Transplant International

Transplant International ISSN 0934-0874

ORIGINAL ARTICLE

Analysis of independent microarray datasets of renal biopsies identifies a robust transcript signature of acute allograft rejection

Pierre Saint-Mezard,[1] Céline C. Berthier,[2]* Hai Zhang,[1†] Alexandre Hertig,[3] Sergio Kaiser,[1] Martin Schumacher,[1] Grazyna Wieczorek,[1] Marc Bigaud,[1] Jeanne Kehren,[1] Eric Rondeau,[3] Friedrich Raulf[1] and Hans-Peter Marti[2,3]

1 Novartis Institutes for BioMedical Research, Basel, Switzerland
2 Division of Nephrology and Hypertension, Inselspital, Bern, Switzerland
3 INSERM U702, Hôpital Tenon, Paris, France

Keywords
acute rejection, biopsies for clinical indication, comparative analysis, gene expression, kidney, molecular diagnostic, transplantation.

Correspondence
Friedrich Raulf PhD, Novartis Institutes for BioMedical Research, Novartis Pharma AG, Novartis Campus, WSJ-386.6.09, CH-4056 Basel, Switzerland. Tel.: +41 (61) 324 68 80; fax: +41 (61) 324 35 76; e-mail: friedrich.raulf@novartis.com

***Present address:**
Department of Internal Medicine – Nephrology, University of Michigan, Ann Arbor, MI, USA

†Present address:
Hoffmann-La Roche Ltd, Basel, Switzerland

Received: 9 June 2008
Revision requested: 24 July 2008
Accepted: 30 September 2008

doi:10.1111/j.1432-2277.2008.00790.x

Summary

Transcriptomics could contribute significantly to the early and specific diagnosis of rejection episodes by defining 'molecular Banff' signatures. Recently, the description of pathogenesis-based transcript sets offered a new opportunity for objective and quantitative diagnosis. Generating high-quality transcript panels is thus critical to define high-performance diagnostic classifier. In this study, a comparative analysis was performed across four different microarray datasets of heterogeneous sample collections from two published clinical datasets and two own datasets including biopsies for clinical indication, and samples from non-human primates. We characterized a common transcriptional profile of 70 genes, defined as acute rejection transcript set (ARTS). ARTS expression is significantly up-regulated in all AR samples as compared with stable allografts or healthy kidneys, and strongly correlates with the severity of Banff AR types. Similarly, ARTS were tested as a classifier in a large collection of 143 independent biopsies recently published by the University of Alberta. Results demonstrate that the 'in silico' approach applied in this study is able to identify a robust and reliable molecular signature for AR, supporting a specific and sensitive molecular diagnostic approach for renal transplant monitoring.

Introduction

Despite clinical application of potent new immuno-suppressive and -regulatory drugs in numerous combinations, acute allograft rejection (AR) remains a common and serious post-transplantation complication [1]. It is also a risk factor for late graft dysfunction [2], a relentlessly progressive process characterized histologically by a gradual increase in interstitial fibrosis and tubular atrophy [3]. A very powerful predictive factor in adults and children is the number of acute rejection episodes that cannot be treated successfully. Strategies to detect and treat AR as early as possible and prior to the occurrence of irreversible structural lesions would increase graft survival.

However, current monitoring and diagnostic modalities are ill-suited to the diagnosis of acute rejection at an early stage. Procedures to diagnose allograft rejection generally depend upon laboratory parameters like serum creatinine or proteinuria, and histologic assessment of graft biopsies

(i.e. interstitial inflammation and tubulitis), which are not very specific and are often found in stable patients [4]. There is a need for more specific and objective diagnostic tools that are also able to provide mechanistic information on the underlying pathologic events.

Gene-expression profiling might provide criteria to refine the histology-based Banff classification [5] of renal allograft pathology as marker transcripts should be detectable very early after the pathologic process starts, preceding the occurrence of histologically visible lesions [6]. Furthermore, the early molecular status might have a prognostic impact on AR and renal function prediction. Over the last 6 years, high-density cDNA and oligonucleotide microarrays have been used to classify and predict allograft rejection [7]. Recently, Mueller *et al.* [8] developed a new analysis concept using pathogenesis-based transcript sets (PBTs) defined in a mouse model, to translate microarray results into measurement of biologic processes in human renal transplant biopsies. The study suggests that the use of PBTs for diagnostic purpose is likely to become a robust and reliable new quantitative tool for assessing rejection and inflammation. The PBTs used in that study were mainly defined in mouse models and confirmed in human cell culture. The definition of new transcripts sets directly based on human biopsies data may provide a possible refinement and/or enhancement of this methodology.

Then, we decided to perform a new gene-expression study on 47 renal biopsies for clinical indication from an unselected cohort of 45 patients (Table 1). Histological biopsy grading according to Banff '97 [5] revealed different types and grades of AR and chronic allograft nephropathy (CAN). We focused on the analysis of AR, and extended the analysis of the AR biopsies from our collection by a comparative analysis across three additional datasets, including heterogeneous clinical sample collections from two published studies [9,10] and samples from an AR renal allograft nonhuman primate (NHP) model [11,12]. Such a comparative analysis of large-scale transcriptomic datasets has been demonstrated previously as a powerful approach in the analysis of multiple cancer studies [13]. This strategy allowed us to identify a consistent gene signature for AR. The ability of this signature to identify AR accurately was evaluated in each of the four independent datasets and further validated in 143 biopsies from the recent Edmonton study [8].

