

Use of copy number variation polymorphism to assess cellular and cell-free DNA chimerism following transplantation

John Breis Whitlam

orcid.org/0000-0003-2202-1462

Doctor of Philosophy

June 2019

Cytomolecular Diagnostics Research Group

Murdoch Children's Research Institute

Royal Children's Hospital

and

Department of Medicine

Austin Health

University of Melbourne

Submitted in total fulfilment of the degree.

ABSTRACT

A state of genetic chimerism is created by allogeneic transplantation, the quality and quantity of which may have diagnostic significance. The monitoring of cellular macrochimerism following allogeneic hematopoietic cell transplantation (HCT) is standard of care. It provides an assessment of engraftment, prognostic information, and guides therapeutic interventions. Separately, the discovery of cell-free DNA (cfDNA) chimerism following solid organ transplantation has raised the possibility that increases in donor genotype graft-derived cfDNA may be biomarker for rejection.

As a result of these and other developments, the clinical and research requirements for chimerism analysis in the transplantation context exceeds the capabilities offered by existing methodologies. To address this, a panel of 39 droplet digital PCR (ddPCR) assays was developed. These targeted highly polymorphic copy number variation (CNV) loci strategically selected using a population genomics approach to maximise the probability of a donor one copy, recipient zero copy genotype combination (and vice versa). These informative combinations create a negative background against which absolute quantification of the one copy genotype can be performed using ddPCR.

Analytical validation of the developed chimerism analysis method for absolute quantification of cfDNA was performed. All studied assays performed linearly across the range <6-1280 copies/mL. Limit of blank was 0 copies/mL, limit of detection was 6 copies/mL, and limit of quantification was 8 copies/mL. A method for genotyping donor and recipient informative CNV loci from chimeric cfDNA was developed and conceptually validated using donor and recipient genomic DNA.

The developed chimerism analysis method was used to evaluate the diagnostic validity of absolute measurements of plasma graft-derived cfDNA and total cfDNA, as well as the relative measure of graft fraction, for the diagnosis of rejection. 61 indication biopsy-matched samples were studied. For the diagnosis of antibody mediated rejection, the receiver-operator characteristic area under the curves of graft-derived cfDNA and graft fraction were 0.91 (95% CI 0.82-0.98) and 0.89 (95% CI 0.79-0.98), respectively. Both measures did not diagnose borderline or type 1A cellular mediated rejection.

Analytical validation was also performed for the quantification of cellular chimerism in blood following allogeneic HCT. CNV chimerism results were compared against established fluorescence in situ hybridization, single nucleotide polymorphism, and short tandem repeat-based methods with excellent correlation. Using 30 ng of DNA, the limit of detection was 0.05% chimerism and the limit of quantification was 0.5% chimerism. Linear performance was established between <0.05%-100% chimerism. High informativity was seen with a median of

four informative markers detectable per individual in 39 recipients and 43 donor genomes studied. The strength of this approach was exemplified in a multi-donor case involving four genomes, three of which were related.

In conclusion, this thesis describes the development and analytical validation of a novel, highly informative test for cellular and cfDNA chimerism analysis based upon highly polymorphic CNV and absolute quantification via ddPCR. In addition to its use in the diagnosis of KT rejection and detection of chimerism after allogeneic HCT, the method holds promise as a means to study chimerism more broadly and in a variety of compartments.

DECLARATION

This thesis comprises only original work towards the PhD except where indicated in the preface. Due acknowledgement has been made in the text to all other material used. The thesis is fewer than the maximum word limit in length, exclusive of tables, maps, bibliographies and appendices.

John Whitlam

PREFACE

Work towards the thesis that was carried out in collaboration with others

Ling Ling, our laboratory research assistant, taught me the laboratory techniques and assisted with some of the experiments. Collaborative contributions for published work are summarised below.

Work towards the thesis that has been submitted for other qualifications

None

Work towards the thesis that was carried out prior to enrolment in the degree

None

Third party editorial assistance in preparation of the thesis

None

Contributions of all persons involved in any multi-authored publications

Chapter 5

- Authors: Whitlam JB¹⁻³, Ling L², Skene A⁵, Kanellis J⁶, Ierino FL^{3,7}, Slater HR^{2,8}, Bruno DL⁴, Power DA^{1,3}
- Affiliations:
 1. Department of Nephrology, Austin Health, Melbourne, Victoria, Australia
 2. Cyto-molecular Diagnostics Research, Murdoch Children's Research Institute, Royal Children's Hospital, Melbourne, Victoria, Australia
 3. Department of Medicine, University of Melbourne, Melbourne, Victoria, Australia
 4. Victorian Clinical Genetics Services, Murdoch Children's Research Institute, Melbourne, Victoria, Australia
 5. Department of Anatomical Pathology, Austin Health, Melbourne, Victoria, Australia
 6. Department of Nephrology, Monash Health and Centre for Inflammatory Diseases, Department of Medicine, Monash University, Melbourne, Victoria, Australia
 7. Department of Nephrology, St Vincent's Health, Melbourne, Victoria, Australia
 8. Department of Paediatrics, The University of Melbourne, Melbourne, Victoria, Australia
- Author contributions: J.B.W. and J.K. were responsible for recruitment and sample collection. J.B.W. and L.L. performed the laboratory work. J.B.W. performed the data analysis and wrote the manuscript. D.L.B. and H.S. provided scientific advice/oversight and edited the paper. Authors D.A.P., J.K., F.L.I. and A.S. provided clinical advice/oversight and edited the paper.

Chapter 6

- Authors: John B Whitlam^{1,2}, Ling Ling¹, Michael Swain⁵, Tom Harrington¹, Oksana Mirochnik⁶, Ian Brooks¹, Sara Cronin¹, Jackie Challis¹, Vida Petrovic¹, Damien L Bruno¹, Francoise Mechinaud⁴, Rachel Conyers⁴, Howard Slater^{1,3}.
- Affiliations:
 1. Cyto-Molecular Diagnostic Research Group, Victorian Clinical Genetics Services and Murdoch Childrens Research Institute, Royal Children's Hospital, Parkville, Victoria, 3052, Australia.
 2. Department of Nephrology, Austin Health, Heidelberg, Victoria, 3084, Australia.
 3. Department of Paediatrics, University of Melbourne, Parkville, Victoria, 3052, Australia.
 4. Children's Cancer Centre, Royal Children's Hospital, Parkville, Victoria, 3052, Australia.
 5. Laboratory Services, Royal Children's Hospital, Parkville, Victoria, 3052, Australia.
 6. Pathology West Institute of Clinical Pathology and Medical Research, Westmead Hospital, Westmead, New South Wales, 2145, Australia.
- Author Contributions: F.M., R.C. and H.S. were responsible for recruitment and sample collection. J.B.W., L.L., M.S., T.H., O.M., I.B., S.C., J.C., and V.P. performed the laboratory work. J.B.W. performed the data analysis and wrote the manuscript. D.L.B. and H.S. provided scientific advice/oversight and edited the paper. Authors F.M. and R.C. provided clinical advice/oversight and edited the paper.

Publication status of all chapters presented in article format

- Chapter 5 published by the American Journal of Transplantation on 12 October 2018.
- Chapter 6 published by Experimental Haematology in May 2017.

Sources of funding

Project funding and organisational support

- Roche Organ Transplant Research Foundation 529172731
- Victorian Government Operational Infrastructure Support Program
- Murdoch Children's Research Institute
- Austin Health Department of Nephrology

Stipend support

- Kidney Health Australia Post-Graduate Research Scholarship SCH16 141516
- Australian Government National Health and Medical Research Council Post-Graduate Scholarship GNT1114931

ACKNOWLEDGEMENTS

This work is dedicated to transplant recipients, who take a leap of faith and cope with significant uncertainty, at this frontier of medicine.

I would like to thank my supervisors, Prof David Power, Dr Damien Bruno, Dr Alison Skene, and A/Prof Peter Hughes, for their thoughtful counsel regarding this work. A/Prof Howard Slater, while not a formal supervisor, was an invaluable source of ideas and experience.

Cyto-molecular Diagnostics Laboratory Head, Dr David Godler, was gracious in sharing his lab and equipment, and took no prisoners when it came to lunchtime table tennis. This work would not be possible without the support of the laboratory staff. Research assistant, Ling Ling, taught me all of the relevant laboratory techniques with good humour, patience, and efficiency. Her assistance with experiments and troubleshooting was invaluable. Tom Harrington, Ranier Arolanda, and Marina Katerelos assisted with sample processing and handling. Marieke Veenendaal assisted with recruitment and sample collection.

I acknowledge the contribution of our collaborators and funding bodies, described further in the preface. In particular, I am grateful for the financial support afforded me by Kidney Health Australia and the National Health and Medical Research Council.

I am incredibly grateful for the support, tolerance, and patience of my wife, Sophie. Thank you for "holding the fort" while I worked on weekends and evenings, and for your emotional support and brutal feedback (where required).

My daughters, Grace and Evelyn, were both born during this PhD. Grace, the eldest, spent many weekends playing at the Royal Children's Hospital playground or adjacent park (supervised, of course) while I worked in the laboratory. Evelyn, 7 months old as I write this, is still learning to sleep through the night, and has kept me company during late night writing. You have both been a joyful, energetic, contextualising presence that has grounded me through this process.

Finally, I am grateful for the support and guidance of my parents, Christine and George, who have encouraged me academically throughout my life. I wish I could have spent more of the thesis writing period in Maleny, amongst the trees and the quiet, free from distraction – but life happened.

TABLE OF CONTENTS

Abstract.....	ii
Declaration.....	iv
Preface.....	v
Acknowledgements.....	vii
Table of Contents	viii
List of Tables	xiii
List of Figures.....	xvi
Abbreviations.....	xxv
Chapter 1 Literature Review.....	1
1.1 Chimerism	1
1.2 Chimerism Analysis	2
1.2.1 Common Methods of Chimerism Analysis in Transplantation	2
1.2.2 Population Genomics Reveals New Targets for Chimerism Analysis	3
1.2.3 Digital PCR Provides True Absolute Quantification	4
1.3 Cellular Chimerism	5
1.3.1 Allogeneic Hematopoietic Cell Transplantation	5
1.3.2 Mixed Bone Marrow Chimerism and Tolerance.....	6
1.3.3 Pregnancy-Associated Microchimerism.....	7
1.3.4 Transplant-Associated Microchimerism	8
1.4 Cell-Free DNA Chimerism	8
1.4.1 Cell-Free DNA	8
1.4.2 Non-Transplant Settings.....	10
1.4.3 Allogenic Hematopoietic Cell Transplantation	10
1.4.4 Solid Organ Transplantation	11
1.5 Kidney Transplantation	21
1.5.1 The Problem of Kidney Transplant Rejection.....	21
1.5.2 Current Strategies for Rejection Surveillance and Diagnosis	23
1.5.3 Non-Invasive Rejection Biomarkers	25

1.5.4	Clinical Translation of Rejection Biomarkers	27
1.6	Summary & Aims.....	28
Chapter 2	Assay Design	30
2.1	Introduction	30
2.2	Methods.....	31
2.2.1	CNV Assay Development	31
2.2.2	Non-Polymorphic Assay Development.....	34
2.3	Results	35
2.3.1	CNV Assay Development	35
2.3.2	Non-Polymorphic Assay Development.....	49
2.4	Discussion	51
2.4.1	CNV Assay Development	51
2.4.2	Non-Polymorphic Assay Development.....	53
2.4.3	Summary	54
Chapter 3	Informativity Genotyping.....	55
3.1	Introduction	55
3.2	Methods.....	57
3.2.1	Participants	57
3.2.2	Recipient Blood Collection and Processing	57
3.2.3	Donor Blood Collection and Processing	60
3.2.4	Recipient Urine Collection and Processing.....	61
3.2.5	Droplet Digital PCR.....	63
3.2.6	Conversion of Results	64
3.2.7	Blood Genotyping	64
3.2.8	Recipient Genotyping Analysis.....	65
3.2.9	Donor Genotyping Analysis.....	65
3.2.10	Urine Genotyping.....	66
3.2.11	Bioinformatics and Statistical Methods	66
3.3	Results	66

3.3.1	Blood	66
3.3.2	Urine.....	72
3.4	Discussion	75
3.4.1	Blood	75
3.4.2	Urine.....	77
3.4.3	Algorithmic Genotyping	79
3.4.4	Summary	80
Chapter 4	Analytical Validation	82
4.1	Introduction	82
4.2	Methods.....	83
4.2.1	Study Participants.....	84
4.2.2	Sample Collection, Processing and DNA Extraction	84
4.2.3	One-Copy Genotypes	85
4.2.4	Zero-Copy Genotypes	87
4.2.5	DNA Fragmentation.....	88
4.2.6	Experimental Design	89
4.2.7	Accepted Droplets	97
4.2.8	Droplet Cluster Centroids.....	97
4.2.9	Droplet Thresholding	99
4.2.10	Conversion of Results to Plasma Equivalent.....	102
4.2.11	No Template and Zero-Copy Controls.....	102
4.2.12	One-Copy Dilution Series	103
4.2.13	Two-Copy Dilution Series	106
4.2.14	Statistical Significance	107
4.3	Results	107
4.3.1	One-Copy Genotypes	107
4.3.2	Zero-Copy Genotypes	114
4.3.3	DNA Fragmentation.....	116
4.3.4	Experimental Design	120

4.3.5	Accepted Droplets	127
4.3.6	Droplet Cluster Centroids.....	129
4.3.7	Droplet Thresholding	134
4.3.8	No Template Control.....	139
4.3.9	Zero-Copy Control	143
4.3.10	One-Copy Dilution Series	150
4.3.11	Two-Copy Dilution Series	168
4.4	Discussion	172
4.4.1	Combinatorial Modelling and DNA Fragmentation.....	172
4.4.2	Dilution Series and Template Volume	173
4.4.3	Cluster Centroids and Droplet Thresholding.....	175
4.4.4	Assay Performance Assessment.....	177
4.4.5	Summary	178
Chapter 5	Kidney Transplant Rejection Diagnosis.....	180
5.1	Abstract	180
5.2	Main Body Text	180
5.2.1	Introduction	180
5.2.2	Materials and Methods.....	181
5.2.3	Results	184
5.2.4	Discussion	194
5.3	Author Contributions.....	197
5.4	Acknowledgments/Funding	197
5.5	Disclosures	197
5.6	Supplemental Material	197
5.6.1	Supplemental Methods.....	197
5.6.2	Supplemental Results	200
Chapter 6	Hematopoietic Cell Transplantation Chimerism	210
6.1	Abstract	210
6.2	Main Body Text	210

6.2.1	Introduction	210
6.2.2	Materials and Methods	212
6.2.3	Results	216
6.2.4	Discussion	222
6.2.5	Conclusion.....	223
6.3	Funding	224
6.4	Disclosure of Conflicts of Interest.....	224
6.5	Supplementary Material	224
6.5.1	Droplet Digital PCR Plots	224
6.5.2	CNV Loci	225
6.5.3	Sensitivity Analysis.....	226
6.5.4	Assay Linear Model Characteristics.....	227
6.5.5	Assay Performance.....	228
Chapter 7	Conclusions and Future Directions.....	229
7.1	Cell-Free DNA Chimerism Analysis	230
7.2	Cellular DNA Chimerism Analysis.....	232
References		235
Appendix A	Selected UCSC Genome Browser Tracks	261
Appendix B	Assay UCSC Genome Browser Tracks	268

LIST OF TABLES

Table 1-1: Published original articles concerning plasma cfDNA chimerism following solid organ transplantation.	16
Table 1-2: Published original articles concerning urine cfDNA chimerism following solid organ transplantation.	20
Table 2-1: Parameters used for the UCSC Genome Browser LiftOver tool.	31
Table 2-2: UCSC Genome Browser annotation tracks used for visual inspection of CNV loci.	32
Table 2-3: Modifications to default parameters for assay design using the IDT PrimerQuest Tool.	32
Table 2-4: Parameters for melting temperature calculation using IDT OligoAnalyzer 3.1.	33
Table 2-5: Parameters used for UCSC Genome Browser In-Silico PCR.	34
Table 2-6: GRCh37/hg19 genome coordinates of 41 CNV loci identified for assay development.	35
Table 2-7: CNV locus allele frequency within source high resolution data sets.	37
Table 2-8: Forward primer characteristics for each CNV locus.	40
Table 2-9: Reverse primer characteristics for each CNV locus.	41
Table 2-10: Probe characteristics for each CNV locus.	42
Table 2-11: Assay oligonucleotide non-specificity identified by BLAT search.	46
Table 2-12: CNV assay in silico PCR amplicon characteristics.	46
Table 2-13: Repeating elements located within CNV assays.	48
Table 2-14: Assays positioned within known genes.	48
Table 2-15: Overlapping CNV loci and assay genomic coordinates.	49
Table 2-16: Overlapping CNV loci assay linkage.	49
Table 2-17: <i>ACE</i> and <i>NPPB</i> non-polymorphic assay characteristics.	50
Table 3-1: Plasma cfDNA extraction reagent volumes based upon the total volume of sample in each tube.	60
Table 3-2: Buffer ACL and carrier RNA master mix composition used for plasma cfDNA extraction.	60
Table 3-3: Urine cfDNA extraction reagent volumes based upon the total volume of sample in each tube.	63
Table 3-4: Buffer ACL and carrier RNA master mix composition used for urine cfDNA extraction.	63
Table 3-5: PCR reaction components for the genotyping studies.	64
Table 3-6: PCR cycling conditions for the genotyping studies.	64
Table 3-7: Simple genotyping definitions applied to plasma cfDNA sample assays.	65

Table 3-8: Contingency table of transplant recipient cfDNA recipient genotyping performance compared against genotyping performed directly on recipient genomic DNA.....	69
Table 3-9: Contingency table of transplant recipient cfDNA recipient copy number genotyping performance compared against genotyping performed directly on recipient genomic DNA.....	70
Table 3-10: Contingency table of transplant recipient cfDNA graft informative genotyping performance compared against genotyping performed directly on donor genomic DNA	71
Table 4-1: Covaris S220 parameters for fragmentation of genomic DNA targeting a peak size of ~150-160 bp.	89
Table 4-2: Sample and template volume model conditions.	90
Table 4-3: Total droplet number model conditions.....	91
Table 4-4: Parameters used to model clinical sample equivalency of dilution series concentrations.	91
Table 4-5: Genotypes and assay pairs for each of the three validation study groups.....	92
Table 4-6: PCR reaction components for the analytical validation studies.....	95
Table 4-7: PCR cycling conditions for the analytical validation study.....	96
Table 4-8: Combinatorial algorithm modelling of maximal recipient one-copy marker coverage using 25 participants and various group sizes.	109
Table 4-9: Assay pairings by participant indicating the variation of each assay from the one-copy median on confirmatory genotyping	113
Table 4-10: Combinatorial algorithm modelling of maximal recipient zero-copy marker coverage of each one-copy participant genotype using 25 participants and various group sizes.....	114
Table 4-11: The most efficient zero-copy participant allocation to act as a zero-copy controls for each one-copy participant assay pair.....	115
Table 4-12: Modelled equivalent sample concentrations and positive droplet number for dilution series PCR and template concentrations.	125
Table 4-13: Digital PCR quantification of one-copy markers in diluted fragmented cellular DNA and the equivalent neat concentration.	126
Table 4-14: Distribution of assay pairs between digital PCR plates for each validation study and the corresponding zero-copy controls for each plate.....	127
Table 4-15: Digital PCR reaction wells with <10,000 accepted droplets.	128
Table 4-16: Channel 1 and channel 2 amplitudes for double negative, FAM-positive and HEX-positive droplet cluster centroids for each assay pair within the validation study.	129
Table 4-17: Double negative cluster centroid channel 1 and channel 2 amplitudes for dilution series, zero-copy control and no template control replicates of each assay pair.	131
Table 4-18: Validation study assays with outlier lower amplitude positive cluster centroids.	133
Table 4-19: Threshold amplitudes for three thresholding techniques based upon fractional distances between the respective positive cluster and double negative cluster.....	134

Table 4-20: Frequency of no template control false positive results by assay and threshold technique.	140
Table 4-21: Limit of blank parameters and results for assays run against no template control samples consisting of distilled water.....	142
Table 4-22: Frequency of zero-copy control false positive results by assay and threshold technique.	145
Table 4-23: Zero-copy control false positive results, using the third-distance threshold technique.	149
Table 4-24: Limit of blank parameters and results for assays run against zero-copy control samples.....	149
Table 4-25: One-copy dilution series linear model parameters.....	155
Table 4-26: Outlier assay linear model parameters after validation study results were pooled.	157
Table 4-27: The maximum concentration at which replicate failure was observed for each CNV assay (proportion of replicates detecting target < 1).	159
Table 4-28: Limit of detection for each CNV assay, using zero-copy control limit of blank. .	160
Table 4-29: Limit of quantification for each of the studied CNV assays and contributing parameters.	167
Table 4-30: <i>ACE</i> one-copy dilution series linear model parameters	170
Table 5-1: Participant characteristics.....	185
Table 5-2: Sample characteristics.....	186
Table 5-3: Diagnostic performance characteristics of gd-cfDNA and graft fraction for the diagnosis of aAMR and ABMR.	191
Table 5-4: Participant characteristics for the aAMR groups.	200
Table 5-5: Sample characteristics for the aAMR groups.....	201
Table 6-1: Genomic locations, sizes and population frequencies of the 38 CNV panel.	225
Table 6-2: Proportion of tested marker replicates that were positive at each chimeric fraction.	226
Table 6-3: Linear model characteristics and correlation coefficient values from marker performance in the sensitivity analysis spiking series.....	227
Table 7-1: Diagnostic validity studies of cell-free DNA chimerism analysis in kidney transplantation.	231

LIST OF FIGURES

Figure 1-1: Conceptual model of “-omic” technologies with respect to the Central Dogma of Biology.....	25
Figure 1-2: A natural history model of rejection indicating the conceptual stage targeted by various diagnostic biomarker approaches.	25
Figure 2-1: Chromosomal frequency of CNV loci identified for assay development.	36
Figure 2-2: Size distribution of CNV loci identified for assay development.....	36
Figure 2-3: Distribution of copy number allelic frequency amongst the CNV loci selected for assay development.	39
Figure 2-4: Distribution of oligonucleotide size for CNV assay forward primers, reverse primers, and probes.	43
Figure 2-5: Distribution of oligonucleotide melting temperatures for CNV assay forward primers, reverse primers, and probes.....	44
Figure 2-6: Distribution of oligonucleotide GC content for CNV assay forward primers, reverse primers, and probes.	44
Figure 2-7: CNV assay distribution of the absolute difference in melting temperature between forward and reverse primers.....	45
Figure 2-8: CNV assay distribution of the difference between probe melting temperature and the mean of the primer melting temperatures.....	45
Figure 2-9: Distribution of CNV assay amplicon size.	47
Figure 3-1: CNV assay results for transplant recipient chimeric cfDNA extracted from plasma	67
Figure 3-2: Within groups sum of squares plotted by k-means cluster number for two samples (SID 1 and SID 5)	68
Figure 3-3: CNV assay results from recipient genomic DNA (top panel) presented against cfDNA CNV assay results for the same recipient (bottom panel)	69
Figure 3-4: CNV assay results from donor genomic DNA (top panel) presented against cfDNA (middle panel) and recipient genomic DNA (bottom panel) CNV assay results corresponding to the same recipient.....	71
Figure 3-5: Graft-derived cfDNA, total cfDNA, and graft fraction for plasma and urine samples collected simultaneously from 34 kidney transplant recipients	73
Figure 3-6: CNV assay results from urine cfDNA (fourth panel) presented against recipient genomic DNA (first panel), plasma cfDNA (second panel), and donor genomic DNA (third panel) CNV assay results corresponding to the same recipient	73
Figure 3-7: Plasma (top panel) and urine (bottom panel) cfDNA CNV assay panel results from the same recipient demonstrating the spectrum of allocated genotypes for both samples	74

Figure 4-1: Schematic representation of the analytical validation experiments, demonstrating how each validation study group was studied in series.....	84
Figure 4-2: Schematic overview of the validation study laboratory work.	95
Figure 4-3: Digital PCR 96-well plate layout for validation study assay pairs.....	95
Figure 4-4: 2D heat maps were used to aid visual approximation of cluster centroids, in this case a double negative droplet cluster.....	98
Figure 4-5: Cluster thresholding technique applied to a 2D amplitude plot of 28 wells <i>en bloc</i>	99
Figure 4-6: Tight thresholding technique applied to a 2D amplitude plot of 28 wells <i>en bloc</i>	100
Figure 4-7: Half-distance thresholding technique applied to a 2D amplitude plot of 28 wells <i>en bloc</i>	101
Figure 4-8: Third-distance thresholding technique applied to a 2D amplitude plot of 28 wells <i>en bloc</i>	101
Figure 4-9: Quarter-distance thresholding technique applied to a 2D amplitude plot of 28 wells <i>en bloc</i>	102
Figure 4-10: Distribution of total number of markers genotyped per individual of the 25 individuals used for combinatorial modelling.....	108
Figure 4-11: Frequency of non-genotyped markers amongst the 25 participants used for combinatorial modelling.	108
Figure 4-12: Frequency of recipient one-copy markers amongst the 25 participants used for combinatorial modelling.	109
Figure 4-13: The distribution of recipient one-copy markers covered by all combinations of different genotypes, grouped by the number of individuals in the combination (group size)... ..	110
Figure 4-14: Confirmatory genotyping of 38 markers using cellular DNA from 3 participants selected for use in the analytical validation experiments.	111
Figure 4-15: CNV one-copy marker coverage and frequency amongst participants 71, 146 and 148.....	112
Figure 4-16: Variation of all CNV markers from the one-copy median at the time of confirmatory genotyping using cellular DNA.....	112
Figure 4-17: Variation of duplicated CNV markers from the one-copy median at the time of confirmatory genotyping using cellular DNA.....	112
Figure 4-18: Frequency of recipient zero-copy markers amongst the 25 participants used for combinatorial modelling.	114
Figure 4-19: Confirmatory genotyping of zero-copy control participants.	116
Figure 4-20: TapeStation gel of three extracted plasma cfDNA samples (S39, S43, S30) against a High Sensitivity D1000 Ladder.....	117
Figure 4-21: TapeStation electropherograms demonstrating fragmentation profiles of three plasma cfDNA samples (S39, S43, S30) within lower and upper boundary markers.....	118

Figure 4-22: TapeStation electropherograms demonstrating fragmentation characteristics of ultrasonicated cellular DNA from one-copy genotype participants 71, 146 and 148 within lower and upper boundary markers	118
Figure 4-23: TapeStation electropherograms demonstrating fragmentation characteristics of ultrasonicated cellular DNA from zero-copy genotype participants 97, 114 and 133 within lower and upper boundary markers	119
Figure 4-24: Modelled number of expected positive droplets from an accepted droplet number of 15,000 as a function of sample concentration using a variety of template DNA volumes across the sample concentration range	122
Figure 4-25: Modelled number of expected positive droplets as a function of sample concentration using a variety of total droplet numbers across the sample concentration range	124
Figure 4-26: Histogram of accepted droplet number for all validation study digital PCR reaction wells.	128
Figure 4-27: Distribution of accepted droplets compared between validation studies.	128
Figure 4-28: Distribution of accepted droplets compared between assay pairs within each plate of the validation studies.....	129
Figure 4-29: Location of heat-map assisted, visually-defined approximate cluster centroids for 19 assay pairs.	130
Figure 4-30: Distribution of channel amplitudes for assay pair double negative cluster centroids comparing dilution series, zero-copy control and no template control replicates.	132
Figure 4-31: Histograms of cluster centroid channel 1 (left column) and channel 2 (right column) amplitude.....	133
Figure 4-32: Assay design characteristics compared between a group of assays with outlier lower amplitude positive cluster centroids and those with normal amplitude positive cluster centroids.	133
Figure 4-33: 2D plots of CNV 29A/18A demonstrating the effect of droplet rain on cluster margin identification using the cluster threshold technique.	135
Figure 4-34: Histograms of the number of FAM- and HEX- positive droplets using the tight thresholding technique, grouped by the number of wells simultaneously selected for thresholding	136
Figure 4-35: 2D droplet plot for CNV assay pair 23A/17A selected <i>en bloc</i> , demonstrating how the quarter-distance threshold technique transects the FAM-positive cluster.....	137
Figure 4-36: The relationship between HEX-positive cluster channel 2 amplitude, FAM-positive cluster channel 2 amplitude and proximity of the quarter-distance defined channel 2 threshold to the FAM-positive cluster.....	138
Figure 4-37: The false positive rate of 126 no template control wells grouped by positive droplet threshold technique.	139

Figure 4-38: The concentration of no template control results grouped by thresholding technique.	140
Figure 4-39: Assay design characteristics compared between a group of assays with at least one no template control false positive result, and those with no false positive results.	141
Figure 4-40: Centroid amplitudes and centroid separation (positive-negative) compared between a group of assays with no template control false positive results, and those with no false positive results.	142
Figure 4-41: The proportion of 144 zero-copy control results that were false positive, grouped by threshold technique.	144
Figure 4-42: The concentration of false positive zero-copy control results grouped by threshold technique.	145
Figure 4-43: Dot plot of the number of false positive droplets per false positive replicate well for each assay, grouped by threshold technique.	145
Figure 4-44: Assay design characteristics compared between a group of assays with at least one false positive result, and those with no false positive results.	147
Figure 4-45: Centroid amplitudes and centroid separation (positive-negative delta) compared between assays with at least one false positive result and those with no false positive results.	148
Figure 4-46: Assay design characteristics compared between assays with a false positive result at the third-distance threshold technique, and those with no false positive results.	148
Figure 4-47: Centroid amplitudes and centroid separation (positive-negative delta) compared between assays with a false positive result at the third-distance threshold technique, and those with no false positive results.	149
Figure 4-48: Box plot of CNV assay concentration as a proportion of the target concentration at that particular dilution step, grouped by validation study.	152
Figure 4-49: The effect of threshold technique on the proportion of one-copy target concentration achieved across all dilution steps.	153
Figure 4-50: Validation study 1 CNV assay replicate results and linear models.	153
Figure 4-51: Validation study 2 CNV assay replicate results and linear models.	154
Figure 4-52: Validation study 3 CNV assay replicate results and linear models.	155
Figure 4-53: Summary of linear model characteristics by validation study.	157
Figure 4-54: Assay design characteristics compared between a group of assays with outlier linear model slope values and those of non-outlier assays.	157
Figure 4-55: Tile plot indicating the proportion of assay replicates that detected target (positive assay result) across validation study dilution step concentrations.	158
Figure 4-56: Distribution of intra-assay coefficient of variation (CV) grouped by median validation study dilution step concentration.	162

Figure 4-57: Intra-assay coefficient of variation (CV) plotted against assay median concentration.	162
Figure 4-58: Distribution of intra-assay coefficient of variation (CV) calculated using assays detecting target only, grouped by median validation study dilution step concentration	163
Figure 4-59: Intra-assay coefficient of variation for each CNV assay, incorporating detectable replicate results from each dilution step ≥ 1 copy/ μL	164
Figure 4-60: Inter-assay coefficient of variation (CV) plotted against the median validation study dilution step concentration	165
Figure 4-61: Inter-assay coefficient of variation (CV) plotted against median validation study dilution step concentration	165
Figure 4-62: Distribution of inter-assay coefficient of variation (CV) calculated using assays detecting target only, grouped by median validation study dilution step concentration	166
Figure 4-63: Proportionate variation of assay replicate median concentration compared to dilution step median concentration, grouped by assay	167
Figure 4-64: Box plot of <i>ACE</i> assay concentration as a proportion of the target concentration at that particular dilution step, grouped by validation study.	169
Figure 4-65: Validation study <i>ACE</i> assay replicate results and linear models.	169
Figure 4-66: <i>ACE</i> intra-assay coefficient of variation (CV) by median <i>ACE</i> concentration....	170
Figure 4-67: Intra-assay coefficient of variation (CV) plotted against <i>ACE</i> assay median concentration	171
Figure 5-1: Study sample flow diagram.	184
Figure 5-2: Sample measurements of gd-cfDNA (left panel), total cfDNA (middle panel) and calculated graft fraction (right panel) grouped by biopsy diagnosis.	189
Figure 5-3: Sample measurements of gd-cfDNA (left panel), total cfDNA (middle panel) and calculated graft fraction (right panel) grouped by the presence (positive) or absence (negative) of ABMR.	189
Figure 5-4: Receiver-operator characteristic curves for gd-cfDNA (black line), graft fraction (light gray line), and creatinine (dark gray dashed line) in the diagnosis of aAMR (left panel), and composite aAMR and caAMR (ABMR, right panel)	190
Figure 5-5: The distribution of gd-cfDNA measurements for each Banff lesion and score....	192
Figure 5-6: The distribution of calculated graft fraction results for each Banff lesion and score.	193
Figure 5-7: Correlation of serum creatinine concentration with gd-cfDNA, total cfDNA and graft fraction.	194
Figure 5-8: Bio-Rad QuantaSoft Analysis Pro 2D plots from two ddPCR wells, each containing two CNV PCR assays and cfDNA from the sample represented in Supplemental Figure 5-10B.	197

Figure 5-9: Bio-Rad QuantaSoft Analysis Pro 1D plots representing the channel 1 (plot A) and channel 2 (plot B) amplitudes of droplets for the wells containing CNV assay 03B and 06B across six samples (x-axis, divided by dashed vertical lines) on the same ddPCR plate.	198
Figure 5-10: DdPCR concentrations for each of 31 CNV assays and the ACE 2-copy positive control from two different participants	199
Figure 5-11: Histogram plots of the number of assays genotyped as graft informative, recipient 1-copy and recipient 2-copy per sample.....	203
Figure 5-12: Distribution of gd-cfDNA concentrations grouped by the number of graft informative markers per sample	203
Figure 5-13: Distribution of the number of graft informative (left panel) and recipient 1-copy (right panel) assays per sample, grouped by relatedness between donor and recipient	204
Figure 5-14: The number of graft informative assays per sample in related donor-recipient pairs, grouped by relationship	204
Figure 5-15: The relationship between inter-assay coefficient of variation and median assay concentration for that genotype.....	205
Figure 5-16: Correlation between gd-cfDNA and graft fraction.....	205
Figure 5-17: Distribution of cell-free DNA measurements for samples taken before or after their corresponding biopsy	207
Figure 5-18: Distribution of cell-free DNA measurements for negative biopsy samples taken before or after their corresponding biopsy	207
Figure 5-19: Distribution of cfDNA measurements by ABMR classification and time since transplantation	207
Figure 5-20: Distribution of cfDNA measurements from negative biopsy cases over time since transplantation	208
Figure 6-1: Measurement linearity and precision across the chimeric range of three representative target markers (CNV5, CNV14 and CNV39)	217
Figure 6-2: Intra- (top panel) and inter- (lower panel) assay variation of 26 CNV markers measured in triplicate across the spiking series indicated	217
Figure 6-3: Inter-assay variation of 8 informative donor CNV markers and 5 informative recipient CNV markers at each time point during longitudinal sampling from a clinical case.....	218
Figure 6-4: Comparison of chimerism measurements using the CNV method with (left panel) an X/Y FISH method using a spiked series of male and female cells; (middle panel) an SNP/rtPCR method using DNA from 15 longitudinally collected blood samples from 4 transplant recipients; (right panel) an STR-based method using a spiked series.....	218
Figure 6-5: Histograms of method informativity in 39 transplant recipients and 43 donors ...	220
Figure 6-6: Concentrations of 15 selected CNVs in blood collected from the recipient immediately before the second transplant (Tx#2).....	220

Figure 6-7: Chimeric fraction before and after Tx#2 using unique informative CNVs for donor (two CNVs), the aunt (three CNVs) and the recipient (one CNV graphed).....	221
Figure 6-8: 1D ddPCR plots of one copy markers at 10% chimerism demonstrating separation of positive clusters from negative clusters for selected representative assays under standard conditions	224
Figure 6-9: Individual marker performance across the spiked chimeric range.	228
Figure 7-1: Royal College of Pathologists Australia chimerism analysis Quality Assurance Program results from 2017,.....	233
Figure A-1: Numerous common SNPs and repeating elements precluded assay development in the CNV 35 locus.	261
Figure A-2: The CNV 37 locus, demonstrating that the entire CNV 40 locus sits within this locus.	261
Figure A-3: Amplicon of assay CNV 15A demonstrating the position of SNP rs111271716 within the probe binding region.	262
Figure A-4: Overlapping CNV locus 3 (locus_3) and CNV locus 25 (locus_25) demonstrating the position of their assays (assay_3, assay_25) with respect to the overlapping loci.	263
Figure A-5: Overlapping CNV 11 locus (locus_11) and CNV 21 locus (locus_21) demonstrating the position of assay amplicons (assay_11, assay_21) with respect to the overlapping loci.....	264
Figure A-6: Overlapping CNV 37 locus (locus_37) and CNV 40 locus (locus_40) demonstrating the position of assay 37A (assay_37) with respect to the overlapping loci.....	265
Figure A-7: Position of the <i>ACE</i> assay within the <i>ACE</i> gene relative to common SNPs, repeating elements and structural variation.....	266
Figure A-8: <i>ACE</i> assay spanning the junction of exon 9/25 and intron 9/24 of the <i>ACE</i> gene..	266
Figure A-9: Position of the <i>NPPB</i> assay within intron 2/2 of the <i>NPPB</i> gene relative to common SNPs and repeating elements.	266
Figure B-1: The CNV 1 locus and position of the CNV 01B assay with respect to genes, common SNPs, structural variation and repeating elements.	268
Figure B-2: The CNV 2 locus and position of the CNV 02B assay with respect to genes, common SNPs, structural variation and repeating elements.	268
Figure B-3: The CNV 3 locus and position of the CNV 03B assay with respect to genes, common SNPs, structural variation and repeating elements.	269
Figure B-4: The CNV 4 locus and position of the CNV 04B assay with respect to genes, common SNPs, structural variation and repeating elements.	270
Figure B-5: The CNV 5 locus and position of the CNV 05B assay with respect to genes, common SNPs, structural variation and repeating elements.	271
Figure B-6: The CNV 6 locus and position of the CNV 06B assay with respect to genes, common SNPs, structural variation and repeating elements.	271

Figure B-7: The CNV 7 locus and position of the CNV 07B assay with respect to genes, common SNPs, structural variation and repeating elements.	272
Figure B-8: The CNV 8 locus and position of the CNV 08B assay with respect to genes, common SNPs, structural variation and repeating elements.	272
Figure B-9: The CNV 9 locus and position of the CNV 09B assay with respect to genes, common SNPs, structural variation and repeating elements.	273
Figure B-10: The CNV 10 locus and position of the CNV 10C assay with respect to genes, common SNPs, structural variation and repeating elements.	273
Figure B-11: The CNV 11 locus and position of the CNV 11A assay with respect to genes, common SNPs, structural variation and repeating elements.	274
Figure B-12: The CNV 12 locus and position of the CNV 12A assay with respect to genes, common SNPs, structural variation and repeating elements.	275
Figure B-13: The CNV 13 locus and position of the CNV 13A assay with respect to genes, common SNPs, structural variation and repeating elements.	276
Figure B-14: The CNV 14 locus and position of the CNV 14A assay with respect to genes, common SNPs, structural variation and repeating elements.	277
Figure B-15: The CNV 15 locus and position of the CNV 15A assay with respect to genes, common SNPs, structural variation and repeating elements.	278
Figure B-16: The CNV 16 locus and position of the CNV 16A assay with respect to genes, common SNPs, structural variation and repeating elements.	279
Figure B-17: The CNV 17 locus and position of the CNV 17A assay with respect to genes, common SNPs, structural variation and repeating elements.	280
Figure B-18: The CNV 18 locus and position of the CNV 18A assay with respect to genes, common SNPs, structural variation and repeating elements.	281
Figure B-19: The CNV 19 locus and position of the CNV 19B assay with respect to genes, common SNPs, structural variation and repeating elements.	282
Figure B-20: The CNV 20 locus and position of the CNV 20A assay with respect to genes, common SNPs, structural variation and repeating elements.	283
Figure B-21: The CNV 21 locus and position of the CNV 21A assay with respect to genes, common SNPs, structural variation and repeating elements.	284
Figure B-22: The CNV 22 locus and position of the CNV 22A assay with respect to genes, common SNPs, structural variation and repeating elements.	285
Figure B-23: The CNV 23 locus and position of the CNV 23A assay with respect to genes, common SNPs, structural variation and repeating elements.	286
Figure B-24: The CNV 24 locus and position of the CNV 24A assay with respect to genes, common SNPs, structural variation and repeating elements.	287

Figure B-25: The CNV 25 locus and position of the CNV 25A assay with respect to genes, common SNPs, structural variation and repeating elements.	288
Figure B-26: The CNV 26 locus and position of the CNV 26A assay with respect to genes, common SNPs, structural variation and repeating elements.	288
Figure B-27: The CNV 27 locus and position of the CNV 27A assay with respect to genes, common SNPs, structural variation and repeating elements.	289
Figure B-28: The CNV 28 locus and position of the CNV 28A assay with respect to genes, common SNPs, structural variation and repeating elements.	290
Figure B-29: The CNV 29 locus and position of the CNV 29A assay with respect to genes, common SNPs, structural variation and repeating elements.	290
Figure B-30: The CNV 30 locus and position of the CNV 30A assay with respect to genes, common SNPs, structural variation and repeating elements.	291
Figure B-31: The CNV 31 locus and position of the CNV 31A assay with respect to genes, common SNPs, structural variation and repeating elements.	291
Figure B-32: The CNV 32 locus and position of the CNV 32A assay with respect to genes, common SNPs, structural variation and repeating elements.	292
Figure B-33: The CNV 33 locus and position of the CNV 33A assay with respect to genes, common SNPs, structural variation and repeating elements.	292
Figure B-34: The CNV 34 locus and position of the CNV 34A assay with respect to genes, common SNPs, structural variation and repeating elements.	293
Figure B-35: The CNV 36 locus and position of the CNV 36A assay with respect to genes, common SNPs, structural variation and repeating elements.	294
Figure B-36: The CNV 37 locus and position of the CNV 37A assay with respect to genes, common SNPs, structural variation and repeating elements.	294
Figure B-37: The CNV 38 locus and position of the CNV 38A assay with respect to genes, common SNPs, structural variation and repeating elements.	295
Figure B-38: The CNV 39 locus and position of the CNV 39A assay with respect to genes, common SNPs, structural variation and repeating elements.	295
Figure B-39: The CNV 41 locus and position of the CNV 41A assay with respect to genes, common SNPs, structural variation and repeating elements.	296

ABBREVIATIONS

ABMR	Composite diagnosis of acute antibody mediated rejection and chronic active antibody mediated rejection
AMR	Antibody mediated rejection
aAMR	Acute antibody mediated rejection
AUC	Area under the curve
aCMR	Acute cellular mediated rejection
bCMR	Borderline cellular mediated rejection
caAMR	Chronic active antibody mediated rejection
BLAT	BLAST-like Alignment Tool
Bp	Base pairs
cfDNA	Cell-free DNA
CI	Confidence interval
CMR	Cellular mediated rejection
CND	Copy number deletion
CNI	Calcineurin inhibitor
CNV	Copy number variation / copy number variant
CV	Coefficient of variation
ddPCR	Droplet digital PCR
DNA	Deoxyribonucleic acid
DSA	Donor specific antibody
FAM	6-carboxyfluorescein
FISH	Fluorescent in-situ hybridisation
gd-cfDNA	Graft-derived cell-free DNA
HCT	Hematopoietic cell transplantation
HR	Hazard ratio
HSCT	Hematopoietic stem cell transplantation
HEX	Hexachlorofluorescein
HLA	Human leukocyte antigen
indel	Insertion-deletion polymorphism
IQR	Inter-quartile range
KT	Kidney transplant
miRNA	MicroRNA
mRNA	Messenger RNA
NOS	Not otherwise specified
NTC	No template control

OR	Odds ratio
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
qrtPCR	Quantitative real-time PCR
RNA	Ribonucleic acid
ROC	Receiver operating characteristic
rtPCR	Real-time PCR
SD	Standard deviation
SNP	Single nucleotide polymorphism
STR	Short tandem repeat
UCSC	University of California Santa Cruz

CHAPTER 1 LITERATURE REVIEW

1.1 CHIMERISM

Human genetic chimerism has traditionally been defined as a scenario in which an individual possesses DNA of more than one distinct genotype due to the fusion of multiple zygotes in a single embryo.[1] Since the first reported case of human genetic chimerism in 1962, over 30 cases meeting this definition have been reported in the literature.[2, 3] Most were identified via blood type, human leukocyte antigen (HLA), or sex chromosome discrepancies, and many expressed phenotypically ambiguous genitalia.[3]

The availability of techniques that permit the genetic discrimination of cells and DNA, and the study of chimerism more broadly, has led to the understanding that chimerism exists more commonly as a natural state, probably arising from pregnancy and intra-uterine events.[4] Transplacental movement of cells and DNA occurs from mother to child, and *vice versa*. [5] The cellular transfer includes mature cells and stem cells that may engraft and establish durable low level chimerism that persists post-partum.[6, 7] In this scenario, genetic sequences specific to the mother and absent from the child's genome are detectable in the child, and *vice versa*. [8] This cellular chimerism appears to contribute to a variety of immunological sequelae, including autoimmunity, allosensitization, and tolerance.[9-11] Additionally, the detection of foetal DNA circulating freely in the maternal plasma has permitted the development of non-invasive pre-natal testing, from which foetal genetic or chromosomal variation can be detected without invasive sampling of foetal cells.[12]

Chimerism may also be created artificially by transplantation and infused cellular therapies, such as stem cell or cytotoxic T cell therapies.[13-15] In these scenarios, DNA specific to the donor genome and absent from the recipient's genome may be detectable in the recipient.[16, 17] Whilst bone marrow cellular chimerism has been extensively studied as a mechanism for immunologic tolerance induction for solid organ transplantation, the detection of donor specific DNA sequences in plasma and urine has raised the possibility that quantification of these sequences may provide an indication of graft health.[18-20]

Consequently, a broader concept of human chimerism has emerged that is independent of a specific embryologic aetiology, but reflects common, dynamic, physiologically and clinically significant states that are the subject of investigation in a broad range of disciplines.

1.2 CHIMERISM ANALYSIS

1.2.1 Common Methods of Chimerism Analysis in Transplantation

Chimerism analysis, the detection and quantification of the chimeric state, underpins the study of chimerism. The detection of chimerism relies upon assay of specific DNA sequences or other forms of variation (“markers”) that provide discriminative capability between individuals. These markers are ideally frequent and highly variable in the population, which maximises the probability that they may discriminate individuals (“informativity”) in any particular chimeric arrangement.

Historically, chimerism analysis in transplantation has relied upon sex mismatch. Fluorescent in-situ hybridisation (FISH), PCR with qualitative amplicon gel electrophoresis, or quantitative real-time PCR (qPCR) have been used to detect Y chromosome markers indicative of male donor chimerism in female recipients.[16, 17, 20-22] Numerous Y-chromosome markers have been used, including genes located on the Y chromosome or Y chromosome specific microsatellite regions. While this approach has the advantages of specificity and quantification against a negative background (absence of Y chromosome markers in the absence of male donor chimerism), the utility of this form of chimerism analysis is limited to the male donor-female recipient transplant scenario.

The sex mismatch requirement has been avoided by approaches that target other sources of human genomic variation. Repeating elements, such as short tandem repeats (STR), were a form of genetic variation that was recognised early as potentially useful for chimerism analysis.[14, 22-24] Repeating elements, including STR, consist of repeating sequence motifs of DNA. If non-chimeric genotyping identifies that the number of repeats is discrepant between individuals, this could be used as an informative marker for chimerism analysis. PCR and gel electrophoresis of amplicon length, or qPCR of motif copy number, have been used to achieve relative quantification of the chimeric fraction after calculations involving certain assumptions. The key assumptions are that the DNA present in the sample reflects the corresponding cell content, that donor and recipient alleles should amplify in accordance with that ratio and the STR allele of the individual, and that homozygosity and heterozygosity can be accounted for.[24, 25] A critical part of this approach, therefore, was that recipient and donor genotypes are established correctly *a priori*. Methodologically, the need to discriminate one genotype from another also introduced sensitivity limitations compared to Y-chromosome markers that are quantified against a negative background.[22]

Polymorphic genes were also targeted in early chimerism analysis approaches. Transplant investigators were well aware of the highly polymorphic nature of the HLA genes, and so allele-specific PCR assay of HLA genes was used to identify chimerism following pre-transplant

serological or DNA genotyping.[20, 24, 26] In this way, sequence specific assays run on qrtPCR could be used to quantify donor informative HLA markers against a negative background.[26] Other genes targeted in chimerism analysis included apolipoprotein E (*APOE*), tyrosine hydroxylase (*THO1*) and collagen 4 (*COL4A5*)[22].

Use of polymorphic genes as informative markers for chimerism analysis avoided the requirement for a particular sex-mismatch, but were not widely adopted. Possible reasons included the need to develop gene allele-specific assays, the relatively low allelic frequencies for some alleles, the donor and recipient genotyping requirement, and potential discordance between HLA typing methods and genotyping methods. The genomics revolution, and the recognition of much more abundant and readily targeted genetic variation, also shifted the focus from gene-specific chimerism analysis to variant-specific chimerism analysis. This conceptual shift, and new multi-analyte technologies that permitted the more sensitive and precise study of genetic variation, has led to a new generation of chimerism analysis tools and opportunities.

1.2.2 Population Genomics Reveals New Targets for Chimerism Analysis

The International Human Genome Project, and the emergence of technologies that permitted large scale DNA sequencing at falling cost, led to the establishment of the human genome assembly and later projects to collate and analyse large volume genetic sequencing data from multiple individuals.[27, 28] The bioinformatic comparison of genetic sequence data from a large number of individuals within and between populations (population genomics) permitted the description, categorisation, and study of genomic variation.[29, 30]

A well-recognised and studied form of variation had been microscopic structural variants. These rare, large (>3 megabases), variants included chromosomal rearrangements and aneuploidies that could be studied under the microscope using techniques such as fluorescent in-situ hybridisation.[31] Population genomic analyses, such as those arising from the International HapMap Project and 1000 Genomes Project, identified ubiquitous non-structural variants include single nucleotide polymorphisms (SNPs), repetitive elements, and small (< 1 kilobase) insertions and deletions (indels), inversions and duplications.[31-33] The sheer number of SNPs, and the historical associations made between sequence variation and disease, meant that SNPs garnered significant interest and it was thought that this was the most common form of genetic variation in humans.[34]

Subsequent advances, such as genome-scanning array technologies and improved bioinformatic approaches, led to the improved detection and characterisation of submicroscopic structural variants.[31, 35-38] This form of variation, comprising 1 kilobase-3 megabase inversions and copy number variants (CNV, which including insertions, deletions, and duplications), had been poorly studied with earlier prior approaches that focussed on sequence variation rather than copy

number, and were confounded by the pre-amplification of DNA required for sequencing.[31] CNV has subsequently been found to account for the majority of human genetic variation, with broad functional and pathological implications.[38]

The catalogue of genomic variation that was created permits the development of chimerism analysis methods that target this variation, independent of whether it is clinically significant or located within genes. Additionally, the allelic frequency of variation was estimated from the populations studied, and this enables the strategic selection of highly polymorphic variants that are likely to be most informative in most chimeric scenarios. In particular, highly informative panels developed in this way may permit discrimination in related, well immunologically matched, or multiple donor transplant arrangements, a limitation of prior approaches.

1.2.3 Digital PCR Provides True Absolute Quantification

Digital PCR is a nucleic acid quantification technique based upon the principles of limiting dilution and Poisson statistical analysis.[39] If a sample containing target is distributed into multiple partitions containing independent target-specific PCR reactions, the proportion of these partitions that qualitatively amplify target is reflective of the underlying concentration of target in the sample. The prefix “digital” therefore refers to the binary significance of these qualitatively analysed partitions: either there is amplification of target (“positive” partition), or not (“negative” partition). Amplification of target is assessed using fluorescent indicators, for example TaqMan hydrolysis probes or fluorescent intercalating dyes.[40]

The distribution of target amongst partitions is random, and Poisson statistical analysis is used to account for the possibility that a partition may contain more than one starting copy of target.[41, 42] This probability is increased as the proportion of positive partitions increases, and can be reduced by ensuring the sample undergoing quantification has suitably low concentration. Alternatively, the dynamic range and discriminative resolution is directly related to the total number of partitions, and so approaches that generate a large number of partitions also reduce the probability that a partition may contain more than one starting copy of target. The enumeration of positive and negative partitions, and the calculation of concentration, is achieved using relatively simple mathematics.

Sykes et al. first proposed the fundamental principles of digital PCR in 1992.[39] Early platforms employing chips with microfluidic channels (Fluidigm) and plates with hydrophilic and hydrophobic surfaces (Life Technologies) implemented these principles, but the number of partitions was suboptimal for high sensitivity and precision applications.[43, 44] The next generation of platforms employed microfluidics to create thousands (Bio-Rad Laboratories) to millions (RainDance Technologies) of uniformly sized 1 nL (Bio-Rad Laboratories) or 5 pL (RainDance Technologies) water-in-oil immersion droplets.[45, 46] Each droplet is a partition

containing the necessary components of the PCR reaction, with or without template sequence depending on the random distribution of target and the target concentration. PCR is performed in all droplets *en masse*, after which the fluorescence intensity of each droplet is measured prior to enumeration of positive and negative droplets.

Digital PCR therefore permits the absolute quantification of target sequences across a large dynamic range without the need for external reference calibrators. This is in contrast to qrtPCR, which relies upon the relation of cycle thresholds in PCR reactions containing target sample to a standard curve for quantification.[47] Furthermore, pre-amplification of DNA prior to quantification is not required for digital PCR, and appears to introduce bias to the quantification result.[48] Additionally, as the quantification of target in digital PCR is based upon a qualitative assessment of the PCR reaction in the partition (positive or negative), factors affecting PCR reaction efficiency are less important than they are to qrtPCR, in which PCR reaction efficiency is central to the measurement.[49] Finally, compartmentalisation of the PCR reaction reduces confounding from contamination and background signal.[50]

Analytical validation studies and methodological comparisons have observed higher sensitivity, precision and reproducibility using digital PCR for absolute quantification compared to qrtPCR.[48, 51-53] Digital PCR has a higher discriminative resolution, being able to detect discriminate < 1.2 fold differences in target concentrations, compared to 1.25-1.5 fold differences using qrtPCR under optimal conditions.[51, 54] These characteristics make digital PCR an appealing quantification platform for chimerism analysis, permitting the absolute measurement of chimerism targets at lower concentrations and with greater precision than previously achieved.

1.3 CELLULAR CHIMERISM

The study of cellular chimerism in transplantation has focussed on a variety of compartments, cell types, and settings.

1.3.1 Allogeneic Hematopoietic Cell Transplantation

In allogeneic hematopoietic cell transplantation (HCT), a recipient's diseased bone marrow is ablated and reconstituted by hematopoietic stem cells from a healthy donor. This form of transplantation is performed for a variety of malignant and non-malignant indications, including leukaemia, lymphoproliferative disorders, multiple myeloma, bone marrow failures, haemoglobinopathies, immune deficiencies, and inherited diseases of metabolism.[55]

Using data from the Australasian Bone Marrow Transplant Recipient Registry collected between 2005-2013, the last published review of HCT in Australia and New Zealand found that allogeneic donors accounted for 32% of activity.[56] Half of these transplants used donors related to the recipient, and half were preceded by a reduced intensity conditioning regimen.[56]

Successful allogeneic HCT is accompanied by a shift from non-chimeric recipient haematopoiesis to chimeric haematopoiesis. Chimerism analysis may be performed on bone marrow or peripheral blood cells, and the cell population subject to chimerism analysis may be further subdivided by techniques such as flow cytometry.[57] Typically, chimerism analysis is performed less invasively on circulating blood cells. The chimerism of circulating cells is, therefore, assumed to be reflective of bone marrow chimerism, and so complete engraftment should be associated with complete donor chimerism in circulating cells. “Complete” or “full donor” chimerism is defined by >95% donor chimerism, whereas “mixed chimerism” is defined as 5-95% donor chimerism.[58]

The periodic assessment of peripheral blood chimerism is a routine part of clinical care following allogeneic haematopoietic cell transplantation.[57] Hematopoietic chimerism in this setting is a dynamic condition, and the clinical significance of this state varies depending on the underlying condition being treated, the conditioning regimen, and therapeutic objectives.[57] The achievement of complete chimerism may not be the objective of allogeneic HCT. One example is the use of reduced intensity conditioning with the objective of mixed chimerism for graft-versus-tumour effects in recipients not suitable for conventional conditioning due to comorbidities.[59] Another approach might permit the existence of T cell mixed chimerism due to associations with tolerance against alloantigens and a reduced risk of graft-versus-host disease.[60]

Chimerism analysis may be used to measure the adequacy of myeloablative conditioning and engraftment [61], predict the risks of graft rejection or graft-versus-host disease [62], diagnose minimal residual disease and relapse [63], and guide or monitor therapeutic interventions, which may include adjustment of immunosuppression or donor lymphocyte infusions.[58] The latter are lymphocytes obtained from the original stem cell donor that are infused into the recipient to shift the mixed chimeric state toward complete donor chimerism, which may combat graft rejection and promote graft-versus-tumour effects for the management of relapse.[64-67]

A review of methods for chimerism analysis following allogeneic HCT is presented in Chapter 6.

1.3.2 Mixed Bone Marrow Chimerism and Tolerance

Mixed bone marrow chimerism also has relevance to solid organ transplantation, due to its tolerogenic effect on shared alloantigens expressed by solid organ transplants from the same donor.

This phenomenon was first described before the concept of chimerism had been established, in a case of durable mixed bone marrow chimerism arising from stem cell transfer through the shared circulatory systems of twin bovine embryos.[68] In a report from 1945, Owen observed two

populations of red blood cell antigens shared between twins, and failure of each twin to transmit to any progeny one population of red blood cell antigens.[68] Furthermore, despite the detection of two different populations of red cell antigens in both twins, haemolysis did not occur.

The immunological tolerance of transplanted kidneys via mixed bone marrow chimerism from allogeneic HCT was subsequently demonstrated in animals and humans.[69-72] Following removal of recipient T cells, the establishment of mixed chimerism, particularly of antigen presenting cells, appears to promote central and peripheral deletion tolerance to both donor and recipient alloantigens.[18, 73, 74] The result is a significant reduction in the requirement of immunosuppression to prevent rejection, and concurrently reduced risks of graft-versus-host disease.[75]

1.3.3 Pregnancy-Associated Microchimerism

Naturally occurring chimerism commonly arises from pregnancy, and this appears to have implications for transplantation. Transplacental movement of cells and DNA occurs from mother to child, and *vice versa*. [5] The cellular transfer includes stem cells, which may engraft and establish durable low level chimerism (“microchimerism”) that persists post-partum.[6, 7] Maternal and foetal microchimerism has been detected all stages of life, in a variety of organs and cell types, and in association with health and disease.[4]

Maternal regulatory T cells are thought to play a role in tolerance induction against maternal, but not paternal, alloantigens in progeny.[76-78] This is consistent with clinical observations that (i) graft survival of sibling donor kidneys expressing non-inherited maternal antigens is significantly better than those kidneys expressing non-inherited paternal antigens, and (ii) broadly sensitized individuals are less likely to develop HLA antibodies against non-inherited maternal antigens.[79, 80] Similarly, allogeneic HCT donors exposed to non-inherited maternal antigens early in life confer a lower risk of graft-versus-host disease in sibling recipients expressing those non-inherited maternal antigens.[81] The association has also been demonstrated in animal models, although establishing the causal relationship and mechanism underpinning maternal microchimerism and tolerance remains incomplete.[9] In certain circumstances exposure to non-inherited maternal antigens may be associated with allosensitization, leading to a more complex model in which both the quality and quantity of exposure to maternal non-inherited antigens plays a role.[82]

In contrast to the tolerance-inducing effects of maternal microchimerism, foetal microchimerism in the mother during and following pregnancy may contribute to allosensitization against foetal paternal antigens, which are not expressed on trophoblast cells at the mother-child placental interface.[83] This pregnancy-associated sensitization occurs in 45-76% of pregnancies

(depending on definition of sensitization used), and may be associated with a higher risk of post-transplantation rejection, as well as reduced access to transplantation.[83]

1.3.4 Transplant-Associated Microchimerism

Solid organ transplantation, in the absence of concurrent allogeneic HCT, also appears to be associated with the establishment of microchimerism, presumably as a result of transfer and engraftment of progenitor or stem cells from the donor kidney to the recipient, and *vice versa*.

Donor microchimerism in peripheral blood cells appears to be common, occurring in more than half of kidney transplant recipients.[24, 84] The quantity of donor microchimerism varies but generally tends to be lost over time.[22, 85] The risk of kidney transplant rejection also appears to be lower in recipients with detectable donor microchimerism, although results have been conflicting and there may be sensitivity limitations arising from the chimerism analysis methods employed and the qualitative analyses performed.[24, 84, 86, 87]

Recipient microchimerism detectable within the donated allograft has also been reported. Using FISH and immunohistochemistry techniques targeting Y-chromosome, blood group and HLA-mismatched markers, Lagaaij et al. demonstrated endothelial cell recipient chimerism in kidney allografts experiencing rejection, particularly vascular rejection, perhaps reflecting reparative mechanisms in response to injury.[88] Tubular cell recipient chimerism has also been observed using laser microdissection from kidney allograft biopsies and STR or Y-chromosome FISH based analysis, with no apparent histological or functional correlates, perhaps indicating recipient precursor cell contribution of cell turnover.[89]

1.4 CELL-FREE DNA CHIMERISM

1.4.1 Cell-Free DNA

Cell-free DNA (cfDNA) is fragmented, degraded, double-stranded DNA that can be detected in extra-cellular body fluids such as plasma, urine and ascites.[90] This DNA may be nuclear, mitochondrial or microbial origin.[91, 92] Structurally, it exists predominantly as DNA bound to protein in the form of nucleosomes and virtosomes, but may also include DNA bound to vesicles in the form of apoptotic bodies, micro-particles, micro-vesicles and exosomes.[90] These structural interactions are thought to prevent complete degradation of DNA by endonuclease activity.[93] Despite significant interest in the potential clinical applications of cfDNA, knowledge with respect to the natural history (origin, variation over time, fragmentation and clearance) of cfDNA is incomplete.

CfDNA is thought to be produced as a result of active secretion, necrosis and apoptosis.[94-97] When the size distribution of cfDNA is examined, it demonstrates a characteristic ladder pattern

with a 10bp and 140-165bp periodicity reflective of nuclease cleavage of genomic DNA at accessible sites on the nucleosome.[98] This pattern of DNA fragmentation has been associated with histone occupancy of DNA and endonuclease activity targeting non-occupied DNA during apoptosis and active secretion.[97, 99] Additionally, there may be a longer smear pattern comprising larger, random fragment lengths that has been attributed to necrosis.[95, 100] This framework almost certainly over-simplifies the production of cfDNA, which likely reflects a complex interplay of related and unrelated processes affecting multiple tissues and cell types that varies in health and disease.

Irrespective of the mechanisms of production, studies that have examined the plasma fraction of Y chromosome markers in sex-mismatched bone marrow transplantation indicate that the majority of circulating cfDNA is hematopoietic in origin.[101, 102] This is supported by deep sequencing analyses, which have found that the predominance of cfDNA exhibits an epigenetic footprint matching hematopoietic lineages.[103] CfDNA levels also appear to be correlated with body mass index in pregnant women, and turnover of adipose tissue has been examined as another possible source of cfDNA.[104, 105]

Studies of placental cfDNA clearance post-partum suggest rapid clearance of cfDNA from the plasma, with a half-life of 16-60 minutes.[106-109] While cfDNA is detectable in the urine, it does not appear to be a major route of cfDNA elimination, accounting for 0.2-19% of clearance of post-partum foetal DNA in one study.[20, 106, 108, 110] CfDNA levels are not higher in dialysis patients.[111] The fragmentation profile of foetal DNA post-partum does not change, arguing against endonuclease activity as a key elimination pathway.[106] Early animal studies suggest clearance by the liver and spleen.[108] This is supported by a genome-wide association study finding that cfDNA levels in young adult Finns were associated with SNPs localised to the UGT1 polypeptide A1 (*UGT1A1*) gene region.[112] This glucuronidation enzyme is predominantly expressed in the liver.[113] Given that cfDNA chimerism assessments performed in the context of transplantation are likely to include states of organ failure, further clarification of cfDNA clearance mechanisms is required to ensure that organ-specific effects can be accounted for.

The total amount of cfDNA detectable in plasma varies significantly between individuals and over time, and may be affected by a variety of physiological events, in addition to pathology. Total cfDNA levels may increase with age and as a result of exercise and smoking.[94, 114-116] Total cfDNA levels are also increased in patients with malignancy, inflammatory conditions and critical illness, and some have proposed prognostic or staging applications.[117-119] In cancer patients, individuals with active lupus and during placental pathology, the fragmentation profile of cfDNA also changes, suggesting alterations to cfDNA production or clearance.[100]

In is on this highly variable background of physiologic cfDNA production by the hematopoietic system, supplemented by physiologic stressors and pathology, that chimeric states are detected and quantified. These properties of the cfDNA baseline have been identified as potential confounders to the assessment of chimerism in a relative sense (i.e. results expressed as a proportion of total cfDNA), as fluctuation in the total cfDNA concentration may obscure or pronounce variation in the chimeric component.[120]

1.4.2 Non-Transplant Settings

In pregnancy, plasma cfDNA was found to include DNA of placental/foetal origin in addition to DNA of maternal origin.[12] This finding has led to rapid development of a range of non-invasive prenatal diagnostic tests for genetic mutations and foetal aneuploidy which are increasingly becoming standard of care and reduce reliance upon invasive amniocentesis and chorionic villus sampling.[121] Additionally, quantification of cfDNA and the fraction arising from the placenta (foetal fraction), is being investigated as an indicator of placental health that precedes clinical events such as pre-eclampsia.[122]

The concept of tumour-derived cfDNA has also been studied in cancer diagnostics and surveillance.[123] Approaches have included genotyping circulating tumour-derived cfDNA for cancer-specific mutations, studies targeting cancer-specific methylation and microsatellite alterations, and fragmentation studies examining the size and distribution of total cfDNA in malignancy.[124] With increasing availability of targeted biologic therapies, the promise of non-invasive serial “liquid biopsy” and simultaneous genotypic analysis of tumour-specific cfDNA continues to drive research in this field.[118]

1.4.3 Allogenic Hematopoietic Cell Transplantation

There has been very limited study of the nature and clinical role of cfDNA chimerism analysis after HCT.

Duque-Afonso et al. studied the characteristics of cfDNA in 110 recipients of allogenic HCT using STR-PCR.[125] Plasma cfDNA chimerism was detected in 41/110 patients, while complete donor cellular chimerism was present in 107/110 (all with mixed cellular chimerism had relapsed disease). Plasma recipient cfDNA chimerism was lost with time in stable recipients, which is consistent with the hypothesis that the majority of plasma cfDNA is produced by the hematopoietic system. Recipient cfDNA chimerism increased during episodes of graft-versus-host disease, presumably as a result of recipient tissue injury, but numbers were insufficient to evaluate diagnostic or prognostic performance. Low level cfDNA chimerism may have also have been missed in this cohort, due to the limited sensitivity of this method.

In a study of 84 adults who underwent HCT for leukaemia, Aljurf et al. observed that recipients with 100% donor chimerism in circulating leucocytes but < 90% donor chimerism in plasma cfDNA were at higher risk of relapse.[126] This indicates that relapse may be detected sooner via loss of cfDNA, compared to cellular DNA, donor chimerism. This was a small retrospective observational study that requires further validation.

1.4.4 Solid Organ Transplantation

Published original articles concerning chimerism analysis in the plasma and urine of solid organ transplant recipients are summarised in Table 1-1 and Table 1-2.

Detection of graft-derived cfDNA (also called donor-derived cfDNA) in plasma was first demonstrated by Lo et al. in 1998 using Y chromosome markers in female recipients of male kidney and liver transplants.[127] Following this, a number of groups replicated this sex-mismatch approach on plasma from kidney, liver, pancreas, pancreatic islet and heart transplants.[21, 22, 128-132] Studies in sex-mismatched kidney transplant recipients also detected graft-derived cfDNA in urine using this approach.[20, 130, 133-137] Assayed Y chromosome markers include *Y1.7*, *Y1.8*, *SRY*, *TSPY1* and *DYZ1/2*. Other early approaches included short tandem repeat and microsatellite loci (*STR at Chr21*, *STR at PA2*, *cP450*, *D1S80*, *ChY microsatellite locus*) and direct assay of human leucocyte antigen mismatches (*HLA DRB1* and other custom designed assays for each donor-recipient pair).[22, 26, 134, 135, 138, 139] Many of these early studies were qualitative and designed to demonstrate proof of concept, reporting low or undetectable levels in the absence of rejection and elevated levels in the presence of rejection. FISH and qrtPCR were the predominant platforms.

Only one of these studies examined quantified plasma cfDNA with respect to the diagnosis of rejection in human kidney transplant recipients. Garcia Moreira et al. quantified total cfDNA (haemoglobin beta, *HBB*) in 100 and graft-derived cfDNA (testis specific protein Y-linked, *TSPY1*) in 17 female kidney transplant recipients during the first 3 months after transplantation.[130] The normal range for total cfDNA was 120-1800 genomic equivalents (GE)/mL. Using a cut-off plasma total cfDNA level of 12,000 GE/mL, acute rejection was distinguished from non-acute rejection with a sensitivity of 89% and specificity of 85%. Specificity was particularly affected by infection and total cfDNA performed better when procalcitonin was used concurrently, increasing specificity to 98%. BK nephropathy was not reported as an observed complication. Given the small sample size in the graft-derived cfDNA quantification group and only three episodes of rejection, the diagnostic performance of graft-derived cfDNA not specifically determined, however, qualitative review suggested it was similar to that of total cfDNA.

Another study by Sigdel et al. examined the diagnostic performance of absolutely quantified urine graft-derived cfDNA for the diagnosis of rejection in female recipients of male kidney transplants.[137] Urine sex-determining region (SRY) concentration was measured using microfluidic digital PCR and log2 transformed then normalised against urine creatinine concentration as a measure of graft-derived cfDNA. Levels of graft-derived cfDNA were significantly higher in recipients with acute rejection and BK nephropathy compared with normal biopsy histopathology (acute rejection 20.5, BK 20.3, normal 2.4 copies/uL Cr; acute rejection vs normal p -value < 0.0001). This approach could not distinguish BK from acute rejection nor calcineurin inhibitor toxicity from normal. No additional pathologies were observed so the specificity of the assay was not tested further. Area under the curve for acute allograft injury (including acute rejection and BK) was 0.8 (95% CI 0.67-0.93, p -value < 0.0006).

The early identification of a relationship between graft-derived cfDNA levels and graft pathologies led to the theory that graft-derived cfDNA is liberated from the allograft as a result of graft cell apoptosis and necrosis.[130] Graft-derived cfDNA may therefore rise prior to allograft dysfunction, permitting earlier diagnosis, and may provide prognostic and treatment response information. The promise of clinical utility, combined with the genomics revolution and access to technologies enabling the specific detection of small quantities of cfDNA, has led to the development of a “second-generation” of chimerism analysis methodologies exploiting human genetic variation. These are best considered grouped by quantification platform, digital PCR or next generation sequencing.

1.4.4.1 Digital PCR Approaches

There are five published digital PCR cfDNA chimerism analysis methodologies targeting human genetic variation, all of which employ the Bio-Rad QX200 droplet digital PCR (ddPCR) platform. A study comparing the performance of digital PCR systems concluded that the Bio-Rad system performed best for chimerism analysis.[140]

Two digital PCR approaches target sequence variation. The Chronix Biomedical assay targets 41 highly polymorphic SNPs to calculate the graft chimeric fraction, which is subsequently multiplied by a separate digital PCR measurement of total cfDNA to calculate the graft-derived cfDNA concentration in absolute terms.[141] This is the most studied digital PCR chimerism analysis method. After proof-of-principle application to kidney, liver and heart transplant recipients, further diagnostic validity studies have been conducted in longitudinal cohorts of liver and kidney transplant recipients.[120, 142, 143] Another sequence variation approach was proposed by Zou et al. targeting 8 HLA-DR alleles in a proof-of-principle study in lung transplant recipients to calculate the graft chimeric fraction.[144] In both of these implementations, cfDNA

undergoes pre-amplification prior to quantification, and both primary measures are, therefore, relative.

The remaining three digital PCR chimerism analysis methods target deletion variants. The first approach, the development and validation of which is presented in this thesis, targets 31 highly polymorphic CNV loci which uniquely permits the measurement of graft-derived cfDNA and total cfDNA, from which graft fraction is calculated. Proof-of-principle was published by my supervisors in 2014 using a 10 CNV locus assay panel and qrtPCR in kidney transplant recipients and pregnancy, following which diagnostic validity has been established in kidney transplant recipients using the newly developed methodology presented in this thesis.[145, 146] Ironically, a second approach was developed concurrently by a different group in our institution targeting 9 indel polymorphisms using a probe-free approach, permitting the measurement of graft-derived cfDNA but not total cfDNA in a proof-of-principle cohort of 3 liver transplant recipients.[147] A third 10 indel approach was demonstrated in a proof-of-principle study in heart transplant recipients using a probe-based approach.[148] These indel approaches include the separate (non-chimeric) genotyping of recipient and donor before chimerism analysis can be performed, in contrast to the CNV approach presented in this thesis.

Both indel and CNV approaches detect variations in the copy number of target DNA, specifically selecting informative markers for the graft that are absent in the recipient genome. This “negative background” plays to the strengths of digital PCR, maximising signal-to-noise ratio with improved sensitivity at low levels of target. Additionally, both approaches do not rely upon pre-amplification of DNA, which may introduce biases and precludes true absolute quantification.[52] In contrast, both approaches targeting sequence variation employ pre-amplification and must discriminate the donor signal from the recipient signal at the same locus to calculate the relative measure of graft chimeric fraction. Whilst this is readily achieved using digital PCR, it can be better achieved using next generation sequencing, albeit at greater cost and complexity.

1.4.4.2 Sequencing Approaches

The falling cost and growing availability of high throughput sequencing has led to a number of chimerism approaches based upon this technology. Broadly speaking, these approaches may target highly polymorphic bi-allelic SNPs selected strategically on the basis of population genomics, or they may target SNPs identified as discrepant specifically in the donor-recipient arrangement under study. No sequencing approaches targeting other genetic variation have been described.

