Validation of single nucleotide polymorphisms associated with acute kidney injury in bone marrow transplant recipients

Laurel K. Willig

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2013

Committee:

Timothy A. Thornton

Sangeeta Hingorani

Program Authorized to Offer Degree:

Genetic Epidemiology

University of Washington

Abstract

Validation of single nucleotide polymorphisms associated with acute kidney injury in transplant recipients

Laurel K. Willig

Chair of Supervisory Committee:

Timothy A. Thornton, Ph.D.

Biostatistics

Background: Rates of acute kidney injury (AKI) as high as 75% have been reported in hematopoietic stem cell transplant (HSCT) patients. Previous studies have identified single nucleotide polymorphisms (SNP) associated with AKI. However, few studies, have examined genetic associations with AKI in HSCT recipients.

Methods: We perform a case-control analysis in a sample of 795 HSCT recipients who developed AKI and 1052 HSCT recipients without AKI. 18 SNPs that were previously identified as candidates for AKI were tested for association using multivariate logistic regression models.

Results: We identified three genetic polymorphisms associated with developing AKI. The T allele in rs3024495 and the C allele in rs1800896 of interleukin 10 (IL-10) both decreased the odds of developing AKI (OR=0.75, p=0.0029 and OR=0.84, p=0.011 respectively). The C allele for rs4540055 in toll-like receptor 1 (TLR1) increased the odds of developing AKI (OR=1.84, p=0.0016). Both rs3034495 and rs4540055 remained significant after adjustment for covariates.

Conclusions: Genetic polymorphisms in IL-10 and TLR1 were associated with AKI development in HSCT recipients.

Acknowledgments

I would like to express my deepest gratitude to all those who made my thesis possible. Thank you to Wenhong Fan, Barry Storer, John Hansen and the rest of John Hansen's laboratory at Fred Hutchinson Cancer Research Center for allowing me access to their research repository, contributing to the research design, and for their help in compiling the needed data. I appreciate all the input provided by the Pediatric Nephrology Division at Seattle Children's Hospital in the design and implementation of the following study. I am especially grateful to Sangeeta Hingorani, MD, MPH for agreeing to be on my thesis committee, for all her guidance in formulating a project, and for reviewing the literature and editing my thesis. To Tim Thornton, PhD, thank you so much for all of your help in designing my research and interpreting the data, guiding me through the research process and through the genetic epidemiology program. Finally, thank you to the Genetic Epidemiology Department for all the teaching over the past three years.

Table of Contents

List of Tables and Figures

Introduction

Hematopoietic stem cell transplant (HSCT) recipients often develop acute kidney injury (AKI) in the acute post transplant period. Reported AKI incidence in HSCT patients varies but some studies suggest up to 75% of HSCT patients will develop some degree of kidney injury following transplantation[1,2]. HSCT patients who develop AKI, especially those who require dialysis have a higher mortality rate than those who do not[1-3]. In addition, studies demonstrate that the AKI predisposes HSCT recipients to long term renal insufficiency[4].

AKI following HSCT is likely multifactorial. The medications and preconditioning regimens may cause intrinsic damage to the kidney. In addition, many HSCT patients develop complications such as venocclusive disease[3,5], infection[3] and acute graft versus host disease (AGVHD) post transplant which may result in ischemia and direct intrinsic damage to the kidney[3]. Finally, endothelial injury may lead to development of thrombotic microangiopathy (TMA) within the kidney.

Previous studies of AKI in hospitalized patients have found significant associates with certain genetic polymorphisms and AKI development[6]. These studies have identified genes that are involved in a diverse array of pathways which have some biologic plausibility for contributing to AKI. Most of these studies have focused on non-HSCT patients who have undergone surgery or developed sepsis. However, the development of AKI in HSCT patients potentially has similar mechanistic underpinnings such as inflammation and ischemia.

We conducted an association study of AKI in a sample of HSCT recipients from the Fred Hutchinson Cancer Research Center. We analyzed candidate single nucleotide polymorphisms (SNPs) for AKI that were previously identified in non-HSCT samples to identify genetic associations that replicate in HSCT recipients.

Methods

Literature Search for Candidate SNPs

We performed a comprehensive PubMed search using the terms “acute kidney injury” “gene” and “gene polymorphism” to identify studies published by Jan 1, 2011 that reported an association between AKI and genetic polymorphisms. These manuscripts were manually examined for subject matter, type and strength of association. We included only studies whose subjects included adult hospitalized patients as this demographic most closely resembled our study cohort. SNPs with associations p values were less than 0.05 in at least one of the studies were included in our analysis. Studies that did not report a p-value that met this significance threshold were excluded from our analysis. In addition, we looked at four previously unpublished SNPs which potentially mediate the development of AKI and are currently being investigated at the University of Washington’s Kidney Research Institute. We excluded studies that reported associations with alternative genetic variants such as deletions, microsatellites, or variable number tandem repeats due to limitations of our SNP genotyping array to provide informative analysis. A total of 19 SNPs met our criteria (Table 1).

The angiotensinogen (AGT) SNP was unable to be properly identified for further study. This SNP was found to be significantly associated with AKI in the cardiac bypass population[7]. AGT plays an important role in blood pressure control as part of the renin-angiotensin system. Recent studies have linked urinary levels of AGT to increased risk of developing AKI [8]. Thus, this SNP may be associated with AKI in HSCT recipients and would have been interesting to study. SNPs in linkage disequilibrium (LD) with the candidate SNPs as defined by an R^2 value >0.8 were identified using SNP Annotation and Proxy Search (SNAP)[9] and Genome Variation Server[10]. These SNPs were used to support associations identified by the candidate SNPs.

Study Design and Sample

This study included a total of 3371 hematopoietic stem cell transplant recipients receiving transplants from 1992 to 2012 who received allogeneic HCT after myeloablative preconditioning at the Fred

Hutchinson Cancer Research Center and Seattle Cancer Care Alliance. All recipient DNA samples were collected before HCT according to approved research protocols. Project specific IRB approval was obtained for the use of these samples. The sample population was divided into two cohorts due to change in genotype platform that occurred during the study. Cohort 1 patients enrolled in the study from 1992-2004 and were genotyped using the Affymetrix GeneChip®Genome-Wide Human SNP Array 5. Cohort 2 patients enrolled in the study from 2005 to 2012 and were genotyped using the Illumina HumanOmni 1.2M Quad®.

Exclusion criteria included those patients without a baseline creatinine, those with a missing indication for dialysis, and those with a baseline creatinine >4. Additionally, participants were excluded if their genotyping information was not available to investigators. There were 48 individuals who were excluded from our analysis due to missing baseline creatinine, missing indication for dialysis, or baseline creatinine >4. Of the remaining 3323, 104 had missing genotype information and were also excluded from the analysis. Figure 1 illustrates study participation groups and exclusions. Summary statistics for age, gender, average grade of AGVHD, and type of donor transplant were similar between the two cohorts (Table 2). The final study sample for the genetic association analysis of AKI consisted of 3218 individuals, with 1399 individuals from Cohort 1 and 1820 individuals from Cohort 2.

Measurements of Outcomes and Variables

Clinical and demographic data were collected on participants at multiple time points during the study. Baseline creatinine was defined as the creatinine obtained at admission for HSCT. Participants were classified as developing AKI if they had a doubling of their baseline creatinine sustained for seven days during the first 100 days post transplant. Our AKI definition was the equivalent of acute kidney injury network (AKIN) stage 2 criteria[11] or the Injury stratification using the RIFLE criteria used by the Acute Dialysis Quality Initiative[12]. The controls were the sample individuals who did not meet this criterion. Available demographic covariates included age, gender, unrelated vs related donor, and average grade of graft versus host disease as previously described[13]. Age was modeled as a continuous variable and the other three variables were categorical. Principle components analysis using previously reported methods was performed with ~120,000 common SNPs genotyped on both arrays to adjust for population

stratification[14]. This analysis was performed by the FHCRC and the first four PCAs were used as covariates in the analysis.

