Exome Sequencing and Prediction of Long-Term Kidney Allograft Function

Authors: L. Mesnard[1,7], T. Muthukumar[2,3], M.Burbach[2], C.Li[2], H. Shang[4], D. Dadhania[2,3], J. R. Lee[2,3], V. K. Sharma[2], J. Xiang[4] , C. Suberbielle[8] , M. Carmagnat[8] N. Ouali[7], E. Rondeau[7], J.J. Friedewald[5], M. M. Abecassis[6], M. Suthanthiran*[2,3], F. Campagne*[1]

Affiliations:

[1] The HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Weill Cornell Medical College, New York, NY, United States of America; Department of Physiology and Biophysics, The Weill Cornell Medical College, New York, NY, United States of America

[2] Division of Nephrology and Hypertension, Weill Cornell Medical College, New York, NY, United States of America

[3] Department of Transplantation Medicine, New York Presbyterian Hospital, New York, NY, United States of America

[4] Genomics Core Facility Weill Cornell Medical College

[5] Northwestern University Feinberg School of Medicine, Chicago, IL, United States of America

[6] Comprehensive Transplant Center, Northwestern University Feinberg School of Medicine, Chicago, IL, United States of America

[7] INSERM UMR1155 et Service des Urgences Néphrologiques et Transplantation Rénale, APHP, Hôpital Tenon, 75020 Paris, France

[8] Laboratoire d'histocompatibilité Hôpital Saint Louis APHP, Paris, France

*To whom correspondence should be addressed: Fabien Campagne, Ph.D. (fac2003@campagnelab.org), Manikkam Suthanthiran, M.D, PhD (msuthan@med.cornell.edu).

One Sentence Summary: Prediction of long-term kidney graft function with exome sequencing

Abstract (<250 words)

Current strategies to improve graft outcome following kidney transplantation consider information at the HLA loci. Here, we used exome sequencing of DNA from ABO compatible kidney graft recipients and their living donors to determine recipient and donor mismatches at the amino acid level over entire exomes. We estimated the number of amino acid mismatches in transmembrane proteins, more likely to be seen as foreign by the recipient's immune system, and designated this tally as the allogenomics mismatch score (AMS). The AMS can be measured prior to transplantation with DNA for potential donor and recipient pairs. We examined the degree of relationship between the AMS and post-transplantation kidney allograft function by linear regression. In a discovery cohort, we found a significant inverse correlation between the AMS and kidney graft function at 36 months post-transplantation (n=10 recipient/donor pairs; 20 exomes) (r^2>=0.57, P<0.05). The predictive ability of the AMS persists when the score is restricted to regions outside of the HLA loci. This relationship was validated using an independent cohort of 24 recipient donor pairs (n=48 exomes) (r^2>=0.39, P<0.005). In an additional cohort of living and mostly intra-familial recipient/donor pairs (n=19, 38 exomes), we validated the association after controlling for donor age at time of transplantation. Finally, a model that controls for donor age, HLA mismatches and time post-transplantation yields a consistent AMS effect across these three independent cohorts (P<0.05). Taken together, these results show that the AMS is a strong predictor of long-term graft function in kidney transplant recipients.

53 **Introduction**

54

55 Survival of patients afflicted with End Stage Renal Disease (ESRD) is superior following kidney
56 transplantation compared to dialysis therapy. The short-term outcomes of kidney grafts have
57 steadily improved since the early transplants (performed in the 1960s) with refinements in
58 immunosuppressive regimens, use of DNA-based HLA typing, and better infection prophylaxis
59 *(1–3)*. Despite these advances, data collected across the USA and Europe show that 40-50% of
60 kidney allografts fail within ten years of transplantation *(4)*. This observation strongly suggests
61 that as yet uncharacterized factors, including genomic loci, may adversely impact long-term
62 post-transplantation outcomes.

63 Observational studies have demonstrated the importance of matching for the HLA-determined
64 proteins on kidney graft outcome. Therefore, in many countries, including the USA, donor
65 kidney allocation algorithms includes consideration of HLA matching. With widespread
66 incorporation of HLA matching in organ allocation decisions, it has become clearer that HLA
67 mismatching represents an important risk factor for kidney allograft failure but fails to fully
68 account for the invariable decline in graft function and failure in a large number of cases over
69 time. Indeed, only a 15% survival difference exist at 10 years post transplantation between the
70 fully matched kidneys and the kidneys mismatched for both alleles at the HLA-A, B and DR loci
71 *(5)*. These observations suggest that mismatches at non-HLA loci in the genome could influence
72 long-term graft outcomes. The current clinical practice of prescribing life-long
73 immunosuppressive therapy to recipients of fully HLA matched donor kidneys, but not to
74 recipients of monozygotic identical twin kidneys, also suggests a role for non-HLA related
75 genomic factors on graft outcome.

76 While tests of allelic frequencies are a hallmark of genetic research, transplantation has none of
77 the Mendelian characteristics for which genetic tests have been developed. Therefore, the
78 assumption of the Mendelian transmission model seems inadequate to develop predictors of graft
79 function following transplantation. Indeed, previous attempts at using this methodology have
80 identified small genotype effects on graft function in cohorts of hundreds of transplant patients,
81 but often could not be replicated in independent cohorts (reviewed in *(6)*).
82 In this work, we present a new method to estimate the genomic compatibility between the organ
83 graft recipient and donor. This approach, designated as allogenomics in this communication,
84 considers the entire coding sequence of both recipient and donor genomes, as determined by
85 exome sequencing. The allogenomics concept makes it possible to estimate a quantitative
86 compatibility score between the genomes of a recipient and potential donor and is calculated
87 from genotypes and genome annotations available before transplantation. The allogenomics
88 approach does not assume a Mendelian inheritance model but integrates the unique features of
89 transplantation such as the existence of two genomes in a single individual and the recipient's
90 immune system mounting an immune response directed at antigens displayed by the donor
91 kidney. In this report, we show that this new concept helps predict long-term kidney transplant
92 function from the genomic information available prior to transplantation.
93

Results

The allogenomic concept and the allogenomics mismatch score (AMS)

The allogenomics concept is the hypothesis that interrogation of the coding regions of the entire genome for both the organ recipient and organ donor DNA can identify the number of incompatible amino-acids (recognized as non-self by the recipient) that inversely correlates with long-term graft function post transplantation. **Fig. 1A** is a schematic illustration of the allogenomics concept. Because human autosomes have two copies of each gene, we consider two possible alleles in each genome of a transplant pair. To this end, we estimate allogenomics score contributions between zero and two, depending on the number of different amino acids that the donor genome encodes for at a given protein position. **Fig. 1B** shows the possible allogenomics score contributions when the amino acids in question are either an alanine, or a phenylalanine or an aspartate amino acid. The allogenomics mismatch score (AMS) is a sum of amino acid mismatch contributions. Each contribution represents an allele coding for a protein epitope that the donor organ may express and that the recipient immune system could recognize as non-self (see Equation 1 and 2 in **Fig. 1C** and Materials and Methods and full description in the supplementary appendix).

We have developed and implemented a computational approach to estimate the allogenomics mismatch score from genotypes derived for pairs of recipient and donor genomes. (See Materials and Methods for a detailed description of this approach and its software implementation, the allogenomics scoring tool, available at http://allogenomics.campagnelab.org.) Our approach is designed to consider the entire set of protein positions measured by a genotyping assay, or

118 restrict the analysis to a subset of positions *P* in the genome. In this study, we focus on the subset
119 of genomic sites P that encode for amino acids in transmembrane proteins.
120

121 ***The AMS correlates with post-transplantation graft function in living donor-kidney recipient***
122 ***pairs***
123
124
125 In order to test the allogenomics hypothesis, we isolated DNA from 10 kidney graft recipients
126 and their living donors (Discovery Cohort), performed whole exome sequencing and analyzed
127 genotype data for these recipient/donor genome pairs (10 pairs, 20 exomes). These patients were
128 a subset of patients enrolled in a multicenter Clinical Trial in Organ Transplantation-04 (CTOT-
129 04) study of urinary cell mRNA profiling, from whom tissue/cells were collected for future
130 mechanistic studies *(7)*. **Table 1** provides demographic and clinical information about the
131 patients included in the Discovery Cohort. Exome data were obtained with the Illumina TrueSeq
132 exome enrichment kit v3, covering 62Mb of the human genome. Primary sequence data analyses
133 were conducted with GobyWeb *(8)* (data and analysis management), Last *(9)* (alignment to the
134 genome) and Goby *(10)* (genotype calls).
135 Kidney graft function is a continuous phenotype and is clinically evaluated by measuring serum
136 creatinine levels or using estimated glomerular filtration rate (eGFR)*(11)*. In this study, kidney
137 graft function was evaluated at months 12, 24, 36 and/or 48 following transplantation using
138 serum creatinine levels and eGFR, calculated using the 2011 MDRD *(12)* formula. We examined
139 whether the allogenomics mismatch score is associated with post-transplantation allograft
140 function.
141

142 We found positive linear associations between the allogenomics mismatch score and serum
143 creatinine level at 36 months post transplantation (r^2 adj.=0.78, P=0.002, n=10, at 36 months) but
144 not at 12 or 24 months following kidney transplantation (**Fig. 2A, B, C**). We also found a
145 negative linear relationship between the score and eGFR at 36 months post transplantation (r^2
146 adj.=0.57, P=0.02) but not at 12 or 24 months following kidney transplantation (**Fig. 2D, E, F**).
147 These findings suggest that in the discovery cohort the allogenomics score is predictive of long-
148 term graft function, but perhaps not short-term function.
149

150 ***The AMS correlates with graft function in a second, independent cohort of non-related living***
151 ***kidney donor/recipient pairs***
152
153 We sought to validate the observation that the AMS is associated with post-transplantation
154 kidney graft function by testing the association in an independent cohort of kidney transplant
155 patients. To this end, we sequenced DNA collected from 24 additional kidney donor/recipient
156 pairs (see **Table 1** for information about subjects included in the Validation cohort). DNA
157 sequencing was performed using the Agilent Haloplex assay covering 37Mb of the coding
158 sequence of the human genome. We called the genotypes and estimated the AMS as described
159 for the discovery cohort (see Methods). **Fig. 3** shows that, as observed with the Discovery
160 cohort, the AMS correlates progressively better with kidney graft function at longer times
161 following transplantation. At 36 months post transplantation, a small to moderate positive
162 association was observed between the allogenomics mismatch score and the serum creatinine
163 level (r^2 adj. 0.139, P=0.049) (**Fig. 3C**) and eGFR (r^2 adj. 0.078, P=0.11) (**Fig. 3G**). The
164 association between the score and graft function was stronger and reached significance at 48
165 months post transplantation for both creatinine level (r^2 adj. 0.394, P<0.01) (**Fig. 3D**), and eGFR