The first method to use this platform was developed by a team at Stanford University with funding from the National Institutes of Health. This technique employs pre-transplant whole-genome

sequencing of donor and recipient using massively parallel sequencing to identify, on average, 53,423 discordant SNPs specific to the donor-recipient arrangement from which to measure graft fraction during the post-transplant period (Genome Transplant Dynamics).[131] Diagnostic validity was subsequently studied in heart transplant recipients, and proof-of-principle prognostication and diagnosis of infection was examined in lung transplant recipients.[149-151]

The second method to use this platform was developed by CareDx to target 266 strategically selected highly polymorphic bi-allelic SNPs that, combined with the assumption that the graft fraction is typically low, circumvents requirement of non-chimeric donor and recipient genotyping (AlloSure). After a comprehensive analytical validation, diagnostic validity was studied in kidney transplant recipients.[152, 153] The company was able to use this preliminary data to obtain Centers for Medicare and Medicaid Services coverage of the assay in the United States at a reimbursement cost of US\$2841 per test, with Medicare coverage extending to 80% of kidney transplant recipients in the United States (CareDx Corporate Presentation, March 2018). From the announcement of Medicare coverage of the test to Jun 2019, the CareDx share price has risen 3600%. Subsequent to the introduction of cost-subsidised testing, internal and external diagnostic validation studies have been performed in kidney transplant recipients.[154, 155] More recently, diagnostic validity has also been studied in heart transplantation.[156]

Natera, another American diagnostics company, has recently published analytical validation and a diagnostic validity study in kidney transplantation of its own method for sequencing-based post-transplant chimerism analysis using 13962 bi-allelic SNPs that undergo targeted pre-amplification (Prospera).[157, 158] A similar targeted pre-amplification approach was also proposed by Gielis et al. using 1027 SNPs and studied in kidney transplant recipients.[159, 160] These approaches achieve greater within and between run precision using massively multiplex PCR rather than non-targeted amplification during standard library preparation.

These sequencing-based techniques achieve chimerism analysis via donor and recipient informative SNP discrimination in pre-amplified cfDNA using complex modelling and bioinformatics approaches. They have the advantages of high informativity, throughput, sensitivity, and precision. The complexity of the methodology means that turnaround time is likely to be longer than digital PCR, which is substantially simpler and less costly. Additionally, these methods are subject to pre-amplification, sequencing and bioinformatics bias. Finally, the primary measure of these techniques is graft fraction, which may be subject to confounding from total cfDNA as previously discussed.

1.4.4.3 Diagnostic Validity Studies

CfDNA chimerism analysis has been studied in kidney, heart, lung and liver transplantation. Diagnostic validity studies support the concept that graft-derived cfDNA may be used to diagnose

allograft injury, however, the diagnostic thresholds and clinical significance of this analyte appear to differ between transplanted organs and the histopathological classification of rejection applied to that organ. With the exception of the CareDx Allosure test in kidney transplantation, to date no other chimerism analysis diagnostic validity studies have been subject to internal or external validation using a different cohort from the original diagnostic validity study.

1.4.4.3.1 Heart

Using the Genome Transplant Dynamics approach, De Vlamincx et al. prospectively studied 65 heart transplant recipients longitudinally.[149] Graft fraction measured in plasma samples taken concurrently with endomyocardial biopsy permitted the diagnosis of moderate and severe acute cellular rejection or antibody mediated rejection with a receiver operating characteristic (ROC) area under the curve (AUC) of 0.83. If mild cellular rejection was included, the AUC dropped to 0.75. The graft fraction correlated with rejection severity.

More recently, CareDx conducted a prospective longitudinal study of 740 heart transplant recipients from 26 American transplant centres receiving Allosure graft fraction testing.[161] Of 841 endomyocardial biopsies performed in this cohort, there were 17 diagnoses of moderate to severe acute cellular rejection, and 18 diagnoses of antibody mediated rejection. Plasma graft fraction sampled on the day of biopsy differentiated moderate and severe acute cellular rejection, and antibody mediated rejection, from normal biopsies. Mild acute cellular rejection could not be discriminated from normal biopsies. The ROC AUC for diagnosis of all rejection was 0.64.

1.4.4.3.2 Lung

Genome Transplant Dynamics graft fraction testing has also been performed in a prospective multicentre cohort study of 106 lung transplant recipients. The average graft fraction during the first three months after transplantation was highly variable, ranging between 0.1% and 9.9%. Recipients in the top tertile of graft fraction had a hazard ratio for allograft failure, defined as severe chronic lung allograft dysfunction, retransplantation, or death from respiratory failure, of 6.6 (95% CI 1.1-19.9, $p = 0.001$) as well as lower peak forced expiratory volume in 1 second measurements. Fewer than half had detectable clinical complications during the early post-transplant period. The diagnostic capability of measurements for diagnosis of rejection was not assessed. The authors proposed that the average graft fraction in the first three months was a reflection of early allograft injury that begat subsequent allograft injury.

1.4.4.3.3 Liver

Schutz et al. conducted a prospective study using the Chronix Biomedical ddPCR method to evaluate graft fraction for assessment of graft injury in 115 liver transplant recipients at three German transplant centres. Compared to 282 samples collected during periods of stability, 31

samples coinciding with biopsy proven-acute rejection exhibited significantly elevated graft fraction. The AUC for rejection diagnosis was 0.97 (95% CI 0.93 – 0.1), and a graft fraction diagnostic threshold of 10% was proposed (sensitivity 90%, specificity 93%). Interestingly, concurrently measured cellular microchimerism in circulating leucocytes was less than or equal to 0.068%.

1.4.4.3.4 Kidney

Kidney transplantation has been a major focus in the development and validation of cfDNA chimerism analysis methods, primarily in the plasma compartment. A number of groups have recently published prospective diagnostic validity studies, including our own (Table 1-1). With exception of the Chronix Biomedical ddPCR method, these chimerism analysis methods were developed and validated concurrently with this PhD. To better contextualise this work in time, an overview of the diagnostic significance of cfDNA chimerism analysis in kidney transplantation is therefore presented in Chapters 5 and 7.

The level of interest in urinary cfDNA chimerism analysis has been more muted, despite its ease of collection (Table 1-2). This may reflect knowledge gaps regarding the integrity, storage and preservation of cfDNA collected from urine.[162] Additionally, urine concentration and volume vary significantly, and may be affected by kidney function. Garcia Moreira et al. observed, using Y-chromosome markers at qrtPCR, that urine cfDNA chimerism was more poorly associated with pathology than that in plasma.[130] Using microfluidic chip digital PCR and Y-chromosome markers, Sigdel et al. identified higher levels of urine graft-derived cfDNA in rejection and BK nephropathy, and could not discriminate between these diagnosis.[137] Perhaps as a result of these findings, no formal diagnostic validity studies have been performed on urine, despite the numerous “second generation” chimerism tests presently under investigation in plasma.

The significant interest in cfDNA chimerism in the kidney transplant context may reflect the prevalent population of kidney transplant recipients and incident activity, the rate of rejection in kidney transplantation and its implication for long term outcomes, and the inadequacies of existing strategies for the detection of rejection.

Table 1-1: Published original articles concerning plasma cfDNA chimerism following solid organ transplantation.

Publication	Organ	Discriminative Target	Genotyping <i>a priori</i>	Non-Discriminative Target	Method	Measure	Outcome
Lo et al. 1998 [16]	K, Li	Y-chromosome gene (<i>DYS14</i>)	-		PCR & gel electrophoresis	Qualitative	Demonstrate chimerism
Lui et al. 2003 [21]	K, Li, H	Y-chromosome gene (<i>SRY</i>)	-	β -globin gene (<i>HBB</i>)	qrtPCR	Quantitative, fraction of total cfDNA	Demonstrate that allograft accounts for a small amount

							of total cfDNA
Fu et al. 2006 [129]	K	Y-chromosome gene (<i>SRY</i> , <i>DYZ-1</i>)	-		PCR & gel electrophoresis	Qualitative	Detect chimerism in a longitudinal cohort, compare detectable chimerism with long-term graft outcomes
Gadi et al. 2006 [26]	K+P	HLA genes (<i>DRB1*01, *11; DQA1*01, *03; DQB1*02, *03; DRB4*01; B*44</i>)	Y	β -globin gene (<i>HBB</i>)	qPCR	Quantitative, GE/mL plasma	Measure gd-cfDNA in a longitudinal cohort and compare to biopsy in a subgroup
Garcia Moreira et al. 2009 [130]	K	Y-chromosome gene (<i>TSPY1</i>)	-	β -globin gene (<i>HBB</i>)	qPCR	Quantitative, GE/mL plasma	Measure gd-cfDNA in a longitudinal cohort and compare to clinical events
Snyder et al. 2011 [131] (Genome Transplant Dynamics)	H	Y-chromosome sequence-specific probes	-	chromosome 1 sequence-specific probes	Microfluidic digital PCR	Quantitative, fraction of total cfDNA	Proof of principle demonstration of chimerism analysis using a new approach, measure graft fraction in a longitudinal cohort and compare to clinical events including biopsy
		~25,000 non-strategic SNPs	Y (recipient and donor informative homozygous)		Shotgun sequencing	Quantitative, fraction of total SNP reads	
Beck et al. 2013 [141] (Chronix Method)	K, Li, H	41 strategic SNPs	Y (recipient homozygous, donor informative heterozygous or homozygous)		Droplet digital PCR	Quantitative, fraction of total cfDNA	Proof of principle demonstration of chimerism analysis using a new approach
Bruno et al. 2014 [145] (MCRI-1 Method)	K	10 strategic CNVs	Y (recipient)		qPCR	Quantitative, GE/mL plasma	Proof of principle demonstration of chimerism analysis using a new approach
De Vlamincx et al. 2014 [149] (Genome Transplant Dynamics)	H	~50,000 non-strategic SNPs	Y (recipient and donor informative homozygous)		Shotgun sequencing	Quantitative, fraction of total cfDNA	Measure graft fraction in a longitudinal cohort, assess diagnostic validity for rejection in biopsy matched samples
Macher et al. 2014 [132]	Li	Y-chromosome gene (<i>SRY</i>)	-	β -globin gene (<i>HBB</i>)	qPCR	Quantitative, ng/mL plasma	Measure gd-cfDNA in a longitudinal cohort and compare to clinical events
Oellerich et al. 2014 [142]	Li	41 strategic SNPs	Y (recipient homozygous, donor		Droplet digital PCR	Quantitative, fraction of total cfDNA	Measure gd-cfDNA and tacrolimus

(Chronix Method)			informative heterozygous or homozygous)				levels in a longitudinal cohort, and relate the measures
De Vlaminck et al. 2015 [150] (Genome Transplant Dynamics)	Lu	~50,000 non-strategic SNPs	Y (recipient and donor informative homozygous)		Shotgun sequencing	Quantitative, fraction of total cfDNA	Proof-of-principle in lung transplant cohort, concurrent virome sequencing to detect infection
Adamek et al. 2016 [163]	K	12 indels, Y-chromosome gene (SRY)	Y (recipient and donor PCR & gel electrophoresis amplicon size)	β -globin gene (HBB)	qPCR	Quantitative, GE/mL (total and graft), fraction of total cfDNA	Proof of principle demonstration of chimerism analysis using a new approach
Gordon et al. 2016 [164]	Li, H	124 SNPs	N (automatic calls based on mixture modelling)		Next generation sequencing	Quantitative, fraction of total cfDNA	Proof of principle SNP genotyping based on mixture modelling
Grskovic et al. 2016 [152] (CareDx AlloSure)	K	266 strategic SNPs	N		Next generation sequencing	Quantitative, fraction of total cfDNA	Analytical validation of method
Macher et al. 2016 [132]	Li	Rhesus gene (RH)	Y	β -globin gene (HBB)	qPCR	Quantitative, copies/mL plasma	Proof of principle demonstration of chimerism analysis using a new approach
Bloom et al. 2017 [153] (CareDx Allosure)	K	266 strategic SNPs	N		Next generation sequencing	Quantitative, fraction of total cfDNA	Diagnostic validity of method
Goh et al. 2017 [147]	Li	9 indels, probe-independent (amplicon length)	Y (high-resolution melting analysis)		Droplet digital PCR	Quantitative, copies/mL plasma	Proof of principle demonstration of chimerism analysis using a new approach
Lee et al. 2017 [140]	K	30 SNPs	Y		Multiple digital PCR systems	Quantitative, fraction of total cfDNA	Comparison of digital PCR platforms for chimerism analysis, Bio-Rad QX200 performed best
Schutz et al. 2017 [143] (Chronix Method)	Li	41 strategic SNPs	Y (recipient homozygous, donor informative heterozygous or homozygous)		Droplet digital PCR	Quantitative, fraction of total cfDNA	Diagnostic validity of method
Zou et al. 2017 [144]	Lu	8 HLA-DR alleles (DRB1*01, *03, *04, *07, *08, *11, *13, *15)	Y		Droplet digital PCR	Quantitative, fraction of total cfDNA	Proof of principle demonstration of chimerism analysis using

							a new approach
Gala-Lopez et al. 2018 [165]	I	Beta-cell specific epigenetic pattern		NS	Next generation sequencing	Quantitative, fraction of total cfDNA, multiplied against total to get graft GE/mL	Proof of principle demonstration of chimerism analysis using a new approach
Gielis et al. 2018 [159] (Belgian Method)	K	1027 SNPs	Y (recipient and donor)		Single tube highly multiplexed PCR followed by next-generation sequencing	Quantitative, fraction of total cfDNA	Proof of principle demonstration of chimerism analysis using a new approach, description of natural history in a longitudinal cohort
Jordan et al. 2018 [154] (CareDx Allosure)	K	266 strategic SNPs	N		Next generation sequencing	Quantitative, fraction of total cfDNA	Observational study relating gd-cfDNA, biopsy and DSA testing for diagnosis of AMR, internal validation of method
Whitlam et al. 2019 [146] (MCRI-2 Method)	K	31 CNV	N		Droplet digital PCR	Quantitative, GE/mL (total and graft), fraction of total cfDNA	Diagnostic validity of method
Oellerich et al. 2019[120] (Chronix Method)	K	41 strategic SNPs	Y using urine (recipient homozygous, donor informative heterozygous or homozygous)	Y, not specified	Droplet digital PCR	Quantitative, fraction of total cfDNA, multiplied against total to get graft GE/mL	Diagnostic validity of method
Shen et al. 2019 [166]	K	6200 SNPs, infection sequences	Y		Next generation sequencing	Quantitative, fraction of total cfDNA	Longitudinal cohort, study of kinetics / variation by transplant type
Agbor-Enoh et al. 2019 [151] (Genome Transplant Dynamics)	Lu	~50,000 non-strategic SNPs	Y (recipient and donor informative homozygous)		Shotgun sequencing	Quantitative, fraction of total cfDNA	Longitudinal cohort, relationship between graft fraction in the first 3 months and allograft failure
Altug et al. 2019 [157] (Natera Prospera)	K	13962 SNPs	N (probability model)		Massively multiplexed PCR followed by sequencing	Quantitative, fraction of total cfDNA	Analytical validation of method
Gielis et al. 2019 [160] (Belgian Method)	K	1027 SNPs	Y (recipient and donor)		Single tube highly multiplexed PCR followed by next-generation sequencing	Quantitative, fraction of total cfDNA	Diagnostic validity of method
Huang et al. 2019 [155]	K	266 strategic SNPs	N		Next generation sequencing	Quantitative, fraction of total cfDNA	External diagnostic validation of method

(CareDx Allosure)							
Khush et al. 2019 [161]	H	266 strategic SNPs	N		Next generation sequencing	Quantitative, fraction of total cfDNA	Diagnostic validity of method
(CareDx Allosure)							
Macher et al. 2019 [148]	H	10 indels	Y (recipient blood and donor biopsy)	β -globin gene (<i>HBB</i>)	Droplet digital PCR	Quantitative, GE/mL (total and graft)	Proof of principle demonstration of chimerism analysis using a new approach
Sigdel et al. 2019 [158]	K	13962 SNPs	N (probability model)		Massively multiplexed PCR followed by sequencing	Quantitative, fraction of total cfDNA	Diagnostic validity of method
(Natera Prospera)							

GE = genomic equivalents, indel = insertion-deletion polymorphism, MALDI-TOF MS = matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, Organs (H = heart, I = islet, K = kidney, Li = liver, Lu = lung, P = pancreas), STR = short tandem repeats, SNP = single nucleotide polymorphism

Table 1-2: Published original articles concerning urine cfDNA chimerism following solid organ transplantation.

Publication	Organ	Discriminative Target	Genotyping <i>a priori</i>	Non-Discriminative Target	Method	Measure	Outcome
Zhang et al. 1999 [20]	K	Y-chromosome gene (<i>SRY</i>)	-	β -globin gene (<i>HBB</i>)	qPCR	Quantitative, GE/mmol creatinine	Demonstrate chimerism, fluctuation in rejection
Zhong et al. 2001 [133]	K	Y-chromosome gene (<i>SRY</i>)	-	β -globin gene (<i>HBB</i>)	qPCR	Quantitative, DNA/mL urine	Demonstrate chimerism in transplantation but not pregnancy
Li et al. 2003 [135]	K	4 STRs genotyped	Y		Nested PCR & gel electrophoresis	Qualitative	Proof of principle demonstration of chimerism analysis using a new approach
Zhang et al. 2003 [136]	K	Y-chromosome genes (<i>SRY</i> , <i>DYZ-1</i>); HLA DR gene (<i>DRB1</i>)	Y		Nested PCR & gel electrophoresis	Qualitative	Demonstration of chimerism during rejection with absence of chimerism after treatment
Li et al. 2006 [167]	K	20 SNPs	Y		MALDI-TOF MS	Qualitative	Proof of principle demonstration of chimerism analysis using a new approach
Garcia Moreira et al. 2009 [130]	K	Y-chromosome gene (<i>TSPY1</i>)	-	β -globin gene (<i>HBB</i>)	qPCR	Quantitative, GE/mL plasma	Measure gd-cfDNA in a longitudinal cohort and compare to clinical events; "use of urine cfDNA did not add relevant information"

Sigdel et al. 2013 [137]	K	Y-chromosome gene (SRY)	-		Microfluidic digital PCR	Quantitative, copies/ μ g creatinine	Measure chimerism of biopsy-matched urine samples to assess diagnostic validity of donor-derived cfDNA
Lee et al. 2017 [140]	K	30 SNPs	Y		Multiple digital PCR systems	Quantitative, fraction of total cfDNA	Comparison of digital PCR platforms for chimerism analysis, Bio-Rad QX200 performed best

GE = genomic equivalents, indel = insertion-deletion polymorphism, MALDI-TOF MS = matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, Organs (H = heart, I = islet, K = kidney, Li = liver, Lu = lung, P = pancreas), STR = short tandem repeats, SNP = single nucleotide polymorphism

1.5 KIDNEY TRANSPLANTATION

Kidney transplantation (KT) is the treatment of choice for most individuals with end-stage renal disease. Quality of life, survival and healthcare costs are all improved with KT compared with maintenance dialysis.[168] The prevalent population of KT recipients is rising and there are currently more than 10,000 functioning kidney transplants in Australia and New Zealand.[169] Despite modern immunosuppression regimes, long-term graft survival remains a problem. 10 year graft survival in Australia and New Zealand is 50-72% while 20 year graft survival is 20-30%.[169] A leading factor in determining long-term graft function and survival is immunological rejection.[170]

1.5.1 The Problem of Kidney Transplant Rejection

Approximately 20% of kidney transplant recipients experience rejection.[169] Progressive graft damage leading to eventual graft failure is correlated with the number of rejection episodes during the first year after transplantation.[171] Timely detection and treatment of rejection is, therefore, an important goal of post-transplant management. Rejection can occur at any time after transplantation but is commonest in the first twelve months.

The classification of rejection is based upon histological features defined by the Banff Working Classification of Renal Allograft Pathology.[172, 173] Diagnostic application of the classification is facilitated through a consensus grading scheme comprising semi-quantitatively assessed “Banff scores” representing specific histopathological lesions denoted by letters.[172] Rejection is diagnosed and classified based upon specific combinations of these lesions and supporting clinical findings. Broadly, rejection may manifest histopathologically as cellular mediated rejection or antibody mediated rejection.

Cellular mediated rejection (CMR) was the first pattern of rejection included in the Banff classification.[174] It is characterised by a predominantly CD4+ or CD8+ T lymphocyte interstitial cellular infiltrate (i, scored on extent) and related inflammation of the renal tubules (tubulitis, t, scored on the number of mononuclear cells in the renal tubules) and arteries (intimal arteritis, v, scored on severity).[175] The relative scoring of each of these lesions further sub classifies CMR into type I (tubulointerstitial rejection without arteritis), type II (vascular rejection) and type III (severe rejection).[175] The incidence and prognosis decrease with each successive CMR type.[176]

Whilst the histologic pattern of inflammation defined as CMR is in routine clinical use, the pathophysiology and underlying immunologic mechanisms remain under investigation. A proposed conceptual model identifies recipient effector T cells encountering donor antigens on either recipient or donor antigen presenting cells in the allograft as the inciting immunologic event.[177] Chemokine-mediated increased endothelial permeability and effector CD4+ and CD8+ T cell recruitment ensues.[177] Critically, interstitial inflammation appears to trigger dedifferentiation of tubular cells rather than apoptosis or necrosis.[178] Apoptosis can be seen in infiltrating recipient mononuclear cells and it may have prognostic significance.[179, 180] Tubular dedifferentiation is characterised by reduced expression of genes associated with normal cellular function and increased expression of embryonic, cell cycle and injury genes.[181, 182] This process appears to occur independently of B lymphocytes, alloantibodies and effector T lymphocyte cytotoxic mechanisms including CD103/E-cadherin binding, perforin and granzyme.[183, 184] Tubular dedifferentiation is eventually associated with loss of cellular polarity and reduced expression of cell-cell adhesion molecules, disrupting integrity of the epithelium and basement membrane, and allowing lymphocyte entry into the tubular lumen (tubulitis).[185] The mechanism by which this process leads to graft dysfunction and fibrosis is unclear. One hypothesis proposes that the immune response perturbs tubular function, leading to reduced glomerular filtration and, later, tubular atrophy and fibrosis.[177] CMR is also thought to play a role in the development of alloantibodies, increasing the risk of antibody mediated rejection.[186]

Antibody mediated rejection (AMR) was added to the Banff classification in 2003 following technological advances that allowed detection of serum donor-specific antibodies (DSAs) and evidence of antibody-mediated complement activation in tissue.[187] The characteristic lesion includes microvascular injury at the glomerulus (glomerulitis, g, graded on extent) or peritubular capillary (peritubular capillaritis, ptc, scored on the number of luminal inflammatory cells) as well as evidence of antibody activity at the vascular endothelium and the presence of DSAs.[187] Additional patterns of injury have been described, including thrombotic microangiopathy, acute tubular injury and arteritis, but these are less common.[188]

Like CMR, the mechanisms of allograft injury in AMR remain under investigation. Pre-existing or *de novo* DSAs have been implicated as the inciting agent.[189] Whilst these are usually directed against major histocompatibility complex molecules, other targets such as the minor histocompatibility antigens, ABO blood group antigens, angiotensin II receptors and endothelial antigens have been identified.[190-193] In acute AMR, direct injury to graft microvascular endothelium is thought to be mediated by antibody-antigen complex-mediated complement activation and antibody-dependent cell-mediated cytotoxicity.[194] Antibody-antigen complex-mediated classical complement pathway activation leads to the production of proinflammatory molecules and endothelial cell activation in addition to cell lysis by the membrane attack complex.[194] Whilst the role of the membrane attack complex has only been demonstrated in lung transplant AMR, successful KT AMR treatment and prophylaxis with anti-C5 monoclonal antibody (eculizumab) support a role in KT AMR as well.[195-198] In antibody-dependent cell-mediated cytotoxicity, bound antibody interacts with low-affinity Fc IgG receptors (CD16) on the surface of natural killer cells and macrophages, signalling release of perforin and granzyme B that induce apoptosis of the target cell.[199] The relative contributions of each of these injurious processes to allograft damage in AMR is not known. Further, the relationship between pathophysiology in acute AMR and chronic AMR remains subject to investigation.

CMR rates have improved significantly since the introduction of modern immunosuppressive regimes and this has correlated with improved 1 year graft survival but not 10 year graft survival.[200, 201] AMR is increasingly diagnosed and recognised as a key factor in long-term graft survival.[186, 202, 203] This may in part be reflective of poorer efficacy of current rejection prophylaxis strategies in suppressing the immunologic mechanisms of AMR, particularly the core role of antibody. Additionally, due to cost and toxicity, effective AMR interventions are difficult to implement prophylactically over the long term.[201] Consequently, early diagnosis of rejection remains an important management strategy, both to prevent sensitisation and *de novo* DSA production in CMR, as well as to limit graft damage and further *de novo* DSA production in AMR.

1.5.2 Current Strategies for Rejection Surveillance and Diagnosis

Regular measurement of serum creatinine is used to monitor KT health, however, this has limited sensitivity and specificity.[204] In the first 12 months post-KT, rejection is identified in up to 40% of fixed time point routine KT biopsies performed when serum creatinine levels are normal (so-called “protocol biopsies”).[205-207] This subclinical rejection is associated with poorer graft outcomes and early treatment appears to be beneficial.[208-211] Furthermore, as an indirect marker of glomerular filtration rate, it is expected that sufficient tissue injury must occur to reduce renal function before serum creatinine rises, limiting opportunities for early intervention. When serum creatinine does increase acutely, the diagnosis of rejection needs to be distinguished from

a variety of other conditions such as drug toxicity, acute tubular necrosis, intercurrent illness, dehydration and viral infections (such as BK virus nephropathy).

Existing diagnostic classification criteria for rejection mandates histopathological examination of KT biopsy tissue.[174] KT biopsy is typically performed under ultrasound guidance using 14G-18G needle biopsy equipment.[212, 213] Frequent sampling is limited by procedural risks, patient discomfort and resource implications. Risks include haemorrhage requiring intervention (up to 6%), sampling error and damage to the graft necessitating removal (<1%).[214, 215] Sampled kidney tissue is processed and examined by a pathologist in accordance with the Banff Working Classification of Renal Allograft Pathology described briefly above.[174]

Whilst the Banff Classification has facilitated inquiry into the pathophysiology of rejection and stratification of appropriate treatment modalities, limitations exist with respect to its application as a diagnostic test. Clinical application is by means of a semi-quantitative lesion scoring system that is subject to suboptimal reproducibility.[216-218] In a study involving 22 major transplant centres in Europe, 11 sets of 5 KT biopsy cases were reviewed over a two year period by 23 pathologists. For most histologic features studied, inter-observer kappa statistics were less than 0.5 (moderate agreement) including a kappa statistic of 0.17 for the tubulitis score (t) and kappa statistic of 0.193 for glomerulitis score (g).[219] Furthermore, a follow-up study performed after comprehensive numerical and graphical feedback to participants failed to demonstrate meaningful improvement in inter-observer agreement. The kappa statistic for the glomerulitis score (g) was 0.195 (improvement $p=0.980$) and for the tubulitis score (t) 0.214 ($p=0.056$).[216]

The Banff Classification is also limited by the lack of formal integration of emerging molecular and genomics data and dependence on histopathological features that are not alloimmune-specific.[174] This leads to occasional challenges with respect to polypathology, such as polyomavirus BK nephropathy and CMR, which can both manifest interstitial inflammation (i) and tubulitis (t). The presence of SV40 immunohistochemical staining for BK nephropathy fails to exclude concurrent CMR as well as being possibly confounded by inter-laboratory differences in quality of SV40 staining (kappa statistic 0.24).[220, 221]

Deficiencies of the current system for rejection surveillance and diagnosis has meant that development of a cost-effective, non-invasive, sensitive and quantitative test for detecting graft damage has been identified as a high research priority in transplantation medicine.[222] Such a rejection monitoring test would have a major impact on clinical practice and patient quality of life by preventing unnecessary transplant biopsies and allowing improved management of the balance between suppressing rejection and limiting toxicities associated with therapy.[210, 211]

1.5.3 Non-Invasive Rejection Biomarkers

Concurrent with the emergence of multi-analyte, high sensitivity quantification platforms in genomics, advances in transcriptomics, proteomics and metabolomics have led to the identification of numerous other non-invasive biomarkers for the detection and diagnosis of rejection in plasma and urine (Figure 1-1). In contrast to cfDNA chimerism analysis, which appears to measure graft injury, the majority of other rejection biomarkers under study relate to host alloimmune activity (Figure 1-2).[223]

Figure 1-1: Conceptual model of “-omic” technologies with respect to the Central Dogma of Biology.

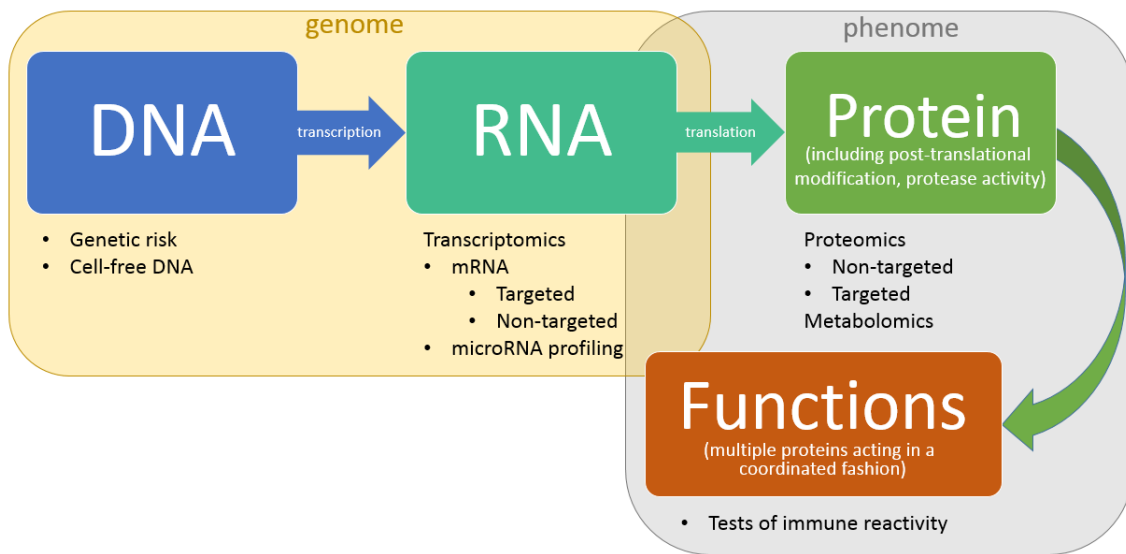
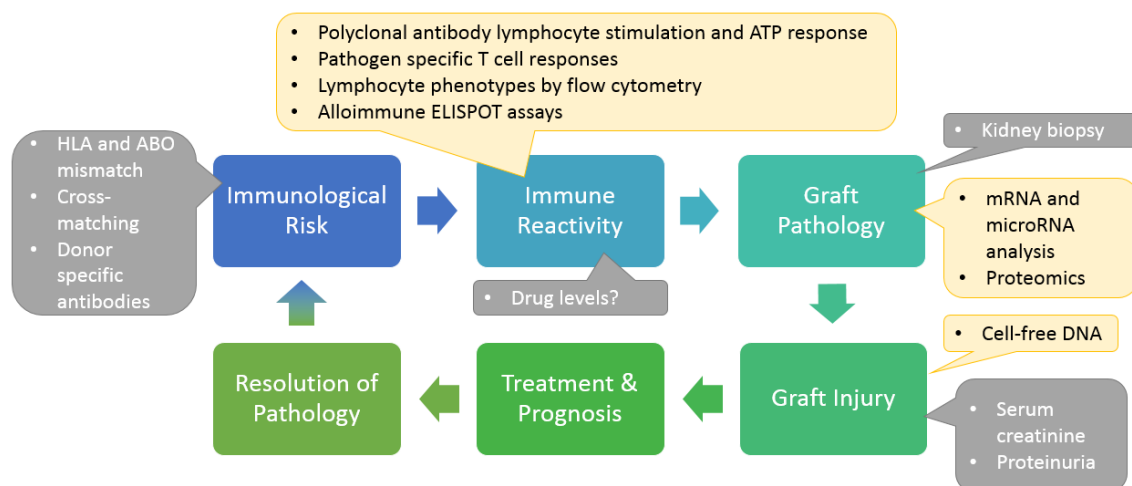


Figure 1-2: A natural history model of rejection indicating the conceptual stage targeted by various diagnostic biomarker approaches.



Targeted studies of peripheral blood lymphocyte messenger RNA (mRNA) expression identified genes *GZMB*, *PRF1* and *FASLG* encoding granzyme B, Fas ligand and perforin, respectively, as promising candidate rejection markers. However, variable results in validation studies and

difficulties distinguishing viral polypathology from alloimmunity has reduced interest in these markers.[224-227] Gene array based approaches appear more promising, with tissue studies demonstrating unique mRNA transcriptional signatures specific for rejection subtypes.[228, 229] Correlation-based step-wise algorithmic analysis of peripheral blood lymphocyte array data has led to development of gene set quantitative real-time PCR (qRT-PCR) based assays, which have performed well in initial validation studies.[230, 231]

Given the intimate contact between urine and the KT allograft, much work has been done on urinary-cell mRNA expression signals. Targeted studies have included grouped performance of *GZMB* and *PRF1* (rejection sensitivity 79-83%, specificity 77-83%) and grouped performance of *FOXP3*, *TNFRSF4*, *TNFSF4* and *PDCD1* (rejection sensitivity 95%, specificity 92%).[232, 233] Interestingly, a separate study of nine candidate genes including *GZMB* and *PRF1* did not include them in the best-fit model, which was able to identify ACR 20 days before confirmatory biopsy in the study population, a finding reproduced in an external validation data set.[234] No studies to date have examined the ability of urinary-cell mRNA expression to distinguish rejection subtypes and the molecular stability of mRNA may hamper diagnostic application.

An alternative approach has been the study of lymphocyte inflammatory chemokines in urine. Specific chemokines studied have included interferon-gamma and interferon-gamma mediated graft-produced chemokines such as CXCL9 and CXCL10.[204, 235] Urinary CXCL9 appears to have prognostic significance for rejection or graft dysfunction and another chemokine, CCL2, appears to predict future interstitial fibrosis and tubular atrophy as well as late allograft loss when measured at 6 months.[236-238] A key limitation of this approach is overlap in the immunologic responses to viral infection and the allograft, resulting in a failure of these assays to distinguish between BK nephropathy and rejection.

Non-targeted microRNA expression studies have also been performed in tissue, plasma and urine of KT recipients. In tissue and plasma, three markers thought to be derived from infiltrating lymphocytes (miRNA-142-5p, miRNA-155 and miRNA-223), were associated with rejection in a sample of patients without confounding polypathology.[235] In urine, reduced miRNA-210 levels predicted allograft dysfunction and acute CMR, normalising after successful treatment.[239] Further prospective studies are required.

Finally, enzyme-linked immunospot (ELISPOT) assays that detect alloimmune T lymphocyte interferon-gamma production have been identified as prognostic biomarkers for acute rejection and graft dysfunction.[240, 241] Notably, risk of rejection predicted by these assays did not correlate with traditional metrics of immunological mismatch between donor and recipient. Studies on the diagnostic utility of this approach for rejection have not been performed.

Despite proof-of-principle and diagnostic validity studies, with the exception of CareDx Allosure discussed earlier, none are presently reimbursed for use in routine clinical practice. International adoption of transplant rejection diagnostic biomarkers has also been poor.

1.5.4 Clinical Translation of Rejection Biomarkers

KT is a complex physiological state characterised by significant inter-individual variation. Kidney function is affected by organ quality, delayed graft function, achieved baseline allograft function, allograft pathology, residual native kidney urine output, and coexistent requirements for dialysis or plasma exchange. The alloimmune response is also highly variable, reflecting the degree of immunological matching, pre-existing and *de novo* allosensitization, immunosuppressive treatment regimen, treatment of rejection, age, and other sources of biological variation. Immunosuppression, particularly the treatment of rejection, may modify the immune response in ways that affect biomarker results without necessarily modifying outcomes.

In order that this variation is adequately accounted for, large international multi-centre diagnostic validation and clinical utility studies are required. Unfortunately, large prospective multicentre studies are hampered by cost and challenges standardising assay methods between centres.[223] It is, therefore, most likely that such evaluations will be performed after clinical translation using commercial testing arrangements subsidised by insurers and/or health services on the basis of diagnostic validity data.

Two development scenarios seem likely. First, as long as the diagnosis and classification of rejection remains dependent upon histopathological features (existing Banff classification), assays which allow earlier and specific identification of alloimmune graft injury should be prioritised, so as to facilitate prompt clinically useful KT biopsy. Second, emerging technology that allows sensitive and specific classification of alloimmune responses as well as detection of other common graft pathologies may supplant the histopathological Banff criteria for classification of allograft pathology and significantly reduce dependence on KT biopsy.

Given the variety of both immunological and non-immunological pathology identified on KT histopathology, it seems likely that any diagnostic approach that fulfils scenario two will involve either multiple methodologies or a panel/array approach. mRNA expression-based qPCR panels that presently discriminate between existing rejection classifications appear most promising in this regard, however, further prospective studies and assay development that allows for polypathology and non-immunologic pathology is required.

In the meantime, a biomarker that fulfils the first development scenario appears more achievable and translatable to widespread clinical practice. Notably, none of the biomarkers evaluated in KT rejection specifically measure graft cell injury or death.[223] In addition to facilitating early

diagnosis through appropriate KT biopsy, graft cell injury biomarkers would provide an alternative perspective to existing biomarkers which focus on the immunological response to the graft. Finally, graft cell injury biomarkers may offer insights into the natural history and pathophysiology of cellular damage in rejection, which is presently lacking and poorly represented on histopathological examination in the acute phase.

1.6 SUMMARY & AIMS

Allogeneic transplantation establishes chimerism within the recipient that is detectable in a variety of cellular and cell-free compartments. The quality and quantity of chimerism, whether naturally-induced during pregnancy, or artificially-induced as a result of transplantation, may be reflective of or contribute to a variety of transplantation-related sequelae.

In allogeneic HCT, quantification of chimerism in circulating leucocytes has been a cornerstone of care as a measure of engraftment, relapse, and graft rejection for some time.[57] Separately, the establishment of mixed bone marrow chimerism was associated with development of immunological tolerance to solid organ transplantation from the same donor, and maternal microchimerism in progeny is associated with tolerance against maternal antigens. [4, 9, 75] More recently, chimerism has been detected in cfDNA in plasma and urine of solid organ transplant recipients, and variations in cfDNA chimerism may reflect graft events.[242]

In kidney transplantation, the impact of rejection on long term graft outcomes remains a big issue.[186, 202, 203] Existing strategies for the surveillance of rejection are suboptimal.[204] A number of avenues for the development of graft rejection biomarkers are under investigation, and the majority of candidate biomarkers arising from these assess the recipient immune response against the allograft. In contrast to these, graft-derived cfDNA released from graft cells undergoing apoptosis and necrosis may be a biomarker specific to graft injury.[153]

Analysis of the many forms of chimerism relevant to transplantation relies upon sensitive, precise and highly informative quantification. These applications exceed the capabilities offered by traditional chimerism analysis methods based upon sex-mismatch, and qualitative PCR or quantitative real time PCR techniques. The genomics revolution has brought about a detailed understanding of human genomic variation, and new platforms for the interrogation and discrimination of genetic variation are now available.

This work seeks to:

- Develop a novel chimerism analysis test based upon a recently characterised form of genomic variation, CNV, and the improved sensitivity and precision of absolute quantification via ddPCR, for sex-independent post-transplant chimerism analysis of cell-free and cellular DNA.

- Demonstrate analytical and diagnostic validity of cfDNA chimerism analysis using this new methodology for the absolute quantification of graft-derived cfDNA following kidney transplantation, in so doing testing the hypothesis that immunological injury to the allograft is associated with increased quantities of graft-derived cfDNA in the plasma.
- Demonstrate analytical and diagnostic validity of cellular chimerism analysis using this new methodology for the relative quantification of chimeric fraction following allogeneic HCT, in so doing demonstrating proof-of-principle application for monitoring engraftment, mixed chimerism, and relapse.

CHAPTER 2 ASSAY DESIGN

2.1 INTRODUCTION

To maximise the study of chimerism between two (or more) genetically non-identical individuals in a transplant arrangement, an assay panel design process was undertaken using CNV loci previously identified to exhibit zero copy number allelic frequencies of between 0.3 and 0.7.[145] These CNV loci had been curated from Centre d'Etude du Polymorphisme Humain data within two high resolution HapMap data sets.[38, 243]

Genome-wide distribution of targeted CNV loci was sought to minimise unequal confounding from genomic variation. Sex chromosomes were avoided to permit sex-mismatch independent chimerism analysis. Genic, particularly exonic, regions were avoided to minimise the potential that assay results would confer diagnostic information other than that required for chimerism analysis. Repetitive regions with low complexity were avoided to ensure sufficiently unique sequence was available for assay development.

Multiple assays targeting different maximally informative CNV loci were designed to comprise an assay panel. Maximising the number of informative markers per individual in a transplant arrangement was sought to improve redundancy and provide multiple independent assessments of chimerism status, in light of other genetic variation and variation in analyte quality and laboratory measurement.

Our group has previously published modelling of the cumulative proportion of informative assays expected in un-related and related individuals based upon panel size and population allelic frequencies.[145] That work identified panel sizes of 10, 20 and 30 markers are predicted to be associated with a cumulative proportion of at least 3 informative markers in 46%, 90% and 99% of un-related pairs, respectively. For first-degree related pairs, at least one informative marker would be expected 94% of the time.

Within the selected highly polymorphic CNV loci, a process of assay design was undertaken to maximise oligonucleotide specificity for the target CNV locus, and ensure reliable performance on the ddPCR platform. Due to the fragmented, degraded nature of cfDNA, each amplicon was designed to be short, albeit without compromising assay function.

This chapter details the process of CNV locus assessment and ddPCR assay development. The previously published ten-assay quantitative real-time PCR panel was completely redesigned. Additionally, new non-polymorphic assays were developed to serve as controls.

2.2 METHODS

2.2.1 CNV Assay Development

2.2.1.1 *Locus Selection*

A data set of 41 CNV loci previously identified to be suitable for chimerism analysis assay development was obtained.[145] Genome coordinates were converted from the March 2006 (NCBI36/hg18) human assembly to the Feb 2009 (GRCh37/hg19) human assembly using the University of California Santa Cruz (UCSC) Genome Browser LiftOver tool (Table 2-1).[27, 244]

Table 2-1: Parameters used for the UCSC Genome Browser LiftOver tool.

Parameter	Value
Minimum ratio of bases that must remap	0.95
BED 4 to BED 6 Options	
Allow multiple output regions	Off
Minimum hit size in query	0
Minimum chain size in target	0
BED 12 Options	
Minimum ratio of alignment blocks or exons that must map	1
If thickStart/thickEnd is not mapped, use the closest mapped base	Off

Allelic frequencies for each CNV locus were obtained from the high resolution data sets used as part of prior work.[243, 245] Zero, one and two copy number genotypes were reviewed, and the summed frequencies of these genotypes were confirmed to approach 1. This effectively excluded CNV loci with significant allelic frequencies of three or more copy number. Additionally, CNV loci displaying two copy number frequencies of less than 0.2 were prioritised to maximise the informative potential of the panel.

2.2.1.2 *Identification of Sequences for Assay Development*

The UCSC Genome Browser was used to visually inspect each CNV locus using the respective Feb 2009 (GRCh37/hg19) human assembly genome coordinates.[244] Table 2-2 details the annotation tracks used to perform the inspection. Regions within each CNV locus that were free from common SNPs, repeating elements and genes were preferentially selected for assay development. DNA sequence from these selected regions was then exported using the UCSC Genome Browser “get DNA” tool.

In circumstances where not all suitability criteria could be met, compromises were made prioritising from greatest to least: (i) absence of common SNPs, (ii) absence of genes, and (iii) freedom from repeating elements.

Table 2-2: UCSC Genome Browser annotation tracks used for visual inspection of CNV loci.

Track Name	Full Name	Contents
Base Position	Genome Base Position	Genomic coordinates for the displayed region, and the base composition if resolution permits.
UCSC Genes[246]	UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)	Gene predictions based on data from RefSeq, GenBank, CCDS, Rfam, and the tRNA Genes track. Includes both protein-coding genes and non-coding RNA genes.
RefSeq Genes[247]	RefSeq Genes	Known human protein-coding and non-protein-coding genes taken from the NCBI RNA reference sequences collection (RefSeq).
Publications[248]		Text mined references to biomedical articles in which visualised sequence or chromosome locations are mentioned.
Common SNPs(147)[249]	Simple Nucleotide Polymorphisms (dbSNP build 147) Found in $\geq 1\%$ of Samples	SNPs with $\geq 1\%$ minor allele frequency (MAF), mapping only once to reference assembly.
DGV Struct Var[250]	Database of Genomic Variants: Structural Variation (CNV, Inversion, In/del)	Structural Variant Regions: annotations that have been generated from one or more reported structural variants at the same location. Supporting Structural Variants: the sample-level reported structural variants.
RepeatMasker[251]	Repeating Elements by RepeatMasker	Interspersed repeats and low complexity DNA sequences, categorised into one of ten classes.

2.2.1.3 Primer and Probe Design

DNA sequence for assay design was imported into the IDT PrimerQuest Tool (Integrated DNA Technologies, Baulkham Hill, NSW, Australia).[252] Default design parameters for quantitative PCR were used (2 Primers + Probe) with a few modifications (Table 2-3) designed to reduce the amplicon length and increase the melting temperature delta between primer and probes to between 4-8 °C. Melting temperatures were confirmed using IDT OligoAnalyzer 3.1 (Integrated DNA Technologies, Baulkham Hill, NSW, Australia) with modifications based upon the Bio-Rad Digital Droplet PCR Applications Guide (Table 2-4).[253] Multiple candidate primer and probe sequences identified by the IDT PrimerQuest Tool were then evaluated for suitability.

Table 2-3: Modifications to default parameters for assay design using the IDT PrimerQuest Tool.

Criteria	Modified	Minimum	Optimal	Maximum
Primer Criteria				
Primer T _m (°C)	Modified	59	61	63
Primer GC (%)	Default	35	50	65
Primer Size (nucleotides)	Modified	17	22	25
Probe Criteria				
Probe T _m (°C)	Modified	66	67	68
Probe GC (%)	Default	40	50	60
Probe Size (nucleotides)	Unchanged	20	24	30

Amplicon Criteria				
Amplicon Size (bp)	Modified	60	80	100

Table 2-4: Parameters for melting temperature calculation using IDT OligoAnalyzer 3.1.

Parameter	Value
Target type	DNA
Oligo Conc	0.3 uM
Na ⁺ Conc	50 mM
Mg ⁺⁺ Conc	3.8 mM
dNTPs Conc	0.8 mM

2.2.1.4 Evaluation of Primer and Probe Suitability

Each assay (comprising forward primer, reverse primer, and probe) identified by the IDT PrimerQuest Tool underwent additional assessments of suitability. Each primer and probe sequence was assessed for unique specificity within the human genome using the UCSC Genome Browser implementation of the BLAST-like Alignment Tool (BLAT) with default settings.[254] BLAT score, match span, match identity, match contiguity and match location were all considered.

In-silico PCR was then performed using assay primer sequences and the UCSC Genome Browser In-Silico PCR Tool to confirm amplicon sequence specificity (Table 2-5).[244] Amplicon product sequence from the in-silico PCR was confirmed to be specific to the target CNV locus and wherever possible within a region free from common SNPs and repeating elements.

Assays that were non-specific, in particular those with 100% identical full-span contiguous non-specificity, were unsuitable. Likewise, assays affected by common SNPs or sited within regions of highly repeating elements were avoided. Prospective assays identified by the IDT PrimerQuest Tool for each CNV locus identified for assay development were sequentially examined until a suitable assay was identified. Within larger CNV loci, multiple areas identified to be suitable for assay development were evaluated as required to increase the number of assay options available for assessment.

In scenarios where assay development options were limited and not all suitability criteria could be met, criteria were relaxed prioritising the following characteristics in order from greatest priority to least: (i) amplicon specificity, (ii) oligonucleotide binding sites free from common SNPs, (iii) assay oligonucleotide non-specificity limited to one oligonucleotide from the assay, (iv) oligonucleotide characteristics (size, melting temperature), (v) oligonucleotide specificity characteristics (score, location), and (vi) presence of interspersed repeats and low complexity DNA sequences within the assay.

Table 2-5: Parameters used for UCSC Genome Browser In-Silico PCR.

Parameter	Value
Assembly	Hg19 assembly
Target	Genome assembly
Max product size	4000 bp
Minimum perfect match	15 bp
Minimum good match	15 bp
Flip reverse primer	False

2.2.1.5 Probe Selection

To facilitate multiplex PCR assays, approximately half of the developed assays were designed to include a 5' 6-FAM™ ZEN™ Double-Quenched Probe (Integrated DNA Technologies, Baulkham Hill, NSW, Australia). The remaining half of developed assays were designed to include a 5' HEX™ ZEN™ Double-Quenched Probe (Integrated DNA Technologies, Baulkham Hill, NSW, Australia).

2.2.1.6 Naming Convention

CNV loci were numbered sequentially on the basis of their order in the original data set (CNV loci 1-41). Developed assays were named to match the respective CNV locus and expressed as a minimum of two digits followed by a character suffix (CNV assays 01B-41A). The character suffix represents iterative versions of assay development. This largely reflects redesign of the real-time PCR primers and probes originally developed for the first 10 CNV loci (CNV assays 01B-10B), but in some instances redesign of assays exhibiting poor performance on arrival to the lab after manufacture (CNV assays 10C and 19B).

2.2.1.7 Summary Statistics

Summary statistics are expressed as median (inter-quartile range) unless otherwise specified.

2.2.2 Non-Polymorphic Assay Development

2.2.2.1 Locus Selection

Agenic regions not affected by common genetic polymorphism were sought as sites for non-polymorphic “positive control” assay development. Specifically, the selected region had to be present at both alleles (i.e. two copy, one on each chromosome) in the vast majority of the population. Random researcher-directed browsing using the UCSC Genome Browser was unsuccessful at identifying such a site as agenic regions are subject to frequent extensive genetic polymorphism. Consequently, assay development occurred from researcher-selected gene regions not affected by common genetic polymorphism using methods described in section 2.2.1.2.

2.2.2.2 Assay Design and In-Silico Evaluation

Primers and probe for regions identified for assay development were developed and evaluated using the methods described in sections 2.2.1.3 and 2.2.1.4 respectively. Two assays for two different non-polymorphic regions were developed. Both assays were designed to incorporate a 5' 6-FAM™ ZEN™ Double-Quenched Probe (Integrated DNA Technologies, Baulkham Hill, NSW, Australia).

2.3 RESULTS

2.3.1 CNV Assay Development

2.3.1.1 Locus Size and Location

Human assembly Feb 2009 (GRCh37/hg19) genome coordinates for 41 CNV loci previously identified to be suitable for chimerism analysis assay development are presented in Table 2-6.

Chromosome frequency is plotted in Figure 2-1. All chromosomes are represented with exception of chromosomes 3, 11, 12, 17, 19, 21 and the sex chromosomes. Chromosomes 1 and 2 were relatively over-represented, each comprising 6/41 (15%) of the CNV loci identified for assay development.

The majority of CNV loci (35/41, 85%) were less than 5,000 bp long, with a median size of 1,829 bp (Figure 2-2). CNV locus 35 (CNV ID 933) on chromosome 6 had a size of 142,219 bp and was the only CNV locus larger than 45,000 bp long.

Table 2-6: GRCh37/hg19 genome coordinates of 41 CNV loci identified for assay development.

CNV #	CNV ID	Chromosome	Start	End	Size (bp)
1	CNVR376.1	chr1	158867528	158870050	2522
2	CNVR7344.1	chr18	54946737	54948704	1967
3	CNVR4014.1	chr8	112294084	112296293	2209
4	CNVR6084.1	chr14	22881700	22882230	530
5	CNVR6676.1	chr16	25340110	25343094	2984
6	CNVR3319.1	chr7	24038230	24040059	1829
7	CNVR5850.1	chr13	39057318	39060191	2873
8	CNVR3753_full	chr8	4122900	4124927	2027
9	CNVR7301.1	chr18	35306075	35306647	572
10	CNVR4886.1	chr10	108030313	108032528	2215
11	CNVR1138.3	chr2	220052107	220054328	2221
12	CNVR3451.1	chr7	73829106	73831236	2130
13	CNVR4596.1	chr10	4708559	4710493	1934
14	CNVR801_full	chr2	54565628	54567472	1844
15	CNVR5853_full	chr13	39933627	39935387	1760

16	CNVR842.1	chr2	76773726	76775435	1709
17	CNVR431.1	chr1	187464835	187466483	1648
18	CNVR4203.1	chr9	17910038	17911633	1595
19	CNVR952.1	chr2	121797526	121798787	1261
20	CNVR1037.1	chr2	159959794	159960963	1169
21	CNVR1138.2	chr2	220053192	220054328	1136
22	CNVR5923.1	chr13	72845782	72846891	1109
23	CNVR8147.1	chr22	35645445	35646472	1027
24	CNVR3004.1	chr6	95193393	95194392	999
25	1343*	chr8	112294279	112295260	981
26	CNVR381.1	chr1	159648762	159649629	867
27	CNVR4374.1	chr9	89154952	89155776	824
28	CNVR2799.1	chr6	18402120	18402881	761
29	CNVR6540.1	chr15	99574727	99575462	735
30	CNVR6357.1	chr15	39744398	39744844	446
31	158	chr1	210081613	210083984	2371
32	CNVR358.1	chr1	152555610	152590091	34481
33	CNVR217.1	chr1	72766282	72811969	45687
34	CNVR7808.1	chr20	21286267	21288529	2262
35	933	chr6	32431552	32573771	142219
36	554	chr4	10214156	10235268	21112
37	CNVR2469.1	chr5	57323526	57333742	10216
38	CNVR3009.1	chr6	100034554	100035273	719
39	CNVR3609.1	chr7	148073126	148076330	3204
40	801	chr5	57326027	57333533	7506
41	CNVR2304.1	chr5	1924604	1925111	507

Figure 2-1: Chromosomal frequency of CNV loci identified for assay development.

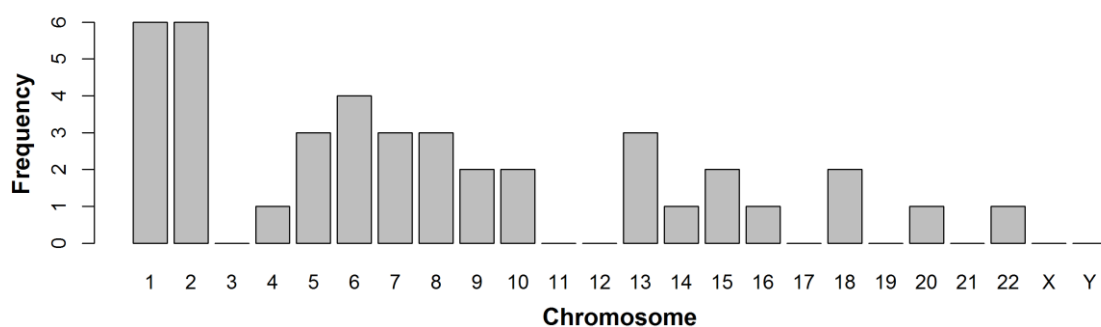
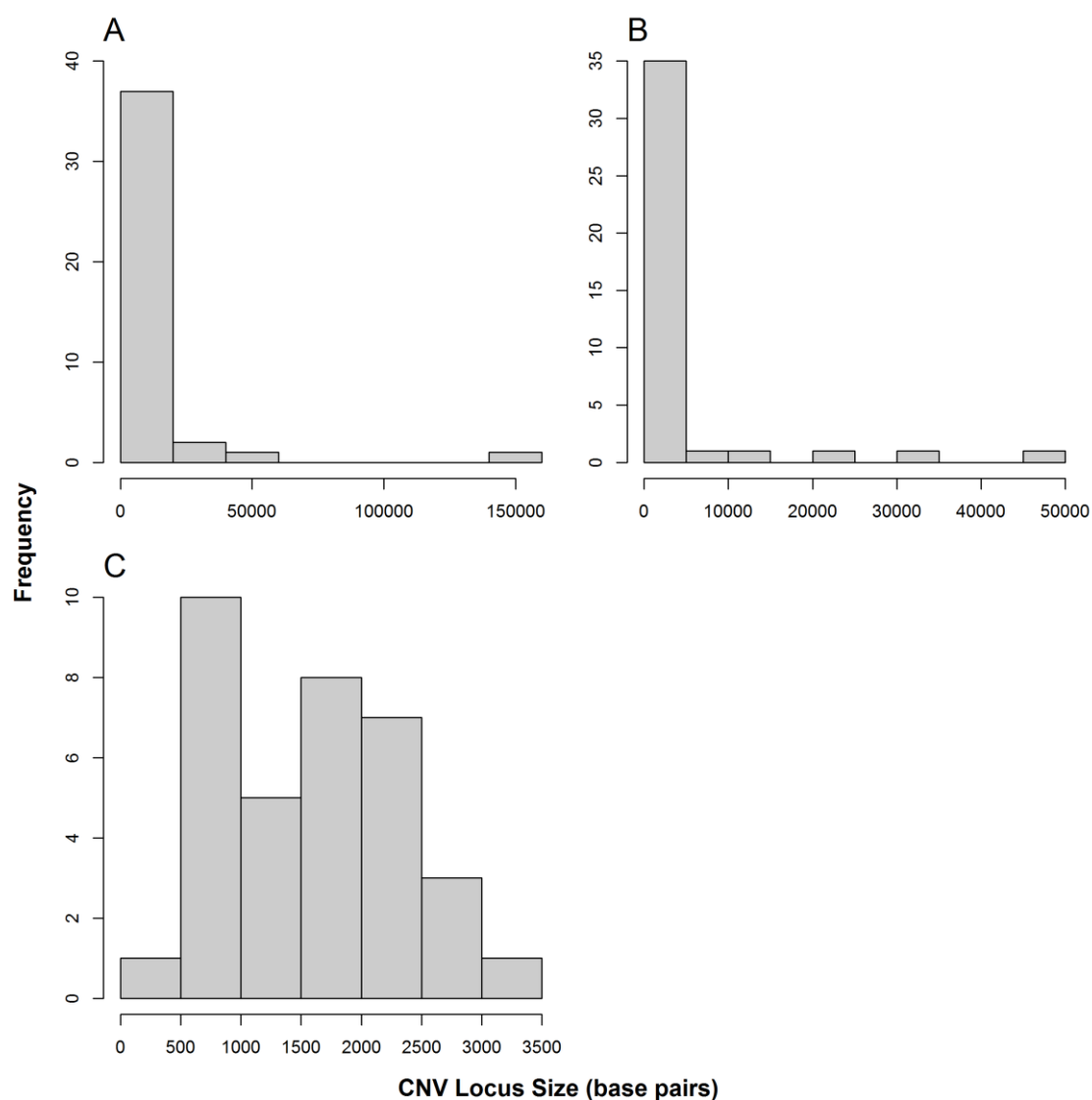


Figure 2-2: Size distribution of CNV loci identified for assay development. Panels A, B and C represent CNV loci size frequencies at resolutions of 0-150000, 0-50000 and 0-3500 bp length

respectively.



2.3.1.2 Locus Allelic Frequencies

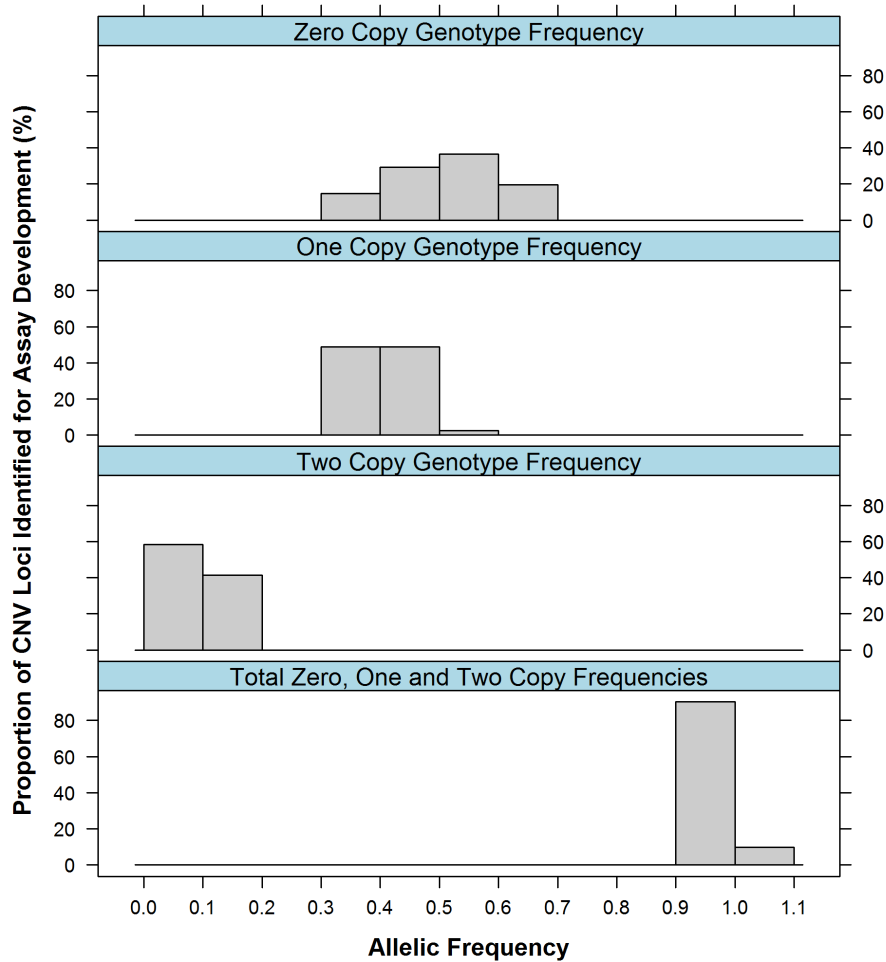
Allelic frequencies for each CNV locus are summarised in Table 2-7. The median zero copy allelic frequency was 0.52 (0.43-0.59), one-copy allelic frequency was 0.41 (0.36-0.43), and two copy allelic frequency was 0.08 (0.05-0.13). The total of zero, one and two copy allelic frequencies had a median of 1 (1-1). The distribution of genotype frequencies is plotted as a proportion of all CNV loci selected for assay development in Figure 2-3.

Table 2-7: CNV locus allele frequency within source high resolution data sets.

CNV #	Zero-Copy Frequency	One-Copy Frequency	Two-Copy Frequency	Total
1	0.51	0.42	0.07	1.00
2	0.52	0.40	0.08	1.00
3	0.50	0.41	0.08	0.99
4	0.46	0.43	0.11	1.00

5	0.45	0.47	0.08	1.00
6	0.47	0.38	0.15	1.00
7	0.43	0.43	0.14	1.00
8	0.43	0.43	0.13	0.99
9	0.42	0.45	0.12	0.99
10	0.41	0.42	0.17	1.00
11	0.40	0.43	0.17	1.00
12	0.55	0.33	0.12	1.00
13	0.61	0.31	0.08	1.00
14	0.59	0.37	0.04	1.00
15	0.55	0.33	0.12	1.00
16	0.60	0.33	0.08	1.01
17	0.62	0.33	0.04	0.99
18	0.61	0.36	0.03	1.00
19	0.53	0.36	0.10	0.99
20	0.50	0.43	0.07	1.00
21	0.39	0.42	0.18	0.99
22	0.41	0.55	0.04	1.00
23	0.45	0.43	0.12	1.00
24	0.40	0.47	0.13	1.00
25	0.57	0.38	0.05	1.00
26	0.60	0.35	0.05	1.00
27	0.56	0.36	0.07	0.99
28	0.41	0.44	0.15	1.00
29	0.51	0.43	0.06	1.00
30	0.52	0.42	0.07	1.01
31	0.38	0.45	0.17	1.00
32	0.39	0.46	0.15	1.00
33	0.39	0.46	0.15	1.00
34	0.52	0.42	0.07	1.00
35	0.55	0.38	0.07	1.00
36	0.55	0.33	0.12	1.00
37	0.61	0.36	0.02	1.00
38	0.62	0.38	0.00	1.00
39	0.66	0.31	0.03	1.00
40	0.67	0.32	0.02	1.00
41	0.67	0.33	0.00	1.00

Figure 2-3: Distribution of copy number allelic frequency amongst the CNV loci selected for assay development.



2.3.1.3 Primer and Probe Characteristics

39 assays were designed for the 41 CNV loci identified as suitable for assay development. PCR assays could not be developed for CNV locus 35 due to extensive SNP and repeating element frequency across the entire 142,219 base pair (bp) locus (Appendix A, Figure 1). Assays were not developed for CNV locus 40, as the entire locus was sited within the CNV locus 37 (Appendix A, Figure 2).

Forward and reverse primer sequences and basic characteristics are summarised in Table 2-8 and Table 2-9, respectively. Median forward and reverse primer size was 21 (20-23) bp and 21 (20-24) bp, respectively (Figure 2-4). Median melting temperatures were 62.9 (61.5-63.2) °C and 62.8 (61.5-63.3) °C, respectively (Figure 2-5). No forward and reverse primer pairs had an absolute melting temperature discrepancy of greater than or equal to 2 °C. Median forward and reverse primer GC content was 45 (41-47) % and 45 (38-50) %, respectively (Figure 2-6).

Probe sequences and basic characteristics are summarised in Table 2-10. Median probe size was 25 (24-26) bp and median melting temperature was 68.3 (67.6-69) °C. The median difference between the probe melting temperature and the mean of the forward and reverse primer melting temperatures was 5.6 (4.6-6.7) °C (Figure 2-8). CNV 25A had a probe melting temperature only 2.45 °C higher than the mean primer melting temperatures. This was the only assay in which probe melting temperature was less than 3 °C higher than the mean of the primer melting temperatures. Median probe GC content was 50 (45-54) %.

Table 2-8: Forward primer characteristics for each CNV locus. All binding sites are on the sense (+) strand.

CNV #	Forward Primer Sequence	Size (bp)	T _m (°C)	GC (%)
01B	AAAGTCAACTTCTCTGTGCGA	21	63.2	43
02B	TGTCGTTTCTAAGGATGGGAAAG	23	63.4	43
03B	ATGCCTGGCTTCATGTTACT	20	63.1	45
04B	GTGTGGACATTGGTTCCTTT	20	61.9	45
05B	TTTCTTCCAGCAGAGCTCAC	20	63.2	50
06B	TCACTCTTCCCACCCTCTTAT	21	63.3	48
07B	CCCAACTCAATTTCAAGCTG	20	60.6	45
08B	ACCTATCAACTTGGTTACCCTAAA	24	62.8	38
09B	TGTCCCAATGCACAGACTT	19	62.9	47
10C	CTTTCGGTTCTCTCCAACCA	20	62.9	50
11A	CACAAATCTGATGACCTCTCTACT	24	62.8	42
12A	AGCCTCAATTCAGGACCTA	19	61.3	47
13A	TGTCATGTGCAATAGAGACTGTT	23	63.4	39
14A	ACTGGCTTCCAAATGTGACTAA	22	63.5	41
15A	ACTGACACACTCAAGCTATGG	21	62.9	48
16A	AGCAAATGATGATGTTGTATGAGC	24	63.2	38
17A	GCCATGCCTCCCACATT	17	63.5	59
18A	TCACAGGTCCTACCAGCTTTA	21	63.7	48
19B	CACACCAAATTGTAAAGAGTTACC	24	61.5	38
20A	CCAGCGTCTTCTCTTATGTCAG	22	63.2	50
21A	TTTGGTGACTGATACATGGCT	21	63	43
22A	ATATCCTCTGGATTTGACAGTG	22	60.9	41
23A	TTCCAAAGACACAAGGCAGTA	21	63.1	43
24A	CCAAGTATGATGTTATTTGTGGGA	24	62.3	38
25A	CATATTGAGATCCAGAAACAACCTT	25	62.5	36
26A	CCTTACTGACACTCTCTGTTTCTT	24	63.1	42
27A	CAGGGAGAGAGTGTAGAAAGAATAG	25	62.6	44
28A	ACACTCCAATTCCAACCTTACC	22	63	45
29A	GCCTGTGCCTACCTGTTTAATA	22	63.3	45
30A	GAGAAATCCATCTTCAGGTCAG	22	61.3	45
31A	TACACCAGTGGAAGGACTAT	20	61	45
32A	GAGGAATGAGAGACAACACTTC	22	61.5	45

33A	GGCAGCAATAATCACCTCTGA	21	63.4	48
34A	TTCTTCCCTTTCCAGACTCA	20	61.7	45
36A	CTAGAACCTTCCATCCTTAATGT	23	61	39
37A	GCTTTGTTGGGAGAGTTAGA	20	60.7	45
38A	AAAGCAGCTCAACCTCTTAAT	21	61.4	38
39A	CCGACTTTCTACTTATTCCACA	22	60.9	41
41A	CCCAAACCTCCTCTCATCTTTC	21	61.2	48

Table 2-9: Reverse primer characteristics for each CNV locus. All binding sites are on the anti-sense (-) strand.

CNV #	Reverse Primer Sequence	Size (bp)	Tm (°C)	GC (%)
01B	TTATTCTTCCACCATTATCCCTCAT	25	63.1	36
02B	GTTTCATGAAGCTGCCTCTCT	21	63.5	48
03B	AATTACTGACCTAGACCTTTCAAGA	25	62.8	36
04B	TTAGGGCAAACAAGATGAGAAA	22	61.7	36
05B	GCCAGCCCTGGATGAAT	17	62.9	59
06B	CATCTCAAGAAAGAGGGCTCAA	22	63.3	45
07B	CCAAAGATCCAATTAGCACATC	22	60.8	41
08B	TTTGTGTGGGAATCTTTATCTGTG	24	62.6	38
09B	AAACAGGGTGGGCCTAAA	18	62.5	50
10C	AACAACCTTCCCTGAGACGTG	20	63.1	50
11A	AGGGAGGTTGAATTTATGTCTCA	23	62.8	39
12A	CTCCTTGGAGAGCTTGTTATC	21	61	48
13A	TTCTGGCCAGGGCAATTC	18	63.6	56
14A	TAACCTGCAGCTCTGGGA	18	63.7	56
15A	AGCATAAATCTCTGCTCCCTTT	22	63.1	41
16A	TAAATGCCAGCACCTGCTA	20	65.1	50
17A	TGGCCCTAACAGTTTGGATATG	22	63.4	45
18A	GACGAAAGATGTCTTAATTTGGTTCAG	27	63.7	37
19B	GTCCTGGTAACAAATCCAGTTAATA	25	62	36
20A	ACAAGCCAACAACGTACAGG	20	63.9	50
21A	CCTAGCTACTATTCCCTCTAGAACA	25	63.3	44
22A	ATGTTGAGAGCATTGGACTT	20	61	40
23A	GGACATTCTGATGGGAGAAA	21	62.9	48
24A	CACAAGTAGATTGAGATTAGAAGGG	25	61.8	40
25A	AGAAATAAACTGAGGCCTGGATA	23	62.6	39
26A	GAGCACCCAGTAGGAATAGAATTAG	25	63.1	44
27A	GCCTGATCTCTAAACACTGGAG	22	63	50
28A	GAGAAATCAGAGGCCCAAAGA	21	63.1	48
29A	AAATCCACAGCCAGTACCTTC	21	63.3	48
30A	AAAGACCCTTCACACTTTGTT	21	61.8	38
31A	CCCTACTAATGACATCCAGAATAA	24	60.9	38
32A	CCTTCTGTAGCCTCTTCCTA	20	61.1	50
33A	CAATCTGATAAGCTTTCTAGGGTAGT	26	63.2	38

34A	GAAAGCCAGAGAATGAGAGAG	21	61	48
36A	GCTCCTACATGGCTACAAAG	20	61.2	50
37A	CAATTCATTGAGAAGTGCCAA	21	60.7	38
38A	TGAAAGATTCTCTGAGTTATTAGCC	24	61.8	38
39A	GCTGTATAAACCTCCACAG	20	60.9	50
41A	AAGGCTGATTAGCCGTAATG	20	61.2	45

Table 2-10: Probe characteristics for each CNV locus. Probe sense (+) and anti-sense (-) binding sites are specified.

CNV #	Sense	Probe Sequence	Size (bp)	Tm (°C)	GC (%)
01B	+	AATACACCAGCACCAGGGCT	20	67.1	55
02B	+	TGAAGTGCTTCTCCTCCCTGATCCA	25	69.4	52
03B	+	TCTCCCTCTGCCATAATATAGAGCA	26	67.1	46
04B	-	TCTGTGTATACACAAGAGGGAGCTATGC	28	67.9	46
05B	-	TTGCCTCCCACACATCTTCACTAGC	25	68.8	52
06B	+	TGTACAATGACTCAGTGGCTCATCAAC	27	67.3	44
07B	+	CCAATGTGACATCACTGAATTCAGAGGT	28	67.5	43
08B	-	ACACCATGGACCTATAGAAATAGCTGGC	28	68.3	46
09B	-	TGCATTCACCGTAAGGGAACAGAGC	25	69.2	52
10C	+	CCCATCCTCTCTTAGCACAACCATGA	26	68.3	50
11A	+	ACTGTTAAGTATGATGGCTCTCCTCCTC	26	67.8	50
12A	+	AATCTTAGGCACCCATCCACTGCC	24	69.1	54
13A	-	TCAAGGAAAGTATATGGACCACATCCAGAC	30	68.1	43
14A	+	ACAGTGCAACAGCCCCTTGGC	20	68.8	60
15A	+	TTAAGCGCTCTCCATAACCACGCA	24	69	50
16A	+	CACAGGCAGCTATTTCCATGCGTG	24	68.5	54
17A	+	TCAACATGAGATTTGAGCAGGGACA	25	67.2	44
18A	+	ACAACACCAGGCAGGTTGACAT	22	67.7	50
19B	+	AGACACAGGATTAGAGAAGATGCCT	25	65.9	44
20A	+	CCAGCCTGTGCCGATGTTCT	20	67.7	60
21A	+	TTGATGCATGCCTTTGCTTGTCCA	24	68.8	46
22A	+	TCCTTGTCTTCTGGCATGACAGG	24	68.7	54
23A	+	ACCCTCCTTACCTCCTTTCACCCA	24	69.4	54
24A	+	TCATAGATGCTCTTTATCCAGTGCAGT	27	66.6	41
25A	+	TCTAGGTAGGATCGTTGTACTCTAACA	27	65	41
26A	+	TCTGTTGCTTCTGCAACTTGAGC	24	68.3	50
27A	+	AAGAACCAGGGTCTGGGAGTGAG	23	68	57
28A	+	ACTCAGCTGTGTAATTGGGCACATGA	26	69.1	46
29A	+	TCTGAGTCTGGGTTCACTAGCCCA	24	69.1	54
30A	+	TGTCCATCATCCAATGTTGATTGAGC	26	66.5	42
31A	-	TGTAGGCCATTATCCCTGTTGGCC	24	69	54
32A	+	TCATAGGGAGACCATGAGTTGCTGC	25	68.5	52
33A	+	CCCAAAGGACCCATTTGGCTGAAT	24	68.3	50
34A	-	AAATACATGAGTCTAAGAGGACCAGTATCC	30	66.1	40

36A	+	TGATGACGCTGTTGGTGAAGGAGA	24	68.5	50
37A	+	TAGCAGCCCATGTCTCCAAACCAA	24	69	50
38A	+	AGACAGTAAACCAAATTCCACATCCCAGT	29	68.8	41
39A	+	CGAACCCTGTTCTACGTGTGAGGC	25	69	56
41A	+	AAACCATGCCACATCCTGACGTTT	24	68.2	46

Figure 2-4: Distribution of oligonucleotide size for CNV assay forward primers, reverse primers, and probes.