Patients and methods

Patients and samples for microarray studies

All patients at Hôpital Tenon, Paris, France, undergoing a renal allograft biopsy because of clinical indication (between February 2003 and September 2004) were included in the study. In addition, a few patients from Hôpital Bicêtre, Paris, and Hôpital Pellegrin, Bordeaux, France, were recruited. Seventy-five renal core biopsies were obtained by a 16-gauge biopsy needle. About 2/3 of the respective biopsy was processed for histopathology and the remaining was collected immediately in RNAlater (Ambion, Austin, TX, USA). Two independent patholo-

Table 1. Demographic and clinical characteristics of the five groups of the Paris biopsies for clinical indications according to histologic analysis.

Parameter	AR	CAN + AR	Borderline	NR	CAN
No. biopsies	8	7	3	7	22
No. patients*	8	6	2	7	22
Recipient age (years)	43.9 ± 10.9	41.7 ± 7.6	34.6 ± 10.2	43.1 ± 8.7†	46.9 ± 12.2†
Recipient gender (*n*, % male)	6 (75%)	2 (33%)	1 (50%)	6 (86%)	15 (68.2%)
Donor age (years)	36.3 ± 8.3†	39.0 ± 19.8	46.5 ± 0.7	45.2 ± 15.4	43.4 ± 17.1†
No. HLA mismatches	2.8 ± 1.6†	2.7 ± 2.1†	3.5 ± 0.7	1.8 ± 1.5†	2.9 ± 1.3†
No. historic AR episodes (% patients with ≥1)	75	83.3	100	14.3	36.4
Time of biopsy (months post-Tx)	28.1 ± 51.1	52.1 ± 48.3	3.4 ± 4.9	25.1 ± 51.4	83.2 ± 64.8
No. patients with CNI toxicity (histology)	1	0	0	1	3
No. patients on a CNI-free regimen	1	0	0	0	2
Serum creatinine (μmol/l)	253.0 ± 109.3	461.1 ± 317.3	223.0 ± 40.7†	160.0 ± 44.4	281.6 ± 204.6
GFR MDRD-calculated (ml/min/1.73 m^2)	33.2 ± 18.3†	19.2 ± 15.0	31.5 ± 2.6	43.0 ± 15.5	27.6 ± 13.5†

CAN, chronic allograft nephropathy; AR, acute rejection; NR, no rejection (but clinically not stable); GFR, glomerular filtration rate; Tx, transplantation.

All patients were on standard CNI triple regimen: 68% cyclosporine and 32% tacrolimus-based, mainly combined with mycophenolate and corticosteroids.

Ethnicity: 60% Caucasian, 18% African, 22% other origin.

Included are all samples used for signature generation (see Table S1 for individual data).

*Five patients had two sequential biopsies.

†Some values unknown.

gists classified the biopsies according to Banff '97 categories (see Table S1) [5]. Table 1 displays a summary of demographic and clinical characteristics of the five sample groups for all biopsies finally included in microarray analyses after filtering with stringent quality controls of total RNA and GeneChip raw data, leading to exclusion of 36% of samples. Nineteen control samples originated from nephrectomy specimens of patients with normal kidney function suffering from solid renal cancer, using nonaffected cortex at maximal distance. Thirteen samples matched all quality criteria.

Microarray hybridization

Total RNA was extracted using RNeasy according to the manufacturer's protocol (Qiagen, Hilden, Germany), quantified by ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA), and quality-controlled by Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA). Only biopsies resulting in high-quality total RNA were further analyzed. Fifty nanograms of total RNA were subjected to Affymetrix 2-cycle cDNA amplification, labeling, and hybridization to Human Genome U133 Plus 2.0 arrays containing 54 120 probe sets representing >47 000 different transcripts (Affymetrix, Santa Clara, CA, USA). The resulting Affymetrix raw data (CEL files) are available in the Gene-expression Omnibus microarray data repository (GEO) under GSE9493 (http://www.ncbi.nlm.nih.gov/geo) and are referred to in this study as Paris dataset.

Further microarray datasets

Nonhuman primate samples

Cynomolgus monkey (*Macaca fascicularis*) kidney allografts and controls were collected at necropsy from a life-supporting AR model [12] with histopathologic assessment of rejection (Table 2). Total RNA from renal cortex was extracted as described above but processed without amplification using Affymetrix standard protocol and HG-U133A Gene Chips.

Publicly available clinical datasets

Two public microarray datasets were downloaded from GEO. One analysis (Stanford) performed on a customized cDNA-spotted two channel microarray platform, is available under the accession number GSE343 [9]. Another analysis (Cleveland) performed on the Affymetrix HG-U95A version 2 oligonucleotide platform is available under the accession number GSE1563 [10]. These two public datasets were compared with above mentioned NHP and clinical 'Paris' datasets. Finally, this comparative analysis includes a total number of 36 AR, three borderline, 32 nonrejecting and 34 nontransplanted control samples (Table 2). We used the Edmonton study [8] as a validation dataset. CEL files were downloaded from http://transplants.med.ualberta.ca.

Table 2. Synopsis of gene-expression datasets and samples used for bioinformatic comparative analyses.