Sample Preparation, Genotyping, Quality Control and Imputations

Genomic DNA specimens were extracted from blood mononuclear cells or EBV transformed B lymphocyte cells using the Puregene kit, Qiagen, Valencia, CA. In Cohort 1, the Affymetrix Service Laboratory (ASL; Santa Clara, CA) performed amplification and hybridization using the Affymetrix GeneChip®Genome-Wide Human SNP Array 5.0. Data quality was assessed via three different methods: the Affymetrix Bayesian Robust Linear Model with Mahalanobis distance classifier (BRLMM) based "QC call rate", the clustering call rate and a PCR-based ABO and XY genotyping based sample verification method. In Cohort 2, the Fred Hutchinson genomic core facility performed amplification and hybridization using the Illumina HumanOmni 1.2M Quad® (Illumina, Inc., San Diego, CA). Data quality was assessed using the Illumina GenomeStudio software®.

The candidate SNP genotype determination algorithm has been described previously[13]. Original candidate SNPs that were not genotyped on one or both arrays were imputed according to 1000 Genomes Project Phase 1 SNPs using the software IMPUTE v2. For each unknown SNP in a data set, IMPUTE v2 calculates a marginal probability for each possible genotype at that location using hidden Markov Models and the information provided by the surrounding observed haplotypes and genotypes of the unknown genotype's flanking regions. Probability distributions for each called genotype are also provided to help correct errors.[15,16] If none of the possible genotypes at an imputed locus had a maximum posterior genotype probability exceeding 0.8, then the SNP was excluded from the association study with AKI (Tables 2 &3).

All genotyped and imputed SNPs that violated Hardy-Weinberg equilibrium with P value <0.001, had a minor allele frequency (MAF) <0.001 or had a <90% call rate were excluded from the analyses. Note that the groups of SNPs passing these quality control measures differed between Cohort 1 and Cohort 2 with more SNPs satisfying the criteria in Cohort 2 (Tables 2 & 3). Only six validation SNPs passed quality

control filters in Cohort 1 while all validation SNPs passed quality metrics in Cohort 2. This discrepancy remained when including SNPs in high linkage disequilibrium with the validation SNPs (Table 4).

Due to the discrepancy noted between the two cohorts, further analysis of the SNPs was performed. A logistic regression association analysis comparing Cohort 1 and Cohort 2 was performed using PLINK software[17,18], where one cohort was designated to be cases and the other was the control. The analysis was restricted only to those SNPs passing quality control in both cohorts. This comparison suggested a systematic difference existed between the two cohorts (Figure 2). Given that more validation SNPs passed quality control measures in Cohort 2 (Tables 2 & 3) and concerns for a systematic difference amongst the participants, samples from Cohort 1 were excluded from the association analysis with AKI.

Statistical Analysis

Summary demographic statistics for the two separate cohorts as well as the combined cohort were calculated in R. Exploratory logistic regression for non-genetic factors including age, gender, type of donor transplant, and average grade of AGVHD was also done using R[19]. Logistic regression for each SNP was performed in PLINK[17,18]. Logistic regression was used to find odds ratio and 95% confidence levels for each of the SNPs analyzed. All odds ratios were adjusted for age, gender, type of donor transplant, and average grade of graft vs host disease as these characteristics may be associated with development of AKI. The first four principle components were included in the model to account for population structure within this population[13]. LocusZoom [20]and R software[19] were used for all graphical representations.

Results

Exploratory logistic regression of covariates showed a significant association between gender and AKI (OR=0.43, p=<0.001). While 50% of all females were cases, only 40% of males were cases. Average grade of AGVHD also was associated with AKI (OR:0.1.17, p=<0.001). The other variables were not statistically associated with the outcome but were included in the model due to the a priori hypothesis that they may be confounders. Three SNPs in two genes were associated with AKI in the sample HSCT Cohort 2 recipients (Table 6). These included rs3024495 and rs1800896 in interleukin 10 (IL-10) and rs4540055 in toll-like receptor 1 (TLR1). When controlling for age, gender, average AGVHD, and type of transplant, each additional T allele in rs3024495 decreased the odds of AKI development (OR=0.75, p=0.0029). For rs1800896, a similar decrease in AKI odds was seen for each additional C allele (OR=0.84, p=0.011). This second association does not remain significant when employing a Bonferroni correction. The decreased odds for AKI associated with these SNPs is consistent with results of the study by Wattanatham[21]. SNPs in high LD with these two SNPs show similar statistical associations with AKI development (Figures 5 and 6).

Each copy of the C allele for rs4540055 significantly increases risk of AKI in HSCT patients when controlling for age, gender, average AGVHD, and type of transplant (OR=1.84, p=0.0016). This SNP has previously been unreported in the literature for association with AKI. This SNP is not in high LD with any of our genotyped SNPs because there were no SNPs in high LD identified in SNAP[9] or GVS[10].

Discussion

Interleukin 10 (IL-10) is an anti-inflammatory cytokine produced by virtually all inflammatory cells with monocytes being the primary producers. IL-10 regulates the immune system to prevent immune mediated damage to the host[22]. Several studies have indicated that IL-10 plays a renoprotective role [23]. IL-10 reduces gene expression of proteins in the fibrosis pathway. It also decreases the migration of inflammatory cells during renal injury. Reduction of IL-10 in mice models of tubulointerstitial disease results in increased fibrosis and renal injury[24]

Genetic polymorphisms in IL-10 has been previously associated with the development of autoimmune diseases such as Crohn's disease[25,26]. Our study examined rs3024495, rs1800896, and rs1800872 because Wattanatham originally found an association of haplotype CGG corresponding to rs 1800872, rs1800896, and rs3024495 and increased risk of renal dysfunction in critically ill patients with sepsis who developed pneumonia [21]. Our study looked at the separate additive effects of these three SNPs and discovered that each T allele at rs3024495 and each C allele at rs1800896 decreased the odds of AKI. While our single SNP association analysis is not the same as the analysis performed in the original study, the association results for these two SNPs are consistent with the results by Wattanatham[21]. The association between rs1800896 and AKI does not remain significant when using a strict Bonferroni correction. However, other studies have suggested that rs1800896 is associated with kidney injury. In a recent study by Chang et al, rs1800896 genotypes AG and GG were associated with increased risk of contrast nephropathy, a form of acute kidney injury, when compared to AA genotype [27] . This result is consistent with the C allele being protective. No other studies have examined the association between rs3024495 and kidney injury.

We identified a novel polymorphism rs4540055 in the TLR1 gene that appears to be associated with AKI in HSCT patients. This particular SNP was not examined in any of the previous validations studies but preliminary unpublished data from the Kidney Research Institute at the University of Washington suggested that it might be related to AKI development in other populations. The toll-like receptors (TLR) family recognizes conserved microbial components and triggers the host's immune system. TLR1 is a

cell surface molecule that recognizes lipopeptides on certain infectious agents. TLR1 is expressed in the renal tubular cells during injury [28]. Other genetic studies have linked polymorphisms in TLR1 to organ dysfunction[29], urinary tract infections[30] and IgA nephropathy[31]. These disease processes often have a similar endpoint in inflammation, which is thought to be a key mediator of AKI.

Our association between TLR1's rs4540055 and AKI coupled with our association of IL-10 polymorphisms with AKI is interesting as other evidence suggests that the two genes are involved in the same pathway. TLRs regulate IL-10 expression during times of infection[32]. Additionally, previous reports of genetic polymorphisms in TLR1 linked different genotypes to differential levels of IL-10 production[33]. Further examination of possible interactions between single nucleotide polymorphisms in these two genes and AKI may be informative.

