

1 **Title**

2 Multiomic analyses uncover immunological signatures in kidney transplantation

3

4 **Author list**

5 Claire Tinel<sup>1,2,3</sup>, Alexis Varin<sup>2</sup>, Dany Anglicheau<sup>4,5</sup>, Jasper Callemeyn<sup>1,6</sup>, Jetty De Loor<sup>1</sup>, Wilfried

6 Gwinner<sup>7</sup>, Pierre Marquet<sup>8</sup>, Marion Rabant<sup>4,9</sup>, Virginia Sauvaget<sup>4</sup>, Elisabet Van Loon<sup>1,6</sup>,

7 Baptiste Lamarthée<sup>1,2,Φ,\*</sup>, Maarten Naesens<sup>1,6,Φ</sup>

8

9 **Affiliations**

10 <sup>1</sup>Department of Microbiology, Immunology and Transplantation, Nephrology and Kidney  
11 Transplantation Research Group, KU Leuven, Leuven, Belgium

12 <sup>2</sup>University of Franche-Comté, Inserm UMR Right, Établissement Français du Sang,  
13 Besançon, France

14 <sup>3</sup>Département de Néphrologie et Kidney Transplantation, Dijon University Hospital, University  
15 of Bourgogne, Dijon, France

16 <sup>4</sup>Département de Néphrologie et Kidney Transplantation, Necker Hospital, Paris, France

17 <sup>5</sup>University Paris Cité, Inserm U1151, Necker Enfants-Malades Institute, Paris, France

18 <sup>6</sup>Department of Nephrology and Kidney Transplantation, University Hospital Leuven, Leuven,  
19 Belgium

20 <sup>7</sup>Department of Nephrology, Hannover Medical School, Hannover, Germany

21 <sup>8</sup>Department of Pharmacology and Transplantation, University of Limoges, Inserm U1248,  
22 Limoges University Hospital, Limoges, France

23 <sup>9</sup>Département de Pathologie, Necker Hospital, Paris, France

24

25

26 <sup>Φ</sup>These authors contributed equally

27

28 **NOTE:** This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.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

29 Baptiste Lamarthée, PhD  
30 Phone: +33 81 61 56 15  
31 Rue du Docteur Girod, 25000 Besançon, France  
32 Email: [baptiste.lamarthee@inserm.fr](mailto:baptiste.lamarthee@inserm.fr)  
33  
34

35 **Abstract**

36 Identifying biomarkers in kidney transplant patients is essential for early detection of rejection,  
37 personalized treatment and improved overall outcomes. It improves our ability to monitor the  
38 health of the transplanted organ and tailor interventions to the specific needs of each patient.  
39 Here we compiled a multicenter, multiomic dataset of the kidney transplant landscape. Using  
40 multi-omics factor analysis (MOFA), we sought to uncover sources of biological variability in  
41 patients' blood, urine and allograft at the epigenetic and transcriptomic levels. MOFA reveals  
42 multicellular immune signatures characterized by distinct monocyte, natural killer and T cell  
43 substates explaining a large proportion of inter-patient variance. We also identified specific  
44 factors that reflect allograft rejection, complement activation or induction treatment. Factor 1  
45 mainly explained the molecular variations in patients' circulation and discriminated antibody-  
46 mediated rejection from T-cell mediated rejection. Factor 2 captured some of the molecular  
47 variation occurring within the allograft and associated with complement/monocytes crosstalk.  
48 Factor 4 captured the impact of ATG induction. These data provide proof-of-concept of  
49 MOFA's ability to reveal multicellular immune profiles in kidney transplantation, opening up  
50 new directions for mechanistic, biomarker and therapeutic studies.

51

52

## 53 **Introduction**

54

55           Kidney transplantation is the treatment of choice for patients with end-stage renal  
56 disease, offering a significant improvement in quality of life and a reduction in morbidity and  
57 mortality. However, the long-term success of renal transplantation depends on many factors,  
58 such as immunological compatibility between donor and recipient, optimal management of  
59 immunosuppressive therapies and early detection of any signs of rejection or post-transplant  
60 complications. In kidney transplantation, both immune and non-immune mechanisms  
61 contribute to the progression of histological lesions and scarring of the kidney graft. These  
62 lesions, resulting from complex interactions between immune cells, soluble molecules and  
63 graft architecture, compromise graft function and long-term survival.

64

65           The objectives of the BIOMARGIN study were to discover biomarkers of renal allograft  
66 lesions derived from biopsies, blood and/or urine. To achieve these objectives, the  
67 BIOMARGIN consortium employed an ambitious strategy involving a succession of clinical  
68 studies in 4 hospitals in three European countries for the discovery, confirmation and validation  
69 of biomarkers in which transplant patients provided blood, urine and biopsy samples. Large-  
70 scale supervised explorations, also known as "omics", of these biological samples separately  
71 using state-of-the-art analytical technologies led to the discovery of several biomarkers or  
72 insights in kidney transplantation<sup>1-5</sup>.

73

74           In addition, recent advances in omics technologies have led to unprecedented efforts  
75 to characterize the molecular changes underlying the pathophysiology of a wide range of  
76 complex human diseases such as coronary syndrome recently<sup>6</sup>. The combination of different  
77 omics technologies, called "multi-omics" analyses, has been proposed to decipher the  
78 molecular mechanisms involved in complex diseases. These analyses can be classified into  
79 supervised and unsupervised methods. The aim of supervised methods is to predict one or  
80 more conditions related to a sample, although over-fitting can be an issue. In contrast,

81 unsupervised methods explore the data by analyzing correlations between samples in order  
82 to condense the large volume of data into a reduced number of factors which, in turn, could be  
83 associated with clinical features. MOFA2 (Multiple Omics Factor Analysis version 2) is an  
84 unsupervised statistical approach developed to explore and integrate multiple omics data  
85 sources such as genomics, transcriptomics, proteomics and metabolomics<sup>7</sup>. In the present  
86 report, we used MOFA2 to uncover specific biological signatures associated with kidney  
87 transplant phenotypes and outcomes by identifying complex patterns and relationships  
88 between different molecular variables such as mRNA and miRNA from biopsies, blood and  
89 urine.

90

## 91 **Results**

### 92 ***MOFA application on BIOMARGIN datasets***

93 We collected data from blood, biopsy and urine samples of 131 kidney transplant recipients  
94 comprising six data types (also called views): 1 blood-derived epigenome (miRNA expression),  
95 2 blood-derived transcriptomes (mRNA quantified by MicroArray, mRNA quantified by RNA  
96 sequencing), 1 biopsy-derived epigenome (miRNA expression), 1 biopsy-derived  
97 transcriptome (mRNA quantified by MicroArray) and 1 urine-derived selected gene set (mRNA  
98 quantified by RT-qPCR). In blood, a total of 58828 genes were measured by RNAseq, 54675  
99 genes were detected by MicroArray and 758 miRNAs were measured by RT-qPCR. In biopsy,  
100 54613 genes were assessed by MicroArray and 758 miRNAs were measured by RT-qPCR. In  
101 urine, 34 genes were measured by RT-qPCR (**Figure 1a**). Given these 6 data matrices with  
102 measurements of multiple omics data types across sample sets or partially overlapping sample  
103 sets, MOFA infers an interpretable low-dimensional data representation of factors. These  
104 learned factors capture the main sources of variation between views, facilitating the  
105 unsupervised identification of continuous molecular gradients or discrete sample subgroups  
106 (**Figure 1a**). In order to integrate the various omics data with MOFA, we constructed 131  
107 multiomics profiles by matching samples in the 6 views. It should be noted that only 26.7%  
108 (N=35) of the 131 samples were profiled with all types of omics data mainly due to the limited

109 number of samples with mRNA data for urine (**Figure 1b**), but such a scenario of missing  
110 values is not uncommon in multidimensional cohort studies and MOFA is designed to cope  
111 with it<sup>7</sup>.

112

113 As it is recommended to perform a stringent selection of features before creating the MOFA  
114 object, we performed an initial selection of features from the Blood RNA-seq, Blood miRNAs  
115 and Biopsy miRNAs, keeping only the top 25% of genes with the greatest variance (First  
116 features selection). Integrating these features with MOFA resulted in an over-representation  
117 of transcriptome views at the expense of epigenome views (**Figure S1a**). Secondly, to ensure  
118 that epigenome views were not under-represented when fitting MOFA, we removed weakly  
119 expressed features from transcriptomic views and filtered out the least expressed genes  
120 (Second features selection). This selection resulted in a more balanced representation of  
121 blood-derived miRNAs, but did not improve the representation of biopsy-derived miRNAs (%  
122 explained variance < 5%, **Figure S1a**). Thirdly, we followed a strategy recommended by the  
123 MOFA authors to adjust the number of transcriptome features, selecting the 5000 genes with  
124 the greatest variability measured by standard deviation but also taking into account all  
125 epigenomic features. We observed that estimating MOFA with a greater number of epigenomic  
126 features led to a significant increase in the percentage of variance explained in the epigenome  
127 and to a more balanced final features selection. (**Figure S1a and Figure 1c**).

