. Reference intervals for serum cystatin C  
622 and serum creatinine in an adult sub-Saharan African population. *BMC Clin Pathol*.  
623 2019;19:4. doi:10.1186/s12907-019-0086-7
- 624 45. Dourado M, Cavalcanti F, Vilar L, Cantilino A. Relationship between Prolactin, Chronic  
625 Kidney Disease, and Cardiovascular Risk. *Int J Endocrinol*. 2020;2020:9524839.  
626 doi:10.1155/2020/9524839
- 627 46. Park JM, Kim MY, Kim TH, Min DK, Yang GE, Ahn YH. Prolactin regulatory element-  
628 binding (PREB) protein regulates hepatic glucose homeostasis. *Biochim Biophys Acta Mol*  
629 *Basis Dis*. 2018;1864(6 Pt A):2097-2107. doi:10.1016/j.bbadis.2018.03.024
- 630 47. Gaudet P, Livstone MS, Lewis SE, Thomas PD. Phylogenetic-based propagation of  
631 functional annotations within the Gene Ontology consortium. *Brief Bioinform*.  
632 2011;12(5):449-462. doi:10.1093/bib/bbr042
- 633 48. Mohammed-Ali Z, Cruz GL, Dickhout JG. Crosstalk between the unfolded protein response  
634 and NF-κB-mediated inflammation in the progression of chronic kidney disease. *J Immunol*  
635 *Res*. 2015;2015:428508. doi:10.1155/2015/428508

- 636 49. Chen G, Paka L, Kako Y, Singhal P, Duan W, Pillarisetti S. A Protective Role for Kidney  
637 Apolipoprotein E: REGULATION OF MESANGIAL CELL PROLIFERATION AND  
638 MATRIX EXPANSION. *J Biol Chem*. 2001;276(52):49142-49147.  
639 doi:10.1074/JBC.M104879200
- 640 50. Kahraman S, Kiykim AA, Altun B, et al. Apolipoprotein E gene polymorphism in renal  
641 transplant recipients: effects on lipid metabolism, atherosclerosis and allograft function. *Clin*  
642 *Transplant*. 2004;18(3):288-294. doi:10.1111/j.1399-0012.2004.00162.x
- 643 51. Hernández D, Salido E, Linares J, et al. Role of apolipoprotein E epsilon 4 allele on chronic  
644 allograft nephropathy after renal transplantation. *Transplant Proc*. 2004;36(10):2982-2984.  
645 doi:10.1016/j.transproceed.2004.10.038
- 646 52. Cofán F, Cofan M, Rosich E, et al. Effect of apolipoprotein E polymorphism on renal  
647 transplantation. *Transplant Proc*. 2007;39(7):2217-2218.  
648 doi:10.1016/j.transproceed.2007.06.011
- 649 53. Czaplińska M, Ćwiklińska A, Sakowicz-Burkiewicz M, et al. Apolipoprotein e gene  
650 polymorphism and renal function are associated with apolipoprotein e concentration in  
651 patients with chronic kidney disease. *Lipids Health Dis*. 2019;18(1):1-9.  
652 doi:10.1186/s12944-019-1003-x
- 653 54. Hsu CC, Kao WHL, Coresh J, et al. Apolipoprotein E and progression of chronic kidney  
654 disease. *J Am Med Assoc*. 2005;293(23):2892-2899. doi:10.1001/jama.293.23.2892
- 655 55. Saito T, Ishigaki Y, Oikawa S, Yamamoto TT. Etiological significance of apolipoprotein E  
656 mutations in lipoprotein glomerulopathy. *Trends Cardiovasc Med*. 2002;12(2):67-70.  
657 doi:10.1016/S1050-1738(01)00148-7
- 658 56. Sienski G, Narayan P, Bonner JM, et al. APOE4 disrupts intracellular lipid homeostasis in  
659 human iPSC-derived glia. *Sci Transl Med*. 2021;13(583):eaaz4564.  
660 doi:10.1126/scitranslmed.aaz4564
- 661 57. Liu CC, Kanekiyo T, Xu H, Bu G. Apolipoprotein E and Alzheimer disease: risk,  
662 mechanisms, and therapy. *Nat Rev Neurol*. 2013;9(2):106-118.  
663 doi:10.1038/nrneurol.2012.263
- 664 58. Saddiki H, Fayosse A, Cognat E, et al. Age and the association between apolipoprotein E  
665 genotype and Alzheimer disease: A cerebrospinal fluid biomarker-based case-control study.  
666 *PLoS Med*. 2020;17(8):e1003289. doi:10.1371/journal.pmed.1003289
- 667 59. Mahley RW, Rall SC. Apolipoprotein E: far more than a lipid transport protein. *Annu Rev*  
668 *Genomics Hum Genet*. 2000;1:507-537. doi:10.1146/annurev.genom.1.1.507
- 669 60. Yao X, Gordon EM, Figueroa DM, Barochia AV, Levine SJ. Emerging Roles of  
670 Apolipoprotein E and Apolipoprotein A-I in the Pathogenesis and Treatment of Lung  
671 Disease. *Am J Respir Cell Mol Biol*. 2016;55(2):159-169. doi:10.1165/rcmb.2016-0060TR

