

Received: 2019.05.27
Accepted: 2019.08.20
Published: 2019.10.04

Single Nucleotide Polymorphisms of Ubiquitin-Related Genes were Associated with Allograft Fibrosis of Renal Transplant Fibrosis

Authors' Contribution:
Study Design A
Data Collection B
Statistical Analysis C
Data Interpretation D
Manuscript Preparation E
Literature Search F
Funds Collection G

ABCDEF 1 **Zeping Gui***
AB 2 **Wencheng Li***
CD 1 **Shuang Fei**
CD 3 **Miao Guo**
CD 1 **Hao Chen**
CD 1 **Li Sun**
CD 1 **Zhijian Han**
CD 1 **Jun Tao**
CD 1 **Xiaobin Ju**
CD 1 **Haiwei Yang**
CD 3 **Ji-Fu Wei**
CDFG 1 **Ruoyun Tan**
CDG 1 **Min Gu**

1 Department of Urology, The First Affiliated Hospital with Nanjing Medical University, Nanjing, Jiangsu, P.R. China
2 Department of Urology, Nanjing First Hospital, Nanjing Medical University, Nanjing, Jiangsu, P.R. China
3 Research Division of Clinical Pharmacology, The First Affiliated Hospital with Nanjing Medical University, Nanjing, Jiangsu, P.R. China

* Zeping Gui and Wencheng Li are co-first authors

Corresponding Authors:
Source of support:

Min Gu, e-mail: lancetgu@aliyun.com, Ruoyun Tan, e-mail: tanruoyun112@vip.sina.com
Departmental sources

Background: Interstitial fibrosis and tubular atrophy (IF/TA) have been recognized as crucial factors contributing to graft loss resulting from chronic renal allograft injuries. Recent studies have indicated a significant association between the progression of organ fibrosis and single nucleotide polymorphisms (SNPs) found on certain genes. Our research sought to understand these potential associations and detect the potential impact of SNPs on ubiquitin-related genes related to allograft fibrosis in kidney transplant recipients.


Material/Methods: There were 200 patients enrolled in this study, from which samples were extracted for total DNA. Targeted next-generation sequencing was used to detect SNPs on 9 genes (*FBXL21*, *PIAS1/2*, *SUMO1/2/3/4*, *UBE2D1*, and *UBE2I*). Minor allele frequency (MAF) and Hardy-Weinberg equilibrium (HWE) tests were used and followed by linkage disequilibrium analysis. General linear models (GLM) were used to identify significant confounding factors. Finally, multiple inheritance models and haplotype analyses were conducted to explore associations between SNPs and the degree of the severity of renal allograft fibrosis.

Results: In total, 144 SNPs were identified in targeted sequencing. After filtering based on results from MAF and HWE tests, 15 tagger SNPs were selected for further analyses of associations. GLMs indicated that the administration of sirolimus significantly contributed to the degree of severity of allograft fibrosis ($P=0.011$). After adjusting for confounding factors and applying a Bonferroni correction, multiple inheritance model analyses indicated that the recessive model of rs644731 of the *PIAS2* gene was significantly correlated with the occurrence of IF/TA ($P=0.01$). Furthermore, single-locus based analysis of rs644731 did not indicate that it had a positive influence on IF/TA in a degree-dependent manner. Finally, linkage disequilibrium analysis revealed 3 haplotypes all lacking significant correlation with respect to the IF/TA experimental cohort.

Conclusions: We are the first to reveal that mutations of rs644731 in the *PIAS2* gene were significantly correlated with the progression of IF/TA in kidney transplant recipients.

MeSH Keywords: Kidney Transplantation • Polymorphism, Single Nucleotide • Ubiquitins

Full-text PDF: <https://www.annalsoftransplantation.com/abstract/index/idArt/917767>

 2793

 10

 1

 18



Background

Kidney transplantation is a meritorious treatment of choice for most patients experiencing end-stage renal disease. Although great progress has been made in the effort to use immunosuppressive agents and HLA matching to address renal allograft fibrosis, the onset and progression of this affliction remains a major determining factor which impacts short-term function and long-term outcome for recipients and their allograft status [1]. The pathogenic dynamics of renal allograft fibrosis are driven by a suite of factors including genetic risk, immune system and inflammatory responses, and oxidative stressors, all which may cause subsequent injury of glomeruli, renal interstitium, and tubules [2]. Although previous research has investigated related roles of polymorphisms, accurate genetic *loci* were insufficient to provide a reliable description related to renal fibrosis after transplantation. Furthermore, gene mutations in renal epithelial cells can lead to the activation of epithelial-mesenchymal transition (EMT) inducing renal fibrosis, ultimately indicating that genetic factors are important considerations in the dynamics of pathogenesis of EMT and renal interstitial fibrosis [3].

Since susceptibility genes can provide insights into principal pathological mechanisms, genetic-based analyses of disease can be powerfully insightful for exploring its pathogenesis. The appearance of fibrosis is mainly caused by the emergence of mechanocytes while the degree of fibrosis varies among individuals as a function of polymorphisms within regulatory regions of genes that play roles in transcriptional activation. Thus far, numerous studies have provided support for the importance of the role of single nucleotide polymorphisms (SNPs) in this process, including for coding, noncoding intron, and promoter regions as part of a mix of a larger suite of a number of genes associated with fibrosis disease. It has been reported that rs58542926 on the *TM6SF2* gene is related to hepatic fibrosis [4] and reported that rs738409 on the *PNPLA3* gene is related to liver allograft fibrosis [5]. Furthermore, the chronic renal disease variant rs4730751 is found on the *CAV1* gene and is related to arterial fibrosis [6]. Additionally, the IL-18-607A/C (rs1946518) promoter polymorphism is reported to be correlated with IgA nephropathy and subsequent renal fibrosis [7].

Consistent with these results, a previous study undertaken by several authors of this manuscript determined that tumor necrosis factor (TNF)- α induced EMT via the TNF- α /Akt/Smurf2 signaling pathways [8]. The *Smurf2* gene encodes the E3 ubiquitin-protein ligase Smurf2 in humans, which provides theoretical support for correlations between ubiquitin-related genes and renal fibrosis. Thus, we sought to investigate the association and influences of SNPs with and upon ubiquitin-related genes related to renal allograft interstitial fibrosis. We examined 1 cohort enrolled from our single renal transplant center.

Material and Methods

Ethics statements

Study design, patient enrollment, and procedural protocols were reviewed and approved by the local Ethics Committee of the First Affiliated Hospital of Nanjing Medical University (2016-SR-029). All kidney transplant recipients confirmed their understanding of procedures, protocols, and risks as described and provided through written informed consent. The procedures in our study abided by the ethical standards of the Declarations of Helsinki and Istanbul. All kidney transplant recipients received transplants from donors who had experienced cardiac death

Study design and population

We used a single-center, retrospective, cohort study based approach to explore the influence of SNPs on the transforming growth factor beta (TGFB) signaling pathway and genes (including *FBXL21*, *PIAS1/2*, *SUMO1/2/3/4*, *UBE2D1*, and *UBE2I*) related to the progression of allograft fibrosis in renal transplant recipients. Two hundred kidney transplant recipients, who received renal transplants from February 1, 2015 to September 1, 2018 at the renal transplantation center of the First Affiliated Hospital of Nanjing Medical University, were enrolled in this study. The average follow-up time with patients in our research was 1555 ± 1054 days, and all follow-ups fell between 3 and 5 years after transplantation, and no significant graft failure or decline of renal function was observed. Patients with rejection episodes and delayed graft function (DGF) after transplantation were all eliminated from further inclusion in the study groups. Specifics for inclusion and exclusion criteria were described in methods outlined in greater detail in a previous related study [9].

Clinical data including age, gender, height, and immunosuppressive protocols were independently determined by one of the authors, Zeping Gui.