Fifteen out of the eighteen SNPs we considered with previous associations with AKI were not associated with AKI in this HSCT population. A number of differences exist between our study population and previous studies, which may explain the non-replication of these SNPs. Our study population likely varies greatly from the previous study populations in terms of overall health, medications administered and medical complications such as liver dysfunction that predisposes to kidney injury. In addition, we severely alter their immune system and inflammatory processes. Thus, many SNPs associated with inflammatory cascades and previously linked to AKI in other diseases may not be the primary mediators of AKI in our population. Our population also has the unique situation of having two genomes, the donor's present in all cells produced by the bone marrow and their own in all other cells. In our study, we examined only the HSCT recipient’s genotype not the donor’s genotype, potentially confounding previously associated genetic polymorphisms especially those involved in pathways that rely on both local and systemic signaling.

Mechanistic studies demonstrate that inflammation is an important mediator of AKI development.[34,35] Both the local environment and systemic circulating factors interact to comprise this inflammation cascade. Renal tubular epithelial cells and renal endothelial cells both produce chemokines and cytokines which attract systemic inflammatory cells. In our population, the renal tubular cells and endothelial cells have the HSCT recipients’ DNA whereas especially after engraftment, the systemic

inflammatory cells would contain donor DNA. IL-10 was significantly associated with AKI development in our study. IL-10 production is locally secreted by local macrophages and dendritic cells in the kidney in addition to systemic inflammatory cells[24,36]. Thus, the association may have remained intact even when excluding donor genotype from analysis. TLR1 receptors are present on renal tubular cells and are thought to work locally to drive the inflammation cascade[31]. Thus, recipient genetic polymorphisms in TLR1 might be expected to play a role in AKI. However, several of the other genes in this study are expressed locally in the kidney and thus one might expect to see validation with these SNPs even when focusing on HSCT recipient's genotype suggesting that there are other reasons for failed validation.

Increased phenotypic variability likely contributes to failure of replication due to increased subjectivity of the case definition. Previous attempts to validate SNPs in other disease phenotypes with a large degree of variability such as acute GVHD support this idea[13]. AKI definitions especially those from early studies were not standardized. Recent attempts to improve standardization of AKI have led to two different classification systems, RIFLE[12] and AKIN [11]. Our AKI definition was the equivalent of AKIN stage 2 criteria[11] or the injury stratification using the RIFLE criteria[12]. However, most of the studies we were trying to validate were completed prior to the development of these standardized tools. Many of the studies focused on more severe types of AKI which may explain the failure of many of the candidate SNPs to replicate. Perhaps, these SNPS are more predictive of severity of disease instead of development of disease.

Interestingly, when examined separately, both cohorts seemed to have a greater number of insignificant p values than one would expect by chance for the number of SNPs examined, suggesting that our association results may be conservative. This finding may be an artifact of the high LD of the SNPs in our study. However, several possible sources of variability remain unaccounted for in our study which could impact the association results. We found an association between both gender and AGVHD and AKI, both previously identified as risk factors[1]. Likely, other known risk factors for AKI that we do not include in our analysis may be associated with AKI development in this population. For example, we have included little information about preexisting conditions which may increase risk of AKI such as diabetes[37] and hypertension[38]. We also have no data on whether patients had infectious complications or developed

venoocclusive disease during their first 100 days post transplant. We also may not be able to detect SNPs with small effect size due to the noise created by the heterogeneous nature of our study participants.

While individual cohort results suggest conservative estimates for our validation SNPs, a comparison of the controls from both cohorts suggests that a nonrandom systematic difference in genotyping exists. Several recent studies of other array technologies suggest that a large degree of discordance exists amongst the platforms[39]. Newly designed arrays target areas previously found to provide poor data quality. Thus one would expect that Cohort 2's genotyping quality might be improved given this cohort included all the later enrollees. Also, each cohort was analyzed using separate genotyping algorithms. Again, the algorithm employed for genotyping in cohort 2 likely has included technologic discoveries potentially improving genotype identification. Our genotyping may have improved had we used multiple algorithms for genotype calling. Using multiple algorithms for variant calling has shown to improve sensitivity of variant detection for several platforms[40]. Use of the same algorithms throughout the sample and/or multiple algorithms to improve quality of calls may have helped dissipate the differences observed between the two groups.

Genotyping platforms likely contributed to the differences between our two cohorts and also might explain why we failed to find associations between certain candidate SNPs and AKI. Several different types of genotyping platforms were employed in the previous studies examining SNP associations with AKI. We also used two different platforms for our sample. Differing sequencing techniques likely results in some variation in the genotype results and may lead to confounding. Additionally, we had to impute several of the validation SNPs as they were not directly genotyped on our assays. Imputation methods are increasingly common in recent genetic literature due to all the different platforms available and have largely shown to reliably produce genotype calls at untyped SNPS[15,41]. However, all imputation methods involve certain assumptions and are affected by the amount of coverage provided on the genotype platform. Thus, imputation introduces another level of variation into analysis and may confound results, leading both to the systematic difference between our two cohorts and also may have led to our inability to identify associations between certain candidate SNPs and AKI.

Summary and Conclusions

In conclusion, SNPs in IL-10 and TLR1 were associated with AKI development and may suggest a role for IL-10 and TLR1 in mediating AKI in HSCT patients. Other SNPs in these genes as well as genes upstream and downstream in each of these pathways should be examined. Future studies of this combined cohort need to be carefully examined for systematic differences between genotyping technologies and biases that may be introduced by the use of different genotyping algorithms on different subgroups within the sample. Finally, exploration of how donor genotype affects AKI development in HSCT is future research that should be conducted for greater insight into the genetic risk factors for AKI in HSCT patients.

Table 1: SNPs identified by literature search for validation					
Gene	**RS#**	**Chr**	**Pos**	**Alleles**	**Publications**
IL10	3024495	1	206942413	C/T	Wattanatham et al 2005[21]
IL10	1800872	1	206946407	T/G	Wattanatham et al 2005[21]
IL10	1800896	1	206946897	T/C	Wattanatham et al 2005[21]
TLR1	4540055	4	38803255	A/C	Himmelfarb unpublished
TLR1	5743551	4	38807654	T/C	Himmelfarb unpublished
IL6	1800797	7	22766221	A/G	Stafford-Smith et al 2005[7]
IL6	1800796	7	22766246	G/C	Stafford-Smith et al 2005[7]
IL6	1800795	7	22766645	C/G	Gaudino et al 2002[42], Stafford-Smith et al 2005[7]
EPO	1617640	7	100317298	C/A	Popov et al 2010[43]
NOS3	2070744	7	150690079	C/T	Popov et al 2009[44]
NOS3	1799983	7	150696111	T/G	Stafford-Smith et al 2005[7]
TIRAP	611953	11	126163691	A/G	Himmelfarb unpublished
TIRAP	625413	11	126164349	T/C	Himmelfarb unpublished
PNMT	876493	17	37824545	G/A	Alam et al 2010[45]
PNMT	5638	17	37826249	A/G	Alam et al 2010[45]
APOE	429358	19	45411941	T/C	Chew et al 2000[46], MacKensen et al 2004[47], Stafford-Smith et al 2005[7]
APOE	7412	19	45412079	C/T	Chew et al 2000[46], MacKensen et al 2004[47], Stafford-Smith et al 2005[7]
HIF1A	11549467	14	62207575	A/G	Kolyada et al 2009[48]
AGT**	?	1	842	T/A	Stafford-Smith et al 2005[7]

**Unable to find the exact identity of this SNP in order to perform validation.

Table 2: Summary demographics for each cohort		
Covariate	**Cohort 1** (n=1399)	**Cohort 2** (n=1820)
Age	37.47 (s.d. 14.26)	43.86 years (s.d. 15.93)
Gender (males)	834 (58.8%)	1095(59.3%)
Number of unrelated donor transplant	682 (48%)	1060 (57%)
Average grade of acute graft vs host disease	2.95 (sd:1.05)	1.75 (sd:1.12)
AKI	669 (47%)	795 (43%)

