

Diagnostic Applications of Microarrays and Gene Expression in Transplants and Native Organs

by

Katelynn Sveanne Madill-Thomsen

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Medicine

University of Alberta

© Katelynn Sveanne Madill-Thomsen, 2020

## **Abstract**

There is a major unmet need for improved accuracy and precision in the diagnoses of transplant rejection and diseases causing tissue injury. Diagnoses relying on histologic assessments and visual assessments demonstrate significant variation between expert observers (as represented by low kappa values) and cannot assess many biological processes because they do not produce histologic changes. Arbitrarily assigned rules determined by consensus may or may not reflect the true disease phenotype, and the lack of objective diagnostic information presents a challenge to the clinician who is managing the patient's care and making therapeutic decisions. Risks of over- or under-treatment can be serious: many therapies for transplant rejection or for primary diseases are expensive and carry risk for significant adverse effects. Improved diagnostic methods could alleviate healthcare costs by preventing treatment errors, increase treatment efficacy, and ultimately improve outcomes.

Molecular diagnostic assessments using microarrays combined with machine learning algorithms for interpretation have shown promise for increasing diagnostic precision via probabilistic assessments, recalibrating standard-of-care (SOC) diagnostic methods, and clarifying ambiguous cases. This approach can use ensembles of algorithms to increase stability and can provide novel mechanistic insights. These features would benefit biopsy-dependent areas of medicine like transplantation or management of inflammatory diseases such as ulcerative colitis (UC).

The analyses described in this thesis are based on the hypothesis that new molecular systems for biopsy interpretation (i.e. the Molecular Microscope Diagnostic System 'MMDx') would provide insights on disease processes and highly reproducible results from a comparatively small amount of tissue, and would constitute a general approach that could be useful in many new areas of medicine; both in transplantation and in diseases in native organs.

Analyses first focused on establishing the reproducibility and robustness of the techniques used in MMDx, and its relationship to SOC approaches currently in use. The effects of tissue heterogeneity on the MMDx output was studied using kidney transplant cortex and medulla biopsy samples. The frequency and pattern of discrepancies within MMDx-Kidney (between expert observers), within histology (comparing the SOC diagnosis to a diagnosis assigned by an algorithm strictly following Banff guidelines), and finally between MMDx and histology were studied. Once the MMDx-Kidney test was well-defined, the

MMDx system was translated into liver transplants where biopsy assessment is more challenging. MMDx-Liver was assessed with regards to its performance for diagnoses of both T cell-mediated rejection (TCMR) and various forms of injury. Although antibody-mediated rejection (ABMR) remains a controversial diagnosis in liver transplantation, previously annotated rejection-associated transcripts (RATs) were used to search the liver biopsy population for an ABMR phenotype analogous to that in kidney and heart. TCMR in liver biopsies was defined using archetypal analysis (AA), and the molecular findings were compared to SOC histology features. Injury was described in a liver biopsy population also using AA, and a classifier was developed for estimating histologic steatohepatitis (AUC=0.84). Finally, an MMDx system was developed for biopsies of a native organ disease: colon biopsies from patients with diagnosed UC. This was done to assess if methods developed for transplant biopsies have the potential for diagnostics of primary diseases in native tissue. These analyses found that there are multiple immune processes involved in UC disease activity: a dominant inflammatory process mediated by innate immunity and an underlying subtle T cell process. UC was examined using AA, revealing heterogeneity in the biopsy population that was not related to the SOC Endoscopic Mayo Score. This finding suggested that the SOC is not capturing all the information describing disease activity.

This thesis explores a diagnostic system that would fulfill unmet needs in transplanted organs (kidney, liver) as well as in native organs (UC). Main findings indicate that MMDx has major implications for kidney transplant biopsies but can also be expanded and translated for use in transplanted livers (see Chapters 6 and 7). The MMDx approach has translational potential in understanding native organ biopsies, for the diagnosis of disease, tissue injury, and loss of function e.g. UC. If molecular diagnostic systems e.g. MMDx are responsibly developed with proper statistical and machine-learning techniques, appropriately validated, and well-defined in terms of observer guidelines for diagnoses, they have the potential to address the current unmet needs for precise and accurate assessments that clinicians are requesting to improve patient care.

## **Preface**

This thesis represents an original work by Katelynn S. Madill-Thomsen. The research projects included in this document (including the INTERCOMEX, INTERLIVER, and IBD Microarray Projects), of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board as follows: Project Name “INTERCOMEX”, PRO00022226, NCT01299168; Project Name “Diagnostic and Therapeutic Applications of Microarrays in Inflammatory Bowel Disease”, PRO00001994, NCT pending; and Project Name “INTERLIVER”, PRO00022225, NCT03193151.

The research conducted for this thesis forms part of an international research collaboration, led by Dr. Philip F. Halloran at the University of Alberta. The biopsy reference sets referred to in chapters 3-9 were collected with the assistance of all international collaborators from the INTERCOMEX, INTERLIVER, and IBD-Microarray projects, under the direction of Dr. Philip F. Halloran and with the assistance of Anna Hutton, Dr. Martina Mackova, Vido Ramassar, Dr. Robert Polakowski, Dr. Konrad Famulski, Dr. Brendan Halloran, and Katelynn Madill-Thomsen.

Chapter 3 of this thesis has been published as Madill-Thomsen et al, “The Effect of Cortex/Medulla Proportions on Molecular Diagnoses in Kidney Transplant Biopsies: Rejection and Injury Can Be Assessed in Medulla”, American Journal of Transplantation, vol. 17, 2117-2128. I was responsible for the data collection and analysis as well as the manuscript composition. P. Halloran was the supervisory author and contributed study design and manuscript edits. R. Wiggins provided guidance in experimental design and data analysis. F. Eskandary and G. Bohmig provided the samples used in the study, sample labels, and assistance with data analysis and manuscript editing.

Chapter 4 of this thesis has been published as Madill-Thomsen et al, “Discrepancy analysis comparing molecular and histology diagnoses in kidney transplant biopsies”, American Journal of Transplantation, (in press). I was responsible for data collection and analysis as well as manuscript composition. P. Halloran was the supervisory author and contributed study design and manuscript edits. A. Perkowska-Ptasinska provided professional advice, manuscript edits, and samples for the study. F.

Eskandary, G. Einecke, G. Gupta provided samples, associated data and manuscript edits. The MMDx-Kidney Study Group provided samples and data for analysis.

Chapter 6 of this thesis has been published as Madill-Thomsen et al, "The molecular diagnosis of rejection in liver transplant biopsies: First results of the INTERLIVER study", *American Journal of Transplantation*, (in press). I was responsible for data collection and analysis, manuscript composition, and participated in the development of the study design. P. Halloran was the supervisory author and contributed study design, manuscript edits, and analytical direction. A. Perkowska-Ptasinska, G. McCaughan, and S. Feng contributed samples, manuscript comments, and professional advice. The INTERLIVER Study Group provided samples and data for analysis.

## **Dedication**

To Ben, for patience and perseverance, and pushing me to keep going. I will always appreciate your support and sacrifice in letting me take the time to accomplish something like this! To Melanie and Doug for encouraging me to do whatever I wanted 'when I grew up' and always being so proud of everything I did. To Tallin, Jerran, Julia and Lydia for humor and support at every stage, and for celebrating every small or big victory as I went along. To my grandparents, Harold and Leona (Pompa and Grammie) for always making me feel like I could do anything I set my mind to (and for being excited for me every time I went to Boston, no matter how many times I went – 'But did you go to Cheers?'). To my extended family for always asking what I was up to (regardless of how complex it was to explain) and for being huge fans of all the presentations, papers, and abstracts. To **all** the staff of ATAGC and TSI for their support, time, and advice. To Jeff for his amazing statistical guidance and unending patience in teaching me code (I owe you so many nougat bars), and Jess for her unending magic with Microsoft word, Excel, and Adobe (you're truly a wizard!). And finally - to PFH aka 'The Godfather', for getting me – a basic scientist – through this program in just over 3 years despite my lack of any prior medical, computational, or biological background. Thank you for helping me achieve my doctoral dreams! You are the Captain Holt to my Amy Santiago and the Michael Scott to my Pam Beesly. I will always be grateful that you saw some potential in me.

## **Acknowledgements**

I would like to acknowledge the many individuals who contributed knowledge, advice, training, edits, and supervision that was crucial to the completion of the interdisciplinary work contained in this thesis. The interdisciplinary nature of this work would not be possible without a large team contributing expertise from a wide array of professional backgrounds, and while I cannot name everyone who submitted samples, expertise, or advice to these projects I would like to thank everyone who was involved in the INTERCOM/INTERCOMEX, INTERLIVER, and IBD-Microarray projects as this work would not have been possible without their contributions.

My supervisor Philip F. Halloran as primary supervisor contributed assistance with study design, data analysis, manuscript preparation for all published works, training in immunology and microarray technologies, and edits. He led the team on funding requests, project design, and established the enormous network of collaborators who submitted biopsies to our studies. None of the work contained in this thesis would have been possible without his expertise, hard work, and team leadership.

My supervisory committee members, for their contributions to these projects and to my education throughout this program. Luis Hidalgo provided mentorship in basic immunology, and thesis edits. Brendan Halloran provided clinical mentorship and insights into the practical components of biopsy collection and clinical SOC, while collecting and contributing biopsies for the IBD-Microarray project.

The team members of both the Alberta Transplant Applied Genomics Center (ATAGC) and Transcriptome Sciences Incorporated (TSI) contributed advice, training, work, expertise, and analysis. These studies require an interdisciplinary team, and ATAGC and TSI truly exemplify what can be accomplished through cooperative work from group of individuals with diverse skill sets. Jeff Venner was a co-author on the work done in inflammatory bowel disease and this cooperative partnership resulted in these analyses. Anna Hutton contributed training and education in all benchwork and lab techniques. All biopsies were processed by Anna Hutton, Martina Mackova and Vido Ramassar. Jeff Reeve contributed training, guidance, and supervision in statistical techniques and programming tasks. Konrad Famulski provided training in pathway analysis, sample selection, and data interpretation. Michael Parkes contributed data analysis advice, assistance with data set creation, and expertise on code assembly. Rob

Polakowski and Jessica Chang managed communications and sample flow from all collaborators. Shane Pon helped collect and manage the clinical and histologic data provided by the centers.

All the international collaborators in the INTERCOMEX, INTERLIVER, and IBD-Microarray projects for submitting biopsies to these studies, taking the time for phone calls to review results and reports, submitting feedback on reports and manuscripts, and supporting these studies. Without this assistance in building the data sets, these analyses would not be possible.

All the funding agencies that have supported the work of ATAGC and TSI throughout the years: Genome Canada, Genome Alberta, the University of Alberta, the University of Alberta Hospital Foundation, Roche Molecular Systems, Hoffman-La Roche Canada Ltd., Alberta Innovates, Alberta Advanced Education and Technology, the Roche Organ Transplant Research Foundation, the Kidney Foundation of Canada, Astellas Canada, the Canada Research Chair in Transplant Immunology, the Muttart Chair in Clinical Immunology, the Mendez Foundation, the Canadian Foundation for Innovation, and the National Research Council Industrial Research Program.

## Table of Contents

### CHAPTER 1: INTRODUCTION

1.1 THE UNMET DIAGNOSTIC NEEDS IN TRANSPLANTS AND NATIVE ORGANS.....	2
1.2 BACKGROUND .....	2
1.2.1 A review of kidney transplantation.....	3
1.2.1.1 The kidney and its function.....	3
1.2.1.2 Renal transplant rejection.....	4
1.2.1.2.1 Renal transplant ABMR.....	4
1.2.1.2.2 Renal transplant TCMR.....	5
1.2.1.3 Renal transplant injury.....	7
1.2.1.4 The Banff guidelines for kidneys.....	7
1.2.1.5 Clinical correlates of kidney function.....	8
1.2.1.6 Molecular studies of kidney rejection and injury to date.....	9
1.2.1.7 Tissue heterogeneity and gene expression in kidneys.....	10
1.2.1.8 Current limitations of clinical and histologic practice.....	11
1.2.2 A review of liver transplantation.....	11
1.2.2.1 The liver and its function.....	11
1.2.2.2 Liver transplantation.....	12
1.2.2.3 Transplant rejection in liver.....	13
1.2.2.4 ABMR in liver transplants.....	13
1.2.2.5 Transplant injury in livers.....	14
1.2.2.6 Clinical correlates of liver function.....	15
1.2.2.7 Liver injury and non-alcoholic fatty liver disease (NAFLD).....	16
1.2.2.8 The Banff guidelines for liver transplants.....	17
1.2.2.9 Molecular studies to date.....	18
1.2.2.10 Limitations in the SOC for liver transplant biopsy interpretation.....	18
1.2.3 A review of UC – a form of inflammatory bowel disease.....	19
1.2.3.1 Overview of UC in the colon.....	19
1.2.3.2 UC disease mechanisms.....	19
1.2.3.3 Endoscopic correlates of UC disease activity.....	20
1.2.3.4 Molecular studies in UC.....	20
1.2.3.5 Limitations of the current SOC in UC.....	21
1.2.4 A review of machine learning and artificial intelligence in medicine.....	21

1.2.4.1 Supervised analyses with microarray data .....	22
1.2.4.1.1 Machine learning algorithms .....	22
1.2.4.1.2 Class comparisons for microarray data .....	23
1.2.4.2 Unsupervised analysis with microarray data .....	23
1.2.4.3 Probabilities and continuous data in medicine .....	24
1.4 TABLES .....	25
1.5 FIGURES .....	27

## CHAPTER 2: METHODS

2.1 OVERVIEW .....	32
2.1.1 Samples used for histologic and microarray assessment .....	32
2.2 BIOPSY COLLECTION .....	33
2.2.1 Renal transplant biopsy collection .....	33
2.2.1.1 Standard renal transplant biopsy collection .....	33
2.2.1.2 Cortex-medulla renal biopsy cohort collection .....	33
2.2.1.3 Renal biopsy biological replicate collection .....	34
2.2.1.4 Renal biopsy technical replicate biopsy collection .....	34
2.2.2 Liver biopsy collection .....	34
2.2.3 UC sample collection .....	35
2.3 PATIENT POPULATIONS AND DEMOGRAPHICS .....	35
2.3.1 Kidney population .....	35
2.3.2 Liver population .....	35
2.3.3 UC population .....	36
2.4 ASSESSMENT OF HUMAN TISSUE BIOPSIES BY MICROARRAY .....	37
2.4.1 Samples assessed by microarray .....	37
2.4.2 RNA extraction, labeling, hybridization, and microarray processing .....	37
2.4.2.1 RNA extraction .....	37
2.4.2.2 Quality and quantity assessments of extracted RNA .....	38
2.4.2.3 Microarray processing .....	38
2.4.3 Microarray data preparation .....	39
2.5 DATA ANALYSIS .....	39
2.5.1 PBT definition and methods .....	39

2.5.2 Principal component analysis (PCA) definition and methods.....	40
2.5.2.1 PCA in kidneys.....	40
2.5.2.2 PCA in livers .....	41
2.5.2.3 PCA in UC.....	41
2.5.3 AA definition and methods.....	41
2.5.3.1 AA in Kidneys.....	42
2.5.3.2 AA in Livers .....	42
2.5.3.3 AA in Colon .....	43
2.5.4 Definition of a molecular classifier and CV .....	43
2.5.5 Developing an ensemble system.....	43
2.5.6 Data preprocessing.....	44
2.5.7 Transcript filtering and selection .....	44
2.5.8 Overrepresentation analysis .....	45
2.5.9 Moving average time course analysis .....	46
2.5.10 Discrepancy analyses.....	46
2.5.10.1 Discrepancies within MMDx.....	46
2.5.10.2 Discrepancies within histology.....	47
2.5.10.2 Discrepancies between MMDx and histology.....	48
2.6 ASSESSMENT OF HUMAN ALLOGRAFT BIOPSIES .....	48
2.6.1 Human biopsy pathology for kidneys (Banff Classification System) .....	48
2.6.1.1 Diagnostic classifications.....	49
2.6.2 Human biopsy pathology for liver (Banff working classification system).....	50
2.6.3 Human biopsy pathology and endoscopic assessment for UC .....	51
2.7 FUNCTIONAL ASSESSMENTS FOR CLINICAL DIAGNOSES.....	52
2.7.1 Assessments of kidney function .....	52
2.7.2 Assessments of liver function .....	52
2.7.3 Assessments of colonic function.....	52

## **SECTION I - ANALYSES IN KIDNEY TRANSPLANTS**

## **CHAPTER 3: THE EFFECT OF TISSUE HETEROGENEITY ON MOLECULAR DIAGNOSIS IN RENAL TRANSPLANT BIOPSIES - REJECTION AND INJURY CAN BE ASSESSED IN MEDULLA**

3.1 OVERVIEW AND HYPOTHESIS.....	55
3.2 BIOPSY COLLECTION AND PROCESSING.....	55
3.2.1 Cortex-medulla cohort preparation and processing.....	56
3.2.2 Biological replicate pair collection and processing.....	57
3.2.3 Technical replicate pair collection and processing.....	57
3.3 HISTOLOGIC ASSESSMENT OF THE BIOPSIES.....	57
3.4 MMDx ASSESSMENT OF THE BIOPSIES.....	57
3.5 DEVELOPING A MOLECULAR ESTIMATE OF PROPORTION CORTEX.....	58
3.5.1 Renal structure.....	58
3.5.2 The development of a classifier-based prediction of biopsy composition.....	58
3.5.2.1 Gene expression differences between cortex and medulla.....	59
3.5.2.2 Podocin/NPHS2 expression in renal tissue.....	59
3.5.2.3 Estimating proportion cortex using NPHS2.....	59
3.5.2.4 Expression of podocin/NPHS2 in a reference set of renal biopsies (N=1208).....	60
3.5.2.5 The MMDx 'percent cortex' classifier score.....	60
3.6 EFFECT OF THE PERCENT CORTEX SCORE ON MMDx READINGS.....	60
3.6.1 Effect of percent cortex estimates on the reference set MMDx readings.....	60
3.6.2 Effect of proportion cortex on the molecular rejection and injury scores.....	61
3.6.3 MMDx assessment of cortex-medulla pairs and biological replicate pairs by multiple classifiers.....	62
3.7 INTERPRETATION OF THE RESULTS.....	62
3.8 TABLES.....	67
3.9 FIGURES.....	72

## **CHAPTER 4: DISCREPANCY ANALYSIS BETWEEN MOLECULAR AND HISTOLOGIC ASSESSMENTS OF REJECTION IN RENAL TRANSPLANTATION**

4.1 OVERVIEW AND HYPOTHESIS.....	82
4.2 BIOPSY POPULATION.....	83
4.3 STATISTICAL AND CLASSIFICATION METHODS FOR DISCREPANCY ANALYSIS.....	83
4.3.1 Computational methods for analyses.....	83
4.3.2 Converting platforms to six diagnostic classes.....	84
4.3.3 'Clear' versus 'boundary' discrepancy classifications.....	84

4.4 ASSESSING DISCREPANCIES WITHIN MMDx .....	85
4.4.1 Reproducibility of raw MMDx scores .....	85
4.4.2 Assignment of alternative MMDx sign-out comments .....	85
4.4.3 Discrepancies within MMDx sign-out comments .....	86
4.5 DISCREPANCIES BETWEEN MMDx SIGN-OUTS AND HISTOLOGIC DIAGNOSIS .....	87
4.5.1 Assessing discrepancies as a proportion of histology categories .....	87
4.5.2 Assessing discrepancies as a proportion of MMDx sign-out categories .....	87
4.6 VISUALIZING HISTOLOGY-MMDx DISCREPANCIES .....	87
4.7 COMPARING CLEAR, BOUNDARY, AND MIXED DISCREPANCIES: HISTOLOGY-MMDx VERSUS INTRA-MMDx.....	88
4.7.1 Expressing overall agreement through kappa values .....	88
4.8 DISCREPANCIES IN PREVIOUSLY FLAGGED PROBLEMATIC SCENARIOS .....	89
4.8.1 Agreement in problematic scenarios described by kappa values .....	89
4.9 INTERPRETATION OF THE RESULTS.....	89
4.10 TABLES .....	93
4.11 FIGURES .....	106

**CHAPTER 5: INTRA-HISTOLOGY DISCREPANCIES IN RENAL TRANSPLANTATION -  
COMPARISONS BETWEEN THE SOC ASSESSMENTS AND AN ALGORITHM STRICTLY  
FOLLOWING BANFF GUIDELINES FOR REJECTION DIAGNOSES**

5.1 OVERVIEW AND HYPOTHESIS.....	111
5.2 BIOPSY POPULATION.....	112
5.3 DEVELOPMENT OF THE AUTOBANFF ALGORITHM .....	112
5.4 STATISTICAL ANALYSES .....	113
5.5 COMPARING SOC HISTOLOGY DIAGNOSES TO AUTOBANFF ASSIGNMENTS .....	113
5.5.1 Discrepancies between ExpertBanff and AutoBanff.....	113
5.6 DISCREPANCIES IN PREVIOUSLY NOTED PROBLEMATIC AREAS.....	113
5.6.1 TCMR lesions in BK.....	114
5.6.2 Rejection lesions in scarred biopsies .....	114
5.6.3 ABMR lesions with negative or ambiguous DSA.....	114
5.6.4 V>0 lesions in rejection.....	115
5.6.5 Discrepancies in biopsies with tubulitis lesions.....	116

5.7 RANDOM FORESTS FOR VARIABLE IMPORTANCE IN THE PREDICTION OF DISCREPANCIES .....	116
5.8 RELATIONSHIPS BETWEEN INTRAHISTOLOGY DISCREPANCIES AND MMDx PHENOTYPES	117
5.8.1 Visualizing IH discrepancies within the context of MMDx diagnoses .....	117
5.8.2 IH discrepancies as a contributor to molecular-histology discrepancies .....	117
5.9 INTERPRETATION OF RESULTS .....	118
5.10 TABLES .....	123
5.11 FIGURES .....	138

## **SECTION II: ANALYSES IN LIVER TRANSPLANTS**

### **CHAPTER 6: ASSESSING TRANSPLANT REJECTION IN LIVER BIOPSIES - THE DEVELOPMENT OF MMDX FOR DIAGNOSTICS WITHIN A LIVER TRANSPLANT POPULATION**

6.1 OVERVIEW AND HYPOTHESIS .....	145
6.2 BIOPSY POPULATION .....	146
6.3 ABMR AND TCMR RAT TRANSCRIPT SEPARATION IN LIVER BIOPSIES .....	146
6.3.1 Derivation of the RAT Transcript Set .....	146
6.3.2 PCA visualization of RAT Expression in Kidneys, Hearts and Livers. ....	147
6.4 AA OF LIVER REJECTION USING RAT TRANSCRIPT EXPRESSION .....	148
6.4.1 AA in livers for unsupervised analysis .....	148
6.4.2 Selection of the rejection-based clusters for AA in the liver transplant population.....	148
6.4.3 Visualization of the liver transplant rejection AA clusters in PCA.....	148
6.4.4 Characterization of the AA clusters .....	148
6.4.4.1 PBTs selected for analysis .....	148
6.4.4.2 PBT expression in the RAT-based AA clusters .....	149
6.4.5 Transcripts correlated with each RAT AA score .....	149
6.5 CHARACTERIZING THE PRINCIPAL COMPONENTS (PCS) IN LIVER TRANSPLANT REJECTION .....	150
6.6 TIME COURSE OF MOLECULAR AA AND PCA SCORES IN THE LIVER TRANSPLANT POPULATION .....	150
6.7 CORRELATIONS WITH THE SOC HISTOLOGIC AND CLINICAL DATA .....	151
6.7.1 Correlations between molecular scores and histologic features .....	151
6.7.2 Relationships between molecular scores and liver biochemistry measurements .....	151

6.7.3 Relationships between molecular scores and histologic rejection .....	152
6.8 DIAGNOSTIC INTERPRETATION OF MMDx-LIVER REPORTS.....	152
6.9 INTERPRETATION OF THE RESULTS.....	152
6.10 TABLES .....	157
6.11 FIGURES .....	172

**CHAPTER 7: ASSESSING INJURY PHENOTYPES IN A LIVER TRANSPLANT POPULATION USING MMDX - A NEW METHOD OF DETECTING STEATOHEPATITIS**

7.1 OVERVIEW AND HYPOTHESIS.....	181
7.2 BIOPSY POPULATION AND DEMOGRAPHICS .....	182
7.3 PCA OF INJURY IN LIVER TRANSPLANTS .....	182
7.4 AA OF LIVER INJURY PHENOTYPES .....	184
7.4.1 AA group selection.....	184
7.4.2 Characterization of injury in the four AA groups .....	185
7.4.3 Transcripts associated with each injury archetype score .....	185
7.4.4 Injury AA group assignment versus previously defined rejection AA group assignment .....	186
7.5 TIME COURSE OF MOLECULAR SCORES AND FEATURES .....	187
7.6 CORRELATIONS BETWEEN INJURY ARCHETYPES AND CLINICAL FEATURES .....	187
7.7 OVERREPRESENTATION ANALYSIS OF ARCHETYPE GROUP TOP TRANSCRIPT LISTS .....	188
7.8 LIVER FUNCTION TESTS AND INJURY ARCHETYPE GROUPS.....	188
7.9 DEVELOPMENT OF A MOLECULAR CLASSIFIER FOR STEATOHEPATITIS .....	189
7.9.1 Steatohepatitis classifier performance versus other molecular features .....	189
7.9.2 Top transcripts associated with steatohepatitis by t-test .....	190
7.9.3 Injury PCA visualization of steatohepatitis versus no steatohepatitis as called by histology and by the molecular classifier .....	190
7.9.4 Logistic regression between the Msgt0 classifier, PC3 score, and steatohepatitis.....	190
7.9.5 Overrepresentation of genes associated with steatohepatitis .....	191
7.10 INTERPRETATION OF RESULTS .....	191
7.11 TABLES .....	196
7.12 FIGURES .....	214

**SECTION III: DEVELOPING MMDX FOR NATIVE ORGANS – MOLECULAR DIAGNOSTICS IN ULCERATIVE COLITIS**

## **CHAPTER 8: THE MOLECULAR LANDSCAPE OF ULCERATIVE COLITIS DISEASE ACTIVITY**

8.1 OVERVIEW AND HYPOTHESIS .....	223
8.2 STUDY POPULATION AND DEMOGRAPHICS .....	224
8.2.1 Biopsy and data collection .....	224
8.2.2 Biopsy processing .....	225
8.2.3 Data processing .....	226
8.3 TOP TRANSCRIPTS EXPRESSED IN UC BIOPSIES .....	226
8.3.1 Volcano plot of transcript expression in UC .....	226
8.3.2 Inflammasome and Toll-like receptor transcripts .....	227
8.3.3 Transcripts associated with the NLRP3 inflammasome .....	227
8.3.4 Transcripts associated with epithelial stem cell memory damage .....	227
8.4 EXPRESSION OF UC DISEASE ACTIVITY TRANSCRIPTS IN CULTURED HUMAN CELL LINES .....	227
8.5 EXPRESSION OF UC DISEASE ACTIVITY-ASSOCIATED TRANSCRIPTS IN TCMR BIOPSIES .	228
8.6 EXPRESSION OF T CELL-SPECIFIC AND T CELL PROCESS-ASSOCIATED TRANSCRIPTS IN UC .....	229
8.6.1 T cell-specific transcripts and UC disease activity .....	229
8.6.2 TCMR-associated transcripts and UC disease activity. ....	229
8.7 OVERLAP BETWEEN TOP 300 UC ACTIVITY-ASSOCIATED AND TOP 300 T CELL PROCESS- ASSOCIATED TRANSCRIPTS .....	230
8.8 OVERREPRESENTATION ANALYSIS: UC ACTIVITY-ASSOCIATED VERSUS TCMR-ASSOCIATED TRANSCRIPTS .....	230
8.9 INTERPRETATION OF RESULTS .....	231
8.10 TABLES .....	235
8.11 FIGURES .....	248

## **CHAPTER 9: ASSESSING HETEROGENEITY IN ULCERATIVE COLITIS USING ARCHETYPAL ANALYSIS**

9.1 OVERVIEW AND HYPOTHESIS .....	252
9.2 BIOPSY POPULATION AND DEMOGRAPHICS .....	252
9.3 ESTABLISHING UC ARCHETYPE GROUPS .....	252
9.3.1 AA initial input and determining the number of the groups .....	253
9.3.2 Characterizing the AA groups .....	253

9.3.3 Evidence for epithelial stem cell memory damage in AA groups .....	253
9.4 RELATIONSHIPS BETWEEN MOLECULAR AND CONVENTIONAL FEATURES .....	254
9.5 CROSS TABULATIONS OF ARCHETYPE GROUP CHARACTERISTICS.....	254
9.5.1 Mayo score improvement in archetype groups.....	255
9.6 LOGISTIC REGRESSION ANALYSIS PREDICTING FUTURE IMPROVEMENT IN MAYO SCORE .....	255
9.7 MONITORING UC DISEASE RESPONSE TO THERAPY USING AA .....	256
9.8 INTERPRETATION OF RESULTS .....	256
9.9 TABLES .....	260
9.10 FIGURES .....	270
<b>SECTION IV: CONCLUSIONS</b>	
CONCLUSIONS AND CUMULATIVE SIGNIFICANCE .....	273
REFERENCES.....	276

## LIST OF TABLES

### CHAPTER 1

Table 1.1 Endoscopic Mayo Scoring system for UC. Adapted from E. Paine, Colonoscopic evaluation in ulcerative colitis. Gastroenterology Report (2014) <sup>157</sup> .....	26
--	----

### CHAPTER 2

### CHAPTER 3

Table 3.1 Demographics of the patients and the biopsy sets .....	68
Table 3.2 Top 30 probe sets regarding fold change between cortex and medulla using a paired t-test....	69
Table 3.3 Chi square test statistical values for Figures 3.6 (A-I) and 3.7 (A-C) <sup>1</sup> .....	70
Table 3.4 Consistency in the molecular diagnosis between paired cortex-medulla samples and in biological replicates.....	71

### CHAPTER 4

Table 4.1 Participating centers in the INTERLIVER study as of December, 2019 .....	94
Table 4.2 Demographics and clinical features of the 1679 biopsy cohort .....	95
Table 4.3 Histologic diagnoses and DSA status in the 1679 cohort.....	96
Table 4.4 Classifications for molecular-histology diagnoses comparisons.....	97
Table 4.5 Agreement in 6 classes between Expert2MMDx and Expert1MMDx sign-out comments .....	98
Table 4.6 All clear discrepancies between alternative molecular sign-out comments ordered from most to least frequent .....	99
Table 4.7 Agreement in 6 classes between the AutoMMDx and Expert1MMDx sign-out comments .....	100
Table 4.8 Relating histologic rejection diagnoses to Expert1MMDx rejection sign-out comments (6 classes, N=1679) .....	101
Table 4.9 Histologic 'No Rejection' subclasses (N=939/1679).....	102
Table 4.10 All clear histology-Expert1MMDx discrepancies ordered from most to least frequent (N=315) .....	103
Table 4.11 All histology-MMDx, AutoMMDx/Expert1MMDx, and Expert2MMDx/Expert1MMDx discrepancies classified as clear, boundary, or mixed-related (N=624, N=149, and N=108 respectively) .....	104
Table 4.12 Kappa values for histology-Expert1MMDx and Expert2MMDx-Expert1MMDx disagreement within previously flagged problematic scenarios (N=1679) .....	105

### CHAPTER 5

Table 5.1 Translating SOC histology diagnoses into six classes to permit comparison with AutoBanff and MMDx diagnoses. ....	124
Table 5.2 Frequency of recorded canonical lesion and feature values in the 1679 cohort. ....	125
Table 5.3 Detail on the 21 cases with TMA in 1679. ....	126
Table 5.4 Agreement between AutoBanff and ExpertBanff in N=1679 (6 classes).....	127
Table 5.5 All clear IH discrepancies ordered from most to least frequent (N=202). ....	128
Table 5.6 All boundary IH discrepancies ordered from most to least frequent (N=237).....	129
Table 5.7 The eight most frequent ( $\geq 20$ ) IH discrepancies out of 439 total discrepancies, classified as clear or boundary. ....	130
Table 5.8 BK virus versus TCMR diagnoses by ExpertBanff and/or AutoBanff (N=190).....	131
Table 5.9 ci 0/1 versus 2/3 lesions in rejection diagnoses by ExpertBanff and/or AutoBanff.....	132
Table 5.10 DSA positive, DSA negative, or DSA ambiguous/not recorded versus ABMR diagnoses by ExpertBanff and/or AutoBanff (N=380) .....	133
Table 5.11 $v>0$ vs $v=0$ lesions in rejection diagnoses by ExpertBanff and/or AutoBanff .....	134
Table 5.12 Cases with $v>0$ lesions called NR by ExpertBanff (N=9).....	135
Table 5.13 Tubulitis lesions (t) 0/1 versus 2/3 lesions versus diagnoses of any rejection (ABMR, TCMR, Mixed) by ExpertBanff and/or AutoBanff (N=593) .....	136
Table 5.14 Summary of previously noted problematic guidelines that contribute to IH discrepancies.....	137

## CHAPTER 6

Table 6.1 Participating Collaborators in the INTERLIVER study .....	158
Table 6.2 INTERLIVER patient and biopsy characteristics .....	159
Table 6.3 Laboratory test data and DSA for all biopsies in the INTERLIVER study (N=235) .....	160
Table 6.4 Mean pathogenesis-based transcript (PBT) set scores in biopsies grouped according to their highest Rejection (RAT4AA) archetype score .....	161
Table 6.5 Top 10 transcripts correlated with the RAT4A $R_{1normal}$ , $R_{2TCMR}$ , $R_{3injury}$ , and $R_{4late}$ scores. ....	162
Table 6.6 Top 30 unique transcripts associated with RAT4A Normal Archetype Score 1, sorted by Spearman correlation coefficient (negative to positive). ....	163
Table 6.7 Top 30 unique transcripts associated with RAT4A TCMR Archetype Score 2, sorted by Spearman correlation coefficient (positive to negative). ....	164
Table 6.8 Top 30 unique transcripts associated with RAT4A Injury Archetype Score 3, sorted by Spearman correlation coefficient (positive to negative). ....	165
Table 6.9 Top 30 unique transcripts associated with RAT4A Late Injury Archetype Score 4, sorted by Spearman correlation coefficient (positive to negative). ....	166
Table 6.10 Correlations between pathogenesis-based transcript (PBT) set scores and Rejection (RAT4AA) PCA scores. ....	167
Table 6.11 Correlations between RAT4A archetype scores, PC scores and high scores for histologic features ( $>0$ ) in liver biopsies (N=235).....	168

Table 6.12 Laboratory test data for biopsies in INTERLIVER by RAT4A archetype group.....	169
Table 6.13 Top 30 unique genes associated with histologic acute rejection >0 by adjusted p value. ....	170
Table 6.14 Characterizing the relationship between histologic and molecular diagnoses in the liver biopsy population (N=235).....	171

## CHAPTER 7

Table 7.1 Participating Collaborators in the INTERLIVER study .....	197
Table 7.2 INTERLIVER patient and biopsy characteristics .....	198
Table 7.3 Laboratory test data and DSA for all biopsies in the INTERLIVER study (N=337) .....	199
Table 7.4 Pathogenesis-based transcript sets <sup>A,B</sup> (PBTs) used in liver analyses .....	200
Table 7.5 Correlations between pathogenesis-based transcript (PBT) set scores and Injury (INJ4AA) PCA scores.....	201
Table 7.6 Mean pathogenesis-based transcript (PBT) set scores in biopsies grouped according to their highest Injury (INJ4AA) archetype score and correlations between PBT set scores and archetypes scores (N=337) .....	202
Table 7.7 Top 10 transcripts correlated with the Injury4A I1, I2 <sub>early-mild</sub> , I3 <sub>severe</sub> , and I4 <sub>late</sub> scores .....	203
Table 7.8 Top 30 unique transcripts associated with Injury4AA Archetype Score 1 (I1 minimal), sorted by Spearman correlation coefficient. ....	204
Table 7.9 Top 30 unique transcripts associated with Injury4AA Archetype Score 2 (I2 early-minimal), sorted by Spearman correlation coefficient.....	205
Table 7.10 Top 30 unique transcripts associated with Injury4AA Archetype Score 3 (I3 severe), sorted by Spearman correlation coefficient. ....	206
Table 7.11 Top 30 unique transcripts associated with Injury4AA Archetype Score 4 (I4 fibrosis), sorted by Spearman correlation coefficient. ....	207
Table 7.12 Crosstab of previously defined Rejection AA classes vs. Injury AA classes (N=235).....	208
Table 7.13 Correlations between injury archetype scores, PC scores and histologic rejection in liver biopsies (N=337).....	209
Table 7.14 Top 10 GO Terms from Injury AA group top 300 transcripts by Spearman correlation coefficient absolute value.....	210
Table 7.15 Laboratory test data for biopsies in INTERLIVER by RAT4A or Injury4A archetype group ...	211
Table 7.16 Top 10 genes increased and decreased with histologic steatohepatitis ordered by adjusted p value.....	212
Table 7.17 Top 10 overrepresented GO Terms from steatohepatitis top 300 transcripts (increased and decreased) by Spearman correlation coefficient absolute value .....	213

## CHAPTER 8

Table 8.1 Demographics of the patients and the biopsies for the molecular landscape of UC (N=71 biopsies from 61 patients).....	236
---	-----

Table 8.2 Demographics of the patients and biopsies for the screening colon biopsy cohort (N=16 biopsies from 16 patients).....	237
Table 8.3 Top 30 transcripts increased in UC biopsies with Endoscopic Mayo Score >1 vs. ≤1 (IQR filtered list), aligned by p-value .....	238
Table 8.4 Top 30 transcripts decreased in UC biopsies with Endoscopic Mayo Score >1 vs. ≤1, aligned by p-value.....	239
Table 8.5 UC activity-associated top 30 genes (redundant, by p-value) found in the TCMR-associated IQR filtered list in various biopsy populations, ordered by rank of transcript in kidney biopsies .....	240
Table 8.6 T cell process-associated/T cell-specific gene expression in UC colon and screening colon biopsies (IQR filtered). .....	241
Table 8.7 Expression of Kidney TCMR-associated Top 30 genes (redundant, by p value) found in the UC IQR filtered list (Endoscopic Mayo Score >1 vs. ≤1) in various biopsy populations, ordered by rank of transcript in kidney biopsies .....	242
Table 8.8 Gene Ontology – Biological Process terms from top 300 UC-associated transcripts (Endoscopic Mayo scores >1 versus ≤1) <sup>a</sup> .....	243
Table 8.9 Gene Ontology – Cellular compartment pathway analysis on top 300 UC-associated transcripts (Endoscopic Mayo scores >1 vs. ≤1) <sup>a</sup> .....	244
Table 8.10 Gene Ontology – Molecular function terms from top 300 UC-associated transcripts (Endoscopic Mayo scores >1 vs. ≤1) <sup>a</sup> .....	245
Table 8.11 KEGG pathway analysis on top 300 UC-associated transcripts (Endoscopic Mayo scores >1 vs. ≤1) <sup>a</sup> .....	246
Table 8.12 Comparisons of top 10 overrepresented terms in TCMR-associated top 300 <sup>a</sup> transcripts in kidney biopsies compared to top ten overrepresented terms from top 300 <sup>a</sup> UC activity-associated transcripts in UC biopsies .....	247

## CHAPTER 9

Table 9.1 Top 30 transcripts associated with Archetype group 1 (redundant list).....	261
Table 9.2 Top 30 transcripts associated with Archetype group 2 (redundant list).....	262
Table 9.3 Top 30 transcripts associated with Archetype group 3 (redundant list).....	263
Table 9.4 Archetype scores and Spearman correlations between SOC and molecular features .....	264
Table 9.5 Biopsy endoscopic Mayo scores tallied per archetype group.....	265
Table 9.6 UC patients with initial Endoscopic Mayo score 2/3 in archetype groups A1, A2, and A3 classified by Mayo score response on a follow up biopsy .....	266
Table 9.7 Summary of recorded treatment for biopsies from the A2 and A3 archetype groups .....	267
Table 9.8 Detailed treatment data on all archetype 2/3 biopsies with high Mayo scores (2/3) used in the logistic regression analysis (Table 5).....	268
Table 9.9 Comparative results from all logistic regression models assessing predictive value (response vs non-response as represented by Mayo Scores) of Endoscopic Mayo score, Archetype Scores (S1, S2, S3), Archetype group membership (A1, A2, A3), and PC1, sorted by p-value.....	269

## LIST OF FIGURES

### CHAPTER 1

Figure 1.1 T cell signaling involved in transplant T cell-mediated rejection.....	28
Figure 1.2 Test set predictions when error is intentionally introduced through label flipping. ....	29
Figure 1.3 Various cross-validation methods used to assign machine learning scores to a sample population.....	30

### CHAPTER 2

### CHAPTER 3

Figure 3.1 Diagram showing the sampling strategy for the technical replicate pairs, biological replicate pairs, and cortex-medulla pairs. ....	73
Figure 3.2 Volcano plot of fold change between cortex and medulla versus negative log of adjusted p-value with false discovery rate. ....	74
Figure 3.3 Predicted proportion of cortex in a sample histologically called cortex or medulla. ....	75
Figure 3.4 Boxplot showing log of <i>NPHS2</i> expression in medulla and cortex samples as established by histology (A) and the predicted proportion cortex distribution across all samples (B). ....	76
Figure 3.5 Density plot of <i>NPHS2</i> expression in 1190 nonbisected biopsy cores.....	77
Figure 3.6 Scatter plots with predicted proportion cortex (y-axis) versus molecular ABMR, TCMR, and rejection scores in the biological replicate set, the cortex and medulla set, and the reference set (x-axis). ....	78
Figure 3.7 Investigating the relationship between injury (IRRAT score) and proportion of cortex. ....	79
Figure 3.8 Reproducibility plots of the molecular scores of rejection (TCMR, ABMR, all rejection) and acute kidney injury (IRRAT) in the technical and biological replicates and in the cortex–medulla divided pairs.....	80

### CHAPTER 4

Figure 4.1 Reproducibility of PC scores, archetype scores, and main classifier scores in MMDx technical replicate pairs. ....	107
Figure 4.2 Clear histology-MMDx discrepancy scenarios where MMDx would provide a useful second opinion.....	108
Figure 4.3 Visualizing discrepancies in molecular-based PCA as shown on the MMDx report. ....	109

### CHAPTER 5

Figure 5.1 The algorithm for assigning AutoBanff diagnoses using the eight canonical rejection lesion scores assigned by the center pathologists. ....	139
--	-----

Figure 5.2 Random forest analysis using the eight canonical rejection features to predict IH discrepancies. ....	140
Figure 5.3 Distribution of the AutoBanff versus ExpertBanff discrepancies in biopsies, based on their Molecular Microscope® (MMDx) phenotype.....	141
Figure 5.4 Venn diagram showing the overlap between MH discrepancies and IH discrepancies in the context of ABMR and TCMR.....	142

## CHAPTER 6

Figure 6.1 Venn Diagram of the Rejection-associated transcripts (RATs). ....	173
Figure 6.2 PCA of the rejection-associated transcripts (RATs) in kidneys, hearts and livers. ....	174
Figure 6.3 Scree plot from this archetypal analysis showing the residual sum of squares (y-axis) versus the potential number of archetypes in a model (x-axis). ....	175
Figure 6.4 Unsupervised Rejection-based analysis of 235 (218 patients) liver transplant biopsies.....	176
Figure 6.5 Moving average of Rejection (RAT4AA) archetypes (A) and PCA scores (B). ....	177
Figure 6.6 Establishing a balanced cutoff for the $R2_{TCMR}$ score in predicting histologic acute rejection (TCMR) with the highest possible combined sensitivity/specificity. ....	178
Figure 6.7 Example of the Molecular Microscope® Diagnostic System report on a new liver biopsy sample. ....	179

## CHAPTER 7

Figure 7.1 Factor maps showing the pathogenesis-based transcript sets (PBTs) used as initial input for the liver injury analyses and their relationship to principal components 1 (PC1), 2 (PC2), and 3 (PC3). ....	215
Figure 7.2 Unsupervised Injury-based analysis of 337 (311 patients) liver transplant biopsies.....	216
Figure 7.3 Moving average plots showing relationships between time post-transplant (TxBx), PBT scores, Injury PCA scores, and clinical laboratory test values. ....	217
Figure 7.4 AUCs for various molecular features and scores predicting binarized clinical steatohepatitis. ....	218
Figure 7.5 Distribution of biopsies called A-B) steatohepatitis by SOC, and C-D) steatohepatitis by the Msgt0 molecular steatohepatitis lda-based classifier. ....	219
Figure 7.6 Optimal cutoff established using the cutpointR function in R. ....	220

## CHAPTER 8

Figure 8.1 Molecular landscape of ulcerative colitis as shown by a volcano plot.....	249
Figure 8.2 Heatmap of expression of top 30 UC activity-associated transcripts in an existing panel of cell lines. CD4+ and CD8+ T cells and NK cells lacked expression of the UC activity-associated transcripts. ....	250

## CHAPTER 9

Figure 9.1 Archetypal analysis of biopsies taken from patients with diagnosed UC and varying levels of disease as represented by the endoscopic Mayo score.....	271
Figure 9.2 Changes in archetype composition over time in serial biopsies taken from 10 different UC patients with changing therapies.....	272

## ABBREVIATIONS

AA	Archetypal analysis
ABMR/AMR	Antibody-mediated rejection
ABMR-RATs	Antibody-mediated rejection transcripts
AKI	Acute kidney injury
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AMAT1	Alternative macrophage associated transcripts
ANOVA	Analysis of variance
APC	Antigen-presenting cell
AST	Aspartate aminotransferase
ATAGC	Alberta Transplant Applied Genomics Center
AUC	Area-under-the-curve
BAT	B cell associated transcripts
BD	Borderline
BK	Polyoma virus nephropathy
BP	Biological Process
CC	Cellular Compartment
CD55	Complement decay-accelerating factor
cDNA	Complementary deoxyribonucleic acid
CEGIIR	Center of Excellence for Gastrointestinal Inflammation and Immunity Research
CEL	File format of raw microarray file prior to preprocessing
CFB	Complement factor B
cIRIT	Cardiac injury and repair-induced transcripts
cRNA	Complementary ribonucleic acid
CT1	Colon transcripts – set 1
CT2	Colon transcripts – set 2

CTLA4	Cytotoxic T-lymphocyte-associated protein 4
CTLA4Ig	Cytotoxic T-lymphocyte-associated protein 4 immunoglobulin
CV	Cross-validation
DAMP	Damage-associated molecular pattern transcripts
DC	Dendritic cell
DSA	Donor-specific antibody
ECD	Extended criteria donor
eDSAST	Endothelial DSA-selective transcripts
EE	Everything else
eGFR	Estimated GFR
ENDATs	Endothelial cell associated transcripts
EpSCs	Epithelial stem cells
ESRD	End-stage renal disease
FDR	False discovery rate
FFPE	Formalin-fixed paraffin-embedded
GFR	Glomerular filtration rate
GI	Gastrointestinal
GO	Gene Ontology
GRITs	Gamma-interferon and rejection induced transcripts
HLA	Human leukocyte antigen
HUVECs	Human umbilical vein endothelial cells
IBD	Inflammatory bowel disease
IBDU	Inflammatory bowel disease unknown
ICC	Intraclass correlation coefficient
IH	Intra-histology
IHC	Immunohistochemistry
IF	Immunofluorescence
IFNG	Interferon gamma

Ig	Immunoglobulin
IGT	Immunoglobulin transcripts
IQR	Interquartile range
IRB	Institutional Review Board
IRRAT/IRRAT30	Injury-repair associated transcripts
IRITD1	Injury and rejection induced transcripts – Early
IRITD3	Injury and rejection induced transcripts – Intermediate
IRITD5	Injury and rejection induced transcripts – Late
KT1	Kidney transcripts – set 1
KT2	Kidney transcripts – set 2
LivGST_UP	Liver transcripts – set 1
MAP	Mitogen activated protein
MCAT	Mast cell associated transcripts
MF	Molecular Function
MH	Molecular-histology
MHC	Major histocompatibility complex
MMDx	Molecular Microscope Diagnostic System
MMF	Mycophenylate mofetil
mRNA	Messenger ribonucleic acid
Msgt0	Molecular steatohepatitis greater than 0 (Ica-based classifier)
NA	Not available
NAFLD	Non-alcoholic fatty liver disease
NASH	Non-alcoholic steatohepatitis
NK	Natural killer
NKB	Natural killer cell transcript burden
NPHS2	Podocin
NR	No rejection
pABMR	Possible antibody-mediated rejection

PBTs	Pathogenesis-based transcript sets
PC1	Principal component 1
PC2	Principal component 2
PC3	Principal component 3
PCA	Principal component analysis
PTC	Peritubular capillaritis
pTCMR	Possible T cell-mediated rejection
QCAT	Quantitative cytotoxic T lymphocyte-associated transcripts
QCMAT	Quantitative constitutive macrophage-associated transcripts
RAI	Rejection activity index
RATs	Rejection-associated transcripts
Rej-RATs	All-rejection-associated transcripts
RIN	RNA Integrity Number
RMA	Robust Multiarray Averaging
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RPTECs	Renal proximal tubule epithelial cells
RTW	Response to wounding
SCC	Spearman correlation coefficient
SOC	Standard-of-care
TCMR	T cell-mediated rejection
TCMR-RATs	Transcripts associated with T cell-mediated rejection
TCR	T cell receptor
TG	Transplant glomerulopathy
TLR	Toll-like receptor
TMA	Thrombotic microangiopathy
TNF	Tumor necrosis factor
TRAC	T cell receptor alpha

TRBC	T cell receptor beta
TRDC	T cell receptor delta
TRGC	T cell receptor gamma
TxBx	Time post-transplant
UC	Ulcerative colitis

**CHAPTER 1**  
**INTRODUCTION**

### 1.1 THE UNMET DIAGNOSTIC NEEDS IN TRANSPLANTS AND NATIVE ORGANS

In many ways, biopsy-based diagnostic practices have remained unchanged over the past several decades. It is widely acknowledged that subjective diagnostics based on a single opinion, visual assessment, and translation to a final diagnosis using arbitrary guidelines chosen by consensus have serious drawbacks, and present opportunities for improvement.<sup>1-5</sup> Histologic diagnoses of disease based on visual assessment of tissue biopsies has inherent shortcomings and can prevent patients from being properly diagnosed and thus appropriately treated. These high levels of interobserver disagreement were addressed by Furness et al<sup>2</sup>: "...international variation in histologic grading is large, under-recognized, difficult to improve, and almost certainly of major clinical relevance. Urgent steps are needed to improve this area of practice". Histology has other inherent limitations such as the inability to assess recent injury,<sup>6-8</sup> and also shares with all diagnostic systems the challenge of balancing the risks of under- versus over-diagnosis in low grade or ambiguous cases by histology.<sup>6,9</sup> An additional concern with these diagnostic methods is "data drift": rules originally formulated in past years or decades are now being applied in a different case mix that changes prior probabilities.<sup>6,10</sup> Providing diagnostic options that are reproducible, objective, accurate, precise, and probabilistic would allow clinicians to place their patients along disease continuums, rather than 'binning' patients which introduces potential diagnostic errors.

Histologic systems are the SOC for the diagnosis of rejection and disease in transplanted organs, but visual assessments are also used in diagnostics for native organs e.g. diagnosing inflammatory bowel disease (IBD) in the colon. These problems are present in all areas of medicine where histology is used, including transplanted organs and in native organ diseases that rely on scoping technology for clinical decision-making. Improved diagnostic systems would avoid subjective or arbitrary guidelines, express the findings as continuous data/probabilities, and allow for higher inter-observer agreement between experts by incorporating ensemble-based diagnoses assigned by machine learning algorithms.<sup>6</sup>

### 1.2 BACKGROUND

These analyses focused on thoroughly understanding the precision and stability of the established Molecular Microscope Diagnostic system developed for kidney transplants (MMDx-Kidney),

## **Chapter 1: Introduction**

translating the system for liver transplants where the SOC data is less reliable, and finally moving MMDx analyses into a native organ disease – ulcerative colitis (UC) – to examine the potential applications beyond transplantation.

### **1.2.1 A review of kidney transplantation**

#### **1.2.1.1 The kidney and its function**

The kidney is a relatively small organ (mean length 11 cm, mean weight 40-175 g) that functions to filter the blood supply, maintain blood fluid and salt homeostasis, and regulate blood pressure.<sup>11</sup> Blood enters the kidney via the renal artery, which branches into progressively smaller vessels as it moves into the kidney tissue until eventually reaching the nephrons via capillaries.<sup>12</sup> Typical human kidney parenchymal tissue is composed of 800,000-1,000,000 nephrons. Each nephron consists of a glomerulus (contained in the renal corpuscle), tubule (composed of specialized epithelial cells capable of reabsorption and secretion), renal artery capillary, and renal vein capillary. The endothelial wall of the glomerulus allows very small molecules (e.g. waste, fluid) to move into the proximal convoluted tubule, while larger molecules (e.g. proteins, blood cells) remain in the blood vessel.<sup>12</sup> This passive filtration is mediated by podocytes surrounding the glomerulus with their foot processes (podocytes will be referred to further in Chapter 3 as a marker for kidney cortex tissue). Fluid in the tubule, along with some minerals and nutrients, is reabsorbed into the adjoining peritubular capillaries.<sup>12</sup> Remaining filtrate enters the loop of Henle to moderate salt concentrations in the interstitium surrounding the tubule. In the distal tubule, the filtrate is further modified using active transport by either absorbing or secreting sodium, potassium, calcium, phosphate, hydrogen, and ammonium. The final filtrate exits the distal tubule and moves to the collecting ducts where water molecules can be reabsorbed. The collecting ducts move the final waste out of the nephron to the bladder via the ureter for excretion, while filtered blood returns to the rest of the body through the renal vein.<sup>12</sup> In this way, the circulating blood is filtered to remove waste products, but blood volume and salt homeostasis is maintained. Filtration in the nephrons is performed by many transporters and channels, many of which are classified as solute carriers (SLCs).<sup>13</sup> Transcripts for SLCs

## **Chapter 1: Introduction**

will be referred to in Chapters 6 and 7, and the loss of these transcripts typically signals a loss of parenchymal function, in kidneys but also in other parenchymal tissue.

In addition to its roles in filtration, the kidney is also responsible for generating several hormones and enzymes (e.g. renin), while acting as a target for hormones produced by other organs (e.g. angiotensinogen produced by the liver).<sup>11</sup> One of the most critical endocrine systems is the renin-angiotensin-aldosterone system (RAAS), which regulates blood pressure, vascular resistance, and fluid/electrolyte balance.<sup>14</sup> Many kidney diseases that lead to end-stage renal disease (ESRD) and kidney loss are endocrine-related (e.g. diabetes, where the high blood sugar levels damage renal capillaries leading to proteinuria, buildup of renal wastes, and renal failure requiring dialysis or kidney transplant).<sup>15</sup> Deterioration in function is affected by the number of functioning neurons, and can be measured by changes in estimated GFR (eGFR, see 1.2.1.5). Significant changes in eGFR from the expected baseline values for that patient will often prompt a biopsy, as these symptoms can be representative of a loss of parenchymal integrity and normal tissue function.

### **1.2.1.2 Renal transplant rejection**

If a native kidney reaches ESRD, the patient will require dialysis until a transplanted kidney becomes available from a live donor (LD) or a deceased donor (DD). Potential transplants are carefully matched based on blood type and measurable donor-specific antibody data (DSA, describing the patient's preformed antibodies against a panel of selected human leukocyte antigens (HLAs) that represent the potential donor population).<sup>16</sup> HLA matching is important for avoiding transplant rejection. Renal transplant rejection has two forms: antibody-mediated rejection (ABMR), and T cell-mediated rejection (TCMR). The kidney can be biopsied for a number of indications related to dysfunction or proteinuria.

#### **1.2.1.2.1 Renal transplant ABMR**

## Chapter 1: Introduction

In ABMR, antigen-presenting cells (APCs) from the graft migrate to secondary lymphoid organs where they display the antigens to naïve T cells.<sup>17</sup> The activated T cells include effector T cells and helper T cells. Follicular helper T cells plus donor alloantigens trigger B cell activation and the subsequent development of plasma cells. These plasma cells migrate to the spleen, and eventually to the bone marrow where they become long-lived and produce antibody against the donor antigens. DSAs are termed 'de novo' if generated post-transplant and 'preformed' if they exist pre-transplant as a result of sensitization.<sup>17</sup> Antibodies produced by the plasma cells move through the circulatory system and eventually infiltrate the donor tissue of the graft, where they can damage the tissue, causing inflammation, loss of function, proteinuria, and eventually graft failure.<sup>17</sup> Tissue damage in the microcirculation of the allograft can occur as a result of complement activation via the classical pathway (where antibodies bound to antigens are bound by C1q, eventually leading to the formation of the cylindrical membrane attack complex and cell lysis),<sup>18</sup> natural killer (NK) cell activity triggered by Fc receptor recognition of the Fc portions of the alloantibodies bound to graft endothelium,<sup>19,20</sup> or potentially other mechanisms.<sup>21</sup>

### 1.2.1.2.2 Renal transplant TCMR

In TCMR, antigens are presented on APCs to the naïve T cells in lymphoid organs, causing T cell activation and clonal expansion of effector T cells against the donor tissue. Effector T cells, activated macrophages, B cells, and plasma cells infiltrate the donor tissue, causing an intense inflammatory immune response against the graft.<sup>17</sup> This response is characterized by interstitial inflammation, interferon gamma (IFNG) effects, tubulitis, an increase in chemokine expression, changes in the permeability of the capillaries and extracellular matrix, and deterioration of parenchymal function.<sup>17,22</sup> This T cell-mediated inflammation is essential under normal circumstances to purge infecting agents and neoplastic cells; but must be suppressed in transplanted patients to prevent the immune system from responding to the foreign graft tissue as a potential threat. TCMR disease activity is primarily mediated through inflammatory mechanisms (e.g. products of activated T cells and macrophages).<sup>17,23</sup> Effector T cells have cytotoxic molecules in their granules (*GNLY*, *PRF1*, *GZMA*, *CD95L* etc.)<sup>24</sup> but experiments in gene-deficient mice have not shown that these are required for TCMR.<sup>25</sup>

## Chapter 1: Introduction

Several T cell-signaling pathways are involved in the development and maintenance of TCMR. APCs process antigens detected in their environment, producing peptides characteristic of those antigens. These peptides occupy the groove of a major histocompatibility complex (MHC) in the endoplasmic reticulum of the APC. The MHC protein is then transported to the cellular surface where it is displayed to T cells.<sup>17</sup> The antigen-MHC complex interacts with the T cell receptor (TCR), producing signal 1.<sup>17</sup> A costimulatory signal, 'signal 2', is required for naive T cell activation, although memory T cells have a lesser requirement for signal 2. CD28 on the T cell surface interacts with CD80/86 molecules on the APC, triggering costimulatory signals.<sup>17</sup> T cell signaling from signal 1 and signal 2 can transmit via the calcium-calcineurin, RAS-mitogen-activated protein (MAP) kinase, or nuclear factor- $\kappa$ B (nf-  $\kappa$ B) pathways (**Figure 1.1**).<sup>17</sup> Once these pathways are activated, they trigger messenger ribonucleic acid (mRNA) transcription for cytokines (e.g. IL-2, IL-15).<sup>17</sup> Engagement of IL-2 and other cytokines with their receptors constitutes signal 3, which triggers clonal expansion via the mTOR and other pathways.<sup>17</sup>

In gene expression experiments using microarrays, the molecular landscape of TCMR highlighted transcripts associated with immune checkpoints on T cells e.g. *CTLA4*, *CD80*, *CD28*; chemokine signaling from macrophages and dendritic cells e.g. *ADAMDEC1* and *CXCL13*; immunoglobulin superfamily members e.g. *SIRPG*; cytokines involved in innate or adaptive immune signaling e.g. *IFNG*; inflammasome activation e.g. *AIM2*; and T cell activation e.g. *BTLA*.<sup>23</sup>

Immunosuppression is essential to prevent T cell activation and thus transplant rejection. Patients can be given initial induction therapy prior to the transplant, and then are managed post-transplant with combinations of immunosuppressive agents (e.g. mycophenylate mofetil 'MMF', tacrolimus or cyclosporine, azathioprine, basilixumab, and anti-thymocyte globulin 'ATG').<sup>17</sup> These agents also check the emergence of de novo DSA and ABMR. While these regimens are not completely protective against ABMR, they prevent most cases of TCMR if properly adhered to by the patient and given at an appropriate dose. TCMR often occurs in instances where patients have been under-immunosuppressed (e.g. because of virus infections) or were non-adherent with their medications.<sup>26</sup>

## **Chapter 1: Introduction**

### **1.2.1.3 Renal transplant injury**

While transplants can be damaged by a rejection episode, there are other sources of injury. Damage can occur as a result of donation-implantation (ischemia-reperfusion type injury). The injury is greater if the graft is taken from a DD or extended criteria donor (ECD) than from an LD. Grafts can be exposed to infection (e.g. polyoma virus infection, pyelonephritis), which can cause inflammation and tissue necrosis. In transplant injury, damage to parenchymal cells of the renal epithelium and endothelium induce a response to wounding (RTW) in the tissue with dedifferentiation, loss-of-function, loss of solute carriers, as well as activating innate immunity. Eventually, atrophy-fibrosis will follow if the nephrons do not recover.

The process of transplantation can result in ischemia-reperfusion injury. Ischemia begins at brain death, where increased intracranial pressure, decreased cardiac output, tachycardia and increased blood pressure as a result of the Cushing reflex, and systemic vascular resistance eventually culminate in hypotension and cardiac arrest.<sup>27</sup> This physiological process leads to the release of free-radicals and proinflammatory cytokines, activating the innate immune system and potentially triggering early adaptive immune damage.<sup>27</sup> Extensive cold-storage of the donated organ (cold-ischemic time, or total time of the donated organ on ice) prior to implantation further contributes to tissue damage.<sup>27</sup> The effector phase of ischemic injury is termed 'reperfusion injury', and involves cellular regeneration and repair, apoptosis, autophagy, and necrosis of the tissue.<sup>27</sup> It is critical to manage ischemia-reperfusion and reduce cold-ischemic times to minimize tissue injury during implantation and reduce the incidence of acute kidney injury (AKI) post-transplant. Ischemia-reperfusion injury and associated AKI have been linked to both short and long-term functionality of the graft, making this stage of transplantation critical for the survival of the graft.<sup>27-31</sup> Improving outcomes for transplant patients remains an area of high interest, thus the quantification and accurate description of injury and RTW must persist as an important target for new diagnostic tools.

### **1.2.1.4 The Banff guidelines for kidneys**

## Chapter 1: Introduction

Management of renal transplants requires a classification system to describe disease and injury in the tissue. Histology in kidney transplants follows the Banff guidelines, which are referred to in supervised analyses and for discrepancy analyses in Section I of this thesis. This system uses a series of categories and scores to assign diagnoses regarding disease or transplant rejection, and a set of guidelines to translate the lesion scores and other features (i.e. DSA, C4d) into a diagnosis. These rules are reviewed every two years, but the changes implemented as the population changes over time are minimal with virtually no change in the rules regarding the major lesions and features that define the two major forms of transplant rejection: ABMR and TCMR.<sup>22,32,33</sup>

Unfortunately, many of the features and lesions used in the Banff guidelines are non-specific, and do not necessarily relate to one disease or type of rejection versus another.<sup>34-40</sup> This process of choosing non-specific features that relate to a latent disease phenotype by consensus can lead to incorrect diagnoses, over- or under-treatment, and missed opportunities for improving patient health and well-being. Although TCMR has always been associated with t- and i-lesions, these lesions are non-specific and associated with other changes.<sup>36-38</sup> V-lesions are inherently problematic in the current Banff guidelines: v>0 always leads to a diagnosis of TCMR even though v-lesions have been shown to be associated with ABMR and other inflammatory conditions.<sup>26,34,41</sup> The relationship between DSA (DSA) and ABMR has been strictly regimented throughout all versions of the Banff guidelines and yet the number of cases with symptoms and features consistent without detectable DSA has been noted and expressed in several publications.<sup>39,40,42</sup> This lack of selectivity and specificity in diagnoses, extensive variation in lesion and feature assignment between experts, and variation in the assignment of diagnoses present serious problems for the clinicians.<sup>2,3</sup>

### 1.2.1.5 Clinical correlates of kidney function

Kidney function can be assessed by several measurements. The most common clinical measures of function are the serum creatinine levels and glomerular filtration rate (GFR). These values are affected by episodes of rejection and/or injury.

## Chapter 1: Introduction

Serum creatinine levels in the patient, as measured from a blood test (serum), are typically constant if there is stable renal function.<sup>43</sup> Degradation of creatine and phosphocreatine in skeletal muscle is the main source for creatinine in the bloodstream, where it travels to the kidney and is filtered by glomeruli.<sup>43</sup> Creatinine removed by glomeruli is mostly excreted into the urine as waste, although up to 15% of creatinine is secreted by the renal tubules under normal conditions, and can also be increasingly excreted by the gastrointestinal (GI) tract in cases of renal failure or ESRD.<sup>43</sup> Most modern methods of clinical creatinine measurement are derived from the Jaffe method described in 1886, and currently is most often measured by enzymatic or Jaffe-based methods.<sup>44</sup> Both methods are based on colorimetry, which introduces a base level of variability to the serum creatinine results.<sup>45</sup>

GFR is a main feature used in the diagnosis of chronic kidney diseases. It is based on an algorithm which assesses the patient's serum creatinine levels versus creatinine excreted in the urine. There are many alternate calculations for estimated GFR (eGFR), e.g. the Cockcroft and Gault formula:

$$\text{eGFR (mL/min)} = [(140 - \text{age}) \times \text{Wt} / (0.814 \times \text{S.Cr in } \mu\text{mol/L})] \times (0.85 \text{ if female})^{46}$$

Many GFR equations correct for patient age, weight, and gender, but not for donor age or gender (which can affect the size of the transplanted kidney). The problem is kidney size and age: for example, eGFR will be influenced by the transplantation of a size-mismatched organ between the donor and recipient. If the donor's body mass is much smaller than the recipient, the kidney will likely be smaller and have a reduced filtration capacity even if the function is completely normal with healthy parenchyma.

### 1.2.1.6 Molecular studies of kidney rejection and injury to date

To date, molecular studies have focused on gene expression correlates for immunological mechanisms in ABMR, TCMR, and RTW. Very early studies of renal transplant biopsies using microarrays identified subtypes of rejection in the population<sup>47</sup> or stratified transplants by incurred allograft damage,<sup>48</sup> but failed to include ABMR in the population and were never validated, limiting applicability in clinical practice. Another study found that otherwise relatively healthy patients with biopsy results consistent with fibrosis and inflammation expressed transcripts shared by cognate and innate immune

## Chapter 1: Introduction

pathways.<sup>49</sup> Tubulointerstitial damage was shown to induce detectable molecular changes in a population of protocol biopsies, although this excluded biopsies with ABMR or TCMR (classified as indication biopsies).<sup>50</sup> Integrins, including ITGB6, correlated with progression to ESRD in several studies, showing that gene expression can be used in renal transplant prognosis.<sup>51,52</sup> Clinical trials have been done to assess molecular correlates for progressive fibrotic injury and risk of allograft loss,<sup>7</sup> while other studies have examined the progression from acute to chronic renal injury as related to ischemia-reperfusion.<sup>53</sup>

More recent studies have focused on gene expression related to diagnoses of rejection, using biopsies, blood, or urine samples.<sup>6,35,36,54–66</sup> These studies have found that gene expression is highly conserved in transplant rejection,<sup>6,35,36,57,65,66</sup> can distinguish between TCMR and ABMR,<sup>35,66</sup> can determine disease stage and severity,<sup>35,66,67</sup> is highly reproducible and stable compared to its histologic counterpart,<sup>6,10</sup> and can increase the precision and accuracy in diagnoses.<sup>6,10</sup> Diagnostic tests developed for blood and urine are only suitable for screening as they distinguish sick versus well without specificity for disease when the patient is sick. Biopsy-based tissue gene expression provides disease specificity (e.g. distinguishing TCMR from ABMR), plus increased reliability because the gene expression is diffuse and stable. Machine learning algorithms and clustering analyses used to process the resulting data from microarray-based gene expression studies can overcome a level of unreliable sample labeling in the clinical phenotype used to train in supervised analyses, and unsupervised analyses can process the data without the need for a clinical phenotype – allowing the molecular data to stand alone.<sup>6</sup>

### 1.2.1.7 Tissue heterogeneity and gene expression in kidneys

The Banff guidelines are unable to assess kidney biopsy tissue with no glomeruli present (medullary tissue).<sup>22</sup> This challenge results in some biopsies being insufficient for histology grading, and the inability of the pathologist to provide an accurate or reliable diagnosis.<sup>64</sup>

Molecular-based diagnostic systems (such as those that rely on the expression of mRNA extracted from tissue) do not rely on the presence of glomeruli in a biopsy in order to make an assessment. Since gene expression is diffuse across the tissue, rejection and injury should be

## **Chapter 1: Introduction**

measurable regardless of tissue heterogeneity. This was assessed in a series of experiments that will be discussed specifically in later sections (see Chapter 3).

### **1.2.1.8 Current limitations of clinical and histologic practice**

In kidney transplants requiring biopsies, precision (reproducibility), accuracy (representation of the true disease states), and standardization are critical for diagnosis and management.<sup>6,68-70</sup> Categorical diagnostic systems like the Banff guidelines lack the required granularity to express all biopsy heterogeneity, and can also carry large inter-observer variation (as represented by low to moderate kappa values),<sup>2,3</sup> observer bias, and can lead to incorrect diagnoses as a result of non-specific lesions assigned to specific diagnostic categories.<sup>35-40,71-74</sup> Many of these limitations can be addressed using molecular diagnostic systems like those discussed throughout this thesis.

## **1.2.2 A review of liver transplantation**

### **1.2.2.1 The liver and its function**

The liver is a large (average weight 1.2-1.5 kg) organ situated in the upper right quadrant of the abdomen. It is composed of a fibrous capsule covering the organ, and two major lobes: the large right lobe and smaller left lobe divided by ligaments; and two minor lobes: the quadrate lobe and caudate lobe. The liver receives blood from the hepatic artery and the portal vein. Each lobe is comprised of many 'hepatic lobules' - functional units throughout the tissue which are made up of multiple hepatic cells separated by vascular channels (hepatic sinusoids) surrounding a central vein.<sup>75</sup> Blood circulates from the gastrointestinal tract into the liver through the hepatic portal vein, where nutrients and oxygen are transferred to the hepatic cells and sinusoids.<sup>75</sup> Specialized phagocytic macrophages called Kupffer cells reside in the sinusoid endothelium and remove foreign material and bacteria from the blood in the portal vein.<sup>75</sup> Blood is transferred from the sinusoids to the central vein of the hepatic lobule, eventually exiting through the hepatic vein.<sup>75</sup>

## **Chapter 1: Introduction**

The liver is also responsible for creating bile for the gastrointestinal tract. Hepatic lobules contain bile canaliculi; small vessels that carry secretions from the hepatic cells to the bile ductules.<sup>75</sup> These ductules converge forming the larger hepatic ducts, and finally to form the common hepatic duct.<sup>75</sup> Bile moves from the common hepatic duct to the cystic duct, is stored in the gallbladder, and finally secreted to aid digestions.<sup>75</sup>

The liver assists with carbohydrate metabolism and regulation of blood glucose levels by polymerizing glucose to glycogen.<sup>75</sup> The liver contributes to lipid metabolism by oxidizing fatty acids to synthesize lipoproteins, phospholipids, and cholesterol.<sup>75</sup> Excess carbohydrate molecules are stored as fats, transported to adipose tissue for storage.<sup>75</sup> The liver also deaminates amino acids to produce urea, converts amino acids to other amino acids, and synthesizes plasma proteins; all contributing to protein metabolism.<sup>75</sup> Glycogen, iron, and Vitamin A, Vitamin D, and Vitamin B12 are stored in liver tissue.<sup>75</sup> Damaged red blood cells are removed from the bloodstream by the liver macrophages, and toxins (i.e. alcohol, some drugs) are filtered out of the bloodstream.<sup>75</sup>

### **1.2.2.2 Liver transplantation**

Liver transplantation is a life-saving intervention for patients with irreversible liver disease.<sup>76</sup> The number of liver transplants per year has been steadily increasing, especially in geographic areas where cirrhosis and obesity-related non-alcoholic fatty liver disease (NAFLD) are common.<sup>77</sup> In 2018, 7526 liver transplants were done in the United States, making liver the second most commonly transplanted organ behind kidneys.<sup>76</sup> Even as the number of transplanted livers increases, the waitlist continues to grow, and in 2013 alone 1767 patients died while waiting for a transplant while another 1223 patients were removed from the list because they had become too sick to undergo a transplant operation.<sup>76</sup> This intense and growing demand for liver transplants makes appropriate management of existing liver transplants crucial for maintaining the health and function of the graft and for preserving patient survival and well-being. The putative tolerogenic properties in the liver suggest that the immunosuppression crucial for the long-term survival of many other grafts (i.e. kidney, heart) may be less critical in the management of liver transplants. Because the liver allograft has the potential to be long-lived (up to 8-9 decades in the case of

## **Chapter 1: Introduction**

many pediatric liver transplant patients<sup>78</sup>), reduction or removal of immunosuppression has been a focus of many studies. The long-term safety of these practices remains unknown, as liver transplant rejection is still an inconsistent and challenging diagnosis, reducing the reliability of clinical endpoints that involve this phenotype.

### **1.2.2.3 Transplant rejection in liver**

Transplant rejection in livers is a complex and unreliable diagnosis and focuses mainly on acute cellular rejection (TCMR) of the allograft. Although chronic ABMR has been described in recent versions of the Banff guidelines for liver histology,<sup>79</sup> there is no consensus for the presence and prevalence of ABMR in the liver population, or to support the use of the described lesions and features in the diagnosis of ABMR. TCMR follows a similar signaling pathway to that in kidney transplants (see section 1.2.1.2). The presence or absence of ABMR in the liver population is discussed in the following section (1.2.2.3).

Liver TCMR can develop early (acute cellular rejection, <6 months post-transplant) or late (chronic cellular rejection, >6 months post-transplant), with the suggestion that the former is related to direct alloantigen presentation while the latter is dependent on indirect alloantigen presentation.<sup>79</sup> Acute TCMR is characterized by more inflammatory bile duct damage; portal inflammation by lymphocytes, macrophages, and eosinophils; and comparatively few necro-inflammatory-type interface activity.<sup>79-81</sup> Alternatively, chronic or 'late' TCMR is characterized by a decrease in lymphocytic cholangitis, low-grade interface and perivenular necro-inflammatory activity, and increasingly homogenous lymphoplasmacytic and histiocytic infiltrates.<sup>79-81</sup> Banff guidelines for liver TCMR will be referred to further in Chapter 6, where they were used in a supervised analysis of rejection.

### **1.2.2.4 ABMR in liver transplants**

The presence of ABMR in the liver transplant population remains highly controversial. It has been suggested that ABMR is not present in liver transplants for a number of reasons: dilution of antibody binding with respect to a large endothelial surface due to overall organ size; Kupffer cell clearance of

## Chapter 1: Introduction

activated complement, platelet aggregates, immune complexes, and DSA; variation in the hepatic microvascular class II DSA producing fewer DSA class II targets; expression of Fc receptors and phagocytic activity on Kupffer and liver sinusoidal endothelial cells; and hepatic regenerative abilities independent of the development of fibrosis or with reverse fibrosis.<sup>79</sup> Because the studies undertaken in this thesis did not find evidence for ABMR, and because DSA and C4d measurements are still not SOC for liver transplantation, ABMR in livers is not a focus of this thesis. Further information can be found in the literature regarding the many perspectives on ABMR in liver allografts.<sup>79,82-84</sup>

### 1.2.2.5 Transplant injury in livers

While ischemia-reperfusion injury in transplanted livers is similar to that in other transplanted organs, the liver is the only transplanted organ that can be implanted as a portion of a full organ which then regenerates into a full liver in the recipient. In typical live donation, the donor undergoes right hepatectomy, and the tissue is implanted in the recipient. After three months, reconstitution of the donor tissue is approximately 80%  $\pm$ 13%, while reconstitution of the recipient tissue is approximately 93%  $\pm$ 18%.<sup>85</sup> Long-term post-transplant complications for the donor and recipient are not well-understood.<sup>76,86</sup> The initial insult during transplantation (including the partial hepatectomy) injures the tissue, as does the cold-ischemic time pre-transplant and reperfusion post-transplant. Injury can occur from inflammation triggered by the release of reactive oxygen species (ROS) post-anastomoses.<sup>28,87</sup> It has been suggested that a neutrophilic inflammatory response is mediated by Kupffer cells and T cells, causing neutrophil infiltration to the post-transplant liver tissue.<sup>87</sup> Pro-inflammatory responses in ischemia-reperfusion may be signaled by Toll-like receptors (TLR) on Kupffer cells.<sup>87</sup> The liver graft is exposed to more damaging processes during live donation versus deceased donation; all live donation liver transplants (LDLT) are partial grafts dissected while circulation is maintained, while few deceased donor liver transplants (DDLT) are transplanted as a split graft.<sup>88</sup> Despite this difference in implantation injury, ten-year outcomes between recipients did not significantly differ between LDLT and DDLT.<sup>88</sup>

Injury can also be caused by disease processes occurring after liver transplantation, such as steatohepatitis, cirrhosis, fibrosis, and infection.<sup>89</sup> Steatohepatitis in particular was considered in the

## Chapter 1: Introduction

following analyses (see Chapter 7) as a form of injury that could be assessed by gene expression and machine learning, as prior gene expression studies were encouraging.<sup>90</sup> Steatohepatitis involves inflammation, tissue damage, and fat deposition, but these non-specific features can be challenging to assess reproducibly by the SOC.<sup>91,92</sup> Steatohepatitis is an increasingly reported indication for liver transplant, and has been associated with poor waitlist survival.<sup>93</sup> Post-transplant injury assessment remains an important area of focus for those managing long-term liver transplants, and will be discussed in greater detail in later sections of this thesis.

While all transplanted organs incur injury during the implantation process, the liver transplant process is unique in all LD and some DD in that a portion of the liver is donated. The well-documented unique regenerative properties of the liver make this possible, and also result in a different response-to-wounding than what is seen in other organ transplants (i.e. kidney, heart, and lung).<sup>85,94–97</sup> This response and its associated gene expression will be further explored in this thesis.

### 1.2.2.6 Clinical correlates of liver function

Liver function can be described through routine biochemistry results, mainly the measured levels of serum albumin, bilirubin, alkaline phosphatase (ALP), alanine aminotransferase (ALT), and aspartate aminotransferase (AST).<sup>98,99</sup> These values are affected by episodes of rejection and/or injury.

Because albumin is synthesized by the liver, low levels (<35-50 g/L) can indicate dysfunction of the liver synthetic ability.<sup>98</sup> The half-life of albumin (20-30 days) makes it better suited for assessing chronic conditions, as albumin levels may remain misleadingly high during periods of acute disease.<sup>100</sup> Bilirubin is generated mainly during hemoglobin metabolism when red blood cells are destroyed (80%), with a small contribution from other heme proteins (20%).<sup>98</sup> It is possible for serum bilirubin levels to remain normal (2-18  $\mu\text{mol/L}$ ) in some cases of liver injury, although high levels coupled with increased ALT and AST can indicate hepatitis or liver cirrhosis, or cholestatic disorders when coupled with high ALP.<sup>98</sup> ALP levels can be increased from normal range (20-130 units/L depending on the measuring assay) by the accumulation of bile or decreased with hepatocellular injury.<sup>98</sup> ALT is related to

## Chapter 1: Introduction

hepatocellular injury; levels in the higher-than-normal range (7-53 units/L depending on the assay) can signal alcoholic hepatitis, or in extreme cases (e.g. 1000 units/L) viral hepatitis, ischemic hepatitis, or toxic liver injury.<sup>98</sup> Finally, AST (another test for hepatocellular injury, normal range 11-47 units/L depending on the assay) can indicate liver damage if ALT, bilirubin, and ALP are also increased, and very high levels (e.g. 1000 units/L) can indicate viral hepatitis, ischemic hepatitis, or toxic liver injury.<sup>98</sup> Because AST is less specific than other biochemical measurements, elevated ALP in isolation from other biochemical scores may be unrelated to liver function and may instead reflect cardiac or muscle disease.<sup>98</sup>

### 1.2.2.7 Liver injury and non-alcoholic fatty liver disease (NAFLD)

NAFLD is the most common cause of abnormal liver function tests<sup>101</sup> (estimated prevalence 20% in the overall population and up to 70% in populations with diagnosed diabetes).<sup>102</sup> NAFLD is asymptomatic, thus a diagnosis usually results from an incidental abnormal liver biochemistry tests or by steatosis detected by imaging.<sup>103</sup> Typical NAFLD-associated biochemistry would include mildly elevated ALT, although ~80% of patients have ALT levels in the normal range (<40 IU/L in males and <31 IU/L in females), and ALT levels fall as fibrosis progresses towards cirrhosis.<sup>103</sup> Criteria for the diagnosis of NAFLD include an ALT level higher than that of the AST result, using an AST:ALT ratio of <0.8 (or >0.8 with advanced disease).<sup>104</sup> However, studies have found that ALT values do not correlate with liver histology and are not useful in diagnosing NAFLD or in determining disease severity.<sup>91,105</sup> Despite these shortcomings, clinicians still rely on liver biochemistry to identify NAFLD in patients, leading to potentially incorrect diagnoses of their patients and missed opportunities for therapeutic intervention.<sup>103</sup>

Non-alcoholic steatohepatitis (NASH, or simply steatohepatitis), a severe form of NAFLD characterized by hepatocyte damage and inflammation,<sup>93</sup> occurs in 3-5% of the general population.<sup>103</sup> As a primary disease, steatohepatitis is associated with poor waitlist survival.<sup>93</sup> Growing waitlist demands have promoted interest in ECD livers, but grafts with >60% steatosis will not be considered for transplantation while those with 30-60% steatosis are associated with poor post-transplant outcomes.<sup>106</sup> Fat pattern in the donated organ is another concern, as organs with macrovesicular fat are less tolerant of implantation injury than those with microvesicular fat.<sup>107</sup> Liver transplantation for steatohepatitis-

## Chapter 1: Introduction

associated liver failure resolves the issue of the associated cirrhosis and fibrosis, but cannot solve the underlying metabolic dysregulation or comorbidities therefore post-transplant recurrence is a concern.<sup>93</sup> There is no therapeutic intervention proven to treat steatohepatitis; although weight loss of >10% can reverse up to 90% steatosis and 45% fibrosis.<sup>108</sup> Immunosuppressive medications can lead to or worsen risk factors for post-transplant NAFLD and steatohepatitis, resulting in a recommendation for reduced immunosuppression.<sup>93,109</sup> Since a diagnosis of steatohepatitis has been associated with increased liver-related and cardiovascular mortality, diagnostic accuracy is important for appropriate patient management and predicting patient outcomes.<sup>103</sup> Steatohepatitis is discussed further in Chapter 7.

### 1.2.2.8 The Banff guidelines for liver transplants

Transplantation diagnoses in liver can follow the Banff guidelines for liver transplants (similar to its counterpart in kidney transplantation) but this is not universal among international centers. Pathologists vary in their use of the Banff guidelines. The lack of international consensus on the essential components for diagnosing acute rejection, chronic rejection, injury, or infection lead to high levels of variability among liver pathologists.<sup>4,110</sup> Kappa values for pathology related to TCMR range from low to moderate (0.15-0.62),<sup>111</sup> and are affected by the presence of multiple diseases or infections.<sup>110</sup>

The Banff guidelines for liver transplants describe criteria for acute and chronic TCMR, plasma cell-rich rejection, and acute and chronic ABMR ('AMR', although there is lengthy discussion within the Banff report regarding the controversy surrounding this diagnosis).<sup>79</sup> ABMR diagnostic criteria focus on both ordinal scored lesions e.g. portal inflammation scores in the criteria for TCMR, but also include descriptions of a visual assessment that do not include an ordinal numerical score e.g. 'portal and/or perivenular plasma cell-rich infiltrates with easily recognizable periportal/interface and/or perivenular necro-inflammatory activity usually involving a majority of portal tracts and/or central veins'<sup>79</sup> as a requirement for the diagnosis of plasma cell-rich rejection). This mixture of grades plus text descriptions make diagnostic reproducibility a significant challenge.

## **Chapter 1: Introduction**

Typical TCMR is described by the Banff guidelines as indeterminate, mild, moderate, or severe by its rejection activity index (RAI) score.<sup>79,112</sup> The RAI score for TCMR (referenced further in Chapter 6) is based on a sum of grades describing the extent of portal vein inflammation (0-3, describing lymphocytic, neutrophilic, and eosinophilic infiltrate causing inflammation involving and eventually expanding the hepatic triads and periportal parenchyma), bile duct inflammation damage (0-3, describing the level of infiltration and degenerative changes i.e. nuclear polymorphism, disordered polarity, cytoplasmic vacuolization and eventually focal luminal disruption of the bile duct epithelial tissue), and venous endothelial damage (0-3, describing the subendothelial lymphocytic infiltration involving the portal and/or hepatic venules and the extent of perivenular inflammation and necrosis).<sup>79,112</sup> Like ABMR guidelines, the TCMR guidelines incorporate a mixture of ordinal scoring plus text descriptions e.g. the global assessment criteria, contributing to reduced reproducibility.

### **1.2.2.9 Molecular studies to date**

To date, gene expression and molecular studies in liver have been promising. Biopsies taken from a seemingly homogenous population of stable, long-term pediatric liver transplant recipients with a record of normal test results showed that some patients exhibited signs of chronic graft injury or gene expression patterns associated with TCMR.<sup>113</sup> Gene expression profiles from patients with long-term liver transplants (median time 13 years post-transplant) and subclinical histologic abnormalities showed that some patients had profiles similar to that of TCMR, and high expression was related to progressive damage to the liver transplant.<sup>114</sup> Gene expression studies in liver transplants found that the expression of several genes was significantly modified by ischemia-reperfusion injury.<sup>115,116</sup> Existing studies of gene expression profiling in liver transplants have also been encouraging.<sup>90,114,117-124</sup> While these studies have shown promise for the use of gene expression-based diagnostics in liver transplants, definitive research is needed to develop these tools for use in the assessment of rejection and injury in liver allografts.<sup>95</sup>

### **1.2.2.10 Limitations in the SOC for liver transplant biopsy interpretation**

## **Chapter 1: Introduction**

Patient management using the current histology interpretation of liver transplant biopsies can be significantly impaired due to documented shortcomings: challenges in determining the source of tissue injury or loss of tissue function,<sup>110,125,126</sup> low interobserver agreement,<sup>4,110,127,128,128,129</sup> and uncertainty concerning ABMR in liver transplants.<sup>79,82–84</sup> Most histology reports for liver transplant biopsies are text-based and descriptive, and do not come to a clear diagnosis of either TCMR or ABMR.<sup>79</sup> The putative tolerogenic properties of the liver raise the possibility of a reducing immunosuppression to alleviate side effects,<sup>130–134</sup> but this cannot be tested properly without a precise and accurate monitoring system for diagnosing rejection. Diagnosing steatohepatitis/NASH remain complicated by the need for biopsy, inaccuracy of biochemistry values, and difficulties separating steatohepatitis from simple steatosis.<sup>103</sup>

### **1.2.3 A review of UC – a form of inflammatory bowel disease**

#### **1.2.3.1 Overview of UC in the colon**

UC is a chronic inflammatory disease of uncertain etiology and pathogenesis and varying activity, with suggested roles for both innate immunity and the adaptive T cell autoimmunity. UC occurs in the colonic and rectal epithelium, which can produce ulcers. The disease moves through phases of active disease and remission, and this cyclical nature poses therapeutic challenges. Symptoms of UC include abdominal pain, frequent bowel movements (that may contain blood, white blood cells, or mucus), nausea, loss of appetite, fatigue, and fever.<sup>135</sup> Ulceration, increased tissue friability, and chronic inflammation can lead to chronic bleeding, producing anemia in some patients.<sup>135</sup> Absorbing nutrients from food and liquids can be difficult. Severe complications or outcomes of UC include toxic megacolon or the need for colectomy.<sup>135</sup>

#### **1.2.3.2 UC disease mechanisms**

While UC is often discussed as an autoimmune disease, the involved immunology is complex and idiopathic. It is still unclear how the adaptive and innate immune system interact to create the diverse and heterogeneous patient population within UC. Many lines of evidence implicate a role for cognate T cell-mediated autoimmunity in UC. Mucosal T cells from UC patients exposed to IL-2 respond differently from controls,<sup>136</sup> and epithelial cells from UC patients stimulate CD4<sup>+</sup> T cells differently from normal epithelial cells.<sup>137</sup> Patients treated with T cell checkpoint inhibitors as immunotherapy for cancer often develop a UC-

## Chapter 1: Introduction

like colitis.<sup>138,139</sup> Genetic susceptibility loci for UC map in the HLA region, which controls many cognate cell responses,<sup>140,141</sup> and cytotoxic T cell transcripts for perforin and granzyme A are increased in the intestinal mucosa of UC.<sup>142,143</sup> Immunosuppressive drugs that control TCMR in organ transplantation (azathioprine, cyclosporine, and tacrolimus) are effective in some cases of UC, though efficacy is inconsistent.<sup>144–153</sup> However, the drugs that form the mainstay of UC management (anti-inflammatory biologics such as infliximab and adalimumab) have not been effective in preventing or treating TCMR in organ transplants, suggesting that there is more to UC than simply cognate T cell processes. Costimulatory blockade with cytotoxic T lymphocyte 4-immunoglobulin (CTLA4Ig, i.e. abatacept and belatacept) has no demonstrable efficacy in UC,<sup>154,155</sup> despite their success in preventing transplant rejection<sup>155</sup> and treating selected autoimmune diseases.<sup>156</sup>

### 1.2.3.3 Endoscopic correlates of UC disease activity

The Endoscopic Mayo score and Total Mayo score are primary measures of disease activity. The Endoscopic Mayo score is composed of an ordinal set of scores from 0-3, describing the severity of the disease as assessed during a scoping appointment (**Table 1.1**).<sup>157</sup> Patients with endoscopic Mayo score 0-1 are considered to have minimally active disease or to be in remission, while patients with an endoscopic Mayo score of 2-3 have active disease requiring intervention.