Samples dataset	AR	CAN + AR	Borderline	CAN	CAV	NR	Control	Reference
Paris	8	7	3	22		7	13	This paper†
Stanford*	15	–	–	–		15	–	[9]
Cleveland*	7	–	–	–		10	9	[10]
NHP	6	–	–	–	8	–	12	This paper
Total	36	7	3	22	8	32	34	

NHP, nonhuman primate; CAV, chronic allograft vasculopathy.
*Number of datasets as available in GEO; no distinction possible from GEO annotation between the eight controls and seven nonrejecting samples in the Stanford dataset.
†The Paris dataset is available in GEO under accession number GSE9493.

It is important to mention that two different GeneChip platforms generated the data in this comparative analysis, the oligonucleotide-based Affymetrix technology and the cDNA-spotted Stanford microarray.

Data analysis

Generation of microarray-derived data

A single weighted mean expression level for each gene along with a *P*-value indicating reliable transcript detection was derived using the Microarray Suite 5.0 software (MAS5; Affymetrix). Data were scaled for each array (target intensity of 150). For further analysis, the cell intensity (CEL) files were subjected to the Robust Multichip Analysis normalization. Several quality control measures on each array were assessed, including review of the scanned image for significant artifacts, background and noise measurements that differ significantly from other chips, average of present and absent calls. Furthermore, arrays failing two out of the three following criteria were excluded from the study: the 3′–5′ ratio of the intensities for glyceraldehyde-3-phosphate dehydrogenase (ratio ≤ 4), average of calls (≥40% of present calls), and scaling factor (≥2 or ≤0.5).

Class comparison

Analysis of the microarray raw data was performed using GeneSpring GX7.3 (Agilent Technologies). Genes differentially expressed among different sample classes were identified using a one-way ANOVA ($P \leq 0.05$), with or without a false discovery rate ≤5% (Benjamini and Hochberg FDR) and additional cutoff based on twofold change

between groups. The differential expression of 20 individual genes was confirmed by inventoried quantitative real-time PCR assays (Applied Biosystems, Foster City, CA, USA) (see Figure S1).

Functional analysis of expression data

Gene regulatory networks were generated using MetaCore analytical suite version 4.2 build 8168 (GeneGo, St Joseph, MI, USA). MetaCore is a web-based suite for functional analysis of experimental data in the context of manually curated human protein interactions, canonic pathways, and knowledge-based ontologies of cellular processes, diseases, and toxicology. The database includes over 160 000 human protein interactions and metabolic reactions. The 70 genes from the ARTS were subjected to enrichment analysis in GO biologic processes (Table 4). Both enrichment analysis and calculation of statistical significance of networks are based on *P*-values, which are defined as the probability of a given number of genes from the input list to match a certain number of genes in the ontology's folder, or the probability of the network's assembly from a random set of nodes (genes) the same size as the input list. The ARTS genes were further used to build networks using the Analyze networks (transcription regulation) algorithm that generate sub-networks centered on transcription factors (Fig. 1). The sub-networks were scored and prioritized based on relative enrichment with the data from input list and saturation with 'canonic pathways' using *P*-values and *z*-scores as statistical metrics. Legend describing symbols in the network could be found at http://www.genego.com/metacore.php.

Score-based classification

To assess the validity of this study's ARTS gene list, we applied it in the context of a recent report [8] proposing the use of PBT-based scores for the evaluation of ongoing biologic processes during renal allograft rejection. In brief, gene expression was defined as fold change versus controls. Gene expression within the ARTS list was summarized as a score by calculating the geometric mean of fold changes across all probe sets in the ARTS list. Using linear discriminant analysis, a score classifier was built based on histopathology and on the retrospective diagnosis of episodes. Two classes were defined in each case as described [8]: rejection (T-cell-mediated, antibody-mediated, mixed) and nonrejection (all other cases). Classifier accuracy (proportion of samples correctly classified), sensitivity, specificity and positive and negative predictive values were estimated based using leave-one-out cross-validation (LOOCV) and 10-fold cross-validation.

Results

Patient demographics

All patients included in the Paris dataset gave written informed consent, and the local ethical commission approved the study. In the present analysis, we focused on transcriptomic profiles of biopsies for clinical indication from patients with AR ($n = 8$), AR on top of CAN ($n = 7$) or borderline cases ($n = 3$) compared with 13 biopsies from nephrectomized nontransplanted control patients representing normal healthy kidney cortex.