128

129 To obtain a MOFA model with this final features selection, we trained MOFA 100 times. We  
130 chose the MOFA model with the lowest absolute Evidence Lower Bound (ELBO) value to strike  
131 a balance between model complexity and explanatory power (**Figure S1b**). MOFA identified 8  
132 factors that were largely orthogonal, capturing independent sources of variation. Cumulatively,  
133 the 8 factors explained 38% of the variation in blood-RNAseq data, 61% in blood-MicroArray  
134 data, 27% in blood-epigenome data, 42% in biopsy-transcriptome data, 23% in biopsy-  
135 epigenome data and 2.9% in urine data (**Figure 1c**). Of these, Factor 2 and Factor 3 were  
136 active in most assays (**Figure 1d**), indicating extensive roles in multiple molecular layers in

137 both the transplanted kidney and the circulation. In addition, other factors such as Factor 1 and  
138 Factor 4 were specific to blood-related views. The fact that most of the factors explain the  
139 variance between several views indicates that the data are very highly correlated between  
140 views, and that it is possible to identify common patterns between different omics views. In  
141 contrast, Factor 6, Factor 7 and Factor 8 were mainly active in only one data modality which  
142 diminishes their interest in this multiomics approach (**Figure 1d**). We next estimated the  
143 correlation between factors and observed correlations  $R^2$  below 0.4 indicating that each factor  
144 captures independent and unique sources of variations (**Figure S1c**).

145

#### 146 ***MOFA identifies important clinical markers in kidney transplantation***

147 To understand how factors relate to kidney transplant phenotypes, we assessed the  
148 relationship between factor loadings and different kidney transplant outcome parameters  
149 (**Figure 1e**). Patient characteristics are shown in **Supplemental Table 2**. Intriguingly, some  
150 factors correlated exclusively with Banff histological lesions or complement activation: Factor  
151 2 was positively correlated with C4d deposition in peritubular capillaries (“C4d positivity”; P  
152 value<0.001) while Factor 5 was associated with C4d positivity ( $R^2<0$ , P value<0.0001), tubular  
153 atrophy (“ct”;  $R^2>0$ , P value<0.01), interstitial fibrosis (“ci”;  $R^2>0$ , P value<0.001) and arteriolar  
154 hyalinosis (“ah”;  $R^2>0$ , P value<0.01). In contrast, Factor 3 and Factor 8 correlated uniquely  
155 with certain causes of renal failure (interstitial nephritis,  $R^2>0$ , P value<0.001 and hypertension,  
156  $R^2<0$ , P value <0.0001 respectively). Factor 4 was strongly associated with  
157 immunosuppression: it was negatively correlated with ATG induction (P value <0.0001) and  
158 positively correlated with tacrolimus immunosuppression (P value <0.01). In addition, certain  
159 factors were associated with multiple parameters of varying natures. Factor 7 was negatively  
160 associated with diabetes as a cause of renal failure (P value<0.01), creatinine (P  
161 value<0.0001), anti-HLA DSA (P value<0.01), g lesions (P value<0.01) and v lesions (P  
162 value<0.00001). Factor 1 correlated with eGFR ( $R^2<0$ , Pearson correlation P value<0.01), anti-  
163 HLA DSA ( $R^2<0$ , P value<0.001), transplant glomerulopathy (“cg”;  $R^2<0$ , P value<0.01) and  
164 total inflammation (“ti”;  $R^2>0$ , P value<0.01). Lastly, Factor 6 was associated with

165 glomerulonephritis as a cause of kidney failure ( $R^2>0$ , Pearson correlation  $P$  value $<0.01$ ),  
166 eGFR ( $R^2>0$ ,  $P$  value $<0.01$ ), erythrocyte titer ( $R^2>0$ ,  $P$  value $<0.001$ ), leukocyte titer ( $R^2<0$ ,  $P$   
167 value $<0.001$ ) and ti ( $R^2>0$ ,  $P$  value $<0.01$ ).

168

### 169 ***Factor 1 discriminates antibody-mediated from T-cell mediated rejection***

170 Factor 1 mainly explained the molecular variations in blood of kidney transplant recipients  
171 (**Figure 1d**) and was negatively correlated with anti-HLA DSA status. Intriguingly, this blood-  
172 related factor was also associated with cg and ti lesions in the graft (**Figure 1e**). To confirm  
173 these correlations, we assessed the global distribution of Factor 1 loading (**Figure 2a**), and  
174 then explored the distribution of Factor 1 loading according to cg and ti lesion scores. We  
175 observed a significant decrease in Factor 1 loading in cases with a cg $>0$  score (**Figure 2b**),  
176 while Factor 1 loading increased significantly in cases with a ti $>1$  score (**Figure 2c**). We also  
177 confirmed the significant decrease in Factor 1 loading in the presence of anti-HLA DSA (**Figure**  
178 **2d, Supplemental Table 3**). We also performed univariate logistic regressions and reported  
179 the performance of Factor 1 to discriminate HLA-DSA status and the presence of ti lesions and  
180 cg lesions by (**Figure S2a**). Given the increase in Factor 1 loading with ti lesions and the  
181 decreases in the presence of HLA-DSA and cg lesions, we tested whether Factor 1 loading  
182 differed between cases with antibody-mediated rejection (AMR) and T-cell mediated rejection  
183 (TCMR). To this end, we stratified patients according to different phenotypes (AMR, TCMR,  
184 Normal and IFTA) and we observed a significant difference between the TCMR and AMR  
185 group suggesting that Factor 1 may discriminate TCMR from AMR (**Figure 2e**).

186 Features contributing more than 0.9 of the absolute Factor 1 loading mainly derived from blood  
187 MicroArray data ( $D= 11$ , **Figure 2f**). The other molecules contributing to each view have also  
188 been represented in **Figure S2b**. The significant contribution of the blood MicroArray data view  
189 to the Factor 1 is reflected in the heatmap, where a clear separation was observed between  
190 blood samples (**Figure 2g**). In fact, unsupervised clustering of blood samples using these 11  
191 top features resulted in two main clusters: samples with HLA-DSA were preferentially grouped  
192 in Cluster A in which the 11 features were overexpressed (two-tailed Fisher's exact test,  $P$ -

193 value = 0.001). For the grouping of cases according to AMR or TCMR phenotypes, the Fisher  
194 exact test P-values were 0.0562 and 0.1676 respectively. We then used a publicly available  
195 blood RNASeq dataset GSE120649<sup>8</sup> (**Figure S2c**) to validate whether these 11 features in  
196 blood MicroArray and the 3 features contributing more than 0.9 of the absolute Factor 1 loading  
197 in blood RNA-Seq would be differentially expressed in AMR and TCMR. Interestingly, 12/14  
198 features were found increased in AMR in this dataset (**Figure S2d**), suggesting that Factor 1  
199 could reflect the different immune responses of the two types of rejection in circulating cells.  
200 In order to map at the single-cell level the different renal and/or immune cells that could express  
201 the top features of Factor 1, we reintegrated 46 publicly available scRNA-Seq datasets from  
202 kidney transplant patients, including transcriptomes from both circulating blood cells (PBMC)  
203 and cells derived from kidney biopsies (**Figure S3a**). We were thus able to obtain 150,876  
204 transcriptomes which passed quality check (**Figure S3b**) and, using canonical markers  
205 (**Figure S3c**), to identify 23 clusters corresponding to all renal and circulating cell populations  
206 (**Figure S4a**). For further granularity, a subset corresponding only to circulating cells was  
207 selected and 29 clusters were identified and automatically annotated in an unsupervised  
208 manner (**Figure S4b**). We then formed a signature corresponding to the top features  
209 explaining Factor 1 in the blood and we observed that the top features of Factor 1 were not  
210 centralized in a single immune cell population, but instead scattered across myeloid cells, T,  
211 B and NK lymphocytes, and even granulocytes (**Figure S4c**). More specifically, *ZNF267* is  
212 preferentially expressed by neutrophils, while *MOB1B* and *TVP23B* are derived from basophils  
213 and progenitors. *CD69*, *SUCO*, *SMIM15* and *AGL* are expressed by lymphocytes, *ABCB10* by  
214 monocytes and *DMXL1*, *CLK1* and *RSL24D1* by B lymphocytes and plasmablasts. Factor 1  
215 thus corresponds to a multicellular immune profile that differs between TCMR and AMR in  
216 patients' blood.