- 672 61. Miwa K, Tanaka M, Okazaki S, et al. Chronic kidney disease is associated with dementia  
673 independent of cerebral small-vessel disease. *Neurology*. 2014;82(12):1051-1057.  
674 doi:10.1212/WNL.0000000000000251
- 675 62. Huang H, van Dullemen LFA, Akhtar MZ, et al. Proteo-metabolomics reveals compensation  
676 between ischemic and non-injured contralateral kidneys after reperfusion. *Sci Rep*.  
677 2018;8(1):8539. doi:10.1038/s41598-018-26804-8
- 678 63. Zhao H, Ji Q, Wu Z, et al. Destabilizing heterochromatin by APOE mediates senescence.  
679 *Nature Aging*. 2022;2(4):303-316. doi:10.1038/s43587-022-00186-z
- 680 64. Julian BA, Gaston RS, Brown WM, et al. Effect of Replacing Race with Apolipoprotein L1  
681 Genotype in Calculation of Kidney Donor Risk Index. *Am J Transplant*. 2017;17(6):1540.  
682 doi:10.1111/AJT.14113
- 683 65. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of  
684 ageing. *Nat Rev Genet*. 2018;19(6):371-384. doi:10.1038/s41576-018-0004-3
- 685 66. Meier F, Brunner AD, Frank M, et al. diaPASEF: parallel accumulation–serial fragmentation  
686 combined with data-independent acquisition. *Nat Methods*. 2020;17(12):1229-1236.  
687 doi:10.1038/s41592-020-00998-0
- 688 67. Perez-Riverol Y, Csordas A, Bai J, et al. The PRIDE database and related tools and  
689 resources in 2019: improving support for quantification data. *Nucleic Acids Res*.  
690 2019;47(D1):D442-D450. doi:10.1093/nar/gky1106