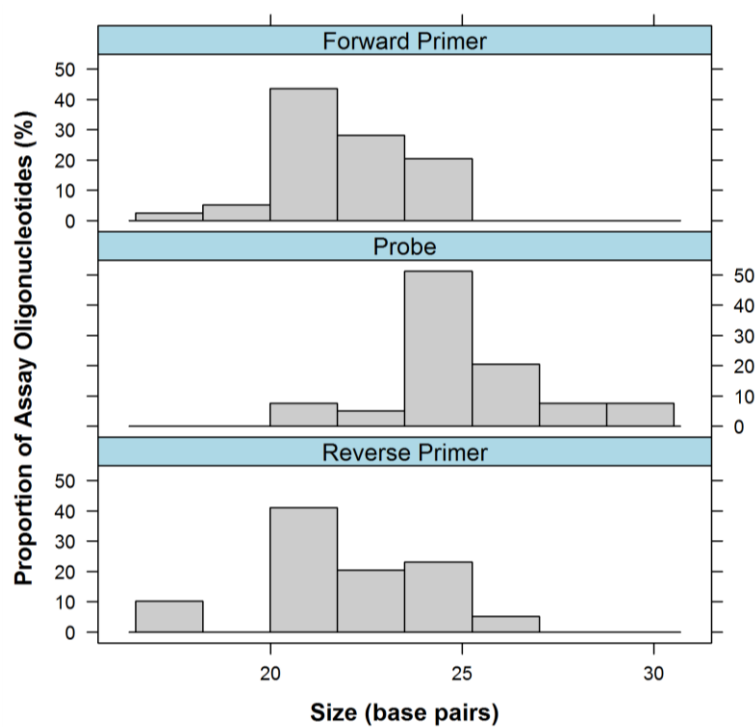


Figure 2-5: Distribution of oligonucleotide melting temperatures for CNV assay forward primers, reverse primers, and probes.

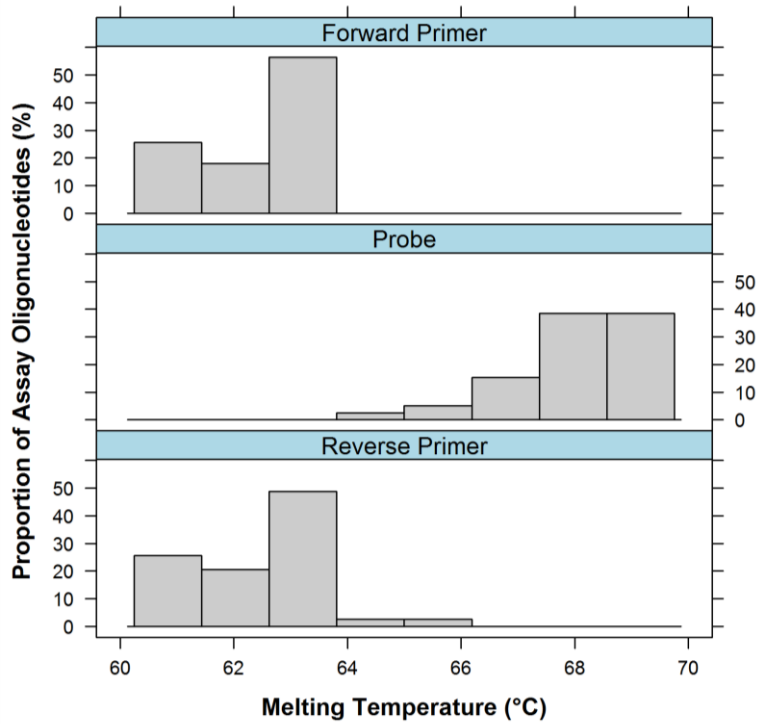


Figure 2-6: Distribution of oligonucleotide GC content for CNV assay forward primers, reverse primers, and probes.

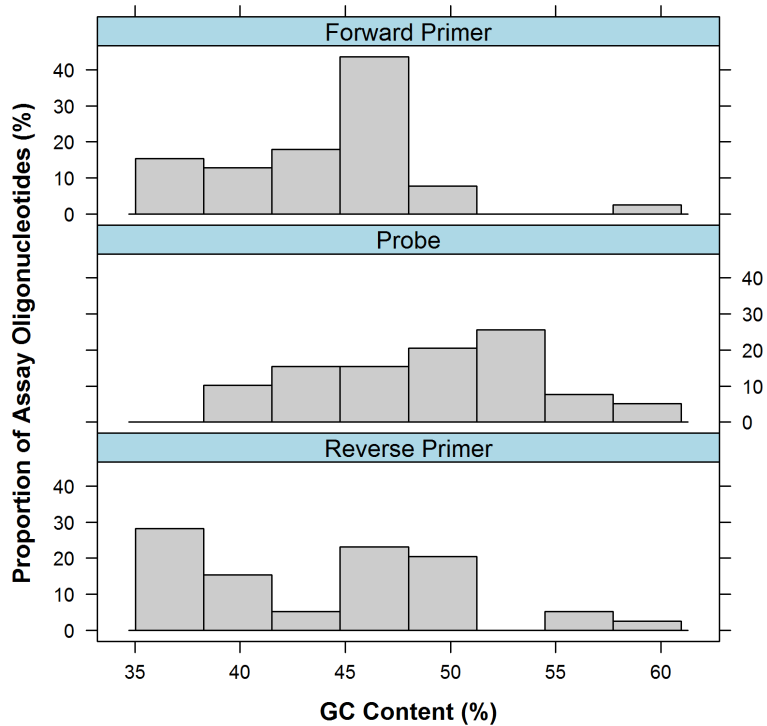


Figure 2-7: CNV assay distribution of the absolute difference in melting temperature between forward and reverse primers.

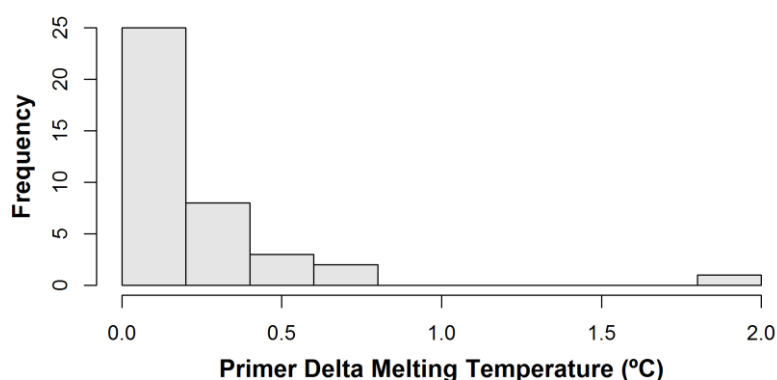
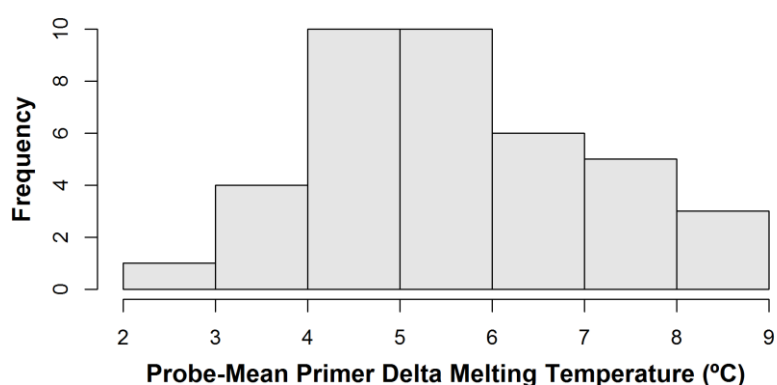


Figure 2-8: CNV assay distribution of the difference between probe melting temperature and the mean of the primer melting temperatures.



2.3.1.4 Assay and Amplicon Specificity

Oligonucleotide non-specificity was detected for 11/117 (9.4%) of CNV assay oligonucleotides, including 3/39 (7.7%) forward primers, 2/39 (5.1%) reverse primers, and 5/39 (12.8%) probes (Table 2-11). Only one assay, CNV 24A, had more than one non-specific oligonucleotide. The CNV 24A forward primer and probe had non-specific binding sites on chromosomes different from each other as well as the target CNV locus. With exception of CNV 17A, no oligonucleotide non-specific match had a BLAT score matching the intended target region. In all cases, this was due to shorter match spans or non-contiguous region matches.

In-silico PCR of assay primers identified only one PCR product for each assay (Table 2-12). Median amplicon size was 74 (70-80) bp (Figure 2-9). For all assays, amplicon sequence was confirmed to be specific to the corresponding target CNV locus. Appendix B displays UCSC Browser Tracks of all CNV loci demonstrating the position of the CNV assay within the locus as well as the presence of genes, common SNPs, structural variation and repeating elements.

Table 2-11: Assay oligonucleotide non-specificity identified by BLAT search.

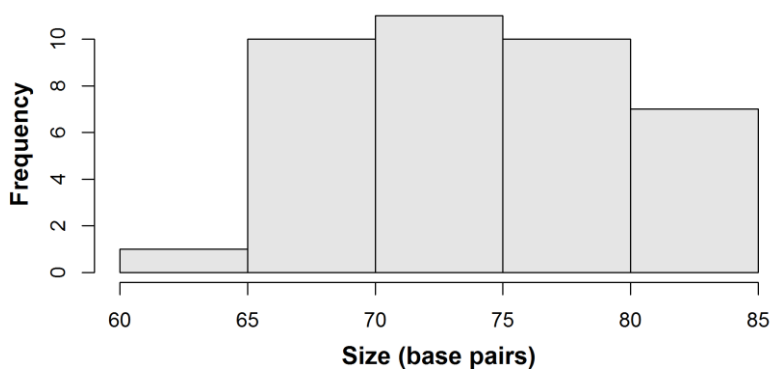
CNV #	Oligonucleotide		BLAT Non-Specific Match				
	Type	Size (bp)	Score	Identity	Span (bp)	Sense	Hg19 Genomic Coordinates
19B	Forward Primer	24	21	100%	21	-	chr1:185873560-185873580
24A	Forward Primer	24	20	100%	20	-	chr10:113481478-113481497
25A	Forward Primer	25	21	100%	25	+	chr4:8540109-8540129
01B	Reverse Primer	25	22	96%	24	-	chr10:26171244-26171267
01B	Reverse Primer	25	21	100%	21	+	chrX:62535357-62535377
27A	Reverse Primer	22	20	100%	20	-	chr4:78533949-78533968
02B	Probe	25	24	100%	263	+	chr2:156956839-156957101
07B	Probe	28	20	100%	20	-	chr12:25122766-25122785
07B	Probe	28	20	100%	20	-	chr1:217411698-217411717
07B	Probe	28	20	100%	20	-	chr1:17871654-17871673
07B	Probe	28	20	100%	20	+	chr2:61053363-61053382
07B	Probe	28	20	100%	20	+	chr18:29747253-29747272
10C	Probe	26	21	100%	21	+	chr13:20859382-20859402
17A	Probe	25	111 matches. Median score 23 (21-23), range 20-25.				
24A	Probe	27	24	96.3	308	-	chrX:54303222-54303529
24A	Probe	27	22	100	22	-	chr7:103857398-103857419

Table 2-12: CNV assay in silico PCR amplicon characteristics.

CNV #	Size (bp)	Hg19 Genomic Coordinates
01B	67	chr1:158868859-158868925
02B	74	chr18:54947989-54948062
03B	75	chr8:112295274-112295348
04B	78	chr14:22881932-22882009
05B	70	chr16:25341111-25341180
06B	72	chr7:24038988-24039059
07B	73	chr13:39058310-39058382
08B	80	chr8:4123338-4123417
09B	67	chr18:35306218-35306284
10C	74	chr10:108030958-108031031
11A	77	chr2:220052182-220052258
12A	72	chr7:73829609-73829680
13A	72	chr10:4709028-4709099
14A	69	chr2:54567127-54567195
15A	73	chr13:39934554-39934626
16A	68	chr2:76775067-76775134
17A	76	chr1:187466293-187466368
18A	70	chr9:17910175-17910244
19B	80	chr2:121798072-121798151
20A	63	chr2:159960331-159960393
21A	80	chr2:220054109-220054188

22A	66	chr13:72846230-72846295
23A	70	chr22:35646043-35646112
24A	83	chr6:95193814-95193896
25A	80	chr8:112294966-112295045
26A	83	chr1:159649196-159649278
27A	83	chr9:89155557-89155639
28A	78	chr6:18402416-18402493
29A	81	chr15:99575326-99575406
30A	72	chr15:39744438-39744509
31A	69	chr1:210082901-210082969
32A	74	chr1:152577254-152577327
33A	76	chr1:72798259-72798334
34A	84	chr20:21288000-21288083
36A	83	chr4:10232748-10232830
37A	71	chr5:57325054-57325124
38A	85	chr6:100035186-100035270
39A	76	chr7:148074400-148074475
41A	70	chr5:1924762-1924831

Figure 2-9: Distribution of CNV assay amplicon size.



2.3.1.5 Common SNPs and Repeating Elements

Common SNPs were present in 1/117 (0.9%) oligonucleotide binding sites. The probe binding site of assay CNV 15A included a common SNP, rs111271716, at the third last nucleotide (Appendix A, Figure 3). The allelic frequency for this SNP is 5.4% A and 94.6% G in dbSNP build 147.[249] No other assays were affected by common SNPs.

Interspersed repeats and low complexity DNA sequences were present in 12/39 (30%) CNV assays, affecting most or all of the assay region in 8/12 (67%). A variety of repeat classes were responsible and no single repeat class predominated (Table 2-13).

Table 2-13: Repeating elements located within CNV assays.

CNV #	Repeat Class	Distribution
03B	DNA repeat elements	Reverse primer
05B	Long terminal repeat elements	Entire assay
07B	DNA repeat elements	Entire assay
10C	Long terminal repeat elements	Entire assay
11A	Long interspersed nuclear elements	Forward primer, 1 nucleotide probe
12A	Long terminal repeat elements	Forward primer, 3 nucleotides probe
13A	Long interspersed nuclear elements	Entire assay
17A	Long terminal repeat elements	Most of assay
18A	Short interspersed nuclear elements	Forward primer, 3 nucleotides probe
24A	Long interspersed nuclear elements	Entire assay
27A	Long interspersed nuclear elements	Entire assay
28A	Short interspersed nuclear elements	Most of assay

2.3.1.6 Genes

9/39 (23%) CNV assays were located within genes, summarised in Table 2-14. In all cases, the gene spanned the entire CNV locus and could not be avoided during assay development. All assays were designed within introns. All genes were protein coding genes with the exception of the *LINC00704* gene (CNV13A).

Table 2-14: Assays positioned within known genes.

CNV #	Gene	Intron
04B	T-cell receptor alpha chain C region (<i>TCRA</i>)	2/5
08B	CUB and Sushi multiple domains 1 (<i>CSMD1</i>)	3/69
13A	Long intergenic non-protein coding RNA 704 (<i>LINC00704</i>)	1/3
14A	Chromosome 2 open reading frame 73 (<i>C2orf73</i>)	2/4
15A	Lipoma HMGIC fusion partner (<i>LHFP</i>)	3/3
20A	Tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 1 (<i>TANC1</i>)	4/11
28A	Ring finger protein 144B (<i>RNF144B</i>)	2/7
34A	5'-3' exoribonuclease 2 (<i>XRN2</i>)	1/28
39A	Contactin associated protein-like 2 (<i>CNTNAP2</i>)	21/23

2.3.1.7 Overlapping CNV loci

6/41 (15%) CNV loci were found to overlap in pairs (i.e. one CNV locus overlapped another CNV locus). These are summarised in Table 2-15. Of 3 pairs of overlapping loci, 2 pairs underwent assay development due to an error in the CNV evaluation process.

The relative positions of CNV loci 3 and 25, and their respective assays, are plotted in Appendix A, Figure 4. The CNV 25 locus sits entirely within the CNV 3 locus. Consequently, the assay for CNV 25A sits within the CNV 3 locus as well as the CNV 25 locus. The CNV 03B assay does

overlap with the CNV 25 locus. Therefore, CNV 25A assay results will be linked to the CNV locus 3 genotype as well as uniquely assaying CNV locus 25.

The relative positions of CNV loci 11 and 21, and their respective assays, are plotted in Appendix A, Figure 5. The CNV 21 locus sits entirely within the CNV 11 locus. Consequently, the assay for CNV 21A also sits within the CNV 11 locus. The CNV 11A assay does not sit within the CNV 21 locus. Therefore, CNV 21A assay results will be linked to the CNV locus 11 genotype as well as uniquely assaying CNV locus 21.

The relative positions of CNV loci 37 and 40, as well as assay CNV 37A, are plotted in Appendix A, Figure 6. The CNV 40 locus sits entirely within the CNV 37 locus. Consequently, any assay developed for CNV 40 would also sit within the CNV 37 locus. As this was identified before assay development, no assay was developed for this locus. The CNV 37A assay does not sit within the CNV 40 locus and is therefore not linked to the CNV 40 genotype.

CNV loci assay linkage is summarised in Table 2-16.

Table 2-15: Overlapping CNV loci and assay genomic coordinates.

Overlapping Locus A			Overlapping Locus B		
CNV #	Locus Hg19 Coordinates	Assay Hg19 Coordinates	CNV #	Locus Hg19 Coordinates	Assay Hg19 Coordinates
3	chr8:112294084-112296293	chr8:112295274-112295348	25	chr8:112294279-112295260	chr8:112294966-112295045
11	chr2:220052107-220054328	chr2:220052182-220052258	21	chr2:220053192-220054328	chr2:220054109-220054188
37	chr5:57323526-57333742	chr5:57325054-57325124	40	chr5:57326027-57333533	No assay developed

Table 2-16: Overlapping CNV loci assay linkage.

CNV #	Locus Linkage
03B	No linkage
25A	Linked to CNV locus 3 polymorphism
11A	No linkage
21A	Linked to CNV locus 11 polymorphism
37A	No linkage

2.3.2 Non-Polymorphic Assay Development

2.3.2.1 ACE Assay

Primers and a probe were developed for a region spanning the junction of exon 9/25 and intron 9/24 of the angiotensin I converting enzyme (*ACE*) gene (Online Mendelian Inheritance of Man no. 106180) (Appendix A, Figures 7 and 8). Assay characteristics are presented in Table 2-17.

The primers and probe were confirmed to be specific to the target region and in-silico PCR confirmed one product of 66 bp size specific to the target region.

This region was free of common SNPs and repeating elements, however, two structural variation polymorphisms have been reported at this locus and could not be avoided. The first variation, nsv962360, is a 15,962 bp duplication at position chr17:61544669-61560630 that was identified in 10/97 deeply sequenced great ape and human genomes as part of a study of diversity and rates of copy number and single nucleotide variation across the hominid phylogeny.[255] Notably, only 10/97 genomes in this study were human and all correspond to the samples containing the duplication variant. The second variation, nsv1055722, is a 35,987 gain at position chr17:61555666-61591652 detected by SNP array in 1/29,084 children with developmental delay undergoing CNV morbidity mapping.[256] The site of this assay is not located near the well-studied *ACE* I/D gene polymorphism (rs1799752) described by Tiet et al. and associated with a variety of clinical outcomes.[257]

2.3.2.2 *NPPB* Assay

Primers and a probe were also developed for a region spanning intron 2/2 of the natriuretic peptide precursor B (*NPPB*) gene (Online Mendelian Inheritance of Man no. 600295) (Appendix A, Figure 9). Assay characteristics are presented in Table 2-17. The primers and probe were confirmed to be specific to the target region and in-silico PCR confirmed one product of 76 bp size specific to the target region. This region was free of common SNPs, repeating elements and structural variation.

Table 2-17: *ACE* and *NPPB* non-polymorphic assay characteristics.

Parameter	ACE assay	NPPB assay
Forward Primer		
Sequence	GGTCTTTAGTGGGCGTACC	GCTCTCTGCCCTTGAGACTAT
Sense	+	+
Size (bp)	19	21
Tm (°C)	62.8	62.5
GC (%)	58	48
Reverse Primer		
Sequence	TCCTCTCACCGAAGATACCA	CCAGTGGTGCAGAAAGAGA
Sense	-	-
Size (bp)	20	19
Tm (°C)	62.9	62.7
GC (%)	50	53
Probe		
Sequence	CCTTCCCGCTACAACCTCGACTGG	AGAGGAAACCAGGAGGAAATGTTTGGT

Sense	+	+
Size (bp)	24	27
Tm (°C)	68.8	68.8
GC (%)	58	44
Primer ΔT_m (°C)	0.1	0.2
Probe-Primer ΔT_m (°C)	6	6.2
Amplicon		
Hg19	chr17:61560478-61560543	chr1:11917911-11917986
Size (bp)	66	76

2.4 DISCUSSION

This chapter details the assessment of 41 highly polymorphic CNV loci, and the design of 39 assays within suitable CNV loci, with the objective of developing a functional and informative ddPCR assay panel for chimerism analysis in a transplantation context.

2.4.1 CNV Assay Development

2.4.1.1 *Locus Assessment*

The CNV loci identified for assay development were distributed throughout the genome with a slight predominance on chromosomes 1 and 2. Chromosomes 1 and 2 each contain approximately 8% of cellular DNA, therefore, it is consistent that 15% of the CNV loci identified for assay development were located on these chromosomes.

The median CNV locus size of 1,829 bp was similar to high resolution population-level characterisation of CNV (median 729 bp, mean 8,000 bp).[258] The CNV locus selection process was biased towards smaller CNV loci, as these are generally less likely to impact genic sequences and have better sequence uniqueness as a result of the underlying aetiological genetic mechanisms.[258]

Multiplying the average allelic frequencies of the panel, the probability of a particular genotype combination can be estimated. For a two individual transplant arrangement (one donor and one recipient), the probability that any individual is one-copy for a particular CNV locus while the other is zero-copy for that locus is 0.21 (0.16-0.25). The probability that one individual is two-copy for a particular CNV locus and the other is zero-copy for that locus is 0.04 (0.02-0.08). These combinations of CNV genotype amongst individuals in a transplant arrangement are considered “informative” for the individual expressing the non-zero-copy genotype. Quantification techniques for an assay at this locus can be used to provide information regarding the degree of chimerism for that particular individual after transplantation.

In the case of two-copy informative genotypes, these CNV loci can be distinguished on the basis that assays targeting these loci will produce quantitatively double the amount of PCR product than the more common one-copy informative genotype. This may be demonstrated (i) via genotyping of a non-chimeric sample sourced from that individual, or (ii) via simple comparison of assay concentration patterns from multiple informative assay results for that individual in a chimeric sample. Non-chimeric samples may include DNA extracted from a sample obtained before transplantation, or DNA extracted from cells in a typically non-chimeric compartment (e.g. buccal mouth swab or recipient leucocytes following solid organ transplantation).

Increasing the number of individuals in a transplant arrangement reduces the likelihood that any particular assay will be informative for a single individual in the arrangement. For example, in a three individual transplant arrangement (e.g. sequential transplantation, or dual donor allogeneic bone marrow transplantation), the probability of a one-copy informative result (remaining individuals zero-copy) is 0.11 (0.07-0.15) and the probability of a two-copy informative result is 0.02 (0.01-0.05), using the average allelic frequencies of the panel. In such scenarios, or when the individuals in a transplant arrangement are related to one another, a larger panel size will be needed to ensure sufficient informative markers can be identified. As an example, a median of 3.3 and 4.3 informative assays could be expected in a three individual transplant arrangement using a panel size of 30 and 39 markers respectively.

2.4.1.2 *Assay Design*

Assays suitable for ddPCR were developed for most of the CNV loci identified as suitable for chimerism analysis.

Amplicon length was minimised to a median of 74 (70-80) bp to optimise quantification of fragmented, degraded cfDNA. Assay oligonucleotide characteristics were consistent with recommendations for ddPCR assay design.[253] Within the limitations of Blat, sequence non-specificity was minimal, and if present, limited to only one assay oligonucleotide in all but one instance.[254] Additionally, where 100% sequence identity non-specificity was present, this typically did not extend the entire length of the oligonucleotide. Consequently, Blat scores, which penalise mismatches, were relatively low.[259] No assays had non-specific amplicons, and all assay-specific amplicons were confirmed to localise to the target CNV locus.

Careful assay design avoided the presence of common SNPs in all but one oligonucleotide binding sequence. Repeating elements were more difficult to avoid, and were present in the entirety of 6/39 (15%) assays, although this did not affect assay specificity. Other forms of larger structural variation, such as insertion/deletion polymorphism and CNV, also overlapped with assay sites (Appendix B). Given the ubiquity of such variation, this could not be avoided. The probability is that the degree of polymorphism at these sites is low, however three overlapping highly

polymorphic CNV loci were identified. In one instance, assay design for the overlapping locus was not performed (CNV locus 40), however two other assays (CNV assays 25A and 21A) were developed in error and targeted highly polymorphic CNV loci located within or overlapping another highly polymorphic CNV locus.

These assays sited within two overlapping highly polymorphic CNV (“nested assays”) will therefore be subject to polymorphism of the targeted CNV locus in addition to polymorphism of the CNV locus overlapping that target sequence, assuming no other factors affect the relationship between these CNV loci. If the overlapping CNV locus is zero-copy, so too will be the nested locus. Thus, the zero-copy allelic frequencies of these nested assays will be increased and the probability that they will be informative is reduced. In instances where the overlapping CNV locus is a one-copy genotype, the nested assays will provide additional discriminative and quantitative value for chimerism analysis. In instances where the overlapping CNV locus is a two-copy genotype, there is potential for confounding when the nested assay targets a non-zero-copy genotype. In such instances, there may be artefactual two-copy, three-copy or four-copy results. For this reason, use of these nested assays beyond the present work should be avoided.

A quarter of assays were designed within intronic regions of known genes. Whilst this is not anticipated to affect the performance of the assay with respect to chimerism analysis *per se*, it was considered undesirable as the result of the assay may confer additional diagnostic information and pose ethical or logistical dilemmas in future. Variants affecting intronic regions may affect gene expression via alterations to gene regulatory elements and splice sites, and there remain significant knowledge gaps with relation to the role and impact of intronic polymorphism.[260] At the time of CNV locus selection, CNV without known associations to human disease were selected, however the selection process did not take into account other sources of genetic variation within these sequences. Whether genic or agenic, it is impossible to be certain that all assay sites will be of no clinical significance currently or in the future. Thus, these genic assay sites are not a contraindication to clinical use of the assay. Like genomics-based diagnostics, pre-specification of the scope of the assay and diagnostic outcomes will be important in any application of this test.

2.4.2 Non-Polymorphic Assay Development

In addition to the development of assays targeting highly polymorphic loci for the purposes of chimerism analysis, two assays were developing targeting conserved non-polymorphic loci for use as control assays. Identification of agenic, non-polymorphic regions was very difficult, and so assays were developed targeting sequence within two known genes, *ACE* and *NPPB*. Both assays target regions presently subject to no-low level polymorphism and are expected to exhibit a two-copy genotype in the vast majority of the population. The oligonucleotides comprising these assays met the same design characteristics of the rest of the CNV assay panel.

2.4.3 **Summary**

In summary, 39 assays targeting highly polymorphic CNV and 2 assays targeting non-polymorphic gene sequences were developed. Oligonucleotide characteristics matched those recommended for ddPCR, and the *in silico* selection and evaluation process of the CNV loci promised utility for chimerism studies in related non-identical and non-related transplant arrangements of two or more individuals.

Analytical validation in fragmented and non-genomic DNA is required to establish the diagnostic performance characteristics of each individual assay and the panel as a whole. The allelic frequencies of the CNV panel, and their diagnostic utility for chimerism analysis after transplantation, require further confirmation in clinical studies using post-transplantation chimeric samples.

CHAPTER 3 INFORMATIVITY GENOTYPING

3.1 INTRODUCTION

The developed CNV assay panel targets specific sequences subject to highly heterozygous CNV. When run against chimeric cfDNA, each CNV assay concentration is a function of the genotypes of the recipient and the donor at that locus, and the proportionate contribution of the recipient and graft to the total cfDNA concentration. Identification of these recipient/donor genotype combinations is therefore a key step in the CNV assay panel methodology.

The quantification of graft-derived cfDNA requires that the results of graft informative CNV assays are distinguished from CNV assays of other recipient/donor genotype combinations. Graft informative CNV assays target CNV loci that are one- or two- copy in the donor (graft) genome, but are zero-copy in the recipient genome (“recipient negative/donor positive”). This provides a negative background against which graft-derived cfDNA can be quantified in absolute terms. Given this recipient/donor genotype combination, any detectable DNA sequence specific to these CNV loci in recipient body fluids must have been produced as the result of graft cfDNA release, either by apoptosis, necrosis, active secretion or another undefined mechanism.

For the purpose of graft-derived cfDNA quantification, the primary genotyping distinction is between graft informative markers and those which are one- or two- copy within the recipient genome (“recipient positive”). Recipient positive assays may also be present as one- or two- copy within the donor genome (“recipient positive/donor positive”). In these circumstances, summation of the respective contributions of recipient and graft to circulating cfDNA may occur, and the distinction of the graft contribution may be less clear than it might otherwise be against the negative background of the recipient negative/donor positive genotype combination.

The final key genotype combination is one in which both the recipient genome and the donor genome are zero-copy for a particular CNV locus (“recipient negative/donor negative”). In this instance, a CNV assay targeting this locus should consistently return an undetectable concentration result, and further testing of the assay longitudinally over time will yield no useful diagnostic information, with the exception of acting as a negative control.

A key challenge, therefore, to any post-transplant chimerism assessment is the establishment of the appropriate genotype for each marker and for each individual in the transplant arrangement (recipient and donors). Of particular importance for the quantification of graft-derived cfDNA is the identification of graft informative markers. Early methods that relied upon sex-mismatch (Y chromosome markers in female recipients of male donor organs) or HLA-mismatch were able to exploit readily evident or pre-existing information to identify markers that were informative for

the graft, but at the expense of abundance and discriminatory capability.[16, 20, 22, 26, 128-130, 136, 137]

The use of highly polymorphic genomic variation (SNPs, in-del polymorphisms, CNV) for post-transplant chimerism assessment has the advantage of abundant and discriminatory assay targets, with the disadvantage that donor and recipient genotypes may not be readily evident. Many approaches, therefore, relied upon separate pre-transplant donor and recipient genotyping.[131, 141, 159, 163, 261] Such approaches involved running the assay methodology against non-chimeric DNA typically extracted from donor and recipient cells before transplantation (genomic DNA). In this manner, the genotype of all markers was established for donor and recipient *a priori*.

A disadvantage of this approach is the requirement for donor genetic material, which poses challenges in deceased donor transplantation, particularly between geographically disparate health services. Unless the role of chimerism genotyping is performed by the tissue typing laboratory using a standardised methodology, existing deceased donation systems would need to be modified to permit sample collection and transmission to a transplant centre-specific third party chimerism laboratory for DNA genotyping. Another approach might include separation of tissue from the organ prior to implantation, with later DNA extraction via the implanting transplant centre pathology service.

In the preliminary application of our CNV-based post-transplant cfDNA chimerism test using quantitative real-time PCR, we avoided these donor genotyping challenges by typing the recipient only. In this approach, the CNV assay panel was run against recipient genomic DNA extracted from buffy coat leucocytes in the post-transplant sample. The assumption was that leucocyte cellular microchimerism was negligible, if present at all, and that the vast majority of circulating leucocytes were of recipient origin. Recipient positive markers were identified and, by definition, could not be graft informative. Thus, any recipient negative but detectable markers must be graft informative.

Genotyping using recipient positive markers is a key benefit of copy number-based chimerism assessment. This is because these approaches rely upon a “negative background” (recipient zero-copy allele) against which graft informative markers are quantified. This can be contrasted to SNP-based chimerism assessment, where multiple SNP genotypes must be distinguished from one another, and no “negative background” exists. In these instances, recipient and donor genotyping was more fundamental to chimerism testing, although some bioinformatic approaches have been developed to address this problem.[152, 164]

The observation, by our group and others, that post-transplant graft-derived cfDNA typically accounts for a small proportion (<5%) of total plasma cfDNA raised the possibility that CNV recipient and graft informative genotyping could be performed without separate non-chimeric genotyping of donor and/or recipient. We hypothesised that recipient and donor genotyping could be conducted on the basis of assay concentration alone. Such an approach would simplify the assay methodology and may improve access to testing, reduce costs and reduce turnaround time.

The aim of this work is to study the patterns of CNV assay concentrations in cfDNA samples, genotype on the basis of these patterns, and establish the performance of concentration-based assay genotyping using recipient and donor genomic DNA genotypes as the gold standard.

3.2 METHODS

3.2.1 Participants

Adult kidney transplant recipients participating in a translational study involving the development and clinical application of the presented CNV-based chimerism assay panel were included in this study. All participants gave written informed consent, and the study was approved by the Austin Health Human Research Ethics Committee.

Ethics committee approval was also obtained from the Australian Red Cross Blood Service and the Australian organ donation service, DonateLife, to access stored adult deceased donor DNA from 20 donors to recipients included in the CNV-based chimerism assay panel study.

3.2.2 Recipient Blood Collection and Processing

3.2.2.1 *Phlebotomy*

Blood was collected in 10 mL Streck Cell-Free DNA BCT® (Streck, Omaha, NE, USA). Where possible, venesection was performed in accordance with Streck Phlebotomy Best Practices and Streck Cell-Free DNA BCT® Instructions For Use.[262, 263] Phlebotomy was ideally performed using a large antecubital vein and a 21 gauge needle attached to a BD Vacutainer® one-use holder (Becton, Dickinson and Company, Franklin Lakes, NJ, USA). Two full tubes were sought. After venesection was completed, the tubes were gently inverted multiple times.

The protocol permitted any method of venesection for sample collection. When phlebotomy was difficult or impossible, samples were collected with variations including a 23 gauge needle, a needle attached to a 20 mL syringe with later transfer to Streck Cell-Free DNA BCT®, collection from a smaller wrist or hand vein, collection of less than 20 mL total sample, and collection from an intravenous cannula, arteriovenous fistula or central line. Where collection was performed from an intravenous cannula or central line, collection after an intravenous infusion was avoided, and 10 mL blood was discarded prior to sample collection.

Collection time and date was recorded. Samples were stored and transported at room temperature until plasma separation was performed.

3.2.2.2 Plasma Separation and Storage

Blood fractionation and plasma separation was performed within 7 days of sample collection. Time and date of plasma separation was recorded.

Blood was poured from Streck Cell-Free DNA BCT® into Falcon™ 15 mL Conical Centrifuge Tubes (Corning Life Sciences, Tewksbury, MA, USA). These were then centrifuged for 10 minutes at 1,600 x g using the Heraeus Megafuge 1.0R general-purpose centrifuge (ThermoFisher Scientific, Waltham, MA, USA). The supernatant plasma was then pipetted into 1.5 mL Eppendorf Tubes® (Eppendorf South Pacific, North Ryde, New South Wales, Australia) (or equivalent). These were then centrifuged for 10 minutes at 16,000 x g. The supernatant plasma was again pipetted into new 1.5 mL Eppendorf Tubes® (or equivalent). These were labelled with the sample number and “plasma” to indicate that cfDNA extraction had not yet been performed. Presence or absence of haemolysis was recorded. Separated plasma was then stored at -80 °C.

At all sample transfer stages, mixing of sample collected from the same patient at the same time point in multiple Streck Cell-Free DNA BCT® was permitted unless otherwise specified.

3.2.2.3 Cell Pellet Storage

The packed cell pellet remaining after the first centrifugation step and plasma separation described in section 3.2.2.2 was stored in selected cases. This included (i) the first blood sample for an individual, and (ii) scenarios where additional cellular DNA from an individual was required or anticipated. Taking care not to disturb the cell pellet, it was stored at -10 °C. The tube was labelled with the participant number and sample number.

3.2.2.4 Leucocyte Genomic DNA Extraction

Genomic DNA was extracted from buffy coat leucocytes using the Nucleobond Blood Extraction Kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's instructions. Quantification of extracted DNA was performed using the NanoDrop ND-1000 spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA) according to manufacturer's instructions.

3.2.2.5 Plasma Cell-Free DNA Extraction

CfDNA was extracted from plasma using the QIAamp® Circulating Nucleic Acid Kit (QIAGEN GmbH, Hilden, Germany) and a protocol derived from the QIAamp® Circulating Nucleic Acid Handbook October 2013 version.[264] All buffers and QIAGEN products described in this section are sourced from the QIAamp® Circulating Nucleic Acid Kit.

Frozen plasma samples were allowed to equilibrate to room temperature. Plasma was pipetted from 1.5 mL Eppendorf Tubes® (or equivalent) into Falcon™ 50 mL Conical Centrifuge Tubes (Corning Life Sciences, Tewksbury, MA, USA) labelled with the sample number. Samples of total volume greater than 5 mL were split into two 50 mL Falcon™ Tubes and labelled with a suffix (A, B, C) following the sample number. Phosphate-buffered saline (PBS) was added to each tube to bring the total volume to the nearest whole mL. Plasma samples less than 3 mL were made up to 3 mL minimum using PBS. QIAGEN Proteinase K was added to each tube based upon the total sample volume (Table 3-1). A buffer ACL and carrier RNA in buffer AVE master mix was created based on the number of tubes of similar sample volume to be extracted in each specific batch (Table 3-2). This master mix was then added to each tube according to the total sample volume (Table 3-1). The solution was pulse-vortexed for 30 seconds and incubated at 60 °C for 30 minutes.

After incubation was complete, buffer ACB was added to the lysate in each tube according to the sample volume (Table 3-1). The lysate-Buffer ACB mixture was pulse-vortexed for 15 seconds and then incubated on ice for 5 minutes. Each tube of lysate-Buffer ACB mixture was then poured into a tube extender attached to a QIAamp® Mini column that had been labelled with the corresponding sample number in advance. Each QIAamp® Mini column was attached via VacConnector to the QIAvac 24 Plus.

All lysates were simultaneously drawn through the QIAamp® Mini column under vacuum pressure over 15-20 minutes. After this was complete for all columns, vacuum pressure was released and 600 µL of Buffer ACW1 was added to the column. Vacuum pressure was reapplied until all solution was drawn through the column, after which vacuum pressure was released. This process was repeated for 750 µL of Buffer ACW2 followed by 750 µL of 100% ethanol.

The QIAamp® Mini columns, placed into clean 2 mL collection tubes, were then centrifuged for 3 minutes at 20,000 x g. The columns were then transferred into clean 2 mL collecting tubes and incubated with lids open at 56 °C for 10 minutes.

The QIAamp® Mini columns were then transferred to clean 1.5 mL elution tubes that had been labelled with the corresponding sample number. 50 µL of distilled water was applied to the middle of each column membrane. The column lid was then closed and the column was incubated at room temperature for 3 minutes. After this, the column-1.5 mL elution tube assembly was centrifuged for 1 minute at 20,000 x g. The eluent was then reapplied to the middle of the respective column membrane and, after a 3 minute room temperature incubation, centrifuged again for 1 minute at 20,000 x g.

The volume of the eluent for each column was measured and recorded. For samples with total volumes >5 mL that had been split and run through two columns, eluent from the two columns was combined at this point into the A eluent tube unless otherwise specified. Tubes were labelled with the sample number and suffix as appropriate prior to storage at -80 °C.

Table 3-1: Plasma cfDNA extraction reagent volumes based upon the total volume of sample in each tube (after being brought to the nearest whole mL with PBS).

Reagent	Volume of reagent added to tube		
	3 mL Sample	4 mL Sample	5 mL Sample
QIAGEN Proteinase K	300 μ L	400 μ L	500 μ L
Buffer ACL and carrier RNA in buffer AVE	2.4 mL	3.2 mL	4 mL
Buffer ACB	5.4 mL	7.2 mL	9 mL

Table 3-2: Buffer ACL and carrier RNA master mix composition used for plasma cfDNA extraction based on number of tubes needing extraction and the total volume of sample in each tube (after being brought to nearest whole mL with PBS)

Number of Tubes for Extraction	Buffer ACL (mL)			Carrier RNA in Buffer AVE (μ L)
	3 mL Sample	4 mL Sample	5 mL Sample	
1	2.6	3.5	4.4	5.6
2	5.3	7.0	8.8	11.3
3	7.9	10.6	13.2	16.9
4	10.6	14.1	17.6	22.5
5	13.2	17.6	22.0	28.1
6	15.8	21.1	26.4	33.8
7	18.5	24.6	30.8	39.4
8	21.1	28.2	35.2	45.0
9	23.8	31.7	39.6	50.6
10	26.4	35.2	44.0	56.3
11	29.0	38.7	48.4	61.9
12	31.7	42.2	52.8	67.5

3.2.3 Donor Blood Collection and Processing

Buffy coat leucocyte donor DNA extraction was performed using tissue typing samples collected during the workup process prior to deceased organ donation.

Limited information is available regarding the methods of sample processing and DNA extraction, in part because the samples used for this work were collected and extracted over a 20-year period and a variety of DNA extraction methodologies were employed by the clinical laboratory over that time. More recent DNA extraction was performed using the QIAcube (QIAGEN GmbH, Hilden, Germany) automated DNA extraction system, according to manufacturer instructions.

DNA was shared at a concentration of approximately 15-20 ng/ μ L. Quantification of extracted DNA was performed using the NanoDrop ND-1000 spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA) according to manufacturer's instructions.

3.2.4 Recipient Urine Collection and Processing

3.2.4.1 *Urine Collection*

Urine was collected in a 70 mL sterile urine specimen jar. Voiding participants were asked to collect a mid-stream urine sample. Catheterised participants had urine collected from the catheter access port and not the drainage bag or bottle. A minimum of 10 mL urine was sought.

Collection time and date was recorded. Urine was stored on ice immediately after collection and at 4 °C until processing could commence.

3.2.4.2 *Urine Processing and Storage*

Urine was processed as soon as practicable after sample collection. This was ideally within 1 hour. Time and date of urine processing was recorded. Presence of macroscopic haematuria was recorded.

Urine was transferred by sterile pipette from the 70 mL sterile urine specimen jar into 10 mL BD Vacutainer® plastic whole blood spray-coated K2EDTA tubes (Becton, Dickinson and Company, Franklin Lakes, NJ, USA). The K2EDTA tubes were gently inverted multiple times before being centrifuged at 4 °C for 10 minutes at 1,600 x g using the Heraeus Megafuge 1.0R general-purpose centrifuge (ThermoFisher Scientific, Waltham, MA, USA).

The urine supernatant was filtered through a 5 µm syringe filter before being pipetted into 1.5 mL Eppendorf Tubes® (or equivalent). These tubes were then centrifuged at 4 °C for 10 minutes at 16,000 x g using the Eppendorf 5415R micro-centrifuge (Eppendorf South Pacific, North Ryde, New South Wales, Australia). The supernatant urine was pipetted into new 1.5 mL Eppendorf Tubes® (or equivalent). These were labelled with the sample number and “urine” to indicate that cfDNA extraction had not yet been performed. Processed urine was then stored at -80 °C.

3.2.4.3 *Urine Cell-Free DNA Extraction*

CfDNA was extracted from urine using the QIAamp® Circulating Nucleic Acid Kit (QIAGEN GmbH, Hilden, Germany) and a protocol derived from the QIAamp® Circulating Nucleic Acid Handbook October 2013 version.[264] All buffers and QIAGEN products described in this section are sourced from the QIAamp® Circulating Nucleic Acid Kit.

Frozen urine samples were allowed to equilibrate to room temperature. Urine was pipetted from 1.5 mL Eppendorf Tubes® (or equivalent) into Falcon™ 50 mL Conical Centrifuge Tubes (Corning Life Sciences, Tewksbury, MA, USA) labelled with the sample number. Samples of total volume greater than 4 mL were split into two 50 mL Falcon™ Tubes and labelled with a suffix (A, B, C) following the sample number. PBS was added to each tube to bring the total volume to the nearest whole mL. Urine samples less than 3 mL were made up to 3 mL minimum using PBS. QIAGEN Proteinase K was added to each tube based upon the total sample volume

(Table 3-3). A buffer ACL and carrier RNA in buffer AVE master mix was created based on the number of tubes of similar sample volume to be extracted in each specific batch (Table 3-4). This master mix was then added to each tube according to the total sample volume (Table 3-3). The solution was pulse-vortexed for 30 seconds and incubated at 60 °C for 30 minutes.

After incubation was complete, buffer ACB was added to the lysate in each tube according to the sample volume (Table 3-3). The lysate-Buffer ACB mixture was pulse-vortexed for 15 seconds and then incubated on ice for 5 minutes. Each tube of lysate-Buffer ACB mixture was then poured into a tube extender attached to a QIAamp® Mini column that had been labelled with the corresponding sample number in advance. Each QIAamp® Mini column was attached via VacConnector to the QIAvac 24 Plus.

All lysates were simultaneously drawn through the QIAamp® Mini column under vacuum pressure over 15-20 minutes. After this was complete for all columns, vacuum pressure was released and 600 µL of Buffer ACW1 was added to the column. Vacuum pressure was reapplied until all solution was drawn through the column, after which vacuum pressure was released. This process was repeated for 750 µL of Buffer ACW2 followed by 750 µL of 100% ethanol.

The QIAamp® Mini columns, placed into clean 2 mL collection tubes, were then centrifuged for 3 minutes at 20,000 x g. The columns were then transferred into clean 2 mL collecting tubes and incubated with lids open at 56 °C for 10 minutes.

The QIAamp® Mini columns were then transferred to clean 1.5 mL elution tubes that had been labelled with the corresponding sample number. 50 µL of distilled water was applied to the middle of each column membrane. The column lid was then closed and the column was incubated at room temperature for 3 minutes. After this, the column-1.5 mL elution tube assembly was centrifuged for 1 minute at 20,000 x g. The eluent was then reapplied to the middle of the respective column membrane and, after a 3-minute room temperature incubation, centrifuged again for 1 minute at 20,000 x g.

The volume of the eluent for each column was measured and recorded. For samples with total volumes >4 mL that had been split and run through two columns, eluent from the two columns was combined at this point into the A eluent tube unless otherwise specified. Tubes were labelled with the sample number and suffix as appropriate prior to storage at -80 °C.

Table 3-3: Urine cfDNA extraction reagent volumes based upon the total volume of sample in each tube (after being brought to the nearest whole mL with PBS).

Reagent	Volume of reagent added to tube		
	2 mL Sample	3 mL Sample	4 mL Sample
QIAGEN Proteinase K	250 μ L	375 μ L	500 μ L
Buffer ACL and carrier RNA in buffer AVE	2 mL	3 mL	4 mL
Buffer ACB	5.4 mL	7.2 mL	9 mL

Table 3-4: Buffer ACL and carrier RNA master mix composition used for urine cfDNA extraction based on number of tubes needing extraction and the total volume of sample in each tube (after being brought to nearest whole mL with PBS)

Number of Tubes for Extraction	Buffer ACL (mL)			Carrier RNA in Buffer AVE (μ L)
	2 mL Sample	3 mL Sample	4 mL Sample	
1	2.2	3.3	4.4	5.6
2	4.4	6.6	8.8	11.3
3	6.6	9.9	13.2	16.9
4	8.8	13.2	17.6	22.5
5	11.0	16.5	22.0	28.1
6	13.2	19.8	26.4	33.8
7	15.4	23.1	30.8	39.4
8	17.6	26.4	35.2	45.0
9	19.8	29.7	39.6	50.6
10	22.0	33.0	44.0	56.3

3.2.5 Droplet Digital PCR

The reagents used for each PCR well are summarised in Table 3-5. Each well included a FAM- and HEX- based assay (duplex arrangement). The two-copy non-polymorphic control assay *ACE* was run at least once for each sample. Some samples were run using 2.5 μ L of template DNA, while others were run using 5 μ L of template DNA, depending on sample volume and availability. CfDNA extraction eluent was added neat, whereas genomic DNA was diluted to a concentration of approximately 5 ng/ μ L prior to addition to the PCR well.

Droplet generation was performed using the Bio-Rad QX200 AutoDG Automated Droplet Generator, according to manufacturer instructions. A Bio-Rad C/S1000 thermal cycler was used for PCR. Cycling conditions are specified in Table 3-6. Plates were held at 4 °C while awaiting the droplet reading.

Droplets were read using the Bio-Rad QX200 Droplet Reader and absolute quantification experiment settings with ddPCR Supermix for Probes (no dUTP) specified. Channel 1 and channel 2 were specified as unknown for all assay wells.

Bio-Rad QuantaSoft software (version 1.7) was used for droplet reading and analysis. Positive droplet thresholds were set *en bloc* for the same assays within each plate, one third of the distance between the double negative cluster and the respective channel positive cluster.

Table 3-5: PCR reaction components for the genotyping studies.

PCR Reagent	Concentration	Volume per Well (μL)		Final Concentration
ddPCR Supermix for Probes (No dUTP)	2x	12.5		1x
FAM-Assay Primers	10 uM	2.25		900 nM
FAM-Assay Probe	10 uM	0.625		250 nM
HEX-Assay Primers	10 uM	2.25		900 nM
HEX-Assay Probe	10 uM	0.625		250 nM
Distilled Water	0 uM	4.25	1.75	0 uM
DNA	Variable	2.5	5	Variable
Total		25		

Table 3-6: PCR cycling conditions for the genotyping studies. Ramp rate was 2 °C/sec.

Cycling Step	Temp (°C)	Time	Number of Cycles
Enzyme activation	95	10 min	1
Denaturation	94	30 sec	40
Annealing/extension	60	1 min	
Enzyme deactivation	98	10 min	1
Hold	4	Infinite	2

3.2.6 Conversion of Results

Droplet digital PCR results for genomic DNA were expressed as copies per well. Droplet digital PCR results for cfDNA samples were converted to sample equivalent concentrations using the following equation, the derivation of which is discussed in detail in Chapter 4:

$$C_{sample} = \frac{C_{PCR} \times V_{elution} \times (V_{template} + V_{MM})}{V_{sample} \times V_{template} \times k} \quad 3-1$$

where C_{sample} is the concentration of target within the collected sample in copies/mL, C_{PCR} is the concentration of target within the PCR reaction in copies/μL as measured by the quantification methodology, $V_{elution}$ is the volume of eluent used during the cfDNA extraction process in μL, $V_{template}$ is the volume of extracted template cfDNA used in the PCR reaction in μL, V_{MM} is the volume of all other PCR reaction components in μL, V_{sample} is the volume of sample used for cfDNA extraction in mL, and k is an unquantified extraction yield variable. The volume of PCR reaction components V_{MM} includes the summed volumes of distilled water, primers, probes and ddPCR Supermix.

3.2.7 Blood Genotyping

Genotyping for each CNV assay was performed separately for recipient cfDNA, recipient genomic DNA, and/or donor genomic DNA. Genotyping of each DNA source was performed blinded to the corresponding genotype allocations from other DNA sources. For example, genotypes for cfDNA were applied without knowledge of or reference to the genotypes of corresponding assays in recipient and donor genomic DNA.

Genotypes were allocated with consideration of (i) concentration, (ii) clustering with other assays of similar concentration, (iii) expected allelic frequencies of each genotype, and (iv) concentration with respect to the two-copy non-polymorphic control assay *ACE*. Recipient and donor genomic DNA were genotyped as zero-copy, one-copy or two-copy. CfDNA was genotyped using a simple convention based upon the three key genotypes of significance for transplant chimerism assessment (Table 3-7).

Table 3-7: Simple genotyping definitions applied to plasma cfDNA sample assays.

Simple Genotype	Included Recipient Genotypes	Included Donor Genotypes
Graft Informative ("Graft")	Zero-Copy	One-Copy Two-Copy
Recipient Positive ("Recipient")	One-Copy Two-Copy	Zero-Copy One-Copy Two-Copy
Recipient Negative/Donor Negative ("Zero")	Zero-Copy	Zero-Copy May also include donor positive assays below the limit of detection

3.2.8 **Recipient Genotyping Analysis**

The recipient genotype of a particular CNV assay run against recipient cfDNA was compared against the genotype of the corresponding CNV assay run against recipient genomic DNA. The comparison only included CNV assays run against samples collected from the same recipient. In some instances, more than one cfDNA sample was required to ensure complete genotype coverage of the assay panel.

First, the performance of cfDNA for distinguishing recipient positive markers from recipient negative markers was evaluated. Sensitivity and specificity were calculated. Then, excluding recipient negative markers, the performance of cfDNA for quantification of recipient copy-number for each CNV locus was evaluated. The true positive rate for one-copy and two-copy genotypes was calculated. In both analyses, genomic DNA was considered the gold standard.

3.2.9 **Donor Genotyping Analysis**

The donor genotype of CNV assays run against donor genomic DNA was compared against corresponding assays run against recipient cfDNA and genotyped as graft informative. Recipient positive markers were excluded.

The performance of cfDNA for distinguishing graft informative markers from zero-copy markers was evaluated. The sensitivity and specificity were calculated. Donor genomic DNA was considered the gold standard.

3.2.10 **Urine Genotyping**

Genotyping for each CNV assay run against urine cfDNA was performed with reference to plasma cfDNA genotyping results. Genotypes were allocated with consideration of (i) concentration, (ii) clustering with other assays of similar concentration, (iii) expected allelic frequencies of each genotype, (iv) plasma cfDNA recipient genotype, and (v) concentration with respect to the two-copy non-polymorphic control assay *ACE*. Each allocated genotype consisted of the recipient genotype copy number prefixed by an “R” and the donor/graft genotype copy number prefixed by a “G” (e.g. R2G1). The distribution of allocated genotypes amongst studied samples was described.

Graft-derived cfDNA was calculated as the median concentration of graft informative assays. Total cfDNA was calculated by halving the *ACE* assay concentration for urine samples, and was calculated as the median recipient one-copy concentration for plasma samples. Graft fraction was calculated by dividing the graft-derived cfDNA concentration by the total cfDNA concentration.

3.2.11 **Bioinformatics and Statistical Methods**

Genotyping, data management, and statistical analyses were performed using R (R Foundation for Statistical Computing; <https://www.R-project.org>).[265] Non-parametric distributions were compared using the Kruskal-Wallis rank sum test. A p -value < 0.05 was considered significant.

3.3 **RESULTS**

3.3.1 **Blood**

3.3.1.1 *Cell-Free DNA Genotyping*

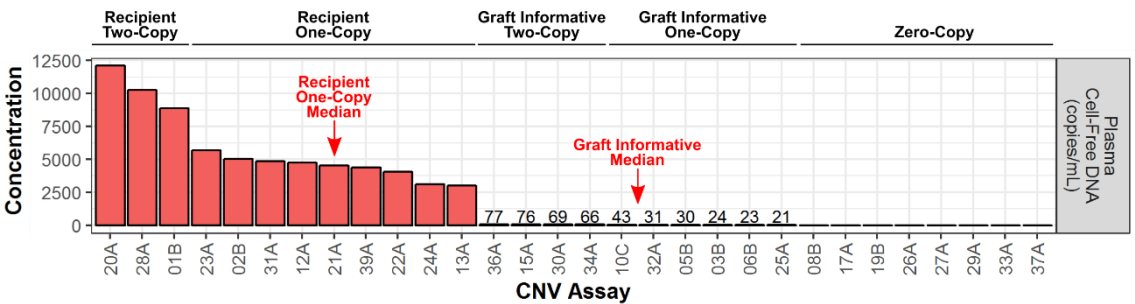
Arrangement of cfDNA CNV assay results by concentration revealed a characteristic pattern comprising 4-5 clusters of assays (Figure 3-1). A small cluster of assays with the highest concentrations were genotyped as recipient two-copy. A larger cluster of assays manifesting intermediate concentrations approximately half that of the recipient two-copy concentrations (or the *ACE* assay result) were genotyped as recipient one-copy. If there was only one cluster of high concentration assays, these assays were genotyped as recipient one-copy on the basis that copy-number frequencies made this genotype most likely. This was confirmed on the basis of the *ACE* assay result, which was approximately double the recipient one-copy assay concentrations. It was not possible to discriminate recipient positive/donor positive assays from recipient positive/donor negative assays. All recipient positive assays exhibited some inter-assay variation, and no clear cluster separation was observed within recipient one-copy or recipient two-copy genotypes.

CNV assays with undetectable results were genotyped as zero-copy assays. CNV assays with detectable results that were less than 10% of the concentration of the recipient one-copy CNV

assays were genotyped as graft informative. In instances when the median graft informative assay concentration was high and numerous graft informative markers were available for review, graft informative markers could be copy number genotyped (as in Figure 3-1). In the majority of instances, however, graft informative markers were present at low concentration and it was impossible to establish clear clusters of graft informative copy number genotypes with certainty. Consequently, graft informative markers were genotyped as graft informative only, without copy number assignment.

For the purpose of the diagnostic application of the CNV assay panel, the median concentration of the graft informative markers was used as a measure of graft-derived cfDNA, and the median concentration of recipient one-copy markers was used as a measure of total cfDNA (marked in Figure 3-1). Graft fraction, or the proportion of total cfDNA that is graft-derived, was calculated as the median graft informative CNV assay concentration divided by the median recipient one-copy CNV assay concentration.

Figure 3-1: CNV assay results for transplant recipient chimeric cfDNA extracted from plasma. Sorting assays by concentration permitted genotyping based on clusters of similar concentration, as specified. The concentration of low positive assay concentrations is printed for clarity. The median of the recipient one-copy CNV assays (a measure of total cfDNA concentration) and the median of the graft informative CNV assays (a measure of graft-derived cfDNA concentration) are indicated.



3.3.1.2 Algorithmic Genotyping

All cfDNA genotyping, including this work, was allocated manually by the researcher. Two proof-of-principle algorithmic methods of genotyping were developed to establish the feasibility of automatic genotyping.

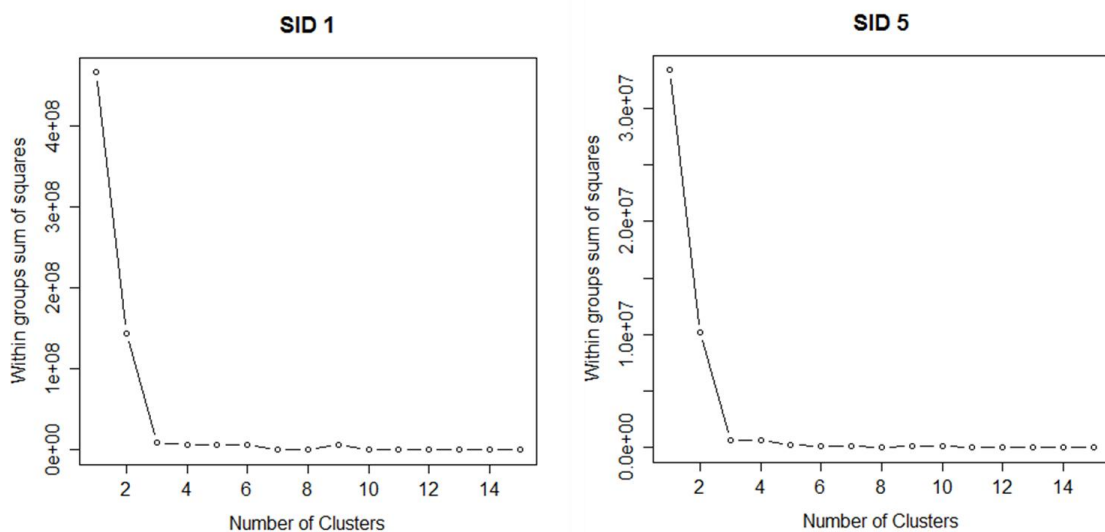
In the first approach, CNV assays 0.75-1.25x the concentration of the two-copy non-polymorphic *ACE* assay were genotyped as recipient two-copy, and CNV assays 0.25-0.75x the concentration of the *ACE* assay were genotyped as recipient one-copy. CNV assays that were detectable and less than 0.05x the concentration of the *ACE* assay were genotyped as graft informative and those

that were undetectable were given the zero genotype. CNV assay results between 0.05-0.25x the *ACE* assay were flagged for review.

In the second approach, genotyping was performed using k-means cluster analysis. Undetectable CNV assays were excluded from the cluster analysis, leaving 3-4 possible genotypes depending on whether graft informative genotypes could be copy number genotyped. The cluster number was defined on the basis of the within group sum of squares deflection point (Figure 3-2).

Both algorithmic genotyping approaches were empirically equivalent to manual genotyping, however comprehensive evaluation and adoption of algorithmic genotyping was not performed as part of this work.

Figure 3-2: Within groups sum of squares plotted by k-means cluster number for two samples (SID 1 and SID 5). The data included in these analyses excluded undetectable CNV assays (zero genotype). Both samples exhibited a clear “elbow” deflection in within groups sum of squares at a cluster number of 3, corresponding to recipient two-copy, recipient one-copy, and graft informative genotype clusters.



3.3.1.3 Recipient Genotyping

Recipient genotyping was performed using genomic DNA from 10 recipients. An example case is presented in Figure 3-3. All recipients were genotyped for 38 CNV assays, so the total number of genotyped results was 380. The genomic DNA genotype distribution was recipient zero-copy (48%), recipient one-copy (40%), and recipient two-copy (12%).

Corresponding recipient cfDNA marker genotyping was available for all but one assay (379/380, 99.7%). The missing data point arose because of a failed assay. The cfDNA genotype distribution was recipient zero-copy (49%), recipient one-copy (40%) and recipient two-copy (11%).

Analysis of recipient genotyping performance thus included 379 paired results. Of these, 195 (51%) were genotyped as recipient positive using recipient genomic DNA, and the remainder (49%) were genotyped as recipient negative. The sensitivity of cfDNA genotyping for recipient positive markers was 195/195 (100%), and the specificity was 184/184 (100%) (Table 3-8).

Recipient copy number genotyping performance was assessed using the 195 markers that were genotyped as recipient positive using recipient genomic DNA. Of these, 150 (77%) were genotyped as recipient one-copy, and the remainder (23%) were genotyped as recipient two-copy. CfDNA genotyping correctly identified 150/150 (100%) of recipient one-copy genotypes, and 43/45 (96%) of recipient two-copy genotypes (Table 3-9). There were two cfDNA assays that had been genotyped as recipient one-copy, but were recipient two-copy on recipient genomic DNA analysis. In these two instances, both assays were clearly part of a recipient one-copy assay cluster but both had high concentrations for the cluster.

Figure 3-3: CNV assay results from recipient genomic DNA (top panel) presented against cfDNA CNV assay results for the same recipient (bottom panel). The CNV assays are ordered by the cfDNA assay concentration. The comparison indicates that recipient two-copy and one-copy markers in genomic DNA cluster together in cfDNA, permitting recipient copy number genotyping from cfDNA. The comparison also demonstrates that 10 graft informative markers present at low concentrations in cfDNA are zero-copy in the recipient genome. In this instance, it is assumed that this cfDNA is being released from graft cells.

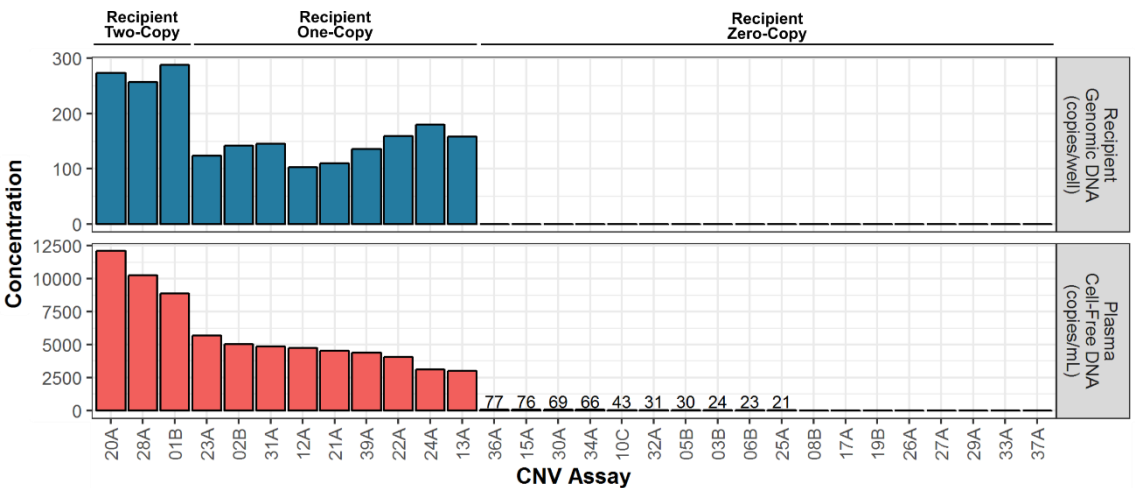


Table 3-8: Contingency table of transplant recipient cfDNA recipient genotyping performance compared against genotyping performed directly on recipient genomic DNA.

Recipient Cell-Free DNA		Recipient Genomic DNA		Total
		One- or Two- Copy	Zero-Copy	
	Recipient	195	0	195
	Undetectable or Graft	0	184	184
	Total	195	184	379

Table 3-9: Contingency table of transplant recipient cfDNA recipient copy number genotyping performance compared against genotyping performed directly on recipient genomic DNA. Only CNV assays that were detectable in recipient genomic DNA are presented.

		Recipient Genomic DNA		Total
		One-Copy	Two-Copy	
Recipient Cell-Free DNA	Recipient One-Copy	150	2	152
	Recipient Two-Copy	0	43	43
	Total	150	45	195

3.3.1.4 Donor Genotyping

Donor genotyping was performed using genomic DNA from 19 donors. Access to 20 donor samples had been approved, and one donor sample was not requested from the supplying laboratory in the event that this allocation was needed to resolve a complex genotyping problem, but such a scenario did not eventuate.

The median number of assays per donor was 31 (IQR 31-34). The total number of genotyped results was 650. Genomic DNA genotype distribution was donor zero-copy (43%), donor one-copy (42%), and donor two-copy (15%).

Corresponding recipient cfDNA marker genotyping was available for a median of 31 (IQR 29-31) markers per donor. The total number of assays for which donor and recipient cfDNA genotyping was available was 558 (86%). Incomplete coverage arose due to assay and panel upgrades leading to mismatched runs, incomplete sample to run the full assay panel, failed extraction, and other laboratory error. CfDNA genotypes were zero (25%), recipient (56%), and graft (19%).

The graft informative genotyping analysis excluded 313 CNV assays that were recipient positive, leaving a total of 245 CNV assay genotypes were available for comparison. Recipient positive markers were excluded as donor genotyping could not be discerned for recipient positive markers (Figure 3-4). Of the recipient negative markers, 131/245 (53%) were donor positive (donor one-copy or donor two-copy) and 114/245 (47%) were donor negative (donor zero-copy). The sensitivity of cfDNA genotyping for graft informative markers was 100/131 (76%) and the specificity was 108/114 (95%) (Table 3-10).

There were 31 CNV assays that were unable to detect target in recipient cfDNA, despite the corresponding CNV loci being present in donor genomic DNA. These samples tended to have lower graft-derived cfDNA concentrations than samples without false negative results (9 (IQR 7-13) versus 27 (IQR 18-44) copies/mL, $W = 14$, $p = 0.005$).

There were 6 CNV assays that detected target in recipient cfDNA, and were genotyped as graft informative markers, when such loci were undetectable in donor genomic DNA. In these cases,

the median number of positive droplets was 3 (IQR 2-4) and the median sample concentration was 11 (IQR 9-14) copies/mL. 5/6 had positive droplet amplitudes that were consistent with true positive droplet clusters, and 1/6 had a solitary positive droplet with an amplitude between a third and half the distance between the negative and positive cluster amplitudes.

Figure 3-4: CNV assay results from donor genomic DNA (top panel) presented against cfDNA (middle panel) and recipient genomic DNA (bottom panel) CNV assay results corresponding to the same recipient. The CNV assays are ordered by the cfDNA assay concentration. The comparison demonstrates that detectable CNV assays at low concentrations genotyped as graft-informative using cfDNA are indeed one-copy or two-copy in donor genomic DNA and zero-copy in recipient genomic DNA. The comparison also shows how recipient one-copy markers genotyped in cfDNA are a combination of recipient positive/donor negative markers (e.g. CNV assays 02B, 31A, and 12A), and recipient positive/ donor positive markers (e.g. CNV assays 23A, 21A, and 24A). In this example, recipient positive/donor positive markers did not distinctly cluster to permit discrimination from recipient positive/donor negative markers.

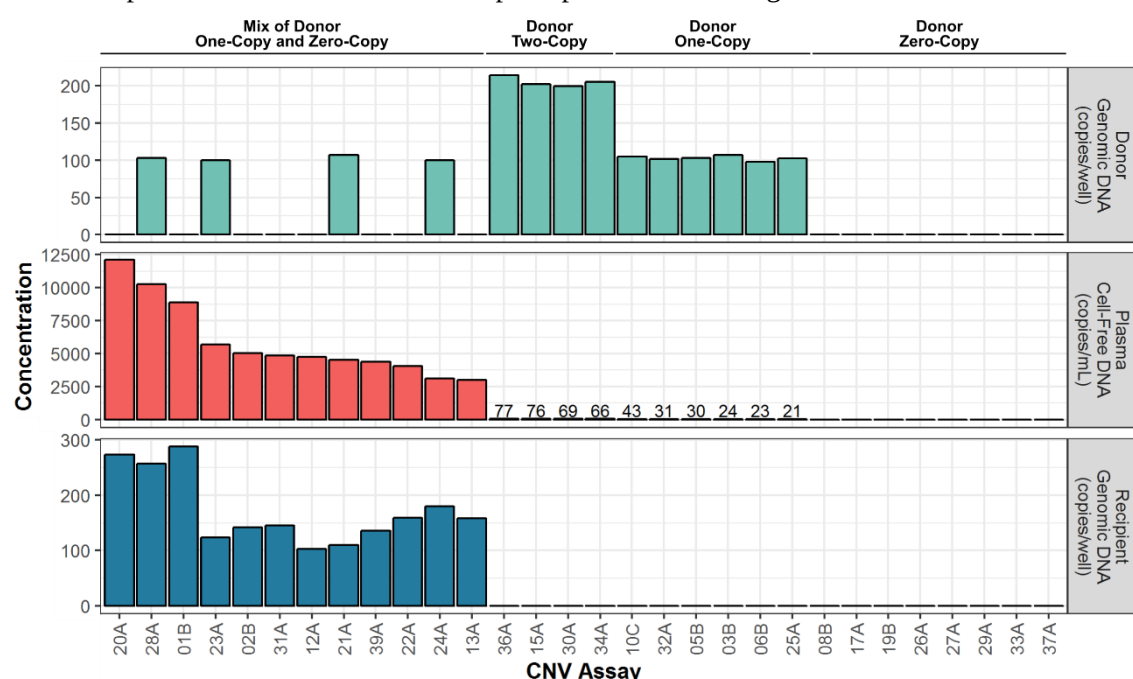


Table 3-10: Contingency table of transplant recipient cfDNA graft informative genotyping performance compared against genotyping performed directly on donor genomic DNA. Note that recipient positive cfDNA assays were excluded.

		Donor Genomic DNA		Total
		One- or Two- Copy	Zero-Copy	
Recipient Cell-Free DNA	Graft Informative	100	6	106
	Undetectable	31	108	139
	Total	131	114	245

3.3.2 Urine

3.3.2.1 *Cell-Free DNA Genotyping*

Urine cfDNA genotyping was performed for 31 CNV assays from each of 34 unique recipient samples. The total number of assay results considered was 1054.

The total cfDNA concentration was lower, and the graft derived cfDNA concentration was higher in urine compared with plasma samples (Figure 3-5). The median graft-derived cfDNA concentration in the urine samples was 150 (IQR 81-338) copies/mL, compared to 9 (IQR 7-15) copies/mL in the simultaneously collected plasma samples ($\chi^2(1, N = 62) = 43, p < 0.0001$). The total cfDNA concentration was 643 (IQR 253-3254) copies/mL in urine, compared with 2038 (IQR 1061-5386) copies/mL in plasma ($\chi^2(1, N = 62) = 7, p < 0.01$). Consequently, the median graft fraction was 27% (IQR 7-39%) in urine compared to 0.5% (IQR 0.2-0.9%) in plasma ($\chi^2(1, N = 62) = 39, p < 0.0001$).

As a result of these relative concentration differences, recipient one-copy markers could not be used as a surrogate measure of one-copy total cfDNA concentration, as they are in plasma. This was because recipient-derived cfDNA no longer comprised the vast majority of the total cfDNA concentration. Instead of this measure, half of the *ACE* non-polymorphic two-copy control assay concentration was used for the measurement of the total cfDNA concentration in urine samples.

The higher graft fraction had a number of implications for genotyping. Firstly, donor copy number could be defined readily for graft informative markers (Figure 3-6). Secondly, it was very difficult to distinguish recipient positive markers from recipient negative/donor positive markers with certainty. Therefore, for the majority of urine samples, genotyping required reference to plasma cfDNA genotyping of recipient positive markers. Finally, summation of recipient positive/donor positive CNV markers had to be accounted for, as this significantly affected CNV assay concentration. The result of these observations was that the number of allocated distinct genotypes for urine samples was greater (Figure 3-7).

The distribution of graft-informative genotypes was recipient zero-copy/graft one-copy in 178/1054 (R0G1, 17%), recipient zero-copy/graft two-copy in 57/1054 (R0G2, 5%), and recipient zero-copy/graft copy number uncertain but detectable in 34/1054 (R0G?, 3%). The total proportion of graft informative CNV markers was therefore 25%, with a median of 8 (IQR 5-10) graft informative markers per sample. In comparison, the number of detected graft informative markers in concurrently collected plasma samples was 5 (IQR 4-8).

The distribution of recipient positive genotypes was recipient one-copy in 443/1054 (R1G-, 42%) and recipient two-copy in 114/1054 (R2G-, 11%). The total proportion of recipient positive CNV markers was therefore 53%. Zero genotype markers (R0G0) accounted for 206/1054 (20%) and

the remaining 2% of markers could not have their genotypes resolved in absence of donor and recipient genomic DNA genotype results.

Figure 3-5: Graft-derived cfDNA, total cfDNA, and graft fraction for plasma and urine samples collected simultaneously from 34 kidney transplant recipients. All results are presented on a logarithmic axis.

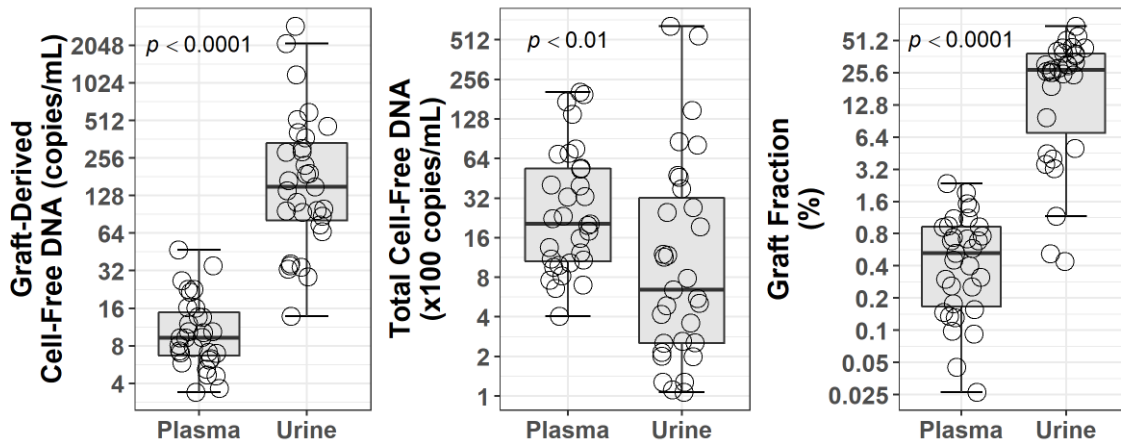


Figure 3-6: CNV assay results from urine cfDNA (fourth panel) presented against recipient genomic DNA (first panel), plasma cfDNA (second panel), and donor genomic DNA (third panel) CNV assay results corresponding to the same recipient. The CNV assays are ordered by the plasma cfDNA assay concentration.