The Endoscopic Mayo score is a subscore of the Total Mayo Score, which includes other factors like the physician's global assessment, stool frequency, and rectal bleeding for a total score ranging from 0-12 (0 = mild disease/remission, 12=severe active disease).

### 1.2.3.4 Molecular studies in UC

Molecular analyses of UC biopsies have shown increased expression of transcripts and/or proteins for decay-accelerating factor CD55,<sup>158,159</sup> complement related proteins,<sup>158</sup> and calprotectin (a dimer of calcium proteins S100A8 and S100A9 highly expressed in myeloid cells).<sup>160,161</sup> Microarrays or RNA-sequencing have shown many expressed transcripts in UC compared to controls, e.g. from screening colonoscopies.<sup>162–169</sup> Differentially expressed genes varied between analyses, probably reflecting details of

## **Chapter 1: Introduction**

the comparators,<sup>162–164</sup> including other forms of IBD (i.e. Crohn's disease or inflammatory bowel disease undetermined (IBDU)).

Microarrays have been used to analyze transcripts in UC biopsies, revealing a large-scale disturbance involving inflammatory cells, parenchymal injury and dedifferentiation with similarities to the transcript sets in TCMR of organ transplants, the prototypic T cell-mediated disease.<sup>64,66,170–172</sup> Expression of these transcript sets derived in rejecting kidney transplants correlated with the endoscopic Mayo score and the presence of lamina propria lymphoplasmacytic infiltrate on histology in colon biopsies,<sup>172</sup> compatible with a relationship between cognate T cell activity and UC.

### **1.2.3.5 Limitations of the current SOC in UC**

Patient management in UC could benefit from a biopsy-based molecular test, where the clinician could predict in advance which therapy would be required to induce remission and to monitor response to therapy. There is significant heterogeneity in the UC patient population, and patients with similar clinical presentations and identical Mayo scores often do not respond to the same therapeutic regimen. This creates a challenge for the clinician managing the patient's disease and forces therapeutic decisions to be made using trial-and-error. Given the extensive cost and adverse side-effects of many UC therapies (e.g. biologic therapies), improved management is imperative.

### **1.2.4 A review of machine learning and artificial intelligence in medicine**

The concept of 'deep medicine' and use of artificial intelligence, statistics, and machine learning has gained prominence in medical diagnostics over recent years.<sup>1</sup> The ability to predict outcomes, provide probabilities, and offer objective data is especially relevant in areas of medicine where the SOC is unreliable.<sup>1</sup> If properly developed with regards to over- or underfitting, machine learning algorithms can provide high quality, robust, and precise diagnostic data.<sup>6</sup> Results produced from machine learning algorithms are also more likely to be correct with regards to the true latent disease phenotype than their SOC counterparts that rely on visual assessments or consensus-driven diagnostic guidelines.<sup>6</sup> This accuracy is a result of the data-driven approach, the use of continuous numbers, and the ability of

## Chapter 1: Introduction

machine learning algorithms to correct a certain amount of error in the training set (**Figure 1.2**).<sup>6</sup> While algorithms cannot replace clinicians, the objective data produced by AI can increase the quality of diagnoses.<sup>1</sup> Analyses for the development of medical diagnostics can take two forms: supervised or unsupervised.

### 1.2.4.1 Supervised analyses with microarray data

Supervised analysis uses sample labels to train an algorithm. In this way, a combination of numerical data (e.g. transcript expression from a set of biopsies) is informed by its labels (i.e. clinical, histologic, or endoscopic data in these analyses). Specific patterns or results in the data that correspond to a feature can be detected and their predictive potential evaluated.

#### 1.2.4.1.1 Machine learning algorithms

Machine learning algorithms are a form of supervised analysis. A data set of samples plus their accompanying SOC information is divided into two groups for cross-validation (CV). Groups can be equal (e.g. a 50/50 split of the data set for training and testing, respectively) or unequal (e.g. train on the entire set minus one, test in the one sample as done in 'leave one out' methods, **Figure 1.3**). Splitting groups evenly offers the advantage of testing in a larger set of samples and thus obtaining a better estimate of algorithm performance, while training in a larger set of samples offers the advantage of giving the algorithm more initial information to base its predictions on. Ideal algorithms will both 1) be able to assign highly accurate values within the existing data set, showing that it fits the current data closely, and 2) be able to assign highly accurate results in a new data set in which it was not trained. However, these two principles are in conflict (bias-variance tradeoff).<sup>173</sup> An algorithm cannot fit the current data set too closely (overfit, biased) in a population subset or it may underperform and misrepresent future subsets with a different case mix. Likewise, an algorithm must not be so loosely fit to the current data that it has not 'learned' the appropriate information to make the designated prediction in the current or future dataset (underfit, excessive variance). The bias and variance must be balanced in order to produce an algorithm that performs well in the current data and is usable in new, unknown populations.

## Chapter 1: Introduction

Machine learning algorithms are diverse (e.g. linear discriminant analysis, random forests), and each method approaches a problem from a different perspective. Several algorithms can be combined by taking the mean or median value of all estimates (called a consensus or ensemble approach).<sup>6,174–179</sup> Statistical and machine learning literature discusses the increases in accuracy and stability when consensus approaches are used; the use of an average of multiple ‘observers’ or algorithms is more likely to be correct than any single observer.<sup>180</sup>

In cases where a single machine learning method is used, there is typically no a priori method of determining which algorithm will be best for the data set or the question asked (the ‘No Free Lunch’ theorem).<sup>181</sup> Instead, the choice of machine learning algorithm is conventionally based on practicality (e.g. ease of use, availability of software, etc.). No machine learning method will perform better than all other methods in all possible datasets.

### 1.2.4.1.2 Class comparisons for microarray data

Class comparisons can reveal distinct patterns in the data associated with a given SOC classification. Data sets are divided into binary ‘disease positive’ and ‘disease negative’ classes, and the appropriate t-test applied to the transcript data. Transcripts that differ significantly between the two classes are highlighted. This method can determine which transcripts are highly associated with particular SOC features. If biological mechanisms are pre-annotated for their associations with specific transcripts (as is done in pathogenesis-based transcript sets ‘PBTs’), top transcripts can link SOC classes with common biological mechanisms of disease (e.g. inflammation, fibrosis).

### 1.2.4.2 Unsupervised analysis with microarray data

Unsupervised analysis does not use labels to inform the data (e.g. using only molecular data with no SOC labels). This can be achieved using clustering methods (i.e. AA or k-means clustering). Samples are grouped based on similarities found in the data, which may or may not correspond to groups assigned by other labels. Some forms of unsupervised analysis provide only group information, failing to

## Chapter 1: Introduction

recognize that some biopsies are closer to group boundaries than others (e.g. standard k-means). Other forms recognize the relationship between samples and assigned groups while still assigning distinct clusters (e.g. AA, fuzzy k-means clustering). The advantage of unsupervised analysis is especially apparent in data sets where the gold standard class labels are unreliable, and the signal-to-noise ratio is too low for supervised analysis.

### 1.2.4.3 Probabilities and continuous data in medicine

Medicine has historically focused on categorical data (e.g. diagnoses). The final decision to treat the patient or not to treat them is binary, thus distinct classes and categories of patients have been regarded as easier to interpret. Although many SOC classifications remain categorical, probabilities are becoming more common in medical testing with the advantage of conveying the 'confidence' or predicted accuracy of the test results. For example, a diagnostic report can state a result in terms of grades 0-3, but this fails to recognize samples that were actually on the borders of those categories (near cutoffs). Probabilities instead provide a continuous value stating how likely it is that the given sample had a specific feature. While many SOC diagnostic tests still use categorical data, clinicians can and do use probabilities in their practice.<sup>182-184</sup> This allows the clinician to be aware of their level of confidence in their decision to treat the patient or not (the cost of a negative error versus a positive error), rather than simply making that decision because the patient's diagnostic information met a given threshold.

## **Chapter 1: Introduction**

### **1.4 TABLES**

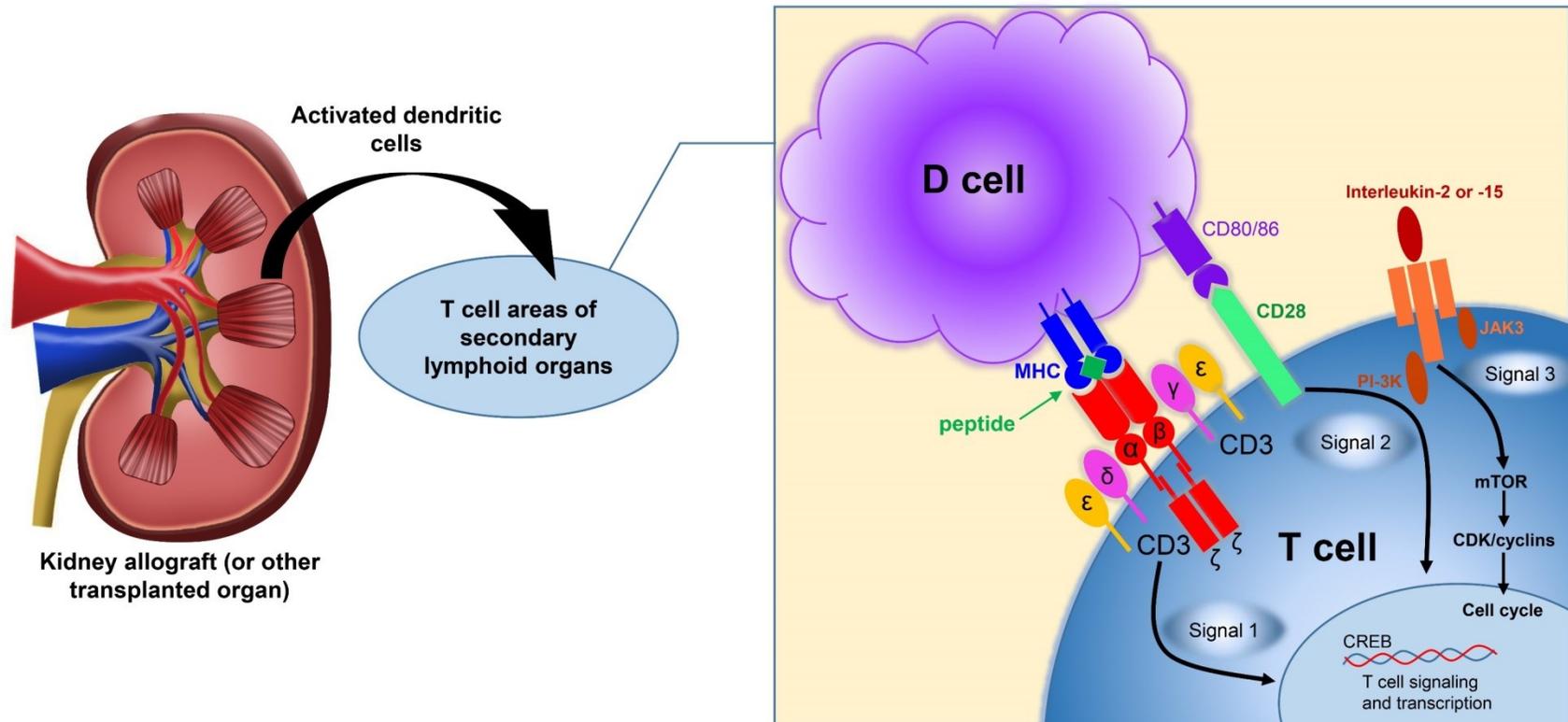
## Chapter 1: Introduction

**Table 1.1** Endoscopic Mayo Scoring system for UC. Adapted from E. Paine, Colonoscopic evaluation in ulcerative colitis. Gastroenterology Report (2014)<sup>157</sup>

<b>Endoscopic Mayo score value</b>	<b>Interpretation</b>
<b>0</b>	Normal mucosa/UC remission. No disease activity.
<b>1</b>	Mild UC. Mild friability, reduced vascular pattern, mucosal erythema.
<b>2</b>	Moderate UC. Friability, erosions, complete loss of vascular pattern, significant erythema.
<b>3</b>	Severe UC. Friability, erosions, complete loss of vascular pattern, significant erythema, plus ulcerations and spontaneous bleeding.

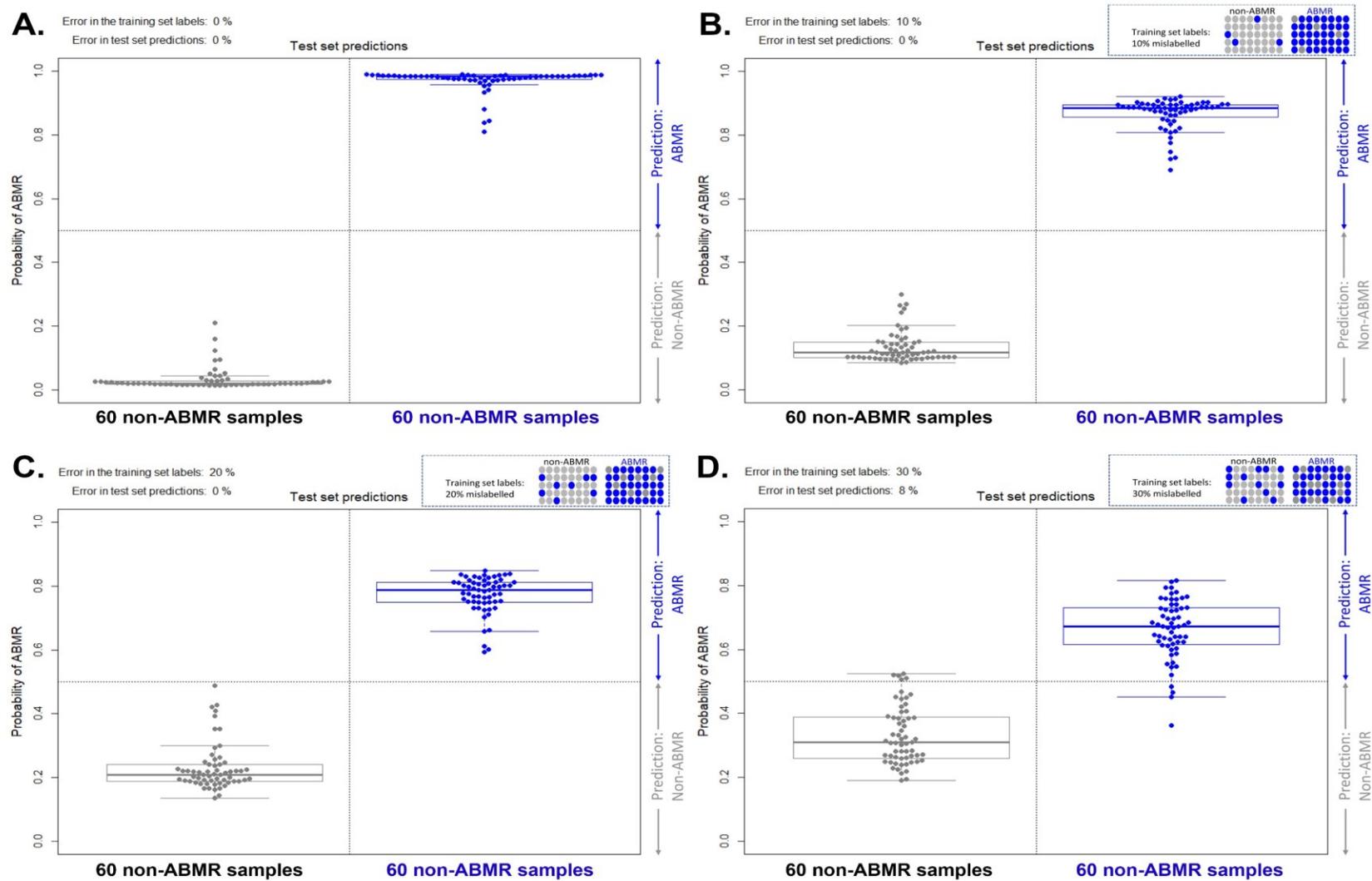
## **Chapter 1: Introduction**

### **1.5 FIGURES**

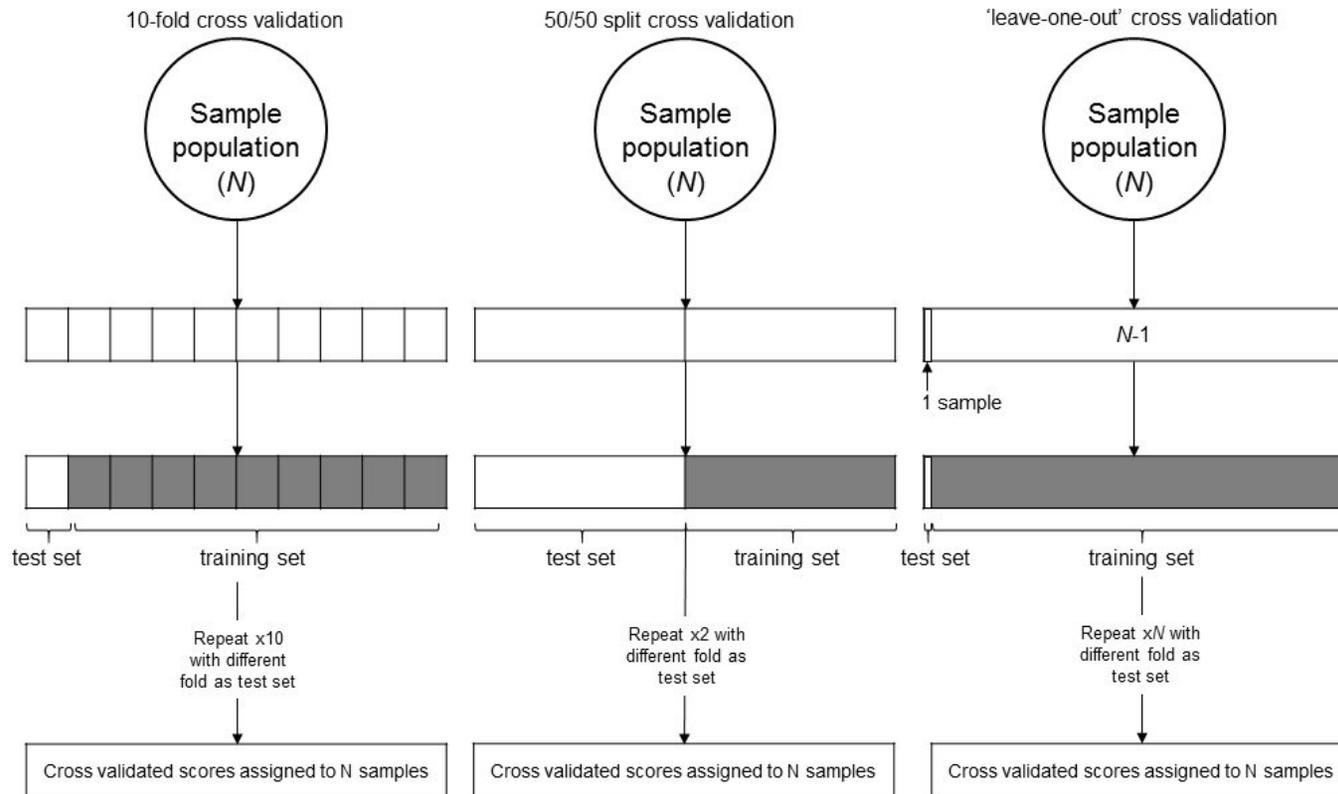


**Figure 1.1 T cell signaling involved in transplant T cell-mediated rejection.** Activated dendritic cells move from the transplant to T cell areas of secondary lymphoid organs, where they encounter T cells. Through a combination of signals 1 and 2, the T cells can be activated and triggered for clonal expansion. Transcription and subsequent production of IL-2 and IL-15 constitute signal 3, and promote clonal expansion through the activation of the mTOR pathway and other pathways.

## Chapter 1: Introduction



**Figure 1.2 Test set predictions when error is intentionally introduced through label flipping.** Test set predictions separating ABMR samples from non-ABMR samples are shown when the training set contains A) 0%, B) 10%, C) 20%, and D) 30% error introduced intentionally by inverting labels prior to training. In each case, the true ABMR and non-ABMR labels applied by the algorithm separate, with the error in the test set predictions substantially less than that introduced in the training set.



**Figure 1.3 Various cross-validation methods used to assign machine learning scores to a sample population.** A sample population with size  $N$  is split into a training set and test set (size varies with the methodology chosen). The algorithm is trained in one set and tested in the other to avoid training and testing in the same population (overfitting). The end product is a population where every sample has an assigned score, and no sample score was assigned by an algorithm that was trained on that particular sample.

## **CHAPTER 2**

## **METHODS**

## Chapter 2: Methods

### 2.1 OVERVIEW

These analyses used gene expression data collected from tissue biopsies by microarray analysis and examine the relationship between that data and various forms of disease and injury. Several analyses assess the stability and precision of this gene expression data and the associated molecular classifiers and scores. All samples were collected under approved study protocols. RNA was extracted from each sample, then processed using manufacturer's protocols or adapted protocols for microarray analysis. Quality and quantity of RNA, complementary deoxyribonucleic acid (cDNA), and microarray results were measured and annotated for all samples. All statistical analysis was done in the R computing language with appropriate packages and versions as described.<sup>185</sup> Sample populations and experimental methods are described in the subsequent sections.

#### 2.1.1 Samples used for histologic and microarray assessment

All samples included were collected under study protocols that were approved by the Institutional Review Board (IRB) at the University of Alberta (INTERCOMEX: Pro00022226, INTERLIVER: Pro00022225, and IBD: Pro00029109 and Pro00044450). Samples were collected between 2004 and 2020. Clinical data was collected and stored in a laboratory data management system (Genologics Geneus SDMS, later migrated to REDCap). These studies are based on mRNA extracted from tissue biopsies retrieved from transplanted kidneys, transplanted livers, and native colons biopsied during endoscopy for UC. All samples were collected prospectively in their respective studies (INTERCOMEX NCT01299168, INTERLIVER NCT03193151, and the IBD Microarray Project (Clinicaltrials.gov registration pending)). Samples were collected during routine SOC biopsy retrieval and placed immediately in a cryovial containing a solution of RNA<sup>later</sup><sup>™</sup> or storage equivalent to preserve the RNA integrity. Biopsies were stored at 4°C if not immediately shipped, or at ambient temperature if shipped the same day. Shipment was carried out via courier at ambient temperature over 24-36 hours. Samples batched over a period of time and sent in a single shipment were frozen at -20°C to -80°C and shipped via courier on dry ice. All biopsies were obtained from consented patients who met the respective study inclusion criteria.

### 2.2 BIOPSY COLLECTION

All biopsies were collected at participating international centers with IRB approval and patient consent. Biopsies were suspended in RNA/ater™ solution immediately after collection, then shipped via courier to the ATAGC laboratory in Edmonton, AB, Canada. Specific biopsy collection details are given in the following sections.

#### 2.2.1 Renal transplant biopsy collection

##### 2.2.1.1 Standard renal transplant biopsy collection

Kidney needle biopsies were collected at participating international centers under ultrasound guidance using spring-loaded needles. Depending on the individual center IRB approvals and sample availability, samples were submitted as an additional full core or as a portion of a core (average biopsy size 3 mm). If a portion of a core was submitted, this piece was separated at the center immediately after biopsy (within ~20 minutes of biopsy collection) then placed immediately in RNA/ater™ solution. A small subset of biopsies used AllProtect as the storage material (N=115). Biopsies stored in AllProtect were found to be the same as those stored in RNA/ater™ in all analyses.

##### 2.2.1.2 Cortex-medulla renal biopsy cohort collection

The cortex–medulla comparison cohort included 26 renal allograft needle biopsies (2 partial cores each), 3 unpaired cortex, and 1 medulla samples from 26 recipients, performed for graft dysfunction and/or proteinuria within the INTERCOMEX study ([www.clinicaltrials.gov](http://www.clinicaltrials.gov), NCT01299168) between June and October 2015. Biopsies were provided within budgetary constraints and per project/ethics protocols for submission to the INTERCOMEX study. Specimens were selected if they contained sufficient material for a comprehensive evaluation of both conventional morphology and region-specific molecular gene expression patterns. Biopsies were performed under ultrasound guidance using a 16- or 18-gauge needle. Immediately after biopsy, one core was evaluated by microscopy; the approximate number of

## **Chapter 2: Methods**

glomeruli was determined in 15 biopsies. This core was separate from those sent for routine assessment (histology, immunochemistry, and electron microscopy). The core was divided into two pieces (1–3-mm length), designated cortex and medulla by its morphological appearance including the presence of one or more glomeruli (median 2.5 glomeruli per cortical specimen, interquartile range 2.25–3, range 1–10) versus medulla showing the presence of medullary rays without glomeruli. Immediately after counting the number of glomeruli, specimens for molecular workup were suspended in *RNAlater*<sup>™</sup> solution and were immediately shipped at room temperature.

### **2.2.1.3 Renal biopsy biological replicate collection**

Biopsies divided in half without assessing proportion cortex for use as biological replicates were at least 4 mm in length and were selected initially based on size and diagnosis from the samples in the study. The biopsy core was then cut evenly in half, and both halves were processed separately as “B1” and “B2.” Reports and sample quality data were generated for both samples, and molecular scores compared and documented.

### **2.2.1.4 Renal biopsy technical replicate biopsy collection**

Technical replicates were prepared by dividing the RNA extracted from a single biopsy into two aliquots and processing the aliquots in parallel by two technicians on separate chips. Reports and sample quality data were produced for both samples.

## **2.2.2 Liver biopsy collection**

Liver biopsies were collected at participating international centers as needle biopsies percutaneously under ultrasound guidance, transjugularly with contrast and X-ray guidance, or laparoscopically under anesthetic. A small subset of biopsies from Polish centers were collected laparoscopically or during the surgery as wedge biopsies (N=8). Depending on the individual center IRB approvals and sample availability, samples were submitted as an additional full core or as a portion of a

## **Chapter 2: Methods**

core (average biopsy size 4.5 mm). If a portion of a core was submitted, this piece was separated immediately after biopsy (within ~20 minutes of biopsy collection).

### **2.2.3 UC sample collection**

UC biopsy samples were collected during endoscopy for indication or follow-up in patients diagnosed with UC. All UC biopsies were collected circumferentially, with four bites collected at the most inflamed portion of the colon as determined by the endoscopic Mayo score. All four bites were placed in a cryovial of RNA $\text{later}^{\text{TM}}$  and stored at  $-20^{\circ}\text{C}$  (for intermediate storage <1 week),  $-80^{\circ}\text{C}$  (for long term storage), or at  $4^{\circ}\text{C}$  if immediately processed. All four bites were processed as a single sample and assessed on one microarray, except for a subset of 16 samples used as biological replicates for the assessment of sample-to-sample heterogeneity. These 16 samples were extracted in sets of 2 bites and assessed on two microarrays as paired samples.

## **2.3 PATIENT POPULATIONS AND DEMOGRAPHICS**

### **2.3.1 Kidney population**

The majority of renal transplants in the largest renal population used (N=1679) were from deceased donors (N=949, 67%). All biopsies were collected for indication. Mean recipient age at biopsy was 51 years (range 8-91); 726 (55%) were male and 953 (45%) were female. Mean donor age was 44 years (range 1-85). Median time post-transplant (TxBx) over all biopsies was 650 days (range 1-12371). Common indications for biopsy included slow deterioration of renal function, rapid deterioration of renal function, investigation of proteinuria/rejection/BK virus nephropathy/serum creatinine levels, and stable but impaired graft function. Specific details of populations used in each analysis are given in the respective chapters (see Chapters 3, 4, and 5).

### **2.3.2 Liver population**

## Chapter 2: Methods

Two populations were used in the liver analysis: N=235 for liver rejection studies and N=337 for liver injury studies.

In the liver rejection N=235 population, most biopsies were collected for indication or follow-up (N=173, 74%), and a small number were collected for protocol/surveillance/follow-up (N=38, 16%). Biopsies were collected from 217 patients. Mean recipient age was 50 years (range 2-71); 107 (49%) were male and 110 (51%) were female. Median TxBx was 967 days (range 0-11676). DSA is not SOC in liver transplantation, thus only 17 DSA values were measured within this population (See Chapter 6).

In the liver injury N=337 population, most biopsies were collected for indication (N=238, 83%), and a small number collected for protocol/surveillance/follow-up (N=50, 18%). Biopsies were collected from 311 patients. Mean recipient age was 50 years (range 2-71); 146 (49%) were male and 149 (51%) were female. Median TxBx was 904 days (range 0-12569). DSA was collected in 33 cases (see Chapter 7).

Specific details of populations used in each analysis are given in the respective chapters.

### 2.3.3 UC population

The majority of colitis biopsies collected for these analyses were collected at the University of Alberta (N=53), with a smaller subset collected at Cedars-Sinai hospital in Los Angeles, USA (N=18). Mean patient age was 39 years (range 19-66) although this did not include the Cedars-Sinai samples as no demographical information could be collected due to IRB approvals and patient confidentiality (18 not available 'NA'). Biopsies were collected from 25 (47%) male patients and 28 (53%) female patients. Median time from index biopsy to the follow-up was 68.6 months (mean 60.9). Estimated mean disease duration was 124 months. Most patients included in the study were given one or more treatment regimens at the time of the biopsy. Demographics are given in Chapter 8 and 9.

A small subset of 16 biopsies were collected from the University of Alberta (N=11) and Cedars-Sinai Hospital (N=5) during screening colonoscopies and used as control samples. All patients had no history of UC. Mean patient age was 58 years; 5 (45%) were male and 6 (55%) were female. All biopsies

## Chapter 2: Methods

were collected for indications unrelated to inflammatory bowel diseases (family history of colon cancer, history of other cancers, history of bleeding, or other indications), and most screenings resulted in polyp removal or no abnormal histology findings. No screening biopsy samples were taken from polyp tissue. Demographics for the screening colon samples are given in Chapters 8 and 9.

## 2.4 ASSESSMENT OF HUMAN TISSUE BIOPSIES BY MICROARRAY

### 2.4.1 Samples assessed by microarray

Microarray analysis using PrimeView GeneChip arrays requires clean, high quality, biotin labeled cRNA. Samples preferably contain adequate tissue and minimal blood, fatty or scarred tissue as this may impact the quality and amount of RNA obtained. RNA was obtained from tissue biopsies stored in RNA $later$ <sup>TM</sup> (or similar RNA preservation material), or snap-frozen. Contact with other solutions (e.g. saline) or an extensive length of time between sample collection and placing the sample in RNA $later$ <sup>TM</sup> or freezing was discouraged at all sample collection sites.

### 2.4.2 RNA extraction, labeling, hybridization, and microarray processing

#### 2.4.2.1 RNA extraction

RNA was extracted from all tissue biopsies using an established TRIzol-chloroform method.<sup>66</sup> Tissue biopsies were removed from RNA $later$ <sup>TM</sup>, then placed in a sterile round bottom microcentrifuge tube with 0.5 mL TRIzol reagent (Invitrogen), and a stainless-steel bead. Samples were disrupted in the tube for 3 minutes at 30 Hz using a QIAGEN TissueLyser. Chloroform (0.1 mL) was added to the tube, followed by centrifugation for 15 minutes at >8,000g and 4°C. The aqueous layer was removed and placed in a new tube. Ethanol (100%) was added at 0.7xVolume to the tube and mixed. The mixture was transferred to the RNEasy MinElute spin column (RNeasy Micro Kit, QIAGEN), then centrifuged for 15 seconds at >8,000g and 4°C. The flow-through was discarded, 0.5 mL of the RW1 buffer added to the column, and centrifugation repeated. The flow-through was discarded again, 0.5 mL of RPE buffer added, and centrifugation repeated. The flow-through was discarded a last time, 0.5 mL 80% ethanol added to

## Chapter 2: Methods

the tube, and a final centrifugation done for 2 minutes at  $>8,000g$  and  $4^{\circ}C$ . The spin column was removed and placed in a new 1.5 mL Eppendorf tube, and 14  $\mu L$  of RNAse free water added followed by a 1-minute incubation at ambient temperature. The tube was then centrifuged at full speed to elute the final RNA solution.

### 2.4.2.2 Quality and quantity assessments of extracted RNA

Final sample concentration was measured by UV absorbance using a Nanodrop 2000. RNA quality was assessed by the Agilent 2100 Bioanalyzer, and the 260/230 and 260/280 ratios recorded. Samples were selected for an alternate 1-day labeling if the concentration was  $>80$  ng/ $\mu L$  (only done in kidney biopsies as an established conversion set for liver is still pending), and the standard 2-day labeling protocols per manufacturer's instructions if the concentration was  $<80$  ng/ $\mu L$ . Samples were excluded if they did not produce a minimum of 12  $\mu g$  of labeled RNA.

Quality was assessed for all extracted total RNA using the RNA Integrity Number (RIN). In kidney RNA, the average RIN was 7.9, in livers the average RIN was 8.2, and in UC the average RIN was 7.2. Mean quantity of RNA from extracted kidney biopsy samples was 3.18 $\mu g$  (1 core or partial core), in livers 9.98 $\mu g$  (1 core or partial core), and in UC 21.6 $\mu g$  (2-4 bites). 99% of biopsies submitted to each study were successfully labeled and suitable for microarray analysis.

### 2.4.2.3 Microarray processing

All samples were labeled using the Affymetrix/Thermo Fisher Scientific 3' IVT Plus Labeling Kit and analyzed using PrimeView GeneChip microarrays. Starting RNA samples were hybridized with a primer for 15 minutes. Reverse transcription for first strand cDNA synthesis was completed in 0.5 hour for 1-day and 2 hours for 2-day labeling. The second strand of cDNA was synthesized in 1 hour, followed by a step to denature the enzymes at  $65^{\circ}$  over 10 minutes. The resulting cDNA was amplified over 2 hours for 1-day labeling or over 16 hours for 2-day labeling. The anti-sense cRNA labeled with biotin per kit instructions. Biotin-labeled cRNA was cleaned using columns or provided beads, fragmented, and

## **Chapter 2: Methods**

hybridized over 16 hours to the PrimeView GeneChip array. Arrays were washed and stained using an Affymetrix Fluidics station. All arrays were scanned using the Affymetrix GeneArray Scanner and processed using GeneChip Command Console software per manufacturer's protocols.

### **2.4.3 Microarray data preparation**

The Ratio G method was used to minimize batch effects and differences between 1- and 2-day labeled samples.<sup>186</sup> The resulting output was a matrix of numbers obtained from the .CEL file (1 row per probe set, 1 column per sample) which contains expression measurements for all probes included on the GeneChip array.

## **2.5 DATA ANALYSIS**

### **2.5.1 PBT definition and methods**

PBTs were developed and annotated in previous analyses in human cell lines, mouse experimental models, and human transplant biopsies,<sup>187</sup> and are associated with biological mechanisms in rejection and injury (<https://www.ualberta.ca/medicine/institutes-centres-groups/atagc/research/gene-lists>). PBT values represent the mean fold difference in expression values of that set of transcripts in a population of biopsies compared to a selected control group. Control groups varied according to the organ and the analysis.

PBTs are defined as follows: ABMR-associated transcripts (ABMR-RATs),<sup>171</sup> alternative macrophage-associated transcripts (AMATs),<sup>188</sup> B cell-associated transcripts (BATs),<sup>65</sup> cardiac injury and repair-induced transcripts (cIRITs),<sup>189</sup> colon transcripts (CT1 and CT2),<sup>172</sup> damage-associated molecular pattern-associated transcripts (DAMPs),<sup>190</sup> endothelial DSA-selective transcripts (eDSASTs),<sup>191</sup> endothelial cell-associated transcripts (ENDATs),<sup>192</sup> gamma-interferon and rejection associated transcripts (GRITs),<sup>193</sup> immunoglobulin transcripts (IGTs),<sup>65</sup> injury and rejection induced transcripts – intermediate (IRITD3) and – late (IRITD5),<sup>194</sup> injury-repair associated transcripts (IRRATs),<sup>195</sup> kidney transcripts (KT1, KT2),<sup>196</sup> mast cell-associated transcripts (MCATs),<sup>197</sup> quantitative CTL-associated

## Chapter 2: Methods

transcripts (QCATs),<sup>24</sup> quantitative constitutive macrophage-associated transcripts (QCMATs),<sup>188</sup> ABMR+TCMR+all-rejection associated transcripts (RATs),<sup>198</sup> all-rejection associated transcripts (Rej-RATs),<sup>171</sup> TCMR-associated transcripts (TCMR-RATs),<sup>171</sup> and fibrillar collagen transcripts (FICOLs). Analyses in this thesis selected PBTs from this list based on the specific focus of that analyses, e.g. rejection, injury, or inflammation.

### 2.5.2 Principal component analysis (PCA) definition and methods

PCA is a dimensionality reduction technique used to plot multidimensional data into a smaller number of dimensions (e.g. two or three). In these analyses, samples were plotted according to their gene expression data. Dimensions represent variation within the data set based on the initial input, with the primary dimension (principal component 1, 'PC1') representing the source of the most variation in the data set. The second dimension represents the second highest source of variation (principal component 2 'PC2'), and so on with principal component 3 'PC3'.

PCA in these analyses was used to represent rejection or injury dimensions in biopsy populations. Input varied depending on analyses and was typically the rejection-associated transcript sets (RATs), or various PBTs relating to types of injury previously annotated in kidney and mouse studies. These PCAs were used to visualize biopsy reference set populations with respect to transplant rejection (ABMR, TCMR, Mixed rejection) and tissue injury (i.e. early acute injury, late chronic injury, fibrosis, atrophy, inflammation-associated injury with infiltrate).

PCA was used in each organ (kidney, liver, and colon) to visualize the biopsy population (by archetypal groups, by diagnoses, or by other classifications).

#### 2.5.2.1 PCA in kidneys

All kidney PCA plots were generated using the 'FactoMineR' R package<sup>199</sup> and the following seven molecular classifier scores as input: the classifiers for predicting diagnoses of ABMR and TCMR, as well as classifiers for predicting high grades of the g-, ptc-, cg-, i-, and t-lesions.

## Chapter 2: Methods

### 2.5.2.2 PCA in livers

PCA with respect to transplant rejection in liver biopsies was based on a population of 235 biopsies and their expression of the RATs (input as the union of ABMR-RATs, TCMR-RATs, Rejection-RATs and IQR filtered with a cutoff of 0.35). Principal component 1 (PC1), principal component 2 (PC2) and principal component 3 (PC3) represented most of the variance in the population and were used to generate the PCA plots. All PC based analyses were done in base R.<sup>185</sup>

PCA with respect to injury in liver was based on a population of 337 biopsies and their expression of transcripts associated with the following injury-associated PBTs: AMATs, DAMPs, IGTs, IRITD3s, IRITD5s, and QCMATs. These PBTs were selected for their relationship to acute early injury, atrophy-fibrosis, and ischemia-reperfusion-related injury from donation-implantation.

### 2.5.2.3 PCA in UC

PCA in colon was done using the top 300 transcripts increased in a class comparison between biopsies taken from patients with a high endoscopic Mayo score (2 or 3) versus biopsies from patients with a low endoscopic Mayo score (0 or 1) as input ('UC activity-associated transcripts').

### 2.5.3 AA definition and methods

AA is a form of unsupervised analysis that identifies a limited number ( $n$ ) of theoretical idealized extreme phenotypes called archetypes ( $A_n$ ) and assigns each biopsy  $n$  scores to describe its proximity to each archetype. AA uses only gene expression as input and does not use accompanying phenotypic information or clinical labels. AA was performed using the 'archetypes' package for R version 3.6.2.<sup>200</sup> AA was chosen as an analytical method because it allows biopsies to be assigned to groups (in rejection-based analyses =  $R_n$ ; in injury-based analyses =  $I_n$ ), while also preserving the uniqueness of each biopsy through a set of archetype scores ( $R_n$  and  $I_n$  scores). Biopsies were given a set of archetype scores that

## **Chapter 2: Methods**

define their relationship to the four idealized archetypes. Biopsies were assigned to groups based on their highest score in each case.

AA offers an advantage over other clustering methods (e.g. standard k-means clustering) as it not only assigns biopsies to distinct clusters but also includes proportional representation of each archetype for each biopsy. AA can be visualized using PCA, allowing for new biopsies to be shown in the context of the reference set by projecting them into the plot.

AA was used in each organ to define clusters in the population based on pure molecular data (unsupervised analysis). Number of clusters in each case was selected based on the scree plot (which shows the residual sum of squares (y-axis) versus the potential number of archetypes in a model (x-axis)), combined with domain-specific clinical knowledge of the prevalent phenotypes in each organ population and past experience with AA in other studies. Biopsies were assigned to clusters based on highest score in every analysis. AA was visualized using PCA. Top transcripts by p-value were analyzed by correlation with archetype scores and mean expression in each archetype cluster.

### **2.5.3.1 AA in Kidneys**

AA in kidneys was based on biopsy classifier scores in a population of 1208 biopsies as input. AA was generated using the same seven inputs as the PCA.

### **2.5.3.2 AA in Livers**

AA focused on liver rejection was done using the RAT transcripts as initial input. Prior to AA, the 600 RAT transcripts were condensed to 453 RAT transcripts after removing duplicates, 453 RAT transcripts remained. IQR filtering was performed with a cutoff of 0.35 to remove low variance transcripts, producing a final set of 417 transcripts. AA was used to assign scores to each biopsy from the 235 population and group biopsies based on their expression of this subset of RATs.

## **Chapter 2: Methods**

AA focused on liver injury used the same PBT input as the PCA. AA was used to assign scores to each biopsy in the 337 population and group biopsies based on this PBT expression.

### **2.5.3.3 AA in Colon**

AA used the top 300 transcripts increased in endoscopic Mayo score 0/1 versus 2/3 (the 'UC activity-associated transcripts') as initial input.

### **2.5.4 Definition of a molecular classifier and CV**

Molecular machine learning classifiers are derived based on specific styles of algorithms. The algorithm is first trained on a set of samples to learn which labels are associated with which resulting data, then tested in a separate set of samples to avoid overfitting. Minimizing bias error (error originating in the learning algorithm) and variance (response to minor variations in the training set) is an essential concern of all supervised machine learning, aka the bias-variance trade-off.

Although CV can be done through a variety of methods, these analyses used ten-fold CV exclusively. In this method, the samples are randomly separated into ten equal-sized groups or folds. Classifiers are trained on 9 of the folds, tested in the one remaining fold, and this process repeated ten times until all ten folds have been used for testing. One result is produced for each biopsy. In this way, an algorithm can 'learn' from a population what a sample label should be based on the inputs – the predictive variables (see Figure 1.3 in Introduction).

For new samples, a single classifier using an existing, locked data set (e.g. the N=1208 data set for kidney transplant biopsies) is used to predict the score of the new samples. The main purpose of CV is to assess how accurate the locked classifier is going to be for future data.

### **2.5.5 Developing an ensemble system**

## Chapter 2: Methods

Using multiple molecular classifiers can provide the final observer with multiple estimates of the predicted label for a new sample (defined as an 'ensemble' of classifiers). Ensembles provide the advantage of a consensus opinion with multiple classifiers providing estimates rather than multiple expert observers (a method which is impractical and rarely available). An ensemble in machine learning is defined as a collection of algorithms or classifiers approaching the same population and the same set of labels from slightly differing perspectives. Different types of classifiers tend to make different types of 'mistakes', but their consensus tends to be better than most individual classifier or opinions.<sup>180</sup> Ensembles offer increased stability as a weighted average of multiple opinions, which has been shown to be more accurate than a single opinion.<sup>180</sup>

For ensemble classifiers used in kidney transplant analyses, CV predictions of clinical features or diagnoses were made using 12 different machine learning methods: linear discriminant analysis (lda), regularized discriminant analysis (rda), mixture discriminant analysis (mda), flexible discriminant analysis (fda), gradient boosting machine (gbm), radial support vector machine (SVMR), linear support vector machine (SVML), random forest (rf), C5.0, neural networks (nnet), Bayes glm (bayesglm), and generalized linear model elastic-net (glmnet). This method has been established for kidney transplant biopsies.<sup>6,35</sup> The median is used as a stable estimate of the net results of all methods and is shown as the final molecular score on MMDx reports.<sup>6,201</sup>

Liver transplant classifiers were single classifiers only, using lda.

### 2.5.6 Data preprocessing

Data resulting from microarray experiments were processed to normalize the data set, remove batch effects (for example, effects based on 1-day labeling versus 2-day labeling). Normalization of data files involved robust multichip averaging (RMA) using the RefPlus<sup>202</sup> package from Bioconductor.<sup>203</sup> All raw data is stored as log<sub>2</sub> expression values.

### 2.5.7 Transcript filtering and selection

## Chapter 2: Methods

Interquartile range (IQR) filtering was used to remove low-variance transcripts in some analyses (described in relevant chapters).

Top transcript selection was based on two methods: t-tests and Spearman correlations. T-tests used gene expression data from two selected sub-populations (nominally defined as 'disease' and 'normal' for the purposes of the R code) to establish the transcripts most differentially expressed between classes. Transcripts most significantly elevated or reduced in the disease population versus normal were considered for future study, and the top 300-1000 were used in further analyses. Most t-tests were done as Bayesian t-tests calculated in R by the fdr adjustment method in the R Bioconductor "limma" package.<sup>204</sup> This process is robust and preferred over a standard t-test as it combines transcript specific variance and variance over the entire population of transcripts. This reduces the risk of obtaining a significant result by chance, which can occur in analyses that rely on transcript-specific values alone in studies with a smaller sample size. Alternatively, lists of top transcripts could be obtained by correlation with a selected continuous score (molecular or histology-based).

Lists of top transcripts by t-tests were sorted by p value. Lists of top transcripts by correlations were sorted first by p value to establish if the most significant genes had increased expression (positive correlation) or decreased expression (negative correlation). If top p values were dominated by positive correlation coefficients, the list was sorted by descending order of the correlation coefficient to avoid ordering transcripts by very small p values that were smaller than the threshold output from R (threshold =  $10^{-16}$ ), and vice versa if the top p values were dominated by negative correlation coefficients.

### 2.5.8 Overrepresentation analysis

Overrepresentation analysis was used to examine the top transcripts in several studies. Top 300 transcripts increased were used as input in each case unless otherwise specified. In analyses using transcripts associated with archetype groups, the top 300 transcripts decreased for the 'normal' phenotype and top 300 increased in disease phenotypes was used, as the primary interest was in these transcripts. Overrepresentation analyses in kidney and liver was done in R using the "enrichGO" function

## Chapter 2: Methods

from the “clusterProfiler” package.<sup>205</sup> Terms from the Biological Process (BP), Molecular Function (MF), and Cellular Compartment (CC) ontologies were used. Overrepresentation analyses in UC were done using the DAVID tool<sup>206</sup> and included KEGG pathways.

### 2.5.9 Moving average time course analysis

Moving average plots were generated in R version 3.5.1 using the ‘zoo’ package.<sup>207</sup> Window sizes fluctuated between analyses and was based on line smoothness and population size.

### 2.5.10 Discrepancy analyses

To allow comparisons between or within diagnostic platforms, all output from MMDx or histology was converted into six major rejection diagnostic classes: ABMR, possible ABMR (pABMR), TCMR, possible TCMR (pTCMR), Mixed rejection (Mixed), and No rejection (NR, see Chapters 4 and 5).

Discrepancies in each analysis were described as either ‘clear’ or ‘boundary’. Clear discrepancies were those involving unambiguous classes, e.g. ABMR versus NR discrepancies. Boundary discrepancies were those involving boundaries: possible rejection e.g. ABMR versus pABMR. Discrepancies involving mixed rejection diagnoses are impacted by even more boundaries i.e. between mixed and ABMR and between mixed and TCMR, and thus were given their own category of ‘mixed discrepancy’ in some analyses.

#### 2.5.10.1 Discrepancies within MMDx

MMDx assessment of rejection projects the prediction assigned by the ensemble of classifiers into PCA space, expresses the individual measurements on the report, then is signed out by an expert.<sup>6</sup> Official MMDx recorded diagnoses (Expert 1) were all signed out by one observer (PFH) based on established guidelines.<sup>6,35,66</sup>

## Chapter 2: Methods

Full text official MMDx diagnoses were converted into six classes using a set of guidelines. 'Full-', 'Severe-', 'Early-', 'Probable-', or 'Late ABMR' was converted to simply 'ABMR'. 'Full-', 'Severe-', 'Early-', 'Probable-', or 'Late TCMR' was converted to 'TCMR'. 'Possible-', and 'Cannot rule out ABMR' were converted to 'pABMR'. 'Possible-', and 'Cannot rule out TCMR' were converted to 'pTCMR'. Any MMDx diagnosis of confirmed ABMR and TCMR simultaneously was called 'Mixed'. Diagnoses excluding rejection (including abnormal with some other cause) were called 'NR'.

To assess discrepancies between alternative MMDx assessments, Expert 1 diagnoses<sup>66</sup> were compared to automated MMDx diagnoses ('AutoMMDx') assigned by the published random forest-derived algorithm predicting Expert 1.<sup>6</sup> AutoMMDx uses all variables on the MMDx report as initial input, predicts ABMR and TCMR with probabilities between 0.0 and 1.0, and assigns these probabilities to categories using cutoffs: if the ABMR/TCMR probability is  $\leq 0.4$ , non-ABMR or non-TCMR; if  $> 0.4$  and  $\leq 0.8$ , possible ABMR 'pABMR'/possible TCMR 'pTCMR'; if  $> 0.8$  the biopsy is called ABMR/TCMR.<sup>6</sup> Expert 1 diagnoses were also compared to a second expert observer (KMT, Expert 2).<sup>6</sup> Alternative sign-outs (AutoMMDx and Expert 2) were assigned independently and without knowledge of the original Expert 1 diagnosis.

### 2.5.10.2 Discrepancies within histology

To convert histologic diagnoses to six classes, all diagnoses that were not related to rejection were converted to 'No Rejection/NR'. ABMR and TCMR diagnoses were left as-is and assigned to those respective categories. 'ABMR suspicious' and 'TG' diagnoses were combined into the possible ABMR 'pABMR' category, as has been done in previous publications. 'Borderline' diagnoses were assigned to the pTCMR category. Diagnoses of both ABMR and TCMR by histology were assigned to the mixed rejection 'Mixed' category.

To estimate the how much interobserver variation or deviations from Banff guidelines contributed to discrepancies between MMDx and histology, an algorithm 'AutoBanff' was generated to strictly follow canonical rejection Banff guidelines. Algorithm development and the comparison between the local SOC

## **Chapter 2: Methods**

as reported on clinical research forms by each participating center and the output from the algorithm is described further in Chapter 5. To allow for comparisons with SOC histology and with MMDx, output from AutoBanff was limited to the six classes previously described.

### **2.5.10.2 Discrepancies between MMDx and histology**

Discrepancy analysis comparing MMDx and histology used the common six classes, and compared the Expert 1 sign-out to the SOC histology diagnosis. Expert 2 sign-outs were used in some analyses as a secondary comparator.

## **2.6 ASSESSMENT OF HUMAN ALLOGRAFT BIOPSIES**

### **2.6.1 Human biopsy pathology for kidneys (Banff Classification System)**

The Banff guidelines<sup>22,32</sup> are the current SOC for diagnosis of rejection, recurrent primary renal diseases, inflammatory conditions, and some forms of injury or loss of function in transplanted renal tissue. The system uses a series of categorical or descriptive lesions and features to reach a diagnosis. All biopsies in these analyses were assessed by the local pathologist nominally following Banff as per SOC from formalin-fixed paraffin-embedded (FFPE) sections. C4d staining was done as immunofluorescence (IF) or immunohistochemistry (IHC) per SOC at each center and reported during the study as positive or negative. DSA was assessed per SOC at all centers and reported as positive class II (DSA II), positive class I (DSA I), positive class II and class I (DSA I/II), negative (DSA Negative), DSA not done, or DSA unknown if the result was inconclusive or not recorded. DSA results were collected at each participating center, per SOC. Samples included in the study were considered adequate for pathology unless otherwise noted, with the exclusion of a set of medulla-only samples that were submitted for the study on renal biopsy heterogeneity.<sup>64</sup>

Lesions and features related to a diagnosis of ABMR include glomerulitis (g-score), peritubular capillaritis (ptc-score), glomerular basement membrane double contours (cg-score), C4d positivity representing complement activation, and DSA positivity representing alloantibodies generated by the

## Chapter 2: Methods

recipient plasma cells against the donor tissue. Glomerulitis (g-score) represents the degree of inflammation within renal glomeruli or glomerulitis defined as the proportion of glomeruli with complete or partial occlusion of the glomerular capillary by endothelial cell enlargement and leukocyte (polymorphonuclear and mononuclear cell) infiltration and is described in grades from 0-3. The ptc-score describes the amount of inflammation within peritubular capillaries caused by luminal infiltrating cells (polymorphonuclear and mononuclear leukocytes) and is determined from the most affected areas by grades 0-3. Glomerular basement membrane double contours or multilamination (cg-score) is based on the most affected glomerulus and is expressed as grades 0-3. C4d is evaluated as the percent of endothelial cells in peritubular capillaries and medullary vasa recta that produce a linear, circumferential staining pattern, and is expressed as grades 0-3 although the data recorded for these studies were restricted to either C4d positive or negative. DSA positivity as measured using Luminex beads at each center was reported by the tissue typing laboratory per SOC.

Lesions and features related to a diagnosis of TCMR include those describing interstitial inflammation (i-score), tubulitis (t-score), and intimal arteritis (v-score, also called endothelialitis or endarteritis). Interstitial inflammation (i-score) is described as the degree of inflammation only in non-scarred areas of the renal cortex (inflammation in scarred and non-scarred areas is captured in the total inflammation or 'ti'-score, however not all centers record this information as SOC), and is captured in grades 0-3. The t-score represents the extent of inflammation seen as mononuclear cells in the basolateral aspect of the epithelial cells in cortical tubules and is graded as 0-3. The v-score is inflammation defined as the presence of inflammatory cells in the subendothelial space of 1 or more arteries and is graded as 0-3.

### 2.6.1.1 Diagnostic classifications

Diagnostic classifications within the Banff system for renal transplants include rejection classes (ABMR, TCMR, or mixed rejection which is the presence of both ABMR and TCMR in varying degrees simultaneously), borderline rejection describing cases with weak but non-zero i- and t-score grades, and ABMR-suspicious classes describing the presence of ABMR-like lesions and features but without DSA

## Chapter 2: Methods

positivity or missing other elements of typical ABMR. BK nephritis is determined by viral load in the biopsy. Interstitial fibrosis with tubular atrophy (IFTA) was defined by the interstitial fibrosis (ci) score and tubular atrophy (ct) scores. Other possible non-rejection diagnoses included post-transplant lymphoproliferative disorder, calcineurin inhibitor toxicity, acute tubular injury, recurrent primary renal diseases (e.g. recurrent glomerulonephritis), de-novo glomerulopathy (excluding transplant glomerulopathy 'TG'), pyelonephritis, and drug-induced interstitial nephritis.

### 2.6.2 Human biopsy pathology for liver (Banff working classification system)

Although the Banff working classification system has been developed for use in liver transplants, it is not routinely followed by pathologists assessing liver biopsies. This contributes to the extensive levels of inter-observer variation recorded in liver pathology.<sup>4,128,208</sup>

Current versions of the Banff classification system for liver transplants<sup>209</sup> includes mostly descriptive, text-based features related to acute TCMR, chronic TCMR, plasma-rich rejection, acute ABMR, acute antibody-mediated injury, and chronic antibody-mediated injury.

TCMR diagnoses involve portal inflammation, bile duct inflammation damage, and venous endothelial inflammation. Portal inflammation describing infiltration of the triads by lymphocytes, neutrophils, or eosinophils (or all), and is expressed as grades 0-3. Bile duct inflammation damage describes inflammatory cells and associated damage in the ducts (from mild reactive changes to most or all ducts showing degenerative changes or foal luminal disruption) expressed as grades 0-3. Venous endothelial inflammation describes lymphocytic infiltration involving portal and/or hepatic venules and associated with perivenular hepatocyte necrosis and is graded 0-3. Chronic TCMR is described as involving two of the following: senescence related changes in small bile ducts, occasional loss in portal tracts, inflammation and perivenular fibrosis with or without focal obliteration, intimal inflammation with or without luminal compromise, inflammation damage, mural fibrosis, transition hepatitis with spotty necrosis of hepatocytes, and cholestasis. Plasma-cell rich rejection is described as portal and/or perivenular

## **Chapter 2: Methods**

plasma cell infiltrates in portal tracts or central veins, lymphocytic cholangitis, and a primary disease other than autoimmune hepatitis.

While there is still significant debate surrounding the presence or absence of ABMR in liver transplants, the current criteria for ABMR involve: C4d deposition (grades 0-3); and histologic grades 1-3 describing 1) portal microvascular endothelial cell enlargement in portal veins, capillaries, and inlet venules; 2) monocytic, eosinophilic, or neutrophilic microvasculitis/capillaritis; and 3) marked capillary dilation, microvascular inflammation, and focal microvascular disruption with fibrin deposition. Acute ABMR is described as a histopathological pattern of injury consistent with acute ABMR, plus positive serum DSA, diffuse C4d staining, and reasonable exclusion of other causes for these findings. Given the indeterminate nature of ABMR in liver allografts, the Banff criteria for ABMR in these analyses are not a major focus and will not be discussed further.

### **2.6.3 Human biopsy pathology and endoscopic assessment for UC**

Although biopsy histologic grading is not SOC for colon biopsy assessment, biopsies retrieved during endoscopy for UC or other forms of inflammatory bowel disease can be graded on the basis of activity-related scores, chronicity scores, and descriptors of architectural distortion.<sup>210</sup> Activity was described by neutrophil-mediated injury to the epithelium: neutrophils in crypt lumens (crypt abscesses), lamina propria neutrophils, and intraepithelial neutrophils (cryptitis). Chronicity was described by crypt distortion, basal lymphoplasmacytosis, or left colon Paneth cell metaplasia. Architectural distortion was reflected by crypt shortening or branching (branched crypts). Lamina propria lymphoplasmacytic infiltrate can also be assessed, and has been shown to be related to T-cell activity in colonic tissue.<sup>172</sup>

Endoscopic assessment of colon biopsies was assigned via the endoscopic Mayo score, which focuses on visual assessment of inflammation and ulceration. Grades were assigned according to severity and extent of the disease: 0-grade if no inflammation or ulceration is visible; 1-grade if mild inflammation is visible with some tissue friability, erythema, and a decrease in vascularity; 2-grade if severe inflammation is visible, with tissue friability, noticeable erythema, absent vascular patterns, and

## **Chapter 2: Methods**

visible erosions; and 3-grade if severe inflammation is present with visible ulceration, spontaneous bleeding, and severe friability.<sup>157</sup>

### **2.7 FUNCTIONAL ASSESSMENTS FOR CLINICAL DIAGNOSES**

#### **2.7.1 Assessments of kidney function**

Renal function is defined by the eGFR and associated serum creatinine levels.

#### **2.7.2 Assessments of liver function**

Liver function is measured by a series of laboratory biochemical tests, including serum albumin, bilirubin, ALP, ALT, and AST. Normal liver function is characterized by a high albumin score relative to a lower bilirubin, ALP, ALT, and AST. Impaired liver function, due to rejection, injury, or infection, can be associated with a lower albumin value, and higher bilirubin, ALP, ALT, and AST.

#### **2.7.3 Assessments of colonic function**

Colon function cannot be directly assessed, but effects of impaired function include changes in calprotectin levels measured from fecal or tissue samples, or changes in Endoscopic Mayo score. Function can also be assessed by the total Mayo score, which includes the Endoscopic Mayo subscore and adds stool frequency per day, most severe rectal bleeding scaled from 0-3, and a physician global assessment. These analyses focused on the Endoscopic Mayo score as a SOC representation of function and parenchymal integrity.

## **SECTION I**

# **ANALYSES IN KIDNEY TRANSPLANT BIOPSIES**

## **CHAPTER 3**

### **THE EFFECT OF TISSUE HETEROGENEITY ON MOLECULAR DIAGNOSIS IN RENAL TRANSPLANT BIOPSIES: REJECTION AND INJURY CAN BE ASSESSED IN MEDULLA**

**This chapter was published**

Madill-Thomsen KS, Wiggins RC, Eskandary F, Bohmig GA, Halloran PF. The effect of cortex/medulla proportions on molecular diagnoses in kidney transplant biopsies: rejection and injury can be assessed in medulla. *Am J Transplant*. 2017 Aug;17(8):2117-2128. Used with permission of Wiley Publishing.

## Chapter 3: Kidney Heterogeneity

### 3.1 OVERVIEW AND HYPOTHESIS

Limitations in existing diagnostic methods have triggered a strong interest in molecular phenotyping of kidney transplant biopsies as a new dimension in disease understanding. MMDx was developed for translating gene expression measurements into diagnostic assessment.<sup>211</sup> Like histology, a molecular biopsy assessment system requires consideration of the effect of sample adequacy. For example, when histologically assessing kidney transplant biopsies, an adequate specimen must have at least 10 glomeruli and 2 arteries,<sup>33</sup> usually requiring at least 2 cores. The sample requirements, including the proportion of cortex in the biopsy core, were unknown for molecular phenotyping. Biopsies processed to date, acquired in consented studies limited by institutional review boards, were typically relatively small segments of single biopsy cores (mean 3 mm), and stabilized immediately to prevent mRNA degradation without assessing the proportion of cortex, presenting an area of ambiguity for molecular measurements.

The present analyses were initiated to learn the effect of the proportion cortex on the fidelity of molecular readings, and whether rejection and injury can be assessed molecularly in medulla. This required us to develop a system for estimating proportion of cortex in a core, and to use this estimate to assess the relationship between proportion of cortex and molecular readings. A set of kidney transplant biopsies were obtained that were divided by a nephrologist (GAB) into cortex and medulla pieces before stabilization, based on visual assessment (light microscopy) of the presence of glomeruli as the indicator of cortex and medullary rays as the indicator of medulla. The goal was to define the top transcripts distinguishing cortex from medulla, develop a molecular estimate of the proportion of cortex, and incorporate this knowledge into MMDx molecular diagnostic reports to establish the effect of cortex/medulla proportions on the MMDx output. The relationships were also examined between estimated proportion of cortex and various previously published molecular scores including TCMR,<sup>36</sup> ABMR,<sup>42,212</sup> all rejection (ABMR, TCMR, or mixed rejection),<sup>213</sup> and AKI.<sup>195</sup> To facilitate interpretation of MMDx readings on paired cortex-medulla samples, the reproducibility of MMDx readings in technical and biological replicates was assessed.

### 3.2 BIOPSY COLLECTION AND PROCESSING

## Chapter 3: Kidney Heterogeneity

All biopsies were collected in the INTERCOMEX study [Clinicaltrials.gov NCT01299168](https://clinicaltrials.gov/ct2/show/study/NCT01299168). Demographics for the biological replicates, and cortex medulla divided pairs, and the kidney biopsy reference set (n = 1208 biopsies from 1045 patients) are shown in **Table 3.1**. All biopsies were for clinical indications, including investigation of newly discovered DSA. Routine protocol biopsies in patients with low risk were not included. The demographics of the 37 biopsies chosen for division into two halves as biological replicates and the 26 biopsies selected for division into cortex and medulla are also shown in **Table 3.1**.

### 3.2.1 Cortex-medulla cohort preparation and processing

The cortex–medulla comparison cohort included 26 renal allograft needle biopsies (2 partial cores each), 3 unpaired cortex, and 1 medulla samples from 26 recipients, performed for graft dysfunction and/or proteinuria within the INTERCOMEX study ([www.clinicaltrials.gov](https://www.clinicaltrials.gov/ct2/show/study/NCT01299168), NCT01299168) between June and October 2015. Biopsies were provided per project/ethics protocols for submission to the INTERCOMEX study. All specimens were processed if they contained sufficient material for a comprehensive evaluation of both conventional morphology and region-specific molecular gene expression patterns. Biopsies were performed under ultrasound guidance using a 16- or 18-gauge needle. Immediately after biopsy, one core was evaluated by microscopy; the approximate number of glomeruli was determined in 15 biopsies. This core was separate from those sent for routine assessment (histology, immunochemistry, and electron microscopy). The core was divided into two pieces (1–3-mm length), designated cortex and medulla by its morphological appearance including the presence of one or more glomeruli (median 2.5 glomeruli per cortical specimen, interquartile range 2.25–3, range 1–10) versus medulla showing the presence of medullary rays without glomeruli.

Immediately after counting the number of glomeruli, specimens for molecular workup were suspended in RNAlater™ and immediately shipped at room temperature.

Paired cortex/medulla sample processing included RNA extraction and microarray analysis on Affymetrix GeneChip arrays. Purified total RNA was labeled with the 30 IVT Plus kit (Affymetrix, Santa

### **Chapter 3: Kidney Heterogeneity**

Clara, CA) and hybridized to PrimeView microarrays (Affymetrix) according to manufacturer protocols published at [www.affymetrix.com](http://www.affymetrix.com). Microarray data were preprocessed using RMA.<sup>202</sup>

Resulting .CEL files were processed in R, and an automated report was generated. Processing time from extraction to reporting was ~48 h.

Reports for paired cortex and medulla biopsies were signed out simultaneously, and the classifier scores and gene expression measurements compared. Completed reports with a sign-out and comments were returned to the participating center.

#### **3.2.2 Biological replicate pair collection and processing**

Biopsies divided in half without assessing proportion cortex for use as biological replicates were at least 4 mm in length and were selected initially based on size and diagnosis from the samples in the study. The biopsy core was then cut evenly in half, and both halves were processed (**Figure 3.1**).

#### **3.3.3 Technical replicate pair collection and processing**

Technical replicate pairs were generated from a single aliquot of extracted total RNA. Two small aliquots of total RNA were taken and processed by two differing technicians on two microarrays (**Figure 3.1**).

### **3.3 HISTOLOGIC ASSESSMENT OF THE BIOPSIES**

Histologic evaluation was done on formalin-fixed paraffin-embedded sections. For C4d staining, a polyclonal anti-C4d antibody (BI-RC4D; Biomedica, Vienna, Austria) was used. Rejection features were graded and scored according to the Banff 2013 guidelines.<sup>214</sup>

### **3.4 MMDx ASSESSMENT OF THE BIOPSIES**

## Chapter 3: Kidney Heterogeneity

The output from the microarray is expressed in terms of 30 different classifiers and gene set scores and interpreted by a single observer (PFH) on the basis of the molecular classifier and gene set scores, without considering the conventional phenotype (e.g. histology, HLA antibody findings).<sup>211</sup> Thus, the results, like histology, are not based on any one result but on a combination of results.

### 3.5 DEVELOPING A MOLECULAR ESTIMATE OF PROPORTION CORTEX

#### 3.5.1 Renal structure

The kidney is a small organ (average 150g) that is composed of three major layers of tissue: medulla, cortex, and renal capsule (**Figure 3.1**, also see Introduction). These tissues can be distinguished by the presence or absence of glomeruli (cortex only), medullary rays (medulla), and fibrous structure (capsule). SOC histology assessment is reliant upon the cortical tissue as many lesions require the presence of glomeruli to accurately grade the biopsy. Any or all of these tissues may be found in a standard renal biopsy (see **Figure 3.1**).

#### 3.5.2 The development of a classifier-based prediction of biopsy composition

In order to assess if the molecular scores and features were affected by each biopsy's unique cortex-medulla composition, we first developed a gene-expression-based method for assessing how much cortex versus medulla tissue was in each biopsy.

Proportion cortex was molecularly determined using expression of the glomerular podocyte-specific transcript *NPHS2* (podocin). *NPHS2* is expressed exclusively in glomerulus and thus in cortical tissue and is not known to be regulated in disease states. Therefore, *NPHS2* expression is directly correlated to the proportion of cortex in a sample. A logistic regression equation was calculated based on samples submitted and microscopically determined to be either medulla or cortex and used to calculate proportion of cortex in unknown samples. Since our main goal was to identify samples with little or no cortex, cutoffs for proportion cortex were chosen arbitrarily as 0.2 in some experiments.

## Chapter 3: Kidney Heterogeneity

### 3.5.2.1 Gene expression differences between cortex and medulla

Gene expression was compared between paired cortex and medulla samples (**Figure 3.2**), expressed as fold change (y-axis) in cortex versus medulla, versus the p-value based on a paired t-test comparing cortex and medulla samples, in 26 cortex–medulla pairs (x-axis). The probe sets most differentially expressed are labeled.

**Table 3.2** ranks the top 30 differentially expressed probe sets by p-value. All of the top 30 had higher expression in cortex, as did 339 of the 408 probesets (83%) that differed between cortex and medulla with false discovery rate (FDR) <0.0001. No medulla-selective genes were among the top 30.

*NPHS2*/podocin was the most differentially expressed gene, both by p-value (FDR =  $6.2 \times 10^{-7}$ ) and fold change (average 28.6-fold higher in the cortex samples compared to the medulla samples). For this reason, we selected *NPHS2* expression as the basis for estimating the proportion of cortex in biopsy samples.

### 3.5.2.2 Podocin/NPHS2 expression in renal tissue

*NPHS2* expression is restricted to the cortex, in particular to the glomerular podocyte, and is relatively stable in its expression.<sup>215</sup> *NPHS2* is shown in **Figure 3.3** to define cortexness in a histologically determined cortex or medulla sample as well as the other 29 of the top 30 probe sets in PCA.

### 3.5.2.3 Estimating proportion cortex using NPHS2

We studied the expression of *NPHS2* in samples divided and separately processed as cortex or medulla (**Figure 3.4A**). Samples designated medulla had much lower *NPHS2* expression than cortex samples.

To establish an equation estimating the proportion of cortex in a biopsy, we assumed that the divided samples were either cortex or medulla as labeled. Four biopsies were included that were found by

## Chapter 3: Kidney Heterogeneity

light microscopy during preparation of the divided cortex-medulla cores to be either all cortex (N=3) or all medulla (N=1).

The distribution of predicted proportion based on *NPHS2* expression is shown in **Figure 3.4B**. The predicted proportion of cortex (and *NPHS2* expression) was high in most cortex samples and low in most medulla samples.

### 3.5.2.4 Expression of podocin/*NPHS2* in a reference set of renal biopsies (N=1208)

The distribution of *NPHS2* expression was analyzed in the reference set of 1190 intact biopsies (i.e. not divided into cortex and medulla). **Figure 3.5** shows the density plot distribution of *NPHS2* expression, compared to the actual cortex (top) and medulla (bottom) samples used to generate the measurement. A large proportion (89%) of the reference set biopsies had higher than 50% proportion cortex and overlapped the cortex samples. A small proportion of biopsies had low expression of *NPHS2*.

### 3.5.2.5 The MMDx ‘percent cortex’ classifier score

Estimates in new biopsies of the percentage of cortex tissue versus medullary tissue present in the sample were generated using the equation described by the curve in **Figure 3.4B**. All biopsies were henceforth assigned a percent cortex score based this equation.

## 3.6 EFFECT OF THE PERCENT CORTEX SCORE ON MMDx READINGS

### 3.6.1 Effect of percent cortex estimates on the reference set MMDx readings

The distribution in molecular scores across the high cortex and low cortex samples was compared to establish if the molecular ABMR, TCMR, and rejection scores were affected by predicted proportion cortex in a sample (y-axis in **Figure 3.6**). The biological replicates, the cortex/medulla pairs, and the reference set minus cortex-medulla samples are shown separately.

### Chapter 3: Kidney Heterogeneity

Positive ABMR molecular scores (right of the dotted vertical line, cutoff =0.2) were found in samples with both high and low proportion cortex (above and below dotted horizontal line, respectively, cutoff =0.2). No significant difference in ABMR scores was found in the biological replicates, the cortex/medulla subset, or the reference set minus cortex-medulla samples, respectively i.e. the likelihood of a positive molecular score was not consistently different in samples with high or low cortex content (**Figure 3.6A, 3.6B, 3.6C**). The results for the molecular TCMR scores (**Figure 3.6D, 3.6E, 3.6F**) and rejection scores (**Figure 3.6G, 3.6H, and 3.6I**) were similar. The statistical results from Chi square tests for these data are shown in **Table 3.3**.

Similar analyses were performed to determine if a relationship existed between the scores of the AKI transcripts (IRRAT scores). **Figure 3.7** shows the distributions of the predicted proportion cortex (y-axis) versus the AKI score. There was no relationship between the predicted proportion of cortex and molecular AKI. The statistical results from Chi square tests for **Figure 3.7** are summarized in **Table 3.3**.

#### 3.6.2 Effect of proportion cortex on the molecular rejection and injury scores

In an independent approach, the difference in four molecular scores (TCMR, ABMR, rejection, and AKI transcripts (IRRATs)) was compared between cortex and medulla to the difference seen between technical replicates or biological replicates. (The cortex and medulla segments were usually smaller than the other cores used for assessment, potentially increasing the sampling error.) The difference in the molecular scores for TCMR, ABMR, all rejection, and IRRAT in the technical replicates (upper panels), biological replicates (middle panels), and cortex-medulla pairs (lower panels) is shown in **Figure 3.8**.

The difference between the two scores can be seen on the y-axis versus the mean of the two scores (x-axis). The difference in the molecular TCMR, ABMR, and rejection scores between technical replicates was minimal (**Figure 3.8**, upper panels), and the difference between biological replicates (**Figure 3.8**, middle panels) was similar to that between technical replicates. The difference between cortex and medulla segments of a divided core (**Figure 3.8**, lower panels) was greater than in the biological replicates.

## Chapter 3: Kidney Heterogeneity

However, the positive-negative classifier calls for TCMR, ABMR, or rejection (indicated by the horizontal line) were usually in agreement.

### 3.6.3 MMDx assessment of cortex-medulla pairs and biological replicate pairs by multiple classifiers

Diagnostic assessment of the microarray analysis of a biopsy in the MMDx system uses multiple classifier scores for each rejection diagnosis and is interpreted by a single observer (PFH) independent of the histology and DSA status.<sup>66</sup> The diagnoses considered for this analysis were ABMR and TCMR. The consistency of diagnoses in paired samples of cortex and medulla was compared to the biological replicate set (**Table 3.4**). The data were divided into three groups based on their molecular report diagnostic sign-out: agreement; agreement with a difference in scale i.e. 'severe TCMR' versus 'moderate TCMR'); and disagreement.

Agreement between paired samples for the cortex-medulla set (23/26, 88%) was similar to the agreement between paired samples in the biological replicate set (32/37, 86%). These agreement values correlated more closely than the interobserver agreement usually recorded for histology assessments.<sup>2,3</sup>

## 3.7 INTERPRETATION OF THE RESULTS

The present study addressed the question of whether molecular methods could assess the proportion of cortex in a biopsy core and how the relative proportions of cortex versus medulla in a biopsy core would affect the molecular diagnosis of rejection and injury using the MMDx system. Genes were identified whose expression was different in cortex and medulla and the top example, *NPHS2*/podocin, was used as a marker to determine the proportion of each biopsy that was cortex versus medulla, and the effect of proportion cortex in 1190 biopsies on their interpretation in the MMDx system. In a smaller subset of paired biopsies, MMDx signatures were directly compared between technical replicates, biological replicates, and cortex-medulla pairs. The conclusion from both these data sets was that MMDx signatures were not impacted in a major way by whether the sample was largely medulla or largely cortex, although

### Chapter 3: Kidney Heterogeneity

89% of samples were >50% cortex. Cortex and medulla samples are less concordant than either biological or technical replicates but this generally did not affect the interpretation. This is important because sample collection for tissue RNA analysis usually specifies that the sample be immediately placed into RNA preservative without estimating the proportion cortex in order to avoid compromising RNA integrity. Thus, MMDx can estimate the proportion cortex in each biopsy when assessing kidney transplant biopsies using microarrays and can read rejection and injury information even when the biopsy is entirely medulla. Finally, in a qualitative sense, MMDx finds that medulla does undergo typical rejection processes, which are currently not being assessed by the histology diagnostic system that does not enable diagnoses of rejection in medulla. This suggests that molecular methods will be able to read other tissues that are currently not assessable by the current histology guidelines such as bronchial mucosa of lung transplants, and that all donor tissue probably undergoes rejection and injury.

In conclusion, molecular changes of rejection and injury can be assessed in medulla and show general agreement with cortex of the same specimen. Furthermore, molecular scores vary more between cortex and medulla than expected from biological replicates, but this generally will not affect disease interpretation by the classifiers and molecular scores. There are certain limitations to these conclusions, due to the small number of samples with certain combinations of features, e.g. samples with TCMR that also have a low proportion cortex. The potential for type II error limits the strength of inferences that can be drawn from the statistical findings, though additional studies will follow as more samples become available.

The use of *NPHS2*/podocin mRNA to estimate proportion of cortex for each biopsy sample is supported by current knowledge of *NPHS2*/podocin biology. The data presented here establish that very low *NPHS2*/podocin mRNA is mainly due to high medulla content. There are three reasons why a transplant kidney biopsy sample might have low expression of *NPHS2*: (i) the sample is comprised largely of medulla (containing no glomeruli); (ii) the expression of *NPHS2* is depressed by injury and/or inflammation; and (iii) many glomeruli are sclerotic and may have lost *NPHS2* mRNA.<sup>216</sup> Biopsies with TCMR had moderately reduced *NPHS2* mRNA (by about 50%) but not to the very low levels characteristic of medulla (28-fold lower than cortex). *NPHS2*/podocin mRNA was reviewed in previous mouse kidney transplant microarray studies, revealing that TCMR and AKI reduced expression of *NPHS2* by a maximum of 50-70% (data not

### Chapter 3: Kidney Heterogeneity

shown), similar to the reductions in other transcripts characteristic of well differentiated kidney tissue.<sup>196,217</sup> In addition, data on human biopsies with extensive atrophy-fibrosis (histologic ci scores >1) was studied to see if these biopsies had lost *NPHS2*. This analysis found relatively little loss compared to kidney transplants with minimal atrophy-fibrosis (ci<1), supporting the utility of *NPHS2* as a guide to low proportion cortex even in the presence of atrophy-fibrosis.

Podocyte loss does occur in glomerular sclerosis and in glomerular inflammation/injury such as in transplant glomerulopathy or recurrent glomerular disease,<sup>218</sup> but the loss of *NPHS2* mRNA may be partially offset by compensatory hypertrophy and increased expression in the remaining glomeruli.

The use of gene expression to reflect proportion of cortex versus medulla is not dependent on the assumption that gene expression will not change in disease states, but only that the genes specific for cortex are relatively well preserved in disease states. Many genes identified as preferentially expressed in cortex are well-known to be central to glomerular function because (a) their mutations are associated with inherited glomerular diseases including diffuse mesangial sclerosis, congenital nephrotic syndrome and focal and segmental glomerulosclerosis (*NPHS1*, *NPHS2*, *PTPRO*, *PLCE1*);<sup>215</sup> (b) circulating antibodies directed against the protein are associated with glomerular disease, namely membranous nephropathy (*PLA2R1*);<sup>219</sup> (c) they are key to renin-angiotensin system regulation through the juxta-glomerular apparatus (*REN*, renin);<sup>220</sup> or (d) they have otherwise been identified as highly expressed by glomerular podocytes (*CLIC5*, *PODXL*).<sup>221,222</sup> Tubulo-interstitial processes such as AKI or TCMR may cause some reduction in expression of the genes typical of the functioning kidney but not complete loss.<sup>196,223</sup> Thus, for the purposes of identifying samples that are primarily medulla, very low *NPHS2*/podocin is reliable, but it is important to remain cognizant of potential disease-related loss of expression of *NPHS2* or other cortex genes when diffuse inflammatory diseases such as TCMR are operating.

Note that some samples labeled as medulla had high expression of *NPHS2* and some labeled cortex had relatively low expression. This is not unexpected given the small size of the samples and the imperfect separation between cortex and medulla, where the precise boundary is difficult to establish by visual inspection using light microscopy.

### Chapter 3: Kidney Heterogeneity

Two features emerged for the TCMR score in relationship to proportion of cortex estimates. First, there were very few high TCMR scores in samples with the highest proportion cortex (0.9-1.0), suggesting that TCMR may somewhat reduce *NPHS2* expression, as it does with other kidney transcripts as shown in mouse kidney allografts with TCMR.<sup>196,217</sup> Nevertheless, the majority of TCMR was in samples with proportion cortex estimates of 0.5-0.9, indicating that any reduction of *NPHS2* expression due to TCMR or other diseases did not approach the very low levels observed in medulla. Second, although high TCMR scores are often recorded in samples with very low cortex content (i.e. medulla), there were fewer positive TCMR scores in Reference Set samples with very low proportion cortex, although the p value (0.03) was of only borderline significance. The possibility that TCMR is underrepresented in samples with very low proportion of cortex cannot be excluded.

While the proportion of cortex has no major effect on the performance of the molecular scores, the difference between cortex and medulla pairs was greater than between technical or biological replicates (within the limits of the power of this sample size of 1208 biopsies) inviting a caveat when diagnosing rejection molecularly in pure medulla. The number of biopsies in the reference set with almost pure medulla samples as estimated by *NPHS2* expression was small: less than 10% of biopsies had <0.5 estimated proportion of cortex, and fewer still had less 10% cortex. In the future as new classifiers to estimate rejection are developed, the effect of proportion cortex can be tested for each and possibly included as a variable in the algorithm.

Based on this analysis, molecular AKI changes (as estimated by the IRRAT score) are similar in cortex and medulla. IRRAT molecular scores were distributed similarly in the cortex-medulla set, the biological replicates set, and the reference set. It was previously reported (in an earlier version of the reference set) that assessment of injury should be molecular because this correlates with function whereas histologic estimates of acute tubular injury do not.<sup>195</sup> These analyses add the conclusion that AKI can be molecularly detected in medulla. This is not surprising, given that the top genes expressed in acutely injured kidneys are often expressed in other injured tissues and in cancers, reflecting the tendency of tissues to lose their differentiated features and become more similar after injury. For example, many AKI transcripts are also increased in biopsies from injured heart transplants.<sup>198</sup>

### Chapter 3: Kidney Heterogeneity

While these studies give reassurance about the reliability of molecular assessment on the biopsies available for the studies, the ideal biopsy size for molecular interpretation cannot be estimated because of the limited cores available for this research study. We have not seen an effect of biopsy size on MMDx if the RNA is high quality (as measured by RIN >7). The average biopsy was only 3 mm in length i.e. a fraction of one core, far below the amount of tissue used for histology assessment. As molecular studies become routine and cores of greater length are available, the relationship between biopsy size and stability of the molecular scores should continue to be explored, particularly for TCMR, which is sometimes patchy in histologic assessment. Nevertheless, the ability of the MMDx system to assess small tissue samples is an advantage over histology, provided that histologic assessment of glomerular diseases (e.g. recurrent glomerulonephritis) is not required. Having said that, the expression of mRNA for a number of important glomerular transcripts in the microarray readout raises the possibility that conclusions about glomerular diseases may eventually be inferred directly from core biopsies without micro dissecting the glomeruli.

In conclusion, *NPHS2* can be used to estimate proportion cortex for MMDx purposes. *NPHS2* is subject to small loss of expression in disease but the cortex-medulla differences override these. Rejection and injury do occur in medulla and can be read molecularly, and the proportion cortex does not seem to have a major influence on the ability to read rejection and injury. The ability of MMDx to read small amounts of tissue, regardless of cortex or medulla content, could make biopsies safer if the clinician believes that histologic assessment is not necessary.

**3.8 TABLES**

**CHAPTER 3: Assessing applications of microarrays in renal transplant diagnostics – tissue heterogeneity**

**Table 3.1** Demographics of the patients and the biopsy sets

	<b>Biological replicates (N = 37 pairs from 37 patients)</b>	<b>Cortex - Medulla (N = 26 pairs + 4 unpaired from 26 recipients)</b>	<b>Reference set (N = 1208, from 1045 patients)</b>
<b>Patient characteristics</b>			
Mean recipient age (years)	52 (22-77) (1 NA)	53 (29-71) (1 NA)	50 (9 - 91)
Recipient Gender (% male)	69% (2 NA)	58%	53%
<b>Primary Disease</b>			
Diabetic nephropathy	7	2	180
Glomerulonephritis/vasculitis	5	5	47
Interstitial nephritis/pyelonephritis	5	4	25
Polycystic kidney disease	2	3	120
Others	10	3	788
Unknown etiology	5	9	48
Mean donor age (years)	50	53	43
Donor gender (% male)	53%	58% (3 NA)	48% (347 NA or blank)
Donor type (% deceased donor transplant)	65%	77%	65%
<b>Biopsy characteristics</b>			
Median and mean time from transplant to biopsy	1959 (905) days	944 (62) days	592 (1553) days
Range	26.2 years	17.1 years	31.4 years
Primary non-function	2	8	10
Rapid deterioration of function	8	2	211
Slow deterioration of function	5	9	217
Stable impaired graft function	0	0	79
Investigate proteinuria	4	3	185
Follow-up from previous biopsy	5	3	91
Others	13	5	415
<b>Conventional biopsy diagnosis</b>			
ABMR	12	10	215
ABMR suspicious	0	4	24
AKI	0	14	96
Borderline	1	2	109
Interstitial fibrosis and tubular atrophy (IFTA)	4	6	145
Normal/NOMOA (No major abnormalities, No rejection)	6	14	274
TCMR	3	2	87
Mixed	1	2	41
Other or N/A	10	2	217
<b>ABMR, antibody-mediated rejection; AKI, acute kidney injury; NA, not available; TCMR, T cell-mediated rejection</b>			

**CHAPTER 3: Assessing applications of microarrays in renal transplant diagnostics – tissue heterogeneity**

**Table 3.2** Top 30 probe sets regarding fold change between cortex and medulla using a paired t-test

P Value	Adjusted P Value	Gene	Name	Cortex	Medulla	Fold	PBTs
1.85E-11	6.2E-07	NPHS2	nephrosis 2, idiopathic, steroid-resistant (podocin)	1221	43	28.62	KT1
2.62E-11	6.2E-07	FGF1	fibroblast growth factor 1 (acidic)	470	95	4.95	KT1
3.76E-11	6.2E-07	ST6GALNAC3	ST6 (alpha-N-acetyl-neuraminy-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 3	50	26	1.9	
9.4E-11	1.16E-06	FGF1	fibroblast growth factor 1 (acidic)	69	40	1.72	KT1
2.87E-10	2.4E-06	NPHS1	nephrosis 1, congenital, Finnish type (nephrin)	106	22	4.9	
2.9E-10	2.4E-06	ZDHHC14	zinc finger, DHHC-type containing 14	131	217	0.6	
4.7E-10	2.93E-06	NTNG1	netrin G1	52	21	2.47	
4.93E-10	2.93E-06	KLK7	kallikrein-related peptidase 7	115	47	2.43	HT1
5.33E-10	2.93E-06	KLK6	kallikrein-related peptidase 6	100	40	2.47	
6.18E-10	3.06E-06	CLIC5	chloride intracellular channel 5	95	28	3.35	
7.13E-10	3.21E-06	PTPRO	protein tyrosine phosphatase, receptor type, O	246	79	3.1	
8.27E-10	3.41E-06	MME	membrane metallo-endopeptidase	275	51	5.42	KT1
1.18E-09	3.96E-06	PLA2R1	phospholipase A2 receptor 1, 180kDa	226	106	2.14	
1.28E-09	3.96E-06	PLCE1	phospholipase C, epsilon 1	101	40	2.5	
1.28E-09	3.96E-06	PTPRO	protein tyrosine phosphatase, receptor type, O	95	33	2.85	
1.29E-09	3.96E-06	CYP3A5	cytochrome P450, family 3, subfamily A, polypeptide 5	135	27	5.09	
1.41E-09	3.96E-06	CYP3A5	cytochrome P450, family 3, subfamily A, polypeptide 5	307	79	3.9	
1.44E-09	3.96E-06	NOX4	NADPH oxidase 4	587	78	7.48	ENDAT
1.71E-09	4.24E-06	CLIC5	chloride intracellular channel 5	899	146	6.16	
1.71E-09	4.24E-06	TNNT2	troponin T type 2 (cardiac)	81	24	3.34	
1.83E-09	4.3E-06	PTPRO	protein tyrosine phosphatase, receptor type, O	296	57	5.2	
2.01E-09	4.3E-06	NOX4	NADPH oxidase 4	979	130	7.54	ENDAT
2.02E-09	4.3E-06	PLCXD3	phosphatidylinositol-specific phospholipase C, X domain containing 3	181	61	2.97	HT1
2.09E-09	4.3E-06	PODXL	podocalyxin-like	2064	884	2.33	ENDAT
2.43E-09	4.73E-06	HPGD	hydroxyprostaglandin dehydrogenase 15-(NAD)	378	102	3.72	KT1
2.48E-09	4.73E-06	NOX4	NADPH oxidase 4	1193	162	7.35	ENDAT
2.78E-09	5.1E-06	HPGD	hydroxyprostaglandin dehydrogenase 15-(NAD)	303	75	4.06	KT1
2.89E-09	5.11E-06	NOX4	NADPH oxidase 4	1588	221	7.18	ENDAT
3.23E-09	5.16E-06	NTNG1	netrin G1	43	18	2.38	
3.23E-09	5.16E-06	REN	renin	389	38	10.2	KT1

**PBTs - pathogenesis-based transcript sets; KT1 - kidney transcripts; ENDAT - endothelial transcripts; fdr - false discovery rate.**

**CHAPTER 3: Assessing applications of microarrays in renal transplant diagnostics – tissue heterogeneity**

**Table 3.3** Chi square test statistical values for Figures 3.6 (A-I) and 3.7 (A-C)<sup>1</sup>

Figure #	Description	Distribution of molecular scores in quadrants				p value
		Top right quadrant (C+M+)	Top left quadrant (C+M-)	Bottom right quadrant (C-M+)	Bottom left quadrant (C-M-)	
Figure 6A	ABMR – Bio replicates	25	39	5	5	0.51
Figure 6B	ABMR – Cortex/Medulla	9	25	11	11	0.07
Figure 6C	ABMR – Reference Set	354	745	30	61	0.88
Figure 6D	TCMR – Bio replicates	13	51	2	8	0.98
Figure 6E	TCMR – Cortex/Medulla	4	30	5	17	0.28
Figure 6F	TCMR – Reference Set	166	933	6	85	0.03
Figure 6G	Rejection – Bio replicates	39	25	7	3	0.58
Figure 6H	Rejection – Cortex/Medulla	10	24	12	10	0.06
Figure 6I	Rejection – Reference Set	534	565	38	53	0.21
Figure 7A	IRRATs – Cortex/Medulla	21	13	14	8	0.89
Figure 7B	IRRATs – Bio replicates	39	25	6	4	0.96
Figure 7C	IRRATs – Reference Set	456	643	37	54	0.88

ABMR, antibody-mediated rejection; IRRATs, acute kidney injury transcripts; TCMR, T cell-mediated rejection

<sup>1</sup> The cutoffs in figure 1 divide the data into high cortex (“C+”) versus low cortex (“C-”) and molecular score positive (“M+”) or negative (“M-”), giving four quadrants: C+M+, C+M-, C-M+, and C-M-.