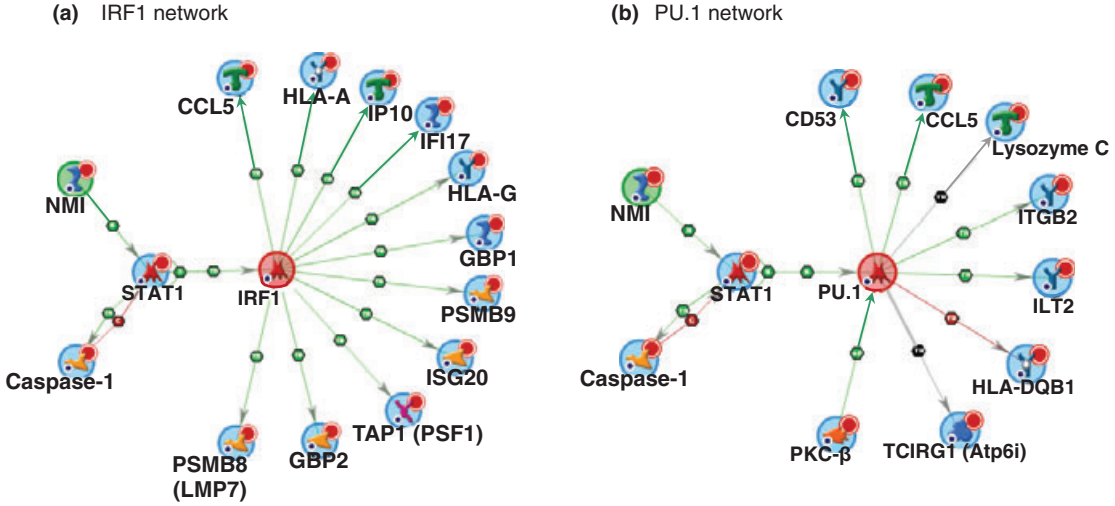

Figure 1 Transcription Regulatory network analysis indicates importance of IRF1 and PU.1 in acute rejection. Protein interaction networks were generated with the 70 genes identifies by the comparative analysis. Cluster a and b are the most significant networks with a respective *P*-value of 4.5×10^{-34} and 1.1×10^{-26}. The networks show that the input genes (blue circles) were transcriptionally regulated via IRF1 and PU.1. Colors of the lines indicate inhibition (red), activation (green), and no clear link (gray). Remaining details of network are as described in Methods.

Comparative analysis

The Stanford- and Cleveland datasets were re-analyzed and compared with the results obtained from the Paris collection. This comparative analysis was extended to include gene-expression profiles from a NHP model of acute renal allograft rejection using the Affymetrix HG-U133A GeneChip. Table 2 summarizes the total number of samples included in the final comparative analysis.

As recently described in a multicenter study of microarray performance [14], we decided to select genes differentially expressed between AR and control samples satisfying a fold change >2 and a *t*-test or ANOVA *P*-value < 0.05 in each individual dataset (see supplement files for method details). This strategy identifies 790 probesets in Stanford, 1420 probesets in Cleveland, 2182 probesets in Paris and 2636 probesets in the NHP dataset, which are significantly dysregulated in AR. These four independent probeset lists were then intersected. Table 3 shows the successive overlaps between the NHP and the three human datasets. This strategy revealed a common transcriptional profile of 70 genes represented by 116 Affymetrix probe sets that are systematically up-regulated in all the AR samples compared with controls or nonrejected kidneys (Table S2). This gene set is further defined as ARTS.

Biomedical relevance of ARTS

These 70 genes were uploaded into MetaCore and analyzed for biologic processes over-represented. Table 4 shows that ARTS genes are mainly associated to antigen presenting cells, T-cell activation as well as the interferon (IFN)-γ pathways. In addition, we performed a transcription regulation analysis to identify sub-networks of genes centered on transcription factors. The MetaCore algorithm selected IRF1, PU.1, STATs and NF-κB as the transcription factors that regulate most significantly the genes represented in the ARTS. A visualization of those networks is presented in Fig. 1. STAT1 and IRF1 are key transcription factors of the IFN-γ pathway, while PU.1 is a specific transcription factor involved in the differentiation or activation of macrophages or B cells. These results show that genes selected by a bioinformatics approach are biologically relevant and related to the pathogenesis of acute rejection.

Table 3. Intersections of acute rejection signatures among four datasets.

	Stanford	Cleveland	Paris	NHP
Stanford	790	289	236	341
Stanford/Cleveland			138	190
Stanford/Cleveland/Paris				116

Table 4. GO enrichment analysis identifies relevant biologic processes for acute rejection.

Biologic processes	max[−log(*P*)]*
Immune_antigen presentation	43.037
Immune_phagosome	27.466
Interferon gamma signaling	15.046
Immune_TCR signaling	10.002
Cell adhesion	7.922
Chemotaxis	7.315

*Biologic processes were ranked according to *P*-values.

The ARTS correlates with the Banff-based severity of AR

First, we evaluated the ability of the 70 genes from the ARTS to discriminate AR from nonrejected (or nontransplanted control) samples in the four independent datasets. A principal component analysis was applied as a visualization method across the four independent experiments to illustrate the relative position of each group among each other (Fig. 2). ARTS expression profile is able to associate each sample to its clinical diagnosis in all dataset tested, except the Paris dataset. In this study, one AR and three borderline samples displayed a very similar expression profile than the control biopsies, showing discrepancy with the pathologic diagnosis. Similar results were obtained with unsupervised hierarchical clustering methods (data not shown).

We then used PBT-based scores method for evaluation of AR in clinical diagnostic samples from the Paris dataset. To increase the number of samples per group, we included seven additional AR samples with additional histologic signs of CAN (AR + CAN). Figure 3 shows that the ARTS scores correlate with the Banff score of AR in the Paris dataset. We observed that the Banff score is positively associated with the ARTS scores and that the 70 genes from the ARTS show a gradual increasing expression across the different Banff scores, from the absence of signal in the control group towards a strong expression in AR score II.