Investigative biopsies were performed for all transplant recipients enrolled in our study, and histological analyses were conducted by 2 independent nephrologists (Hao Chen and Li Sun) through the use of hematoxylin-eosin (HE), periodic acid Schiff (PAS), Masson, and immunohistological staining following guidelines in Banff 2015 [10]. Allograft fibrosis severity and type/grade were scored using metrics related to the degree of interstitial infiltration and intimal arteritis and by following guidelines in Banff 2015 [10].

Immunosuppressive protocols

Immunosuppressive protocols for all patients included 3 or 4 differently composed treatments of drugs: cyclosporin A or

tacrolimus, combined with mycophenolate mofetil (MMF) and prednisone, with or without sirolimus which dosage schedules were adjusted according to serum creatinine levels and drug concentrations. Each patient was treated with immunotherapy on the fourth day before, and fourth day after surgery. Detailed information and methodologies for immunosuppressive agent schedules can be found in a previous related study [9].

Sample collection and TS

Detailed information and methodologies for sample collection and TS can be found in a previous related study [9]. Peripheral blood (2 mL) of each patient was used for DNA extraction. We quantitatively analyzed genomic DNA (gDNA) concentration and purity and assessed gene integrity by using agarose gel electrophoresis. We selected from a randomized pool containing upstream and downstream oligonucleotides and gDNA hybrids specific to target regions of interest. We then fragmented gDNA and amplified the adapter-ligated DNA by selective, limited-cycle polymerase chain reaction (PCR). We denatured captured libraries and loaded them into an Illumina cBot instrument following manufacturer protocols. Subsequently, we analyzed sequencing data based around available data for the human reference sequence UCSC hg19 assembly (NCBI build 37.2), using Genome Analysis Tool Kit, Picard Software, and dbSNP 132. We also detected putative somatic variant cells with 2 separate programs: MuTect 1.1.5 and VarScan 2.3.6.

Statistical analysis

Data are presented as mean±standard deviation (SD), except where stated otherwise. We explored minor allele frequency (MAF) and Hardy-Weinberg equilibrium (HWE) by using R package genetics (genetics: Population Genetics, R package version 1.3.8.1.). Linkage disequilibrium (LD) blocks were analyzed using Haploview version 4.2 (Broad Institute, Cambridge, MA, USA). General linear models (GLMs) were used to determine the importance and influence of clinical variables on AR (acute rejection). We used the R Statistics Package SNPassoc (SNPs-based whole genome association studies; R Package Version 1.9-2.) to examine 5 multiple inheritance models based on the different treatments of sirolimus concentrations. These 5 models included: codominant model 1 (major allele homozygotes versus heterozygotes), codominant model 2 (major allele homozygotes versus minor allele homozygotes), dominant model (major allele homozygotes versus minor allele homozygotes plus heterozygotes), recessive model (major allele homozygotes plus heterozygotes versus minor allele homozygotes), overdominant model (heterozygotes versus major allele homozygotes plus minor allele homozygotes), and a log-additive model (major allele homozygotes versus heterozygotes versus minor allele homozygotes). We used chi-square analyses to examine the levels of variance and compare Banff

Table 1. Basic demographics of patients in this cohort.

| Characteristics | Value |
|------------------------------------|------------|
| Case number | 200 |
| Age (years, mean±SD) | 44.59±4.09 |
| Gender (Male/Female) | 119/81 |
| PRA before renal transplant (%) | 0.00 |
| Primary/Secondary renal transplant | 200 |
| Primary renal transplant | 200 |
| Secondary renal transplant | 0 |
| HLA mismatching | 4.56±0.34 |
| Type of donor | 200 |
| Living-related | 24 |
| DCD | 176 |
| Administration of sirolimus (%) | 12.00 |
| IFTA (n) | 69 |
| Mild | 32 |
| Moderate | 24 |
| Severe | 13 |

SD – standard deviations; PRA – panel reactive antibody; DCD – donor after cardiac death; IFTA – interstitial fibrosis and tubular atrophy.

scores when considering 2 or 3 of the selected most important genotypes. All data in our study were analyzed using SPSS Software Version 13.0 (SPSS Inc., Chicago, IL, USA) and $P < 0.05$ was considered statistically significant.

Results

Patient demographics

A total of 200 patients were enrolled including 119 males, and 81 females. A greater proportion (12%) of patients were treated with sirolimus and none of the panel reactive antibody was found in this cohort before transplantation. Demographics for this patient group are presented in Table 1.

Tagger SNP selection

A total of 15 SNPs were identified. We extracted ubiquitin-related genes and determined the levels of genetic association between 9 associated gene SNPs (*FBXL21*, *PIAS1/2*, *SUMO1/2/3/4*, *UBE2D1*, and *UBE2I*) as well as measures of allograft fibrosis. Our use of reference and alternating alleles helped to support a robust approach and the resultant observational evidence used for genotype analyses (Supplementary Table 1).

Table 2. Influence of confounding factors on the outcomes of allograft fibrosis by general linear model in this cohort.

| Confounding factors | F value | P value |
|---------------------------------|---------|---------|
| Gender | 0.313 | 0.7547 |
| Age | 0.429 | 0.6686 |
| Weight | 1.175 | 0.2414 |
| ISD protocol | 1.557 | 0.1210 |
| Duration after renal transplant | 0.429 | 0.6686 |
| Administration of Sirolimus | 2.477 | 0.0134 |

ISD – immunosuppressive drugs; DGF – delayed graft function.

We defined common variants as those with MAF >0.05, and we set a threshold of 0.05 for HWE values. HWE, MAF, and LD analyses revealed 15 tagger SNPs (rs7283639, rs2838697, rs13050872, rs9306116, rs237025, rs237024, rs237023, rs73288305, rs74377516, rs75362994, rs3737448, rs644731, rs113887072, rs72915074, and rs2066913) that were deemed as statistically frequent SNPs (tSNPs) (MAF >0.05), whereas remainders identified were rare (Supplementary Table 2).

Confounding factor analysis and multiple inheritance model analysis

Based upon relative strength of their association with allograft fibrosis we added markers sequentially as continuous variables

to a model using the adjusted confounding factors. In this cohort, confounding factors of patients who were administrated sirolimus were significant ($P=0.011$) in predicting the incidence of fibrosis, compared with other factors that were not ($P>0.05$) based upon GLM results. In sum, this suggested that sirolimus had an influence on the outcomes of allograft fibrosis (Table 2).

After applying a Bonferroni correction to adjust the different sirolimus treatments (adjusted P value=0.005) we conducted multiple inheritance model analyses. Synonymous SNP rs644731 was significantly associated with allograft fibrosis [Table 3; odds ratio (OR)=4.42; 95% confidence interval (CI)=1.32–13.64, $P=0.01$ recessive model; OR=6.57, 95% CI=2.99–14.45, $P=1.54 \times 10^{-7}$ dominant model; OR=3.5, 95% CI=1.77–6.89, $P=0.00017$ overdominant model; OR=5.95, 95% CI=2.66–13.3, $P=5.08 \times 10^{-7}$ codominant model; and OR=4.28, 95% CI=2.34–7.83, $P=1.73 \times 10^{-7}$ additive model], suggesting that the risk of allograft fibrosis was strongly correlated with the rs644731 locus, compared with other non-significant tagger SNPs ($P>0.005$; Supplementary Table 3).