**CHAPTER 3: Assessing applications of microarrays in renal transplant diagnostics – tissue heterogeneity**

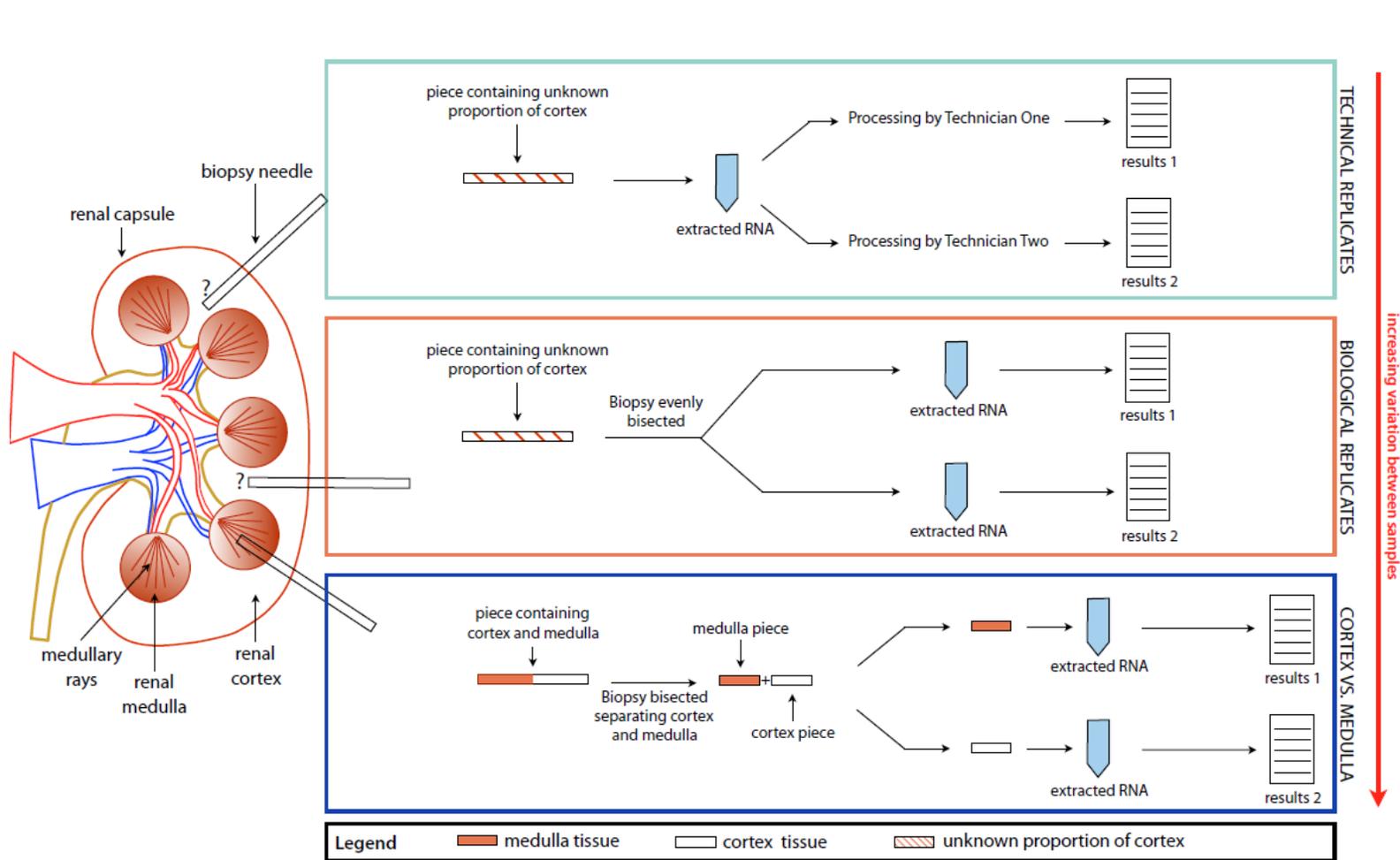
**Table 3.4** Consistency in the molecular diagnosis between paired cortex-medulla samples and in biological replicates.

<b>26 pairs of cortex and medulla MMDx reports</b>		
<b>Number of cortex/medulla pairs</b>	<b>Agreement/Disagreement</b>	<b>Description</b>
23	Agreement	Comparable diagnosis with no substantive change
3	Disagreement	Disagreement on a major diagnostic point (i.e. 'No TCMR' versus 'TCMR')
<b>37 pairs of biological replicate MMDx reports</b>		
<b>Number of Biological Replicate pairs</b>	<b>Agreement/Disagreement</b>	<b>Description</b>
32	Agreement	Comparable diagnosis with no substantive change
4	Disagreement	Disagreement on a major diagnostic point (i.e. 'No TCMR' versus 'TCMR')

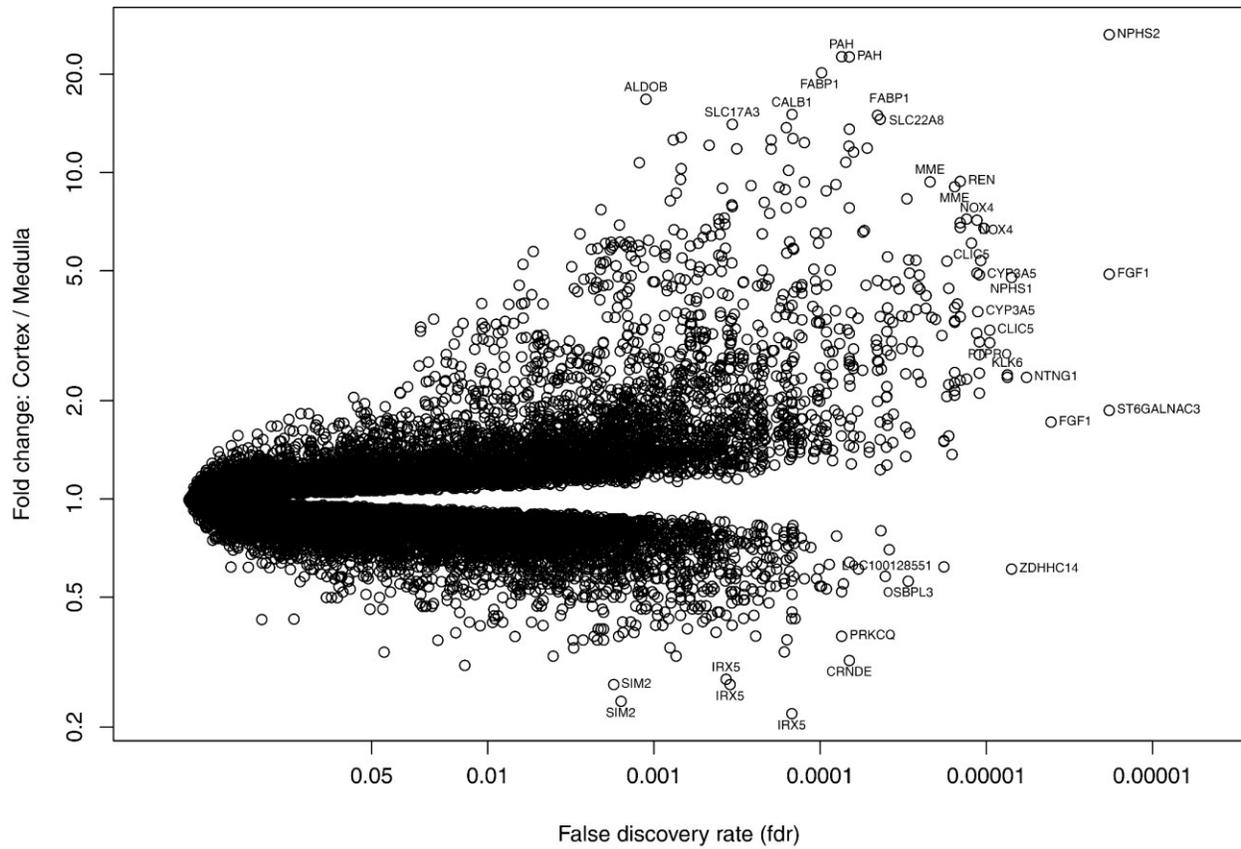
ABMR, antibody-mediated rejection; MMDx, molecular microscope diagnostic system; TCMR, T cell-mediated rejection.  
 1Agreement defined as either perfect agreement (presence/absence of type of rejection and agreement in scale, i.e. "severe ABMR, no TCMR" in both diagnostic signouts), or agreement on presence/absence of rejection with difference in scale (i.e. "severe ABMR, No TCMR" and "moderate ABMR, no TCMR").  
 2Disagreement defined as a pair of samples with one having rejection and the other lacking rejection (i.e. "no ABMR, moderate TCMR" and "moderate ABMR, moderate TCMR").  
 3Ambiguous samples included those with histological screen failures, ambiguous histology and molecular reads, or samples too damaged/lacking in quality for a proper diagnostic read. Agreement could not be determined in these cases as no distinct diagnostic prediction was possible.

**3.9 FIGURES**

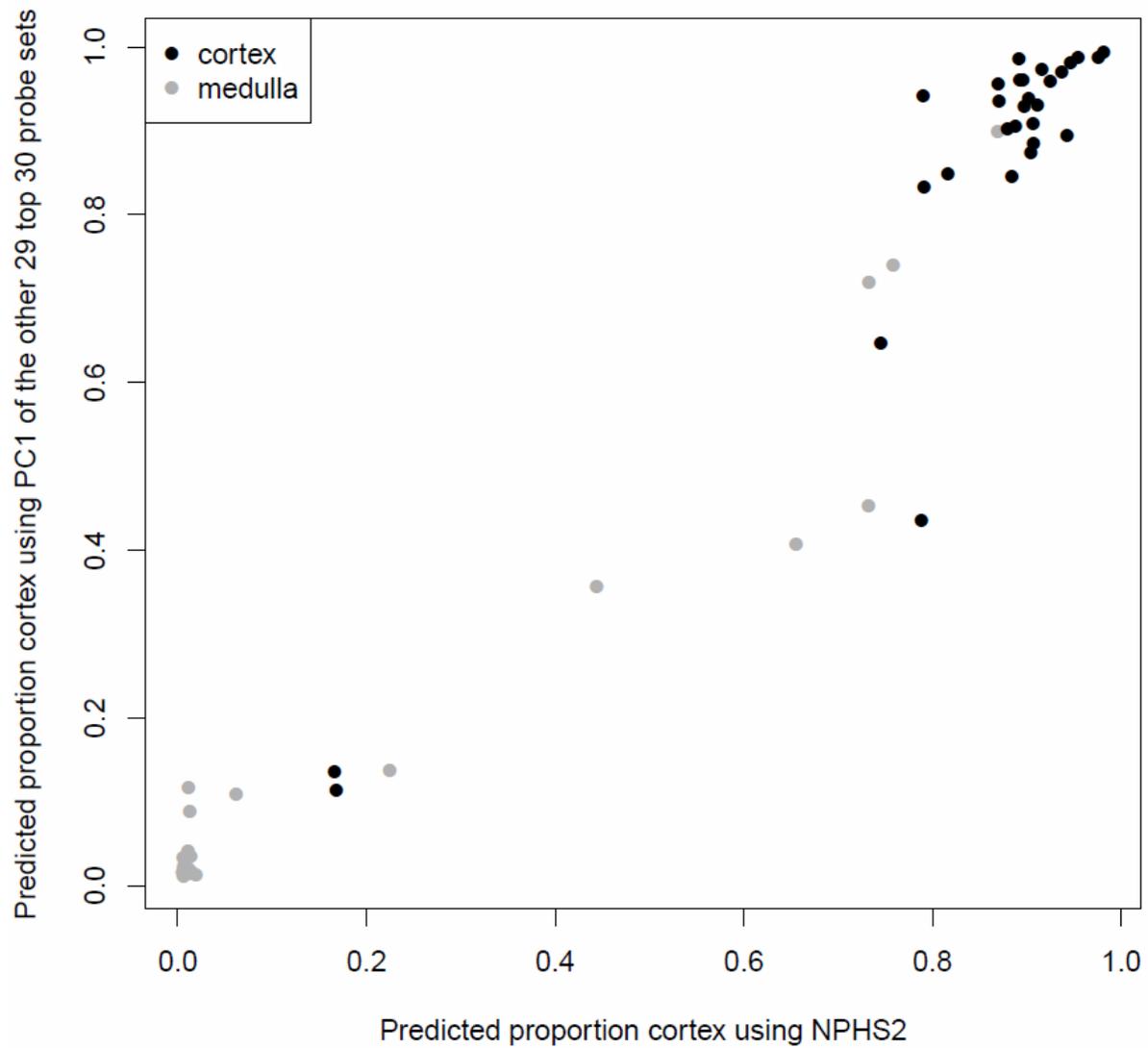
**CHAPTER 3:** Assessing applications of microarrays in renal transplant diagnostics – tissue heterogeneity



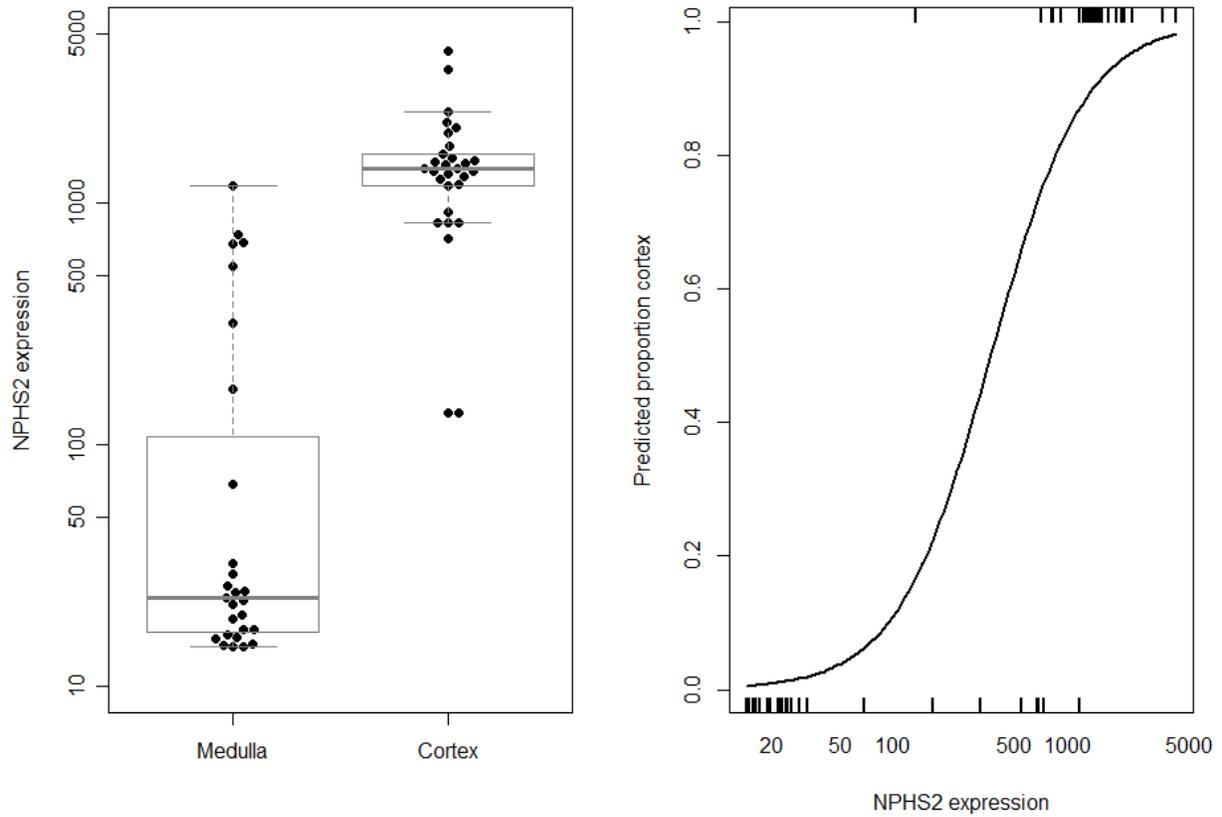
**Figure 3.1** Diagram showing the sampling strategy for the technical replicate pairs, biological replicate pairs, and cortex-medulla pairs.



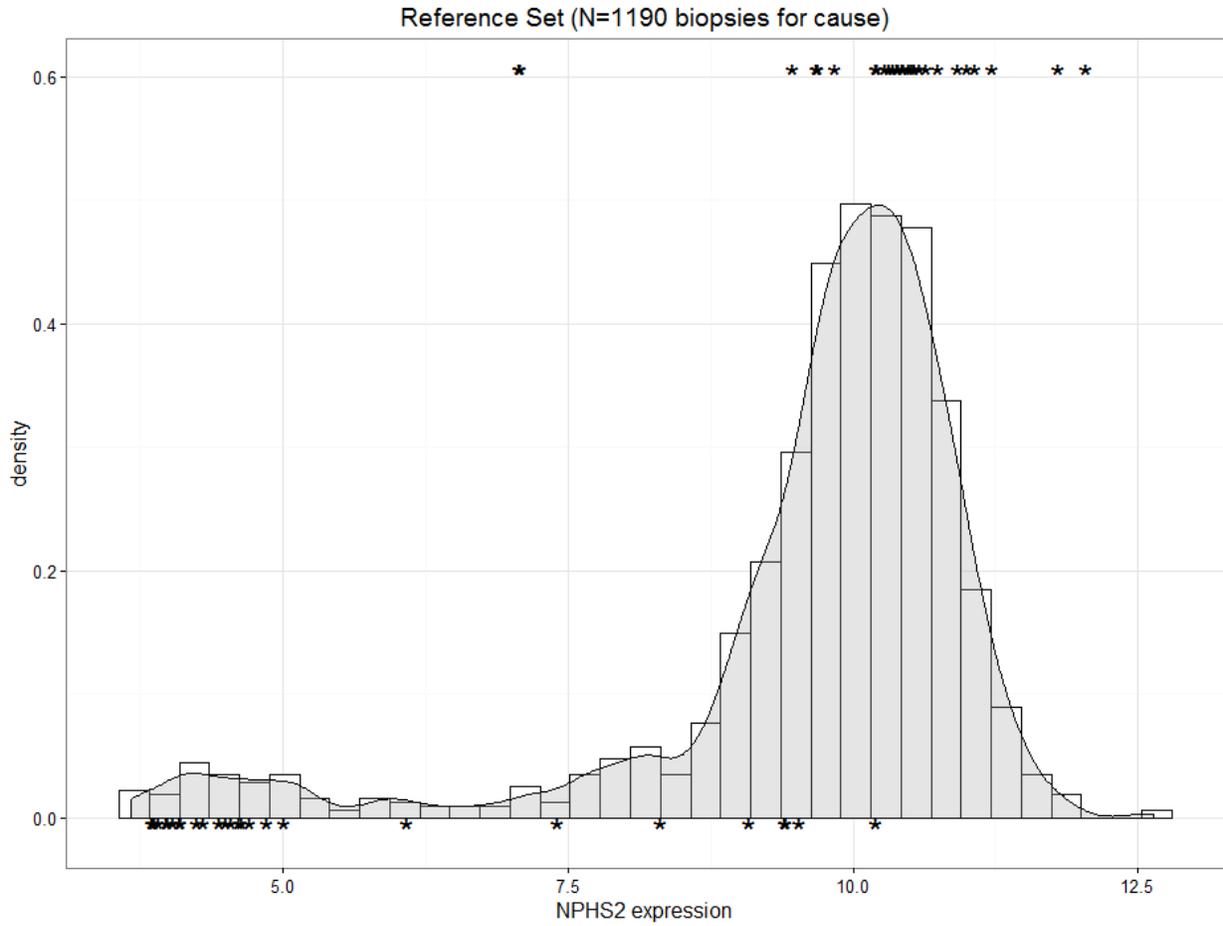
**Figure 3.2** Volcano plot of fold change between cortex and medulla versus negative log of adjusted p-value with false discovery rate. NPHS2 had the highest association and fold change between cortex and medulla of 55,000 probe sets. A selection of highly significant probe sets distinguishing cortex from medulla is labeled.



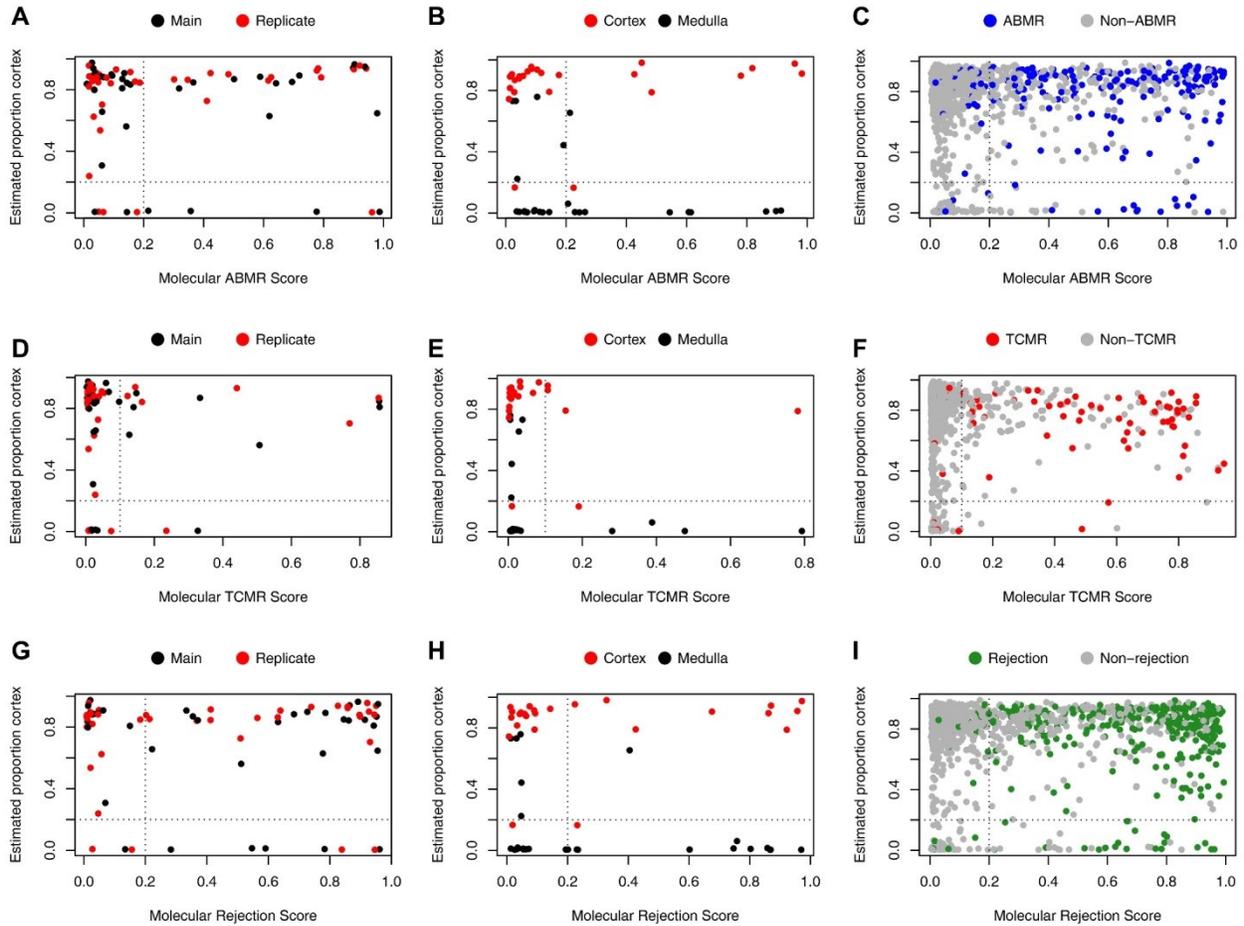
**Figure 3.3 Predicted proportion of cortex in a sample histologically called cortex or medulla.** Predictions were made using either the principal component 1 score based on 29 of the top 30 probe sets (excluding NPHS2) (y-axis) or using NPHS2 expression alone (x-axis).



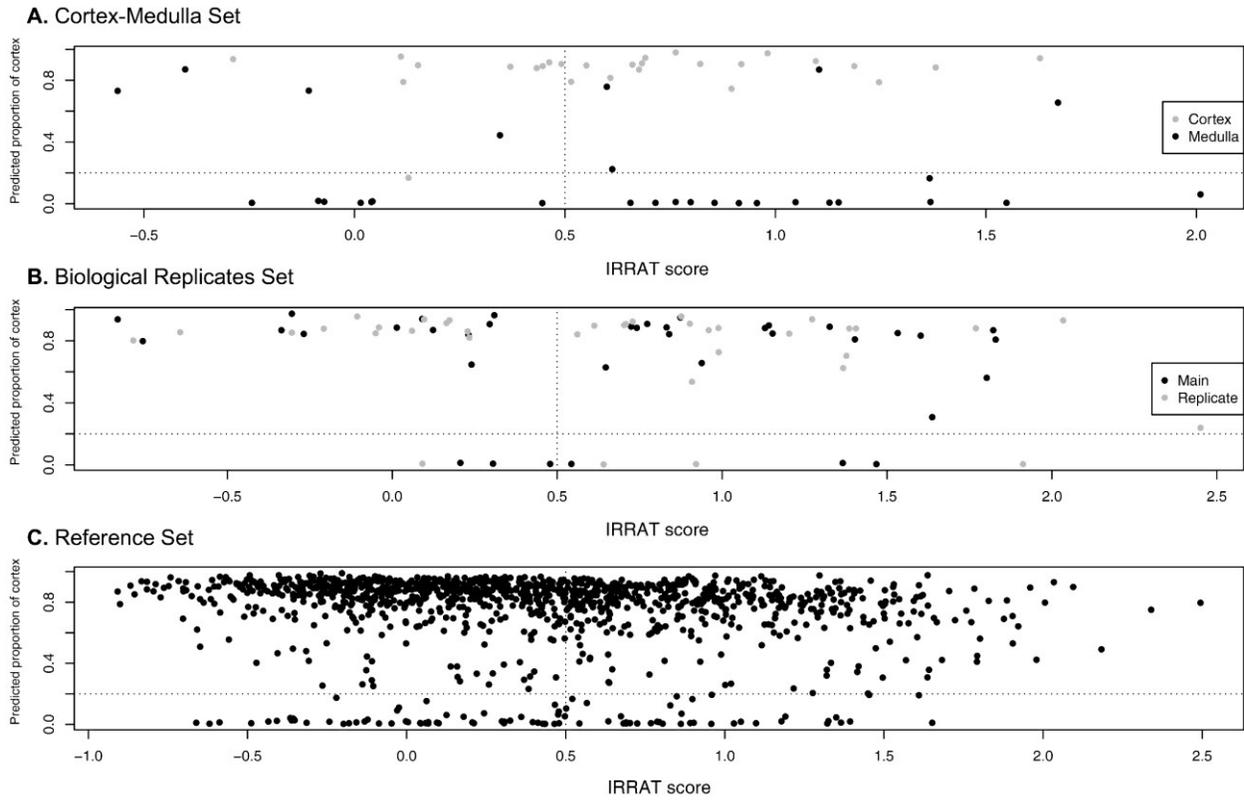
**Figure 3.4** Boxplot showing log of *NPHS2* expression in medulla and cortex samples as established by histology (A) and the predicted proportion cortex distribution across all samples (B). Box shows the interquartile range, horizontal bar—median and whiskers— 1.5 x standard deviation.



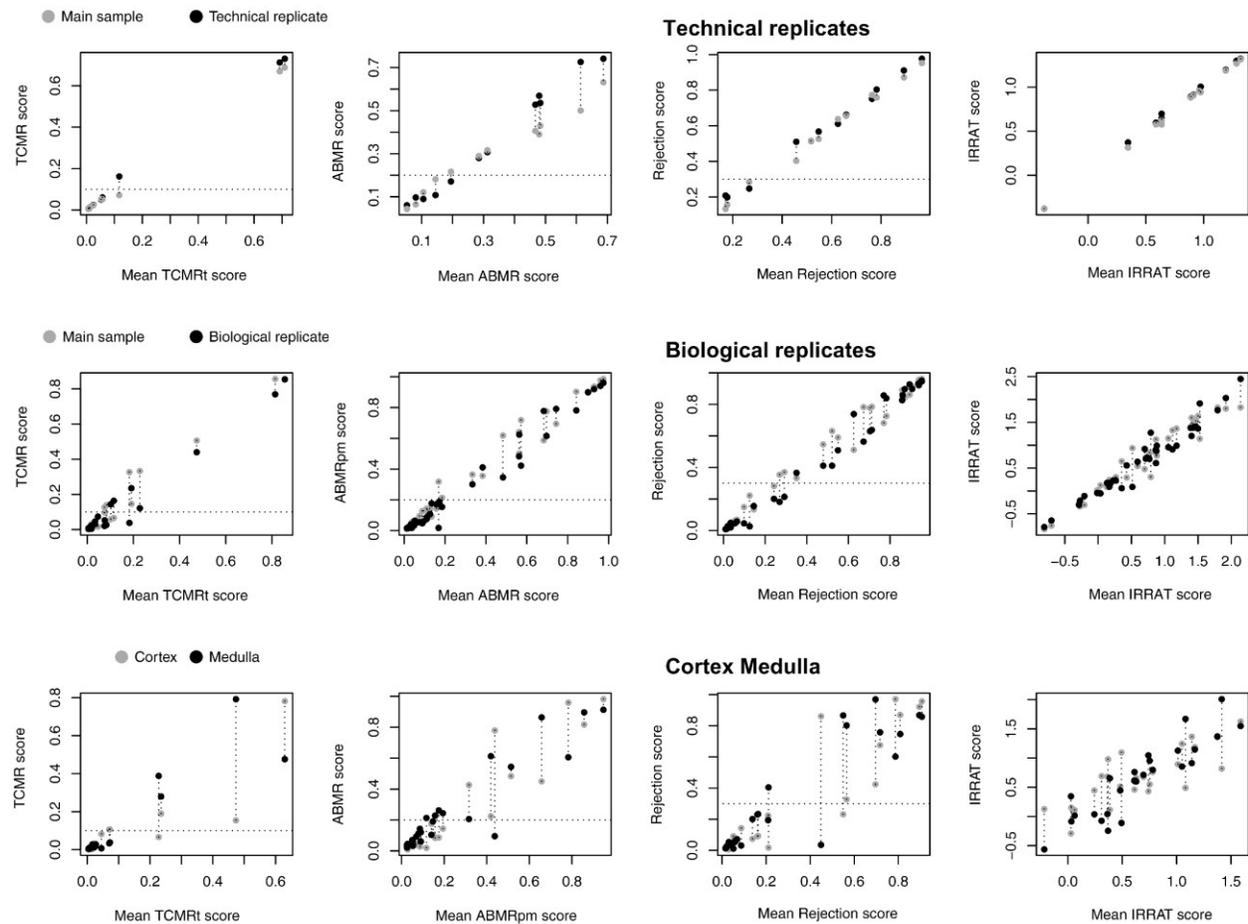
**Figure 3.5** Density plot of *NPHS2* expression in 1190 nonbisected biopsy cores. Black symbols show the distribution of *NPHS2* expression values in cortex and medulla divided pair samples, respectively.



**Figure 3.6 Scatter plots with predicted proportion cortex (y-axis) versus molecular ABMR, TCMR, and rejection scores in the biological replicate set, the cortex and medulla set, and the reference set (x-axis).** Vertical dotted line indicates the positive/ negative cutoffs for the molecular scores; horizontal dotted line indicates the 0.2/0.8 split for the proportion of cortex. A) ABMR - biological replicates, B) ABMR - cortex/medulla, C) ABMR - reference set, D) TCMR - biological replicates, E) TCMR - cortex/medulla, F) TCMR - reference set, G) Rejection - biological replicates, H) Rejection - cortex/medulla, I) Rejection - reference set. Refer to Table 3.3 for further details. ABMR, antibody-mediated rejection; TCMR, T cell-mediated rejection.



**Figure 3.7 Investigating the relationship between injury (IRRAT score) and proportion of cortex.** Vertical dotted line indicates the arbitrary positive/negative cutoff for the molecular IRRAT scores; horizontal dotted line indicates an arbitrary 0.2/0.8 split for the proportion of cortex. A) IRRATs - cortex/medulla, B) IRRATs - biological replicates, C) IRRATs - reference set. Refer to Table 3.3 for further details. IRRAT, acute kidney injury transcripts.



**Figure 3.8** Reproducibility plots of the molecular scores of rejection (TCMR, ABMR, all rejection) and acute kidney injury (IRRAT) in the technical and biological replicates and in the cortex–medulla divided pairs. Dotted horizontal line indicates the positive/negative cutoffs for the molecular scores. The y-axes are the scores for the two samples compared for the TCMR, ABMR, and Rejection classifiers (a number between 0 and 1.0) or the geometric mean of the expression of the AKI transcripts (IRRATs). The x-axis is the mean of classifier or IRRAT scores. The TCMR score is the TCMRt classifier score. The ABMR score is the ABMRpm classifier score. The rejection score is the rejection classifier score. The IRRAT score is used to represent kidney injury. ABMR, antibody-mediated rejection; AKI, acute kidney injury; IRRAT, acute kidney injury transcripts; TCMR, T cell-mediated rejection.

## **CHAPTER 4**

### **DISCREPANCY ANALYSIS BETWEEN MOLECULAR AND HISTOLOGIC ASSESSMENTS OF REJECTION IN RENAL TRANSPLANTATION**

**This chapter has been published**

Madill-Thomsen KS, Perkowska-Ptasinska A, Bohmig G, Eskandary F, Einecke G, Gupta G, Halloran PF, et al. Discrepancy analysis comparing molecular and histology diagnoses in kidney transplant biopsies. Am J Transplant. 2019 Dec; (epub ahead of print). Used with permission of Wiley Publishing.

## CHAPTER 4: Discrepancy analysis comparing molecular and histologic diagnoses

### 4.1 OVERVIEW AND HYPOTHESIS

In kidney transplants requiring biopsies, precision (reproducibility), accuracy (representation of the true disease states), and standardization across centers and countries are critical for the diagnosis of rejection.<sup>6,68–70</sup> Unfortunately, histology diagnoses are associated with high levels of interobserver disagreement, as summarized by Furness et al: “...international variation in histologic grading is large, under-recognized, difficult to improve, and almost certainly of major clinical relevance. Urgent steps are needed to improve this area of practice”.<sup>2</sup> Histology also has inherent limitations such as the inability to assess recent injury.<sup>6–8</sup> An additional concern in the Banff consensus is “data drift”: rules originally formulated in 1991 are now being applied in a different case mix that changes prior probabilities. These issues are compounded by the challenges faced by all diagnostic systems such as balancing the risks of under- versus over-diagnosis.<sup>6,9</sup>

Precision, accuracy, and standardization in kidney transplant diagnostics can be improved by incorporating central molecular biopsy assessments that measure rejection and injury changes.<sup>224</sup> The high technical and biological reproducibility of the molecular scores assigned by MMDx were previously demonstrated (see Chapter 3).<sup>35,64,71,211</sup> In MMDx, ensembles of algorithms provide a stable probabilistic estimate of diagnoses based exclusively on molecular findings and are more likely to be correct than histology as previously discussed.<sup>6</sup> MMDx uses continuous scores rather than categories (grades or binary classes) and employs machine learning-derived algorithms, which can overcome errors.<sup>225,226</sup> These advantages are significant for the diagnoses of rejection and injury in kidney transplants, although MMDx cannot currently make a diagnosis of specific glomerular diseases such as glomerulonephritis and diabetic nephropathy that are defined histologically.

The emergence of an optimized central MMDx system<sup>69</sup> creates an opportunity to compare MMDx to SOC histology in a discrepancy analysis, a valuable method for addressing guiding improvements in diagnostic systems without assuming that either system is always correct.<sup>227–230</sup> It was previously noted that discrepancies between MMDx and SOC histologic diagnosis are common in certain scenarios e.g. arteritis,<sup>34</sup> tubulitis,<sup>35–38</sup> ambiguous ABMR cases,<sup>39,40</sup> and BK nephropathy,<sup>35,72–74</sup> but it remains unclear how many discrepancies are explained by these scenarios. The present study aimed to

## **CHAPTER 4: Discrepancy analysis comparing molecular and histologic diagnoses**

perform discrepancy analysis between the official MMDx sign-out ('Expert1MMDx')<sup>6</sup> and SOC histology, and to provide possible guidance to clinicians about how to utilize these platforms. The main research question of these analyses focused on the relationships with SOC histology, therefore central reading by a single pathologist was not considered. Moreover, there is no evidence that such central review is more accurate than the local SOC diagnoses, which can access more data from the clinical environment.

### **4.2 BIOPSY POPULATION**

These analyses studied 1679 prospectively collected indication biopsies from international centers, obtained with consent under local IRB approved protocols during the INTERCOMEX study (**Table 4.1, Table 4.2**). This study population was previously described in another publication.<sup>6</sup> Previous versions of MMDx systems used 1208/1679 biopsies to train classifiers,<sup>66</sup> but classifiers were regenerated in the 1679 population for these analyses.<sup>6</sup> All biopsies and biopsy processing methods have been previously described in published literature.<sup>6</sup> Histology diagnoses were assigned per the SOC by the pathologists at each center following Banff guidelines and were collected for this study as reported by each center without modification or exclusion, with the sole exception of the reclassification of pre-2013 biopsies with ABMR using the new C4d-negative ABMR class.<sup>42</sup> CEL files are available on the Gene Expression Omnibus website (GSE124203).

Histologic rejection (ABMR, TCMR, Mixed, and pABMR e.g. ABMR suspicious or TG) was diagnosed in 612/1679 biopsies (36%), with 8% borderline or pTCMR (128/1679). DSA assessed by the center SOC protocols was positive in 573 biopsies (34%) (**Table 4.3**). BK virus nephropathy was diagnosed by histology in 55 biopsies (three of which were diagnosed as 'TCMR/BK' by the center, **Table 4.3**).

### **4.3 STATISTICAL AND CLASSIFICATION METHODS FOR DISCREPANCY ANALYSIS**

#### **4.3.1 Computational methods for analyses**

## CHAPTER 4: Discrepancy analysis comparing molecular and histologic diagnoses

All analyses were done using functions in the base R system (version 3.5.1)<sup>185</sup>. Unweighted kappa values to assess interobserver agreement in previously flagged problematic scenarios (negative or ambiguous DSA, v>0 lesion scores, BK virus, ci 2/3 lesion scores, and t 2/3 lesion scores, all discussed and explored in prior publications) were calculated using the 'cohen.kappa' function of the 'psych' package<sup>231</sup>. Intraclass correlation coefficients (ICCs) were calculated using the 'ICC' function of the 'psych' package<sup>231</sup>. Bar plots were generated using the 'barplot' function in base R.

The PCA is generated using the 'FactoMineR' R package<sup>199</sup> and the following seven molecular classifier scores as input: the classifiers for predicting diagnoses of ABMR and TCMR, as well as classifiers for predicting high grades of the g-, ptc-, cg-, i-, and t-lesions.

### 4.3.2 Converting platforms to six diagnostic classes

To allow comparisons between diagnoses assigned by histology or MMDx, output from both platforms was converted into six classes (**Table 4.4**). All diagnoses that were not related to rejection were converted to 'NR'. ABMR and TCMR diagnoses were left as-is and assigned to those respective categories. 'ABMR suspicious' and 'TG' diagnoses were combined into the 'pABMR' category, as has been done in previous publications. 'Borderline' diagnoses were assigned to the 'pTCMR' category. Diagnoses of both ABMR and TCMR by histology were assigned to the 'Mixed' category.

### 4.3.3 'Clear' versus 'boundary' discrepancy classifications

Discrepancies between assigned diagnoses (either within MMDx or histology-MMDx) were classified as 'clear' if involving unambiguous classes as e.g. ABMR versus NR. Discrepancies were designated as 'boundary' if involving boundaries or cutoffs e.g. ABMR versus pABMR. Those discrepancies involving mixed rejection diagnoses are impacted by multiple boundaries and were described as 'mixed', e.g. between mixed and ABMR and between mixed and TCMR.

## CHAPTER 4: Discrepancy analysis comparing molecular and histologic diagnoses

### 4.4 ASSESSING DISCREPANCIES WITHIN MMDx

MMDx assessment of rejection outputs the ensemble of classifiers as a point in PCA space, expresses the individual measurements on the report, then is signed out by an expert.<sup>6</sup> MMDx reports also contain a set of classifier scores (e.g. ABMRpm, TCMRt, Rejection Score classifier) and AA results showing the proportion of the biopsy belonging to each archetype cluster. A diagnosis of molecular rejection is determined by visualizing the biopsy in relationship to the reference set and assessing both its position in data space in tandem with its molecular scores with respect to assigned cutoffs (e.g. TCMRt>0.1, ABMRpm>0.2). Variation within MMDx diagnoses was assumed to be attributable to the interpretation of the set guidelines around diagnostic class boundaries between expert observers, i.e. between TCMR and Mixed Rejection 'Mixed', or between pABMR and ABMR. The stability of various scores and features on the report was also assessed to determine if this was a relevant factor contributing to intra-MMDx discrepancies.

#### 4.4.1 Reproducibility of raw MMDx scores

Reproducibility of the raw scores was assessed by comparisons of the MMDx data from technical replicate pairs. Plots of the scores within each pair demonstrated that all pairs had consistently similar scores for PC1, PC2, PC3 (**Figure 4.1**) and therefore high reproducibility on the report figure (which plots the unknown sample in PCA space relative to the reference set). Reproducibility of the report scores was assessed by comparing MMDx reports from replicate pairs run on the same biopsy. All pairs had similar scores for PC1, PC2, and PC3 with correlations of approximately 0.99 (**Figure 1A-D**), indicating high reproducibility of the report figure which shows the relationship of the new biopsy to the known reference set based on its PC1, PC2, and PC3 scores. Correlations were also high (ICCs >0.90) for other major MMDx scores, including archetype scores and classifier scores (**Figure 1D**). The automated MMDx output is almost identical between technical replicates.

#### 4.4.2 Assignment of alternative MMDx sign-out comments

## CHAPTER 4: Discrepancy analysis comparing molecular and histologic diagnoses

To assess discrepancies between alternative MMDx assessments, we compared the Expert1MMDx diagnoses<sup>66</sup> to automated MMDx diagnoses ('AutoMMDx') assigned by the previously described random forest-derived algorithm predicting Expert1MMDx, and to alternative diagnoses assigned by another expert observer 'Expert2MMDx'.<sup>6</sup>

Official MMDx recorded diagnoses ('Expert1MMDx') were all signed out by one observer (PFH) based on previously described ensembles of estimates.<sup>6,35,66</sup> Expert1MMDx and Expert 2 followed published guidelines for predicting ABMR, TCMR, Mixed, and No Rejection.<sup>6</sup> Alternative Expert2 signouts were assigned independently and without knowledge of the original Expert1MMDx diagnosis.

AutoMMDx uses all variables on the MMDx report as initial input, predicts ABMR and TCMR with probabilities between 0.0 and 1.0, and assigns these probabilities to categories using cutoffs: if the ABMR/TCMR probability is  $\leq 0.4$ , non-ABMR or non-TCMR; if  $> 0.4$  and  $\leq 0.8$ , pABMR/pTCMR; if  $> 0.8$  the biopsy is called ABMR/TCMR. Expert1MMDx diagnoses were also compared to a second expert observer (KMT, 'Expert2MMDx').<sup>6</sup> AutoMMDx was developed and described in a separate published study.<sup>6</sup>

### 4.4.3 Discrepancies within MMDx sign-out comments

MMDx reports provide a sign-out comment interpreting the rejection category for the biopsy. This categorical interpretation is assigned by an expert or by the automated MMDx random forest model ('AutoMMDx').<sup>6</sup> We studied discrepancies between alternative MMDx sign-outs: those assigned by Expert 1<sup>6,66</sup> versus those assigned independently by another expert, Expert 2; and those assigned by Expert 1 versus AutoMMDx. Experts 1 and 2, following the same guidelines but allowed to consider all scores on the report<sup>6</sup>, showed 108 (6%) discrepancies, of which 1.4% involved clear categories (**Table 4.5**). The most frequent clear discrepancies were 11 biopsies Expert 1 called ABMR and Expert 2 called NR (**Table 4.6**). Expert 1 versus AutoMMDx had 149 discrepancies (9%) (**Table 4.7**), of which 1.6% were clear, most frequent being 10 biopsies called TCMR by Expert 1 and no rejection by AutoMMDx (**Table 4.6**). Overall, the sign-out comments between either Expert1 and Expert2, or Expert 1 and AutoMMDx were highly

## **CHAPTER 4: Discrepancy analysis comparing molecular and histologic diagnoses**

correlated, with all discrepancies located around boundaries. Discrepancies between MMDx sign-out comments are summarized in **Table 4.6**.

### **4.5 DISCREPANCIES BETWEEN MMDx SIGN-OUTS AND HISTOLOGIC DIAGNOSIS**

#### **4.5.1 Assessing discrepancies as a proportion of histology categories**

Histology and MMDx (Expert 1) sign-out categories disagreed in 624/1679 biopsies (37%, **Table 4.8**). As a percentage of histology categories (columns in **Table 4.8**), discrepancies were numerous in the definite categories: 22% of ABMR, 61% of TCMR and 26% of no rejection. Discrepancies were much more frequent in the ambiguous categories: 97% of ABMR suspicious; 98% of TG; and 98% of borderline. Note that although based on tubulitis, borderline was more frequently ABMR (24/128) or no rejection (83/128) than TCMR (9/128) and was seldom recognized by MMDx as weak TCMR (pTCMR 3/129). Mixed histology diagnoses had 61% discrepancy with MMDx.

Biopsies called 'BK virus positive' by histology were frequently called TCMR by MMDx (15/52 cases) and occasionally ABMR (4/52 cases, **Table 4.9**).

#### **4.5.2 Assessing discrepancies as a proportion of MMDx sign-out categories**

Among MMDx diagnoses (the rows in **Table 4.8**), the most numerous discrepancies with histology were 249/509 (49%) MMDx ABMR biopsies not called ABMR by histology. Moreover, while MMDx and histology called similar numbers of TCMR (123 and 139, respectively), there was extensive disagreement: 84/139 histologic TCMR biopsies were not called TCMR by MMDx, and 68/123 of MMDx TCMR biopsies were not called TCMR by histology. Boundary MMDx biopsies – pABMR and pTCMR – were less frequent than in histology, and seldom agreed with histology pABMR (ABMR suspicious or TG) or pTCMR (borderline).

### **4.6 VISUALIZING HISTOLOGY-MMDx DISCREPANCIES**

## CHAPTER 4: Discrepancy analysis comparing molecular and histologic diagnoses

We plotted the proportion of histologic biopsy diagnoses per MMDx diagnostic category (**Figure 2A**).<sup>6,66</sup> Although some MMDx categories had high agreement with histology (e.g. MMDx-NR), several unambiguous categories were frequently discrepant (e.g. MMDx-ABMR or MMDx-TCMR, where a large percentage of biopsies were called NR by MMDx).

In ambiguous histology categories (e.g. pABMR and pTCMR, **Figure 2B**), MMDx typically gave an unambiguous sign-out comment of ABMR, TCMR, or no rejection. The results illustrate that MMDx usually gave unambiguous diagnoses in biopsies ambiguous by histology.

### 4.7 COMPARING CLEAR, BOUNDARY, AND MIXED DISCREPANCIES: HISTOLOGY-MMDx VERSUS INTRA-MMDx

**Table 4.11** shows that clear discrepancies were much more frequent for histology-MMDx (19%) than intra-MMDx (1.4%-1.6%). The most common clear histology-MMDx discrepancies were 131 biopsies that MMDx signed out as ABMR and that histology considered no rejection (**Table 4.11**). Boundary and mixed histology-MMDx discrepancies were also more frequent than intra-MMDx discrepancies.

Histology-MMDx discrepancies were distributed diffusely in the population and included many clear-cut cases by MMDx. This finding indicated that many discrepancies occurred in biopsies that were unambiguous by MMDx and unrelated to interpretation around boundaries (**Figure 4.3A** and **4.3B**). In contrast, discrepancies within MMDx were limited to differences in boundary interpretations among observers (**Figure 4.3C-F**).

#### 4.7.1 Expressing overall agreement through kappa values

To quantify and compare agreement, we calculated unweighted kappa values (higher kappa values indicating more agreement) for histology-MMDx and intra-MMDx (**Table 4.11**). Kappa values within histology-MMDx were 0.41, in keeping with estimates for inter-pathologist variability.<sup>3</sup> Kappa values for intra-MMDx agreement were much higher at 0.85-0.89, indicating high agreement between sign-out categories.

#### **4.8 DISCREPANCIES IN PREVIOUSLY FLAGGED PROBLEMATIC SCENARIOS**

Five previously flagged problematic scenarios – ambiguous or negative DSA in ABMR, tubulitis lesions, v-lesions, positive BK virus, and fibrosis – were assessed for their relative contribution to the total discrepancies. They accounted for 195 of the 624 (31%) of the histology-MMDx discrepancies.

##### **4.8.1 Agreement in problematic scenarios described by kappa values**

We confirmed that these histology scenarios contributed to discrepancies by calculating the kappa values for histology-MMDx agreement compared to other biopsies (**Table 4.12**). Kappa values were lower (indicating more discrepancies) when DSA was negative or ambiguous versus positive; when BK was positive versus negative; when tubulitis was present versus absent; and when arteritis was present versus absent. Scarring (as represented by the ci lesion) slightly increased discrepancies.

Of interest, problematic histology biopsies were not associated with increased MMDx ambiguity as measured by kappa values (**Table 4.12**).

#### **4.9 INTERPRETATION OF THE RESULTS**

The present study addressed the agreement between MMDx report scores, between alternative MMDx sign-outs, and between MMDx and histology in a collection of unselected indication biopsies collected from multiple centers. We found 99% correlation between scores in the replicate MMDx analyses, but some variation between alternative MMDx sign-out interpretations in biopsies near boundaries. Between MMDx and histology we found 37% discrepancies, many in unambiguous cases unrelated to boundaries. Ambiguous histology categories - ABMR suspicious, TG, and borderline - showed very high discrepancies with MMDx, but histology TCMR and mixed rejection were also often discrepant. Many MMDx ABMR biopsies were not called ABMR by histology. Previously flagged problematic histology scenarios - ambiguous or negative DSA in ABMR-related cases, BK nephritis, v>0, t>1, or ci>1 lesion scores - accounted for 31% of histology-MMDx discrepancies. However, these biopsies

## CHAPTER 4: Discrepancy analysis comparing molecular and histologic diagnoses

did not associate with increased discrepancies within MMDx interpretations, indicating that MMDx can offer clarity in histologically challenging situations. The results provide an overview of the relative utility of adding MMDx to various histology diagnoses, particularly when histology is ambiguous (e.g. borderline) or problematic.

When MMDx is ordered to clarify the histologic assessment, the present data describes the expected results for each histology category. For example, because MMDx is independent of DSA results, it can resolve ambiguous ABMR when the DSA is not clear or the histology conflicts with the clinical opinion. MMDx will also be useful in resolving ambiguities around mixed rejection, which presents major challenges in terms of therapeutic decisions. Many diagnostic lesions of rejection are somewhat non-specific: tubulitis can occur in ABMR (and in other glomerular diseases) as well as in TCMR and AKI;<sup>71</sup> ptc-lesions can occur in TCMR as well as ABMR and AKI;<sup>34</sup> and v-lesions occur in TCMR and ABMR but also in injury. Hence it is not surprising to see MMDx ABMR diagnosed as histology mixed or TCMR, and vice versa. Assigning arbitrary guidelines to nonspecific histologic features cannot eliminate these ambiguities, but MMDx can help resolve these scenarios. The presence of two diseases always presents challenges in diagnostic systems: the criteria for diagnosing a second disease can be obscured or complicated when the tissue is already disturbed by a previous disease.

MMDx clarifies histologic borderline tubulitis because it almost always assigns a clear diagnosis such as no rejection, ABMR, or TCMR. Borderline is often assumed to be mild TCMR or a precursor to clinical rejection with poorer outcomes in the literature.<sup>232-239</sup> However, the low frequency of MMDx pTCMR in histology borderline means that low-grade tubulitis is usually not mild TCMR. In 128 cases of borderline (i.e. tubulitis), many were actually ABMR or definite TCMR or no rejection. Note that these results in unselected indication biopsies would differ from studies in selected biopsies with a different case-mix. When the Banff classification was developed, most biopsies in the population were taken early after transplant and TCMR was extremely common, giving tubulitis a high prior probability of being TCMR. However, interpretation of lesions such as tubulitis, intimal arteritis, and ptc-lesions depends on the prior probabilities,<sup>2,3</sup> which have changed markedly since the 1990s. Comparisons of histology with MMDx can be useful for recalibrating the interpretation of these lesions in the contemporary case mix.

## CHAPTER 4: Discrepancy analysis comparing molecular and histologic diagnoses

Histology-MMDx discrepancies could be reduced by changing the histology convention for biopsies with BK virus by which pathologists do not diagnose TCMR in BK-positive kidney biopsies even when the lesions are diagnostic. MMDx identifies all biopsies with TCMR, whether or not they have BK diagnosed by histology. In addition, the mRNA isolated for MMDx can be tested for BK viral mRNA.<sup>73</sup> Biopsies with BK can be associated with TCMR-like histologic and molecular changes<sup>72</sup> because of changes induced by the virus or because of true rejection triggered by the management of BK, which involves reduction in immunosuppression. A major threat to graft survival in BK infection is rejection, either TCMR or later de novo DSA ABMR.<sup>74,240</sup> This may resolve when full immunosuppression is restored, or it may persist and require treatment.<sup>73,240</sup> The legitimate concern that T cell responses against the virus might mimic TCMR can be offset by a simple caveat in the diagnostic report such as: “TCMR-like changes are present that could represent true TCMR or viral effects. TCMR-like changes in the context of BK may not require treatment and may resolve when full immunosuppression is restored but increasing nephron damage or evolution to de novo DSA and ABMR”.

Sampling differences are sometimes invoked to explain histology-MMDx discrepancies, but they are unlikely to contribute greatly since both platforms are receiving adequate tissue for assessment. Biopsies ruled inadequate for either histology or MMDx were excluded from these analyses. MMDx samples are much smaller than histology samples (average 3mm), but molecular changes are more diffuse than histology changes. MMDx has high agreement in two pieces from the same kidney and can read medulla, and does not require a minimum number of arteries and glomeruli.<sup>64</sup>

Returning to the admonition by Furness et al<sup>2,3</sup> that “urgent steps are needed” to reduce clinically important errors due to inter-pathologist variation, the present study suggests some steps. First, save a small sample in RNA/later™ solution for problem-solving. Second, send ambiguous samples for molecular assessment, since they usually represent molecularly unambiguous states that will impact management. Third, forcing the pathologist to dichotomize difficult scenarios such as v-lesions and tubulitis is less useful than recognizing the differential diagnosis and seeking to address it. MMDx emphasizes the position of the new biopsy in the reference set as the basis for assigning the probability of various categories, and histology could also benefit from probabilistic expression that communicates uncertainty

## CHAPTER 4: Discrepancy analysis comparing molecular and histologic diagnoses

to the clinician.<sup>182,183,241</sup> Fourth, algorithm-based support for diagnoses would also reduce inter-observer discrepancies and has previously been recommended for generating the differential diagnosis.<sup>2,3,242</sup> Candidate logistic regression equations for rejection have been published.<sup>34,39</sup> Fifth, the Banff guidelines can be recalibrated and simplified against independent assessments such as MMDx. The current guidelines were developed starting in 1991 and many have changed minimally. Classifications become more complex over time, and complexity increased inter-observer variation. This could incorporate new approaches such as considering DSA results in probabilistic terms.<sup>34,36,39,40,72</sup> Other possibilities would be difficult in practice: for example, ensembles of multiple independent opinions or votes would reduce the inherent variation caused by reliance on single observers<sup>2,174-179</sup> but this is impractical in SOC assessments.

The take-away message from this study is that MMDx can offer clarification and differential diagnosis of probable states for the pathologist or clinician facing scenarios with known uncertainty in indication biopsies, without assuming that MMDx is always correct. When histology diagnoses ABMR suspicious or TG, there is a high probability that MMDx will produce a clear diagnosis of ABMR or no rejection. When histology diagnoses borderline, MMDx is usually no rejection, ABMR, or TCMR. Histologic TCMR is MMDx TCMR less than half the time and contains cases of molecular ABMR and mixed rejection. Histologic mixed is not usually molecular mixed, but often pure ABMR or TCMR. Even the SOC histology diagnosis of NR contains MMDx ABMR and TCMR. Retaining a portion of all biopsies in RNA<sup>later</sup>™ solution for addressing problematic scenarios would go a long way towards reducing clinical uncertainty.

## **CHAPTER 4:** Discrepancy analysis comparing molecular and histologic diagnoses

### **4.10 TABLES**

**CHAPTER 4:** Discrepancy analysis comparing molecular and histologic diagnoses**Table 4.1** Participating centers in the INTERLIVER study as of December, 2019.

<b>Center</b>	<b>Principal investigators</b>	<b>Samples contributed</b>
<b>Birmingham, AL, USA</b>	Roslyn Mannon	28
<b>Barcelona, Spain</b>	Daniel Serón and Joana Sellarés	71
<b>Bronx, NY, USA</b>	Enver Akalin	33
<b>Manchester, UK</b>	Declan de Freitas and Michael Picton	39
<b>Baltimore, USA</b>	Jonathan Bromberg and Matt Weir	50
<b>Berlin, Germany</b>	Klemens Budde and Timm Heinbokel	12
<b>Hannover, Germany</b>	Gunilla Einecke	67
<b>Harrisburg, PA, USA</b>	Harold Yang and Seth Narins	12
<b>Baltimore, MD, USA</b>	Jonathan Bromberg and Matt Weir	22
<b>Detroit, MI, USA</b>	Milagros Samaniego-Picota	1
<b>Paris, France</b>	Carmen Lefaucheur, Alexandre Loupy	212
<b>Poland</b>	Marek Myslak and Agnieszka Perkowska-Ptasinska	2
<b>San Antonio, TX, USA</b>	Adam Bingaman	83
<b>St Louis, MO, USA</b>	Daniel Brennan and Andrew Malone	20
<b>Minneapolis, MN, USA</b>	Bertram Kasiske	6
<b>Edmonton, AB, CAD</b>	Philip F Halloran	487
<b>Minneapolis, MN, USA</b>	Arthur Matas	76
<b>Madison, WI, USA</b>	Arjang Djamali	10
<b>Vienna, Austria</b>	Georg Böhmig and Farsad Eskandary	198
<b>Richmond, VA, USA</b>	Gaurav Gupta	250
<b>TOTAL</b>		<b>1679</b>

**CHAPTER 4:** Discrepancy analysis comparing molecular and histologic diagnoses

**Table 4.2** Demographics and clinical features of the 1679 biopsy cohort.

<b>Patient Demographics (n=1448)</b>		
<b>Mean recipient age (range)</b>		51 (8 – 91)
<b>Recipient Gender Male (% male)</b>		726 (55%)
<b>Ethnicity</b>	Caucasian	619
	Black	189
	Other	145
	Not available <sup>a</sup>	495
<b>Primary Disease</b>	Diabetic nephropathy (DN)	215
	Hypertension / large vessel disease	110
	Glomerulonephritis / vasculitis (GN)	47
	Interstitial nephritis / pyelonephritis	26
	Polycystic kidney disease	128
	Others	640
	Unknown etiology	282
<b>Mean donor age (range)</b>		44 (1 – 85)
<b>Donor gender (% male)</b>		413 (34%)
<b>Donor type (% deceased donor transplants)</b>		949 (67%)
<b>Latest kidney status (% of total)</b>	Functioning graft	1011 (70%)
	Graft failure/return to dialysis	206 (14%)
	Patient death with functioning graft	88 (6%)
	Mean (median) follow-up (functioning grafts) in days	651 (285)
<b>Biopsy data (n=1679)</b>		
<b>Median time of biopsy post-transplant (TxBx) in days (range)</b>		650 (1 – 12371)
<b>Early biopsies (&lt; 1 year) (% total)</b>		709 (42%) <sup>b</sup>
<b>Late biopsies (≥ 1 year) (% total)</b>		966 (57%) <sup>b</sup>
<b>Indication for biopsy (% of total)</b>	Primary non-function	10 (1%)
	Rapid deterioration of graft function	292 (17%)
	Slow deterioration of graft function	307 (18%)
	Stable impaired graft function	92 (5%)
	Investigate proteinuria/rejection/BK/creatinine	247 (15%)
	Delayed graft function	74 (4%)
	Others	617 (37%)
	Indication not specified	40 (2%)
<sup>a</sup> Some centers preferred not to identify ethnicity		
<sup>b</sup> Four biopsies had no date of transplant		

**CHAPTER 4:** Discrepancy analysis comparing molecular and histologic diagnoses

**Table 4.3** Histologic diagnoses and DSA status in the 1679 cohort.

<b>Rejection N = 612 (36%)</b>	<b>ABMR-related</b>	ABMR	333 (20%)
		Transplant glomerulopathy (TG)	51 (3%)
		ABMR suspected	33 (2%)
	Mixed (TCMR plus ABMR)	56 (3%)	
	TCMR <sup>a</sup>	139 (8%)	
<b>Borderline</b>			128 (8%)
<b>No rejection N = 939 (56%)</b>	AKI		117 (7%)
	BK		52 (3%)
	Diabetic Nephropathy		24 (1%)
	Glomerulonephritis		108 (6%)
	IFTA not otherwise specified		193 (11%)
	No major abnormalities (NOMOA)		371 (22%)
	Others <sup>b</sup>		74 (4%)
<b>Patient HLA antibody status at DSA at time of biopsy (N = 1679)</b>			
DSA positive			597 (34%)
DSA negative, PRA positive			268 (15%)
DSA negative, PRA unknown			115 (7%)
PRA negative/DSA negative or not done			495 (28%)
DSA/PRA not done			253 (14%)

<sup>a</sup> Three biopsies had histology diagnoses of both TCMR and BK virus - we have categorized these as TCMR in this table and throughout the paper.

<sup>b</sup> Others includes calcineurin inhibitor toxicity, C4d deposition without morphologic evidence for active rejection, donor origin vascular disease, pyelonephritis, systemic infection/diarrhea, and bacterial infection.

**CHAPTER 4:** Discrepancy analysis comparing molecular and histologic diagnoses

**Table 4.4** Classifications for molecular-histology diagnoses comparisons.

	<b>Common classes</b>	<b>MMDx</b>	<b>Histology</b>
<b>Common classes</b>	ABMR	ABMR	ABMR
	pABMR	Possible ABMR 'pABMR' <sup>a</sup>	Suspected ABMR, TG
	Mixed	Mixed <sup>b</sup>	Mixed <sup>b</sup>
	TCMR	TCMR	TCMR
	pTCMR	Possible TCMR 'pTCMR'	Borderline
	No rejection	No rejection	No rejection

<sup>a</sup> ABMR-related included histology suspected ABMR and transplant glomerulopathy because these groups had a relatively high frequency of MMDx ABMR.

<sup>b</sup> Includes possible mixed.

**CHAPTER 4:** Discrepancy analysis comparing molecular and histologic diagnoses

**Table 4.5** Agreement in 6 classes between Expert2MMDx and Expert1MMDx sign-out comments.

		Expert2MMDx(6 classes)						Row totals	# discrepancies (%)
		ABMR	pABMR	Mixed	TCMR	pTCMR	NR		
Expert1MMDx (6 classes)	ABMR	481	16*	1 <sup>◇</sup>	0	0	11**	509	28/509 (6%)
	pABMR	6*	24	1 <sup>◇</sup>	0	1**	20*	52	28/52 (53%)
	Mixed	4 <sup>◇</sup>	0	58	7 <sup>◇</sup>	0	0	69	11/69 (16%)
	TCMR	2**	0	3 <sup>◇</sup>	105	5*	8**	123	18/123 (15%)
	pTCMR	0	1**	0	4*	11	5*	21	10/21 (48%)
	NR	2**	10*	0	0	1*	892	905	13/905 (1%)
<b>Column totals</b>		<b>495</b>	<b>51</b>	<b>63</b>	<b>116</b>	<b>18</b>	<b>936</b>	<b>1679</b>	<b>108</b>
<b># discrepancies (%)</b>		<b>14/495 (3%)</b>	<b>27/51 (53%)</b>	<b>5/63 (8%)</b>	<b>11/116 (9%)</b>	<b>7/18 (38%)</b>	<b>44/936 (5%)</b>	<b>108</b>	

\*\* Clear discrepancy between Expert1MMDx and Expert2MMDx (N=25).

\* Boundary discrepancy between Expert1MMDx and Expert2MMDx (N=67).

◇ Mixed-related discrepancy between Expert1MMDx and Expert2MMDx (N=16).

**CHAPTER 4:** Discrepancy analysis comparing molecular and histologic diagnoses

**Table 4.6** All clear discrepancies between alternative molecular sign-out comments ordered from most to least frequent.

<b>AutoMMDx versus Expert1MMDx discrepancies</b>			
<b>Discrepancy class</b>	<b>Assigned sign-out comments</b>		<b>Number of biopsies</b>
	<b><u>Expert1MMDx</u></b>	<b><u>AutoMMDx</u></b>	
<b>Clear discrepancies (N=23)</b>	TCMR	NR	10
	TCMR	pABMR	5
	ABMR	NR	4
	pTCMR	pABMR	3
	NR	ABMR	1
<b>Expert1MMDx and Expert2MMDx discrepancies</b>			
<b>Discrepancy class</b>	<b>Assigned sign-out comments</b>		<b>Number of biopsies</b>
	<b><u>Expert1MMDx</u></b>	<b><u>Expert2MMDx</u></b>	
<b>Clear discrepancies (N=25)</b>	ABMR	NR	11
	TCMR	NR	8
	TCMR	ABMR	2
	NR	ABMR	2
	pTCMR	pABMR	1
	pABMR	pTCMR	1

**CHAPTER 4:** Discrepancy analysis comparing molecular and histologic diagnoses

**Table 4.7** Agreement in 6 classes between the AutoMMDx and Expert1MMDx sign-out comments.

		AutoMMDx (6 classes)						Row totals	# discrepancies per row (%)
		ABMR	pABMR	Mixed	TCMR	pTCMR	NR		
Expert1MMDx (6 classes)	ABMR	477	28*	0	0	0	4**	509	32/509 (6%)
	pABMR	17*	29	0	0	0	6*	52	23/52 (44%)
	Mixed	13 <sup>◇</sup>	2 <sup>◇</sup>	36	18 <sup>◇</sup>	0	0	69	33/69 (48%)
	TCMR	0	5**	0	97	11*	10*	123	26/123 (21%)
	pTCMR	0	3**	0	0	13	5*	21	8/21 (38%)
	NR	1**	26*	0	0	0	878	905	27/905 (3%)
Column totals		508	93	36	115	24	903	1679	149
# discrepancies per column (%)		31/509 (6%)	64/93 (69%)	0/36 (0%)	18/115 (16%)	11/24 (46%)	25/903 (2%)	149	

\*\* Clear discrepancy between ExpertMMDx and AutoMMDx (N=23).

\* Boundary discrepancy between ExpertMMDx and AutoMMDx (N=93).

◇ Mixed-related discrepancy between ExpertMMDx and AutoMMDx (N=33).

**CHAPTER 4:** Discrepancy analysis comparing molecular and histologic diagnoses

**Table 4.8** Relating histologic rejection diagnoses to Expert1MMDx rejection sign-out comments (6 classes, N=1679)

			Histologic diagnoses (6 classes)								No Rejection	Row totals	# discrepancies per row (%)
			Rejection-related=740						Mixed (N=56)	TCMR related (N=267)			
			ABMR-related (N=417)			TCMR	pTCMR (Borderline)						
			ABMR	pABMR									
		ABMR suspicious		TG									
Expert1MMDx sign-out comments (6 class)	ABMR-related N=561	ABMR	260	17*	32*	25 <sup>◇</sup>	20**	24**	131**	509	249/509 (49%)		
		Possible ABMR	12*	1	1	1 <sup>◇</sup>	3**	5**	29*	52	50/52 (96%)		
	Mixed		6 <sup>◇</sup>	2 <sup>◇</sup>	1 <sup>◇</sup>	22	25 <sup>◇</sup>	4 <sup>◇</sup>	9 <sup>◇</sup>	69	47/69 (68%)		
	TCMR related N=144	TCMR	5**	0	1**	5 <sup>◇</sup>	55	9*	48**	123	68/123 (55%)		
		Possible TCMR	0	0	0	1 <sup>◇</sup>	8*	3	9*	21	18/21 (86%)		
	No rejection		50**	13*	16*	2 <sup>◇</sup>	28**	83*	713	905	192/905 (21%)		
<b>Column totals</b>			<b>333</b>	<b>33</b>	<b>51</b>	<b>56</b>	<b>139</b>	<b>128</b>	<b>939</b>	<b>1679</b>	<b>624/1679 (37%)</b>		
<b># discrepancies per column (%)</b>			<b>73/333 (22%)</b>	<b>32/33 (97%)</b>	<b>50/51 (98%)</b>	<b>34/56 (61%)</b>	<b>84/139 (60%)</b>	<b>125/128 (98%)</b>	<b>226/939 (26%)</b>	<b>624/1679 (37%)</b>			

NOTE. pTCMR and pABMR were ignored in definite ABMR or TCMR respectively. 26 Expert1MMDx TCMR/Histology NR were BK virus positive.

\*\* Clear discrepancies between Expert1MMDx and histology (N=315).

\* Boundary discrepancies between Expert1MMDx and histology (N=228).

◇ Mixed discrepancies between Expert1MMDx and histology (N=81).

**CHAPTER 4:** Discrepancy analysis comparing molecular and histologic diagnoses

**Table 4.9** Histologic ‘No Rejection’ subclasses (N=939/1679).

		<i>BK<sup>a</sup></i>	<i>AKI / No major abnormalities</i>	<i>Abnormal<sup>b</sup></i>	Row totals
<b>MMDx diagnoses (6 class)</b>	<b>ABMR</b>	<b>4</b>	<b>60</b>	<b>67</b>	<b>131</b>
	<b>Possible ABMR</b>	<b>3</b>	<b>15</b>	<b>11</b>	<b>29</b>
	<b>Mixed</b>	<b>1</b>	<b>1</b>	<b>7</b>	<b>9</b>
	<b>TCMR</b>	<b>26</b>	<b>5</b>	<b>17</b>	<b>48</b>
	<b>Possible TCMR</b>	<b>3</b>	<b>1</b>	<b>5</b>	<b>9</b>
	<b>No rejection</b>	<b>15</b>	<b>406</b>	<b>292</b>	<b>713</b>
	<b>Column totals</b>	<b>52</b>	<b>488</b>	<b>399</b>	<b>939</b>

<sup>a</sup> Histology class ‘BK’ excludes three cases of ‘TCMR/BK’ by pathology. These cases were counted once as TCMR rather than as No rejection.

<sup>b</sup> Abnormal histology class includes diabetic nephropathy, glomerulonephritis, IFTA, and other renal diseases.

**CHAPTER 4:** Discrepancy analysis comparing molecular and histologic diagnoses

**Table 4.10** All clear histology-Expert1MMDx discrepancies ordered from most to least frequent (N=315).

	Assigned Diagnoses		Number of biopsies
	<u>Expert1MMDx</u>	<u>Histology</u>	
<b>Clear discrepancies (N=315)</b>	ABMR	NR	131
	NR	ABMR	50
	TCMR	NR	48
	NR	TCMR	28
	ABMR	pTCMR	24
	ABMR	TCMR	20
	TCMR	ABMR	5
	pABMR	pTCMR	5
	pABMR	TCMR	3
	TCMR	pABMR	1

**CHAPTER 4:** Discrepancy analysis comparing molecular and histologic diagnoses

**Table 4.11** All histology-MMDx, AutoMMDx/Expert1MMDx, and Expert2MMDx/Expert1MMDx discrepancies classified as clear, boundary, or mixed-related (N=624, N=149, and N=108 respectively)

	<b>Discrepancies (% of all 1679 biopsies)</b>			
	<b><u>Histology versus Expert1MMDx</u></b>	<b><u>Intra-MMDx</u></b>		
		<b><u>Mean Intra-MMDx</u></b>	<b><u>AutoMMDx versus Expert1MMDx</u></b>	<b><u>Expert2MMDx versus Expert1MMDx</u></b>
<b>Clear discrepancies</b>	<b>315 (19%)</b>	<b>24 (1.4%)</b>	<b>23 (1.4%)</b>	<b>25 (1.5%)</b>
<b>Mixed discrepancies</b>	<b>81 (5%)</b>	<b>25 (1.5%)</b>	<b>33 (2%)</b>	<b>16 (1%)</b>
<b>Boundary discrepancies</b>	<b>228 (14%)</b>	<b>80 (5%)</b>	<b>93 (6%)</b>	<b>67 (4%)</b>
<b>COLUMN TOTAL</b>	<b>624 (37%)</b>	<b>129 (8%)</b>	<b>149 (9%)</b>	<b>108 (6%)</b>
<b>Overall kappa value</b>	<b>0.41</b>	<b>0.87</b>	<b>0.85</b>	<b>0.89</b>

**CHAPTER 4:** Discrepancy analysis comparing molecular and histologic diagnoses

**Table 4.12** Kappa values for histology-Expert1MMDx and Expert2MMDx-Expert1MMDx disagreement within previously flagged problematic scenarios (N=1679).

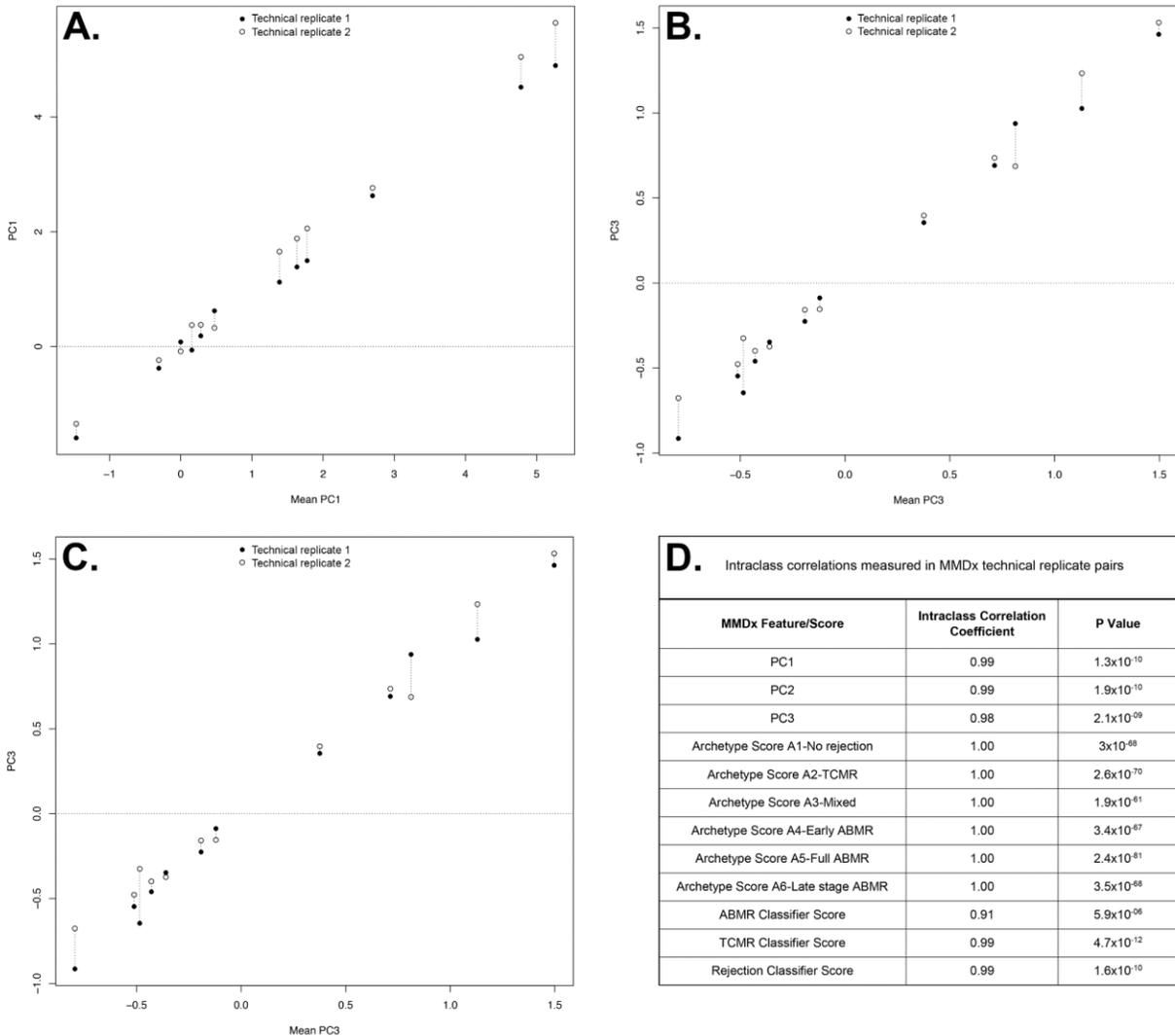
Histology versus Expert1MMDx		Expert2MMDx versus Expert1MMDx	
DSA negative <sup>A</sup> (N=808)	DSA positive (N=573)	DSA negative <sup>A</sup> (N=808)	DSA positive (N=573)
0.30	0.46	0.88	0.89
BK negative (N=1624)	BK positive (N=55)	BK negative (N=1624)	BK positive (N=55)
0.43	0.33	0.90	0.78
t 0/1 lesions (N=1371)	t 2/3 lesions (N=257)	t 0/1 lesions (N=1371)	t 2/3 lesions (N=257)
0.42	0.12	0.89	0.88
ci 0/1 lesions (N=1021)	ci 2/3 lesions (N=589)	ci 0/1 lesions (N=1021)	ci 2/3 lesions (N=589)
0.42	0.37	0.90	0.88
v=0 lesions (N=1482)	v>0 lesions (N=77)	v=0 lesions (N=1482)	v>0 lesions (N=77)
0.40	0.19	0.89	0.91
<sup>A</sup> DSA not done (N=298), Kappa value = 0.36		<sup>A</sup> DSA not done (N=298), Kappa value = 0.89	

**NOTE.** AutoMMDx versus Expert1MMDx had similar kappa values to Expert2MMDx versus Expert1MMDx.

## **CHAPTER 4:** Discrepancy analysis comparing molecular and histologic diagnoses

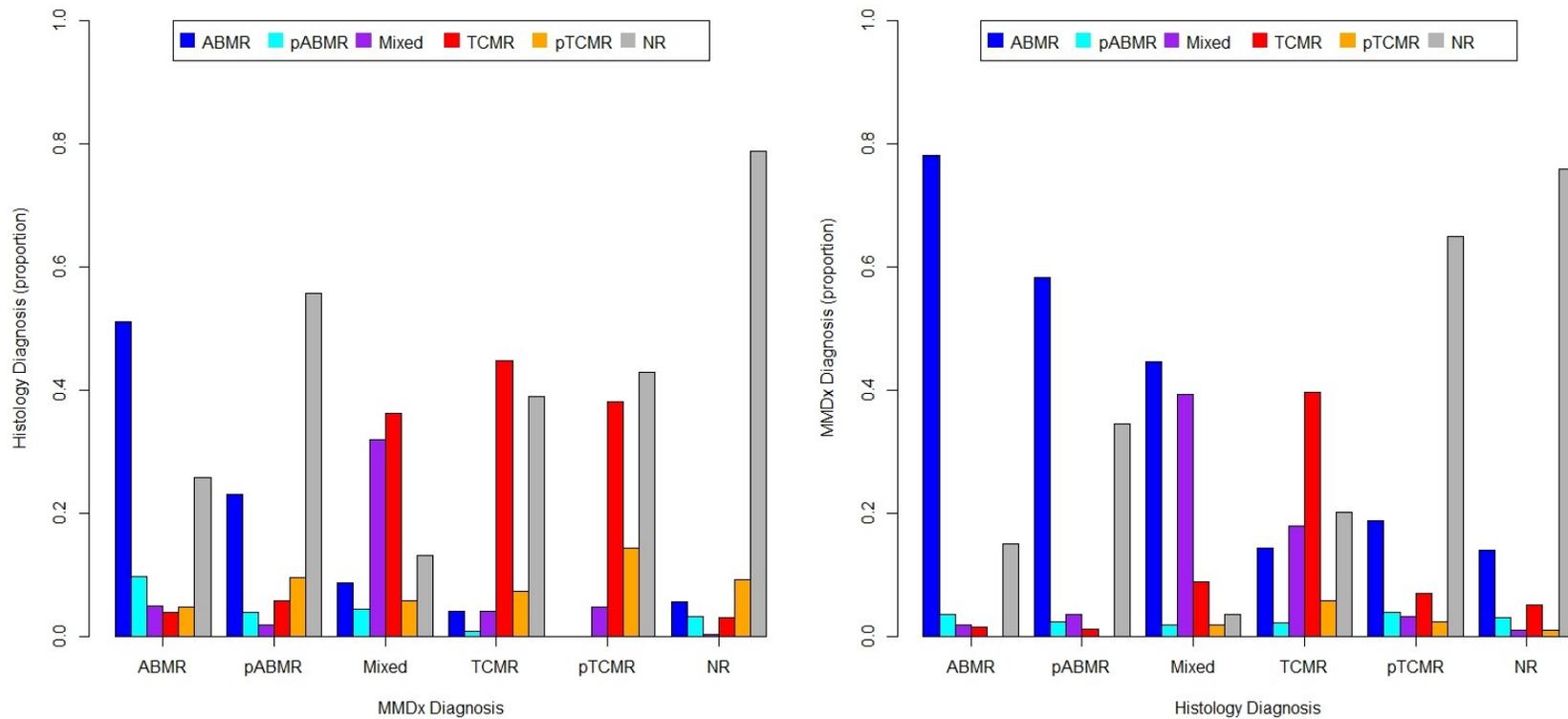
### **4.11 FIGURES**

## CHAPTER 4: Discrepancy analysis comparing molecular and histologic diagnoses

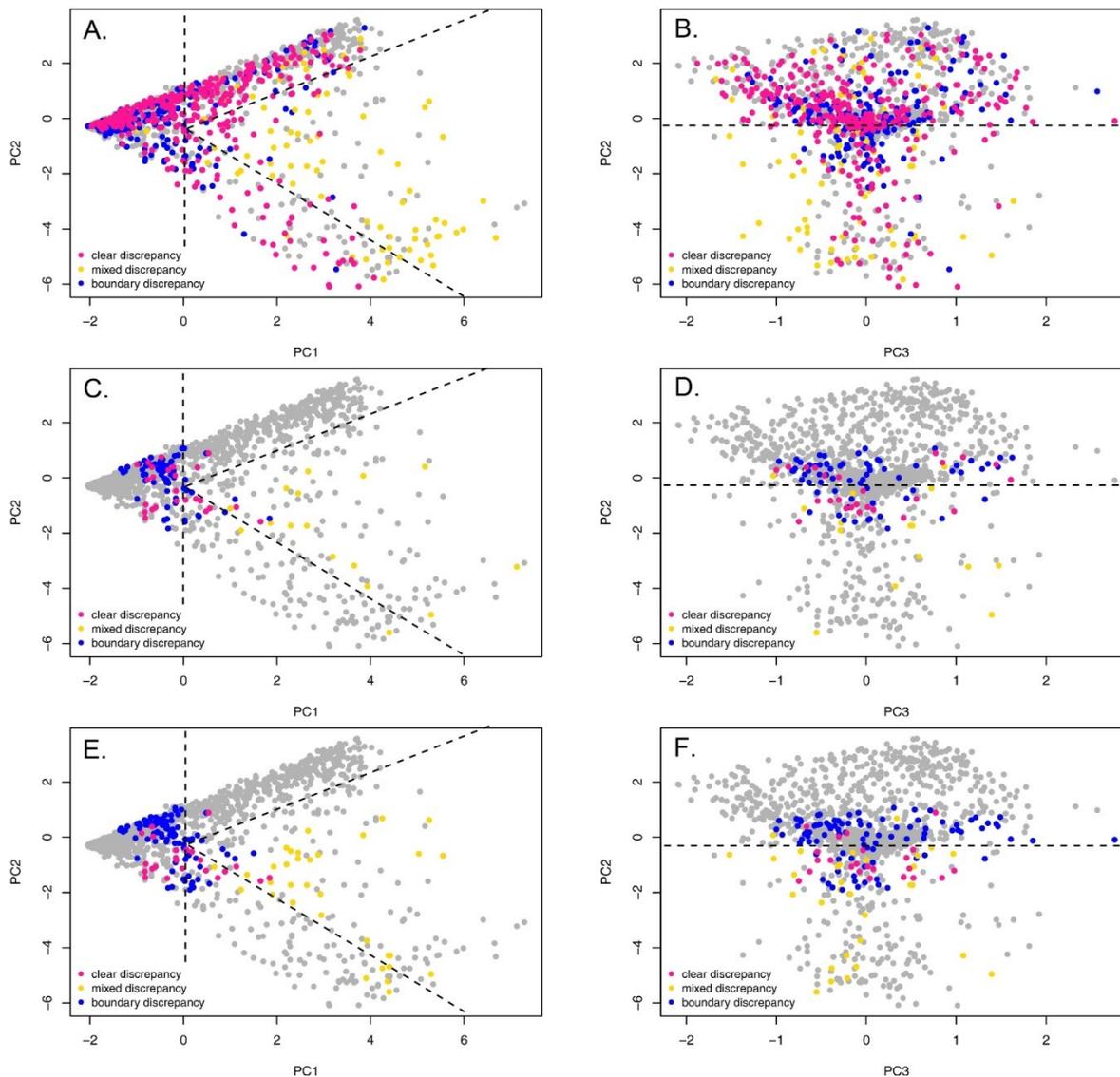


**Figure 4.1 Reproducibility of PC scores, archetype scores, and main classifier scores in MMDx technical replicate pairs.** Technical replicate pairs are joined by a dashed line. The actual score of each replicate is plotted against the mean score of the pair in A) showing PC1 scores, B) showing PC2 scores, and C) showing PC3 scores. Intraclass correlations calculated for the main MMDx scores (Panel D) are high, even in PC and classifier scores where the values are not bounded (archetype scores are bounded between 0-1 giving less room for variability). Abbreviations: ABMR – antibody-mediated rejection, PC1 – principal component 1, PC2 – principal component 2, PC3 – principal component 3, TCMR – T cell-mediated rejection.

**CHAPTER 4: Discrepancy analysis comparing molecular and histologic diagnoses**



**Figure 4.2 Clear histology-MMDx discrepancy scenarios where MMDx would provide a useful second opinion.** Panel A shows the overall mapping of all histology diagnoses per MMDx diagnostic category in the 1679 population. Panel B shows the mapping of MMDx diagnostic sign-out comments to histology diagnoses in the 1679 population.



**Figure 4.3 Visualizing discrepancies in molecular-based PCA as shown on the MMDx report.**

Histology molecular discrepancies are shown in A) PC2 versus PC1, B) PC3 versus PC1. In contrast, intra-MMDx (Expert1 versus Expert2) discrepancies are shown in C) PC2 versus PC1, D) PC3 versus PC1. Expert1MMDx versus AutoMMDx is shown in Panels E) PC2 versus PC1 and F) PC2 versus PC3. Dashed lines show diagnostic boundaries between common classes (i.e. rejection and no rejection, ABMR/Mixed, Mixed/TCMR, etc.) Clear discrepancies are colored pink, mixed-related discrepancies are colored yellow, and boundary discrepancies are colored blue. Abbreviations: PC1 – principal component 1, PC2 – principal component 2, PC3 – principal component 3.

