Identification and Characterisation of Genetic Lesions that Predispose to and Gene Expression Patterns that Contribute to Myeloid Malignancies

PARVATHY VENUGOPAL
Bachelor of Science (Zoology with Medical Microbiology)
Master of Science (Regenerative Medicine)

A thesis submitted for the degree of Doctor of Philosophy

School of Biological Sciences
Discipline of Genetics and Evolution

Faculty of Sciences
The University of Adelaide, South Australia

March 2016

Table of Contents

Abstract

Acute Myeloid Leukaemia (AML) is a heterogeneous disease caused by multiple genetic lesions. Our laboratory focuses on understanding the genetics of both inherited and acquired haematopoietic malignancies. In this thesis, I have investigated both inherited and acquired genetic changes that contribute to myeloid malignancies.

One of the key factors regulating haematopoiesis is GATA2, a zinc finger transcription factor. Germline mutations in *GATA2* have been associated with several clinical phenotypes such as myelodysplastic syndrome (MDS)/AML, immunodeficiency disorders (MonoMAC syndrome, DCML deficiency, congenital neutropenia, NK cell deficiency, aplastic anaemia) and Emberger syndrome. Moreover, several somatic mutations in *GATA2* have been reported in MDS/AML. Intriguingly, missense somatic and germline mutations reported to date are mutually exclusive, and several clinical phenotypes are associated with specific mutations. We generated a zinc finger 2 (ZF2) mutant allelic series representing a range of clinical phenotypes to investigate how each mutation effects transactivation, DNA binding, protein structure, protein partner interactions and *in vitro* differentiation. Specific GATA2 mutations perturb the interactions and functions in distinct ways that are beginning to explain differences in observed clinical phenotypes.

We performed gene expression analysis of 91 selected MDS/AML genes, including *GATA2*, on 166 well annotated primary AML samples, bone marrow mononuclear cells (BMMNC) and CD34 controls. Correlation analyses of *GATA2* expression levels with expression of other genes and other mutational and clinical data, was performed to help identify genetic aberrations that cooperate with abnormal levels of *GATA2* in AML.

Statistical correlations of expression levels of various other genes with outcome and mutation status were also identified.

One such correlation was reduced *GATA2* expression with oncogenic RAS mutations. A pilot study was carried out to evaluate and optimise an NRAS G12D-induced leukaemia model. All mice transplanted with mutant *NRAS* G12D rapidly developed haematopoietic disease post-transplantation whereas the control group did not. Based on these pilot studies, we have initiated transplantation experiments in a conditional *GATA2* knockout model to investigate the requirement of *GATA2* in NRAS G12D induced myeloid disease. Recipient mice continue to be monitored, but are yet to develop disease.

We also identified gene expression patterns of prognostic significance in AML and narrowed down a combination of three genes that are highly predictive of outcome. We devised a strategy integrating these genes into currently used risk stratification strategies and significantly improved risk stratification of AML patients at diagnosis.

Among syndromes that predispose to MDS/AML, is Diamond Blackfan Anaemia (DBA), a congenital disorder characterised by red blood cell deficiency. The underlying genetic cause of DBA in a child was identified using whole genome sequencing (WGS), targeted massively parallel sequencing (MPS) and high density SNP array. A complex scenario of germline and somatic aberrations were identified in two genetic loci that helped to explain the clinical features seen in the patient and the progression of this disease. These have led to the discovery of a mechanism by which spontaneous remissions occur in DBA patients.

Together, these studies have given us valuable insights into malignant myeloid disease biology and offer potential applications in improving therapeutic approaches in AML patients.

STATEMENT

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968. The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed,

Parvathy Venugopal

31 March 2016

Acknowledgements

I wish to thank my supervisors Hamish Scott, Chris Hahn and Manuela Klingler-Hoffmann for providing guidance and support throughout my PhD. Thank you for taking time to give me valuable feedback on my thesis in spite of being busy with deadlines and a million other things. The opportunity to work in the Department of Genetics and Molecular Pathology has been invaluable and I have learnt so much in my time here.

I would like to express my sincere gratitude to Hamish for giving me this opportunity and always encouraging me to step out of my comfort zone. I think that is some of the best advice I will ever get. I admire his ability to see the big picture and his brutal honesty which gets the best out of you. He is an incredible individual who is passionate about making a difference in people's lives through his work.

I'm extremely grateful to Chris for his invaluable assistance and encouragement over the last three years. Chris personifies patience, always making time and explaining things as many times and in as many ways as it would take for one to understand. I cannot emphasise enough on the integral role he has played in my PhD. Thank you Chris!

My sincere thanks to Manuela for helping me get started in the lab and giving me that first push I needed to ease into a work environment/culture that was completely new to me. Her role in helping me settle in my early days was immense and I am grateful for that.

I would like to thank all the Scott lab members- Milly, Pete (special thanks for all his help with looking after the mice), Anna, Chan-Eng, Young, Bergithe, Bradley, Jesse and Alicia for their friendship, company and support in all things scientific and non-scientific. Thank you Milly, Alicia and Jesse, for ensuring I didn't go hungry by buying me lunch or chocolate in both solid and liquid form.

Special thanks to Richard D'Andrea, Anna Brown, Mahmoud Bassal, Justine Marum, Jim Manavis, John Finnie, Andreas Schreiber, Joel Geoghegan, David Lawrence, Jinghua (Frank) Feng, Michele Grimbaldeston, Dave Yip and all our collaborators for their inputs and contributions to various aspects of my work. Thanks to the patients and their families for their willingness to participate in research without which my thesis wouldn't have been complete. Special thanks to Justine Marum and Ranjini Vaidyanathan for helping

me make sense of random forests and other complex statistical methods used in Chapter 4. All members of the Genetics and Molecular Pathology department, the ACRF genome facility, the animal facility and the Detmold facility deserve special mention for their help and their kind and encouraging words on some of my not-so-great days.