Urine cfDNA CNV assays were more representative of the donor genotype compared to plasma cfDNA CNV assays. Consequently, graft copy number genotypes could be readily discerned. In this example, one- and two-copy CNV loci apparent in donor genomic DNA are readily identified in urine cfDNA, where they are more difficult to discern in plasma cfDNA, particularly when a recipient positive/donor positive combination exists. Summation of recipient positive/donor positive markers was also more apparent in urine cfDNA. In this example, CNV assays 28A, 23A and 24A have a higher urine concentration than copy number equivalent recipient positive/donor negative markers.

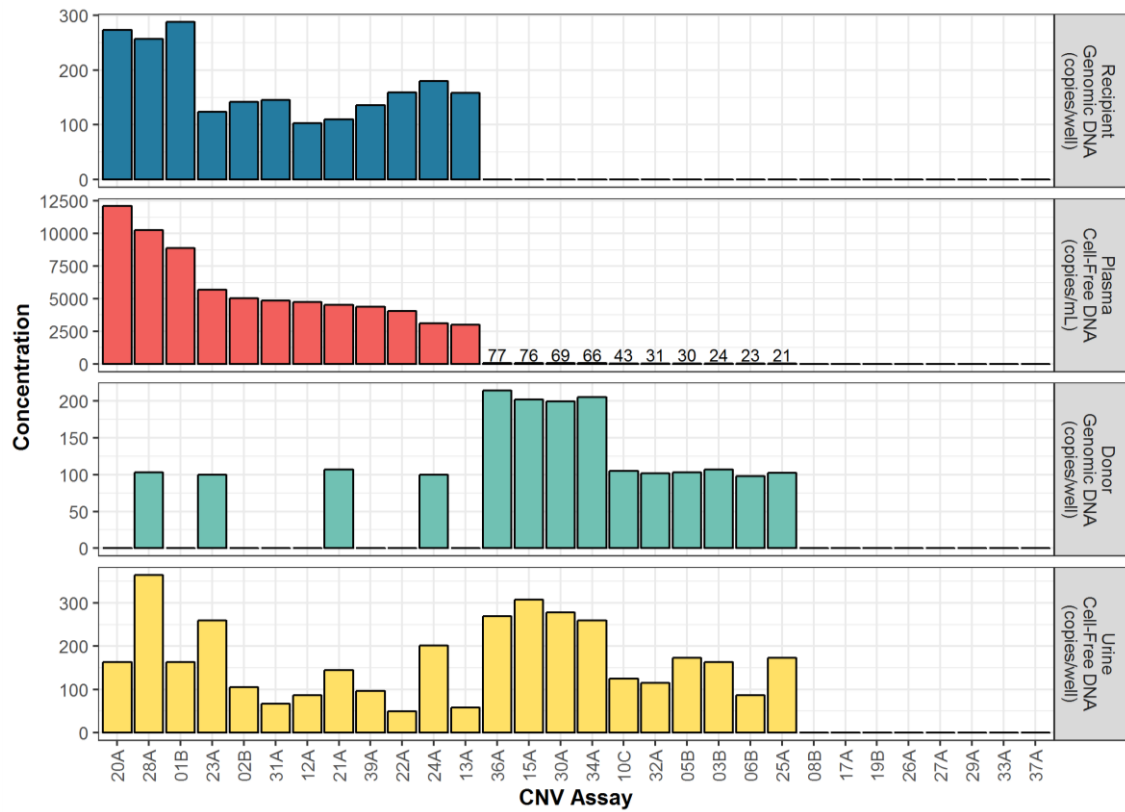
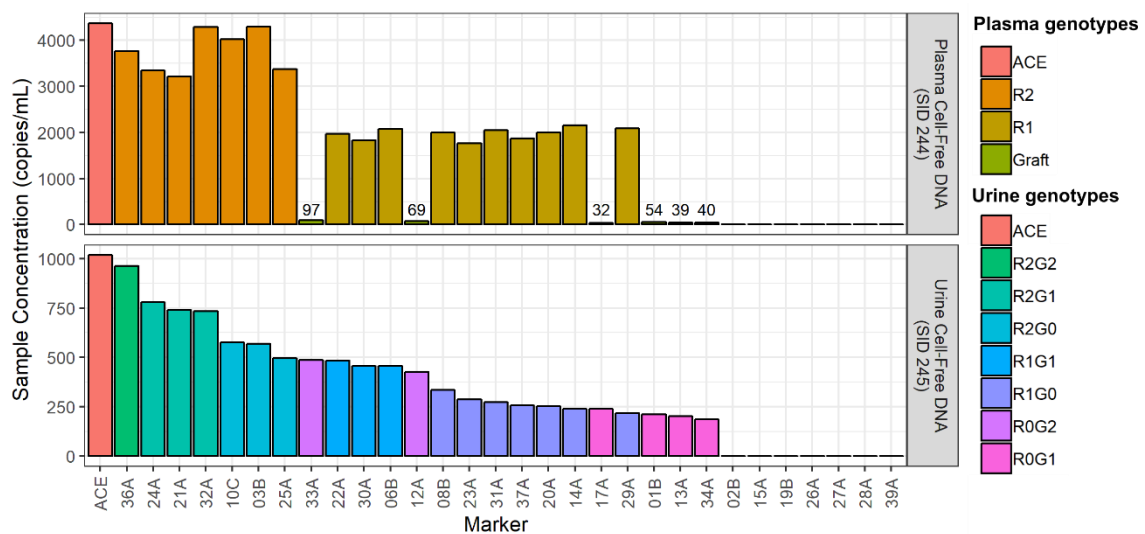


Figure 3-7: Plasma (top panel) and urine (bottom panel) cfDNA CNV assay panel results from the same recipient demonstrating the spectrum of allocated genotypes for both samples. *ACE* non-polymorphic two-copy control assay results are also presented for clarity (red bar, first assay on left). CNV assays are ordered by urine cfDNA concentration. Genotyping of urine CNV assays required consideration of more genotypes given the greater graft fraction and larger summation effect of recipient and donor copy number genotypes.



3.4 DISCUSSION

3.4.1 Blood

This chapter presents a conceptual validation of CNV assay genotyping for chimeric plasma cfDNA, and introduces a methodological basis for the identification and quantification of graft-derived and total cfDNA, as well as the concept of graft fraction. Using non-chimeric recipient and donor genomic DNA as a gold standard, identification of simplified genotypes relevant for the discrimination and quantification of these metrics in plasma was achieved with satisfactory accuracy.

3.4.1.1 *Genotyping*

Genotyping was performed by recognition and labelling of a characteristic pattern of CNV assay clusters that reflect (i) the allelic frequencies of the CNV loci underpinning the assay panel, (ii) the predictable factorial effect of copy number on assay concentration, and (iii) the summation of the relative contribution of donor and recipient genotype to the chimeric CNV assay result. Notably, underpinning the genotyping method is the fact that graft-derived cfDNA comprises a small proportion of total cfDNA in the plasma of kidney transplant recipients. This low graft fraction is manifest by relatively higher concentrations of recipient positive markers in plasma, and very low concentrations of graft informative (recipient negative/donor positive) markers.

This chimeric genotyping method is the simplest approach described in the literature, and reflects the value of biallelic copy number deletion in creating a negative background against which to identify graft informative markers. Two other indel polymorphism-based chimeric cfDNA assays require *a priori* genotyping of donor and recipient, although this requirement may be avoided by following an approach similar to that presented.[147, 163] More complex bioinformatic approaches have been developed for SNP-based sequencing chimerism assessment.[152, 164] In these implementations, allelic frequency models and assumptions based upon low graft fraction also form the basis for the genotyping bioinformatic pipeline.

The observed genotyping results matched the allelic frequencies predicted *in silico*. In donor and recipient genomic DNA, averaged observed allelic frequencies were zero-copy in 46%, one-copy in 41%, and two-copy in 13%. *In silico* predicted allelic frequencies were zero-copy 52%, one-copy 41%, and two-copy 8%. Plasma and urine cfDNA genotype combinations also matched those predicted on the basis of *in silico* individual probabilities. The averaged observed frequency of the recipient positive genotype was 54% (predicted 49%), graft-informative genotype was 22% (predicted 25%), and zero-copy genotype was 23% (predicted 27%).

Sensitivity for the detection of graft informative markers in plasma cfDNA was not as high as it was for recipient positive markers. This tended to occur when the plasma graft-derived cfDNA

concentration was lower, suggesting that CNV assay “drop-out” (false negative results) occurs when target sequence is below the limit of detection for the assay. This is supported by observation that urine cfDNA genotyping, with a higher graft fraction, identified more informative markers than concurrently collected plasma samples.

3.4.1.2 Graft-Derived Cell-Free DNA Measurement

The low concentration of graft informative markers and variation between these markers also contributed to difficulty copy number genotyping graft informative markers. To account of this, the calculation of graft-derived cfDNA concentration was based upon the median of all graft informative CNV assay concentrations. Given the allelic frequencies of the CNV loci, this effectively ensures that the median concentration represents a one-copy graft informative CNV assay. This probability nears certainty if more than two graft informative markers are present, however it is acknowledged that there is a small probability ($\approx 10\%$) of confounding if the number of informative markers is two or less. This scenario appears to be most common when the graft derived cfDNA concentration is low, and so the absolute impact on the graft derived cfDNA concentration of using a two-copy graft-informative marker is likely to be small.

The possibility that two-copy graft informative assays may be included in the determination of the median graft-derived cfDNA concentration also raises the problem that this measure of central tendency may be shifted upwards by two-copy results. In the cases where numerous informative markers are present, the probability that this shift will be materially significant is low given the allelic predominance of the one-copy genotype. Additionally, the effect is likely to be less pronounced with the median compared to the mean.

3.4.1.3 Total Cell-Free DNA Measurement

In contrast, the higher concentration of recipient positive markers meant that copy number genotyping for these marker clusters was easily achieved. The median of the recipient one-copy markers was used as a measure of total cfDNA. Given that most genetic sequences not subject to CNV are present in two-copy, and given that the allelic frequencies of the assay panel maximise the likelihood that graft informative markers are present in one-copy, a one-copy measure of total cfDNA was used for the calculation of the graft fraction. This was sought to ensure comparability with other methods which may assess chimerism in the absence of allelic deletion.

Notably, the plasma genotyping methodology proposed in this work does not specifically seek to identify CNV markers that are recipient positive/donor positive. Recipient one-copy/donor one-copy combinations theoretically provide the most complete one-copy measure of total cfDNA. The probability of this combination is 17% whereas the probability of a recipient one-copy/donor zero-copy combination is 21%, and the probability of a recipient one-copy/donor two-copy

combination is 3%. Consequently, just under half of recipient one-copy markers provide an ideal assessment of total cfDNA. Using the median recipient one-copy concentration raises the possibility that recipient one-copy/donor zero-copy markers may skew downwards the total cfDNA estimate if not accounted for. This does not appear to be required, nor achievable, given that the low graft fraction is unlikely to meaningfully change the magnitude of the total-cell free DNA estimate, and observed variation between markers genotyped as recipient one-copy prevent the discrimination of recipient positive/donor positive markers from recipient positive/donor negative markers.

An alternative approach to estimation of one-copy total cfDNA would have been use of half the *ACE* two-copy non-polymorphic control assay concentration. Other groups use the human beta-globin gene (*HBB*) to this end [130, 132, 163]. *ACE* was not used in the first instance given that (i) the option of not running the *ACE* assay for each sample needed to be maintained in event that the methodology was simplified, (ii) complete application of the CNV assay panel to each sample ensured multiple independent measures of recipient one-copy markers as compared to only one measure via the *ACE* assay, and (iii) a single *ACE* assay result is more prone to summative bias as a solitary recipient positive/donor positive marker than multiple CNV assays.

3.4.2 Urine

An exploratory study of chimeric urine cfDNA genotyping was also presented. This work identified that urine cfDNA consists of a higher proportion of graft-derived cfDNA than plasma cfDNA. This finding is consistent with other reports.[130, 140]

The higher graft fraction and concentration of graft-derived cfDNA provided an opportunity to copy number genotype the allograft, which was difficult using plasma cfDNA. Whilst the argument has been made that plasma graft-derived cfDNA does not need copy number genotyping due to the allelic frequencies of the CNV loci and the method employed for calculation of graft-derived cfDNA concentration, urine cfDNA genotyping provides an opportunity to resolve complex genotyping scenarios without requiring access to donor genomic DNA. Such scenarios might include closely related donor-recipient pairs, in whom the number of discriminative graft-informative markers may be reduced, and for whom graft informative copy number genotyping may avoid confounding by two-copy CNV loci.

The higher graft fraction in urine cfDNA led to challenges identifying recipient positive markers, given that in some instances recipient negative/donor positive markers exhibited concentrations equivalent to or exceeding recipient positive/donor negative markers. Consequently, genotyping of urine cfDNA assays was dependent on reference to plasma cfDNA genotypes. As a better reflection of the recipient genotype, plasma cfDNA genotyping worked complementarily with urine cfDNA genotyping to resolve donor and recipient copy number genotypes. In this manner,

genotyping using genomic DNA extracted from recipient and donor buffy coat leucocytes may be avoided altogether, in the uncommon scenarios where this may be required. Such an approach has recently been reported using a SNP-based chimerism assay in which urine genotyping of the graft has supplanted genotyping using donor genomic DNA.[120]

Further validation of recipient positive/donor positive copy number genotyping is required. Unfortunately, the study design did not specifically overlap simultaneously collected urine and blood samples with donor genomic DNA samples for the same recipients. Graft copy number genotyping using urine cfDNA should be assessed against the gold standard of donor genomic DNA. This is particularly the case for recipient positive/donor positive combinations. Validation for recipient copy number genotyping is not required, as there is sufficient evidence from the presented work that plasma cfDNA recipient copy number genotyping is equivalent to that performed using recipient genomic DNA, and recipient genotyping in urine cfDNA was defined by plasma cfDNA.

Assessment of total cfDNA concentration using CNV assays was also complicated by the higher graft fraction. Unlike plasma cfDNA, in which the graft fraction was typically <1% of total cfDNA, the median urinary cfDNA graft fraction in the presented cohort was 27%. Consequently, neither recipient positive markers nor donor positive markers alone could be used as a measure of total cfDNA. The presented work used half the *ACE* two-copy non-polymorphic control assay concentration as a measure of one-copy total cfDNA concentration. An alternative approach using CNV assays alone could include use of the recipient one-copy/donor one-copy (R1D1) genotype assay concentration and/or half the recipient two-copy/donor two-copy (R2D2) genotype assay concentration. These should provide a similar measure of total cfDNA to half the *ACE* concentration. Figure 3-7 exemplifies this point: CNV 36A (R2D2) approximates the *ACE* concentration, and CNVs 22A, 30A and 06B approximate half the *ACE* and CNV 36A concentrations. A limitation of this approach is the allelic probability of these combinations, which are 17% for R1D1 and 0.6% for R2D2. Using a 31 CNV assay panel, this equates to a high probability of five R1D1 CNV assays but only a 20% chance of one R2D2 assay. Given the strengths of multiple independent CNV assay measurements of total cfDNA concentration, consideration should be given to adopting an R1D1 approach to urine total cfDNA measurement in future work, once graft copy number genotyping validation work has been completed.

The mechanisms underpinning the origin of urinary cfDNA remain under investigation, but it is hypothesised that cfDNA enters the urine via glomerular filtration, tubular active secretion, and apoptosis and necrosis of cells lining the urinary tract.[19, 137, 266, 267] Given that the graft is responsible for only a small proportion of plasma cfDNA but appears to be responsible for a much larger proportion of urine cfDNA, urine graft-derived cfDNA must arise predominantly from the

tubular cells rather than glomerular filtration alone. Pathologies affecting the allograft may modify the mechanisms underpinning urinary cfDNA production, and consequently graft-derived cfDNA, total cfDNA, or graft fraction may have biomarker potential separate or complementary to plasma cfDNA.

The urine is a convenient sample medium for biomarker development that is accessed non-invasively and is typically relatively abundant. Despite this, the study of urine cfDNA has a number of challenges.

Urine storage and processing requirements are more sensitive and laborious than plasma. Whilst it is acknowledged that there are a number of commercial products which promise to preserve urine cfDNA at room temperature, the sample processing methodology used for this work required immediate refrigeration and cell separation to avoid contamination and degradation of cfDNA. This method was chosen as it permitted cfDNA extraction using the same commercial product used for plasma cfDNA extraction. Additionally, extraction methodologies for urinary cfDNA have been less rigorously evaluated.

With exception of the graft fraction, absolute measures of graft-derived and total cfDNA are likely to be affected by variation in urine concentration arising from homeostatic activities. This is evident in the large variation in graft-derived and total cfDNA concentration observed in the urine samples of this cohort. Unless the magnitude of variation arising from pathology is so great that urine concentration does not significantly affect the discrimination of pathology from non-pathology, correction of absolute measures for urine concentration is likely to be required. This may be achieved by concurrent measurement of urine creatinine, recognising the inter-individual variation in creatinine arising from variable body habitus, muscle mass, and kidney function.[137] An alternative approach could involve study of graft fraction alone, which as a relative measure, would be less affected by variation in urine concentration.

Given the scope of this thesis, and the challenges and limitations arising from the study of urine cfDNA as a diagnostic biomarker, subsequent chapters will focus only on plasma cfDNA. The additional utility of urine cfDNA is an area for future work.

3.4.3 **Algorithmic Genotyping**

Proof of principle algorithmic genotyping has been demonstrated using simple mathematical methods. These approaches are based upon simple cluster discrimination techniques, and scope exists to improve this with more sophisticated methods that also incorporate quality assurance.

Both approaches employed in this work have the limitation that cluster number needs to be pre-specified, or that further algorithmic assessment of within groups sum of squares is required. Whilst this may be suitable for research applications, a more sophisticated approach should be

sought in event of diagnostic translation. Optimised cluster number genotyping would permit discrimination of two-copy graft informative CNV assays when such assays are evident, and allow automated genotyping of urine cfDNA results, which have a greater number of possible genotype combinations, some of which may not be present in all samples.

Given the early investigational phase of the assay panel, researcher review of all assay results as part of manual genotyping was felt to be an important quality assurance step. High throughput automated genotyping as part of diagnostic application should include quality assurance assessments that consider allelic frequency of the genotypes, inter-assay variation within genotype clusters, outlier results, and consistency with respect to non-polymorphic control assays if present.

More sophisticated automated genotyping would also permit the collection of genotyping data longitudinally over serial samples from the same recipient. Such an approach could use samples with higher graft fraction/graft-derived cfDNA concentration to identify the full complement of graft-informative CNV assays for a particular recipient-donor transplant arrangement. These scenarios are likely to occur early in the post-transplant period, when ischaemia reperfusion injury may be associated with higher levels of graft-derived cfDNA, and during periods of rejection.[120, 159] Having established a history of consistent genotyping, CNV assay false negative or false positive results could be identified and corrected for. This could be particularly relevant if the graft-derived cfDNA concentration falls below the limit of detection of the CNV assay panel, or for the identification of assay contamination. This genotype record could be supplemented by additional, more definitive, genotyping results, such as those from urine cfDNA or genomic DNA.

These considerations promise to not only improve the efficiency, turnaround time, and cost of CNV assay panel testing, but also ensure improved quality assurance and genotyping performance of serial sampling in a high throughput diagnostic environment. Consequently, the identification of clinically relevant recipient-donor genotype combinations is not a barrier to translation of this methodology to clinical practice.

3.4.4 **Summary**

This work demonstrated that the identification of four simplified donor-recipient genotype combinations based upon CNV assay concentration clustering can be performed accurately, and permitted measurement of plasma graft-derived cfDNA, total cfDNA and graft fraction from chimeric post-kidney transplant blood samples without *a priori* genotyping of recipient and donor.

Genotyping performance appears to be affected by low levels of graft-derived cfDNA, in which graft informative markers “drop out”, presumably because the limit of detection is exceeded at these low concentrations. The effect of this phenomenon on graft-derived cfDNA results is likely to be low, but could be avoided by if graft informative assays are established within the immediate post-transplant period when graft-derived cfDNA concentrations are higher. This work also identified some inter-assay variation within each genotype CNV assay cluster that requires further evaluation in a controlled manner.

Whilst plasma cfDNA exhibited a very low graft fraction, urine cfDNA was found to have a higher graft fraction, which permitted better identification and copy number genotyping of donor positive markers. This observation suggests complementary roles for plasma and urine cfDNA as surrogate means to establish recipient and donor genotypes without needing to genotype either individual via a non-chimeric compartment. Whilst such an approach was not required in routine sample genotyping, it may be helpful in more complex genotyping settings.

CHAPTER 4 ANALYTICAL VALIDATION

4.1 INTRODUCTION

The preceding chapters detail (i) the selection and development of 41 PCR assays targeting highly heterozygous CNV loci, and (ii) a conceptual validation demonstrating that these CNV assays can be used, based upon specific CNV genotype combinations in the recipient and donor(s), to perform absolute quantification of graft-derived and total-cell free DNA in chimeric post-transplant plasma and urine samples. The fundamental hypothesis underlying the application of this technique to chimeric cfDNA is that immunologic injury to an allograft may be detectable via increases in graft-derived cfDNA. To test this, the assay panel and ddPCR quantification platform must be able to measure this analyte with appropriate robustness.

Droplet digital PCR is the gold standard method for copy number quantification.[45, 46, 268-272] The targeting of sequences specific to and sited within CNV loci permits quantification of CNV abundance, the result being factorial concentration increments consistent with copy number.[45] Quantitative real-time PCR can also quantify CNV, albeit against a relative standard and with potential confounding arising from machine, laboratory, PCR assay, template, and reporter dye variability.[273] Next-generation sequencing is also inferior to ddPCR for CNV analysis. It is limited by relative quantification, which arises in part from the need to pre-amplify target DNA before sequencing can be performed. Additionally, noise, bias and coverage limitations confound sequencing-based assessment, which are incompletely addressed with newer bioinformatic approaches.[274-277] Given that multiple CNV assays form part of a discriminative panel, an assay-agnostic gold standard quantification platform (ddPCR) was favoured for this work.[53, 278]

Preliminary work, including the genotyping studies presented in the previous chapter, has demonstrated that plasma graft-derived cfDNA concentrations are typically low, ranging from undetectable to 75 copies/mL. This is similar to the findings of Adamek et al., the only other group to quantify graft-derived cfDNA in kidney transplantation in absolute terms.[163] Using 12 indel markers and qrtPCR, 4 kidney transplant recipients had serum graft-derived cfDNA concentrations of between 0.48-47 GE/mL in the first 3 days after transplantation. At these concentrations, the probability that more than one target sequence will be partitioned into any particular ddPCR droplet is low. Consequently, each positive droplet reflects one copy of target sequence, and discrimination between different concentrations is limited by the fractional abundance of positive droplets (“resolution”). Such small variation is unlikely to be detectable using quantitative real-time PCR.[51, 54]

No absolute graft-derived cfDNA quantification method employing ddPCR has undergone analytical validation to date. Analytical validation studies of methods that employ pre-amplification before quantification, as well as those using sequencing or qrtPCR-based methodologies, are of limited utility in informing the design of this absolute quantification method.

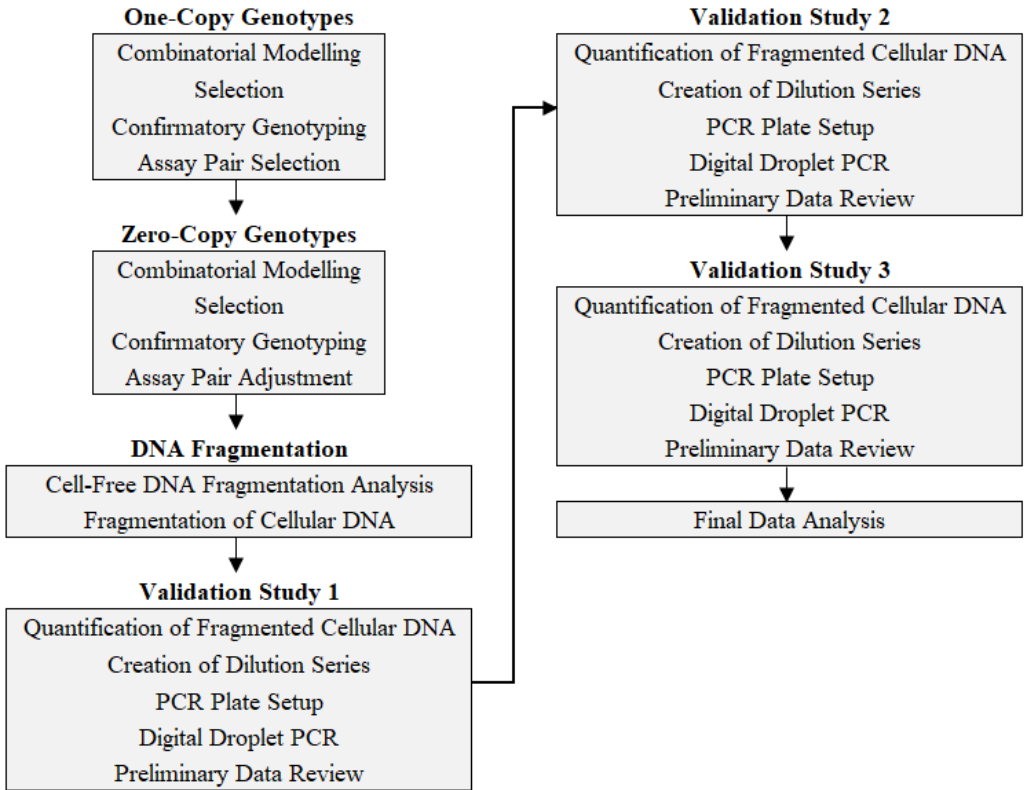
The majority of the work conducted on the analytical performance of ddPCR in quantitation of circulating DNA relates to rare event detection in oncology, in which small amounts of circulating tumour DNA with a targeted oncogenic mutation must be distinguished from a background of abundant non-tumour wild-type DNA sequence.[279] Likewise, applications of ddPCR to CNV analysis in circulating DNA have predominantly focussed on circulating tumour DNA, where CNV plays a role in oncogenesis and subtle-fold copy number discrimination needs to occur due to the background of the wild-type copy number.[45, 279, 280] In both of these instances, ddPCR performs very well, although the results of these studies are not generalizable to the CNV-based cfDNA chimerism approach proposed. Unlike these studies in oncology, which may use a relative approach to quantification (e.g. tumour/wild-type), the targeted sequence in the current chimerism method occurs on a negative background strategically created by the combination of a donor positive genotype on a recipient negative background. Not only does this theoretically improve the signal-to-noise ratio in the quantification methodology, it permits absolute quantification of the target expressed as a concentration per volume of sample.[103]

The aim of this study is to assess the analytical validity of the CNV assay panel and *ACE* for chimeric cfDNA analysis, with the objective of defining sensitivity, limit of detection, limit of blank, intra-assay variation, inter-assay variation, and the limit of quantification. This assessment will shed light on two assay performance issues identified during the prior genotyping study. The first was slight variation noted in the cfDNA concentrations determined when assaying a single genotype with two or more different CNV sequences. The second was the occurrence of apparent false negative results at low target concentrations, suggesting that the limit of detection for the assay may have been reached.

4.2 METHODS

The analytical validation study comprised a number of experimental steps, which are overviewed in Figure 4-1.

Figure 4-1: Schematic representation of the analytical validation experiments, demonstrating how each validation study group was studied in series.



4.2.1 Study Participants

DNA (cfDNA and genomic DNA) used for this analytical validation work was obtained from blood samples submitted by adult participants of our cfDNA solid organ transplantation assay development study. Institutional ethics approval was obtained for this study, and all participants gave written informed consent.

Each participant is identified in this work using a participant identification number (PID, P prefix to a number), which is common to all samples submitted by the participant in the course of the study, each of which was identified by a separate sample identification number (SID, S prefix to a number).

The term “participant genotype” is used to refer to the unique combination of individual assay genotypes present within that individual’s genomic DNA. By definition, this genotype should be consistent across samples.

4.2.2 Sample Collection, Processing and DNA Extraction

The methods employed for blood sample collection, processing, and DNA extraction are presented in Chapter 3.

4.2.3 **One-Copy Genotypes**

4.2.3.1 *Combinatorial Modelling*

A combinatorial algorithm was written in R (R Foundation for Statistical Computing; <https://www.R-project.org>) to model combinations of participant genotypes with the aim of determining genotype combinations that permitted (i) study of the maximal number of CNV assays *in toto*, (ii) study of the maximal number of CNV assays per individual, facilitating measurement of inter-assay variation amongst the largest possible groups of assays, and (iii) achievement of the prior criteria in the fewest number of experiments.[281]

Marker genotypes from 25 individuals that had undergone CNV assay testing using plasma cfDNA were considered by the algorithm. Recipient one-copy assays were identified for each individual. These were common (by definition, based on allelic frequency) positive assays that expressed a concentration approximately half of the *ACE* concentration. Markers that were recipient two-copy, informative for the graft, or recipient zero-copy, were ignored. Group sizes of 1-6 individuals in combination were considered.

All possible combinations of the 25 individuals within the permitted group size were obtained by the algorithm. The genotypes of each of the individuals within each combination was then considered. The total number of unique CNV markers for which at least one individual in the combination was recipient one-copy was counted. Markers where more than one individual within a combination was recipient one-copy were not treated differently from markers in which only one individual within a combination was recipient one-copy.

Combinations of the smallest group size which permitted coverage of the largest number of CNV assays were then selected for further review.

4.2.3.2 *Genotype Selection*

Genotype combinations with maximal one-copy genotype coverage identified by combinatorial modelling were assessed against a number of criteria. The number of unique CNV assays covered per individual was maximised. Group size was minimised. Sufficient cellular DNA was confirmed to be available for all individuals within the combination to perform all experiments, including (i) confirmatory genotyping, (ii) pre-fragmentation quantification, (iii) fragmentation optimisation, (iv) ddPCR quantification post-fragmentation, (v) creation of dilution series, (vi) repeat experiments if necessary, and (vii) surplus for experimental error.

Digital PCR plate capacity and experimental plate layout options were considered. An optimal multiplex ratio of CNV assays with FAM-probes versus CNV assays with HEX-probes within each individual of the combination was sought. Preference was given to combinations in which

each individual within the group had a CNV assay FAM:HEX ratio equal to or slightly below 50%, reflecting preference for 1:1 matched FAM:HEX assays, or at most one un-matched CNV HEX-probe assay. This enabled use of the *ACE* FAM-labelled non-polymorphic control assay in a duplexed experimental design. Combinations containing individuals with one un-matched CNV FAM-probe assay were avoided as this would result in inefficiencies arising from two unmatched FAM simplex wells per assay run (the unmatched CNV FAM-probe assay and *ACE*).

4.2.3.3 *Confirmatory Genotyping*

Confirmatory genotyping was performed using cellular DNA extracted from stored buffy coat leucocytes belonging to each of the individuals in the selected high coverage combination. Droplet digital PCR was performed for all 38 CNV assays in multiplex. The method for the ddPCR is presented in Chapter 3.

Undetectable assays were genotyped as zero-copy. The bulk of detectable assays, which were approximately half the concentration of the highest concentration markers, were genotyped as one-copy. The highest concentration markers were genotyped as two-copy. The marker genotypes determined from cellular DNA were compared against marker genotypes determined from cfDNA prior to the combinatorial modelling. One-copy markers destined for use in the analytical validation experiments were confirmed to be one-copy markers.

For each sample of genomic DNA to undergo confirmatory genotyping, the median one-copy marker concentration was calculated. The difference between each marker concentration and the one-copy median was determined. This degree of variation was expressed as a percentage of the median one-copy marker concentration for the respective individual. Additionally, some markers were present in a one-copy genotype in more than one individual. The mean variation of these duplicated assays from the one-copy median of their respective samples was also calculated. These measures of variation were considered during assay pair selection.

4.2.3.4 *Duplexed Assay Pair Selection*

One-copy CNV loci present in multiple individuals within each combination were identified. Removal of duplicated one-copy CNV loci was performed based upon (i) optimisation of the CNV assay FAM:HEX ratio for that particular individual as described in section 4.3.1.2, and (ii) consideration and minimisation of variation of the respective markers from the recipient one-copy median value for the individual as calculated in section 4.3.1.3.

After distribution of the CNV assays amongst one-copy genotypes, FAM-probe and HEX-probe assay pairings were selected. Key considerations for assay pairing were (i) pairing of assays in a way that minimised the number of individual zero-copy genotypes required to act as negative controls, (ii) existence of both FAM and HEX probes for a few assays, and (iii) pairing of assays

by degree of polar variation from the one-copy median. With respect to this latter consideration, FAM-probe assays that tended to under-estimate the median one-copy concentration on confirmatory genotyping were paired against HEX-probe assays that tended to over-estimate the median one-copy genotype. This was done to assist with the troubleshooting of aberrant results, if present.

4.2.4 **Zero-Copy Genotypes**

4.2.4.1 *Combinatorial Modelling and Genotype Selection*

A combinatorial algorithm was written in R (R Foundation for Statistical Computing; <https://www.R-project.org>) to identify participants with zero-copy genotypes that could efficiently act as zero-copy controls for the one-copy assay pairings identified in section 4.2.3.4.[281]

Marker genotypes from 25 participants that had undergone CNV assay testing using plasma cfDNA were considered by the algorithm. Recipient zero-copy markers or graft-informative markers were identified for each participant. Markers that were recipient two-copy or recipient one-copy were ignored.

Participants that were zero-copy for both assays of each duplexed one-copy assay pair were identified. All possible combinations of these participants within a group size of 3 or less were obtained. The assay pair zero-copy coverage of these participant combinations was then determined for each one-copy participant genotype. In the event that not all assay pairs defined for a particular one-copy participant genotype could be covered by 3 or fewer zero-copy participant genotypes, FAM-HEX assay pairings for the one-copy participant were adjusted. The zero-copy participant combinations that afforded the broadest zero-copy coverage at the smallest group size were then selected for manual review and confirmatory genotyping.

The zero-copy participant combination associated with maximal coverage of assay pairs for each one-copy participant, at the smallest group size, was selected for use. Additional considerations included (i) selection of zero-copy participant genotypes that could act as a negative control for multiple one-copy participant genotypes, and (ii) confirmation that sufficient cellular DNA was available to complete planned experiments.

4.2.4.2 *Confirmatory Genotyping*

Confirmatory genotyping was performed using cellular DNA extracted from buffy coat leucocytes of selected zero-copy control participants. Droplet digital PCR was performed for all negative control CNV assay pairs in duplex. The method for the ddPCR is presented in Chapter 3. One-copy positive control markers were also selected for each zero-copy participant.

Undetectable markers were genotyped as zero-copy. Detectable markers were genotyped according to their cfDNA genotype, in this instance one-copy, and acted as positive controls. Zero-copy markers destined for use as negative controls in analytical validation experiments were confirmed to have a zero-copy genotype.

4.2.5 **DNA Fragmentation**

4.2.5.1 *Fragmentation Profile of Cell-Free DNA Samples*

DNA electrophoresis using the Agilent 2200 TapeStation (Agilent Technologies, Santa Clara, CA, USA) was performed on cfDNA extracted from processed plasma samples as described in Chapter 1. Three samples with amongst the highest ACE concentrations studied to date were chosen to ensure maximal cfDNA concentration for fragment characterisation.

A positive control and size ladder was created by adding 2 μ L High Sensitivity D1000 Ladder (Agilent Technologies, Santa Clara, CA, USA) to 2 μ L High Sensitivity D1000 Sample Buffer (Agilent Technologies, Santa Clara, CA, USA). Sample preparation consisted of adding 2 μ L cfDNA sample to 2 μ L High Sensitivity D1000 Sample Buffer. Electrophoresis was performed using the Agilent 2200 TapeStation and High Sensitivity D1000 ScreenTape (Agilent Technologies, Santa Clara, CA, USA) according to manufacturer instructions.

Analysis was performed using the TapeStation Analysis Software, versions A.01.05 (SR1) and A.02.01 (SR1) (Agilent Technologies, Santa Clara, CA, USA). Peaks within the sizing range of 25-1,500 bp were identified and characterised.

4.2.5.2 *Fragmentation of Cellular DNA*

To mimic the fragmented nature of cfDNA, cellular DNA from one-copy and zero-copy control individuals underwent sonic fragmentation prior to use in analytical validation experiments.

Cellular DNA samples first underwent quantification using the NanoDrop 2000 spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA) according to manufacturer instructions using nucleic acid, double-stranded DNA defaults. A working solution of approximately 25 ng/ μ L or 50 ng/ μ L cellular DNA was made prior to fragmentation using distilled water. The targeted dilution concentration was selected on basis of the number of assay pairs per one-copy or zero-copy genotype and the quantity of DNA available. Working solution concentrations were confirmed prior to fragmentation using the NanoDrop 2000 spectrophotometer.

DNA fragmentation was performed using the Covaris S220 focused-ultrasonicator (Covaris, Woburn, MA, USA) with reference to manufacturer instructions for DNA shearing. Default

parameters were adjusted through a process of trial and error to obtain fragmented DNA with a peak size of approximately 150-160 bp consistently between samples (Table 4-1).

Achieved fragmentation characteristics were confirmed using the Agilent 2200 TapeStation and High Sensitivity D1000 ScreenTape as described in section 4.2.5. For this confirmatory DNA electrophoresis, fragmented DNA samples were diluted to 1-3 ng/ μ L to ensure concentrations were low enough to adequately identify the High Sensitivity D1000 Sample Buffer lower and upper limit markers.

Table 4-1: Covaris S220 parameters for fragmentation of genomic DNA targeting a peak size of ~150-160 bp.

Parameter	Value
Volume	130 μ L
Duty factor	10%
Peak incident power	175 watts
Cycles per burst	200
Treatment time	820-860 seconds
Temperature	8 °C
Sample volume	130 μ L

4.2.6 Experimental Design

4.2.6.1 *Assay Modelling*

4.2.6.1.1 Model Development

A mathematical model was created to better understand the relationships between (i) modifiable clinical and experimental variables, (ii) ddPCR dynamic range, (iii) droplet number, and (iv) assay target concentration in the PCR well. The objective was to use this model to achieve a practical balance between optimising platform and assay performance with clinical considerations such as cfDNA concentration and sample volume. The model incorporated methodology-specific calculations, as well as published equations used by the Bio-Rad QuantaSoft software to convert negative and total droplet number into a concentration.[253]

4.2.6.1.2 Standard Model Conditions

Four sample target concentration ranges were tested in each model. These ranged from (i) 0-10 copies/mL, (ii) 0-50 copies/mL, (iii) 0-50,000 copies/mL and (iv) 0-300,000 copies/mL. These concentration ranges were chosen to assess (i) sensitivity at very low concentrations, (ii) performance across a clinically significant graft-derived cfDNA range, (iii) performance across a clinically significant total cfDNA range, and (iv) the effect of saturation of total droplets at extremely high concentrations, respectively.

For all model runs, droplet volume $V_{droplet}$ was fixed at 0.001 μL (1nL), the standard droplet volume for the Bio-Rad droplet generators. Positive droplet number outputs were conservatively rounded down to the nearest whole droplet number.

4.2.6.1.3 Modelling Sample and Template Volume

Three models were constructed using Equation 4-16 to examine the effect of template DNA volume (the volume of extracted cfDNA used per well) on the expected number of positive droplets from clinical samples. The template volume used per well is limited by the desired number of PCR reactions and the template volume available for use. The template volume available for use is a function of the sample volume subject to extraction and the eluent volume used during the extraction process.

For post-transplant cfDNA monitoring, a panel size of 31 CNV assays and *ACE* assay distributed across 16 wells (2 columns on a ddPCR plate) was proposed. Consequently, minimum total template volumes were 46 μL , 92 μL and 138 μL for template volumes of 2.5 μL , 5 μL and 7.5 μL per well respectively, incorporating a 15% conservative surplus. Given that extraction of cfDNA is limited to a maximum of 5 mL per adsorption column and that each column is eluted in 50 μL of eluent, template volumes larger than 2.5 μL require that sample volumes greater than 5 mL are collected and extracted via multiple adsorption columns.

Additionally, the minimum volume of PCR reagents within each well is approximately 18.3 μL . Consequently, template volumes above 6.7 μL are associated with an increase in the total PCR reaction volume above 25 μL . Below this template volume, increasing volumes of DNA are offset by equivalent reductions in the volume of distilled water required to make up the total reaction volume to 25 μL .

These relationships between template volume, total PCR reaction volume, required total template volume, eluent volume and sample volumes have informed the model conditions summarised in Table 4-2.

Table 4-2: Sample and template volume model conditions.

Variable	2.5 μL Model	5 μL Model	7.5 μL Model
Template volume $V_{template}$	2.5 μL	5 μL	7.5 μL
Volume of sample V_{sample}	5 mL	10 mL	15 mL
Elution volume $V_{elution}$	50 μL	100 μL	150 μL
Volume of PCR reagents V_{MM}	22.5 μL	20 μL	18.3 μL
Total number of droplets N_{total}	15,000		

4.2.6.1.4 Modelling Droplet Number

Three models were developed using Equation 4-16 to examine the effect of total droplet number on the expected number of positive droplets from clinical samples across the concentration ranges specified.

Template volume, sample volume, elution volume and PCR reaction volume were fixed across the models (Table 4-3). A template volume of 5 μL was chosen based upon the results of the prior modelling (section 4.2.6.1.3).

Table 4-3: Total droplet number model conditions.

Variable	10K Model	15K Model	20K Model
Template volume V_{template}	5 μL		
Volume of sample V_{sample}	10 mL		
Elution volume V_{elution}	100 μL		
Volume of PCR reagents V_{MM}	20 μL		
Total number of droplets N_{total}	10,000	15,000	20,000

4.2.6.2 *Dilution Series Design*

A titration series of one-copy genotype fragmented genomic DNA was designed. Consideration was given to (i) ensuring coverage of a clinically-relevant concentration range, including detailed coverage of concentrations of diagnostic significance and at least partial coverage of concentrations seen for recipient one-copy markers, (ii) minimising the number of titration steps to ensure sufficient DNA availability to cover all necessary replicates and assay pairs, (iii) ensuring the proposed number of titration steps could fit within a single digital PCR plate where possible including provision of zero-copy and no template negative controls, and (iv) ease of creation of the dilution series to minimise the risk of carrier error at each step.

Equivalent clinical sample concentrations, expected positive droplet numbers and equivalent PCR reaction concentrations were modelled against the starting template DNA concentration using Equations 4-11, 4-12 and 4-15. Parameters used for modelling are specified in Table 4-4.

Table 4-4: Parameters used to model clinical sample equivalency of dilution series concentrations.

Parameter	Value
Template volume	5 μL
PCR reagent volume	20 μL
Total number of droplets	15,000
Elution volume	100 μL
Sample volume	10 mL

4.2.6.3 Experimental Grouping

Creation of each dilution series was split into three separate groups based upon the three one-copy genotypes studied (Table 4-5). The study groups were run in series (Figure 4-1). Each group was named “validation study” followed by a suffix representing the chronological order in which the experiments were performed. Given the strong dependence of the experimental plan on laboratory skill and the expected learning curve relating to this, validation experiments were performed starting with the one-copy genotype with the fewest assay pairs progressing to the one-copy genotype with the most assay pairs. All methodological steps subsequently described, with the exception of the final analysis, were thus conducted separately for each validation study group.

Table 4-5: Genotypes and assay pairs for each of the three validation study groups.

Assay Pair	One-Copy Dilution Series Genotype	FAM-Assay	HEX-Assay	Zero-Copy Control Genotype
Validation Study 1				
1	148	31A	01B	114
2	148	25A	30A	114
3	148	03B	20A	114
4	148	ACE	34A	114
Validation Study 2				
1	146	29A	18A	148
2	146	38A	32A	148
3	146	36A	26A	148
4	146	10C	39A	148
5	146	14A	41A	148
6	146	05B	24A	133
7	146	ACE	28A	133
Validation Study 3				
1	71	23A	17A	97
2	71	21A	06B	97
3	71	12A	37A	97
4	71	27A	08B	97
5	71	19B	13A	114
6	71	02B	22A	114
7	71	33A	15A	114
8	71	ACE	04B	114

4.2.6.4 Fragmented Cellular DNA Quantification

Fragmented cellular DNA from one-copy and zero-copy genotype participants was diluted with distilled water to approximately 2-5 ng/μL using pre-fragmentation NanoDrop and post-fragmentation TapeStation concentrations as a guide. This target template concentration ($C_{template}$) was calculated to approximate 150-375 copies/μL. The conversion was based upon a 6.6 pg

quantity of DNA per diploid genome, which is converted from pg to ng and doubled to account for the one-copy targets used in the experiment. This doubling is required as twice the amount of DNA is required to achieve 150 copies/ μ L of a one-copy target than that required to achieve the same concentration of a two-copy target.

$$C_{template} \left(\frac{copies}{\mu L} \right) = \frac{C_{template} \left(\frac{ng}{\mu L} \right)}{\left(\frac{6.6 \times 2}{1000} \right)} \quad 4-1$$

$C_{template}$ is equivalent to $C_{elution}$ in Equation 4-9, but is named differently to reflect the use of fragmented cellular DNA rather than eluted cfDNA in these calculations.

Diluted fragmented cellular DNA for each genotype was run against one-copy markers using ddPCR. The average one-copy marker concentration was used to calculate the template concentration of the diluted fragmented cellular DNA sample. The post-fragmentation pre-dilution concentration (the neat fragmented concentration) was also calculated from the average one-copy marker concentration ($C_{diluted}$, copies/mL) and the volumes of neat fragmented sample used for dilution (V_{neat} , μ L) and final volume of the diluted sample ($V_{diluted}$, μ L).

$$C_{neat} = \frac{C_{diluted} \times V_{diluted}}{V_{neat}} \quad 4-2$$

4.2.6.5 Dilution Series Creation

Droplet digital PCR one-copy marker concentrations were used to create a starting solution approximating 128 copies/ μ L for each one-copy and zero-copy genotype(s) within the respective validation study. Distilled water was used as a diluent. Special attention was paid towards ensuring sufficient volume was available for creation of the dilution series and replicates without unnecessary waste of DNA in the event that the experimented needed to be repeated.

For one-copy genotype DNA, half of the starting solution was aliquoted into a 1.5 mL Eppendorf Tube (or equivalent) and diluted with an equal volume of distilled water. Half of this diluted starting solution was aliquoted into another 1.5 mL Eppendorf Tube and diluted with an equal volume of distilled water. The process was repeated to create the 8 steps of the dilution series. The tubes were labelled with the participant number, “fragmented diluted” and the expected dilution concentration.

Dilution series and zero-copy control DNA for use in imminent validation study experiments was stored at -20 °C. Neat fragmented DNA samples were stored at -80 °C.

4.2.6.6 *PCR Plate Setup*

PCR reagents and the prepared one-copy genotype dilution series, as well as the zero-copy genotypes corresponding to assay pairs within the validation study underway, were equilibrated to 4 °C. All ddPCR plates for a particular validation study were created on the same day, one plate at a time. Each subsequent plate was created after the preceding plate commenced PCR (Figure 4-2). PCR reagents and DNA were stored on ice or at 4 °C when not immediately in use. They were equilibrated with room temperature, vortexed for 30 seconds and briefly spun down prior to use.

Two assay pairs from the validation study underway, each pair comprising a FAM- and a HEX-assay run in duplex, were selected in sequence. Where possible, assay pairs with the same zero-copy control genotype were selected together. A master solution comprising ddPCR Supermix for Probes (No dUTP), forward and reverse primers, probes and distilled water was created for each assay pair (Table 4-6). The master solution was distributed into 4 tubes of an 8 well strip of 0.2 mL tubes. The first four tubes (A-D) corresponded to the first assay pair for that plate, and the second four tubes (E-H) corresponded to the second assay pair for that plate, consistent with the proposed plate layout (Figure 4-3). From this 8 well strip, an 8 channel pipette was used to distribute the assay pair master solution into all 10 columns of the digital PCR plate. 20 µL master solution was aliquoted per well.

50 µL of each one-copy dilution step was transferred into sequential tubes of an 8 well strip of 0.2 mL tubes. From this 8 well strip, an 8 channel pipette was used to distribute the dilution series into 8 rows of the digital PCR plate, consistent with the proposed plate layout (Figure 4-3). 5 µL of DNA was aliquoted per well.

For the negative controls, 5 µL of no template control (distilled water) or 5 µL of assay pair-matched zero-copy genotype DNA was pipetted using a single channel pipette into the appropriate wells (Figure 4-3).

A foil seal was applied to the completed digital PCR plate using the Bio-Rad PX1 PCR Plate Sealer at 180 °C for 10 seconds. The completed plate was labelled, vortexed for 1 minute, spun down for 30 seconds at 1000 x g, and then transferred to a PCR machine.

Figure 4-2: Schematic overview of the validation study laboratory work. The preparation of PCR plates occurred two assay pairs at a time, repeating until all assay pairs for the particular validation study had been completed (*).

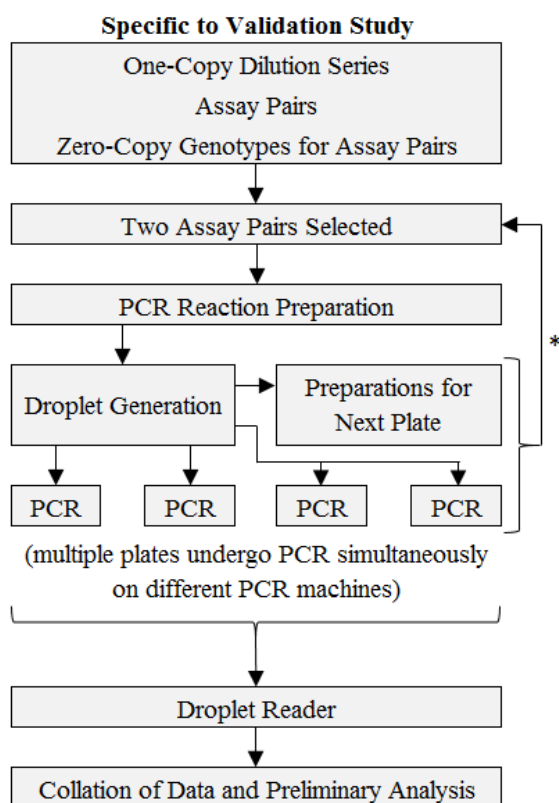


Table 4-6: PCR reaction components for the analytical validation studies.

PCR Reagent	Concentration	Volume per Well (µL)	Final Concentration
ddPCR Supermix for Probes (No dUTP)	2x	12.5	1x
FAM-Assay Primers	10 uM	2.25	900 nM
FAM-Assay Probe	10 uM	0.625	250 nM
HEX-Assay Primers	10 uM	2.25	900 nM
HEX-Assay Probe	10 uM	0.625	250 nM
Distilled Water	0 uM	1.75	0 uM
DNA	Variable	5	Variable
Total		25	

Figure 4-3: Digital PCR 96-well plate layout for validation study assay pairs, demonstrating that 4 replicates for the proposed dilution series and negative controls fit within one plate. Columns are labelled numerically, rows are labelled alphabetically. Note that whilst the one-copy dilution

series template for a particular plate is the same irrespective of assay pair, the zero-copy negative control template may differ between assay pairs.

		1	2	3	4	5	6	7	8	9	10	11	12
PAIR 1	A	128	64	32	16	8	4	2	1	NTC	0-Copy		
	B	128	64	32	16	8	4	2	1	NTC	0-Copy		
	C	128	64	32	16	8	4	2	1	NTC	0-Copy		
	D	128	64	32	16	8	4	2	1	NTC	0-Copy		
PAIR 2	E	128	64	32	16	8	4	2	1	NTC	0-Copy		
	F	128	64	32	16	8	4	2	1	NTC	0-Copy		
	G	128	64	32	16	8	4	2	1	NTC	0-Copy		
	H	128	64	32	16	8	4	2	1	NTC	0-Copy		

4.2.6.7 Droplet Digital PCR

Droplet generation was performed using the Bio-Rad QX200 AutoDG Automated Droplet Generator, according to manufacturer instructions. While this was running, the laboratory workspace and researcher were readied for preparation of the next digital PCR plate.

When droplet generation was complete, a foil seal was applied to the new plate containing generated droplets using the Bio-Rad PX1 PCR Plate Sealer at 180 °C for 10 seconds. The completed plate was labelled and transferred to the PCR thermal cycler. Two Bio-Rad C/S1000 and two Bio-Rad T100 thermal cyclers were used to permit immediate concurrent PCR of each plate as it was completed. The PCR cycling conditions are specified in Table 4-7. Plates were held at 4 °C while awaiting droplet reader availability.

Plates were transferred to the Bio-Rad QX200 Droplet Reader in order of PCR completion. Bio-Rad QuantaSoft software (version 1.7) was used to complete well setup with the aid of template files designed to reflect the PCR plate setup design. Droplets were read using absolute quantification experiment settings with ddPCR Supermix for Probes (no dUTP) specified. Channel 1 and channel 2 were specified as unknown for all assay wells.

When all droplets had been read, preliminary review of the data confirmed no gross failure of the droplet reader or laboratory work prior to the next plate being loaded.

Table 4-7: PCR cycling conditions for the analytical validation study. Ramp rate was 2 °C/sec for all steps.

Cycling Step	Temp (°C)	Time	Number of Cycles
Enzyme activation	95	10 min	1
Denaturation	94	30 sec	40
Annealing/extension	60	1 min	
Enzyme deactivation	98	10 min	1
Hold	4	Infinite	2

4.2.7 **Accepted Droplets**

The distribution of total accepted droplets for each digital PCR well was described. Outlier wells were identified as falling below the 1st quartile minus 1.5x the inter-quartile range. Variation in accepted droplet numbers between validation studies was compared using the Kruskal-Wallis rank sum test with post-hoc pairwise comparisons using the Wilcoxon rank sum test with Bonferroni adjustment.

To assess the contribution of plate and assay pair to variation in accepted droplet number, random-effects models with varying intercepts were constructed with validation study, plate and assay pair as multilevel nested group effects. Models were compared using analysis of variance and the Akaike information criterion with a correction for finite sample size (AICc). The delta-AICc was calculated as the difference in AICc between sequential levels of the nested group effect variables.

4.2.8 **Droplet Cluster Centroids**

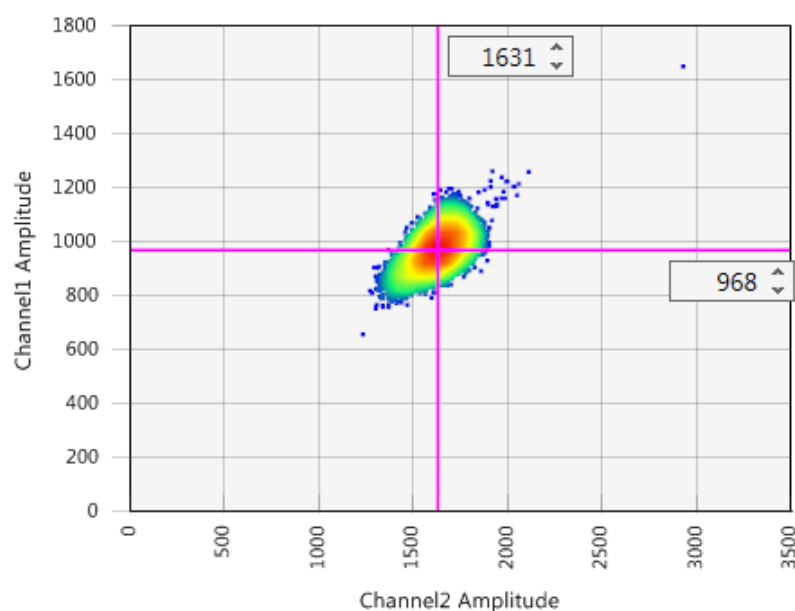
4.2.8.1 *Cluster Centroid Identification*

Droplet data was reviewed using Bio-Rad QuantaSoft Analysis Pro software (version 1.0.596).

Visual inspection of 2D heat maps was used to define approximate cluster centroids for each assay pair *en bloc* for all replicates of the dilution series (Figure 4-4). A cluster centroid was defined as the geometric centre of a cluster. Cluster centroids were obtained for the channel 1 and channel 2 negative (double negative) cluster, channel 1 positive and channel 2 negative (FAM-positive) cluster, and channel 1 negative and channel 2 positive (HEX-positive) cluster for each assay pair. Channel 1 positive and channel 2 positive (double positive) cluster centroid amplitudes were not recorded as they were usually more loosely distributed. The relative channel 1 and channel 2 distribution of cluster centroids for each cluster type (double negative, FAM-positive and HEX-positive) was described.

To assess consistency in visual centroid approximation, 8 replicates of ACE/28A from plate D of validation study 2 underwent cluster centroid estimation separately as two groups consisting of 4 replicates each. The coefficient of variation (CV, standard deviation/mean) for channel 1 and 2 amplitudes of each cluster centroid was calculated.

Figure 4-4: 2D heat maps were used to aid visual approximation of cluster centroids, in this case a double negative droplet cluster.



4.2.8.2 Double Negative Cluster Centroid Analysis

Visual inspection of 2D heat maps of zero-copy and no template controls for each assay pair was used to define approximate double negative cluster centroids for these groups. The distribution of double negative cluster centroid amplitudes was described for the dilution series, zero-copy and no template control groups.

To evaluate for variation in centroid distribution based on the template used, linear models were constructed for channel 1 and channel 2 amplitude across the groups. One-way analysis of variance tables were created for each model. Additionally, coefficient of variation (standard deviation/mean) was calculated for channel 1 and channel 2 amplitudes across the groups within each assay pair.

4.2.8.3 Positive Cluster Centroid Analysis

The distribution of CNV assay pair positive cluster centroid channel 1 and 2 amplitudes was described. FAM-positive cluster channel 1 amplitudes and HEX-positive cluster channel 2 amplitudes were compared using the Wilcoxon signed rank test. As a non-polymorphic two-copy control marker that was run in each validation study, *ACE* positive cluster centroids were evaluated separately from the remaining CNV assays. The distribution of *ACE* positive cluster centroids was described. Coefficient of variation (standard deviation/mean) was calculated for channel 1 and 2 amplitudes of *ACE* positive clusters across the validation studies.

Outlier low amplitude positive cluster centroids were identified as cluster centroids with (i) channel amplitudes falling below the 1st quartile minus 1.5x the inter-quartile range (Q1-IQR), and/or (ii) channel amplitude z-scores of less than -1.96. Assay design characteristics for assays with outlier low amplitude positive droplet cluster centroids were compared against those with non-outlier amplitude positive droplet cluster centroids. Assay design characteristics assessed included oligonucleotide melting temperatures, size, GC content, primer melting temperature delta, mean primer-probe melting temperature delta and amplicon size. Assay specificity was also reviewed. The statistical significance of variation of assay characteristics between non-outlier and outlier low amplitude positive droplet cluster centroids was assessed using the Wilcoxon rank sum test with continuity correction.

4.2.9 **Droplet Thresholding**

Positive droplet thresholding was performed on all wells for a particular assay pair *en bloc* using Bio-Rad QuantaSoft Analysis Pro software (version 1.0.596), blinded to the concentration result of each well. Within each assay pair, the same threshold was applied to the dilution series, zero-copy control and no template control.

Five different thresholding techniques were applied to each assay pair, and concentration data for each thresholding technique was exported for analysis:

- In the cluster thresholding technique, the threshold line mode was used to define channel 1 and channel 2 amplitude thresholds based tightly on positive droplet clusters, as visually assessed by the researcher (Figure 4-5).
- In the tight thresholding technique, the threshold cluster mode was used to lasso droplets outside of the double negative droplet cluster and assign them to the positive droplet cluster corresponding with the location of the droplet (Figure 4-6).
- The remaining thresholding techniques used threshold line mode limits based on a proportion of the distance between the double negative droplet cluster and the respective (FAM- or HEX-) positive droplet cluster for that channel. The half-, third- and quarter-distance thresholds were defined based on 50%, 33.3% and 25% of this distance respectively (Figure 4-7, Figure 4-8 and Figure 4-9).

Figure 4-5: Cluster thresholding technique applied to a 2D amplitude plot of 28 wells *en bloc* (CNV 29A and 18A dilution series with all replicates). Threshold lines are defined by the bounds of the positive droplet clusters which are visually assessed by the researcher.

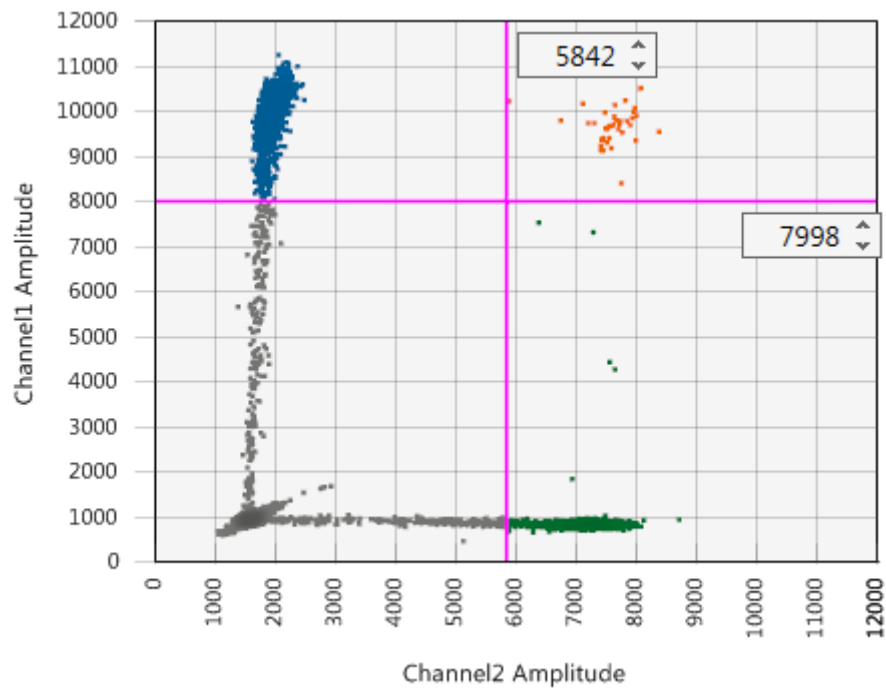


Figure 4-6: Tight thresholding technique applied to a 2D amplitude plot of 28 wells *en bloc* (CNV 29A and 18A dilution series with all replicates). Droplets outside of the negative droplet cluster were assigned to each respective positive cluster using the threshold cluster lasso tool.

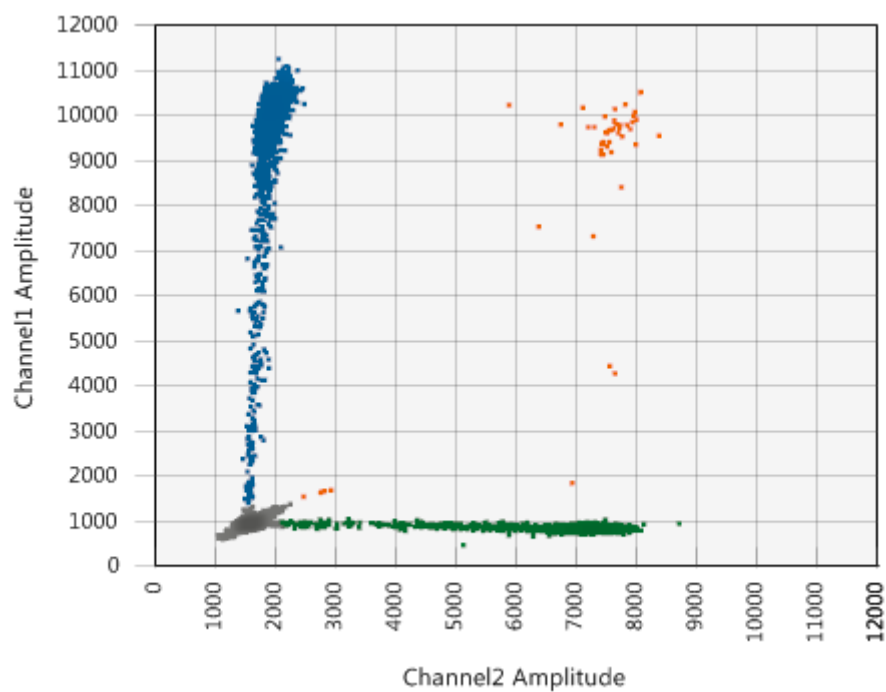


Figure 4-7: Half-distance thresholding technique applied to a 2D amplitude plot of 28 wells *en bloc* (CNV 29A and 18A dilution series with all replicates). Threshold lines were defined at half the distance between the double negative and respective (FAM- or HEX-) positive droplet cluster.

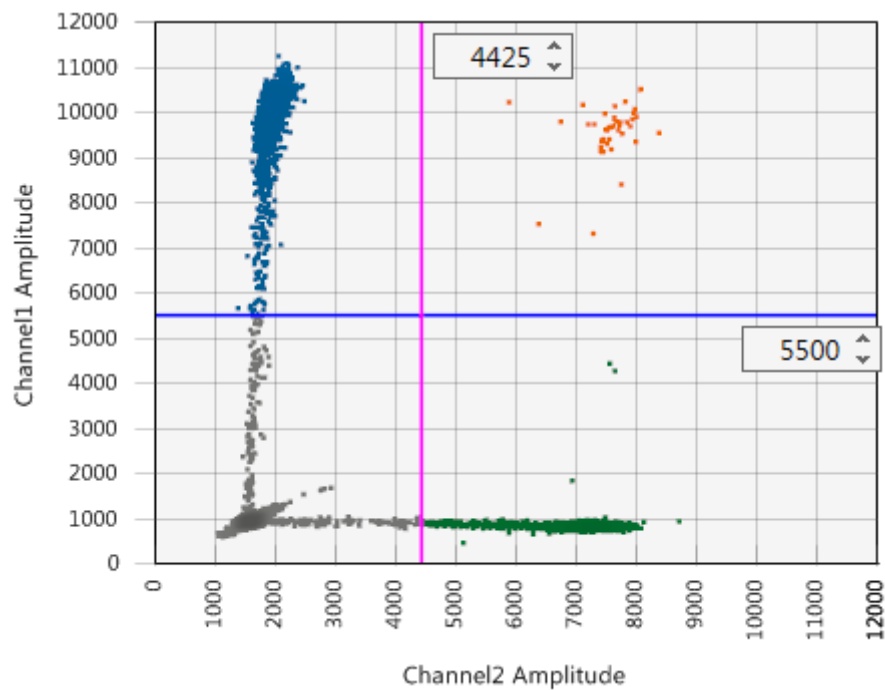


Figure 4-8: Third-distance thresholding technique applied to a 2D amplitude plot of 28 wells *en bloc* (CNV 29A and 18A dilution series with all replicates). Threshold lines were defined at a third of the distance between the double negative and respective (FAM- or HEX-) positive droplet cluster.

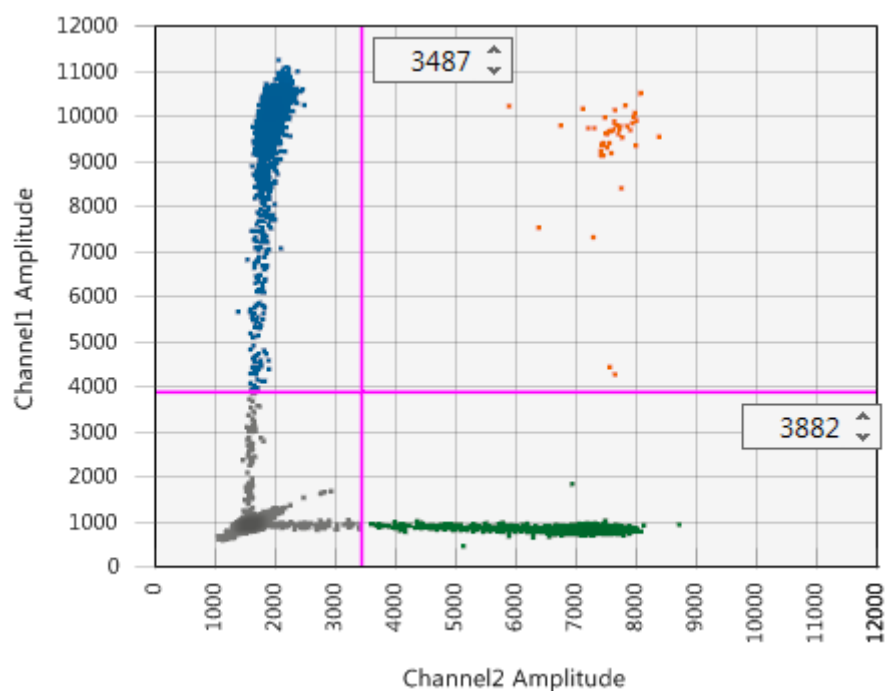
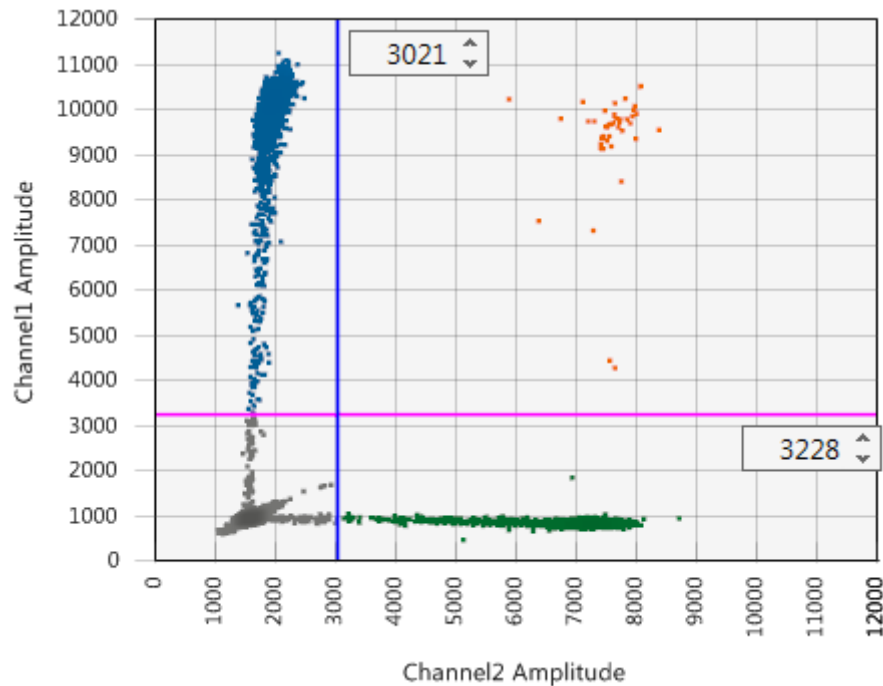


Figure 4-9: Quarter-distance thresholding technique applied to a 2D amplitude plot of 28 wells *en bloc* (CNV 29A and 18A dilution series with all replicates). Threshold lines were defined at a quarter of the distance between the double negative and respective (FAM- or HEX-) positive droplet cluster.



4.2.10 Conversion of Results to Plasma Equivalent

Droplet digital PCR results were converted to plasma equivalent concentrations using equation 4-12. As the DNA used in these validation experiments was not extracted cfDNA, standard modelled variables were used (Table 4-4).

4.2.11 No Template and Zero-Copy Controls

Data from no template control (distilled water) and zero-copy control wells was analysed separately. Results arising from wells with less than 10,000 accepted droplets were excluded.

False positive results were defined as those with at least one positive droplet (equivalent to a concentration greater than 0 copies/ μ L). All concentration results were converted from a digital PCR well concentration to an equivalent template concentration on the basis of 5 μ L template (distilled water or zero-copy control DNA) volume and a total PCR reaction volume of 25 μ L.

For zero-copy control wells, given that the *ACE* assay measured a non-polymorphic target present in all individuals, *ACE* results were not expected to be 0 copies/ μ L and so were excluded from

calculation of the false positive result rate. *ACE* results were included in the no template control analysis as this target sequence was expected to be absent in distilled water.

The false positive result rate was calculated as the proportion of all results that were false positive.

4.2.11.1 *ACE Results*

Non-polymorphic *ACE* assay concentrations for each validation study zero-copy control were converted into equivalent template concentrations on the basis of a 5 µL template volume and total PCR reaction volume of 25 µL. The median *ACE* template concentration was described for each validation study zero-copy participant DNA template.

4.2.11.2 *Threshold Technique*

A 5-sample test for equality of proportions without continuity correction was used to compare false positive result rates between threshold techniques. The chi-square test for trend in proportions was used to assess the relationship between threshold technique and false positive result rate. The relationship between threshold technique and false positive result concentration was tested with Spearman's rank correlation rho. Approximate *p*-values were used for tied data.

4.2.11.3 *Assay Design Characteristics*

Assay design characteristics and cluster centroid amplitudes for false positive assays were compared against the remaining tested assays using the two-tailed Wilcoxon rank sum test with continuity correction. Assay design characteristics assessed included oligonucleotide melting temperatures, size, GC content, primer melting temperature delta, mean primer-probe melting temperature delta and amplicon size. A Fisher's Exact Test was used to compare the false positive rate of FAM- compared to HEX- assays. Cluster centroid amplitudes for channels corresponding with the respective assay probe dye were used in the analysis (i.e. channel 1 amplitude for FAM assays and channel 2 amplitude for HEX assays).

4.2.11.4 *Limit of Blank (No Template Control)*

The limit of blank for each assay was calculated using the approach proposed by Armbruster et al., which is based upon the Clinical and Laboratory Standards Institute EP17 guideline:[282]

$$\text{limit of blank} = \text{mean}_{\text{blank}} + (1.645 \times \text{standard deviation}_{\text{blank}}) \quad 4-3$$

4.2.12 **One-Copy Dilution Series**

CNV assay results with one-copy targets were included in this analysis. Third-distance threshold technique results were used. *ACE* results were excluded given its two-copy genotype. Results arising from wells with less than 10,000 accepted droplets were excluded.

4.2.12.1 Achieved Concentration and Linearity

Target dilution step concentrations using the third-distance threshold technique were compared against the median concentration of all assays at that dilution step, grouped by validation study. The result was expressed as the proportion of the targeted concentration achieved. The consistency of the dilution steps within each validation study was compared using the Kruskal-Wallis rank sum test.