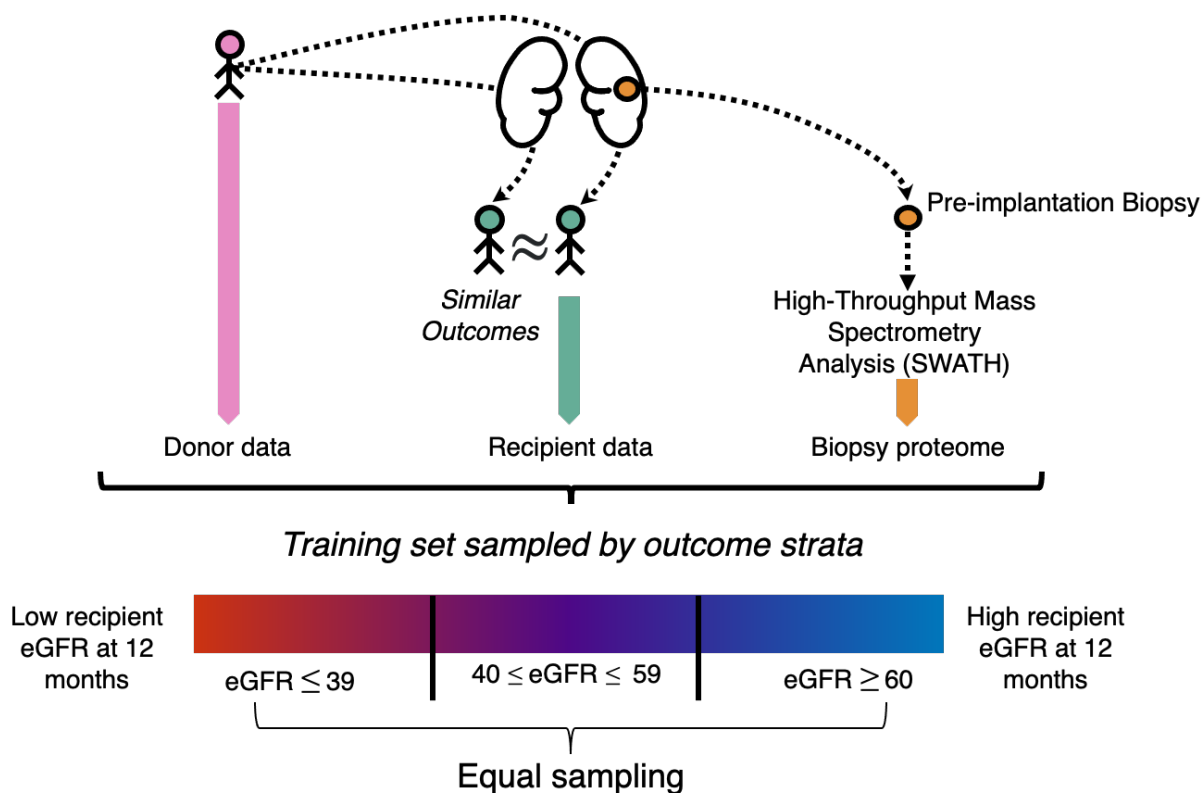
691

## 692 **ACKNOWLEDGMENTS**

693 We thank the UK QUOD Consortium and NHS Blood and Transplant UK Registry; in particular  
694 Sheba Ziyenge, Lewis Simmonds and Dr Sarah Cross for the clinical samples and metadata  
695 analyzed in this study. We thank Dr Sergei Maslau and Mr Tomas Surik for their support on the  
696 QUOD sample selection.

697 We thank members of the Discovery Proteomics Facility within the TDI Mass Spectrometry  
698 Laboratory for expert help with mass spectrometry analysis.