I am grateful to my parents for letting me chase my dreams. I would like to thank my father for being supportive of unconventional career (and partner) choices, my mother for being my super-cool pillar of support and my sister for backing me in all my decisions. I thank my family and friends for being proud of my achievements and I will be forever indebted to my grandfather for encouraging me to ask questions and introducing me to the fascinating world of science.

Most importantly, I would like to express my everlasting gratitude to my partner, Vishal, for his unconditional love and support over the last decade and for patiently putting up with me during these last few months. I fall short of words to express how valuable his contribution to this entire process has been. Thank you for keeping me sane!

List of Figures

List of Tables

Abbreviations

ADBA	Australian Diamond Blackfan Anaemia
AGM	Aorta-gonad mesonephros
ALT	Alanine transaminase
AML	Acute Myeloid Leukaemia
APS	Ammonium persulfate
BFU-E	Burst forming unit-erythroid
BMCMCs	Bone marrow-cultured mast cells
BMMNC	Bone marrow mononuclear cell
BMT	Bone marrow transplant
BSA	Bovine serum albumin
CD	Circular dichroism
CFU-G	Colony forming unit-granulocyte
CFU-GEMM	Colony forming unit-granulocyte, erythroid, macrophage, megakaryocyte
CFU-GM	Colony forming unit-granulocyte, monocyte
CFU-M	Colony forming unit-monocyte
CML-BC	Chronic myeloid leukemia in blast crisis
CMML	Chronic myelomonocytic leukaemia
CMV	Cytomegalovirus
CN	Cytogenetically normal
Co-IP	Co-immunoprecipitation
DBA	Diamond Blackfan Anaemia
DCML	Dendritic cell, monocyte, B and NK lymphoid
DEL	Deletion
DEPC	Diethylpyrocarbonate
DFS	Disease Free Survival
DNP	Dinitro-phenyl-albumin
DSS	Disease Specific Survival
EDTA	Ethylenediaminetetraacetic acid
EFS	Event Free Survival
ELN	European LeukaemiaNet
EMSA	Electromobility shift assay

EV	Empty vector
FAB	French-American-British
FACS	Fluorescence Activated Cell Sorting
FBS	Foetal Bovine Serum
G-CSF	Granulocyte-colony stimulating factor
GFP	Green Fluorescent Protein
GMP	Granulocyte Monocyte Progenitor
GOF	Gain-of-function
H&E	Haematoxylin and Eosin
HEK	Human Embryonic Kydney
HEPES	N-2-hydroxyethylpiperazine-N-2-ethane sulfonic acid
Het	Heterozygous
HSC	Haematopoietic Stem Cell
HSPC	Hematopoietic stem and progenitor cells
ID	Immunodeficiency disorders
IFC	Integrated Fluidic Circuit
IL-6	Interleukin 6
ITC	Isothermal titration calorimetry
JMML	Juvenile myelomonocytic leukaemia
LB	Luria Broth
LBI	*LAPTM4B/BSPRY/IDH1*
LCL	Lymphoblastic cell line
LOF	Loss-of-function
LOH	Loss of heterozygosity
LSK	$Lin^{-}Sca1^{+}c\text{-}Kit^{+}$
MDS	Myelodysplastic Syndrome
MLL	Mixed-lineage leukemia
MNC	Mononuclear cell
MOI	Multiplicity of infection
MonoMAC	Monocytopenia with *Mycobacterium avium* complex
MPD	Myeloproliferative disease
MPS	Massively Parallel Sequencing
MRI	Magnetic resonance imaging
MSCV	Murine stem cell virus

Mut	Mutant
NGS	Next Generation Sequencing
NK	Natural Killer
NLS	Nuclear Localisation Signal
NMR	Nuclear magnetic resonance
NSCLC	Non-Small Cell Lung Cancer
OS	Overall Survival
PAGE	Polyacrylamide gel electrophoresis
PBMNC	Peripheral blood mononuclear cell
PBS	Phosphate buffered saline
PBS-T	PBS-Tween 20
p-NAG	p-Nitrophenyl-N-Acetyl-β-D-Glucosaminide
qRT-PCR	Quantitative Real Time Polymerase Chain Reaction
RA	Refractory anaemia
RAEB	Refractory anaemia with excess blasts
RAEB-T	Refractory anaemia with excess blasts in transformation
RARS	Refractory anaemia with ring sideroblasts
rfsrc	Random forests for survival, regression and classification
RIPA	Radio-Immunoprecipitation Assay
RP	Ribosomal proteins
SACRB	South Australian Cancer Research Biobank
SCF	Stem cell factor
SDS	Shwachman Diamond Syndrome
SDS	Sodium dodecyl sulphate
SNP	Single nucleotide polymorphism
STA	Specific target amplification
TALL	T cell lymphoblastic leukaemia
t-AML	Therapy related AML
TGE	Tris-glycine-EDTA
TNF	Tumour Necrosis Factor
TPO	Thrombopoietin
UPD	Uniparental disomy
VIMP	Variable importance
WB	Western blot

WEMSA	Western blotting-electromobility shift assay
WES	Whole exome sequencing
WGS	Whole genome sequencing
WHO	World Health Organization
WT	Wild type
ZF1	Zinc finger 1
ZF2	