## **CHAPTER 5**

### **INTRA-HISTOLOGY DISCREPANCIES IN RENAL TRANSPLANTATION: COMPARISONS BETWEEN THE SOC ASSESSMENTS AND AN ALGORITHM STRICTLY FOLLOWING BANFF GUIDELINES FOR REJECTION DIAGNOSES**

## **5.1 OVERVIEW AND HYPOTHESIS**

In assessing biopsies, diagnostic systems select features to be measured, usually features that are not completely specific, and develop rules for using these features to assign diagnoses. Banff consensus guidelines<sup>22,32</sup> for diagnosing rejection in renal transplant biopsies specify how eight predetermined canonical rejection features lead to diagnoses: for TCMR, i-, t-, and v-lesions; and for ABMR, ptc-, g-, and cg-lesions, plus complement factor C4d staining and circulating DSA. The system acknowledges ambiguous ‘boundary’ categories: “borderline” (BD) for TCMR, and “ABMR suspicious” and “transplant glomerulopathy” for ABMR. (A recent proposal for diagnosing chronic active TCMR based on inflammation in scarred areas<sup>32,243–245</sup> was not used for the present cases). Other features must be considered occasionally e.g. thrombotic microangiopathy (TMA).

Histology manifests considerable imprecision as measured by interobserver disagreement<sup>2,3,36</sup> due to the need for subjective assessments, as reflected in low kappa values.<sup>2,3,42,246</sup> Accuracy depends on how well the guidelines capture the true disease states. Some guidelines are probably inaccurate<sup>66</sup> because they were of necessity derived from prevailing expert opinions at a time of limited understanding of the true disease processes. In addition, sampling differences arise from uneven distribution of the disease features, changes in the population over time (data drift), and inherent variability in the tissue itself e.g. cortex versus medulla. To date, it has been unclear how deviations from Banff guidelines affect agreement between molecular diagnoses assigned by MMDx and SOC diagnoses assigned by a local pathologist.

The present study was designed to understand the relationship between histology diagnoses in established centers that follow the Banff guidelines and the Banff guidelines i.e. how pathologists actually apply the system in SOC biopsy assessments. Diagnoses recorded by pathologists per SOC in experienced international centers as captured in the multicenter INTERCOMEX study (“ExpertBanff”) were compared to diagnoses assigned by strictly applying the Banff guidelines to recorded canonical rejection features using an automated computer algorithm (“AutoBanff”). Scenarios where disagreement was common were examined. A simulated strict application of the Banff guidelines was also compared to the optimized microarray-based Molecular Microscope<sup>®</sup> Diagnostic System (MMDx),<sup>6</sup> to estimate if this

## **CHAPTER 5: Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm**

practice would improve MMDx-histology agreement. The hypothesis was that strict application of the guidelines to the observed lesions using the AutoBanff algorithm would improve agreement between SOC ExpertBanff and the external MMDx diagnoses.

### **5.2 BIOPSY POPULATION**

The biopsy population used in these analyses was previously described in Chapter 4. Please see Tables 4.1, 4.2, and 4.3 for details.

### **5.3 DEVELOPMENT OF THE AUTOBANFF ALGORITHM**

The algorithm for AutoBanff was developed in the R computing language and applied to all 1679 biopsies (**Figure 5.1**). This algorithm was based on Banff 2017 guidelines<sup>32</sup> converted to a 'decision tree' format, where a series of subsequent decisions based on data results in a final diagnosis (**Figure 5.1**). The algorithm functions with a decision-making hierarchy; in the order  $A \rightarrow B \rightarrow C \rightarrow D$  (**Figure 5.1**) to allow for the diagnosis of some types of rejection that would rely on a previous decision (i.e. a diagnosis of ABMR requires prior knowledge about the presence or absence of TCMR in the biopsy). Thus, AutoBanff is a programmed algorithm built to strictly apply Banff 2017 guidelines to eight canonical rejection features recorded for each biopsy, as a simulation of a single expert pathologist strictly applying Banff guidelines. AutoBanff generates one of six different diagnoses: ABMR, pABMR, TCMR, pTCMR, Mixed Rejection, and NR.<sup>6</sup> The SOC histology diagnoses (ExpertBanff) were translated into these six classes (as was done in Chapter 4),<sup>6</sup> permitting comparisons among ExpertBanff, AutoBanff, and MMDx (Table 5.1). The frequency of these lesions and features within the 1679 population is summarized in **Table 5.2**.

Some factors were not available or not used for the AutoBanff algorithm i.e. TMA. Overall, these conditions were rare in the population and did not have a substantial effect on the accuracy of the algorithm (**Table 5.3**).

## **5.4 STATISTICAL ANALYSES**

All analyses were done using functions in the base R system (version 3.5.1),<sup>185</sup> while PCA plots were generated using the extension package FactoMineR.<sup>199</sup> Random forests were run with the 'rfsrc' function in the randomForestSRC<sup>247</sup> package, using the default parameters except for importance = "permute", ntree = 5000, nsplit = 1, and na.action= "na.impute".

## **5.5 COMPARING SOC HISTOLOGY DIAGNOSES TO AUTOBANFF ASSIGNMENTS**

As done in Chapter 4, discrepancies between ExpertBanff, AutoBanff, and MMDx were called 'boundary' if they involved ambiguous classes (discrepancies occurring along cutoffs, e.g. NR versus pABMR, pTCMR versus TCMR). Otherwise, discrepancies were classified as 'clear' discrepancies between distinct classes (e.g. ABMR versus NR).

### **5.5.1 Discrepancies between ExpertBanff and AutoBanff**

ExpertBanff was compared to AutoBanff in **Table 5.4**. The agreement is indicated along the diagonal, and cells with more than 20 discrepant biopsies are bolded. Overall, diagnoses were discrepant in 439/1679 biopsies (26%). Intra-histology (IH) discrepancies with  $\geq 20$  counts were classified in **Table 5.4** as either clear (\*) or boundary (\*\*).

Clear discrepancies accounted for 202/439 (46%): of which the most frequent was 59 ExpertBanff ABMR called AutoBanff NR (**Table 5.5**). Boundary discrepancies accounted for 237/439 (54%), of which 137/237 (57%) were related to ABMR diagnoses (ABMR/pABMR/NR) (**Table 5.6**). The most frequent boundary discrepancies were 77 ExpertBanff NR cases called AutoBanff pTCMR.

The most frequent clear and boundary discrepancies ( $\geq 20$ ) are summarized in **Table 5.7**.

## **5.6 DISCREPANCIES IN PREVIOUSLY NOTED PROBLEMATIC AREAS**

## **CHAPTER 5: Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm**

Problematic histology guidelines were previously flagged in several publications and in Chapter 4: BK virus in the presence of TCMR,<sup>72-74,201</sup> ci lesion scores = 2/3 (scarring),<sup>248</sup> DSA ambiguous/negative in the context of ABMR,<sup>39,40</sup> v-lesions > 0 in rejection,<sup>34</sup> and tubulitis in ABMR.<sup>37,38,42,201</sup> These scenarios were examined to see if they were also problematic for IH agreement.

Discrepancies in previously flagged problematic scenarios (BK virus nephropathy, ci 2/3 lesions, negative or ambiguous DSA, v- and t-lesions) were assessed using 2x2 chi-squared analyses, focusing on the overall disagreement versus positive agreement.

### **5.6.1 TCMR lesions in BK**

Biopsies called TCMR by ExpertBanff and/or AutoBanff were examined. In 15 biopsies identified as BK virus positive, TCMR was not diagnosed by ExpertBanff despite sufficient lesions. Thus, the presence of BK virus increases the rate of discrepancies, reflecting pathologists' reluctance to diagnose TCMR in BK-positive kidneys ( $p=6 \times 10^{-8}$ , **Table 5.8**), although a strict reading of the guidelines does dictate exclusion of such cases.

### **5.6.2 Rejection lesions in scarred biopsies**

The effect of scarring on the frequency of discrepancies in biopsies with rejection was assessed. A subpopulation of 179 biopsies called TCMR and 366 biopsies called ABMR by ExpertBanff and/or AutoBanff was selected to represent biopsies with clear rejection (**Table 5.9**). In both cases, the frequency of discrepancies was similar in scarred versus unscarred biopsies (TCMR:  $p=0.36$ , ABMR:  $p=0.80$ ). Thus, scarring did not significantly affect the frequency of discrepancies in biopsies with rejection (ABMR or TCMR). This is likely because scarring affects lesion scores but does not affect guideline applications.

### **5.6.3 ABMR lesions with negative or ambiguous DSA**

## CHAPTER 5: Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

The effect of DSA status on the rate of discrepancies in biopsies called ABMR by ExpertBanff and/or AutoBanff was assessed (**Table 5.10**). ExpertBanff diagnosed ABMR in 60 cases recorded as DSA negative, while AutoBanff was not permitted to diagnose ABMR in any of these cases. (Although Banff 2017 guidelines introduced caveats for ABMR when DSA is negative, including the substitution of C4d positivity or ‘thoroughly validated molecular test [results]’<sup>132</sup> these caveats still contain ambiguities and were not programmable). ExpertBanff diagnosed ABMR in 35 cases where the recorded DSA was unclear i.e. the DSA was recorded as ‘pending’, ‘not done’, ‘missing’, etc. Again, AutoBanff did not diagnose ABMR in any of these cases. ExpertBanff frequently diagnosed ABMR in situations where AutoBanff was restricted to NR or pABMR. The presence of negative or ambiguous DSA in the context of ABMR features increased the discrepancies ( $p < 2 \times 10^{-16}$ ), generally in the direction that pathologists are assigning ABMR diagnoses despite ambiguous or negative DSA. It should be noted that many of these assignments were made before the release of the updated Banff 2017 guidelines and were not based on the guideline change but rather the pathologist’s confidence in the ABMR diagnosis despite the DSA result.

### 5.6.4 V>0 lesions in rejection

The effect of v-lesions >0 on the frequency of discrepancies in biopsies called rejection (TCMR, ABMR, or Mixed rejection) by ExpertBanff and/or AutoBanff was assessed (**Table 5.11**).

In the subset of 181 biopsies called TCMR and 65 biopsies called Mixed rejection by ExpertBanff and/or AutoBanff, the frequency of discrepancies was unaffected by v>0 lesions ( $p=0.80$  and  $p=0.11$  respectively, **Table 5.11**).

Within a subset of 348 biopsies called ABMR by ExpertBanff and/or AutoBanff, the frequency of discrepancies was increased by v>0 lesions ( $p=0.001$ , **Table 5.11**). It was also noted that in 10 biopsies with v>0 lesions, ExpertBanff diagnosed ABMR while AutoBanff diagnosed TCMR or Mixed (as per Banff 2017 guidelines, which state that v>0 lesions must be called TCMR, either alone or in the context of accompanying ABMR).

## **CHAPTER 5: Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm**

In 551 biopsies called ABMR, TCMR, or Mixed rejection (representing an 'all rejection' cohort) by Expert and/or AutoBanff,  $v>0$  lesions did not affect the frequency of discrepancies ( $p=0.56$ , **Table 5.11**). Of interest, ExpertBanff did not diagnose rejection in 9 cases with  $v>0$  lesions: in 5 of these cases, MMDx and ExpertBanff agreed that there was no rejection (**Table 5.12**).

Thus,  $v$ -lesions did not affect the frequency of IH discrepancies in all biopsies with any rejection (the cohort of ABMR, TCMR, or Mixed), but there were difficulties in deciding how to apply  $v$ -lesions in the context of ABMR features. Note that Banff guidelines do not currently acknowledge that  $v$ -lesions can be caused by injury.<sup>22,32</sup>

### **5.6.5 Discrepancies in biopsies with tubulitis lesions.**

The effect of tubulitis lesions ( $t$  lesions = 0/1 versus  $t$  lesions = 2/3) on the rate of discrepancies in all biopsies diagnosed as rejection by ExpertBanff and/or AutoBanff ( $N=593$ ) was assessed. Tubulitis did not significantly affect the frequency of discrepancies ( $p=0.08$ , **Table 5.13**).

Findings regarding the effect of previously flagged problematic scenarios on the frequency of IH discrepancies are summarized in **Table 5.14**.

## **5.7 RANDOM FORESTS FOR VARIABLE IMPORTANCE IN THE PREDICTION OF DISCREPANCIES**

Random forest analysis was used to compare all eight canonical rejection features plus  $ci$ -lesions and BK virus status to determine their relative importance for the prediction of discrepancies (**Figure 5.2**). All 1679 biopsies were incorporated into the analysis in **Figure 5.2A**. In all biopsies, the important features were mostly ABMR-related ( $ptc$ ,  $DSA$ ,  $g$ ,  $cg$ ), but  $t$ -lesions and BK status also contributed (**Figure 5.2A**). **Figure 5.2B** shows a similar analysis using only biopsies with a rejection diagnoses (ABMR, TCMR, Mixed diagnosed by ExpertBanff and/or AutoBanff,  $N=595$ ). In rejection biopsies,  $DSA$  was the dominant factor predicting discrepancies (**Figure 5.2B**).

## CHAPTER 5: Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

The frequencies of lesion scores (and missing values per feature) are summarized in **Table 5.2**. Some variables were frequent but unimportant (i-lesions) and others were infrequent but important (BK status).

### 5.8 RELATIONSHIPS BETWEEN INTRAHISTOLOGY DISCREPANCIES AND MMDx PHENOTYPES

#### 5.8.1 Visualizing IH discrepancies within the context of MMDx diagnoses

PCA was used to visualize the relationship between IH discrepancies and the MMDx phenotype of the biopsies. The MMDx phenotype was represented in a plot of principal component 2 (PC2) versus principal component 1 (PC1), based on seven previously published molecular classifier scores.<sup>6,6,66</sup> IH discrepant biopsies were then colored to show their distribution against the molecular phenotype of all biopsies (represented by grey dots) (**Figure 5.3**). The IH discrepancies were distributed over the whole range of molecular rejection-related phenotypes. Discrepancies were particularly frequent in the areas of molecular ABMR and molecular no rejection. Therefore, clear discrepancies were present in many cases with unequivocal MMDx diagnoses.

#### 5.8.2 IH discrepancies as a contributor to molecular-histology discrepancies

IH discrepancies, as reflections of problematic biopsies, were theorized to account for most molecular-histology (MH) discrepancies i.e. how many MH discrepant cases were also IH discrepant. MH discrepancies were previously defined by a comparison between the expert diagnosis using MMDx and the ExpertBanff diagnosis (See Chapter 4). A Venn diagram examined the overlap between the MH discrepancies and IH discrepancies (**Figure 5.4**). In ABMR MH discrepancies, only 19% (61/322) were also IH discrepancies (**Figure 5.4**). Of 152 TCMR MH discrepancies, only 40 (26%) were also IH discrepancies (**Figure 5.4**). The majority of disagreements between histology and MMDx occur in biopsies where AutoBanff and ExpertBanff agree, i.e. both are discrepant with MMDx.

Contrary to the hypothesis, IH discrepancies were not a major source of MH discrepancies. ExpertBanff more closely predicted the MMDx diagnosis than AutoBanff. ExpertBanff disagreed with

## CHAPTER 5: Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

MMDx in 624/1679 cases (37%, clear and boundary discrepancies), while AutoBanff disagreed with MMDx in 712/1679 cases (42%, clear and boundary discrepancies) ( $p < 0.05$ ). Pathologist deviations from the canonical feature guidelines *increase* agreement with MMDx, disproving the original hypothesis.

### 5.9 INTERPRETATION OF RESULTS

The present study explored the extent of discrepancy between pathologists' diagnosis and Banff guidelines, and whether rigorous adherence to the Banff guidelines for interpreting lesions scores would improve agreement with MMDx. This was an important analysis to undertake, given that discrepancy analysis between two platforms and between observers within one platform can inform on the ease of use, accuracy with regards to the true clinical phenotype, and overall guideline clarity. Diagnostic systems are consistently evolving and improving, and understanding how any system is used practically is imperative to continue improving the final diagnosis and resulting patient management. The AutoBanff algorithm for assigning diagnoses based on eight canonical rejection features compared to recorded ExpertBanff diagnoses showed 26% disagreement, nearly half of which were clear discrepancies and not simply boundary issues. IH discrepancies were higher in some previously flagged scenarios, particularly in assessing ABMR with negative or ambiguous DSA, interpreting TCMR lesions in biopsies with BK virus, and interpreting v-lesions in the context of ABMR features. Random forest analysis confirmed the importance of ABMR as a source of discrepancies: 4/5 of the top variables predicting discrepancies were ABMR-related and DSA was the dominant variable for predicting discrepancies in biopsies with rejection. These findings suggest that a diagnosis of ABMR without positive DSA or with v-lesions, and of TCMR in BK infections should be targets for future evolution of the guidelines. Contrary to the original hypothesis, strict application of Banff guidelines by AutoBanff decreased agreement with MMDx, compared to expert pathologists. Moreover, there was no strong relationship between IH discrepancies and disagreements between local pathologists and MMDx. The conclusion is that histology-MMDx agreement would not be improved by rigorously applying the Banff guidelines to lesion scores, although pathologists may still find utility in seeing the rigorous AutoBanff score before making their final assessment based on all available

## CHAPTER 5: Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

sources of information, and kappa values within histology would likely be increased with guideline refinement.

The experience of expressing the Banff guidelines as a rigorous algorithm was useful and probably should be used to test proposed changes to the guidelines. This analysis suggests that each new change in the Banff guidelines should be created as a clear algorithm or decision tree to avoid ambiguity or unintended consequences. Some Banff 2017 guidelines could not be incorporated into a programmed algorithm as they contained too many ambiguous or contradictory elements (e.g. TMA, electron microscopy results, peritubular capillary basement membrane multilayering).<sup>32</sup>

Clear IH discrepancies are a particular source of concern because they impact care. In the 439 IH discrepancies, 47% were clear e.g. ABMR versus NR. Efforts should be directed at simplifying the guidelines and ensuring that they properly represent the targeted phenotype, starting with ABMR. While boundary discrepancies are expected in any diagnostic system that uses cutoffs, the proportion of clear discrepancies is troubling and suggests that some areas need special attention.

The results of these analyses suggest that standard practices for histologic interpretation of TCMR-like lesions in BK nephropathy need to be clarified, and the presence of a TCMR-like state explicitly recorded with appropriate caveats. BK virus nephropathy is often accompanied by typical histologic and molecular features of TCMR,<sup>72-74</sup> which in many cases resolve when the BK clears and full immunosuppression is restored. Pathologists have been reluctant to record these lesions as compatible with TCMR, reflecting justifiable concern about the possibilities of 1) harmful treatment and 2) that the lesions are triggered by the virus infection, not alloimmunity. These results argue that the TCMR-like state must be clearly noted. TCMR does not always require treatment, but the clinician should know of possible smoldering alloimmunity if function fails to recover as the virus infection resolves. BK management requires reducing immunosuppression and risking alloimmunity, including TCMR that could evolve to ABMR and graft loss.<sup>72-74</sup> In the present study, 15/37 cases with BK nephropathy fit the diagnostic Banff criteria for TCMR (**Figure 5.1A**) but were not diagnosed as TCMR, despite the absence of a clear statement to this effect in the guidelines.<sup>22,32</sup> Although the significance of a TCMR-like process in BK does not imply the need for therapy, and may or may not reflect actual alloimmunity, the presence of TCMR-

## CHAPTER 5: Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

like criteria should be reported to the clinician and must be acknowledged as the disease phenotype evolves.

Atrophy-scarring did not increase the frequency of IH discrepancies, but it poses a challenge for the diagnosis of rejection because active atrophy-scarring triggers inflammation that can simulate rejection, and because scarring reduces the amount of unscarred tissue available to score the canonical i-lesions and t-lesions. In our analyses, biopsies with ci 2/3 lesion scores and biopsies with ci 0/1 lesion scores had similar rates of IH discrepancies. But the struggle to define rejection in biopsies with extensive fibrosis continues as exemplified by the attention being directed to inflammation in scarred areas.<sup>243,244,249</sup> The issue of defining rejection in previously damaged tissue will continue to be an important area.

Negative or ambiguous DSA is a vexing problem that plays a major role in the IH discrepancies. DSA assessments are highly variable between and within laboratories and between batches of kits used for DSA testing and varies in complement binding and titre. DSA is present in many patients with no ABMR phenotype, and some patients have ABMR phenotypes with no detectable DSA, all of which poses a profound challenge to those writing diagnostic guidelines. Pathologists in the present study seem to be acknowledging these challenges by frequently diagnosing ABMR in cases outside the guidelines: Banff 2015-2017 requires DSA (with few exceptions) for a diagnosis of ABMR,<sup>22</sup> but 95 cases in our population were diagnosed as ABMR with either negative or unclear DSA. This shows that the pathologists believed that the total evidence indicated ABMR despite the absence of a clear positive DSA test. It is impossible to write guidelines that cover the full complexity of DSA phenotyping, and it would be advantageous to move to a probabilistic approach to DSA measurements that incorporate quantity, specificity, and other features such as effector system activation (e.g. complement binding). Nevertheless, many cases of unequivocal molecular ABMR with no DSA or measurable HLA antibodies have been documented,<sup>39</sup> a scenario that is now confirmed.<sup>39,250</sup> This finding also raises an interesting question: if the pathogenic antibody in ABMR is not detectable in some cases, is the measured DSA not the relevant antibody causing ABMR? The characterization of HLA-antibody negative ABMR will undoubtedly be an important ongoing scientific and clinical issue.

## CHAPTER 5: Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

Endarteritis lesions (v-lesions) on IH discrepancies did not affect the frequency of IH discrepancies overall but within ABMR they posed challenges to the pathologists. V-lesions are ambiguous: TCMR, ABMR, and other types of endothelial injury can produce v-lesions.<sup>26,34,36,41,212,214,251–253</sup> Many biopsies with isolated v-lesions<sup>34,213</sup> lack molecular TCMR, but continue to be considered TCMR in the Banff guidelines for 2015 and 2017.<sup>22,32</sup> Donation-implantation injury to the endothelium can manifest as v-lesions, but Banff continues to designate v-lesions as primarily TCMR<sup>254</sup> or in specific circumstances ABMR<sup>214</sup> without acknowledging other causes. The interpretation of v-lesions should be clarified by molecular assessments, because resolving ambiguity has implications for appropriate treatment.

The AutoBanff program could not incorporate some details, but relatively few biopsies are affected by these details and most IH discrepancies cannot be attributable to these omissions. These factors could not be included because the data was too often not available, or because the guidelines required a level of subjectivity that could not be programmed (as previously discussed). Excluded factors would affect only a small portion of the IH discrepancies (TMA was only present in 21/1679 biopsies) and the effect of eliminating them from the algorithm is probably small.

The finding that ExpertBanff had better agreement with MMDx than AutoBanff indicates that pathologists' judgment or other sources of information/prior probabilities adds value to the final diagnosis. Guidelines should be over-ruled when the pathologists' professional judgment determines that other sources of information are available. While reducing IH discrepancies is desirable, the long-term goal is to create algorithms that eliminate clear discrepancies and deal with the uncertainties around boundaries in a manner that is optimally useful for guiding management. More importantly, some lesions such as proliferative arteriopathy are not adequately distinguished and could add important granularity to a new set of Banff algorithms. The results suggest that pathologists nominally following Banff guidelines are actually making professional overall assessments, and in doing so they are more likely to be in agreement with a histology-independent assessment, MMDx. The local pathologists have access to details in their local clinical environment such as clinicians' suspicions, and use the Banff consensus as guidelines, which of course is how they were intended. The best estimates, whether conventional or

**CHAPTER 5:** Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

molecular, will always be those that incorporate all valid information. These results do not preclude the results from Chapter 4, which suggested that there is much room for improvement in the Banff guidelines. Increasing emphasis on molecular assessments will help clarify many of these issues that are challenges to pathologists using Banff guidelines, such as v-lesions in ABMR, DSA negative ABMR, and BK with TCMR lesions, and can be used to test new Banff histology algorithms. Based on these findings, it is recommended that MMDx data be considered in ongoing improvements to the Banff guidelines for renal transplantation.

**5.10 TABLES**

**CHAPTER 5:** Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

**Table 5.1** Translating SOC histology diagnoses into six classes to permit comparison with AutoBanff and MMDx diagnoses

	<b>SOC histology classes (ExpertBanff)</b>	<b>Six classes for comparison with AutoBanff/MMDx**</b>
Common classes		ABMR
	Suspected ABMR, TG	Possible ABMR ('pABMR')
	Mixed <sup>▲</sup>	Mixed <sup>▲</sup>
	TCMR	TCMR
	Borderline	Possible TCMR ('pTCMR')
	No rejection	No rejection

\*ABMR-related included histology suspected ABMR and transplant glomerulopathy because these groups had a relatively high frequency of MMDx ABMR.

<sup>▲</sup>Includes possible mixed

\*\*As given in Chapter 4.

**Table 5.2** Frequency of recorded canonical lesion and feature values in the 1679 cohort.

Lesion	Counts per score				
	0	1	2	3	Missing
ptc	1057	230	219	111	62
g	1090	293	154	88	54
cg	1273	140	96	112	58
t	1120	251	151	106	51
v	1482	55	18	4	120
i	1023	277	133	95	151
ci	390	631	396	193	69
Feature	Positive	Negative	Missing or Ambiguous		
DSA	573	808	298		
C4d	156	769	754		
BK*	55	1624	0		

\*BK was either given as a diagnosis (positive) or the biopsy was given another diagnosis (negative). Three biopsies were given the diagnosis 'TCMR/BK' in this population and were counted as BK positive for this table.

**CHAPTER 5:** Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

**Table 5.3** Detail on the 21 cases with TMA in 1679.

Case #	DSA	BK	Lesion scores							C4d	TxBx (days)	MMDx Diagnosis (6 class)	ExpertBanff Diagnosis (6 class)	ExpertBanff Diagnosis (multiclass)	AutoBanff Diagnosis (6 class)	Discrepancy (1 = IH discrepancy, 0= IH agreement)
			g	cg	i	ci	t	v	ptc							
1	-	+	NA	NA	3	0	2	NA	NA	1	164	TCMR	NR	TMA, ATN	TCMR	1
2	+	-	1	3	0	1	0	0	0	0	2828	NR	NR	TMA, CNIT	pABMR	1
3	-	-	2	3	0	2	0	0	0	+	4085	ABMR	ABMR	ABMR, CNIT, TMA	pABMR	1
4	+	-	0	0	0	0	0	0	0	0	2594	NR	NR	TMA, CNIT	NR	0
5	-	-	1	1	0	0	0	0	0	0	50	NR	NR	TMA	NR	0
6	-	-	1	0	0	2	0	NA	0	0	1528	NR	NR	TMA, IFTA	NR	0
7	+	NA	0	0	0	0	0	0	0	0	2868	NR	NR	TMA	NR	0
8	+	NA	2	2	0	2	0	NA	2	2	1632	ABMR	ABMR	ABMR, TMA	ABMR	0
9	NA	NA	0	0	0	0	0	0	0	0	51	NR	NR	Possible TMA	NR	0
10	NA	NA	0	0	0	0	0	0	0	0	7	ABMR	NR	TMA	NR	0
11	-	-	1	0	0	2	0	0	0	0	382	NR	ABMR	ABMR, TMA	NR	1
12	NA	NA	0	0	0	0	0	0	0	0	11	ABMR	NR	TMA	NR	0
13	NA	NA	0	3	3	3	1	0	2	0	2640	ABMR	NR	TMA	pTCMR	1
14	NA	-	0	0	2	3	1	0	0	0	3297	NR	NR	TMA, AIN	pTCMR	1
15	-	NA	1	1	0	1	0	0	1	0	3029	NR	NR	TMA	NR	0
16	-	NA	0	0	NA	2	0	0	0	-	795	NR	NR	TMA	NR	0
17	NA	NA	0	0	0	1	0	0	0	0	70	ABMR	NR	TMA, GN	NR	0
18	+	NA	0	0	0	0	0	0	0	1	24	NR	NR	TMA, AKI	NR	0
19	NA	NA	0	0	0	0	0	0	0	3	10	NR	ABMR	ABMR, TMA	NR	1
20	NA	NA	0	0	1	1	0	0	0	0	10	NR	NR	TMA, AKI	NR	0
21	+	NA	1	0	0	0	0	0	0	3	9	NR	ABMR	ABMR, TMA	ABMR	0

NA – the data for cells marked 'NA' was unknown, missing, not done, or not recorded by the center.

**NOTE.** Shaded cells in the discrepancy column are discrepant. 7/21 cases were IH discrepant.

**CHAPTER 5:** Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

**Table 5.4** Agreement between AutoBanff and ExpertBanff in N=1679 (6 classes)

		ExpertBanff diagnoses						
		ABMR	Mixed	NR	pABMR	pTCMR	TCMR	TOTAL
AutoBanff diagnoses	ABMR	209	3	30*	8	3	3	256
	Mixed	6	23	3	0	0	7	39
	NR	59*	2	769	39**	5	2	876
	pABMR	49**	2	40**	21	8	2	122
	pTCMR	5	3	77**	15	110	17	227
	TCMR	5	23*	20*	1	2	108	159
	TOTAL	333	56	939	84	128	139	1679
Fraction (%) of ExpertBanff discrepant with AutoBanff		124/333 (37%)	33/56 (59%)	170/939 (18%)	63/84 (75%)	18/128 (14%)	31/139 (22%)	439/1679 (26%)

NOTE. Grey cells along the diagonal represent agreement between ExpertBanff and AutoBanff diagnoses.

\*Clear discrepancies with  $\geq 20$  counts.

\*\*Boundary discrepancies with  $\geq 20$  counts.

**CHAPTER 5:** Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

**Table 5.5** All clear IH discrepancies ordered from most to least frequent (N=202).

<b>Assigned Diagnoses</b>		<b>Frequency in 202 count (% of 202)</b>
<b><u>ExpertBanff</u></b>	<b><u>AutoBanff</u></b>	
ABMR	NR	59 (29%)
NR	ABMR	30 (15%)
Mixed	TCMR	23 (11%)
NR	TCMR	20 (10%)
pABMR	pTCMR	15 (7%)
pTCMR	pABMR	8 (4%)
TCMR	Mixed	7 (3%)
ABMR	Mixed	6 (3%)
ABMR	TCMR	5 (2%)
ABMR	pTCMR	5 (2%)
Mixed	ABMR	3 (1%)
Mixed	pTCMR	3 (1%)
NR	Mixed	3 (1%)
pTCMR	ABMR	3 (1%)
TCMR	ABMR	3 (1%)
Mixed	NR	2 (1%)
Mixed	pABMR	2 (1%)
TCMR	NR	2 (1%)
TCMR	pABMR	2 (1%)
pABMR	TCMR	1 (0.5%)

**Table 5.6** All boundary IH discrepancies ordered from most to least frequent (N=237)

<b>Assigned Diagnosis</b>		<b>Frequency in 237 count (% of 237)</b>
<b><u>ExpertBanff</u></b>	<b><u>AutoBanff</u></b>	
NR	pTCMR (borderline)	77 (32%)
ABMR	pABMR	49 (21%)
NR	pABMR	40 (17%)
pABMR	NR	39 (16%)
TCMR	pTCMR	17 (7%)
pABMR	ABMR	8 (3%)
pTCMR	NR	5 (2%)
pTCMR	TCMR	2 (1%)

**CHAPTER 5:** Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

**Table 5.7** The eight most frequent ( $\geq 20$ ) IH discrepancies out of 439 total discrepancies, classified as clear or boundary.

Diagnosis		Frequency in 439 count (%)	Clear or boundary
ExpertBanff	AutoBanff		
NR	pTCMR	77 (18%)	boundary
ABMR	NR	59 (13%)	clear
ABMR	pABMR	49 (11%)	boundary
NR	pABMR	40 (9%)	boundary
pABMR	NR	39 (9%)	boundary
NR	ABMR	30 (7%)	clear
Mixed	TCMR	23 (5%)	clear
NR	TCMR	20 (5%)	clear

**Table 5.8** BK virus versus TCMR diagnoses by ExpertBanff and/or AutoBanff (N=190)

	<b>E = A</b>	<b>E ≠ A</b>	<b>TOTAL</b>	<b>E+ / A-</b>	<b>E- / A+</b>
<b>BK virus positive</b>	1	17	<b>18</b>	2	15
<b>BK virus negative</b>	107	65	<b>172</b>	29	36
<b>TOTAL</b>	<b>108</b>	<b>82</b>	<b>190</b>	31	51

Chi squared 2x2 agreement versus disagreement in shaded cells p value =  $1 \times 10^{-5}$ .  
 E = ExpertBanff. A = AutoBanff. A positive diagnosis ("TCMR") is marked by '+', while '-' denotes a negative diagnosis ("not TCMR").

**CHAPTER 5:** Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

**Table 5.9** ci 0/1 versus 2/3 lesions in rejection diagnoses by ExpertBanff and/or AutoBanff

ci 0/1 versus 2/3 lesions versus diagnoses of TCMR by ExpertBanff and/or AutoBanff (N=179) Chi-squared 2x2 agreement versus disagreement in shaded cells p value = 0.27					
	<b>E = A</b>	<b>E ≠ A</b>	<b>TOTAL</b>	<b>E+ / A-</b>	<b>E- / A+</b>
<b>ci lesion score = 0 or 1</b>	65	43	<b>108</b>	14	29
<b>ci lesion score = 2 or 3</b>	36	35	<b>71</b>	14	21
<b>TOTAL</b>	<b>101</b>	<b>78</b>	<b>179</b>	28	50

ci 0/1 versus 2/3 lesions versus diagnoses of ABMR by ExpertBanff and/or AutoBanff (N=366) Chi-squared 2x2 agreement versus disagreement in shaded cells p value = 0.83					
	<b>E = A</b>	<b>E ≠ A</b>	<b>TOTAL</b>	<b>E+ / A-</b>	<b>E- / A+</b>
<b>ci lesion score = 0 or 1</b>	106	85	<b>191</b>	62	23
<b>ci lesion score = 2 or 3</b>	100	75	<b>175</b>	52	23
<b>TOTAL</b>	<b>206</b>	<b>160</b>	<b>366</b>	114	46

E = ExpertBanff. A = AutoBanff. A positive diagnosis ("ABMR") is marked by '+', while '-' denotes a negative diagnosis ("not ABMR").

**CHAPTER 5:** Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

**Table 5.10** DSA positive, DSA negative, or DSA ambiguous/not recorded versus ABMR diagnoses by ExpertBanff and/or AutoBanff (N=380)

	<b>E = A</b>	<b>E ≠ A</b>	<b>TOTAL</b>	<b>E+ / A-</b>	<b>E- / A+</b>
<b>DSA positive</b>	209	76	<b>285</b>	29	47
<b>DSA negative</b>	0	60	<b>60</b>	60	0
<b>DSA ambiguous or not recorded</b>	0	35	<b>35</b>	35	0
<b>TOTAL</b>	<b>209</b>	<b>171</b>	<b>380</b>	124	47

Chi squared 2x2 agreement versus disagreement in shaded cells p value <  $2 \times 10^{-16}$ .

E = ExpertBanff. A = AutoBanff. A positive diagnosis ("ABMR") is marked by '+', while '-' denotes a negative diagnosis ("not ABMR").

**CHAPTER 5:** Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

**Table 5.11** v>0 versus v=0 lesions in rejection diagnoses by ExpertBanff and/or AutoBanff

v>0 versus v=0 lesions and TCMR diagnoses by both ExpertBanff and AutoBanff (N=181).  
Chi squared 2x2 agreement versus disagreement in shaded cells p value = 0.80.

	<b>E = A</b>	<b>E ≠ A</b>	<b>TOTAL</b>	<b>E+ / A-</b>	<b>E- / A+</b>
<b>v lesion score = 0</b>	71	56	<b>127</b>	27	29
<b>v lesion score = 1, 2 or 3</b>	32	22	<b>54</b>	3	19
<b>TOTAL</b>	<b>103</b>	<b>78</b>	<b>181</b>	30	48

v>0 versus v=0 lesions and Mixed diagnoses by both ExpertBanff and AutoBanff (N=65)  
Chi squared 2x2 agreement versus disagreement in shaded cells p value = 0.11

	<b>E = A</b>	<b>E ≠ A</b>	<b>TOTAL</b>	<b>E+ / A-</b>	<b>E- / A+</b>
<b>v lesion score = 0</b>	7	24	<b>31</b>	20	4
<b>v lesion score = 1, 2 or 3</b>	15	19	<b>34</b>	8	11
<b>TOTAL</b>	<b>22</b>	<b>28</b>	<b>65</b>	28	15

v>0 versus v=0 lesions and ABMR diagnoses by both ExpertBanff and AutoBanff (N=348)  
Chi squared 2x2 agreement versus disagreement in shaded cells p value = 0.001

	<b>E = A</b>	<b>E ≠ A</b>	<b>TOTAL</b>	<b>E+ / A-</b>	<b>E- / A+</b>
<b>v lesion score = 0</b>	191	147	<b>338</b>	102	45
<b>v lesion score = 1, 2 or 3</b>	0	10	<b>10</b>	10	0
<b>TOTAL</b>	<b>191</b>	<b>112</b>	<b>348</b>	112	45

v>0 versus v=0 lesions and all ABMR/TCMR/Mixed diagnoses by ExpertBanff and/or AutoBanff (N=551)  
Chi squared 2x2 agreement versus disagreement in shaded cells p value = 0.56

	<b>E = A</b>	<b>E ≠ A</b>	<b>TOTAL</b>		
<b>v lesion score = 0</b>	269	205	<b>474</b>		
<b>v lesion score = 1, 2 or 3</b>	47	30	<b>77</b>		
<b>TOTAL</b>	<b>316</b>	<b>235</b>	<b>551</b>		

E = ExpertBanff. A = AutoBanff. A positive diagnosis ("Rejection") is marked by '+', while '-' denotes a negative diagnosis ("not Rejection").

**CHAPTER 5:** Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

**Table 5.12** Cases with v>0 lesions called NR by ExpertBanff (N=9)

Case #	DSA	Lesion scores							C4d	IRRATs score	Time of Biopsy post-transplant (days)	MMDx Diagnosis	ExpertBanff Diagnosis (6 class)	ExpertBanff Diagnosis (multiclass)	AutoBanff Diagnosis
		g	cg	i	ci	t	v	ptc							
1	+	2	0	0	3	1	1	0	0	0.636686	1066	Mixed	NR	IFTA	Mixed
2	-	NA	NA	3	NA	3	1	NA	0	1.453987	272	TCMR	NR	BK	TCMR
3	-	0	0	1	2	3	1	0	NA	0.218061	183	NR	NR	BK	TCMR
4	-	0	NA	0	1	1	1	0	0	0.392809	92	NR	NR	BK	TCMR
5	NA	0	0	0	1	2	1	0	1	0.124151	427	NR	NR	Normal	TCMR
6	-	0	1	1	1	1	2	0	0	0.846598	1860	Mixed	pABMR	TG	TCMR
7	+	0	0	0	NA	0	1	3	3	0.114186	790	ABMR	NR	Normal	Mixed
8	+	1	0	NA	1	0	1	1	0	0.300222	45	NR	NR	Normal	Mixed
9	-	0	0	1	1	1	1	1	0	1.993172	12	NR	pTCMR	Borderline	TCMR

NA – The feature was not done or not reported by the center for that biopsy.

**CHAPTER 5:** Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

**Table 5.13** Tubulitis lesions (t) 0/1 versus 2/3 lesions versus diagnoses of any rejection (ABMR, TCMR, Mixed) by ExpertBanff and/or AutoBanff (N=593)

	<b>E = A</b>	<b>E ≠ A</b>	<b>TOTAL</b>
<b>t lesion score = 2 or 3</b>	122	71	<b>193</b>
<b>t lesion score = 0 or 1</b>	213	172	<b>385</b>
<b>TOTAL</b>	<b>335</b>	<b>243</b>	<b>578</b>

Chi-square 2x2 shaded cells p value = 0.08.  
E = ExpertBanff. A = AutoBanff.

**CHAPTER 5:** Discrepancy analysis comparing SOC histologic diagnoses and the AutoBanff algorithm

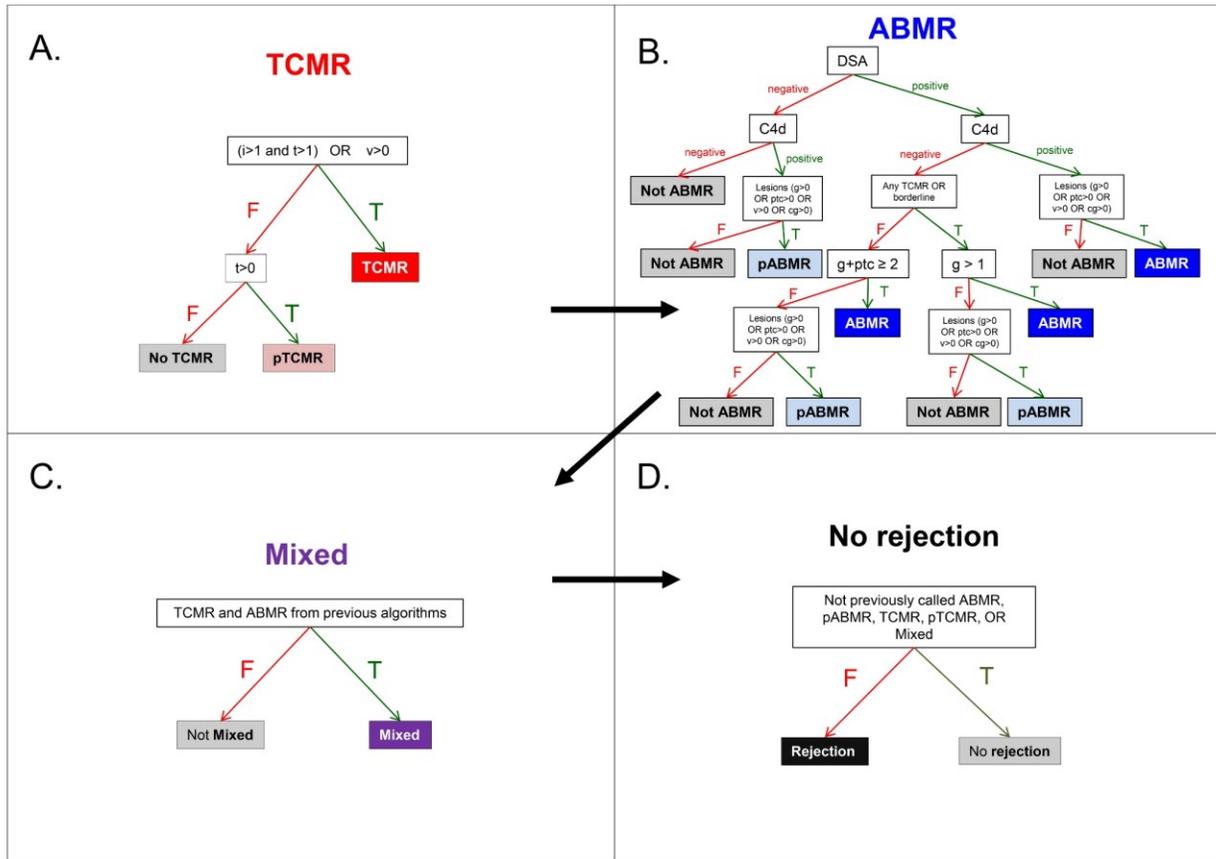
**Table 5.14** Summary of previously noted problematic guidelines that contribute to IH discrepancies

Condition	p-value
Positive BK virus in TCMR cases increases the frequency of discrepancies	1x10 <sup>-5</sup>
ci 2/3 lesions in rejection (TCMR, ABMR) biopsies does not change the frequency of discrepancies	TCMR = 0.27 ABMR = 0.83
DSA negative/not recorded/ambiguous in ABMR cases increases the frequency of discrepancies	<2x10 <sup>-16</sup>
v-lesions in all rejection biopsies (by ExpertBanff and/or AutoBanff) does not affect the frequency of discrepancies v-lesions in ABMR biopsies increases the frequency of discrepancies *v-lesions in TCMR biopsies does not affect the frequency of discrepancies *v-lesions in Mixed rejection biopsies does not affect the frequency of discrepancies	All rejection = 0.56 ABMR = 0.001 *TCMR = 0.80 *Mixed = 0.11
*t 2/3 lesions in all rejection biopsies (by ExpertBanff and/or AutoBanff) does not affect the frequency of discrepancies	*0.08

\* Indicates data not shown.

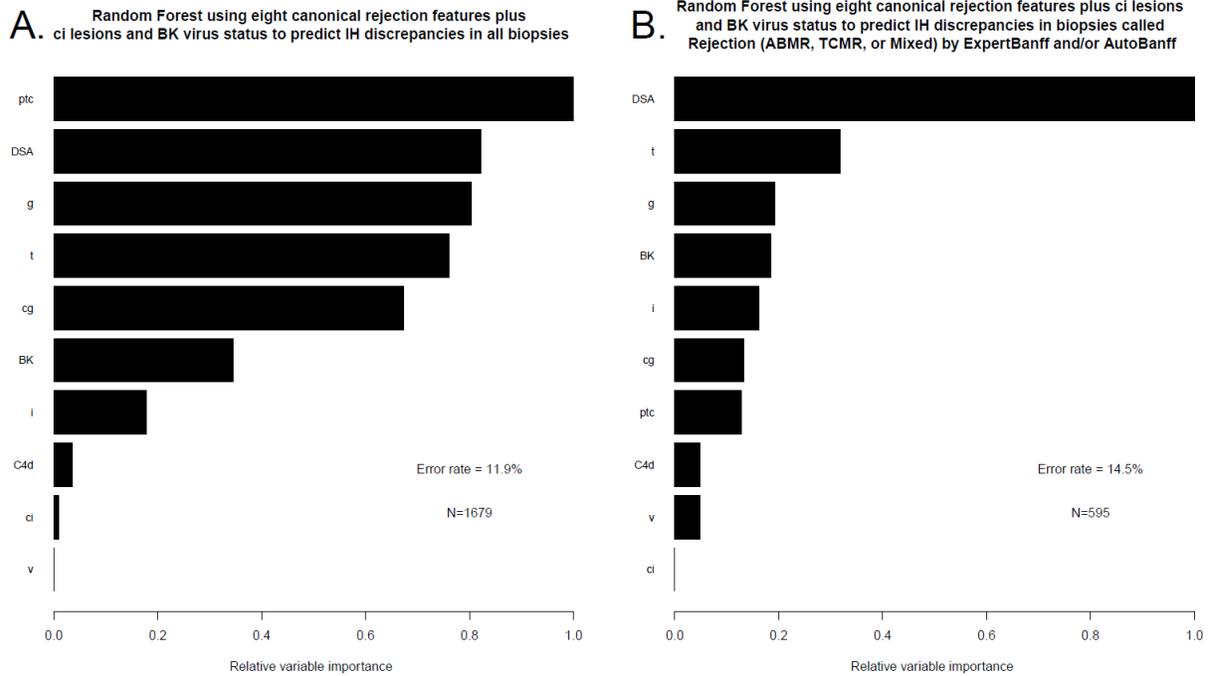
## **CHAPTER 5: Kidney IH Discrepancies**

### **5.11 FIGURES**

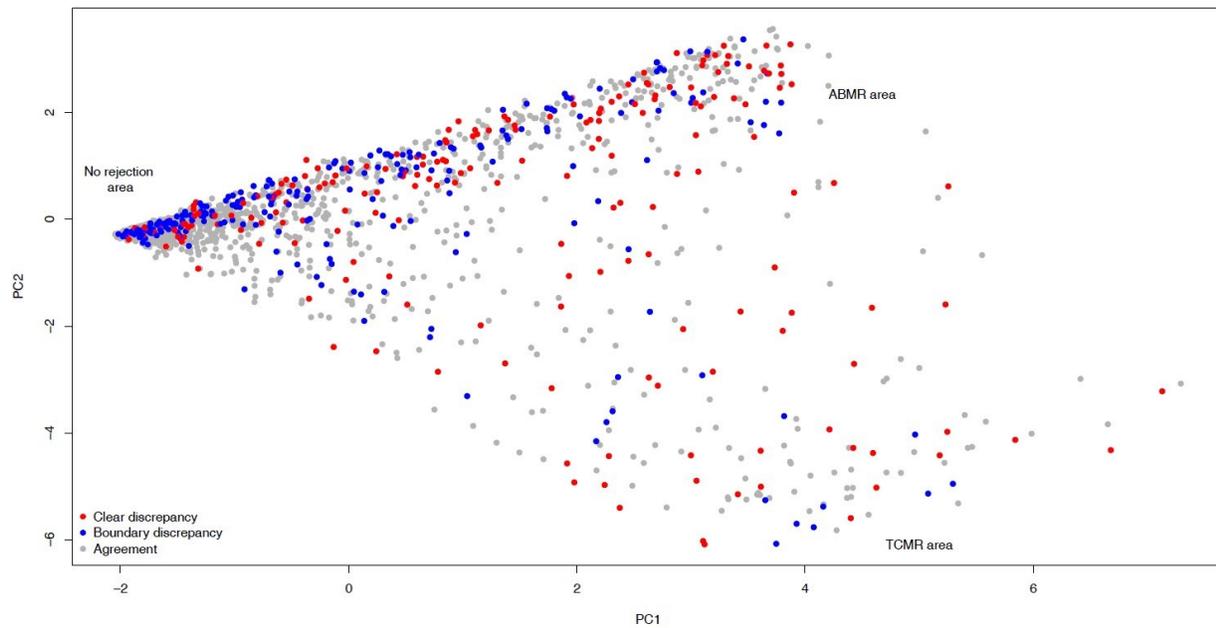


**Figure 5.1** The algorithm for assigning AutoBanff diagnoses using the eight canonical rejection lesion scores assigned by the center pathologists. Algorithm was based on the most recent iteration of the Banff guidelines (2017 Meeting Report, see References). The algorithm works by elimination, and moves through panels A, B, C, and D in that order.

## CHAPTER 5: Kidney IH Discrepancies



**Figure 5.2 Random forest analysis using the eight canonical rejection features to predict IH discrepancies.** Variable importance indicates that ABMR-related lesions contribute most to discrepancies between ExpertBanff and AutoBanff.



**Figure 5.3** Distribution of the AutoBanff versus ExpertBanff discrepancies in biopsies, based on their Molecular Microscope® (MMDx) phenotype. Note that many discrepancies are in areas that are molecularly unequivocal as NR, ABMR, and TCMR.

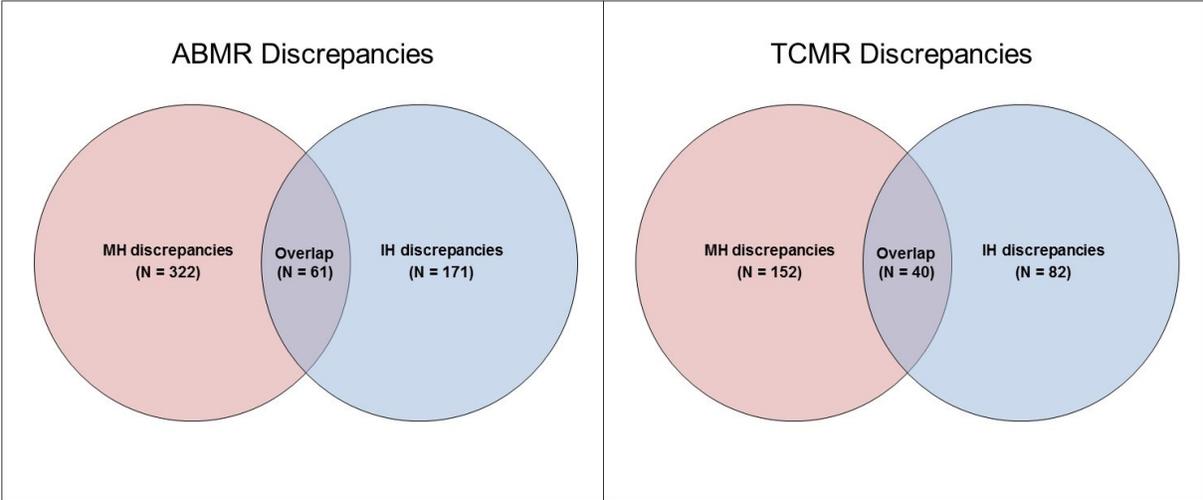


Figure 5.4 Venn diagram showing the overlap between MH discrepancies and IH discrepancies in the context of ABMR and TCMR. Most MH discrepancies are not IH discrepancies.

## **SECTION II**

### **ANALYSES IN LIVER TRANSPLANT BIOPSIES**

## **CHAPTER 6**

### **ASSESSING TRANSPLANT REJECTION IN LIVER BIOPSIES: THE DEVELOPMENT OF MMDX FOR DIAGNOSTICS WITHIN A LIVER TRANSPLANT POPULATION**

**This chapter was published**

Madill-Thomsen KS, Halloran PF, et al. The molecular diagnosis of rejection in liver transplant biopsies: First results of the INTERLIVER study. *Am J Transplant*. 2020 Feb (epub ahead of print). Used with permission of Wiley Publishing.

## CHAPTER 6: Liver rejection

### 6.1 OVERVIEW AND HYPOTHESIS

Diagnosis of rejection in liver transplantation remains an important issue in clinical management.<sup>63,119,255,256</sup> The current SOC for liver biopsy diagnoses is histology, generally following Banff guidelines.<sup>209</sup> Histology is based on pattern recognition by experts, and assessments differ between observers.<sup>4,125,127–129,257</sup> Reported kappa values for pathology related to TCMR are low to moderate (0.15–0.62),<sup>111</sup> especially when comorbidities are present,<sup>257</sup> leaving an unmet need for improvement in precision. Moreover, the diagnosis and prevalence of ABMR in liver transplants remain controversial.<sup>82–84,209</sup> Liver transplants present unique challenges because of their tolerogenic properties, inviting clinicians to consider reducing immunosuppression.<sup>130–134</sup> However, this practice requires a precise and accurate system for diagnosing rejection.<sup>117,118,258–260</sup> Liver function test abnormalities are associated with rejection but cannot distinguish TCMR from other diseases such as steatohepatitis.<sup>91,92</sup>

Molecular measurement of gene expression using microarrays coupled with machine learning has the potential to improve the assessment of transplant biopsies by overcoming the limitations of conventional diagnostics.<sup>1</sup> Previous data-driven molecular technologies – the Molecular Microscope<sup>®</sup> Diagnostic System (MMDx) - in kidney,<sup>6,64,66</sup> heart,<sup>198,261</sup> and lung<sup>262–264</sup> transplant biopsies have produced objective diagnoses that often differ from histology,<sup>36,66,171,262</sup> and could improve the assessment of liver transplant biopsies.<sup>63,265</sup> A number of factors argue that MMDx is more accurate than histology:<sup>6</sup> e.g. use of continuous quantitative measurements,<sup>6</sup> low sampling error,<sup>64</sup> high reproducibility,<sup>64</sup> and lack of requirements for specific tissue elements such as glomeruli, cortex, or portal triads. MMDx predicts outcomes better than histologic assessments.<sup>71,212,249,266–268</sup> Many major rejection and injury features initially described in kidney transplants<sup>19,66,211</sup> are shared with heart<sup>261,269–271</sup> and lung transplants,<sup>263,272</sup> indicating that rejection changes are not organ-specific and are likely shared by liver transplants. Existing studies have established the potential utility of molecular assessment of liver biopsies.<sup>117–121,123,124,273–275</sup>

These analyses aimed to develop an MMDx system for diagnosing rejection in liver transplant biopsies using microarray measurements of gene expression plus machine learning to interpret the output in terms of injury and rejection. Unsupervised analysis used expression of rejection-associated transcripts (RATs, derived in kidney transplants<sup>276</sup> and validated in heart and lung) to evaluate the potential for supervised analyses based on histology diagnoses.

### 6.2 BIOPSY POPULATION

The biopsy population used in these analyses consisted of 235 biopsies prospectively collected from 217 liver transplant patients in 10 international centers during the INTERLIVER study (ClinicalTrials.gov NCT03193151, **Table 6.1**). All biopsies were stabilized in RNA<sup>later</sup><sup>™</sup> and shipped at ambient temperature to the ATAGC laboratory for analysis per established protocols.<sup>66</sup> Clinical data were reported by the participating centers per the SOC. Since DSA testing is not SOC at most centers, a majority of biopsies were not accompanied by DSA assessment. Histologic classifications were assigned locally at the centers.

Indication biopsies accounted for 72% of the population (170/235), with the remainder taken for various protocols (**Table 6.2**). Median TxBx was 962 days (range 0-11676 days). DSA testing per center practice was performed in only 17 biopsies as per center practice (10 positive, **Table 6.3**). RNA extracted from these biopsies (average length 4.5 mm, range 0.5-14mm of 16- or 18-gauge core) was high yield (average 9.91µg RNA, range 0.28-73µg) and high quality (average RIN number 8.2). Biological replicates showed high reproducibility on the molecular report (data not shown but similar to published results in kidney).<sup>64</sup> Technical replicates in other transplanted organs confirmed the reliability of the techniques used in the biopsy processing.<sup>64</sup>

### 6.3 ABMR AND TCMR RAT TRANSCRIPT SEPARATION IN LIVER BIOPSIES

#### 6.3.1 Derivation of the RAT Transcript Set.

RATs were derived in kidney transplants and were validated in heart<sup>171,198,277</sup> and lung transbronchial and mucosal biopsies.<sup>262,264</sup> The RATs are the union of the top 200 Affymetrix probe sets associated with the three class comparisons based on histology labels: all-rejection (ABMR, TCMR, or Mixed) versus everything else (EE), ABMR versus EE, and TCMR versus EE.<sup>278</sup> (All-rejection and TCMR transcripts were also identified in mouse TCMR models.)<sup>211</sup> After removing duplicates, 453 RAT transcripts remained. Some transcripts were identified by more than one algorithm, producing six classes

## CHAPTER 6: Liver rejection

of RATs shown in **Figure 6.1**: ABMR-selective (blue), TCMR-selective (red), all-rejection (green): ABMR/all-rejection (cyan), TCMR/all-rejection (orange), and ABMR/TCMR/ and all-rejection (black). IQR filtering was performed with a cutoff of 0.35 to remove low variance transcripts, producing a final set of 417 transcripts.

### 6.3.2 PCA visualization of RAT Expression in Kidneys, Hearts and Livers.

PCA was used to plot the entire population in terms of the expression of RAT genes. PC1, 2, and 3 comprised 66%, 7%, and 3% of the variation respectively.

**Figure 6.2** shows the correlation of each RAT transcript with PC1, PC2, and PC3 for the liver biopsies, compared to published analyses for kidney and heart biopsies,<sup>171</sup> adapted to permit comparison with liver biopsies. PCs for Figure 6.2 were established based on the RAT expression within each organ (kidney N=1526, heart N=1320, and liver N=235). The location of the transcripts in this factor map were used to establish an appropriate interpretation of the actual biopsies in PCA: for example, if TCMR-RATs are located towards the right side of the plot, biopsies located to the right will likely have molecular TCMR.

In each organ population, Rejection-RATs and TCMR-RATs were strongly associated with PC1. In kidneys (**Figure 6.2A/6.2B**) and hearts (**Figure 6.2C/6.2D**), the ABMR-RATs (blue) and ABMR-all-rejection RATs (cyan) vertically separated from the TCMR-RATs (red) in PC2, indicating distinct ABMR and TCMR rejection states. ABMR/all-rejection RATs (cyan) were particularly instructive because they are strongly associated with ABMR and separate from TCMR-RATs (red) in kidney and heart. There was no ABMR-TCMR separation of the RATs in liver biopsies (**Figure 6.2E/6.2F**), this finding is particularly clear in PC2 versus PC3 (**Figure 6.2F**).

RATs not trained in hearts still separated TCMR from ABMR in PC2 in the heart population, thus the lack of separation between ABMR-RATs and TCMR-RATs in liver biopsies was particularly instructive. Assuming that liver ABMR is molecularly similar to heart and kidney ABMR, this finding argues against any sizeable ABMR population in this liver biopsy cohort. Instead, this result suggests that ABMR in liver transplantation is rare, is very different from the ABMR seen in kidney and heart transplants, or has increasingly subtle signals – possibly due to the tolerogenic properties of the liver.

## **6.4 AA OF LIVER REJECTION USING RAT TRANSCRIPT EXPRESSION**

### **6.4.1 AA in livers for unsupervised analysis.**

AA was used to assign scores to each biopsy and separate groups based on their expression of RATs. AA was performed using the ‘archetypes’ package for R version 1.1.463.<sup>200</sup> Biopsies were given a set of four archetype scores (R1, R2, R3, and R4 scores) that define their relationship to the four idealized archetypes. Per convention, biopsies were assigned to groups (R1, R2, R3, and R4) based on their highest score.

### **6.4.2 Selection of the rejection-based clusters for AA in the liver transplant population.**

**Figure 6.3** is a scree plot showing the residual sum of squares (y-axis) versus the potential number of archetypes in a model (x-axis). Combined with our expectations based on AA in kidney, heart, and lung,<sup>201,261,262</sup> we selected four archetypes (highlighted in red): no rejection or injury, rejection (specifically TCMR), early mild injury from donation-implantation, and late changes potentially representing atrophy-fibrosis. Because injury activates innate immunity, which shares mechanisms with adaptive immunity, RAT expression is expected to identify some injured biopsies.

### **6.4.3 Visualization of the liver transplant rejection AA clusters in PCA.**

Each biopsy was assigned to an archetype group: R1, N=129, R2, N=37, R3, N=61, and R4, N=8.

**Figure 6.4** shows the four archetypes and groups of biopsies distributed in PCA, colored according to their assigned archetype group: PC2 versus PC1 (**Figure 6.4A**) and PC2 versus PC3 (**Figure 6.4B**). Early biopsies taken within two weeks of transplant (triangle symbols) were exclusively in the upper regions of PC2. PC1 separated R2 from R1. PC2 separated R3 biopsies, many of which were early post-transplant. PC3 separated R4 from everything else.

### **6.4.4 Characterization of the AA clusters.**

#### **6.4.4.1 PBTs selected for analysis**

## CHAPTER 6: Liver rejection

PBTs associated with biological mechanisms in rejection and injury were previously annotated in human cell lines, mouse experimental models, and human transplant biopsies.<sup>187</sup> PBTs have been extensively used in liver transplant studies.<sup>279,280</sup> PBT values represent the mean fold change in expression compared to a selected control group. Biopsies assigned to the R1 group by highest archetype score were used as the control group.

For these analyses we selected the following PBTs: ABMR-RATs,<sup>171</sup> BATs,<sup>65</sup> cIRITs,<sup>281</sup> DAMPs,<sup>282</sup> eDSASTs,<sup>191</sup> ENDATs,<sup>192</sup> GRITs,<sup>193</sup> IGTs,<sup>65</sup> IRITD3 and IRITD5,<sup>194</sup> IRRATs,<sup>195</sup> MCATs,<sup>197</sup> QCATs,<sup>24</sup> RATs,<sup>171</sup> Rej-RATs,<sup>171</sup> TCMR-RATs,<sup>171</sup> and FICOLs.

### 6.4.4.2 PBT expression in the RAT-based AA clusters

In **Table 6.4**, the median TxBx of each RAT-based group differed between archetype groups: the earliest in R3 (99 days), and R2 (214 days). R1 was much later at 2534 days and R4 at 3117 days. R1 biopsies lacked expression of transcripts associated with rejection and injury compared to the other three groups. R2 biopsies had the highest expression of TCMR-selective (TCMR-RATs), effector T cell (QCATs), and rejection-related PBTs (IFNG-induced GRITs and all-rejection-RATs), and increased expression of injury-related PBTs, particularly DAMPs.

R3 biopsies had increased expression of the injury-induced PBTs compared to R1 (e.g. IRRATs), with no expression of rejection-related PBTs. R4 biopsies had increased injury-related, moderate elevation of rejection-related, and high expression of endothelium-related and atrophy-fibrosis-related PBTs (IGTs, BATs, and MCATs). Note that the IGTs, BATs, and MCATs have previously been associated with a fibrotic phenotype in liver transplants.<sup>113,275</sup>

Based on these characteristics, the archetypal groups were named 'R1<sub>normal</sub>' for the relatively normal phenotype, 'R2<sub>TCMR</sub>', for the group with high expression of TCMR-like transcripts, 'R3<sub>injury</sub>' for the early group with injury, and 'R4<sub>late</sub>' for the group with transcripts characteristic of atrophy-fibrosis.

### 6.4.5 Transcripts correlated with each RAT AA score.

The top 10 transcripts correlated with each archetype score are summarized in **Table 6.5**. Based on previous experience, the hypothesis was that transcripts lower in R1 and increased in R2-R4

## CHAPTER 6: Liver rejection

compared to the other groups would be the most revealing.<sup>66,171,201</sup> The top 30 unique transcripts by Spearman correlation coefficient in R1-R4 are compiled in **Tables 6.6-6.9**.

The R1<sub>normal</sub> score was associated with low expression of transcripts previously annotated for associations with transplant injury or rejection (e.g. *TRIP1*, *MIR21*, and *IFNGR1*). Top transcripts correlated with the R2<sub>TCMR</sub> score were typical of TCMR i.e. mainly all-rejection- and TCMR-associated s (e.g. *STAT1*, *CXCL11*, and *GBP5*), and all were IFNG-inducible, a hallmark of transplant rejection. Top transcripts correlated with the R3<sub>injury</sub> score reflected injury, including hypoxia-inducible factor *EGLN1*, and many were previously annotated as increased in recent injury models (e.g. *PVR*, *PTPN11*, and *SERPINA3*). Top transcripts correlated with the R4<sub>late</sub> score were previously annotated in injury models (e.g. *JAM2*).

### 6.5 CHARACTERIZING THE PRINCIPAL COMPONENTS (PCS) IN LIVER TRANSPLANT REJECTION

PC1 correlated with rejection (e.g. TCMR-RATS, QCATs, GRITs) and parenchymal injury (e.g. DAMPs). PC2 correlated with early TxBx, mild IFNG effects, and mild to moderate elevation of injury parameters (e.g. IRITs). PC3 correlated with endothelial-related and late injury-related PBTs (e.g. IGTs, BATs, MCATs), moderate elevation of many injury-related PBTs (e.g. IRRATs) but without rejection-related PBTs (e.g. TCMR-RATs) (**Table 6.10**).

### 6.6 TIME COURSE OF MOLECULAR AA AND PCA SCORES IN THE LIVER TRANSPLANT POPULATION

Relationships between molecular scores and TxBx were explored using moving average plots (**Figure 6.5A**). Moving average plots were generated in R version 3.5.1 using the 'zoo' package.<sup>207</sup> A window size of 75 was used for each graph. Plots were left-aligned and scores were normalized on all 235 biopsies.

Average standardized R3<sub>injury</sub> and R2<sub>TCMR</sub> scores were highest in early biopsies and fell steadily to low levels by 2000 days. Average standardized R1<sub>normal</sub> scores were initially low but rose steadily to be the dominant score in biopsies after 1000 days, thus later biopsies frequently had a relatively normal

## CHAPTER 6: Liver rejection

phenotype as they recovered from early donation-implantation injury and as TCMR became less common. High  $R4_{\text{late}}$  scores became more common in later years post-transplant.

Like the  $R2_{\text{TCMR}}$  score, PC1 declined slowly over time, likely related to the resolution of early injury and the decline in TCMR. PC2, which is associated with recent injury such as that occurring in donation-implantation, declined sharply over time like the  $R3_{\text{injury}}$  score. PC3, reflecting late injury, increased steadily like the  $R4_{\text{late}}$  score (**Figure 6.5B**).

### 6.7 CORRELATIONS WITH THE SOC HISTOLOGIC AND CLINICAL DATA

#### 6.7.1 Correlations between molecular scores and histologic features

SOC histology diagnoses of rejection in these centers was usually reported as lesion grades plus text commentary rather than discrete classes (rejection versus no rejection). For comparison with MMDx, biopsies were classified as histologic rejection (TCMR) versus NR, on the basis of their summed portal, bile duct, and venous inflammation.<sup>112,209</sup> Biopsies were considered positive for histologic rejection if the sum of these grades was  $>0$  (alternate analyses used  $>1$  or  $>2$ ). Biopsies with inadequate information were excluded (N=15).

Correlations were studied between the rejection archetype scores, PCA, and the recorded histologic features (**Table 6.11**). Relationships with SOC histologic features were considered significant if their Spearman correlation coefficient values were above the arbitrarily assigned cutoff of 0.2. All correlations above this cutoff had significant p values ( $p < 0.001$ ).

PC1 and  $R2_{\text{TCMR}}$  scores correlated with TCMR-related lesions (portal/bile duct inflammation), and histologic diagnoses of CMV hepatitis. PC2 and  $R3_{\text{injury}}$  scores, which are features of early biopsies, negatively correlated with fibrosis and positively correlated with recurrent HCV. PC3 and  $R4_{\text{late}}$  scores, features of late biopsies, positively correlated with fibrosis.

#### 6.7.2 Relationships between molecular scores and liver biochemistry measurements

In **Table 6.12**, the  $R1_{\text{normal}}$  group had the highest albumin values and the lowest bilirubin, AST, ALT, and ALP. The  $R2_{\text{TCMR}}$  and  $R3_{\text{injury}}$  groups both had abnormal mean values:  $R2_{\text{TCMR}}$  had the highest

## CHAPTER 6: Liver rejection

ALP, and R3<sub>injury</sub> had the highest bilirubin and AST. The R4<sub>late</sub> group had the lowest albumin, highest bilirubin, and highest ALP.

### 6.7.3 Relationships between molecular scores and histologic rejection

Given the associations of the R2<sub>TCMR</sub> group with histologic rejection features, we performed t-tests comparing gene expression in biopsies with and without histologic rejection as defined by the >0 threshold. The top 30 differentially expressed genes are shown in **Table 6.13** (adjusted p value range 0.004-0.001). Transcripts associated with the R2<sub>TCMR</sub> score had much smaller p-values (**Table 6.7**,  $p=1.7 \times 10^{-84} - 1.2 \times 10^{-56}$ )

A supervised classifier based on histology rejection lesions >0 was developed using linear discriminant analysis (lda)-based machine learning. The AUC for the molecular rejection scores derived through 10-fold CV for predicting histologic rejection was only 0.57, lower than that for the unsupervised R2<sub>TCMR</sub> archetype score (0.70, AUCs significantly different,  $p=0.002$ ). The optimized cutoff for the R2<sub>TCMR</sub> score of 0.2 was established using the 'cutpointr' package in R,<sup>283</sup> to balance sensitivity and specificity and generate a binary prediction per the requirements of AUC performance measurements (**Figure 6.6**).

## 6.8 DIAGNOSTIC INTERPRETATION OF MMDx-LIVER REPORTS

MMDx diagnoses were assigned by one expert (PFH) using a first-generation MMDx-Liver report (**Figure 6.7**). The relationship between the molecular sing-out comments and histologic rejection is analyzed in **Table 6.14**. Use of the >2, >1, or >0 cutoff increased sensitivity but decreased specificity e.g. increased the MMDx no rejection called histologic no rejection from 33% to 80% but also decreased the MMDx rejection called histologic rejection from 89% to 76%, and did not significantly change the balanced accuracy (0.60, 0.63, and 0.62 respectively).

## 6.9 INTERPRETATION OF THE RESULTS

This analysis aimed to develop a first-generation MMDx system for liver transplant rejection using RAT expression, unsupervised PCA, and AA, and examine its relationship to SOC histology assessments

## CHAPTER 6: Liver rejection

in the center. Small pieces of liver transplant biopsy cores placed in RNA/ater™ solution gave excellent yields of high-quality RNA. In RAT-based PCA, PC1 was highly associated with TCMR-like transcripts, and PC2 correlated with early parenchymal injury. AA identified four archetype groups: R1<sub>normal</sub>, R2<sub>TCMR</sub>, R3<sub>injury</sub>, and R4<sub>late</sub>, each with distinctive top transcripts, PBT associations, time courses, and histology associations. No distinct molecular ABMR group was detected in this unselected biopsy population using strategies that detected ABMR in kidney and heart transplants. Liver function biochemistry was most normal in the R1<sub>normal</sub> biopsies and comparatively abnormal in other groups. High R3<sub>injury</sub> and R2<sub>TCMR</sub> scores were common early and became infrequent at later times, while R4<sub>late</sub> scores progressively increased with time and correlated with transcripts associated with atrophy-fibrosis. R2<sub>TCMR</sub> correlated with typical TCMR-like histology changes, and top transcripts reflected IFNG effects e.g. such as *TAP1*, *CXCL9*, and *CXCL11*, hallmarks of TCMR in kidney, heart, and lung transplants and previously documented in liver transplant rejection.<sup>113,275,280</sup> The R4<sub>late</sub> group was distinct, with increased histologic fibrosis and elevated IGTs, MCATs, and BATs, similar to changes in long-term pediatric liver transplants with interface inflammation.<sup>113</sup> Comparison between molecular findings and histology showed significant associations but extensive discrepancies. Therefore, RAT expression permits MMDx to detect TCMR, injury, and a late inflamed fibrosis state in small pieces of liver transplant biopsy cores.

The R1<sub>normal</sub> score identified the biopsies with the absence of inflammation, injury, and fibrosis features and became the dominant phenotype with late TxBx. The predominance of high R1<sub>normal</sub> scores at late time points and high R2<sub>TCMR</sub> and R3<sub>injury</sub> scores at earlier times underscores the unique time course of liver transplants compared with other transplanted organs, recovering from early injury and rejection to achieve long term stability.

R2<sub>TCMR</sub> was associated with typical TCMR transcripts and were relatively early post-transplant (median days post-transplant = 214), consistent with literature showing that the highest risk of TCMR was within the first month post-transplant with an incidence of 20-40%.<sup>284</sup> Thus, R2 and PC1 scores could be used to assess the probability of rejection in a liver biopsy. The TCMR phenotype disappeared over time, perhaps reflecting T cell exhaustion as suggested for other transplants.<sup>23,71</sup>

The R3<sub>injury</sub> biopsies were associated with injury-related PBTs, had the earliest median TxBx, and strongly expressed hypoxia-inducible factor (HIF) *EGLN*, of interest given the role of HIFs in liver

## CHAPTER 6: Liver rejection

donation-implantation injury.<sup>285</sup> In contrast, R4<sub>late</sub> biopsies were strongly associated with IGTs and BATs and MCATs, reflecting the mild inflammatory infiltrate common in tissues undergoing atrophy-fibrosis. As previously mentioned, the R4<sub>late</sub> group resembled biopsies with interface inflammation,<sup>113</sup> with high expression of IGTs, MCATs, and BATs as well as a low level of TCMR-like inflammation. Liver fibrosis has been associated with B cell activity in other studies.<sup>286,287</sup> The question remains whether this is a smoldering cognate alloimmune response (TCMR) or a late response-to-wounding. This group was small but minor increases in R4 score in R1/R2/R3 biopsies were common. Full exploration of late fibrosis-inflammation characteristics in liver transplants will require further analyses incorporating a wider range of transcripts than simply the RATs.

While the incidence of ABMR in liver transplants remains unresolved, the present study does not support the concept of a common liver transplant ABMR state analogous to that in heart and kidney transplants. Liver ABMR phenotypes have been contentious since the beginnings of liver transplantation. Current versions of MMDx cannot make the claim of being able to identify ABMR in troubled liver transplants, because the diagnosis was not made in these centers, and because the unsupervised search in Figure 6.2 did not reveal separation of ABMR-RAT transcripts from TCMR-RAT transcripts. Most of the centers contributing biopsies to these analyses do not believe that an ABMR phenotype is identifiable in their population, at least not with any frequency. As approved by the IRBs, the study was precluded from changing the SOC in participating centers thus could not dictate that DSA testing be done if this was not the standard in each center. Previously annotated endothelial transcripts associated with ABMR in kidney transplant biopsies were elevated in R4<sub>late</sub> biopsies, but this probably reflects angiogenesis in grafts undergoing fibrogenesis and actively forming atrophy-fibrosis. DSA testing is not SOC in liver transplantation, but more DSA data would be very welcome to help solve the ABMR issue. However, DSA can be misleading: the presence of DSA may simply be a marker of an alloresponse.<sup>288</sup> DSA appears with increasing frequency as TxBx increased in all organ transplant populations and is often not responsible for a phenotype: troubled transplants with late TCMR will be associated with DSA even if the phenotype is T cell-driven. Late TCMR occurring during nonadherence or in immunosuppressive withdrawal studies (“chronic rejection”) may be DSA-positive,<sup>289,290</sup> and does not necessarily indicate a distinctive ABMR state. Moreover, early ABMR reported in ABO-incompatible and highly sensitized

## CHAPTER 6: Liver rejection

patients is now very uncommon and no cases were suspected in this cohort, and analyses can make no statement about phenotypes that did not exist in the study cohort. The molecular phenotype of liver ABMR remains an unresolved topic of strong interest, given the interest in immunosuppressive withdrawal and tolerance,<sup>258,259,291</sup> and failure to find it by the tests applied in this study may simply be due to under-representation in the present biopsy cohort. Liver ABMR phenotypes have been contentious since the beginnings of liver transplantation. The next steps will be to develop classifiers based on cases the collaborators consider definite ABMR, but this will require separate analyses and larger study enrolment.

While sampling error is a concern in all biopsy-based testing, prior analyses have shown that the effect of sampling is low on MMDx assessments (see Chapter 3).<sup>64</sup> Biological replication in liver biopsies was also high, confirming that biopsies of the size and yield used in this study are appropriate for reliable MMDx assessment.

The relationship between histologic rejection (>0, >1, or >2 classifications) and MMDx was weaker than that seen in other organs (i.e. kidney or heart transplants). However, this was expected given the known limitations of liver pathology and noise within the rejection diagnoses. The choice of threshold for histologic rejection (0, 1, or 2) did not significantly alter the balanced accuracy, suggesting that the severity of recorded lesions did not significantly impact overall association with the MMDx rejection diagnoses. SOC histology assessment in these centers does not typically include a summary sign-out. Since these analyses focused purely on the relationships between MMDx and unmodified SOC, only the available information was used, and further interpreted as was necessary for analyses requiring binary classes. Future analyses may include a classifier based on a more definitive style of diagnostic classification if this information can be obtained.

The limitations of this first generation MMDx study include its relatively small number of samples, and the lack of information regarding DSA and infectious complications. Follow-up times after biopsy remain relatively short, with too few failures or re-transplants to analyze survival or clinical outcomes at present. The ongoing INTERLIVER study should resolve these issues for future analyses. Supervised analysis was limited by SOC text-based reports.

In conclusion, liver transplant biopsy assessment can be successfully approached using genome-wide discovery and correlates with the current SOC histology classification of biopsies. While distinct

## **CHAPTER 6:** Liver rejection

TCMR, injury, and late fibrosis states emerged, no distinct ABMR state was identified using approaches that shared distinct ABMR syndromes in kidney and heart transplants. There were considerable discrepancies as expected given the limited reproducibility in histology diagnoses: molecular tests cannot agree with histology more than histology agrees with itself. However, the ongoing INTERLIVER study provides an opportunity to study the discrepancies to calibrate MMDx readings and recalibrate histology readings. Further studies now focus on understanding more phenotypes such as the molecular identification of steatohepatitis as well as the resolution of the question of liver ABMR.

## **CHAPTER 6: Liver rejection**

### **6.10 TABLES**

## CHAPTER 6: Liver rejection

**Table 6.1** Participating Collaborators in the INTERLIVER study

Name	Institution	Location	# of biopsies contributed			
Amar Gupta	Baylor University Medical Center	Dallas, TX, USA	1			
Goran Klintmalm						
Marwan Abouljoud	Henry Ford Hospital	Detroit, MI, USA	23			
Iman Francis						
Dilip Moonka						
Rosa Miquel	King's College London	London, UK	*			
Alberto Sanchez-Fueyo						
Grzegorz Piecha	Medical University of Silesia	Katowice, Poland	8			
Marta Gryczman	Pomeranian Medical University	Szczecin, Poland	36			
Krzysztof Jurczyk						
Joanna Mazurkiewicz						
Marek Myślak						
Marta Wawrzynowicz-Syczewska						
Samir Zeair						
Aldo Montano-Loza	University of Alberta	Edmonton, AB, Canada	6			
Martina Brozynski	University of California San Francisco	San Francisco, CA, USA	2			
Sandy Feng						
Monique Koenigsberg						
David Bowen	University of Sydney	Sydney, NSW, Australia	11			
Fiona Guan						
Ken Liu						
Avik Majumdar						
Geoff McCaughan						
Simone Strasser						
Tatiana Tsoutsman						
Michael Akyeampong	Virginia Commonwealth University	Richmond, VA, USA	38			
Jeanette Amery						
Chandra Bhatti						
Johanna Christensen						
Adrian Cotterell						
Megan Gray						
Becky Hickey						
Aamir Khan						
Marlon Levy						
Trevor Reichman						
Amit Sharma						
Vanessa Taylor						
Michal Ciszek				Warsaw Medical University	Warsaw, Poland	110
Dominika Dęborska-Materkowska						
Magdalena Durlik						
Bartosz Foroniewicz						
Michal Grat						
Krzysztof Mucha						
Agnieszka Perkowska-Ptasinska						
Olga Tronina						
Krzysztof Zieniewicz						

NOTE. Sample submitted from King's College, London were not included in the current datalock but will be included in subsequent datalocks.

**Table 6.2** INTERLIVER patient and biopsy characteristics

Patient characteristics	Patients N = 218
<b>Recipient sex (% total)</b>	
<i>Male</i>	111 (51%)
<i>Female</i>	107 (49%)
<b>Recipient age at transplant (median, range)</b>	
	50 (2-71)
<b>Primary disease (% total)<sup>A</sup></b>	
<i>Alcoholic Liver Disease</i>	30 (14%)
<i>Autoimmune hepatitis</i>	20 (9%)
<i>Hepatitis B</i>	12 (6%)
<i>Hepatitis C</i>	35 (16%)
<i>Hepatocellular carcinoma</i>	20 (9%)
<i>Non-alcoholic Steatohepatitis</i>	14 (6%)
<i>Primary Biliary Cholangitis</i>	15 (7%)
<i>Primary Sclerosing Cholangitis</i>	23 (11%)
<i>Other</i>	26 (12%)
<i>Missing</i>	48 (22%)
Biopsy characteristics	Biopsies N = 235
<b>Days (median, range) from transplant to biopsy (TxBx)</b>	
	962 (0 - 11676)
<b>Immunosuppression at biopsy (% total)</b>	
<i>Corticosteroids</i>	1 (<1%)
<i>Cyclosporine</i>	2 (1%)
<i>Tacrolimus</i>	48 (20%)
<i>Missing</i>	185 (79%)
<b>Indication for biopsy (% total)</b>	
<i>Indication: clinician concerned about graft function</i>	170 (72%)
<i>Follow-up from previous biopsy</i>	2 (1%)
<i>Protocol/surveillance</i>	29 (12%)
<i>Missing</i>	34 (14%)

<sup>A</sup>Some patients fell under multiple categories

**CHAPTER 6:** Liver rejection**Table 6.3** Laboratory test data and DSA for all biopsies in the INTERLIVER study (N=235)

<b>Laboratory Tests</b>	<b>Values in all biopsies</b> Mean (median, range)
<i>Albumin (g/dL)</i>	3.9 (4.0, 1.7-5.4) N=200
<i>Bilirubin (mg/dL)</i>	2.2 (0.9, 0.1-36.0) N=226
<i>AST (IU/L)</i>	124.2 (45.0, 11-5779) N=225
<i>ALT (IU/L)</i>	18.2 (68, 8-1781) N=226
<i>ALP (IU/L)</i>	203 (109, 29-1863) N=225
<b>DSA at biopsy</b>	<b># of results</b> (% of all known results)
<i>Positive</i>	10 (4%)
<i>Negative</i>	7 (3%)
<i>Not tested</i>	218 (93%)

NOTE. Missing DSA values included those not provided by the center or were not measured not within  $\pm 6$  months of the biopsy.

**CHAPTER 6:** Liver rejection

**Table 6.4** Mean pathogenesis-based transcript (PBT) set scores in biopsies grouped according to their highest Rejection (RAT4AA) archetype score

	R1 <sub>normal</sub> (n = 129)	R2 <sub>TCMR</sub> (n = 37)	R3 <sub>injury</sub> (n = 61)	R4 <sub>late</sub> (n = 8)
<b>Median time of biopsy post-transplant (in days)</b>	2534	214	99	<b>3117</b>
<b>PBT</b>	<b>Mean PBT score in each archetype group<sup>1</sup></b>			
	R1 <sub>normal</sub>	R2 <sub>TCMR</sub>	R3 <sub>injury</sub>	R4 <sub>late</sub>
<b>TCMR-related transcripts</b>				
<i>TCMR-RAT – TCMR-associated RATs</i>	1 ± 1.25	<b>2.20</b> ± 1.33	0.99 ± 1.31	1.69 ± 1.33
<i>QCAT – Cytotoxic T cell associated transcripts</i>	1 ± 1.29	<b>2.17</b> ± 1.39	0.89 ± 1.36	1.78 ± 1.25
<b>Rejection-related</b>				
<i>GRIT – Interferon gamma-inducible transcripts</i>	1 ± 1.19	<b>1.84</b> ± 1.23	1.12 ± 1.33	1.39 ± 1.21
<i>Rejection-RATs – rejection associated RATs</i>	1 ± 1.26	<b>2.12</b> ± 1.28	0.95 ± 1.37	1.38 ± 1.33
<b>Endothelium-related transcripts</b>				
<i>eDSAST – Endothelium-expressed DSA-selective transcripts</i>	1 ± 1.34	0.84 ± 1.26	0.92 ± 1.56	<b>3.28</b> ± 1.90
<i>ENDAT – Endothelial cell-associated transcripts</i>	1 ± 1.17	1.17 ± 1.22	1.09 ± 1.26	<b>2.13</b> ± 1.29
<b>Late injury-related transcripts (atrophy-fibrosis)</b>				
<i>IGT – Immunoglobulin transcripts</i>	1 ± 1.97	1.38 ± 2.56	0.71 ± 2.01	<b>4.05</b> ± 2.78
<i>BAT – B cell-associated transcripts</i>	1 ± 1.20	1.32 ± 1.29	0.97 ± 1.22	<b>2.00</b> ± 1.25
<i>MCAT – Mast cell-associated transcripts</i>	1 ± 1.42	0.78 ± 1.36	0.80 ± 1.50	<b>3.12</b> ± 1.96
<b>Recent injury-related transcripts</b>				
<i>FICOL – fibrillar collagen-associated transcripts</i>	1 ± 1.43	1.49 ± 1.61	1.32 ± 1.69	<b>4.61</b> ± 1.98
<i>DAMP – Damage-associated molecular pattern transcripts</i>	1 ± 1.35	<b>1.48</b> ± 1.71	1.44 ± 1.94	1.44 ± 2.00
<i>IRRAT – Injury/repair associated transcripts (human kidney)</i>	1 ± 1.34	1.48 ± 1.43	1.33 ± 1.66	<b>2.64</b> ± 1.64
<i>IRITD3 – tissue injury and repair associated transcripts</i>	1 ± 1.15	1.27 ± 1.18	1.27 ± 1.26	<b>1.88</b> ± 1.30
<i>IRITD5 – tissue injury and repair associated transcripts</i>	1 ± 1.18	1.42 ± 1.24	1.19 ± 1.26	<b>2.19</b> ± 1.30
<i>cIRIT – cardiac injury and repair induced transcripts</i>	1 ± 1.12	1.39 ± 1.17	1.24 ± 1.19	<b>1.52</b> ± 1.22

<sup>1</sup>Score represents the mean fold difference in PBT expression between biopsies in each archetype group and the R1 biopsies as a control. Biopsies were grouped according to their highest of the four archetype scores. The highest score in each row is bolded.

**CHAPTER 6: Liver rejection**

**Table 6.5** Top 10 transcripts correlated with the RAT4A R1<sub>normal</sub>, R2<sub>TCMR</sub>, R3<sub>injury</sub>, and R4<sub>late</sub> scores.

R1 <sub>normal</sub>			R2 <sub>TCMR</sub>			R3 <sub>injury</sub>			R4 <sub>late</sub>		
Gene Symbol	Correlation with R1 <sub>normal</sub>	PBT Annotations	Gene Symbol	Correlation with R2 <sub>TCMR</sub>	PBT Annotations	Gene Symbol	Correlation with R3 <sub>injury</sub>	PBT Annotations	Gene Symbol	Correlation with R4 <sub>late</sub>	PBT Annotations
<i>TRIP12</i>	-0.79	-	<i>TAP1</i>	0.90	GRIT, REJ-RAT, TCMR-RAT	<i>EGLN1</i>	0.56	-	<i>JAM3</i>	0.76	-
<i>MIR21</i>	-0.79	-	<i>PSMB9</i>	0.89	GRIT, REJ-RAT, TCMR-RAT	<i>SHROOM3</i>	0.53	-	<i>ELK3</i>	0.75	-
<i>IFNGR1</i>	-0.79	-	<i>STAT1</i>	0.87	GRIT, IRRAT, TCMR-RAT	<i>PVR</i>	0.52	CT1, IRRAT950	<i>GYPC</i>	0.74	-
<i>FCGR1A</i>	-0.78	GRIT, IRRAT, RAT, TCMR-RAT	<i>PSMB8</i>	0.87	GRIT, REJ-RAT	<i>ELMOD2</i>	0.52	-	<i>RAB34</i>	0.74	-
<i>ACTR3</i>	-0.78	cIRIT, IRRAT, LivGST_UP	<i>GBP5</i>	0.87	GRIT, REJ-RAT	<i>PTPN11</i>	0.51	IRRAT950	<i>VIM</i>	0.74	cIRIT, IRIT, IRRAT, LivGST
<i>ACOT9</i>	-0.78	LivGST_UP	<i>PSMB10</i>	0.85	GRIT, REJ-RAT, TCMR-RAT	<i>SERPINA3</i>	0.51	IRITD3, IRRAT30, IRRAT950	<i>DPYSL3</i>	0.74	IRIT
<i>TMEM165</i>	-0.78	IRRAT	<i>GBP1</i>	0.85	GRIT, REJ-RAT	<i>SDC4</i>	0.50	-	<i>SPARC</i>	0.73	IRIT
<i>CAP1</i>	-0.78	LivGST_UP	<i>CXCL11</i>	0.85	GRIT, REJ-RAT	<i>BCAP29</i>	0.50	-	<i>LUM</i>	0.73	IRIT
<i>DNM1L</i>	-0.78	-	<i>CXCL9</i>	0.84	GRIT, REJ-RAT	<i>SLC25A33</i>	0.50	--	<i>ADGRA2</i>	0.73	-
<i>TMEM50A</i>	-0.78	-	<i>CTSS</i>	0.84	GRIT, IRIT, IRRAT, REJ-RAT	<i>SRXN1</i>	0.50	-	<i>LTBP2</i>	0.73	-

NOTE. Top transcripts were selected based on highest negative value of the Spearman correlation coefficient in R1, and highest positive value in R2-R4.