699 FIGURES AND TABLES

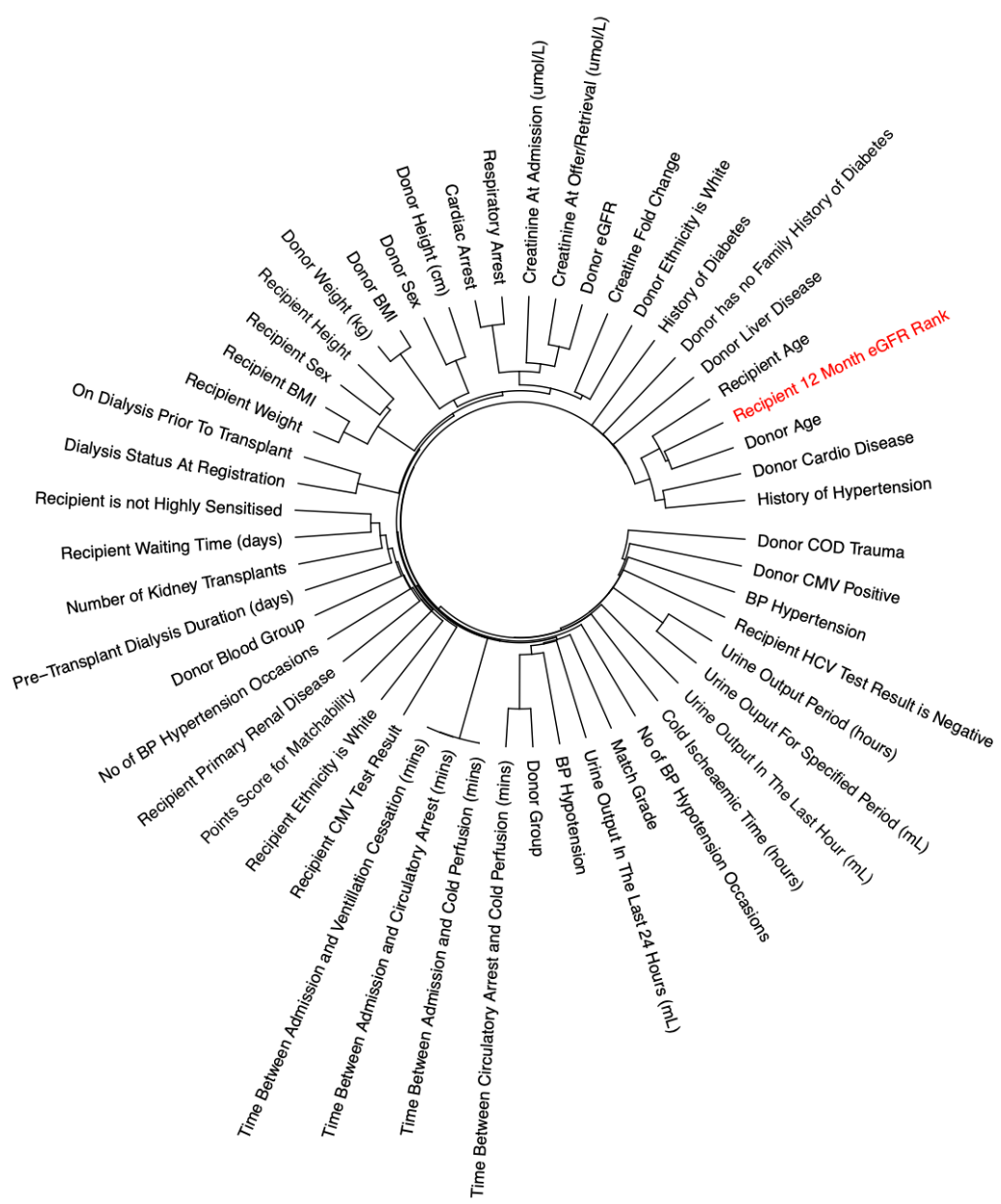


700 **Figure 1: Experimental design to discover donor kidney proteome associations with transplant outcome**  
701 One kidney from each donor pair was biopsied at the back table. Donor kidney samples were selected randomly  
702 from pairs where both recipients had similar outcomes. The biopsy samples were subjected to proteomic analysis to  
703 yield a snapshot of the organ proteome before transplantation. We analyzed donor characteristics and clinical  
704 variables, recipient characteristics and protein abundances in a combined model against outcome.

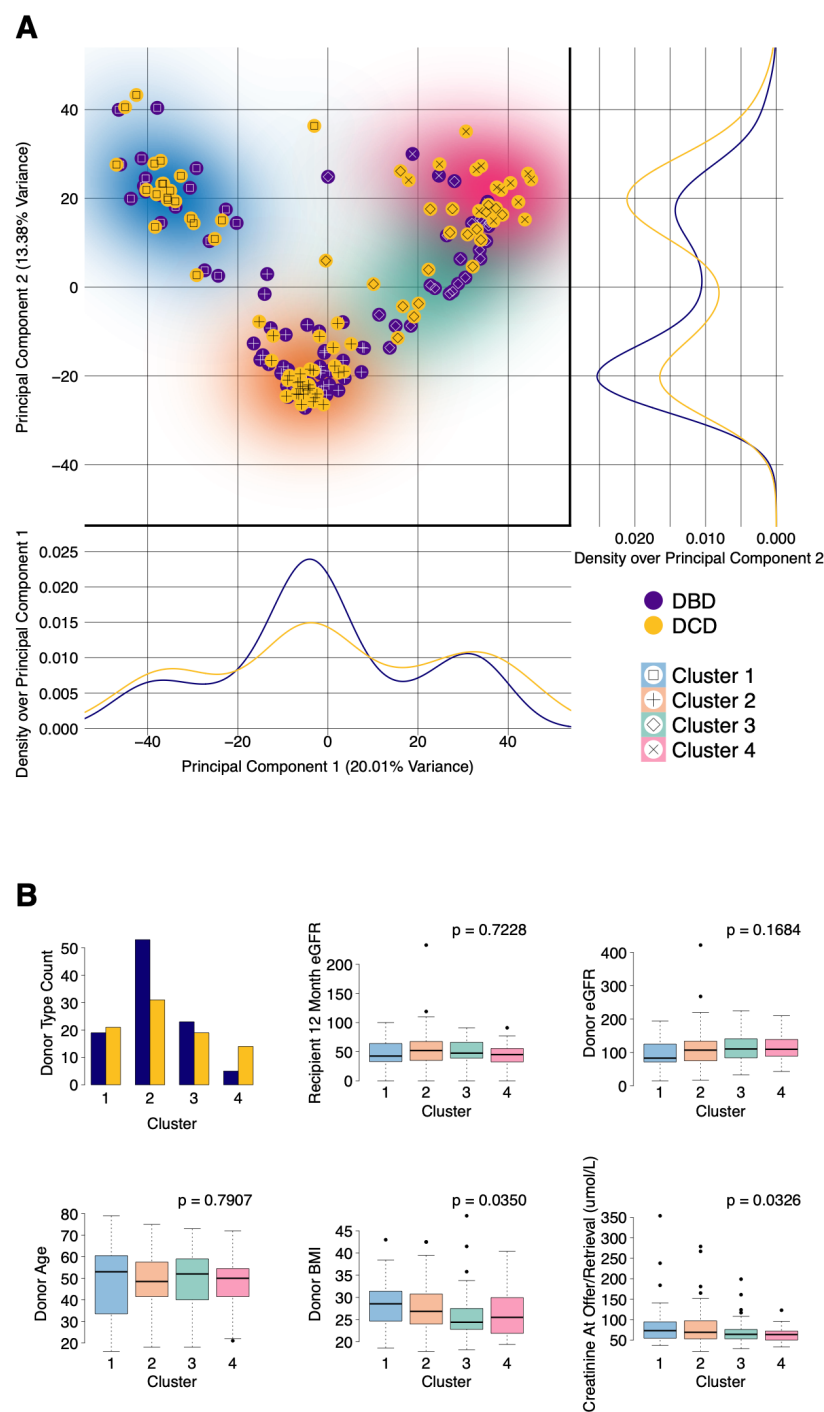
Donor Type	DBD			DCD		
	1 <sup>st</sup> : Suboptimal ( <i>eGFR</i> ≤39)	2 <sup>nd</sup> : Intermediate (40≤ <i>eGFR</i> ≤59)	3 <sup>rd</sup> : Good ( <i>eGFR</i> ≥60)	1 <sup>st</sup> : Suboptimal ( <i>eGFR</i> ≤39)	2 <sup>nd</sup> : Intermediate (40≤ <i>eGFR</i> ≤59)	3 <sup>rd</sup> : Good ( <i>eGFR</i> ≥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 per 1.73 m <sup>2</sup> )						
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