The effect of threshold technique on one-copy dilution step concentrations was studied. The proportion of the targeted concentration that was achieved was compared between threshold techniques of progressively lower amplitude. The lower amplitude threshold value was subtracted from the higher amplitude threshold value, and the difference expressed as an absolute delta value.

For each CNV assay, a linear model was created of observed versus expected assay concentration, using third-distance threshold technique data. The expected assay concentration ($C_{expected}$) was calculated based upon the median starting concentration (dilution step zero, $C_{starting}$) of the respective validation study, and the two-fold dilution series study design (where *dilution* is the dilution step).

$$C_{expected} = \frac{C_{starting}}{2^{dilution}} \quad 4-4$$

Linearity was assessed using the slope of the linear model and Pearson's correlation coefficient (r). Linear model parameters were compared using the Kruskal-Wallis rank sum test with post-hoc comparisons using Dunn's test.

Multiple linear regression was performed to investigate whether linear model slope and correlation coefficient was predicted by assay design characteristics including oligonucleotide melting temperatures, size, GC content, primer melting temperature delta, and amplicon size.

Linear model parameter outlier assays were identified using the distribution of linear model parameter values from all studied CNV assays (i.e. the data was not grouped by validation study). Outlier CNV assays were identified when they met one of the following rules (i) z-score greater than 1.96/less than -1.96, (ii) parameter value less than the 1st quartile minus 1.5x the inter-quartile range (Q1-IQR), or (iii) parameter value greater than the 3rd quartile plus 1.5x the inter-quartile range (Q3+IQR). These latter two criteria are collectively termed the "IQR test".

4.2.12.2 Limit of Detection

The proportion of assay replicates detecting target at each dilution step was calculated. Detection of target was defined as the presence of at least one positive droplet, equivalent to a concentration of greater than 0 copies/μL.

The limit of detection was calculated using Equation 4-5. [282] The assay-specific limit of blank was defined by the zero-copy control analysis. The low concentration sample standard deviation was calculated using the lowest concentration at which the assay had at least two replicates detecting target.

$$\begin{aligned} \text{limit of detection} \\ &= \text{limit of blank} + (1.645 \\ &\times \text{standard deviation}_{\text{low concentration sample}}) \end{aligned} \quad 4-5$$

4.2.12.3 Intra-Assay Variation

Intra-assay coefficient of variation (standard deviation/mean) was calculated for each CNV assay at each dilution step. False negative (0 copy/μL) replicate results were excluded from this analysis. Outlier assays at each dilution step were identified using the IQR test.

A linear model was constructed relating the median assay concentration and intra-assay coefficient of variation. Correlation was assessed using Pearson's correlation coefficient. The model was used to calculate a 95% prediction interval across the studied concentration range. Stepwise linear regression and repeated measures analysis of variance was used to determine whether the CNV assay itself predicted intra-assay coefficient of variation.

Intra-assay variation calculated with exclusion of false negative results was also performed. Kruskal-Wallis rank sum test was used to compare the distribution of intra-assay variation between CNV assays. Repeated measures analysis of variance was also used to examine the relationship between dependent variable intra-assay coefficient of variation and independent variables including median assay concentration, number of replicates detecting target, and CNV assay.

4.2.12.4 Inter-Assay Variation

Inter-assay coefficient of variation (standard deviation/mean) was calculated for each validation study dilution step using the concentration each assay at that dilution step. The relationship between inter-assay coefficient of variation and dilution step concentration, defined as the median assay concentration at that dilution step was described.

A linear model was constructed relating the median assay concentration and inter-assay coefficient of variation. Correlation was assessed using Pearson's correlation coefficient. The

model was used to calculate a 95% prediction interval across the studied concentration range. Stepwise linear regression and analysis of variance was used to determine whether the validation study predicted inter-assay coefficient of variation.

The proportionate variation of each CNV assay result was calculated by dividing the median of the replicate results by the median result of all CNV assays at that validation study dilution step. The proportionate variation of each CNV assay across all dilution steps was plotted, grouped by dilution step. Repeated measures analysis of variance and multiple regression were used to evaluate the effect of CNV assay, and assay design characteristics, on proportionate variation across the dilution series.

4.2.12.5 *Limit of Quantification*

The limit of quantification was defined as the lowest concentration at or above the limit of detection at which inter- and intra-assay coefficient of variation were > 50%.

4.2.13 **Two-Copy Dilution Series**

ACE assay results with two-copy targets were included in this analysis. Third-distance threshold technique results were used. CNV assay results were excluded. Results arising from wells with less than 10,000 accepted droplets were excluded.

4.2.13.1 *Linearity*

Target dilution step concentrations were compared against the median concentration of ACE replicates at each dilution step, grouped by validation study. The result was expressed as the proportion of the targeted concentration achieved. The consistency of the dilution steps within each validation study was compared using the Kruskal-Wallis rank sum test.

For each validation study, a linear model was created of observed versus expected ACE assay concentration. The expected assay concentration was calculated based upon Equation 4-4. Linearity was assessed using the slope of the linear model and Pearson's correlation coefficient (r).

4.2.13.2 *Limit of Detection*

The proportion of assay replicates detecting target at each dilution step was calculated. Detection of target was defined as the presence of at least one positive droplet, equivalent to a concentration of greater than 0 copies/ μ L.

The limit of detection was calculated using Equation 4-5. [282] The assay-specific limit of blank was defined by the no template control analysis. As a non-polymorphic marker, no zero-copy control template was available by definition. The low concentration sample standard deviation

was calculated using the lowest concentration at which the assay had at least two replicates detecting target.

4.2.13.3 *Intra-Assay Variation*

Intra-assay coefficient of variation (standard deviation/mean) was calculated at each dilution step.

A linear model was constructed relating the median *ACE* assay concentration and the intra-assay coefficient of variation. Correlation was assessed using Pearson's correlation coefficient. The model was used to calculate a 95% prediction interval across the studied concentration range.

4.2.13.4 *Inter-Assay Variation*

Inter-assay coefficient of variation was not calculated as only the *ACE* assay was included in this analysis.

4.2.13.5 *Limit of Quantification*

The limit of quantification was defined as the lowest concentration at or above the limit of detection at which inter- and intra-assay coefficient of variation were > 50%.

4.2.14 **Statistical Significance**

Statistical significance was defined as a p -value < 0.05. In some graphs, for clarity, p -value is expressed using asterisk symbols (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$).

4.3 RESULTS

4.3.1 **One-Copy Genotypes**

4.3.1.1 *Combinatorial Modelling*

38/39 (97%) CNV assays were included in the combinatorial modelling. CNV assay 11A was not included in the modelling as the assay had not been widely run in the individuals in the study group by the time that combinatorial modelling was run.

The extent of assay genotyping varied between the 25 individuals. A median of 37 (34-39) assays had been genotyped per individual (Figure 4-10). In all cases where genotyping was not available for an assay, this was due to the respective assay not having been run on the sample by the time of the combinatorial modelling. This scenario arose because the assays were developed (or re-developed, as needed) over a period of time, in numerical order of assay number. Therefore, higher assay numbers tended to have been run in fewer individuals than lower assay numbers (Figure 4-11).

Of the markers for which genotyping was available, 338/827 (41%) were genotyped as recipient one-copy, 84/827 (10%) were genotyped as recipient two copy, and 405/827 (49%) were

genotyped as recipient zero-copy. The frequency of markers with a recipient one-copy genotype varied (Figure 4-12) with each marker being recipient one-copy in a median of 9 (7-11) participants.

The number of combinations examined and the resultant maximal unique recipient one-copy marker coverage achieved are summarised in Table 4-8. Increasing the group size improved the distribution of recipient one-copy markers covered, although benefits of increasing the group size beyond 5 individuals was limited (Figure 4-13). Complete marker coverage was achieved using a group size of 4 or more.

Figure 4-10: Distribution of total number of markers genotyped per individual of the 25 individuals used for combinatorial modelling.

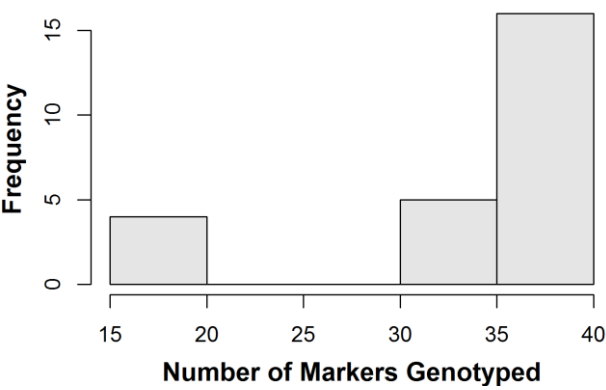


Figure 4-11: Frequency of non-genotyped markers amongst the 25 participants used for combinatorial modelling.

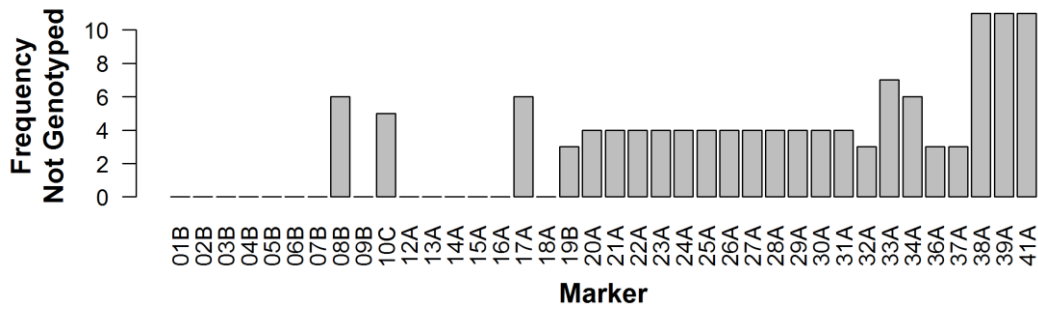


Figure 4-12: Frequency of recipient one-copy markers amongst the 25 participants used for combinatorial modelling.

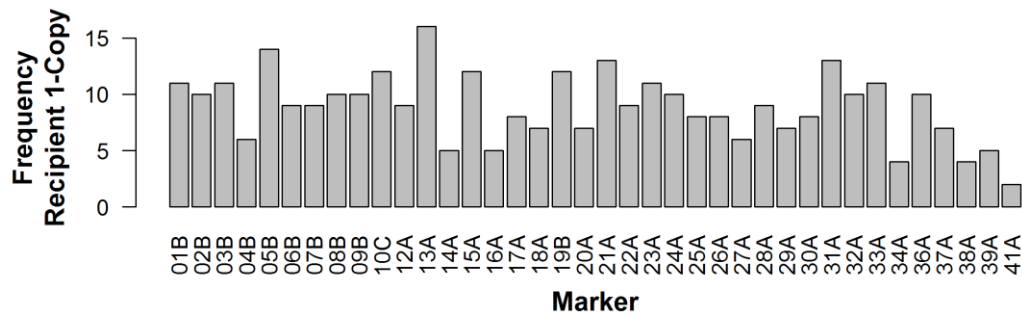
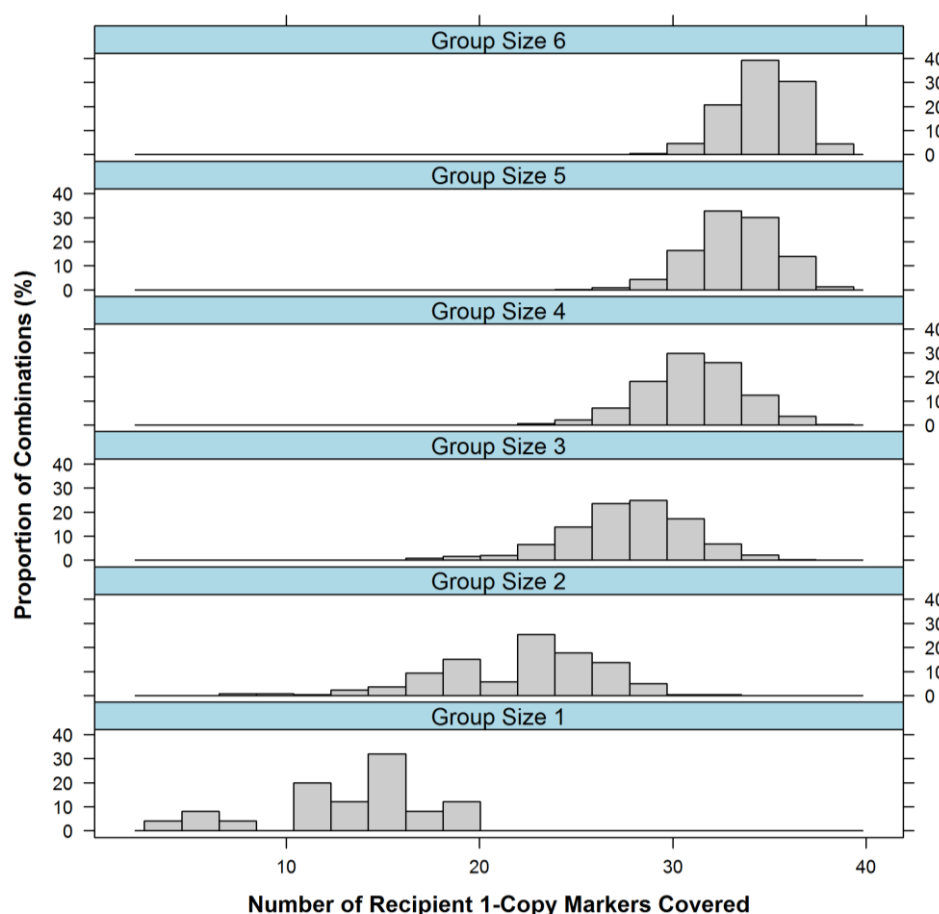


Table 4-8: Combinatorial algorithm modelling of maximal recipient one-copy marker coverage using 25 participants and various group sizes.

Group Size	Combinations Evaluated	Maximal Recipient One-Copy Marker Coverage	Combinations With Maximal Marker Coverage
1	25	20	1
2	300	32	1
3	2,300	37	1
4	12,650	38	21
5	53,130	38	653
6	177,100	38	7,841

Figure 4-13: The distribution of recipient one-copy markers covered by all combinations of different genotypes, grouped by the number of individuals in the combination (group size).



4.3.1.2 Genotype Selection

Genotype combinations consisting of 3 individuals within the top 1% of marker coverage (35 markers or above) were selected for manual review. This group size was considered to represent the best compromise between group size and marker coverage. Incorporating the limitations concerning participant 101 discussed below, combinations comprising paired (group size 2) participant genotypes could only achieve maximal CNV assay coverage of 29. The incremental coverage benefit of group sizes greater than 3 individuals was small.

High coverage trio combinations relied heavily on a number of individuals who were widely genotyped and had high recipient one-copy frequencies. Of the 24 trio combinations with recipient one-copy coverage of 35 markers or more, participant 146 occurred in 20/24 (83%) and participant 101 occurred in 14/24 (58%). As there was insufficient DNA for participant 101, 58% of combinations were excluded. Additionally sampling from this individual was not permitted within the ethics approval obtained for this work.

The highest coverage combination option that didn't include participant 101 met all other selection criteria and proceeded to confirmatory genotyping. This trio consisted of participants 71, 146 and 148.

4.3.1.3 Confirmatory Genotyping

Confirmatory genotyping using cellular DNA from buffy coat leucocytes matched the one-copy genotypes defined using cfDNA for the combinatorial modelling (Figure 4-14).

The combination of participants 71, 146 and 148 provided coverage of 35 markers, missing CNV assays 07B, 09B and 16A. 10/35 (29%) markers were one-copy in two participants, and 3/35 (9%) markers were one-copy in all three participants (Figure 4-15).

Variation amongst the one-copy markers was observed within each sample (Figure 4-16). The inter-quartile range was -3% to 5%, and 90% of markers varied between -11% and +11%. CNV 04B and 41A underestimated the one-copy concentration by 21% and 25% respectively. CNV 33A and 24A overestimated the one-copy concentration by 12% and 17% respectively.

When the variation amongst duplicated one-copy markers was compared across samples, mean variation around the one-copy median was less than 5% for all assays except CNV 17A (+9%), CNV 33A (+9%), and CNV 24A (+16%) (Figure 4-17).

Figure 4-14: Confirmatory genotyping of 38 markers using cellular DNA from 3 participants selected for use in the analytical validation experiments. Two-copy markers are shown to contextualise one-copy results, but were not included in the assay selection.

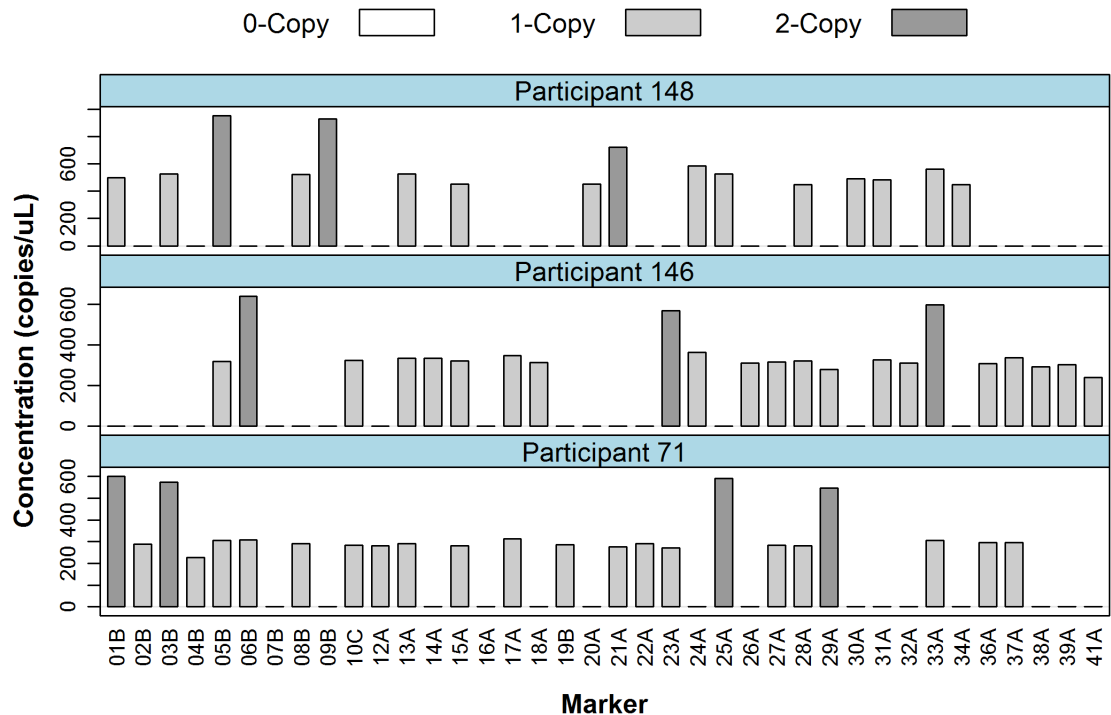


Figure 4-15: CNV one-copy marker coverage and frequency amongst participants 71, 146 and 148.

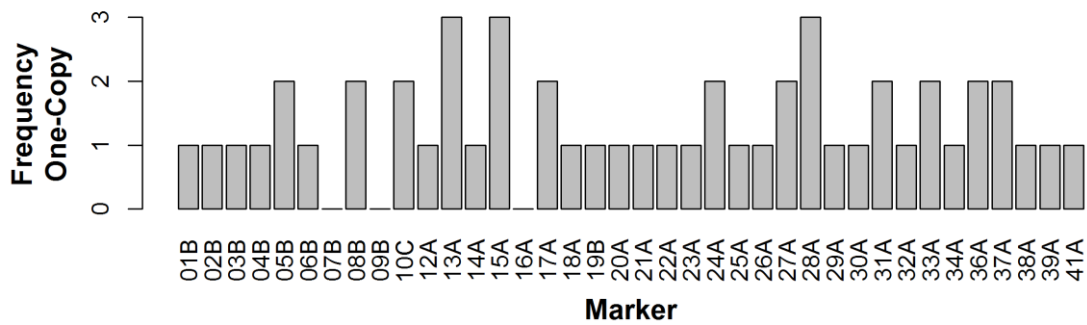


Figure 4-16: Variation of all CNV markers from the one-copy median at the time of confirmatory genotyping using cellular DNA.

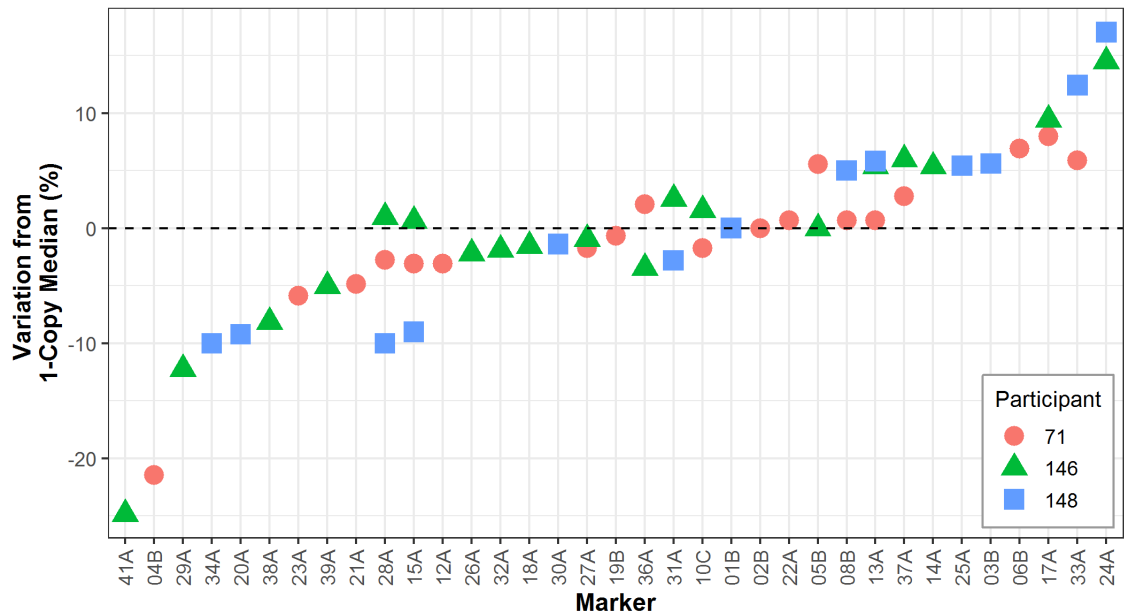
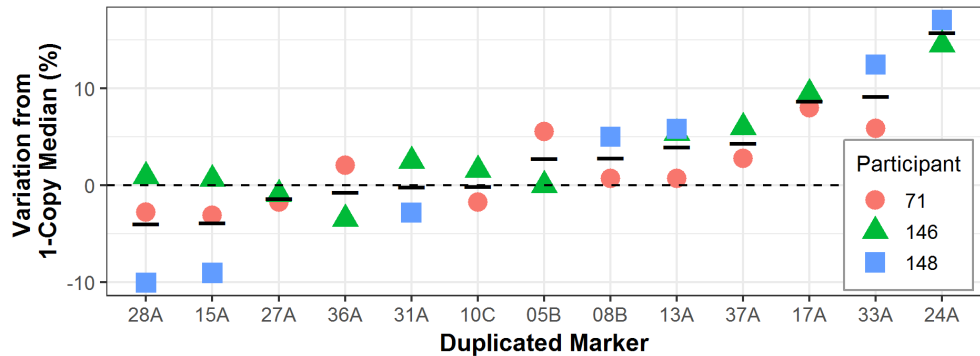


Figure 4-17: Variation of duplicated CNV markers from the one-copy median at the time of confirmatory genotyping using cellular DNA. Dash (-) points indicate the mean variation of the duplicates from the one-copy median of their respective samples.



4.3.1.4 Duplexed Assay Pair Selection

Duplex assay pairings for each participant are specified in Table 4-9.

CNV loci that were one-copy in more than one participant permitted some flexibility in the allocation of CNV assays, permitting (i) achievement of the optimal FAM:HEX assay-probe ratio for each participant, and (ii) selection of CNV assays within each duplicate that varied least from the median one-copy concentration at the time of confirmatory genotyping (Table 4-9A). Pairing of CNV assays which appeared to under-estimate the one-copy median concentration on confirmatory genotyping with those that tended to over-estimate this metric was achieved in two of the participants. For participant 146, this could not be achieved for all assay pairs due to the need to ensure pairings had a compatible zero-copy control (section 4.3.2).

Table 4-9: Assay pairings by participant indicating the variation of each assay from the one-copy median on confirmatory genotyping, the duplicated assays removed (A, strikethrough) and the final assay pairings (B). Asterisked (*) assays (B) indicate pair rearrangements within a participant driven by the need to match a suitable zero-copy control.

A				B				
Variation (%)	FAM Assay	HEX Assay	Variation (%)	Pair #	Variation (%)	FAM Assay	HEX Assay	Variation (%)
Participant 71				Participant 71				
-5.9	23A	17A	8.0	1	-5.9	23A	17A	8.0
-4.8	21A	06B	6.9	2	-4.8	21A	06B	6.9
-3.1	12A	~~05B~~	5.5	3	-3.1	12A	37A	2.8
-1.7	~~10C~~	37A	2.8	4	-1.7	27A	08B	0.7
-1.7	27A	08B	0.7	5	-0.7	19B	13A	0.7
-0.7	19B	13A	0.7	6	0.0	02B	22A	0.7
0.0	02B	22A	0.7	7	5.9	33A	15A	-3.1
2.0	~~36A~~	~~28A~~	-2.8	8		ACE	04B	-22
5.5	~~05B~~	15A	-3.1	Count		7	8	
5.9	33A	04B	-22	Participant 146				
6.9	~~06B~~			1	-12	29A	18A*	-1.6
Count	11	10		2	-8.2	38A	32A*	-1.9
Removed	4	2		3	-3.5	36A	26A*	-2.2
Net	7	8		4	1.6	10C*	39A*	-5.0
Participant 146				5	5.4	14A*	41A*	-25
-12	29A	24A	15	6	0.0	05B*	24A*	15
-8.2	38A	~~17A~~	9.4	7		ACE	28A*	0.9
-3.5	36A	~~37A~~	6.0	Count		6	7	
-0.9	~~27A~~	~~13A~~	5.4	Participant 148				
0.0	05B	28A	0.9	1	-2.8	31A	01B	0.0
1.6	10C	~~15A~~	0.63	2	5.4	25A	30A	-1.4
2.5	~~31A~~	~~05B~~	0.0	3	5.6	03B	20A	-9.2
5.4	14A	18A	-1.6	4		ACE	34A	-10
		32A	-1.9	Count		3	4	
		26A	-2.2					
		39A	-5.0					
		41A	-25					
Count	8	12						
Removed	2	5						
Net	6	7						

Participant 148			
-2.8	31A	~~24A~~	17
5.4	25A	~~13A~~	5.8
5.6	03B	~~08B~~	5.0
12	~~33A~~	01B	0.0
		30A	-1.4
		~~15A~~	-9.0
		20A	-9.2
		~~28A~~	-10
		34A	-10
Count	4	9	
Removed	1	5	
Net	3	4	

4.3.2 Zero-Copy Genotypes

4.3.2.1 Combinatorial Modelling and Genotype Selection

A summary of the genotypes included in the combinatorial model was presented in section 4.3.1.1. Of the markers for which genotyping was available, 405/827 (49%) were genotyped as recipient zero-copy. The frequency of markers with a recipient zero-copy genotype varied, with each marker being recipient zero-copy in a median of 10 (8-12) individuals (Figure 4-18).

The number of paired assay combinations modelled, and the resultant maximal unique recipient zero-copy marker coverage achieved for each one-copy participant genotype, is summarised in Table 4-10. Two participants were identified as permitting the most efficient zero-copy marker coverage for two of the one-copy participant genotypes, whereas one participant was identified to provide complete coverage of the outstanding one-copy participant genotype (Table 4-11).

Figure 4-18: Frequency of recipient zero-copy markers amongst the 25 participants used for combinatorial modelling.

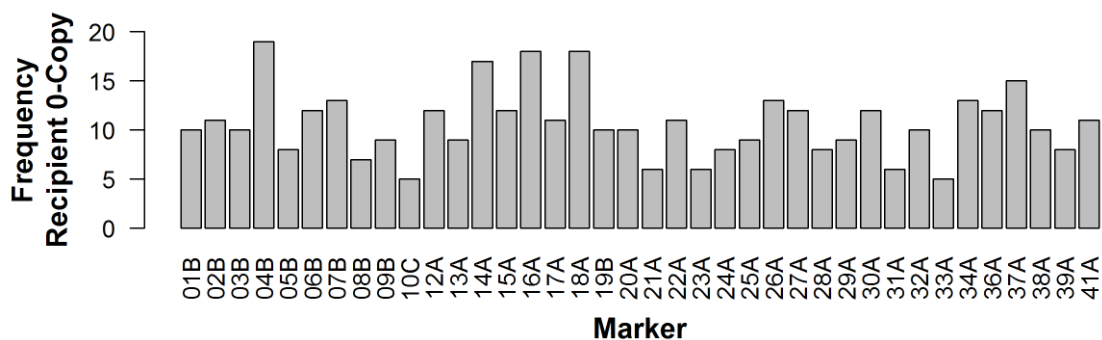


Table 4-10: Combinatorial algorithm modelling of maximal recipient zero-copy marker coverage of each one-copy participant genotype using 25 participants and various group sizes.

Group Size	Paired Assay Combinations Evaluated	Proportion of One-Copy Assays Pairs Covered by Zero-Copy Genotype	Combinations With Maximal Paired Assay Coverage
Participant 71			

1	21	6/8 (75%)	1
2	210	8/8 (100%)	1
3	1330	8/8 (100%)	25
Participant 146			
1	19	5/7 (71%)	1
2	171	7/7 (100%)	1
3	969	7/7 (100%)	31
Participant 148			
1	15	4/4 (100%)	1
2	105	4/4 (100%)	18
3	455	4/4 (100%)	133

Table 4-11: The most efficient zero-copy participant allocation to act as a zero-copy controls for each one-copy participant assay pair.

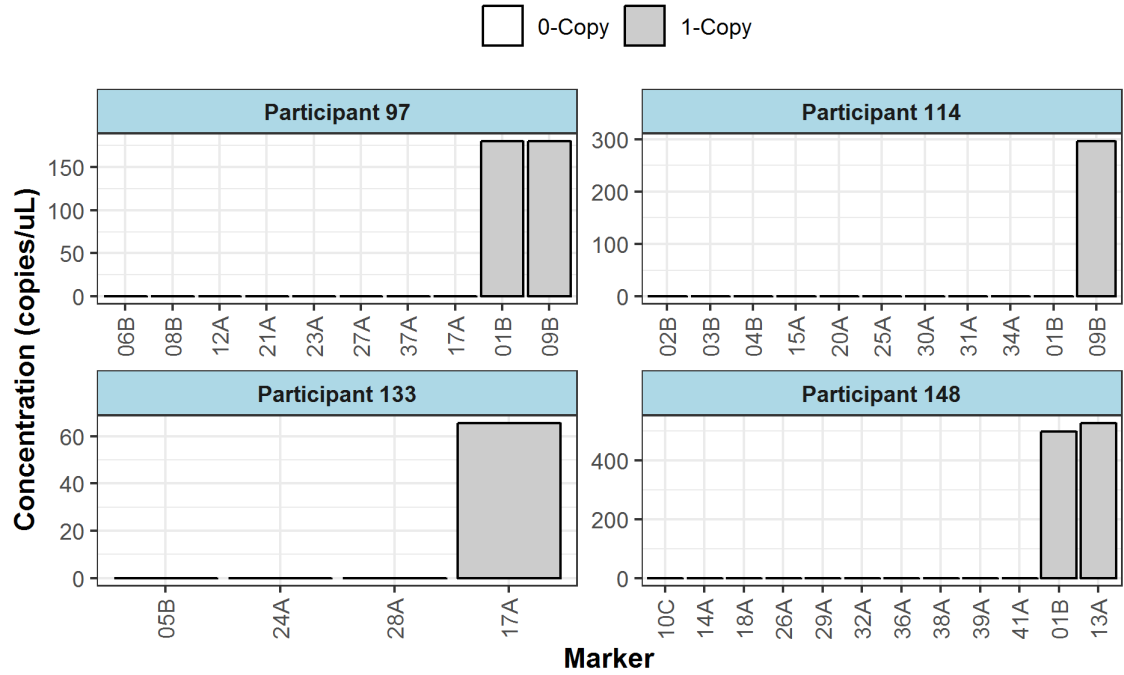
Assay Pair #	FAM Assay	HEX Assay	Zero-Copy Participant
One-Copy Participant 71			
1	23A	17A	97
2	21A	06B	97
3	12A	37A	97
4	27A	08B	97
5	19B	13A	114
6	02B	22A	114
7	33A	15A	114
8	ACE	04B	114
One-Copy Participant 146			
1	29A	18A	148
2	38A	32A	148
3	36A	26A	148
4	10C	39A	148
5	14A	41A	148
6	05B	24A	133
7	ACE	28A	133
One-Copy Participant 148			
1	31A	01B	114
2	25A	30A	114
3	03B	20A	114
4	ACE	34A	114

4.3.2.2 Confirmatory Genotyping

Confirmatory genotyping of cellular DNA from the selected zero-copy participants confirmed detectable target for one-copy positive control markers and undetectable target for all respective zero-copy control assays (Figure 4-19). Participants 97, 114 and 148 had digital PCR assays run

against un-fragmented cellular DNA. To preserve limited DNA for participant 133, genotyping of zero-copy markers was performed as part of DNA quantification after sonic fragmentation, discussed below.

Figure 4-19: Confirmatory genotyping of zero-copy control participants. Detectable markers had been genotyped as recipient one-copy assays using cfDNA.



4.3.3 DNA Fragmentation

4.3.3.1 *Fragmentation Profile of Cell-Free DNA*

Three samples were chosen for fragmentation analysis on the basis of digital PCR plasma *ACE* marker concentrations of approximately 76,600 copies/mL (S39), 30,400 copies/mL (S43) and 13,000 copies/mL (S30). The TapeStation gel and electropherogram are presented in Figure 4-20 and Figure 4-21, respectively. A clear peak was observed in all samples between 100-300 bp, with a mean fragment size of 150-160 bp. In two of the samples (S30 and S43), a broader pattern was also observed starting from 400-500 bp and extending beyond the upper marker at 1,500 bp.

Figure 4-20: TapeStation gel of three extracted plasma cfDNA samples (S39, S43, S30) against a High Sensitivity D1000 Ladder (Lad, sizes in bp marked). Arrow heads (◄) indicate peaks.

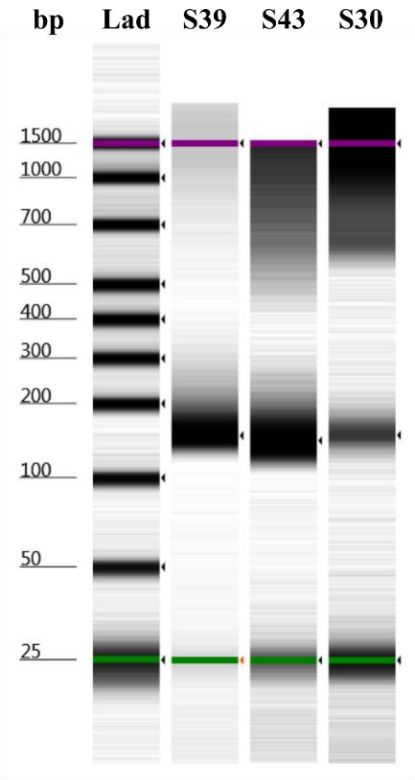
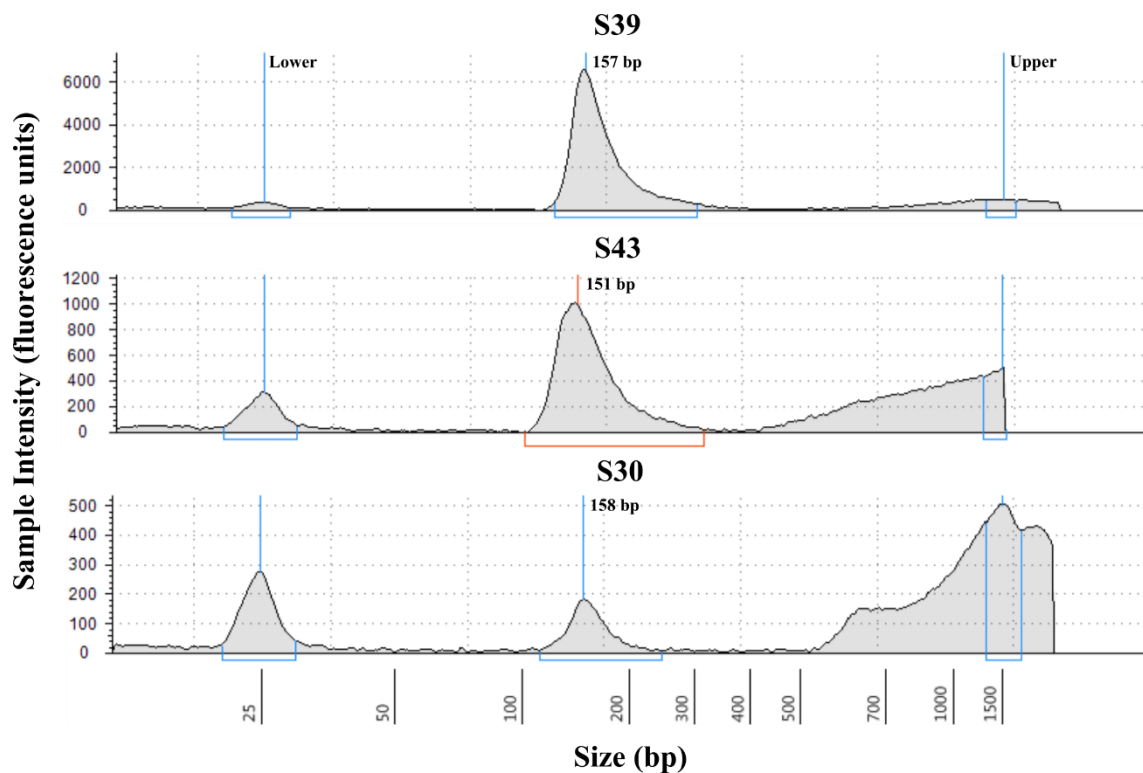


Figure 4-21: TapeStation electropherograms demonstrating fragmentation profiles of three plasma cfDNA samples (S39, S43, S30) within lower and upper boundary markers (marked). The approximate mean fragment size of any peaks is specified.



4.3.3.2 Fragmentation of Cellular DNA

Trial and error was used to determine the optimal ultrasonication protocol for one-copy and zero-copy genotype cellular DNA targeting a peak fragment size of 150-160 bp. Compared with recommended defaults, treatment time for the 130 μ L samples was extended from 430 seconds to between 820-860 seconds. This treatment duration was associated with a rise in the water temperature to approximately 7.3 $^{\circ}$ C. To permit completion of the entire treatment time in one application rather than requiring multiple bursts of treatment separated by periods of cooling, the default maximum water temperature of 7 $^{\circ}$ C was increased to 8 $^{\circ}$ C.

Achieved fragmentation profiles for cellular DNA from one-copy genotype and zero-copy genotype participants are presented in Figure 4-22 and Figure 4-23, respectively. Cellular DNA from participant 148 was used as both a one-copy and zero-copy genotype but underwent fragmentation once, the characteristics of which are displayed in Figure 4-22. All cellular DNA samples were ultrasonicated to achieve an approximate mean peak fragmentation size of 147-164 bp. Peak width was also similar between cellular DNA samples, ranging from 50 bp to 350 bp.

Figure 4-22: TapeStation electropherograms demonstrating fragmentation characteristics of ultrasonicated cellular DNA from one-copy genotype participants 71, 146 and 148 within lower

and upper boundary markers (marked). The approximate mean fragment size of any peaks is specified.

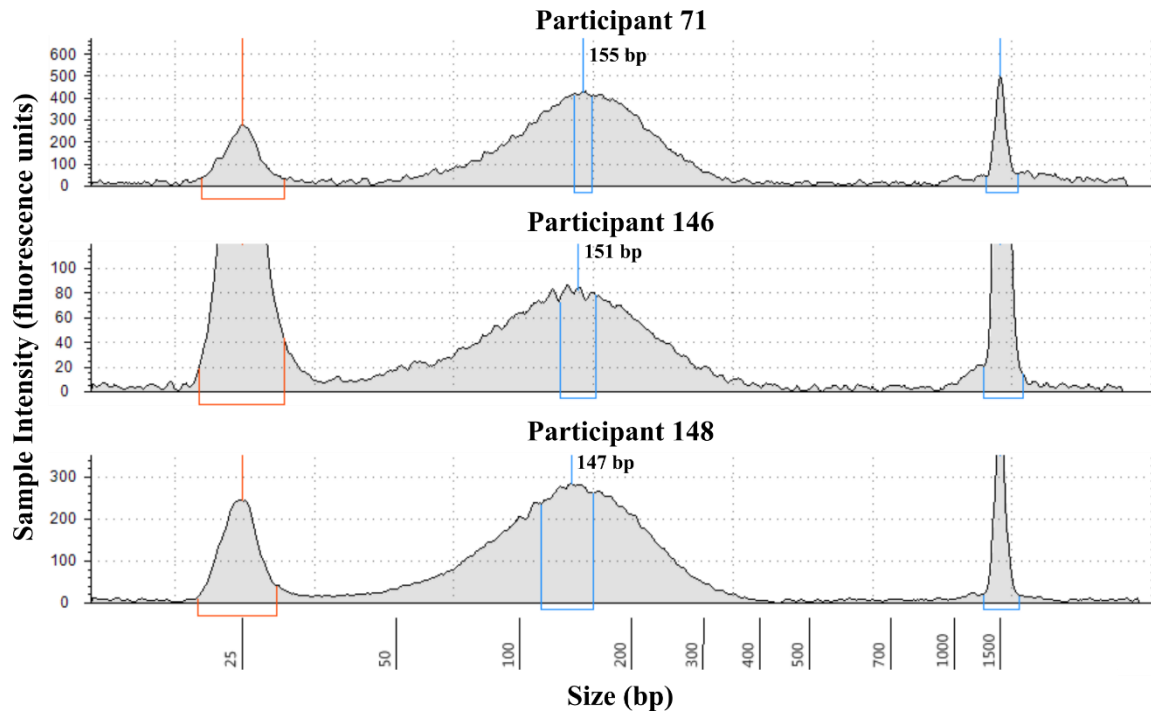
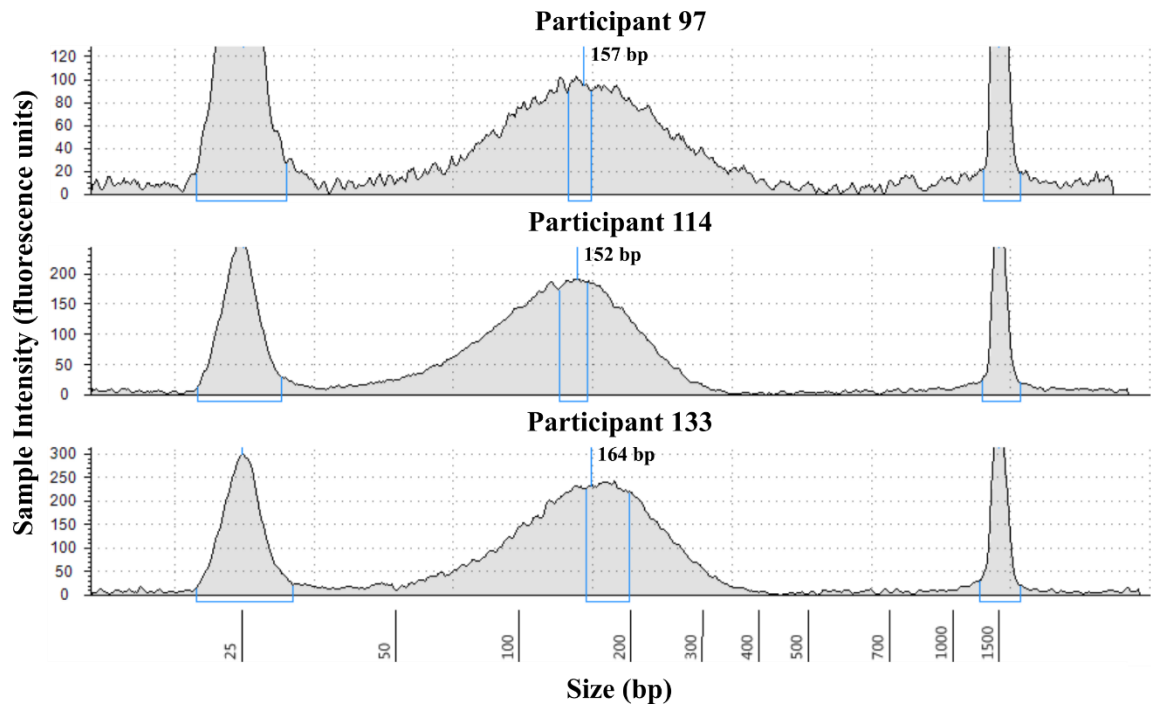


Figure 4-23: TapeStation electropherograms demonstrating fragmentation characteristics of ultrasonicated cellular DNA from zero-copy genotype participants 97, 114 and 133 within lower and upper boundary markers (marked). The approximate mean fragment size of any peaks is specified.



4.3.4 **Experimental Design**

4.3.4.1 *Assay Modelling*

4.3.4.1.1 Model Development

The ddPCR quantification platform provides a measure of the concentration of target sequence within the PCR reaction volume. In the case of clinical samples undergoing cfDNA quantification, the concentration of target within the PCR reaction volume must be converted to the concentration of target within the sample volume. This is desirable as it permits variations in sample volume undergoing extraction to be reflected in the test result. Such variations may arise from (i) patient factors such as haematocrit (affected by hydration status, anaemia or polycythaemia) or low urine output (affected by kidney function, hydration status, voiding function or timing of sample collection), and (ii) collection factors such as ease of venesection, sample loss, need for concurrent testing on the same sample, or limitations related to the volume of sample that may be extracted.

The concentration of target within the PCR reaction volume C_{PCR} in copies/ μL is a function of the number of copies of target within the PCR reaction T_{PCR} and the volume of the PCR reaction V_{PCR} in μL .

$$C_{PCR} = \frac{T_{PCR}}{V_{PCR}} \quad 4-6$$

V_{PCR} is the sum of the volume of template used for the reaction $V_{template}$ in μL and the volume of other reaction components V_{MM} in μL . These other reaction components comprise the volume of ddPCR Supermix, primers, probes and distilled water.

$$V_{PCR} = V_{template} + V_{MM} \quad 4-7$$

T_{PCR} is a product of the concentration of target within the eluent of the cfDNA extraction $C_{elution}$ in copies/ μL and the $V_{template}$.

$$T_{PCR} = C_{elution} \times V_{template} \quad 4-8$$

$C_{elution}$ is a quotient of the number of copies of target within the eluent of the cfDNA extraction $T_{elution}$ and the elution volume $V_{elution}$ in μL . $V_{elution}$ is specified by the cfDNA extraction methodology.

$$C_{elution} = \frac{T_{elution}}{V_{elution}} \quad 4-9$$

$T_{elution}$ is a product of the concentration of target within the clinical sample undergoing cfDNA extraction C_{sample} in copies/mL, the volume of sample used for cfDNA extraction V_{sample} in mL

and an extraction yield variable k . This recognises that cfDNA extraction is presumably not complete and may vary depending on concentration and extraction methodology. The value of k is not known and is assumed to be consistent between clinical samples subject to extraction. For this reason, a value for k is taken as 1 in calculations for practical purposes.

$$T_{elution} = C_{sample} \times V_{sample} \times k \quad 4-10$$

Equations 4-6 to 4-10 can be combined yielding

$$C_{PCR} = \frac{C_{sample} \times V_{sample} \times V_{template} \times k}{V_{elution} \times (V_{template} + V_{MM})} \quad 4-11$$

which can be rearranged for C_{sample}

$$C_{sample} = \frac{C_{PCR} \times V_{elution} \times (V_{template} + V_{MM})}{V_{sample} \times V_{template} \times k} \quad 4-12$$

where C_{sample} is the concentration of target within the collected sample in copies/mL, C_{PCR} is the concentration of target within the PCR reaction in copies/ μ L as measured by the quantification methodology, $V_{elution}$ is the volume of eluent used during the cfDNA extraction process in μ L, $V_{template}$ is the volume of extracted template cfDNA used in the PCR reaction in μ L, V_{MM} is the volume of all other PCR reaction components in μ L, V_{sample} is the volume of sample used for cfDNA extraction in mL, and k is an unquantified extraction yield variable. The volume of PCR reaction components V_{MM} includes the summed volumes of distilled water, primers, probes and ddPCR Supermix.

The function by which the concentration of target is determined from the proportion of total droplets that are negative is

$$C_{PCR} = \frac{-\ln\left(\frac{N_{neg}}{N_{total}}\right)}{V_{droplet}} \quad 4-13$$

where C_{PCR} is the concentration of target in the PCR reaction in copies/ μ L, N_{neg} is the number of negative droplets, N_{total} is the total number of droplets, and $V_{droplet}$ is the volume of each droplet in μ L.[253]

Solving Equation 4-13 for N_{neg} yields

$$N_{neg} = e^{(\ln(N_{total}) - (C_{PCR} \times V_{droplet}))}. \quad 4-14$$

Given that $N_{total} = N_{pos} + N_{neg}$, it follows that

$$N_{pos} = N_{total} - (e^{(\ln(N_{total}) - (C_{PCR} \times V_{droplet}))}) \quad 4-15$$

where N_{pos} is the number of positive droplets.

Equation 4-11 can be substituted into Equation 4-15 to create a function by which variations in experimental variables on positive droplet numbers can be modelled,

$$N_{pos} = N_{total} - (e^{(\ln(N_{total}) - (\frac{C_{sample} \times V_{sample} \times V_{template} \times k}{V_{elution} \times (V_{template} + V_{MM})}) \times V_{droplet}))}) \quad 4-16$$

where N_{pos} is the number of positive droplets in the ddPCR well, N_{total} is the number of total droplets in the ddPCR well, C_{sample} is the concentration of target within the collected sample in copies/mL, V_{sample} is the volume of sample used for cfDNA extraction in mL, $V_{template}$ is the volume of extracted template cfDNA used in the PCR reaction in μ L, $V_{elution}$ is the volume of eluent used during the cfDNA extraction process in μ L, V_{MM} is the volume of all other PCR reaction components in μ L, and k is an unquantified extraction yield variable. The volume of PCR reaction components V_{MM} includes the summed volumes of distilled water, primers, probes and ddPCR Supermix.

4.3.4.1.2 Modelling Sample and Template Volume

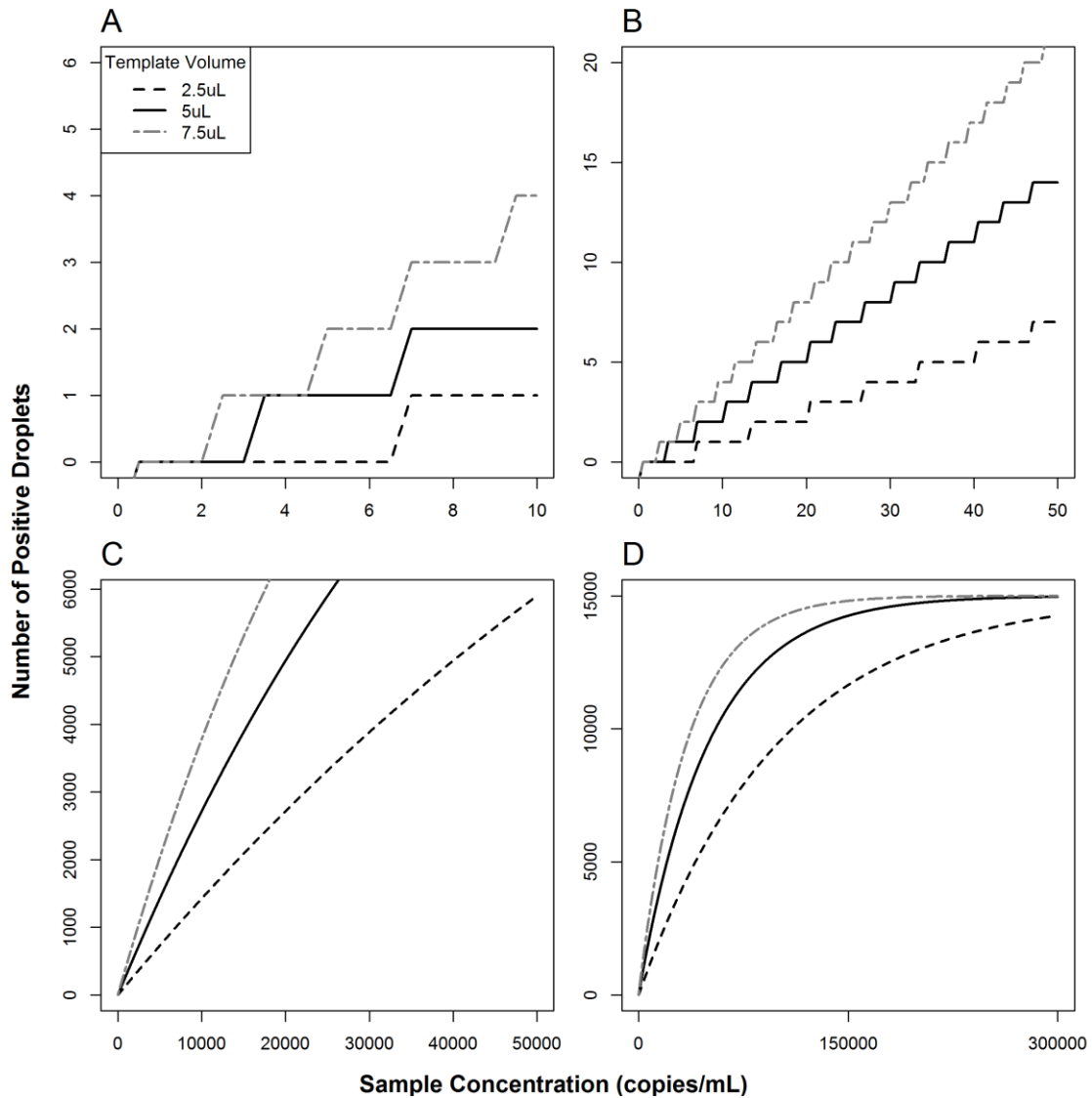
The number of expected positive droplets as a function of the sample concentration for 2.5 μ L, 5 μ L and 7.5 μ L template volume models is graphed in Figure 4-24.

Higher template volumes were associated with improved sensitivity for very low sample target concentrations (Figure 4-24A). Theoretical minimum sample concentrations detectable as 1 positive droplet were 6.7 copies/mL, 3.4 copies/mL, and 2.3 copies/mL using template volumes of 2.5 μ L, 5 μ L, and 7.5 μ L, respectively.

Larger template volumes were also associated with a more rapid increase in positive droplet numbers. A template volume of 7.5 μ L incremented 1 positive droplet for every 2.3 copies/mL increase in sample concentration of target, whereas a template volume of 2.5 μ L incremented 1 positive droplet for every 6.7 copies/mL increase in sample concentration of target. Whilst the relationship between sample concentration of target and positive droplet number was essentially linear in the lower concentration range (Figure 4-24B), linearity was lost at higher sample target concentrations, particularly as saturation of total droplet number is approached (Figure 4-24C and D).

Figure 4-24: Modelled number of expected positive droplets from an accepted droplet number of 15,000 as a function of sample concentration using a variety of template DNA volumes across the

sample concentration range 0-10 copies/mL (A), 0-50 copies/mL (B), 0-50,000 copies/mL (C), and 0-300,000 copies/mL (D).

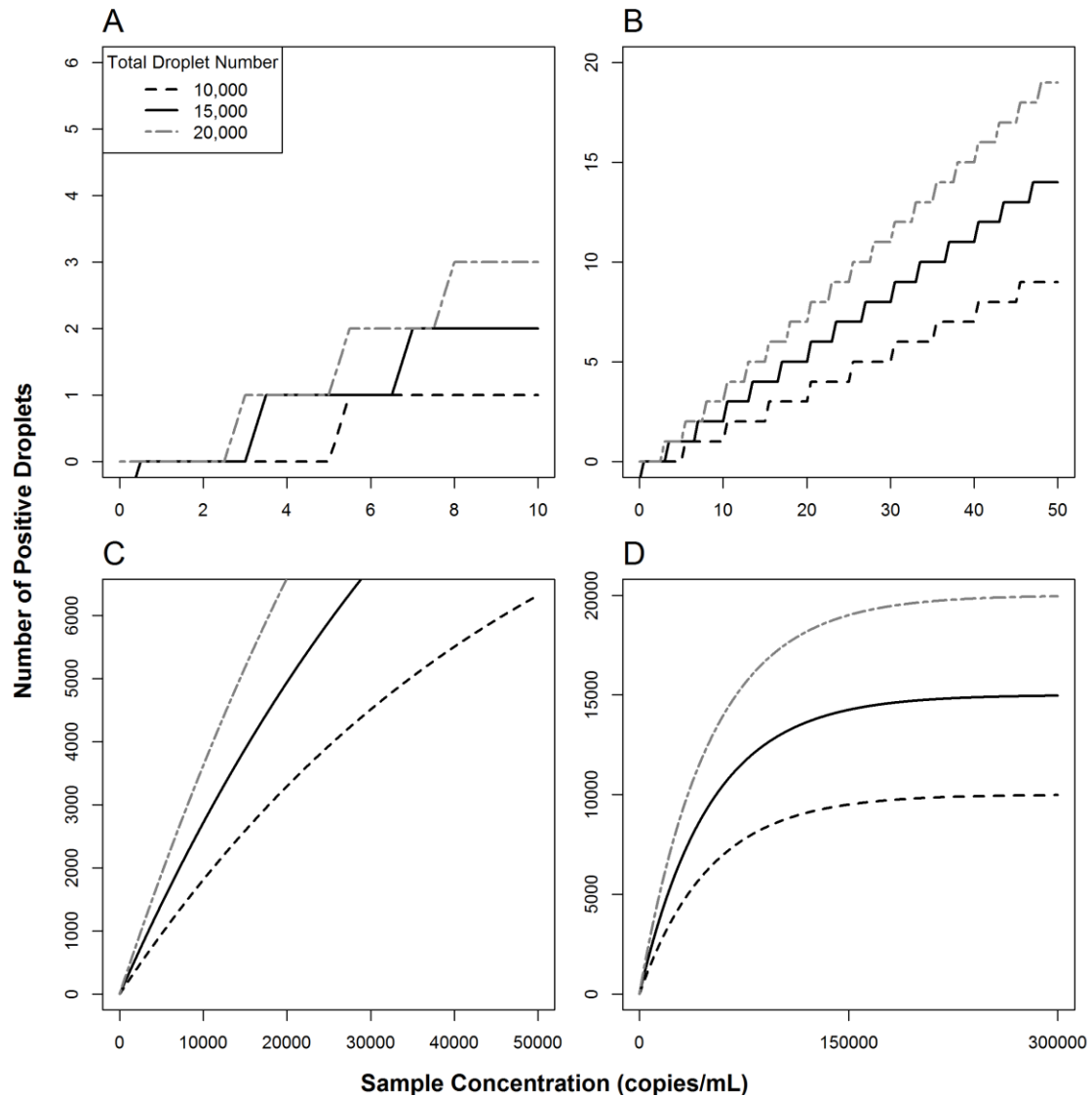


4.3.4.1.3 Modelling Total Droplet Number

The number of expected positive droplets as a function of the sample concentration for 10K, 15K and 20K total droplet number models is graphed in Figure 4-25.

Higher total droplet numbers were associated with improved sensitivity for very low sample target concentrations (Figure 4-25A). The minimum sample concentration theoretically detectable as 1 positive droplet was 5.0 copies/mL, 3.4 copies/mL, and 2.5 copies/mL for 10K, 15K and 20K total droplet number respectively. These minimum detectable sample concentrations were also the gradient of the linear component of the model (Figure 4-25B), which was progressively lost at higher concentrations, particularly >5,000 copies/mL (Figure 4-25C and Figure 4-25D).

Figure 4-25: Modelled number of expected positive droplets as a function of sample concentration using a variety of total droplet numbers across the sample concentration range 0-10 copies/mL (A), 0-50 copies/mL (B), 0-50,000 copies/mL (C), and 0-300,000 copies/mL (D).



4.3.4.2 Dilution Series Design

A two-fold dilution series comprising a starting concentration of 128 copies/ μ L and 7 dilution steps was designed (Table 4-12). These concentrations correspond to modelled equivalent sample concentrations of 10-1,280 copies/ μ L and 3-379 positive droplet droplets out of 15,000 total droplets. The limiting factor for the number of dilution steps was DNA quantity, particularly for genotypes with the most assay pair coverage.

In addition to the dilution series, each assay-pair replicate was designed to run a zero-copy control and no template control (water). The zero-copy control would be added at a template concentration of approximately 128 copies/ μ L, defined by one-copy markers for that genotype.

Four replicates of the proposed dilution series, zero-copy negative control and no template negative control were designed for each assay pair. All dilution series, no template negative control, and zero-copy negative control replicates for two assay pairs fit on one digital PCR 96-well plate (Figure 4-3).

Table 4-12: Modelled equivalent sample concentrations and positive droplet number for dilution series PCR and template concentrations. Each step of the dilution series targets a one-copy assay pair. The no template control and the zero-copy control for the corresponding assay pair should be zero-copy.

Dilution Step	Copy Number of Target	Positive Droplets	PCR Concentration (copies/ μ L)	Template DNA Concentration (copies/ μ L)	Sample Concentration (copies/mL)
0	1	379	25.6	128	1280
1	1	191	12.8	64	640
2	1	96	6.4	32	320
3	1	48	3.2	16	160
4	1	24	1.6	8	80
5	1	12	0.8	4	40
6	1	6	0.4	2	20
7	1	3	0.2	1	10
NTC (water)	0	0	0	0	0
Zero-Copy Control	0	0	0	0	0
	(1)	(379)	(25.6)	(128)	(1280)

4.3.4.3 Fragmented Cellular DNA Quantification

Multiple markers were used to measure the concentration of one-copy targets within diluted fragmented cellular DNA samples for each genotype (Table 4-13). One exception to this was PID 114 from validation study 1, in which zero-copy markers were incorrectly used for one-copy quantification, precluding a concentration measurement using this approach. As a surrogate, the *ACE* non-polymorphic assay concentration (a two-copy target) was halved and used to provide an estimate of concentration. The quantification experiment was not repeated using correct one-copy markers prior to validation study 1 due to concerns regarding the availability of sufficient DNA for validation study 3 and a repeat of validation study 1 if this was required. On completion of validation study 1, correct quantification was obtained prior to validation study 3. The equivalent neat fragmented cellular DNA concentration observed using 7 one-copy markers was 1,540 copies/ μ L compared with the half-*ACE* concentration of 1,860 copies/ μ L.

Table 4-13: Digital PCR quantification of one-copy markers in diluted fragmented cellular DNA and the equivalent neat concentration. Neat fragmented cellular DNA concentration in copies/ μ L was converted to the equivalent concentration in ng/ μ L using Equation 4-1.

PID	Number of One-Copy Markers Measured	Measured One-Copy Concentration of Diluted Fragmented Cellular DNA (copies/ μ L)	Dilution Factor	Equivalent Concentration of Neat Fragmented Cellular DNA (copies/ μ L)	Equivalent Concentration of Neat Fragmented Cellular DNA (ng/ μ L)
Validation Study 1					
148	7	220	0.8333	2,640	34.8
114	0 (used ACE)	155 (ACE 310 / 2)	0.8333	1,860	24.6
Validation Study 2					
146	6	374	0.0990	3,779	49.9
148	5*	196	0.0975	2,006	26.5
133	7	347	0.1727	2,011	26.5
Validation Study 3					
71	15	94.5	0.0238	3,970	52.4
97	8	302	0.0880	3,431	45.3
114	7	305	0.1980	1,540	20.3

* One outlier marker was excluded for the calculation of the one-copy average for this PID. CNV 34A measured concentration at 82.5 copies/ μ L and was excluded.

4.3.4.4 Dilution Series Creation

In validation study 1, there was sufficient volume of diluted fragmented cellular DNA that the average one-copy marker concentration measured by ddPCR could be used to guide further dilution of these samples to reach the starting concentration of approximately 128 copies/ μ L for the dilution series.

In validation studies 2 and 3, the large number of assay pairs necessitated a larger volume of diluted fragmented cellular DNA than was available from the solution quantified using ddPCR. Consequently, the calculated neat fragmented DNA concentration was used to guide creation of starting solution approximating 128 copies/ μ L using neat fragmented DNA and distilled water.

From these starting solutions, dilution series were created as described.

4.3.4.5 PCR Plate Setup

Validation studies 1, 2 and 3 were completed using 2, 4 and 4 plates respectively (Table 4-14). In one case (validation study 2, plate C), two different zero-copy controls were used within a single plate. In all other plates, only one zero-copy control was needed based on selected assay pairs. In validation study 2, plate D, the assay pair ACE/28A was duplicated to fill the plate (in other words, this assay pair had 8 replicates).

Table 4-14: Distribution of assay pairs between digital PCR plates for each validation study and the corresponding zero-copy controls for each plate.

Plate	Assay Pairs	Zero-Copy Control PID
Validation Study 1 (PID 148)		
A	31A/01B, 25A/30A	114
B	03B/20A, ACE/34A	114
Validation Study 2 (PID 146)		
A	29A/18A, 38A/32A	148
B	36A/26A, 10C/39A	148
C	14A/41A, 05B/24A	148, 133
D	ACE/28A, ACE/28A	133
Validation Study 3 (PID 71)		
A	23A/17A, 21A/06B	97
B	12A/37A, 27A/08B	97
C	19B/13A, 02B/22A	114
D	33A/15A, ACE/04B	114

4.3.5 Accepted Droplets

A total of 800 digital PCR reaction wells underwent droplet generation. The median number of accepted droplets per well was 16,076 (14,562-17,635) (Figure 4-26). The lower outlier threshold was 9,953 droplets (rounded up to 10,000 droplets).

4/800 (0.5%) wells had <10,000 accepted droplets (Table 4-15). Two of the low droplet number wells were adjacent one another on plate 3C and assayed CNV pair 19B/13A. Another two wells were non-contiguous but shared plate 1A and included CNV assay 31A/01B.

The number of accepted droplets differed between validation studies ($\chi^2(2, N = 800) = 138, p < 0.0001$). There were fewer accepted droplets in validation study 1 (median 14,927 (13,481-16,196)) compared with validation studies 2 and 3 (median 16,318 (14,949-17,761) versus 16,653 (14,825-18,371), respectively, $p < 0.0001$) (Figure 4-27). There was no difference between validation studies 2 and 3 ($p = 0.06$).

There were multi-level nested group effects on accepted droplet numbers within each validation study (null deviance 14,524) including plate effects (delta-AICc = 524, deviance 13,998, $\chi^2(1) = 526, p < 0.0001$) and assay pair effects (delta-AICc = 24, deviance = 13,972, $\chi^2(1) = 26, p < 0.0001$) (Figure 4-28).

Figure 4-26: Histogram of accepted droplet number for all validation study digital PCR reaction wells.

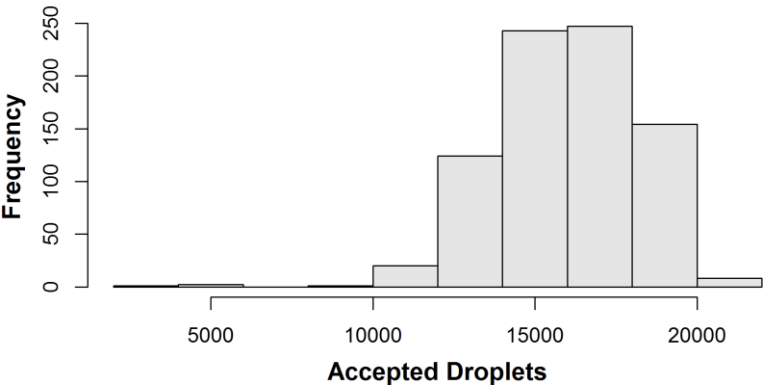


Table 4-15: Digital PCR reaction wells with <10,000 accepted droplets.

Validation	Plate	Well	PID	Dilution	Assay Pair	Accepted Droplets
1	A	A08	148	1	31A/01B	9921
1	A	D01	148	128	31A/01B	2687
3	C	C08	71	1	19B/13A	5797
3	C	C09	Water	0	19B/13A	4345

Figure 4-27: Distribution of accepted droplets compared between validation studies.

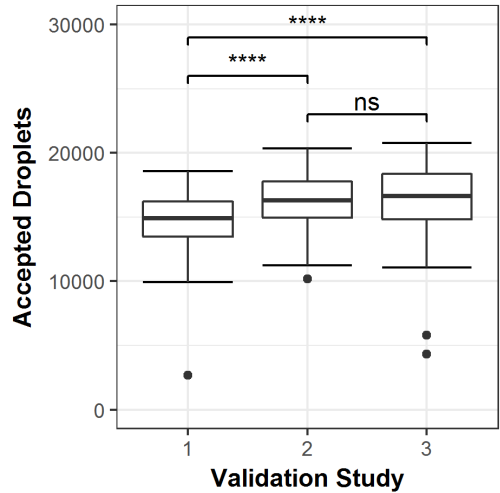
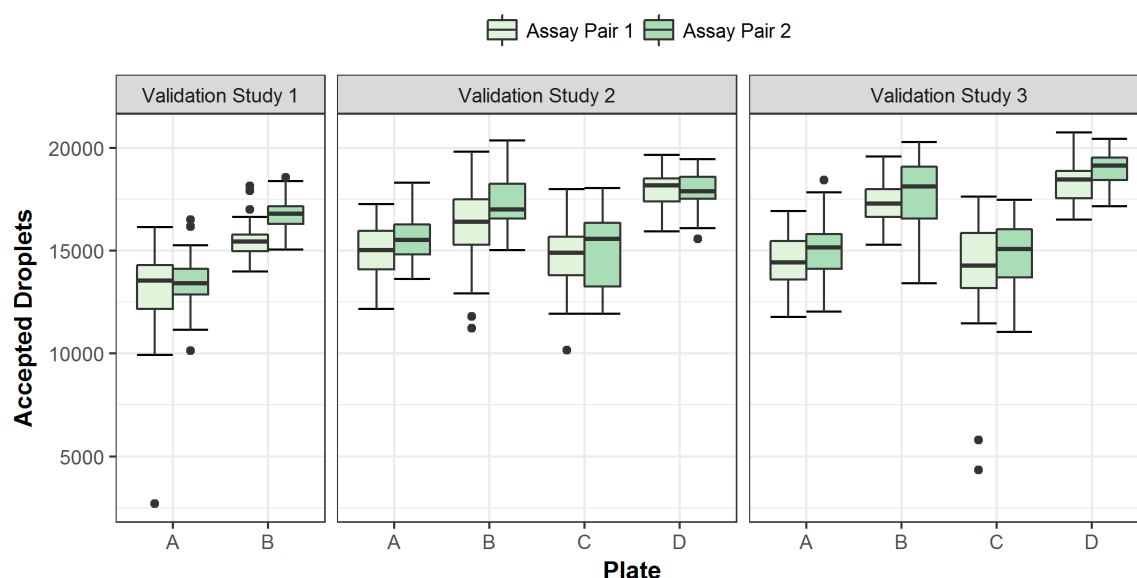


Figure 4-28: Distribution of accepted droplets compared between assay pairs within each plate of the validation studies.



4.3.6 Droplet Cluster Centroids

4.3.6.1 Cluster Centroid Identification

Within each assay pair, clustering of positive and negative droplets appeared to be consistent between wells, forming clear positive and negative clusters on *en bloc* assay pair cluster centroid identification (Figure 4-5 to Figure 4-9). Approximate droplet cluster centroid channel 1 and channel 2 amplitudes were obtained for all 19 assay pairs, comprising 35 CNV and 3 *ACE* assays (Table 4-16).

Channel 1 and channel 2 amplitudes for double negative, FAM-positive and HEX-positive clusters varied minimally when compared between two groups of four replicates of *ACE*/28A that had undergone separate cluster centroid identification. The median amplitude coefficient of variation was 0.86% (0.55-1.67%).

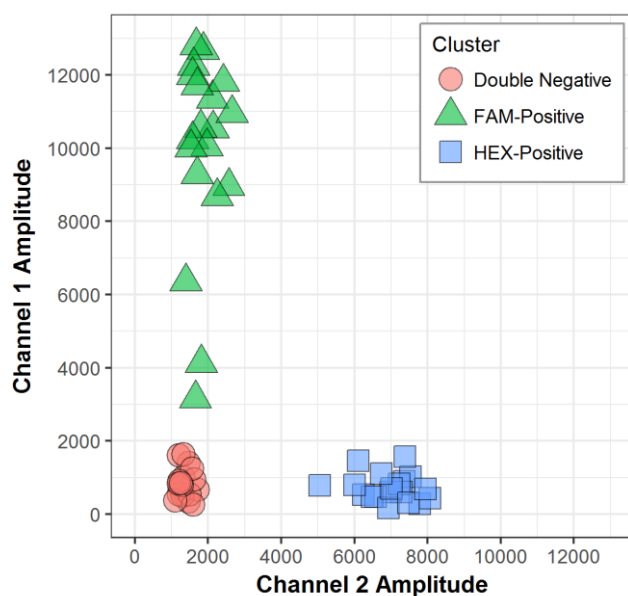
Clustering of double negative, FAM-positive and HEX-positive cluster centroids was observed (Figure 4-29).

Table 4-16: Channel 1 and channel 2 amplitudes for double negative, FAM-positive and HEX-positive droplet cluster centroids for each assay pair within the validation study.

Plate	Assays		Double Negative		FAM-Positive		HEX-Positive	
	FAM	HEX	Channel 1	Channel 2	Channel 1	Channel 2	Channel 1	Channel 2
Validation Study 1								
A	31A	01B	261	1593	3155	1654	169	6933
A	25A	30A	552	1197	10236	1572	479	6488
B	03B	20A	1010	1328	12678	1871	899	7368
B	<i>ACE</i>	34A	852	1248	11710	1703	788	5059

Validation Study 2								
A	29A	18A	956	1618	10043	1963	827	7231
A	38A	32A	1251	1567	8679	2255	1103	6732
B	36A	26A	380	1100	6356	1392	303	7475
B	10C	39A	1618	1201	12801	1675	1569	7386
C	14A	41A	724	1221	10531	2139	602	7285
C	05B	24A	1397	1463	11351	2119	1029	7537
D	ACE	28A	844	1192	12153	1575	706	7998
D	ACE	28A	847	1204	11833	1586	675	7885
Validation Study 3								
A	23A	17A	1634	1322	11809	2422	1462	6104
A	21A	06B	503	1290	10555	1805	479	6587
B	12A	37A	353	1465	4132	1815	280	7781
B	27A	08B	552	1487	10949	2650	429	8070
C	19B	13A	675	1710	8958	2565	602	6985
C	02B	22A	577	1282	9287	1699	534	6243
D	33A	15A	792	1281	10018	1534	706	7012
D	ACE	04B	906	1213	12250	1609	798	6001

Figure 4-29: Location of heat-map assisted, visually-defined approximate cluster centroids for 19 assay pairs.