Consistency testing of the transcript sets

In order to evaluate the robustness and biomedical relevance of ARTS, we used 143 biopsy microarray data from Edmonton [8] as an independent confirmation set. We first calculated the correlations with histologic lesions and clinical diagnosis. Figure 4 shows that ARTS scores correlate with the lesion intensity (or extent) of interstitial inflammation, tubulitis and vasculitis in a very similar way to the PBTs used in Mueller *et al.* [8]. Then, we

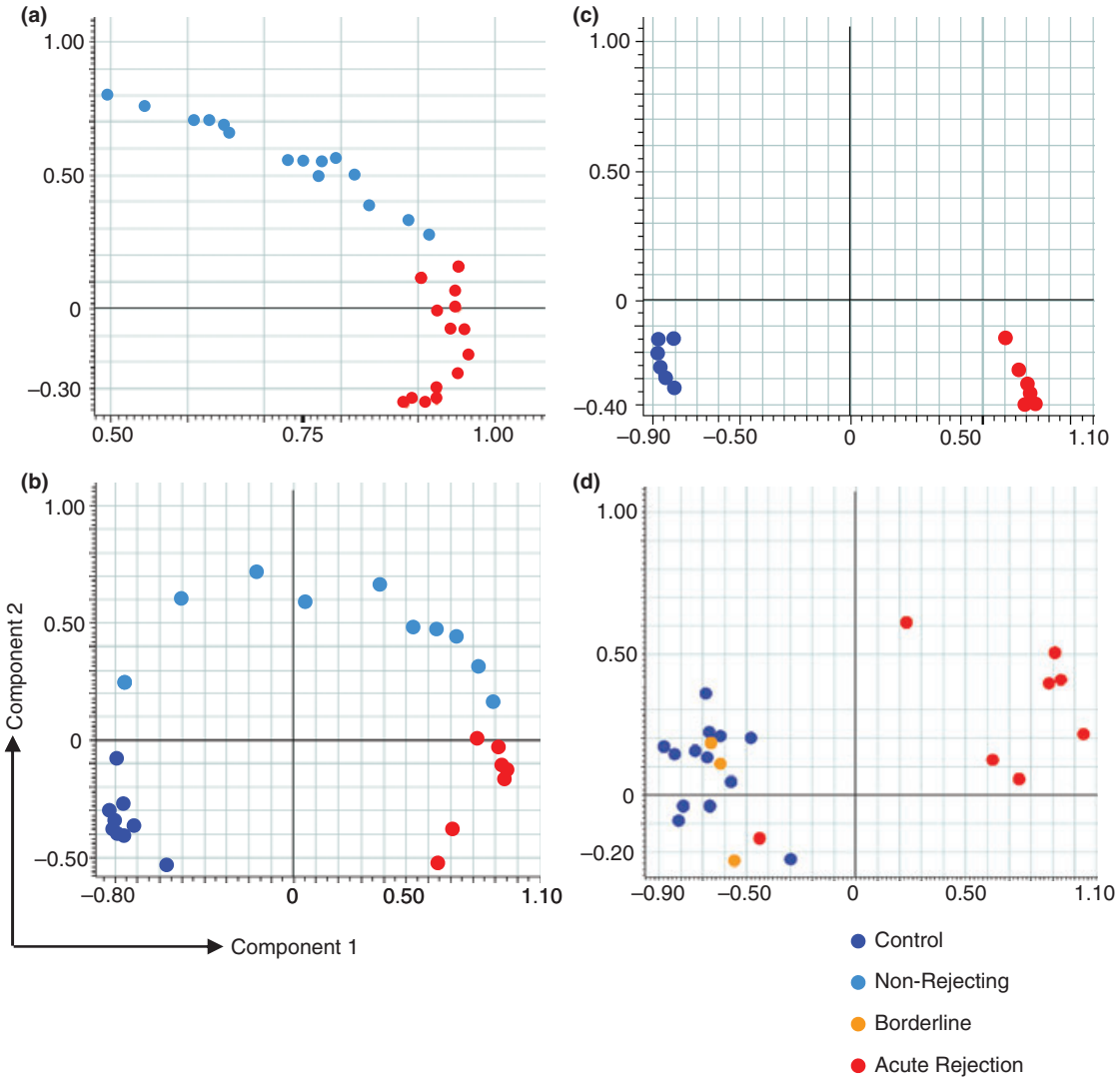

Figure 2 Principal component analysis with the ARTS genes on the four datasets included in the comparative analysis. The Stanford dataset (a, [9]), Cleveland dataset (b, [10]), the NHP dataset (c) and the Paris dataset (d). As the four experiments were not performed using the same technical GeneChip platform, the number of probe sets was different for each analysis, namely 116 probesets for Paris, 129 for NHP, 103 for Cleveland and only 71 for Stanford dataset. This explains in part the different components in each separated analysis. For example, the principal component 1 (*X*-axis) correspond to 76% of the variability in (a), 68% in (b), 89% in (c) and 82% in (d). The ARTS gene list is provided in Table S2. Note that the normal nonrejecting samples in (a) and (b) present a similar scattered profile from controls to AR sample distribution.

tested the diagnostic performances of the ARTS scores. The classifier was built on the same 143 biopsies to determine accuracy in the prediction of histopathology and clinical episode (Table 5). Statistics were calculated with LOOCV and a more stringent 10-fold cross-validation method. Both approaches returned very similar results, indicating the robustness of error estimates. Consistent with our previous results, classification of clinical episode with ARTS display similar performance as the PBTs classifier described [8], with an accuracy of 77% vs. 81% respectively (Table 5).