705 **Table 1: Donor and recipient clinical and demographic variables**

706 Donor kidney associated metadata. Samples are subdivided by donor type and by final assigned outcome tertile.  
 707 Numerical variables are given ± standard deviation. Categorical variables are given alongside percentage of total  
 708 cohort



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



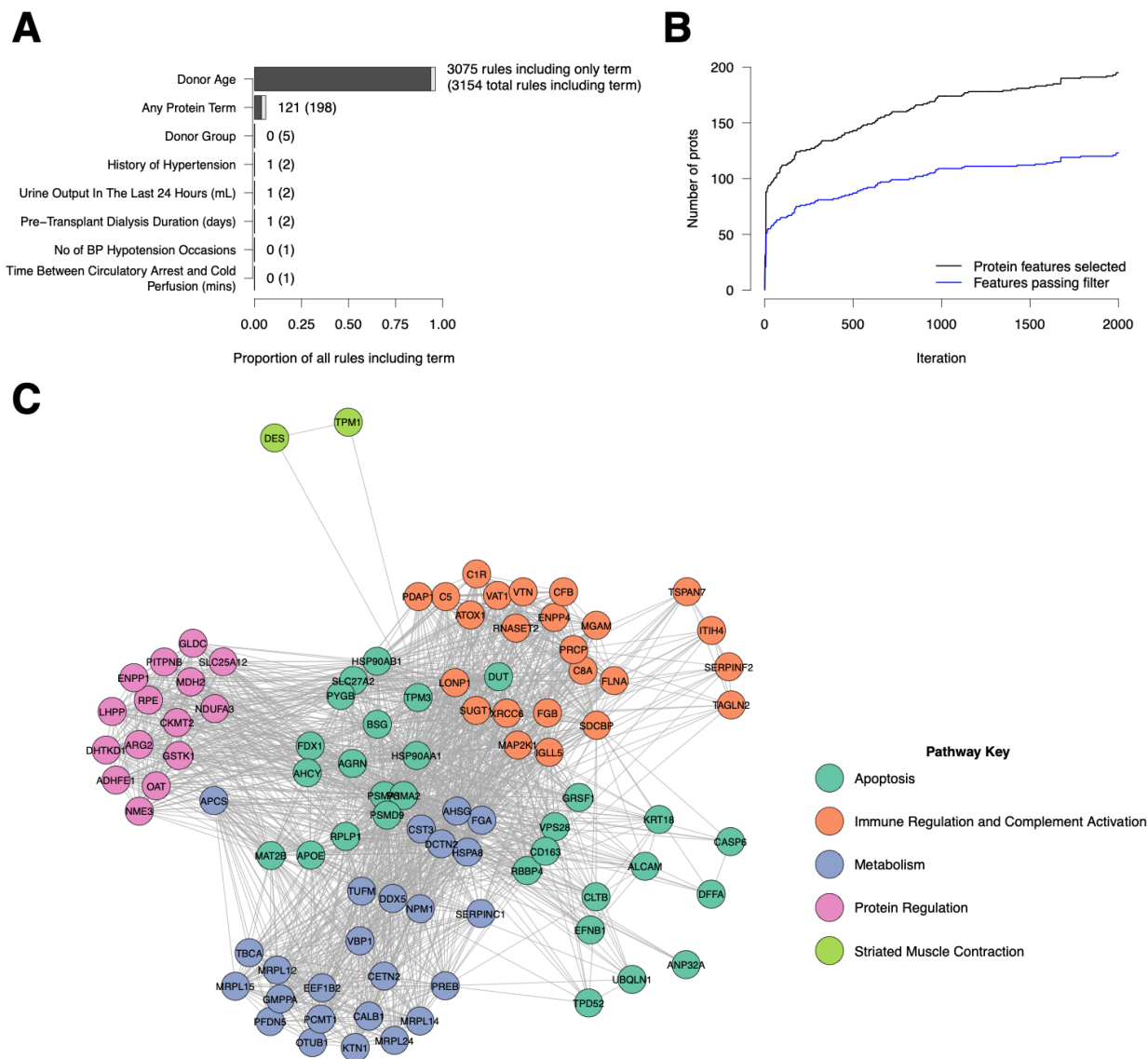
713 **Figure 3: Unbiased analysis of pretransplant kidney proteomes and cluster associations**