217

### 218 ***Factor 2 is associated with complement/monocytes crosstalk***

219 In contrast to Factor 1, Factor 2 captured some of the molecular variation occurring inside the  
220 allograft. Among all the outcomes tested, it was only positively correlated with C4d (Pearson

221 correlation,  $P$ -value $<0.001$ ) (**Figure 1e**). Of note, C4d deposition in peritubular capillaries is  
222 associated with immune reactions directed against the allograft and are the result of activation  
223 of the complement system. To confirm the correlation between C4d positivity and Factor 2, we  
224 assessed the global distribution of Factor 2 loading (**Figure 3a**), and then explored the  
225 distribution of Factor 2 loading according to C4d positivity. We observed a significant increase  
226 in Factor 2 loading in cases with C4d positivity (**Figure 3b**). By stratifying patients according  
227 to Factor 2 median, a trend towards an increase in the number of C4d-positive cases was also  
228 found in patients above the median (**Supplemental Table 3**). The top features contributing to  
229 the absolute Factor 2 loading mainly derived from biopsy MicroArray data ( $D=15$ , **Figure 3c**).  
230 The top features deriving from biopsy MicroArray data and explaining Factor 2 clearly  
231 separated the samples into two groups (**Figure 3d**) in an unsupervised manner. Cases positive  
232 for C4d were grouped showed high expression of the top features *LYZ*, *CALHM6*, *EVI2A*,  
233 *CD52*, *IGSF6*, *C1QB*, *EVI2B*, *BCL2A1*, *C1QC*, *C16orf54*, *CSTA*, *CXCL9*, *CTSS*, *CXCL10* and  
234 *TYROBP*. Interestingly, two of these features *C1QC* and *C1QB* encode the B and C-chains  
235 polypeptide of serum complement subcomponent C1q, which associates with C1r and C1s to  
236 yield the first component of the serum complement system. With regard to other top features  
237 explaining Factor 2 in biopsies, five miRNAs were detected: miR-150, miR-223, miR-1227,  
238 miR-624 and miR-155 (**Figure S5a**). Among the upregulated miRNAs, our group has  
239 previously reported that miR-155 is preferentially expressed by monocytes<sup>4</sup>, suggesting that  
240 Factor 2 may capture the infiltration of blood monocytes/macrophages into the allograft and  
241 the overexpression of complement by these cells. In order to validate the association between  
242 these Factor2 top features and C4d deposition in peritubular capillaries in an external dataset,  
243 MicroArray data from biopsies of kidney transplant patients with and without C4d were  
244 analyzed (**Figure 3e**). Here, the transcriptomes of 23 C4d-negative biopsies were compared  
245 with 16 biopsies showing focal, diffuse or minimal C4d ptc staining (**Figure 3f**). As shown on  
246 the volcano plot, biopsies positive for C4d showed overexpression of both mRNA transcripts  
247 and miR-150, miR-223 and miR-155 explaining Factor 2. In the scRNA-Seq dataset, we then  
248 formed a signature corresponding to the top features explaining Factor 2 in the kidney and

249 observed that this signature was particularly strong in myeloid cells (**Figure S5b**). Strikingly,  
250 macrophages were the main cell population expressing the top features (**Figure S5c**) and  
251 more specifically *C1QC* and *C1QB*, suggesting that Factor 2 captured the crosstalk between  
252 macrophages and complement activation in the allograft after transplantation.

253

#### 254 ***Factor 4 captures the impact of induction prophylaxis***

255 Similar to Factor 1, Factor 4 captured some of the molecular variation in blood samples (**Figure**  
256 **1d**). Factor 4 was strongly and negatively correlated with ATG induction (Pearson correlation,  
257  $P\text{-value} < 0.00001$ ) and positively correlated with tacrolimus immunosuppression protocol  
258 (**Figure 1e**). Evaluating the overall distribution of Factor 4 loading (**Figure 4a**), and then  
259 exploring the distribution of Factor 4 loading according to induction types, we observed that  
260 Factor 4 loading was also reduced in cases that had received ATG as induction compared to  
261 no induction ( $P\text{-value} = 0.0023$ ) or other induction therapy (**Figure 4b, Supplemental Table 5**).  
262 Considering the top 10 features measured by RNAseq in blood in terms of absolute loading,  
263 two clear groups of patients were distinguished in an unsupervised manner: one showing low  
264 *IL1R2* expression and without patients who had received ATG ( $N = 0/21$ , 0%), and a second  
265 with high *IL1R2* levels including 15/53 patients who had received ATG as induction (28%).  
266 Interestingly, among the other top features, expressions of *TRBC2*, *BCL11B*, *CD3G*, *CD247*,  
267 *CD3E*, *STMN3*, *GIMAP7*, *FCMR*, *SRNPN* and *CCR7* were down-regulated in the group  
268 including ATG-treated patients (**Figure 4c-d**). The advantage of MOFA is to uncover  
269 transcriptomic profiles across omics types and tissues. Strikingly, one of the most weighted  
270 features explaining Factor 4 in the biopsy is *IL1R2* (**Figure S6a**).

271

272 Querying the scRNA-Seq dataset, we observed that the other major features were principally  
273 expressed by T cells and NK cells (**Figure S6b**). More specifically, naive T cells  
274 subpopulations as well as regulatory T cells and central CD8 T cells highly expressed Factor  
275 4 top features such as *FCMR*, *CCR7*, *SNRPN* and *BCL11B*. *FCMR* expression was mainly  
276 restricted to B cells subpopulation. In addition, we observed that *IL1R2* is predominantly

277 expressed by and myeloid dendritic cells and classical monocytes in the blood (**Figure S6c**),  
278 suggesting a disturbance in the lymphoid/myeloid cell ratio in both patient groups. To validate  
279 these data, we took advantage of a public GSE10040 dataset corresponding to an *in vitro*  
280 experiment in which PBMCs from healthy volunteers were treated with ATG at 10 $\mu$ g/mL for  
281 24h<sup>9</sup>. After culture, cells were harvested and mRNA extracted for full transcriptomic analysis  
282 (**Figure 4e**). We found that ATG treatment induced the differential expression of 4438 genes  
283 (DEGs) comprising 2327 down-regulated genes and 2112 up-regulated genes. By mapping  
284 the top features of Factor 4 among these 4438 DEGs, we observed that *IL1R2* is strongly  
285 induced by ATG treatment (fold change (FC) >3.9 and Log<sub>10</sub>(pvalue) >6.4) but so is *CCR7*  
286 (FC>0.9, Log<sub>10</sub>(pvalue) >3.3). In contrast, the other features are mainly down-regulated, such  
287 as *SNRPN* (FC<-0.7, Log<sub>10</sub>(pvalue) >3.5), *CD3E* (FC<-0.3, Log<sub>10</sub>(pvalue) >1.7), *FCMR*  
288 (FC<-0.9, Log<sub>10</sub>(pvalue) >2.1), *BCL11B* (FC<-0.5, Log<sub>10</sub>(pvalue) >1.8), *CD247* (FC<-0.3,  
289 Log<sub>10</sub>(pvalue) >1.3), *STMN3* (FC<-0.1, ns) and *GIMPAP7* (FC<-0.1, ns). Of note, *TRCB2* was  
290 not detected and *CD3G* was slightly increased (FC>0.1, ns) (**Figure 4f**). Overall, these results  
291 suggest that ATG treatment induces a strong disruption of the immune system that could be  
292 detected several months after treatment in blood but also in the allograft, leading to a decrease  
293 in the lymphoid compartment in favor of the myeloid compartment.

294

## 295 **Discussion**

296 MOFA2 analysis enables researchers to simultaneously combine and analyze high-  
297 throughput data from different biological sources. This method of analysis is proving  
298 particularly valuable in the field of medical research, where it enables us to better understand  
299 the complexity of the molecular and cellular interactions involved in complex disease  
300 processes. By elucidating these mechanisms, clinicians can better target therapies and  
301 improve long-term kidney success. To our knowledge, this is the first time this type of MOFA  
302 analysis has been conducted in the field of kidney transplantation. Integrating the different  
303 omics layers, MOFA2 delimited 8 different factors in an unsupervised manner. As expected,  
304 these factors captured independent transcriptomic profiles across blood and renal allografts.

305 Surprisingly, we observed very few associations between histology and the factors.  
306 This indicates that rejection phenotypes (according to the Banff classification), but also  
307 rejection severity and phenotype patterns such as rejectionclass are not the main drivers of  
308 molecular heterogeneity within allografts and in circulation. In fact, only Factor 1 reflects  
309 different profiles between HLA-DSA positive and HLA-DSA negative cases.

310 Factor 2 was related to the intragraft crosstalk between monocytes/macrophages and  
311 complement. It was associated with C4d deposition independent of rejection status, suggesting  
312 that this crosstalk is not specific to the AMR process. Indeed, Factor 2 was not significantly  
313 correlated with the presence of HLA-DSA, suggesting that this crosstalk between  
314 macrophages and complement might be due to non-HLA antibodies which could be  
315 detrimental in the context of kidney transplantation<sup>10-12</sup>. In fact, interactions between  
316 monocytes/macrophages and complement factors are multiple and complex, contributing  
317 significantly to the innate immune responses. Infiltrating monocytes and tissue-resident  
318 macrophages can locally synthesize several components of the complement system, including  
319 C1, C2, C3, C4, C5, C6, C7, C8, and C9. In particular, we have previously shown that CD163+  
320 macrophages are the main immune cells expressing complement-associated genes such as  
321 *C1QA*, *C1QB* and *C1QC* among myeloid cells present in the allograft<sup>13</sup>. In turn, certain  
322 complement degradation products, such as the C5a fragment via CD88 or C3a via Toll Like  
323 Receptors or C4d via LILRB2 and LILRB3<sup>14</sup>, act as chemokines, attracting monocytes and  
324 macrophages and modulating their production of cytokines such as TGF- $\beta$ <sup>15</sup>. It is worth noting  
325 that urinary *TGFB1* mRNA was the second most important feature (absolute loading>0.95)  
326 explaining Factor 2, and one could speculate that this urinary detection of *TGFB1* mRNA could  
327 be due to podocyte lesions<sup>16</sup>. At the same time, biopsy-derived miR-155 was also among the  
328 key features explaining Factor 2. Intriguingly, this miRNA was closely linked to podocyte  
329 apoptosis after exposure to TGF- $\beta$ <sup>17</sup>. Altogether our results suggest that Factor 2 reflects the  
330 monocytes/macrophages crosstalk with complement within the allograft and that this crosstalk  
331 may be associated with podocyte lesions.