166 (r^2 adj. 0.284, P=0.02) (**Fig. 3H**), further validating the association in the Validation cohort. In
167 order to test whether models trained on one cohort would generalize to another similar cohort,
168 we trained models on the Discovery cohort and used the fixed model to predict graft function in
169 the Validation cohort. **Fig. S1** shows that such a fixed model does generalize when presented
170 with new recipient-donor pairs, and also exhibited better fit to the longer 48-month time-point
171 compared to the earlier time point (**Fig. S1B** vs. **Fig. S1C**). Similarly, models trained on the
172 Validation cohort generalize to the Discovery cohort (**Fig. S2**). These results establish that the
173 parameters of the models are stable, despite the relatively small numbers of kidney recipient-
174 donor pairs included in the Discovery and Validation cohorts.
175
176 ***The AMS weakly correlates with graft function in a third cohort***
177 In order to further test the strength of the relationship between AMS and graft function we
178 applied the approach to a third independent cohort. To this end, we assembled a new cohort
179 composed mostly of living related kidney donor pairs. We used a peculiarity of the French
180 transplant system, directed by the French national agency for organ procurement transplant,
181 which until recently did not allow living non-related kidney transplantation, to select 19
182 additional pairs from one French center. Demographic and clinical data for this cohort are shown
183 in **Table 1**. As expected in this situation, the range of the AMS is lower than in the other two
184 cohorts, ranging from 349 to 811 in the well-matched cohort (mean 559 ± 147). This range was
185 700 to 1630 (mean 1094 ± 259) in the validation cohort and 994 to 2033 in the discovery cohort
186 (mean 1335 ± 304). These data suggest that the effective range of variation of the AMS is
187 approximately 300-2000 (1700 AMS units).
188 In this third cohort, we did not find the simple association between the AMS and graft function
189 that we observed in the first two cohorts. However, after correction for the strong effect of donor

190 age, we observe a trend in the same direction as the association observed in the first two cohorts
191 at 24, 36 and 48 months when predicting eGFR (e.g., AMS effect estimate=-0.021798 at 36
192 months). The three cohorts combined yield a statistically significant effect of AMS on graft
193 function at 36 months (n=48 pairs, AMS effect estimate=-0.015808, P<0.01) and 60 months
194 (n=13 pairs, AMS effect estimate=-0.04479, P<0.01), and borderline at 48 months (n=32 pairs,
195 AMS effect estimate=-0.016685, P=0.07).

196
197 ***Analysis of the relationship between the AMS and the number of HLA mismatches***
198

199 HLA mismatches are well-described risk factors for kidney graft failure. Therefore given data
200 obtained in the first two cohorts , we next focused our attention on the association between AMS
201 and number of ABDR HLA mismatches. **Fig. 4** presents an analysis where we combined the
202 Discovery and Validation cohorts (32 transplant kidney recipient-donor pairs) and compared the
203 AMS to the number of mismatches at the HLA-A, B and DR loci. We find that the AMS was
204 moderately correlated with the number of mismatch at the HLA loci (**Fig. 4A**, r^2 adj.=0.35,
205 P<0.001). However, the number of HLA mismatches correlates poorly with an AMS estimated
206 from exome data when restricting the sites to the HLA A, B and DR loci (**Fig. 4B**, r^2 adj.=0.09,
207 P=0.047). Furthermore, the AMS estimated outside of the HLA A, B and DR loci still associates
208 significantly with serum creatinine levels (**Fig. 4C**, r^2 adj.=0.36, P<0.001 and eGFR (**Fig. 4D**, r^2
209 adj.=0.18, P=0.011). These data indicate that the ability of the AMS to predict future graft
210 function is mostly independent of the HLA loci.

211

212 ***Final predictive models***
213

214 We fit models across the three combined cohorts to yield final models with fixed parameters:

215 36-month eGFR = 106.311073 -0.015808 AMS -0.749321*Donor_Age

216 48-month eGFR = 108.553088 -0.016685*AMS -0.798154*Donor_Age

217 60-month eGFR = 138.60095 -0.04479*AMS -1.14063 *Donor_Age

218 These equations and parameters are provided to enable testing these models on independent
219 cohorts of living transplant pairs genotyped with exome sequencing. Fit parameter values were
220 estimated with MetaR (train linear model statement, version 1.5.0). We note that the fixed
221 parameters of this model can be sensitive to the exact analysis pipeline used to align reads to the
222 genome and to call genotypes and that an objective test of this model would follow the analysis
223 protocols used to analyze data for this report.

224

225 In the models presented so far, we have considered the prediction of graft function separately at
226 different time points. An alternative analysis would consider time since transplantation as
227 another factor that can influence graft function. This is particularly useful when studying cohorts
228 where graft function was assessed at several distinct time points (e.g., in the French cohort,
229 clinical data describes graft function from 1 to 96 months post transplantation, but few time
230 points have observations for all recipients). To implement this alternative analysis, we fit a
231 mixed linear model of the form eGFR ~ donor age at time of transplant + AMS + T + (1|P),
232 where T is the time post-transplantation, measured in months, and (1|P) a random effect which
233 models separate model intercepts for each donor/recipient pairs. To determine the effect of AMS
234 on eGFR, we compare the fit of models that do or do not include AMS. We find that the effect of
235 AMS is significant (P=0.0042, $\chi 2$= 8.1919, d.f.=1). A similar result is obtained if HLA is used in
236 covariate in the model (i.e., eGFR ~ donor age at time of transplant + AMS + T + HLA + (1|P),
237 comparing model with AMS or without, P= 0.038, $\chi 2$= 4.284, d.f.=1). In contrast, models that

include AMS, but do or do not include the number of ABDR HLA mismatches fit the data equally well (P= 0.60, $\chi 2$= 0.2737, d.f.=1), confirming that the effect of AMS is independent of the number of HLA mismatches.

Table 2 presents confidence intervals for the parameters of the full model (eGFR ~ donor age at time of transplant + AMS + T + HLA + (1|P)) as well as the effective range of the model predictors. The table shows the expected impact of each predictor on eGFR when this predictor is varied over its range, assuming all other predictors are kept constant. For instance, we assume that donor age at time of transplant varies from 20 years old to 80 years old (range: 60). Across this range, eGFR will decrease by an estimated 28 units as the donor gets older. The AMS effect has an effective range of 1,700 and the corresponding eGFR decrease is 19 units.

Table 2. Estimated Model Parameters, 95% Confidence Intervals and Expected Impact on eGFR.

	Estimate				Impact on eGFR			
Model Coefficient	**Fit**	**2.50%**	**97.50%**	**Effective Range**	**Fit**	**2.50%**	**97.50%**	**Note**
Time post transplantation (in months)	-0.2439	-0.32348868	-0.1644768	480	-117.072	-155.2745664	-78.948864	*
Donor age at transplant (in years)	-0.46843	-0.78098518	-0.15179	60	-28.1058	-46.8591108	-9.1074	*
AMS	-0.01141	-0.02217172	-0.000627551	1700	-19.397	-37.691924	-1.06683704	*
HLA_ABDR_mismatches	-0.57055	-2.73627999	1.606998	6	-3.4233	-16.41767994	9.641988	**n.s.**

Impact of genotyping platform on the estimation of the AMS

We studied the impact of the genotyping platform on the estimation of the AMS (**Fig. S3**). Large cohorts of matched recipient and donor DNA are being assembled and genotyped with SNP chip array technology such as the Illumina 660W bead array platform *(13)*. We asked whether such platforms would be appropriate to validate the allogenomics model in large cohorts. **Fig. S3A** documents the number of sites that contribute to the allogenomics score on each platform. **Fig. S3B** indicates that the exome assay captures many more sites with rare polymorphisms (minor allele frequency <5%) than the GWAS array platform. This is expected because exome assays directly sequence an individual DNA, while GWAS platforms are designed with a fixed set of

260 polymorphisms and will not include many of the rare polymorphisms any given individual may
261 carry. **Fig. S3C** compares the correlations measured with the exome assay or that could have
262 been obtained if we had measured the allogenomics mismatch score with the Illumina 660W
263 assay. The weak correlations obtained suggest that GWAS platforms are not ideal for future tests
264 of the allogenomics model.
265

266 **Discussion**

267

268 Several Donor/Recipient matching factors have been identified prior to this study as important
269 for transplantation. For instance, blood group compatibility is a prerequisite unless pre-
270 conditioning of the recipient is undertaken to facilitate blood group incompatible kidney
271 transplantation. While HLA compatibility is a necessary requirement for successful bone marrow
272 transplants, full HLA compatibility is not an absolute prerequisite for all types of transplantations
273 as indicated by the thousands of solid organ transplants performed yearly despite lack of full
274 matching between the donor and recipient at the HLA-A, B and DR loci. In view of better patient
275 survival following transplantation compared to dialysis, kidney transplants are routinely
276 performed with varying degrees of HLA-mismatches, including HLA mismatches for all HLA-
277 class I and II antigens. Although, graft outcomes improve with better HLA-matching, excellent
278 long-term graft outcomes with stable graft function have been observed in patients with full
279 ABDR HLA mismatches. The success of these transplants clearly suggests that factors other than
280 HLA compatibility may influence the long-term clinical outcome of kidney allografts.