Table 3: Validation SNP Metrics for Cohort 1							
Gene	**RS#**	**Genotyped/ Imputation**	**Call Rate**	**Posterior Probability**	**HWE (p value)**	**MAF**	**Pass/Fail**
IL10	3024495	I	0.99	0.99	0.54	0.15	P
IL10	1800872	I	0.89	0.94	0.92	0.25	F
IL10	1800896	I	0.75	0.87	0.94	0.43	F
TLR1	4540055	I	0.96	0.99	<0.001	0.054	F
TLR1	5743551	I	0.94	0.97	0.03	0.31	P
IL6	1800797	I	0.99	0.99	0.75	0.36	P
IL6	1800796	I	0.93	0.96	<0.001	0.11	F
IL6	1800795	G	1.0	1.00	0.94	0.36	P
EPO	1617640	I	0.68	0.85	0.0040	0.37	F
NOS3	2070744	I	0.59	0.81	0.18	0.34	F
NOS3	1799983	I	0.54	0.79	0.79	0.28	F
TIRAP	611953	I	0.99	1.00	0.15	0.21	P
TIRAP	625413	I	0.99	1.00	0.40	0.19	P
PNMT	876493	I	0.14	0.63	<0.001	0.27	F
PNMT	5638	I	0.67	0.84	0.81	0.079	F
APOE	429358	I	0.79	0.90	1.00	0.13	F
APOE	7412	I	0.78	0.88	0.35	0.038	F
HIF1A	11549467	I	1.0	0.99	1 .00	0.00035	F

Table 4: Validation SNP Metrics for Cohort 2							
Gene	**RS#**	**Genotyped/ Imputation**	**Call Rate**	**Posterior Probability**	**HWE (p value)**	**MAF**	**Pass/Fail**
IL10	3024495	I	1.00	1.00	0.27	0.16	P
IL10	1800872	I	1.00	1.00	0.47	0.23	P
IL10	1800896	G	1.00	1.00	0.13	0.48	P
TLR1	4540055	I	0.99	0.99	1.00	0.033	P
TLR1	5743551	G	1.00	1.00	0.054	0.25	P
IL6	1800797	G	1.00	0.99	0.034	0.41	P
IL6	1800796	G	1.00	0.99	0.73	0.049	P
IL6	1800795	I	0.99	0.96	0.018	0.42	P
EPO	1617640	I	1.00	1.00	0.79	0.39	P
NOS3	2070744	I	0.94	0.97	1.00	0.38	P
NOS3	1799983	I	0.94	0.97	0.077	0.31	P
TIRAP	611953	G	1.00	1 .00	1.00	0.22	P
TIRAP	625413	G	1.00	1 .00	1.00	0.22	P
PNMT	876493	G	1.00	1.00	1.00	0.43	P
PNMT	5638	G	1.00	1 .00	0.43	0.021	P
APOE	429358	I	0.99	1.00	0.61	0.14	P
APOE	7412	G	0.99	0.99	0.081	0.074	P
HIF1A	11549467	G	1.00	1.00	1.00	0.0092	P

Table 5: Number of SNPs passing all imputation and quality control metrics		
Cohort	**Validation SNPs Passing Quality Measures (total SNPs=18)**	**All SNPs Passing Quality Measures (total SNPs=533)**
Cohort 1	6 (33%)	323 (61%)
Cohort 2	18 (100%)	489 (92%)

Table 6: Association of Validation SNPs and Acute Kidney Injury in Cohort 2						
Gene	**SNP**	**CHR**	**BP**	**Minor Allele**	**OR**	**p value**
IL-10	rs3024495	1	206942413	T	0.75	0.0029
IL-10	rs1800896	1	206946897	C	0.84	0.011
IL-10	rs1800872	1	206946407	T	1.07	0.44
TLR1	rs4540055	4	38803255	C	1.84	0.0016
TLR1	rs5743551	4	38807654	C	1.07	0.41
NOS3	rs1799983	7	150696111	T	1.10	0.18
NOS3	rs2070744	7	150690079	C	1.08	0.27
IL6	rs1800795	7	22766645	C	0.93	0.29
IL6	rs1800796	7	22766246	C	1.01	0.96
IL6	rs1800797	7	22766221	A	0.93	0.29
EPO	rs1617640	7	100317298	C	0.98	0.78
TIRAP	rs625413	11	126164349	T	0.93	0.26
TIRAP	rs611953	11	126163691	A	1.094	0.31
PNMT	rs876493	17	37824545	G	0.99	0.92
PNMT	rs5638	17	37826249	G	0.76	0.26
APOE	rs429358	19	45411941	C	0.97	0.75
APOE	rs7412	19	45412079	T	1.37	0.17
HIF1A	rs11549467	14	62207575	A	1.00	0.99