**CHAPTER 6: Liver rejection**

**Table 6.6** Top 30 unique transcripts associated with RAT4A Normal Archetype Score 1, sorted by Spearman correlation coefficient (negative to positive).

Spearman correlation coefficient	P value	Gene Symbol	Name	PBT	Expression in archetype groups			
					R1	R2	R3	R4
-0.79	7.6E-52	TRIP12	thyroid hormone receptor interactor 12		425	528	506	535
-0.79	7.7E-52	MIR21	microRNA 21; vacuole membrane protein 1		864	1082	1185	1293
-0.79	6.5E-51	IFNGR1	interferon gamma receptor 1		1022	1502	1479	1963
-0.78	5.8E-50	FCGR1A	Fc fragment of IgG, high affinity Ia, receptor (CD64)	GRIT3, IRRAT950, TCMR-RAT	51	324	117	156
-0.78	2.5E-49	ACTR3	ARP3 actin-related protein 3 homolog (yeast)	cIRIT, IRRAT950, LivGST_UP	1109	1690	1365	1708
-0.78	3.5E-49	ACOT9	acyl-CoA thioesterase 9	LivGST_UP	62	104	96	115
-0.78	4.3E-49	TMEM165	transmembrane protein 165	IRRAT950	99	168	148	213
-0.78	4.9E-49	CAP1	CAP, adenylate cyclase-associated protein 1 (yeast)	LivGST_UP	639	944	802	1119
-0.78	6.5E-49	DNM1L	dynammin 1-like		52	77	76	90
-0.78	1.2E-48	TMEM50A	transmembrane protein 50A		577	830	734	913
-0.77	2.3E-47	AGFG1	ArfGAP with FG repeats 1		150	218	228	221
-0.77	5E-47	YWHAH	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta	IRITD3	526	826	778	1069
-0.77	5.8E-47	RHOA	ras homolog family member A	IRRAT950	3182	4167	3740	4357
-0.77	6.6E-47	HN1	hematological and neurological expressed 1	IRRAT950	162	302	248	307
-0.76	6.4E-46	DDX39A	DEAD (Asp-Glu-Ala-Asp) box polypeptide 39A	cIRIT, IRITD3	185	287	243	308
-0.76	8.7E-46	H2AFY	H2A histone family, member Y		1124	1710	1444	1856
-0.76	1.2E-45	TUBA1C	tubulin, alpha 1c	IRRAT950	1147	1881	1954	2072
0.76	1.3E-45	SMO	smoothened, frizzled class receptor		374	243	213	188
-0.76	1.3E-45	VAMP5	vesicle associated membrane protein 5	GRIT3	377	781	606	635
-0.76	1.4E-45	ANXA2	annexin A2	cIRIT, IRITD3, IRRAT950	400	881	783	1513
-0.76	2.1E-45	TMEM87B	transmembrane protein 87B	IRRAT950	62	93	90	97
-0.76	2.1E-45	ECT2	epithelial cell transforming 2	cIRIT, IRITD5	19	47	37	46
-0.76	6.5E-45	GARS	glycyl-tRNA synthetase		425	547	588	611
-0.75	1.5E-44	ANXA2P2	annexin A2 pseudogene 2	cIRIT, IRITD3, IRRAT950	1157	2361	2140	3772
0.75	1.6E-44	TTC36	tetratricopeptide repeat domain 36	KT1	1489	942	987	560
-0.75	2.5E-44	KPNA2	karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	IRRAT950	351	657	573	740
-0.75	3.5E-44	SLC35F6	solute carrier family 35, member F6		266	356	360	388
-0.75	5.4E-44	RAN	RAN, member RAS oncogene family	cIRIT, IRITD3	1143	1474	1434	1656
-0.75	6.1E-44	SMARCA5	SWI/SNF related, matrix associated, actin dependent regulator of chromatin		190	266	251	344
-0.75	8.8E-44	TPM3	tropomyosin 3	cIRIT	1371	1913	1619	1990

**CHAPTER 6: Liver rejection**

**Table 6.7** Top 30 unique transcripts associated with RAT4A TCMR Archetype Score 2, sorted by Spearman correlation coefficient (positive to negative).

Spearman correlation coefficient	P-value	Gene Symbol	Name	PBT	Expression in archetype groups			
					R1	R2	R3	R4
0.90	1.7E-84	TAP1	transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)	GRIT3, RAT, Rej-RAT, TCMR-RAT	166	453	203	230
0.89	1.3E-81	PSMB9	proteasome subunit beta 9	ABMR-RAT, GRIT3, RAT, Rej-RAT, TCMR-RAT	762	2023	776	881
0.87	1.1E-73	STAT1	signal transducer and activator of transcription 1	GRIT2, GRIT3, IRRAT950, RAT, TCMR-RAT	276	956	386	348
0.87	2.7E-73	PSMB8	proteasome subunit beta 8	GRIT3, RAT, Rej-RAT	1127	2166	1291	1119
0.87	1E-72	GBP5	guanylate binding protein 5	ABMR-RAT, GRIT3, RAT, Rej-RAT	76	332	73	119
0.85	1.7E-67	PSMB10	proteasome subunit beta 10	GRIT3, RAT, Rej-RAT, TCMR-RAT	577	1013	568	620
0.85	8.2E-67	GBP1	guanylate binding protein 1, interferon-inducible	ABMR-RAT, GRIT3, RAT, Rej-RAT	225	770	323	246
0.85	1.7E-66	CXCL11	chemokine (C-X-C motif) ligand 11	ABMR-RAT, GRIT3, RAT, Rej-RAT	20	139	35	38
0.84	1.8E-64	CXCL9	chemokine (C-X-C motif) ligand 9	ABMR-RAT, GRIT1, GRIT3, RAT, Rej-RAT	497	3151	615	1031
0.84	4.3E-64	CTSS	cathepsin S	GRIT3, IRITD5, IRRAT30, RAT, Rej-RAT	1029	2275	1169	1417
0.84	5.3E-64	CXCL10	chemokine (C-X-C motif) ligand 10	ABMR-RAT, GRIT3, RAT, Rej-RAT	394	2821	681	906
0.84	1.4E-63	TAP2	transporter 2, ATP-binding cassette, sub-family B	GRIT2, GRIT3, RAT, Rej-RAT	150	336	173	209
0.84	9.4E-63	SAMD9L	sterile alpha motif domain containing 9-like	GRIT3	63	171	66	99
0.83	2.7E-62	HLA-B	major histocompatibility complex, class I, B	ABMR-RAT, GRIT3, RAT, Rej-RAT	6968	12105	7239	8827
0.83	9.1E-62	CD72	CD72 molecule	BAT, RAT, TCMR-RAT	40	143	42	67
0.83	2.8E-60	CD53	CD53 molecule		329	799	366	761
0.83	1E-59	FCER1G	Fc fragment of IgE, high affinity I, receptor for	IRITD5, LivGST_UP, QCMAT	894	1913	983	1209
0.82	1.3E-59	WARS	tryptophanyl-tRNA synthetase	ABMR-RAT, GRIT3, RAT, Rej-RAT	383	973	445	522
0.82	3.3E-59	HLA-C	major histocompatibility complex, class I, C	ABMR-RAT, GRIT3, RAT, Rej-RAT	1626	3187	1738	1942
0.82	1.7E-58	LILRB1	leukocyte immunoglobulin-like receptor	QCMAT, RAT, Rej-RAT	53	123	58	60
0.82	2.1E-58	FCGR3A	Fc fragment of IgG, low affinity IIIa, receptor (CD16a)	GRIT2, IRRAT950, RAT, Rej-RAT	1356	2984	1668	1505
0.82	2.3E-58	LCP1	lymphocyte cytosolic protein 1 (L-plastin)	cIRIT, IRITD5, IRRAT950	399	811	419	637
0.82	3.8E-58	PTPRC	protein tyrosine phosphatase, receptor type, C	IRRAT30, IRRAT950	344	822	324	738
0.82	1E-57	HLA-A	major histocompatibility complex, class I, A	ABMR-RAT, GRIT2, GRIT3, RAT, Rej-RAT	1841	4112	2065	2520
0.82	1.2E-57	CD86	CD86 molecule	GRIT2, IRRAT950, QCMAT, RAT, TCMR-RAT	106	231	121	159
0.82	2.1E-57	CYBB	cytochrome b-245, beta polypeptide	GRIT2, IRRAT950, RAT, TCMR-RAT	293	595	317	453
0.81	5.3E-57	MARCH1	membrane associated ring finger 1	RAT, TCMR-RAT	53	125	56	92
0.81	5.8E-57	FYB	FYN binding protein	RAT, TCMR-RAT	177	396	162	360
0.81	9E-57	C1QB	complement component 1, q subcomponent, B chain	GRIT3, RAT, TCMR-RAT	2679	5539	2926	3425
0.81	1.2E-56	AIF1	allograft inflammatory factor 1	GRIT2, RAT, Rej-RAT	402	772	399	591

**CHAPTER 6: Liver rejection**

**Table 6.8** Top 30 unique transcripts associated with RAT4A Injury Archetype Score 3, sorted by Spearman correlation coefficient (positive to negative).

Spearman correlation coefficient	P value	Gene Symbol	Name	PBT	Expression in archetype groups			
					R1	R2	R3	R4
0.56	3.4E-21	EGLN1	egl-9 family hypoxia-inducible factor 1		96	105	137	0.56
0.53	1.3E-18	SHROOM3	shroom family member 3		99	98	133	0.53
0.52	1.7E-17	PVR	poliovirus receptor	CT1, IRRAT950	86	106	135	0.52
0.52	1.9E-17	ELMOD2	ELMO/CED-12 domain containing 2		92	112	125	0.52
0.51	2.7E-17	PTPN11	protein tyrosine phosphatase, non-receptor type 11	IRRAT950	350	380	456	0.51
0.51	3.7E-17	SERPINA3	serpin peptidase inhibitor, clade A (alpha-1 antitrypsin, antitrypsin), member 3	IRITD3, IRRAT30, IRRAT950	13952	14583	16113	0.51
0.50	2E-16	SDC4	syndecan 4		1296	1398	1887	0.50
0.50	2.9E-16	BCAP29	B-cell receptor-associated protein 29		113	136	149	0.50
0.50	4E-16	SLC25A33	solute carrier family 25 (pyrimidine nucleotide carrier), member 33		297	440	644	0.50
0.50	4.9E-16	SRXN1	sulfiredoxin 1		185	240	308	0.50
0.49	7.7E-16	ARHGEF5	Rho guanine nucleotide exchange factor 5; rho guanine nucleotide exchange factor 5-like		83	91	122	0.49
0.49	1.4E-15	TWF1	twinfilin actin binding protein 1		280	326	353	0.49
0.49	2E-15	MTCH2	mitochondrial carrier 2	CT1	1076	1128	1265	0.49
0.49	2.2E-15	PLPP6	phospholipid phosphatase 6		241	257	297	0.49
0.48	4.4E-15	CCDC47	coiled-coil domain containing 47		442	452	511	0.48
0.48	6.5E-15	DUSP3	dual specificity phosphatase 3	IRRAT950	173	200	216	0.48
0.47	1.3E-14	VMP1	vacuole membrane protein 1		299	359	388	0.47
0.47	1.4E-14	SLC25A16	solute carrier family 25 (mitochondrial carrier), member 16	CT1	208	227	256	0.47
0.47	1.4E-14	ABCC3	ATP binding cassette subfamily C member 3	IRRAT950	133	131	169	0.47
0.47	1.7E-14	TMCO1	transmembrane and coiled-coil domains 1	CT1	1803	1893	2044	0.47
0.47	2.6E-14	PTGFRN	prostaglandin F2 receptor inhibitor	IRITD5	86	108	126	0.47
0.47	3E-14	F2RL1	coagulation factor II (thrombin) receptor-like 1	IRITD3, IRRAT950	38	40	54	0.47
0.47	3.3E-14	SNX7	sorting nexin 7		123	149	165	0.47
0.47	4.8E-14	SURF4	surfeit 4	cIRIT	710	721	821	0.47
0.47	4.9E-14	P4HA2	prolyl 4-hydroxylase, alpha polypeptide II		143	167	205	0.47
0.47	4.9E-14	TMPRSS2	transmembrane protease, serine 2	KT1	161	164	212	0.47
0.47	5E-14	SGMS2	sphingomyelin synthase 2		35	38	43	0.47
0.46	6.5E-14	CCNC	cyclin C		802	851	975	0.46
0.46	6.5E-14	LETM1	leucine zipper-EF-hand containing transmembrane protein 1		102	104	124	0.46
0.46	6.8E-14	GPX3	glutathione peroxidase 3	KT1, LivGST_UP	1002	1509	1837	0.46

**CHAPTER 6:** Liver rejection

**Table 6.9** Top 30 unique transcripts associated with RAT4A Late Injury Archetype Score 4, sorted by Spearman correlation coefficient (positive to negative).

Spearman correlation coefficient	P-value	Gene Symbol	Name	PBT	Expression in archetype groups			
					R1	R2	R3	R4
0.76	4E-46	JAM3	junctional adhesion molecule 3		89	88	91	217
0.75	2.6E-43	ELK3	ELK3, ETS-domain protein (SRF accessory protein 2)		333	336	331	698
0.74	1.9E-42	GYPC	glycophorin C (Gerbich blood group)		168	220	167	374
0.74	2.7E-42	RAB34	RAB34, member RAS oncogene family		115	143	117	270
0.74	4.4E-42	VIM	vimentin	clRIT, IRITD3, IRRAT950, LivGST_UP	2045	3299	2456	6142
0.74	7E-42	DPYSL3	dihydropyrimidinase-like 3	IRITD3	52	58	55	259
0.73	3.5E-41	SPARC	secreted protein, acidic, cysteine-rich (osteonectin)	IRITD3	1455	1654	1547	4186
0.73	4E-41	LUM	lumican	IRITD5	446	726	615	3323
0.73	1.1E-40	ADGRA2	adhesion G protein-coupled receptor A2		242	256	242	620
0.73	1.3E-40	LTBP2	latent transforming growth factor beta binding protein 2		48	53	54	441
0.73	1.8E-40	AEBP1	AE binding protein 1	IRITD3	200	236	223	758
0.73	5.4E-40	CNRIP1	cannabinoid receptor interacting protein 1		106	109	107	241
0.73	5.8E-40	VASH1	vasohibin 1		96	124	95	178
0.73	1.2E-39	ZNF532	zinc finger protein 532		94	111	93	259
0.72	2.8E-39	PAPLN	papilin, proteoglycan-like sulfated glycoprotein		57	81	63	292
0.72	1E-38	GSTP1	glutathione S-transferase pi 1		296	401	297	673
0.72	1.8E-38	THY1	Thy-1 cell surface antigen		27	37	34	198
0.72	2.3E-38	COL4A2	collagen, type IV, alpha 2	ENDAT, IRITD3, IRRAT950	370	522	477	1723
0.71	5.2E-38	ANTXR1	anthrax toxin receptor 1		128	126	123	587
0.71	3.3E-37	BCL2	B-cell CLL/lymphoma 2	HT1	34	46	37	116
0.71	4.1E-37	EFEMP1	EGF containing fibulin-like extracellular matrix protein 1	ENDAT, IRITD3, IRRAT950	137	295	252	2571
0.71	4.6E-37	FBN1	fibrillin 1	IRITD5	151	215	194	841
0.71	4.7E-37	EFEMP2	EGF containing fibulin-like extracellular matrix protein 2	IRITD5	117	118	119	255
0.71	1.2E-36	GAL3ST4	galactose-3-O-sulfotransferase 4		36	55	47	204
0.70	2.2E-36	CDH11	cadherin 11, type 2, OB-cadherin (osteoblast)	IRITD5	36	51	46	371
0.70	2.3E-36	MMP2	matrix metalloproteinase 2	IRITD5	109	111	121	605
0.70	3.3E-36	CCND2	cyclin D2		147	281	139	615
0.70	5.5E-36	COL6A3	collagen, type VI, alpha 3	IRITD5	309	414	384	1669
0.70	7.3E-36	LBH	limb bud and heart development	LivGST_UP	149	217	140	396
0.70	9.1E-36	KCTD12	potassium channel tetramerization domain containing 12	LivGST_UP	444	788	520	1430

**CHAPTER 6:** Liver rejection

**Table 6.10** Correlations between pathogenesis-based transcript (PBT) set scores and Rejection (RAT4AA) PCA scores.

PBT	Spearman correlation		
	PC1	PC2	PC3
<b>Correlations with TxBx</b>	-0.08	-0.58	<b>0.13</b>
<b>TCMR-related transcripts</b>			
<i>TCMR-RAT – TCMR-associated RATs</i>	<b>0.98</b>	0.07	-0.01
<i>QCAT – Cytotoxic T cell associated transcripts</i>	<b>0.94</b>	-0.09	0.03
<b>Rejection-related</b>			
<i>GRIT – Interferon gamma-inducible transcripts</i>	<b>0.91</b>	0.30	-0.02
<i>Rejection-RATs – rejection associated RATs</i>	<b>0.99</b>	0.05	-0.05
<b>Endothelium-related transcripts</b>			
<i>eDSAST – Endothelium-expressed DSA-selective transcripts</i>	0.03	-0.29	<b>0.69</b>
<i>ENDAT – Endothelial cell-associated transcripts</i>	0.47	0.06	<b>0.70</b>
<b>Late injury-related transcripts</b>			
<i>IGT – Immunoglobulin transcripts</i>	<b>0.44</b>	-0.27	0.36
<i>BAT – B cell-associated transcripts</i>	<b>0.74</b>	-0.09	0.44
<i>MCAT – Mast cell-associated transcripts</i>	0.06	-0.47	<b>0.50</b>
<b>Recent injury-related transcripts</b>			
<i>FICOL – fibrillar collagen-associated transcripts</i>	0.42	0.20	<b>0.59</b>
<i>DAMP – Damage-associated molecular pattern transcripts</i>	<b>0.33</b>	0.19	0.26
<i>IRRAT – Injury/repair associated transcripts (human kidney)</i>	0.49	0.19	<b>0.53</b>
<i>IRITD3 – tissue injury and repair associated transcripts</i>	0.48	0.48	<b>0.50</b>
<i>IRITD5 – tissue injury and repair associated transcripts</i>	<b>0.65</b>	0.33	0.48
<i>cIRIT – cardiac injury and repair induced transcripts</i>	<b>0.63</b>	0.53	0.28

<sup>1</sup>PBT scores represent the mean fold difference in PBT expression between biopsies in each archetype group and the R1 biopsies as a control. Biopsies were grouped according to their highest of the four archetype scores. The highest positive score in each row is bolded.

**CHAPTER 6: Liver rejection**

**Table 6.11** Correlations between RAT4A archetype scores, PC scores and high scores for histologic features (>0) in liver biopsies (N=235)

Histology features	PC1	PC2	PC3	R1 <sub>normal</sub> score	R2 <sub>TCMR</sub> score	R3 <sub>injury</sub> score	R4 <sub>late</sub> score	NA (# of missing values)
Acute rejection: portal inflammation	0.27, p=4E-05	0.01, p=9E-01	-0.07, p=3E-01	-0.09, p=2E-01	0.29, p=9E-06	-0.14, p=3E-02	0.03, p=6E-01	13
Acute rejection: bile duct inflammation	0.21, p=2E-03	-0.05, p=4E-01	-0.06, p=4E-01	-0.04, p=6E-01	0.20, p=4E-03	-0.14, p=3E-02	0.05, p=5E-01	15
Acute rejection: venous inflammation	0.19, p=4E-03	-0.11, p=1E-01	0.00, p=1E+00	-0.00, p=1E+00	0.16, p=2E-02	-0.19, p=5E-03	0.11, p=9E-02	14
Chronic rejection: bile duct degeneration	0.09, p=2E-01	0.00, p=1E+00	0.07, p=3E-01	-0.08, p=2E-01	0.03, p=6E-01	0.01, p=9E-01	0.12, p=8E-02	20
Chronic rejection: focal obliteration	0.15, p=3E-02	0.01, p=9E-01	0.08, p=3E-01	-0.14, p=4E-02	0.09, p=2E-01	-0.03, p=7E-01	0.11, p=9E-02	18
Chronic rejection: cholestasis	0.05, p=5E-01	0.18, p=7E-03	0.17, p=1E-02	-0.26, p=8E-05	0.01, p=9E-01	0.16, p=2E-02	0.17, p=1E-02	18
Chronic rejection: mural fibrosis	0.09, p=2E-01	0.07, p=3E-01	0.07, p=3E-01	-0.16, p=2E-02	-0.02, p=8E-01	0.10, p=1E-01	0.10, p=1E-01	18
Other disease: autoimmune hepatitis	0.14, p=4E-02	0.07, p=3E-01	0.03, p=6E-01	-0.15, p=3E-02	0.12, p=8E-02	0.05, p=5E-01	0.11, p=1E-01	20
Other disease: steatohepatitis	0.01, p=9E-01	0.12, p=1E-01	-0.03, p=7E-01	-0.04, p=6E-01	0.05, p=5E-01	0.11, p=1E-01	-0.02, p=8E-01	54
<u>Other disease: fibrosis</u>	-0.05, p=5E-01	-0.33, p=5E-06	0.31, p=3E-05	0.27, p=3E-04	-0.17, p=3E-02	-0.18, p=1E-02	0.28, p=1E-04	56
Other disease: recurrent HCV	0.14, p=5E-02	0.25, p=2E-04	-0.08, p=2E-01	-0.22, p=1E-03	0.20, p=3E-03	0.18, p=1E-02	-0.05, p=5E-01	21
Other disease: suspected CMV hepatitis	0.22, p=9E-04	0.02, p=8E-01	-0.23, p=8E-04	-0.14, p=4E-02	0.24, p=5E-04	-0.13, p=6E-02	-0.12, p=8E-02	18

**NOTE.** Clinical data was binarized for this analysis. Sum score of portal, bile duct, and venous inflammation grades >0 was considered positive, 0 if negative. Spearman correlation coefficients are given alongside p values approximated from the value of the coefficient. Shaded cells are those considered significant based on the absolute value of the Spearman correlation coefficient >0.2.

**CHAPTER 6: Liver rejection**

**Table 6.12** Laboratory test data for biopsies in INTERLIVER by RAT4A archetype group.

Values in biopsies belonging to designated archetype group		Tests					TxBx in days mean (median, range)
		Albumin g/dL mean (median, range) (N=200)	Bilirubin mg/dL mean (median, range) (N=226)	AST IU/L mean (median, range) (N=226)	ALT IU/L mean (median, range) (N=226)	ALP IU/L mean (median, range) (N=225)	
RAT Archetype groups	<b>R1<sub>normal</sub></b> (N=129)	4.3 (4.3, 3.2-5.4) N=102	1.2 (0.7, 0.1-36.0) N=123	39.0 (27.0, 11-165) N=123	51.8 (32.0, 8-289) N=123	126.4 (95.0, 29-632) N=123	2534 (2152, 58-9169) N=129
	<b>R2<sub>TCMR</sub></b> (N=37)	3.6 (3.6, 2.2-4.7) N=33	2.8 (1.2, 0.2-18.9) N=36	176.2 (95.0, 21-848) N=36	214.9 (146.5, 27-815) N=36	319.9 (189.0, 67-1467) N=35	777 (214, 7-4918) N=37
	<b>R3<sub>injury</sub></b> (N=61)	3.4 (3.5, 1.9-4.7) N=58	3.4 (1.9, 0.2-20.0) N=59	274.5 (70.0, 17-5779) N=58	205.2 (113.0, 9-1781) N=59	273.7 (190.0, 54-1863) N=59	814 (99, 0-5622) N=61
	<b>R4<sub>late</sub></b> (N=8)	3.1 (2.9, 1.7-4.4) N=7	5.2 (1.5, 0.2-26.6) N=8	110.5 (96.5, 19-292) N=8	61.25 (59.0, 20-127) N=8	353.8 (261.0, 166-890) N=8	4807 (3117, 350-11676) N=8

\*Highlighted cells are the highest value in the column.

**NOTE.** ALP measurements showed significant differences between R2 biopsies and normal (R1 biopsies). No statistically significant differences were seen between R2 and other groups for bilirubin, or AST.

**CHAPTER 6: Liver rejection**

**Table 6.13** Top 30 unique genes associated with histologic acute rejection >0 by adjusted p value.

p value	Adjusted p value	Gene Symbol	Gene Name	PBT Annotation	Expression in rejection biopsies	Expression in no rejection biopsies
2.9E-08	0.0004	PTPRC	protein tyrosine phosphatase, receptor type, C	IRRAT30,IRRAT950	178	113
5.0E-08	0.0004	SAMHD1	SAM domain and HD domain 1	GRIT2,GRIT3,IRRAT950	203	150
1.6E-07	0.0006	KLRD1	killer cell lectin-like receptor subfamily D, member 1	ABMR-RAT,RAT,Rej-RAT,TCMR-RAT	96	68
3.9E-07	0.0007	PLEK	pleckstrin	IRRAT950,RAT,Rej-RAT	164	117
4.0E-07	0.0007	CXorf38	chromosome X open reading frame 38		44	36
4.1E-07	0.0007	LYN	v-yes-1 Yamaguchi sarcoma viral related oncogene homolog	clRIT,IRITD5,RAT,Rej-RAT	134	105
4.2E-07	0.0007	CORO1A	coronin, actin binding protein, 1A		210	147
4.7E-07	0.0007	CD48	CD48 molecule		185	120
6.2E-07	0.0007	CCL5	chemokine (C-C motif) ligand 5	GRIT2,RAT,Rej-RAT,TCMR-RAT	244	153
7.8E-07	0.0008	TLR2	toll-like receptor 2	IRRAT950,QCMAT	150	112
8.4E-07	0.0008	FAM26F	family with sequence similarity 26, member F	ABMR-RAT,GRIT3,RAT,Rej-RAT,TCMR-RAT	585	377
9.0E-07	0.0008	SLA	Src-like-adaptor	RAT,TCMR-RAT	161	114
9.3E-07	0.0008	HLA-A	major histocompatibility complex, class I, A	ABMR-RAT,GRIT1,GRIT2,GRIT3,RAT,Rej-RAT	4788	3695
9.5E-07	0.0008	FYB	FYN binding protein	RAT,TCMR-RAT	248	170
9.6E-07	0.0008	HCST	hematopoietic cell signal transducer	RAT,Rej-RAT	181	134
1.0E-06	0.0008	PSMB10	proteasome (prosome, macropain) subunit, betatype, 10	GRIT1,GRIT3,RAT,Rej-RAT,TCMR-RAT	666	534
1.2E-06	0.0008	STK17B	serine		134	93
1.4E-06	0.0009	LCK	lymphocyte-specific protein tyrosine kinase	QCAT,RAT,TCMR-RAT	65	44
1.5E-06	0.0009	SLC7A7	solute carrier family 7	KT2	224	177
1.6E-06	0.0009	PRKCB	protein kinase C, beta		132	101
1.8E-06	0.0009	CST7	cystatin F (leukocystatin)	ABMR-RAT,QCAT,RAT,Rej-RAT	80	61
1.8E-06	0.0009	MS4A7	membrane-spanning 4-domains, subfamily A, member 7	ABMR-RAT,IRRAT950,QCMAT,RAT	395	298
1.8E-06	0.0009	CLEC7A	C-type lectin domain family 7, member A	AMAT1,IRRAT950,LivGST_UP,RAT,Rej-RAT,TCMR-RAT	64	48
2.0E-06	0.0009	HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1	GRIT1,GRIT3,IRRAT950,RAT,Rej-RAT	1222	778
2.0E-06	0.0009	CD53	CD53 molecule		430	308
2.1E-06	0.0009	EMR2	egf-like module containing, mucin-like, hormone receptor-like 2	IRRAT950	39	26
2.2E-06	0.0009	RNF166	ring finger protein 166		70	59
2.6E-06	0.0009	LST1	leukocyte specific transcript 1	ABMR-RAT,IRRAT950,RAT,Rej-RAT	784	641
2.7E-06	0.0009	TLR8	toll-like receptor 8	QCMAT,RAT,Rej-RAT,TCMR-RAT	123	87
2.8E-06	0.0010	COTL1	coactosin-like 1 (Dictyostelium)	LivGST_UP	438	333

**NOTE.** Samples were considered positive for rejection if the sum of their histology grades were >0 for portal inflammation, bile duct inflammation, venous inflammation. Samples with sum histology grade = 0 for these categories were considered 'non-rejection'. 15 samples removed because no information was available, 220 included in analysis (162 rejection, 58 no rejection)

**Table 6.14** Characterizing the relationship between histologic and molecular diagnoses in the liver biopsy population (N=235).

<b>Crosstab of overall histologic rejection versus molecular rejection sign-outs.</b>				
		<b>MMDx-Liver sign-outs (% of column)</b>		<b>Row totals</b>
		<b>No rejection</b>	<b>TCMR</b>	
Overall Histologic acute rejection >0 <sup>a</sup>	<b>No rejection</b>	51 (33%)	7 (11%)	58
	<b>Rejection</b>	104 (67%)	58 (89%)	162
	<b>Column totals</b>	155	65	220
Overall Histologic acute rejection >1 <sup>b</sup>	<b>No rejection</b>	102 (66%)	23 (35%)	125
	<b>Rejection</b>	53 (34%)	42 (65%)	95
	<b>Column totals</b>	155	65	220
Overall Histologic acute rejection >2 <sup>c</sup>	<b>No rejection</b>	124 (80%)	39 (60%)	163
	<b>Rejection</b>	31 (20%)	26 (40%)	57
	<b>Column totals</b>	155	65	220

**Confusion matrix statistics for MMDx Diagnoses predicting the histologic diagnosis in liver transplant acute rejection**

<b>Reference Standard</b>	<b>Diagnostic Test</b>	<b>Sensitivity</b>	<b>Specificity</b>	<b>Positive Predictive Value</b>	<b>Negative Predictive Value</b>	<b>Accuracy</b>	<b>Balanced Accuracy</b>
Histologic rejection lesion score >0 <sup>a</sup>	MMDx Diagnosis <sup>d</sup>	0.36	0.89	0.89	0.33	0.50	0.62
Histologic rejection lesion score >1 <sup>b</sup>	MMDx Diagnosis <sup>d</sup>	0.44	0.82	0.65	0.66	0.65	0.63
Histologic rejection lesion score >2 <sup>c</sup>	MMDx Diagnosis <sup>d</sup>	0.46	0.76	0.40	0.80	0.68	0.60

<sup>a</sup> Based on our algorithm interpreting the acute rejection scores/features, where the presence of any score >0 in portal, bile duct, or venous inflammation classified the biopsy as acute rejection. Samples with missing information were excluded from this analysis (N=15), except where those samples were missing a single score and already clearly met the threshold for histologic rejection sum >0 (N=2).

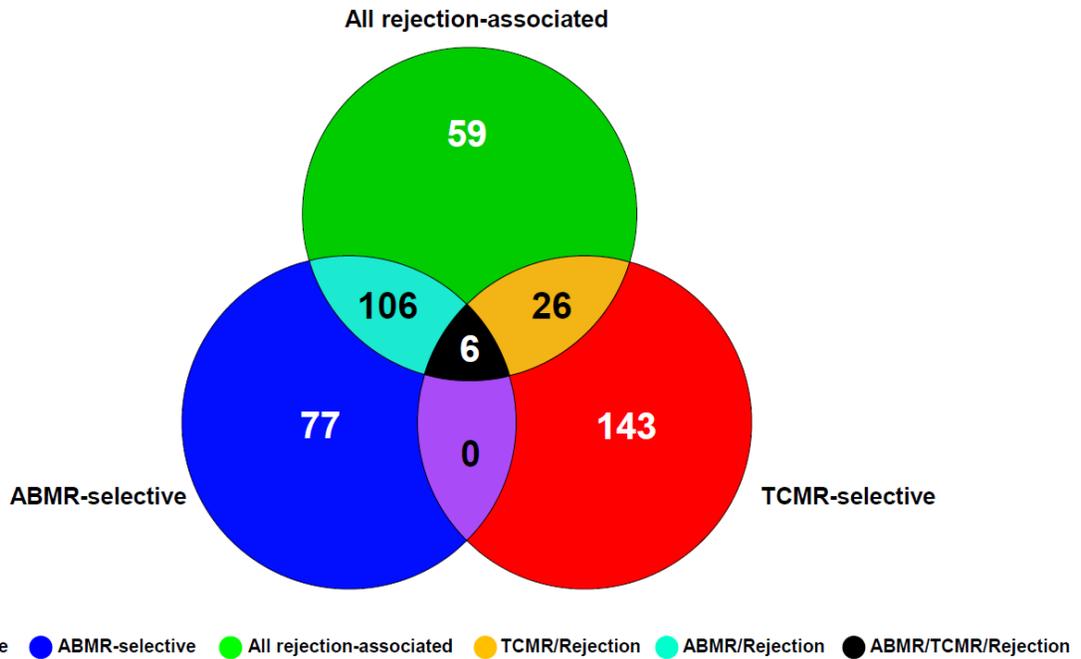
<sup>b</sup> Based on our algorithm interpreting the acute rejection scores/features, where the presence of any score >1 in portal, bile duct, or venous inflammation classified the biopsy as acute rejection. Samples with missing information were excluded from this analysis (N=15), except where those samples were missing a single score and already clearly met the threshold for histologic rejection sum >1 (N=2).

<sup>c</sup> Based on our algorithm interpreting the acute rejection scores/features, where the presence of any score >2 in portal, bile duct, or venous inflammation classified the biopsy as acute rejection. Samples with missing information were excluded from this analysis (N=15), except where those samples were missing a single score and already clearly met the threshold for histologic rejection sum >1 (N=2).

<sup>d</sup> Based on the diagnosis of acute rejection (TCMR) or no rejection by an expert signing out the official MMDx report. Diagnoses were based on position of the biopsy in the report figure, archetypal data, and PBT information provided on page 2.

## **CHAPTER 6: Liver rejection**

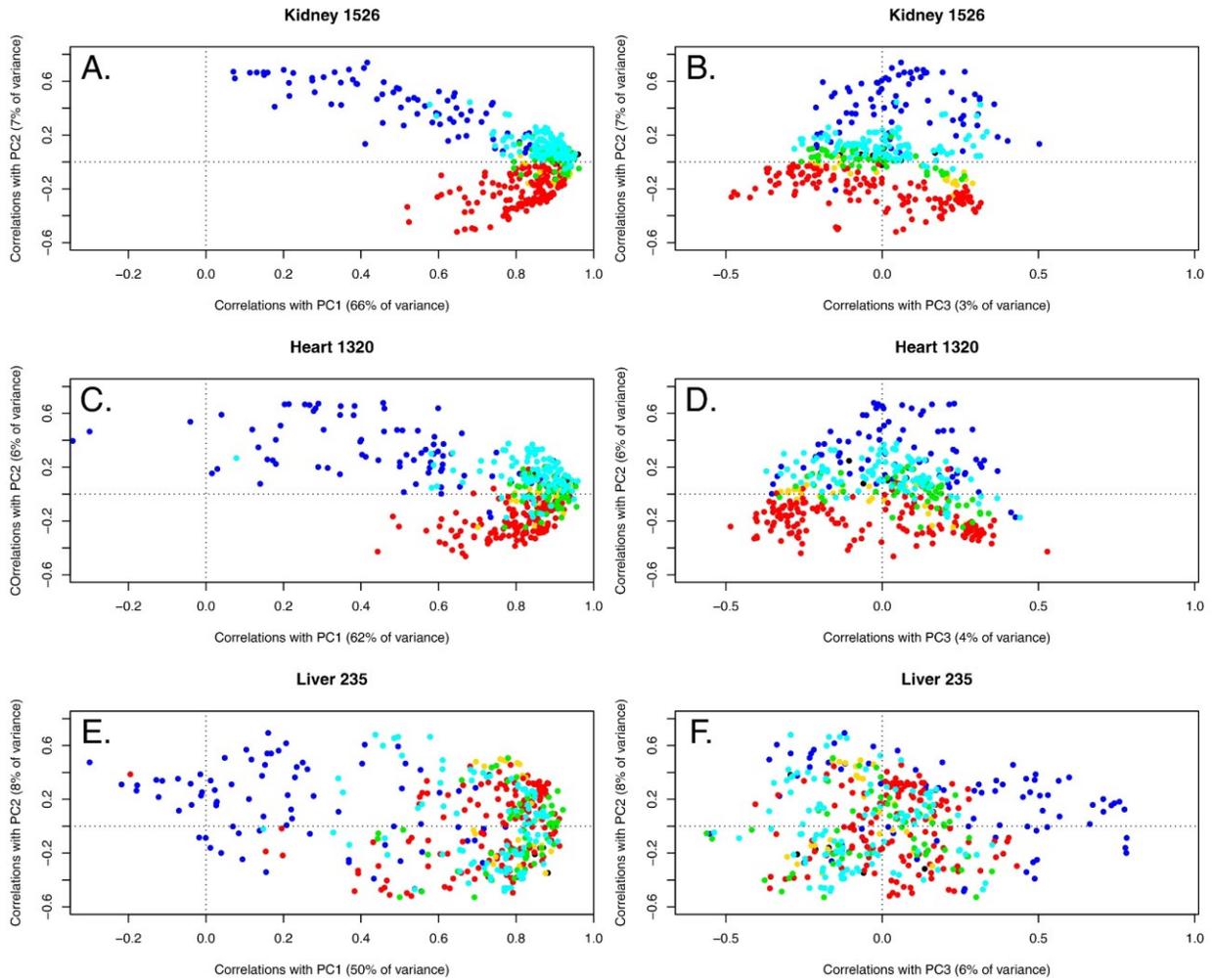
### **6.11 FIGURES**



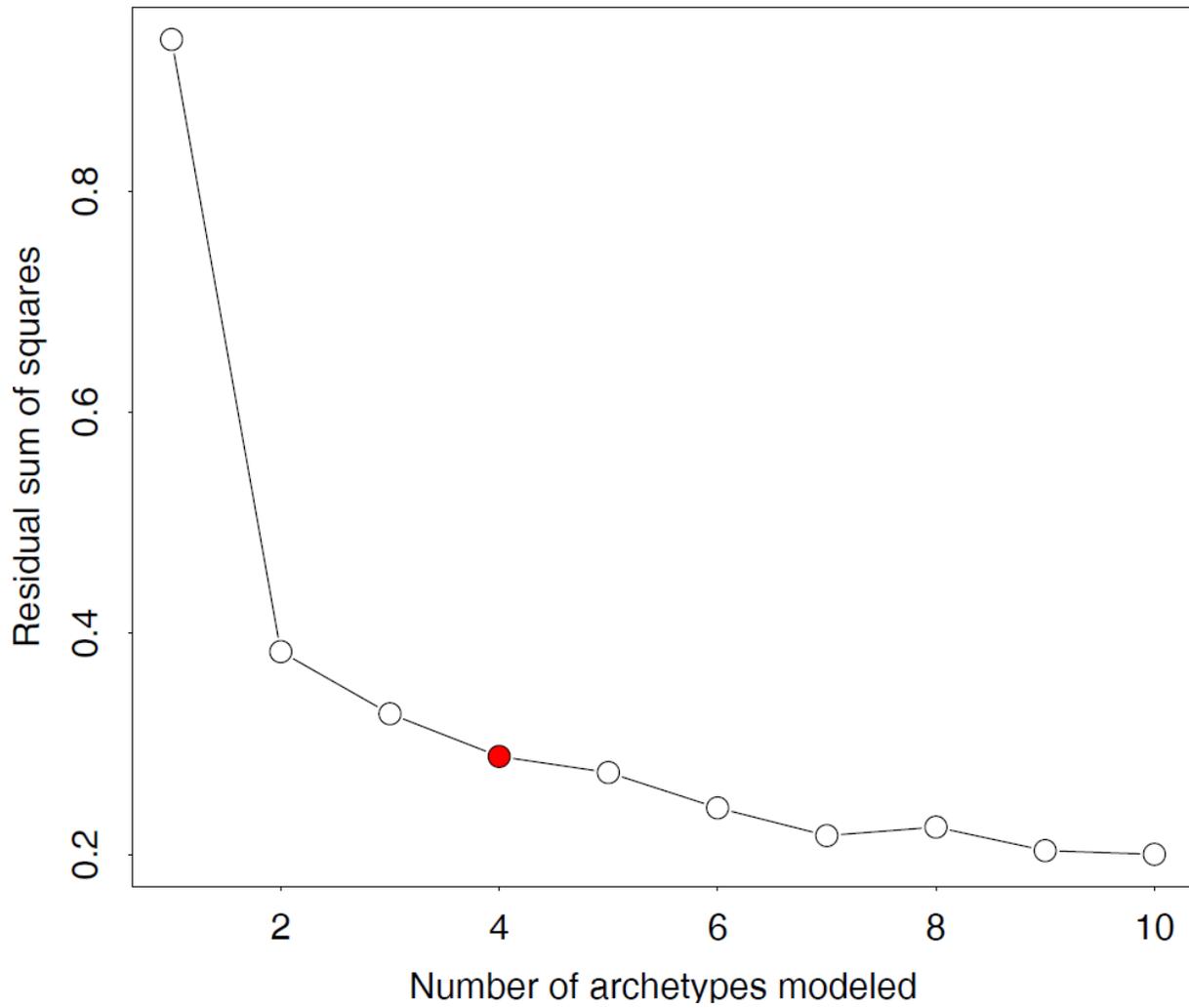
**Figure 6.1 Venn Diagram of the Rejection-associated transcripts (RATs).** RATs were divided into six groups: all-rejection-RATs (green, 'Rejection-RATs'), TCMR-RATs (red), and ABMR-RATs (blue), as well as the overlap groups TCMR/rejection (orange), ABMR/rejection (cyan), and ABMR/TCMR/rejection (black) that were identified by more than one algorithm. The Venn diagram shows the RATs as they are assigned to ABMR, TCMR, and all-rejection ('Rejection'), and the corresponding overlaps (transcripts common to multiple categories).

## CHAPTER 6: Liver rejection

● TCMR-selective ● ABMR-selective ● All rejection-associated ● TCMR/Rejection ● ABMR/Rejection ● ABMR/TCMR/Rejection

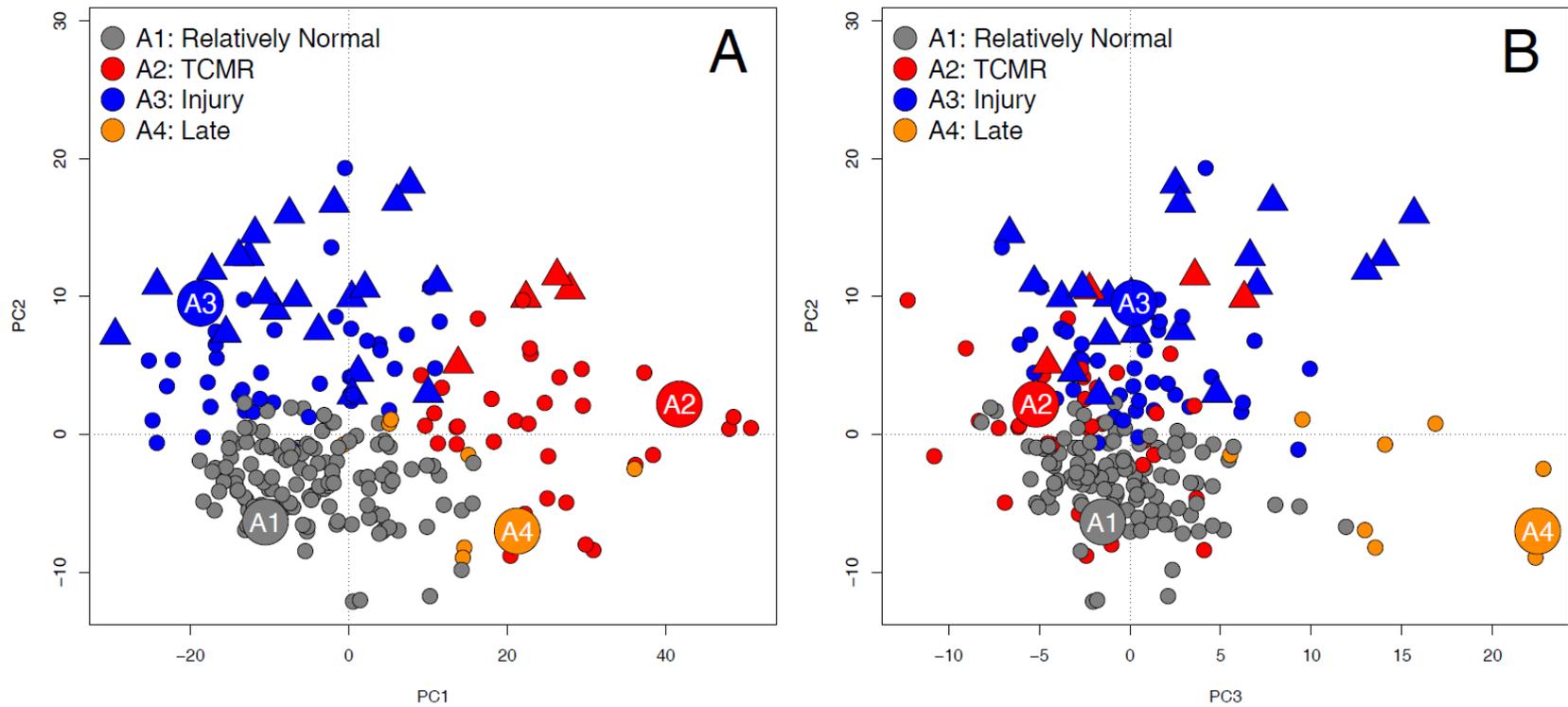


**Figure 6.2** PCA of the rejection-associated transcripts (RATs) in kidneys, hearts and livers. The RATs are shown in kidneys (Panel A showing PC1 versus PC2 and Panel B showing PC3 versus PC2), and in hearts (Panel C showing PC1 versus PC2 and Panel D showing PC3 versus PC2) for comparison. The distribution of RATs in liver biopsies is shown in Panel E (PC1 versus PC2) and Panel F (PC3 versus PC2).



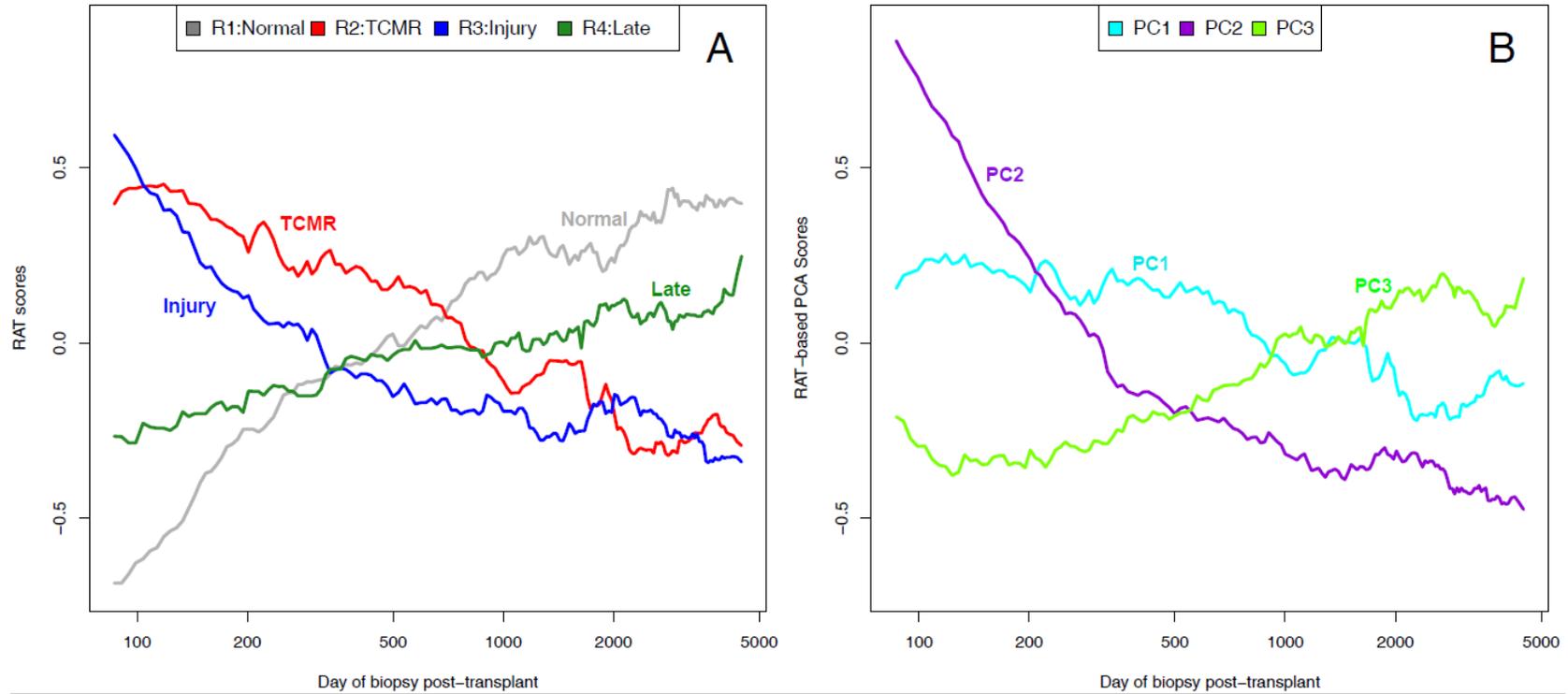
**Figure 6.3** Scree plot from this archetypal analysis showing the residual sum of squares (y-axis) versus the potential number of archetypes in a model (x-axis). The selected archetype model is highlighted in red.

CHAPTER 6: Liver rejection

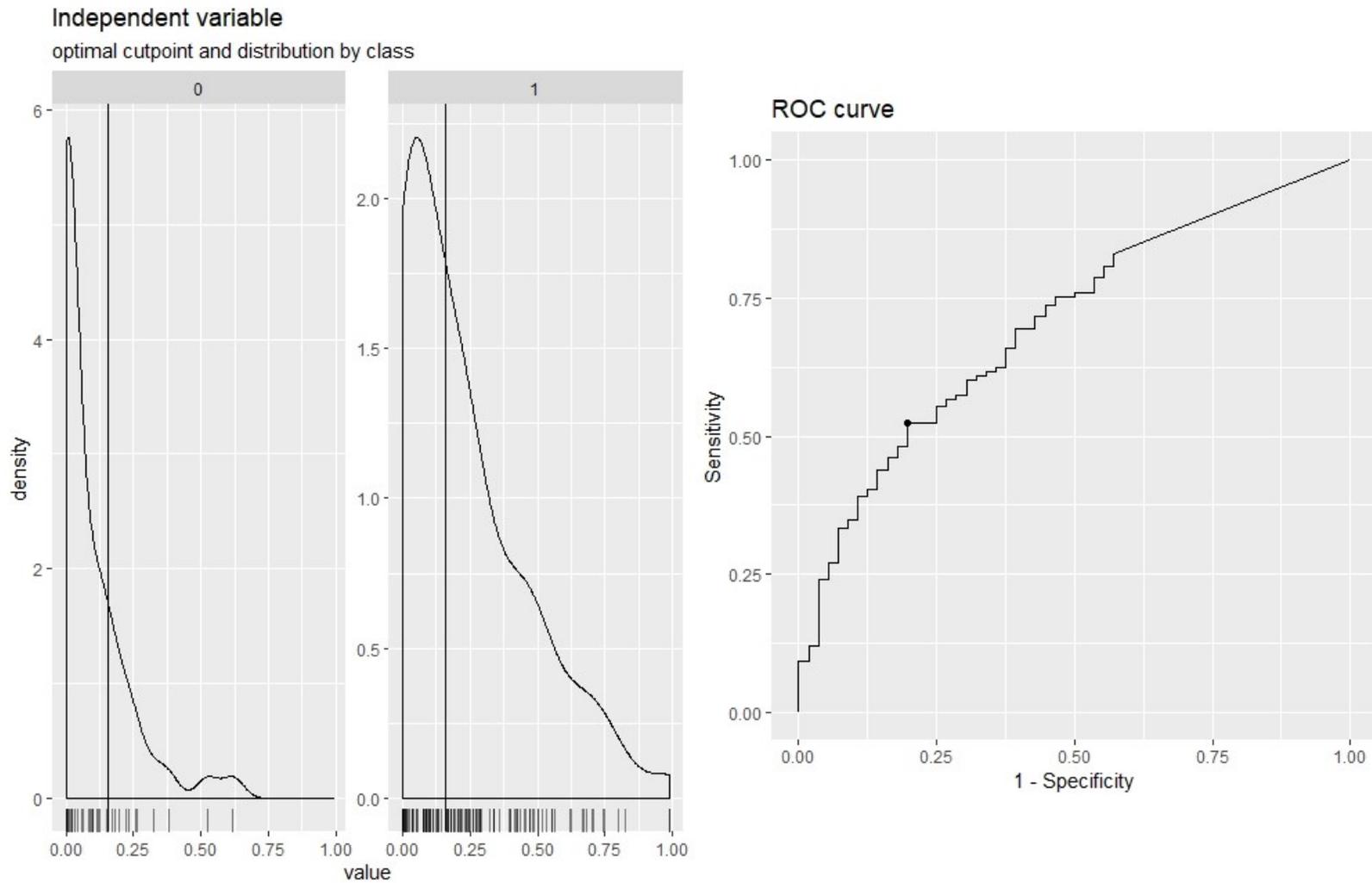


**Figure 6.4 Unsupervised Rejection-based analysis of 235 (218 patients) liver transplant biopsies.** Liver biopsies were separated by their expression of rejection-associated transcripts (RATs) in PCA. Archetypal analysis identified four major phenotypes, or “archetypes”: A1, A2, A3, and A4. Each biopsy was given four archetype scores describing their similarity to each archetype. Biopsies are grouped by their highest archetype score into clusters: R1 (N=129), R2 (N=37), R3 (N=61), and R4 (N=8). Panel B plots principal component 2 versus principal component 1, and panel C plots principal component 2 versus principal component 3. Triangles represent biopsies taken in the first two weeks post-transplant. Groups were named based on characterization by gene expression (see Results).

CHAPTER 6: Liver rejection



**Figure 6.5 Moving average of Rejection (RAT4AA) archetypes (A) and PCA scores (B).** Left-aligned 75-sample (Panel A) and 75-sample (Panel B) moving averages were calculated on all 235 liver biopsies ordered by time of biopsy post-transplant. The x-axis tracks the number of days elapsed between biopsy and transplant on a logarithmic scale. Scores were normalized in the full set of 235 biopsies. Scores were scaled prior to plotting.



**Figure 6.6** Establishing a balanced cutoff for the  $R2_{TCMR}$  score in predicting histologic acute rejection (TCMR) with the highest possible combined sensitivity/specificity. Left panels show the cutoff point marked by a vertical line over the '0' population ('without disease') and the '1' population ('with disease'). The optimized sensitivity and specificity is marked by a '●' on the right hand panel.

CHAPTER 6: Liver rejection

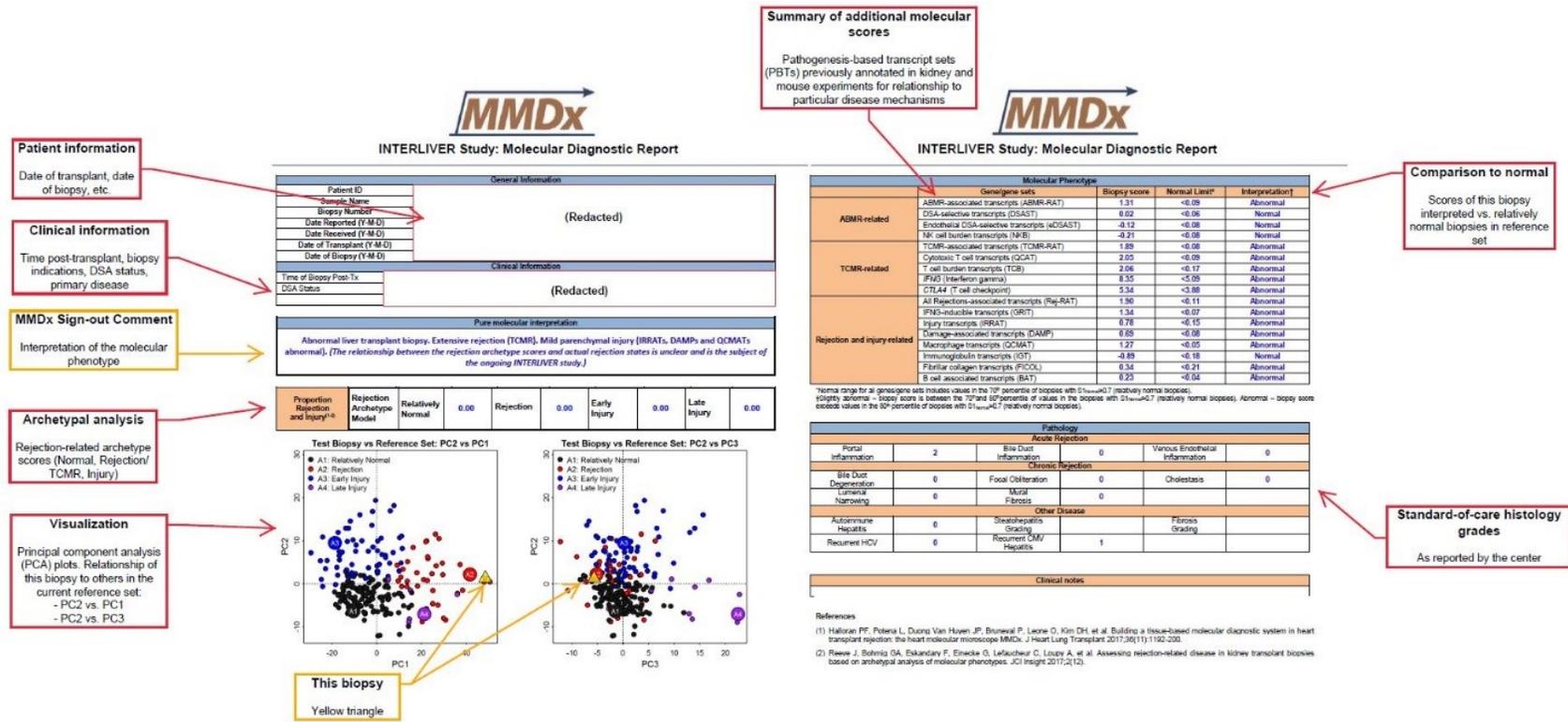


Figure 6.7 Example of the Molecular Microscope® Diagnostic System report on a new liver biopsy sample. The report is based on the reference set of 235 previously characterized liver biopsy samples. Left: page 1 of the report (molecular interpretation of the biopsy results, AA, PCA); right: page 2 of the report (additional molecular PBTs).

## **CHAPTER 7**

### **ASSESSING INJURY PHENOTYPES IN A LIVER TRANSPLANT POPULATION USING MMDX: A NEW METHOD OF DETECTING STEATOHEPATITIS**

**7.1 OVERVIEW AND HYPOTHESIS**

Variation in the assignment of histologic features and in the assignment of clinical diagnoses reduce precision and accuracy<sup>127,129,257</sup> for some injury phenotypes i.e. steatohepatitis. Biochemistry values are not sufficiently specific or sensitive for liver conditions such as NAFLD, steatohepatitis, or advanced fibrosis.<sup>91,92</sup> Recent data-driven approaches using molecular technologies in other transplanted organs indicate that these histology-based diagnostic systems frequently produce an incorrect diagnosis (up to 40-50% in abnormal kidney and heart transplants<sup>36,292</sup> and more in lung biopsies<sup>126</sup>), posing significant risk for harming patients with inappropriate or misinformed treatment.<sup>125</sup> This risk presents a major unmet need for those managing liver transplants.<sup>63,265</sup> Previous studies in heart and lung transplants have also shown that long-term survival in transplanted organs is mainly predicted by the amount and type of injury present in the tissue.<sup>262,293</sup> The detection and quantification of early, acute injury or late-stage fibrotic injury remains an important goal in liver transplantation, especially as liver transplants should function for decades if properly maintained.

Prior studies have shown that MMDx, particularly through machine learning, can overcome a flawed gold standard and provide phenotyping results that are more likely to be clinically accurate,<sup>6</sup> that continuous quantitative numbers derived from data are more useful than categorical data for describing a patient's disease,<sup>6</sup> that MMDx results are robust despite sampling and tissue heterogeneity,<sup>64</sup> and that MMDx is capable of discovering new biological mechanisms and disease features beyond what is available via the SOC. MMDx was developed previously for the diagnosis of acute rejection (TCMR) in a liver biopsy population (see Chapter 6). These results indicated that MMDx would be useful in the clinical management of liver transplants and could be successfully developed for both rejection and injury.

In these analyses, the injury-related phenotypes in a population of liver transplant biopsies were explored using supervised and unsupervised analyses (machine learning classifiers, PCA, AA). The prominence of these molecular injury phenotypes over time, and their relationships to biochemistry, histology, and clinical diagnoses were assessed. Molecular injury phenotypes and principal components were assessed by gene expression and top pathway terms in each group were established by overrepresentation analysis. Finally, a machine learning classifier trained on SOC clinical labels for steatohepatitis versus no steatohepatitis was developed.



## Chapter 7: Liver injury

identified associated immunologic profiles. METHODS: We conducted a cross-sectional study of 157 stable, long-term pediatric recipients of transplanted livers (70 boys; > 6 years old at time of transplantation; mean,  $8.9 \pm 3.46$  years after liver transplantation) who underwent liver biopsy analysis from August 13, 2012, through May 1, 2014. Participants had received livers from a living or deceased donor and had consistently normal results from liver tests. Liver biopsy specimens were scored by a central pathologist; an unsupervised hierarchical cluster analysis of histologic features was used to sort biopsy samples into 3 clusters. We conducted transcriptional and cytometric analyses of liver tissue samples and performed a systems biology analysis that incorporated clinical, serologic, histologic, and transcriptional data. RESULTS: The mean level of alanine aminotransferase in participants was  $27.6 \pm 14.57$  U/L, and the mean level of  $\gamma$ -glutamyl transferase was  $17.4 \pm 7.93$  U/L. Cluster 1 was characterized by interface activity ( $n = 34$ ), cluster 2 was characterized by periportal or perivenular fibrosis without interface activity ( $n = 45$ ), and cluster 3 had neither feature ( $n = 78$ ). We identified a module of genes whose expression correlated with levels of alanine aminotransferase, class II donor-specific antibody, portal inflammation, interface activity, perivenular inflammation, portal and perivenular fibrosis, and cluster assignment. The module was enriched in genes that regulate T-cell-mediated rejection (TCMR) of liver and other transplanted organs. Functional pathway analysis showed overrepresentation of TCMR gene sets for cluster 1 but not clusters 2 or 3. CONCLUSION: In an analysis of biopsies from an apparently homogeneous group of stable, long-term pediatric liver transplant recipients with consistently normal liver test results, we found evidence of chronic graft injury (inflammation and/or fibrosis). Biopsy samples with interface activity had a gene expression pattern associated with TCMR.

title:"Gastroenterology", "DOI": "10.1053/j.gastro.2018.08.023", "ISSN": "1528-

0012", "issue": "6", "journalAbbreviation": "Gastroenterology", "language": "eng", "note": "PMID: 30144432\nPMCID: PMC6279538", "page": "1838-

1851.e7", "source": "PubMed", "title": "Evidence of Chronic Allograft Injury in Liver Biopsies From Long-term Pediatric Recipients of Liver

Transplants", "volume": "155", "author": [{"family": "Feng", "given": "Sandy"}, {"family": "Bucuvalas", "given": "John C."}, {"family": "Demetris", "given": "Anthony

J."}, {"family": "Burrell", "given": "Bryna E."}, {"family": "Spain", "given": "Katherine M."}, {"family": "Kanaparathi", "given": "Sai"}, {"family": "Magee", "given": "John

C."}, {"family": "Ikle", "given": "David"}, {"family": "Lesniak", "given": "Andrew"}, {"family": "Lozano", "given": "Juan J."}, {"family": "Alonso", "given": "Estella

M."}, {"family": "Bray", "given": "Robert A."}, {"family": "Bridges", "given": "Nancy E."}, {"family": "Doo", "given": "Edward"}, {"family": "Gebel", "given": "Howard

M."}, {"family": "Gupta", "given": "Nitika A."}, {"family": "Himes", "given": "Ryan W."}, {"family": "Jackson", "given": "Annette M."}, {"family": "Lobritto", "given": "Steven

J."}, {"family": "Mazariegos", "given": "George V."}, {"family": "Ng", "given": "Vicky L."}, {"family": "Rand", "given": "Elizabeth B."}, {"family": "Sherker", "given": "Averell

H."}, {"family": "Sundaram", "given": "Shikha"}, {"family": "Turmelle", "given": "Yumirle P."}, {"family": "Sanchez-Fueyo", "given": "Alberto"}], "issued": {"date-

parts": [{"2018"}]}, "schema": "https://github.com/citation-style-language/schema/raw/master/csl-citation.json"}<sup>1</sup> biopsies (See Chapter 6).

PCA was used to visualize the biopsy population in terms of expression of injury-associated transcripts. PC1, PC2, and PC3 comprised 60%, 14%, and 10% of the variation within the population, respectively (**Figure 7.1**). PC1 was characterized by an increase in expression related to all injury PBTs, with a negative relationship to time (TxBx) and parenchymal transcripts (represented by KT1s). PC2 was

## Chapter 7: Liver injury

defined by a positive relationship with IGTs and time, but a negative relationship with DAMPs. Finally, PC3 was defined as a positive relationship with macrophage-related and parenchymal PBTs (AMAT1s and QCMATs, and KT1s respectively), and a negative relationship with all other injury PBTs (IGTs, IRITs, IRRATs, DAMPs) and TxBx. These relationships are summarized in **Table 7.5**.

### 7.4 AA OF LIVER INJURY PHENOTYPES

AA was used to separate liver biopsies by their expression of injury-associated transcripts defined by the PBTs (see 7.3). AA was done using the ‘archetypes’ package for R version 1.1.463.<sup>200</sup> Previous experience with AA in both livers and in other transplanted organs was used to select the appropriate number of clusters for this analysis.

Biopsies were given four scores that define their relationship to the four idealized biopsies or ‘archetypes’ representative of the four major groups. Grouping the biopsies allows the relationship between scores and clinical scenarios to be studied.

#### 7.4.1 AA group selection

Unsupervised AA was used to identify clusters representing heterogeneity in the population based on the expression of the previously selected injury-related PBTs. The four-archetype model contained four groups we were expecting to find based on experience in prior kidney, heart, and lung injury analyses: no injury/relatively normal, severe acute injury from rejection episodes or other diseases causing hepatocyte damage, early-stage injury as the liver is damaged post-implantation, and late-stage injury with fibrosis. Per convention, biopsies were then assigned to clusters by the highest AA score, producing the following clusters: I1 (N=149), I2 (N=36), I3 (N=17), and I4 (N=135).

Archetypal clusters were visualized using PCA, colored according to the AA cluster assignment: PC2 versus PC1 (**Figure 7.2A**) and PC2 versus PC3 (**Figure 7.2B**). Early biopsies taken within two weeks of transplant were highlighted with triangle symbols, and were found in all AA clusters, mostly in the lower region of PC2. PC1 separated injured biopsies from uninjured, PC2 separated early-stage injury from late-stage injury, and PC3 separated A2 and A4 from A1 and A3.

### **7.4.2 Characterization of injury in the four AA groups**

PBT transcripts used in the injury model were IQR filtered to select those with an IQR range >0.35 across all 337 livers. Injury-associated transcripts and PBTs were not re-derived in the liver, but have been demonstrated to function in kidney,<sup>267,295</sup> heart,<sup>198</sup> and lung.<sup>262</sup>

Injury-based AA groups were characterized based on their expression of the injury-related PBTs (**Table 7.6**). The median TxBx differed between archetype groups, earliest in I2 (82 days), followed by I3 (962 days), I1 (1051 days) and I4 (1371 days).

The I1 group lacked expression of all injury-associated PBTs compared to the other groups, but had the highest expression of transcripts associated with healthy parenchymal tissue (KT1s, KT2s). Biopsies belonging to the I2 group had the highest expression of inflammation and rejection-related PBTs (TCMR-RATs, QCATs, GRITs, Rejection-RATs) and PBTs related to macrophage infiltration (QCMATs, AMAT1s). The I3 group had the highest expression of endothelium-, late injury/fibrosis-, and recent injury-related transcript sets, as well as transcripts associated with atrophy-fibrosis (IGTs, BATs, and MCATs), and fibrillar collagens (FICOLs). I4 biopsies did not have the highest expression of any PBTs compared to other AA groups, but did have increased expression of all transcript sets compared to I1, including TCMR-related transcripts and IGTs. Based on these characteristics, we titled the groups: I1<sub>minimal</sub> for the relatively normal biopsy group, I2<sub>early-mild</sub> for the group with moderate injury and expression of some rejection-associated transcripts, I3<sub>severe</sub> for the severely injured group, and I4<sub>fibrosis</sub> for the group with atrophy-fibrosis changes related to immunoglobulin expression.

### **7.4.3 Transcripts associated with each injury archetype score**

The top 10 transcripts correlated with each archetype score are summarized in **Table 7.7**. As with previous analyses, primary interest resided in the transcripts strongly decreased in I1<sub>no injury</sub>, and the transcripts most increased in I2<sub>early-mild</sub>- I4<sub>fibrosis</sub>. Top 30 unique transcripts by Spearman correlation coefficient for each of the injury archetype groups are listed in **Tables 7.8-7.11**.

## Chapter 7: Liver injury

I1<sub>minimal</sub> negatively correlated with transcripts previously annotated for their relationships to injury and transplant rejection (e.g. *SEL1L3*, *HLA-DMA*). I2<sub>early-mild</sub> was positively correlated with transcripts annotated as inflammatory (injury- or transplant rejection-related, e.g. *LILRB2*, *CD163*). I3<sub>severe</sub> was positively correlated with transcripts annotated as severe injury-related, many of which were associated with matrix remodeling (e.g. *VCAN*, *COL4A1*). I4<sub>fibrosis</sub> was positively correlated with many immunoglobulin transcripts annotated for their relationship to injury over time (e.g. *IGKC*, *IGHG1*).

### 7.4.4 Injury AA group assignment versus previously defined rejection AA group assignment

To assess the relationships between the previously defined rejection AA groups and the new injury AA groups, we tabulated the biopsies assigned to each group. Since the N=235 population was used in the rejection analysis, the N for this table was limited to the common 235 biopsies between rejection and injury. It was important to consider that rejection AA defined a specific “early injury – no rejection” group, i.e. rejection AA isolates a form of injury independent from rejection episodes. The injury AA model does not isolate rejection into any particular AA group. Therefore, biopsy distribution between rejection and injury AA groups was not expected to be clean. Results are summarized in **Table 7.12**.

R1<sub>normal</sub> (N=129) was distributed mainly between the I1<sub>minimal</sub> (N=74) and I4<sub>fibrosis</sub> (N=54) groups. R2<sub>TCMR</sub> biopsies (N=37) fell mainly into the I2<sub>early-mild</sub> (N=13) and I4<sub>fibrosis</sub> (N=22). R3<sub>injury</sub> biopsies (N=61) were distributed amongst all injury AA groups: 26 biopsies into I1<sub>minimal</sub>, 10 biopsies into I2<sub>early-mild</sub>, 7 biopsies into I3<sub>severe</sub>, and 18 biopsies into I4<sub>fibrosis</sub>. Finally, R4<sub>late</sub> biopsies (N=8) distributed into I3<sub>severe</sub> (N=5) and I4<sub>fibrosis</sub> (N=3), but it was noted that this was an extremely small group so no conclusions should be drawn.

I1<sub>minimal</sub> biopsies (N=101) distributed mainly to R1<sub>normal</sub> (N=74) and R3<sub>injury</sub> (N=26). I2<sub>early-mild</sub> biopsies (N=23) distributed into R2<sub>TCMR</sub> (N=13) and R3<sub>injury</sub> (N=10). I3<sub>severe</sub> biopsies (N=14) were assigned mainly to R3<sub>injury</sub> (N=7) and R4<sub>late</sub> (N=5). Finally, I4<sub>fibrosis</sub> biopsies (N=97) were distributed amongst all four rejection AA groups: 54 biopsies in R1<sub>normal</sub>, 22 biopsies in R2<sub>TCMR</sub>, 18 biopsies in R3<sub>injury</sub> and 3 biopsies in R4<sub>late</sub>.

## **7.5 TIME COURSE OF MOLECULAR SCORES AND FEATURES**

Relationships between molecular scores, molecular features, and TxBx were visualized in moving average plots (**Figure 7.3**). Moving average plots were generated in R version 3.5.1 using the 'zoo' package.<sup>207</sup> Sample size for each graph was chosen based on line smoothness and varied between analyses. All plots were left-aligned and scores were normalized on all 337 biopsies.

PC1 and PC3 decreased over TxBx, while PC2 increased (**Figure 7.3A**, window size=100).

Relationships between PC scores and time were visualized (window sizes=50). While most PBTs (excluding KT1s) increased with PC1 (**Figure 7.3B-D, F-I**), many remained unchanged in PC2 and PC3. DAMPs decreased as PC2 increased (**Figure 7.3B**). As IGT and BAT scores increased, so did the PC2 score (**Figure 7.3E and F**). QCMATs increased slightly as PC3 increased (**Figure 7.3I**).

Standardized laboratory test scores for albumin, bilirubin, AST, ALT, and ALP were assessed for relationships to PC scores (**Figure 7.3J-L**, window size=125). Albumin decreased sharply as PC1 increased (**Figure 7.3J**), while all other biochemistry scores increased. Biochemistry was u-shaped over PC2 and PC3: albumin was low and bilirubin, AST, ALP, and ALT were high at the extreme ends of PC2 and PC3 scores (**Figure 7.3J and K**). These results indicate that biochemically abnormal biopsies were present at either extreme of PC2 and PC3 and the most normal biopsies were at mid-range PC2 or PC3.

## **7.6 CORRELATIONS BETWEEN INJURY ARCHETYPES AND CLINICAL FEATURES**

Relationships were assessed between the injury archetype scores, PCA scores, and the histology features recorded by each center (**Table 7.13**). Correlation coefficients with an absolute value >0.20 were considered significant and are highlighted. All relationships with correlation coefficients >0.20 had significant p values ( $p < 0.05$ ).

PC1 positively correlated with acute rejection and inflammation (portal and venous inflammation), and autoimmune hepatitis. PC2 strongly positively correlated with fibrosis (Spearman correlation

## Chapter 7: Liver injury

coefficient 'SCC'=0.47). PC3 positively correlated with portal inflammation, steatohepatitis, and CMV hepatitis (SCC=0.32, 0.26, and 0.22 respectively).

The I1<sub>minimal</sub> score was negatively correlated with portal and venous inflammation (SCC= -0.27 and -0.24), as well as autoimmune hepatitis (-0.24), fibrosis (-0.21), and CMV hepatitis (-0.21). The I2<sub>early-mild</sub> score was positively correlated with portal inflammation (SCC=0.25) and negatively correlated with fibrosis (SCC=-0.38). The I3<sub>severe</sub> score was positively correlated with autoimmune hepatitis and fibrosis (SCC=0.23 and 0.22), and negatively correlated with steatohepatitis (-0.20). The I4<sub>fibrosis</sub> score was positively correlated with fibrosis (SCC=0.36).

### 7.7 OVERREPRESENTATION ANALYSIS OF ARCHETYPE GROUP TOP TRANSCRIPT LISTS

Overrepresentation analysis was used to study the top transcripts associated with each archetype group (**Table 7.14**). Top 300 transcripts associated with each archetype score were used as input (top 300 decreased for I1<sub>minimal</sub>, top 300 increased in I2-I4, as our primary interest was in these transcripts).

Top transcripts anti-correlated with the I1<sub>minimal</sub> score were mainly associated with ontology terms related to neutrophil activity (e.g. neutrophil activation, neutrophil degranulation) with some terms associated with leukocytes (e.g. T cell activation, leukocyte cell-cell adhesion). Ontology terms overrepresented in the top 300 transcripts associated with the I2<sub>early-mild</sub> score were related to innate immune activity (e.g. neutrophil activation, regulation of innate immune response). Ontology terms overrepresented in the top 300 transcripts associated with the I3<sub>severe</sub> score were related to cellular matrix remodeling and atrophy-fibrosis (e.g. collagen-containing extracellular matrix, collagen trimer). Ontology terms overrepresented by the top transcripts associated with the I4<sub>fibrosis</sub> score were related to lymphocyte activity, mainly due to the concentration of immunoglobulin transcripts in the I4<sub>fibrosis</sub> score transcript list (e.g. T cell activation, regulation of lymphocyte activation).

### 7.8 LIVER FUNCTION TESTS AND INJURY ARCHETYPE GROUPS

## Chapter 7: Liver injury

Relationships were studied between liver function test results as reported by the center and the injury archetype groups. Mean, median, and range values for albumin, bilirubin, AST, ALT, and ALP per injury archetype group are summarized in **Table 7.15**.

Mean albumin values were highest in the I1<sub>minimal</sub> group – consistent with the lack of injury, and lowest in the I3<sub>severe</sub> group consistent with severe injury. Conversely, bilirubin and ALP were both highest in the I3<sub>severe</sub> group, and lowest in the I1<sub>minimal</sub> group. AST and ALT were highest in the I2<sub>early-mild</sub> group and lowest in I1<sub>minimal</sub>.

### 7.9 DEVELOPMENT OF A MOLECULAR CLASSIFIER FOR STEATOHEPATITIS

Linear discriminant analysis (Lda)-based machine learning was used to create a classifier (called ‘Molecular steatohepatitis greater than 0’ or ‘Msgt0’) trained on histologic steatohepatitis grades as reported by the center. Grades were binarized into a positive (with disease) class and a negative (without disease) class, to accommodate the classifier training. The classifier used 10-fold CV (trained on 9 folds, testing in the remaining 1, then repeated 10 times) to assign scores to the 337 population. The top 20 probe sets by p-value that were differentially expressed in the binary phenotypes (selected within each CV training set iteration) were used as classifier input as per previously published protocols.<sup>35</sup>

#### 7.9.1 Steatohepatitis classifier performance versus other molecular features

The Msgt0 classifier predicted histologic steatohepatitis with an AUC of 0.84 (**Figure 7.4A**). This performance was compared to that of other molecular scores for predicting steatohepatitis: the PC1 score (**Figure 7.4B**), PC2 score (**Figure 7.4C**), PC3 score (**Figure 7.4D**), I1<sub>minimal</sub> score (**Figure 7.4E**), I2<sub>early-mild</sub> score (**Figure 7.4F**), I3<sub>severe</sub> score (**Figure 7.4G**), and I4<sub>fibrosis</sub> score (**Figure 7.4H**). The PC3 score predicted steatohepatitis with an AUC of 0.73, a performance much lower than that of the Msgt0 classifier (**Figure 7.4D**). The I2<sub>early-mild</sub> score predicted steatohepatitis with an AUC=0.67 (**Figure 7.4F**), while the I3<sub>severe</sub> score negatively predicted steatohepatitis with an AUC=0.64 (**Figure 7.4G**).

## Chapter 7: Liver injury

### 7.9.2 Top transcripts associated with steatohepatitis by t-test

Top transcripts associated with steatohepatitis (half increased in steatohepatitis, half decreased) were mainly annotated in PBTs for relationships with parenchymal tissue function (e.g. *AASS*, *K1AA1191*, *PEL12*, *SLC12A1*, **Table 7.16**). Some transcripts were annotated for their relationship to injury, e.g. *TMEM154*, *FOS*, *SLC19A1*, *CCAAT*. No pattern based on previous PBT annotations was visible, suggesting that this pattern is novel (and has the potential for insights into the metabolic and inflammatory events in steatohepatitis).

### 7.9.3 Injury PCA visualization of steatohepatitis versus no steatohepatitis as called by histology and by the molecular classifier

PCA was used to visualize the distribution of biopsies called steatohepatitis positive versus negative by either the histologic classification (which was used to train the molecular classifier), or by the molecular classifier (**Figure 7.5**). In the case of the molecular classifier, an optimal cutoff of 0.2 was established using the `cutpointR` function in R (**Figure 7.6**).<sup>283</sup> Red dots represented biopsies called steatohepatitis in either case. The histologic steatohepatitis classification necessarily excluded biopsies for which steatohepatitis grades were not available (N=131). Molecular classifier predictions were available for all 337 biopsies.

The injury PBTs were not expected to be the basis for the identification of steatohepatitis but there were some relationships. While biopsies with histologic steatohepatitis did not cluster in PC1/PC2 (**Figure 7.5A**), they did distribute to the right in PC3 (**Figure 7.5B**). Biopsies called molecular steatohepatitis distributed similarly, with no clustering in PC1/PC2 (**Figure 7.5C**) but a shift towards the right in PC3 (**Figure 7.5D**). The concentration of steatohepatitis positive biopsies in either case shifted towards the right in PC3 was in agreement with the positive relationship seen in other analyses between PC3 and steatohepatitis.

### 7.9.4 Logistic regression between the `Msgt0` classifier, PC3 score, and steatohepatitis

## Chapter 7: Liver injury

Univariate logistic regression showed that both the PC3 score and particularly the Msgt0 classifier were related to histologic steatohepatitis ( $p=0.006$  and  $p=8.2 \times 10^{-14}$ , respectively). When the cross-validated Msgt0 classifier scores and PC3 scores were both combined in a regression model predicting histologic steatohepatitis, the classifier ( $p=1.1 \times 10^{-13}$ ) outperformed the injury PC3 score ( $p=0.008$ ), as expected.

### 7.9.5 Overrepresentation of genes associated with steatohepatitis

Analysis of overrepresented transcripts was done using Gene Ontologies (BP, MF, and CC terms). Analysis used the top 300 unique steatohepatitis transcripts by p-value (derived from a t test of histologic steatohepatitis versus no histologic steatohepatitis). Pathway terms were considered significant if  $p < 0.01$ . All significant terms were ordered by adjusted p value, and the top ten terms summarized in

#### Table 7.17.

Top terms associated with steatohepatitis were related to metabolism, mainly catabolic or metabolic processes. Biosynthetic and carboxylic processes were also noted in top pathways. These results indicate that metabolic dysregulation is a hallmark of steatohepatitis.

## 7.10 INTERPRETATION OF RESULTS

These analyses were intended to explore the molecular phenotypes of injury in a population of SOC liver transplant biopsies. PCA of 337 liver transplant biopsies prospectively collected from 13 international centers using injury-related PBT expression as input revealed that PC1 was characterized by injury versus no injury, PC2 by a positive relationship with immunoglobulin transcripts and time, and PC3 by a positive relationship with macrophage-related and parenchymal PBTs. PC1 and PC3 increased over time, while PC2 decreased. Unsupervised AA of the biopsy population gave 4 distinct groups: I<sub>1minimal</sub>, the relatively normal group; I<sub>2early-mild</sub>, the group with moderate injury and some rejection-like gene expression (e.g. *LILRB2*, *CD163*); I<sub>3severe</sub>, the group with severe injury and very high expression of all injury-related PBTs (e.g. *VCAN*, *COL4A1*); and I<sub>4fibrosis</sub> for the group with moderate injury related to

## Chapter 7: Liver injury

immunoglobulin expression (e.g. *IGKC*, *IGHG1*) at a late TxBx. Many PBT scores increased with PC1 but remained relatively unchanged in PC2 and PC3, with the exception of an increase in IGTs and BATs as the PC2 score increased, and a slight increase in BATs as PC3 increased. PC2 separated IGTs and DAMPs, e.g. early acute versus late stage fibrotic injury. PC3 separated rejection-like injury associated with inflammatory processes from all other injury. Laboratory test values became more abnormal with higher PC1 scores but were abnormal at both extremes of PC2 and PC3. The PC1 and I2<sub>moderate</sub> scores were positively correlated with portal inflammation. The PC2, I3<sub>severe</sub> and I4<sub>late</sub> scores were positively related to histologic fibrosis, while I1<sub>minimal</sub> and I2<sub>early-mild</sub> were negatively related. The PC3 score positively correlated with steatohepatitis, and the I3<sub>severe</sub> score was negatively related. Overrepresentation analysis identified a lack of injury-induced inflammation (neutrophil activity) in I1<sub>minimal</sub>, innate immune activity in I2<sub>early-mild</sub>, cellular matrix remodeling and atrophy-fibrosis in I3<sub>severe</sub>, and fibrosis-related pathways in I4<sub>fibrosis</sub> (likely due to the predominance of immunoglobulin transcripts). A classifier for histologic steatohepatitis (Msgt0) predicted histologic steatohepatitis in this population with a cross-validated AUC of 0.84. Top transcripts associated with the classifier were mainly annotated for relationships to parenchymal tissue function.

Steatohepatitis remains an important yet challenging clinical diagnosis by SOC. Analysis of relationships between molecular and clinical features showed that the PC3 and I2 scores were positively correlated with steatohepatitis, while the I3<sub>severe</sub> score was negatively related. AUCs for these molecular features predicting SOC steatohepatitis were fair, at 0.73 for the PC3 score, 0.67 for the I2<sub>early-mild</sub> score, and 0.64 for the I3<sub>severe</sub> score. Therefore, unknown biopsies with a high PC3 and I2<sub>early-mild</sub> score and low I3<sub>severe</sub> score are more likely to have steatohepatitis. Steatohepatitis is characterized by both mild tissue inflammation and fat deposition – two relatively nonspecific disease features that contribute to the variability in clinical steatohepatitis diagnoses. The I3 score showed a lack of inflammation compared to some other molecular phenotypes (i.e. I2<sub>early-mild</sub>), while the PC3 score showed an increase in PBTs normally related to inflammatory processes (like T cell-mediated rejection). PC3 was also positively associated with macrophage transcripts, but negatively associated with all other selected injury PBTs; possibly a result of the inflammatory processes and parenchymal changes occurring as the PC3 score

## Chapter 7: Liver injury

increases. Further analyses of these molecular features may reveal more insights into the disease mechanisms in clinical steatohepatitis.

Biochemistry results as recorded by the SOC at the centers showed overall abnormalities as PC1 increased, confirming that biopsies with a higher PC1 were at more severe stages of injury, tissue damage, or fibrosis. PC2 and PC3 were both abnormal at either the high or low extreme end of scores in this population and peaked to the most normal patterns (high albumin, low bilirubin, AST, ALT, and ALP) in the middle around 0.0 in PC2 and -0.3 in PC3. Despite this result, the upper end (>0) of either PC2 or PC3 was slightly more abnormal, indicating that biopsies with high PC1, PC2, and PC3 were the most severely injured. This description fit the I4<sub>fibrosis</sub> and some of the I3<sub>severe</sub> group, both of which showed signs of serious molecular injury through relationships to high expression of injury-related PBTs, biochemical abnormalities, and relationships to clinical injury.

Overrepresentation analysis revealed that a lack of markers for myeloid cell recruitment (neutrophil activity pathways) corresponded to a lack of injury, as seen in the I1<sub>minimal</sub> group. Neutrophils have been implicated in other literature as related to an acute response to recent or ongoing liver injury or hepatic stress,<sup>296</sup> including alcohol-induced injury<sup>297</sup> and other forms of inflammatory liver injury.<sup>298</sup> Neutrophil migration can be triggered by distressed hepatic cells and recruited by inflammatory mediators, and cause damage once they accumulate in the sinusoids. Terms in I2<sub>early-mild</sub> were mainly evidence of neutrophil activity and innate immune response, similar to the neutrophil-mediated acute injury noted in the literature. Overrepresentation also identified immunoglobulin transcripts that translated to lymphocyte pathways in I4<sub>fibrosis</sub>. These immunoglobulin transcripts (annotated as BAT and IGT PBTs) were also assessed previously for their association with interface inflammation in liver tissue,<sup>113</sup> a form of inflammation and necrosis of the liver parenchyma. Finally, terms in I3<sub>severe</sub> showed matrix remodeling and collagen organization, evidence of severe parenchymal damage, fibrogenesis, and angiogenesis.

PC scores over time showed that the biopsies in this liver population got less abnormal over time, as PC1 and PC3 decreased and PC2 increased. Biopsies taken closer to the time of transplant were more likely to have high PC1 and PC3 scores, and thus be biochemically abnormal with high DAMPs, IRRATs, IRITs, IGTs, AMATs, and BATs – all markers of tissue injury, damage, and parenchymal

## Chapter 7: Liver injury

dedifferentiation. This is likely due to the noted increased incidence of rejection in the early TxBx, donation-implantation injury, or other forms of acute stress. The chronic and less severe injury represented by PC2 did increase slightly over time, representing the biopsies that accumulate signs of atrophy-fibrosis over several years. Similar results regarding normalness were seen in the rejection analyses (Chapter 6).