4.3.6.2 Double Negative Cluster Centroid Analysis

Channel amplitudes for double negative cluster centroids were identified *en bloc* for (i) dilution series replicates (19 assay pairs), (ii) zero-copy control replicates (19 assay pairs) and (iii) no template control replicates (15 assay pairs) (Table 4-17). The no template control from validation

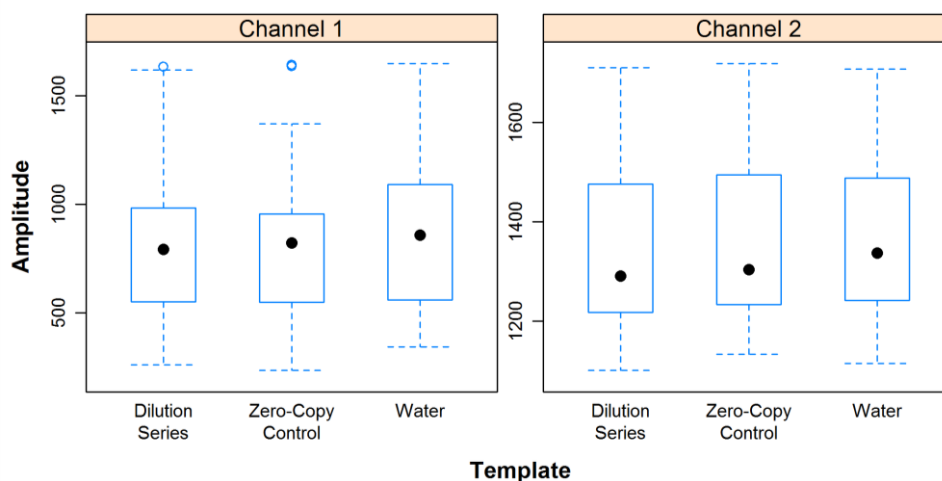
study 1 was excluded as it consisted of a further dilution step (~0.5 copies/μL) rather than distilled water as a result of laboratory error.

Median channel 1 and channel 2 amplitude values for dilution series double negative clusters was 818 (552-970) and 1286 (1211-1470). Across all assay pairs, there was no significant difference in channel amplitude between dilution series, zero-copy control and no template control double negative cluster centroids (channel 1 $p = 0.92$, channel 2 $p = 0.97$) (Figure 4-30). Additionally, within each assay pair, there was minimal variation in assay pair double negative cluster centroid amplitude between the groups. Median channel 1 and channel 2 coefficients of variation were 1.6% (0.98-2.8%) and 1.1% (0.82-1.5%) within each assay pair.

Table 4-17: Double negative cluster centroid channel 1 and channel 2 amplitudes for dilution series, zero-copy control and no template control replicates of each assay pair. The no template control from validation study 1 was excluded (NA, below) as it consisted of a further dilution step (~0.5 copies/μL) rather than distilled water as a result of human error.

Plate	Assays		Dilution Series		Zero-Copy Control		No Template		Coefficient of Variation (%)	
	FAM	HEX	Ch 1	Ch 2	Ch 1	Ch 2	Ch 1	Ch 2	Ch 1	Ch 2
Validation Study 1										
A	31A	01B	261	1593	235	1573	NA	NA	7.4	0.9
A	25A	30A	552	1197	527	1200	NA	NA	3.3	0.2
B	03B	20A	1010	1328	967	1295	NA	NA	3.1	1.8
B	ACE	34A	852	1248	823	1247	NA	NA	2.4	0.1
Validation Study 2										
A	29A	18A	956	1618	945	1604	976	1649	1.6	1.4
A	38A	32A	1251	1567	1191	1540	1205	1540	2.6	1.0
B	36A	26A	380	1100	372	1133	375	1114	1.1	1.5
B	10C	39A	1618	1201	1641	1219	1649	1219	1.0	0.9
C	14A	41A	724	1221	710	1233	718	1251	1.0	1.2
C	05B	24A	1397	1463	1371	1484	1391	1506	1.0	1.4
D	ACE	28A	844	1192	899	1233	897	1225	3.5	1.8
D	ACE	28A	847	1204	872	1256	844	1219	1.8	2.2
Validation Study 3										
A	23A	17A	1634	1322	1636	1336	1644	1345	0.3	0.9
A	21A	06B	503	1290	514	1312	499	1278	1.5	1.3
B	12A	37A	353	1465	352	1467	343	1450	1.6	0.6
B	27A	08B	552	1487	573	1504	543	1469	2.8	1.2
C	19B	13A	675	1710	667	1719	664	1707	0.9	0.4
C	02B	22A	577	1282	578	1276	577	1261	0.1	0.8
D	33A	15A	792	1281	833	1303	858	1337	4.0	2.2
D	ACE	04B	906	1213	899	1224	901	1231	0.4	0.7

Figure 4-30: Distribution of channel amplitudes for assay pair double negative cluster centroids comparing dilution series, zero-copy control and no template control replicates.



4.3.6.3 ACE Positive Cluster Centroid Analysis

ACE was present in 3 validation study assay pairs, and was the only marker run against three different dilution series. The ACE FAM-positive cluster had a median channel 1 amplitude of 11,993 (range 11,710-12,250) and a median channel 2 amplitude of 1,609 (range 1,580-1,703). The coefficient of variation amongst these measurements was 2.3% for channel 1 and 3.9% for channel 2 amplitudes.

4.3.6.4 CNV Assay Positive Cluster Centroid Analysis

19 CNV assay HEX-positive cluster centroids tended to be closer to the double negative cluster than 16 CNV assay FAM-positive clusters (Figure 4-31). The distribution of HEX-positive cluster channel 2 amplitudes (median 7,012, IQR 6,538-7,430) was lower than the distribution of FAM-positive cluster channel 1 amplitudes (median 10,140, IQR 8,888-11,050) ($p < 0.001$).

FAM-positive cluster centroids also tended to exhibit higher channel 2 amplitudes than double zero cluster centroids (Figure 4-31). The median FAM-positive cluster centroid channel 2 amplitude was 1,843 (1,670-2,168), higher than the median double negative cluster centroid channel 2 amplitude of 1290 (1,217-1,476) ($p < 0.0001$). A similar pattern was not seen with HEX-positive clusters in channel 1. The distribution of HEX-positive cluster channel 1 amplitudes (median 691, IQR 479-863) was no different from the distribution of FAM-positive cluster channel 1 amplitudes (median 792, IQR 552-983) ($p \approx 0.38$).

Three FAM assays and one HEX assay tended to have lower amplitude positive droplet cluster centroids (Table 4-18). Three (CNV 31A, 34A and 12A) had z-scores between -2.58 and -1.96 and/or had amplitudes lower than the positive cluster channel Q1-1.5xIQR. These outlier cluster centroids tended to have lower reverse primer melting temperatures (61.0 (IQR 60.1-61.0) vs 63.0

(IQR 61.8-63.3) °C, $W = 90$, $p \approx 0.02$) and smaller forward primer size (20.0 (19.5-20.0) vs 22.0 (IQR 21.0-23.0) bp, $W = 88$, $p \approx 0.02$), there was also a trend towards lower forward primer melting temperature (61.3 (IQR 61.2-61.5) vs 62.9 (IQR 61.5-63.2) °C, $W = 78$, $p \approx 0.08$). Due to rank tied values between the two groups, p -values are approximate. There were no other significant assay design differences between the outlier and non-outlier cluster centroid groups, nor were there known assay non-specificity issues involving these outlier cluster centroid assays.

Figure 4-31: Histograms of cluster centroid channel 1 (left column) and channel 2 (right column) amplitude. HEX-positive clusters (top right panel) tended to be closer to double negative clusters (bottom right panel) than FAM-positive clusters (middle left panel, lower left panel). Additionally, FAM-positive clusters appeared to have higher channel 2 amplitudes (middle right panel) than double negative clusters (bottom right panel). The converse was not seen for HEX-positive clusters (top left panel, bottom left panel).

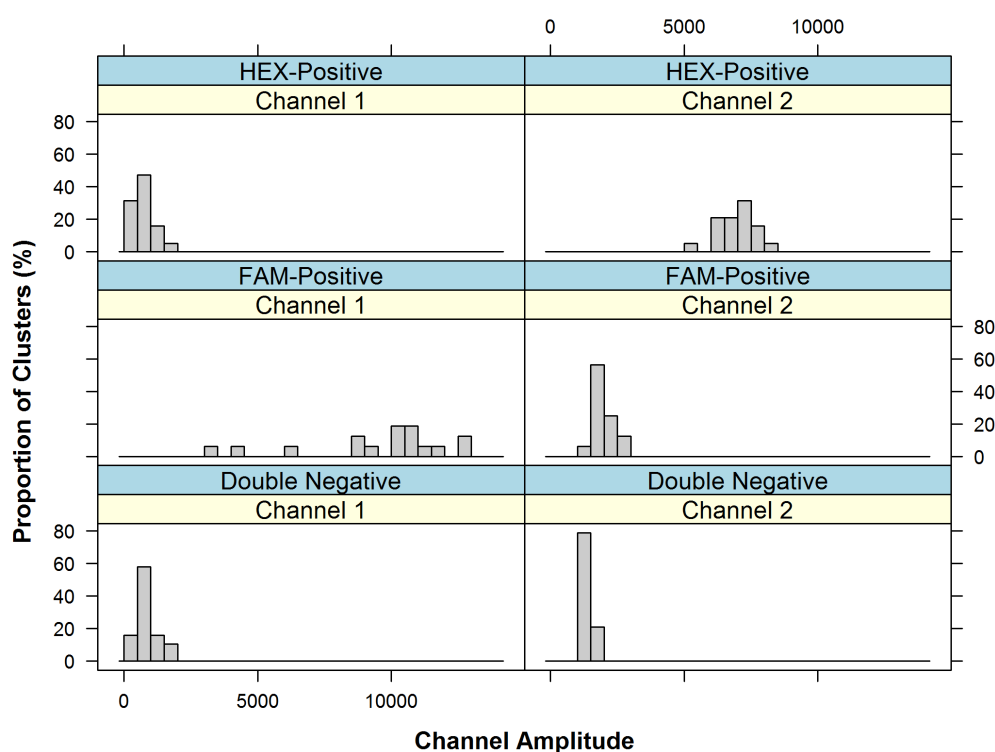
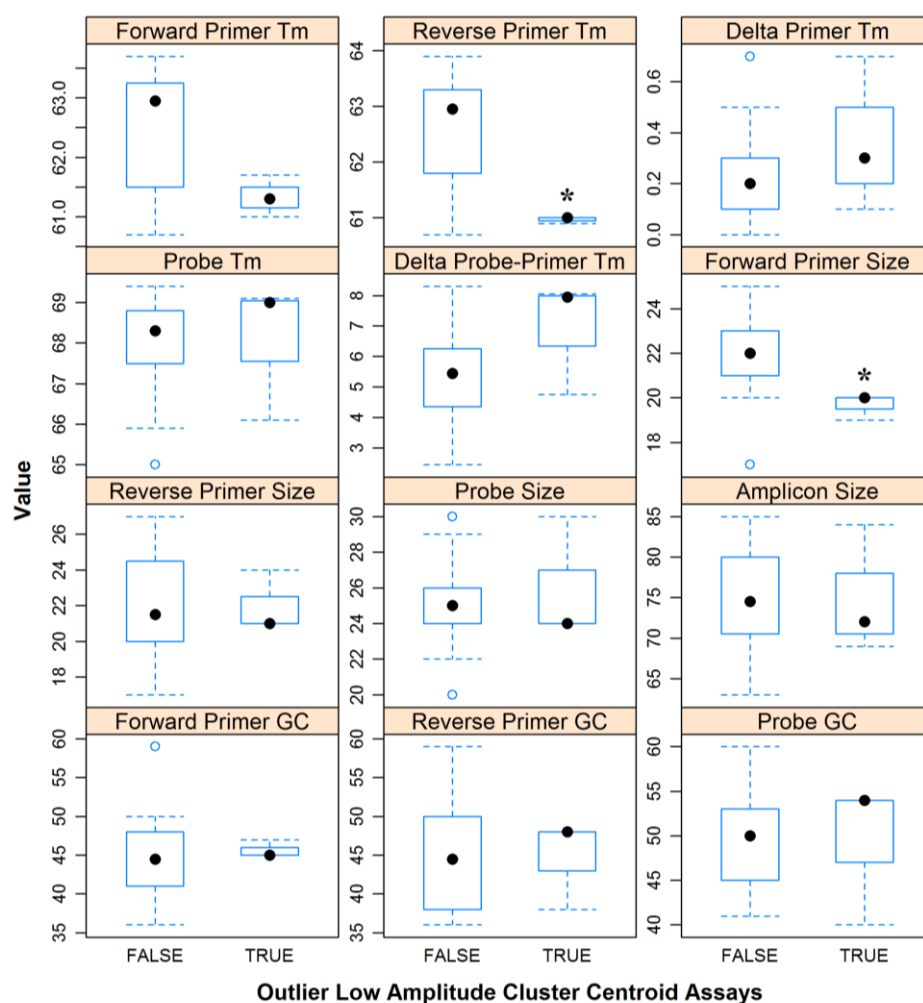


Table 4-18: Validation study assays with outlier lower amplitude positive cluster centroids.

Validation Study	Plate	Assay Pair	Assay	Cluster	Amplitude	Z-Score	IQR Test
1	A	1	31A	FAM-Positive	3155	-2.28	Fail
3	B	1	12A	FAM-Positive	4132	-1.93	Fail
2	B	1	36A	FAM-Positive	6356	-1.12	Pass
1	B	2	34A	HEX-Positive	5059	-2.52	Fail

Figure 4-32: Assay design characteristics compared between a group of assays with outlier lower amplitude positive cluster centroids and those with normal amplitude positive cluster centroids.

Melting temperatures in °C, size in bp and GC content in %. Statistical significance represented with asterisks if significant. Non-significant differences not marked.



4.3.7 Droplet Thresholding

Droplet thresholding was performed as described for all 19 assay pairs using the 5 threshold techniques described, totalling 180 thresholding steps per assay pair. Calculated threshold amplitudes for the half-, third- and quarter-distance techniques are summarised in Table 4-19.

Table 4-19: Threshold amplitudes for three thresholding techniques based upon fractional distances between the respective positive cluster and double negative cluster.

Plate	Assays		Half-Distance		Third-Distance		Quarter-Distance	
	FAM	HEX	Channel 1	Channel 2	Channel 1	Channel 2	Channel 1	Channel 2
Validation Study 1								
A	31A	01B	1708	4263	1225	3371	985	2928
A	25A	30A	5394	3843	3777	2959	2973	2520
B	03B	20A	6844	4348	4895	3339	3927	2838
B	ACE	34A	6281	3154	4468	2517	3567	2201

Validation Study 2								
A	29A	18A	5500	4425	3982	3487	3228	3021
A	38A	32A	4965	4150	3725	3287	3108	2858
B	36A	26A	3368	4288	2370	3223	1874	2694
B	10C	39A	7210	4294	5342	3261	4414	2747
C	14A	41A	5628	4253	3990	3240	3176	2737
C	05B	24A	6374	4500	4712	3486	3886	2982
D	ACE	28A	6499	4595	4610	3458	3671	2894
Validation Study 3								
A	23A	17A	6722	3713	5022	2914	4178	2518
A	21A	06B	5529	3939	3850	3054	3016	2614
B	12A	37A	2243	4623	1611	3568	1298	3044
B	27A	08B	5751	4779	4014	3679	3151	3133
C	19B	13A	4817	4348	3433	3467	2746	3029
C	02B	22A	4932	3763	3477	2934	2755	2522
D	33A	15A	5405	4147	3864	3189	3099	2714
D	ACE	04B	6578	3607	4684	2807	3742	2410

4.3.7.1 Droplet Rain

En bloc 2D views for each assay pair demonstrated numerous intermediate amplitude droplets extending between the positive clusters and double negative cluster (Figure 4-5). Each thresholding method variably dissected droplet rain into positive and negative droplets, ranging from complete inclusion (tight threshold technique) to near-complete exclusion (cluster threshold technique). Droplet rain complicated the cluster thresholding technique as cluster margins were less readily appreciable.

The extent of droplet rain and the impact of it on identification of cluster margins appeared to be related to the number of positive droplets plotted on the 2D plot, and the resolution of the plot display and corresponding plot droplet size. The number of positive droplets plotted was higher when target concentrations were higher (Figure 4-33) and more assay wells underwent thresholding *en bloc* (Figure 4-34).

Figure 4-33: 2D plots of CNV 29A/18A demonstrating the effect of droplet rain on cluster margin identification using the cluster threshold technique. The degree of droplet rain is less when using a single well (B) compared with 4 replicates *en bloc* (A) at approximately 128 copies/ μ L. It is

further reduced in single wells at 8 (C) and 2 (D) copies/ μ L concentration. Figure 4-5 presents this 2D plot when all replicates of the entire assay pair dilution series (32 wells) was selected.

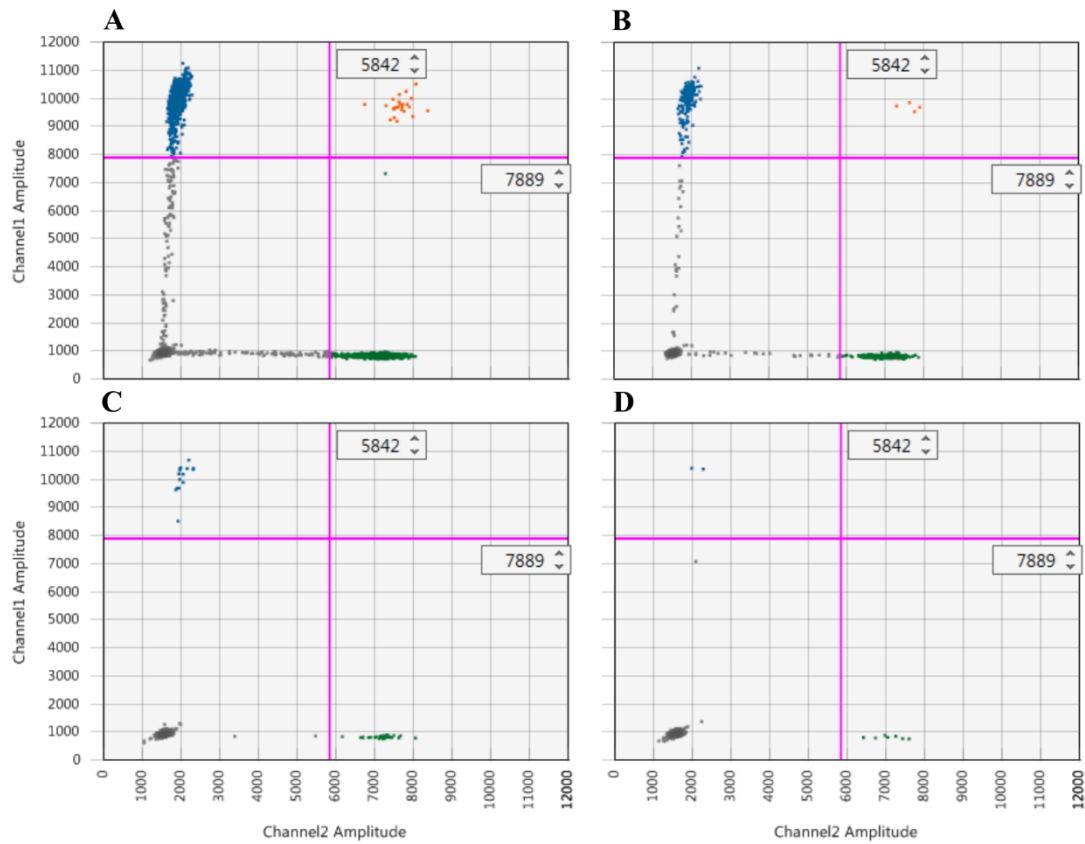
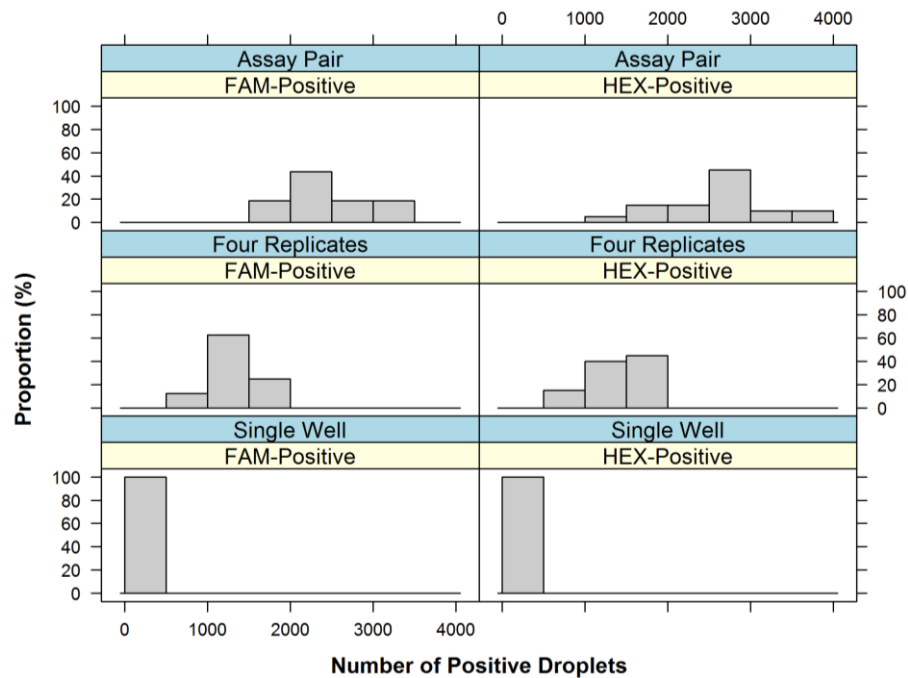


Figure 4-34: Histograms of the number of FAM- and HEX- positive droplets using the tight thresholding technique, grouped by the number of wells simultaneously selected for thresholding:

40 wells (assay pair), 4 wells (four replicates) and single wells. FAM-positive droplets targeting *ACE* were excluded from this analysis.



4.3.7.2 FAM-Positive Cluster Channel 2 Drift

One CNV assay pair (23A/17A) could not undergo quarter-distance thresholding as described due to transection of the FAM-positive cluster by the distance defined channel 2 threshold (Figure 4-35). In this instance, threshold cluster mode was used to manually lasso and correct the misclassified transected FAM-positive cluster. For an additional three assay pairs (*ACE*/34A, 19B/13A and 27A/08B), the quarter-distance threshold abutted the FAM-positive cluster but did not transect it. In these instances, quarter-distance thresholding could be achieved.

The proximity of the quarter-distance channel 2 threshold to the FAM-positive cluster was a function of the FAM-positive cluster centroid channel 2 amplitude and the HEX-positive cluster channel 2 amplitude (Figure 4-36). High FAM-positive cluster and low HEX-positive cluster channel 2 amplitudes were associated with proximity of the quarter-distance defined channel 2 threshold and the FAM-positive cluster. In the one circumstance where both of these conditions were met, the quarter-distance defined threshold transected the FAM-positive cluster.

Figure 4-35: 2D droplet plot for CNV assay pair 23A/17A selected *en bloc*, demonstrating how the quarter-distance threshold technique transects the FAM-positive cluster. This was attributed

to the relatively high FAM-positive cluster channel 2 amplitude and low HEX-positive cluster channel 2 amplitude.

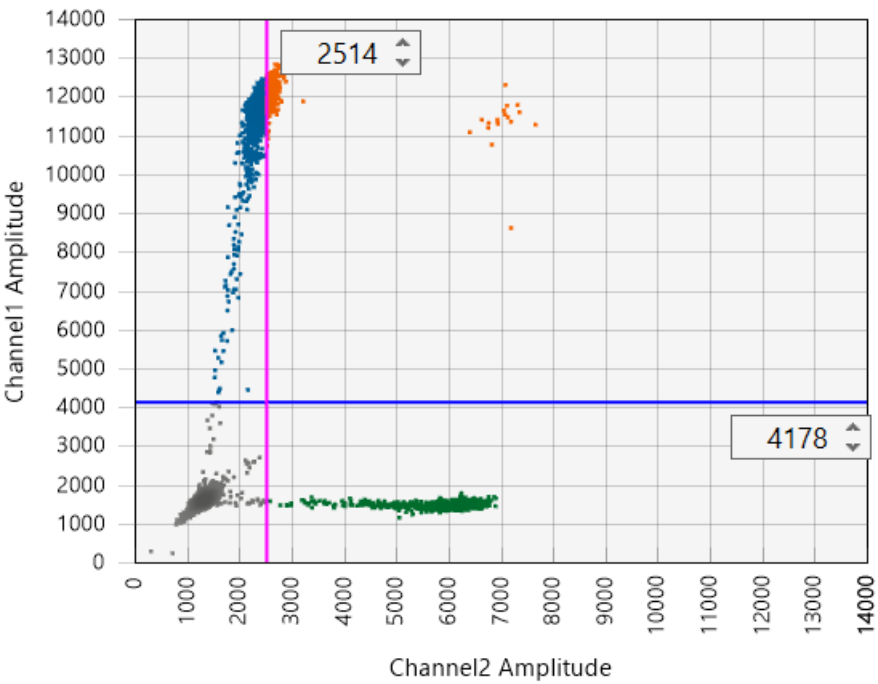
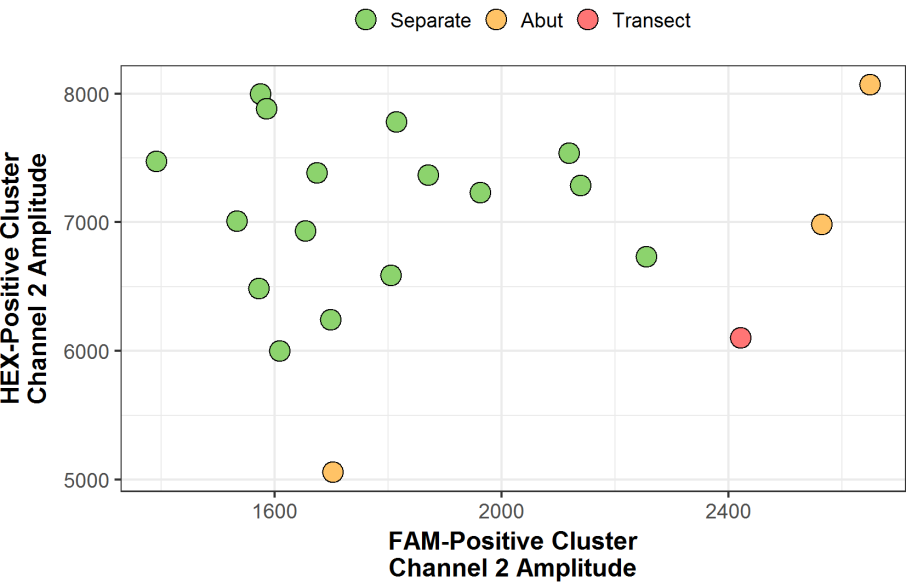


Figure 4-36: The relationship between HEX-positive cluster channel 2 amplitude, FAM-positive cluster channel 2 amplitude and proximity of the quarter-distance defined channel 2 threshold to the FAM-positive cluster (separate to it, abutting it or transecting it).



4.3.8 No Template Control

Due to a laboratory error, no template control wells from validation study 1 contained a further one-copy template dilution step and not distilled water, so were excluded from the analysis. There were 64 no template control wells in validation studies 2 and 3. For each well, there was a FAM result and HEX result. Each assay had 4 replicates with the exception of CNV 28A (8 replicates) and ACE (12 replicates). One well from validation study 3, containing CNV assay pair 19B/13A, was excluded due to an accepted droplet count of 4,345 (both assay results were true negative results). The total number of no template control results was therefore 126.

4.3.8.1 *Threshold Technique*

The false positive result rate varied according to the threshold technique used ($\chi^2(4, N = 5) = 120$, $p < 0.0001$) (Figure 4-37). Higher amplitude thresholding techniques (cluster, half-distance) were associated with a 0/126 (0%) false positive result rate. Lower amplitude thresholding techniques were associated with progressively higher false positive rates. Third-distance, quarter-distance and tight thresholding was associated with false positive result rates of 1/126 (0.8%), 11/126 (8.7%) and 39/126 (31%) respectively. There was also a weak positive monotonic relationship between lower threshold technique amplitudes and the concentration of false positive results ($r_s = 0.3$, $N = 51$, $p \approx 0.04$) (Figure 4-38).

8/29 (28%) assays had no false positive results irrespective of the threshold technique used. For the remaining 21 assays with at least one false positive replicate, most (18/21) of these occurred only when the tight threshold technique had been applied (Table 4-20). Using the third-distance amplitude threshold, only one assay replicate for ACE was false positive.

Figure 4-37: The false positive rate of 126 no template control wells grouped by positive droplet threshold technique. The false positive rate increased when lower droplet amplitude threshold techniques were used.

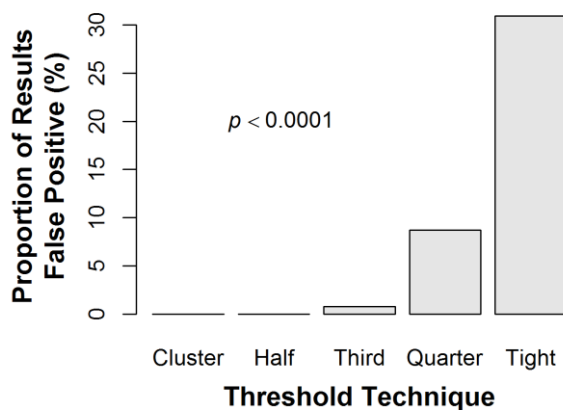


Figure 4-38: The concentration of no template control results grouped by thresholding technique. There was a weak positive monotonic relationship between lower threshold technique amplitudes and the concentration of the false positive results.

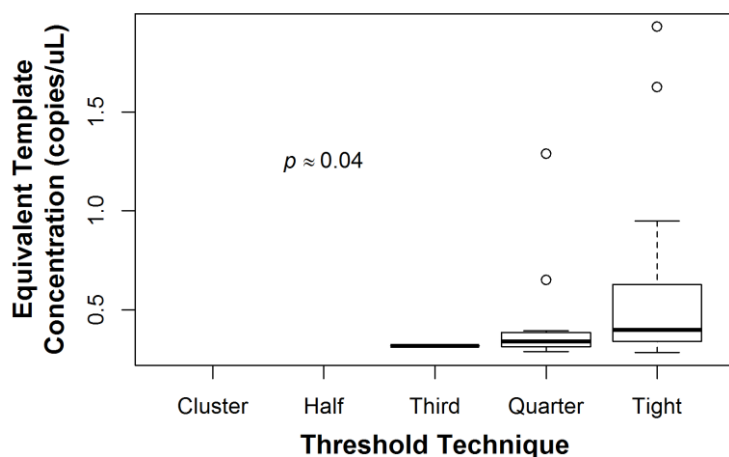


Table 4-20: Frequency of no template control false positive results by assay and threshold technique. Assays with no false positive results are not included.

Assay	Threshold Technique				
	Cluster	Half	Third	Quarter	Tight
ACE	0	0	1	9	12
15A	0	0	0	1	3
32A	0	0	0	1	2
05B	0	0	0	0	2
38A	0	0	0	0	2
28A	0	0	0	0	3
02B	0	0	0	0	1
04B	0	0	0	0	1
06B	0	0	0	0	1
08B	0	0	0	0	1
14A	0	0	0	0	1
17A	0	0	0	0	1
18A	0	0	0	0	1
19B	0	0	0	0	1
21A	0	0	0	0	1
22A	0	0	0	0	1
26A	0	0	0	0	1
27A	0	0	0	0	1
29A	0	0	0	0	1
33A	0	0	0	0	1
36A	0	0	0	0	1

4.3.8.2 Assay Design Characteristics

There was no significant difference in assay design parameters between assays with false positive results and assays without false positive results (Figure 4-39). Additionally, there was no significant difference between the groups in negative or positive cluster centroid amplitude, or the distance between the positive and negative cluster centroid (Figure 4-40).

Figure 4-39: Assay design characteristics compared between a group of assays with at least one no template control false positive result, and those with no false positive results. Melting temperatures in °C, size in bp and GC content in %. No distributions were significantly different between the groups.

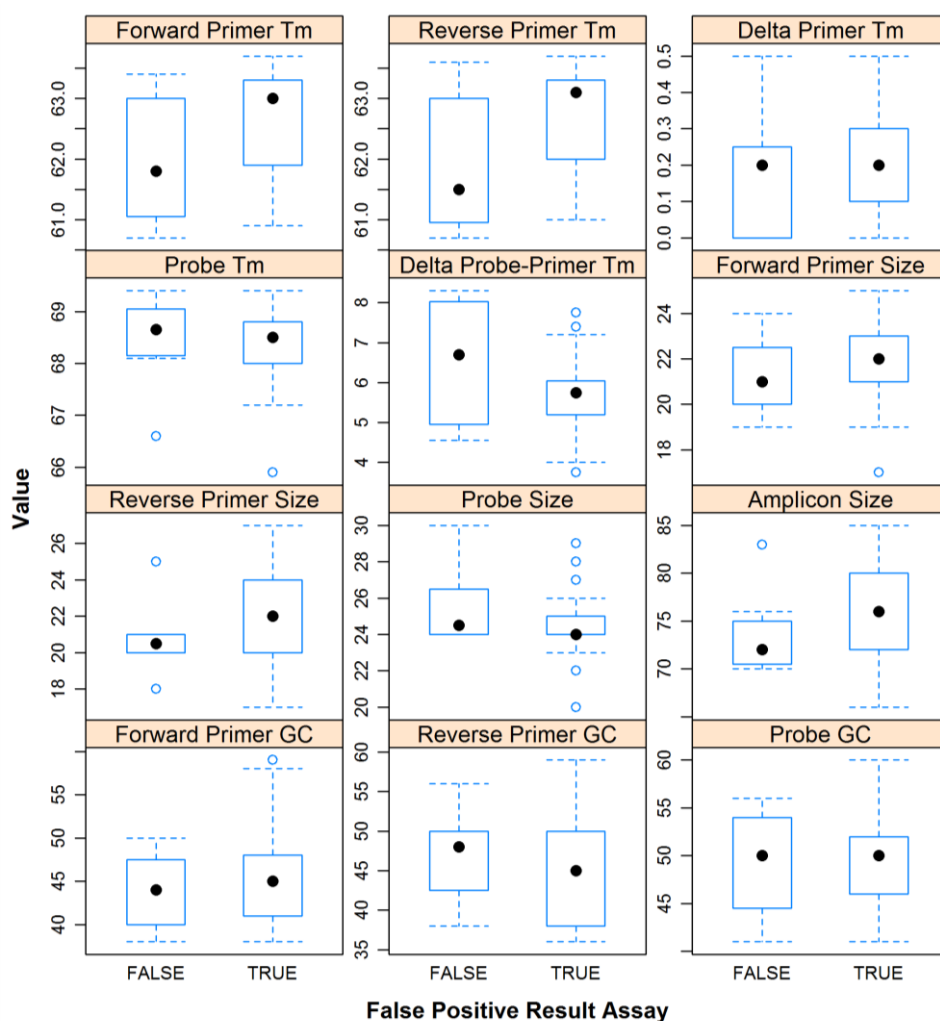
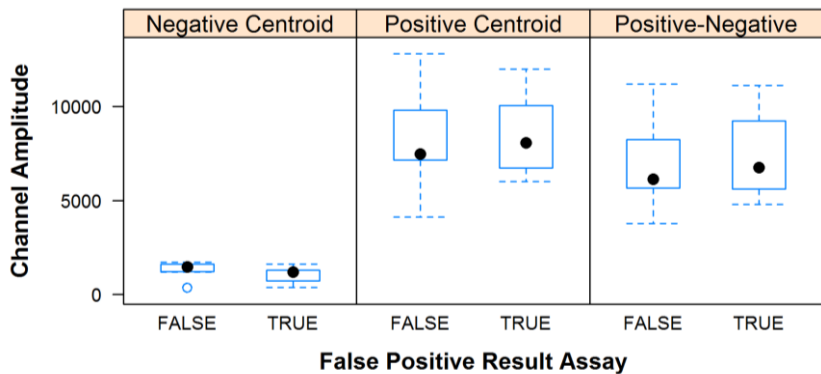


Figure 4-40: Centroid amplitudes and centroid separation (positive-negative) compared between a group of assays with note template control false positive results, and those with no false positive results. No distributions were significantly different between the groups.



4.3.8.3 Limit of Blank (No Template Control)

Using the third-distance threshold technique and no template control samples consisting of distilled water, the false positive rate was 1/126 (0.8%). The only false positive result was 1 of 12 *ACE* replicates, which had 1 false positive droplet, equivalent to a concentration of 0.3 copies/ μ L. Higher thresholds excluded this positive droplet, indicating that it had an amplitude between one third and half the distance between the negative and positive droplet clusters.

The limit of blank was therefore 0 copies/ μ L for the majority (28/35) of the assays except *ACE*, which had a limit of blank of 0.18 copies/ μ L (Table 4-21).

Table 4-21: Limit of blank parameters and results for assays run against no template control samples consisting of distilled water. The limit of blank was 0 copies/ μ L for all assays except *ACE*.

Assay	Number of Replicates	Mean (copies/ μ L)	Standard Deviation (copies/ μ L)	Limit of Blank (No Template Control, copies/ μ L)
02B	4	0	0	0
04B	4	0	0	0
05B	4	0	0	0
06B	4	0	0	0
08B	4	0	0	0
10C	4	0	0	0
12A	4	0	0	0
13A	3	0	0	0
14A	4	0	0	0
15A	4	0	0	0
17A	4	0	0	0
18A	4	0	0	0
19B	3	0	0	0

21A	4	0	0	0
22A	4	0	0	0
23A	4	0	0	0
24A	4	0	0	0
26A	4	0	0	0
27A	4	0	0	0
28A	8	0	0	0
29A	4	0	0	0
32A	4	0	0	0
33A	4	0	0	0
36A	4	0	0	0
37A	4	0	0	0
38A	4	0	0	0
39A	4	0	0	0
41A	4	0	0	0
ACE	12	0.03	0.09	0.18

4.3.9 **Zero-Copy Control**

There were 16 zero-copy control wells in validation study 1, and 32 zero-copy control wells each in validation studies 2 and 3. For each well, there was a FAM-assay result and HEX-assay result. Excluding the *ACE* assay, 35 CNV assays were run against a zero-copy control. Each assay had 4 assay replicates with exception of CNV 28A, which had 8 replicates. There were no zero-copy control wells excluded for an accepted droplet number below 10,000. The total number of zero-copy control well results was therefore 160. The non-polymorphic control assay *ACE* accounted for 16 results, and the remaining 144 results were from CNV assays.

4.3.9.1 *ACE Results*

There were 4 *ACE* results for the PID 114 zero-copy control in validation study 1, 8 *ACE* results for the PID 133 zero-copy control in validation study 2, and 4 *ACE* results for the PID 114 zero-copy control in validation study 3. Due to the distribution of zero-copy control genotypes within each validation study, *ACE* zero-copy control results were not available for the PID 148 zero-copy control in validation study 2 and the PID 97 zero-copy control in validation study 3 (Table 4-14).

The median *ACE* template concentrations for PID 114 validation study 1, PID 133 validation study 2 and PID 114 validation study 3 were 255 (IQR 249-262), 130 (IQR 111-132), and 243 (IQR 231-247) copies/μL respectively.

4.3.9.2 Threshold Technique

The false positive result rate varied according to the threshold technique used ($\chi^2(4, N = 5) = 54$, $p < 0.0001$) (Figure 4-41). Higher amplitude thresholding techniques (cluster, half-distance) were associated with a 3/144 (3%) false positive result rate. Lower amplitude thresholding techniques were associated with progressively higher false positive rates. Third-distance, quarter-distance and tight thresholding was associated with false positive result rates of 5/144 (4%), 8/144 (6%) and 29/144 (20%) respectively. There was no significant relationship between threshold technique and the false positive result concentration ($r_s = -0.13$, $N = 48$, $p \approx 0.38$) (Figure 4-42).

16/35 (46%) CNV assays had no false positive results, irrespective of the threshold technique used. For the remaining 19 assays with at least one false positive replicate, most (13/19) of these were present only when the tight threshold technique had been applied (Table 4-22).

All assays with the exception of assay 03B had a maximum of one false positive droplet at threshold technique amplitudes higher than the tight threshold technique, and thresholds higher than the third-distance threshold technique excluded this false positive droplet (Figure 4-43). This indicates that the false positive droplet had an amplitude between one third to half of the distance between the positive and negative clusters.

CNV assay 03B had 3 of 4 replicates detecting 1-2 droplets at all threshold amplitudes including the cluster amplitude. This indicates that the amplitude of these false positive droplets is the same as true positive droplets, raising the possibility of contamination of the PCR reaction with DNA that was not zero-copy for the CNV 3 locus.

Figure 4-41: The proportion of 144 zero-copy control results that were false positive, grouped by threshold technique. Lower droplet amplitude threshold techniques were associated with a higher false positive rate.

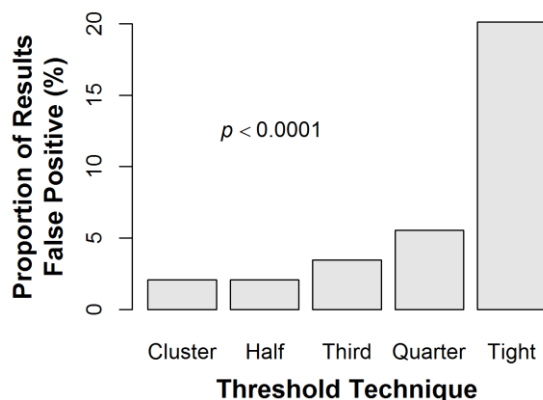


Figure 4-42: The concentration of false positive zero-copy control results grouped by threshold technique. There was no significant relationship between threshold techniques of declining threshold amplitude and the false positive result concentration.

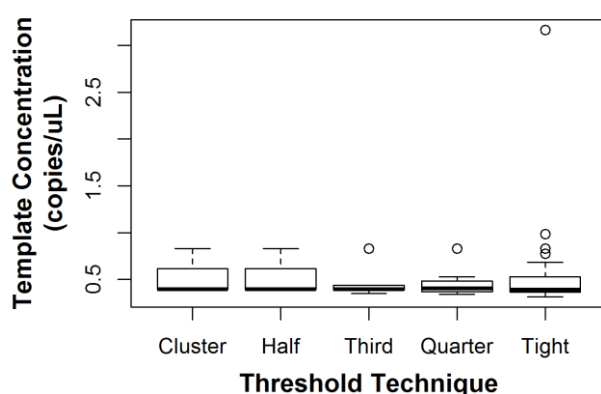
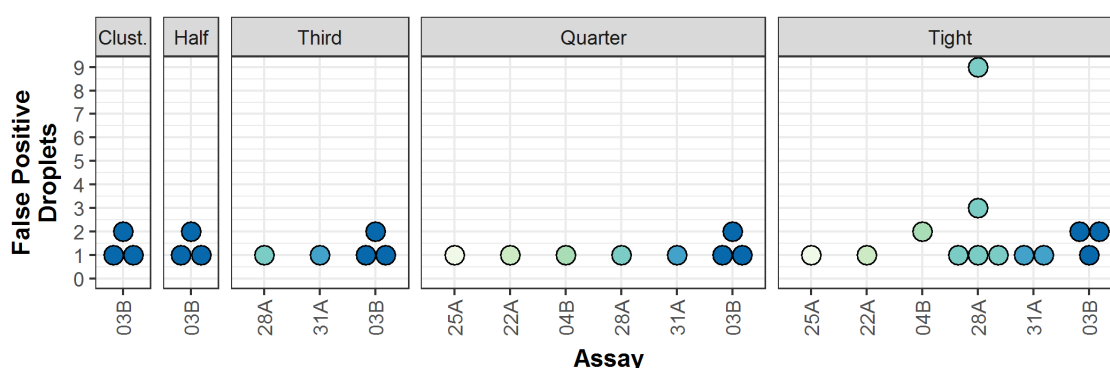


Table 4-22: Frequency of zero-copy control false positive results by assay and threshold technique. Assays with no false positive results are not shown.

Assay	Threshold Technique				
	Cluster	Half	Third	Quarter	Tight
03B	3	3	3	3	3
31A	0	0	1	1	2
28A	0	0	1	1	5
04B	0	0	0	1	1
22A	0	0	0	1	1
25A	0	0	0	1	1
26A	0	0	0	0	2
34A	0	0	0	0	2
36A	0	0	0	0	2
05B	0	0	0	0	1
12A	0	0	0	0	1
13A	0	0	0	0	1
14A	0	0	0	0	1
15A	0	0	0	0	1
20A	0	0	0	0	1
24A	0	0	0	0	1
27A	0	0	0	0	1
29A	0	0	0	0	1
32A	0	0	0	0	1

Figure 4-43: Dot plot of the number of false positive droplets per false positive replicate well for each assay, grouped by threshold technique. Only false positive results are displayed. With the

exception of assay 03B, all other assays had only one false positive droplet at threshold amplitudes higher than the tight threshold technique. (Clust., cluster threshold technique.)



4.3.9.3 Assay Design Characteristics

CNV assays with any false positive result (03B, 04B, 22A, 25A, 28A and 31A) were compared against those with no false positive result (Figure 4-44). There was a lower reverse primer GC content in CNV assays with a false positive result (median GC 39% (37-40%) versus 45% (38-50%), $W = 133$, $p \approx 0.047$). There were no other significant assay design characteristic differences between these groups. The rate of false positive results did not differ between HEX and FAM assays (OR 0.87, 95% CI 0.09-7.17, $p = 1$). Additionally, there was no significant difference between the groups in negative or positive cluster centroid amplitude, or the distance between the positive and negative cluster centroid (Figure 4-45).

CNV assay design characteristics for those assays associated with a false positive result at the third-distance threshold technique (03B, 28A and 31A) were compared against the remaining assays, excluding *ACE*. There were no significant differences in assay design parameters between these groups (Figure 4-46). The rate of false positive results did not differ between HEX and FAM assays (OR 0.40, 95% CI 0.01-8.41, $p = 0.58$). Additionally, there was no significant difference between the groups in negative or positive cluster centroid amplitude, or the distance between the positive and negative cluster centroid (Figure 4-47).

Figure 4-44: Assay design characteristics compared between a group of assays with at least one false positive result, and those with no false positive results. Melting temperatures in °C, size in bp and GC content in %. Statistical significance represented with asterisks.

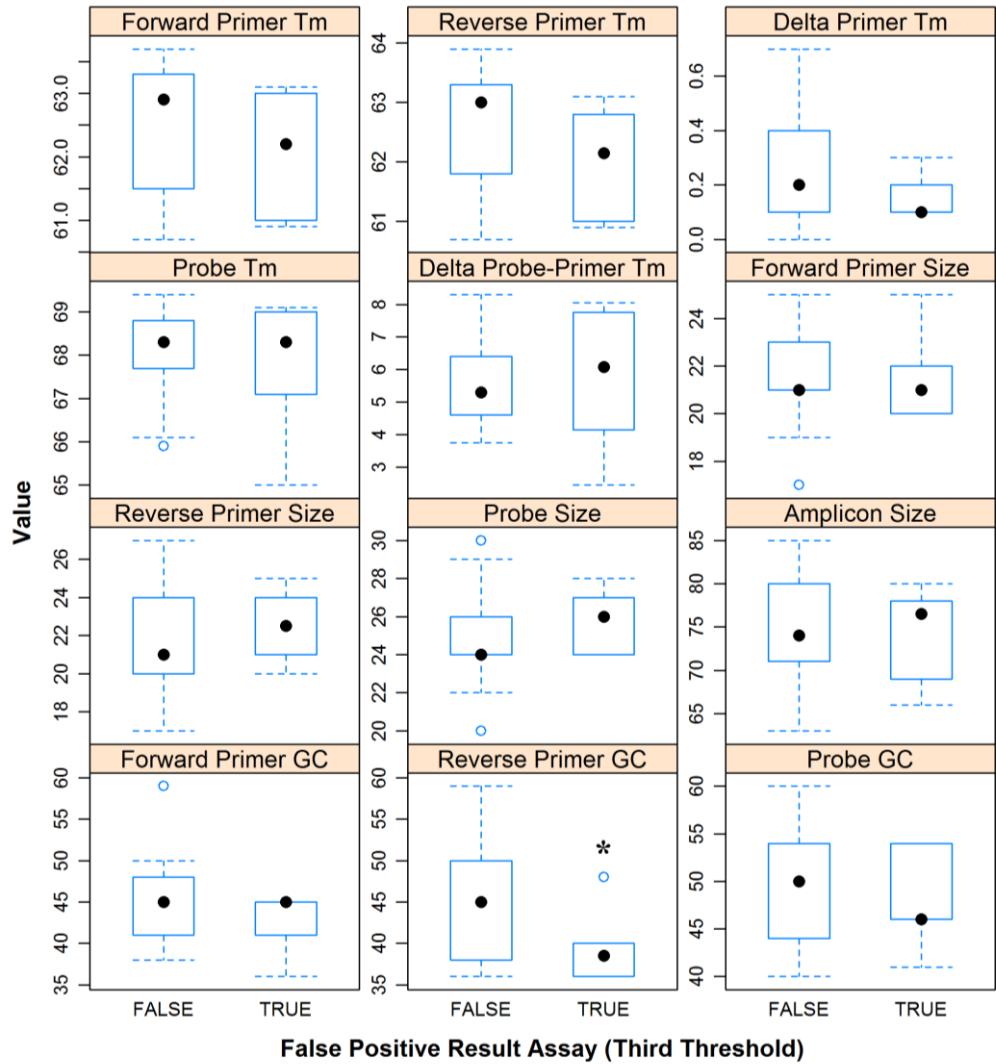


Figure 4-45: Centroid amplitudes and centroid separation (positive-negative delta) compared between assays with at least one false positive result and those with no false positive results. No distributions were significantly different between the groups.

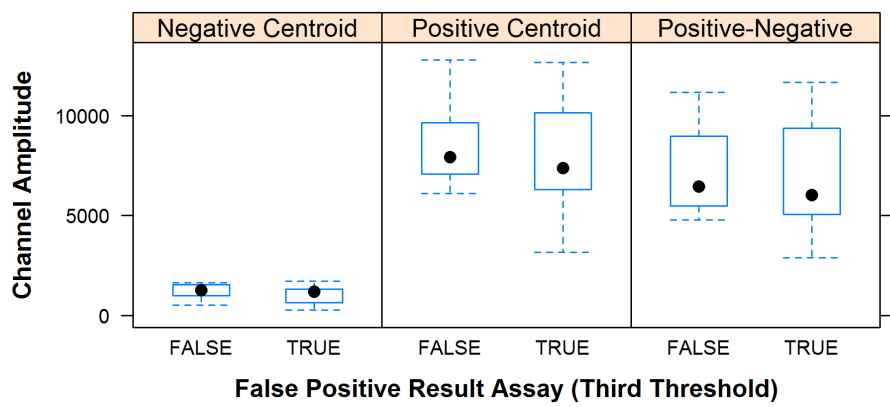


Figure 4-46: Assay design characteristics compared between assays with a false positive result at the third-distance threshold technique, and those with no false positive results. Melting temperatures in °C, size in bp and GC content in %. No distributions were significantly different between the groups.

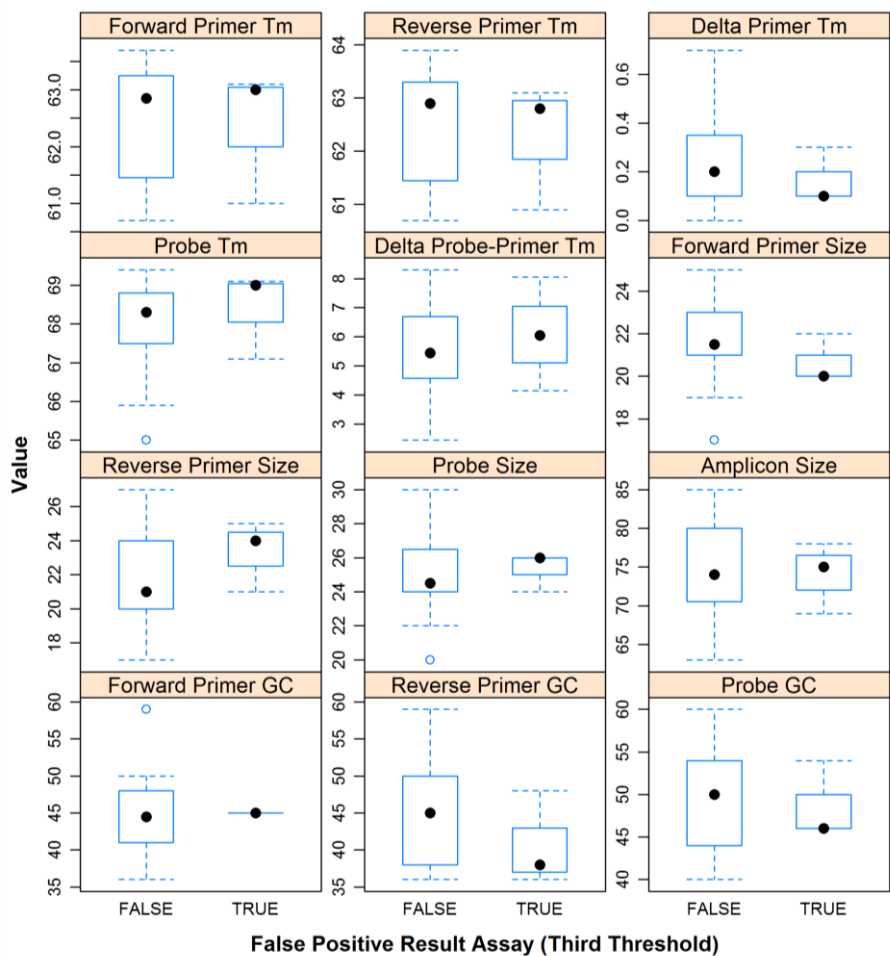
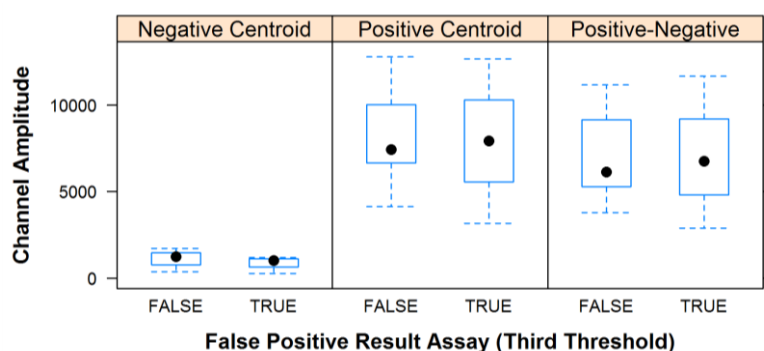


Figure 4-47: Centroid amplitudes and centroid separation (positive-negative delta) compared between assays with a false positive result at the third-distance threshold technique, and those with no false positive results. No distributions were significantly different between the groups.



4.3.9.4 Limit of Blank (Zero-Copy Control)

Using the third-distance threshold technique and zero-copy DNA template, the false positive rate was 5/144 (4%). The majority (4/5) of the false positive results consisted of 1 false positive droplet, equating to a concentration of 0.4 copies/ μ L (plasma equivalent concentration 4 copies/mL) (Table 4-23). The remaining false positive result consisted of 2 false positive droplets in one replicate of assay CNV 03B, equivalent to a concentration of 0.8 copies/ μ L. CNV assay 03B accounted for the majority (3/5) of the zero-copy control false positive results.

The limit of blank was 0 copies/ μ L for the majority (32/35) of CNV assays (Table 4-24). For CNV assay 28A and 31A, the limit of blank was 0.25 and 0.47 copies/ μ L, respectively. The limit of blank for CNV assay 03B was 0.96 copies/ μ L.

Table 4-23: Zero-copy control false positive results, using the third-distance threshold technique.

Validation Study	Plate	Well	CNV Assay	Positive Droplets	Template Concentration
1	B	B10	03B	2	0.8
1	B	C10	03B	1	0.4
1	B	D10	03B	1	0.4
2	D	B10	28A	1	0.4
1	A	A10	31A	1	0.4

Table 4-24: Limit of blank parameters and results for assays run against zero-copy control samples. The limit of blank was 0 copies/ μ L for all assays except 03B, 28A and 31A.

Assay	Number of Replicates	Mean (copies/ μ L)	Standard Deviation (copies/ μ L)	Limit of Blank (Zero-Copy Control, copies/ μ L)
01B	4	0	0	0
02B	4	0	0	0
03B	4	0.40	0.34	0.96
04B	4	0	0	0

05B	4	0	0	0
06B	4	0	0	0
08B	4	0	0	0
10C	4	0	0	0
12A	4	0	0	0
13A	4	0	0	0
14A	4	0	0	0
15A	4	0	0	0
17A	4	0	0	0
18A	4	0	0	0
19B	4	0	0	0
20A	4	0	0	0
21A	4	0	0	0
22A	4	0	0	0
23A	4	0	0	0
24A	4	0	0	0
25A	4	0	0	0
26A	4	0	0	0
27A	4	0	0	0
28A	8	0.04	0.12	0.25
29A	4	0	0	0
30A	4	0	0	0
31A	4	0.11	0.22	0.47
32A	4	0	0	0
33A	4	0	0	0
34A	4	0	0	0
36A	4	0	0	0
37A	4	0	0	0
38A	4	0	0	0
39A	4	0	0	0
41A	4	0	0	0

4.3.10 **One-Copy Dilution Series**

The experimental plan included 8 dilution series steps per validation study. Due to a laboratory error, validation study 1 included 9 dilution steps as the no template control well contained a further dilution step (0.5 copies/ μ L) rather than distilled water. This additional dilution step was included in the one-copy dilution series analysis.

There were 653 wells containing one-copy dilution series assays. 3/653 (0.5%) wells were excluded from analysis due to accepted droplet numbers <10,000. These were validation study 1, plate A, wells A08 (9,925 droplets) and D01 (2,687 droplets) containing assay pair 31A/01B, and validation study 3, plate C, well C08 (5,797 droplets) containing assay pair 19B/13A.

Consequently, 1,174 CNV assay results were available for analysis.

4.3.10.1 Achieved Concentrations and Linearity

The median starting concentrations for the dilution series of validation studies 1, 2 and 3 were 113, 128 and 117 copies/ μ L respectively.

Achieved dilution step concentrations were consistently slightly below target for validation study 1, consistently at target for validation study 2, and inconsistently below target for validation study 3 (Figure 4-48). The median proportionate variation of achieved versus targeted concentration was 0.88 (0.70-1.01), 0.96 (0.78-1.15), and 0.46 (0.17-0.83) for validation studies 1-3 respectively. Excluding target concentrations ≤ 1 copy/ μ L, the distribution of proportionate variation of achieved versus targeted concentration was consistent across dilution steps for validation study 1 ($\chi^2(6, N = 194) = 3, p = 0.86$) and validation study 2 ($\chi^2(6, N = 392) = 12, p = 0.08$), but not for validation study 3 ($\chi^2(6, N = 420) = 336, p < 0.0001$). Validation study 3 could not be repeated due to insufficient DNA.

Lower threshold amplitudes were associated with variable, but generally small, increases in one-copy assay concentrations and the proportion of target concentration achieved (Figure 4-49). The difference was greatest moving from the cluster to half-distance threshold technique (3.1% (IQR 2.4-4.0%)), and from the quarter-distance to the tight threshold technique (2.0% (IQR 1.7-2.4%)). There was a slight increase moving from the half-distance to the third-distance technique (1% (IQR 0.0-1.7%)), but effectively no increase moving from the third-distance technique to the quarter-distance technique (0.0% (IQR 0.0-0.6%)). The third-distance threshold technique was used for all data presented, unless specified otherwise.

Plots of replicate data at each dilution step and corresponding assay linear models are presented in Figure 4-50 (validation study 1), Figure 4-51 (validation study 2), and Figure 4-52 (validation study 3). Linear model characteristics are presented in Table 4-25 and summarised in Figure 4-53. The median slope was 1.0 (0.9-1.1) copies/ μ L, intercept -1.1 (-3.0--0.1) copies/ μ L, and Pearson's r 0.996 (0.994-0.997). There was statistically significant variation in intercept between the validation studies ($\chi^2(2, N = 35) = 27, p < 0.001$). Post-hoc testing identified this was caused by validation study 3 ($p < 0.001$). There was no statistically significant variation between validation studies for slope ($\chi^2(2, N = 35) = 0.4, p = 0.82$), Pearson's r ($\chi^2(2, N = 35) = 4.7, p = 0.10$) and Pearson's R^2 ($\chi^2(2, N = 35) = 5.0, p = 0.08$).

Multiple regression found assay design characteristics explained 34% of the variance (adjusted R^2) in assay slope ($F(11,23) = 2.60, p = 0.03$). The assumptions of normality and homoscedasticity were confirmed. Mean primer-probe melting temperature delta was excluded as it varied co-linearly with primer and probe melting temperatures. Significant predictors in the model were

amplicon size ($\beta = -0.025$, $p = 0.002$) and probe GC content ($\beta = -0.028$, $p = 0.03$). There was also a trend towards significance for probe length ($\beta = -0.049$, $p = 0.06$). Assay design characteristics did not predict variance in assay Pearson's r ($F(11,23) = 0.57$, $p = 0.83$).

Outlier assay linear model parameters are summarised in Table 4-26. CNV assay 34A (validation study 1) and 04B (validation study 3) were outlier linear model slope assays. These assays had a longer probe size compared to non-outlier assays (29 (28.5-29.5) versus 24 (24-26), $W = 3$, $p \approx 0.03$) (Figure 4-54). There were trends towards lower reverse primer melting temperature ($p \approx 0.14$), shorter forward primer size ($p \approx 0.09$), longer amplicon size ($p \approx 0.13$), and lower probe GC content ($p \approx 0.11$) in the outlier group. Within the assay panel developed, CNV 34A had the highest delta primer melting temperature (tied with CNV 20A), longest probe size (tied with CNV 13A), and lowest probe GC content. CNV 04B had the lowest reverse primer GC content (tied with CNVs 01B, 03B and 19B).

Figure 4-48: Box plot of CNV assay concentration as a proportion of the target concentration at that particular dilution step, grouped by validation study.

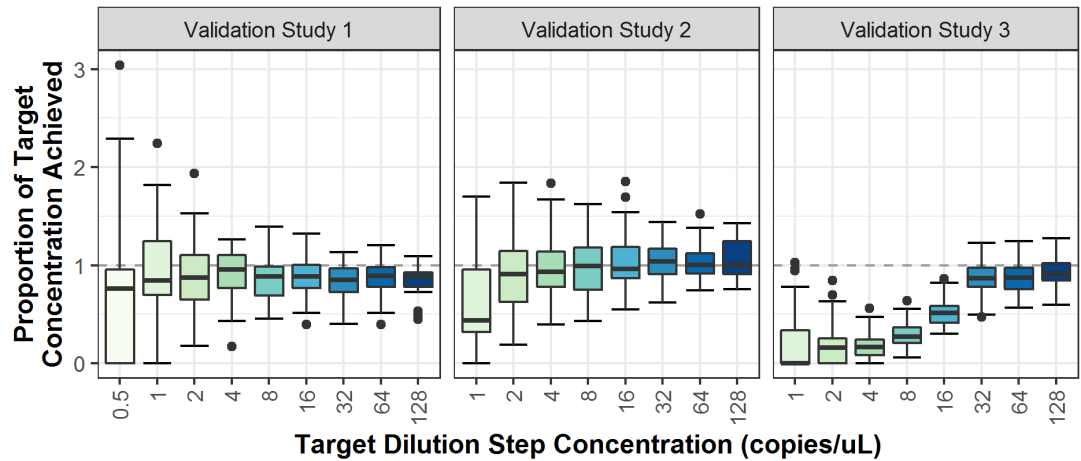


Figure 4-49: The effect of threshold technique on the proportion of one-copy target concentration achieved across all dilution steps. Reducing the threshold amplitude variably increased the assay concentration, which was most pronounced moving from the cluster to half threshold technique.

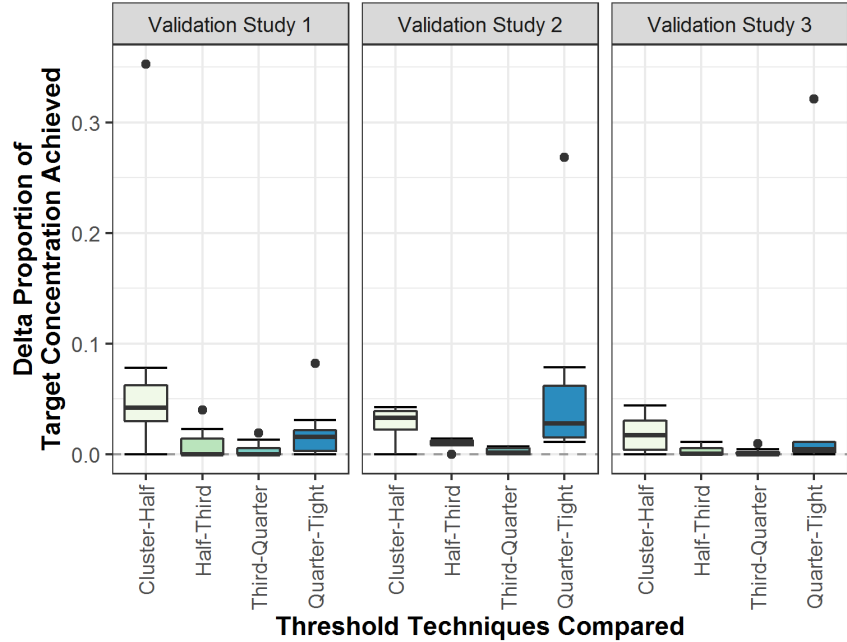


Figure 4-50: Validation study 1 CNV assay replicate results and linear models. Both axes are based on a Log2+1 scale.

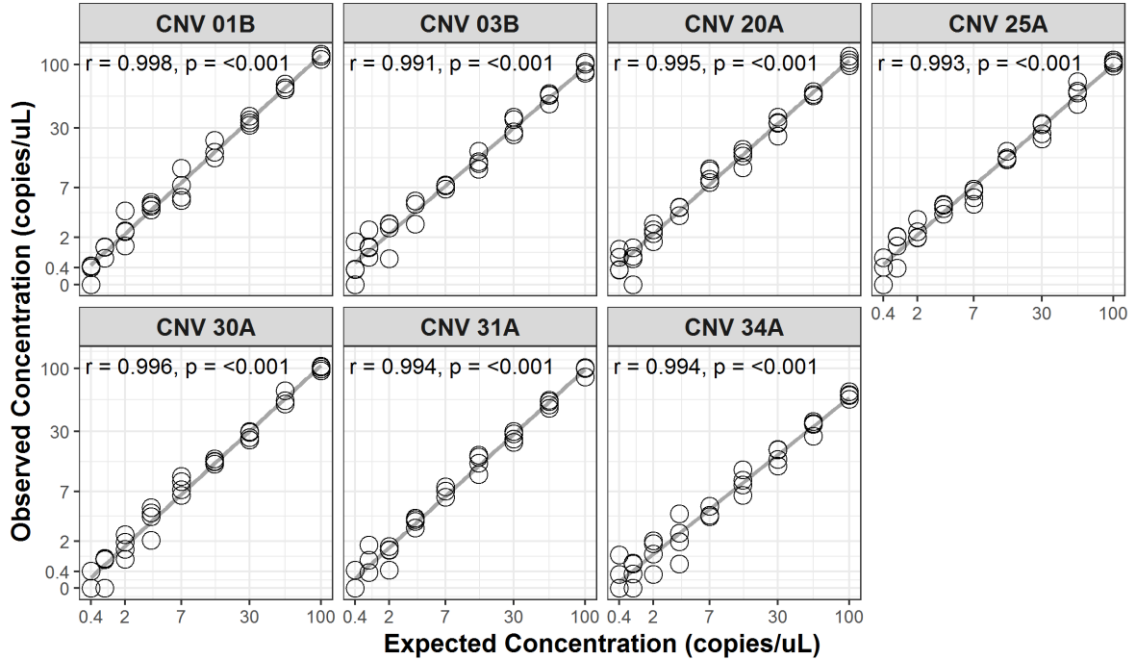


Figure 4-51: Validation study 2 CNV assay replicate results and linear models. Both axes are based on a Log2+1 scale.

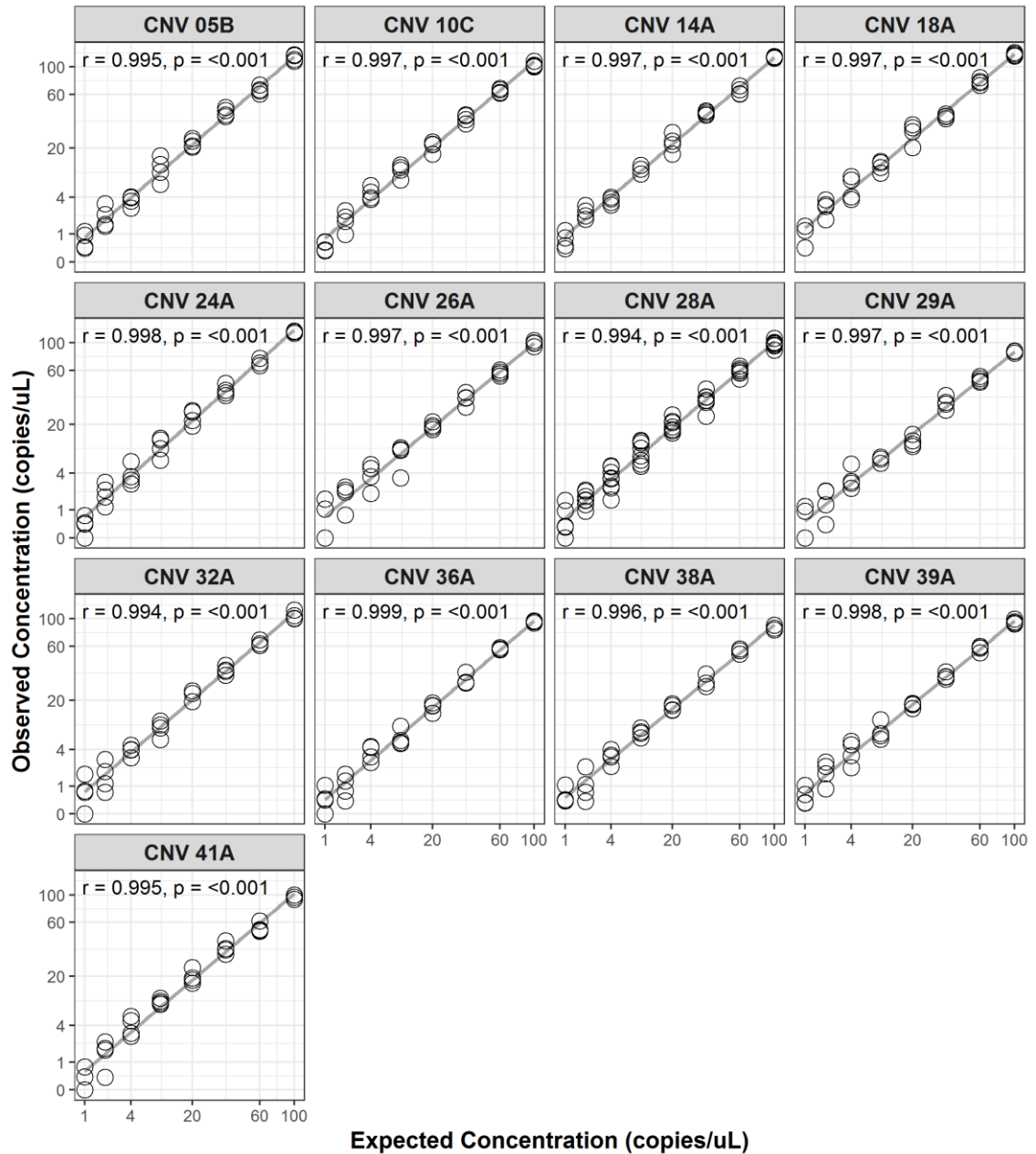


Figure 4-52: Validation study 3 CNV assay replicate results and linear models. Both axes are based on a Log2+1 scale.

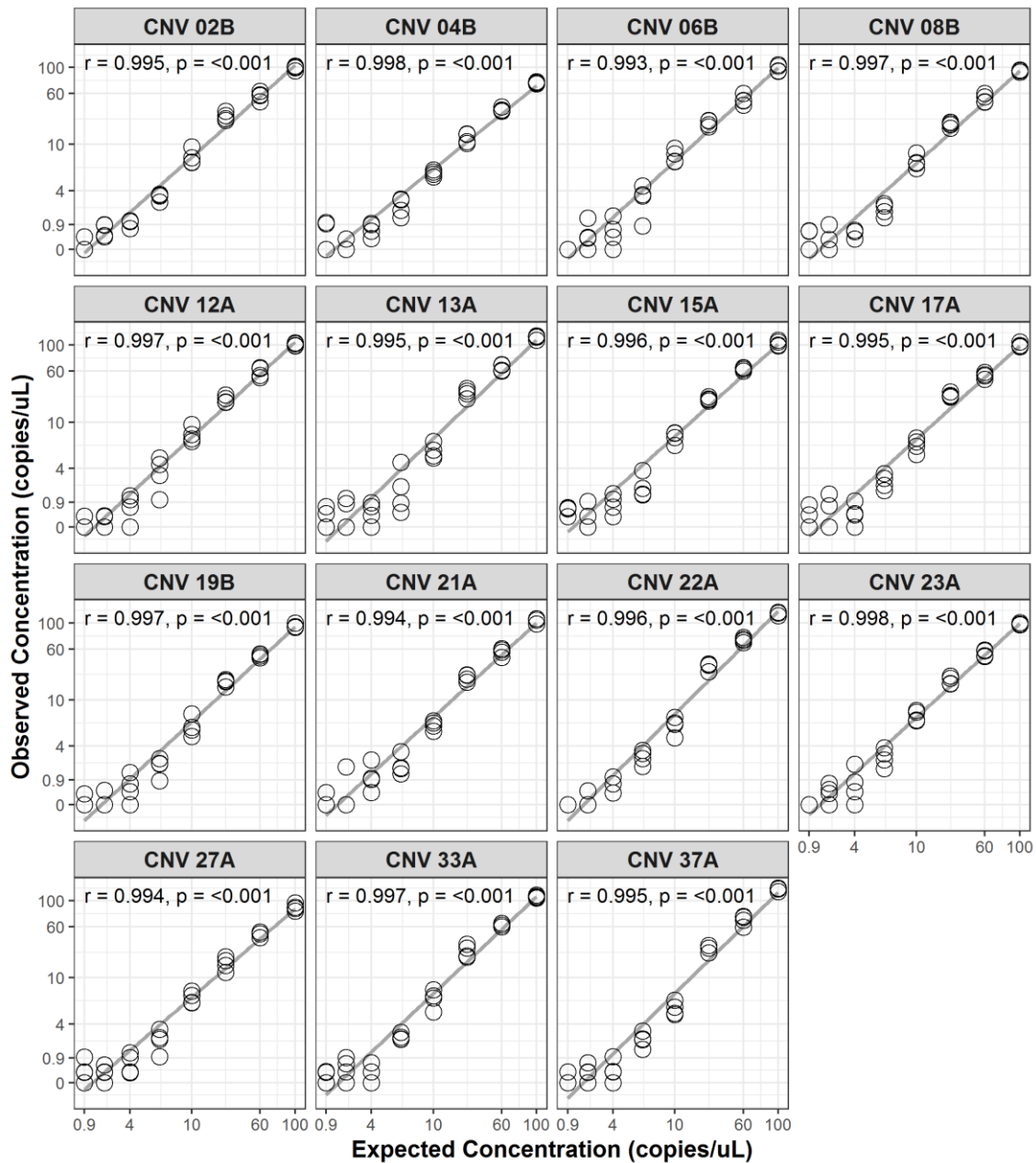


Table 4-25: One-copy dilution series linear model parameters, including Pearson's correlation coefficient (r) and Pearson's coefficient of determination (R²).