714 A: Unbiased analysis of proteomic data by k-means clustering. Sample separation by Principal Component Analysis.

715 Top Left: Samples were assigned to four clusters by k-means. Bottom & Right: There was a difference in the  
 716 distribution of DBD and DCD donors across clusters, with the DBD donors being more heavily concentrated in  
 717 Cluster 2 (+ symbol; orange shading), and DCD in Cluster 4 (x symbol; pink shading)

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





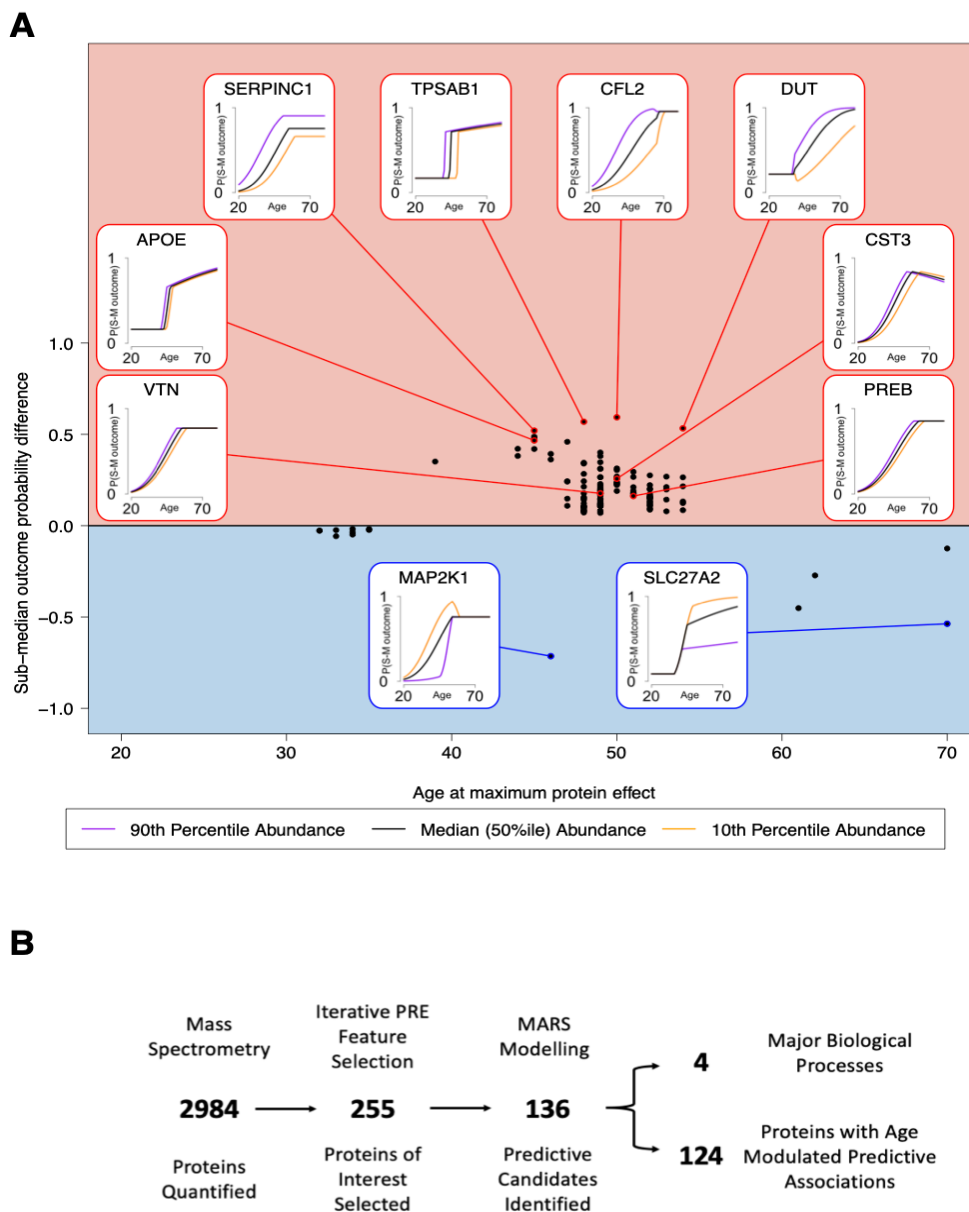
722 **Figure 4: Age and combined age:protein related associations link to construction of age-modulated immune**  
 723 **metabolic biological networks**