Furthermore, we selected 3 genotypes and compared differences between the degrees of severity of interstitial fibrosis and tubular atrophy (IF/TA) for each of the 3 groups. Results indicated no significant differences ($P=0.38$) for degrees of IF/TA including mild, moderate, and severe among the CC, CT, and TT genotypes (Table 4). However, we observed that when one T allele appeared (CT genotype), a moderate degree of increase in the severity of IF/TA was noted than in comparison

Table 3. Results of multiple inheritance models in rs644731 adjusted by the administration of sirolimus in 5 models.

| rs644731 | OR | Lower 95% CI | Upper 95% CI | P value* |
|--------------------|------|--------------|--------------|----------|
| Recessive model | 4.24 | 1.32 | 13.64 | 0.01 |
| Dominant model | 6.57 | 2.99 | 14.45 | 1.54E-07 |
| Overdominant model | 3.5 | 1.77 | 6.89 | 0.000169 |
| Codominant model | 5.95 | 2.66 | 13.3 | 5.08E-07 |
| Additive model | 4.28 | 2.34 | 7.83 | 1.73E-07 |

OR – odds ratio; CI – confidential interval. * Associations were considered significant if P value is less than 0.005 (Bonferroni corrected- P value).

Table 4. Distributions and analysis of rs644731 in patients with IFTA.

| Genotype | IFTA degree | | | P value |
|----------|-------------|----------|--------|---------|
| | Mild | Moderate | Severe | |
| CC | 10 | 5 | 3 | 0.38 |
| CT | 18 | 17 | 6 | |
| TT | 4 | 2 | 4 | |

IFTA – interstitial fibrosis and tubular atrophy.

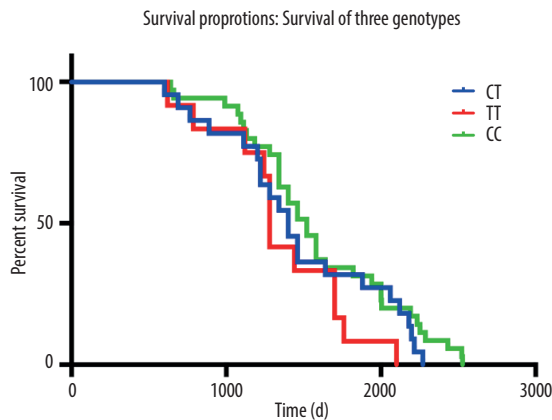


Figure 1. Results of Linkage disequilibrium and haplotypes of all detected single nucleotide polymorphisms.

to the no T allele group. Furthermore, samples with 2 T alleles (TT genotypes) showed a significant increase in the percentage of the degree of severe IF/TA and this was greater than when compared to the other 2 treatments. This result indicated a tendency for the degree of severity of IF/TA to rise with an increase in the number T alleles.

We also performed time-dependent analyses to explore relationships between 3 genotypes and the time from transplantation to biopsy. Results from this survival analysis indicated no significant differences among 3 genotypes which was consistent with previous findings supportive of the idea that duration after renal transplant is not significantly associated with allograft fibrosis (Table 2, Supplementary Table 4).

LD Analysis and haplotype analysis

Haplotype structure and plots of pairwise LD and gene structure are displayed in Figure 1. Each gene was categorized into 3 haplotypes: block 1 (SNPs 4–5: rs7283639, rs2838697), block 2 (SNPs 8–10: rs73288305, rs74377516, rs75362994), and block 3 (SNPs 13–14: rs237025, rs237024). Associations between haplotype frequencies are given in Supplementary Table 5.

Haplotype association analyses results using the 3 blocks are listed in Supplementary Table 6. Associations with allograft fibrosis were not statistically significant for block 1 [likelihood ratio (LR)=2.74, $P=0.74$], block 2 (LR=3.46, $P=0.18$), or for block 3 (LR=0.77, $P=0.68$).

Discussion

Previous studies have only focused on a few allograft Ubiquitin-related gene variants, which may be insufficient to capture the

full effects of susceptibility related genes. Although such hypotheses have become long-established through randomized trials, evidence for renal allograft fibrosis has been less definite. Our approach was retrospective in targeting the main genes associated with allograft fibrosis. We accordingly used multiple inheritance models and haplotype analyses to test the hypothesis that sigSNP might be an underlying locus related to IF/TA severity after transplantation. We presented the novel finding that the rs644731 variant of the *PIAS2* gene exhibited a statistically significant association with renal allograft fibrosis.

A summary of a genome-wide association study concluded that approximately 80% of trait-associated SNPs are located in non-coding regions [11]. Results from the Encyclopedia of DNA Elements Consortium (ENCODE) attributed important regulatory functions to these noncoding intronic *loci* within the human genome [12]. Rs644731 is an intron of the *PIAS2* gene which was found to lack significant linkage disequilibrium. However, rs644731 expression was high for the LD region between an intron SNP (rs737448) and an exon SNP (rs113887072) and might be strongly linked with a potential functional locus exerting a molecular influence on the dynamics of *PIAS2* related gene transcription. Alternative splicing of introns within a gene can act to introduce greater variability in protein sequences translated from a single gene and result in more than just a single unique precursor mRNA transcript with accordingly multiple associated functions. The dynamics of the control of alternative RNA splicing involves a complex network of signaling molecules that respond to a wide range of intracellular and extracellular stimuli [13]. Correspondingly, some introns can enhance expression for the gene containing them through a process known as intron-mediated enhancement (IME) [14]. Further experiments should seek to identify the dynamics of IME that cause a resultant enhancement of expression. So far, one of the most common and best understood mechanisms and approaches is to move the intron upstream from the starting point of transcription, thus removing it and its influence from the final transcript product. If such a change indicates that the intron cannot or has no longer enhanced expression, then inclusion of the intron in the transcript is a vitally important consideration, and the intron may help to induce or solely cause IME [15,16]. Rs644731 is potentially such a type of an intron and may be important in the pathways of regulation of transcription and gene expression, as well as in the sequential steps of the allograft fibrosis pathway. This finding is compatible with the hypothesis that the same locus, such as one intron, can have a similar regulatory type effect on genetic expression. However, we did not find significant differences between the degree of IF/TA (mild, moderate, and severe) and the varied sigSNP genotypes. It is worth noting that with a higher number of T alleles, that the percentage of cases with severe IF/TA displayed an obvious increase. This tendency of increases with T alleles may

illustrate that the difference might be significant and should be followed up with experimentation using an expanded sample size. Thus, we helped to determine that this locus may represent a potential therapeutic target which may possibly help to reduce the progression of IF/TA in patients after renal transplantations.

Haplotype analyses suggested that the association between locus blocks and allograft fibrosis was not statistically significant. However, our results did indicate a high-risk tendency for patients with renal IF/TA. This possibly indicates that a specific haplotype might be related to higher levels of *PIAS2* gene transcription activity. The resultant higher levels of transcription might play an important role in the pathway leading to fibroblast proliferation and the progression of fibrosis. Accordingly, we also investigated other ubiquitin-related genes, but failed to identify an association between gene mutation and the progression of IF/TA. However, results from a previous related study by authors of this manuscript found that having a 1 or 2 copies of the risk allele appeared to significantly increase the ubiquitin-related gene *Smurf2* transcript level in biopsied tissues from the allograft treatment group [17]. *Smurf2* mRNA expression increased with the expression of TGF- β 1 in the early stages of renal fibrosis and development and it is possible that various methods to induce and enhance *Smurf2* expression could be used to some extent in order to slow or prevent the progression of allograft fibrosis [18]. Other ubiquitin-related genes were insignificant in our related analyses which may have resulted because the signal pathways associated

with them are yet to be clearly identified. With an increased sample size more potential *loci* might be detected based on an expanded next-generation sequencing approach building from the research we completed herein.

Conclusions from our study while important are still limited in many respects. For example, despite that we adjusted for confounding factors related to the development and progression of allograft fibrosis, sample size was a key factor likely to have impacted our final results and conclusions. Furthermore, current statistically based epidemiological literature does not unanimously concur on when and how to make corrections for confounding factors. Thus, additional studies are needed to clarify roles of genetic polymorphisms on ubiquitin-related genes in renal transplantation patients.