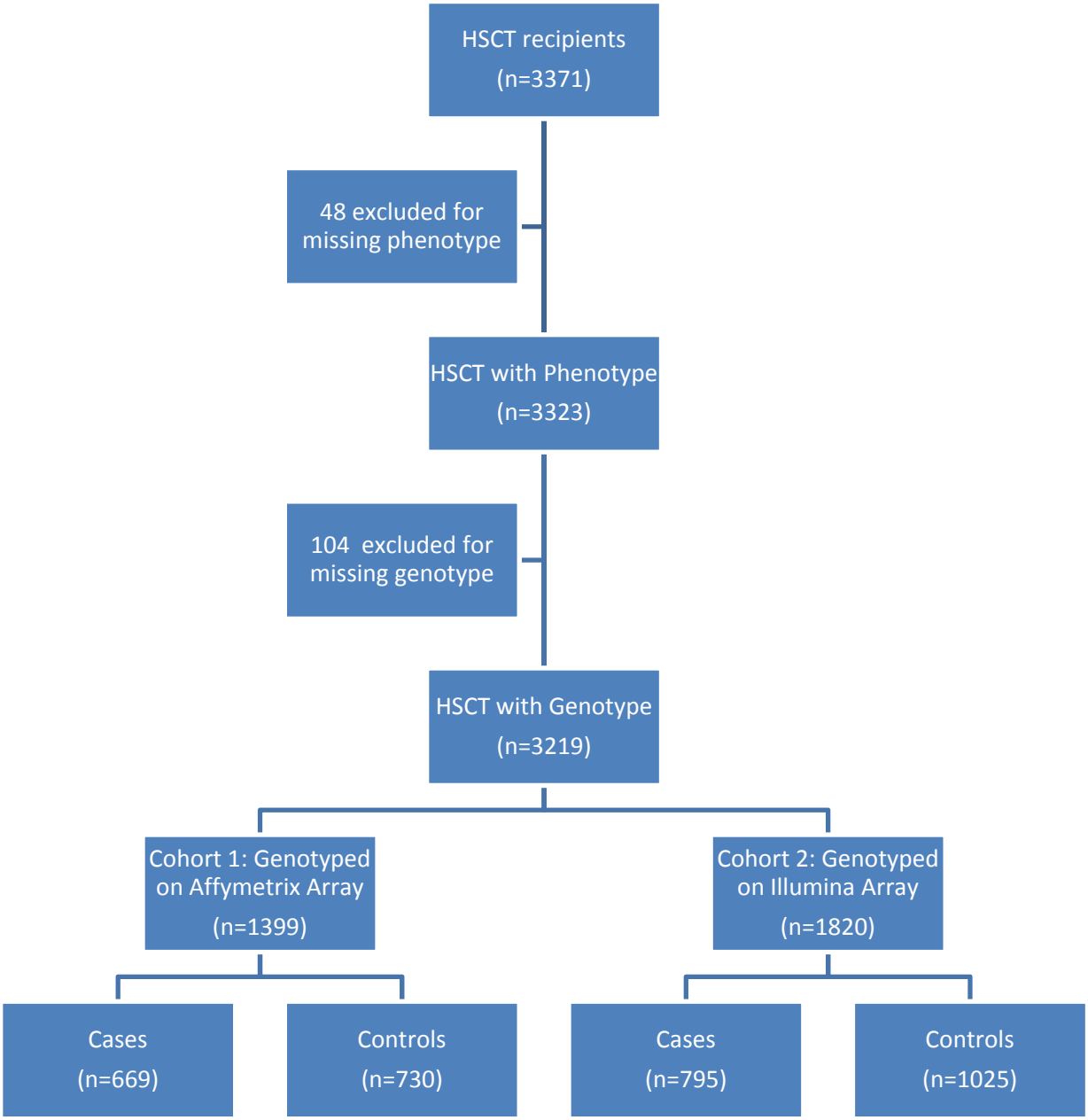

Figure 1: Enrollment flow diagram for HSCT recipients in the study

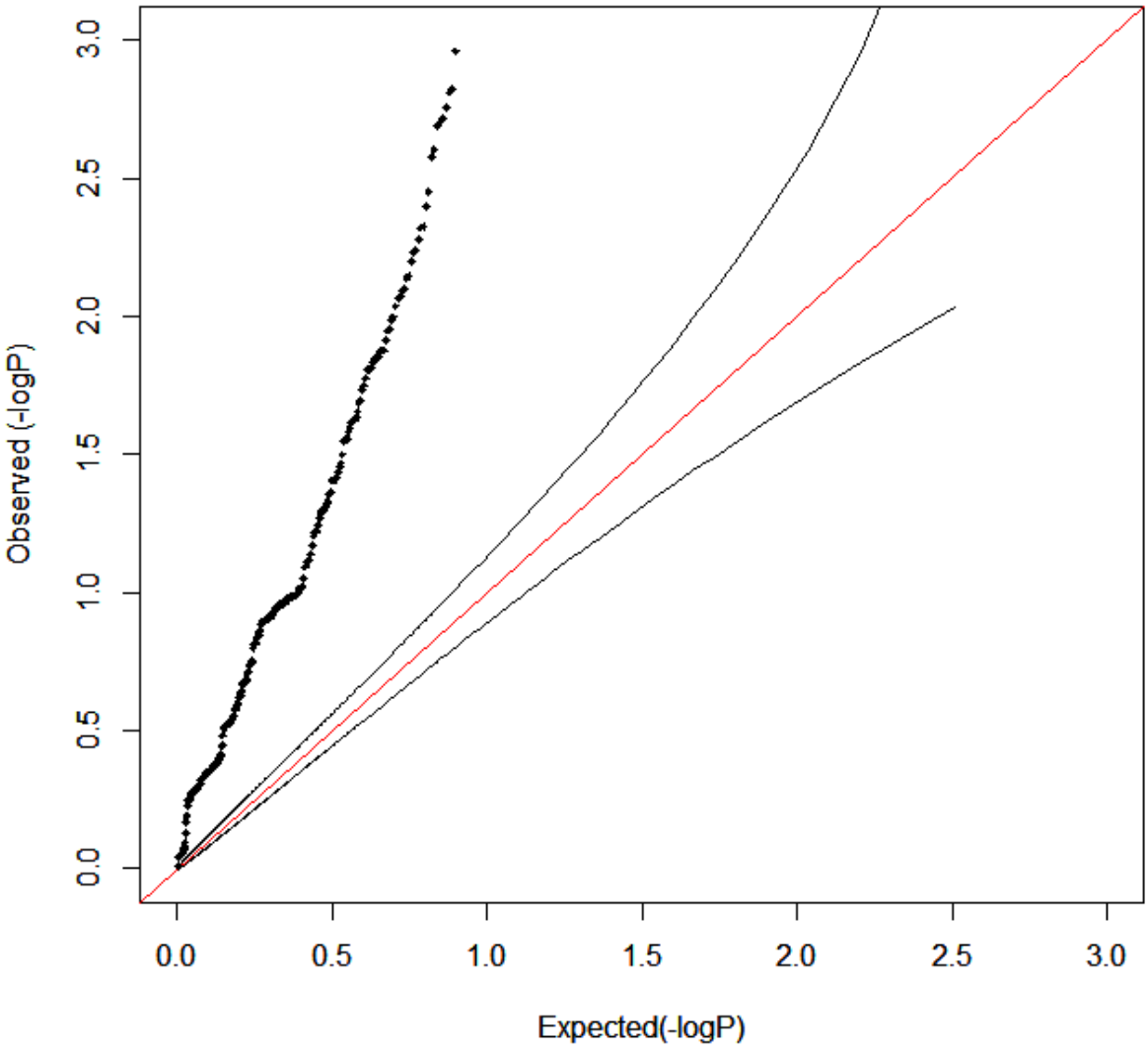

Figure 2: Cohort 1 vs Cohort 2 Controls: Normal QQ Plot of P Values

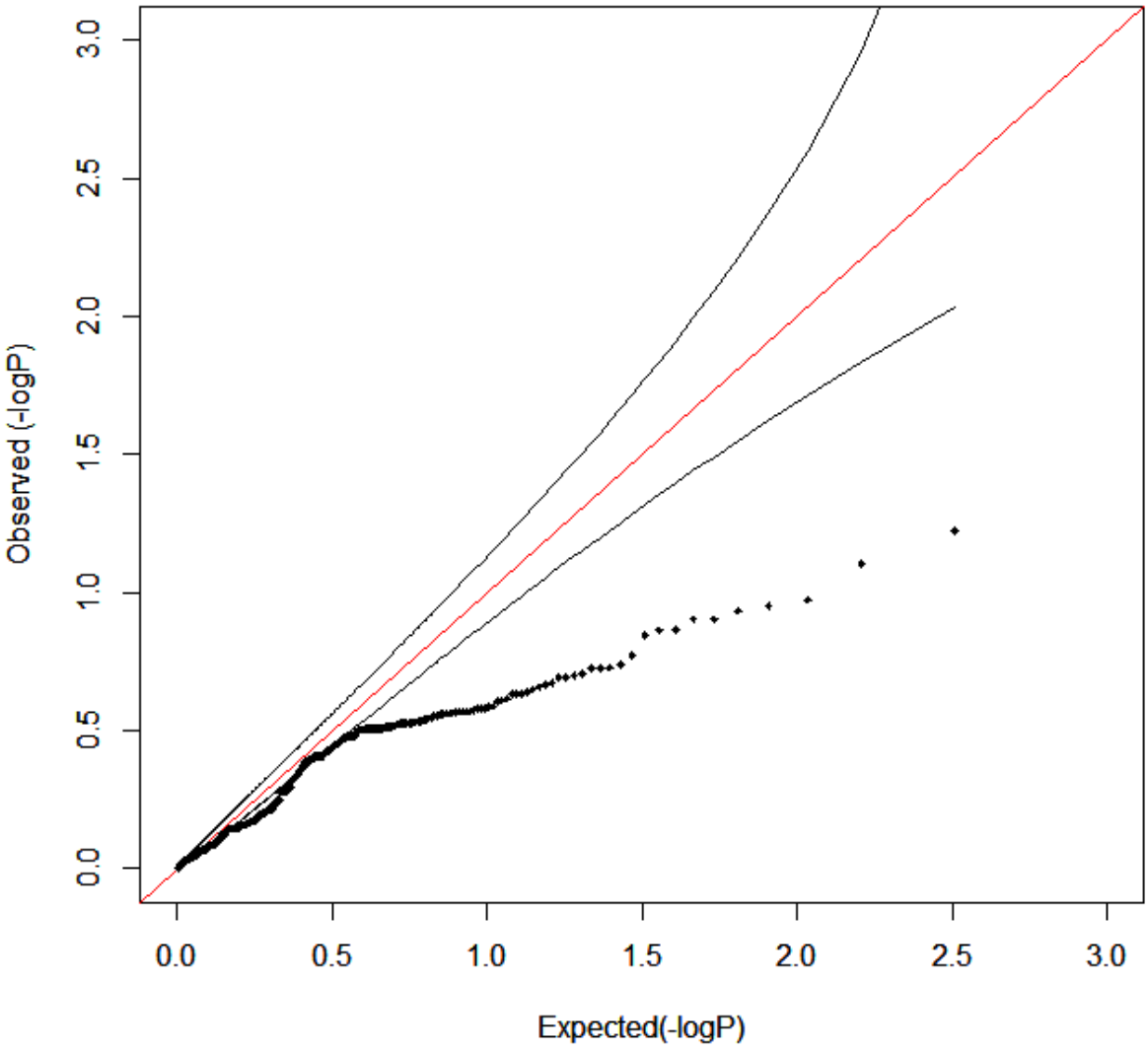

Figure 3: SNP Association of AKI in Cohort 1: Normal QQ Plot of P Values

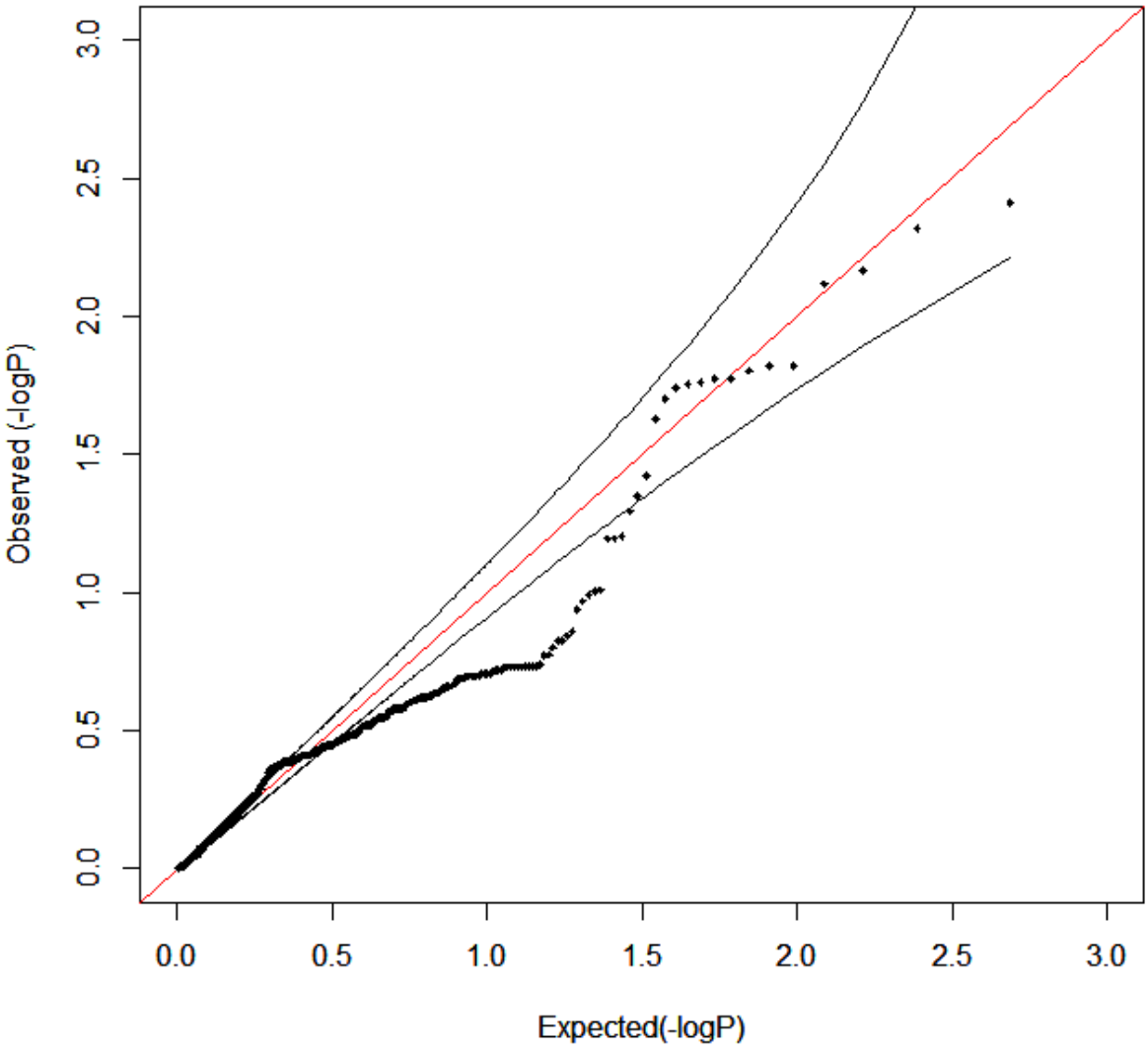

Figure 4: SNP Association of AKI in Cohort 2: Normal QQ Plot of P Values

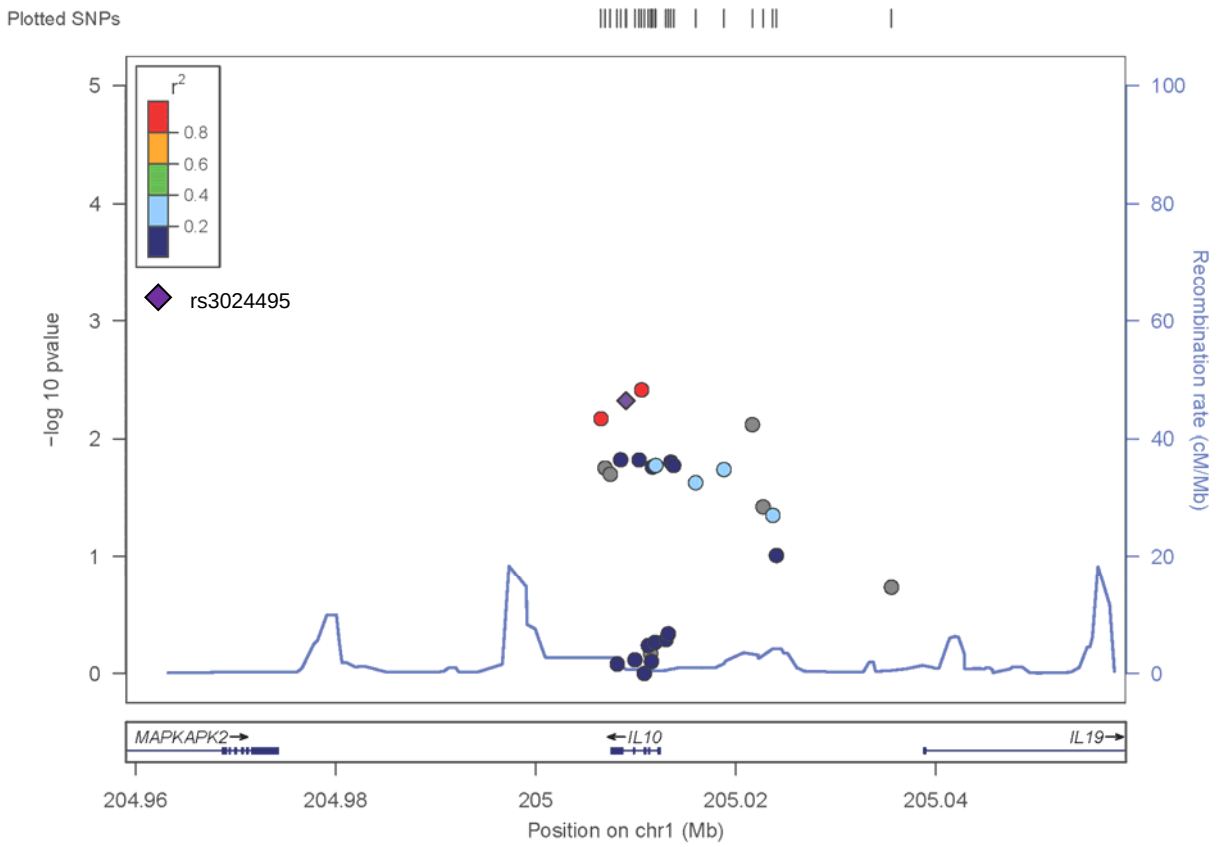

Figure 5: Association results for rs3024495 and surrounding SNPs

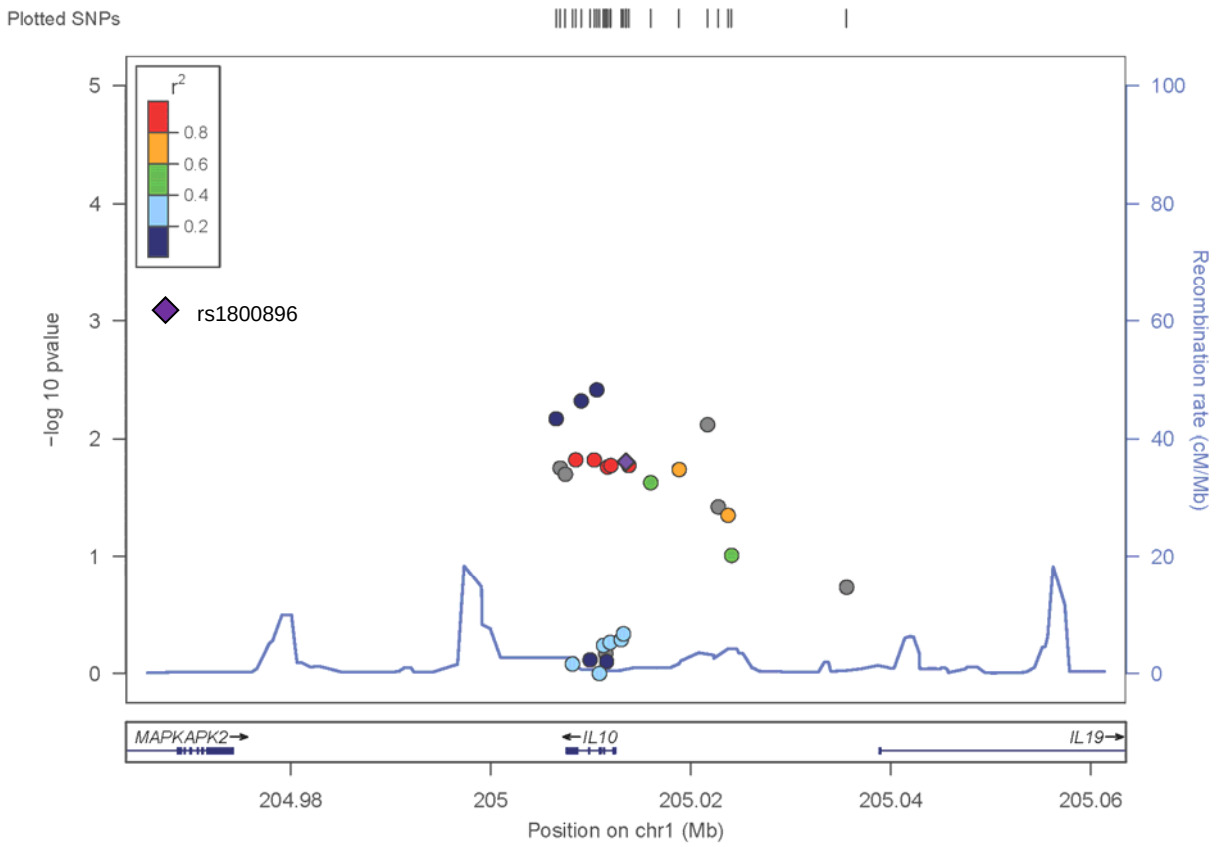

Figure 6: Association results for rs1800896 and surrounding SNPs

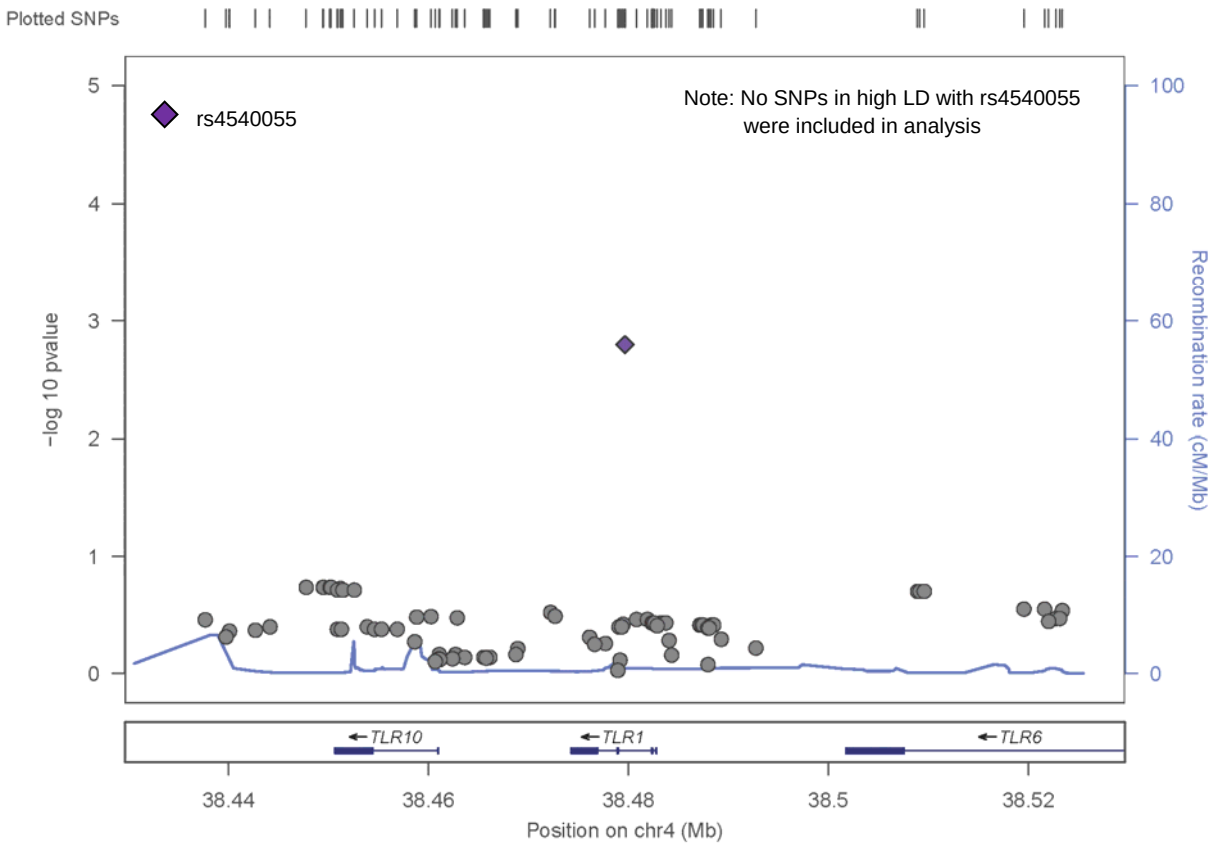

Figure 7: Association results for rs4540055 and surrounding SNPs

Bibliography

1. Helal, I. *et al.* Acute renal failure following allogeneic hematopoietic cell transplantation: incidence, outcome and risk factors. *Saudi J Kidney Dis Transpl* **22**, 437-43 (2011).
2. Kagoya, Y., Kataoka, K., Nannya, Y. & Kurokawa, M. Pretransplant predictors and posttransplant sequels of acute kidney injury after allogeneic stem cell transplantation. *Biol Blood Marrow Transplant* **17**, 394-400 (2011).
3. Liu, H. *et al.* A multicenter, retrospective study of acute kidney injury in adult patients with nonmyeloablative hematopoietic SCT. *Bone Marrow Transplant* **45**, 153-8 (2010).
4. Kist-van Holthe, J.E. *et al.* Prospective study of renal insufficiency after bone marrow transplantation. *Pediatr Nephrol* **17**, 1032-7 (2002).
5. Letourneau, I. *et al.* Acute renal failure in bone marrow transplant patients admitted to the intensive care unit. *Nephron* **90**, 408-12 (2002).
6. Lu, J.C. *et al.* Searching for genes that matter in acute kidney injury: a systematic review. *Clin J Am Soc Nephrol* **4**, 1020-31 (2009).
7. Stafford-Smith, M. *et al.* Association of genetic polymorphisms with risk of renal injury after coronary bypass graft surgery. *Am J Kidney Dis* **45**, 519-30 (2005).
8. Alge, J.L. *et al.* Urinary angiotensinogen predicts adverse outcomes among acute kidney injury patients in the intensive care unit. *Crit Care* **17**, R69 (2013).