281

282 Case-control designs are appropriate when studying phenotypes that are expected to be
283 associated with genotypes that follow a Mendelian inheritance mechanism. We reason that
284 transplant patients are not ideal subjects for this experimental design. Indeed, patients who
285 received a kidney transplant have two genomes in their body: their germline DNA, and the DNA
286 of the donor. In cases when the transplanted kidney was from an unrelated donor (e.g., organs
287 from deceased donors), it is clear that a Mendelian genetic transmission mechanism is not at
288 play. Importantly, even in cases where the donor is one of the parents of the transplant recipient
289 (familial, living related transplant), the genome of the parent will break the assumptions of

290 Mendelian inheritance. Because the transplant recipient has two genomes after transplant, it is
291 not appropriate to assume that genomic markers can be identified when assuming a Mendelian
292 inheritance process. Yet, this assumption has been made in most of the transplantation genomic
293 studies published to date.
294
295 The allogenomics concept that we present in this manuscript postulates a different mechanism
296 for the development of the immune response in the transplant recipient: immunological and
297 biophysical principles strongly suggest that alleles present in the donor genome, but not in the
298 recipient genome, will have the potential to produce epitopes that the recipient immune system
299 will recognize as non-self. This reasoning explains why the allogenomics score is not equivalent
300 to the genetic measures of allele sharing distance that have been used to perform genetic
301 clustering of individuals *(14)*.
302
303 Results presented in this manuscript suggest that allogenomic mismatches in proteins expressed
304 at the surface of donor cells could explain why some recipients' immune systems mount an
305 attack against the donor organ, while other patients tolerate the transplant for many years, when
306 given similar immunosuppressive regimens. If the results of this study are confirmed in
307 additional independent transplant cohorts (renal transplants, solid or hematologic transplants),
308 they may prompt the design of prospective clinical trials to evaluate whether allocating organs to
309 recipients with a combination of low allogenomics mismatch scores and low HLA mismatch
310 scores improves long term graft outcome. A positive answer to this question could profoundly
311 impact the current clinical and regulatory framework for assigning organs to ESRD patients.
312

313 In this study, we introduced the allogenomics concept to quantitatively estimate the
314 histoincompatibility between an organ donor from an ABO compatible living donor and recipient
315 outside of the HLA locus. We tested the simplest model derived from this concept to calculate an
316 allogenomics mismatch score (AMS). We demonstrated that the AMS, which can be estimated
317 before transplantation, helps predict post-transplantation kidney graft function. Interestingly, the
318 strength of the correlation increases with the time post transplantation, an observation we make
319 in both the discovery and validation cohorts. When testing the approach in a cohort of closely
320 related pairs (transplantation mostly from related living donors), we observed that controlling for
321 donor age at time of transplant was necessary to validate the association between eGFR and
322 AMS. Donor age at time of transplant is a well-documented predictor of long-term graft function
323 *(15)*. An alternative model where we control for donor age and time post-transplantation found a
324 consistent effect of AMS across all three cohorts.
325
326 We chose to focus this study on living ABO compatible (either related or non-related) donors
327 because kidney transplantation can be planned in advance and because differences in cold
328 ischemia times and other covariates common in deceased donor transplants are negligible when
329 focusing on living donors, especially in a small cohort.
330
331 The selection criteria for cadaveric donors include consideration of HLA matching and of the age
332 of the donor. Compared to related donors we expect that the range of the AMS will be
333 comparable to that in the discovery cohort (composed in majority of unrelated donors),
334 Therefore, we expect that the allogenomics approach would also be predictive in cadaveric
335 deceased donors, given that the AMS score should be high in these pairs. However, since many

336 factors can independently influence graft function after transplantation from a cadaveric donor
337 (e.g. cold ischemia time), potentially much larger cohorts may be required in such settings to
338 achieve sufficient power to adequately control for the covariates relevant to cadaveric donors and
339 to detect the allogenomics effect.
340
341 While several case-control studies have been conducted with large organ transplant cohorts, the
342 identification of genotype/phenotype associations has been limited to the discoveries of
343 polymorphisms with small effect, that have been reviewed in *(16)*, which have often not been
344 replicated *(17–19)*. Such studies have observed small average effects measured across groups of
345 transplants when our study is measuring moderate effects in individual transplants. Rather than
346 focusing on specific genomic sites, the allogenomics concept sums contributions of many
347 mismatches that can impact protein sequence and structure and could yield an immune response
348 in the recipient.
349
350 While we have not attempted to optimize the set of sites considered to estimate the allogenomics
351 mismatch score, it is possible that reduced and more focused subsets could increase the
352 predictive ability of the score. For instance, the AMS could be applied to look for clusters of
353 allogenomic mismatch sites in genes outside of the HLA loci. Such studies would require larger
354 cohorts and may enable the discovery of loci enriched in allogenomics mismatches responsible
355 for a part of the recipient alloresponse against yet unsuspected donor antigens. Their discovery
356 might foster the development of new immunosuppressive agents targeting the expression of these
357 immuno-dominant epitopes.

358 On the other hand, it is also possible that most polymorphisms that contribute to the score have a
359 low frequency in the population (e.g., minor allele frequency less than 5%), which would make
360 the identification of common sites of mismatches unlikely.
361
362 In this study, we propose several linear models that use the AMS to predict graft function after
363 transplantation. A model that controls for donor age and time post transplantation, and includes
364 the AMS, achieves an r^2 of 0.22 and a RMSE of 15.49 when predicting MDRD eGFR. This
365 model makes it possible to envision predicting graft function at an arbitrary future time after
366 transplantation using genomic information available before transplantation. However, several
367 multi-center independent validations will be essential to establish if prospective clinical trials are
368 warranted. We recommend focusing such validation efforts on transplant pairs that are not
369 familial, where the AMS effect appears to be maximal. We distribute the software that we
370 developed to estimate the allogenomics mismatch score to facilitate further studies by others (see
371 http://allogenomics.campagnelab.org).
372
373

374 **Materials and Methods**

375

376 The study was reviewed and approved by the Weill Cornell Medical College Institutional
377 Review Board (protocol #1407015307 "Predicting Long-Term Function of Kidney Allograft by
378 Allogenomics Score", approved 09/09/2014). The second study made on the French cohort was
379 approved by the Comité de Protection des Personnes (CPP), Ile de France 5, (02/09/2014). Codes
380 were used to ensure donor and recipient anonymity. All subjects gave written informed consent.
381 Living donor ABO compatible kidney transplantations were performed according to common
382 immunological rules for kidney transplantation with a mandatory negative IgG T-cell and B-cell
383 complement-dependent cytotoxicity cross-match.

384 **Whole exome sequencing and genotyping**

385 Briefly, genotypes of donors and recipients were assayed with exome sequencing (Illumina
386 TruSeq enrichment kit for the Discovery Cohort and Agilent Haloplex kit for the Validation
387 cohort and the French cohort. Reads were aligned to the human genome with the Last *(9)* aligner
388 integrated as a plugin in GobyWeb *(8)*. Genotype calls were made with Goby *(10)* and GobyWeb
389 *(8)*. Prediction of polymorphism impact on the protein sequence were performed with the Variant
390 Effect Predictor *(20)*. Genes that contain at least one transmembrane segment were identified
391 using Ensembl Biomart *(21)*.

392

393 **Discovery cohort: Transplant recipients and DNA samples**

394 We selected 10 kidney transplant recipients from those who had consented to participate in the
395 Clinical Trials in Organ Transplantation-04 (CTOT-04), a multicenter observational study of
396 noninvasive diagnosis of renal allograft rejection by urinary cell mRNA profiling. We included

397 only the recipients who had a living donor transplant and along with their donors, had provided
398 informed consent for the use of their stored biological specimens for future research. Their
399 demographic and clinical information is shown in Table 1. DNA was extracted from stored
400 peripheral blood using the EZ1 DNA blood kit (Qiagen®) based on the manufacturer's
401 recommendation.
402
403 **Discovery cohort: Whole exome sequencing**
404 DNA was enriched for exome regions with the TruSeq exome enrichment kit v3. Sequencing
405 libraries were constructed using the Illumina TruSeq kit DNA sample preparation kit. Briefly,
406 1.8 µg of genomic DNA was sheared to average fragment size of 200 bp using the Covaris E220
407 (Covaris, Woburn, MA, USA). Fragments were purified using AmpPureXP beads (Beckman
408 Coulter, Brae, CA, USA) to remove small products (<100 bp), yielding 1 µg of material that was
409 end-polished, A-tailed and adapter ligated according to the manufacturer's protocol. Libraries
410 were subjected to minimal PCR cycling and quantified using the Agilent High Sensitivity DNA
411 assay (Agilent, Santa Clara, CA, USA). Libraries were combined into pools of six for solution
412 phase hybridization using the Illumina (Illumina, San Diego, CA, USA) TruSeq Exome
413 Enrichment Kit. Captured libraries were assessed for both quality and yield using the Agilent
414 High Sensitivity DNA assay Library Quantification Kit. Sequencing was performed with six
415 samples per lane using the Illumina HiSeq 2000 sequencer and version 2 of the sequencing-by-
416 synthesis reagents to generate 100 bp single-end reads (1×100SE).
417
418 **Validation cohort: Transplant recipients and DNA samples**
419 We studied 24 kidney transplant recipients who had a living donor transplant at the NewYork-
420 Presbyterian Weill Cornell Medical Center. This was an independent cohort and none of the

421 recipients had participated in the CTOT-04 trial. Recipients were selected randomly based on the
422 availability of archived paired recipient-donor DNA specimens obtained at the time of
423 transplantation at our Immunogenetics and Transplantation Laboratory. The Institutional Review
424 Board at Cornell approved the study. DNA extraction from peripheral blood was done using the
425 EZ1 DNA blood kit (Qiagen®) based on the manufacturer's recommendation.
426
427 **French cohort: Transplant recipients and DNA samples**
428 We studied 19 kidney transplant recipients who had a living donor transplant at Tenon Hospital.
429 This represented a third independent cohort. Recipients were selected randomly based on the
430 availability of archived paired recipient-donor DNA specimens obtained either at the Laboratoire
431 d'histocompatibilité, Hôpital Saint Louis APHP, Paris or during patient's follow-up between
432 October 2014 and January 2015. DNA extraction from peripheral blood was done using the
433 Nucleospin blood L kit (Macherey-Nagel®) based on the manufacturer's recommendation.
434
435 **Validation and French cohorts: Whole exome sequencing**
436 The Validation and French cohorts were both assayed with the Agilent Haloplex exome
437 sequencing assay. The Haloplex assay enriches 37 Mb of coding sequence in the human genome
438 and was selected for the validation cohort because it provides a strong and consistent exome
439 enrichment efficiency for regions of the genome most likely to contribute to the allogenomics
440 contributions in protein sequences. In contrast, the TrueSeq assay (used for the Discovery
441 Cohort) enriches 63Mb of sequence and includes regions in untranslated regions (5' and 3'
442 UTRs), which do not contribute to allogenomics scores and therefore do not need to be
443 sequenced to estimate the score. Libraries were prepared as per the Agilent recommended