332

333 Surprisingly, Factor 4 was significantly associated with the type of induction regimen,  
334 suggesting that ATG therapy can affect patients' immune profiles longer term after  
335 transplantation. It is clear that ATG makes patients more susceptible to infections, both in the  
336 short and long term<sup>18</sup>. This increased sensitivity may persist even after ATG treatment has  
337 been discontinued. ATG induction has also been associated with a higher risk of long-term  
338 malignancies compared with anti-CD25 induction<sup>19</sup>. The top-weighted genes of Factor 4  
339 suggest an imbalance in the patient's immune system, with a decrease in the T cell  
340 compartment in favor of the myeloid compartment both in the circulation and in the allograft.

341  
342 This study has several limitations. The datasets used in this study are relatively small, which  
343 may limit the ability to capture the full spectrum of variability and complexity present in larger,  
344 more diverse populations. Indeed, our study primarily involves Western European centers with  
345 a predominantly Caucasian population. This demographic constraint may limit the applicability  
346 of our results to other ethnic groups and geographic regions, potentially reducing the overall  
347 generalizability of our conclusions. In addition, the analysis in this study is fully dependent on  
348 the MOFA2 algorithm. While MOFA2 is a powerful tool for integrative analyses, relying  
349 exclusively on this algorithm means that other integrative approaches might yield different  
350 factors and insights. This dependence highlights the need for comparative studies using  
351 alternative methodologies to validate our findings. Although we conducted systematic  
352 validation on external datasets for each factor of interest, our study lacks in-depth mechanistic  
353 studies to confirm the identified factors with greater certainty.

354  
355 In conclusion, MOFA2 analysis represents a major advance in the integration of omics data to  
356 understand kidney transplantation. Our study highlights the significant associations between  
357 factors and clinical parameters. This highlights MOFA as an innovative approach to dissect  
358 multicellular immune profiles with mechanistic and clinical implications in kidney  
359 transplantation. Furthermore, our study, as a globally available resource, provides new targets

360 for large-scale MOFA-based experimental studies and biomarker assays, and helps prioritize  
361 new candidate targets for immunomodulatory interventions in kidney transplantation.

362

## 363 **Methods**

### 364 *Patient population and data collection*

365 The present study is part of the Reclassification using OmiCs integration in KidnEy  
366 Transplantation (ROCKET) project, which is based on the BIOMarkers of Renal Graft INjuries  
367 (BIOMARGIN) study (ClinicalTrials.gov number NCT02832661). Patients were included  
368 prospectively in four European transplant centers between June 2011 and March 2017  
369 (University Hospitals Leuven, Belgium; Medizinische Hochschule Hannover, Germany; Centre  
370 Hospitalier Universitaire Limoges, France, and Hôpital Necker Paris, France). In all four clinical  
371 centers, protocol renal allograft biopsies were performed 3, 12 and sometimes 24 months after  
372 transplantation, in accordance with local practice, in addition to clinically indicated biopsies  
373 (biopsies at the time of graft dysfunction). In parallel, blood and urine samples were collected  
374 at the same times. All adult patients who had received a single renal allograft at these  
375 institutions and provided written informed consent for this study were eligible. This consent  
376 adheres to the Declaration of Istanbul. Ethics committee XI #13016, Paris, France for Necker  
377 Hospital gave ethical approval for this work. Ethics committee #S55598, Leuven, Belgium for  
378 Leuven Hospital gave ethical approval for this work. Ethics committee #6475, Hannover,  
379 Germany for Hannover Medical School gave ethical approval for this work. Ethics committee  
380 #DC-2010-1075, Limoges, France for Limoges Hospital gave ethical approval for this work.  
381 Recipients of combined transplantations were excluded. All transplantations were  
382 complement-dependent cytotoxicity cross-match negative. The study protocol was approved  
383 by institutional review boards and national regulatory agencies (where applicable) at each  
384 clinical center. The BIOMARGIN study was divided into three phases. Only data from the first  
385 exploratory phase are used in the present report. In this discovery phase, blood samples were  
386 used for an epigenome analysis (miRNA expression [E-MTAB-9595]<sup>5</sup>) and two transcriptome  
387 analyses (mRNA quantified by MicroArray [GSE129166]<sup>1</sup>, and bulk RNA sequencing

388 [GSE175718]<sup>3</sup>). In parallel, biopsy samples were used for an epigenome analysis (miRNA  
389 expression [GSE179772]<sup>4</sup>) and a transcriptome analysis (mRNA quantified by MicroArray  
390 [GSE147089]<sup>2</sup>). With the exception of urinary mRNA, each dataset is publicly available, and  
391 details of RNA extractions and RNA expression analysis are extensively detailed in the  
392 corresponding reports.

393

#### 394 *Urinary mRNA quantification*

395 For mRNA quantified in urine, samples were centrifuged for 20 min at 2,000g at 4°C. The cell  
396 pellet was resuspended in 700µL of PBS and centrifuged for 5 min on a tabletop centrifuge at  
397 maximum speed. PBS was discarded and cells were resuspended in 500µL RLT buffer  
398 (Qiagen, Courtaboeuf, France) in a cryotube before being frozen at -80°C. MessengerRNA  
399 was extracted from the pellet using the RNeasy mini kit (Qiagen) and reverse transcribed into  
400 cDNA using TaqMan® reverse transcription reagents (Applied Biosystems). We used in-house  
401 designed oligonucleotide primers and fluorogenic probes to measure mRNA levels of  
402 ribosomal RNA 18S, *ACTA2*, *ENG*, *CD14*, *CD3E*, *CD46*, *CFB*, *CXCL13*, *CXCL9*, *CXCL10*,  
403 *IL2RA*, *CDH1*, *FN1*, *FOXP3*, *GZMB*, *GAPDH*, *HGF*, *PRF1*, *PSMB9*, *PSMB10*, *SLC12A1*,  
404 *TGFB1*, *TLR4* and *VIM* as detailed in Supplementary Table 1 or commercial assays  
405 (Thermofisher) for *CTSS* (Hs00175407\_m1), *GNLY* (Hs00246266\_m1), *FCGR3A*  
406 (Hs00275547\_m1), *ISG20* (Hs00158122\_m1), *KLRD1* (Hs00233844\_m1), *MMP7*  
407 (Hs01042796\_m1), *MMP9* (Hs00957562\_m1), *NKG2D* (Hs00183683\_m1), *NKG7*  
408 (Hs01120688\_g1), *RUNX3* (Hs00231709\_m1), and *UPK1A* (Hs01086736\_m1). PCR analysis  
409 was performed in two steps, a pre-amplification step as described previously<sup>20</sup> (Veriti 96-Well  
410 Thermal Cycler, Applied Biosystems) followed by mRNA measurement with a Viia 7 Real-Time  
411 PCR system (Thermofisher). *GAPDH* was used as housekeeping gene<sup>21</sup> for normalization  
412 using the delta Ct method. The normalized expression was log transformed before MOFA  
413 integration.

#### 414 *Data processing and filtering*

415 Both blood- and biopsy-derived miRNA expressions were determined as previously  
416 described<sup>4,5</sup> using two small-nucleolar RNAs for normalization: RNU44 and RNU48. The  
417 normalized expressions were log transformed before MOFA integration. Blood- and biopsy-  
418 derived transcriptomes quantified by MicroArray were subjected to Robust Multichip Average  
419 (RMA) normalization as previously described<sup>1,2</sup>. These datasets were then curated and  
420 annotated using the Biological Interpretation Of Multiomics EXperiments (BIOMEX) workflow<sup>22</sup>  
421 before MOFA integration. Raw counts corresponding to the blood-derived transcriptome  
422 quantified by bulk RNA sequencing were filtered to exclude weakly expressed genes (with <5  
423 counts in 50% of the samples). After filtration, the data were subjected to normalization by  
424 variance stabilizing transformation (VST) using the R package DamiRseq (2.1.0)<sup>23</sup> prior to  
425 MOFA integration.

426

#### 427 *Multiple Omics Factor Analysis version 2*

428 To integrate the six datasets, we used the R package MOFA2 (1.8.0). MOFA is an  
429 unsupervised machine learning method that identifies latent factors that capture biological  
430 sources of variability in multi-omics datasets. It should be noted that clinical covariates were  
431 not used to train the model. Data, model and learning options were left default. MOFA was run  
432 with 100 iterations in 'slow' convergence mode to ensure model convergence; the final model  
433 converged after 61 iterations. Interpretation of the factors is analogous to that of the principal  
434 components and the relationship between clinical covariates and MOFA factors was analyzed  
435 a posteriori.

436

#### 437 *Clinicopathological diagnosis of acute rejection subtypes*

438 Two kidney biopsy cores were obtained using a 14-gauge needle under sonographic guidance.  
439 One biopsy core was fixed in formalin and embedded in paraffin for standard histopathological  
440 assessment. Half of the second biopsy core was used for frozen sections and/or electron  
441 microscopy; the remaining half core was used for epigenome and transcriptome analyses. All

442 biopsies were scored according to the internationally standardized Banff 2017 lesion scores<sup>24</sup>.  
443 The follow-up of anti-HLA antibodies and annotation of donor-specific antibodies (DSA) was  
444 systematically monitored in the histocompatibility laboratory referent of each inclusion center.  
445 A diagnostic label was awarded to each biopsy based on the presence and severity of these  
446 histological lesions and on the DSA status, in concordance with the Banff 2017 classification.