Archetype groups were identified that described distinct sets of biopsies within the population. While it was initially hypothesized that one group may represent steatohepatitis, it was unsurprising that this group was not isolated. AA will not identify rare phenotypes as a cluster in smaller populations, and we estimate the incidence of steatohepatitis in this population to be 12% based on the clinical SOC-assigned diagnoses. It was noted that many of the scores had relatively strong positive or negative relationships with steatohepatitis – PC3, I2, and I3 – and these scores may be used to detect biopsies with a higher risk for having clinical steatohepatitis. This objective data, especially in larger populations as the ongoing INTERLIVER study progresses, would be very welcome to improve the precision and reduce interobserver variability in these challenging diagnoses.

The steatohepatitis Msgt0 classifier had excellent cross-validated predictive performance with an AUC of 0.84. While machine learning cannot circumvent all errors in a set of training labels, prior experiments with label training that contained an intentional amount of error showed that machine learning is capable of ‘re-labeling’ samples post-training. Even if labels contain errors, the classifier is capable of reassigning these biopsies to the correct label and producing a better set of biopsy assignments than what was originally available. This is evidence suggesting that, although the steatohepatitis classifier is trained on undoubtedly noisy labels as assigned by the SOC, the performance of the classifier may be an improvement over the SOC. This is a feature of machine learning; it is capable of overcoming and correcting errors in the training set. Further analysis in a larger population will be done as the INTERLIVER study expands.

Some limitations of these analyses include the relatively small sample size, and reliance on the SOC for supervised analysis that contains known levels of noise or interobserver variation. However, it is important to note that a clinically relevant test must relate to the SOC and be able to operate within less-

## Chapter 7: Liver injury

than-ideal SOC conditions with labels and data from international centers. Therefore, we chose to focus on SOC as opposed to undertaking an expensive task of central review. In published analyses, critical central review has never been shown to be superior to SOC assessments, and in fact the central reviewers did not agree.<sup>36,42</sup> In the case where two reviewers disagree, it is impossible to resolve who is correct. Focusing on the SOC was the best method for these analyses and has worked in kidney, heart and lung previously.

Precision, accuracy, and objective diagnostic information is a major unmet need in the management of liver transplants, and past analyses in kidney, heart and lung have shown significant associations between tissue injury and graft survival. While these analyses are done in a recent population without available outcome data, similar findings are expected as the INTERLIVER study expands. MMDx is more precise in other transplanted organs than histology and will achieve this in liver transplants as well. The development of molecular classifiers and features capable of predicting steatohepatitis or fibrosis in an unknown biopsy would be valuable to the clinician who is balancing risks and attempting to manage the transplant potentially over several decades. These analyses suggest that it may be beneficial to retain a portion of a core in RNA*later*<sup>TM</sup> for molecular analysis via MMDx to improve diagnostics, monitor the patient in real time by RNA expression, and determine their long-term risk of permanent tissue damage to the allograft. With these methods, precision can be increased, and clinicians can have more confidence in their decision-making.

## **Chapter 7: Liver injury**

### **7.11 TABLES**

## Chapter 7: Liver injury

**Table 7.1** Participating Collaborators in the INTERLIVER study

Name	Institution	Location	# of biopsies contributed
Stephan Gray	University of Maryland	Baltimore, MD, USA	5
Seth Karp and Roman Perri	Vanderbilt University	Nashville, TN, USA	1
Jorge Reyes	University of Washington	Seattle, WA, USA	2
Amar Gupta	Baylor University Medical Center	Dallas, TX, USA	3
Goran Klintmalm			
Marwan Abouljoud	Henry Ford Hospital	Detroit, MI, USA	29
Iman Francis			
Dilip Moonka			
Rosa Miquel	King's College London	London, UK	5
Alberto Sanchez-Fueyo			
Grzegorz Piecha	Medical University of Silesia	Katowice, Poland	17
Marta Gryczman	Pomeranian Medical University	Szczecin, Poland	53
Krzysztof Jurczyk			
Joanna Mazurkiewicz			
Marek Myślak			
Marta Wawrzynowicz-Syczewska			
Samir Zeair			
Aldo Montano-Loza	University of Alberta	Edmonton, AB, Canada	7
Martina Brozynski	University of California San Francisco	San Francisco, CA, USA	15
Sandy Feng			
Monique Koenigsberg			
David Bowen	University of Sydney	Sydney, NSW, Australia	22
Fiona Guan			
Ken Liu			
Avik Majumdar			
Geoff McCaughan			
Simone Strasser			
Tatiana Tsoutsman			
Michael Akyeampong	Virginia Commonwealth University	Richmond, VA, USA	55
Jeanette Amery			
Chandra Bhatti			
Johanna Christensen			
Adrian Cotterell			
Megan Gray			
Becky Hickey			
Aamir Khan			
Marlon Levy			
Trevor Reichman			
Amit Sharma			
Vanessa Taylor	Warsaw Medical University	Warsaw, Poland	123
Michał Ciszek			
Dominika Dęborska-Materkowska			
Magdalena Durlik			
Bartosz Foronczewicz			
Michał Grat			
Krzysztof Mucha			
Agnieszka Perkowska-Ptasinska			
Olga Tronina			
Krzysztof Zieniewicz			

**Chapter 7: Liver injury**

**Table 7.2** INTERLIVER patient and biopsy characteristics

<b>Patient characteristics</b>	<b>Patients N = 311</b>
<b>Recipient sex (% total)</b>	
<i>Male</i>	146 (49%)
<i>Female</i>	149 (51%)
<b>Recipient age at transplant (median, range)</b>	50 (2-71)
<b>Primary disease (% total)<sup>A</sup></b>	
<i>Alcoholic Liver Disease</i>	43 (14%)
<i>Autoimmune hepatitis</i>	24 (8%)
<i>Hepatitis B</i>	14 (5%)
<i>Hepatitis C</i>	46 (15%)
<i>Hepatocellular carcinoma</i>	25 (8%)
<i>Non-alcoholic Steatohepatitis</i>	17 (5%)
<i>Primary Biliary Cholangitis</i>	18 (6%)
<i>Primary Sclerosing Cholangitis</i>	29 (9%)
<i>Other</i>	41 (13%)
<i>Missing</i>	73 (23%)
<b>Biopsy characteristics</b>	<b>Biopsies N = 337</b>
<b>Days (median, range) from transplant to biopsy (TxBx)</b>	904 (0 -12569)
<b>Immunosuppression at biopsy (% total)</b>	
<i>Corticosteroids</i>	3 (<1%)
<i>Cyclosporine</i>	6 (2%)
<i>Tacrolimus</i>	77 (23%)
<i>Missing</i>	251 (74%)
<b>Indication for biopsy (% total)</b>	
<i>Indication: clinician concerned about graft function</i>	238 (83%)
<i>Follow-up from previous biopsy</i>	5 (2%)
<i>Protocol/surveillance</i>	45 (16%)
<i>Missing</i>	49 (15%)

<sup>A</sup>Some patients fell under multiple categories

**Chapter 7: Liver injury**

**Table 7.3** Laboratory test data and DSA for all biopsies in the INTERLIVER study (N=337)

<b>Laboratory Tests</b>	<b>Values in all biopsies</b> Mean (median, range)
<i>Albumin (g/dL)</i>	3.8 (4.0, 1.7-5.4) N=269
<i>Bilirubin (mg/dL)</i>	2.3 (0.9, 0.3-36.0) N=299
<i>AST (IU/L)</i>	117.6 (46.0, 11-5779) N=298
<i>ALT (IU/L)</i>	126 (67, 8-1781) N=226
<i>ALP (IU/L)</i>	204 (121, 29-1863) N=298
<b>DSA at biopsy</b>	<b># of results</b> (% of all known results)
<i>Positive</i>	25 (7%)
<i>Negative</i>	8 (2%)
<i>Not tested</i>	304 (90%)

NOTE. Missing DSA values included those not provided by the center, or instances where the test was not done within a relevant time period of the biopsy ( $\pm 7$  days).

## Chapter 7: Liver injury

**Table 7.4** Pathogenesis-based transcript sets<sup>A,B</sup> (PBTs) used in liver analyses

Relationship to Injury	PBT	Description	Detail
Increased in injury	IRRAT	Injury-repair response associated transcripts	Transcript set estimating kidney transplant injury, developed in early transplants <sup>1</sup>
	IRITD3	Injury and rejection induced transcripts – intermediate TxBx	Human orthologues of mouse genes induced by non-immune kidney injury in isografts, peaking ~day 3 post-transplant in mouse kidney transplants <sup>2</sup>
	IRITD5	Injury and rejection induced transcripts – late TxBx	Human orthologue of mouse genes induced by non-immune kidney injury in isografts, peaking ~day 5 post-transplant in mouse kidney transplants <sup>2</sup>
	DAMP	Damage-associated molecular pattern transcripts	Literature-based damage-associated molecular pattern (DAMP) transcripts annotated as markers of cellular stress <sup>3,4</sup>
Decreased in injury	KT1	Kidney transcripts - Set 1	Human orthologues of genes with high expression in normal human kidney tissue <sup>5</sup>
	KT2	Kidney transcripts - Set 2	Human orthologues of genes with high expression in normal human kidney tissue <sup>5</sup>
Macrophage infiltration	QCMAT	Quantitative Constitutive Macrophage-Associated Transcripts	Transcripts with high expression in human primary macrophages, not inducible by IFNG, and high correlation with amounts of macrophage RNA in a sample <sup>6</sup>
	AMAT	Alternative Macrophage Associated Transcripts	Alternative activation of macrophages in mouse model of ischemic acute kidney injury <sup>6</sup>
Increased in atrophy-fibrosis	IGT	Immunoglobulin transcripts	Time-dependent increase in injured tissue reflect plasma cell infiltrate <sup>7</sup>
	BAT	B cell-associated transcripts	Transcripts with 5x expression in B cells compared to other immune cells <sup>7</sup>
	MCAT	Mast cell-associated transcripts	Transcripts highly correlated with scarring (fibrosis) in allograft biopsies <sup>8</sup>
Increased in acute injury	FICOL	Fibrillar collagens	Increase in expression of transcripts reflecting fibrillar collagens in response to wounding.
Increased in rejection/inflammation	TCMR-RAT	T cell-mediated rejection-associated transcripts	Transcripts associated with TCMR in kidney transplants, high expression in CD8+CTL cells <sup>9</sup>
	Rej-RAT	General transplant rejection-associated transcripts	Transcripts increased in ABMR/TCMR biopsies, developed in human kidney biopsies <sup>10</sup>
	QCAT	Quantitative CTL-Associated Transcripts	Transcripts with highest expression in CD8+CTL cells compared to normal tissue or other immune cells, developed in kidney transplant biopsies <sup>9</sup>
	GRIT	Gamma-IFN and rejection-induced transcripts <sup>1</sup>	Human orthologues of IFNG-inducible mouse genes, initially developed in mouse kidney transplants <sup>11</sup>

A. <https://www.ualberta.ca/medicine/institutes-centres-groups/atagc/research/gene-lists>

B. The gene sets were empirically derived in human cell lines, human transplants, and mouse models. They reflect biological processes relevant to rejection and injury. The gene set score is a ratio between mean gene set expression in each cluster and a set of four nephrectomies.

\*Denotes PBTs incorporated in the Injury 4AA and Injury PCA model, but shown for reference in Figure 1.

### References

- Famulski KS, de Freitas DG, Kreepala C, Chang J, Sellares J, Sis B, et al. Molecular phenotypes of acute kidney injury in human kidney transplants. *Journal of the American Society of Nephrology*. 2012;23(5):948-58.
- Famulski KS, Broderick G, Einecke G, Hay K, Cruz J, Sis B, et al. Transcriptome analysis reveals heterogeneity in the injury response of kidney transplants. *Am J Transplant*. 2007;7(11):2483-95.
- Land WG, Agostinis P, Gasser S, Garg AD, and Linkermann A. Transplantation and Damage-Associated Molecular Patterns (DAMPs). *Am J Transplant*. 2016;16(12):3338-61.
- Heil M, and Land WG. Danger signals - damaged-self recognition across the tree of life. *Front Plant Sci*. 2014;5:578.
- Einecke G, Broderick G, Sis B, Halloran PF. Early loss of renal transcripts in kidney allografts: relationship to the development of histologic lesions and alloimmune effector mechanisms. *American Journal of Transplantation*. 2007;7(5):1121-30.
- Famulski KS, Einecke G et al. Defining the canonical form of T-cell-mediated rejection in human kidney transplants. *Am. J. Transplant*. 2010;10(4):810-20
- Einecke G, Reeve J, Mengel M, Sis B, Bunnag S, Mueller TF, et al. Expression of B cell and immunoglobulin transcripts is a feature of inflammation in late allografts. *American Journal of Transplantation*. 2008;8(7):1434-43.
- Mengel M, Reeve J, Bunnag S, Einecke G, Sis B, Mueller T, Kaplan B, Halloran PF. Molecular correlates of scarring in kidney transplants: the emergence of mast cell transcripts. *Am. J. Transplant*. 2009;9(1):169-78.
- Hidalgo LG, Einecke G, Allanach K, Mengel M, Sis B, Mueller TF, Halloran PF. The transcriptome of human cytotoxic T cells: measuring the burden of CTL-associated transcripts in human kidney transplants. *Am J. Transplant*. 2008;8(3):637-46.
- Halloran PF, et al. Building a tissue-based molecular diagnostic system in heart transplant rejection: The heart Molecular Microscope Diagnostic (MMDx) System. *J Heart Lung Transplant*. 2017;36(11):1192-1200
- Famulski KS, Einecke G, Reeve J, et al. Changes in the transcriptome in allograft rejection: IFN-gamma-induced transcripts in mouse kidney allografts. *Am. J. Transplant*. 2006;6(6):1342-54.

Chapter 7: Liver injury

**Table 7.5** Correlations between pathogenesis-based transcript (PBT) set scores and Injury (INJ4AA) PCA scores

PBT/feature	Spearman correlation		
	PC1	PC2	PC3
TxBx	-0.25	<b>0.47</b>	-0.23
<b>Increased in injury</b>			
<i>DAMP – Damage-associated molecular pattern transcripts*</i>	0.49	<b>-0.50</b>	-0.12
<i>IRRAT – Injury/repair associated transcripts (human kidney)*</i>	<b>0.84</b>	0.03	-0.13
<i>IRITD3 – tissue injury and repair associated transcripts*</i>	<b>0.90</b>	0.02	0.05
<i>IRITD5 – tissue injury and repair associated transcripts*</i>	<b>0.90</b>	0.18	0.14
<i>clRIT – cardiac injury and repair induced transcripts</i>	<b>0.86</b>	-0.05	0.29
<b>Fibrillar collagens</b>			
<i>FICOL – fibrillar collagen-associated transcripts</i>	<b>0.70</b>	0.23	-0.02
<b>Normal parenchymal tissue</b>			
<i>KT1 – kidney parenchymal transcripts 1</i>	-0.56	-0.06	-0.18
<i>KT2 – kidney parenchymal transcripts 2</i>	-0.66	-0.03	-0.05
<b>Macrophage infiltration</b>			
<i>QCMAT – macrophage associated transcripts*</i>	<b>0.82</b>	0.05	0.59
<i>AMAT1 – alternative macrophage activation-associated transcripts*</i>	<b>0.89</b>	0.04	0.42
<b>Increased in atrophy-fibrosis</b>			
<i>IGT – Immunoglobulin transcripts*</i>	0.45	<b>0.76</b>	-0.12
<i>BAT – B cell-associated transcripts</i>	<b>0.73</b>	0.46	0.08
<i>MCAT – Mast cell-associated transcripts</i>	0.04	<b>0.47</b>	-0.34
<b>TCMR-related transcripts</b>			
<i>TCMR-RAT – TCMR-associated RATs</i>	<b>0.73</b>	0.24	<b>0.51</b>
<i>QCAT – Cytotoxic T cell associated transcripts</i>	<b>0.65</b>	0.31	<b>0.41</b>
<b>Rejection-related</b>			
<i>GRIT – Interferon gamma-inducible transcripts</i>	<b>0.75</b>	0.10	<b>0.49</b>
<i>Rejection-RATs – rejection associated RATs</i>	<b>0.68</b>	0.18	<b>0.54</b>
<b>Endothelium-related transcripts</b>			
<i>ENDAT – Endothelial cell-associated transcripts</i>	<b>0.72</b>	0.21	-0.15

<sup>1</sup> PBT scores represent the mean fold difference in PBT expression between biopsies in each archetype group and the previously defined R1 biopsies as a relatively control. Biopsies were grouped according to their highest of the four archetype scores. The highest absolute score in each row is bolded.

\* Indicates a PBT used in the AA Injury Model and in the PCA.

Chapter 7: Liver injury

**Table 7.6** Mean pathogenesis-based transcript (PBT) set scores in biopsies grouped according to their highest Injury (INJ4AA) archetype score and correlations between PBT set scores and archetypes scores (N=337)

	I1 <sub>minimal</sub> (n = 149)	I2 <sub>early-mild</sub> (n = 36)	I3 <sub>severe</sub> (n = 17)	I4 <sub>fibrosis</sub> (n = 135)
<b>Median time of biopsy post-transplant (in days)</b>	1051	<u>82</u>	962	<b>1371</b>
	<b>Mean PBT score in each archetype group<sup>1</sup> (± SD)</b>			
PBT	I1 <sub>minimal</sub> (n = 149)	I2 <sub>early-mild</sub> (n = 36)	I3 <sub>severe</sub> (n = 17)	I4 <sub>fibrosis</sub> (n = 135)
<b>Increased in injury</b>				
<i>DAMP – Damage-associated molecular pattern transcripts*</i>	1.05 (±1.20)	1.36 (±1.20)	<b>1.40 (±1.23)</b>	0.99 (±1.13)
<i>IRRAT – Injury/repair associated transcripts (human kidney)*</i>	0.96 (±1.31)	1.45 (±1.32)	<b>2.74 (±1.34)</b>	1.11 (±1.24)
<i>IRIT3 – tissue injury and repair associated transcripts*</i>	1.00 (±1.12)	1.22 (±1.10)	<b>1.60 (±1.14)</b>	1.09 (±1.11)
<i>IRIT5 – tissue injury and repair associated transcripts*</i>	0.97 (±1.13)	1.24 (±1.14)	<b>1.68 (±1.30)</b>	1.11 (±1.14)
<i>cIRIT – cardiac injury and repair induced transcripts</i>	0.99 (±1.10)	1.22 (±1.09)	<b>1.32 (±1.11)</b>	1.07 (±1.09)
<b>Fibrillar collagens</b>				
<i>FICOL – fibrillar collagen-associated transcripts</i>	0.95 (±1.48)	1.42 (±1.55)	<b>3.78 (±2.07)</b>	1.29 (±1.51)
<b>Normal parenchymal tissue</b>				
<i>KT1 – kidney parenchymal transcripts 1</i>	0.98 (±1.13)	0.89 (±1.09)	0.76 (±1.26)	0.93 (±1.13)
<i>KT2 – kidney parenchymal transcripts 2</i>	0.97 (±1.24)	0.82 (±1.17)	0.58 (±1.48)	0.91 (±1.20)
<b>Macrophage infiltration</b>				
<i>QCMAT – macrophage associated transcripts*</i>	0.94 (±1.15)	<b>1.58 (±1.23)</b>	1.34 (±1.32)	1.19 (±1.20)
<i>AMAT1 – alternative macrophage activation-associated transcripts*</i>	0.92 (±1.25)	<b>1.68 (±1.23)</b>	1.82 (±1.37)	1.25 (±1.23)
<b>Increased in atrophy fibrosis</b>				
<i>IGT – Immunoglobulin transcripts*</i>	0.76 (±1.34)	0.80 (±1.37)	<b>1.78 (±1.91)</b>	<b>1.34 (±1.58)</b>
<i>BAT – B cell-associated transcripts</i>	0.93 (±1.08)	1.06 (±1.14)	<b>1.38 (±1.23)</b>	<b>1.09 (±1.15)</b>
<i>MCAT – Mast cell-associated transcripts</i>	0.87 (±1.46)	0.68 (±1.28)	<b>1.90 (±2.40)</b>	<b>0.99 (±1.54)</b>
<b>TCMR-related transcript sets</b>				
<i>TCMR-RAT – TCMR-associated RATs</i>	0.90 (±1.18)	<b>1.61 (±1.46)</b>	1.39 (±1.39)	1.27 (±1.32)
<i>QCAT – Cytotoxic T cell associated transcripts</i>	0.86 (±1.24)	<b>1.51 (±1.59)</b>	1.38 (±1.51)	1.27 (±1.38)
<b>Rejection-related transcript sets</b>				
<i>GRIT – Interferon gamma-inducible transcripts</i>	0.94 (±1.19)	<b>1.49 (±1.30)</b>	1.38 (±1.22)	1.20 (±1.23)
<i>Rejection-RATs – rejection associated RATs</i>	0.86 (±1.28)	<b>1.69 (±1.49)</b>	1.32 (±1.35)	1.29 (±1.37)
<b>Endothelium-related transcript sets</b>				
<i>ENDAT – Endothelial cell-associated transcripts</i>	0.96 (±1.13)	1.09 (±1.15)	<b>1.60 (±1.28)</b>	1.04 (±1.13)

<sup>1</sup> Score represents the mean fold difference in PBT expression between biopsies in each archetype group and previously defined R1 biopsies as a relatively normal control. Biopsies were grouped according to their highest of the four archetype scores. The highest mean score in each row is bolded.

\* Indicates a PBT used in the AA Injury Model and in the PCA.

Chapter 7: Liver injury

Table 7.7 Top 10 transcripts correlated with the Injury4A I1, I2<sub>early-mild</sub>, I3<sub>severe</sub>, and I4<sub>late</sub> scores

I1 <sub>minimal</sub>			I2 <sub>early</sub>			I3 <sub>severe</sub>			I4 <sub>fibrosis</sub>		
Gene Symbol	Correlation with I1 <sub>normal</sub>	PBT Annotations	Gene Symbol	Correlation with I2 <sub>early</sub>	PBT Annotations	Gene Symbol	Correlation with I3 <sub>severe</sub>	PBT Annotations	Gene Symbol	Correlation with I4 <sub>late</sub>	PBT Annotations
<i>CD53</i>	-0.856	-	<i>FCGR3A</i>	0.643	Rej-RAT	<i>VCAN</i>	0.753	IRRAT30	<i>IGKC</i>	0.604	IGT
<i>KLHL6</i>	-0.846	-	<i>FPR2</i>	0.607	LivGST_UP	<i>LXN</i>	0.737	IRITD3	<i>IGHG3</i>	0.598	IGT
<i>DOCK2</i>	-0.833	IRRAT950	<i>LILRB2</i>	0.607	QCMAT	<i>EFEMP1</i>	0.733	IRITD3	<i>IGHG1</i>	0.596	IGT
<i>SEL1L3</i>	-0.833	IRRAT950	<i>MOB1A</i>	0.593	cIRIT	<i>ANXA1</i>	0.721	IRITD3	<i>IGK</i>	0.587	IGT
<i>EVI2A</i>	-0.832	IRRAT30	<i>HAVCR2</i>	0.589	IRRAT950	<i>PMP22</i>	0.717	-	<i>IGKV1-5</i>	0.577	IGT
<i>CECR1</i>	-0.830	-	<i>NFAM1</i>	0.586	-	<i>COL1A2</i>	0.716	FICOL	<i>IGKV1-27</i>	0.574	IGT
<i>HLA-DMA</i>	-0.826	Rej-RAT	<i>VERSUSIG4</i>	0.585	IRRAT950	<i>COL6A3</i>	0.708	IRITD5	<i>IGKV3-20</i>	0.569	IGT
<i>LAPTM5</i>	-0.826	IRRAT950	<i>FCGR1A</i>	0.584	TCMR-RAT	<i>COL4A1</i>	0.707	IRITD3	<i>IGKV1-39</i>	0.567	IGT
<i>HLA-DMB</i>	-0.826	Rej-RAT	<i>ADGRE2</i>	0.580	IRRAT950	<i>TUBB6</i>	0.702	IRITD3	<i>IGKV3-11</i>	0.564	IGT
<i>OSBPL3</i>	-0.819	-	<i>CD163</i>	0.579	AMAT1	<i>CDH11</i>	0.699	IRITD5	<i>IGLV1-41</i>	0.534	IGT

**Chapter 7: Liver injury**

**Table 7.8** Top 30 unique transcripts associated with Injury4AA Archetype Score 1 (I1 minimal), sorted by Spearman correlation coefficient.

Gene Symbol	Name	PBT	Spearman Correlation Coefficient	P value	Expression in archetype groups			
					I1	I2	I3	I4
CD53	CD53 molecule		-0.856	3.8E-98	272	717	670	480
KLHL6	kelch-like family member 6		-0.846	1E-93	24	45	59	37
DOCK2	dedicator of cytokinesis 2	IRRAT950	-0.833	3.2E-88	92	202	189	147
SEL1L3	sel-1 suppressor of lin-12-like 3 (C. elegans)	IRRAT950	-0.833	4.8E-88	220	352	507	379
EVI2A	ecotropic viral integration site 2A	IRRAT30	-0.832	7.3E-88	41	100	124	74
CECR1	cat eye syndrome chromosome region, candidate 1		-0.830	5E-87	280	650	530	499
HLA-DMA	major histocompatibility complex, class II, DM alpha	Rej-RAT	-0.826	1.8E-85	309	753	693	606
LAPTM5	lysosomal protein transmembrane 5	IRRAT950	-0.826	2.3E-85	506	1258	1240	800
HLA-DMB	major histocompatibility complex, class II, DM beta	Rej-RAT	-0.826	2.7E-85	573	1347	1098	1050
OSBPL3	oxysterol binding protein-like 3		-0.819	7.3E-83	32	71	87	50
CCDC109B	coiled-coil domain containing 109B		-0.817	2.9E-82	49	135	189	87
CLEC7A	C-type lectin domain family 7, member A	AMAT1	-0.814	5.1E-81	92	256	206	161
MARCH1	membrane associated ring finger 1	RAT	-0.813	7.6E-81	121	311	269	202
CYBA	cytochrome b-245, alpha polypeptide	cIRIT	-0.812	2.6E-80	1005	2033	1860	1484
HLA-DRA	major histocompatibility complex, class II, DR alpha	Rej-RAT	-0.812	3E-80	5028	8444	8090	7813
DOCK8	dedicator of cytokinesis 8		-0.811	8.1E-80	152	346	341	244
WIPF1	WAS/WASL interacting protein family, member 1	RAT	-0.809	2E-79	53	118	120	81
ITGA4	integrin alpha 4	IRRAT950	-0.808	8.5E-79	74	158	178	128
FYB	FYN binding protein	RAT	-0.806	1.9E-78	145	357	340	249
TFEC	transcription factor EC	QCMAT	-0.806	3.4E-78	25	77	56	43
FAM105A	family with sequence similarity 105, member A	CT1	-0.802	7.6E-77	29	53	53	39
CD74	CD74 molecule, major histocompatibility complex, class II invariant chain	Rej-RAT	-0.801	1.3E-76	1237	2561	2409	2199
TNFAIP8	tumor necrosis factor, alpha-induced protein 8		-0.800	1.8E-76	54	137	179	91
HLA-DPB1	major histocompatibility complex, class II, DP beta 1	Rej-RAT	-0.800	2.8E-76	434	1029	1070	830
GLIPR1	GLI pathogenesis-related 1	IRRAT950	-0.800	3.5E-76	48	95	141	77
RGS10	regulator of G-protein signaling 10		-0.799	5.2E-76	91	204	250	159
PTPRC	protein tyrosine phosphatase, receptor type, C	IRRAT30	-0.797	3.3E-75	265	638	610	462
SYK	spleen tyrosine kinase	cIRIT	-0.796	6.1E-75	74	127	133	103
LYZ	lysozyme	QCMAT	-0.795	8.9E-75	981	2654	3388	2103

**Chapter 7: Liver injury**

**Table 7.9** Top 30 unique transcripts associated with Injury4AA Archetype Score 2 (I2 early-minimal), sorted by Spearman correlation coefficient.

Gene Symbol	Name	PBT	Spearman Correlation Coefficient	P value	Expression in archetype groups			
					I1	I2	I3	I4
FCGR3A	Fc fragment of IgG, low affinity IIIa, receptor (CD16a)	Rej-RAT	0.643	1.1E-40	1245	3016	1782	1819
FPR2	formyl peptide receptor 2	LivGST_UP	0.607	2.5E-35	30	103	74	38
LILRB2	leukocyte immunoglobulin-like receptor, subfamily B	QCMAT	0.607	3E-35	110	374	232	181
MOB1A	MOB kinase activator 1A	cIRIT	0.593	2E-33	383	599	542	452
HAVCR2	hepatitis A virus cellular receptor 2	IRRAT950	0.589	6.4E-33	35	80	55	48
NFAM1	NFAT activating protein with ITAM motif 1		0.586	1.9E-32	44	72	53	54
VERSUSIG4	V-set and immunoglobulin domain containing 4	IRRAT950	0.585	2.5E-32	346	801	517	435
FCGR1A	Fc fragment of IgG, high affinity Ia, receptor (CD64)	TCMR-RAT	0.584	3E-32	52	343	203	113
ADGRE2	adhesion G protein-coupled receptor E2	IRRAT950	0.580	1.1E-31	26	87	59	37
CD163	CD163 molecule	AMAT1	0.579	1.3E-31	902	1777	1353	1137
IFI30	interferon, gamma-inducible protein 30	GRIT3	0.579	1.5E-31	1447	4017	2869	2362
FCGR1B	Fc fragment of IgG, high affinity Ib, receptor (CD64)	RAT	0.573	7.6E-31	11	30	22	15
SLC7A7	solute carrier family 7 (amino acid transporter light chain, y+L system), member 7	KT2	0.571	1.7E-30	167	358	235	233
LILRB1	leukocyte immunoglobulin-like receptor, subfamily B	QCMAT	0.566	5.6E-30	50	125	73	71
HPSE	heparanase	GRIT2	0.564	1.2E-29	31	76	54	49
NCF2	neutrophil cytosolic factor 2	QCMAT	0.560	3.3E-29	43	105	92	61
LILRA2	leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 2		0.559	4.1E-29	41	84	55	53
SAMSN1	SAM domain, SH3 domain and nuclear localization signals 1		0.559	4.6E-29	118	291	240	172
CYBB	cytochrome b-245, beta polypeptide	RAT	0.558	6E-29	87	197	162	127
MFSD1	major facilitator superfamily domain containing 1	cIRIT	0.557	8E-29	698	1049	957	811
NABP1	nucleic acid binding protein 1		0.556	8.6E-29	118	230	202	151
C3AR1	complement component 3a receptor 1	LivGST_UP	0.556	1E-28	203	506	463	330
ADORA3	adenosine A3 receptor; transmembrane and immunoglobulin domain containing 3		0.555	1.2E-28	126	330	178	143
CD300A	CD300a molecule		0.554	1.6E-28	276	493	359	362
SIGLEC9	sialic acid binding Ig-like lectin 9		0.551	3.9E-28	53	85	60	63
NLRC4	NLR family, CARD domain containing 4		0.549	6.1E-28	14	31	22	19
CD86	CD86 molecule	QCMAT	0.549	6.9E-28	98	221	153	148
LILRA6	leukocyte immunoglobulin-like receptor, subfamily A	QCMAT	0.548	8.2E-28	91	229	192	127
STX11	syntaxin 11	Rej-RAT	0.548	9.3E-28	82	187	164	110

**Chapter 7: Liver injury**

**Table 7.10** Top 30 unique transcripts associated with Injury4AA Archetype Score 3 (I3 severe), sorted by Spearman correlation coefficient.

Gene Symbol	Name	PBT	Spearman Correlation Coefficient	P value	Expression in archetype groups			
					I1	I2	I3	I4
VCAN	versican	IRRAT30	0.753	7.9E-63	18	31	225	23
LXN	latexin	IRITD3	0.737	6.2E-59	14	24	84	19
EFEMP1	EGF containing fibulin-like extracellular matrix protein 1	IRITD3	0.733	4.2E-58	85	269	1650	176
ANXA1	annexin A1	IRITD3	0.721	2.1E-55	241	459	1031	343
PMP22	peripheral myelin protein 22		0.717	2.1E-54	187	302	862	233
COL1A2	collagen, type I, alpha 2	FICOL	0.716	3.7E-54	263	490	1719	413
COL6A3	collagen, type VI, alpha 3	IRITD5	0.708	1.6E-52	291	418	1345	365
COL4A1	collagen, type IV, alpha 1	IRITD3	0.707	2.7E-52	480	780	2385	674
TUBB6	tubulin, beta 6 class V	IRITD3	0.702	2.1E-51	242	405	1045	303
CDH11	cadherin 11, type 2, OB-cadherin (osteoblast)	IRITD5	0.699	8.7E-51	33	47	223	44
TUBA1A	tubulin, alpha 1a	IRITD5	0.692	2E-49	246	491	1001	348
S100A6	S100 calcium binding protein A6	IRITD3	0.690	5E-49	344	626	1347	483
GEM	GTP binding protein overexpressed in skeletal muscle	ENDAT	0.688	1.6E-48	16	23	95	22
PDLIM3	PDZ and LIM domain 3		0.687	2.1E-48	39	62	252	47
COL4A2	collagen, type IV, alpha 2	IRITD3	0.687	2.3E-48	344	498	1512	460
VIM	vimentin	IRITD3	0.686	3.1E-48	1806	3210	5269	2699
FBN1	fibrillin 1	IRITD5	0.686	4.2E-48	141	218	625	192
BTG2	BTG family, member 2		0.684	7.1E-48	42	76	152	64
PNMA1	paraneoplastic Ma antigen 1		0.684	8E-48	42	68	145	53
SOX4	SRY box 4		0.683	1.3E-47	89	122	336	127
IGFBP7	insulin like growth factor binding protein 7	ENDAT	0.682	1.6E-47	26	31	56	30
FAM60A	family with sequence similarity 60, member A		0.678	9.4E-47	95	128	196	118
LUM	lumican	IRITD5	0.678	1.3E-46	375	637	2391	672
C1orf198	chromosome 1 open reading frame 198		0.668	5.6E-45	173	245	646	227
LAMA2	laminin, alpha 2		0.668	7.4E-45	109	139	363	134
SEPT7	septin 7		0.667	1.1E-44	1231	1409	2094	1349
QPCT	glutaminy-peptide cyclotransferase	IRRAT950	0.665	1.9E-44	21	39	63	27
GJA1	gap junction protein alpha 1	IRITD3	0.665	2.1E-44	117	176	541	153
THBS2	thrombospondin 2	IRITD5	0.665	2.1E-44	89	210	866	184

**Chapter 7: Liver injury**

**Table 7.11** Top 30 unique transcripts associated with Injury4AA Archetype Score 4 (I4 fibrosis), sorted by Spearman correlation coefficient.

Gene Symbol	Name	PBT	Spearman Correlation Coefficient	P value	Expression in archetype groups			
					I1	I2	I3	I4
IGKC	immunoglobulin kappa constant	IGT	0.604	6.2E-35	97	102	237	211
IGHG3	immunoglobulin heavy constant gamma 3 (G3m marker)	IGT	0.598	4.2E-34	412	462	2356	1569
IGHG1	immunoglobulin heavy constant gamma 1 (G1m marker)	IGT	0.596	1E-33	597	671	3308	2173
IGK	immunoglobulin kappa locus	IGT	0.587	1.3E-32	1067	1375	6613	4060
IGKV1-5	immunoglobulin kappa variable 1-5	IGT	0.577	2.7E-31	198	235	1106	702
IGKV1-27	immunoglobulin kappa variable 1-27	IGT	0.574	5.6E-31	198	240	1235	785
IGKV3-20	immunoglobulin kappa variable 3-20; immunoglobulin kappa variable 3D-20	IGT	0.569	2.6E-30	28	41	207	119
IGKV1-39	immunoglobulin kappa variable 1-39 (gene/pseudogene)	IGT	0.567	4.5E-30	122	150	831	516
IGKV3-11	immunoglobulin kappa variable 3-11; immunoglobulin kappa variable 3D-11	IGT	0.564	1.2E-29	31	45	291	149
IGLV1-41	immunoglobulin lambda variable 1-41 (pseudogene)	IGT	0.534	3.1E-26	90	97	368	191
TNFRSF17	tumor necrosis factor receptor superfamily, member 17		0.528	1.4E-25	11	11	27	19
IGKV2-28	immunoglobulin kappa variable 2-28; immunoglobulin kappa variable 2D-28	IGT	0.522	6.4E-25	26	32	130	81
IGLL5	immunoglobulin lambda-like polypeptide 5	IGT	0.520	9E-25	10	13	85	25
JCHAIN	joining chain of multimeric IgA and IgM		0.518	1.4E-24	239	245	959	496
SEL1L3	sel-1 suppressor of lin-12-like 3 (C. elegans)	BAT	0.514	4.1E-24	109	163	213	177
IGLV1-40	immunoglobulin lambda variable 1-40	IGT	0.504	4.2E-23	36	35	77	59
ANKRD36BP2	ankyrin repeat domain 36B pseudogene 2		0.502	6.7E-23	8	8	21	13
IGLJ3	immunoglobulin lambda joining 3	IGT	0.497	2.1E-22	30	30	66	53
IGLV3-10	immunoglobulin lambda variable 3-10	IGT	0.492	6.1E-22	34	32	85	56
MZB1	marginal zone B and B1 cell-specific protein		0.489	1.2E-21	18	21	69	34
IGLL3P	immunoglobulin lambda-like polypeptide 3, pseudogene	IGT	0.489	1.2E-21	19	20	65	36
GZMK	granzyme K	QCAT	0.486	2.3E-21	107	192	191	215
EOMES	eomesodermin	Rej-RAT	0.477	1.6E-20	67	91	69	94
FCRL5	Fc receptor-like 5	BAT	0.477	1.6E-20	19	19	28	25
SLAMF7	SLAM family member 7	Rej-RAT	0.469	7.9E-20	93	245	168	198
THEMIS	thymocyte selection associated		0.466	1.4E-19	23	39	48	41
LY9	lymphocyte antigen 9	BAT	0.465	1.7E-19	22	23	27	28
CD3D	CD3d molecule, delta (CD3-TCR complex)	QCAT	0.463	2.4E-19	17	23	29	24
LAX1	lymphocyte transmembrane adaptor 1		0.462	3.4E-19	18	24	39	27
HLA-DPA1	major histocompatibility complex, class II, DP alpha 1	Rej-RAT	0.455	1.2E-18	3230	5816	5515	5623

**Chapter 7: Liver injury**

**Table 7.12** Crosstab of previously defined Rejection AA classes versus Injury AA classes (N=235)

		<b>Injury AA classes</b>			
		<b>I1<sub>minimal</sub></b> (N=101)	<b>I2<sub>early-mild</sub></b> (N=23)	<b>I3<sub>severe</sub></b> (N=14)	<b>I4<sub>fibrosis</sub></b> (N=97)
<b>Rejection AA classes</b>	<b>R1<sub>normal</sub></b> (N=129)	74	0	1	54
	<b>R2<sub>TCMR</sub></b> (N=37)	1	13	1	22
	<b>R3<sub>injury</sub></b> (N=61)	26	10	7	18
	<b>R4<sub>late</sub></b> (N=8)	0	0	5	3

**Chapter 7: Liver injury**

**Table 7.13** Correlations between injury archetype scores, PC scores and histologic rejection in liver biopsies (N=337)

Histology features	PC1	PC2	PC3	I1 <sub>minimal</sub> score	I2 <sub>early-mild</sub> score	I3 <sub>severe</sub> score	I4 <sub>fibrosis</sub> score
Acute rejection: portal inflammation	0.20, p=3E-03	-0.01, p=9E-01	0.32, p=2E-06	-0.27, p=6E-05	0.25, p=2E-04	0.01, p=9E-01	0.11, p=9E-02
Acute rejection: bile duct inflammation	0.13, p=5E-02	0.02, p=7E-01	0.16, p=2E-02	-0.14, p=4E-02	0.14, p=4E-02	0.06, p=4E-01	0.06, p=3E-01
Acute rejection: venous inflammation	0.20, p=3E-03	0.09, p=2E-01	0.18, p=1E-02	-0.24, p=3E-04	0.11, p=1E-01	0.13, p=6E-02	0.12, p=7E-02
Chronic rejection: bile duct degeneration	0.03, p=7E-01	0.09, p=2E-01	0.01, p=8E-01	-0.06, p=4E-01	-0.02, p=7E-01	0.04, p=6E-01	0.08, p=3E-01
Chronic rejection: focal obliteration	0.10, p=2E-01	0.13, p=6E-02	0.04, p=6E-01	-0.18, p=1E-02	-0.00, p=1E+00	0.04, p=6E-01	0.18, p=1E-02
Chronic rejection: cholestasis	0.17, p=2E-02	0.05, p=5E-01	0.04, p=6E-01	-0.12, p=9E-02	0.05, p=5E-01	0.17, p=2E-02	0.02, p=8E-01
Chronic rejection: mural fibrosis	0.10, p=2E-01	0.17, p=2E-02	-0.09, p=2E-01	-0.13, p=8E-02	-0.15, p=4E-02	0.10, p=2E-01	0.08, p=3E-01
Other disease: autoimmune hepatitis	0.21, p=2E-03	0.18, p=1E-02	0.09, p=2E-01	-0.24, p=4E-04	-0.04, p=6E-01	0.23, p=8E-04	0.08, p=2E-01
Other disease: steatohepatitis	-0.07, p=4E-01	-0.06, p=4E-01	0.26, p=8E-04	0.00, p=1E+00	0.16, p=3E-02	-0.20, p=8E-03	0.03, p=7E-01
<u>Other disease: fibrosis</u>	0.07, p=3E-01	0.47, p=1E-10	-0.13, p=1E-01	-0.21, p=7E-03	-0.38, p=3E-07	0.22, p=5E-03	0.36, p=2E-06
Other disease: recurrent HCV	0.13, p=5E-02	0.01, p=8E-01	0.09, p=2E-01	-0.13, p=6E-02	0.09, p=2E-01	0.10, p=1E-01	0.03, p=7E-01
Other disease: suspected CMV hepatitis	0.10, p=1E-01	0.04, p=6E-01	0.22, p=1E-03	-0.21, p=1E-03	0.09, p=2E-01	-0.06, p=4E-01	0.10, p=1E-01

NOTE. Clinical data is binary: 1 if positive, 0 if negative. Spearman correlation coefficients are given alongside p values approximated from the value of the coefficient.

Chapter 7: Liver injury

Table 7.14 Top 10 GO Terms from Injury AA group top 300 transcripts by Spearman correlation coefficient absolute value

Rank	I1 <sub>minimal</sub>				I2 <sub>early-mild</sub>				I3 <sub>severe</sub>				I4 <sub>fibrosis</sub>			
	ONT	GO ID	Pathway Description	P value	ONT	GO ID	Pathway Description	P value	ONT	GO ID	Pathway Description	P value	ONT	GO ID	Pathway Description	P value
1	BP	GO:0042119	neutrophil activation	2.62E-20	BP	GO:0002283	neutrophil activation involved in immune response	2.06E-30	CC	GO:0062023	collagen-containing extracellular matrix	3.61E-29	BP	GO:0042110	T cell activation	2.65E-41
2	BP	GO:0042110	T cell activation	1.24E-19	BP	GO:0042119	neutrophil activation	4.71E-30	BP	GO:0030198	extracellular matrix organization	8.78E-27	BP	GO:0051249	regulation of lymphocyte activation	2.65E-35
3	BP	GO:0007159	leukocyte cell-cell adhesion	2.11E-19	BP	GO:0043312	neutrophil degranulation	2.32E-29	CC	GO:0031012	extracellular matrix	3.25E-25	BP	GO:0002429	immune response-activating cell surface receptor signaling pathway	4.2E-34
4	BP	GO:0043312	neutrophil degranulation	1.08E-18	BP	GO:0002446	neutrophil mediated immunity	7.18E-29	BP	GO:0043062	extracellular structure organization	1.35E-24	BP	GO:0002768	immune response-regulating cell surface receptor signaling pathway	5.47E-34
5	BP	GO:0002283	neutrophil activation involved in immune response	1.29E-18	CC	GO:0030667	secretory granule membrane	2.31E-18	MF	GO:0005201	extracellular matrix structural constituent	1.19E-20	BP	GO:0050851	antigen receptor-mediated signaling pathway	2.36E-30
6	BP	GO:0002446	neutrophil mediated immunity	2.48E-18	BP	GO:0002768	immune response-regulating cell surface receptor signaling pathway	1E-15	CC	GO:0044420	extracellular matrix component	2.15E-14	BP	GO:0030098	lymphocyte differentiation	2.29E-27
7	BP	GO:0050863	regulation of T cell activation	6.63E-16	BP	GO:0045088	regulation of innate immune response	7.78E-15	CC	GO:0005604	basement membrane	1.6E-13	BP	GO:0051251	positive regulation of lymphocyte activation	1.23E-26
8	CC	GO:0042613	MHC class II protein complex	2.78E-15	CC	GO:0101002	ficolin-1-rich granule	1.55E-14	CC	GO:0005581	collagen trimer	1.54E-12	BP	GO:0002696	positive regulation of leukocyte activation	3.36E-26
9	BP	GO:1903039	positive regulation of leukocyte cell-cell adhesion	7.7E-15	BP	GO:0002429	immune response-activating cell surface receptor signaling pathway	1.24E-13	BP	GO:0030199	collagen fibril organization	2.2E-12	BP	GO:0050867	positive regulation of cell activation	9.93E-26
10	BP	GO:0042119	neutrophil activation	2.62E-20	BP	GO:0002283	neutrophil activation involved in immune response	2.06E-30	CC	GO:0062023	collagen-containing extracellular matrix	3.61E-29	BP	GO:0042110	T cell activation	2.65E-41

**Chapter 7: Liver injury**

**Table 7.15** Laboratory test data for biopsies in INTERLIVER by RAT4A or Injury4A archetype group

Values in biopsies belong to designated archetype group		Tests					TxBx in days mean (median, range)
		Albumin g/dL mean (median, range) (N=200)	Bilirubin mg/dL mean (median, range) (N=226)	AST IU/L mean (median, range) (N=226)	ALT IU/L mean (median, range) (N=226)	ALP IU/L mean (median, range) (N=225)	
Injury groups	<b>I1</b> minimal (N=149)	4.0 (4.1, 1.9-5.4) N=118	1.2 (0.7, 0.03-25.00) N=134	56.0 (31.5, 11-607) N=134	87.5 (42.5, 8-765) N=134	144.9 (101.0, 38-611) N=134	1841 (1051, 0-12569) N=149
	<b>I2</b> early-mild (N=36)	3.3 (3.3, 2.2-4.2) N=27	3.4 (0.25, 17.0) N=28	393.9 (95.0, 32-5779) N=28	311.5 (211.0, 36-1781) N=28	324.4 (189.0, 67-1467) N=27	438 (82, 0-5419) N=36
	<b>I3</b> severe (N=17)	3.1 (3.1, 1.7-4.3) N=16	7.0 (3.6, 0.5-31.1) N=17	311.9 (107, 19-1689) N=17	197.5 (112.0, 9-653) N=17	370.0 (252.0, 61-998) N=17	2078 (962, 0-10851) N=17
	<b>I4</b> fibrosis (N=135)	3.9 (4.0, 2.2-5.3) N=108	2.5 (1.0, 0.05-36.0) N=120	94.0 (58.0, 11-598) N=119	116.0 (73.5, 9-815) N=120	219.3 (124.0, 29-1863) N=120	2177 (1371, 8-11676) N=134

\*Highlighted cells are the highest value in the column

Chapter 7: Liver injury

**Table 7.16** Top 10 genes increased and decreased with histologic steatohepatitis ordered by adjusted p value

	P value	Adjusted p value	Gene Symbol	Gene Name	PBT Annotation	Expression in Steatohepatitis Biopsies	Expression in I2/I3/I4 Biopsies without Steatohepatitis	Expression in I1 Biopsies without Steatohepatitis
Top 10 genes with decreased expression in steatohepatitis biopsies	7.63E-10	1.26E-05	PPM1K	protein phosphatase, Mg2+		57	89	111
	1.05E-08	5.08E-05	CNPY3	canopy 3 homolog (zebrafish)		143	193	214
	2.07E-07	2.89E-04	AASS	aminoadipate-semialdehyde synthase	KT1	87	145	191
	5.40E-07	5.96E-04	FOS	FBJ murine osteosarcoma viral oncogene homolog	IRRAT30	66	328	328
	6.34E-07	6.56E-04	RBM33	RNA binding motif protein 33		116	140	150
	1.71E-06	1.42E-03	SH3PXD2A	SH3 and PX domains 2A		855	1203	1507
	2.07E-06	1.59E-03	KIAA1191	KIAA1191	KT1	1295	1553	1582
	2.91E-06	2.01E-03	SDS	serine dehydratase		86	190	268
	3.98E-06	2.35E-03	PELI2	pellino E3 ubiquitin protein ligase family member 2	HT1	36	51	51
	6.40E-06	3.07E-03	CEBPD	CCAAT	IRITD3	990	1470	1383
Top 10 genes with increased expression in steatohepatitis biopsies	4.72E-09	3.91E-05	TMEM154	transmembrane protein 154	BAT	64	42	40
	1.42E-08	5.08E-05	GOSR2	golgi SNAP receptor complex member 2		94	73	73
	1.53E-08	5.08E-05	SLC12A1	solute carrier family 12	KT1	18	11	10
	2.18E-08	6.01E-05	LOC100506538	uncharacterized LOC100506538		290	221	223
	6.90E-08	1.53E-04	FGF21	fibroblast growth factor 21		70	35	37
	1.13E-07	2.09E-04	ZBTB33	zinc finger and BTB domain containing 33	LT3	503	395	328
	1.38E-07	2.29E-04	H2AFY2	H2A histone family, member Y2		93	73	66
	2.09E-07	2.89E-04	ACOT12	acyl-CoA thioesterase 12	KT1	455	250	292
	1.09E-06	9.99E-04	FKBP9	FK506 binding protein 9, 63 kDa		35	28	27
	3.28E-06	2.17E-03	INHBE	inhibin, beta E		2199	969	1186

The biopsy population with the highest expression in each row is highlighted.

**NOTE.** P values as determined in a t test of 'steatohepatitis' biopsies versus 'non-steatohepatitis' biopsies. 'Non-steatohepatitis' biopsies were those with steatohepatitis grades =0, while 'steatohepatitis' biopsies were those with grades >0.

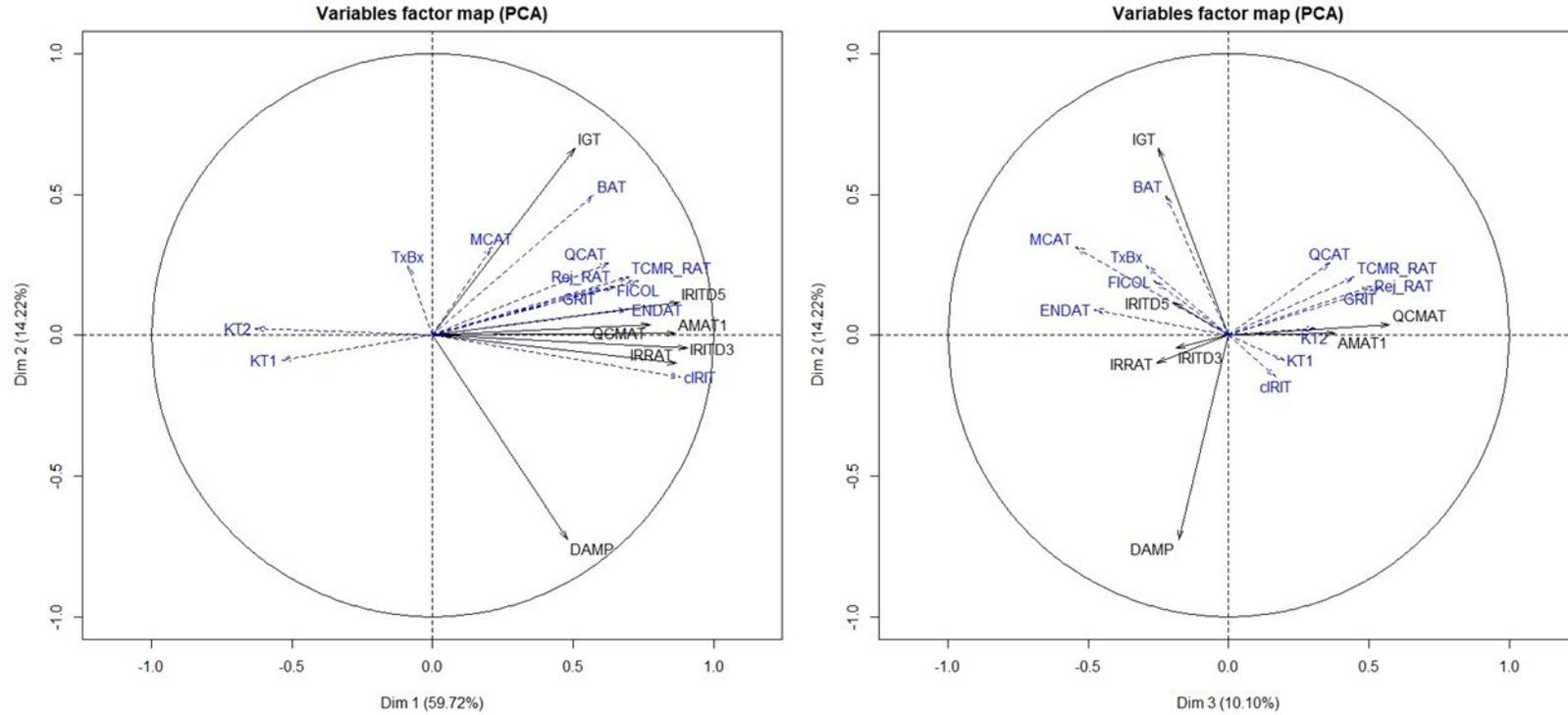
**Chapter 7: Liver injury****Table 7.17 Top 10 overrepresented GO Terms from steatohepatitis top 300 transcripts (increased and decreased) by Spearman correlation coefficient absolute value**

	<b>ONTOLOGY</b>	<b>GO ID</b>	<b>Pathway Description</b>	<b>P value</b>
<b>1</b>	BP	GO:0062012	regulation of small molecule metabolic process	1.59E-07
<b>2</b>	BP	GO:0009063	cellular amino acid catabolic process	8.72E-07
<b>3</b>	BP	GO:0006109	regulation of carbohydrate metabolic process	3.22E-06
<b>4</b>	BP	GO:0043470	regulation of carbohydrate catabolic process	4.51E-06
<b>5</b>	BP	GO:1901606	alpha-amino acid catabolic process	1.17E-05
<b>6</b>	BP	GO:0044282	small molecule catabolic process	1.2E-05
<b>7</b>	BP	GO:0006520	cellular amino acid metabolic process	1.5E-05
<b>8</b>	BP	GO:0006536	glutamate metabolic process	1.87E-05
<b>9</b>	BP	GO:0006110	regulation of glycolytic process	2.17E-05
<b>10</b>	BP	GO:0046394	carboxylic acid biosynthetic process	2.31E-05

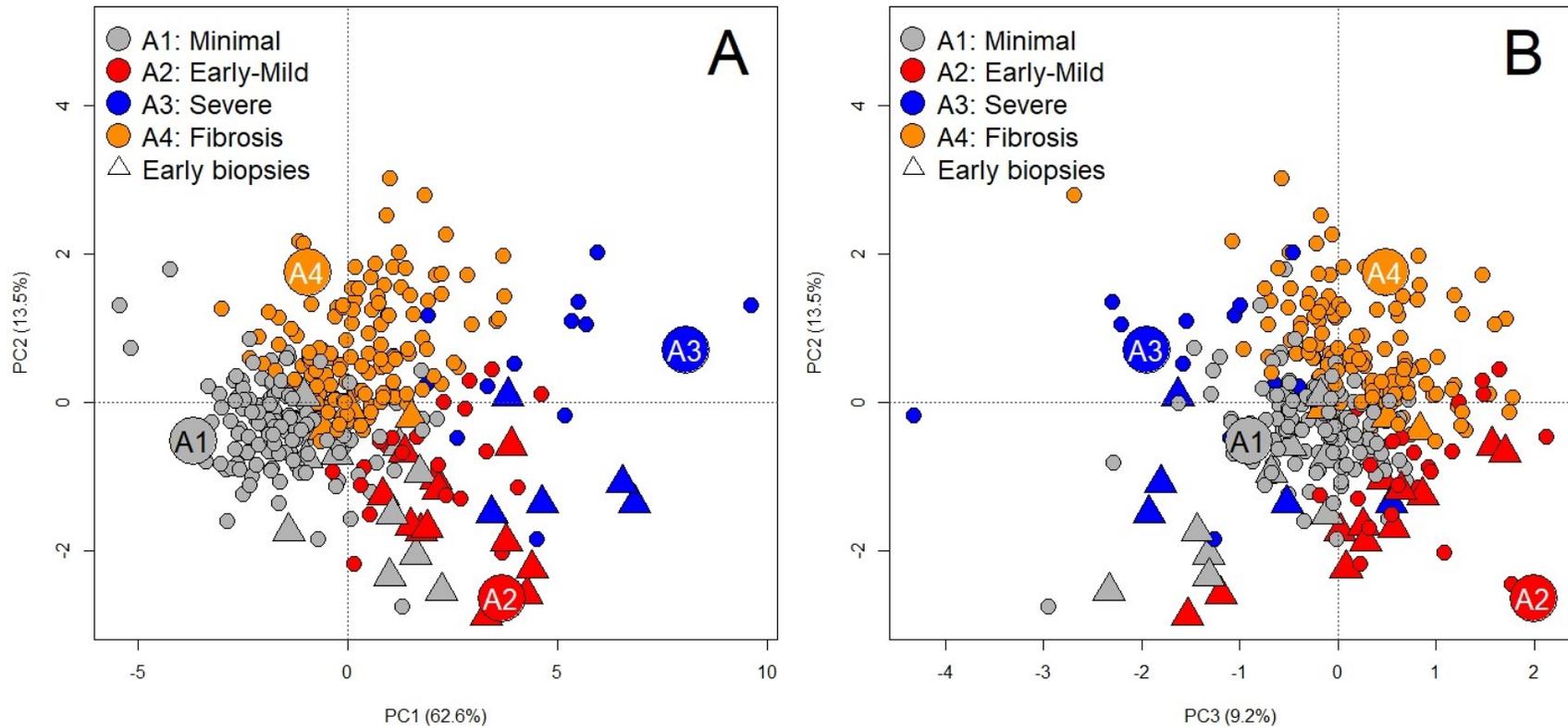
## **Chapter 7: Liver injury**

### **7.12 FIGURES**

**Chapter 7: Liver injury**

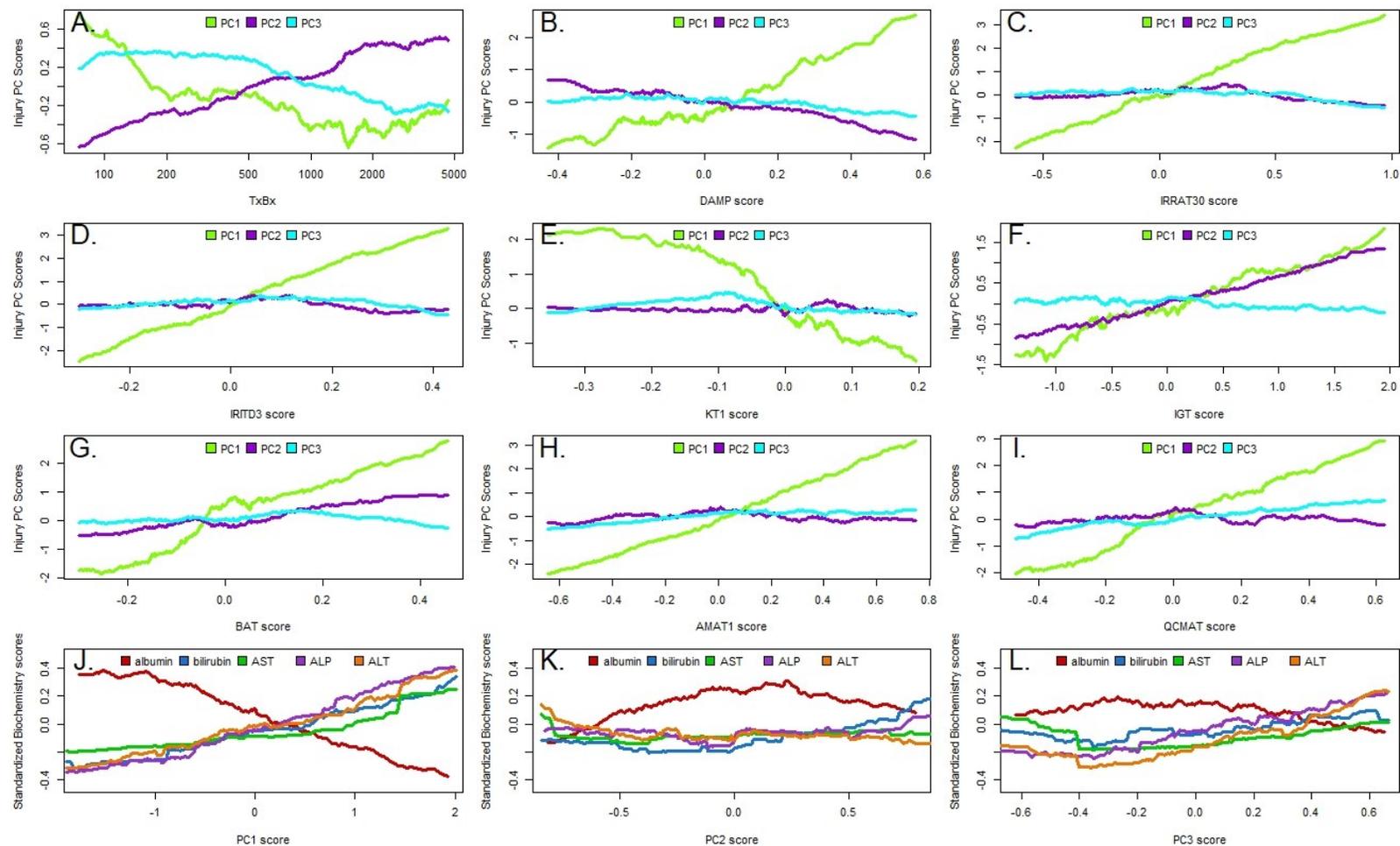


**Figure 7.1** Factor maps showing the pathogenesis-based transcript sets (PBTs) used as initial input for the liver injury analyses and their relationship to principal components 1 (PC1), 2 (PC2), and 3 (PC3). PBTs shown in black were included in the model. Some PBTs or features (e.g. TxBx, KT1s) were not included in the model and are shown as quantitative supplementary variables in blue with dotted lines. PC1 represented injury versus no injury, PC2 represented acute injury versus fibrotic injury (DAMPs versus IGTs), and PC3 represented macrophage transcripts (AMAT1s and QCMATs) versus others.



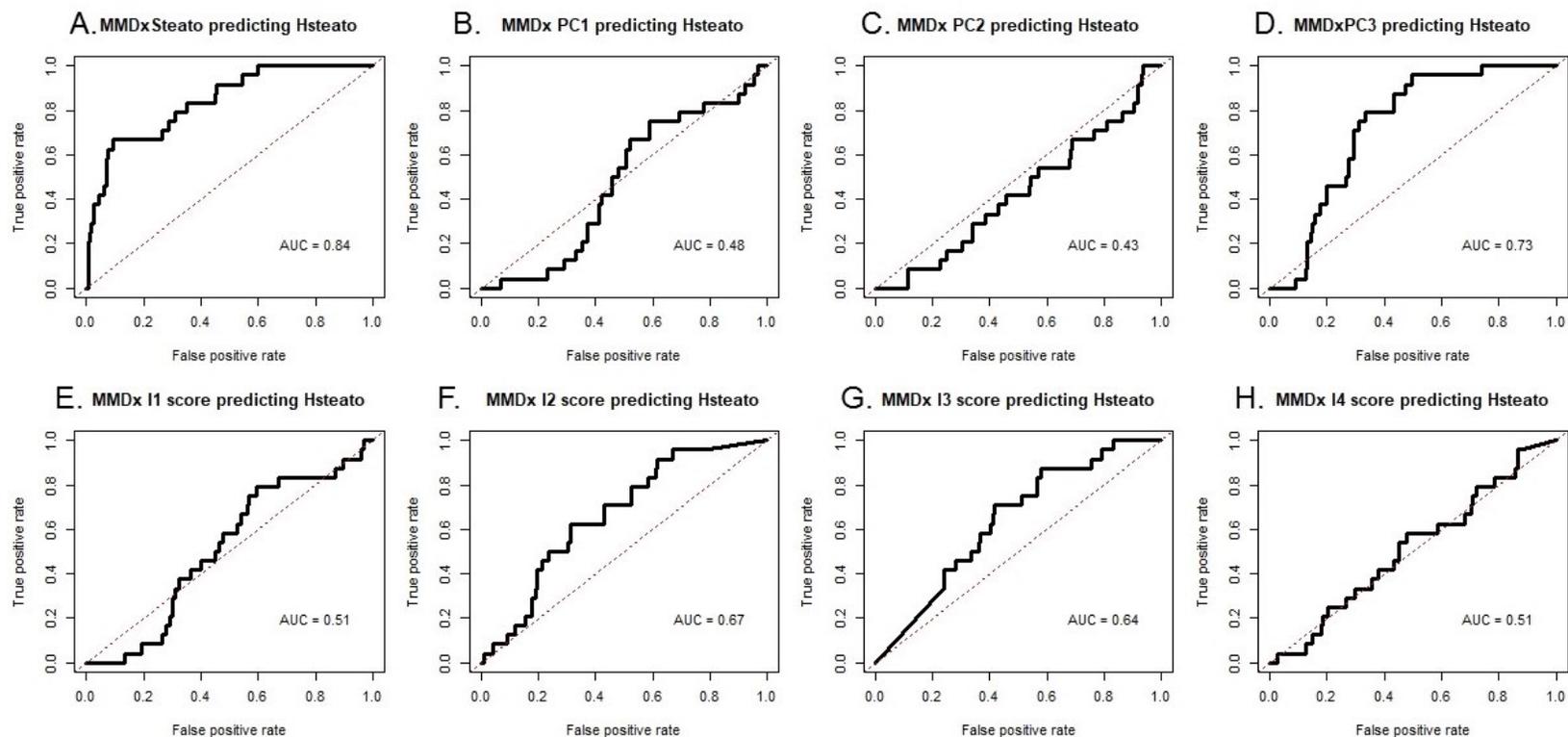
**Figure 7.2 Unsupervised Injury-based analysis of 337 (311 patients) liver transplant biopsies.** Liver biopsies were separated by their injury-associated pathogenesis-based transcript sets (PBTs) scores in PCA. Archetypal analysis identified four major phenotypes, or “archetypes”: minimal injury (A1), early-mild injury (A2), severe injury (A3), and atrophy-fibrosis injury (A4). Each biopsy was given four archetype scores describing their similarity to each archetype. Biopsies are grouped by their highest archetype score. Panel A plots principal component 2 versus principal component 1, and panel B plots principal component 2 versus principal component 3. Triangles represent biopsies taken in the first two weeks post-transplant.

## Chapter 7: Liver injury



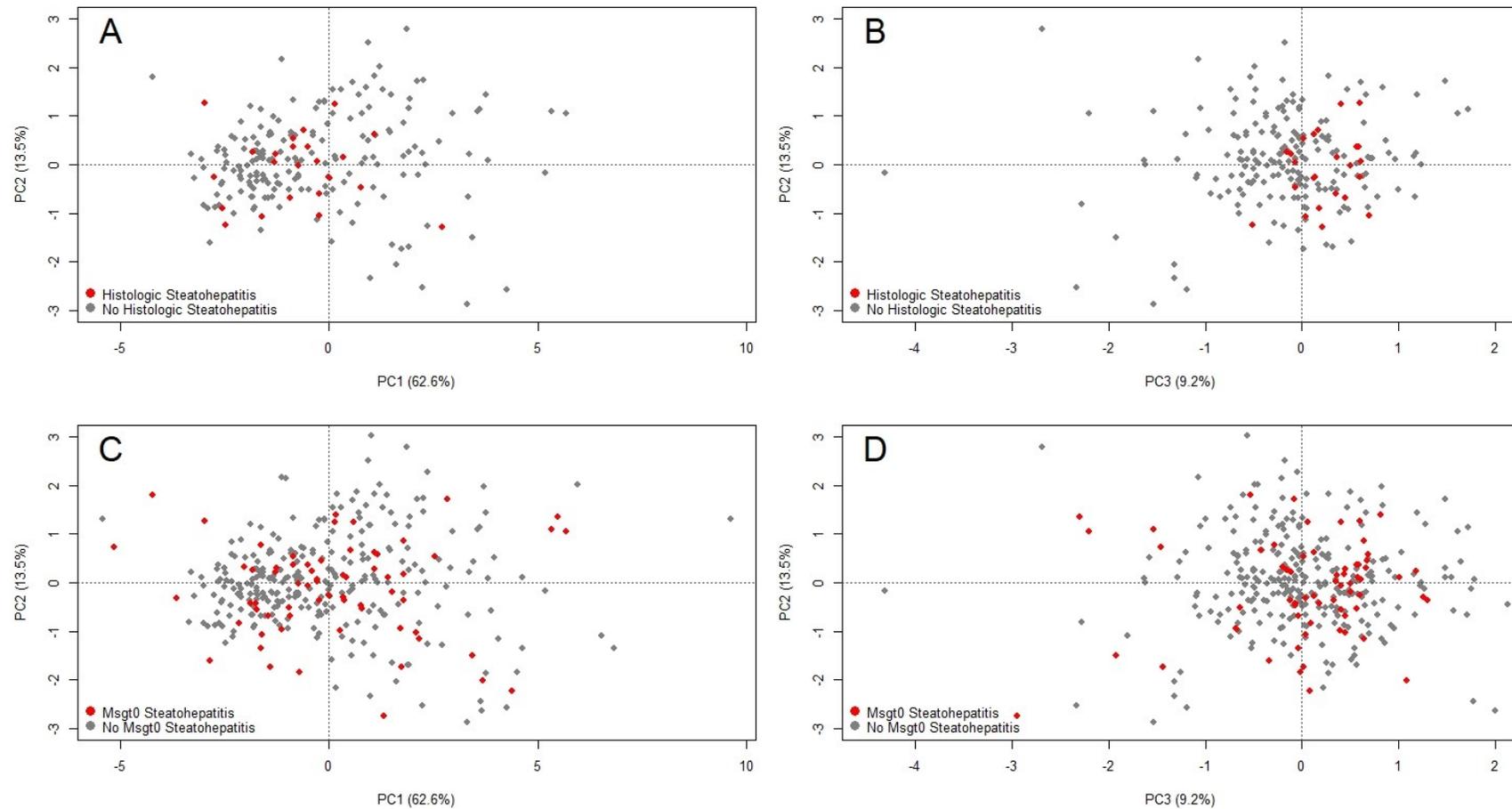
**Figure 7.3 Moving average plots showing relationships between TxBx, PBT scores, Injury PCA scores, and clinical laboratory test values.** Relationships between A) PC scores and TxBx, B) PC scores and DAMP scores, C) PC scores and IRRAT scores, D) PC scores and IRIT scores, E) PC scores and KT1 scores, F) PC scores and IGT scores, G) PC scores and BAT scores, H) PC scores and AMAT1 scores, I) PC scores and QCMAT scores, J) biochemistry scores and PC1 score, K) biochemistry scores and PC2 score, and L) biochemistry scores and PC3 score. All moving averages used sliding window sizes selected based on smoothness: A=100, B-I=50, J-L=125. Biochemistry scores were standardized prior to plotting.

## Chapter 7: Liver injury

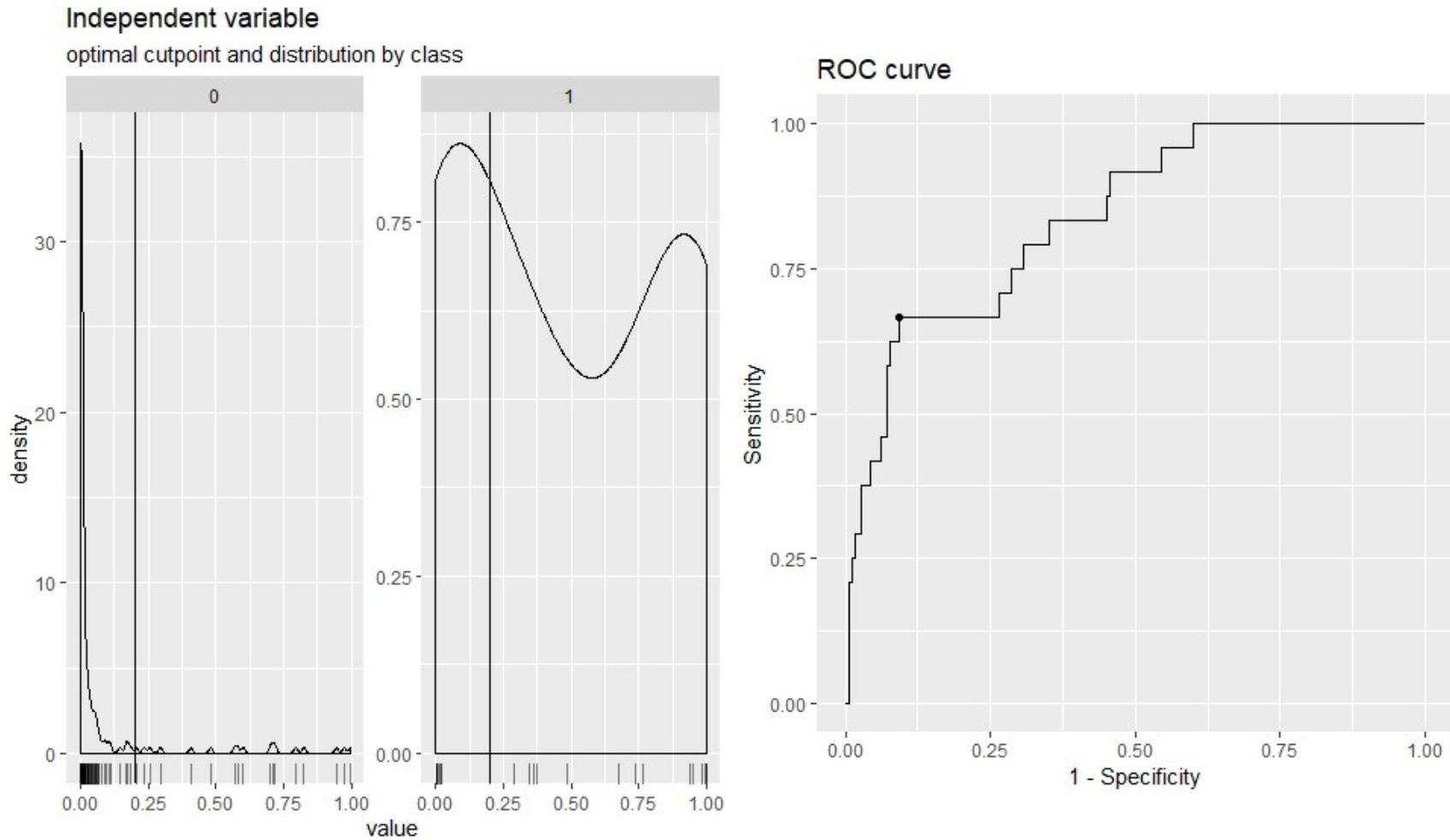


**Figure 7.4 AUCs for various molecular features and scores predicting binarized clinical steatohepatitis.** The following molecular features were assessed for their ability to predict the known clinical labels for steatohepatitis in the given N=337 population. Molecular values were available for all 337 samples, clinical steatohepatitis was positive in 24 biopsies, negative in 182 biopsies, and 131 biopsies were excluded as no steatohepatitis SOC data was available. Note that the I3 score is a negative predictor of steatohepatitis, but by convention the AUC is flipped to be displayed as a positive value.

## Chapter 7: Liver injury



**Figure 7.5 Distribution of biopsies called A-B) steatohepatitis by SOC, and C-D) steatohepatitis by the Msgt0 molecular steatohepatitis Ida-based classifier.** Since PC3 was noted to have a positive correlation with steatohepatitis, biopsies with steatohepatitis were expected to cluster to the right in PC3 but have no particular clustering pattern in PC1 or PC2. Findings in both SOC and molecularly determined steatohepatitis models were consistent with this hypothesis.



**Figure 7.6 Optimal cutoff established using the cutpointR function in R.** The cutoff was optimized based on the classifier raw scores predicting binarized histologic steatohepatitis (Y/N). Optimization balances sensitivity and specificity to find the point with the fewest false positives and false negatives. Above the cutoff was called 'molecular steatohepatitis positive' and below the cutoff was called 'molecular steatohepatitis negative'.

## **SECTION III**

### **DEVELOPING MMDX FOR NATIVE ORGANS – MOLECULAR DIAGNOSTICS IN ULCERATIVE COLITIS**

## **CHAPTER 8**

### **THE MOLECULAR LANDSCAPE OF ULCERATIVE COLITIS DISEASE ACTIVITY**

### 8.1 OVERVIEW AND HYPOTHESIS

UC is a chronic inflammatory disease of uncertain etiology and pathogenesis and varying activity, with suggested roles for both innate immunity and the adaptive T cell autoimmunity. Molecular analyses of UC biopsies have shown increased expression of transcripts and/or proteins for decay-accelerating factor CD55,<sup>158,159</sup> complement related proteins,<sup>158</sup> and calprotectin (a dimer of calcium-binding proteins S100A8 and S100A9 highly expressed in myeloid cells).<sup>160,161</sup> Microarrays or RNA-sequencing have shown many transcripts expressed in UC biopsies compared to controls e.g. screening colonoscopies.<sup>162–169</sup> Differentially expressed genes varied between analyses, probably reflecting details of the comparators<sup>162–165</sup> including other forms of IBD i.e. Crohn's disease or IBDU. Though MMDx was originally developed for diseases in transplanted kidneys, findings and methods were successfully transferred to hearts<sup>261</sup>, lungs<sup>262,299</sup>, and livers (see Section II). With this consideration, it was hypothesized that MMDx could be expanded for diagnostics and disease assessment in native organs as well; namely UC in colonic biopsy tissue. These analyses may lead to a more complete understanding of the disease mechanisms in UC, much as they have in transplantation.

Many lines of evidence implicate a role for cognate T cell-mediated autoimmunity in UC. Mucosal T cells from UC patients exposed to IL-2 respond differently from controls,<sup>300</sup> and epithelial cells from UC patients stimulate CD4<sup>+</sup> T cells differently from normal epithelial cells.<sup>137</sup> Patients treated with T cell checkpoint inhibitors as immunotherapy for cancer often develop a UC-like colitis.<sup>138,139,301</sup> Genetic susceptibility loci for UC map in the HLA region, which controls many cognate cell responses,<sup>140,141</sup> and cytotoxic T cell transcripts for perforin and granzyme A are increased in the intestinal mucosa of UC.<sup>142,143</sup> We previously used microarrays to analyze transcripts in UC biopsies, and showed a large-scale disturbance involving inflammatory cells, parenchymal injury and dedifferentiation with similarities to the transcript sets in TCMR of organ transplants, the prototypic T cell-mediated disease.<sup>64,66,170–172</sup> Expression of these transcript sets derived in rejecting kidney transplants correlated with the endoscopic Mayo score and the presence of lamina propria lymphoplasmacytic infiltrate on histology in colon biopsies,<sup>172</sup> compatible with a relationship between cognate T cell activity and UC. Immunosuppressive drugs that control TCMR in organ transplantation (azathioprine, cyclosporine, and tacrolimus) are effective in some cases of UC, though efficacy is inconsistent.<sup>144–150,152,153</sup> However, the drugs that form the mainstay of UC

## CHAPTER 8: Molecular landscape of ulcerative colitis

management (anti-inflammatory biologics such as infliximab and adalimumab) have not been effective in preventing or treating TCMR in organ transplants, suggesting that there is much more to UC than simply cognate T cell processes. Costimulatory blockade with CTLA4Ig (i.e. abatacept and belatacept) has no demonstrable efficacy in UC,<sup>154,155</sup> despite their success in preventing transplant rejection<sup>155</sup> and treating selected autoimmune diseases.<sup>302</sup> These findings pose a paradox: if cognate T cell autoimmunity is significant in UC, why does the response to therapy differ so much between TCMR and UC?<sup>144–150,152,153</sup>

These analyses explored the relationship of UC activity to cognate effector T cell activation (represented by kidney transplant TCMR) by mapping the Mayo score associated transcripts and comparing them to the transcripts associated with organ transplant TCMR.<sup>36</sup> Kidney transplant TCMR was used as a comparator because TCMR is a sterile process and not exposed to the microbiome, and features and transcript patterns of TCMR (and thus T cell activity) are highly conserved across different organs.<sup>171</sup> The present study explored the molecular changes and pathways associated with UC activity as defined by the endoscopic Mayo score (scores 2/3 versus 0/1). The objective of these analyses was to understand whether episodes of significant UC disease activity could be explained by a cognate T cell-mediated inflammatory process, similar to TCMR episodes in organ transplants. These findings will be helpful in the translation of the MMDx system into UC, with the goals of both new understanding of disease mechanisms and the development of a new test to guide therapy.

### 8.2 STUDY POPULATION AND DEMOGRAPHICS

The study included 71 prospectively collected biopsies from 61 consenting patients at the University of Alberta Hospital (Edmonton, Canada, 53 biopsies) and Cedars-Sinai Hospital (Los Angeles, USA, 18 biopsies) (**Table 8.1**), and 16 control colon biopsies (collected during screening colonoscopies, **Table 8.2**) all characterized by SOC conventional methods.<sup>172</sup> Gene expression in the biopsies was measured using Affymetrix GeneChip human PrimeView arrays.<sup>303</sup> Some of these biopsies were included and described in an earlier publication.<sup>172</sup>

#### 8.2.1 Biopsy and data collection

## CHAPTER 8: Molecular landscape of ulcerative colitis

All biopsies were collected during SOC endoscopy as per existing protocols at Center of Excellence for Gastrointestinal Inflammation and Immunity Research (CEGIIR, University of Alberta Hospital in Edmonton, Canada) and at Cedars-Sinai Hospital (Los Angeles, USA). 16 control samples from 16 patients were collected during screening colonoscopies and were included in some analyses as non-UC comparators. Biopsies were collected segmentally from the ascending, transverse, descending colon and rectum, placed into RNA $\text{later}^{\text{TM}}$ , and stored at  $-20^{\circ}\text{C}$  for isolation of RNA. Biopsies were also collected and given histologic scores by a gastrointestinal pathologist, and blood samples were sent for laboratory testing including CRP, albumin, and routine chemistry. Biopsies from the most inflamed colonic segments as determined by endoscopy were selected for the microarray analysis, since these would coincide with the endoscopic Mayo score assigned at that scope date.

Demographics (age, sex, date of diagnosis, disease duration), medications at the time of biopsy, and endoscopic data (extent of disease, endoscopic Mayo score) after biopsy collection were retrospectively assessed for all patients (**Table 8.1**). Endoscopic Mayo score was chosen to classify biopsies as disease active versus disease inactive for all analyses.

### 8.2.2 Biopsy processing

Biopsies were processed for microarray analysis as described in a prior publication.<sup>36</sup> Purified total RNA was labeled with the IVT Express labeling kit (Affymetrix, Santa Clara, CA) and hybridized to human PrimeView arrays (Affymetrix) according to manufacturer protocols available on ThermoFisher.com.<sup>303</sup> Microarrays were scanned, 'CEL' files were obtained using GeneChip Operating Software (Affymetrix), and robust multiarray averaging was used to normalize the CEL files.<sup>36</sup> IQR filtering was done to reduce the total number of probe sets from 49,495 to 12,359 in UC and to 12,354 in TCMR, removing non-informative, low variance probe sets. The kidney transplant biopsies, demographics, and CEL files were previously described.<sup>170</sup> CEL files are available on the Gene Expression Omnibus website (accession number GSE107202).

IRB approval was obtained through CEGIIR (consent obtained for collection and storage of samples, PRO:00001994). All patients were over the age of 18 and consented at the time of tissue collection, and all patient data were de-identified prior to analyses.

### 8.2.3 Data processing

Statistical analysis and graphics were done in the “R” software package, version 3.3.2<sup>185</sup> with various libraries from Bioconductor 3.2,<sup>203</sup> and in Microsoft Excel version 16 (Redmond, WA), and probe set lists were IQR filtered. Significance of probe set expression is given as unadjusted p-values (Bayesian t-test), except in cases where the FDR is specified.

### 8.3 TOP TRANSCRIPTS EXPRESSED IN UC BIOPSIES

Transcripts with high (2/3) versus low (0/1) endoscopic Mayo scores in all UC biopsies were examined (N=71). Variance filtering yielded 12,359 IQR-filtered transcripts and these were ranked by their association (p-value) with the Mayo score. The top 30 non-redundant transcripts increased in biopsies with high Mayo score are shown in Table 8.3, and those decreased in expression in biopsies with high Mayo score are shown in Table 8.4. Top transcripts increased with higher Mayo scores were mainly annotated for their expression in parenchymal tissue and in myeloid cells, relationships to inflammation, and inducibility by IFNG. Top transcripts decreased with higher Mayo scores were all expressed primarily in parenchymal cells and associated with loss of functionality in tissue or tissue dedifferentiation (e.g. *SLC36A1*).

#### 8.3.1 Volcano plot of transcript expression in UC

**Figure 7.1** is a volcano plot showing association strength versus fold change for each of the 12,359 transcripts remaining after IQR variance filtering of the UC disease activity-associated transcripts. Transcripts of special interest are annotated in the legend and appropriately colored in the figure. The top transcript associated with UC activity was complement factor B (*CFB*). Complement inhibitor decay-accelerating factor *CD55* was another feature strongly associated with increased UC disease activity: all six probe sets were highly associated ( $p=1.8 \times 10^{-8}$ - $5.3 \times 10^{-7}$ ). Calprotectin transcripts *S100A8* and *S100A9* were also strongly associated ( $p=7.2 \times 10^{-7}$  and  $5.3 \times 10^{-6}$ ). *IFNG* was moderately associated with UC activity ( $p=5.3 \times 10^{-4}$ ), and many IFNG-inducible transcripts such as apolipoproteins (*APOL1* and *APOL2*) and chemokines (*CXCL9*, *10*, and *11*) were highly induced and had moderate to strong associations with UC

## CHAPTER 8: Molecular landscape of ulcerative colitis

activity. Tumor necrosis factor (*TNF*) alpha was weakly associated ( $p=3.1 \times 10^{-2}$ ). Solute carrier transcripts representing parenchymal function were decreased in UC.<sup>172</sup>

### 8.3.2 Inflammasome and Toll-like receptor transcripts

Inflammasome transcripts and TLRs showed variable association with UC activity ( $p=7.2 \times 10^{-7}$ - $9.8 \times 10^{-1}$ ). The strongest UC activity-associated inflammasome transcript was caspase 1 (*CASP1*,  $p=7.2 \times 10^{-7}$ ), which is also IFNG inducible.<sup>304</sup> The TLR with the strongest association with UC activity was *TLR5* ( $p=5.1 \times 10^{-5}$ ).

### 8.3.3 Transcripts associated with the NLRP3 inflammasome

Because the *NLRP3* inflammasome had been recently implicated in UC pathogenesis in a murine model,<sup>305</sup> we specifically assessed the *NLRP3* transcripts for their association with UC activity. *NLRP3* probe sets showed little variance in these biopsies and did not pass the IQR filter. In the non-IQR-filtered data (not shown), *NLRP3* transcripts were not associated with Mayo score ( $p=5.8 \times 10^{-2}$  and  $1.7 \times 10^{-1}$ , p-value rank 18210 and 25959, respectively).

### 8.3.4 Transcripts associated with epithelial stem cell memory damage

It has been reported that epithelial stem cells (EpSCs) can be reprogrammed by injury and confer potential epithelial memory through chromatin changes.<sup>306,307</sup> Three transcripts in the UC activity-associated IQR-filtered list have reported associations with chromatin domains that are open in reprogrammed EpSCs: interferon-inducible protein 2 (*AIM2*; two probe sets,  $p=8.1 \times 10^{-5}$  and  $p=8.3 \times 10^{-5}$ ); interleukin 1 beta (*IL-1 $\beta$* ;  $p=7.4 \times 10^{-5}$ ); and *CASP1* (seven probe sets,  $p=10^{-7}$ - $10^{-5}$ ).

## 8.4 EXPRESSION OF UC DISEASE ACTIVITY TRANSCRIPTS IN CULTURED HUMAN CELL LINES

Expression of the top UC activity-associated transcripts (activity represented by the endoscopic Mayo score 2/3 versus 0/1) in the biopsies was studied in a previously published panel of cultured effector CD8<sup>+</sup> and CD4<sup>+</sup> T cells, NK cells, B cells, macrophages, monocytes, dendritic cells (DCs), human umbilical vein endothelial cells (HUVECs), and renal proximal tubule epithelial cells (RPTECs, (**Figure 7.2**)).<sup>170</sup>

## CHAPTER 8: Molecular landscape of ulcerative colitis

Technical details of cell isolation, purification, and culture methods are published.<sup>36,170</sup> Cultured cells were analyzed using HG-U133 plus 2.0 arrays, thus gene symbols of the top transcripts from the UC analyses (PrimeView arrays) were translated to HG-U133 plus 2.0 array probe set IDs using the NetAffx™ Batch Query function (<http://www.affymetrix.com>). Heatmaps and raw expression data of the cell panel were used to empirically assign a principal cellular origin for that transcript, supplemented by the published literature. A transcript was defined as IFNG-inducible if at least one type of cell in the cell lines demonstrated a minimum increase of 100% in expression between unstimulated and stimulated cells.