444 protocol. Sequencing was performed on an Illumina 2500 sequencer with the 100bp paired-end
445 protocol recommended by Agilent for the Haloplex assay. Libraries were multiplexed 6 per lane
446 to yield approximately 30 million PE reads per sample.
447
448 **Minor Allele Frequencies of the AMS Sites**
449 We determined the minor allele frequency of sites used in the calculation of the allogenomics
450 mismatch score using data from the NHLBI Exome Sequencing Project (ESP) release
451 ESP6500SI-V2. We downloaded the data file ESP6500SI-V2-SSA137.protein-hgvs-
452 update.snps_indels.txt.tar.gz and extracted MAF in the European American population (EA) and
453 in the African American population (AA) *(22)*. The ESP measured genotypes in a population of
454 6,503 individuals across the EA and AA populations using an exome-sequencing assay*(22)*. This
455 resource made it possible to estimate MAF for most of the variations that are observed in the
456 subjects included in our discovery and validation cohort.
457
458 **Overlap with EVP variants**
459 Of 12,457 sites measured in the validation cohort with an allogenomics contribution strictly
460 larger than zero (48 exomes, sites with contributions across 24 clinical pairs of transplants),
461 9,765 (78%) have also been reported in EVP (6,503 exomes).
462
463 **Sequence Data Analysis**
464 Illumina sequence base calling was performed in the Weill Cornell Genomics Core Facility.
465 Sequence data in FASTQ format were converted to the compact-reads format using the Goby
466 framework [14]. Compact-reads were uploaded to the GobyWeb*(8)* system and aligned to the
467 1000 genome reference build for the human genome (corresponding to hg19, released in

468 February 2009) using the Last *(9, 23)* aligner (parallelized in a GobyWeb *(8)* plugin). Single
469 nucleotide polymorphisms (SNPs) and small indels genotype were called using GobyWeb with
470 the Goby *(24)* discover-sequence-variants mode (parameters: minimum variation support=3,
471 minimum number of distinct read indices=3) and annotated using the Variant Effect Predictor
472 *(20)* (VEP version 75-75.7) from Ensembl. The data were downloaded as a Variant Calling
473 format *(25)* (VCF) file from GobyWeb *(8)* and further processed with the allogenomics scoring
474 tool (see http://allogenomics.campagnelab.org).
475
476 **Estimation of the Allogenomics Mismatch Score (AMS)**
477 The allogenomics mismatch score $\Delta(r,d)$ is estimated for a recipient r and donor d as the sum of
478 score mismatch contributions (see Fig. 1C, Equation 1).
479
480 **Equation 1** (reproduced from Fig. 1C).

481 $$\Delta(r,d) = \sum_{p \in P} \sigma_p [G_{rp} = genotype(r,p), G_{dp} = genotype(d,p)]$$

482 Contributions are observed for each polymorphic site p in a set P, where P is determined by the
483 genotyping assay and analysis methods, and can be further restricted (*e.g.*, to polymorphisms
484 within genes that code for membrane proteins). Score mismatch contributions $\sigma_p(G_{rp},G_{dp})$ are
485 calculated using the recipient genotype G_{rp} and the donor genotype G_{dp} at the polymorphic site p.
486 Here, we consider that a genotype can be represented as a set of alleles that were called in a
487 given genome. For instance, if a subject has two alleles at one polymorphic site, and we denote
488 each allele A or B, the genotype at p is represented by the set {A,B}. This representation is
489 general and sufficient to process polymorphic sites with single nucleotide polymorphisms or
490 insertion/deletions.

491

492 Equation 2 describes how the individual score mismatch contributions are calculated at a
493 polymorphic site of interest.

$$\delta_p(G_{rp}, G_{dp}) = \sum_{a \in G_{dp}} \begin{cases} 0 \text{ if } a \in G_{r,p} \\ 1 \text{ otherwise} \end{cases}$$

494 **Equation 2** (reproduced from Fig. 1C).

495 A contribution of 1 is added to the score for each polymorphic site where the donor genome has
496 an allele (a_{dp}) that is not also present in the recipient genome. When both donor and recipient
497 genome are called at polymorphic site P, no contribution is added. For example, assuming a
498 genomic site where the donor genome has two alleles, *i.e.*, G_{dp}={A,B}, and the recipient genome
499 is homozygote with G_{rp}={A}. In this case, (G_{rp},G_{dp})=1. Fig. 1B presents additional examples of
500 donor and recipient genotypes and indicates the resulting score contribution (the subscript *p* is
501 omitted for conciseness). Score contributions are summed across all polymorphism sites in the
502 set *P* to yield the allogenomic mismatch score (see Fig. 1C Equation 1).

503

504 **Selection of Informative Polymorphisms**

505 The selection of the set of polymorphic sites *P* is important to the effectiveness of the approach.
506 In the current method, we select exonic polymorphic sites that are (1) predicted to create non-
507 synonymous change in a protein sequence, (2) are located in a gene that codes for one or more
508 membrane proteins (defined as any protein with at least one predicted transmembrane segment,
509 information obtained from Biomart *(21)*, Ensembl database 75). Additional filters can be applied
510 to restrict *P*, which may lead to improved prediction of transplant clinical endpoints.
511 Constructing additional filters will require the study of a larger training set of matched recipient

512 and donor genotypes, which currently does not exist. It is possible that such study will indicate
513 that other criteria than (2) also lead to predictive scores.

514 **Implementation: the Allogenomics Scoring Tool**

515 We developed the *allogenomics scoring tool* to process genotypes in the VCF format and
516 produce allogenomics mismatch scores for specific pairs of genomes in the input file. The
517 *allogenomics scoring tool* was implemented in Java with the Goby framework and is designed to
518 read VCF files produced by Goby and GobyWeb. The source code of the allogenomics scoring
519 tools is distributed for academic and non-commercial purposes at
520 http://allogenomics.campagnelab.org. The following command line arguments were used to
521 generate the estimates described in this manuscript and can be run from the
522 Allogenomics_Package file provided in supplementary. The genotype input file(s) necessary to
523 reproduce these results (GobyWeb tags: JEOHQUR (2.3GB), YOOLWXH (83MB)) cannot be
524 distributed through dbGAP (http://www.ncbi.nlm.nih.gov/gap), or an equivalent archive, because
525 the consent form signed by the CTOT-04 participants is not compatible with such distribution of
526 the subject information. A copy of the VCF files can be provided to the editors of the journal
527 upon request should they wish to make it available to the reviewers upon condition of
528 confidentiality during peer-review.

529 Pre-requisite to running the command lines: (1) You must have the Java runtime environment
530 installed on your computer (the software has been tested with version 1.6) (2) You must define
531 the environment variable ALLO to the location where you have downloaded the distribution of
532 the allogemomics scoring tool (3). You must obtain the input VCF files and place them under:
533 ${ALLO}/VCF_files_input/JEOHQUR-stats.vcf.gz ${ALLO}/VCF_files_input/YOOLWXH-
534 stats.vcf.gz.

535 Estimating allogenomics mismatch scores on the Discovery cohort:

```
536 java -Xmx4g -jar allogenomics-1.1.7-scoring-tool.jar \
537         --input ${ALLO}/VCF_files_input/JEOHQUR-stats.vcf.gz \
538         -p ${ALLO}/Pair_files/Discovery_cohort.pairs.tsv \
539         -a Annotation_files/All_protein_coding_Ensembl_75.gtf \
540         --output ${ALLO}/Output/TM-Discovery.tsv \
541         --output-format TSV --only-non-synonymous-coding --vep \
542         --consider-indels --minimum-depth 10 --max-depth 500 \
543         -t ${ALLO}/Annotation_files/TrM-Transcript_Ensembl_75.tsv \
544         --clinical
```

545

546 Estimating allogenomics mismatch scores on the Validation cohort:

```
547 java -Xmx4g -jar allogenomics-1.1.7-scoring-tool.jar \
548         --input ${ALLO}/VCF_files_input/JEOHQUR-stats.vcf.gz \
549         -p ${ALLO}/Pair_files/Validation_cohort.pairs.tsv \
550         -a ${ALLO}/Annotation_files/All_protein_coding_Ensembl_75.gtf \
551         --output ${ALLO}/Output/TM-Validation.tsv \
552         --output-format TSV --only-non-synonymous-coding --vep \
553         --consider-indels --minimum-depth 10 --max-depth 500 \
554         -t ${ALLO}/Annotation_files/TrM-Transcript_Ensembl_75.tsv \
555         --clinical --measured-sites SitesHaloplexExome.tsv
```

556 Estimating allogenomics mismatch scores on the French cohort:

```
557 java -Xmx4g -jar allogenomics-1.1.7-scoring-tool.jar \
```

```
558         --input ${ALLO}/VCF_files_input/YOOLWXH-stats.vcf.gz \
559         -p Pair_files/French_cohort.pairs.tsv \
560         -a ${ALLO}/Annotation_files/All_protein_coding_Ensembl_75.gtf  \
561         --output ${ALLO}/Output/TM-French_cohort.tsv \
562         --output-format TSV --only-non-synonymous-coding \
563         --vep --consider-indels --minimum-depth 10 \
564         --max-depth 500 \
565         -t ${ALLO}/Annotation_files/TrM-Transcript_Ensembl_75.tsv --clinical --no-dash
566
```

567 <u>Estimating allogenomics mismatch scores on merged discovery and validation cohorts:</u>

```
568 java -Xmx4g -jar allogenomics-1.1.7-scoring-tool.jar \
569         --input ${ALLO}/VCF_files_input/JEOHQUR-stats.vcf.gz\
570         -p Pair_files/Discovery+Validation_cohort.pairs.tsv \
571         -a ${ALLO}/Annotation_files/All_protein_coding_Ensembl_75.gtf \
572         --output ${ALLO}/Output/TM-Discovery+Validation.tsv \
573         --output-format TSV --only-non-synonymous-coding \
574         --vep --consider-indels --minimum-depth 10 \
575         --max-depth 500 \
576         -t ${ALLO}/Annotation_files/TrM-Transcript_Ensembl_75.tsv --clinical
577
```