447

#### 448 *Single-cell RNA-Sequencing (scRNA-Seq) validation*

449 We reintegrated 46 publicly available scRNA-Seq datasets corresponding to kidney transplant  
450 biopsies or peripheral blood mononuclear cells (PBMCs) from patients presenting allograft  
451 rejection or not, whose raw data were downloaded from various repositories: E-MTAB-11450  
452 <sup>25</sup>, E-MTAB-12051 <sup>26</sup>, GSE140989 <sup>27</sup>, GSE145927 <sup>28</sup>, GSE171374 <sup>29</sup>, PRJNA974568 <sup>30</sup>. Seurat  
453 R package (v5.0.1) was used to read, create and merge all raw counts matrices into a single  
454 object, which was subjected to the following QC parameters: number of features (genes)  
455 between 300 and 10,000 per cell and percent of total counts from mitochondrial genes, as  
456 defined by the prefix « MT- », below 10% for cells from PBMCs datasets, and below 25% for  
457 cells from kidney transplant biopsies datasets.

458 A total of 151,862 cells and 40,353 genes successfully passed QC. Data were log-normalized,  
459 scaled, the top 2000 variable features as well as the top 50 Principal Component (PC)  
460 dimensions were calculated prior to integration using Reciprocal Principal Component Analysis  
461 (RPCA) method, with GSM4339779 and pbmc5 as references and all other datasets as  
462 queries on a supercomputer (Mésocentre de calcul de Franche-Comté). Uniform Manifold  
463 Approximation and Projection (UMAP) dimensionality reduction was calculated from the top  
464 50 PC dimensions, and a resolution of 0.47 was used for unsupervised clustering. scDbIFinder  
465 R package (v1.16.0) was used for doublets discrimination (as defined by clusters composed  
466 of two or more different cells captured in a droplet and sequenced). Kidney-derived clusters  
467 annotation was made using already-described canonical markers<sup>25-36</sup>. PBMC-derived clusters  
468 automatic annotation was performed using SingleR (v2.4.1) and celldex (v1.12.0) R packages

469 and MonacolmuneData as reference dataset <sup>37</sup>. Seurat, scCustomize <sup>38</sup> (v2.0.1) and  
470 RightSeuratTools <sup>39</sup> (v1.0.1) R packages were used for data visualization.

471

#### 472 *Whole Transcriptomic external validations*

473 Various external datasets were used to validate the top features explaining the factors  
474 determined by MOFA.

##### 475 a) External validation of Factor 1 top-weighted genes

476 To validate the potential for discrimination between AMR and TCMR patients by the top-  
477 weighted Factor-1-related genes, an external blood-derived transcriptomic dataset,  
478 GSE120649<sup>40</sup>, was used. This dataset includes bulk RNAseq analysis of whole blood cells  
479 isolated from 6 patients with histologically verified AMR and 4 patients with histologically  
480 verified TCMR after kidney transplantation. In brief, the BIOMEX pipeline was used to  
481 determine differentially expressed genes between the two patient groups.

##### 482 b) External validation of Factor 2 top-weighted genes

483 To validate the potential for discrimination between C4d negative biopsies and C4d positive  
484 biopsies by the top-weighted Factor-2-related genes, an external kidney allograft-derived  
485 transcriptomic dataset, E-GEOD-38262<sup>41</sup>, was used. This dataset includes 92 microarray  
486 datasets corresponding to kidney transplant biopsies from patients divided into seven groups  
487 classified according to their histopathological scores, whose raw data (CEL files) were  
488 downloaded from ArrayExpress (<https://www.ebi.ac.uk/biostudies/arrayexpress/>). oligo  
489 (1.66.0), hugene10sttranscriptcluster.db (8.8.0), hugene10stprobeset.db (8.8.0) and  
490 pd.hugene.1.0.st.v1 (3.14.1) R packages were used to read raw data and normalize  
491 expression, as well as to convert genes names from Affymetrix HuGene 1.0 st Probe IDs into  
492 ENSEMBL IDs. The expression matrix was subsequently analyzed using the BIOMEX pipeline.  
493 Here we selected only samples of interest corresponding to biopsy transcriptomes showing  
494 minimal, diffuse or focal C4d deposition in the absence of glomerular disease (N=25) or  
495 absence of C4d and glomerular disease (N=17). After PCA analysis, 3 outlier samples were

496 excluded (GSM937601, GSM937665, GSM937666) and genes differentially expressed  
497 between the 2 groups were calculated.

#### 498 c) External validation of Factor 4 top-weighted genes

499 To validate the differential expression of top-weighted genes of Factor 4 according to ATG  
500 treatment, an external transcriptomic dataset GSE10040<sup>42</sup>, was used. This dataset includes a  
501 MicroArray analysis of PBMC isolated from healthy volunteers and in vitro treated with or  
502 without ATG (10µg/mL), whose raw data (CEL files) were downloaded from Gene expression  
503 Omnibus (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE10040>). affy (1.70.0),  
504 hgu133plus2.db (3.13.0), and limma (3.48.3) R packages were used to read raw data and  
505 normalize expression, as well as to annotate genes into ENSEMBL IDs. The expression matrix  
506 was subsequently analyzed using the BIOMEX pipeline. In brief, the BIOMEX pipeline was  
507 used to determine which genes were differentially expressed between the two experimental  
508 groups.

509

#### 510 *Statistical analysis*

511 We report descriptive statistics using mean and standard deviation (or median and interquartile  
512 range for skewed distributions) for continuous variables or numbers, and percentages for  
513 discrete variables, for the full cohort and for the rejection subgroups. We used R Studio  
514 (2023.06.1+524) and GraphPad Prism (v10; GraphPad Software, San Diego, CA, United  
515 States) for statistical analysis and data presentation. The volcano plots were constructed using  
516 the R packages ggplot2 (3.4.1) and ggrepel (0.9.3) and were used to annotate the genes of  
517 interest.

518

#### 519 **Data availability**

520 All data presented in this present study are derived from the BIOMARGN study and are already  
521 publicly available, with the exception of urinary gene expression, which can be shared upon  
522 reasonable request. Epigenome analyses were performed using the following datasets: blood-  
523 derived miRNA expression E-MTAB-9595 and biopsy-derived miRNA expression GSE179772.

524 Blood-derived transcriptome analyses were performed using the following datasets:  
525 MicroArray-quantified mRNA GSE129166 and bulk RNA sequencing GSE175718. Biopsy-  
526 derived transcriptome analysis was performed using the following dataset: GSE147089.

## 527 **Code availability**

528 Code for MOFA2 is available at <https://biofam.github.io/MOFA2/>.

529

## 530 **References**

- 531 1. Van Loon, E. *et al.* Development and validation of a peripheral blood mRNA assay for  
532 the assessment of antibody-mediated kidney allograft rejection: A multicentre,  
533 prospective study. *EBioMedicine* **46**, 463–472 (2019).
- 534 2. Callemeyn, J. *et al.* Transcriptional Changes in Kidney Allografts with Histology of  
535 Antibody-Mediated Rejection without Anti-HLA Donor-Specific Antibodies. *J Am Soc*  
536 *Nephrol* **31**, 2168–2183 (2020).
- 537 3. Van Loon, E. *et al.* Biological pathways and comparison with biopsy signals and cellular  
538 origin of peripheral blood transcriptomic profiles during kidney allograft pathology.  
539 *Kidney Int* (2022) doi:10.1016/j.kint.2022.03.026.
- 540 4. Tinel, C. *et al.* Integrative Omics Analysis Unravels Microvascular Inflammation-  
541 Related Pathways in Kidney Allograft Biopsies. *Front Immunol* **12**, 4595 (2021).
- 542 5. Tinel, C. *et al.* Modulation of Monocyte Response by MicroRNA-15b/106a/374a  
543 During Antibody-mediated Rejection in Kidney Transplantation. *Transplantation* **107**,  
544 1089–1101 (2023).
- 545 6. Pekayvaz, K. *et al.* Multiomic analyses uncover immunological signatures in acute and  
546 chronic coronary syndromes. *Nat Med* **30**, 1696–1710 (2024).
- 547 7. Argelaguet, R. *et al.* Multi-Omics Factor Analysis—a framework for unsupervised  
548 integration of multi-omics data sets. *Mol Syst Biol* **14**, (2018).
- 549 8. Matz, M. *et al.* The regulation of interferon type I pathway-related genes RSAD2 and  
550 ETV7 specifically indicates antibody-mediated rejection after kidney transplantation.  
551 *Clin Transplant* **32**, e13429 (2018).
- 552 9. Feng, X. *et al.* Rabbit ATG but not horse ATG promotes expansion of functional CD4  
553 CD25 high FOXP3 regulatory T cells in vitro. doi:10.1182/blood-2008.
- 554 10. Delville, M. *et al.* Early acute microvascular kidney transplant rejection in the absence  
555 of anti-HLA antibodies is associated with preformed IgG antibodies against diverse  
556 glomerular endothelial cell antigens. *Journal of the American Society of Nephrology*  
557 **30**, 692–709 (2019).
- 558 11. Lebraud, E., Eloudzeri, M., Rabant, M., Lamarthée, B. & Anglicheau, D. Microvascular  
559 Inflammation of the Renal Allograft: A Reappraisal of the Underlying Mechanisms.  
560 *Front Immunol* **13**, (2022).
- 561 12. Lamarthée, B. *et al.* CRISPR/Cas9-Engineered HLA-Deleted Glomerular Endothelial  
562 Cells as a Tool to Predict Pathogenic Non-HLA Antibodies in Kidney Transplant  
563 Recipients. *Journal of the American Society of Nephrology* **32**, 3231–3251 (2021).