Discussion

This work identifies a consistent molecular signature for renal acute allograft rejection in biopsies. Gene expression of the signature genes was compared with Banff histopathologic diagnosis. Several previous studies have successfully applied a transcriptomic approach to characterize acute rejection in kidney biopsies. However, the heterogeneity of microarray platforms and various data analysis methods complicates the identification of robust and consistent signatures across experiments as well as

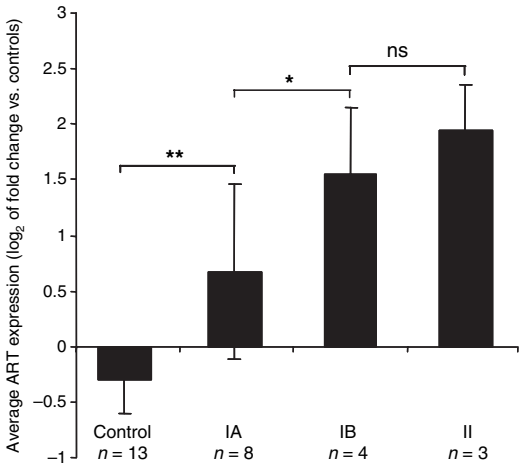

Figure 3 Association between ARTS scores and Banff '97 diagnosis in the Paris dataset. ARTS scores were calculated for each sample and compare with Banff 97 types of acute rejection. Numerical values represented in this study are the average ARTS scores of each diagnostic group (log 2 values ± SD). The *P*-values were calculated by *t*-test. *<0.05, **<0.01. ns = non-significant.

the relatively low numbers of available samples per center. This study takes advantage of previously published transcript profiling data sets of renal biopsies from acute rejecting patients, to propose a comparative analysis together with the newly generated Paris microarray dataset. This new investigation of the data revealed a unique expression pattern of 70 genes, consistently overexpressed in AR across three independent datasets. In addition, this panel of genes was overexpressed in the context of a monkey model of acute renal allograft rejection that closely resembles the clinical situation.

The set of genes generated based on an 'in silico' approach appears to be relevant to the pathogenesis of AR. Most of the genes constituting the ARTS are associated to antigen presenting cells (e.g. HCK, CD52, CD163, RUNX3, MHCs), cytotoxic T lymphocytes (CTLs; e.g. CD8, granzyme A, LCK, LCP1, RAC2), or IFN-γ responses (GBP1, CCL5, CXCL9, IFI30, ISG20, STAT1, PSMB-8, -9, -10, WARS). The relationship between the ARTS and IFN-γ signaling is further supported by a functional transcriptional analysis using MetaCore, which identifies the most significant regulatory networks centered on STAT1 and IRF1. In this respect, the ARTS genes present similarities with the IFN-γ (GRITs) and T cells (CATs) PBTs previously defined in a mouse model of kidney acute rejection [15]. For instance, 23 human ARTS genes were found present in the GRITs gene list. Surprisingly, we identified the IL-10 receptor alpha strongly associated with AR severity. This observation may support a recent study showing that AR was almost fivefold more likely in patients homozygous for a particular IL-10 polymorphism [16]. This striking result might have some important implication for the prognosis of AR.

The ARTS is able to discriminate AR from control in Stanford, Cleveland and the NHP datasets. In addition, the ARTS is strongly correlated with the Banff type of rejection in the Paris dataset. From a clinical perspective, it is important to know that the ARTS data set clearly differentiates between controls and AR type IA. Therefore, the ARTS may allow a very sensitive detection of AR, which could improve the diagnostic of biopsies and related patient treatment. For instance, Fig. 2d shows that one AR type IA samples display a similar profile as control samples, despite an intra-graft cellular infiltrate detected by histology. In this particular case, the patient was subjected to two consecutive biopsies. The first biopsy was diagnosed as borderline, suspicious changes, and the patient received a stronger immunosuppressive treatment. Two months later, the second biopsy was diagnosed as AR IA. However, our gene-expression analysis identified this biopsy as normal. We could hypothesize in this study that the increase in immunosuppressive treatment down-regulated the activity of the immune infiltrate and the expression of the genes associated with the IFN-γ pathway or T-cell activation. This observation suggests that ARTS reflect treatment efficacy more quickly or in a more sensitive way than histology. Because the ARTS consists of a limited number of genes, a PCR-based approach, which shows similar performance than the microarray technology in our hands (Fig. S1) and in the literature [17], could be considered for confirmatory investigations and development of a diagnostic tool.

In a manner very similar to that of CATs and GRITs, the ARTS scores correlate to interstitial inflammation and tubulitis lesions in the Edmonton collection of samples and demonstrate equivalent classifier performance. ARTS might be more sensitive for distinguishing *i*0 versus *i*1 lesions, but did not discriminate *t*1 versus *t*2 lesions or *v*1 versus *v*2 lesions. Hence, these results support previous findings and conclusions from Mueller *et al.* [8], suggesting that microarray results could be used as independent and reproducible measures to provide an objective alternative to the potential weak points of the current classification system. Taken together, our data show that ARTS, CATs or GRITs could be used as sensitive means to monitor the immune system activation and inflammation during organ transplantation.