578 Estimating allogenomics mismatch score limited to Illumina GeneChip660W loci on the
579 validation cohort:

```
580 java -Xmx4g -jar allogenomics-1.1.7-scoring-tool.jar \
```

```
--input ${ALLO}/VCF_files_input/JEOHQUR-stats.vcf.gz \
-p ${ALLO}/Pair_files/Validation_cohort.pairs.tsv \
-a ${ALLO}/Annotation_for_660W/Human660W_Gene_Annotation_hg19-ilmn.tsv \
--output ${ALLO}/Output/TM-Validation_Illumina660W.tsv \
--output-format TSV --only-non-synonymous-coding --vep \
--consider-indels --minimum-depth 10 --max-depth 500 \
-t ${ALLO}/Annotation_for_660W/TM-as-gene-names_for_Illumina660W.tsv \
--clinical --measured-sites sites-660W.tsv
```

Statistical Analyses

Analyses were conducted with either JMP Pro version 11 (SAS Inc.) or metaR (http://metaR.campagnelab.org). Figures 2, 3 and 4, as well as SF1B, SF1C, SF2B, SF3C were constructed with metaR analysis scripts and edited with Illustrator CS6 to increase some font sizes or adjust the text of some axis labels. The model that includes the time post-transplantation as a covariate was constructed in metaR and JMP. The R implementation of train linear model uses the lm R function. This model was executed using the R language 3.1.3 (2015-03-09) packaged in the docker image fac2003/rocker-metar:1.4.0 (https://hub.docker.com/r/fac2003/rocker-metar/). Models with random effects were estimated with metaR 1.5.1 and R (train mixed model and compare mixed models statements, which use the lme4 R package*(26)*). Comparison of fit for models with random effects was obtained by training each model alternative with REML=FALSE an performing an anova test, as described in *(27)*.

604

605
606
607

608 **References and Notes:**

609 1. G. Opelz, T. Wujciak, B. Dohler, S. Scherer, J. Mytilineos, HLA compatibility and
610 organ transplant survival. Collaborative Transplant Study, *Rev Immunogenet* **1**, 334–
611 342 (1999).

612 2. N. Angaswamy, V. Tiriveedhi, N. J. Sarma, V. Subramanian, C. Klein, J. Wellen, S.
613 Shenoy, W. C. Chapman, T. Mohanakumar, Interplay between immune responses to
614 HLA and non-HLA self-antigens in allograft rejection., *Hum. Immunol.* **Nov;74**, 1478–
615 1485 (2013).

616 3. J. R. Leventhal, J. M. Mathew, D. R. Salomon, S. M. Kurian, M. Suthanthiran, A.
617 Tambur, J. Friedewald, L. Gallon, J. Charette, J. Levitsky, Y. Kanwar, M. Abecassis, J.
618 Miller, Genomic biomarkers correlate with HLA-identical renal transplant tolerance, *J*
619 *Am Soc Nephrol* **24**, 1376–1385 (2013).

620 4. A. Gondos, B. Döhler, H. Brenner, G. Opelz, Kidney graft survival in Europe and the
621 United States: strikingly different long-term outcomes., *Transplantation* **95**, 267–74
622 (2013).

623 5. N. Sasaki, A. Idica, The HLA-matching effect in different cohorts of kidney transplant
624 recipients: 10 years later., *Clin. Transpl.* , 261–82 (2010).

625 6. P. J. Phelan, P. J. Conlon, M. A. Sparks, Genetic determinants of renal transplant
626 outcome: where do we stand?, *J Nephrol* (2014), doi:10.1007/s40620-014-0053-4.

627 7. M. Suthanthiran, J. E. Schwartz, R. Ding, M. Abecassis, D. Dadhania, B. Samstein,
628 S. J. Knechtle, J. Friedewald, Y. T. Becker, V. K. Sharma, N. M. Williams, C. S. Chang,
629 C. Hoang, T. Muthukumar, P. August, K. S. Keslar, R. L. Fairchild, D. E. Hricik, P. S.
630 Heeger, L. Han, J. Liu, M. Riggs, D. N. Ikle, N. D. Bridges, A. Shaked, Urinary-cell
631 mRNA profile and acute cellular rejection in kidney allografts, *N Engl J Med* **369**, 20–31
632 (2013).

633 8. K. C. Dorff, N. Chambwe, Z. Zeno, M. Simi, R. Shaknovich, F. Campagne, N. J.
634 Provart, Ed. GobyWeb: Simplified Management and Analysis of Gene Expression and
635 DNA Methylation Sequencing Data, *PLoS One* **8**, e69666 (2013).

636 9. S. M. Kielbasa, R. Wan, K. Sato, P. Horton, M. C. Frith, Adaptive seeds tame
637 genomic sequence comparison, *Genome Res* **21**, 487–493 (2011).

638 10. F. Campagne, K. C. Dorff, N. Chambwe, Robinson J.T., J. P. Mesirov, Compression
639 of structured high-throughput sequencing data, *PLoS One* **8**, e79871 (2013).

640 11. U. Poge, T. Gerhardt, H. Palmedo, H. U. Klehr, T. Sauerbruch, R. P. Woitas, MDRD
641 equations for estimation of GFR in renal transplant recipients, *Am J Transpl.* **5**, 1306–
642 1311 (2005).

643 12. A. S. Levey, N. Tangri, L. A. Stevens, Classification of chronic kidney disease: a
644 step forward., *Ann. Intern. Med.* **154**, 65–7 (2011).

645 13. C. Franklin, M. H. Fuentes, J. Mollon, I. R. Mesa, E. Perucha, P. Conlon, N.
646 Anyanwu, S. H. Sacks, M. E. Weale, N. Soranzo, G. Lord, in *American Conference of*
647 *Human Genetics*, (2012).

648 14. J. L. Mountain, L. L. Cavalli-Sforza, Multilocus Genotypes, a Tree of Individuals, and
649 Human Evolutionary History, *Am. J. Hum. Genet.* **61**, 705–718 (1997).

650 15. L. S. P. de la Vega, A. Torres, H. E. Bohorquez, J. K. Heimbach, J. M. Gloor, T. R.
651 Schwab, S. J. Taler, S. L. Nyberg, M. B. Ishitani, M. Prieto, others, Patient and graft
652 outcomes from older living kidney donors are similar to those from younger donors
653 despite lower GFR, *Kidney Int.* **66**, 1654–1661 (2004).

654 16. A. S. Goldfarb-Rumyantzev, N. Naiman, Genetic prediction of renal transplant
655 outcome, *Curr Opin Nephrol Hypertens* **17**, 573–579 (2008).

656 17. R. P. O'Brien, P. J. Phelan, J. Conroy, P. O'Kelly, A. Green, M. Keogan, D. O'Neill,
657 S. Jennings, C. Traynor, J. Casey, M. McCormack, R. Conroy, A. Chubb, S. Ennis, D.
658 C. Shields, G. L. Cavalleri, P. J. Conlon, A genome-wide association study of recipient
659 genotype and medium-term kidney allograft function, *Clin Transpl.* **27**, 379–387 (2013).

660 18. M. Varagunam, M. M. Yaqoob, B. Dohler, G. Opelz, C3 polymorphisms and allograft
661 outcome in renal transplantation, *N Engl J Med* **360**, 874–880 (2009).

662 19. R. Lv, X. Hu, Y. Bai, H. Long, L. Xu, Z. Liu, X. Li, H. Huang, L. Wang, B. Ying,
663 Association between IL-6 -174G/C polymorphism and acute rejection of renal allograft:
664 evidence from a meta-analysis, *Transpl Immunol* **26**, 11–18 (2012).

665 20. W. McLaren, B. Pritchard, D. Rios, Y. Chen, P. Flicek, F. Cunningham, Deriving the
666 consequences of genomic variants with the Ensembl API and SNP Effect Predictor,
667 *Bioinformatics* **26**, 2069–2070 (2010).

668 21. S. Haider, B. Ballester, D. Smedley, J. Zhang, P. Rice, A. Kasprzyk, BioMart Central
669 Portal--unified access to biological data, *Nucleic Acids Res* (2009) (available at
670 http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstra
671 ct&list_uids=19420058&query_hl=16).

672 22. Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP)**2014** (available
673 at (URL: http://evs.gs.washington.edu/EVS/)).

674 23. M. C. Frith, M. Hamada, P. Horton, Parameters for accurate genome alignment,
675 *BMC Bioinformatics* **11**, 80.

24. F. Campagne, K. C. Dorff, N. Chambwe, J. T. Robinson, J. P. Mesirov, F. Lisacek, Ed. Compression of Structured High-Throughput Sequencing Data, *PLoS One* **8**, e79871 (2013).

25. P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, The variant call format and VCFtools, *Bioinformatics* **27**, 2156–2158 (2011).

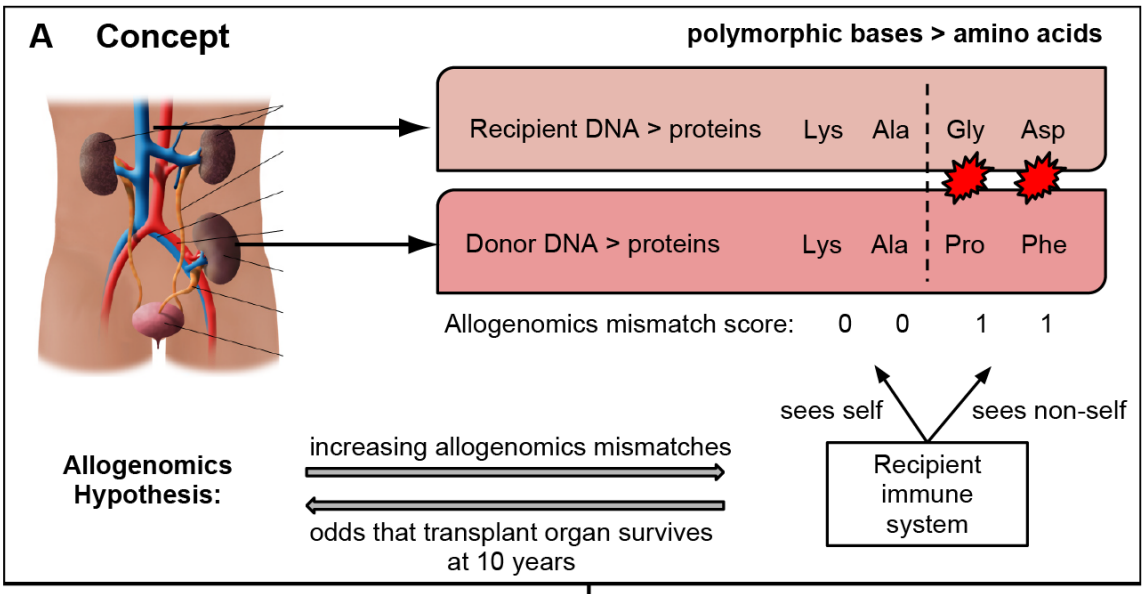
26. D. Bates, M. Maechler, B. M. Bolker, S. Walker, Fitting Linear Mixed-Effects Models using {lme4} (2015) (available at http://arxiv.org/abs/1406.5823).