UC activity-associated transcripts were most highly expressed in myeloid inflammatory cells (monocytes, macrophages and/or DCs, e.g. *CD55*, *CXCL3*), IFNG-treated macrophages (e.g. *CFB*, *CHI3L*), and in IFNG-treated RPTECs and/or HUVECs (e.g. *SAA1*, *SAA2*, *RTEL*, *APOL1*, *APOL2*, *CXCL1*). All top 30 UC activity transcripts were poorly expressed in lymphocytes (T cells, B cells, and NK cells, Figure 7.2), and 16 of 30, including *CD55*, showed high expression (>100) in normal colon (Table 8.3, screening colon samples).

**Table 8.4** summarizes the cell panel expression of the Mayo-associated transcripts decreased in biopsies with a high endoscopic Mayo score. All transcripts decreased with increasing Mayo score were primarily expressed in parenchymal cells.

### 8.5 EXPRESSION OF UC DISEASE ACTIVITY-ASSOCIATED TRANSCRIPTS IN TCMR BIOPSIES

To contrast transcripts highly associated with UC disease activity against transcripts highly associated with a sterile cognate T cell process like kidney transplant TCMR, we compared rankings in both transcript lists. The top 30 UC activity-associated transcripts were not highly ranked by p-value in the TCMR top transcript list<sup>278</sup> (**Table 8.5** summarizes the top 30 UC activity-associated transcripts as they would rank in TCMR). Some top UC activity-associated transcripts did not pass the IQR filter in the T cell process-associated analysis because they lack expression in normal kidney tissue or kidney tissue with TCMR (“NA” in **Table 8.5**). While UC and TCMR share some transcript expression patterns, top transcripts in UC and TCMR were not the same, confirming major mechanistic differences between these two processes.

## **8.6 EXPRESSION OF T CELL-SPECIFIC AND T CELL PROCESS-ASSOCIATED TRANSCRIPTS IN UC**

The expression and representation of selected T cell-specific and TCMR-associated transcripts was assessed in the UC biopsy population.

### **8.6.1 T cell-specific transcripts and UC disease activity.**

Representative T cell-specific transcripts (empirically derived from those available on the PrimeView arrays) *CD3D*, *CD2*, and T cell receptors alpha (*TRAC*), beta (*TRBC*), delta (*TRDC*), and gamma (*TRGC*) in UC were studied for their expression in UC-active biopsies (Mayo 2/3), UC-inactive biopsies (Mayo 0/1), and screening colon biopsies. All selected TCMR-associated/T cell-specific transcripts are primarily expressed in effector T cells (with the exception of *TRDC*, which was expressed in both NK cells and T cells) *CD3D*, *CD2*, *TRAC*, and *TRBC* were increased with UC activity, while *TRGC* and *TRDC* were not (**Table 8.6**). In summary, expression of TCMR transcripts increased with UC activity, though they were not ranked by p-value near the top of UC activity-associated transcript lists, in agreement with previous work.<sup>172</sup>

### **8.6.2 TCMR-associated transcripts and UC disease activity.**

The top 30 TCMR transcripts<sup>278</sup> consisted of T cell activation transcripts, macrophage activation transcripts, and IFNG-inducible transcripts (**Table 8.7**). These transcripts are not organ specific but are expressed both in kidney and heart tissue with transplant TCMR and are related to a sterile cognate T cell process rather than tissue-related.<sup>171,277</sup> Of the top 30 genes in kidney TCMR, 20 of these transcripts were associated with UC (p-value <0.05), including transcripts typical of activated effector T cells - *CTLA4* (p=8.7x10<sup>-6</sup>), *ICOS* (p=2.6x10<sup>-4</sup>), *IFNG* (p=5.3x10<sup>-4</sup>), and *LAG3* (p=2.2x10<sup>-4</sup>) - and transcripts associated with activated macrophages e.g. *CD84* (p=4.5x10<sup>-2</sup>) (complete IQR-filtered list not shown).

*IFNG*-inducible T cell process transcripts<sup>170,278</sup> were moderately to strongly associated with UC activity: e.g. *APOL1* and *APOL2* (p=4.2x10<sup>-8</sup>-1.6x10<sup>-7</sup>, respectively), guanylate binding protein (*GBP5*, p=2.1x10<sup>-5</sup>) and chemokines *CXCL9*, *CXCL10*, and *CXCL11* (*CXCL9* p=7.5x10<sup>-6</sup>, *CXCL10* p=3.0x10<sup>-4</sup>,

## CHAPTER 8: Molecular landscape of ulcerative colitis

and *CXCL11* (three probe sets)  $p=2.5 \times 10^{-5}$ - $1.0 \times 10^{-4}$ ). T cell cytotoxic molecules were also associated: *GZMB*, *PRF1*, and *LAG3*.

Some TCMR-associated transcripts did not make the IQR filter in UC. These transcripts had little or no expression in UC or normal colon (*CD8B*, *S1PR4*). The lack of expression of CD8B in colon suggests that the effector T cells in UC are likely expressing the CD8A homodimer and not the beta chain (raw expression values of *CD8A* in UC are 302 in Mayo 2/3, 264 in Mayo 0/1, and 207 in screening colon biopsies).

Given these results, T cell-specific transcripts and TCMR-associated transcripts are increased significantly in UC disease, but are not the transcripts most prominently associated with UC activity.

### 8.7 OVERLAP BETWEEN TOP 300 UC ACTIVITY-ASSOCIATED AND TOP 300 T CELL PROCESS-ASSOCIATED TRANSCRIPTS

Of the top 300 UC activity-associated (data not shown) and top 300 TCMR-associated (data not shown), only 22 transcripts were shared, and 17/22 were IFNG-inducible. Only one transcript (*APOL2*) was shared between the top 30 lists. Most overlapping transcripts between UC activity and TCMR were IFNG-inducible either in macrophages or in parenchymal cells. A few transcripts were expressed in monocytes and macrophages (*FPR1*), NK cells (*FCGR3A/B*), or T cells (*CTLA4*, *GZMB*).

### 8.8 OVERREPRESENTATION ANALYSIS: UC ACTIVITY-ASSOCIATED VERSUS TCMR-ASSOCIATED TRANSCRIPTS

Overrepresentation of UC activity-associated transcripts in GO terms (Biological Process, Cellular Compartment, and Molecular Function) and in KEGG pathways was analyzed using the DAVID tool.<sup>308</sup> A similar analysis was done on the top transcripts increased in a comparison of TCMR versus all other diagnoses in human kidney transplants.<sup>278</sup> Top GO terms in both analyses were compared to one another using GO term names.

Top GO terms and KEGG pathways ( $p$ -value  $< 0.05$ ) from DAVID analysis of the UC activity associated transcripts are summarized in **Tables 8.8-8.11**. Top 10 GO terms enriched between the 300 UC activity-associated transcripts and the top 300 TCMR-associated transcripts<sup>170,278</sup> for each grouping

## CHAPTER 8: Molecular landscape of ulcerative colitis

were compared (**Table 8.12**). Of 40 possible pathway overlaps, only seven were shared. These did not reflect molecular sharing: most of the transcripts in overlapping GO terms and KEGG pathways differed between UC and TCMR, as only 17 of 206 unique transcripts overlapped, 13 of which were IFNG-inducible (data not shown). Aside from the IFNG effects, there was little evidence of activation of a significant cognate T cell-mediated process in UC.

### 8.9 INTERPRETATION OF RESULTS

While our findings are consistent with a cognate T cell process playing a role in UC, disease activity cannot be explained by this process alone. Our study analyzed molecular expression patterns in UC colonic biopsies to produce a novel definition of the molecular landscape of UC specifically related to disease activity represented by the endoscopic Mayo score, and how this activity compares to a sterile cognate T cell-mediated process like transplant TCMR. The top 30 UC activity transcripts were involved in complement regulation (e.g. *CFB*, *CD55*) and macrophages (e.g. *S100A8* and *S100A9*). These transcripts were highly expressed in epithelial cells (RPTECs), endothelial cells (HUVECs), and in activated macrophages (e.g. *CHI3L1*), but not T cells. Transcripts reflecting a cognate T cell process were associated with UC disease activity but were never prominent in a class comparison of high versus low Mayo score, and there was little overlap in pathways overrepresented by the top UC activity-associated and the TCMR-associated transcripts suggesting that effector (or effector-memory) T cells although present are not heavily involved disease activity. UC activity was primarily characterized by expression changes in parenchymal cells and myeloid cells, whereas TCMR was dominated by effector T cell activation, although both processes expressed IFNG-inducible transcripts.

Transplant TCMR acts as a non-organ-specific model for a cognate (i.e. antigen-specific) human T cell process. Kidney and heart TCMR are very similar to one another, demonstrating that molecular features are highly conserved among different tissues with active T cell-mediated diseases.<sup>171</sup> Both kidney and heart TCMR are characterized by the prominence of cognate effector T cells interacting with APCs (through the immunological synapse) in a sterile tissue compartment, and by the expression of effector T cell transcripts (e.g. *CTLA4*, *ICOS*, *ADAMDEC1*, *IFNG*), IFNG-inducible transcripts (e.g. *CXCL9/10/11*), and antigen-presenting cell transcripts (e.g. *CD80/86*, *CXCL13*, *SLAMF8*).<sup>170,171</sup> UC displayed associations

## CHAPTER 8: Molecular landscape of ulcerative colitis

with some of these transcripts typical of a cognate T cell-mediated process, such as *CTLA4*,<sup>309</sup> and many of these T cell process-associated and T cell-specific transcripts showed increased expression in high Mayo (2/3) score biopsies when compared to Mayo 0/1 or screening colon. However, most of the top transcripts associated with disease activity in UC were expressed in activated myeloid cells and stressed epithelial cells that are not prominent in TCMR.<sup>170</sup> In addition, the source of IFNG in UC is not necessarily effector T cells, since NK cells are also a potent producer of IFNG. While lymphoplasmacytic infiltrate in UC biopsies correlates with the molecular disturbance<sup>172</sup> and has been shown to be a predictor of relapse in patients,<sup>310</sup> the present findings indicate that cognate effector T cell activity alone is not sufficient to explain disease activity. These findings suggest that there are multiple immunologic driving processes occurring in UC.

The pathogenesis of UC is going to require some new concepts such as stem cell reprogramming. If UC represents both a cognate T cell process and an independent inflammatory process, this would require links between these processes, and recent experimental findings offer a possible model. Injury to the epithelium can be remembered through reprogramming of EpSCs (through changes in chromatin accessibility patterns in genomic regions associated with inflammation), rendering the epithelium hypersusceptible to otherwise innocuous inflammatory stimuli.<sup>306,307</sup> Of interest, certain transcripts correlating with UC activity are among those encoded by genes in open chromatin domains induced by damage in EpSCs (e.g. *AIM2*, *IL-1 $\beta$* , and *CASP1*).<sup>306,307</sup> This finding would be compatible with the association between UC activity and inflammasome transcripts, including strong associations of *CASP1*, *CASP10*, *CASP6* and *IL18* with UC activity.<sup>311</sup> It is possible that a cognate T cell-mediated autoimmunity may facilitate UC activity<sup>142,312-314</sup> via alterations in EpSCs, which then retain memory of this damage and program the epithelium to be vulnerable to typically inoffensive local influences such as the microbes of the lumen and their products, triggering an inflammatory response-to-wounding and wound healing.<sup>195</sup> In this model, a wave of cognate T cell-mediated autoimmunity creates a long term memory through chromatin changes in mucosal EpSCs that persists long after the cognate process has abated or been successfully treated.<sup>306,307,315</sup> Future work includes analyses to separate these processes (inflammation and T cell effects as individual mechanisms), examining evidence for the degrees of these processes present in a varied UC population.

## CHAPTER 8: Molecular landscape of ulcerative colitis

A binary model for UC would have implications for clinical management in that optimizing therapy for a person in relapse may have to define the relative role of each element: the cognate T cell autoimmune process and the self-sustaining inflammation-injury process. Diverse mechanisms of relapse and resistance to therapy could explain some of the limitations and inconsistent efficacy of empirically-derived therapies for UC, and why such therapies differ strikingly in their efficacy from a process like organ transplant TCMR. The difference in efficacy between biologic monotherapy and combination therapy (biologic and thiopurine or methotrexate) is generally held to be on the basis of decreased immunogenicity towards the biologics<sup>316</sup>, but our data raise the possibility that the suppression of a cognate process in addition to the inflammatory suppression offered by the biologic therapy may be required for some patients to establish prolonged, deep remission. Treatments may be more effective if we treat the processes that drive the disease activity in each individual case, and incorporate anti-inflammatory and immunosuppressive therapies accordingly, rather than adopting a pure monotherapy versus combination therapy approach.

The strong association of the endoscopic Mayo score with *CD55*, *CFB*, and other complement transcripts invites us to consider the possibility of a central local role for complement proteins, possibly activated by the microbial environment, either in provoking inflammation or regulating the inflammatory environment. *CD55* is a membrane-bound regulatory molecule for the complement system which could be playing a protective role in injured epithelium. It is possible that *CD55* is most associated with clinical activity (Mayo 2/3) because it is the epithelium attempting to control complement-related damage.<sup>158</sup> *CD55* is expressed in monocytes, macrophages, DCs, and endothelium, but it has prominent expression on the luminal surface of intestinal epithelial cells. *CD55* deficiency leads to protein-losing enteropathy, which responds to inhibition of complement component 5 (eculizumab).<sup>159,317</sup> This observation argues for roles for both complement interaction with the gut microbiota and in normal homeostasis of the gut epithelium.

Limitations to our study include a limited understanding of the expression patterns present in normal colon tissue and a pure observational approach. Another limitation in this analysis is the nature of TCMR itself, as a process in a sterile organ (while UC is not). Future directions include seeking additional evidence to show the role of the molecules analyzed, specifically for cognate T cell recognition separate from the inflammatory process. Being able to further delineate the different inflammatory signals within individual

## CHAPTER 8: Molecular landscape of ulcerative colitis

patients could potentially lead to better outcomes by tailoring of the treatment plan to address all active processes while avoiding ineffective and potentially hazardous overtreatment.

The selected platform for transcript analysis in this case was microarrays, rather than RNA sequencing, because microarrays are capable of generating large amounts of transcript data in a short time, can be done in real time with prospectively collected biopsies, were compatible with our microarray-based reference set of kidney and heart biopsies (~2600 biopsies), were cost-effective, and provided all necessary information for the class comparison. Above all, they can be standardized to allow comparisons with future studies, unlike RNA sequencing which is difficult to standardize in part due to a dependency on library preparation.

While these studies support a role for cognate  $\alpha\beta$  (not  $\gamma\delta$ ) T cell recognition in UC, the molecular phenotype associated with active UC indicates that this mechanism is not solely responsible for the pathology seen in UC, and that a second main inflammatory stimulus is acting that determines activity. While we expected to find predominant T cell activity, empirical analyses indicate a dominant inflammation and inflammasome activation with primary expression of the top UC activity transcripts in epithelial and myeloid cell lines. Among the possibilities for the linkage between these two processes, the recent demonstration of EpSC reprogramming raises the possibility that a cognate T cell-mediated autoimmune process may induce long term reprogramming of the EpSC and make the epithelium hypersensitive to a subsequent inflammatory process, which may create the conditions and symptoms seen in UC. In future iterations of MMDx for UC, it may be important to assess the disease by the inflammatory component, and the presence of ongoing cognate T cell recognition separately, as this may affect the diagnosis and appropriate therapy.

## **CHAPTER 8: Molecular landscape of ulcerative colitis**

### **8.10 TABLES**

## CHAPTER 8: Molecular landscape of ulcerative colitis

**Table 8.1** Demographics of the patients and the biopsies for the molecular landscape of UC (N=71 biopsies from 61 patients)

<u>Patient characteristics</u>	
Mean patient age (years)	39.0 (range 19-66, 18 NA)
Patient Gender (% male)	47% male (18 NA)
<u>Diagnosis</u>	
UC	71
<u>Centers</u>	
University of Alberta Hospital - Edmonton	53
Cedars-Sinai Hospital – Los Angeles	18
<u>Biopsy characteristics</u>	
Median time from index to last known follow-up in months (mean)	68.6 (60.95, 28 NA)
Estimated mean disease duration in months	124.3 (21 NA)
<u>Endoscopic Mayo Scores of the biopsy</u>	
0	16
1	20
2	21
3	14
<u>Treatment regimen at the time of the biopsy<sup>a</sup></u>	
5ASA	33 (18 NA)
Imuran	19 (18 NA)
Biologic	15 (18 NA)
Prednisone	8 (18 NA)
Cortifoam	5 (18 NA) <sup>b</sup>

5ASA, 5-aminosalicylic acid; NA, not available; UC, ulcerative colitis.

<sup>a</sup> Many patients were on a combination of these therapies, with 1-4 different therapies administered at the same time. Samples from Cedars-Sinai had unavailable treatment information, accounting for the 18 NA in each category.

<sup>b</sup> Cortifoam was not applicable in all cases due to the topical nature of the treatment.

## CHAPTER 8: Molecular landscape of ulcerative colitis

**Table 8.2** Demographics of the patients and biopsies for the screening colon biopsy cohort (N=16 biopsies from 16 patients)

<b><u>Patient characteristics</u></b>	
Mean patient age (years)	<b>58 (5 NA)</b>
Patient Gender (% male)	<b>45% (5 NA)</b>
<b>Diagnosis</b>	
No UC	<b>16</b>
<b>Centers</b>	
<i>University of Alberta Hospital - Edmonton</i>	<b>11</b>
<i>Cedars-Sinai Hospital – Los Angeles</i>	<b>5</b>
<b><u>Biopsy characteristics</u></b>	
Median time from index to last known follow-up in months (mean)	<b>68.21</b>
<b>Colonoscopy Indication</b>	
<i>Family history of colon cancer</i>	<b>4</b>
<i>History of other cancers</i>	<b>3</b>
<i>History of bleeding</i>	<b>2</b>
<i>Other</i>	<b>2</b>
<i>Not available</i>	<b>5</b>
<b>Colonoscopy findings</b>	
<i>Polyp removal</i>	<b>5</b>
<i>No abnormal histology</i>	<b>5</b>
<i>Other</i>	<b>1</b>
<i>Not available</i>	<b>5</b>

NA, not available; UC, ulcerative colitis.

**CHAPTER 8: Molecular landscape of ulcerative colitis**

**Table 8.3** Top 30 transcripts increased in UC biopsies with Endoscopic Mayo Score >1 versus ≤1 (IQR filtered list), aligned by p-value

Probe Set ID	Title	Gene	Fold change	p-value	Adjusted p-value	Expression in biopsies			Expression in cell lines
						Mayo 2/3	Mayo 0/1	Screening colon	Principal Expression <sup>a</sup>
11733725_a_at	complement factor B	CFB	2.6	3.5x10 <sup>-9</sup>	3.4x10 <sup>-5</sup>	1160	443	275	MMDC**/weak IFNG
11723303_at	PDZK1 interacting protein 1	PDZK1IP1	2.6	8.1x10 <sup>-9</sup>	3.4x10 <sup>-5</sup>	1444	549	475	Parenchymal/RPTECs
11756316_a_at	chitinase 3 like 1	CHI3L1	11.5	8.2x10 <sup>-9</sup>	3.4x10 <sup>-5</sup>	406	35	16	MMDC
11730214_at	dual oxidase 2	DUOX2	8.7	1.3x10 <sup>-8</sup>	3.7x10 <sup>-5</sup>	1058	121	79	Parenchymal
11755084_x_at	CD55 molecule (Cromer blood group)	CD55	3.0	1.8x10 <sup>-8</sup>	3.7x10 <sup>-5</sup>	2464	827	535	Parenchymal/ MMDC
11725000_at	dual oxidase maturation factor 2	DUOXA2	7.3	3.6x10 <sup>-8</sup>	3.7x10 <sup>-5</sup>	845	116	60	Parenchymal
11718564_at	gap junction protein alpha 4	GJA4	1.6	3.8x10 <sup>-8</sup>	3.7x10 <sup>-5</sup>	53	32	31	Parenchymal
11737791_s_at	apolipoprotein L1 / apolipoprotein L2	APOL1 / APOL2	2.4	4.2x10 <sup>-8</sup>	3.7x10 <sup>-5</sup>	797	330	211	IFNG inducible (HUVECs)
11724541_s_at	von Willebrand factor	VWF	2.4	4.3x10 <sup>-8</sup>	3.7x10 <sup>-5</sup>	1093	459	357	Endothelial (HUVECs)
11742449_a_at	apolipoprotein L1	APOL1	1.8	4.5x10 <sup>-8</sup>	3.7x10 <sup>-5</sup>	386	212	155	IFNG in HUVECs, RPTECs; many others
11719366_s_at	C-X-C motif chemokine ligand 1	CXCL1	6.5	5.0x10 <sup>-8</sup>	3.7x10 <sup>-5</sup>	1947	298	182	IFNG inducible
11716639_a_at	collagen type IV alpha 1 chain	COL4A1	2.5	5.2x10 <sup>-8</sup>	3.7x10 <sup>-5</sup>	460	183	135	Parenchymal/HUVECs/RPTECs
11732566_at	solute carrier family 6 member 14	SLC6A14	8.6	7.3x10 <sup>-8</sup>	4.5x10 <sup>-5</sup>	305	36	18	Parenchymal
11763250_x_at	C-X-C motif chemokine ligand 1 / ligand 2	CXCL1 / 2	3.6	8.5x10 <sup>-8</sup>	5.0x10 <sup>-5</sup>	406	113	89	IFNG inducible, macrophages, HUVECs
11742862_a_at	adhesion G protein-coupled receptor F5	ADGRF5	2.1	9.6x10 <sup>-8</sup>	5.0x10 <sup>-5</sup>	245	115	91	Parenchymal/HUVECs
11756072_s_at	serum amyloid A1 / serum amyloid A2 / serum amyloid A2-A4	SAA1 / SAA2 / SAA2-SAA4	9.5	1.0x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	387	41	19	IFNG inducible
11744718_a_at	collagen triple helix repeat containing 1	CTHRC1	3.9	1.3x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	155	40	30	Parenchymal
11720440_at	olfactomedin like 2B	OLFML2B	2.2	1.4x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	142	65	59	Macrophages
11718832_a_at	lysyl oxidase like 2	LOXL2	1.9	1.5x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	189	98	65	HUVECs and RPTECs
11720020_x_at	regulator of telomere elongation helicase 1	RTEL1	2.4	1.6x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	285	120	80	Ambiguous
11752035_a_at	nitric oxide synthase 2	NOS2	3.5	1.7x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	183	53	45	IFNG inducible
11716281_a_at	collagen type XVIII alpha 1 chain	COL18A1	2.1	1.8x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	274	133	118	RPTECs/HUVECs/weak IFNG inducible
11739315_at	kinase insert domain receptor	KDR	1.5	1.8x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	197	130	119	HUVECs/Parenchymal
11739134_a_at	collagen type V alpha 2 chain	COL5A2	2.2	1.8x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	453	206	185	HUVECs
11756068_a_at	beta-site APP-cleaving enzyme 2	BACE2	1.8	2.0x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	1583	890	675	HUVECs/RPTECs/Parenchymal
11742889_at	F-box protein 6	FBXO6	1.6	2.0x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	219	138	120	MMDC, ubiquitous
11726124_at	transcobalamin I	TCN1	3.8	2.0x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	119	32	23	RPTECs/monocytes
11715239_x_at	interferon induced transmembrane protein 3	IFITM3	1.8	2.4x10 <sup>-7</sup>	5.2x10 <sup>-5</sup>	2506	1417	1063	Ubiquitous/weak IFNG
11736477_a_at	elongation factor for RNA polymerase II 2	ELL2	2.1	2.4x10 <sup>-7</sup>	5.2x10 <sup>-5</sup>	475	226	203	MMDC, ubiquitous
11728477_at	C-X-C motif chemokine ligand 3	CXCL3	4.9	2.6x10 <sup>-7</sup>	5.2x10 <sup>-5</sup>	775	158	119	IFNG inducible
<b>Column Means<sup>b</sup></b>						<b>455</b>	<b>148</b>	<b>107</b>	

IFNG, interferon gamma; HUVECs, human umbilical vein endothelial cells; MMDC, monocytes, macrophages, and dendritic cells; RPTECs, renal proximal tubule epithelial cells; UC, ulcerative colitis.

NOTE: Screening colon samples were not used to generate the top genes, but we show the expression of the top genes from UC biopsies in screening colon samples as a representation of healthy colon tissue.

<sup>a</sup> Expression values had to be over 100 in the cell lines to be considered for principal expression. Parenchymal expression was determined either by expression in cultured parenchymal cells (RPTECs), expression in normal kidney, or from published literature.

<sup>b</sup> Column means are geometric.

## CHAPTER 8: Molecular landscape of ulcerative colitis

**Table 8.4** Top 30 transcripts decreased in UC biopsies with Endoscopic Mayo Score >1 versus ≤1, aligned by p-value

Probe Set ID	Title	Gene	Fold change	p-value	Adjusted p-value	Expression in colon biopsies			Expression in cell lines
						Mayo 2/3	Mayo 0/1	Screening colon	Principal Expression <sup>a</sup>
11724537_a_at	ATP binding cassette subfamily G member 2 (Junior blood group)	ABCG2	0.4	4.1x10 <sup>-8</sup>	3.7x10 <sup>-5</sup>	32	83	77	Parenchymal
11758175_s_at	solute carrier family 36 member 1	SLC36A1	0.6	6.9x10 <sup>-8</sup>	4.5x10 <sup>-5</sup>	238	428	496	Parenchymal
11719006_at	ANKH inorganic pyrophosphate transport regulator	ANKH	0.7	1.1x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	42	60	61	Parenchymal
11754623_x_at	choline phosphotransferase 1	CHPT1	0.6	1.4x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	1893	2916	3181	Parenchymal
11723376_a_at	GRINL1A complex locus 1 / myocardial zonula adherens protein	GCOM1 / MYZAP	0.6	1.7x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	32	53	51	Parenchymal
11741952_x_at	progesterin and adipoQ receptor family member 5	PAQR5	0.5	1.9x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	68	145	161	Parenchymal
11736513_s_at	N-myristoyltransferase 2	NMT2	0.7	1.9x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	126	190	191	Parenchymal
11747276_a_at	A-kinase anchoring protein 1	AKAP1	0.7	1.9x10 <sup>-7</sup>	5.0x10 <sup>-5</sup>	1714	2570	2824	Parenchymal
11753916_at	long intergenic non-protein coding RNA 526	LINC00526	0.6	2.1x10 <sup>-7</sup>	5.1x10 <sup>-5</sup>	144	222	223	Parenchymal
11758546_x_at	vitamin D receptor	VDR	0.6	2.4x10 <sup>-7</sup>	5.2x10 <sup>-5</sup>	592	997	1086	Parenchymal
11757885_at	V-set and immunoglobulin domain containing 10	VERSUSIG10	0.6	2.5x10 <sup>-7</sup>	5.2x10 <sup>-5</sup>	151	272	293	Parenchymal
11730907_a_at	ATP binding cassette subfamily B member 1	ABCB1	0.4	2.6x10 <sup>-7</sup>	5.2x10 <sup>-5</sup>	184	481	445	Parenchymal
11723377_a_at	myocardial zonula adherens protein	MYZAP	0.5	2.9x10 <sup>-7</sup>	5.2x10 <sup>-5</sup>	155	286	304	Parenchymal
11762517_at	cordon-bleu WH2 repeat protein	COBL	0.6	2.9x10 <sup>-7</sup>	5.2x10 <sup>-5</sup>	71	117	125	Parenchymal
11716412_s_at	aldehyde dehydrogenase 18 family member A1	ALDH18A1	0.7	3.0x10 <sup>-7</sup>	5.2x10 <sup>-5</sup>	1204	1779	1653	Parenchymal
11722546_a_at	dehydrogenase/reductase 11	DHRS11	0.5	3.0x10 <sup>-7</sup>	5.2x10 <sup>-5</sup>	347	702	774	Parenchymal
11723599_a_at	PH domain and leucine rich repeat protein phosphatase 2	PHLPP2	0.5	3.1x10 <sup>-7</sup>	5.2x10 <sup>-5</sup>	344	747	868	Parenchymal
11730778_a_at	hydroxysteroid 17-beta dehydrogenase 11	HSD17B11	0.6	3.1x10 <sup>-7</sup>	5.2x10 <sup>-5</sup>	770	1211	1193	Parenchymal
11733595_x_at	pleckstrin homology and RhoGEF domain containing G6	PLEKHG6	0.6	3.2x10 <sup>-7</sup>	5.2x10 <sup>-5</sup>	229	396	406	Parenchymal
11717417_s_at	ras homolog family member U	RHOU	0.5	3.5x10 <sup>-7</sup>	5.3x10 <sup>-5</sup>	343	737	798	Parenchymal
11725601_at	soosondawah ankyrin repeat domain family member A	SOWAHA	0.4	3.7x10 <sup>-7</sup>	5.3x10 <sup>-5</sup>	170	406	490	Parenchymal
11736979_at	fibrous sheath interacting protein 1	FSIP1	0.6	3.7x10 <sup>-7</sup>	5.3x10 <sup>-5</sup>	22	39	47	Parenchymal
11744698_x_at	tumor protein p63 regulated 1 like	TPRG1L	0.7	3.7x10 <sup>-7</sup>	5.3x10 <sup>-5</sup>	570	796	858	Parenchymal
11763271_at	peroxisomal biogenesis factor 11 alpha	PEX11A	0.6	4.0x10 <sup>-7</sup>	5.5x10 <sup>-5</sup>	182	317	353	Parenchymal
11735331_a_at	RUN domain containing 3B	RUNDC3B	0.4	4.0x10 <sup>-7</sup>	5.5x10 <sup>-5</sup>	54	121	136	Parenchymal
11721773_at	WSC domain containing 1	WSCD1	0.5	4.0x10 <sup>-7</sup>	5.5x10 <sup>-5</sup>	44	84	84	Parenchymal
11747457_x_at	UDP glucuronosyltransferase family 1 member A1, A3-10	UGT1A1, UGT1A3-10	0.5	4.3x10 <sup>-7</sup>	5.5x10 <sup>-5</sup>	249	494	439	Parenchymal
11717192_x_at	troponin C2, fast skeletal type	TNNC2	0.6	4.3x10 <sup>-7</sup>	5.5x10 <sup>-5</sup>	45	70	73	Parenchymal
11736803_at	fibroblast growth factor 9	FGF9	0.7	4.4x10 <sup>-7</sup>	5.5x10 <sup>-5</sup>	28	41	47	Parenchymal
11755123_x_at	lysophosphatidylcholine acyltransferase 3	LPCAT3	0.7	4.5x10 <sup>-7</sup>	5.5x10 <sup>-5</sup>	601	866	792	Parenchymal
<b>Column Means<sup>b</sup></b>						172	306	321	

UC, ulcerative colitis.

NOTE: Screening colon samples were not used to generate the top genes, but we show the expression of the top genes from UC biopsies in screening colon samples as a representation of healthy colon tissue.

<sup>a</sup> Expression values had to be over 100 in the cell lines to be considered for principal expression. Parenchymal expression was determined either by expression in cultured parenchymal cells (RPTECs), expression in normal kidney, or from published literature.

<sup>b</sup> Column means are geometric.

**CHAPTER 8: Molecular landscape of ulcerative colitis**

**Table 8.5** UC activity-associated top 30 genes (redundant, by p-value) found in the TCMR-associated IQR filtered list in various biopsy populations, ordered by rank of transcript in kidney biopsies

Probe Set ID	Gene Name	Gene	Expression in kidney transplants versus colon biopsies								
			Values in the Kidney TCMR-associated IQR filtered list			Kidney biopsies			Colon biopsies		
			Fold Change	p-value	Adjusted p-value	TCMR	No TCMR	Normal Kidney	Mayo 2/3	Mayo 0/1	Normal Colon
11733725_a_at	complement factor B	CFB	1.7	3.1x10 <sup>-12</sup>	3.4x10 <sup>-11</sup>	1107	643	457	1160	443	275
11723303_at	PDZK1 interacting protein 1	PDZK1IP1	1.0	8.6x10 <sup>-1</sup>	8.9x10 <sup>-1</sup>	6943	6815	10497	1444	549	475
11756316_a_at	chitinase 3 like 1	CHI3L1	1.5	1.8x10 <sup>-3</sup>	4.6x10 <sup>-3</sup>	151	100	367	406	35	16
11730214_at	dual oxidase 2	DUOX2	NA	NA	NA	NA	NA	NA	1058	121	79
11755084_x_at	CD55 molecule (Cromer blood group)	CD55	1.1	7.4x10 <sup>-3</sup>	1.6x10 <sup>-2</sup>	222	196	213	2464	827	535
11725000_at	dual oxidase maturation factor 2	DUOXA2	NA	NA	NA	NA	NA	NA	845	116	60
11718564_at	gap junction protein alpha 4	GJA4	NA	NA	NA	NA	NA	NA	53	32	31
11737791_s_at	apolipoprotein L1 / apolipoprotein L2	APOL1 / APOL2	1.8	6.3x10 <sup>-16</sup>	1.2x10 <sup>-14</sup>	594	331	198	797	330	211
11724541_s_at	von Willebrand factor	VWF	1.2	9.8x10 <sup>-2</sup>	1.5x10 <sup>-1</sup>	505	419	152	1093	459	357
11742449_a_at	apolipoprotein L1	APOL1	1.5	4.4x10 <sup>-14</sup>	6.1x10 <sup>-13</sup>	450	301	228	386	212	155
11719366_s_at	C-X-C motif chemokine ligand 1	CXCL1	1.5	6.3x10 <sup>-4</sup>	1.8x10 <sup>-3</sup>	164	107	106	1947	298	182
11716639_a_at	collagen type IV alpha 1 chain	COL4A1	1.2	1.7x10 <sup>-2</sup>	3.3x10 <sup>-2</sup>	798	671	398	460	183	135
11732566_at	solute carrier family 6 member 14	SLC6A14	NA	NA	NA	NA	NA	NA	305	36	18
11763250_x_at	C-X-C motif chemokine ligand 1 / ligand 2	CXCL1 / 2	NA	NA	NA	NA	NA	NA	406	113	89
11742862_a_at	adhesion G protein-coupled receptor F5	ADGRF5	0.7	3.8x10 <sup>-10</sup>	3.2x10 <sup>-9</sup>	705	979	1378	245	115	91
11756072_s_at	serum amyloid A1 / serum amyloid A2 / serum amyloid A2-A4	SAA1 / SAA2 / SAA2-SAA4	2.5	1.6x10 <sup>-6</sup>	7.8x10 <sup>-6</sup>	131	52	16	387	41	19
11744718_a_at	collagen triple helix repeat containing 1	CTHRC1	2.6	1.6x10 <sup>-7</sup>	9.5x10 <sup>-7</sup>	238	93	16	155	40	30
11720440_at	olfactomedin like 2B	OLFML2B	1.6	9.0x10 <sup>-7</sup>	4.6x10 <sup>-6</sup>	181	116	58	142	65	59
11718832_a_at	lysyl oxidase like 2	LOXL2	1.1	6.5x10 <sup>-2</sup>	1.0x10 <sup>-1</sup>	114	101	66	189	98	65
11720020_x_at	regulator of telomere elongation helicase 1	RTEL1	NA	NA	NA	NA	NA	NA	285	120	80
11752035_a_at	nitric oxide synthase 2	NOS2	NA	NA	NA	NA	NA	NA	183	53	45
11716281_a_at	collagen type XVIII alpha 1 chain	COL18A1	NA	NA	NA	NA	NA	NA	274	133	118
11739315_at	kinase insert domain receptor	KDR	NA	NA	NA	NA	NA	NA	197	130	119
11739134_a_at	collagen type V alpha 2 chain	COL5A2	1.4	1.3x10 <sup>-3</sup>	3.4x10 <sup>-3</sup>	144	100	56	453	206	185
11756068_a_at	beta-site APP-cleaving enzyme 2	BACE2	1.1	3.4x10 <sup>-1</sup>	4.2x10 <sup>-1</sup>	527	499	718	1583	890	675
11742889_at	F-box protein 6	FBXO6	NA	NA	NA	NA	NA	NA	219	138	120
11726124_at	transcobalamin I	TCN1	NA	NA	NA	NA	NA	NA	119	32	23
11715239_x_at	interferon induced transmembrane protein 3	IFITM3	1.1	1.4x10 <sup>-2</sup>	2.7x10 <sup>-2</sup>	3422	3047	2106	2506	1417	1063
11736477_a_at	elongation factor for RNA polymerase II 2	ELL2	1.2	1.3x10 <sup>-4</sup>	4.5x10 <sup>-4</sup>	170	138	186	475	226	203
11728477_at	C-X-C motif chemokine ligand 3	CXCL3	1.2	8.2x10 <sup>-3</sup>	1.7x10 <sup>-2</sup>	13	11	15	775	158	119
			<b>Column Mean Expression<sup>a</sup></b>			<b>339</b>	<b>249</b>	<b>196</b>	<b>456</b>	<b>148</b>	<b>107</b>

NA, not available.

NOTE. Columns were compared using t tests (two sided, Welch's corrected). Mayo 2/3 to Mayo 0/1 had p=1.7x10<sup>-9</sup>. Mayo 0/1 to screening colon had p=0.04. Screening colon to normal kidney had p=0.26. Transcripts marked 'NA' did not pass the IQR filter of the UC-associated transcript list, and thus no data was available.

<sup>a</sup>Column means are geometric.

**CHAPTER 8: Molecular landscape of ulcerative colitis**

**Table 8.6** T cell process-associated/T cell-specific gene expression in UC colon and screening colon biopsies (IQR filtered).

Probe Set ID	Title	Gene	Fold Change	Adjusted p value	Mayo 2/3	Mayo 0/1	Screening Colon
11733736_a_at	CD2 molecule	CD2	1.3	$3.8 \times 10^{-2}$	279	214	189
11723264_a_at	CD3d molecule	CD3D	1.4	$4.5 \times 10^{-2}$	423	306	232
11758882_at	CD3d molecule	CD3D	1.0	$9.2 \times 10^{-1}$	91	89	77
11759629_a_at	T-cell receptor alpha constant	TRAC	1.4	$7.8 \times 10^{-2}$	604	441	379
11763233_x_at	T-cell receptor alpha constant	TRAC	1.3	$1.1 \times 10^{-1}$	538	414	347
11761960_x_at	T-cell receptor alpha constant	TRAC	1.3	$1.3 \times 10^{-1}$	395	305	257
11761959_x_at	T-cell receptor alpha constant	TRAC	1.2	$1.6 \times 10^{-1}$	304	243	200
11762294_x_at	T cell receptor beta constant 1	TRBC1	1.4	$4.4 \times 10^{-2}$	670	495	410
11762287_x_at	T cell receptor beta constant 2	TRBC2	1.3	$4.0 \times 10^{-2}$	1174	894	765
11763557_x_at	T cell receptor beta constant 1 and 2	TRBC1 / TRBC2	1.3	$6.6 \times 10^{-2}$	808	603	502
11761918_x_at	T cell receptor beta constant 1	TRBC1	1.3	$7.9 \times 10^{-2}$	663	509	418
				<b>Mean Expression</b>	<b>541</b>	<b>410</b>	<b>343</b>
11732956_a_at	T cell receptor gamma constant 1 and 2	TRGC1 / TRGC2	1.0	$9.2 \times 10^{-1}$	157	155	130
11763446_s_at	T cell receptor delta constant	TRDC	0.9	$2.7 \times 10^{-1}$	13	15	12
11761790_x_at	T cell receptor delta constant	TRDC	0.9	$4.3 \times 10^{-1}$	45	50	42
11763447_x_at	T cell receptor delta constant	TRDC	1.0	$8.2 \times 10^{-1}$	70	72	66
				<b>Mean Expression</b>	<b>71</b>	<b>73</b>	<b>63</b>

NOTE. Columns were compared using t tests (two sided, Welch's corrected), and done separately for the top (A/B T cells) and bottom (D/G T cells). For the top section (A/B T cells): Mayo 2/3 versus Mayo 0/1 had p-value=0.066. Mayo 0/1 versus screening colon had p-value = 0.27. For the bottom section (D/G T cells): Mayo 2/3 versus Mayo 0/1 had p-value=0.54. Mayo 0/1 versus screening colon had p-value = 0.19.

**CHAPTER 8: Molecular landscape of ulcerative colitis**

**Table 8.7** Expression of Kidney TCMR-associated Top 30 genes (redundant, by p value) found in the UC IQR filtered list (Endoscopic Mayo Score >1 versus ≤1) in various biopsy populations, ordered by rank of transcript in kidney biopsies

Probe Set ID	Gene Name	Gene	Values in the UC disease activity top transcripts list			Expression in biopsies					
			Fold Change	p-value	Adjusted p-value	Kidney biopsies			Colon biopsies		
						TCMR	No TCMR	Normal Kidney	Mayo 2 or 3	Mayo 0 or 1	Screening Colon
11758830_at	Src-like-adaptor	SLA	1.4	2.7x10 <sup>-4</sup>	1.5x10 <sup>-3</sup>	22	11	11	15	11	9
11750942_x_at	janus kinase and microtubule interacting protein 1	JAKMIP1	1.2	2.7x10 <sup>-3</sup>	7.7x10 <sup>-3</sup>	27	13	11	18	14	13
11730637_a_at	cytotoxic T-lymphocyte-associated protein 4	CTLA4	2.4	8.7x10 <sup>-6</sup>	1.8x10 <sup>-4</sup>	66	20	14	120	49	38
11735474_a_at	CD8b molecule	CD8B	NA	NA	NA	34	13	9	NA	NA	NA
11724729_a_at	CD8a molecule	CD8A	NA	NA	NA	120	52	43	NA	NA	NA
11735221_a_at	B and T lymphocyte associated	BTLA	1.3	2.9x10 <sup>-2</sup>	5.2x10 <sup>-2</sup>	65	22	14	58	44	41
11730060_a_at	lymphocyte-activation gene 3	LAG3	1.4	2.2x10 <sup>-4</sup>	1.3x10 <sup>-3</sup>	146	72	55	113	80	62
11730947_at	interferon, gamma	IFNG	1.6	5.3x10 <sup>-4</sup>	2.3x10 <sup>-3</sup>	79	30	22	34	22	19
11747366_x_at	cytotoxic and regulatory T-cell molecule	CRTAM	NA	NA	NA	41	22	16	NA	NA	NA
11731360_a_at	inducible T-cell co-stimulator	ICOS	1.5	2.9x10 <sup>-4</sup>	1.5x10 <sup>-3</sup>	83	38	28	85	56	48
11731764_a_at	CD3g molecule, gamma (CD3-TCR complex)	CD3G	1.1	1.6x10 <sup>-1</sup>	2.2x10 <sup>-1</sup>	126	47	31	148	129	109
11734112_at	T-cell immunoreceptor with Ig and ITIM domains	TIGIT	1.3	4.6x10 <sup>-2</sup>	7.7x10 <sup>-2</sup>	75	29	19	48	37	31
11754881_a_at	signal-regulatory protein gamma	SIRPG	1.3	2.5x10 <sup>-2</sup>	4.6x10 <sup>-2</sup>	64	27	17	51	41	34
11749595_a_at	guanylate binding protein 4	GBP4	1.2	1.2x10 <sup>-2</sup>	2.6x10 <sup>-2</sup>	42	18	14	13	10	8
11750826_x_at	ADAM-like, decysin 1	ADAMDEC1	0.9	3.4x10 <sup>-1</sup>	4.1x10 <sup>-1</sup>	46	13	8	929	1061	1298
11730994_at	sphingosine-1-phosphate receptor 4	S1PR4	NA	NA	NA	65	37	42	NA	NA	NA
11724374_at	basic leucine zipper transcription factor, ATF-like	BATF	1.6	4.1x10 <sup>-4</sup>	1.9x10 <sup>-3</sup>	102	45	30	123	77	57
11719692_a_at	retinoic acid receptor responder (tazarotene induced) 3	RARRES3	1.3	5.6x10 <sup>-3</sup>	1.4x10 <sup>-2</sup>	2382	886	588	841	626	420
11740451_a_at	killer cell lectin-like receptor subfamily D, member 1	KLRD1	NA	NA	NA	176	62	32	NA	NA	NA
11731065_a_at	GTPase, IMAP family member 1	GIMAP1	NA	NA	NA	114	66	55	NA	NA	NA
11723069_at	gamma-aminobutyric acid (GABA) B receptor, 1 /// ubiquitin D	GABBR1 /// UBD	2.9	1.3x10 <sup>-4</sup>	8.7x10 <sup>-4</sup>	1854	417	213	1021	358	174
11739527_a_at	secreted and transmembrane 1	SECTM1	0.8	9.4x10 <sup>-3</sup>	2.1x10 <sup>-2</sup>	709	403	378	430	561	478
11751299_a_at	apolipoprotein L, 2	APOL2	1.2	5.9x10 <sup>-2</sup>	9.4x10 <sup>-2</sup>	193	106	113	85	69	51
11746087_a_at	CD84 molecule	CD84	1.1	6.7x10 <sup>-1</sup>	7.3x10 <sup>-1</sup>	131	32	11	46	43	33
11732425_at	ankyrin repeat domain 22	ANKRD22	1.5	2.7x10 <sup>-3</sup>	7.7x10 <sup>-3</sup>	172	55	34	609	399	280
11742241_a_at	IKAROS family zinc finger 3	IKZF3	1.5	3.0x10 <sup>-2</sup>	5.4x10 <sup>-2</sup>	57	23	13	80	52	44
11720161_at	chemokine (C-X-C motif) ligand 13	CXCL13	3.7	6.6x10 <sup>-4</sup>	2.7x10 <sup>-3</sup>	374	40	14	614	165	284
11750000_a_at	C-type lectin domain family 7, member A	CLEC7A	1.1	1.4x10 <sup>-1</sup>	2.0x10 <sup>-1</sup>	73	29	25	20	17	15
11736581_a_at	glucosaminyl (N-acetyl) transferase 1, core 2	GCNT1	0.8	4.2x10 <sup>-3</sup>	1.1x10 <sup>-2</sup>	34	17	12	88	109	112
11726287_a_at	tryptophanyl-tRNA synthetase	WARS	2.0	4.2x10 <sup>-6</sup>	1.2x10 <sup>-4</sup>	1227	605	442	1482	755	529
<b>Column Mean Expression<sup>a</sup></b>						<b>114</b>	<b>45</b>	<b>31</b>	<b>116</b>	<b>82</b>	<b>69</b>

NA, not available.

NOTE. Columns were compared using t tests (two sided, Welch's corrected). Mayo 2/3 to Mayo 0/1 had p=5.4x10<sup>-4</sup>. Mayo 0/1 to screening colon had p=0.03. Screening colon to normal kidney had p=0.1.2x10<sup>-9</sup>. Transcripts marked 'NA' did not pass the IQR filter of the UC-associated transcript list, and thus no data was available.

<sup>a</sup>Column means are geometric.

**CHAPTER 8: Molecular landscape of ulcerative colitis**

**Table 8.8** Gene Ontology – Biological Process terms from top 300 UC-associated transcripts (Endoscopic Mayo scores >1 versus ≤1)<sup>a</sup>

Term Name	Count	Pop Hits	Fold Enrichment	Benjamini corrected p-value
GO:0030574~collagen catabolic process	18	64	17.62	1.94x10 <sup>-13</sup>
GO:0030198~extracellular matrix organization	25	196	7.99	7.08x10 <sup>-12</sup>
GO:0001525~angiogenesis	24	223	6.74	7.01x10 <sup>-10</sup>
GO:0007155~cell adhesion	31	459	4.23	2.40x10 <sup>-8</sup>
GO:0032496~response to lipopolysaccharide	17	164	6.49	2.80x10 <sup>-6</sup>
GO:0001666~response to hypoxia	17	172	6.19	4.63x10 <sup>-6</sup>
GO:0022617~extracellular matrix disassembly	12	76	9.89	7.67x10 <sup>-6</sup>
GO:0006954~inflammatory response	24	379	3.97	9.33x10 <sup>-6</sup>
GO:0045087~innate immune response	25	430	3.64	2.04x10 <sup>-5</sup>
GO:0050900~leukocyte migration	13	122	6.68	9.82x10 <sup>-5</sup>
GO:0006508~proteolysis	25	500	3.13	2.52x10 <sup>-4</sup>
GO:0051607~defense response to virus	14	165	5.32	3.50x10 <sup>-4</sup>
GO:0006958~complement activation, classical pathway	11	99	6.96	5.27x10 <sup>-4</sup>
GO:0006935~chemotaxis	12	122	6.16	4.93x10 <sup>-4</sup>
GO:0006955~immune response	22	421	3.27	4.68x10 <sup>-4</sup>
GO:0042060~wound healing	10	80	7.83	5.37x10 <sup>-4</sup>
GO:0071230~cellular response to amino acid stimulus	8	47	10.66	8.92x10 <sup>-4</sup>
GO:0045766~positive regulation of angiogenesis	11	115	5.99	0.0015
GO:0070098~chemokine-mediated signaling pathway	9	71	7.94	0.0015
GO:0030199~collagen fibril organization	7	39	11.25	0.0028
GO:0016525~negative regulation of angiogenesis	8	62	8.08	0.0046
GO:0035987~endodermal cell differentiation	6	27	13.92	0.0047
GO:0043434~response to peptide hormone	7	44	9.97	0.0049
GO:0060326~cell chemotaxis	8	65	7.71	0.0055
GO:0043542~endothelial cell migration	6	29	12.96	0.0059
GO:0016477~cell migration	12	172	4.37	0.0066
GO:0030449~regulation of complement activation	6	30	12.53	0.0065
GO:0042127~regulation of cell proliferation	12	185	4.06	0.012
GO:0090023~positive regulation of neutrophil chemotaxis	5	22	14.24	0.022
GO:0007165~signal transduction	35	1161	1.89	0.024
GO:0030335~positive regulation of cell migration	11	184	3.75	0.041
GO:0051591~response to cAMP	6	46	8.17	0.042
GO:0060333~interferon-gamma-mediated signaling pathway	7	71	6.18	0.048

GO, Gene Ontology; UC, ulcerative colitis.

<sup>a</sup>268 genes from the top 300 UC-associated transcripts (Endoscopic Mayo score >1 versus ≤1) were included in this analysis, as they were also found in the 16792 genes in this query. This list was truncated to include only significant Benjamini corrected p-values (≥0.05).

## CHAPTER 8: Molecular landscape of ulcerative colitis

**Table 8.9** Gene Ontology – Cellular compartment terms from top 300 UC-associated transcripts (Endoscopic Mayo scores >1 versus ≤1)<sup>a</sup>

Term Name	Count	Pop Hits	Fold Enrichment	Benjamini corrected p-value
GO:0005615~extracellular space	79	1347	3.82	2.62x10 <sup>-23</sup>
GO:0005576~extracellular region	78	1610	3.15	3.63x10 <sup>-18</sup>
GO:0031012~extracellular matrix	30	296	6.60	1.64x10 <sup>-13</sup>
GO:0005578~proteinaceous extracellular matrix	26	268	6.31	3.49x10 <sup>-11</sup>
GO:0005581~collagen trimer	14	92	9.90	7.58x10 <sup>-8</sup>
GO:0070062~extracellular exosome	81	2811	1.88	4.67x10 <sup>-7</sup>
GO:0005604~basement membrane	12	79	9.89	1.19x10 <sup>-6</sup>
GO:0005788~endoplasmic reticulum lumen	17	192	5.76	1.48x10 <sup>-6</sup>
GO:0009986~cell surface	28	542	3.36	2.30x10 <sup>-6</sup>
GO:0009897~external side of plasma membrane	17	213	5.19	4.99x10 <sup>-6</sup>
GO:0072562~blood microparticle	13	152	5.57	9.39x10 <sup>-5</sup>
GO:0005886~plasma membrane	96	4121	1.52	1.60x10 <sup>-4</sup>
GO:0031093~platelet alpha granule lumen	6	55	7.10	0.030

GO, Gene Ontology; UC, ulcerative colitis.

<sup>a</sup>280 genes from the top 300 UC-associated transcripts (Endoscopic Mayo score >1 versus ≤1) were included in this analysis, as they were also found in the 18224 possible genes in this pathway query. This list was truncated to include only significant Benjamini corrected p-values (≥0.05).

**CHAPTER 8: Molecular landscape of ulcerative colitis**

**Table 8.10** Gene Ontology – Molecular function terms from top 300 UC-associated transcripts (Endoscopic Mayo scores >1 versus ≤1)<sup>a</sup>

Term Name	Count	Pop Hits	Fold Enrichment	Benjamini corrected p-value
GO:0005201~extracellular matrix structural constituent	15	67	14.76	5.05x10 <sup>-10</sup>
GO:0048407~platelet-derived growth factor binding	6	11	35.97	7.58x10 <sup>-5</sup>
GO:0004252~serine-type endopeptidase activity	18	255	4.65	5.51x10 <sup>-5</sup>
GO:0005509~calcium ion binding	31	717	2.85	5.26x10 <sup>-5</sup>
GO:0005178~integrin binding	11	105	6.91	3.97x10 <sup>-4</sup>
GO:0045236~CXCR chemokine receptor binding	5	9	36.63	4.70x10 <sup>-4</sup>
GO:0005518~collagen binding	8	60	8.79	0.00215
GO:0008009~chemokine activity	7	49	9.42	0.005292
GO:0005515~protein binding	160	8785	1.20	0.026901

GO, Gene Ontology; UC, ulcerative colitis.

<sup>a</sup>256 genes from the top 300 UC-associated transcripts (Endoscopic Mayo score >1 versus ≤1) were included in this analysis, as they were also found in the 16881 genes in this query. This list was truncated to include only significant Benjamini corrected p-values (≥0.05).

**CHAPTER 8: Molecular landscape of ulcerative colitis**

**Table 8.11** KEGG pathway analysis on top 300 UC-associated transcripts (Endoscopic Mayo scores >1 versus ≤1)<sup>a</sup>

Term Name	Count	Pop Hits	Fold Enrichment	Benjamini corrected p-value
hsa04610:Complement and coagulation cascades	13	69	9.23	1.80x10 <sup>-6</sup>
hsa05133:Pertussis	12	75	7.84	2.37x10 <sup>-5</sup>
hsa04974:Protein digestion and absorption	12	88	6.68	8.25x10 <sup>-5</sup>
hsa05146:Amoebiasis	12	106	5.55	3.94x10 <sup>-4</sup>
hsa04668:TNF signaling pathway	12	106	5.55	3.94x10 <sup>-4</sup>
hsa05150:Staphylococcus aureus infection	9	54	8.17	4.05x10 <sup>-4</sup>
hsa04512:ECM-receptor interaction	10	87	5.63	0.002
hsa04151:PI3K-Akt signaling pathway	20	345	2.84	0.002
hsa04510:Focal adhesion	14	206	3.33	0.006
hsa04062:Chemokine signaling pathway	13	186	3.43	0.007
hsa05323:Rheumatoid arthritis	9	88	5.01	0.007

KEGG, Kyoto Encyclopedia of Genes and Genomes; UC, ulcerative colitis.

<sup>a</sup> 141 genes from the top 300 UC-associated transcripts (Endoscopic Mayo score >1 versus ≤1) were included in this analysis, as they were also found in the 6910 genes in this query. This list was truncated to include only significant Benjamini corrected p-values (≥0.05).

**CHAPTER 8: Molecular landscape of ulcerative colitis**

**Table 8.12** Comparisons of top 10 overrepresented terms in TCMR-associated top 300<sup>a</sup> transcripts in kidney biopsies compared to top ten overrepresented terms from top 300<sup>a</sup> UC activity-associated transcripts in UC biopsies

#	Biological Process		Cellular Compartment		KEGG pathway		Molecular Function	
	TCMR-associated <sup>b</sup>	UC activity-associated <sup>c</sup>	TCMR-associated <sup>b</sup>	UC activity-associated <sup>c</sup>	TCMR-associated <sup>b</sup>	UC activity-associated <sup>c</sup>	TCMR-associated <sup>b</sup>	UC activity-associated <sup>c</sup>
1	immune response	collagen catabolic process	<i>external side of plasma membrane</i>	extracellular space	Antigen processing and presentation	Complement and coagulation cascades	receptor binding	extracellular matrix structural constituent
2	adaptive immune response	extracellular matrix organization	immunological synapse	<i>extracellular region</i>	Natural killer cell mediated cytotoxicity	Pertussis	receptor activity	platelet-derived growth factor binding
3	regulation of immune response	angiogenesis	integral component of plasma membrane	extracellular matrix	Graft-versus-host disease	Protein digestion and absorption	MHC class I protein binding	serine-type endopeptidase activity
4	<i>inflammatory response</i>	cell adhesion	plasma membrane	proteinaceous extracellular matrix	Cell adhesion molecules (CAMs)	Amoebiasis	<i>protein binding</i>	calcium ion binding
5	T cell costimulation	response to lipopolysaccharide	T cell receptor complex	collagen trimer	Primary immunodeficiency	TNF signaling pathway	<i>chemokine activity</i>	integrin binding
6	positive regulation of T cell proliferation	response to hypoxia	membrane	extracellular exosome	Allograft rejection	<i>Staphylococcus aureus infection</i>	non-membrane spanning protein tyrosine kinase activity	CXCR chemokine receptor binding
7	<i>innate immune response</i>	extracellular matrix disassembly	extrinsic component of cytoplasmic side of plasma membrane	basement membrane	T cell receptor signaling pathway	ECM-receptor interaction	coreceptor activity	collagen binding
8	T cell receptor signaling pathway	<i>inflammatory response</i>	<i>extracellular region</i>	endoplasmic reticulum lumen	Type I diabetes mellitus	PI3K-Akt signaling pathway	GTPase activator activity	<i>chemokine activity</i>
9	T cell activation	<i>innate immune response</i>	MHC class II protein complex	cell surface	<i>Staphylococcus aureus infection</i>	Focal adhesion	transmembrane signaling receptor activity	<i>protein binding</i>
10	cell surface receptor signaling pathway	leukocyte migration	cell-cell junction	<i>external side of plasma membrane</i>	Cytokine-cytokine receptor interaction	Chemokine signaling pathway	SH3/SH2 adaptor activity	

NOTE. Terms overlapping between TCMR and Mayo lists are as follows: Innate immune response and inflammatory response (Biological Process), external side of plasma membrane and extracellular region (Cellular Compartment), Staphylococcus aureus infection (KEGG pathway), chemokine activity and protein binding (Molecular Function).

KEGG, Kyoto Encyclopedia of Genes and Genomes; TCMR, T cell-mediated rejection; UC, ulcerative colitis.

<sup>a</sup>All top 300 transcript lists are non-redundant and IQR filtered.

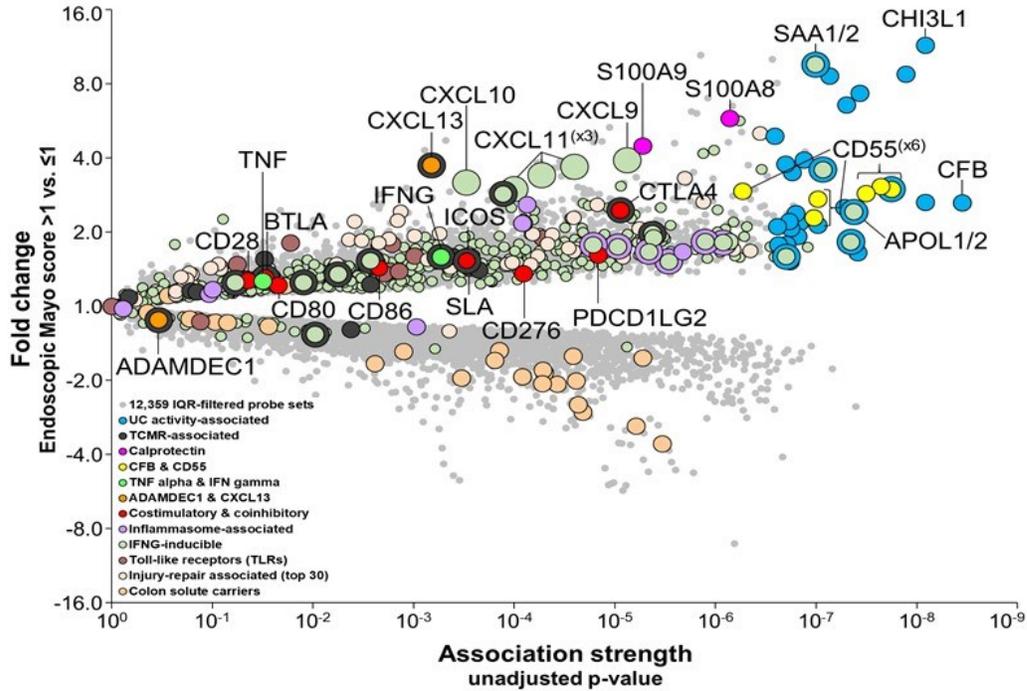
<sup>b</sup>TCMR-associated is TCMR versus everything else (EE, including no rejection, leaving out ABMR, borderline and mixed), based on a non-redundant annotated gene list of top 300 increased genes.

<sup>c</sup>UC-associated pathways are represented by top 300 non-redundant annotated genes increased in high mayo score UC versus low mayo score UC (2/3 versus 0/1).

## **CHAPTER 8: Molecular landscape of ulcerative colitis**

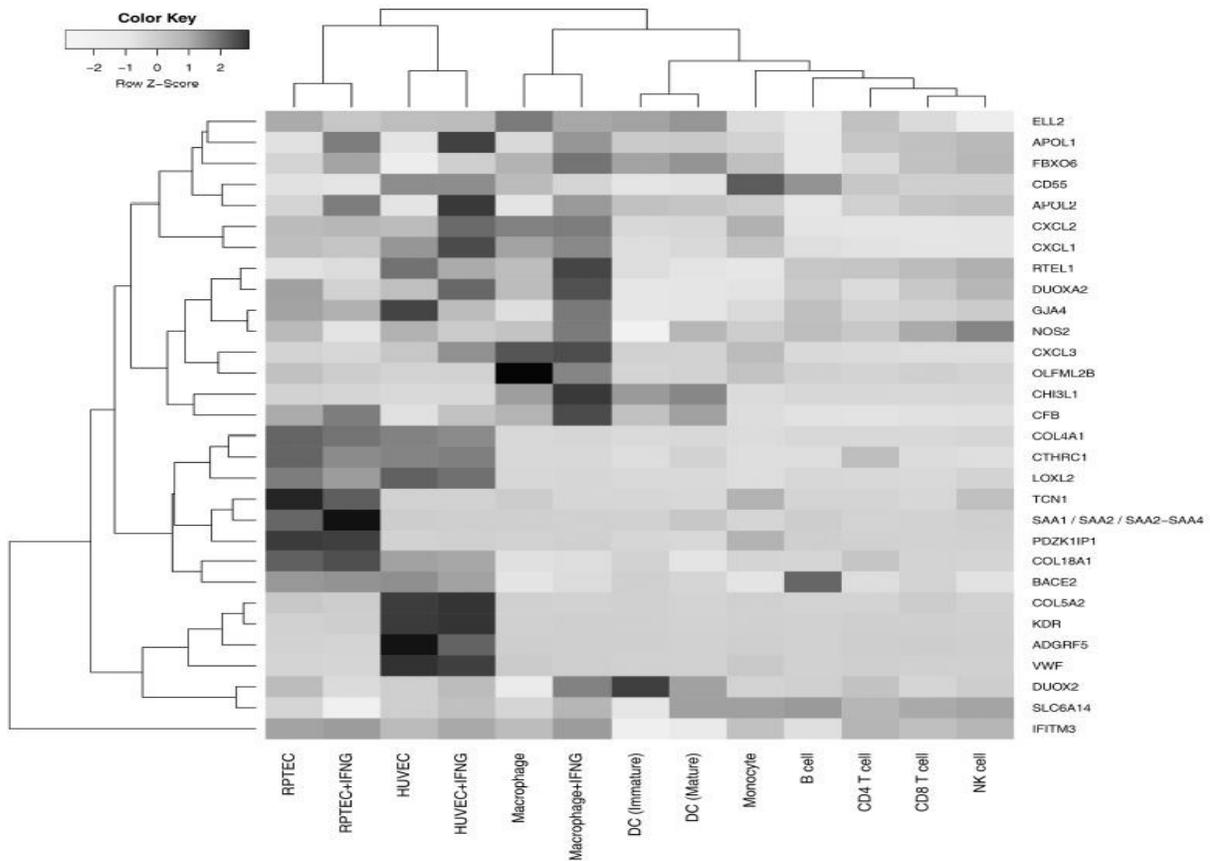
### **8.11 FIGURES**

CHAPTER 8: Molecular landscape of ulcerative colitis



**Figure 8.1 Molecular landscape of ulcerative colitis as shown by a volcano plot.** Probe sets towards the upper right have high association and fold change, indicating a strong relationship with UC (endoscopic Mayo score 2/3 versus 0/1). Probe sets towards the middle and further left have moderate to lower associations and fold change, indicating a lack of relationship with UC activity. Transcripts of interest were annotated. Abbreviations: UC, ulcerative colitis.

## CHAPTER 8: Molecular landscape of ulcerative colitis



**Figure 8.2** Heatmap of expression of top 30 UC activity-associated transcripts in an existing panel of cell lines. CD4<sup>+</sup> and CD8<sup>+</sup> T cells and NK cells lacked expression of the UC activity-associated transcripts. Probe sets were translated from U133 arrays to the PrimeView top 30 UC activity-associated transcripts. U133 non-redundant probe sets were selected using the highest mean expression across all cell types. For color coding, each transcript was assigned a z-score: normalized expression value across all conditions, divided by the standard deviation. Dendrograms were based on Euclidean distance and complete linkage strength. Abbreviations: NK, Natural killer; UC, ulcerative colitis

## **CHAPTER 9**

### **ASSESSING HETEROGENEITY IN ULCERATIVE COLITIS USING ARCHETYPAL ANALYSIS**

## **CHAPTER 9: Archetypal analysis of ulcerative colitis**

### **9.1 OVERVIEW AND HYPOTHESIS**

The limited understanding of disease mechanisms in UC, and the difficulty of predicting patient response to therapy present clinical challenges that could benefit from a precision medicine approach like that offered by the MMDx system. While therapeutic options for UC have expanded,<sup>318</sup> response to therapy remains inconsistent among patients with similar clinical presentation, and an empirical approach to management is required. Lack of precision in therapeutic decisions, the need to administer therapy over long periods before the efficacy can be assessed, and the enormous variation between patient cases, outcomes, and responses present major challenges to clinicians trying to induce remission and avoid colectomy. These approaches have also placed a significant financial burden on the healthcare system, especially in regions where UC is prevalent. Ineffective prediction of the most effective agent,<sup>319–321</sup> coupled with concerns over costs and adverse effects<sup>322,323</sup> have led to an increased interest in disease heterogeneity. It is possible that some form of disease heterogeneity that remains undetected by SOC methods is responsible for the variation in patient presentations, responses, and outcomes.

Diagnostics based on the microarray analysis of tissue biopsies in other organs with inflammatory diseases (such as TCMR of kidney transplants) has been shown to correlate well with the clinical phenotype,<sup>23,36</sup> require less tissue than histology, while providing more information per test and maintaining objectivity and reproducibility.<sup>64,66,171</sup> AA is an analytical method useful for detecting heterogeneity in a population.<sup>201</sup> This heterogeneity may be unrelated to the current SOC clinical phenotype, since AA assigns phenotypes based purely on molecular data. Based on prior experience with AA, it was hypothesized that applying these techniques in a UC biopsy population would lead to phenotypes defined by factors outside of the current SOC.<sup>6,66,171,201,299</sup> These analyses were part of a pilot study aiming to classify patients in a new way, using AA to look for heterogeneity among biopsies with similar endoscopic Mayo scores.

### **9.2 BIOPSY POPULATION AND DEMOGRAPHICS**

Biopsy collection, biopsy processing, and data collection methods and details are all described in Chapter 8, **Tables 8.1-8.2**.

### **9.3 ESTABLISHING UC ARCHETYPE GROUPS**

## CHAPTER 9: Archetypal analysis of ulcerative colitis

AA was done using the “archetypes” package in R,<sup>185</sup> and previously established methods.<sup>201</sup> Assignment of each biopsy to an AA group was done by convention using the highest archetype score.<sup>201</sup> Top transcripts by p-value were analyzed by correlation with archetype scores and mean expression in each archetype group.

### 9.3.1 AA initial input and determining the number of the groups

AA used the top 300 UC activity-associated transcripts as initial input (described previously in Chapter 8). Residual sums of squares of models using different numbers of archetypes (represented as a scree plot) suggested that three was an optimally small number of archetypes that could explain most of the phenotypic diversity among the biopsies (**Figure 9.1A**). Each individual biopsy in the population was assigned three archetype scores (S1, S2, S3 – the three archetype scores sum to 1.0) and its highest score was used to assign it to a group A1 (N=35), A2 (N=25), and A3 (N=11). Biopsies were plotted using PCA to visualize the AA groups, and colored by their group assignment (**Figure 9.1B**). ‘A1’, ‘A2’, and ‘A3’ on the figure designate the location of the ‘idealized’ biopsy typical of each of the archetypal groups.

### 9.3.2 Characterizing the AA groups

Statistical analysis and graphics were done in the “R” software package, version 3.3.2<sup>185</sup> with various libraries from Bioconductor 3.2,<sup>203</sup> and in Microsoft Excel version 15 (Redmond, WA). Significance of probe set expression is given as unadjusted p-values (Bayesian t-test), except in cases where the FDR is specified. Top 30 archetype-associated transcript lists in UC biopsies were compiled and sorted by p-value.

Biopsies in A1 were relatively normal, with lowest expression of transcripts associated with inflammation and injury (e.g. chemokines, *CASP1*, *CD55*) (**Table 9.1**). Biopsies in A2 showed increased expression of transcripts associated with inflammation (e.g. *C1S*) and response to wounding (RTW, e.g. collagens, *FBN1*) (**Table 9.2**). Biopsies in A3 showed increased expression of transcripts associated with inflammation and complement (*CASP1* and 5, *PSMB8* and 10) (**Table 9.3**).

### 9.3.3 Evidence for epithelial stem cell memory damage in AA groups

## CHAPTER 9: Archetypal analysis of ulcerative colitis

Some transcripts (i.e. *AIM2* and *IL-1 $\beta$* ) associated with open chromatin domains introduced during epithelial stem cell reprogramming<sup>306,307</sup> were associated with S2 (*AIM2* rank=1304 and  $p=1.4\times 10^{-17}$ , *IL-1 $\beta$*  rank=4926 and  $p=7.3\times 10^{-11}$ , compared to values in A3: *AIM2* rank=17556 and  $p=0.087$ , *IL-1 $\beta$*  rank 1121 and  $p=7.7\times 10^{-6}$ ).

### 9.4 RELATIONSHIPS BETWEEN MOLECULAR AND CONVENTIONAL FEATURES

Comparison of the archetype groups and scores to histological, clinical, and endoscopic parameters was done using a Bartlett's test for homoscedasticity of variance in each clinical parameter. If  $p<0.05$  for a given parameter, it was treated as heteroscedastic and a Welch's t-test was used to test whether its means in archetype groups A1, A2, and A3 were equal. If the parameter was homoscedastic ( $p\geq 0.05$ ), a simple F-test was used.

AA group means, PC1 and PC2 scores, and AA scores were analyzed for their relationships to clinical, endoscopic, and histological parameters using Spearman correlations and ANOVA (**Table 9.4**). The groups have differing characteristics: analysis of the A1, A2, and A3 group means using ANOVA show differences between the three groups in terms of endoscopic Mayo score at biopsy site, total Mayo score, and lymphoplasmacytic inflammation in the lamina propria, with moderate differences between the three groups in number of crypt abscesses and disease duration. However, A2 and A3 means are not significantly different from each other in terms of clinical or histological measures when compared in a t-test (including endoscopic and total Mayo score means, see footnote). Patient age is similar between A1, A2, and A3. Overall, results show that correlations are found between the molecular parameters (derived from the top 300 UC activity-associated transcript list), and other clinically relevant SOC measures that are not necessarily mayo score-related. Results also suggest that groups A2 and A3 differ based on heterogeneity not captured by SOC.

### 9.5 CROSS TABULATIONS OF ARCHETYPE GROUP CHARACTERISTICS

Cross tabulations of the endoscopic Mayo scores per archetype group are shown in **Table 9.5**, showing no Mayo 0 scores in A2 or A3. A2 and A3 archetype groups are associated with high Mayo scores, but the mean Mayo score between A2 and A3 is not significantly different.

### 9.5.1 Mayo score improvement in archetype groups

A subgroup analysis of biopsies with an initial endoscopic Mayo score of 2 or 3 (representing biopsies with active disease, N=26) was conducted to assess potential differences between archetype groups in terms of response to therapy. Improvement of the endoscopic Mayo score was defined as a reduction from 2-3 on the current biopsy to 0-1 at a future biopsy or endoscopic appointment.

Results are compiled in **Table 9.6**. A higher number of patients from A2 improved (53%) versus A3 (14%). These differences did not appear to be related to therapy applications between archetype groups (**Tables 9.7** and **9.8**). While the A2 was not significantly different than A3 with regards to Mayo score improvement ( $p=0.14$  by Fishers exact test), the numbers are trending in this direction. The available sample size was relatively small (restricted to the biopsies that had active disease and could improve to a lower Mayo score) limiting the conclusions of this analysis. Future iterations of this analysis will be carried out on a larger data set.

### 9.6 LOGISTIC REGRESSION ANALYSIS PREDICTING FUTURE IMPROVEMENT IN MAYO SCORE

It was initially hypothesized that archetype group assignment would be a predictor for response to therapy. In comparisons to the current SOC (endoscopic Mayo score), logistic regression analysis was done using endoscopic Mayo scores, archetype scores (S1, S2, S3), archetype group membership (A1, A2, A3), and PC1 scores on their own or as co-variables in the regression model. Various combinations of these factors were used to create models, compared by p-value and AUC.

The only models associated ( $p<0.05$ ) with this type of response to therapy include archetype group membership or archetype scores as a factor in the model, either in combination with the Mayo score or on their own. The only models capable of predicting a future reduction in endoscopic Mayo score were those including archetype group membership or archetype scores; used alone (model 2,  $p=0.01$ , AUC 0.77) or in conjunction with other factors (models 1, 3, 4, and 5:  $p=0.02$ ,  $p=0.03$ ,  $p=0.03$ ,  $p=0.03s$  and AUCs 0.74, 0.74, and 0.78 respectively). Results from comparing the logistic regression models are summarized in **Table 9.9**.

### **9.7 MONITORING UC DISEASE RESPONSE TO THERAPY USING AA**

Serial biopsies were collected from patients undergoing changes in their UC therapy (multiple sequential biopsies from a single patient range 2-5, collected from 10 patients). These biopsies were chart-reviewed for clinical information, including the endoscopic Mayo score of each biopsy, type of therapy being administered at each biopsy, and dose of therapy.

**Figure 9.2** shows the AA composition for the set of serial biopsies. Since AA scores are proportions summing to 1.0, score values are represented by proportional colors (A1 score as green, A2 score as blue, and A3 score as red). Patient IDs are displayed on the x-axis (numbers 1-10), and the proportion of 1.0 represented by each archetype score is displayed on the y-axis.

In many cases, changes in AA score composition for each biopsy mirrored changes in the endoscopic Mayo score, but in a small number of biopsies the AA composition changed when no shift in endoscopic Mayo score was detected. For example, in Patient #2 where the March 2014, December 2014, and November 2015 biopsies had endoscopic Mayo scores of 2, 3, and 3 respectively, AA composition showed a dynamic shift towards the A3 phenotype which was finally resolved in the last November 2016 biopsy. The AA composition mirrored a non-response and increase in disease severity when Humira was administered. In Patient #8, the AA composition of the March 2014 biopsy was fully A1, then shifted to A3 plus A2 in the May 2015 biopsy, suggesting that perhaps the original therapies (Asacol or Remicade) were a delayed response rather than a non-response, and perhaps should have been continued rather than a switch in therapy to Humira. Regardless, dynamic shifts in AA composition over time and changes in therapy suggest that molecular biopsy interpretation using AA groups can be used to monitor patient progress and may represent 'real-time' changes better than the SOC (which relies on visual healing and tissue recovery – changes that take time even after the molecular signals of distress in the tissue have stopped).

### **9.8 INTERPRETATION OF RESULTS**

Although this is currently a pilot study to generate hypotheses, findings were promising for future work in a larger cohort. Using the previously defined top 300 UC activity-associated transcripts (see Chapter 8), AA revealed heterogeneity within the population that was not captured by the Mayo score. Three main groups of biopsies were defined in the data set, and the mean characteristics differed between each group.

## CHAPTER 9: Archetypal analysis of ulcerative colitis

A2 displayed injury (RTW) transcripts that were not present in A3, and potentially represents UC disease that is more likely to improve compared to A3, despite similar Mayo score distributions between both groups. The top four predictive logistic regression models (in terms of model p-value) emphasized the importance of either the archetype group membership or the archetype scores for predicting subsequent reduction in Mayo score, even when Mayo score was excluded as a variable. Transcripts associated with epithelial stem cell (EpSC) reprogramming<sup>306,307</sup> were found in the transcripts associated with A2. The emergence of parenchymal changes in the archetype groups associated with wounding but not captured by histology is potentially a key finding in predicting and guiding treatment and following response-to-therapy. This has echoes in molecular studies of kidney transplant biopsies, where transcript expression levels predict functional impairment but histologic readings do not,<sup>195</sup> and where parenchymal injury is the dominant feature of prognosis.<sup>266</sup> The only histologic feature associated with these molecular scores was the lymphoplasmacytic infiltrate in the lamina propria, consistent with previous work.<sup>172,310,324</sup> In conclusion, molecular phenotyping in this pilot study shows promise for mapping clinically relevant heterogeneity that is not captured by the SOC.

Molecular heterogeneity in biopsies with high endoscopic Mayo scores (2/3) raises the possibility of relating archetype group membership or scores to clinical measures or endpoints. Not all severe cases (Mayo 2/3) have the same molecular characteristics or appear in the same archetype group, which may relate to responders versus non-responders, chronicity of disease, or potential for remission. While other phenotypic endpoints are getting attention in clinical practice (e.g. mucosal healing, defined as the absence of friability, blood, erosions and ulcers in colonic mucosa),<sup>325</sup> we propose that archetype scores and group membership may provide another clinical endpoint representing remission or healing in the tissue – with more reproducibility and granularity than other available measures. Logistic regression results also demonstrated the utility of molecular measurements for prediction of future reduction in Mayo score, showing that models incorporating archetype membership performed the best in terms of predicting improvement versus no improvement. This suggests that there is heterogeneity within patients with significant inflammation, and that this difference is not currently available via endoscopy or standard histopathology. Two models performed particularly well for the prediction of a future reduction in endoscopic Mayo score: a model using AA cluster membership, AA scores, endoscopic Mayo score, and PC1 (bias

## CHAPTER 9: Archetypal analysis of ulcerative colitis

corrected AUC = 0.82); and a model using only archetype cluster membership (bias-corrected AUC = 0.77). Either of these models could be used to predict the probability of a future reduction in endoscopic Mayo score in an unknown sample without follow-up information. Net reclassification index will be used in future analyses to establish the difference in performance between the two models in a larger data set.

Molecular patterns associated with tissue injury and loss-of-function are additional features capable of providing data not currently accessible to UC clinicians, and which would add to clinical assessment and management. It has been shown that tissue injury is not visually assessable in other tissues.<sup>195</sup> The ability to detect reversible or irreversible injury in the colonic tissue would be useful as a real-time measure of colon tissue function, for which there is currently no available test.

The differences between A2 and A3 may be a result of the RTW seen in the gene lists, and this may be related to a resistance or response to therapy. The contingency table tallying improved versus not improved cases in each archetype group showed a difference in percentage of cases improved between A2 and A3 (53% versus 14%, respectively), though both show signs of significant tissue inflammation and both contain a majority of high Mayo score biopsies with similar distributions. It is possible that the presence of RTW-associated transcripts corresponds to a response to therapy and the tissue undergoing a remodeling process to repair damage, whereas the A3 patients are inflamed without this repair process and thus are subject to continuous damage without recovery. As RTW is a finding not available via endoscopy or histological measurements, this novel measure may prove very useful in distinguishing patients truly resistant to therapy from those who take longer to respond or require alternate therapy regimens.

As discussed previously (see Chapter 8), we have found transcripts associated with EpSC reprogramming (such as *AIM2* and *IL-1 $\beta$* )<sup>306,307,311</sup> in our UC activity-associated data sets, and now in our AA (particularly in A2), suggesting that damage remembered by EpSC may contribute to the overall pathogenesis in UC. More A2 samples had a reduced Mayo score on a subsequent biopsy, thus the degree of EpSC reprogramming may be a contributor to this discrepancy.

Further evidence for inflammasome activation within UC was found in AA (*CASP1*, *CASP10*, *CASP6* and *IL-18* are all most associated with A3, and Mayo score 2/3 versus 0/1 in Chapter 8). This finding supports the model that hypersensitive EpSCs and inflammasome activation together play a role in the disease mechanisms of UC.

## CHAPTER 9: Archetypal analysis of ulcerative colitis

Some limitations to this study include the unavailability of clinical information regarding biopsies from other centers, the instability of mathematical models when N is small, and the heterogeneity of treatment. While the clarity and stability of these results would be improved with a larger data set and improved treatment information that takes a partial response into account, these findings are still promising and suggest clinical utility for molecular measurements in UC diagnosis. Future work includes collecting more serial biopsies, increasing the N of the reference set (with special focus on high Mayo score biopsies and follow-up data), and prospectively seeking paired biopsies from before and after treatment.

In summary, UC has more heterogeneity among patients and between biopsies than is currently defined by clinical standards. PCA and AA may add to the diagnostic assessment of biopsies to give more clarity to the clinician. EpSC damage via a T cell process may make the epithelium hypersensitive, and archetypes may be a way of assessing the presence or absence of this mechanism. The archetype group membership and the RTW patterns with A2 versus A3 show promise in this pilot study for prognosis or diagnosis in similar endoscopic Mayo score patients. Models developed through these analyses can be used to predict the probability of a reduced endoscopic Mayo score at follow-up for an unknown biopsy, providing some additional prognostic information that was not previously available.

## **CHAPTER 9: Archetypal analysis of ulcerative colitis**

### **9.9 TABLES**

**CHAPTER 9:** Archetypal analysis of ulcerative colitis

**Table 9.1** Top 30 transcripts associated with Archetype group 1 (redundant list)

Probe Set ID	Spearman Correlation	p-value	Gene Symbol	Name	PBT	A1	A2	A3	In UC activity-associated 300?
11733725_a_at	-0.945	1.1E-35	CFB	complement factor B		363	1538	1088	1
11719366_s_at	-0.940	5.1E-34	CXCL1	C-X-C motif chemokine ligand 1		173	4022	1916	1
11754114_a_at	-0.927	4.1E-31	CXCL1	C-X-C motif chemokine ligand 1		130	1630	814	0
11763250_x_at	-0.924	1.7E-30	CXCL1	C-X-C motif chemokine ligand 1		79	710	348	1
11744331_a_at	-0.924	1.9E-30	CFB	complement factor B		74	328	252	0
11732566_at	-0.916	4E-29	SLC6A14	solute carrier family 6 member 14		19	728	291	1
11726287_a_at	-0.916	4E-29	WARS	tryptophanyl-tRNA synthetase		635	2122	1133	1
11737146_a_at	-0.912	2.1E-28	SOCS1	suppressor of cytokine signaling 1		45	167	107	1
11756316_a_at	-0.912	2.3E-28	CHI3L1	chitinase 3 like 1		20	1113	220	1
11728701_a_at	-0.912	2.4E-28	CD55	CD55 molecule (Cromer blood group)		624	3717	2464	0
11719218_at	-0.912	2.5E-28	SOCS3	suppressor of cytokine signaling 3		46	404	160	1
11755084_x_at	-0.909	5.2E-28	CD55	CD55 molecule (Cromer blood group)		628	3600	2350	1
11718399_s_at	-0.907	1.1E-27	TGM2	transglutaminase 2	IRIT3	209	909	541	1
11715638_s_at	-0.906	1.8E-27	LPCAT1	lysophosphatidylcholine acyltransferase 1		160	590	388	1
11739471_a_at	0.904	4E-27	MAGI1	membrane associated guanylate kinase		182	86	119	0
11736477_a_at	-0.902	6.1E-27	ELL2	elongation factor for RNA polymerase II 2		187	693	364	1
11752832_x_at	-0.902	6.5E-27	CD55	CD55 molecule (Cromer blood group)		652	3502	2321	0
11728477_at	-0.901	1E-26	CXCL3	C-X-C motif chemokine ligand 3		99	1319	938	1
11726286_a_at	-0.897	3.2E-26	WARS	tryptophanyl-tRNA synthetase		369	1206	674	0
11762060_x_at	-0.897	4.2E-26	CD55	CD55 molecule (Cromer blood group)		520	2703	1788	0
11717732_s_at	-0.896	4.7E-26	PFKFB3	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3		312	1088	600	1
11715239_x_at	-0.896	5.1E-26	IFITM3	interferon induced transmembrane protein 3		1265	3009	2312	1
11744128_x_at	-0.895	6.2E-26	CXCL2	C-X-C motif chemokine ligand 2		128	1514	1005	1
11741885_x_at	-0.895	6.4E-26	CASP1	caspase 1	ENDAT	511	1247	1267	1
11737791_s_at	-0.895	7E-26	APOL1, APOL2	apolipoprotein L1, apolipoprotein L2		288	1068	613	1
11736478_a_at	-0.893	1.2E-25	ELL2	elongation factor for RNA polymerase II 2		22	66	39	0
11732565_at	-0.892	1.5E-25	SLC6A14	solute carrier family 6 member 14		44	1396	657	0
11716589_x_at	-0.892	1.6E-25	TMEM165	transmembrane protein 165		860	1580	1197	1
11757351_a_at	-0.892	1.7E-25	TAP1	transporter 1, ATP binding cassette subfamily B member		1207	2388	1908	0
11738884_x_at	-0.892	2E-25	CASP1	caspase 1	ENDAT	482	1185	1254	0

NOTE. All annotations for overlap between archetype lists was made by matching gene symbols between lists. CFB, WARS, CXCL3, CASP1 overlapped with the top 300 genes archetype list for Archetype 3. LPCAT1 overlapped with the top 300 genes archetype list for Archetype 2.

**CHAPTER 9:** Archetypal analysis of ulcerative colitis

**Table 9.2** Top 30 transcripts associated with Archetype group 2 (redundant list)

Probe Set ID	Spearman Correlation	p-value	Gene Symbol	Name	PBT	A1	A2	A3	In UC activity-associated 300?
11715496_a_at	0.958	2.9E-39	CTSK	cathepsin K	IRIT5	641	2549	995	1
11754368_a_at	0.955	4.1E-38	FBN1	fibrillin 1	IRIT5	230	668	273	1
11741710_a_at	0.953	2E-37	COL6A3	collagen type VI alpha 3 chain	IRIT5	1025	4271	1429	1
11715604_x_at	0.952	3.2E-37	SPARC	secreted protein acidic and cysteine rich	IRIT3	1507	4221	1859	1
11720644_a_at	0.950	1.4E-36	PECAM1	platelet and endothelial cell adhesion molecule 1	ENDAT	139	576	204	1
11743964_a_at	0.948	4E-36	COL6A3	collagen type VI alpha 3 chain	IRIT5	596	2519	801	0
11715651_s_at	0.948	4.4E-36	FSTL1	folliculin like 1	IRIT5	1083	2935	1395	1
11732526_s_at	0.943	1.2E-34	CSGALNACT1	chondroitin sulfate N-acetylgalactosaminyltransferase 1		39	152	53	0
11715603_s_at	0.941	4E-34	SPARC	secreted protein acidic and cysteine rich	IRIT3	1097	2946	1365	0
11723298_s_at	0.940	5.8E-34	AKAP2, PALM2-AKAP2	A-kinase anchoring protein 2, paralemmin 2 - A-kinase anchoring protein 2		506	1206	579	0
11758100_s_at	0.939	1.3E-33	C1S	complement C1s		1297	3807	1643	1
11716203_a_at	0.938	1.9E-33	MGP	matrix Gla protein	ENDAT IRIT5	333	1645	550	1
11718309_at	0.938	2.3E-33	HEG1	heart development protein with EGF like domains 1		235	604	315	0
11718198_at	0.936	4.3E-33	LHFP	LHFPL tetraspan subfamily member 6	IRIT3	197	497	216	1
11763872_x_at	0.936	4.7E-33	SPARC	secreted protein acidic and cysteine rich	IRIT3	517	1391	658	0
11716405_a_at	0.935	8.1E-33	HTRA1	HtrA serine peptidase 1		272	660	296	0
11727223_a_at	0.934	1.2E-32	CSGALNACT1	chondroitin sulfate N-acetylgalactosaminyltransferase 1		61	206	80	1
11720441_x_at	0.933	2E-32	OLFML2B	olfactomedin like 2B		48	206	76	0
11715852_at	0.931	5.7E-32	PDGFRB	platelet derived growth factor receptor beta	ENDAT	124	349	142	1
11716639_a_at	0.931	6.8E-32	COL4A1	collagen type IV alpha 1 chain	ENDAT IRIT3	156	716	281	1
11728768_s_at	0.931	7.2E-32	KLHL5	kelch like family member 5		203	609	258	0
11715501_s_at	0.930	1.2E-31	IGFBP7	insulin like growth factor binding protein 7	ENDAT KT1	1996	5103	2595	1
11749461_a_at	0.929	1.5E-31	CDH11	cadherin 11	IRIT5	248	685	313	0
11715542_s_at	0.929	1.6E-31	THY1	Thy-1 cell surface antigen		92	296	136	1
11721993_at	0.929	1.6E-31	SLC6A6	solute carrier family 6 member 6		145	446	201	0
11716638_s_at	0.928	2.3E-31	COL4A1	collagen type IV alpha 1 chain	ENDAT IRIT3	769	3009	1260	0
11719488_at	0.928	2.4E-31	EDNRA	endothelin receptor type A		50	190	71	0
11726581_a_at	0.927	3.6E-31	GUCY1B3	guanylate cyclase 1 soluble subunit beta		113	410	153	0
11728766_a_at	0.927	4E-31	KLHL5	kelch like family member 5		29	72	35	0
11716941_at	0.92694	4.3E-31	TRIB2	tribbles pseudokinase 2		177	690	258	0

NOTE. All annotations for overlap between archetype lists was made by matching gene symbols between lists. CTSK, PECAM1, COL4A1, and THY1 overlapped with the top 300 genes list for Archetype 3.