- 564 13. Lamarthée, B. *et al.* Transcriptional and spatial profiling of the kidney allograft  
565 unravels a central role for FcyRIII+ innate immune cells in rejection. *Nat Commun*  
566 2022.07.07.22276374 (2023) doi:10.1038/s41467-023-39859-7.
- 567 14. Hofer, J. *et al.* Ig-like transcript 4 as a cellular receptor for soluble complement  
568 fragment C4d. *FASEB J* **30**, 1492–503 (2016).
- 569 15. Gu, H. *et al.* Crosstalk between TGF- $\beta$ 1 and complement activation augments  
570 epithelial injury in pulmonary fibrosis. *FASEB J* **28**, 4223–34 (2014).
- 571 16. Fukuda, A. *et al.* Urinary podocyte and TGF- $\beta$ 1 mRNA as markers for disease activity  
572 and progression in anti-glomerular basement membrane nephritis. *Nephrology*  
573 *Dialysis Transplantation* **32**, 1818–1830 (2017).
- 574 17. Lin, X. *et al.* Role of MiR-155 Signal Pathway in Regulating Podocyte Injury Induced by  
575 TGF- $\beta$ 1. *Cellular Physiology and Biochemistry* **42**, 1469–1480 (2017).
- 576 18. Brennan, D. C. *et al.* Rabbit antithymocyte globulin versus basiliximab in renal  
577 transplantation. *N Engl J Med* **355**, 1967–77 (2006).
- 578 19. Wang, K., Xu, X. & Fan, M. Induction therapy of basiliximab versus antithymocyte  
579 globulin in renal allograft: a systematic review and meta-analysis. *Clin Exp Nephrol* **22**,  
580 684–693 (2018).
- 581 20. Anglicheau, D. *et al.* Discovery and Validation of a Molecular Signature for the  
582 Noninvasive Diagnosis of Human Renal Allograft Fibrosis. *Transplantation* **93**, 1136–  
583 1146 (2012).
- 584 21. Galichon, P. *et al.* Urinary mRNA for the Diagnosis of Renal Allograft Rejection: The  
585 Issue of Normalization. *American Journal of Transplantation* **16**, 3033–3040 (2016).
- 586 22. Taverna, F. *et al.* BIOMEX: an interactive workflow for (single cell) omics data  
587 interpretation and visualization. *Nucleic Acids Res* **48**, W385–W394 (2020).
- 588 23. Chiesa, M., Colombo, G. I. & Piacentini, L. DaMiRseq—an R/Bioconductor package for  
589 data mining of RNA-Seq data: normalization, feature selection and classification.  
590 *Bioinformatics* **34**, 1416–1418 (2018).
- 591 24. Haas, M. *et al.* The Banff 2017 Kidney Meeting Report : Revised diagnostic criteria for  
592 chronic active T cell – mediated rejection , antibody- mediated rejection , and  
593 prospects for integrative endpoints for next- generation clinical trials. *American*  
594 *Journal of Transplantation* **18**, 293–307 (2018).
- 595 25. Van Loon, E. *et al.* Biological pathways and comparison with biopsy signals and cellular  
596 origin of peripheral blood transcriptomic profiles during kidney allograft pathology.  
597 *Kidney Int* **102**, 183–195 (2022).
- 598 26. Lamarthée, B. *et al.* Transcriptional and spatial profiling of the kidney allograft  
599 unravels a central role for FcyRIII+ innate immune cells in rejection. *Nat Commun* **14**,  
600 (2023).
- 601 27. Menon, R. *et al.* Single cell transcriptomics identifies focal segmental  
602 glomerulosclerosis remission endothelial biomarker. *JCI Insight* **5**, (2020).
- 603 28. Malone, A. F. *et al.* Harnessing Expressed Single Nucleotide Variation and Single Cell  
604 RNA Sequencing to Define Immune Cell Chimerism in the Rejecting Kidney Transplant.  
605 *Journal of the American Society of Nephrology* **31**, 1977–1986 (2020).
- 606 29. Cormican, S. *et al.* Chronic Kidney Disease Is Characterized by Expansion of a Distinct  
607 Proinflammatory Intermediate Monocyte Subtype and by Increased Monocyte  
608 Adhesion to Endothelial Cells. *Journal of the American Society of Nephrology* **34**, 793–  
609 808 (2023).

- 610 30. Shi, T. *et al.* Single-cell transcriptomic analysis of renal allograft rejection reveals  
611 insights into intragraft TCR clonality. *Journal of Clinical Investigation* **133**, (2023).
- 612 31. Lake, B. B. *et al.* An atlas of healthy and injured cell states and niches in the human  
613 kidney. *Nature* **619**, 585–594 (2023).
- 614 32. Stewart, B. J. *et al.* Spatio-temporal immune zonation of the human kidney Europe  
615 PMC Funders Group. **365**, 1461–1466 (2019).
- 616 33. Wigerblad, G. *et al.* Single-Cell Analysis Reveals the Range of Transcriptional States of  
617 Circulating Human Neutrophils. *The Journal of Immunology* **209**, 772–782 (2022).
- 618 34. Shankland, S. J., Smeets, B., Pippin, J. W. & Moeller, M. J. The emergence of the  
619 glomerular parietal epithelial cell. *Nature Reviews Nephrology* vol. 10 158–173  
620 Preprint at <https://doi.org/10.1038/nrneph.2014.1> (2014).
- 621 35. Tran, T. *et al.* In Vivo Developmental Trajectories of Human Podocyte Inform In Vitro  
622 Differentiation of Pluripotent Stem Cell-Derived Podocytes. *Dev Cell* **50**, 102-116.e6  
623 (2019).
- 624 36. Schumacher, A. *et al.* Defining the variety of cell types in developing and adult human  
625 kidneys by single-cell RNA sequencing. *npj Regenerative Medicine* vol. 6 Preprint at  
626 <https://doi.org/10.1038/s41536-021-00156-w> (2021).
- 627 37. Xu, W. *et al.* Mapping of  $\gamma/\delta$  T cells reveals V $\delta$ 2+ T cells resistance to senescence.  
628 *EBioMedicine* **39**, 44–58 (2019).
- 629 38. Marsh, S., Salmon, M. & Hoffman, P. samuel-marsh/scCustomize: Version 2.0.1.  
630 Preprint at <https://doi.org/10.5281/zenodo.10161832> (2023).
- 631 39. Varin, A. Alexis-Varin/RightSeuratTools: RightSeuratTools v1.0.1. Preprint at  
632 <https://doi.org/10.5281/zenodo.12518909> (2024).
- 633 40. Matz, M. *et al.* The regulation of interferon type I pathway-related genes RSAD2 and  
634 ETV7 specifically indicates antibody-mediated rejection after kidney transplantation.  
635 *Clin Transplant* **32**, e13429 (2018).
- 636 41. Hayde, N. *et al.* The Clinical and Molecular Significance of C4d Staining Patterns in  
637 Renal Allografts. *Transplantation* **95**, (2013).
- 638 42. Feng, X. *et al.* Rabbit ATG but not horse ATG promotes expansion of functional  
639 CD4+CD25highFOXP3+ regulatory T cells in vitro. *Blood* **111**, 3675–3683 (2008).
- 640

641

642

643 **Acknowledgements**

644 We thank the clinical centers of the BIOMARGIN consortium, the clinicians, surgeons, nursing  
645 staff and patients. Our special thanks go to Ricard Argelaguet, the first author of the original  
646 article describing MOFA, for his precious help with feature selection. We also thank Dr David  
647 Hildeman, Krishna Roskin and Babacar Ndao for their help with scRNA-Seq alignment.

648

649

650 **Author Contributions Statement**

651 C.T., B.L. and M.N. conceived and designed the study. C.T. performed the MOFA analyses.  
652 A.V., C.T and B.L. generated, analysed and interpreted the single cell and bulk transcriptomic  
653 data. C.T, D.A., J.C., H.D.L., W.G., P.M., M.R., V.S., E.V.L. and M.N. included patients and  
654 collected and interpreted the clinical data. C.T., B.L. and M.N. wrote the paper with contribution  
655 from all co-authors.