Conclusions

In conclusion, we found that mutations of rs644731 in the *PIAS2* gene were significantly associated with the risk of allograft fibrosis following renal transplantation. Although additional research is needed, our results support the need for a cautious approach as the dynamics of a multifactorial and multigenic disease like allograft fibrosis after renal transplantation are complex and not yet fully understood. Nevertheless, our study provided novel information and a potential new direction for a comprehensive research-based analysis of the importance of and roles that SNPs play in fibrosis-related diseases.

Supplementary Data

Supplementary Table 1. Detailed information of SNP detected in the Ubiquitin-related genes in the cohort.

| Chromosome | Location | Reference allele | Alternation allele | Gene name | Function | Gene detail | Avsnp144 |
|------------|-----------|------------------|--------------------|-----------|----------|-------------|-------------|
| chr5 | 135272403 | A | C | FBXL21 | Exonic | | rs573196792 |
| chr5 | 135272692 | C | T | FBXL21 | Intronic | | rs150784167 |
| chr5 | 135272823 | G | A | FBXL21 | Intronic | | rs17702049 |
| chr5 | 135273078 | C | A | FBXL21 | Intronic | | rs544746984 |
| chr5 | 135273177 | T | C | FBXL21 | Exonic | | . |
| chr5 | 135273298 | C | A | FBXL21 | Intronic | | rs183239904 |
| chr5 | 135273370 | G | T | FBXL21 | Intronic | | rs10052673 |
| chr5 | 135276049 | G | A | FBXL21 | Intronic | | rs31549 |
| chr5 | 135276205 | G | C | FBXL21 | Exonic | | rs76075237 |
| chr5 | 135276314 | C | T | FBXL21 | Exonic | | rs40986 |
| chr5 | 135276701 | T | C | FBXL21 | Intronic | | rs31548 |
| chr5 | 135276814 | G | A | FBXL21 | Exonic | | rs2066913 |
| chr5 | 135276847 | T | C | FBXL21 | Exonic | | rs31547 |
| chr5 | 135277204 | C | G | FBXL21 | Exonic | | rs530876112 |

| Chromosome | Location | Reference allele | Alternation allele | Gene name | Function | Gene detail | Avsnp144 |
|------------|----------|------------------|--------------------|-----------|----------|---|-------------|
| chr15 | 68346778 | T | G | PIAS1 | Intronic | | . |
| chr15 | 68346778 | T | C | PIAS1 | Intronic | | rs537399079 |
| chr15 | 68348207 | C | T | PIAS1 | Intronic | | . |
| chr15 | 68349843 | C | G | PIAS1 | Intronic | | rs182709174 |
| chr15 | 68349893 | A | G | PIAS1 | Intronic | | rs1489598 |
| chr15 | 68349918 | T | C | PIAS1 | Intronic | | rs79833223 |
| chr15 | 68349925 | G | A | PIAS1 | Intronic | | rs183625509 |
| chr15 | 68378937 | G | A | PIAS1 | Exonic | .PIAS1: NM_016166: exon2: c.G318A: p.S106S | rs145053928 |
| chr15 | 68379138 | C | T | PIAS1 | Intronic | | rs186753486 |
| chr15 | 68434378 | C | T | PIAS1 | Splicing | NM_016166: exon3: c.554+10C>T. | rs750502048 |
| chr15 | 68434379 | G | C | PIAS1 | Intronic | | rs117588299 |
| chr15 | 68434428 | A | G | PIAS1 | Intronic | | . |
| chr15 | 68434744 | G | A | PIAS1 | Intronic | | rs145280358 |
| chr15 | 68439101 | A | G | PIAS1 | Intronic | | rs576273919 |
| chr15 | 68445850 | C | G | PIAS1 | Intronic | | . |
| chr15 | 68445912 | T | A | PIAS1 | Intronic | | rs11633620 |
| chr15 | 68468049 | C | T | PIAS1 | Exonic | .PIAS1: NM_016166: exon10: c.C1244T: p.P415L | . |
| chr15 | 68468098 | A | G | PIAS1 | Exonic | .PIAS1: NM_016166: exon10: c.A1293G: p.G431G | rs113555272 |
| chr15 | 68468136 | C | T | PIAS1 | Intronic | | . |
| chr15 | 68468695 | T | C | PIAS1 | Intronic | | rs12438361 |
| chr15 | 68473443 | T | C | PIAS1 | Intronic | | rs3759823 |
| chr15 | 68473808 | G | A | PIAS1 | Intronic | | rs75673552 |
| chr15 | 68476069 | A | G | PIAS1 | Intronic | | . |
| chr15 | 68479858 | G | C | PIAS1 | Intronic | | . |
| chr15 | 68480086 | A | G | PIAS1 | Exonic | .PIAS1: NM_016166: exon14: c.A1869G: p.E623E | rs191408288 |
| chr15 | 68480226 | A | G | PIAS1 | Utr3 | NM_016166: c.*53A>G. | rs372957210 |
| chr18 | 44395368 | T | C | PIAS2 | Intronic | | rs3737448 |
| chr18 | 44398111 | C | T | PIAS2 | Utr3 | NM_173206: c.*229G>A. | rs10502879 |
| chr18 | 44398183 | A | T | PIAS2 | Utr3 | NM_173206: c.*157T>A. | . |
| chr18 | 44398464 | G | A | PIAS2 | Intronic | | rs149740503 |
| chr18 | 44398557 | A | T | PIAS2 | Intronic | | rs183321210 |
| chr18 | 44401052 | G | A | PIAS2 | Intronic | | rs146442641 |
| chr18 | 44407993 | G | A | PIAS2 | Exonic | .PIAS2: NM_004671: exon11: c.C1437T: p.D479D,PIAS2: NM_173206: exon11: c.C1437T: p.D479D | rs35451178 |
| chr18 | 44409581 | C | A | PIAS2 | Intronic | | rs72907142 |
| chr18 | 44409617 | G | A | PIAS2 | Intronic | | . |
| chr18 | 44416287 | G | A | PIAS2 | Intronic | | . |