9. Johnson, A.D. *et al.* SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* **24**, 2938-9 (2008).
10. Washington, U.o. Genome Variation Server 137, version 8. (2012).
11. Englberger, L. *et al.* Clinical accuracy of RIFLE and Acute Kidney Injury Network (AKIN) criteria for acute kidney injury in patients undergoing cardiac surgery. *Crit Care* **15**, R16 (2011).
12. Bellomo, R. *et al.* Acute renal failure - definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. *Crit Care* **8**, R204-12 (2004).
13. Chien, J.W. *et al.* Evaluation of published single nucleotide polymorphisms associated with acute GVHD. *Blood* **119**, 5311-9 (2012).
14. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-9 (2006).
15. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5**, e1000529 (2009).
16. Howie, B., Marchini, J. & Stephens, M. Genotype imputation with thousands of genomes. *G3 (Bethesda)* **1**, 457-70 (2011).
17. Purcell, S. Plink, version 1.07. (2009).
18. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
19. Team, R.C. R Foundation for Statistical Computing. (Vienna, Austria., 2012).
20. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336-7 (2010).
21. Wattanathum, A., Manocha, S., Groshaus, H., Russell, J.A. & Walley, K.R. Interleukin-10 haplotype associated with increased mortality in critically ill patients with sepsis from pneumonia but not in patients with extrapulmonary sepsis. *Chest* **128**, 1690-8 (2005).
22. Banchereau, J., Pascual, V. & O'Garra, A. From IL-2 to IL-37: the expanding spectrum of anti-inflammatory cytokines. *Nat Immunol* **13**, 925-31 (2012).

23. Westenfelder, C. Programmed anti-inflammatory macrophages protect against AKI and promote repair through trophic actions. *Kidney Int* **81**, 939-41 (2012).
24. Jin, Y. *et al.* Interleukin-10 deficiency aggravates kidney inflammation and fibrosis in the unilateral ureteral obstruction mouse model. *Lab Invest* (2013).
25. Sanchez, R., Levy, E., Costea, F. & Sinnett, D. IL-10 and TNF-alpha promoter haplotypes are associated with childhood Crohn's disease location. *World J Gastroenterol* **15**, 3776-82 (2009).