27. B. Winter, Linear models and linear mixed effects models in R with linguistic applications., **1308**, 42 (2013).

Acknowledgments: We thank Dr. Joseph Schwartz and Dr. Samprit Banerjee for independent critical review of the manuscript. **Funding**: This study was funded from institutional funds (MS and FC). **Author contributions**: Study design and manuscript preparation (LM, TM, MS, FC). Statistical analyses (LM, JRL, FC). Bioinformatics analyses (LM, FC). Wrote the allogenomics scoring tool (FC). Genomics Core, Exome Sequencing (HS, JX). Collating clinical data from hospital systems for CTOT-04 and validation study (TM, MB, CL). HLA Typing, validation cohort (DD), French cohort (CS, MC). Samples and clinical information for Discovery cohort (JJF, MMA). French cohort: samples (LM, NO, ER), clinical information (LM, NO, ER). **Competing interests**: Laurent Mesnard, Thangamani Muthukumar, Manikkam Suthanthiran and Fabien Campagne disclose that they are named inventors in a filed international patent application entitled "A METHOD TO MATCH ORGAN DONORS TO RECIPIENTS FOR TRANSPLANTATION". **Data and materials availability**: de-identified data and analysis software are distributed online. A copy of the genotype files can be provided to the journal upon request should the editors wish to make them available to the reviewers upon condition of confidentiality during peer-review.

704 **Figures**:

705



B Score Contribution Examples

Donor amino-acids	Recipient amino-acids	Allogenomics score contribution δ_p
{Ala,Phe}	{Ala,Phe}	0
{Ala}	{Phe}	1
{Phe}	{Ala,Phe}	0
{Ala,Phe}	{Asp}	2

C Allogenomics Model

$$\Delta(r,d) = \sum_{p \in P} \delta_p(G_{rp}, G_{dp}) \qquad \textbf{Eqn 1}$$

$$\delta_p(G_{rp}, G_{dp}) = \sum_{a \in G_{dp}} \begin{cases} 0 \text{ if } a \in G_{r,p} \\ 1 \text{ otherwise} \end{cases} \qquad \textbf{Eqn 2}$$

706

707 **Figure 1**. **Recipient/Donor incompatibility quantified by exome sequencing and calculation**
708 **of allogenomics mismatch score (AMS). (A)** Hypothesis: Post-transplantation kidney graft
709 function is associated with the number of amino acids coded by the donor genome that the
710 recipient's immune system could recognize as non-self. **(B)** Examples of donor/recipient amino-
711 acid mismatches at one protein position, and resulting contributions to the allogenomics
712 mismatch score. The allogenomics mismatch score is calculated by summing contributions over
713 a set of genomic polymorphisms (see Methods for details). **(C)** Equations for the allogenomics
714 model. Score contributions are summed across all genomic positions of interest (set *P*) to yield
715 the allogenomics score $\Delta(r,d)$. $G_{r,p}$: genotype of recipient *r* at genomic site/position *p*. $G_{d,p}$:
716 genotype of donor *d* at site *p*. Alleles of a genotype are denoted with the letter *a*.

717

718

Characteristic	Discovery cohort	Validation cohort	French Cohort
Number of Transplant Pairs with living donors	10/10	24/24	19/19
Allogenomics mismatch score AMS(SD)[range]	1335(304)[994-2033]	1094(259)[700-1630]	560(147)[349-811]
Clinical factors			
Age			
Donor (SD)	41 (13)	46 (10)	44(16)
Recipient (SD)	48 (10)	51 (13)	38(15)
Living Donor type			
Living related N (AMS) [SD])	4 (1116 [143])	13 (939 [218])	**15(503[108])***
Living unrelated N (AMS) [SD])	6 (1481 [300])	11(1277 [170])	**4(769[41])***
Donor sex			
Male (%)	2 (20%)	8 (33%)	6(32%)
Female (%)	8 (80%)	16 (67%)	13(68%)
Donor Race			
Black (%)	4(40%)	5 (21%)	2(10%)
Non-Black (%)	6(60%)	19 (79%)	17(90%)
Recipient sex			
Male (%)	9 (90%)	13 (54%)	13 (53%)
Female (%)	1 (10%)	11 (46%)	13 (47%)
Recipient Race			
Black (%)	4 (40%)	7 (29%)	2 (10%)
Non-Black (%)	6 (60%)	17 (71%)	17 (90%)
Number of HLA mismatches ABDR (SD)	3.9 (1.91)	3.5 (1.89)	2.5 (1.68)*
Functional Factors			
Number of Patients at 12 months	10	24	17
Serum creatinine level at 12 months mg/dL (SD)	1.51 (0.35)	1.45 (0.41)	1.29 (0.41)
eGFR at 12 months ml/min/1.73m^2 (SD)	54.3(10)	54.3 (16.3)	**61.8 (18.9)***
Number of Patients at 24 months	9	23	19
Serum creatinine level at 24 months mg/dL (SD)	1.36 (0.19)	1.45 (0.49)	1.26 (0.3)
eGFR at 24 months ml/min/1.73m^2 (SD	59 (7.7)	54.85 (15.7)	**59.3 (14.5)***
Number of Patients at 36 months	8	22	19
Serum creatinine level at 36 months mg/dL(SD)	1.62 (0.50)	1.38 (0.40)	1.35 (0.45)
eGFR at 36 months ml/min/1.73m^2 (SD)	53.4 (15)	55.3 (15.9)	56.3 (16.4)
Number of Patients at 48 months	0	16	16
Serum creatinine level at 48 months mg/dL(SD)	-	1.34 (0.43)	1.40 (0.56)
eGFR at 48 months ml/min/1.73m^2 months (SD)	-	57.4 (16.4)	55.7 (18.2)
Patients with an Acute Cellular rejection episode in the first year of transplantation, N (%)	3 (30%)	5 (20%)	2 (10%)
Immunosupression			
Calcineurin Inhibitors, n (%)	9 (90%)	24 (100%)	19 (100%)
Corticosteroids, n (%)	0 (0%)	5 (21%)	**17 (90%)***

719

720 **Table 1. Characteristics of Kidney transplant recipients and their donors.** In bold,
721 characteristics that differ between the French *and* Validation cohorts (*P<0.05 , two tailed t-test).
722

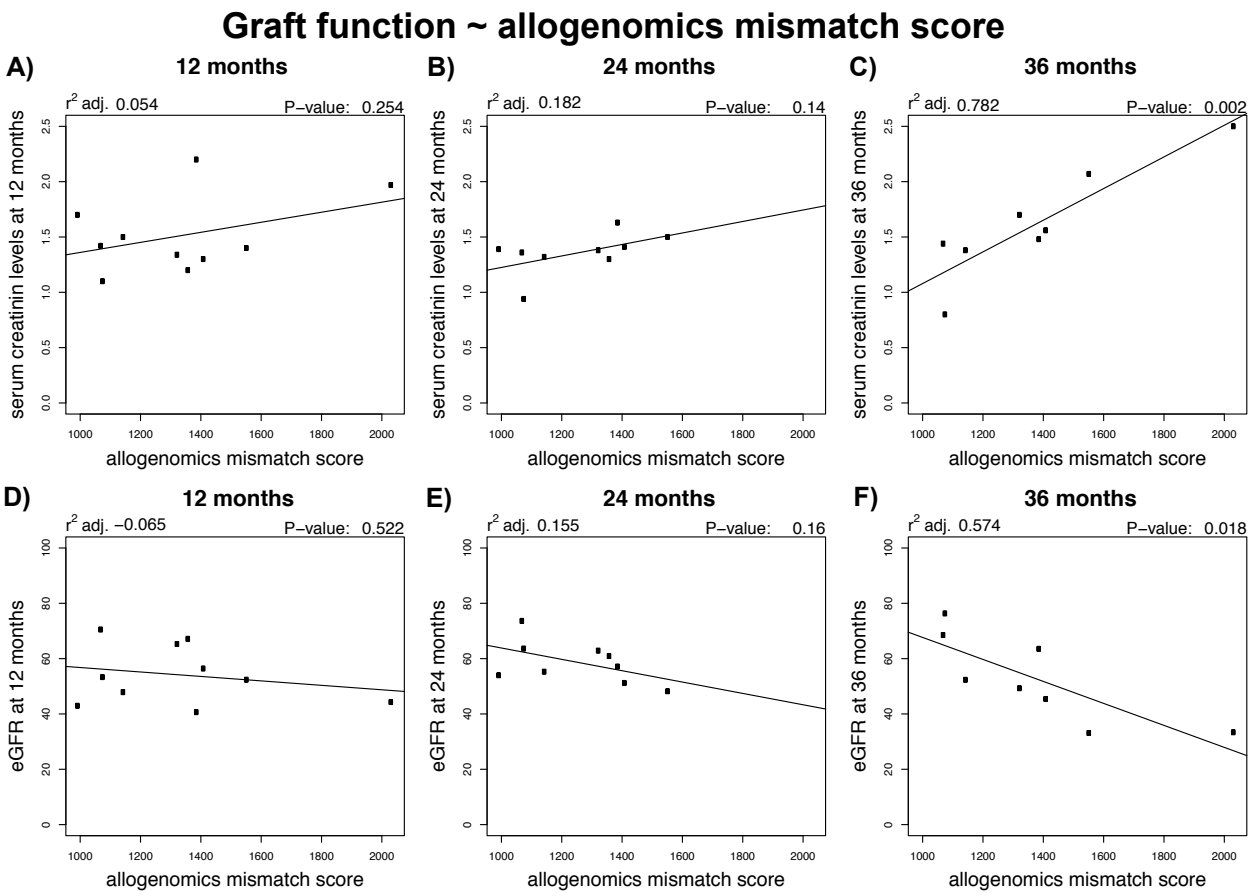


Figure 2. Relationship between the allogenomics mismatch score (AMS) and kidney graft function at 12, 24 or 36 months following transplantation in the Discovery cohort. DNA was isolated from 10 pairs of kidney graft recipients and their living kidney donors (Discovery set). Whole exome sequencing of the donor genomes and recipient genomes was performed and the sequencing information was used to calculate allogenomics mismatch scores based on amino acid mismatches in transmembrane proteins. The panels depict the relationship between the allogenomics mismatch scores and serum creatinine levels at 12, 24 and 36 months post transplantation (Panels A, B and C, respectively) and the relationship between the allogenomics mismatch scores and estimated glomerular filtration rate at 12, 24 and 36 months post transplantation (Panels D, E and F, respectively). Both serum creatinine levels and eGFR correlate in a time-dependent fashion with the allogenomics mismatch score with the strongest correlations being observed at 36 months post-transplantation.