| Chromosome | Location | Reference allele | Alternation allele | Gene name | Function | Gene detail | Avsnp144 |
|------------|-----------|------------------|--------------------|-----------|----------|--|-------------|
| chr18 | 44416608 | C | T | PIAS2 | Intronic | | rs72907148 |
| chr18 | 44423925 | A | G | PIAS2 | Intronic | | rs77040088 |
| chr18 | 44424122 | A | G | PIAS2 | Intronic | | rs188584114 |
| chr18 | 44424939 | G | A | PIAS2 | Intronic | | rs372024400 |
| chr18 | 44424995 | A | G | PIAS2 | Intronic | | rs150885589 |
| chr18 | 44426800 | T | C | PIAS2 | Exonic | .PIAS2: NM_004671: exon6: c.A731G: p.Y244C,PIAS2: NM_173206: exon6: c.A731G: p.Y244C | rs114135676 |
| chr18 | 44426877 | C | T | PIAS2 | Intronic | | rs644731 |
| chr18 | 44435407 | A | G | PIAS2 | Splicing | NM_004671: exon6: c.636-9T>C;NM_173206: exon6: c.636-9T>C. | rs764656048 |
| chr18 | 44470481 | T | C | PIAS2 | Intronic | | rs539782995 |
| chr18 | 44470706 | G | A | PIAS2 | Exonic | .PIAS2: NM_004671: exon2: c.C336T: p.H112H,PIAS2: NM_173206: exon2: c.C336T: p.H112H | rs113887072 |
| chr18 | 44470825 | A | G | PIAS2 | Exonic | .PIAS2: NM_004671: exon2: c.T217C: p.S73P,PIAS2: NM_173206: exon2: c.T217C: p.S73P | . |
| chr18 | 44483961 | T | C | PIAS2 | Intronic | | rs72915074 |
| chr18 | 44496847 | C | T | PIAS2 | Intronic | | rs567309045 |
| chr18 | 44497099 | C | A | PIAS2 | Intronic | | rs530227612 |
| chr18 | 44497124 | C | T | PIAS2 | Intronic | | . |
| chr18 | 44497173 | C | A | PIAS2 | Intronic | | . |
| chr18 | 44497338 | T | G | PIAS2 | Utr5 | NM_004671: c.-30A>C;NM_173206: c.-30A>C. | . |
| chr2 | 203072108 | G | A | SUMO1 | Intronic | | . |
| chr2 | 203072889 | T | C | SUMO1 | Intronic | | . |
| chr2 | 203079307 | T | C | SUMO1 | Intronic | | rs3769817 |
| chr2 | 203096545 | C | A | SUMO1 | Intronic | | rs116081766 |
| chr2 | 203103241 | G | T | SUMO1 | Utr5 | NM_001005781: c.-67C>A;NM_001005782: c.-67C>A;NM_003352: c.-67C>A. | . |
| chr17 | 73177444 | C | A | SUMO2 | Intronic | | rs1471453 |
| chr17 | 73179065 | C | G | SUMO2 | Utr5 | NM_001005849: c.-136G>C;NM_006937: c.-136G>C. | . |
| chr21 | 46226786 | A | G | SUMO3 | Utr3 | NM_001286416: c.*80T>C;NM_006936: c.*80T>C. | rs1051331 |
| chr21 | 46227163 | A | G | SUMO3 | Intronic | | rs2329902 |
| chr21 | 46227968 | T | G | SUMO3 | Intronic | | . |
| chr21 | 46228150 | G | A | SUMO3 | Intronic | | . |
| chr21 | 46228153 | C | T | SUMO3 | Intronic | | . |
| chr21 | 46228154 | G | A | SUMO3 | Intronic | | rs752652207 |
| chr21 | 46228155 | C | T | SUMO3 | intronic | | rs757331290 |

| Chromosome | Location | Reference allele | Alternation allele | Gene name | Function | Gene detail | Avsnp144 |
|------------|-----------|------------------|--------------------|-----------|----------|---|-------------|
| chr21 | 46228157 | T | G | SUMO3 | Intronic | | rs765382230 |
| chr21 | 46228165 | T | C | SUMO3 | Intronic | | rs7283639 |
| chr21 | 46228170 | T | G | SUMO3 | Intronic | | rs188978703 |
| chr21 | 46228243 | T | C | SUMO3 | Intronic | | rs141141907 |
| chr21 | 46228662 | C | T | SUMO3 | Intronic | | . |
| chr21 | 46228930 | C | T | SUMO3 | Intronic | | rs235293 |
| chr21 | 46228945 | C | T | SUMO3 | Intronic | | rs564735586 |
| chr21 | 46228949 | G | C | SUMO3 | Intronic | | . |
| chr21 | 46233836 | C | A | SUMO3 | Exonic | .SUMO3: NM_001286416: exon2: c.G205T: p.V69F | rs2838697 |
| chr21 | 46233863 | G | C | SUMO3 | Exonic | .SUMO3: NM_001286416: exon2: c.C178G: p.L60V | rs13050872 |
| chr21 | 46233866 | T | C | SUMO3 | Exonic | .SUMO3: NM_001286416: exon2: c.A175G: p.S59G | rs9981327 |
| chr21 | 46234079 | T | A | SUMO3 | Intronic | .. | rs9306116 |
| chr6 | 149721690 | G | A | SUMO4 | Exonic | .SUMO4: NM_001002255: exon1: c.G163A: p.V55M | rs237025 |
| chr6 | 149721778 | T | C | SUMO4 | Exonic | .SUMO4: NM_001002255: exon1: c.T251C: p.I84T | rs777445425 |
| chr6 | 149721800 | G | A | SUMO4 | Exonic | .SUMO4: NM_001002255: exon1: c.G273A: p.T91T | rs145312495 |
| chr6 | 149721965 | T | C | SUMO4 | Utr3 | NM_001002255: c.*150T>C | rs237024 |
| chr6 | 149722040 | A | G | SUMO4 | Utr3 | NM_001002255: c.*225A>G | rs237023 |
| chr10 | 60095105 | G | T | UBE2D1 | Intronic | | rs112660736 |
| chr10 | 60121139 | A | G | UBE2D1 | Exonic | .UBE2D1: NM_003338: exon2: c.A66G: p.S22S | rs759280904 |
| chr10 | 60121240 | C | T | UBE2D1 | Intronic | | . |
| chr10 | 60123486 | A | G | UBE2D1 | Intronic | | rs73288305 |
| chr10 | 60123523 | A | G | UBE2D1 | Intronic | | . |
| chr10 | 60124703 | C | T | UBE2D1 | Intronic | | . |
| chr10 | 60127627 | A | G | UBE2D1 | Intronic | | rs531786752 |
| chr10 | 60127639 | G | T | UBE2D1 | Intronic | | rs74377516 |
| chr10 | 60127798 | T | C | UBE2D1 | Intronic | | . |
| chr10 | 60127838 | G | A | UBE2D1 | Intronic | | rs75362994 |
| chr10 | 60128364 | T | C | UBE2D1 | Intronic | | rs3802699 |
| chr10 | 60128583 | A | G | UBE2D1 | Utr3 | NM_001204880: c.*58A>G;NM_003338: c.*58A>G | rs148198083 |
| chr16 | 1363878 | C | T | UBE2I | Intronic | | rs9926183 |
| chr16 | 1363927 | A | G | UBE2I | Intronic | | . |
| chr16 | 1364140 | T | C | UBE2I | Intronic | | rs9941160 |
| chr16 | 1364158 | G | A | UBE2I | Intronic | | . |
| chr16 | 1364251 | C | T | UBE2I | Intronic | | rs201695180 |
| chr16 | 1364281 | T | C | UBE2I | Intronic | | rs4984806 |

| Chromosome | Location | Reference allele | Alternation allele | Gene name | Function | Gene detail | Avsnp144 |
|------------|----------|------------------|--------------------|-----------|----------|--|-------------|
| chr16 | 1364365 | A | G | UBE2I | Exonic | .UBE2I: NM_003345: exon3: c.A138G: p.P46P,UBE2I: NM_194260: exon3: c.A138G: p.P46P,UBE2I: NM_194261: exon3: c.A138G: p.P46P,UBE2I: NM_194259: exon4: c.A138G: p.P46P | rs4610 |
| chr16 | 1365612 | G | A | UBE2I | Intronic | | rs781398317 |
| chr16 | 1365915 | C | T | UBE2I | Intronic | | rs112302601 |
| chr16 | 1365935 | C | G | UBE2I | Intronic | | rs4984807 |
| chr16 | 1365943 | C | T | UBE2I | Intronic | | rs7186045 |
| chr16 | 1365967 | C | T | UBE2I | Intronic | | rs4984808 |
| chr16 | 1369612 | C | T | UBE2I | Intronic | | rs201661304 |
| chr16 | 1369730 | A | G | UBE2I | Intronic | | rs9933497 |
| chr16 | 1369837 | C | T | UBE2I | Intronic | | . |
| chr16 | 1369926 | T | C | UBE2I | Intronic | | rs909915 |
| chr16 | 1370203 | G | A | UBE2I | Exonic | .UBE2I: NM_003345: exon5: c.G252A: p.P84P,UBE2I: NM_194260: exon5: c.G252A: p.P84P,UBE2I: NM_194261: exon5: c.G252A: p.P84P,UBE2I: NM_194259: exon6: c.G252A: p.P84P | rs758216436 |
| chr16 | 1370303 | T | C | UBE2I | Intronic | | rs909916 |
| chr16 | 1370309 | C | G | UBE2I | Intronic | | rs909917 |
| chr16 | 1370383 | C | T | UBE2I | Intronic | | rs148789348 |
| chr16 | 1370575 | G | C | UBE2I | Intronic | | . |
| chr16 | 1370597 | G | C | UBE2I | Intronic | | rs4017786 |
| chr16 | 1370614 | C | G | UBE2I | Intronic | | rs8063770 |
| chr16 | 1370630 | A | G | UBE2I | Intronic | | rs8043720 |
| chr16 | 1370682 | G | A | UBE2I | Intronic | | rs79005361 |
| chr16 | 1370698 | G | A | UBE2I | Intronic | | rs571836605 |
| chr16 | 1370716 | C | T | UBE2I | Intronic | | rs142273742 |
| chr16 | 1370729 | C | A | UBE2I | Intronic | | rs909918 |
| chr16 | 1374513 | G | A | UBE2I | Intronic | | rs2369700 |
| chr16 | 1374524 | A | G | UBE2I | Intronic | | rs761059 |
| chr16 | 1374629 | G | T | UBE2I | Intronic | | . |
| chr16 | 1374656 | A | G | UBE2I | Intronic | | rs761060 |
| chr16 | 1374785 | G | A | UBE2I | Exonic | .UBE2I: NM_003345: exon7: c.G468A: p.A156A,UBE2I: NM_194260: exon7: c.G468A: p.A156A,UBE2I: NM_194261: exon7: c.G468A: p.A156A,UBE2I: NM_194259: exon8: c.G468A: p.A156A | rs762904858 |
| chr16 | 1374818 | A | G | UBE2I | UTR3 | NM_003345: c.*24A>G;NM_194259: c.*24A>G;NM_194260: c.*24A>G;NM_194261: c.*24A>G. | rs8063 |