26. Wang, A.H. *et al.* The effect of IL-10 genetic variation and interleukin 10 serum levels on Crohn's disease susceptibility in a New Zealand population. *Hum Immunol* **72**, 431-5 (2011).
27. Chang, C.F. *et al.* Gene polymorphisms of interleukin-10 and tumor necrosis factor-alpha are associated with contrast-induced nephropathy. *Am J Nephrol* **37**, 110-7 (2013).
28. Anders, H.J., Banas, B. & Schlondorff, D. Signaling danger: toll-like receptors and their potential roles in kidney disease. *J Am Soc Nephrol* **15**, 854-67 (2004).
29. Pino-Yanes, M. *et al.* Common variants of TLR1 associate with organ dysfunction and sustained pro-inflammatory responses during sepsis. *PLoS One* **5**, e13759 (2010).
30. Cheng, C.H., Lee, Y.S., Chang, C.J. & Lin, T.Y. Genetic polymorphisms in Toll-like receptors among pediatric patients with renal parenchymal infections of different clinical severities. *PLoS One* **8**, e58687 (2013).
31. Lee, J.S. *et al.* Toll-like receptor 1 gene polymorphisms in childhood IgA nephropathy: a case-control study in the Korean population. *Int J Immunogenet* **38**, 133-8 (2011).
32. Medzhitov, R. Toll-like receptors and innate immunity. *Nat Rev Immunol* **1**, 135-45 (2001).
33. Strle, K., Shin, J.J., Glickstein, L.J. & Steere, A.C. Association of a Toll-like receptor 1 polymorphism with heightened Th1 inflammatory responses and antibiotic-refractory Lyme arthritis. *Arthritis Rheum* **64**, 1497-507 (2012).
34. Bajwa, A., Kinsey, G.R. & Okusa, M.D. Immune mechanisms and novel pharmacological therapies of acute kidney injury. *Curr Drug Targets* **10**, 1196-204 (2009).
35. White, L.E. & Hassoun, H.T. Inflammatory Mechanisms of Organ Crosstalk during Ischemic Acute Kidney Injury. *Int J Nephrol* **2012**, 505197 (2012).
36. Tadagavadi, R.K. & Reeves, W.B. Endogenous IL-10 attenuates cisplatin nephrotoxicity: role of dendritic cells. *J Immunol* **185**, 4904-11 (2010).
37. Finlay, S. *et al.* Identification of risk factors associated with acute kidney injury in patients admitted to acute medical units. *Clin Med* **13**, 233-8 (2013).