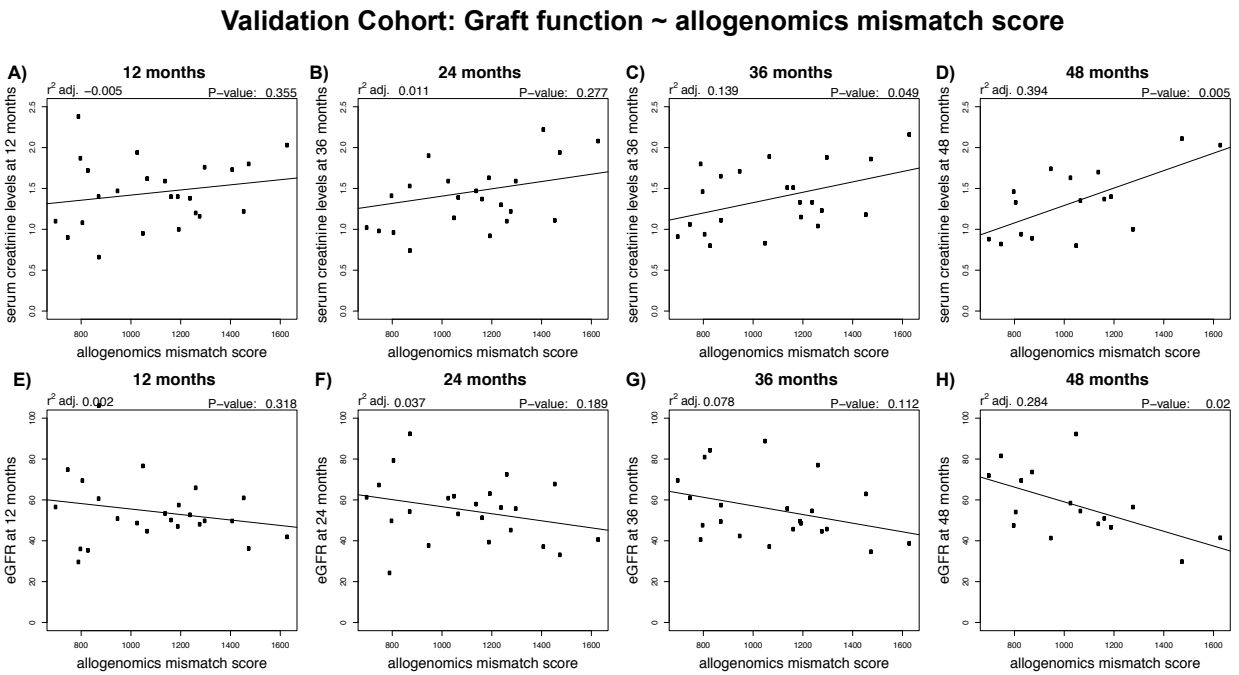


Figure 3. Relationships between allogenomics mismatch scores (AMS) and kidney graft function at 12, 24, 36 or 48 months post transplantation in the Validation Cohort. DNA was isolated from 24 pairs of kidney graft recipients and their living kidney donors (Validation set). Whole exome sequencing of the donor and recipient genomes was performed and the sequencing information was used to calculate allogenomics mismatch scores based on amino acid mismatches in transmembrane proteins. The relationships between allogenomics mismatch scores and serum creatinine levels at 12, 24, 36 and 48 months post transplantation (Panels A, B, C, and D respectively) are shown. In panels E, F, G and H, the relationships between the allogenomics mismatch scores and estimated glomerular filtration rate at 12, 24, 36 and 48 months post transplantation are shown. Both serum creatinine levels and eGFR correlate in a time dependent fashion with the allogenomics mismatch score with the strongest correlations being observed at 48 months post-transplantation.

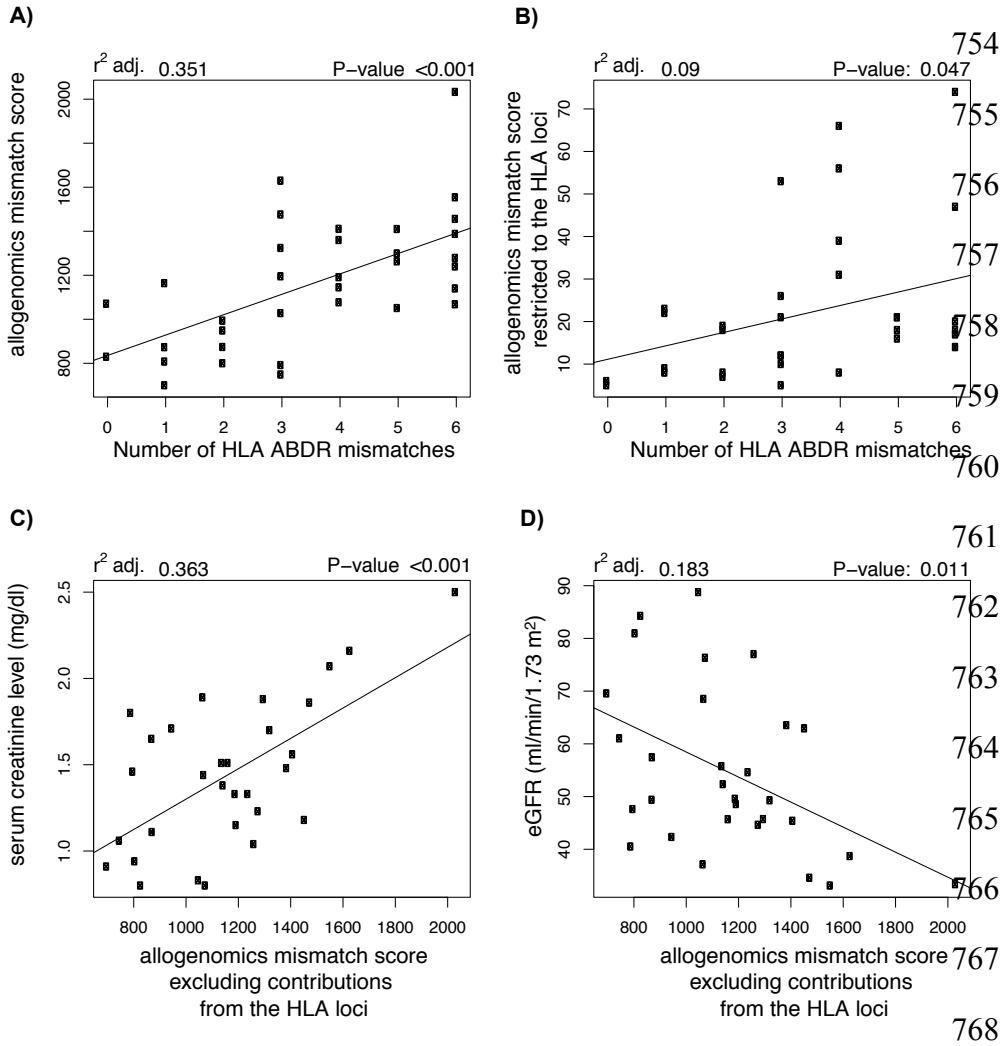


Figure 4. Relationships between allogenomics mismatch scores and the HLA loci. We combined the Discovery and Validation cohorts to examine the relation of the allogenomics mismatch score to the HLA-A, B and DR mismatch between the recipient and the kidney donor. Allogenomics mismatch scores, either calculated over all transmembrane proteins (Panel A), or restricted to the HLA A, B DR loci (Panel B) correlate with the number of mismatches in the A,B, and DR HLA loci for the complete cohort. These correlations however do not explain the association with post-transplantation graft function because when the HLA loci (A/B/ DR/DQ) are excluded from the sites included in the calculation of the allogenomics mismatch score, a significant correlation is still observed between the allogenomics mismatch score and serum creatinine level (Panel C) and eGFR (Panel D) at 36 months post transplantation (similar results are obtained when excluding the A/B/C/DR/DQ/DP HLA loci, data not shown).