Validation Study	CNV Assay	Slope	Intercept	r	95% CI	p-value	R ²
1	01B	1.2	-0.1	0.998	0.995-0.999	<0.001	0.995
3	02B	1.0	-2.1	0.995	0.99-0.998	<0.001	0.990
1	03B	0.9	0.9	0.991	0.983-0.996	<0.001	0.983
3	04B	0.7	-1.9	0.998	0.996-0.999	<0.001	0.997
2	05B	1.2	-1.3	0.995	0.989-0.998	<0.001	0.990

3	06B	1.0	-3.1	0.993	0.986-0.997	<0.001	0.986
3	08B	0.9	-2.8	0.997	0.995-0.999	<0.001	0.995
2	10C	1.1	0.2	0.997	0.994-0.999	<0.001	0.995
3	12A	1.0	-2.9	0.997	0.993-0.998	<0.001	0.993
3	13A	1.3	-5.4	0.995	0.989-0.998	<0.001	0.990
2	14A	1.2	-1.4	0.997	0.993-0.998	<0.001	0.993
3	15A	1.1	-3.3	0.996	0.993-0.998	<0.001	0.993
3	17A	1.0	-3.1	0.995	0.99-0.998	<0.001	0.990
2	18A	1.4	-0.6	0.997	0.994-0.998	<0.001	0.994
3	19B	0.9	-3.7	0.997	0.993-0.998	<0.001	0.993
1	20A	1.1	0.3	0.995	0.991-0.998	<0.001	0.991
3	21A	1.1	-4.4	0.994	0.988-0.997	<0.001	0.989
3	22A	1.3	-4.7	0.996	0.992-0.998	<0.001	0.993
3	23A	1.0	-3.2	0.998	0.995-0.999	<0.001	0.995
2	24A	1.3	-1.6	0.998	0.996-0.999	<0.001	0.996
1	25A	1.1	-0.1	0.993	0.986-0.996	<0.001	0.985
2	26A	1.0	-0.6	0.997	0.993-0.998	<0.001	0.994
3	27A	0.9	-2.5	0.994	0.988-0.997	<0.001	0.989
2	28A	1.0	-0.4	0.994	0.99-0.996	<0.001	0.988
2	29A	0.8	-0.3	0.997	0.995-0.999	<0.001	0.995
1	30A	1.0	-0.1	0.996	0.993-0.998	<0.001	0.993
1	31A	0.9	-0.1	0.994	0.988-0.997	<0.001	0.988
2	32A	1.1	-0.2	0.994	0.988-0.997	<0.001	0.988
3	33A	1.1	-3.9	0.997	0.994-0.999	<0.001	0.994
1	34A	0.6	0.4	0.994	0.988-0.997	<0.001	0.988
2	36A	0.9	-1.0	0.999	0.997-0.999	<0.001	0.998
3	37A	1.4	-6.1	0.995	0.989-0.997	<0.001	0.989
2	38A	0.8	0.3	0.996	0.992-0.998	<0.001	0.993
2	39A	0.9	-0.2	0.998	0.995-0.999	<0.001	0.996
2	41A	0.9	0.2	0.995	0.99-0.998	<0.001	0.990

Figure 4-53: Summary of linear model characteristics by validation study.

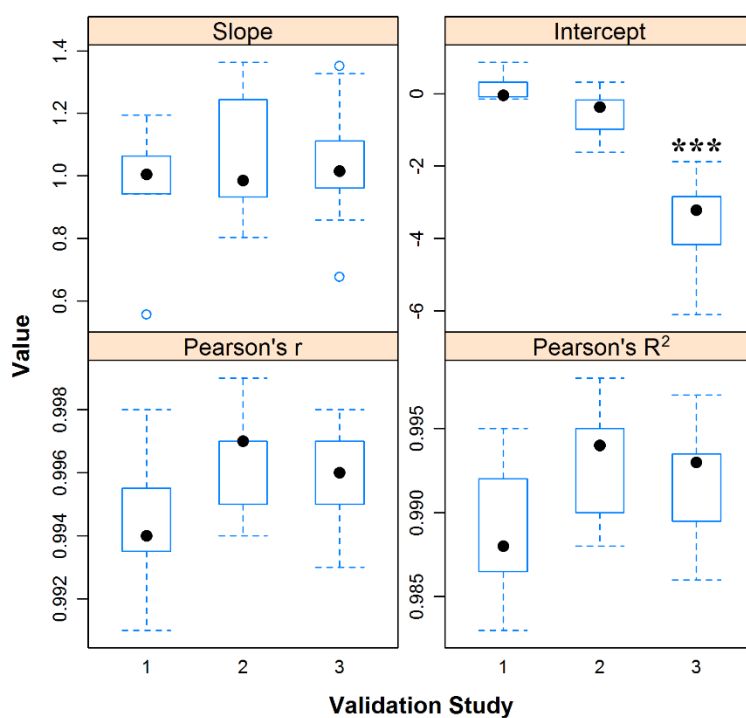
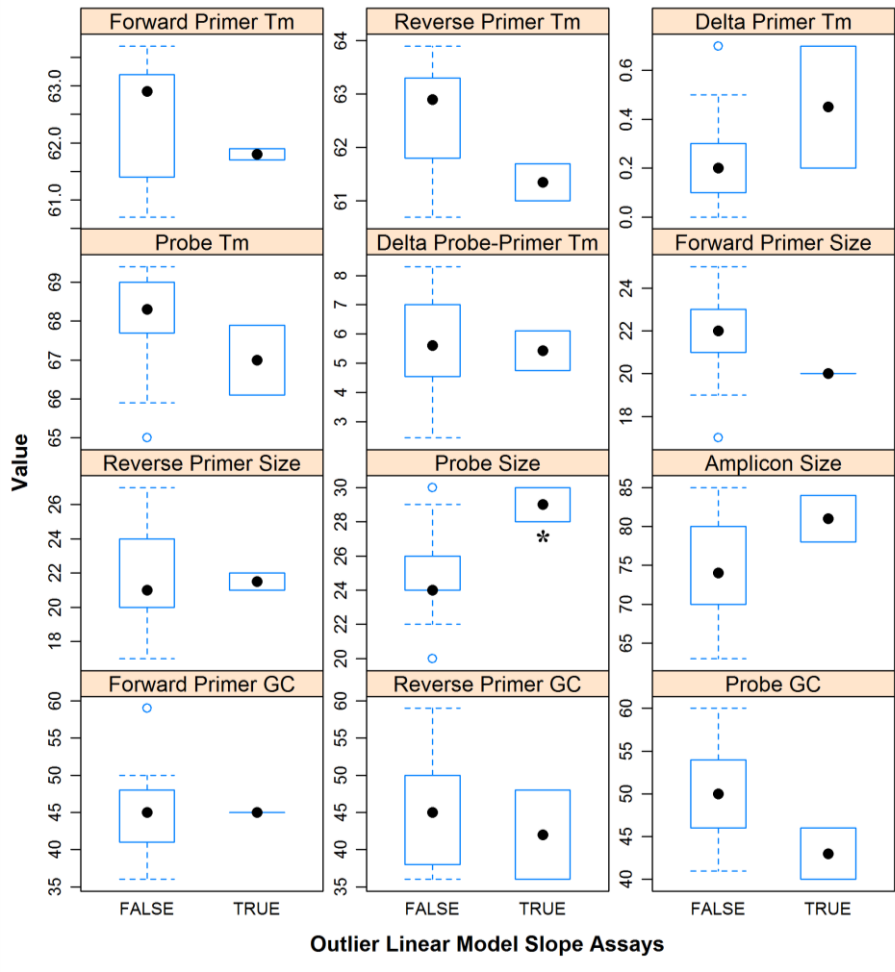


Table 4-26: Outlier assay linear model parameters after validation study results were pooled.

Validation	Assay	Parameter	Result	Z-Score	IQR Test
1	03B	Pearson's r	0.991	-2.65	Pass
1	03B	Pearson's R ²	0.983	-2.44	Pass
1	34A	Slope	0.56	-2.63	Fail
3	04B	Slope	0.68	-1.96	Pass
3	13A	Intercept	-5.4	-2.00	Pass
3	37A	Intercept	-6.1	-2.37	Pass

Figure 4-54: Assay design characteristics compared between a group of assays with outlier linear model slope values and those of non-outlier assays. Melting temperatures in °C, size in bp and

GC content in %. Statistical significance represented with asterisks if significant. Non-significant differences not marked.



4.3.10.2 Limit of Detection

All CNV assay replicates detected target sequence at concentrations above 0.9 copies/ μ L (Figure 4-55). At and below a concentration of 0.6 copies/ μ L, the proportion of all replicates that failed to detect target sequence rose above 50%. The median concentration at which any CNV assay replicate failed to detect target was 0.5 (IQR 0.4-0.6) copies/ μ L (Table 4-27). For the majority of assays 18/29 (62%), this represented failure of one replicate to detect target.

The median limit of detection was 0.6 (IQR 0.3-0.7) copies/ μ L. The low concentration sample standard deviation was calculated at a median assay concentration of 0.6 (IQR 0.6-0.8) copies/ μ L. Individual CNV assay limit of detection results and parameters are presented in Table 4-28.

Figure 4-55: Tile plot indicating the proportion of assay replicates that detected target (positive assay result) across validation study dilution step concentrations. Darker coloured tiles indicate that a proportion of assay replicates at that concentration failed to detect the target (no positive

droplets, negative assay result). The percentages displayed in the bottom right corner indicate the proportion of all assay replicates at the indicated dilution step concentration that detected target (i.e. the sum of the proportions of all assays on that row). The dashed red line indicates the threshold at which the proportion of assay replicates detecting target fell below 50%.

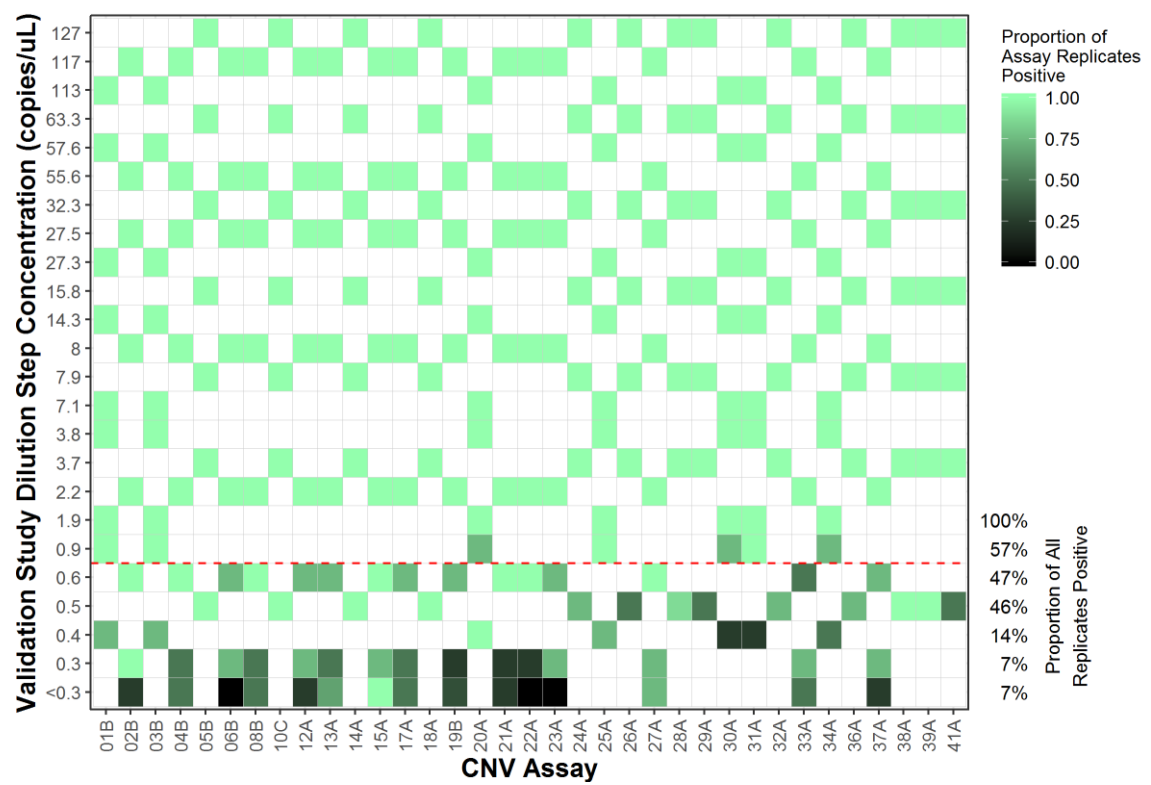


Table 4-27: The maximum concentration at which replicate failure was observed for each CNV assay (proportion of replicates detecting target < 1). Above these concentrations, all assay replicates detected target and generated a result.

CNV Assay	Maximum Concentration of Replicate Failure to Detect Target (copies/μL)	Proportion of Replicates Detecting Target
01B	0.4	0.75
02B	<0.3	0.25
03B	0.4	0.75
04B	0.3	0.5
06B	0.6	0.75
08B	0.3	0.5
12A	0.6	0.75
13A	0.6	0.75
15A	0.3	0.75
17A	0.6	0.75
19B	0.6	0.75

20A	0.9	0.75
21A	0.3	0.25
22A	0.3	0.25
23A	0.6	0.75
24A	0.5	0.75
25A	0.4	0.75
26A	0.5	0.5
27A	0.3	0.75
28A	0.5	0.875
29A	0.5	0.5
30A	0.9	0.75
31A	0.4	0.25
32A	0.5	0.75
33A	0.6	0.5
34A	0.9	0.75
36A	0.5	0.75
37A	0.6	0.75
41A	0.5	0.5

Table 4-28: Limit of detection for each CNV assay, using zero-copy control limit of blank.

CNV Assay	Mean (copies/ μ L)	Standard Deviation (copies/ μ L)	Limit of Blank (copies/ μ L)	Limit of Detection (copies/ μ L)
01B	0.47	0.03	0	0.04
02B	0.66	0.28	0	0.47
03B	0.77	0.65	0.96	2.03
04B	0.99	0.05	0	0.08
05B	0.73	0.36	0	0.59
06B	0.66	0.52	0	0.86
08B	0.62	0.02	0	0.03
10C	0.48	0.18	0	0.29
12A	0.33	0.01	0	0.02
13A	0.56	0.20	0	0.33
14A	0.72	0.36	0	0.59
15A	0.56	0.16	0	0.27
17A	0.58	0.28	0	0.46
18A	1.12	0.48	0	0.79
19B	0.83	0.47	0	0.77
20A	0.67	0.37	0	0.61
21A	1.14	0.79	0	1.29
22A	0.63	0.34	0	0.56
23A	0.53	0.21	0	0.34
24A	0.53	0.19	0	0.31
25A	0.57	0.19	0	0.32

26A	1.34	0.41	0	0.68
27A	0.52	0.37	0	0.60
28A	0.59	0.49	0.25	1.05
29A	1.07	0.16	0	0.26
30A	0.88	0.03	0	0.05
31A	0.92	0.57	0.47	1.40
32A	1.06	0.55	0	0.91
33A	0.32	0.02	0	0.03
34A	0.70	0.49	0	0.81
36A	0.64	0.36	0	0.59
37A	0.44	0.21	0	0.35
38A	0.57	0.33	0	0.55
39A	0.58	0.35	0	0.57
41A	0.58	0.28	0	0.47

4.3.10.3 Intra-Assay Variation

The intra-assay coefficient of variation increased as the median assay concentration fell (Figure 4-56). The majority of assays exhibited an intra-assay coefficient of variation less than 50% at target concentrations above 0.9 copies/ μ L (equivalent to a plasma concentration of 9 copies/mL), and coefficient of variations less than 25% at target concentrations above 3.7 copies/ μ L (equivalent to a plasma concentration of 37 copies/mL). There were two high coefficient of variation outliers: CNV 13A with a coefficient of variation of 93% at the median concentration step of 2.2 copies/ μ L, and CNV 01B with a coefficient of variation of 37% at the median concentration step of 7.1 copies/ μ L.

A significant linear relationship existed between the logarithm of median assay concentration and the logarithm of the coefficient of variation ($r = -0.91$, $p < 0.0001$) (Figure 4-57). Based on this linear model ($F(1,272) = 1298$, $p < 0.0001$), target concentration accounted for 83% of the variance (adjusted R^2) of the intra-assay coefficient of variation. Including CNV assay in the model did not improve performance ($F(34) = 1.1$, $p = 0.33$).

False negative replicates (no positive droplets detectable) contribute to higher intra-assay coefficient of variation at lower concentrations. The relationship between the number of false negative replicates and assay concentration is examined further in Section 4.3.10.2. Excluding false negative replicates from the calculation of the intra-assay variation reduces intra-assay coefficient of variation at concentrations less than 1.0 copies/ μ L (Figure 4-58). The effect of this is that median intra-assay coefficient of variation remained below 50% until a dilution step concentration of 0.7 copies/ μ L.

The median intra-assay coefficient of variation across all dilution steps ≥ 1 copy/ μ L was 16% (IQR 9-34%). The distribution of intra-assay coefficient of variation was not different between CNV assays ($\chi^2(34, N = 284) = 16, p > 0.99$). Repeated measures analysis of variance found that median assay concentration ($F(1,203) = 107, p < 0.0001$) and number of replicates available for calculation of the intra-assay variation ($F(1,203) = 14, p < 0.001$) were factors in the intra-assay coefficient of variation, but there were no CNV assay effects ($F(34,203) = 0.72, p = 0.88$).

Figure 4-56: Distribution of intra-assay coefficient of variation (CV) grouped by median validation study dilution step concentration (ordinal x-axis). The equivalent plasma concentration is not displayed for clarity, but is a factor of 10 times the median validation step concentration. Red vertical dashed lines indicate two thresholds below which the median intra-assay variation was greater than the labelled coefficient of variation.

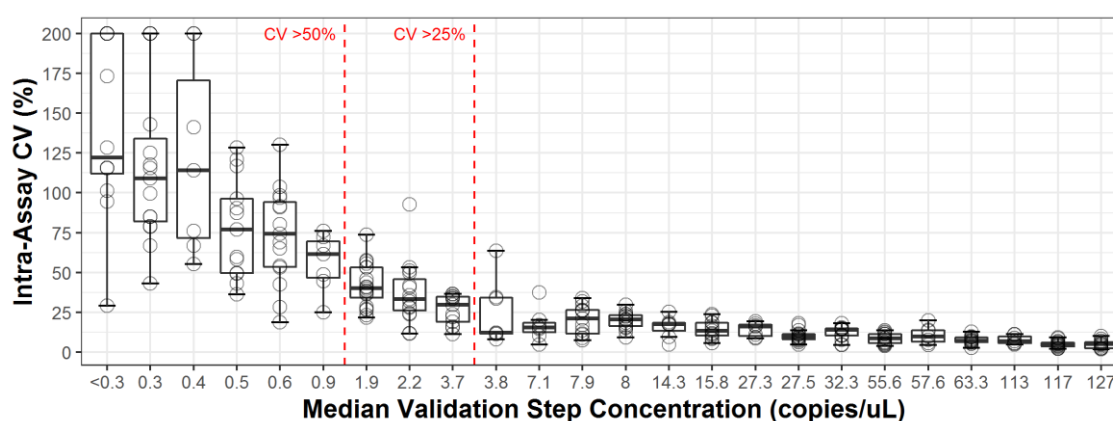


Figure 4-57: Intra-assay coefficient of variation (CV) plotted against assay median concentration. Note log₂ transformation of both axes. Dark solid line indicates linear regression model, grey shading the 95% confidence interval of the model, blue diagonal dashed lines the 95% prediction interval of the model, and red horizontal dashed lines indicate coefficient of variation thresholds as labelled (presented for clarity given logarithmic axis). In addition to the median assay

concentration in the reaction, the equivalent plasma concentration is displayed in parentheses along the x-axis based upon modelling from section 4.3.4.2.

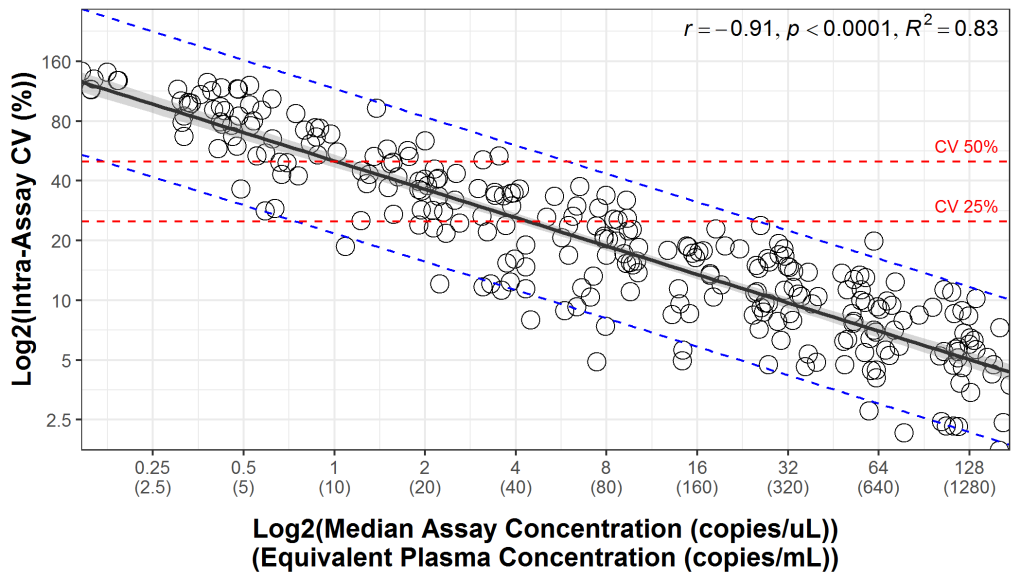


Figure 4-58: Distribution of intra-assay coefficient of variation (CV) calculated using assays detecting target only, grouped by median validation study dilution step concentration (ordinal x-axis). Assays that failed to detect target (false negatives) were excluded from calculation of the intra-assay variation. The equivalent plasma concentration is not displayed for clarity, but is a factor of 10 times the median validation step concentration. Red vertical dashed line indicates a threshold concentration below which the median intra-assay variation was greater than 25%.

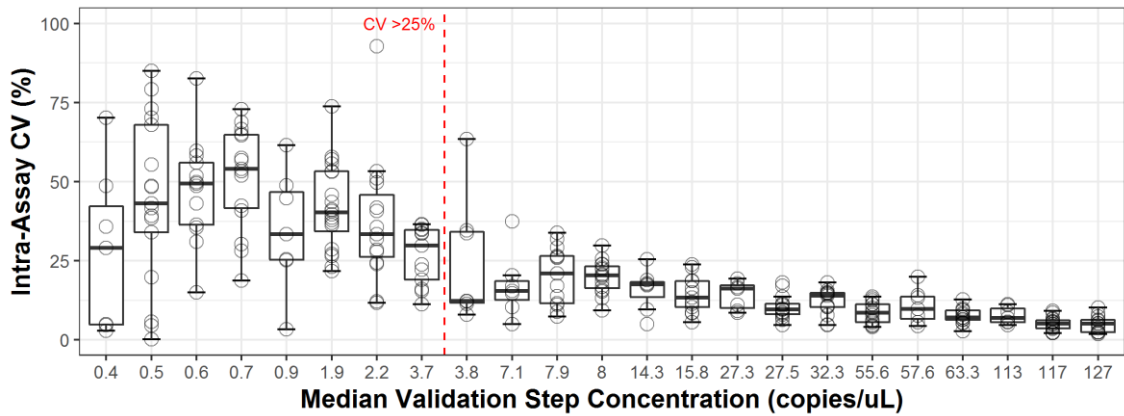
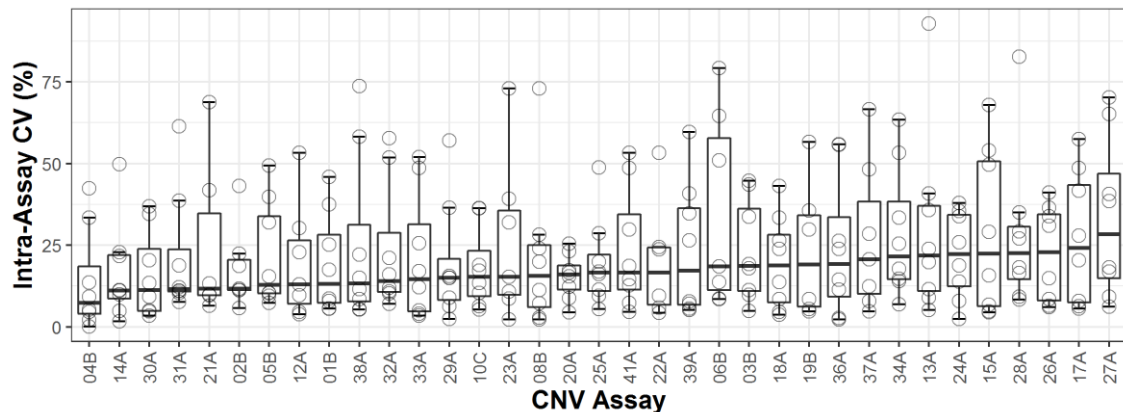


Figure 4-59: Intra-assay coefficient of variation for each CNV assay, incorporating detectable replicate results from each dilution step ≥ 1 copy/ μ L.



4.3.10.4 Inter-Assay Variation

The median inter-assay coefficient of variation was 18% (17-20%) at median validation study dilution step concentrations above 2.2 copies/ μ L (plasma equivalent 22 copies/mL) (Figure 4-60). Below this point, the inter-assay coefficient of variation progressively increased above 25% as the median validation study dilution step concentration fell. Inter-assay coefficient of variation was greater than 50% at concentrations below 0.5 copies/ μ L (plasma equivalent 5 copies/mL).

There was a linear relationship between the logarithm of the concentration and the logarithm of the inter-assay coefficient of variation ($r = -0.83$, $p < 0.0001$) (Figure 4-61). Based on this linear model ($F(1,22) = 49$, $p < 0.0001$), concentration accounted for 68% of the variance (adjusted R^2) of the inter-assay coefficient of variation. Including validation study in the model (and thus the CNV assays studied in each validation study) did not improve model performance ($F(2) = 0.4$, $p = 0.67$).

False negative results (no positive droplets detectable) contributed to a higher coefficient of variation at lower concentrations, particularly below 1 copy/ μ L. Figure 4-62 presents inter-assay coefficient of variation when false negative results have been excluded from calculations. Using this approach, inter-assay coefficient of variation rose above 50% for only one validation study at a dilution step concentration equivalent to 0.5 copies/ μ L.

The proportionate variation of assay results from the dilution step median concentration had an IQR of 0.88-1.15 (Figure 4-63). The distribution of proportionate variation was different between markers ($\chi^2(34, N = 284) = 107$, $p < 0.0001$). Repeated measures analysis of variance identified that this was a CNV assay effect rather than an assay concentration effect ($F(34,214) = 2.25$, $p = 0.001$). Considering the CNV assay and median assay concentration, mixed effects modelling

found CNVs 34A ($\beta = -0.43$, $p = 0.003$), 38A ($\beta = -0.37$, $p = 0.01$), 36A ($\beta = -0.33$, $p = 0.03$), and 18A ($\beta = 0.38$, $p = 0.01$) to be significant factors affecting proportionate variation. Repeated measures analysis of variance of the relationship between assay design characteristics and proportionate variation found amplicon size ($F(1,272) = 7.9$, $p < 0.01$) and forward primer melting temperature ($F(1,272) = 24$, $p < 0.0001$) to be significant factors. Of the designed CNV assay panel, CNVs 34A (84 bp), 36A (83 bp) and 38A (85 bp) had the longest amplicon sizes, and CNV 18A had the highest forward primer melting temperature (63.7 °C).

Figure 4-60: Inter-assay coefficient of variation (CV) plotted against the median validation study dilution step concentration (note ordinal x-axis). The equivalent plasma concentration is not displayed for clarity, but is a factor of 10 times the median validation step concentration. Red vertical dashed lines indicate two thresholds below which the median intra-assay variation was greater than the labelled coefficient of variation.

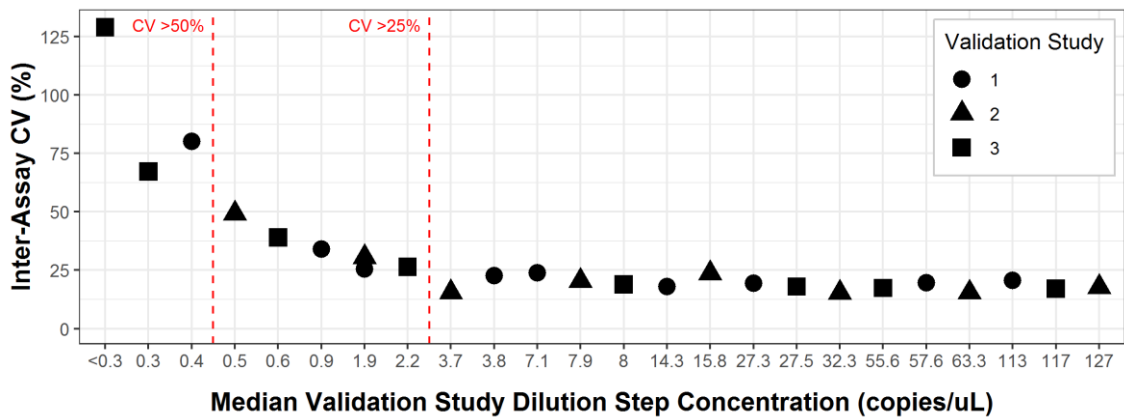


Figure 4-61: Inter-assay coefficient of variation (CV) plotted against median validation study dilution step concentration. Note log2 transformation of both axes. Dark solid line indicates linear regression model, grey shading the 95% confidence interval of the model, blue diagonal dashed lines the 95% prediction interval of the model, and red horizontal dashed lines indicate coefficient of variation thresholds as labelled (presented for clarity given logarithmic axis). In addition to the

assay concentration in the reaction, the equivalent plasma concentration is displayed in parentheses along the x-axis based upon modelling from section 4.3.4.2.

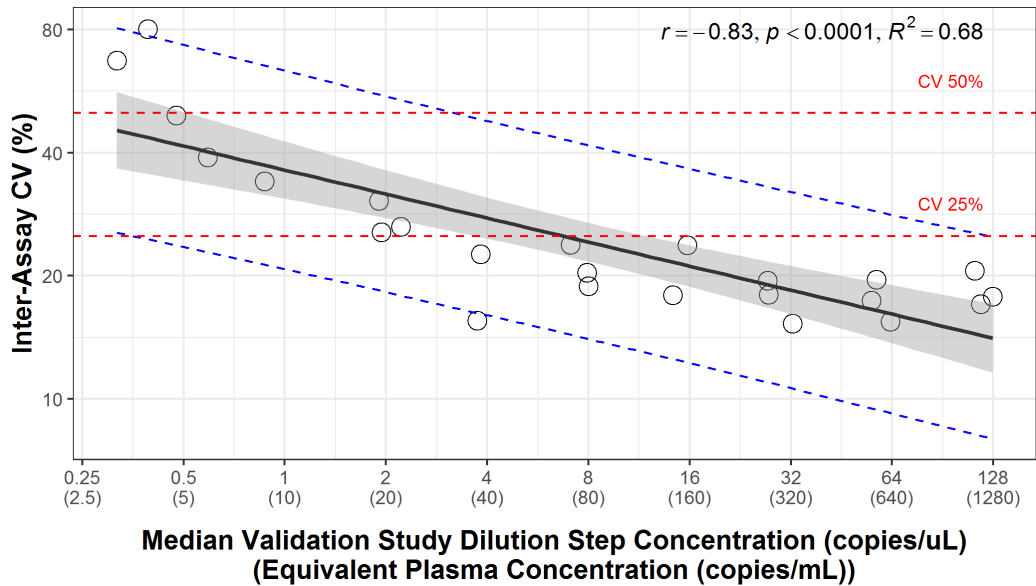


Figure 4-62: Distribution of inter-assay coefficient of variation (CV) calculated using assays detecting target only, grouped by median validation study dilution step concentration (note ordinal x-axis). Assays that failed to detect target (false negatives) were excluded from calculation of the inter-assay variation. The equivalent plasma concentration is not displayed for clarity, but is a factor of 10 times the median validation step concentration. Red vertical dashed lines indicate two thresholds below which the median intra-assay variation was greater than the labelled coefficient of variation.

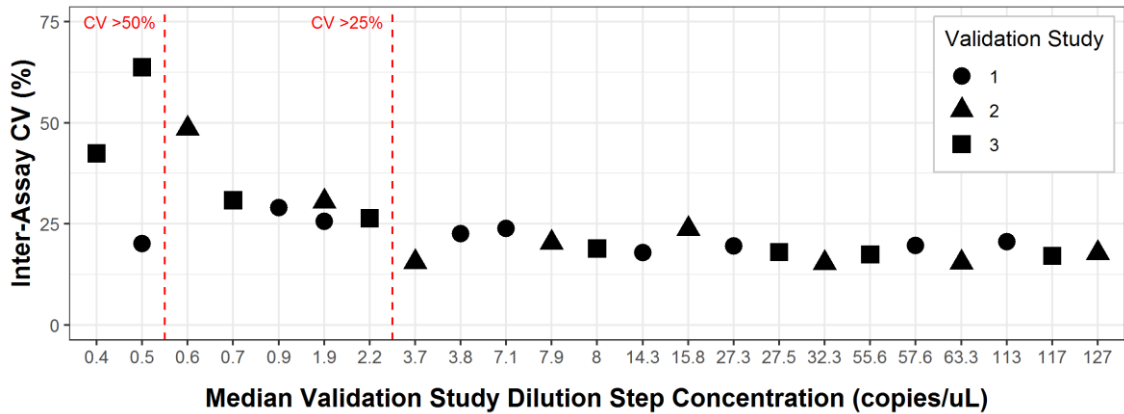
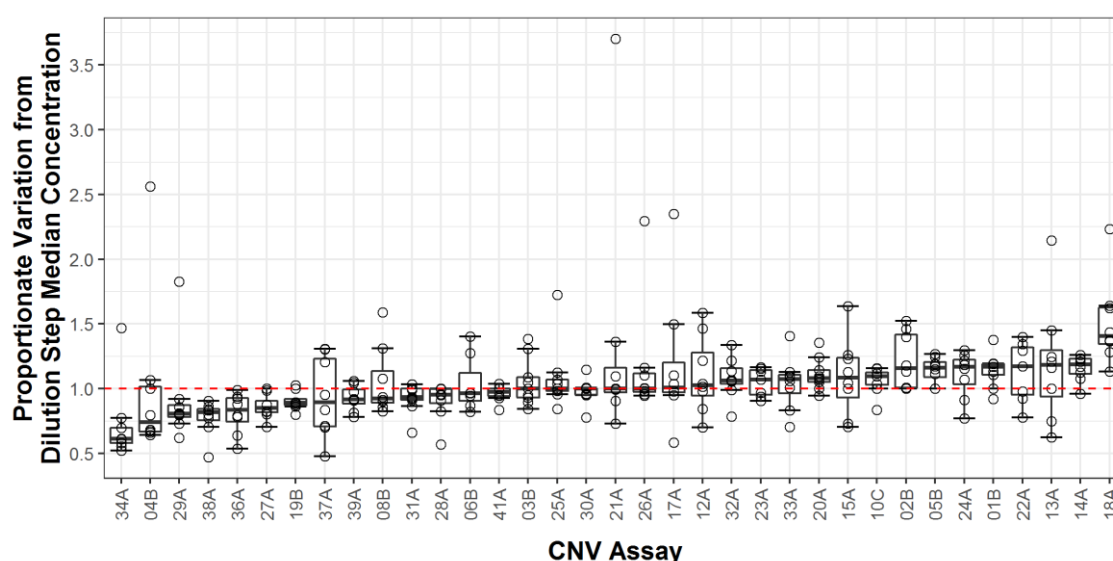


Figure 4-63: Proportionate variation of assay replicate median concentration compared to dilution step median concentration, grouped by assay.



4.3.10.5 Limit of Quantification

The intra-assay coefficient of variation was the determining factor for the limit of quantification in 23/35 (66%) assays, and limit of detection in 9/35 (26%). The median limit of quantification was 0.8 (IQR 0.6-1.4) copies/ μ L (Table 4-29). In four instances, the intra-assay coefficient of variation had a single instance of variation $> 50\%$ and was followed, at lower concentrations, by coefficients of variation consistently $< 50\%$. In these instances, the solitary high reading was still used to define the limit of quantification however another approach may have excluded these readings.

Table 4-29: Limit of quantification for each of the studied CNV assays and contributing parameters. Assay concentration results marked “(once)” indicate assay concentrations below which intra-assay coefficient of variation (CV) fell consistently below 50%.

Validation Study	CNV Assay	Assay Concentration Intra-Assay CV $> 50\%$ (copies/ μ L)	Dilution Step Concentration Inter-Assay CV $> 50\%$ (copies/ μ L)	Limit of Detection (copies/ μ L)	Limit of Quantification (copies/ μ L)
1	01B	< 0.5	0.4	0.04	0.5
3	02B	< 0.7	0.3	0.47	0.7
1	03B	0.4	0.4	2.03	2.0
3	04B	< 0.3	0.3	0.08	0.3
2	05B	< 0.7	< 0.5	0.59	0.7
3	06B	3.1	0.3	0.86	3.1
3	08B	0.6	0.3	0.03	0.6
2	10C	< 0.5	< 0.5	0.29	0.5

3	12A	3.5 (once)	0.3	0.02	3.5
3	13A	1.4 (once)	0.3	0.33	1.4
2	14A	< 0.7	< 0.5	0.59	0.7
3	15A	0.9	0.3	0.27	0.9
3	17A	0.4	0.3	0.46	0.5
2	18A	< 1.3	< 0.5	0.79	1.3
3	19B	0.7	0.3	0.77	0.8
1	20A	0.6	0.4	0.61	0.6
3	21A	1.0	0.3	1.29	1.3
3	22A	0.6	0.3	0.56	0.6
3	23A	0.8 (once)	0.3	0.34	0.8
2	24A	< 0.5	< 0.5	0.31	0.5
1	25A	< 0.5	0.4	0.32	0.5
2	26A	< 1.3	< 0.5	0.68	1.3
3	27A	0.6	0.3	0.60	0.6
2	28A	0.3	< 0.5	1.05	1.1
2	29A	1.7 (once)	< 0.5	0.26	1.7
1	30A	< 0.9	0.4	0.05	0.9
1	31A	0.8	0.4	1.40	1.4
2	32A	1.5	< 0.5	0.91	1.5
3	33A	0.5	0.3	0.03	0.5
1	34A	2.0	0.4	0.81	2.0
2	36A	1.0	< 0.5	0.59	1.0
3	37A	0.3	0.3	0.35	0.4
2	38A	0.9	< 0.5	0.55	0.9
2	39A	0.5	< 0.5	0.57	0.6
2	41A	1.8	< 0.5	0.47	1.8

4.3.11 **Two-Copy Dilution Series**

There were 36, 64, and 32 *ACE* results from 4, 8, and 4 replicates at each dilution step for validation studies 1-3 respectively. No results were excluded due to accepted droplet counts below 10,000. The total number of *ACE* results for analysis was therefore 132.

4.3.11.1 *Achieved Concentrations and Linearity*

The median *ACE* starting concentrations for validation studies 1-3 were 196 (192-202), 243 (233-251), and 230 (222-239) respectively.

Achieved dilution step concentrations were consistently below target for validation study 1, consistently on target for validation study 2, and inconsistently below target for validation study 3 (Figure 4-64). The median proportional variation of achieved versus target concentration was 1.59 (1.51-1.77), 1.83 (1.67-1.95), and 1.13 (0.49-1.65) for validation studies 1-3 respectively. Excluding target concentrations ≤ 1 copies/uL, the distribution of proportionate variation of

achieved versus target concentration was consistent across dilution steps for validation study 1 (χ^2 (6, N = 28) = 7.6, p = 0.27) and validation study 2 (χ^2 (6, N = 56) = 3.7, p = 0.72), but not for validation study 3 (χ^2 (6, N = 28) = 24, p < 0.001). Validation study 3 could not be repeated due to insufficient DNA.

Plots of ACE replicate data at each dilution step and corresponding linear models are presented in Figure 4-65. Linear model characteristics for each validation study are presented in Table 4-30. Across the validation studies, ACE performed linearly across the studied concentration range with a median slope of 1.00 (IQR 0.99-1.01) and Pearson’s r 0.997 (IQR 0.997-0.998).

Figure 4-64: Box plot of ACE assay concentration as a proportion of the target concentration at that particular dilution step, grouped by validation study.

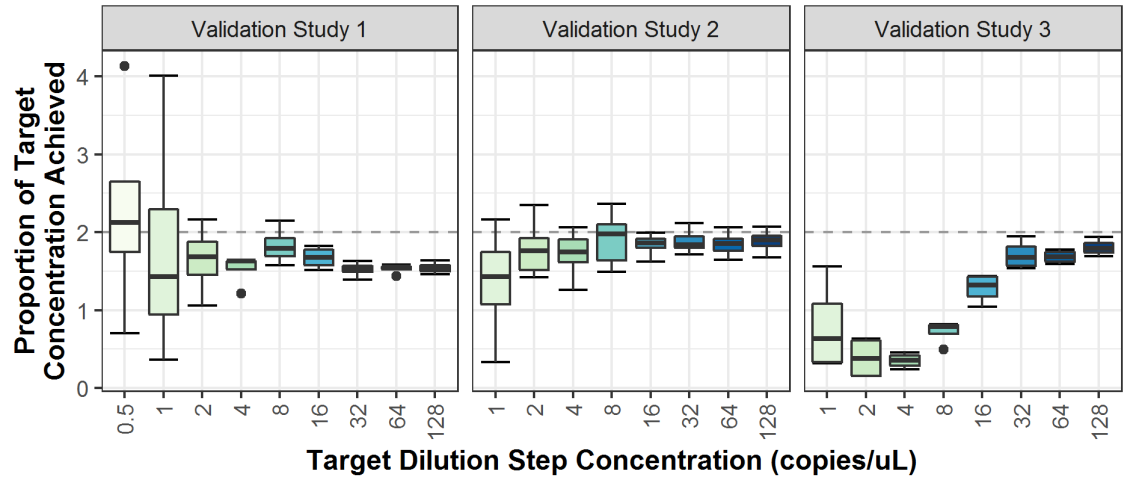


Figure 4-65: Validation study ACE assay replicate results and linear models. Both axes are based on a Log2+1 scale.

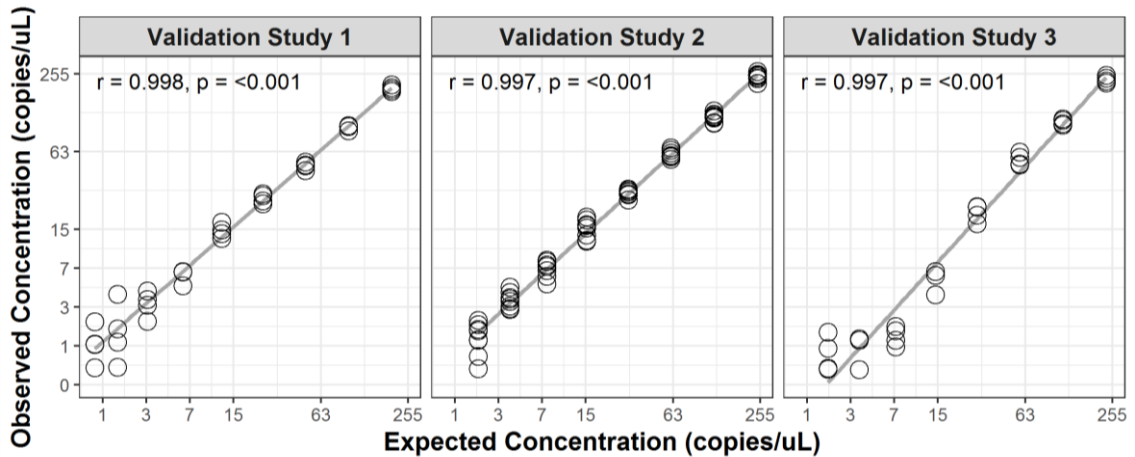


Table 4-30: *ACE* one-copy dilution series linear model parameters, including Pearson's correlation coefficient (r) and Pearson's coefficient of determination (R^2).

Validation Study	Slope	Intercept	r	95% CI	p -value	R^2
1	1.00	0.7	0.998	0.997-0.999	<0.001	0.997
2	0.99	-0.6	0.997	0.995-0.998	<0.001	0.994
3	1.02	-5.6	0.997	0.994-0.999	<0.001	0.994

4.3.11.2 Limit of Detection

All *ACE* replicates in all validation studies detected target down to the lowest concentration dilution step. This concentration was 1.1, 1.4 and 0.6 copies/ μ L for validation studies 1-3 respectively.

At the lowest *ACE* concentration (validation study 3, dilution step 1), the mean concentration was 0.79 copies/ μ L with a standard deviation of 0.59 copies/ μ L. Using the limit of blank determined from no template controls of 0.18 copies/ μ L, the limit of detection for *ACE* was calculated to be 1.1 copies/ μ L.

4.3.11.3 Intra-Assay Variation

The *ACE* intra-assay coefficient of variation increased as the *ACE* concentration reduced (Figure 4-66). The majority of validation study *ACE* assays exhibited intra-assay coefficients of variation below 50% at concentrations above 11 copies/ μ L. The only exception to this was the 1 copy/ μ L dilution step of validation study 1, in which one replicate produced a result of 4 copies/ μ L, resulting in a dilution step coefficient of variation of 86%.

A significant linear relationship existed between the logarithm of the median *ACE* concentration and the logarithm of the coefficient of variation ($r = -0.94$, $p < 0.0001$) (Figure 4-67). Based on this linear model ($F(1,23) = 168$, $p < 0.0001$), *ACE* concentration accounted for 87% of the variance (adjusted R^2) of the intra-assay coefficient of variation.

Figure 4-66: *ACE* intra-assay coefficient of variation (CV) by median *ACE* concentration (ordinal x-axis). The equivalent plasma concentration is not displayed for clarity, but is a factor of 10 times

the *ACE* concentration. Red vertical dashed lines indicate two thresholds below which the median intra-assay variation was consistently greater than the labelled coefficient of variation.

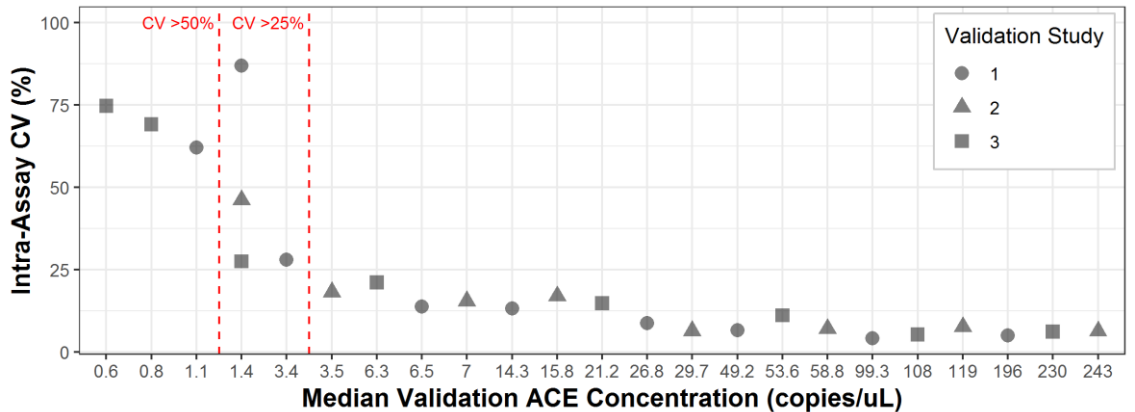
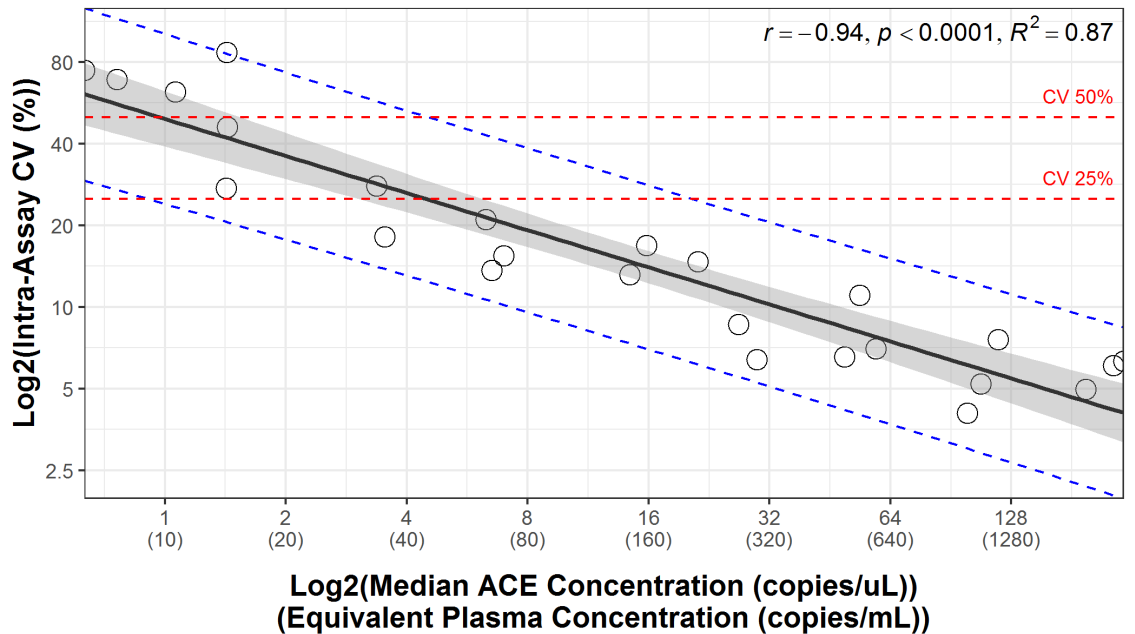


Figure 4-67: Intra-assay coefficient of variation (CV) plotted against *ACE* assay median concentration. Note log2 transformation of both axes. Dark solid line indicates linear regression model, grey shading the 95% confidence interval of the model, blue diagonal dashed lines the 95% prediction interval of the model, and red horizontal dashed lines indicate coefficient of variation thresholds as labelled (presented for clarity given logarithmic axis). In addition to the median assay concentration in the reaction, the equivalent plasma concentration is displayed in parentheses along the x-axis.



4.3.11.4 Limit of Quantification

The limit of quantification for *ACE* was 1.1 copies/μL.

4.4 DISCUSSION

The work presented in this chapter collectively represents an assessment of the analytical validity of the CNV assay panel and *ACE* for the purpose of cfDNA quantification.

4.4.1 Combinatorial Modelling and DNA Fragmentation

Ideally, the analytical validation of an assay panel would include the replicated and repeated assessment of the same sample by all assays comprising the assay panel. This was not possible in this circumstance due to (i) the highly polymorphic nature of sequences targeted by these assays, which are essential for the application of this panel to transplant chimerism assessment, and (ii) the limited quantity of cfDNA present within plasma.

The probability that any particular individual is one-copy for all 41 CNV assays from the developed CNV assay panel is approximately one in ten quadrillion ($1/10^{16}$), and the probability of the same scenario with a zero-copy genotype is one in one trillion ($1/10^{12}$). The solution to this problem was to identify an optimal arrangement of individuals with complementary panel genotypes that permitted study of the broadest number of CNV assays within the fewest number of experiments, thereby maximising the opportunity to compare as many assay results with one another. Three one-copy participant genotypes, and four zero-copy participant genotypes, were identified to permit assessment 35 of the 41 developed CNV assays. Those assays which were not included in these validation study experiments require further study, however it is likely on the basis of the assay development process that performance will be similar to those studied in this chapter.

The concentration of cfDNA extracted from plasma using the presented methodology is approximately 1,500 copies/mL of plasma (Chapter 5), although this varies significantly due to a variety of physiologic and pathologic variables. With a 10% surplus, the number of copies of target sequence required to complete these validation studies was 112,200 copies. To obtain this quantity of cfDNA, 75mL of plasma or 187mL of blood (assuming a haematocrit of 40%) would be required. Whilst technically possible, this was not practical nor was it included in the ethics approval for the study, and so fragmented genomic DNA was used as a surrogate analyte.

As reviewed in Chapter 1, cfDNA consists of fragmented and degraded DNA that is produced predominantly via apoptosis and necrosis in the course of physiologic and pathologic processes.[95, 97, 117, 283] The fragmentation profile of cfDNA in plasma varies depending on the predominant aetiologic mechanism, but typically consists of a number of peaks (“ladder pattern”) with a period of 140-170 bp, attributed to an apoptotic aetiology.[94] There may also be an additional “smear pattern” consisting of longer, more variably fragmented DNA, attributed to a necrotic aetiology.[94]

The fragmentation of genomic DNA used for this analytical validation represents a simplification of the typical pattern of cfDNA. TapeStation electropherograms of extracted cfDNA from clinical samples identified a peak of fragmented DNA at 150-160 bp, and a smear pattern at higher fragment lengths. Periodic peaks were not visible. This may be due to artefact introduced by the extraction method, the aetiology of the cfDNA or because concentrations of these peaks were below the limit of detection.

The fragmentation pattern of genomic DNA obtained using sonic fragmentation matched this observed peak at 150-160 bp, albeit with a wider range of fragment sizes. This is analogous to the predominant peak of cfDNA described by others at a fragment length of 140-170 bp, although given the wider range of fragment sizes arising from sonic fragmentation, some fragments were below the median size of the assay amplicons at 74 bp (up to 10-15%, estimated on the basis of the distribution). The implication of this is that some fragments will have been too small to assay, and it appears this has contributed to increased inter-assay coefficient of variation for assays with higher amplicon lengths. As these assays could not quantify these smaller fragments, they tended to under-estimate the validation dilution step concentration. Given the fragmentation profile of cfDNA is not as broad, the designed CNV assay amplicons should be small enough to avoid this problem in clinical samples.

Fragmentation of DNA may occur at different sites depending on the fragmentation process. Thus, the DNA motifs available for amplicon-based quantification are likely to be different between the sonic fragmentation method employed for this work, and those observed in clinical samples as a result of apoptosis and necrosis, which may further change depending on physiological and pathological events. Furthermore, due to sequence variation between CNV loci and the presence of structurally significant sequences such as repeating elements, the location of shear points and fragment sizes may also vary between CNV loci. The consequence of these considerations is that there is likely to be introduced (from laboratory work) and biological (from aetiology) variation of the analytes of the CNV assays that may contribute to inter-assay variation within and between samples over time. This has implications for the interpretation of variation of CNV assays in clinical samples, but also generalisation of results from this work to the clinical setting.

4.4.2 Dilution Series and Template Volume

The experimental design considered a range of concentrations encountered in clinical samples. The focus was on the graft-derived cfDNA concentration range, from less than 6 copies/mL to 72 copies/mL of plasma in our clinical samples. The dilution series started at a concentration around 1,280 copies/mL of plasma to include concentrations seen in total cfDNA measurements. It should be noted that higher concentrations of total cfDNA up to 20,000 copies/mL have been encountered in clinical samples during states of pathology. While these concentrations were not covered by

this analytical validation study, it is likely that assay performance would be comparable to the higher concentration dilution steps included in this study. This is supported by the modelling of positive droplets that has been conducted, which demonstrates that this concentration range remains well within the dynamic range of the ddPCR platform.

The results of this assay validation work suggest that a key threshold for CNV assay performance exists between an equivalent of 6 copies/mL and 20 copies/mL of plasma. In this range lies the limit of detection, important thresholds of coefficient of variation, and limit of quantification. The clinical work presented later in this thesis also finds that key diagnostic thresholds lie in this range. Due to the design of the dilution series, the number of concentration steps within this range for each validation study was limited. To better define thresholds of assay performance, more detailed study of assay performance in this region is required.

Given the confluence of key analytical and diagnostic thresholds in this low concentration range, consideration could be given to increasing the template DNA concentration to improve the resolution of ddPCR in this range. An additional 2.5 μ L of template DNA (to a total of 7.5 μ L) would be associated with a 32% increase in the rate of droplet positivity in proportion to concentration. Such an approach would improve the discriminative capability of ddPCR from 3.4 copies/mL to 2.3 copies/mL, with improvement of the intra- and inter-assay variation at low concentrations < 20 copies/mL.

A key limitation of a 7.5 μ L template volume is that more sample would be required for cfDNA extraction. 16 PCR wells assaying 31 CNV loci and *ACE* would require 120 μ L of template DNA. Considering that cfDNA from a maximum of 5 mL of sample undergoes extraction in one cfDNA adsorption column, and that this extraction is eluted in 50 μ L, 15 mL of sample would be required. For plasma samples this would be the equivalent of three 9 mL blood collection tubes. Given concerns regarding patient tolerability of this sample volume for serial monitoring, as well as the labour and cost implications of extracting this sample volume through three adsorption columns, a template volume of 7.5 μ L was considered impractical for the purpose of this preliminary work. However, the result of this work suggest that his approach should be revisited, perhaps with rationalisation of the CNV assays to informative assays only, which would mitigate the sample requirement.

A template volume of 2.5 μ L has the inherent advantage that a panel of CNV assays can be run using only 5 mL of sample (one 9 mL blood collection tube). Limitations of this approach include poor resolution and greater confounding from false positive droplets (identified by this work to be uncommon, affecting $\leq 4\%$ of assays) at low concentrations. Higher concentrations of target are not subject to these limitations. Consequently, a template volume of 2.5 μ L would be suitable for genotyping individuals using genomic DNA, or for genotyping recipients using cfDNA. This

template volume would permit genotyping of the CNV assay panel from a single blood collection tube, and could assist with the rationalisation of CNV assays to informative assays only.

An alternative approach to modification of the template volume could involve replicate well summation. In this approach, replicate wells or wells of assays targeting the same genotype are summed and analysed *in toto*, increasing the total droplet number and permitting more granular discrimination of target concentration.[253] Replicated wells would consume more template DNA, but combined analysis of wells targeting the same genotype would not. This approach would require significant adjustments to post-transplant chimeric sample assessment. Either (i) graft informative and recipient one-copy genotypes would need to be established first before manual summation of the wells was performed in the QuantaSoft Analysis Software and relevant concentrations were later calculated, or (ii) the entire pipeline would have to be programmatically accomplished in R. Both adjustments increase the complexity of the bioinformatic component of this work, and for these reasons well summation was considered outside of scope of this preliminary work, but it remains a promising avenue for investigation to address the problem of limiting resolution at low concentrations of target.

4.4.3 **Cluster Centroids and Droplet Thresholding**

A key advantage of ddPCR compared to quantitative real-time PCR is that quantification is based upon a binary measure, the presence or absence of template DNA within a droplet.[45] The identification of whether template DNA is present within a droplet is achieved by detection of fluorescent assay-specific PCR product within the droplet. Thus, whilst PCR reaction efficiency may affect the droplet fluorescence amplitude, it does not typically affect the quantification of target as long as positive droplets are correctly distinguished from negative droplets. This contrasts with quantitative real-time PCR, in which PCR reaction efficiency plays a key role in quantification, and in which definition of the cycle threshold used for quantification may be arbitrary and difficult to define, may vary from experiment to experiment, and substantively impact the result obtained.[273]

This work includes a comprehensive evaluation of positive and negative droplet cluster centroids for each of the CNV assays studied. Clear separation of positive clusters from negative clusters was demonstrated for both FAM- and HEX- labelled assays. The location of these clusters was consistent between replicates and dilution steps, but varied slightly between assays. In the case of the *ACE* assay, coefficient of variation across plates and template DNA was < 4%, and this variation can be attributed to the estimation method used to define cluster centroids. Three assay positive cluster centroids tended to be closer to the negative droplet clusters, possibly due to assay oligonucleotide characteristics, however separation was still sufficient to permit algorithmic thresholding.

Thresholds were more difficult to define when “droplet rain” or “droplet scatter” was present. Such scenarios, typically seen at higher assay concentrations or when multiple wells were selected for thresholding *en bloc*, may lead to droplet misclassification. At higher concentrations, such as those seen in recipient cfDNA genotyping or genomic DNA genotyping, false positive droplets minimally impact results. In circumstances such as graft-derived cfDNA quantification, where the target concentration is low and/or the template volume results in low resolution ddPCR quantification, false positive droplets may confound results by introducing low positive CNV results which may appear to be graft-informative, or by increasing the results of graft-informative assays. If the number of graft informative markers is low, or if there are numerous false positive CNV assays, the median graft-derived cfDNA result may be affected.

No clear guidance is provided for droplet thresholding from Bio-Rad, perhaps because droplet thresholding is assay and application dependent.[253] Flaws exist within the algorithmic defaults used for automated thresholding in the Bio-Rad QuantaSoft analysis software.[284] Given the need for high throughput, ideally algorithmic and objective thresholding, a variety of clearly defined thresholding methods were assessed. Low amplitude threshold methods, particularly those below a third of the distance between the double negative and positive droplet clusters, were associated with an unacceptable rate of false positive results due to droplet misclassification in most instances. Higher threshold techniques, such as the cluster threshold technique, was associated with disparately lower achieved target concentrations than the other methods. The third-distance technique was chosen for the work included in this thesis, in large part because it generates one-copy results dissimilar to the quarter-distance threshold technique, but at the expense of a slightly higher false positive rate than the half-distance threshold technique.

The strengths of ddPCR were exemplified by the limit of blank assessments performed. With the exception of three replicates from CNV assay 03B, the false positive rate was very low (0% for almost all assays using the third-distance threshold technique). CNV 03B exhibited multiple false positive droplets at an amplitude consistent with true positive droplets, suggesting contamination of wells or well-specific PCR reagents (e.g. CNV 03B oligonucleotides) with template DNA that was one-copy for this locus. Interestingly, as discussed in Chapter 2, the CNV 3 locus overlaps with CNV 25. If CNV 03B one-copy template was present in a PCR reagent common to limit of blank wells of both assays, CNV 25A zero-copy control assays should also have been false positive. This was not the case, suggesting that if contamination were present, it was localised to the CNV 03B assay wells only. CNV 25A zero-copy control assays were run on a different plate to CNV 03B assays. Additionally, broader contamination seems unlikely as CNV 03B inter-assay proportionate variation was close to 1, suggesting against persistent over-estimation of one-copy dilution series target as a result of contamination.

4.4.4 **Assay Performance Assessment**

The presented work includes an assessment of assay linear dynamic range, analytical sensitivity (limit of detection), precision (intra-assay variation), accuracy (inter-assay variation), and limit of quantification. Guidelines concerning the measurement, calculation, reporting and assessment of these metrics vary depending on the platform used (ddPCR, chip-based real-time digital PCR, quantitative real-time PCR, sequencing), nature of target (SNP, in-del polymorphism, CNV), nature of template DNA (unprocessed genomic DNA, cfDNA), and application (oncology, genetics, pre-natal diagnostics).[273, 278, 282, 285] There are presently no standards with respect to the analytical validation of transplant chimerism diagnostic tests, although certain characteristics of this application are shared with rare event detection (oncology), copy number detection (genetics), and mosaicism assessment (genetics). Additionally, there is presently no gold standard test for transplant cfDNA chimerism assessment, against which to benchmark assay performance. Finally, there is presently no directly equivalent assay used for chimerism analysis that has been subject to analytical validation against which to compare performance.

For these reasons, and with consideration of the unique challenges faced conducting analytical validation experiments on cfDNA discussed earlier, the presented analytical validation of the CNV assay panel differs from standard approaches in a number respects.

Firstly, the inter-assay coefficient of variation has been calculated on the basis of variation between assays targeting different CNV polymorphisms at the same point in time, rather than repeated measurement of the same assay on the same sample at different points in time. The inter-assay coefficient of variation thus provides an estimate of the variation of multiple independent assays targeting the same analyte (one-copy fragmented DNA) via different target sequences. This approach is arguably the best assessment of accuracy in the absence of a gold standard, but the issue of assay-specific reproducibility is not assessed. Further work to establish the reproducibility inter-assay coefficient of variation would require repeated dilution series assessment over time using a consistent dilution series template, or simultaneous collection of multiple cfDNA blood collection tubes that undergo separate extraction and CNV assay panel runs.

Secondly, the application of intra- and inter-assay coefficient of variation thresholds for limit of quantification were liberal, using a cut-off of 50% or greater. “Functional sensitivity” is generally defined as the concentration at which coefficient of variation is 20%, however it is recognised that predefined goals of bias and precision that are “fit for purpose” of the assay may modify the coefficient of variation cut-off for limit of quantification.[282]

A coefficient of variation threshold of 50% is more appropriate for this CNV assay panel approach as the final result is expected to be calculated from the median of multiple independent

informative markers, each assaying the same analyte via a different target sequence. Every additional independent informative estimate of the target concentration provides proportionately more certainty of the final result. This is particularly important given the fragmented and degraded nature of cfDNA and the non-amplified, absolute (rather than relative) quantification approach employed. This differs from standard assays (such as a troponin assay), which consist of a single estimate of a biologically consistent target, in which parameters of variation would understandably be more important.

Finally, the focus of this analytical validation was on assay performance. This work did not consider pre-analytical factors affecting cfDNA, in particular plasma cfDNA extraction. The yield and quality of cfDNA extraction is subject to variation, in particular between extraction methodologies.[286-288] The QIAamp circulating nucleic acid kit employed for this work appears to be a top performer in this regard, although most assessments performed to date have used moderate concentrations of target sequence.[286, 288]

Given the absolute quantification methodology proposed, extraction yield is expected to play an important role in determining results. The present implementation of the CNV assay assumes consistent extraction efficiency between extraction runs based upon strict adherence to an extraction protocol using the same cfDNA extraction product. Variation from this practice, without correction for extraction yield, would invalidate established diagnostic thresholds (presented later in this thesis).

An alternative approach employed by others involves spiking the plasma sample with a known quantity of artificial DNA not present in the human genome, and quantifying this spike after extraction, correcting any results for extraction efficiency.[120] This has the advantage that it offers an sample-specific correction for yield, albeit failing to evaluate the quality of DNA extracted. The disadvantage is additional complexity introduced into the methodology, which would include careful consideration of the volumes subject to extraction, as well as assumption that the spiked DNA is subject to the same biases introduced by the extraction and ddPCR quantification as other fragmented DNA.

4.4.5 **Summary**

The analytical validation work presented in this chapter indicates that the CNV assay panel performs adequately for the quantification of fragmented DNA down to a concentration of approximately 8 copies/mL plasma. Modifications to the assay methodology, such as an increase of template volume to 7.5 μ L, and incorporation of a spiked sample to correct for extraction efficiency, could be included in further work to improve performance of the assay for diagnostic applications. In such an instance, further analytical validation may be warranted, with greater

focus on the concentration range between 8 copies/mL and 30 copies/mL of plasma, within which lies important assay performance and diagnostic thresholds.

CHAPTER 5 KIDNEY TRANSPLANT REJECTION DIAGNOSIS

5.1 ABSTRACT

Graft-derived cell-free DNA (donor-derived cell-free DNA) is an emerging marker of kidney allograft injury. Studies examining the clinical validity of this biomarker have previously used the graft fraction, or proportion of total cell-free DNA that is graft-derived. The present study evaluated the diagnostic validity of absolute measurements of graft-derived cell-free DNA, as well as calculated graft fraction, for the diagnosis of graft dysfunction. Plasma graft-derived cell-free DNA, total cell-free DNA and graft fraction were correlated with biopsy diagnosis as well as individual Banff scores. 61 samples were included in the analysis. For the diagnosis of antibody mediated rejection, the receiver-operator characteristic area under the curves of graft-derived cell-free DNA and graft fraction were 0.91 (95% CI 0.82-0.98) and 0.89 (95% CI 0.79-0.98), respectively. Both measures did not diagnose borderline or type 1A cellular mediated rejection. Graft fraction was associated with a broader range of Banff lesions, including lesions associated with cellular mediated rejection, while graft-derived cell-free DNA appeared more specific for antibody mediated rejection. Limitations of this study include a small sample size and lack of a validation cohort. The capacity for absolute quantification, and lower barriers to implementation of this methodology recommend it for further study.

5.2 MAIN BODY TEXT

5.2.1 Introduction

An unmet need in kidney transplantation is the ability to identify rejection non-invasively, so permitting frequent testing, facilitating early diagnosis and treatment, as well as reducing the requirement for unnecessary graft biopsy. To address this problem, a wide variety of putative biomarkers are under investigation, many of which employ multi-analyte technologies within the fields of genomics, transcriptomics, proteomics and metabolomics.

Within genomics, the detailed description of common genetic variation between individuals, combined with technologies that permit measurement of low concentrations of target DNA sequence, has permitted distinction and quantification of DNA sequences unique to the allograft within the plasma of the recipient. This has been termed donor-derived cell-free DNA (dd-cfDNA), but it is perhaps more accurately called graft-derived cell-free DNA (gd-cfDNA).

Cell-free DNA (cfDNA) is fragmented, degraded DNA that is detectable in body fluids, including plasma and urine. The majority of detectable cfDNA in plasma is thought to arise from cell-turnover within the haematopoietic system.[101] Total cfDNA levels may rise in a variety of physiologic and pathologic states. These include exercise, malignancy, sepsis, myocardial

infarction, stroke and critical illness.[119, 289] Higher levels of total cfDNA have been associated with increased severity or poorer prognosis in some of these diseases.[119, 289]

In transplantation, it has been proposed that injury to the allograft results in increased release of gd-cfDNA into recipient plasma via graft-cell apoptosis and necrosis. Whilst the clinical validity of gd-cfDNA within kidney transplantation has been explored with variable success for some time, it is only recently that larger scale prospective clinical studies have been undertaken to support diagnostic validity for this biomarker and recommend translation into clinical practice.[19, 153, 261]

To date, there has been one large published prospective clinical study evaluating the diagnostic validity of plasma gd-cfDNA for allograft dysfunction in kidney transplantation. Bloom et al. used single nucleotide polymorphisms (SNPs) and sequencing of cfDNA to calculate the graft fraction (proportion of total cfDNA that is graft-derived) in 102 adult kidney transplant recipients with graft dysfunction, and observed higher levels in AMR and more severe cellular mediated (CMR) rejection.[153] A possible limitation of the method raised by this group is the potential for confounding of graft fraction by concurrent elevations in total cfDNA, an issue also observed in non-invasive prenatal testing.[290]

The present study employed droplet digital PCR and highly heterozygous copy number variation (CNV) DNA sequences to generate a negative background against which absolute quantification of gd-cfDNA could be performed. The objective was to prospectively evaluate the performance of gd-cfDNA for diagnosis of AMR in an adult kidney transplant cohort with allograft dysfunction who were undergoing graft biopsy. To permit comparison with other approaches, graft fraction was also calculated and evaluated alongside gd-cfDNA, and the performance of both methods described.

5.2.2 **Materials and Methods**

5.2.2.1 *Study Design*

A non-experimental prospective cross-sectional study design was used. Study sample size was calculated using a method described by Buderer et al. with a maximum clinically-acceptable two-tailed 95% confidence interval width of 0.1, estimated disease prevalence of 0.3, expected sensitivity of 0.9, and expected specificity of 0.8.[291] Recruitment occurred at two transplanting centres in Melbourne, Australia. Institutional ethics approval was obtained at both sites and all participants provided written informed consent. Adult kidney transplant recipients undergoing standard-of-care allograft biopsy for investigation of graft dysfunction were included in the study. Exclusion criteria were multi-organ transplant, pregnancy, biopsy within 3 weeks of transplant,

and critical illness at the time of biopsy. A separate group of samples were collected concurrently for assay development and validation.

5.2.2.2 Histopathologic Diagnosis

Allograft biopsy was performed according to local practice using ultrasound-guided needle biopsy. Histopathologic examination of biopsy samples was undertaken by institutional transplant pathologists that were blinded to study participation. Allograft pathology was classified and reported according to the Banff schema.[292] Mixed rejection and ‘suspicious for AMR’ diagnoses were classified under the corresponding AMR diagnosis. In addition to Banff diagnoses, individual Banff scores were collected.

5.2.2.3 Plasma cfDNA Extraction

Blood was collected in two 10ml Streck Cell-Free DNA blood tubes (Streck, Omaha, NE, USA) by peripheral venesection within a week of renal biopsy. Samples were stored at room temperature. Plasma separation was performed within 7 days of sample collection by centrifugation for 10 minutes at 1,600 x g, following which the supernatant was centrifuged for 10 minutes at 16,000 x g. Plasma was combined and stored at -80 °C until cfDNA extraction. cfDNA extraction was performed using the QIAamp® Circulating Nucleic Acid Kit (QIAGEN GmbH, Hilden, Germany) according to the manufacturer’s instructions, and an elution volume of 50 µL distilled water per column. A maximum of 5 mL plasma was used per cfDNA extraction column, after which the eluted cfDNA was combined (a single cfDNA pool per participant-biopsy) prior to storage at -80 °C.

5.2.2.4 Assay Design

A total of 38 CNV loci were identified using a population genomics approach to maximise informativity, as described previously.[145, 293] ddPCR primers and Zen double-quenched probes (Integrated DNA Technologies, Baulkham Hills, Australia) were designed within each of the CNV loci in the panel. All targets were within agenic or intronic regions. A non-polymorphic control assay was also developed within the angiotensin I converting enzyme (*ACE*) gene (Online Mendelian Inheritance of Man no. 106180). Analytical validation (manuscript in preparation) limit of blank 0 copies/mL, limit of detection ~6 copies/mL, limit of quantification (inter- and intra-assay coefficient of variation < 50%) ~8-10 copies/mL, and linear performance demonstrated from 8 to 1280 copies/mL.