SNP – single nuclear polymorphism.

Supplementary Table 2. Outcomes of HWE and MAF calculation for all detected SNPs in the cohort.

| Gene name | SNP | Location | MAF | HWE |
|-----------|-------------|-----------|--------|------|
| SUMO1 | . | 203072108 | 0.0025 | 1.00 |
| SUMO1 | . | 203072889 | 0.005 | 1.00 |
| SUMO1 | rs3769817 | 203079307 | 0.1575 | 0.00 |
| SUMO1 | rs116081766 | 203096545 | 0.0025 | 1.00 |
| SUMO1 | . | 203103241 | 0.0025 | 1.00 |
| SUMO2 | rs1471453 | 73177444 | 0.05 | 0.00 |
| SUMO2 | . | 73179065 | 0.0025 | 1.00 |
| SUMO3 | rs1051331 | 46226786 | 0.0025 | 1.00 |
| SUMO3 | rs2329902 | 46227163 | 0.13 | 0.00 |
| SUMO3 | . | 46227968 | 0.0025 | 1.00 |
| SUMO3 | . | 46228150 | 0.0025 | 1.00 |
| SUMO3 | . | 46228153 | 0.0025 | 1.00 |
| SUMO3 | rs752652207 | 46228154 | 0.0025 | 1.00 |
| SUMO3 | rs757331290 | 46228155 | 0.0025 | 1.00 |
| SUMO3 | rs765382230 | 46228157 | 0.005 | 1.00 |
| SUMO3 | rs7283639 | 46228165 | 0.155 | 0.27 |
| SUMO3 | rs188978703 | 46228170 | 0.02 | 0.07 |
| SUMO3 | rs141141907 | 46228243 | 0.0025 | 1.00 |
| SUMO3 | . | 46228662 | 0.0025 | 1.00 |
| SUMO3 | rs235293 | 46228930 | 0.0125 | 1.00 |
| SUMO3 | rs564735586 | 46228945 | 0.0025 | 1.00 |
| SUMO3 | . | 46228949 | 0.0025 | 1.00 |
| SUMO3 | rs2838697 | 46233836 | 0.445 | 0.48 |
| SUMO3 | rs13050872 | 46233863 | 0.085 | 0.15 |
| SUMO3 | rs9981327 | 46233866 | 0.0025 | 1.00 |
| SUMO3 | rs9306116 | 46234079 | 0.4375 | 0.67 |
| SUMO4 | rs237025 | 149721690 | 0.305 | 0.32 |
| SUMO4 | rs777445425 | 149721778 | 0.0025 | 1.00 |
| SUMO4 | rs145312495 | 149721800 | 0.0025 | 1.00 |
| SUMO4 | rs237024 | 149721965 | 0.305 | 0.32 |
| SUMO4 | rs237023 | 149722040 | 1 | NA |
| UBE2D1 | rs112660736 | 60095105 | 0.015 | 0.04 |
| UBE2D1 | rs759280904 | 60121139 | 0.0025 | 1.00 |
| UBE2D1 | . | 60121240 | 0.0025 | 1.00 |
| UBE2D1 | rs73288305 | 60123486 | 0.3025 | 0.24 |
| UBE2D1 | . | 60123523 | 0.0025 | 1.00 |
| UBE2D1 | . | 60124703 | 0.0025 | 1.00 |
| UBE2D1 | rs531786752 | 60127627 | 0.0025 | 1.00 |
| UBE2D1 | rs74377516 | 60127639 | 0.2975 | 0.17 |
| UBE2D1 | . | 60127798 | 0.0025 | 1.00 |
| UBE2D1 | rs75362994 | 60127838 | 0.3025 | 0.24 |

| Gene name | SNP | Location | MAF | HWE |
|-----------|-------------|----------|--------|------|
| UBE2D1 | rs3802699 | 60128364 | 0.0025 | 1.00 |
| UBE2D1 | rs148198083 | 60128583 | 0.0475 | 1.00 |
| UBE2I | rs9926183 | 1363878 | 0.035 | 0.00 |
| UBE2I | . | 1363927 | 0.0025 | 1.00 |
| UBE2I | rs9941160 | 1364140 | 1 | NA |
| UBE2I | . | 1364158 | 0.0025 | 1.00 |
| UBE2I | rs201695180 | 1364251 | 0.0025 | 1.00 |
| UBE2I | rs4984806 | 1364281 | 1 | NA |
| UBE2I | rs4610 | 1364365 | 0.0375 | 0.02 |
| UBE2I | rs781398317 | 1365612 | 0.0025 | 1.00 |
| UBE2I | rs112302601 | 1365915 | 0.0025 | 1.00 |
| UBE2I | rs4984807 | 1365935 | 0.395 | 0.00 |
| UBE2I | rs7186045 | 1365943 | 0.015 | 1.00 |
| UBE2I | rs4984808 | 1365967 | 0.0725 | 0.00 |
| UBE2I | rs201661304 | 1369612 | 0.0075 | 1.00 |
| UBE2I | rs9933497 | 1369730 | 1 | NA |
| UBE2I | . | 1369837 | 0.0025 | 1.00 |
| UBE2I | rs909915 | 1369926 | 0.185 | 0.00 |
| UBE2I | rs758216436 | 1370203 | 0.0025 | 1.00 |
| UBE2I | rs909916 | 1370303 | 1 | NA |
| UBE2I | rs909917 | 1370309 | 1 | NA |
| UBE2I | rs148789348 | 1370383 | 0.0125 | 1.00 |
| UBE2I | . | 1370575 | 0.0025 | 1.00 |
| UBE2I | rs4017786 | 1370597 | 1 | NA |
| UBE2I | rs8063770 | 1370614 | 1 | NA |
| UBE2I | rs8043720 | 1370630 | 1 | NA |
| UBE2I | rs79005361 | 1370682 | 0.0475 | 0.36 |
| UBE2I | rs571836605 | 1370698 | 0.0025 | 1.00 |
| UBE2I | rs142273742 | 1370716 | 0.0075 | 0.01 |
| UBE2I | rs909918 | 1370729 | 1 | NA |
| UBE2I | rs2369700 | 1374513 | 0.005 | 0.00 |
| UBE2I | rs761059 | 1374524 | 0.0175 | 0.00 |
| UBE2I | . | 1374629 | 0.0025 | 1.00 |
| UBE2I | rs761060 | 1374656 | 0.0425 | 0.00 |
| UBE2I | rs762904858 | 1374785 | 0.0025 | 1.00 |
| UBE2I | rs8063 | 1374818 | 0.0475 | 0.01 |
| PIAS1 | . | 68346778 | 0.0025 | 1.00 |
| PIAS1 | rs537399079 | 68346778 | 0.0025 | 1.00 |
| PIAS1 | . | 68348207 | 0.0025 | 1.00 |
| PIAS1 | rs182709174 | 68349843 | 0.0025 | 1.00 |
| PIAS1 | rs1489598 | 68349893 | 0.165 | 0.12 |
| PIAS1 | rs79833223 | 68349918 | 0.0325 | 1.00 |
| PIAS1 | rs183625509 | 68349925 | 0.01 | 1.00 |