656

657 **Competing Interests Statement**

658 The authors declare no conflict of interest regarding this manuscript.

659

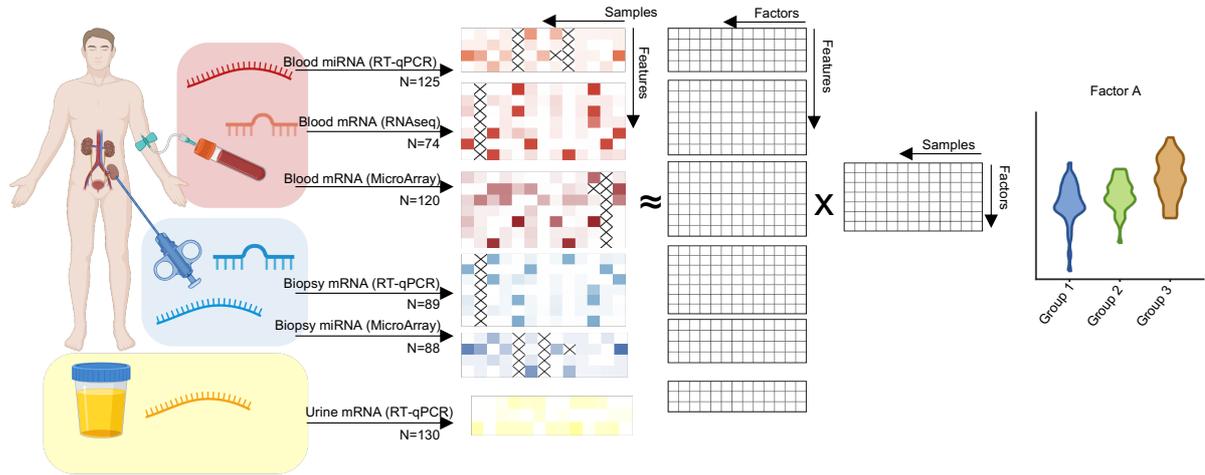
660

661

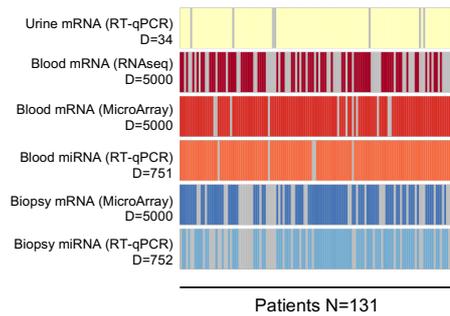
662 **Figures**

a.

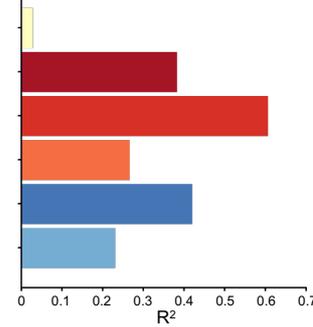
Fig.1



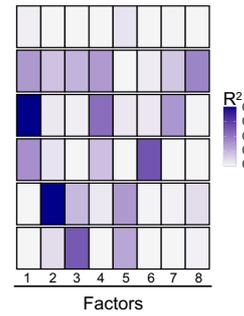
b.



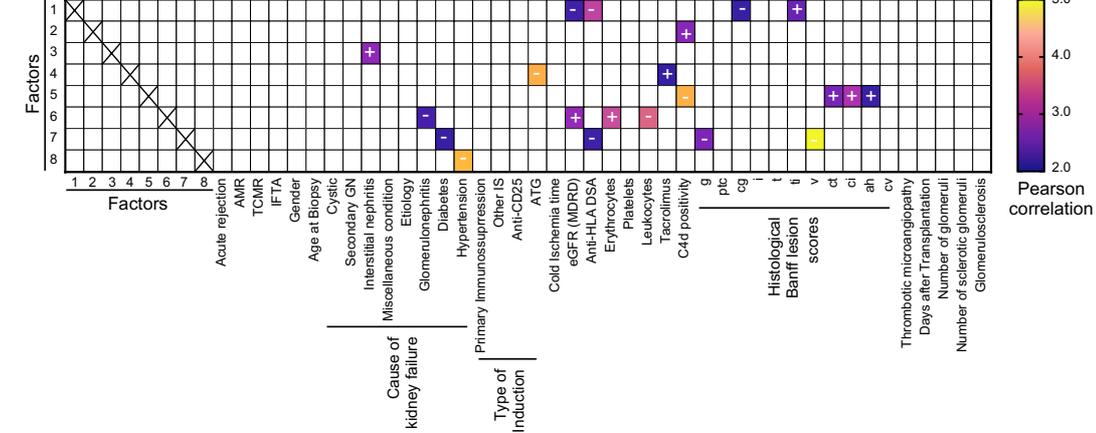
c.



d.



e.



