DISCOVERY OF PERIPHERAL BLOOD AND BIOPSY-BASED MOLECULAR CLASSIFIERS IN BRAZILIAN KIDNEY TRANSPLANT PATIENTS

C. Ventura,[1] S. M. Kurian,[2] T. Gelbart,[2] D. David,[1] F. Agena,[1] M. Abecassis,[3] J. Friedewald,[3] E. David-Neto,[1] D.R. Salomon.[2]

[1]University of Sao Paulo, Sao Paulo, Brazil; [2]The Scripps Research Institute, CA; [3]Northwestern University, Chicago.

INTRODUCTION

Analysis of gene expression in peripheral blood and biopsy samples is increasingly being used in kidney transplant biomarker studies for both disease diagnosis and prognosis. However, these profiles of kidney transplant patients have primarily been studied in European and American populations. Even though knowledge of inter-individual and inter-ethnic variation in gene expression is a prerequisite to select robust biomarkers from a list of candidate genes, very few studies have attempted to characterize such biological variation at a genome wide level.

In this study, we determined both the peripheral blood and biopsy gene expression profiles of Brazilian transplants with Acute Rejection (AR), Acute Dysfunction/No Rejection (ADNR), Chronic Allograft Nephropathy (CAN/IFTA), recurrent and *de novo* Transplant Glomerular Disease (TGDz) and Transplant Excellent (TX). ***This is the first discovery study of molecular phenotypes in Brazilian transplant patients.***

METHODS

A total of 94 precisely histology-phenotyped kidney transplant biopsies and matching peripheral blood samples were collected at the time of biopsy from the transplant clinic at the University of Sao Paulo. All samples were analyzed on Affymetrix HG-U133 Plus PM microarrays using standard manufacturer's protocols. Diagnostic classifiers were determined using a 3-way ANOVA analyses in Partek Genomics Suite. Diagnostic classifiers were constructed using multiple predictive tools such as Nearest Centroids (NC), Diagonal Linear Discriminant Analysis (DLDA), Random Forest and Support Vector Machines (SVM). We used Harrell's method of "optimism corrected" bootstrapping and sampling with replacement to correct for statistical over fitting.

RESULTS

No significant demographic and clinical differences were observed among the phenotypes in this study except a longer time for biopsy in the TGDz group.

*DISCLOSURES: *Kurian, S., Salomon, D. Friedewald, J. Abecassis, M*: Stockholders, TGI.

Demographic and clinical data at biopsy

	ADNR N=10	AR N=23	CAN N=30	TGDz N=14	TX N=17	P
Age ± SD (ys)	41 ± 12	41 ± 13	40 ± 13	48 ± 10	47 ± 14	NS
Male (%)	40%	52%	56%	57%	53%	NS
Non-white race (%)	40%	31%	50%	30%	48%	NS
Time to Bx ± SD (months)	10 ± 8	30 ± 66	62 ± 50	84 ± 83*	49 ± 30	0.01
MDRD (mL/min)	36 ± 11	32 ± 20	29 ± 12	28 ± 16	74 ± 21	NS

For the peripheral blood, from 710 differentially expressed genes (p<0.001), derived from an ANOVA analysis of CAN, TGDz, and TX, classifier signatures with 59-100 probesets were able to distinguish these phenotypes with optimism-corrected AUCs of 0.759-0.815.

Diagnostic metrics for the 3-way CAN, TGDz, TX peripheral blood classifiers

Method	Classifies	Classifiers	% Predictive Accuracy (Single Randomization)	AUC	% Predictive Accuracy (Optimism Corrected 1000 Randomizations)	AUC
SVM	CAN, GLOM, TX	59	100%	1.000	77%	0.815
DLDA	CAN, GLOM, TX	75	83%	0.851	78%	0.793
Nearest Centroid	CAN, GLOM, TX	100	83%	0.850	75%	0.759
Random Forest	CAN, GLOM, TX	100	100%	1.000	73%	0.792

For the biopsies, from 4700 differentially expressed genes (p<0.001), derived from an ANOVA, a 216 probeset classifier was able to distinguish AR from TX with optimism corrected AUCs of 0.870-1.000.

Diagnostic metrics for the 2-way AR, TX biopsy classifier

Method	Classifies	Classifiers	% Predictive Accuracy (Discovery Cohort)	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predicitve Value (%)	AUC
SVM	AR, TX	216	100%	100%	100%	100%	100%	1.000
DLDA	AR, TX	216	89%	77%	96%	93%	87%	0.870
Nearest Centroid	AR, TX	216	91%	77%	100%	100%	87%	0.888
Random Forest	AR, TX	216	100%	100%	100%	100%	100%	1.000

Using 568 differentially expressed probesets from a 3-way ANOVA of CAN vs. TGDz vs. TX, a 106 probeset classifier was obtained, which when tested on the three phenotypes gave optimism-corrected AUCs of 0.833-1.000.

Diagnostic metrics for the 3-way CAN, GLOM, TX biopsy classifier

Method	Classifies	Classifiers	% Predictive Accuracy (Discovery Cohort)	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predicitve Value (%)	AUC
SVM	CAN, GLOM, TX	106	97%	92%	100%	100%	95%	0.958
DLDA	CAN, GLOM, TX	106	82%	68%	100%	100%	86%	0.833
Nearest Centroid	CAN, GLOM, TX	106	83%	77%	100%	100%	85%	0.854
Random Forest	CAN, GLOM, TX	106	100%	100%	100%	100%	100%	1.000

Importantly, we were also able to validate our new 4-way classifier of AR, ADNR, CAN and TX derived originally in a US population from 274 kidney biopsies (Abstract 1668) in this Brazilian cohort with 87-94% Predictive Accuracy, 95-96% Specificity, 91-92% PPV and 85-94% NPV. These results also demonstrate the importance of sample size in constructing classifiers.

CONCLUSION

This is a first discovery study of the identification of molecular phenotypes in Brazilian transplant patients (a very different racial/ethnic mix than European/Caucasian populations). We are in the process of validation of these results in an independently collected cohort of patients with matching demographics from the same transplant center.

This is also the first study to profile blood and biopsies of recurrent and *de novo* transplant glomerular disease (TGDz) and correctly classify it as a distinct molecular entity. We successfully validated our biopsy molecular phenotypes created with a US population in this independent cohort of significantly different racial/ethnic backgrounds. ***These results reveal that there are strong unifying immune mechanisms driving transplant biology despite differences in the racial, ethnic and genetic backgrounds.***