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

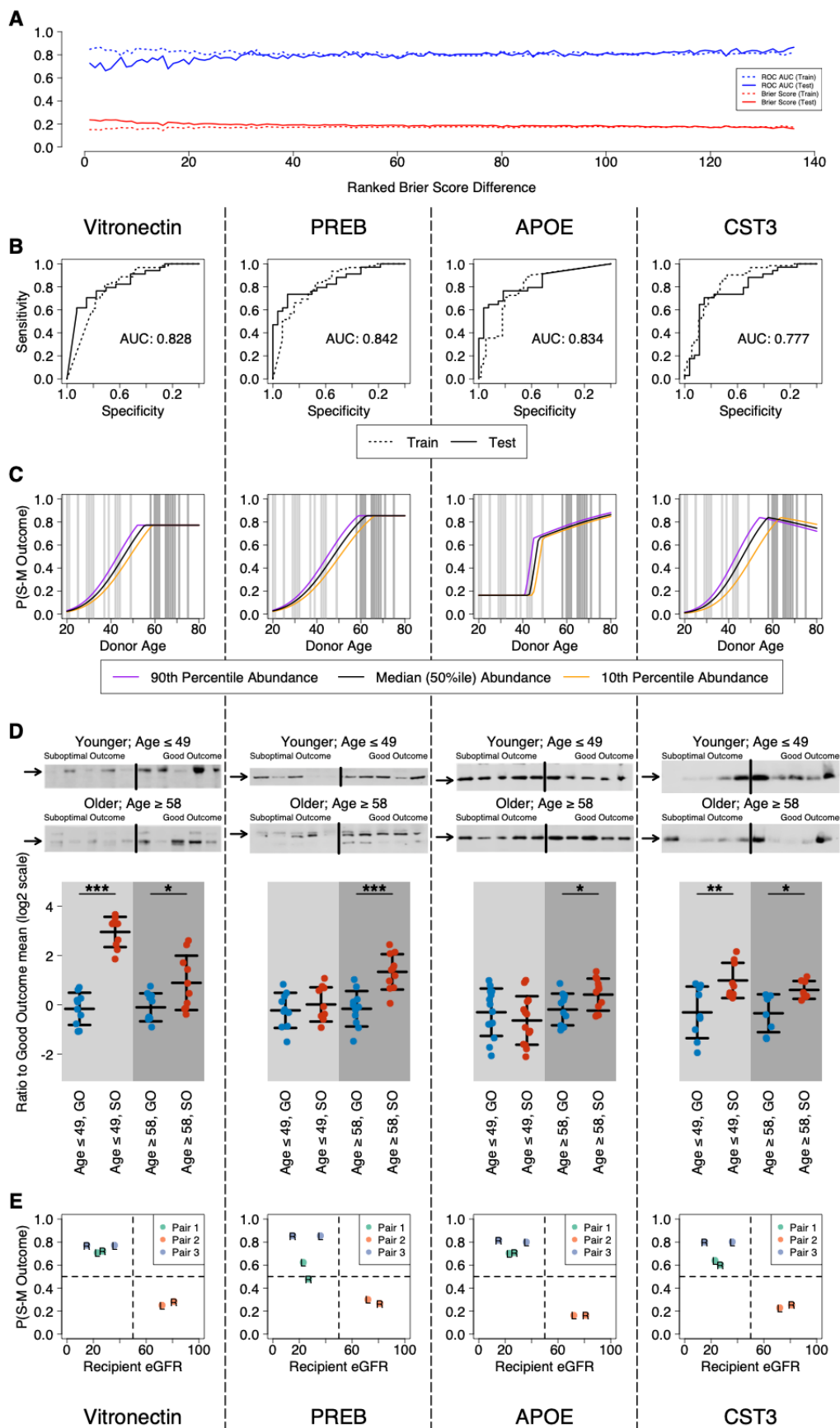
727 B: Cumulative protein features identified at each iteration. Black line: all features identified by feature selection  
 728 approach. Blue line: features passing the secondary filter for predictive power and accuracy.