663

664

665

666

667

668

669

670

671

672

673

674

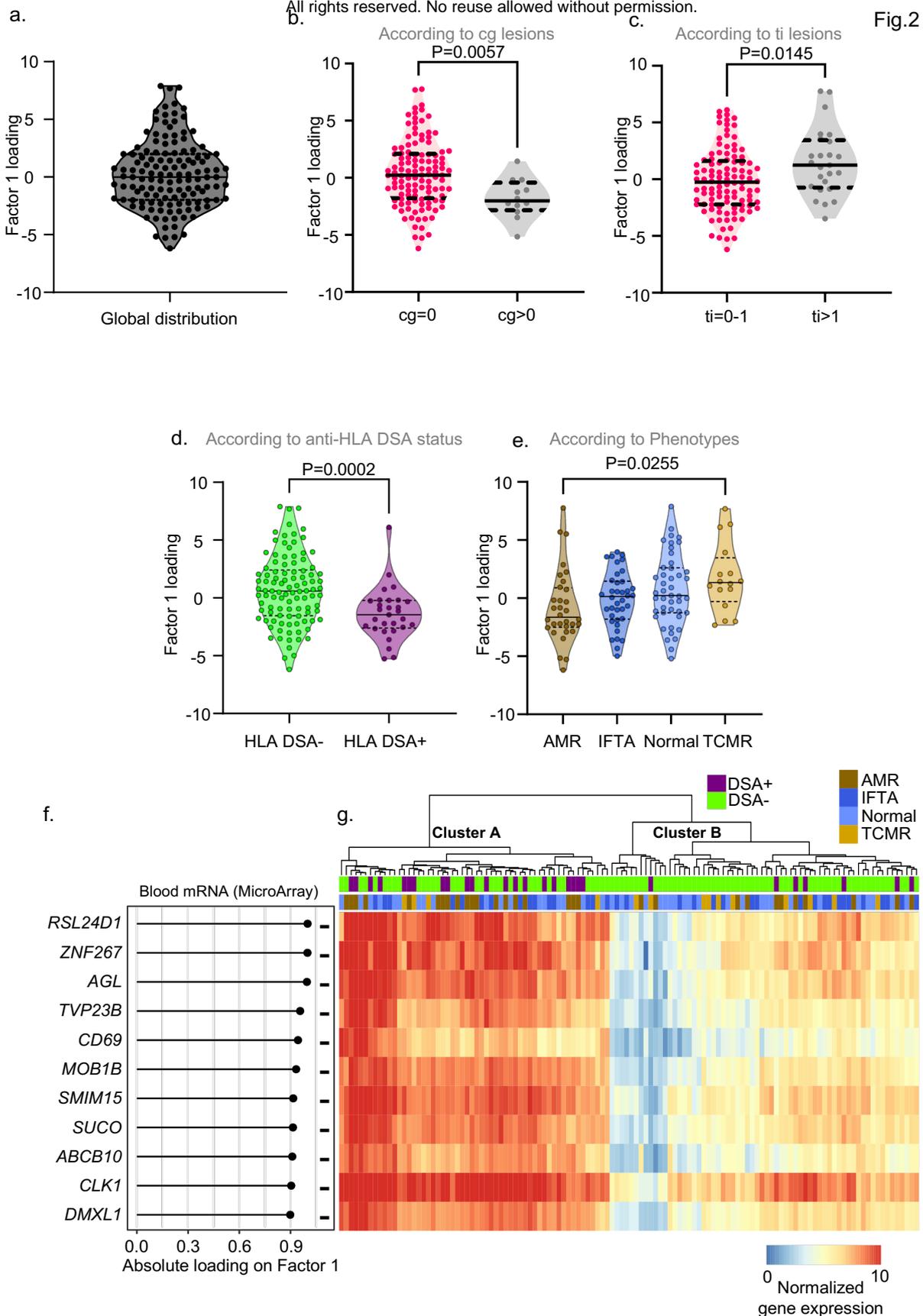
675

676

677

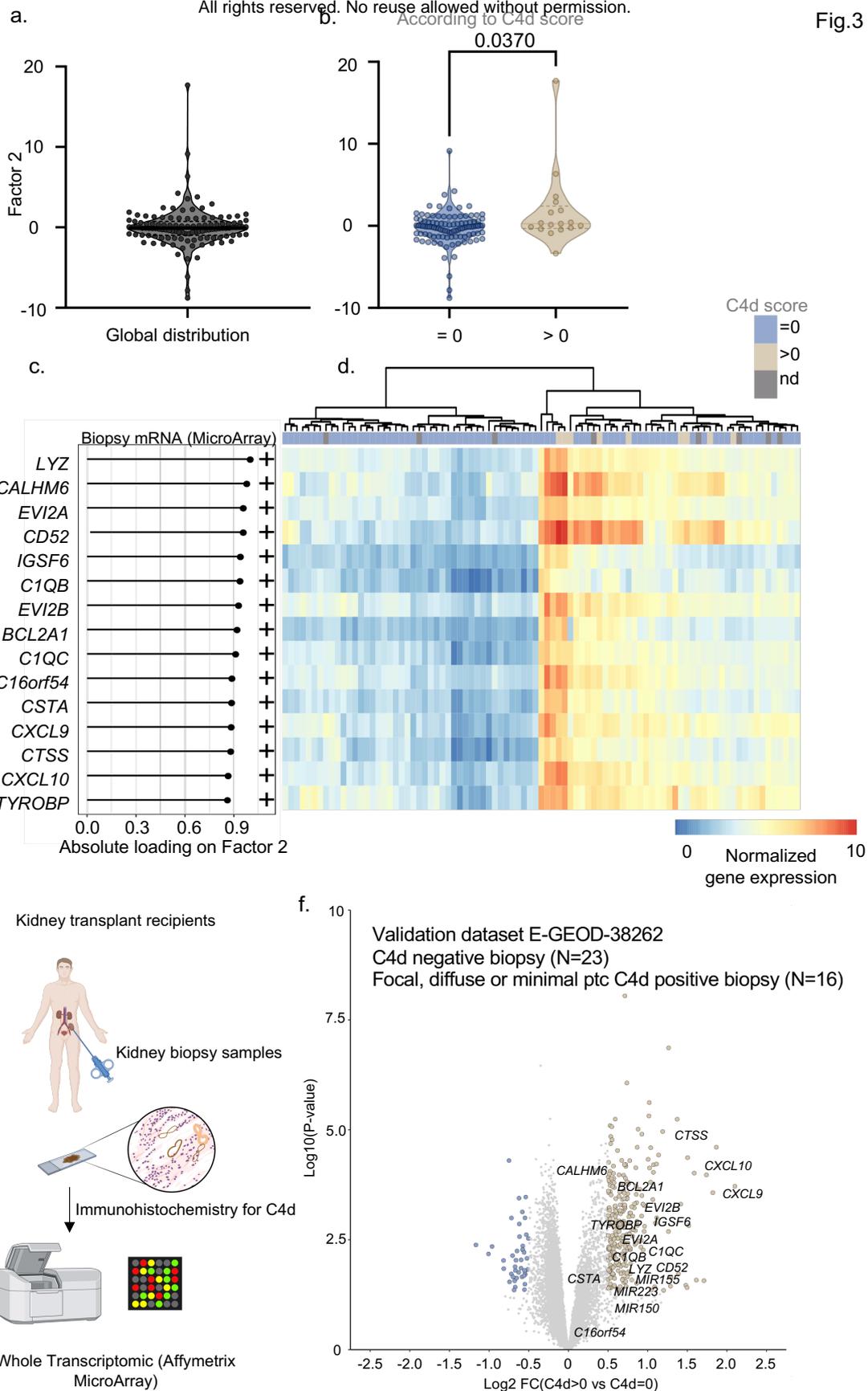
678

**Figure 1: Multi-Omics Factor Analysis application on BIOMARGIN datasets identifies important clinical markers in kidney transplantation**  
 (a) Graphical representation of MOFA's matrix decomposition of each view's data into a product composed of a view-specific factor loadings matrix and a shared latent factor matrix. The number of samples ("N") used for multi-omics factor analysis per data set is indicated. The loading of a given factor can be compared a posteriori between patient groups. (b) Number of features ("D") in each view. A grey bar indicates that the sample is missing in the given omic view (c) Total percentage of variance explained (R<sup>2</sup>) per omic view (d) Percentage of variance explained by each latent factor in the different omic views for the selected MOFA model. (e) Pearson correlation matrix analysis of the 8 latent factors and the various clinical parameters indicated. Positive (+) or negative (-) correlations with the loading of each latent factor are indicated, as well as the log<sub>10</sub> p-value resulting from the Pearson correlation tests. Panel a was created using Biorender.com.

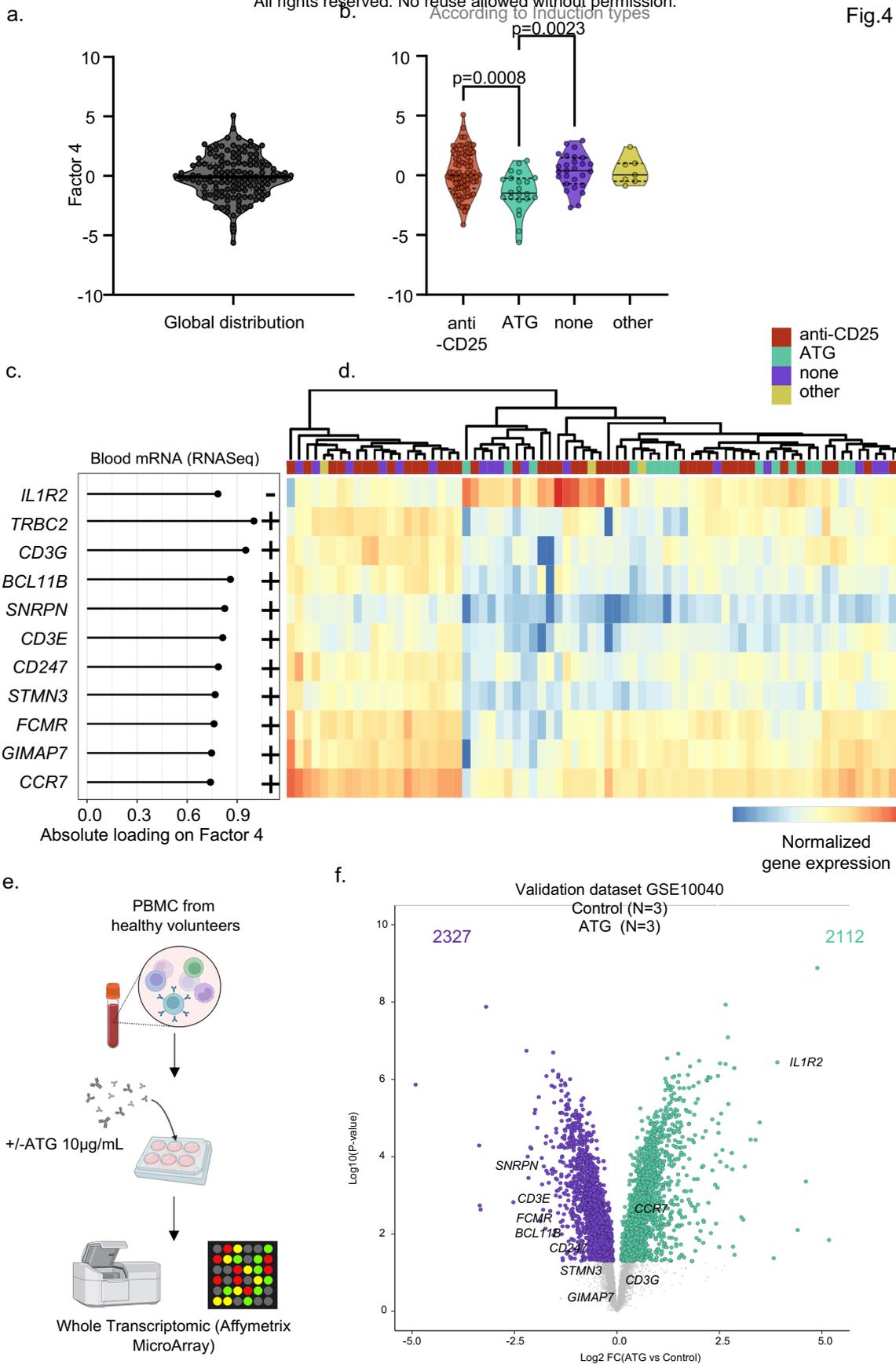


679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

**Figure 2: Factor 1 discriminates antibody-mediated from T-cell mediated rejection**  
(a-e) Violin plots representing the distribution of latent Factor 1 (a) Global distribution (b) According to Banff cg histological lesions. The difference between groups was assessed by a two-tailed Mann-Whitney test (c) According to Banff ti histological lesions. The difference between groups was assessed by a two-tailed Mann-Whitney test (d) According to HLA status. The difference between groups was assessed by a two-tailed Mann-Whitney test (e) According to phenotypes. The difference between groups was assessed by an ordinary one-way ANOVA test and multiple comparisons using the Tukey's test. (f) Lollipop graph shows the top-weighted genes derived from the MicroArray blood dataset in latent Factor 1. (g) Heatmap showing the distribution of the top-weighted genes in the samples. A and B clusters A were determined here in an unsupervised manner.



**Figure 3: Factor 2 is associated with complement/monocytes crosstalk**  
 (a-b) Violin plots representing the distribution of latent Factor 2 (a) Global distribution (b) According to C4d positivity. The difference between groups was assessed by a two-tailed Mann-Whitney test (c) Lollipop graph shows the top-weighted genes derived from the MicroArray biopsy dataset in latent Factor 2. (d) Heatmap showing the distribution of the top-weighted genes explaining Factor 2 in the samples. (e-f) An external validation dataset E-GEOD-38262 was used to confirm the increase of the top-weighted features explaining Factor 2 in C4d positive biopsies. (e) Experimental scheme (f) Volcano plot showing the differentially expressed genes and miRNAs in C4d positive biopsies (N=16) compared to C4d negative biopsies (N=23). The top-weighted features explaining Factor 2 are indicated. Panel e was created using Biorender.com.



695  
696  
697  
698  
699  
700  
701

**Figure 4: Factor 4 captures the impact of induction prophylaxis**

(a-b) Violin plots representing the distribution of latent Factor 4 (a) Global distribution (b) According to induction types. The difference between groups was assessed by a two-tailed Mann-Whitney test (c) Lollipop graph shows the top-weighted genes derived from the RNAseq blood dataset in latent Factor 4. (d) Heatmap showing the distribution of the top-weighted genes of Factor 2 in the samples. (e-f) An external validation dataset GSE10040 was used to confirm the decrease of the top-weighted features explaining Factor 4 in PBMC treated with ATG. (e) Experimental scheme (f) Volcano plot showing the differentially expressed genes in PBMC treated with ATG (N=3) compared to control (N=3). The top-weighted features explaining Factor 4 are indicated. Panel e was created using Biorender.com. ATG, anti thymo globulin ; PBMC, peripheral blood mononuclear cells