However, potential drawbacks of this approach should be considered. For instance, the specificity of such gene signatures might be questioned and their ability to discriminate between rejection and other causes of immune activation such as infection or recurrence of an underlying inflammatory disease remains to be established. The

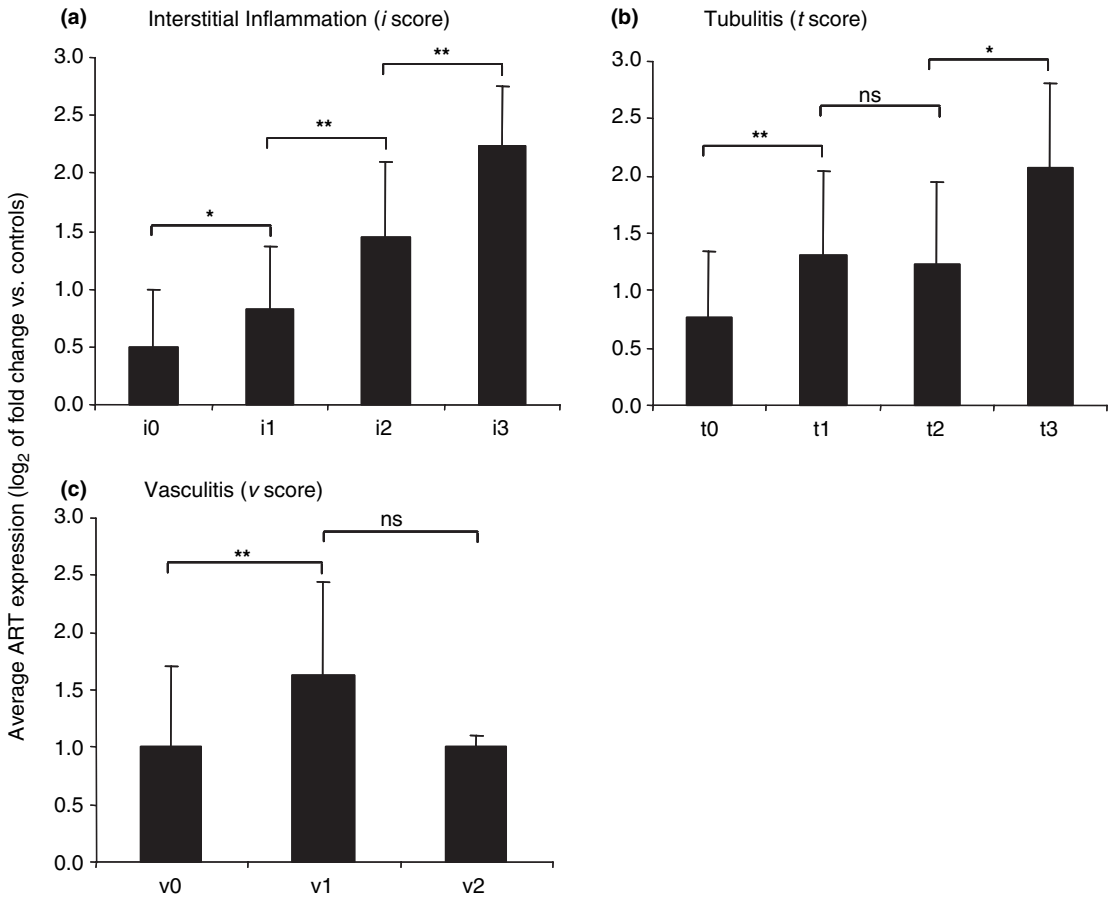

Figure 4 Association between ARTS scores and histopathologic lesion diagnosis in the Edmonton dataset. ARTS scores were calculated for each sample and compared with the degree of interstitial infiltrate (*i*-score) (a), tubulitis (*t*-score) (b), or intimal arteritis (*v*-score) (c). Numerical values represented in this study are the average ARTS scores of each histopathologic scores (0–3) (log 2 values ± SD). The *P*-values were calculated by *t*-test. *<0.05, **<0.01. ns = non-significant.

Table 5. Classifier statistic from cross-validation methods.

Variable to predict	Cross-validation	Accuracy	Sensitivity	Specificity	PPV	NPV
Histopathology	LOOCV	68.1	66.7	68.8	46.4	83.5
Histopathology	10-fold CV	68.1	66.7	68.8	46.4	83.5
Clinical episode	LOOCV	77.0	76.2	77.2	38.1	94.6
Clinical episode	10-fold CV	77.8	76.2	78.1	39.0	94.7

LOOCV, leave-one-out cross-validation; CV, cross-validation; PPV, positive predictive value; NPV, negative predictive value.

next step in the definition of a diagnostic approach would probably require the definition of additional tissue-specific marker-genes or sets as suggested by a recent report of gene-expression study of heart rejection in the context of Chagas infection [18]. In addition, this study is restricted to marker definition of acute cell-mediated rejection (ACR), but markers of antibody-mediated rejection (AMR) are clearly underrepresented in this study and should be further investigated in trials with systematic parallel C4d immunostaining to obtain a fully differential molecular diagnosis, as pathologic findings in renal allograft biopsies show sometimes mixed lesions of ACR and AMR in especially aggressive cases [19].

In conclusion, this study identifies a systematic set of genes for which intra-graft expression levels correlated with renal acute allograft rejection. This relationship is platform- and clinical center-independent and demonstrates a robust consistency with animal models of kidney acute allograft rejection as well as high relevance with the pathophysiology of organ rejection. This 'in silico' approach demonstrates the ability of the PBTs scoring method to support a diagnostic approach and suggest an improvement of sensitivity and accuracy of ambiguous classifications such as borderline. Finally, this work is in accordance with the latest Banff '07 [20] recommendations and should help to promote a consen-

sus on the optimal molecular diagnosis and follow-up of renal AR.