38. Weir, M.R., Aronson, S., Avery, E.G. & Pollack, C.V., Jr. Acute kidney injury following cardiac surgery: role of perioperative blood pressure control. *Am J Nephrol* **33**, 438-52 (2011).
39. Curtis, C. *et al.* The pitfalls of platform comparison: DNA copy number array technologies assessed. *BMC Genomics* **10**, 588 (2009).
40. Kim, S.Y., Kim, J.H. & Chung, Y.J. Effect of Combining Multiple CNV Defining Algorithms on the Reliability of CNV Calls from SNP Genotyping Data. *Genomics Inform* **10**, 194-9 (2012).
41. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G.R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* **44**, 955-9 (2012).
42. Gaudino, M. *et al.* Genetic control of postoperative systemic inflammatory reaction and pulmonary and renal complications after coronary artery surgery. *J Thorac Cardiovasc Surg* **126**, 1107-12 (2003).
43. Popov, A.F. *et al.* Relation between renal dysfunction requiring renal replacement therapy and promoter polymorphism of the erythropoietin gene in cardiac surgery. *Artif Organs* **34**, 961-8 (2010).

44. Popov, A.F. *et al.* The eNOS 786C/T polymorphism in cardiac surgical patients with cardiopulmonary bypass is associated with renal dysfunction. *Eur J Cardiothorac Surg* **36**, 651-6 (2009).
45. Alam, A. *et al.* Phenylethanolamine N-methyltransferase gene polymorphisms and adverse outcomes in acute kidney injury. *Nephron Clin Pract* **114**, c253-9 (2010).
46. Chew, S.T. *et al.* Preliminary report on the association of apolipoprotein E polymorphisms, with postoperative peak serum creatinine concentrations in cardiac surgical patients. *Anesthesiology* **93**, 325-31 (2000).
47. MacKensen, G.B. *et al.* Preliminary report on the interaction of apolipoprotein E polymorphism with aortic atherosclerosis and acute nephropathy after CABG. *Ann Thorac Surg* **78**, 520-6 (2004).
48. Kolyada, A.Y. *et al.* A genetic variant of hypoxia-inducible factor-1alpha is associated with adverse outcomes in acute kidney injury. *Kidney Int* **75**, 1322-9 (2009).