| Gene name | SNP | Location | MAF | HWE |
|-----------|-------------|----------|--------|------|
| PIAS1 | rs145053928 | 68378937 | 0.0125 | 1.00 |
| PIAS1 | rs186753486 | 68379138 | 0.01 | 1.00 |
| PIAS1 | rs750502048 | 68434378 | 0.0025 | 1.00 |
| PIAS1 | rs117588299 | 68434379 | 0.0225 | 1.00 |
| PIAS1 | . | 68434428 | 0.0025 | 1.00 |
| PIAS1 | rs145280358 | 68434744 | 0.01 | 1.00 |
| PIAS1 | rs576273919 | 68439101 | 0.0025 | 1.00 |
| PIAS1 | . | 68445850 | 0.0025 | 1.00 |
| PIAS1 | rs11633620 | 68445912 | 1 | NA |
| PIAS1 | . | 68468049 | 0.0025 | 1.00 |
| PIAS1 | rs113555272 | 68468098 | 0.0225 | 1.00 |
| PIAS1 | . | 68468136 | 0.0025 | 1.00 |
| PIAS1 | rs12438361 | 68468695 | 0.0375 | 0.00 |
| PIAS1 | rs3759823 | 68473443 | 0.0225 | 1.00 |
| PIAS1 | rs75673552 | 68473808 | 0.0025 | 1.00 |
| PIAS1 | . | 68476069 | 0.0025 | 1.00 |
| PIAS1 | . | 68479858 | 0.0025 | 1.00 |
| PIAS1 | rs191408288 | 68480086 | 0.01 | 1.00 |
| PIAS1 | rs372957210 | 68480226 | 0.01 | 1.00 |
| PIAS2 | rs3737448 | 44395368 | 0.0925 | 0.68 |
| PIAS2 | rs10502879 | 44398111 | 0.0025 | 1.00 |
| PIAS2 | . | 44398183 | 0.0025 | 1.00 |
| PIAS2 | rs149740503 | 44398464 | 0.0025 | 1.00 |
| PIAS2 | rs183321210 | 44398557 | 0.005 | 1.00 |
| PIAS2 | rs146442641 | 44401052 | 0.015 | 1.00 |
| PIAS2 | rs35451178 | 44407993 | 0.0025 | 1.00 |
| PIAS2 | rs72907142 | 44409581 | 0.005 | 1.00 |
| PIAS2 | . | 44409617 | 0.0025 | 1.00 |
| PIAS2 | . | 44416287 | 0.0025 | 1.00 |
| PIAS2 | rs72907148 | 44416608 | 0.005 | 1.00 |
| PIAS2 | rs77040088 | 44423925 | 0.01 | 1.00 |
| PIAS2 | rs188584114 | 44424122 | 0.0025 | 1.00 |
| PIAS2 | rs372024400 | 44424939 | 0.0025 | 1.00 |
| PIAS2 | rs150885589 | 44424995 | 0.0025 | 1.00 |
| PIAS2 | rs114135676 | 44426800 | 0.005 | 1.00 |
| PIAS2 | rs644731 | 44426877 | 0.3025 | 0.32 |
| PIAS2 | rs764656048 | 44435407 | 0.0025 | 1.00 |
| PIAS2 | rs539782995 | 44470481 | 0.0025 | 1.00 |
| PIAS2 | rs113887072 | 44470706 | 0.0575 | 0.49 |
| PIAS2 | . | 44470825 | 0.0025 | 1.00 |
| PIAS2 | rs72915074 | 44483961 | 0.005 | 1.00 |
| PIAS2 | rs567309045 | 44496847 | 0.0025 | 1.00 |
| PIAS2 | rs530227612 | 44497099 | 0.0025 | 1.00 |

| Gene name | SNP | Location | MAF | HWE |
|-----------|-------------|-----------|--------|------|
| PIAS2 | . | 44497124 | 0.0025 | 1.00 |
| PIAS2 | . | 44497173 | 0.0025 | 1.00 |
| PIAS2 | . | 44497338 | 0.0025 | 1.00 |
| FBXL21 | rs573196792 | 135272403 | 0.0025 | 1.00 |
| FBXL21 | rs150784167 | 135272692 | 0.0025 | 1.00 |
| FBXL21 | rs17702049 | 135272823 | 0.01 | 1.00 |
| FBXL21 | rs544746984 | 135273078 | 0.0025 | 1.00 |
| FBXL21 | . | 135273177 | 0.0025 | 1.00 |
| FBXL21 | rs183239904 | 135273298 | 0.005 | 1.00 |
| FBXL21 | rs10052673 | 135273370 | 0.0175 | 1.00 |
| FBXL21 | rs31549 | 135276049 | 0.1125 | 0.00 |
| FBXL21 | rs76075237 | 135276205 | 0.0025 | 1.00 |
| FBXL21 | rs40986 | 135276314 | 0.22 | 0.04 |
| FBXL21 | rs31548 | 135276701 | 0.22 | 0.04 |
| FBXL21 | rs2066913 | 135276814 | 0.225 | 0.84 |
| FBXL21 | rs31547 | 135276847 | 0.22 | 0.04 |
| FBXL21 | rs530876112 | 135277204 | 0.005 | 1.00 |

SNP – single nuclear polymorphism; MAF – minor allele frequency; HWE – Hardy Weinberg equilibrium; NA – not available.

Supplementary Table 3. Results of logistic regression adjusted by the administration of sirolimus in non-significant tagger SNPs by five models.

| SNPs | OR | Lower 95% CI | Upper 95% CI | P value |
|------------------------|------|--------------|--------------|---------|
| Recessive model | | | | |
| rs644731 | 4.24 | 1.32 | 13.64 | 0.01 |
| rs3737448 | NA | 0 | NA | 0.08 |
| rs75362994 | 0.53 | 0.17 | 1.6 | 0.24 |
| rs7283639 | 1.95 | 0.42 | 9.03 | 0.40 |
| rs113887072 | 0 | 0 | NA | 0.41 |
| rs1489598 | 2.55 | 0.16 | 41.59 | 0.52 |
| rs13050872 | 0.5 | 0.03 | 8.36 | 0.62 |
| rs2066913 | 0.7 | 0.14 | 3.44 | 0.65 |
| rs2838697 | 0.93 | 0.43 | 2.01 | 0.85 |
| Dominant model | | | | |
| rs2838697 | 0.62 | 0.32 | 1.19 | 0.15 |
| rs2066913 | 0.65 | 0.34 | 1.24 | 0.18 |
| rs3737448 | 1.66 | 0.75 | 3.64 | 0.21 |
| rs237025 | 0.74 | 0.4 | 1.38 | 0.34 |
| rs113887072 | 0.61 | 0.21 | 1.79 | 0.36 |
| rs75362994 | 1.32 | 0.71 | 2.47 | 0.38 |
| rs7283639 | 1.18 | 0.59 | 2.35 | 0.64 |
| rs13050872 | 0.85 | 0.35 | 2.04 | 0.71 |
| rs1489598 | 0.92 | 0.47 | 1.79 | 0.80 |