Authorship

HPM, AH, ER, JK, FR: designed study. CCB: performed RNA extractions of biopsies and collected clinical data. GW, MB: designed and supervised the NHP experiments. PSM, HZ, SK, MS, JK: analyzed the data. PSM, FR, HPM: wrote the paper.

Funding sources

Supported by grants from the Swiss National Science Foundation (NFP-46), a personal grant from INSERM, France, to H.P. Marti (poste orange), and financial support by the Établissement Français des Greffes and the Agence de Biomedicine in France (grants to H.P. Marti).

Acknowledgements

We are grateful to C. Deminière, B. Mougenou and S. Ferlicot for histopathology support, to N. Hartmann and her team for expert Gene Chip processing at the Novartis Genomics Factory, Basel, Switzerland, and to C. Delucis-Bronn for RNA and real-time PCR technical support. This work was supported by a research grant from the Swiss National Science Foundation (NFP-46) and a personal grant from INSERM, France, to H.P. Marti (poste orange).

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1 Correlation analysis between qRT-PCR and Gene Chip data.

Table S1 Main clinical parameters and histologic biopsy grading of the individual samples included in the GeneChip analysis (collection Hôpital Tenon, Paris).

Table S2 AR gene-expression signature of 70 genes identified by comparative analysis of four independent datasets.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

1. Meier-Kriesche HU, Schold JD, Kaplan B. Long-term renal allograft survival: have we made significant progress or is it time to rethink our analytic and therapeutic strategies? *Am J Transplant* 2004; **4**: 1289.
2. Cecka JM. The OPTN/UNOS renal transplant registry. UCLA Tissue Typing Laboratory.
3. Yates PJ, Nicholson ML. The aetiology and pathogenesis of chronic allograft nephropathy. *Transpl Immunol* 2006; **16**: 148.
4. Furness PN, Taub N, Assmann KJM, *et al.* International variation in histologic grading is large, and persistent feedback does not improve reproducibility. *Am J Surg Pathol* 2003; **27**: 805.
5. Racusen LC, Solez K, Colvin RB, *et al.* The Banff 97 working classification of renal allograft pathology. *Kidney Int* 1999; **55**: 713.
6. Einecke G, Melk A, Ramassar V, *et al.* Expression of CTL associated transcripts precedes the development of tubulitis in T-cell mediated kidney graft rejection. *Am J Transplant* 2005; **5**: 1827.
7. Weintraub LA, Sarwal MM. Microarrays: a monitoring tool for transplant patients? *Transplant Int* 2006; **19**: 775.
8. Mueller TF, Einecke G, Reeve J, *et al.* Microarray analysis of rejection in human kidney transplants using pathogenesis-based transcript sets. *Am J Transplant* 2007; **7**: 2712.
9. Sarwal M, Chua MS, Kambham N, *et al.* Molecular heterogeneity in acute renal allograft rejection identified by DNA microarray profiling. *N Engl J Med* 2003; **349**: 125.
10. Flechner SM, Kurian SM, Solez K, *et al.* Kidney transplant rejection and tissue injury by gene profiling of biopsies and peripheral blood lymphocytes. *Am J Transplant* 2004; **4**: 1475.
11. Bigaud M, Wieczorek G, Menninger K, Riesen S, Raulf F, Kehren J. Transcriptional profiling specifically discriminates between chronic and acute rejection in non-human primate (NHP) renal allograft recipients [abstract #547]. *Am J Transplant* 2004; **4**(S8): 308.
12. Wieczorek G, Bigaud M, Menninger K, *et al.* Acute and chronic vascular rejection in nonhuman primate kidney transplantation. *Am J Transplant* 2006; **6**: 1285.
13. Rhodes DR, Yu J, Shanker K, *et al.* Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc Natl Acad Sci USA* 2004; **101**: 9309.
14. Guo L, Lobenhofer EK, Wang C, *et al.* Rat toxicogenomic study reveals analytical consistency across microarray platforms. *Nat Biotechnol* 2006; **24**: 1162.
15. Famulski KS, Einecke G, Reeve J, *et al.* Changes in the transcriptome in allograft rejection: IFN-γ-induced transcripts in mouse kidney allografts. *Am J Transplant* 2006; **6**: 1342.
16. Grinyó J, Vanrenterghem Y, Nashan B, *et al.* Association of four DNA polymorphisms with acute rejection after kidney transplantation. *Transpl Int* 2008; **21**: 879.

17. Allanach K, Mengel M, Einecke G, *et al.* Comparing microarray versus RT-PCR assessment of renal allograft biopsies: similar performance despite different dynamic ranges. *Am J Transplant* 2008; **8**: 1006.
18. Morgun A, Shulzhenko N, Perez-Diez A, *et al.* Molecular profiling improves diagnoses of rejection and infection in transplanted organs. *Circ Res* 2006; **98**: 74.
19. Al-Aly Z, Reddivari V, Moiz A, *et al.* Renal allograft biopsies in the era of C4d staining: the need for change in the Banff classification system. *Transpl Int* 2008; **21**: 268.
20. Solez K, Colvin RB, Racusen LV, *et al.* Banff 07 classification of renal allograft pathology: updates and future directions. *Am J Transplant* 2008; **8**: 753.