782
783

784

785
786
787
788
789
790
791
792
793

794 **Supplementary Materials**
795 Supplementary materials and methods appendix containing:

796 **Fig. S1**. Model trained on the Discovery cohort applied to the Validation cohort

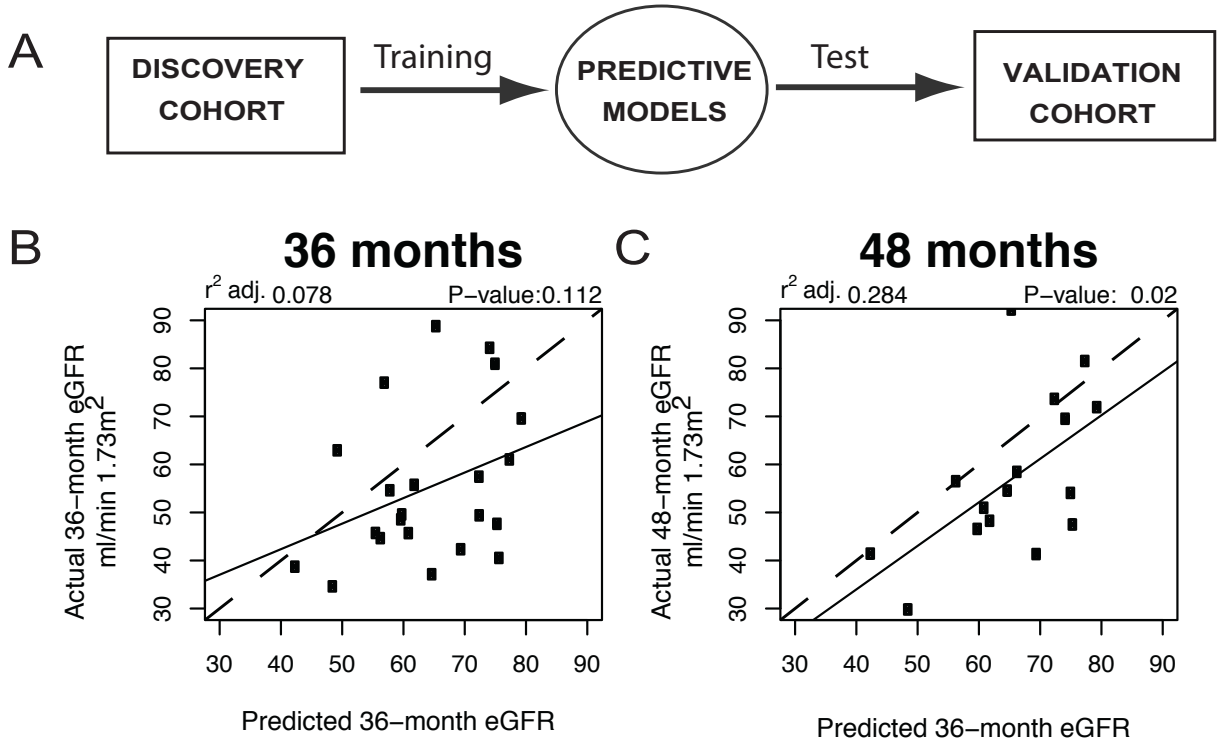
797 **Fig. S2**. Model trained on the Validation cohort applied to the Discovery cohort

798 **Fig. S3**. Effect of genotyping platform on future replication studies.

799

800
801

802 **Supplementary Figures**



803
804 **Figure S1. Model trained on the Discovery cohort applied to the Validation cohort.** A) We
805 trained a model to predict eGFR on the discovery cohort (using eGFR at 36 months) and used the
806 trained, fixed, model to predict eGFR at 36 months and 48 months for recipients of the
807 Validation cohort. The trained model was eGFR= 107.39547- -0.03974*AMS. Correlation
808 between predicted eGFR and observed eGFR on the Validation cohort at 36 (B) and 48 (C)
809 months post transplantation. Dashed lines indicate the diagonal and solid lines the regression
810 lines.
811
812

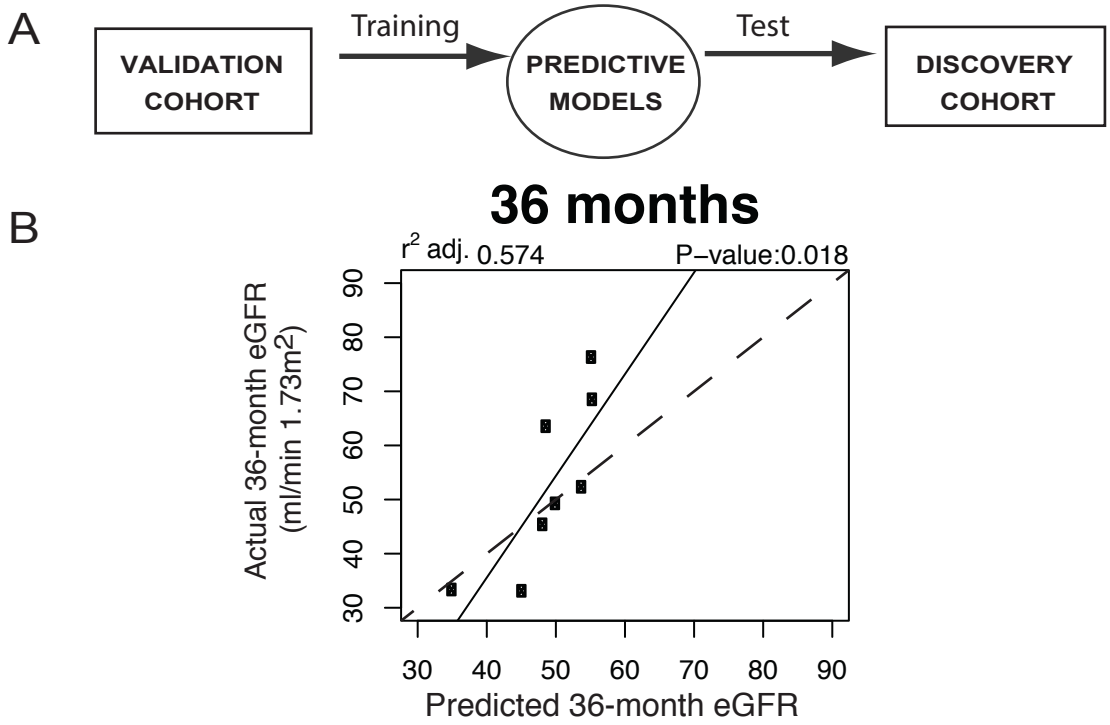


Figure S2. Model trained on the Validation cohort applied to the Discovery cohort. A) We trained models to predict serum creatinine and eGFR on the validation cohort and used the trained, fixed, model to predict serum creatinine and eGFR for recipients of the Discovery cohort. B) Correlation between the eGFR predicted by the fixed model and that observed in the Discovery cohort. The trained model was eGFR= 78.20459 -0.02114*AMS. Dashed line indicates the diagonal and solid line the regression lines.

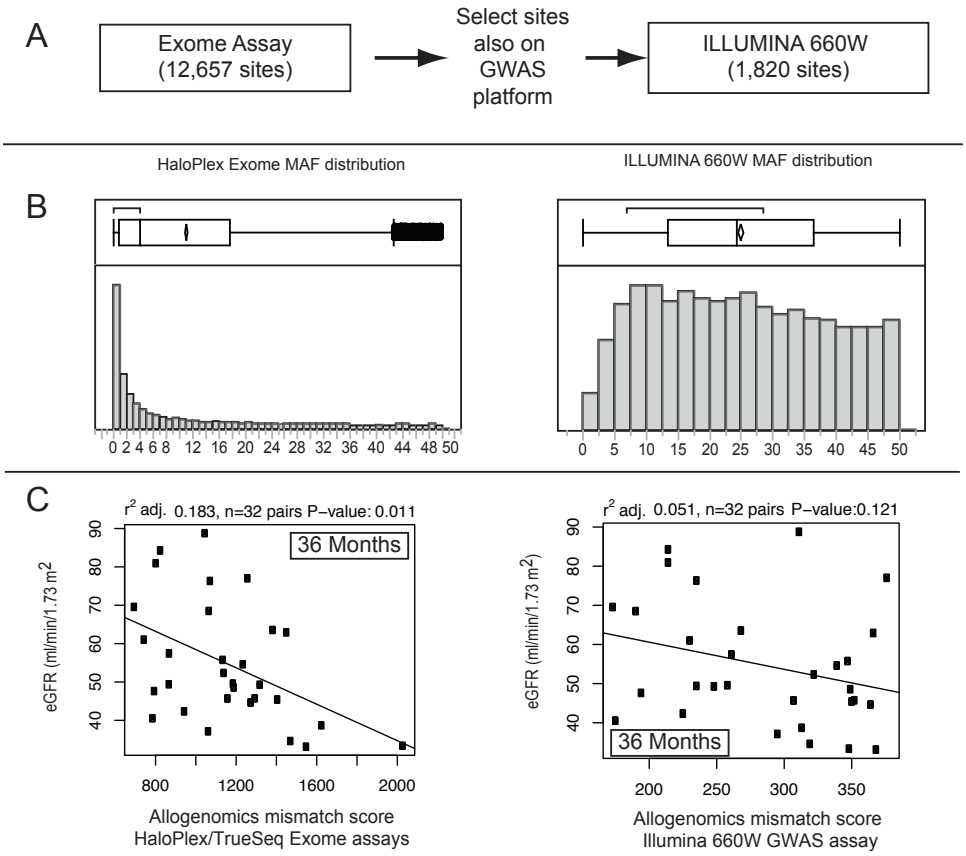


Figure S3. Effect of genotyping platform on future replication studies. In this analysis, we estimate how well the allogenomics mismatch score could be evaluated with the genotyping array technology frequently used in GWAS studies. Analyses are done on the combined Discovery and Validation cohorts (n=32 pairs with 36-month eGFR, 64 exomes). A) The allogenomics mismatch score evaluated with the Illumina TrueSeq or Agilent Haloplex exome platforms takes advantage of 12,657 genomic sites to estimate allogenomic contributions in transmembrane proteins. Only sites where an allogenomics mismatch score contribution different from zero are counted. We filtered the exome genomic sites to exclude sites not found on the Illumina 660W genotyping platform (used in *(13)*). After filtering, the allogenomics score is estimated with 1,820 remaining genomic sites. B) The minor allele frequency (MAF) of the alleles described at each set of genomic sites is shown as a histogram (MAF is estimated from the EVP database, see Methods). Exome sequencing is an assay that directly observes variations in an individual DNA sample. The MAF distributions confirm that exome sequencing helps estimate contributions from many rare (MAF<5%) polymorphisms, whereas the chip genotyping platform estimates the score based on contributions from frequent alleles. C) The scatterplot of the relationship between 36-month eGFR and the score estimated from the exome sites, or the subset of sites also measured by the GWAS platform. While some trend is still visible with sites measured on the GWAS platform, more samples would be needed to reach significance in the combined Discovery and Validation cohorts (n=34 pairs). Note that the magnitude of the scores is smaller on the GWAS platform because fewer contributions are summed. In contrast, the exome assays (Illumina TrueSeq for the Discovery cohort or Agilent Haloplex for the Validation cohort) result in stronger and significant correlations in the same set of samples.