**CHAPTER 9:** Archetypal analysis of ulcerative colitis

**Table 9.3** Top 30 transcripts associated with Archetype group 3 (redundant list)

Probe Set ID	Spearman Correlation	p-value	Gene Symbol	Name	PBT	A1	A2	A3	In UC activity-associated 300?
11742386_a_at	0.828	5.2E-19	CASP1	caspase 1	ENDAT	279	645	799	0
11750553_x_at	0.789	3.1E-16	CASP1	caspase 1	ENDAT	444	1036	1165	0
11741076_a_at	0.770	4.3E-15	CASP1	caspase 1	ENDAT	222	587	652	0
11738884_x_at	0.768	5.2E-15	CASP1	caspase 1	ENDAT	482	1185	1254	0
11763186_at	0.768	5.3E-15	CASP1	caspase 1	ENDAT	386	790	868	0
11733465_at	0.764	9.1E-15	XDH	xanthine dehydrogenase	GRIT1	424	737	1039	0
11732870_a_at	0.761	1.3E-14	CASP1	caspase 1	ENDAT	923	2213	2283	0
11716641_x_at	0.757	2.1E-14	PSMB10	proteasome subunit beta 10	GRIT1	1284	1731	1852	0
11716345_a_at	0.757	2.2E-14	BACE2	beta-site APP-cleaving enzyme 2		250	405	492	0
11741885_x_at	0.750	4.9E-14	CASP1	caspase 1	ENDAT	511	1247	1267	1
11756138_a_at	0.750	5E-14	HK2	hexokinase 2		1008	1322	1925	0
11757278_x_at	0.749	5.6E-14	ASS1	argininosuccinate synthase 1	KT1	1493	2991	4563	0
11733464_at	0.749	5.8E-14	XDH	xanthine dehydrogenase	GRIT1	853	1436	2127	0
11731726_a_at	0.747	7E-14	CASP5	caspase 5		122	301	418	0
11716640_a_at	0.747	7.2E-14	PSMB10	proteasome subunit beta 10	GRIT1	1106	1431	1523	0
11753939_x_at	0.744	1E-13	PSMB10	proteasome subunit beta 10	GRIT1	1054	1400	1472	0
11758486_s_at	0.742	1.4E-13	BACE2	beta-site APP-cleaving enzyme 2		497	976	1067	0
11756614_a_at	0.739	1.9E-13	RNF145	ring finger protein 145		100	135	164	0
11753343_a_at	0.738	2.1E-13	CASP5	caspase 5		121	310	415	0
11758627_s_at	0.737	2.3E-13	ASS1	argininosuccinate synthase 1	KT1	1645	3497	4840	0
11731727_x_at	0.734	3.3E-13	CASP5	caspase 5		132	344	451	0
11715862_x_at	-0.729	5.9E-13	CDIPT	CDP-diacylglycerol--inositol 3-phosphatidyltransferase		257	208	186	0
11732425_at	0.727	7.3E-13	ANKRD22	ankyrin repeat domain 22		329	659	878	0
11715665_a_at	0.724	9.8E-13	PSMB8	proteasome subunit beta 8	GRIT1	1827	2496	2621	0
11755276_a_at	0.723	1.1E-12	GPX2	glutathione peroxidase 2		3324	4899	5978	0
11724829_at	0.716	2.2E-12	STS	steroid sulfatase		48	100	123	0
11736297_a_at	0.716	2.3E-12	SLC5A1	solute carrier family 5 member 1	IRIT5 KT2	192	257	332	0
11717621_a_at	-0.713	3.2E-12	FOXN3	forkhead box N3		230	194	157	0
11724830_at	0.710	4.2E-12	STS	steroid sulfatase		141	336	444	0
11748253_a_at	0.706	7E-12	SLC5A1	solute carrier family 5 member 1	IRIT5 KT2	158	202	265	0

NOTE. All annotations for overlap between archetype lists was made by matching gene symbols between lists. CASP1 and BACE2 overlapped with the top 300 genes archetype list for Archetype 1.

**CHAPTER 9:** Archetypal analysis of ulcerative colitis

**Table 9.4** Archetype scores and Spearman correlations between SOC and molecular features

Type of data	Feature Studied	Mean score for the cluster			ANOVA A1, A2, A3 p-values	Spearman correlations of archetype scores, PC1, and PC2 with the feature studied				
		A1	A2	A3		S1	S2	S3	PC1	PC2
<b>Clinical and endoscopic</b>	Endoscopic Mayo score at biopsy site	0.7	2.3	1.8	<0.0001	<b>-0.69<sup>a</sup></b>	0.62 <sup>a</sup>	0.33 <sup>a</sup>	0.68 <sup>a</sup>	0.22
	Total Mayo score	2.5	7.3	6.4	<0.0001	<b>-0.64<sup>a</sup></b>	0.55 <sup>a</sup>	0.34 <sup>a</sup>	0.63 <sup>a</sup>	0.28 <sup>b</sup>
	Age at biopsy (years)	40.7	33.4	40.4	0.07	<b>0.43<sup>a</sup></b>	-0.28 <sup>b</sup>	-0.37 <sup>a</sup>	-0.39 <sup>a</sup>	-0.30 <sup>b</sup>
	Disease duration (Months)	156.9	73.3	87.7	0.04	0.50 <sup>a</sup>	-0.43 <sup>a</sup>	-0.27 <sup>b</sup>	<b>-0.51<sup>a</sup></b>	-0.14
<b>Histology</b>	Lymphoplasmacytic inflammation in lamina propria	0.7	2.0	1.6	<0.0001	-0.58 <sup>a</sup>	0.55 <sup>a</sup>	0.36 <sup>a</sup>	<b>0.61<sup>a</sup></b>	0.34 <sup>a</sup>
	# of neutrophils in lamina propria	8.8	19.8	21.8	0.05	<b>-0.49<sup>a</sup></b>	0.39 <sup>a</sup>	0.44 <sup>a</sup>	0.47 <sup>a</sup>	0.40 <sup>a</sup>
	# of intraepithelial neutrophils	5.7	11.0	14.5	0.05	<b>-0.46<sup>a</sup></b>	0.36 <sup>a</sup>	<b>0.46<sup>a</sup></b>	0.43 <sup>a</sup>	0.42 <sup>a</sup>
	Number of crypt abscesses	0.1	1.4	0.6	0.02	-0.39 <sup>a</sup>	0.35 <sup>a</sup>	0.26 <sup>a</sup>	<b>0.41<sup>a</sup></b>	0.30 <sup>b</sup>
	Number of branched abscesses	1.7	2.7	2.9	0.15	-0.26 <sup>b</sup>	0.24	0.12	<b>0.28<sup>b</sup></b>	0.12
	Number of biopsy pieces assessed	2.7	2.6	2.8	0.82	0.02	-0.18	0.26 <sup>b</sup>	-0.11	<b>0.30<sup>b</sup></b>

NOTE. A1, A2, and A3 give the means of the rows in samples in the A1, A2, A3 groups respectively. A Bartlett test was performed on each feature studied to determine heteroscedasticity of variances. If Bartlett's  $p < 0.05$ , observations were treated as heteroscedastic and a Welch's test of one-way variance was used. Otherwise, a simple F-test was used. A2 and A3 means were not significantly different (all  $p$  values  $>0.05$ )

<sup>a</sup> P value  $<0.01$

<sup>b</sup> P value  $<0.05$

Bolded text indicates highest absolute value of the Spearman correlation coefficients within the row.

**CHAPTER 9:** Archetypal analysis of ulcerative colitis

**Table 9.5** Biopsy endoscopic Mayo scores tallied per archetype group

		# of Biopsies with each Endoscopic Mayo Score				
		Mayo 0	Mayo 1	Mayo 2	Mayo 3	TOTAL
Archetype Cluster (UC only, no replicates)	A1 <sup>a</sup>	16	13	5	1	35
	A2 <sup>b</sup>	0	3	11	11	24
	A3 <sup>b</sup>	0	4	5	2	12
	TOTAL	16	20	21	14	71

AA, archetypal analysis. A1, Archetype group 1. A2, archetype group 2. A3, archetype group 3.

<sup>a</sup> Entire table, Chi squared, df (38.7, 6), p-value <0.0001

<sup>b</sup> A2 versus A3 Fisher's exact test p-value 0.17

## CHAPTER 9: Archetypal analysis of ulcerative colitis

**Table 9.6** UC patients with initial Endoscopic Mayo score 2/3 in archetype groups A1, A2, and A3 classified by Mayo score response on a follow up biopsy

Archetype Cluster	# Cases Did Not Improve <sup>a</sup>	# Cases Improved <sup>b</sup>
<b>A1</b>	0 (0%)	4 (100%)
<b>A2</b>	6 (38%)	10 (62%)
<b>A3</b>	5 (83%)	1 (17%)

A1, Archetype group 1. A2, archetype group 2. A3, archetype group 3.

NOTE. A second method using an initial group of Endoscopic Mayo 2/3 was used, where lack of improvement was defined as if the score remained the same or was higher on subsequent biopsy, and improvement was a score reduced to any lower endoscopic mayo score on subsequent biopsy. The results were exactly the same as the method shown. Statistical results: Pearson's Chi-squared = 7.222, df = 2, p = 0.027 (A1, A2, A3). Fishers exact t-test for entire table p=0.03. Fishers exact t-test for A2 versus A3 p=0.14

<sup>a</sup> Score of 2/3 on original biopsy, score of 2/3 on subsequent biopsy

<sup>b</sup> Score of 2/3 on original biopsy, score of 0/1 on subsequent biopsy

## CHAPTER 9: Archetypal analysis of ulcerative colitis

**Table 9.7** Summary of recorded treatment for biopsies from the A2 and A3 archetype groups

<b>Archetype Group</b>	<b>5ASA</b>	<b>Imuran</b>	<b>Biologic</b>	<b>Prednisone</b>	<b>Cortifoam</b>
<b>A2 (N=24)</b>	11 (8 N/A)	6 (8 N/A)	5 (8 N/A)	4 (8 N/A)	1 (13 N/A)
<b>A3 (N=12)</b>	6 (4 N/A)	1 (4 N/A)	0 (4 N/A)	0 (4 N/A)	1 (11 N/A)

A2, archetype group 2. A3, archetype group 3.

**CHAPTER 9:** Archetypal analysis of ulcerative colitis

**Table 9.8** Detailed treatment data on all archetype 2/3 biopsies with high Mayo scores (2/3) used in the logistic regression analysis (Table 5)

	Initial Endoscopic Mayo Score	Follow-up Endoscopic Mayo score	responder (0=no, 1=yes)	Summary: Any treatment given?	Treatment types (1 = administered, 0 = not administered)				
					5ASA	Imuran	Biologic	Prednisone	Cortifoam
<b>A2</b>	2	1	1	No	0	0	0	0	0
	2	2	0	Yes	1	0	0	0	0
	3	1	1	Yes	1	1	0	0	1
	3	3	0	No	0	0	0	0	0
	2	3	0	No	0	0	0	0	0
	3	3	0	Yes	0	1	1	1	0
	2	0	1	Yes	1	0	0	0	0
	3	1	1	Yes	1	1	0	0	0
	2	1	1	Yes	1	1	1	1	0
	2	1	1	Yes	1	0	0	0	1
	2	2	0	Yes	0	1	1	0	0
	3	3	0	Yes	1	0	0	0	0
	2	0	1	Yes	0	1	1	1	1
	3	1	1	Yes	1	0	0	0	1
	3	1	1	Yes	1	1	0	0	0
3	1	1	Yes	1	0	0	1	0	
<b>A3</b>	2	2	0	Yes	1	0	0	0	0
	2	3	0	Yes	1	0	0	0	0
	2	2	0	Yes	1	0	0	0	0
	2	3	0	Yes	1	0	0	0	0
	2	3	0	No	0	0	0	0	0
	3	0	1	Yes	1	0	0	0	1

A2, archetype group 2. A3, archetype group 3.

**CHAPTER 9:** Archetypal analysis of ulcerative colitis

**Table 9.9** Comparative results from all logistic regression models assessing predictive value (response versus non-response as represented by Mayo Scores) of Endoscopic Mayo score, Archetype Scores (S1, S2, S3), Archetype group membership (A1, A2, A3), and PC1, sorted by p-value.

Analysis	Model Components	Model p-value	Bias-corrected AUC	Overall significance of the model?
<b>1</b>	<b>Endoscopic Mayo score (initial), Archetype group membership (A1, A2, A3), Archetype scores (S2, S3), PC1</b>	<b>0.01</b>	<b>0.82</b>	<b>Yes</b>
<b>2</b>	<b>Archetype group membership (A1, A2, A3)</b>	<b>0.01</b>	<b>0.77</b>	<b>Yes</b>
<b>3</b>	<b>Endoscopic Mayo score (initial), Archetype group membership (A1, A2, A3)</b>	<b>0.02</b>	<b>0.74</b>	<b>Yes</b>
<b>4</b>	<b>Archetype scores (S2, S3)<sup>a</sup>, PC1</b>	<b>0.03</b>	<b>0.74</b>	<b>Yes</b>
<b>5</b>	<b>Endoscopic Mayo score (initial), archetype scores (S2, S3), PC1</b>	<b>0.03</b>	<b>0.78</b>	<b>Yes</b>
<b>6</b>	Archetype scores (S2, S3)	0.09	0.67	No
<b>7</b>	Endoscopic Mayo score (initial), archetype scores (S2, S3)	0.12	0.69	No
<b>8</b>	Endoscopic Mayo score (initial)	0.31	0.55	No
<b>9</b>	Endoscopic Mayo score (initial), PC1	0.34	0.56	No
<b>10</b>	PC1	0.65	0.43	No

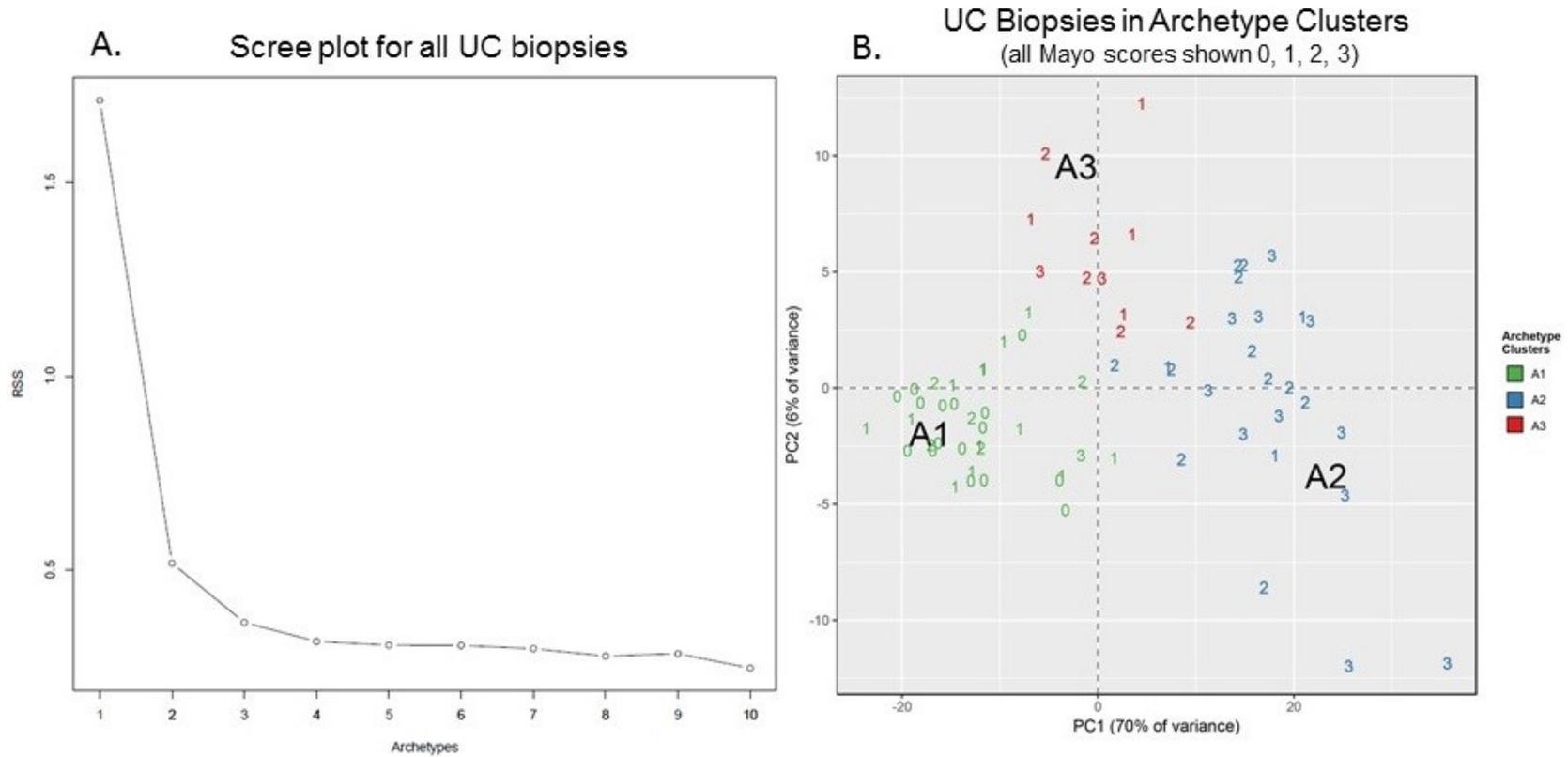
A1, Archetype group 1. A2, archetype group 2. A3, archetype group 3. S2, archetype score 2. S3, archetype score 3. PC1, principal component 1.

Bolded models were significant (p-value <0.05)

<sup>a</sup> S2, S3 indicates that all archetype scores were tested in the model. As all scores sum to 1.0, one score must be assumed for the scores to be included in the analysis.

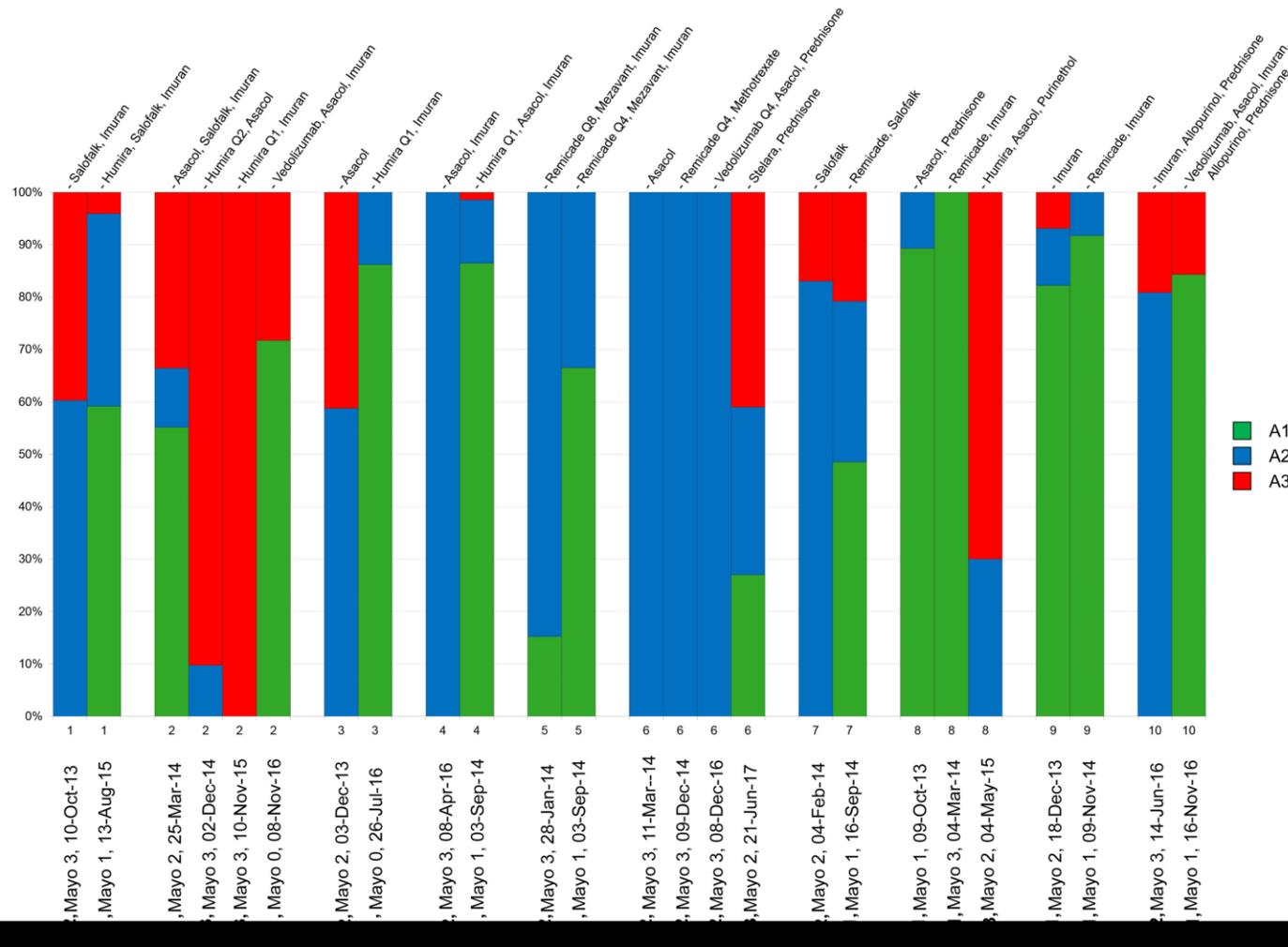
## **CHAPTER 9: Archetypal analysis of ulcerative colitis**

### **9.10 FIGURES**



**Figure 9.1 Archetypal analysis of biopsies taken from patients with diagnosed UC and varying levels of disease as represented by the endoscopic Mayo score.** Panel A shows the scree plot: the residual sum of squares versus the number of groups in a given archetype model. The three-archetype model was selected for these analyses based on the diminishing returns for an increased number of archetypes and on sample size (N). Panel B shows the archetype analysis plotted in principal component analysis space. Biopsies are placed based on their principal component 1 and 2 scores, and colored based on their archetype group assignment. Biopsies are represented by numbers which correspond to their endoscopic Mayo score as assessed by SOC. A1, A2, and A3 represent the idealized ‘biopsy’ that represents each archetype group.

## CHAPTER 9: Archetypal analysis of ulcerative colitis



**Figure 9.2 Changes in archetype composition over time in serial biopsies taken from 10 different UC patients with changing therapies.** Patients are labeled by number along the x-axis, with each biopsy noted for its date, endoscopic Mayo score as assessed by SOC, and archetype group assignment. The y-axis represents the proportion of the 1.0 allotted to each biopsy that is attributed to each of the archetype scores 1, 2, and 3. Therapies and dosages administered at the time of each biopsy are given along the top of the plot. Each bar is colored based on the archetype scores assigned to it.

## **SECTION IV**

### **CONCLUSIONS AND CUMULATIVE SIGNIFICANCE**

#### **SECTION 4: Conclusions and cumulative significance**

The initial hypothesis driving these analyses was that MMDx could address the unmet need for precision and accuracy in biopsy assessment, provide new insights and highly reproducible results from a small amount of tissue, and that the general approach used to develop MMDx-Kidney could be extrapolated to other areas of medicine in transplants and in native organs. This hypothesis was proven to be true throughout the course of the studies described in this thesis.

In kidney transplants, analyses of reproducibility in replicate pairs, including cortex and medulla paired biopsies, showed that MMDx-Kidney was highly reproducible and robust. Discrepancy analyses comparing interobserver variation within MMDx, within histology, and finally between MMDx and histology showed that an MMDx result can help clarify ambiguous cases and recalibrate the SOC to better reflect the true disease state.

The successful translation of the MMDx system into a new liver transplant population with a more challenging SOC histology system showed that MMDx is suitable for other transplants and is capable of detecting gene expression associated with disease from tissue samples regardless of their origin. Identification of TCMR using AA and the development of a high-performing molecular classifier for steatohepatitis indicate that MMDx-Liver will be useful in further delineating the disease phenotypes in a liver transplant population, and will be useful with diagnostic assignment in ambiguous cases.

Finally, in a population of biopsies taken from patients with diagnosed UC, MMDx was used to study disease mechanisms, and explore disease heterogeneity. These results showed that UC disease cannot be explained by one immunological process alone, but likely are caused by both innate and adaptive processes. Disease heterogeneity was not captured by the SOC Mayo scoring but with potential clinical implications for predicting responses was identified by AA in a pilot study, suggesting that AA may be useful in classifying patients with more granularity.

Overall, these experiments show that MMDx represents a general approach to creating next generation biopsy diagnostic systems that are robust, applicable to diverse areas of medicine, have highly reproducible results, and can clarify many ambiguous or challenging cases where the SOC is

**SECTION 4: Conclusions and cumulative significance**

insufficient. The continued development of MMDx systems for other organs or diseases will be of use to clinicians who want to make informed decisions regarding treatment.

## REFERENCES

## REFERENCES

## REFERENCES

1. Topol, E. J. *Deep medicine: how artificial intelligence can make healthcare human again*. (Basic Books, 2019).
2. Furness, P. N. *et al.* International variation in histologic grading is large, and persistent feedback does not improve reproducibility. *Am. J. Surg. Pathol.* **27**, 805–810 (2003).
3. Furness, P. N. & Taub, N. International variation in the interpretation of renal transplant biopsies: Report of the CERTPAP Project(1). *Kidney Int.* **60**, 1998–2012 (2001).
4. Horvath, B. *et al.* Interobserver Variability in Scoring Liver Biopsies with a Diagnosis of Alcoholic Hepatitis. *Alcohol. Clin. Exp. Res.* **41**, 1568–1573 (2017).
5. Daperno, M. *et al.* Inter-observer agreement in endoscopic scoring systems: preliminary report of an ongoing study from the Italian Group for Inflammatory Bowel Disease (IG-IBD). *Dig. Liver Dis. Off. J. Ital. Soc. Gastroenterol. Ital. Assoc. Study Liver* **46**, 969–973 (2014).
6. Reeve, J. *et al.* Generating automated kidney transplant biopsy reports combining molecular measurements with ensembles of machine learning classifiers. *Am. J. Transplant.* **00**, 1–13 (2019).
7. O'Connell, P. J. *et al.* Biopsy transcriptome expression profiling to identify kidney transplants at risk of chronic injury: a multicentre, prospective study. *Lancet Lond. Engl.* **388**, 983–993 (2016).
8. Van Loon, E., Lerut, E. & Naesens, M. The time dependency of renal allograft histology. *Transpl. Int. Off. J. Eur. Soc. Organ Transplant.* **30**, 1081–1091 (2017).
9. Gupta, G. *et al.* Late Antibody-Mediated Rejection in Renal Allografts: Outcome After Conventional and Novel Therapies. *Transplantation* **97**, 1240–1246 (2014).
10. Madill-Thomsen, K. *et al.* Discrepancy Analysis Comparing Molecular and Histology Diagnoses in Kidney Transplant Biopsies. *Am. J. Transplant.* *ajt.15752* (2019) doi:10.1111/ajt.15752.
11. Sahay, M., Kalra, S. & Bandgar, T. Renal endocrinology: The new frontier. *Indian J. Endocrinol. Metab.* **16**, 154–155 (2012).
12. NIH, National Institute of Diabetes and Digestive and Kidney Diseases. Your kidneys & how they work. *U.S. Department of Health and Human Services* <https://www.niddk.nih.gov/health-information/kidney-disease/kidneys-how-they-work#how>.

## REFERENCES

13. Hediger, M. A. *et al.* The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteinsIntroduction. *Pflugers Arch.* **447**, 465–468 (2004).
14. Fountain, J. H. & Lappin, S. L. Physiology, Renin Angiotensin System. in *StatPearls* (StatPearls Publishing, 2019).
15. Diabetes Canada. Managing my diabetes - Kidney Disease. *Diabetes.ca*  
<https://www.diabetes.ca/managing-my-diabetes/preventing-complications/kidney-disease>.
16. Zhang, R. Donor-Specific Antibodies in Kidney Transplant Recipients. *Clin. J. Am. Soc. Nephrol.* **13**, 182–192 (2018).
17. Halloran, P. F. Immunosuppressive drugs for kidney transplantation. *N. Engl. J. Med.* **351**, 2715–2729 (2004).
18. Eskandary, F., Wahrmann, M., Mühlbacher, J. & Böhmig, G. A. Complement inhibition as potential new therapy for antibody-mediated rejection. *Transpl. Int.* **29**, 392–402 (2016).
19. Venner, J. M., Hidalgo, L. G., Famulski, K. S., Chang, J. & Halloran, P. F. The molecular landscape of antibody-mediated kidney transplant rejection: evidence for NK involvement through CD16a Fc receptors. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **15**, 1336–1348 (2015).
20. Parkes, M. D., Halloran, P. F. & Hidalgo, L. G. Evidence for CD16a-Mediated NK Cell Stimulation in Antibody-Mediated Kidney Transplant Rejection. *Transplantation* **101**, e102–e111 (2017).
21. Parkes, M. D., Halloran, P. F. & Hidalgo, L. G. Mechanistic Sharing Between NK Cells in ABMR and Effector T Cells in TCMR. *Am. J. Transplant.* **18**, 63–73 (2018).
22. Loupy, A. *et al.* The Banff 2015 Kidney Meeting Report: Current Challenges in Rejection Classification and Prospects for Adopting Molecular Pathology. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **17**, 28–41 (2017).
23. Venner, J. M. *et al.* Molecular landscape of T cell-mediated rejection in human kidney transplants: prominence of CTLA4 and PD ligands. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **14**, 2565–2576 (2014).

## REFERENCES

24. Hidalgo, L. G. *et al.* The transcriptome of human cytotoxic T cells: measuring the burden of CTL-associated transcripts in human kidney transplants. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **8**, 637–646 (2008).
25. Halloran, P. F. T cell-mediated rejection of kidney transplants: a personal viewpoint. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **10**, 1126–1134 (2010).
26. Sellarés, J. *et al.* Inflammation lesions in kidney transplant biopsies: association with survival is due to the underlying diseases. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **11**, 489–499 (2011).
27. Kosieradzki, M. & Rowiński, W. Ischemia/Reperfusion Injury in Kidney Transplantation: Mechanisms and Prevention. *Transplant. Proc.* **40**, 3279–3288 (2008).
28. Jaeschke, H. & Woolbright, B. L. Current strategies to minimize hepatic ischemia–reperfusion injury by targeting reactive oxygen species. *Transplant. Rev.* **26**, 103–114 (2012).
29. Jaworska, K. *et al.* Both PD-1 ligands protect the kidney from ischemia reperfusion injury. *J. Immunol. Baltim. Md 1950* **194**, 325–333 (2015).
30. Ko, D. S. C. Reducing Ischemia-Reperfusion Injury in Renal Transplantation. *J. Urol.* **194**, 1531–1532 (2015).
31. Gong, L. & Wang, K. [Research progress on preventions of ischemia-reperfusion injury during kidney transplant]. *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi J. Biomed. Eng. Shengwu Yixue Gongchengxue Zazhi* **35**, 817–821 (2018).
32. Haas, M. *et al.* The Banff 2017 Kidney Meeting Report: Revised diagnostic criteria for chronic active T cell-mediated rejection, antibody-mediated rejection, and prospects for integrative endpoints for next-generation clinical trials. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **18**, 293–307 (2018).
33. Racusen, L. C. *et al.* The Banff 97 working classification of renal allograft pathology. *Kidney Int.* **55**, 713–723 (1999).
34. Salazar, I. D. R., López, M. M., Chang, J. & Halloran, P. F. Reassessing the Significance of Intimal Arteritis in Kidney Transplant Biopsy Specimens. *J. Am. Soc. Nephrol.* **26**, 3190–3198 (2015).

## REFERENCES

35. Reeve, J. *et al.* Assessing rejection-related disease in kidney transplant biopsies based on archetypal analysis of molecular phenotypes. *JCI Insight* **2**, (2017).
36. Reeve, J. *et al.* Molecular diagnosis of T cell-mediated rejection in human kidney transplant biopsies. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **13**, 645–655 (2013).
37. Iványi, B., Marcussen, N. & Olsen, S. Tubulitis in primary vascular and glomerular renal disease. *Pathol. Res. Pract.* **191**, 1245–1257 (1995).
38. Berden, A. E. *et al.* Tubular lesions predict renal outcome in antineutrophil cytoplasmic antibody-associated glomerulonephritis after rituximab therapy. *J. Am. Soc. Nephrol. JASN* **23**, 313–321 (2012).
39. Halloran, P. F., Famulski, K. S. & Chang, J. A Probabilistic Approach to Histologic Diagnosis of Antibody-Mediated Rejection in Kidney Transplant Biopsies. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **17**, 129–139 (2017).
40. Reeve, J. & Halloran, P. F. Antibody-Mediated Rejection Can Be Assessed by Histology When DSA Is Unknown or Negative. *Am J Transpl.* **16**, (suppl 3) (2016).
41. Magil, A. *et al.* Renal Biopsy in Acute Allograft Rejection. *Nephron* **26**, 180–183 (1980).
42. Sellarés, J. *et al.* Molecular diagnosis of antibody-mediated rejection in human kidney transplants. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **13**, 971–983 (2013).
43. Traynor, J., Mactier, R., Geddes, C. C. & Fox, J. G. How to measure renal function in clinical practice. *BMJ* **333**, 733–737 (2006).
44. Peake, M. & Whiting, M. Measurement of serum creatinine—current status and future goals. *Clin. Biochem. Rev.* **27**, 173–184 (2006).
45. Delanaye, P., Cavalier, E. & Pottel, H. Serum Creatinine: Not So Simple! *Nephron* **136**, 302–308 (2017).
46. Mula-Abed, W.-A. S., Al Rasadi, K. & Al-Riyami, D. Estimated Glomerular Filtration Rate (eGFR): A Serum Creatinine-Based Test for the Detection of Chronic Kidney Disease and its Impact on Clinical Practice. *Oman Med. J.* **27**, 108–113 (2012).

## REFERENCES

47. Sarwal, M. *et al.* Molecular Heterogeneity in Acute Renal Allograft Rejection Identified by DNA Microarray Profiling. *N. Engl. J. Med.* **349**, 125–138 (2003).
48. Naesens, M. *et al.* Progressive histological damage in renal allografts is associated with expression of innate and adaptive immunity genes. *Kidney Int.* **80**, 1364–1376 (2011).
49. Park, W. D., Griffin, M. D., Cornell, L. D., Cosio, F. G. & Stegall, M. D. Fibrosis with inflammation at one year predicts transplant functional decline. *J. Am. Soc. Nephrol. JASN* **21**, 1987–1997 (2010).
50. Vitalone, M. J. *et al.* Transcriptome changes of chronic tubulointerstitial damage in early kidney transplantation. *Transplantation* **89**, 537–547 (2010).
51. Dosanjh, A. *et al.* Genomic meta-analysis of growth factor and integrin pathways in chronic kidney transplant injury. *BMC Genomics* **14**, 275 (2013).
52. Einecke, G. *et al.* A molecular classifier for predicting future graft loss in late kidney transplant biopsies. *J. Clin. Invest.* **120**, 1862–1872 (2010).
53. Cippà, P. E. *et al.* Transcriptional trajectories of human kidney injury progression. *JCI Insight* **3**, (2018).
54. Ong, S. & Mannon, R. B. Genomic and proteomic fingerprints of acute rejection in peripheral blood and urine. *Transplant. Rev. Orlando Fla* **29**, 60–67 (2015).
55. Lee, J. R. *et al.* Urinary cell mRNA profiles predictive of human kidney allograft status. *Immunol. Rev.* **258**, 218–240 (2014).
56. Roedder, S. *et al.* The kSORT assay to detect renal transplant patients at high risk for acute rejection: results of the multicenter AART study. *PLoS Med.* **11**, e1001759 (2014).
57. Loupy, A. *et al.* Molecular microscope strategy to improve risk stratification in early antibody-mediated kidney allograft rejection. *J. Am. Soc. Nephrol. JASN* **25**, 2267–2277 (2014).
58. Sis, B. *et al.* Endothelial gene expression in kidney transplants with alloantibody indicates antibody-mediated damage despite lack of C4d staining. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **9**, 2312–2323 (2009).
59. O’Riordan, E. *et al.* Bioinformatic analysis of the urine proteome of acute allograft rejection. *J. Am. Soc. Nephrol. JASN* **15**, 3240–3248 (2004).

## REFERENCES

60. Kienzl-Wagner, K., Pratschke, J. & Brandacher, G. Proteomics--a blessing or a curse? Application of proteomics technology to transplant medicine. *Transplantation* **92**, 499–509 (2011).
61. Bohra, R. *et al.* Proteomics and metabolomics in renal transplantation-quo vadis? *Transpl. Int. Off. J. Eur. Soc. Organ Transplant.* **26**, 225–241 (2013).
62. Ling, X. B. *et al.* Integrative urinary peptidomics in renal transplantation identifies biomarkers for acute rejection. *J. Am. Soc. Nephrol. JASN* **21**, 646–653 (2010).
63. Wadström, J. *et al.* Advancing Transplantation: New Questions, New Possibilities in Kidney and Liver Transplantation. *Transplantation* **101 Suppl 2S**, S1–S41 (2017).
64. Madill-Thomsen, K. S., Wiggins, R. C., Eskandary, F., Böhmig, G. A. & Halloran, P. F. The Effect of Cortex/Medulla Proportions on Molecular Diagnoses in Kidney Transplant Biopsies: Rejection and Injury Can Be Assessed in Medulla. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **17**, 2117–2128 (2017).
65. Einecke, G. *et al.* Expression of B cell and immunoglobulin transcripts is a feature of inflammation in late allografts. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **8**, 1434–1443 (2008).
66. Halloran, P. F. *et al.* Real Time Central Assessment of Kidney Transplant Indication Biopsies by Microarrays: The INTERCOMEX Study. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **17**, 2851–2862 (2017).
67. Halloran, P. F., Reeve, J. P., Pereira, A. B., Hidalgo, L. G. & Famulski, K. S. Antibody-mediated rejection, T cell-mediated rejection, and the injury-repair response: new insights from the Genome Canada studies of kidney transplant biopsies. *Kidney Int.* **85**, 258–264 (2014).
68. Jameson, J. L. & Longo, D. L. Precision medicine--personalized, problematic, and promising. *N. Engl. J. Med.* **372**, 2229–2234 (2015).
69. Eskandary, F. *et al.* A Randomized Trial of Bortezomib in Late Antibody-Mediated Kidney Transplant Rejection. *J. Am. Soc. Nephrol. JASN* **29**, 591–605 (2018).
70. Nankivell, B. J. & Alexander, S. I. Rejection of the kidney allograft. *N. Engl. J. Med.* **363**, 1451–1462 (2010).

## REFERENCES

71. Halloran, P. F. *et al.* Disappearance of T Cell-Mediated Rejection Despite Continued Antibody-Mediated Rejection in Late Kidney Transplant Recipients. *J. Am. Soc. Nephrol.* **26**, 1711–1720 (2015).
72. Einecke, G., Reeve, J. & Halloran, P. F. TCMR Is Underdiagnosed in BK Virus Nephropathy: A Tale of 2 Populations [abstract]. *Am J Transpl.* **13**, (suppl 5) (2013).
73. Gupta, G., Reeve, J. & Halloran, P. F. The Complexity of T Cell-Mediated Rejection Scenarios: Identifying BK in Biopsies with TCMR [abstract]. *Am J Transpl.* **19**, (suppl 3) (2019).
74. Drachenberg, C. B. *et al.* Histological Evolution of BK Virus-Associated Nephropathy: Importance of Integrating Clinical and Pathological Findings. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **17**, 2078–2091 (2017).
75. Shier, D., Butler, J. & Lewis, R. *Hole's essentials of human anatomy & physiology.* (McGraw-Hill Education, 2015).
76. Hart, A. *et al.* Predicting Outcomes on the Liver Transplant Waiting List in the United States: Accounting for Large Regional Variation in Organ Availability and Priority Allocation Points. *Transplantation* **100**, 2153–2159 (2016).
77. Kim, W. R. *et al.* OPTN/SRTR 2016 Annual Data Report: Liver. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **18 Suppl 1**, 172–253 (2018).
78. Feng, S. & Bucuvalas, J. Tolerance after liver transplantation: Where are we?: Feng and Bucuvalas. *Liver Transpl.* **23**, 1601–1614 (2017).
79. Demetris, A. J. *et al.* 2016 Comprehensive Update of the Banff Working Group on Liver Allograft Pathology: Introduction of Antibody-Mediated Rejection. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **16**, 2816–2835 (2016).
80. Banff Working Group *et al.* Liver biopsy interpretation for causes of late liver allograft dysfunction. *Hepatol. Baltim. Md* **44**, 489–501 (2006).
81. Banff Working Group on Liver Allograft Pathology. Importance of liver biopsy findings in immunosuppression management: biopsy monitoring and working criteria for patients with operational tolerance. *Liver Transplant. Off. Publ. Am. Assoc. Study Liver Dis. Int. Liver Transplant. Soc.* **18**, 1154–1170 (2012).

## REFERENCES

82. Lee, M. Antibody-Mediated Rejection After Liver Transplant. *Gastroenterol. Clin. North Am.* **46**, 297–309 (2017).
83. Taner, T., Stegall, M. D. & Heimbach, J. K. Antibody-mediated rejection in liver transplantation: current controversies and future directions. *Liver Transplant. Off. Publ. Am. Assoc. Study Liver Dis. Int. Liver Transplant. Soc.* **20**, 514–527 (2014).
84. Dao, M. *et al.* Morphological characterization of chronic antibody-mediated rejection in ABO-identical or ABO-compatible pediatric liver graft recipients. *Liver Transplant. Off. Publ. Am. Assoc. Study Liver Dis. Int. Liver Transplant. Soc.* **24**, 897–907 (2018).
85. Olthoff, K. M. *et al.* Liver regeneration after living donor transplantation: Adult-to-adult living donor liver transplantation cohort study: ADULT-TO-ADULT LDLT COHORT STUDY. *Liver Transpl.* **21**, 79–88 (2015).
86. Sotiropoulos, G. C. *et al.* Long-term follow-up after right hepatectomy for adult living donation and attitudes toward the procedure. *Ann. Surg.* **254**, 694–700; discussion 700-701 (2011).
87. Kupiec-Weglinski, J. W. & Busuttil, R. W. Ischemia and reperfusion injury in liver transplantation. *Transplant. Proc.* **37**, 1653–1656 (2005).
88. Urrunaga, N. H., Rachakonda, V. P., Magder, L. S. & Mindikoglu, A. L. Outcomes of living versus deceased donor liver transplantation for acute liver failure in the United States. *Transplant. Proc.* **46**, 219–224 (2014).
89. Fedoravicius, A. & Charlton, M. Abnormal liver tests after liver transplantation: Abnormal Liver Tests After Liver Transplantation. *Clin. Liver Dis.* **7**, 73–79 (2016).
90. Šeda, O. *et al.* Hepatic Gene Expression Profiles Differentiate Steatotic and Non-steatotic Grafts in Liver Transplant Recipients. *Front. Endocrinol.* **10**, 270 (2019).
91. Mofrad, P. *et al.* Clinical and histologic spectrum of nonalcoholic fatty liver disease associated with normal ALT values. *Hepatol. Baltim. Md* **37**, 1286–1292 (2003).
92. Sorrentino, P. *et al.* Silent non-alcoholic fatty liver disease—a clinical-histological study. *J. Hepatol.* **41**, 751–757 (2004).
93. Bzowej, N. H. Nonalcoholic steatohepatitis: the new frontier for liver transplantation. *Curr. Opin. Organ Transplant.* **23**, 169–174 (2018).

## REFERENCES

94. Forbes, S. J. & Newsome, P. N. Liver regeneration - mechanisms and models to clinical application. *Nat. Rev. Gastroenterol. Hepatol.* **13**, 473–485 (2016).
95. Hashmi, S. K., Baranov, E., Gonzalez, A., Olthoff, K. & Shaked, A. Genomics of liver transplant injury and regeneration. *Transplant. Rev. Orlando Fla* **29**, 23–32 (2015).
96. Olthoff, K. M. Molecular pathways of regeneration and repair after liver transplantation. *World J. Surg.* **26**, 831–837 (2002).
97. Mao, S. A., Glorioso, J. M. & Nyberg, S. L. Liver regeneration. *Transl. Res. J. Lab. Clin. Med.* **163**, 352–362 (2014).
98. BC Cancer. Lab Test Interpretation Table.
99. Cleveland Clinic Center for Continuing Education. Liver Test Interpretation - Approach to the Patient with Liver Disease: A guide to Commonly Used Liver Tests.  
<http://www.clevelandclinicmeded.com/medicalpubs/diseasemanagement/hepatology/guide-to-common-liver-tests/>.
100. U.S. Department of Veterans Affairs. Tests of Liver Function.  
<https://www.hepatitis.va.gov/HEPATITIS/course/index.asp?page=/provider/courses/livertests/livertests-11&>.
101. Clark, J. M., Brancati, F. L. & Diehl, A. M. E. Nonalcoholic fatty liver disease: The most common cause of abnormal liver enzymes in the U.S. population. *Gastroenterology* **120**, A65 (2001).
102. Chalasani, N. *et al.* The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. *Hepatol. Baltim. Md* **55**, 2005–2023 (2012).
103. Dyson, J. K., Anstee, Q. M. & McPherson, S. Non-alcoholic fatty liver disease: a practical approach to diagnosis and staging. *Frontline Gastroenterol.* **5**, 211–218 (2014).
104. Sattar, N., Forrest, E. & Preiss, D. Non-alcoholic fatty liver disease. *BMJ* **349**, g4596 (2014).
105. McPherson, S., Stewart, S. F., Henderson, E., Burt, A. D. & Day, C. P. Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. *Gut* **59**, 1265–1269 (2010).

## REFERENCES

106. Perkins, J. D. Saying 'Yes' to obese living liver donors: short-term intensive treatment for donors with hepatic steatosis in living-donor liver transplantation. *Liver Transplant. Off. Publ. Am. Assoc. Study Liver Dis. Int. Liver Transplant. Soc.* **12**, 1012–1013 (2006).
107. Spitzer, A. L. *et al.* The biopsied donor liver: incorporating macrosteatosis into high-risk donor assessment. *Liver Transplant. Off. Publ. Am. Assoc. Study Liver Dis. Int. Liver Transplant. Soc.* **16**, 874–884 (2010).
108. Vilar-Gomez, E. *et al.* Weight Loss Through Lifestyle Modification Significantly Reduces Features of Nonalcoholic Steatohepatitis. *Gastroenterology* **149**, 367-378.e5; quiz e14-15 (2015).
109. Siddiqui, M. S. & Charlton, M. Liver Transplantation for Alcoholic and Nonalcoholic Fatty Liver Disease: Pretransplant Selection and Posttransplant Management. *Gastroenterology* **150**, 1849–1862 (2016).
110. Regev, A. *et al.* Reliability of histopathologic assessment for the differentiation of recurrent hepatitis C from acute rejection after liver transplantation. *Liver Transplant. Off. Publ. Am. Assoc. Study Liver Dis. Int. Liver Transplant. Soc.* **10**, 1233–1239 (2004).
111. Netto, G. J. *et al.* Interobserver agreement in hepatitis C grading and staging and in the Banff grading schema for acute cellular rejection: the 'hepatitis C 3' multi-institutional trial experience. *Arch. Pathol. Lab. Med.* **130**, 1157–1162 (2006).
112. Banff schema for grading liver allograft rejection: an international consensus document. *Hepatology. Baltim. Md* **25**, 658–663 (1997).
113. Feng, S. *et al.* Evidence of Chronic Allograft Injury in Liver Biopsies From Long-term Pediatric Recipients of Liver Transplants. *Gastroenterology* **155**, 1838-1851.e7 (2018).
114. Londoño, M.-C. *et al.* Molecular profiling of subclinical inflammatory lesions in long-term surviving adult liver transplant recipients. *J. Hepatol.* **69**, 626–634 (2018).
115. Conti, A. *et al.* Wide gene expression profiling of ischemia-reperfusion injury in human liver transplantation. *Liver Transpl.* **13**, 99–113 (2007).
116. de Jonge, J. *et al.* Unique early gene expression patterns in human adult-to-adult living donor liver grafts compared to deceased donor grafts. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **9**, 758–772 (2009).

## REFERENCES

117. Mastoridis, S., Martínez-Llordella, M. & Sanchez-Fueyo, A. Emergent Transcriptomic Technologies and Their Role in the Discovery of Biomarkers of Liver Transplant Tolerance. *Front. Immunol.* **6**, 304 (2015).
118. Vionnet, J. & Sánchez-Fueyo, A. Biomarkers of immune tolerance in liver transplantation. *Hum. Immunol.* **79**, 388–394 (2018).
119. McCaughan, G. W. *et al.* Molecular pathogenesis of liver disease: an approach to hepatic inflammation, cirrhosis and liver transplant tolerance. *Immunol. Rev.* **174**, 172–191 (2000).
120. Okubo, K. *et al.* Identification of Novel and Noninvasive Biomarkers of Acute Cellular Rejection After Liver Transplantation by Protein Microarray. *Transplant. Direct* **2**, e118 (2016).
121. Inkinen, K. *et al.* DNA microarray-based gene expression profiles of cytomegalovirus infection and acute rejection in liver transplants. *Transplant. Proc.* **37**, 1227–1229 (2005).
122. Muffak-Granero, K. *et al.* Study of gene expression profile in liver transplant recipients with hepatitis C virus. *Transplant. Proc.* **40**, 2971–2974 (2008).
123. Lozano, J. J. *et al.* Comparison of transcriptional and blood cell-phenotypic markers between operationally tolerant liver and kidney recipients. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **11**, 1916–1926 (2011).
124. Spivey, T. L. *et al.* Gene expression profiling in acute allograft rejection: challenging the immunologic constant of rejection hypothesis. *J. Transl. Med.* **9**, 174 (2011).
125. Fisher, L. R., Henley, K. S. & Lucey, M. R. Acute cellular rejection after liver transplantation: variability, morbidity, and mortality. *Liver Transplant. Surg. Off. Publ. Am. Assoc. Study Liver Dis. Int. Liver Transplant. Soc.* **1**, 10–15 (1995).
126. Arcasoy, S. M. *et al.* Pathologic interpretation of transbronchial biopsy for acute rejection of lung allograft is highly variable. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **11**, 320–328 (2011).
127. Younossi, Z. M. *et al.* Agreement in pathologic interpretation of liver biopsy specimens in posttransplant hepatitis C infection. *Arch. Pathol. Lab. Med.* **123**, 143–145 (1999).
128. Pournik, O. *et al.* Inter-observer and Intra-observer Agreement in Pathological Evaluation of Non-alcoholic Fatty Liver Disease Suspected Liver Biopsies. *Hepat. Mon.* **14**, e15167 (2014).

## REFERENCES

129. Wiesner, R. H. Is hepatic histology the true gold standard in diagnosing acute hepatic allograft rejection? *Liver Transplant. Surg. Off. Publ. Am. Assoc. Study Liver Dis. Int. Liver Transplant. Soc.* **2**, 165–167 (1996).
130. Clavien, P.-A., Muller, X., de Oliveira, M. L., Dutkowski, P. & Sanchez-Fueyo, A. Can immunosuppression be stopped after liver transplantation? *Lancet Gastroenterol. Hepatol.* **2**, 531–537 (2017).
131. Ascha, M. S., Ascha, M. L. & Hanouneh, I. A. Management of immunosuppressant agents following liver transplantation: Less is more. *World J. Hepatol.* **8**, 148–161 (2016).
132. Terasaki, P. I. Tolerogenic Mechanisms in Liver Transplantation. *SOJ Immunol.* **3**, 01–13 (2015).
133. Moini, M., Schilsky, M. L. & Tichy, E. M. Review on immunosuppression in liver transplantation. *World J. Hepatol.* **7**, 1355–1368 (2015).
134. Lerut, J. & Sanchez-Fueyo, A. An appraisal of tolerance in liver transplantation. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **6**, 1774–1780 (2006).
135. U.S. National Library of Medicine, NIH. Ulcerative Colitis.  
<https://ghr.nlm.nih.gov/condition/ulcerative-colitis>.
136. Ko, J. K. & Auyeung, K. K. Inflammatory bowel disease: etiology, pathogenesis and current therapy. *Curr. Pharm. Des.* **20**, 1082–1096 (2014).
137. Mayer, L. & Eisenhardt, D. Lack of induction of suppressor T cells by intestinal epithelial cells from patients with inflammatory bowel disease. *J. Clin. Invest.* **86**, 1255–1260 (1990).
138. Bertha, M., Bellaguara, E., Kuzel, T. & Hanauer, S. Checkpoint Inhibitor-Induced Colitis: A New Type of Inflammatory Bowel Disease? *ACG Case Rep. J.* **4**, e112 (2017).
139. Geukes Foppen, M. H. *et al.* Immune checkpoint inhibition-related colitis: symptoms, endoscopic features, histology and response to management. *ESMO Open* **3**, e000278 (2018).
140. Ahmad, T., Marshall, S.-E. & Jewell, D. Genetics of inflammatory bowel disease: the role of the HLA complex. *World J. Gastroenterol.* **12**, 3628–3635 (2006).
141. McGovern, D. P. B. *et al.* Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nat. Genet.* **42**, 332–337 (2010).

## REFERENCES

142. Kappeler, A. & Mueller, C. The role of activated cytotoxic T cells in inflammatory bowel disease. *Histol. Histopathol.* **15**, 167–172 (2000).
143. Müller, S. *et al.* Activated CD4+ and CD8+ cytotoxic cells are present in increased numbers in the intestinal mucosa from patients with active inflammatory bowel disease. *Am. J. Pathol.* **152**, 261–268 (1998).
144. Bergan, S. *et al.* Monitored high-dose azathioprine treatment reduces acute rejection episodes after renal transplantation. *Transplantation* **66**, 334–339 (1998).
145. Fraser, A. G., Orchard, T. R. & Jewell, D. P. The efficacy of azathioprine for the treatment of inflammatory bowel disease: a 30 year review. *Gut* **50**, 485–489 (2002).
146. Timmer, A., Patton, P. H., Chande, N., McDonald, J. W. D. & MacDonald, J. K. Azathioprine and 6-mercaptopurine for maintenance of remission in ulcerative colitis. *Cochrane Database Syst. Rev.* CD000478 (2016) doi:10.1002/14651858.CD000478.pub4.
147. Ito, A., Iizuka, B., Omori, T., Nakamura, S. & Tokushige, K. Tacrolimus for Remission Induction and Maintenance Therapy in Patients with Ulcerative Colitis: A Retrospective Evaluation Study. *Gastroenterol. Res. Pract.* **2016**, 5956316 (2016).
148. Matsuoka, K. *et al.* Tacrolimus for the Treatment of Ulcerative Colitis. *Intest. Res.* **13**, 219–226 (2015).
149. Lichtiger, S. *et al.* Cyclosporine in severe ulcerative colitis refractory to steroid therapy. *N. Engl. J. Med.* **330**, 1841–1845 (1994).
150. Kornbluth, A. Cyclosporine versus infliximab for the treatment of severe ulcerative colitis. *Gastroenterol. Hepatol.* **7**, 677–679 (2011).
151. Wu, D., Yang, Z., Zhao, C. & Yao, L. Infliximab versus cyclosporine for severe ulcerative colitis refractory to steroids: A protocol for systematic review and meta-analysis. *Medicine (Baltimore)* **97**, e12657 (2018).
152. Loftus, C. G., Loftus, E. V. & Sandborn, W. J. Cyclosporin for refractory ulcerative colitis. *Gut* **52**, 172–173 (2003).
153. Zenlea, T. & Peppercorn, M. A. Immunosuppressive therapies for inflammatory bowel disease. *World J. Gastroenterol.* **20**, 3146–3152 (2014).

## REFERENCES

154. Sandborn, W. J. *et al.* Abatacept for Crohn's disease and ulcerative colitis. *Gastroenterology* **143**, 62-69.e4 (2012).
155. Hardinger, K. L., Sunderland, D. & Wiederrich, J. A. Belatacept for the prophylaxis of organ rejection in kidney transplant patients: an evidence-based review of its place in therapy. *Int. J. Nephrol. Renov. Dis.* **9**, 139–150 (2016).
156. Buch, M. H., Vital, E. M. & Emery, P. Abatacept in the treatment of rheumatoid arthritis. *Arthritis Res. Ther.* **10**, S5 (2008).
157. Paine, E. R. Colonoscopic evaluation in ulcerative colitis. *Gastroenterol. Rep.* **2**, 161–168 (2014).
158. Andoh, A. *et al.* Increased expression of decay-accelerating factor (CD55) in the inflamed mucosa of patients with ulcerative colitis. *Pathophysiology* **5**, 105–110 (1998).
159. Ozen, A. *et al.* CD55 Deficiency, Early-Onset Protein-Losing Enteropathy, and Thrombosis. *N. Engl. J. Med.* **377**, 52–61 (2017).
160. Leach, S. T. *et al.* Serum and mucosal S100 proteins, calprotectin (S100A8/S100A9) and S100A12, are elevated at diagnosis in children with inflammatory bowel disease. *Scand. J. Gastroenterol.* **42**, 1321–1331 (2007).
161. Foell, D. *et al.* Phagocyte-specific S100 proteins are released from affected mucosa and promote immune responses during inflammatory bowel disease. *J. Pathol.* **216**, 183–192 (2008).
162. Montero-Meléndez, T., Llor, X., García-Planella, E., Perretti, M. & Suárez, A. Identification of novel predictor classifiers for inflammatory bowel disease by gene expression profiling. *PLoS One* **8**, e76235 (2013).
163. Planell, N. *et al.* Transcriptional analysis of the intestinal mucosa of patients with ulcerative colitis in remission reveals lasting epithelial cell alterations. *Gut* **62**, 967–976 (2013).
164. Bjerrum, J. T., Hansen, M., Olsen, J. & Nielsen, O. H. Genome-wide gene expression analysis of mucosal colonic biopsies and isolated colonocytes suggests a continuous inflammatory state in the lamina propria of patients with quiescent ulcerative colitis. *Inflamm. Bowel Dis.* **16**, 999–1007 (2010).
165. Smith, P. J. *et al.* Mucosal transcriptomics implicates under expression of BRINP3 in the pathogenesis of ulcerative colitis. *Inflamm. Bowel Dis.* **20**, 1802–1812 (2014).

## REFERENCES

166. Noble, C. L. *et al.* Regional variation in gene expression in the healthy colon is dysregulated in ulcerative colitis. *Gut* **57**, 1398–1405 (2008).
167. Arijis, I. *et al.* Mucosal gene signatures to predict response to infliximab in patients with ulcerative colitis. *Gut* **58**, 1612–1619 (2009).
168. Taman, H. *et al.* Transcriptomic Landscape of Treatment-Naïve Ulcerative Colitis. *J. Crohns Colitis* **12**, 327–336 (2018).
169. Holgersen, K. *et al.* High-resolution gene expression profiling using RNA sequencing in patients with inflammatory bowel disease and in mouse models of colitis. *J. Crohns Colitis* **9**, 492–506 (2015).
170. Venner, J. M. *et al.* Molecular landscape of T cell-mediated rejection in human kidney transplants: prominence of CTLA4 and PD ligands. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **14**, 2565–2576 (2014).
171. Halloran, P. F. *et al.* Building a tissue-based molecular diagnostic system in heart transplant rejection: The heart Molecular Microscope Diagnostic (MMDx) System. *J. Heart Lung Transplant. Off. Publ. Int. Soc. Heart Transplant.* **36**, 1192–1200 (2017).
172. Halloran, B. *et al.* Molecular Patterns in Human Ulcerative Colitis and Correlation with Response to Infliximab: *Inflamm. Bowel Dis.* **20**, 2353–2363 (2014).
173. Briscoe, E. & Feldman, J. Conceptual complexity and the bias/variance tradeoff. *Cognition* **118**, 2–16 (2011).
174. Diettrich, T. G. Ensemble Methods in Machine Learning. *Or. State Univ.* 1–15.
175. Che, D., Liu, Q., Rasheed, K. & Tao, X. Decision tree and ensemble learning algorithms with their applications in bioinformatics. *Adv. Exp. Med. Biol.* **696**, 191–199 (2011).
176. Oza, N. C. & Tumer, K. Classifier ensembles: Select real-world applications. *Inf. Fusion* **9**, 4–20 (2008).
177. Rokach, L. Ensemble Methods for Classifiers. in *Data Mining and Knowledge Discovery Handbook* (eds. Maimon, O. & Rokach, L.) 957–980 (Springer-Verlag, 2005). doi:10.1007/0-387-25465-X\_45.
178. Re, M. & Valentini, G. Ensemble methods: A review. in *Advances in Machine Learning and Data Mining for Astronomy* 3–27 (Chapman and Hall, 2001).

## REFERENCES

179. Opitz, D. & Maclin, R. Popular Ensemble Methods: An Empirical Study. *J. Artif. Intell. Res.* **11**, 169–198 (1999).
180. Levenson, R. M., Krupinski, E. A., Navarro, V. M. & Wasserman, E. A. Pigeons (*Columba livia*) as Trainable Observers of Pathology and Radiology Breast Cancer Images. *PLOS ONE* **10**, e0141357 (2015).
181. Adam, S. P., Alexandropoulos, S.-A. N., Pardalos, P. M. & Vrahatis, M. N. No Free Lunch Theorem: A Review. in *Approximation and Optimization : Algorithms, Complexity and Applications* (eds. Demetriou, I. C. & Pardalos, P. M.) 57–82 (Springer International Publishing, 2019). doi:10.1007/978-3-030-12767-1\_5.
182. Jain, B. P. Why is diagnosis not probabilistic in clinical-pathological conference (CPCs): Point. *Diagn. Berl. Ger.* **3**, 95–97 (2016).
183. Cahan, A. Diagnosis is driven by probabilistic reasoning: counter-point. *Diagn. Berl. Ger.* **3**, 99–101 (2016).
184. Doust, J. Diagnosis in General Practice. Using probabilistic reasoning. *BMJ* **339**, b3823 (2009).
185. Multiple authors. *R: a language and environment for statistical computing*. (R Foundation for Statistical Computing, 2018).
186. Luo, J. *et al.* A comparison of batch effect removal methods for enhancement of prediction performance using MAQC-II microarray gene expression data. *Pharmacogenomics J.* **10**, 278–291 (2010).
187. ATAGC team (various authors). Core PBT List (HG U219 arrays).
188. Famulski, K. S. *et al.* Defining the canonical form of T-cell-mediated rejection in human kidney transplants. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **10**, 810–820 (2010).
189. Halloran, P. F. *et al.* The molecular phenotype of kidney transplants. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **10**, 2215–2222 (2010).
190. Land, W. G., Agostinis, P., Gasser, S., Garg, A. D. & Linkermann, A. Transplantation and Damage-Associated Molecular Patterns (DAMPs). *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **16**, 3338–3361 (2016).

## REFERENCES

191. Hidalgo, L. G. *et al.* Interpreting NK cell transcripts versus T cell transcripts in renal transplant biopsies. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **12**, 1180–1191 (2012).
192. Einecke, G. *et al.* Antibody-mediated microcirculation injury is the major cause of late kidney transplant failure. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **9**, 2520–2531 (2009).
193. Famulski, K. S. *et al.* Changes in the transcriptome in allograft rejection: IFN-gamma-induced transcripts in mouse kidney allografts. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **6**, 1342–1354 (2006).
194. Famulski, K. S. *et al.* Transcriptome analysis reveals heterogeneity in the injury response of kidney transplants. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **7**, 2483–2495 (2007).
195. Famulski, K. S. *et al.* Molecular phenotypes of acute kidney injury in kidney transplants. *J. Am. Soc. Nephrol. JASN* **23**, 948–958 (2012).
196. Einecke, G., Broderick, G., Sis, B. & Halloran, P. F. Early loss of renal transcripts in kidney allografts: relationship to the development of histologic lesions and alloimmune effector mechanisms. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **7**, 1121–1130 (2007).
197. Mengel, M. *et al.* Molecular correlates of scarring in kidney transplants: the emergence of mast cell transcripts. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **9**, 169–178 (2009).
198. Halloran, P. F. *et al.* Exploring the cardiac response to injury in heart transplant biopsies. *JCI Insight* **3**, (2018).
199. Lê, S., Josse, J. & Husson, F. FactoMineR: AnRPackage for Multivariate Analysis. *J. Stat. Softw.* **25**, 18 (2008).
200. Eugster, M. & Leisch, F. From Spider-Man to Hero - Archetypal Analysis in R. *J. Stat. Softw.* **30**, 1–23 (2009).

## REFERENCES

201. Reeve, J. *et al.* Assessing rejection-related disease in kidney transplant biopsies based on archetypal analysis of molecular phenotypes. *JCI Insight* **2**, (2017).
202. Chang, K.-M., Harbron, C. & South, M. C. *RefPlus: A function set for the Extrapolation Strategy (RMA+) and Extrapolation Averaging (RMA++) methods.* (2018).
203. Huber, W. & Carey, V. J. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* **12**, 115–121 (2015).
204. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
205. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics J. Integr. Biol.* **16**, 284–287 (2012).
206. Huang, D. *et al.* The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol.* **8**, R183 (2007).
207. Zeileis, A. & Grothendieck, G. Zoo: S3 Infrastructure for Regular and Irregular Time Series. *J. Stat. Softw.* **14**, 1–27 (2005).
208. Younossi, Z. M. *et al.* Agreement in pathologic interpretation of liver biopsy specimens in posttransplant hepatitis C infection. *Arch. Pathol. Lab. Med.* **123**, 143–145 (1999).
209. Demetris, A. J. *et al.* 2016 Comprehensive Update of the Banff Working Group on Liver Allograft Pathology: Introduction of Antibody-Mediated Rejection. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **16**, 2816–2835 (2016).
210. DeRoche, T. C., Xiao, S.-Y. & Liu, X. Histological evaluation in ulcerative colitis. *Gastroenterol. Rep.* **2**, 178–192 (2014).
211. Halloran, P. F., Famulski, K. S. & Reeve, J. Molecular assessment of disease states in kidney transplant biopsy samples. *Nat. Rev. Nephrol.* **12**, 534–548 (2016).
212. Halloran, P. F. *et al.* Microarray diagnosis of antibody-mediated rejection in kidney transplant biopsies: an international prospective study (INTERCOM). *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **13**, 2865–2874 (2013).

## REFERENCES

213. Reeve, J. *et al.* Diagnosing rejection in renal transplants: a comparison of molecular- and histopathology-based approaches. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **9**, 1802–1810 (2009).
214. Haas, M. *et al.* Banff 2013 meeting report: inclusion of c4d-negative antibody-mediated rejection and antibody-associated arterial lesions. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **14**, 272–283 (2014).
215. Vivante, A. & Hildebrandt, F. Exploring the genetic basis of early-onset chronic kidney disease. *Nat. Rev. Nephrol.* **12**, 133–146 (2016).
216. Fukuda, A. *et al.* Urine podocin:nephrin mRNA ratio (PNR) as a podocyte stress biomarker. *Nephrol. Dial. Transplant. Off. Publ. Eur. Dial. Transpl. Assoc. - Eur. Ren. Assoc.* **27**, 4079–4087 (2012).
217. Einecke, G. *et al.* Loss of solute carriers in T cell-mediated rejection in mouse and human kidneys: an active epithelial injury-repair response. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **10**, 2241–2251 (2010).
218. Yang, Y. *et al.* The two kidney to one kidney transition and transplant glomerulopathy: a podocyte perspective. *J. Am. Soc. Nephrol. JASN* **26**, 1450–1465 (2015).
219. Timmermans, S. A. M. E. G. *et al.* Evaluation of anti-PLA2R1 as measured by a novel ELISA in patients with idiopathic membranous nephropathy: a cohort study. *Am. J. Clin. Pathol.* **142**, 29–34 (2014).
220. Brunskill, E. W. *et al.* Genes that confer the identity of the renin cell. *J. Am. Soc. Nephrol. JASN* **22**, 2213–2225 (2011).
221. Kerjaschki, D., Sharkey, D. J. & Farquhar, M. G. Identification and characterization of podocalyxin-- the major sialoprotein of the renal glomerular epithelial cell. *J. Cell Biol.* **98**, 1591–1596 (1984).
222. Tavasoli, M. *et al.* Both CLIC4 and CLIC5A activate ERM proteins in glomerular endothelium. *Am. J. Physiol. Renal Physiol.* **311**, F945–F957 (2016).
223. Mueller, T. F. *et al.* Microarray analysis of rejection in human kidney transplants using pathogenesis-based transcript sets. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **7**, 2712–2722 (2007).

## REFERENCES

224. Viklicky, O., Hribova, P. & Brabcova, I. Molecular markers of rejection and tolerance: lessons from clinical research. *Nephrol. Dial. Transplant. Off. Publ. Eur. Dial. Transpl. Assoc. - Eur. Ren. Assoc.* **28**, 2701–2708 (2013).
225. Natarajan, N. I., Ravikumar, P. & Tewari, A. Learning with Noisy Labels. (2013).
226. Rolnick, D., Veit, A., Belongie, S. & Shavit, N. Deep Learning is Robust to Massive Label Noise. *CoRR abs/1705.10694*, (2017).
227. McAdam, A. J. Discrepant analysis: how can we test a test? *J. Clin. Microbiol.* **38**, 2027–2029 (2000).
228. Quinn, G. E. *et al.* Analysis of Discrepancy Between Diagnostic Clinical Examination Findings and Corresponding Evaluation of Digital Images in the Telemedicine Approaches to Evaluating Acute-Phase Retinopathy of Prematurity Study. *JAMA Ophthalmol.* **134**, 1263–1270 (2016).
229. Roulson, J., Benbow, E. W. & Hasleton, P. S. Discrepancies between clinical and autopsy diagnosis and the value of post mortem histology; a meta-analysis and review. *Histopathology* **47**, 551–559 (2005).
230. Woods, S. P., Lovejoy, D. W., Stutts, M. L., Ball, J. D. & Fals-Stewart, W. Comparative efficiency of a discrepancy analysis for the classification of Attention-Deficit/Hyperactivity Disorder in adults. *Arch. Clin. Neuropsychol. Off. J. Natl. Acad. Neuropsychol.* **17**, 351–369 (2002).
231. Revelle, W. *psych: Procedures for Psychological, Psychometric, and Personality Research.* Northwest. Univ. (2017).
232. Kurian, S. M. *et al.* Orthogonal Comparison of Molecular Signatures of Kidney Transplants With Subclinical and Clinical Acute Rejection: Equivalent Performance Is Agnostic to Both Technology and Platform. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **17**, 2103–2116 (2017).
233. Mehta, R., Sood, P. & Hariharan, S. Subclinical Rejection in Renal Transplantation: Reappraised. *Transplantation* **100**, 1610–1618 (2016).
234. Loupy, A. *et al.* Outcome of subclinical antibody-mediated rejection in kidney transplant recipients with preformed donor-specific antibodies. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **9**, 2561–2570 (2009).

## REFERENCES

235. Hoffman, W. *et al.* The Impact of Early Clinical and Subclinical T Cell-mediated Rejection After Kidney Transplantation. *Transplantation* **103**, 1457–1467 (2019).
236. Moreso, F. *et al.* Subclinical rejection associated with chronic allograft nephropathy in protocol biopsies as a risk factor for late graft loss. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **6**, 747–752 (2006).
237. Nankivell, B. J. & Chapman, J. R. The significance of subclinical rejection and the value of protocol biopsies. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **6**, 2006–2012 (2006).
238. Loupy, A. *et al.* Subclinical Rejection Phenotypes at 1 Year Post-Transplant and Outcome of Kidney Allografts. *J. Am. Soc. Nephrol. JASN* **26**, 1721–1731 (2015).
239. Kayser, D. *et al.* Reassessing Subclinical Rejection Using Molecular Features: Typical T Cell Mediated Rejection is Rare in Protocol Biopsies. *Am. J. Transplant.* **s2**, 305 (2009).
240. Dieplinger, G. *et al.* Onset and progression of de novo donor-specific anti-human leukocyte antigen antibodies after BK polyomavirus and preemptive immunosuppression reduction. *Transpl. Infect. Dis. Off. J. Transplant. Soc.* **17**, 848–858 (2015).
241. Jain, B. P. Response to paper on probabilistic diagnosis. *Diagnosis* **4**, 101–101 (2017).
242. Bröcker, V. *et al.* Histopathological and clinical findings in renal transplants with Banff type II and III acute cellular rejection without tubulointerstitial infiltrates. *Virchows Arch. Int. J. Pathol.* **464**, 203–211 (2014).
243. Lefaucheur, C. *et al.* T cell-mediated rejection is a major determinant of inflammation in scarred areas in kidney allografts. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **18**, 377–390 (2018).
244. Nankivell, B. J. *et al.* The causes, significance and consequences of inflammatory fibrosis in kidney transplantation: The Banff i-IFTA lesion. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **18**, 364–376 (2018).
245. Famulski, K. S. & Halloran, P. F. Letter to *AJT* editor re: Nankivell *et al.* *Am. J. Transplant.* **18**, 765–766 (2018).

## REFERENCES

246. Schinstock, C. A. *et al.* Banff survey on antibody-mediated rejection clinical practices in kidney transplantation: Diagnostic misinterpretation has potential therapeutic implications. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **19**, 123–131 (2019).
247. Ishwaran, H. & Kogalur, U. *Random Forests for Survival, Regression and Classification (RF-SRC)*. (2018).
248. Venner, J. M., Famulski, K. S., Reeve, J., Chang, J. & Halloran, P. F. Relationships among injury, fibrosis, and time in human kidney transplants. *JCI Insight* **1**, e85323 (2016).
249. Halloran, P. F. *et al.* Molecular phenotype of kidney transplant indication biopsies with inflammation in scarred areas. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* (2018) doi:10.1111/ajt.15178.
250. Senev, A. *et al.* Histological picture of antibody-mediated rejection without donor-specific anti-HLA antibodies: Clinical presentation and implications for outcome. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **19**, 763–780 (2019).
251. Kones, R., Rumana, U. & Morales-Salinas, A. Confronting the most challenging risk factor: non-adherence. *Lancet Lond. Engl.* **393**, 105–106 (2019).
252. Schroeder, T. J., Weiss, M. A., Smith, R. D., Stephens, G. W. & First, M. R. The efficacy of OKT3 in vascular rejection. *Transplantation* **51**, 312–315 (1991).
253. Bröcker, V. *et al.* Histopathological and clinical findings in renal transplants with Banff type II and III acute cellular rejection without tubulointerstitial infiltrates. *Virchows Arch. Int. J. Pathol.* **464**, 203–211 (2014).
254. Solez, K. *et al.* Banff 07 classification of renal allograft pathology: updates and future directions. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **8**, 753–760 (2008).
255. Neuberger, J. An update on liver transplantation: A critical review. *J. Autoimmun.* **66**, 51–59 (2016).
256. Zarrinpar, A. & Busuttil, R. W. Liver transplantation: past, present and future. *Nat. Rev. Gastroenterol. Hepatol.* **10**, 434–440 (2013).
257. Regev, A. *et al.* Reliability of histopathologic assessment for the differentiation of recurrent hepatitis C from acute rejection after liver transplantation. *Liver Transplant. Off. Publ. Am. Assoc. Study Liver Dis. Int. Liver Transplant. Soc.* **10**, 1233–1239 (2004).

## REFERENCES

258. Savage, T. M. *et al.* Deletion of donor-reactive T cell clones after human liver transplant. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* (2019) doi:10.1111/ajt.15592.
259. Levitsky, J. & Feng, S. Tolerance in clinical liver transplantation. *Hum. Immunol.* **79**, 283–287 (2018).
260. Thomson, A. W. & Knolle, P. A. Antigen-presenting cell function in the tolerogenic liver environment. *Nat. Rev. Immunol.* **10**, 753–766 (2010).
261. Reeve, J. *et al.* Molecular Diagnosis of Rejection Phenotypes in 889 Heart Transplant Biopsies: The INTERHEART Study. *J. Heart Lung Transplant.* **37**, S27 (2018).
262. Halloran, K. M. *et al.* Molecular assessment of rejection and injury in lung transplant biopsies. *J. Heart Lung Transplant. Off. Publ. Int. Soc. Heart Transplant.* **38**, 504–513 (2019).
263. Parkes, M., Halloran, K., Famulski, K. & Halloran, P. Molecular phenotypes of injury and rejection in lung transplant transbronchial biopsies. *Am. J. Transplant.* **18**, 274–275 (2018).
264. Halloran, K. *et al.* Molecular phenotyping of rejection-related changes in mucosal biopsies from lung transplants. *Am. J. Transplant.* *ajt.15685* (2019) doi:10.1111/ajt.15685.
265. Selzner, N., Grant, D. R., Shalev, I. & Levy, G. A. The immunosuppressive pipeline: meeting unmet needs in liver transplantation. *Liver Transplant. Off. Publ. Am. Assoc. Study Liver Dis. Int. Liver Transplant. Soc.* **16**, 1359–1372 (2010).
266. Einecke, G. *et al.* A molecular classifier for predicting future graft loss in late kidney transplant biopsies. *J. Clin. Invest.* **120**, 1862–1872 (2010).
267. Famulski, K. S. *et al.* Kidney Transplants With Progressing Chronic Diseases Express High Levels of Acute Kidney Injury Transcripts: AKI Signal in Transplants With CKD. *Am. J. Transplant.* **13**, 634–644 (2013).
268. G., E., J, R. & P, H. Predictors of Graft Survival at the Time of a Kidney Transplant Indication Biopsy. *Am J Transpl.* **17**, suppl 3 (2017).
269. Halloran, P. *et al.* Development and validation of a molecular microscope diagnostic system (MMDx) for heart transplant biopsies. *Am J Transpl.* **16**, (suppl 3) (2016).
270. Reeve, J. *et al.* Validating the INTERHEART Classifiers for Molecular Diagnosis of Rejection in 558 New Endomyocardial Biopsies. *J. Heart Lung Transplant.* **37**, S303–S304 (2018).

## REFERENCES

271. Parkes, M. *et al.* Molecular Assessment of Heart Transplant Biopsies: Emergence of the Injury Dimension. *Transplantation* **102**, S62–S63 (2018).
272. Halloran, K. *et al.* Molecular Detection of Rejection-like Changes in Proximal Bronchial Mucosal Lung Transplant Biopsies: Initial Findings of the INTERLUNG Study. *J. Heart Lung Transplant.* **37**, S80–S81 (2018).
273. Muffak-Granero, K. *et al.* Study of gene expression profile in liver transplant recipients with hepatitis C virus. *Transplant. Proc.* **40**, 2971–2974 (2008).
274. Šeda, O. *et al.* Hepatic Gene Expression Profiles Differentiate Steatotic and Non-steatotic Grafts in Liver Transplant Recipients. *Front. Endocrinol.* **10**, 270 (2019).
275. Londoño, M.-C. *et al.* Molecular profiling of subclinical inflammatory lesions in long-term surviving adult liver transplant recipients. *J. Hepatol.* **69**, 626–634 (2018).
276. Halloran, P. F. *et al.* Review: The transcripts associated with organ allograft rejection. *Am. J. Transplant.* **18**, 785–795 (2018).
277. Loupy, A. *et al.* Gene Expression Profiling for the Identification and Classification of Antibody-Mediated Heart Rejection. *Circulation* **135**, 917–935 (2017).
278. Halloran, P. F., Venner, J. M. & Famulski, K. S. Comprehensive Analysis of Transcript Changes Associated With Allograft Rejection: Combining Universal and Selective Features. *Am. J. Transplant.* **17**, 1754–1769 (2017).
279. Martínez-Llordella, M. *et al.* Multiparameter immune profiling of operational tolerance in liver transplantation. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **7**, 309–319 (2007).
280. Bonaccorsi-Riani, E. *et al.* Molecular Characterization of Acute Cellular Rejection Occurring During Intentional Immunosuppression Withdrawal in Liver Transplantation. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **16**, 484–496 (2016).
281. Mengel, M. *et al.* The molecular phenotype of heart transplant biopsies: relationship to histopathological and clinical variables. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **10**, 2105–2115 (2010).

## REFERENCES

282. Land, W. G., Agostinis, P., Gasser, S., Garg, A. D. & Linkermann, A. Transplantation and Damage-Associated Molecular Patterns (DAMPs). *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **16**, 3338–3361 (2016).
283. Thiele, C. *cutpointr: Determine and Evaluate Optimal Cutpoints in Binary Classification Tasks*. (2019).
284. Dogan, N. *et al.* Acute allograft rejection in liver transplant recipients: Incidence, risk factors, treatment success, and impact on graft failure. *J. Int. Med. Res.* **46**, 3979–3990 (2018).
285. Dar, W. A., Sullivan, E., Bynon, J. S., Eltzhig, H. & Ju, C. Ischaemia reperfusion injury in liver transplantation: Cellular and molecular mechanisms. *Liver Int. Off. J. Int. Assoc. Study Liver* **39**, 788–801 (2019).
286. Faggioli, F. *et al.* B lymphocytes limit senescence-driven fibrosis resolution and favor hepatocarcinogenesis in mouse liver injury. *Hepatology. Baltim. Md* **67**, 1970–1985 (2018).
287. Novobrantseva, T. I. *et al.* Attenuated liver fibrosis in the absence of B cells. *J. Clin. Invest.* **115**, 3072–3082 (2005).
288. Jucaud, V. *et al.* Prevalence and Impact of De Novo Donor-Specific Antibodies During a Multicenter Immunosuppression Withdrawal Trial in Adult Liver Transplant Recipients. *Hepatology. Baltim. Md* **69**, 1273–1286 (2019).
289. Kaneku, H. *et al.* Donor-specific human leukocyte antigen antibodies of the immunoglobulin G3 subclass are associated with chronic rejection and graft loss after liver transplantation. *Liver Transplant. Off. Publ. Am. Assoc. Study Liver Dis. Int. Liver Transplant. Soc.* **18**, 984–992 (2012).
290. O’Leary, J. G. *et al.* High mean fluorescence intensity donor-specific anti-HLA antibodies associated with chronic rejection Postliver transplant. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **11**, 1868–1876 (2011).
291. Shaked, A. *et al.* Outcomes of immunosuppression minimization and withdrawal early after liver transplantation. *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.* **19**, 1397–1409 (2019).

## REFERENCES

292. Crespo-Leiro, M. G. *et al.* Concordance among pathologists in the second Cardiac Allograft Rejection Gene Expression Observational Study (CARGO II). *Transplantation* **94**, 1172–1177 (2012).
293. Parkes, M. D. *et al.* An integrated molecular diagnostic report for heart transplant biopsies using an ensemble of diagnostic algorithms. *J. Heart Lung Transplant. Off. Publ. Int. Soc. Heart Transplant.* **38**, 636–646 (2019).
294. ATAGC. ATAGC Official Gene Lists. <https://www.ualberta.ca/medicine/institutes-centres-groups/atagc/research/gene-lists>.
295. Einecke, G. *et al.* The Tissue Response to Injury in Kidney Allografts: The Balance between Injury Pathways Determines Outcome after Injury. in 57–57 (2011).
296. Ramaiah, S. K. & Jaeschke, H. Role of neutrophils in the pathogenesis of acute inflammatory liver injury. *Toxicol. Pathol.* **35**, 757–766 (2007).
297. Ramaiah, S. K. & Jaeschke, H. Hepatic neutrophil infiltration in the pathogenesis of alcohol-induced liver injury. *Toxicol. Mech. Methods* **17**, 431–440 (2007).
298. Jaeschke, H. & Hasegawa, T. Role of neutrophils in acute inflammatory liver injury. *Liver Int. Off. J. Int. Assoc. Study Liver* **26**, 912–919 (2006).
299. Halloran, K. *et al.* Molecular Diagnosis of Rejection Phenotypes in Lung Transplant Biopsies: Initial Findings of the INTERLUNG Study. *J. Heart Lung Transplant.* **37**, S80 (2018).
300. Fiocchi, C. Inflammatory bowel disease: etiology and pathogenesis. *Gastroenterology* **115**, 182–205 (1998).
301. Prieux-Klotz, C. *et al.* Immune Checkpoint Inhibitor-Induced Colitis: Diagnosis and Management. *Target. Oncol.* **12**, 301–308 (2017).
302. Schiff, M. Abatacept treatment for rheumatoid arthritis. *Rheumatol. Oxf. Engl.* **50**, 437–449 (2011).
303. Affymetrix Inc. Affymetrix 3' IVT Express Kit Insert.
304. Lin, X. Y., Choi, M. S. & Porter, A. G. Expression analysis of the human caspase-1 subfamily reveals specific regulation of the CASP5 gene by lipopolysaccharide and interferon-gamma. *J. Biol. Chem.* **275**, 39920–39926 (2000).

## REFERENCES

305. Kanneganti, T.-D. Inflammatory Bowel Disease and the NLRP3 Inflammasome. *N. Engl. J. Med.* **377**, 694–696 (2017).
306. Naik, S. *et al.* Inflammatory memory sensitizes skin epithelial stem cells to tissue damage. *Nature* **550**, 475–480 (2017).
307. Dai, X. & Medzhitov, R. Memory beyond immunity. *Nature* **550**, 460–461 (2017).
308. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
309. Liu, J. Z. *et al.* Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat. Genet.* **47**, 979–986 (2015).
310. Feagins, L. A., Melton, S. D., Iqbal, R., Dunbar, K. B. & Spechler, S. J. Clinical implications of histologic abnormalities in colonic biopsy specimens from patients with ulcerative colitis in clinical remission. *Inflamm. Bowel Dis.* **19**, 1477–1482 (2013).
311. Hornung, V. *et al.* AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature* **458**, 514–518 (2009).
312. Harbour, S. N., Maynard, C. L., Zindl, C. L., Schoeb, T. R. & Weaver, C. T. Th17 cells give rise to Th1 cells that are required for the pathogenesis of colitis. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 7061–7066 (2015).
313. Lamb, C. A. *et al.*  $\alpha$ E $\beta$ 7 Integrin Identifies Subsets of Pro-Inflammatory Colonic CD4+ T Lymphocytes in Ulcerative Colitis. *J. Crohns Colitis* **11**, 610–620 (2017).
314. Croitoru, K. & Zhou, P. T-cell-induced mucosal damage in the intestine. *Curr. Opin. Gastroenterol.* **20**, 581–586 (2004).
315. Ungaro, R., Mehandru, S., Allen, P. B., Peyrin-Biroulet, L. & Colombel, J.-F. Ulcerative colitis. *Lancet Lond. Engl.* **389**, 1756–1770 (2017).
316. Moss, A. C. Optimizing the use of biological therapy in patients with inflammatory bowel disease. *Gastroenterol. Rep.* **3**, 63–68 (2015).
317. Kurolap, A. *et al.* Loss of CD55 in Eculizumab-Responsive Protein-Losing Enteropathy. *N. Engl. J. Med.* **377**, 87–89 (2017).

## REFERENCES

318. Ananthakrishnan, A. N. *et al.* Variation in treatment of patients with inflammatory bowel diseases at major referral centers in the United States. *Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc.* **13**, 1197–1200 (2015).
319. Rutgeerts, P. *et al.* Infliximab for Induction and Maintenance Therapy for Ulcerative Colitis. *N. Engl. J. Med.* **353**, 2462–2476 (2005).
320. Nielsen, O. H. & Ainsworth, M. A. Tumor Necrosis Factor Inhibitors for Inflammatory Bowel Disease. *N. Engl. J. Med.* **369**, 754–762 (2013).
321. Järnerot, G. *et al.* Infliximab as Rescue Therapy in Severe to Moderately Severe Ulcerative Colitis: A Randomized, Placebo-Controlled Study. *Gastroenterology* **128**, 1805–1811 (2005).
322. van der Valk, M. E. *et al.* Healthcare costs of inflammatory bowel disease have shifted from hospitalisation and surgery towards anti-TNF $\alpha$  therapy: results from the COIN study. *Gut* **63**, 72–79 (2014).
323. Dretzke, J. *et al.* A systematic review and economic evaluation of the use of tumour necrosis factor-alpha (TNF- $\alpha$ ) inhibitors, adalimumab and infliximab, for Crohn's disease. *Health Technol. Assess.* **15**, (2011).
324. Bitton, A. *et al.* Clinical, biological, and histologic parameters as predictors of relapse in ulcerative colitis. *Gastroenterology* **120**, 13–20 (2001).
325. Mohammed, N. & Subramanian, V. Clinical relevance of endoscopic assessment of inflammation in ulcerative colitis: Can endoscopic evaluation predict outcomes? *World J. Gastroenterol.* **22**, 9324–9332 (2016).