| SNPs | OR | Lower 95% CI | Upper 95% CI | P value |
|---------------------------|------|--------------|--------------|---------|
| Overdominant model | | | | |
| rs75362994 | 1.7 | 0.9 | 3.21 | 0.10 |
| rs2838697 | 0.68 | 0.36 | 1.28 | 0.23 |
| rs2066913 | 0.68 | 0.35 | 1.31 | 0.24 |
| rs237025 | 0.72 | 0.38 | 1.35 | 0.30 |
| rs3737448 | 1.45 | 0.64 | 3.25 | 0.38 |
| rs113887072 | 0.65 | 0.22 | 1.93 | 0.43 |
| rs1489598 | 0.87 | 0.44 | 1.72 | 0.69 |
| rs13050872 | 0.9 | 0.37 | 2.24 | 0.83 |
| rs7283639 | 1.04 | 0.51 | 2.16 | 0.91 |
| Codominant model | | | | |
| rs3737448 | 1.49 | 0.66 | 3.34 | 0.14 |
| rs75362994 | 1.58 | 0.82 | 3.05 | 0.20 |
| rs2838697 | 0.59 | 0.29 | 1.19 | 0.33 |
| rs2066913 | 0.65 | 0.33 | 1.28 | 0.41 |
| rs113887072 | 0.65 | 0.22 | 1.92 | 0.52 |
| rs237025 | 0.71 | 0.37 | 1.36 | 0.58 |
| rs7283639 | 1.08 | 0.52 | 2.25 | 0.69 |
| rs1489598 | 0.88 | 0.45 | 1.75 | 0.76 |
| rs13050872 | 0.89 | 0.36 | 2.22 | 0.86 |
| Additive model | | | | |
| rs3737448 | 1.77 | 0.84 | 3.69 | 0.13 |
| rs2066913 | 0.7 | 0.4 | 1.22 | 0.20 |
| rs2838697 | 0.79 | 0.51 | 1.23 | 0.30 |
| rs113887072 | 0.6 | 0.21 | 1.69 | 0.31 |
| rs237025 | 0.83 | 0.5 | 1.38 | 0.48 |
| rs7283639 | 1.22 | 0.69 | 2.15 | 0.49 |
| rs13050872 | 0.83 | 0.38 | 1.81 | 0.64 |
| rs75362994 | 1.03 | 0.65 | 1.63 | 0.90 |
| rs1489598 | 0.97 | 0.51 | 1.83 | 0.92 |

SNP – single nuclear polymorphisms; OR – odds ratio; CI – confidential interval; NA – not available.

Supplementary Table 4. Results of survival analysis between three genotypes and the time from transplantation to biopsy in this cohort.

| Confounding factors | P value |
|------------------------------------|---------|
| Three genotypes (CT vs. TT vs. CC) | 0.1573 |
| CT vs. TT | 0.2586 |
| CT vs. CC | 0.2833 |
| CC vs. TT | 0.0864 |

Supplementary Table 5. Results of linkage disequilibrium haplotype analysis in this cohort.

| Block 1 | rs7283639 | rs2838697 | Haplotype frequency (%) | |
|---------|------------|------------|-------------------------|-------------------------|
| H1 | G | A | 55.5 | |
| H2 | G | C | 29 | |
| H3 | C | C | 15.5 | |
| Block 2 | rs73288305 | rs74377516 | rs75362994 | Haplotype frequency (%) |
| H1 | A | T | T | 69.7 |
| H2 | T | G | A | 29.8 |
| Block 3 | rs237025 | rs237024 | Haplotype frequency (%) | |
| H1 | A | C | 69.5 | |
| H2 | T | G | 30.5 | |

Supplementary Table 6. Results of haplotype analysis among detected blocks in this cohort.

| Blocks | Likelihood ratio | Test df | P value |
|---------|------------------|---------|---------|
| Block 1 | 2.74 | 5 | 0.74 |
| Block 2 | 3.46 | 2 | 0.18 |
| Block 3 | 0.77 | 2 | 0.68 |

df – degree of freedom.

References:

- Djamali A, Premasathian N, Pirsch JD: Outcomes in kidney transplantation. *Semin Nephrol*, 2003; 23(3): 306–16
- Dussaule JC, Chatziantoniou C: Reversal of renal disease: Is it enough to inhibit the action of angiotensin II? *Cell Death Differ*, 2007; 14(7): 1343–49
- Chan SC, Zhang Y, Shao A et al: Mechanism of fibrosis in HNF1B-related autosomal dominant tubulointerstitial kidney disease. *J Am Soc Nephrol*, 2018; 29(10): 2493–509
- Liu YL, Reeves HL, Burt AD et al: TM6SF2 rs58542926 influences hepatic fibrosis progression in patients with non-alcoholic fatty liver disease. *Nat Commun*, 2014; 5: 4309
- Dunn W, O'Neil M, Zhao J et al: Donor PNPLA3 rs738409 genotype affects fibrosis progression in liver transplantation for hepatitis C. *Hepatology*, 2014; 59(2): 453–60
- Chand S, Edwards NC, Chue CD et al: Caveolin-1 single-nucleotide polymorphism and arterial stiffness in non-dialysis chronic kidney disease. *Nephrol Dial Transplant*, 2016; 31(7): 1140–44
- Yang B, Feng W, Li Y et al: Interleukin 18–607 A/C gene polymorphism is associated with susceptibility to IgA nephropathy in a Chinese Han population. *Appl Immunohistochem Mol Morphol*, 2017; 25(10): 725–30
- Zhao C, Xu Z, Wang Z et al: Role of tumor necrosis factor- α in epithelial-to-mesenchymal transition in transplanted kidney cells in recipients with chronic allograft dysfunction. *Gene*, 2018; 642: 483–90
- Wang Z, Yang H, Si S et al: Polymorphisms of nucleotide factor of activated T cells cytoplasmic 2 and 4 and the risk of acute rejection following kidney transplantation. *World J Urol*, 2018; 36(1): 111–16
- Loupy A, Haas M, Solez K et al: The Banff 2015 Kidney Meeting Report: Current challenges in rejection classification and prospects for adopting molecular pathology. *Am J Transplant*, 2017; 17(1): 28–41
- Hindorff LA, Sethupathy P, Junkins HA et al: Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA*, 2009; 106(23): 9362–67
- Consortium EP: An integrated encyclopedia of DNA elements in the human genome. *Nature*, 2012; 489(7414): 57–74
- Xue C, Zhang H, Lin Q et al: Manipulating mRNA splicing by base editing in plants. *Sci China Life Sci*, 2018; 61(11): 1293–300
- Akua T, Berezin I, Shaul O: The leader intron of AtMHX can elicit, in the absence of splicing, low-level intron-mediated enhancement that depends on the internal intron sequence. *BMC Plant Biol*, 2010; 10: 93
- Laxa M: Intron-mediated enhancement: A tool for heterologous gene expression in plants? *Front Plant Sci*, 2016; 7: 1977
- Gallegos JE, Rose AB: The enduring mystery of intron-mediated enhancement. *Plant Sci*, 2015; 237: 8–15
- Ma L, Li H, Zhang S et al: Emodin ameliorates renal fibrosis in rats via TGF- β 1/Smad signaling pathway and function study of Smurf 2. *Int Urol Nephrol*, 2018; 50(2): 373–82
- Lin CM, Xu J, Yang WT et al: Smurf downregulates echinoid in the amnioserosa to regulate *Drosophila* dorsal closure. *Genetics*, 2017; 206(2): 985–92