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

731



732 **Figure 5: Modelled associations between proteins and kidney transplant outcome change with donor age**  
 733 A: Ages at which the predicted probability of sub-median outcome is most different between the 10<sup>th</sup> percentile and  
 734 90<sup>th</sup> percentile of protein abundance.  $x$  axis: age at which difference is greatest (i.e. when protein has greatest  
 735 effect).  $y$  axis: greatest difference.  
 736 Proteins above  $x=0$  are modelled as having a more negative association with outcome when the protein abundance is  
 737 high, at that donor age. Proteins below  $x=0$  are modelled to have a more positive association with outcome when  
 738 protein abundance is high, at that donor age.  
 739 Proteins with absolute net difference  $> 0.5$  are labelled, as well as the selected proteins VTN, PREB, APOE and  
 740 CST3.  
 741 The inset graphs indicate how the prediction of sub-median outcome (“P(S-M outcome)”;  $y$  axes) changes with  
 742 donor age ( $x$  axes) for labelled proteins. Black trace: prediction at median protein abundance. Purple trace:  
 743 prediction at 90<sup>th</sup> percentile of protein abundance. Orange trace: prediction at 10<sup>th</sup> percentile of protein abundance.  
 744 The corresponding point on the main figure thus indicates the age at which the difference between orange and purple  
 745 lines is greatest.  
 746 B: Summary of feature selection and modelling analysis



748 **Figure 6: Orthogonal validation confirms age-modulated immuno-metabolic proteins predict 12-month**  
749 **transplant outcomes**  
750 A: Validation of models in test dataset. Models are plotted in order of decreasing Brier score (mean squared  
751 prediction error) *difference* between test and train data along the x axis. The lower two traces indicate the Brier score  
752 in train (purple) and test (green) data. The upper two traces indicate the AUC from the corresponding ROC analyses  
753 in train (orange) and test (blue) data.  
754 B-E: Validation of four selected proteins. Left-Right: VTN, PREB, APOE, CST3.  
755 B: Final ROC curves and AUC values for models trained on each protein (and donor age) against test data. The  
756 dotted line indicates the original performance against training data.  
757 C: Change in the prediction of sub-median outcome (“P(S-M Outcome)”; y axes) with donor age (x axes) for each  
758 protein. Black trace: prediction at median protein abundance. Purple trace: prediction at 90<sup>th</sup> percentile of protein  
759 abundance. Orange trace: prediction at 10<sup>th</sup> percentile of protein abundance. (These are the same as the inset graphs  
760 in Figure 5). The light grey and dark grey lines, respectively, indicate the corresponding sampled ages for the  
761 western blots below.  
762 D: Western blots comparing younger (age  $\leq 49$ ) and older (age  $\geq 58$ ) donors between Good Outcome (GO;  
763 eGFR  $\geq 60$ ) and Suboptimal Outcome (SO; eGFR  $\leq 40$ ) outcome tertiles. Top row: representative western blots  
764 (n=5 per group) from comparison of younger donors. Middle row: representative western blots (n=5 per group) from  
765 comparison of older donors. Bottom row: result values for all quantified samples relative to the GO mean. Error bars  
766 indicate  $\pm 1$  standard deviation; the central wider bar indicates mean. Significance stars indicate t-test comparison p-  
767 values (\*\*\*:  $< 0.001$ , \*:  $< 0.05$ ).  
768 E: Predicted outcome for six paired Left (L) and Right (R) kidneys from three donors. x axes: recipient eGFR at 12  
769 months (i.e. actual outcome). y axes: predicted probability of sub-median outcome (“P(S-M Outcome)”) using  
770 models trained on each protein with donor age. Vertical dotted line indicated median outcome (eGFR = 50).  
771 Horizontal dotted line indicates P(Sub-Median Outcome) = 0.5