5.2.2.5 Droplet Digital PCR

Extracted cfDNA was run against the CNV assays using the Bio-Rad QX200 Droplet Digital System (Bio-Rad Laboratories, Gladesville, Australia). 31 of the 38 CNV assays developed as well as the *ACE* non-polymorphic control were run in a duplex arrangement using paired 6-

carboxyfluorescein and hexachloro-fluorescein labelled hydrolysis probes. The PCR reaction consisted of 12.5 μ L of Bio-Rad ddPCR Supermix for Probes (no deoxyuridine triphosphate), 2.25 μ L of 10 μ mol/L forward and reverse primers for both duplexed assays, 0.65 μ L each of the 10 μ mol/L duplexed probes, 2.5-5 μ L of extracted cfDNA and distilled water to make the reaction volume 25 μ L. PCR-cycling conditions were 95 °C for 10 minutes, 44 cycles of 94 °C for 30 seconds, 60 °C for 1 minute, 98 °C for 10 minutes, then an indefinite hold at 4 °C. ddPCR data were processed using the Bio-Rad QuantaSoft software (version 1.6). The method for assigning ddPCR thresholds for droplet positivity, as well as representative ddPCR 1D and 2D plots, are presented in the Supplementary Material, and Supplementary Figure 5-8 and Figure 5-9.

5.2.2.6 *Calculation of Plasma Concentration*

CNV assay concentrations within the PCR reaction volume measured by ddPCR were converted to the equivalent concentration in copies/mL of plasma that underwent cfDNA extraction (Supplementary Material).

5.2.2.7 *Genotyping & cfDNA Measurement*

Genotyping and cfDNA measurement was performed blinded to biopsy diagnosis. Recipient positive markers (CNV loci that are present within the recipient genome) were genotyped based upon their concentration relative to the *ACE* non-polymorphic marker concentration, and other genotypes (Supplemental Figure 5-10). The median recipient one-copy marker concentration was calculated as the total cfDNA concentration. Detectable markers at concentrations less than 10% of total cfDNA concentration were genotyped as graft informative markers (Supplemental Figure 5-10). No donor genomic DNA or non-chimeric recipient genomic DNA was therefore required for genotyping. The median graft informative assay concentration was calculated as the gd-cfDNA concentration. Graft fraction was calculated by dividing the gd-cfDNA concentration by total cfDNA concentration. Inter-assay coefficient of variation (CV) was calculated by dividing the standard deviation by the mean.

5.2.2.8 *Diagnostic Performance Assessment*

Receiver operator characteristic (ROC) area under the curves (AUC) were calculated for gd-cfDNA and graft fraction for the diagnosis of acute AMR (aAMR), and a composite diagnosis of aAMR and chronic active AMR (caAMR). ROC curves were compared between gd-cfDNA and graft fraction using the DeLong method and a one sided significance test. In comparisons with serum creatinine ROC curves, a bootstrap test was employed to compare against curves with an AUC < 0.5. Diagnostic performance parameters were calculated using diagnostic thresholds that were calculated to maximise sensitivity and specificity.[294]

5.2.2.9 Statistical Methods

Non-parametric distributions were compared using the Wilcoxon rank sum test (two groups), and the Kruskal-Wallis rank sum test (multiple groups). Post-hoc analysis for the latter was performed using Dunn's test with Holm-Šidák adjustment. Spearman's rank correlation was used to assess relationships between ordinal and continuous variables. Pearson's product-moment correlation was used to measure relationships between continuous variables. A p value < 0.05 was considered significant. All statistical analyses were performed using R.[281]

5.2.3 Results

5.2.3.1 Study Cohort

Seventy (20mL) blood samples meeting study criteria were collected. Two samples were duplicated by rescheduled biopsies, two samples were from non-diagnostic biopsies, four samples failed cfDNA extraction, and one sample was damaged in transport. Sixty-one sample results, collected from fifty-five patients were, therefore, used for the study (Figure 5-1).

The participant (Table 5-1, Supplemental Table 5-4) and sample (Table 5-2, Supplemental Table 5-5) characteristics are presented. The cohort was predominantly Caucasian, deceased donor recipients, and recipients of first grafts. The graft dysfunction indication for biopsy was creatinine in 92%, proteinuria in 3% and both in 2%. The median creatinine at sample collection was 165 (IQR 120-184) $\mu\text{mol/L}$, and median eGFR was 40 (IQR 29-54) ml/min . Duration since transplant was less than 6 months for 23% of samples, 6-12 months for 20%, 1-5 years for 20%, 5-10 years for 15% and more than 10 years in 23%. At the time of sample collection, the predominant immunosuppression regimen was prednisolone (92%), mycophenolate (82%) and tacrolimus (79%).

ABMR was more likely in participants recruited from the one recruitment centre ($p = 0.02$), recipients of third grafts ($p = 0.03$) and individuals biopsied for creatinine (rather than proteinuria, $p = 0.02$). There were no cohort differences between the aAMR group and no aAMR groups.

Figure 5-1: Study sample flow diagram. Study inclusion and exclusion criteria were met for 70 samples, 9 of which were not included in the clinical study due to duplication and technical issues. The final study cohort included 61 samples from 55 participants. Of these, 13 samples were included in the ABMR cohort and, within this cohort, 9 samples were included in the aAMR cohort. The no aAMR cohort was therefore comprised of 48 no ABMR and 3 caAMR samples.

The no ABMR cohort comprised 48 samples, 28 of which were negative biopsies with no other pathology evident.

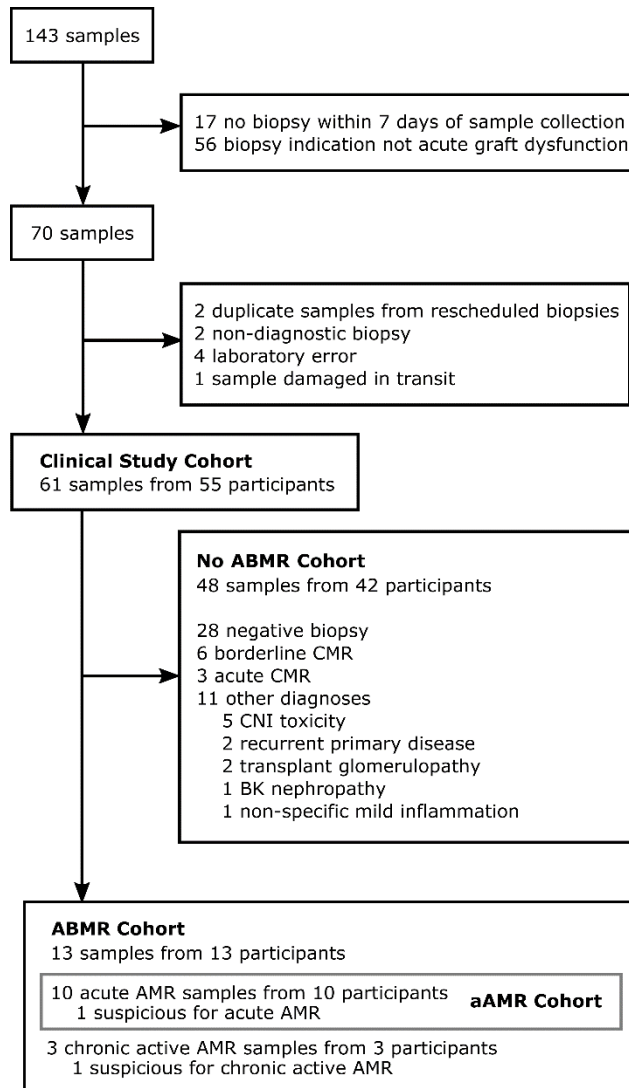


Table 5-1: Participant characteristics.

	Participant Cohort (N=55)	No ABMR (N=42)	ABMR (N=13)	p-value ⁱ
Demographics				
Age at first sample (years)	53 (39-61)	55 (39-61)	49 (40-55)	0.49
Male sex	30 (55%)	23 (55%)	7 (54%)	>0.99
Race				0.32
Caucasian	44 (80%)	34 (81%)	10 (77%)	
Asian	7 (13%)	6 (14%)	1 (8%)	
Polynesian	2 (4%)	1 (2%)	1 (8%)	
Arabic	1 (2%)	0 (0%)	1 (8%)	
African	1 (2%)	1 (2%)	0 (0%)	
Primary disease				0.63
IgA nephropathy	14 (25%)	11 (26%)	3 (23%)	
Reflux nephropathy	8 (15%)	4 (10%)	4 (31%)	
Primary FSGS	6 (11%)	5 (12%)	1 (8%)	
Polycystic kidney disease	6 (11%)	4 (10%)	2 (15%)	
Secondary FSGS	4 (7%)	3 (7%)	1 (8%)	
Diabetes	3 (5%)	3 (7%)	0 (0%)	

Unknown	2 (4%)	2 (5%)	0 (0%)	
ANCA vasculitis	2 (4%)	1 (2%)	1 (8%)	
Hypertension	2 (4%)	2 (5%)	0 (0%)	
Other	8 (15%)	7 (17%)	1 (8%)	
Recruiting hospital				0.02
Hospital 1	49 (89%)	40 (95%)	9 (69%)	
Hospital 2	6 (11%)	2 (5%)	4 (31%)	
Transplant				
Graft number				0.03
First	51 (93%)	40 (95%)	11 (85%)	
Second	2 (4%)	2 (5%)	0 (0%)	
Third	2 (4%)	0 (0%)	2 (15%)	
Donor Type				>0.99
Deceased	40 (73%)	30 (71%)	10 (77%)	
Brain death	19 (35%)	14 (33%)	5 (38%)	
Circulatory death	13 (24%)	11 (31%)	2 (15%)	
Unknown	8 (15%)	5 (12%)	3 (23%)	
Living	15 (27%)	12 (29%)	3 (23%)	
Parent	7 (13%)	6 (14%)	1 (8%)	
Sibling	2 (4%)	2 (5%)	0 (0%)	
Related NOS	1 (2%)	0 (0%)	1 (8%)	
Not related	5 (9%)	4 (10%)	1 (8%)	
Matched donor-recipient sex	25/51 (49%)	17/39 (43%)	8/12 (66%)	0.20
ABO compatible	52 (95%)	40 (95%)	12 (92%)	0.56
DSA at transplant	10/47 (21%)	6/34 (18%)	4/13 (31%)	0.43
HLA mismatch				0.53
5-6	12 (22%)	8 (19%)	4 (31%)	
3-4	25 (45%)	21 (50%)	4 (31%)	
1-2	16 (29%)	12 (29%)	4 (31%)	
0	2 (4%)	1 (2%)	1 (8%)	
Induction agent				0.49
Basiliximab	42 (76%)	32 (76%)	10 (77%)	
None	4 (7%)	4 (10%)	0 (0%)	
OKT3	1 (2%)	1 (2%)	0 (0%)	
Unknown	8 (15%)	5 (12%)	3 (23%)	

Data ranges are presented as median (inter-quartile range).

ⁱ The *p*-value presented is the significance of differences between the ABMR and no ABMR groups. The Wilcoxon rank sum test, Fisher's exact test for count data, and the equality of proportions test were used for comparisons as indicated.

FSGS, focal segmental glomerulosclerosis; ANCA, anti-neutrophil cytoplasmic antibody; NOS, not otherwise specified; DSA, donor specific antibody; HLA, human leukocyte antigen; OKT3, muromonab-CD3

Table 5-2: Sample characteristics.

	Sample Cohort (N=61)	No ABMR (N=48)	ABMR (N=13)	p-value ⁱ
Sample Collection				
Time post-transplant (years)	2.6 (0.6-8.2)	2.3 (0.6-8.3)	4.0 (0.6-8.2)	0.91
Time post-transplant cohorts (years)				0.89
< 0.5	14 (23%)	11 (23%)	3 (23%)	
0.5-1	12 (20%)	10 (21%)	2 (15%)	
1-5	12 (20%)	10 (21%)	2 (15%)	
5-10	9 (15%)	6 (13%)	3 (23%)	
>10	14 (23%)	11 (23%)	3 (23%)	
Collection time from biopsy (days)				0.76
-6 (before biopsy)	1 (2%)	1 (2%)	0 (0%)	

-1	1 (2%)	1 (2%)	0 (0%)	
0 (day of biopsy)	57 (93%)	44 (92%)	13 (100%)	
+1 (after biopsy)	2 (3%)	2 (4%)	0 (0%)	
Clinical Characteristics				
Biopsy indication				0.02
Creatinine	56 (92%)	45 (93%)	11 (85%)	
Proteinuria	3 (3%)	3 (6%)	0 (0%)	
Both	2 (5%)	0 (0%)	2 (15%)	
Baseline creatinine (μmol/L) (N=60) ⁱⁱ	125 (102-156)	121 (102-157)	140 (121-156)	0.39
Baseline eGFR (ml/min) (N=60) ⁱⁱ	52 (42-65)	52 (43-65)	51 (33-65)	0.72
Creatinine (μmol/L)	165 (120-184)	159 (114-185)	168 (147-182)	0.34
eGFR (ml/min)	40 (29-54)	40 (29-54)	36 (29-43)	0.41
Indication urine protein:creatinine (g/mmol) (N=5)	0.21 (0.20-0.35)	0.21 (0.17-0.28)	0.48 (0.34-0.62)	0.80
Urine protein:creatinine (g/mmol) (N=40)	0.03 (0.01-0.09)	0.03 (0.01-0.07) (N=32)	0.09 (0.02-0.19) (N=8)	0.29
Any history of donor specific antibody	18/35 (51%)	11/24 (46%)	7/11 (64%)	0.47
Maintenance immunosuppression				
Prednisolone	56 (92%)	44 (92%)	12 (92%)	>0.99
Mycophenolate	50 (82%)	38 (79%)	12 (92%)	0.43
Azathioprine	6 (10%)	5 (10%)	1 (8%)	>0.99
Tacrolimus	48 (79%)	37 (77%)	11 (85%)	0.72
Ciclosporin	10 (16%)	9 (19%)	1 (8%)	0.67
mTOR	4 (7%)	4 (8%)	0 (0%)	0.57
Drug levels				
Tacrolimus level (μg/L) (N=48)	5.3 (4.2-6.3)	5.2 (4.2-6.0) (N=37)	5.7 (4.1-7.9) (N=11)	0.49
Ciclosporin level (μg/L) (N=10)	318 (207-490)	388 (200-521) (N=9)	227 (N=1)	0.8
Weight (kg) (N=56)	78 (69-85)	78 (68-85) (N=46)	80 (75-85) (N=10)	0.45

Data ranges are presented as median (inter-quartile range).

ⁱ The *p*-value presented is the significance of differences between the ABMR and no ABMR groups. The Wilcoxon rank sum test, Fisher's exact test for count data, and the equality of proportions test were used for comparisons as indicated.

ⁱⁱ One participant had not achieved a baseline creatinine by the time of sample collection.
eGFR, estimated glomerular filtration rate (CKD-EPI); mTOR, mammalian target of rapamycin.

5.2.3.2 Assay Genotyping

Assay genotyping yield and informative marker frequencies, including the effect of donor-recipient relationship and gd-cfDNA concentration, are presented in the Supplementary Material. Median inter-assay CV for graft informative, recipient 1-copy and recipient 2-copy assays was 47% (IQR 36-63%), 14% (IQR 11-19%) and 10% (IQR 7-13%) respectively. The inter-assay CV inversely correlated with the logarithm of the concentration ($r = -0.70$, $N = 159$, $p < 0.0001$) (Supplemental Figure 5-15).

5.2.3.3 Cell-Free DNA and Biopsy Diagnosis

There were 22 biopsies with Banff rejection diagnoses, including 6 borderline CMR (bCMR), 3 acute CMR (aCMR), 3 caAMR and 10 aAMR. Two suspicious for AMR diagnoses were included under the corresponding AMR diagnosis. Non-rejection diagnoses included 1 BK nephropathy, 2 transplant glomerulopathy, 1 non-specific mild inflammation, 5 calcineurin inhibitor (CNI) toxicity, 2 recurrent primary disease, and 28 negative biopsies.

Gd-cfDNA concentration ranged from <6-72 copies/mL and varied between biopsy diagnoses ($\chi^2(5) = 22, p = 0.0006$) (Figure 5-2, left panel). Gd-cfDNA was significantly elevated in aAMR compared to (i) negative biopsy (30 (IQR 23-48) vs 7 (IQR 5-11) copies/mL, $p = 0.0004$) and (ii) bCMR (30 (IQR 23-48) vs 4 (IQR 4-10) copies/mL, $p = 0.01$). There was also a trend toward elevation in aAMR compared to other group diagnoses (30 (IQR 23-48) vs 8 (IQR 6-17) copies/mL, $p = 0.051$). Combining aAMR and caAMR into one entity (ABMR), the distribution of gd-cfDNA concentrations was significantly higher than all other diagnoses combined (25 (IQR 22-35) vs 7 (IQR 5-12), $W = 59, p < 0.0001$) (Figure 5-3, left panel).

Total cfDNA concentration ranged from 177-20542 copies/mL and was not significantly different between biopsy diagnoses ($\chi^2(5) = 105, p = 0.07$) (Figure 5-2, middle panel), or for ABMR diagnosis (1266 (IQR 1073-2038) vs 1671 (IQR 968-4038), $W = 356, p = 0.45$) (Figure 5-3, middle panel). Extreme total cfDNA levels coincided with inflammatory illness, including *E. coli* bacteraemia (20542 copies/mL), haemoptysis (19651 copies/mL), unexplained fevers (17341 copies/mL), and foot inflammation (17341 copies/mL).

Graft fraction ranged from <0.04-7.6% and varied between biopsy rejection diagnoses ($\chi^2(5) = 28, p < 0.0001$) (Figure 5-2, right panel). Graft fraction was elevated in aAMR compared to (i) negative biopsy (2.4% (IQR 0.9-3.0%) vs 0.3% (IQR 0.1-0.6%), $p = 0.0002$), and (ii) bCMR (2.4% (IQR 0.9-3.0%) vs 0.37% (IQR 0.1-0.5%), $p = 0.046$). Graft fraction was also increased in caAMR compared to negative biopsy (2.5% (IQR 1.7-3.2%) vs 0.3% (IQR 0.1-0.6%), $p = 0.043$), and other compared to negative biopsy (1.0% (IQR 0.4-1.5%) vs 0.3% (IQR 0.1-0.6%), $p = 0.048$). For the combined diagnosis of ABMR, the distribution of graft fraction results was higher than all other diagnoses combined (2.5% (IQR 0.8-3.0%) vs 0.4% (IQR 0.1-0.8%), $W = 70, p < 0.0001$) (Figure 5-3, right panel).

There was moderate-strong correlation between gd-cfDNA levels and graft fraction ($r = 0.71, p < 0.0001$) (Supplemental Figure 5-16). Half of the variance in graft fraction was not explained by gd-cfDNA levels ($r^2 = 0.50$).

The distribution of gd-cfDNA and graft fraction results for all samples (Supplemental Figure 5-17), and negative biopsy samples alone (Supplemental Figure 5-18), were not statistically

significantly different if the sample was collected after biopsy compared to before biopsy (Supplementary Material).

With increasing time following transplantation, total cfDNA levels tended to be lower ($r = -0.32$, $p = 0.01$), but not gd-cfDNA ($r = 0.02$, $p = 0.85$) or graft fraction ($r = 0.17$, $p = 0.18$) (Supplemental Figure Figure 5-19). For negative biopsy samples alone, stable gd-cfDNA concentrations ($r = -0.12$, $p = 0.54$) and a decline in total cfDNA concentrations ($r = -0.37$, $p = 0.05$) resulted in a rise in graft fraction with increasing time following transplantation ($r = 0.43$, $p = 0.02$) (Supplemental Figure 5-20).

Figure 5-2: Sample measurements of gd-cfDNA (left panel), total cfDNA (middle panel) and calculated graft fraction (right panel) grouped by biopsy diagnosis. Gd-cfDNA levels were significantly elevated in aAMR compared to negative biopsy, bCMR and other diagnoses. In contrast, graft fraction levels in caAMR, aAMR and other diagnoses were significantly elevated compared to negative biopsies. Hollow circles represent individual results.

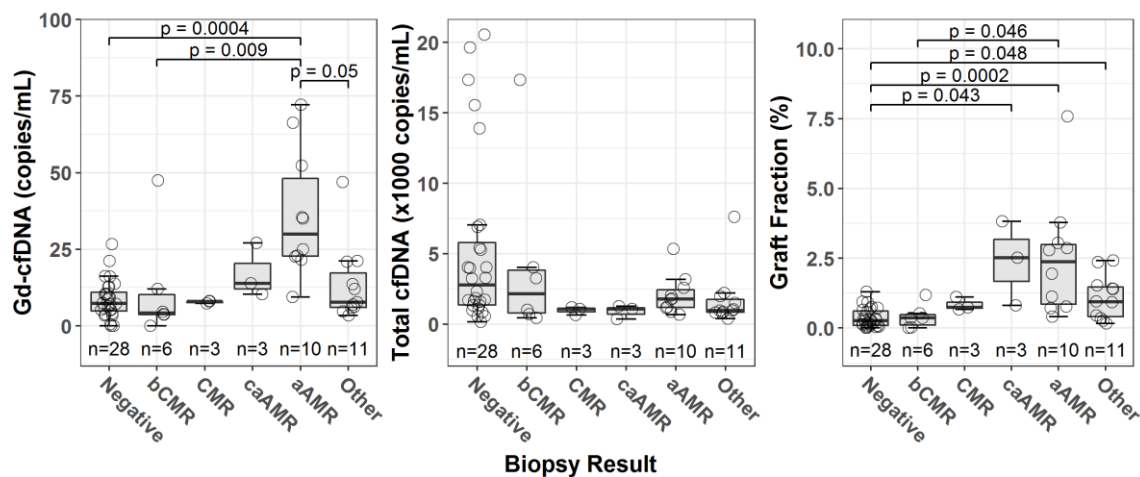
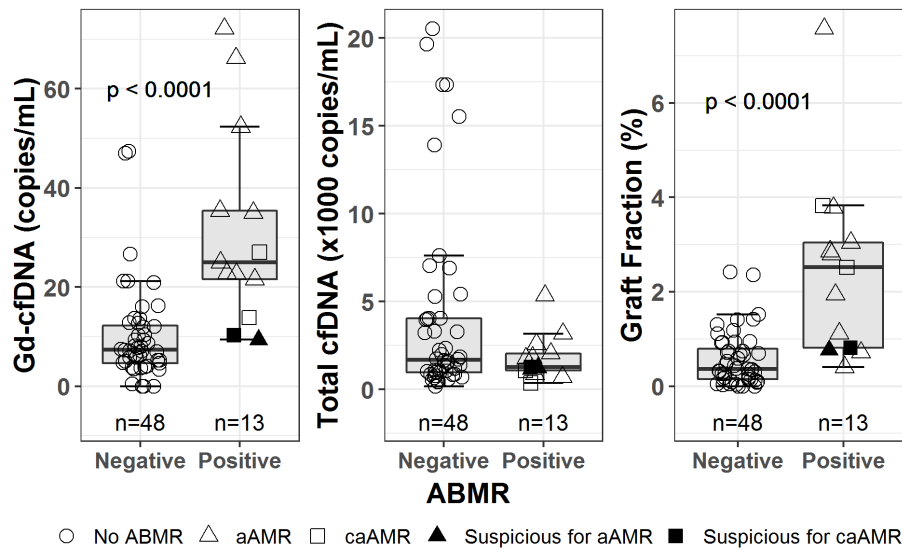


Figure 5-3: Sample measurements of gd-cfDNA (left panel), total cfDNA (middle panel) and calculated graft fraction (right panel) grouped by the presence (positive) or absence (negative) of ABMR. Gd-cfDNA and graft fraction were significantly elevated in ABMR. Points are individual

results and point shapes represent type of AMR. Suspicious for ABMR diagnoses had some of the lowest levels with the ABMR group (black points).



5.2.3.4 Diagnostic Performance

ROC AUC of gd-cfDNA and graft fraction for diagnosis of aAMR were 0.92 (95% CI 0.84-1.00) and 0.85 (95% CI 0.73-0.97), respectively (Figure 5-4, left panel). For the composite diagnosis of ABMR, ROC AUC for gd-cfDNA and graft fraction were 0.91 (95% CI 0.82-0.98) and 0.89 (95% CI 0.79-0.98), respectively (Figure 5-4, right panel). Compared with graft fraction, there was a non-significant trend towards better performance of gd-cfDNA for diagnosis of aAMR (AUC 0.92 vs 0.85, $Z = 1.2$, $p = 0.11$) but not the composite diagnosis of ABMR (AUC 0.91 vs 0.89, $Z = 0.4$, $p = 0.36$).

Using a diagnostic threshold of 21 copies/mL for gd-cfDNA and 0.75% for graft fraction, the diagnostic odds ratio for aAMR was 69 (95% CI 7-630) and 10 (95% CI 2-50), respectively. Diagnostic odds ratios for ABMR were 21 (95% CI 4-109) and 17 (95% CI 3-85) using thresholds of 13 copies/mL and 0.75%, respectively. Sensitivity, specificity, predictive values and likelihood ratios for both measures in diagnosis of aAMR and ABMR are presented in Table 5-3.

Other group diagnoses that met both gd-cfDNA and graft fraction diagnostic cutoffs for ABMR were one case each of diabetic glomerulopathy (21 copies/mL, 2.4%) and mild non-specific chronic inflammation (47 copies/mL, 2.4%), and two cases of transplant glomerulopathy (14 copies/mL, 1.4%; 21 copies/mL, 1.4%). Additional diagnoses meeting graft fraction diagnostic cutoffs but not gd-cfDNA cutoffs for ABMR were two cases of chronic CNI toxicity (0.9%; 0.9%), and one case of BK nephropathy (1.5%).

Figure 5-4: Receiver-operator characteristic curves for gd-cfDNA (black line), graft fraction (light gray line), and creatinine (dark gray dashed line) in the diagnosis of aAMR (left panel), and

composite aAMR and caAMR (ABMR, right panel). ROC AUC point estimates are presented. Compared with graft fraction, there was a non-significant trend towards better performance of gd-cfDNA for diagnosis of aAMR ($p = 0.11$) but not the composite diagnosis of ABMR ($p = 0.36$). Both measures performed better than serum creatinine for both diagnoses ($p \leq 0.001$).

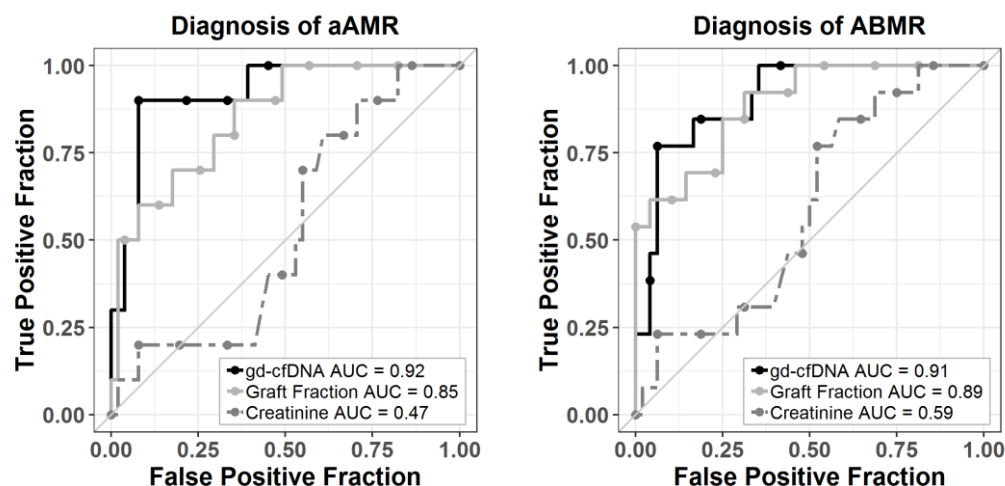


Table 5-3: Diagnostic performance characteristics of gd-cfDNA and graft fraction for the diagnosis of aAMR and ABMR.

	Diagnosis of aAMR		Diagnosis of ABMR	
	gd-cfDNA > 21 copies/mL	Graft Fraction > 0.75%	gd-cfDNA > 13 copies/mL	Graft Fraction > 0.75%
Sensitivity	0.90 (0.55-1.00)	0.80 (0.44-0.97)	0.85 (0.55-0.98)	0.85 (0.55-0.98)
Specificity	0.88 (0.76-0.96)	0.71 (0.56-0.83)	0.79 (0.65-0.90)	0.75 (0.60-0.86)
PPV	0.60 (0.32-0.84)	0.35 (0.16-0.57)	0.52 (0.30-0.74)	0.48 (0.27-0.69)
NPV	0.98 (0.88-1.00)	0.95 (0.82-0.99)	0.95 (0.83-0.99)	0.95 (0.82-0.99)
PLR	7.65 (3.51-16.68)	2.72 (1.61-4.60)	4.06 (2.23-7.39)	3.38 (1.97-5.82)
NLR	0.11 (0.02-0.73)	0.28 (0.08-0.99)	0.19 (0.05-0.70)	0.21 (0.06-0.74)
DOR	68.5 (7.22-630)	9.60 (1.82-50.6)	20.9 (3.97-109)	16.5 (3.19-85.2)

Performance measures presented as value (95% confidence interval).

PPV, positive predictive value; NPV, negative predictive value; PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio.

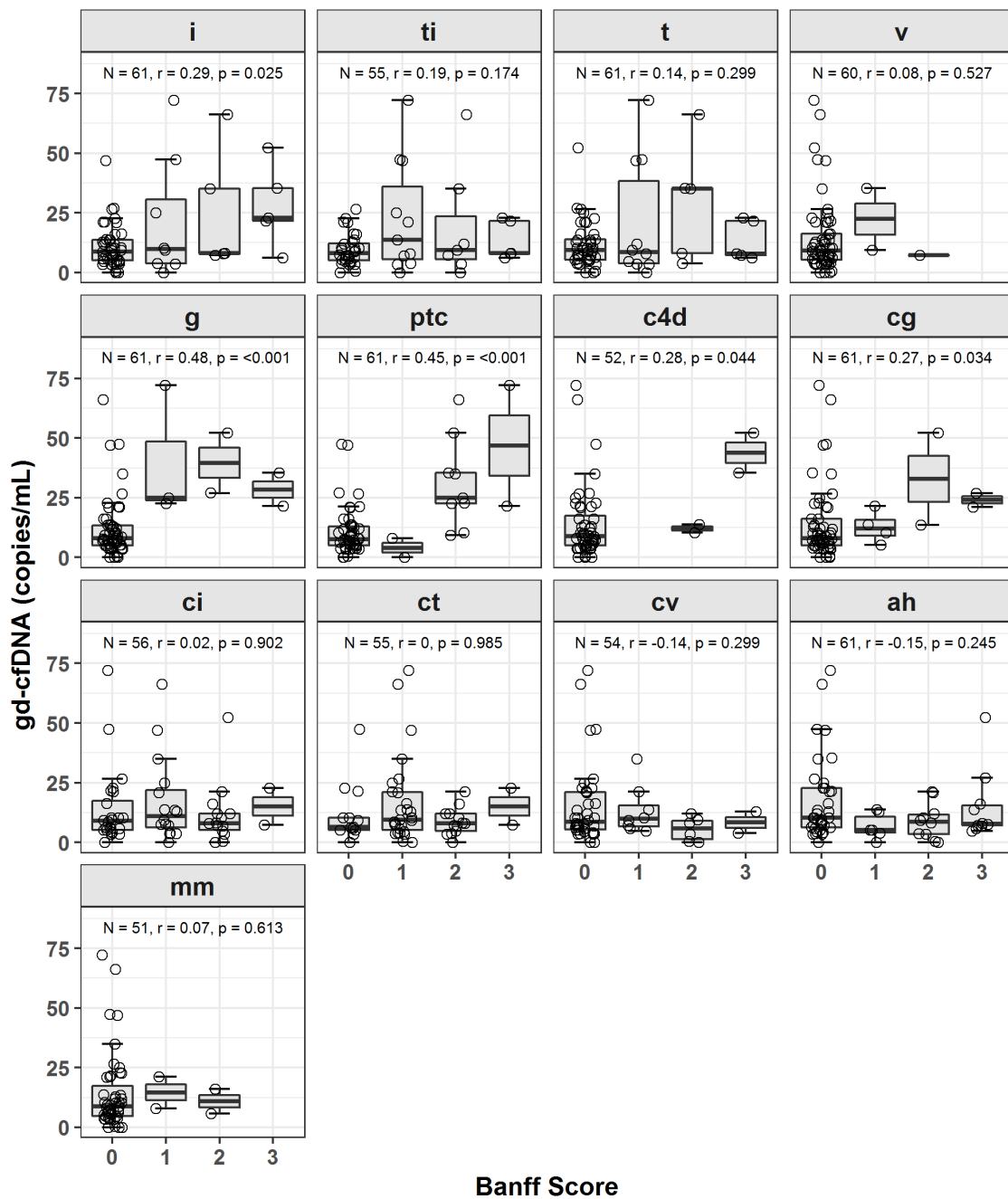
5.2.3.5 Banff Scores

There were positive monotonic relationships between gd-cfDNA and Banff scores for interstitial inflammation ($r = 0.29$, $p = 0.03$), glomerulitis ($r = 0.48$, $p < 0.001$), peritubular capillaritis ($r = 0.45$, $p < 0.001$), c4d staining ($r = 0.28$, $p = 0.04$) and transplant glomerulopathy ($r = 0.27$, $p = 0.03$) (Figure 5-5).

Graft fraction was correlated with more Banff lesions, specifically interstitial inflammation ($r = 0.45$, $p < 0.001$), total inflammation ($r = 0.39$, $p = 0.004$), tubulitis ($r = 0.3$, $p = 0.02$), glomerulitis ($r = 0.38$, $p = 0.002$), peritubular capillaritis ($r = 0.39$, $p = 0.002$), c4d staining ($r = 0.39$, $p =$

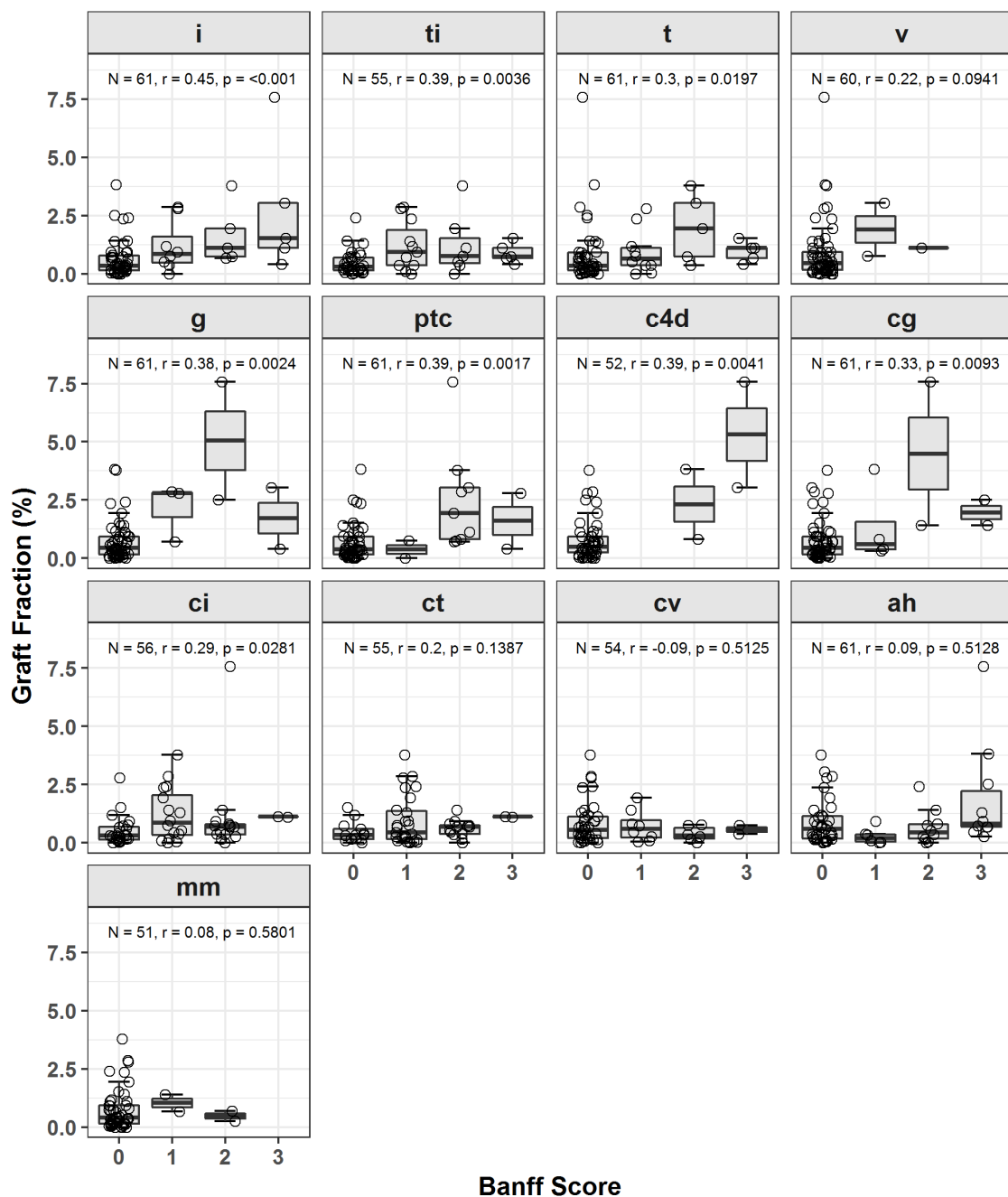
0.004), transplant glomerulopathy ($r = 0.33$, $p = 0.009$), and interstitial fibrosis ($r = 0.29$, $p = 0.03$) (Figure 5-6).

Figure 5-5: The distribution of gd-cfDNA measurements for each Banff lesion and score. Hollow circles are individual results. The number of results included in the analysis (N), Spearman's rho (r), and the significance of the correlation (p) are presented for each Banff lesion. gd-cfDNA positively correlated with interstitial inflammation (i), glomerulitis (g), peritubular capillaritis (ptc), C4d staining (c4d), and transplant glomerulopathy (cg) lesions.



Banff score lesions: i, interstitial inflammation; ti, total inflammation; t, tubulitis; v, arterial inflammation; g, glomerulitis; ptc, peritubular capillaritis; c4d, c4d deposition; ci, intersitital fibrosis; ct, tubular atrophy; cv, arterial intimal thickening; cg, transplant glomerulopathy; ah, arteriolar hyalinosis; mm, mesangial matrix increase.

Figure 5-6: The distribution of calculated graft fraction results for each Banff lesion and score. Hollow circles are individual results. The number of results included in the analysis (N), Spearman's rho (r), and the significance of the correlation (p) are presented for each Banff lesion. Graft fraction positively correlated with interstitial inflammation (i), total inflammation (ti), tubulitis (t), glomerulitis (g), peritubular capillaritis (ptc), C4d staining (c4d), transplant glomerulopathy (cg), and interstitial fibrosis (ci) lesions.

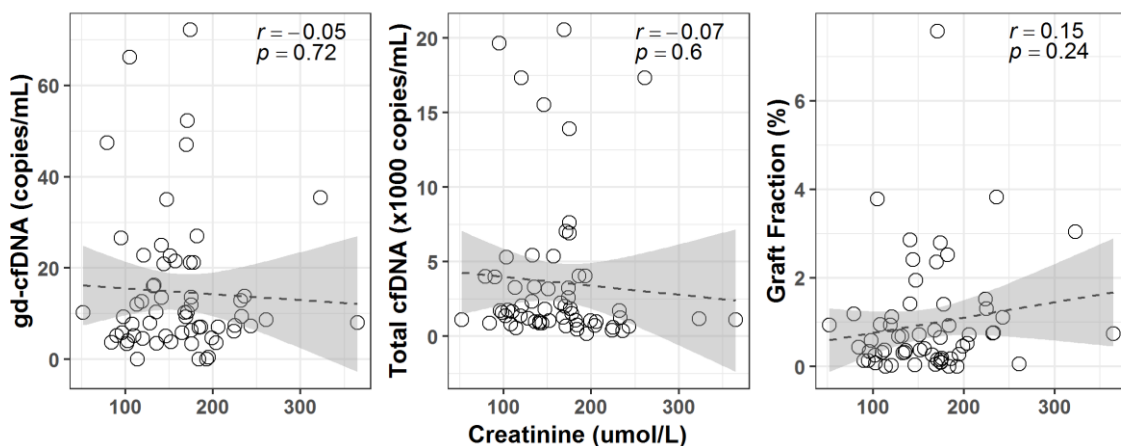


Banff score lesions: i, interstitial inflammation; ti, total inflammation; t, tubulitis; v, arterial inflammation; g, glomerulitis; ptc, peritubular capillaritis; c4d, c4d deposition; ci, interstitial fibrosis; ct, tubular atrophy; cv, arterial intimal thickening; cg, transplant glomerulopathy; ah, arteriolar hyalinosis; mm, mesangial matrix increase.

5.2.3.6 Creatinine

There was no significant correlation between serum creatinine and gd-cfDNA ($r = -0.05$, $p = 0.72$), total cfDNA ($r = -0.07$, $p = 0.60$) and graft fraction ($r = 0.15$, $p = 0.24$) (Figure 5-7). The ROC AUC of serum creatinine for diagnosis of aAMR was 0.47 (95% CI 0.29-0.66), and for diagnosis of ABMR it was 0.59 (95% CI 0.42-0.75) (Figure 5-4). Compared to serum creatinine, gd-cfDNA and graft fraction performed better for diagnosis of aAMR (ROC AUC 0.92 vs 0.47, $D = 5.4$, $p < 0.001$; ROC AUC 0.85 vs 0.47, $D = 3.5$, $p < 0.001$) and ABMR (ROC AUC 0.91 vs 0.59, $Z = 3.0$, $p = 0.001$; ROC AUC 0.89 vs 0.59, $Z = 3.3$, $p < 0.001$).

Figure 5-7: Correlation of serum creatinine concentration with gd-cfDNA, total cfDNA and graft fraction. There was no correlation between serum creatinine concentration and cfDNA levels.



5.2.4 Discussion

This study is the first to examine the assay performance of absolute quantification of gd-cfDNA for the diagnosis of Banff-classified kidney allograft rejection, and compare it with that of graft fraction. This was achieved using highly heterozygous CNV DNA sequences to facilitate mismatches in which the graft DNA contained 1- or 2-copies of the target sequence and the recipient had no copies for that respective sequence. This enabled gd-cfDNA to be quantified without interference from recipient DNA sequences. Compared with methods quantifying the relative proportion of one SNP allelotype against another, the present approach has theoretical advantages with respect to analytical sensitivity and specificity for absolute quantification. Additionally, unlike next generation sequencing-based platforms, there is no pre-amplification of cfDNA, with its potential to introduce bias.

The present study assayed gd-cfDNA and total cfDNA, and calculated from these the graft fraction. It demonstrated that both graft-specific measurements were higher in ABMR compared to non-ABMR and negative biopsies. There was no association between total cfDNA and these diagnoses. There was a non-significant trend towards better performance of gd-cfDNA for diagnosis of aAMR, and an association between gd-cfDNA and the Banff lesions associated with ABMR. The performance of gd-cfDNA and graft fraction for composite diagnosis of ABMR was comparable.

Graft fraction was also associated with Banff lesions associated with ABMR, but unlike gd-cfDNA, demonstrated association with some CMR lesions such as total inflammation, tubulitis and arterial inflammation, as well interstitial fibrosis. This suggests that the relative suppression of total cfDNA in these scenarios, whilst not helpful in isolation, adds additional diagnostic information to the graft fraction. Supporting this, only half of the variance in graft fraction in this study was accounted for by the gd-cfDNA concentration, implying total cfDNA concentration was responsible for the remainder. Thus, whilst gd-cfDNA may be a more specific marker for aAMR, graft fraction could provide more diagnostic value in caAMR and potentially more severe forms of CMR, although confounding of graft fraction by other pathologies (BK nephropathy, CNI toxicity) was observed. Considering both gd-cfDNA and graft fraction together may also provide additional diagnostic information and this requires further evaluation, in particular because study of the effect of CMR, BK nephropathy and CNI toxicity on both measures was limited by small numbers in this study.

The results of the present study support the findings of Bloom et al. in their examination of graft fraction for diagnosis of rejection in a similar cohort using a sequencing-based method and SNP polymorphisms. In their study, diagnostic performance for a composite outcome of type $\geq$ 1B CMR and ABMR was used (“active rejection”). The current study cohort included only two biopsies with type $\geq$ 1B CMR. One biopsy had type IB CMR and concurrent acute AMR with high gd-cfDNA (23 copies/mL) and graft fraction (1.1%); the other had type IIB CMR alone and no elevation in gd-cfDNA (7 copies/mL) but a high graft fraction (1.1%). Similar to the study of Bloom et al., no significant rise in graft fraction (or gd-cfDNA) was observed in type IA CMR and borderline CMR. This may reflect the limited sensitivity of these approaches for milder forms of inflammation, and/or the localisation of the cellular injury further away from the vascular space as compared to AMR. Excluding CMR, both methods appear to offer similar performance with respect to diagnosis of ABMR. The positive predictive value of both analytes for ABMR was relatively low but the negative predictive value was very high, suggesting that these tests perform best as “rule out” tests.

Notably, this study did not observe confounding of the diagnostic performance of graft fraction by concurrent elevation in total cfDNA levels arising from non-alloimmune processes, although the study was not powered to examine this.

The droplet digital PCR method employed for this work can be performed rapidly, with a turnaround time of less than 24 hours. This platform also permits a lower per sample cost and avoids access barriers compared with sequencing-based approaches. These characteristics favour application to serial monitoring in clinical practice as a marker of predominantly antibody-mediated allograft injury. Unlike other transcriptomic or proteomic biomarkers that measure alloimmune activity, the current understanding of cfDNA physiology proposes that gd-cfDNA release reflects graft injury itself (i.e. the end-organ effects of the alloimmune response). The high negative predictive value $\geq 95\%$ could provide utility in monitoring response to treatment after a diagnosis of AMR, where serial biopsy is often employed. The 90% sensitivity of gd-cfDNA for aAMR could provide evidence of relapse in high pre-test probability circumstances. A negative gd-cfDNA test could also modify the risk:benefit of performing surveillance biopsy, particularly in those with a high biopsy risk.

The current study found moderate variation between informative markers at low gd-cfDNA concentrations within the clinically significant range. This finding supports the use of multiple informative markers, which independently measure gd-cfDNA and can be used to improve reliability of the result. Additionally, the method employed in this study assayed all markers from a panel of strategically informative markers across all samples. To maximise efficiency, reduce cost, and improve sensitivity, once informative markers have been identified from the first sample, future implementations could assay only selected informative markers for donor and recipient. This would permit use of higher DNA concentrations per assay reaction, improving sensitivity and precision at the low gd-cfDNA concentrations observed in this study. Such an approach would also reduce the sample collection requirement to 9 ml blood for serial samples.

Key limitations of the present study arise from the small sample size and breadth of pathologies assessed. In particular, the number of cases of CMR, caAMR and BK nephropathy was small. Additionally, the study only assessed transplant recipients undergoing biopsy for cause. This was necessary to maximise the prevalence of rejection on the study cohort. The effect of DSA on the diagnostic performance of cfDNA was not included in the study design, and analysis was not performed due to incomplete DSA data at the time of biopsy. Further validation of the diagnostic thresholds and study of the performance of this biomarker in surveillance, and in serial monitoring during/after AMR treatment is warranted, since results of this work suggest utility in these scenarios.

5.3 AUTHOR CONTRIBUTIONS

J.B.W. and J.K. were responsible for recruitment and sample collection. J.B.W. and L.L. performed the laboratory work. J.B.W. performed the data analysis and wrote the manuscript. D.L.B. and H.S. provided scientific advice/oversight and edited the paper. Authors D.A.P., J.K., F.L.I. and A.S. provided clinical advice/oversight and edited the paper.

5.4 ACKNOWLEDGMENTS/FUNDING

The authors thank Tom Harrington, Ranier Arolanda and Marina Katerelos for their assistance with sample processing and handling. The authors also thank Marieke Veenendaal for assistance with recruitment and sample collection.

The project was supported by a Roche Organ Transplant Research Foundation grant (529172731), and by the Victorian Government's Operational Infrastructure Support Program. Author J.B.W. received stipend support from Kidney Health Australia (SCH16 141516) and the Australian Government National Health and Medical Research Council (GNT1114931). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

5.5 DISCLOSURES

The authors of this manuscript have conflicts of interest to disclose as described by the *American Journal of Transplantation*. D.L.B. and H.R.S. are co-inventors on a patent (WO2013/049892) covering the analytical method used in this study.

5.6 SUPPLEMENTAL MATERIAL

5.6.1 Supplemental Methods

5.6.1.1 *Droplet Digital PCR Thresholding*

Amplitude thresholds were set at one-third of the distance between the negative droplet cluster and the positive droplet cluster for each FAM- (channel 1) and HEX- (channel 2) labelled assay in the duplex arrangement (Figure 5-8). The amplitude threshold was defined by the combined cluster amplitudes for all wells on the same plate containing the same assay combination (Figure 5-9). The same amplitude threshold was therefore used for all assays on the same plate.

Figure 5-8: Bio-Rad QuantaSoft Analysis Pro 2D plots from two ddPCR wells, each containing two CNV PCR assays and cfDNA from the sample represented in Supplemental Figure 5-10B. Each droplet (dot) is plotted against its channel 1 (FAM-labelled probe) and channel 2 (HEX-labelled probe) fluorescent intensity. The crosshair represents the channel 1 (horizontal) and channel 2 (vertical) threshold distinguishing positive from negative droplets. Each positive

droplet is stochastically likely to contain only one copy of target DNA template, and the concentration of each CNV locus in the sample is calculated based upon the proportion of negative to total droplets.

Plot A shows CNV assay 03B (y-axis) and 06B (x-axis); plot B shows CNV assay 29A (y-axis) and 30A (x-axis). There is clear separation between positive clusters and negative clusters. CNV 03B is a graft informative marker, CNV 06B is a recipient 1-copy marker, CNV 29A is not detectable, and CNV 30A is graft informative. For CNV 29A, the threshold amplitude was defined by other samples within the same plate that exhibited detectable target sequence (Figure 5-9).

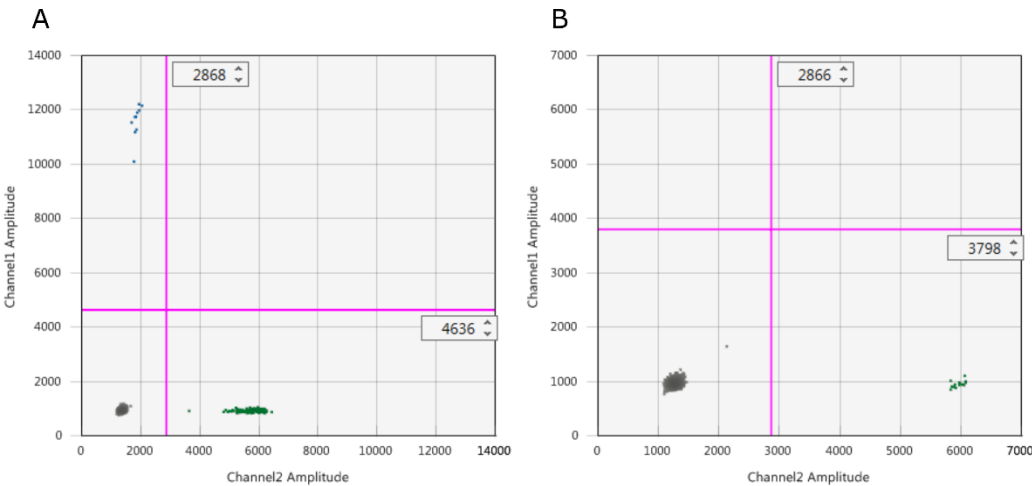
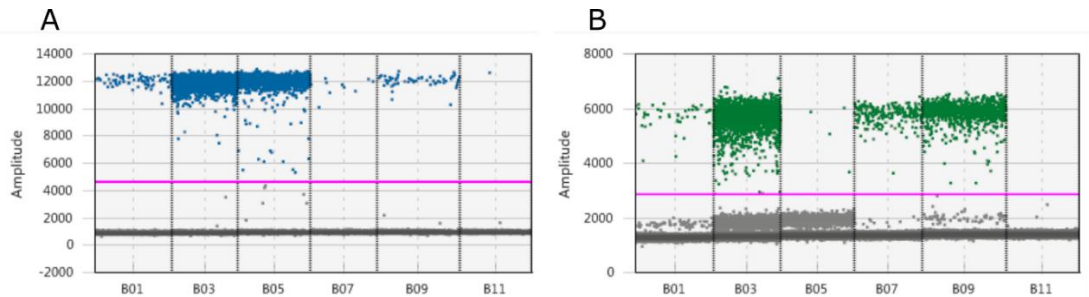


Figure 5-9: Bio-Rad QuantaSoft Analysis Pro 1D plots representing the channel 1 (plot A) and channel 2 (plot B) amplitudes of droplets for the wells containing CNV assay 03B and 06B across six samples (x-axis, divided by dashed vertical lines) on the same ddPCR plate. Supplemental Figure 5-8A is represented as well B07 in this figure. The horizontal line represents the amplitude threshold defining positive from negative droplets. The plots demonstrate co-localisation of the CNV 03B and 06B positive droplet clusters across samples at an amplitude of ~12,000 and ~6,000 respectively, with clear separation from the negative droplet clusters.



5.6.1.2 Conversion of Droplet Digital PCR Results

DdPCR concentration results were converted to a plasma-equivalent concentration using the equation

$$C_{plasma} = \frac{C_{ddPCR} \times V_{elution} \times (V_{cfDNA} + V_{MM})}{V_{plasma} \times V_{cfDNA}}$$

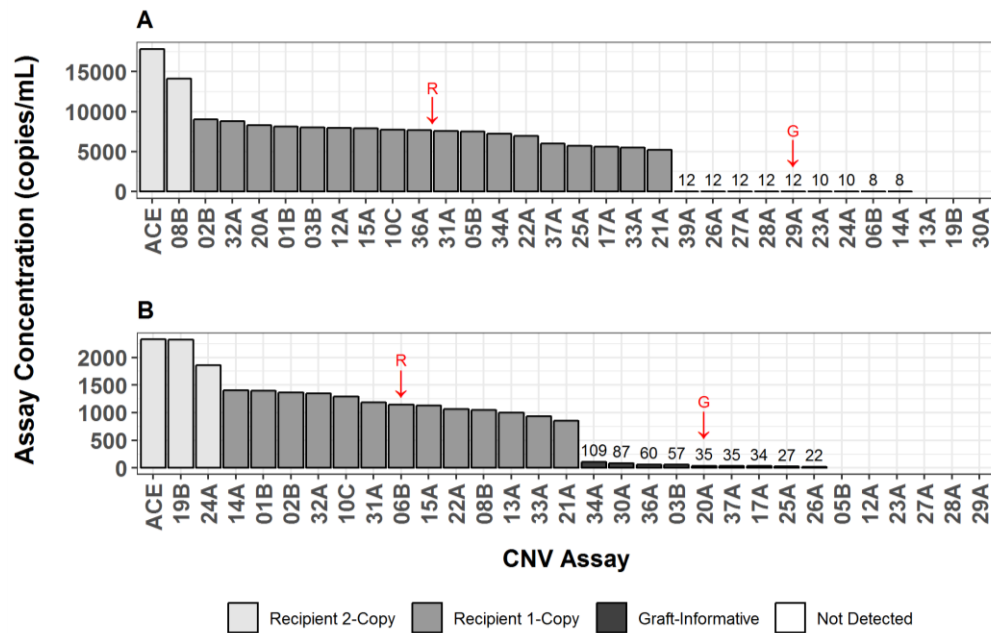
where C_{plasma} is the concentration of target within the plasma in copies/mL, C_{ddPCR} is the concentration of target within the PCR reaction in copies/ μ L, $V_{elution}$ is the volume of eluent used during the cfDNA extraction process in μ L, V_{cfDNA} is the volume of extracted cfDNA used in the PCR reaction in μ L, V_{MM} is the volume of all other PCR reaction components in μ L, and V_{plasma} is the volume of plasma used for cfDNA extraction in mL.

5.6.1.3 CNV Assay Genotyping

Figure 5-10: DdPCR concentrations for each of 31 CNV assays and the ACE 2-copy positive control from two different participants are presented. Note that the concentration of each assay is presented as the equivalent sample concentration (copies/mL plasma).

Sample A (top panel) was collected from an individual exhibiting no rejection but chronic CNI and hypertensive changes. The profile of assay concentrations demonstrates 1 recipient 2-copy assay (excluding ACE), 18 recipient 1-copy assays, 9 graft informative assays, and 3 CNVs that were not detectable. The median graft informative assay concentration (12 copies/mL, “G” arrow) was divided by the median recipient 1-copy assay concentration (7614 copies/mL, “R” arrow) to calculate the graft fraction (0.16%).

Sample B (bottom panel) was collected from an individual with aAMR. The profile of assay concentrations demonstrates 2 recipient 2-copy assays (excluding ACE), 13 recipient 1-copy assays, 9 graft informative assays, and 6 CNVs that were not detectable. The median graft informative assay concentration (35 copies/mL, “G” arrow) was divided by the median recipient 1-copy assay concentration (1143 copies/mL, “R” arrow) to calculate the graft fraction (3.1%). CNV assays 34A and 30A may be 2-copy graft informative markers. Use of the median graft informative assay value, combined with the uncommon allelic frequency of 2-copy CNV loci in the panel, means that if > 3 informative markers are present, the likelihood of incorrectly using a 2-copy CNV assay concentration is low.



5.6.2 Supplemental Results

5.6.2.1 Study Cohort

Table 5-4: Participant characteristics for the aAMR groups.

	Study Cohort (N=55)	No aAMR (N=45)	aAMR (N=10)	p-value ⁱ
Demographics				
Age at first sample (years)	53 (39-61)	55 (42-61)	42 (39-55)	0.13
Male sex	30 (55%)	24 (53%)	6 (60%)	0.74
Race				0.18
Caucasian	44 (80%)	37 (82%)	7 (70%)	
Asian	7 (13%)	6 (13%)	1 (10%)	
Polynesian	2 (4%)	1 (2%)	1 (10%)	
Arabic	1 (2%)	0 (0%)	1 (10%)	
African	1 (2%)	1 (2%)	0 (0%)	
Primary disease				0.67
IgA nephropathy	14 (25%)	11 (24%)	3 (30%)	
Reflux nephropathy	8 (15%)	5 (11%)	3 (30%)	
Primary FSGS	6 (11%)	6 (13%)	0 (0%)	
Polycystic kidney disease	6 (11%)	5 (11%)	1 (10%)	
Secondary FSGS	4 (7%)	3 (7%)	1 (10%)	
Diabetes	3 (5%)	3 (7%)	0 (0%)	
Unknown	2 (4%)	2 (4%)	0 (0%)	
ANCA vasculitis	2 (4%)	1 (2%)	1 (10%)	
Hypertension	2 (4%)	2 (4%)	0 (0%)	
Other	8 (2%)	7 (16%)	1 (10%)	
Recruiting hospital				0.30
Hospital 1	49 (89%)	41 (91%)	8 (80%)	
Hospital 2	6 (11%)	4 (9%)	2 (20%)	
Transplant				
Graft number				0.40

First	51 (93%)	42 (93%)	9 (90%)	
Second	2 (4%)	2 (4%)	0 (0%)	
Third	2 (4%)	1 (2%)	1 (0%)	
Donor Type				0.71
Deceased	40 (73%)	32 (71%)	8 (80%)	
Brain death	19 (35%)	15 (33%)	4 (40%)	
Circulatory death	13 (24%)	11 (24%)	2 (20%)	
Unknown	8 (15%)	6 (13%)	2 (20%)	
Living	15 (27%)	13 (29%)	2 (20%)	
Parent	7 (13%)	6 (13%)	1 (10%)	
Sibling	2 (4%)	2 (4%)	0 (0%)	
Related NOS	1 (2%)	1 (2%)	0 (0%)	
Not related	5 (9%)	4 (9%)	1 (10%)	
Matched donor-recipient sex	25/51 (49%)	19/41 (46%)	6/10 (60%)	0.50
ABO compatible	52 (95%)	43 (95%)	9 (90%)	0.46
DSA at transplant	10/47 (21%)	7/37 (19%)	3/10 (30%)	0.43
HLA mismatch				0.24
5-6	12 (22%)	8 (17%)	4 (40%)	
3-4	25 (45%)	22 (49%)	3 (30%)	
1-2	16 (29%)	14 (31%)	2 (20%)	
0	2 (4%)	1 (2%)	1 (10%)	
Induction agent				0.66
Basiliximab	42 (76%)	33 (73%)	9 (90%)	
None	4 (7%)	4 (9%)	0 (0%)	
OKT3	1 (2%)	1 (2%)	0 (0%)	
Unknown	8 (15%)	7 (16%)	1 (10%)	

Data ranges are presented as median (inter-quartile range).

ⁱ The *p*-value presented is the significance of differences between the aAMR and no aAMR groups. The Wilcoxon rank sum test, Fisher's exact test for count data, and the equality of proportions test were used for comparisons as indicated.

FSGS, focal segmental glomerulosclerosis; ANCA, anti-neutrophil cytoplasmic antibody; NOS, not otherwise specified; DSA, donor specific antibody; HLA, human leukocyte antigen; OKT3, muromonab-CD3.

Table 5-5: Sample characteristics for the aAMR groups.

	Sample Cohort (N=61)	No aAMR (N=51)	aAMR (N=10)	p-value ⁱ
Sample Collection				
Time post-transplant (years)	2.6 (0.6-8.2)	2.7 (0.6-9.9)	1.0 (0.3-4.9)	0.32
Time post-transplant cohorts (years)				0.85
< 0.5	14 (23%)	11 (22%)	3 (30%)	
0.5-1	12 (20%)	10 (20%)	2 (20%)	
1-5	12 (20%)	10 (20%)	2 (20%)	
5-10	9 (15%)	7 (14%)	2 (20%)	
>10	14 (23%)	13 (25%)	1 (10%)	
Collection time from biopsy (days)				0.84
-6 (before biopsy)	1 (2%)	1 (2%)	0 (0%)	
-1	1 (2%)	1 (2%)	0 (0%)	
0 (day of biopsy)	57 (93%)	47 (92%)	10 (100%)	
+1 (after biopsy)	2 (3%)	2 (4%)	0 (0%)	
Clinical Characteristics				
Biopsy indication				0.59
Creatinine	56 (92%)	46 (90%)	10 (100%)	
Proteinuria	3 (3%)	2 (4%)	0 (0%)	
Both	2 (5%)	3 (6%)	0 (0%)	

Baseline creatinine (μmol/L) (N=60) ⁱⁱ	125 (102-156)	122 (104-157)	137 (95-151)	0.84
Baseline eGFR (ml/min) (N=60) ⁱⁱ	52 (42-65)	52 (42-62)	52 (41-66)	0.70
Creatinine (μmol/L)	165 (120-184)	168 (117-185)	154 (143-173)	0.79
eGFR (ml/min)	40 (29-54)	39 (27-54)	42 (32-47)	0.99
Indication urine protein:creatinine (g/mmol) (N=5)	0.21 (0.20-0.35)	0.21 (0.20-0.35)		
Urine protein:creatinine (g/mmol) (N=40)	0.03 (0.01-0.09)	0.03 (0.01-0.08) (N=34)	0.03 (0.01-0.11) (N=6)	0.88
Any history of donor specific antibody	18/35 (51%)	13/26 (50%)	5/9 (56%)	>0.99
Maintenance immunosuppression				
Prednisolone	56 (92%)	46 (90%)	10 (100%)	0.56
Mycophenolate	50 (82%)	40 (78%)	10 (100%)	0.18
Azathioprine	6 (10%)	6 (11%)	0 (0%)	0.58
Tacrolimus	48 (79%)	38 (75%)	10 (100%)	0.10
Ciclosporin	10 (16%)	10 (20%)	0 (0%)	0.19
mTOR	4 (7%)	4 (8%)	0 (0%)	>0.99
Drug levels				
Tacrolimus level (μg/L) (N=48)	5.3 (4.2-6.3)	5.2 (4.2-6.0) (N=38)	6.6 (4.0-8.0) (N=10)	0.35
Ciclosporin level (μg/L) (N=10)	318 (207-490)	318 (207-490) (N=10)		
Weight (kg) (N=56)	78 (69-85)	77 (67-85) (N=48)	83 (78-89) (N=8)	0.14

Data ranges are presented as median (inter-quartile range).

ⁱ The *p*-value presented is the significance of differences between the aAMR and no aAMR groups. The Wilcoxon rank sum test, Fisher's exact test for count data, and the equality of proportions test were used for comparisons as indicated.

ⁱⁱ One participant had not achieved a baseline creatinine by the time of sample collection.

eGFR, estimated glomerular filtration rate (CKD-EPI); mTOR, mammalian target of rapamycin.

5.6.2.2 CNV Assay Genotyping Results

A total of 1,996 PCR results were included in the analysis. Excluding the *ACE* positive control, 44/61 (72%) samples had 31 CNV assay results, 7/61 (11%) samples had 29-30 CNV assay results, and 10/61 (16%) had 32-40 CNV assay results.

At least one graft informative assay was identified in 58/61 (95%) of samples (Supplemental Figure 5-11). The median number of graft informative assays detectable per sample was 5 (IQR 3-7), corresponding to a graft informative assay yield (number of graft informative assays/number of assays tested per sample) of 16% (IQR 8-23%).

The number of graft informative markers appeared higher when the median gd-cfDNA concentration was higher (Supplemental Figure 5-12). There were 3 samples for which no graft informative markers were detectable. Analysis of detectable markers in urine cell-free DNA identified multiple graft informative markers that were undetectable in plasma (unpublished data). In these cases, the plasma gd-cfDNA concentration was below the limit of detection for the assay.

The median number of recipient 1-copy markers was 13 (IQR 11-15), recipient 2-copy markers 3 (IQR 2-5), and zero result assays 10 (IQR 8-12). The proportionate representation of these genotypes from the total number of assays tested per sample was 41% (IQR 35-48), 10% (IQR 5-16%) and 32% (IQR 26-39%) respectively.

Figure 5-11: Histogram plots of the number of assays genotyped as graft informative, recipient 1-copy and recipient 2-copy per sample.

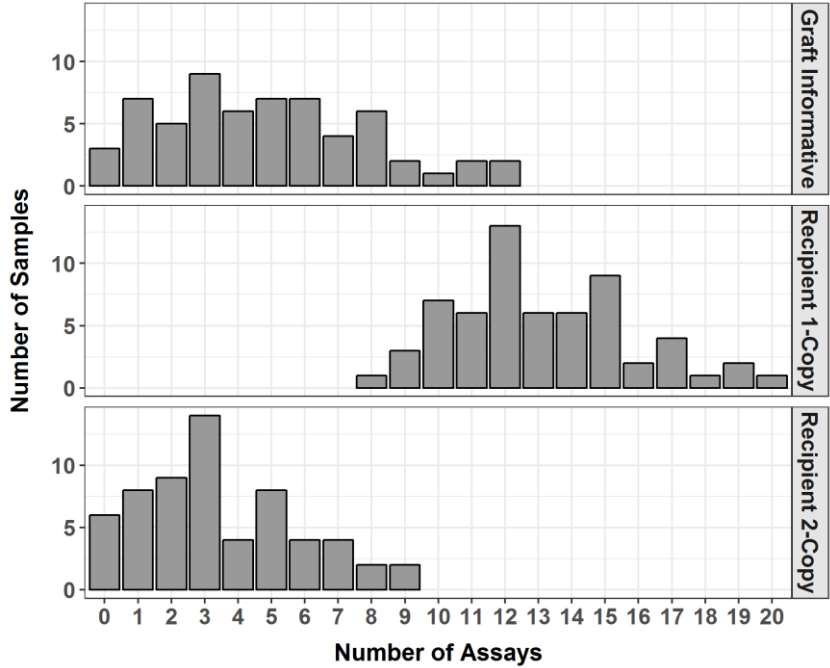
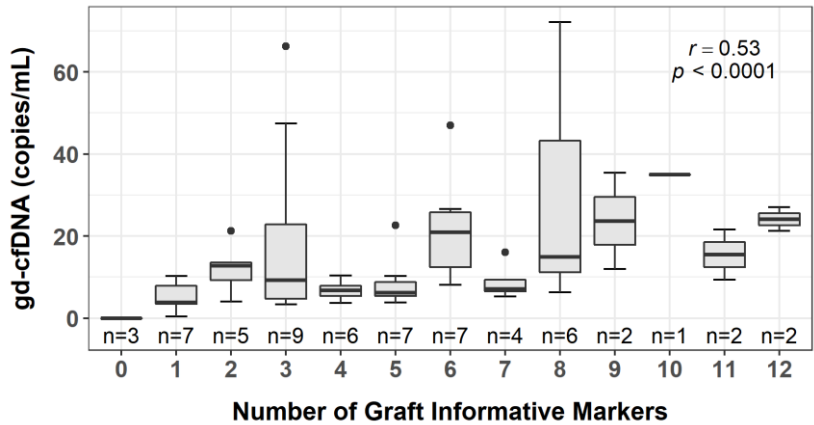


Figure 5-12: Distribution of gd-cfDNA concentrations grouped by the number of graft informative markers per sample. There was a moderate correlation between gd-cfDNA and number of informative markers.



5.6.2.3 Effect of Relatedness on Informativity

10/55 (18%) recipients were related to their donor, accounting for 11/61 (18%) samples (6 participants were biopsied twice over the course of the study).

The number of graft informative assays was lower in related donor-recipient pairs compared to unrelated donor-recipient pairs (3 (IQR 1.5-4.5) vs 5 (IQR 3.0-7.8), $N = 61$, $W = 391$, $p = 0.03$) (Supplemental Figure 5-13, left panel). The number of graft informative assays did not appear to be different amongst related donor-recipient pairs when grouped by relationship type ($\chi^2(2) = 0.42$, $p = 0.81$), however the number of samples in each group was low (Supplemental Figure 5-14).

The number of recipient 1-copy markers was no different between unrelated and related pairs (13 (IQR 11-15) vs 12 (IQR 11.5-14), $N = 61$, $W = 306$, $p = 0.57$) (Supplemental Figure 5-13, right panel).

Figure 5-13: Distribution of the number of graft informative (left panel) and recipient 1-copy (right panel) assays per sample, grouped by relatedness between donor and recipient. The number of graft informative assays was lower in related pairs compared to unrelated pairs. There was no difference in the number of recipient 1-copy markers between related and unrelated pairs.

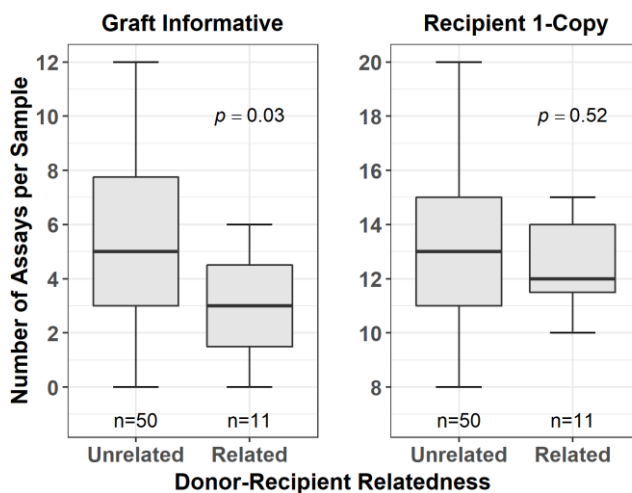
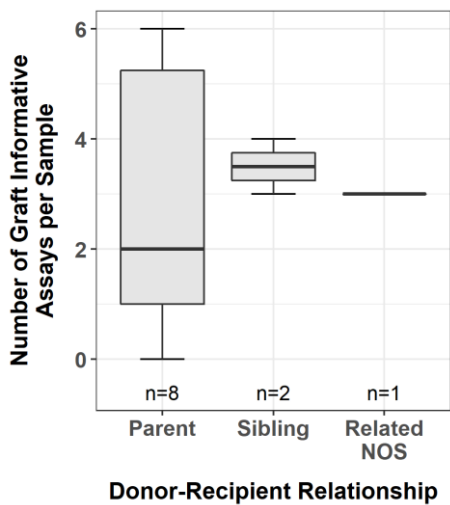


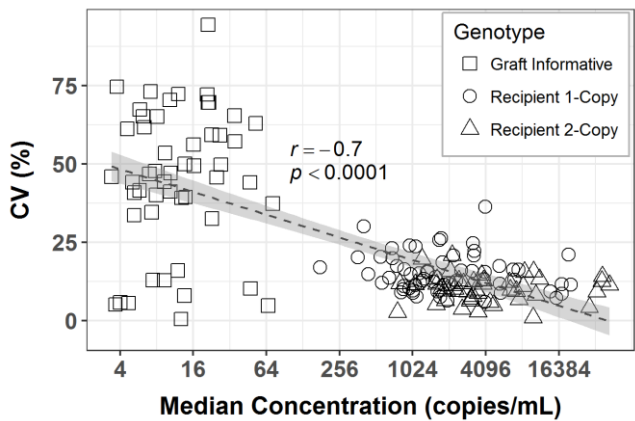
Figure 5-14: The number of graft informative assays per sample in related donor-recipient pairs, grouped by relationship. There was no apparent difference in informativity by relationship,

although the number of cases was small and no non-first degree related donors were captured in the cohort. NOS, not otherwise specified.



5.6.2.4 Inter-Assay Variation

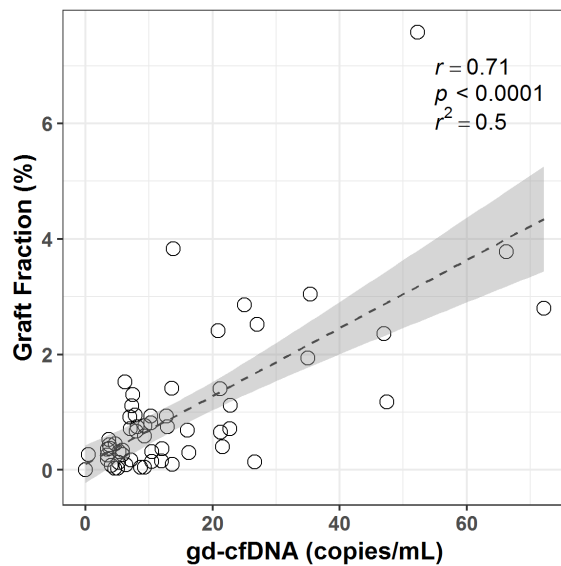
Figure 5-15: The relationship between inter-assay coefficient of variation and median assay concentration for that genotype. A linear model (dashed line) with 95% confidence intervals (grey shading) is plotted. Coefficient of variation reduced as the median genotype concentration increased ($r = -0.7$, $p < 0.0001$). Note Log2 transformation of x-axis.



5.6.2.5 Correlation of Gd-cfDNA and Graft Fraction

Figure 5-16: Correlation between gd-cfDNA and graft fraction. A linear model (dashed line) with 95% confidence interval (grey shading) is plotted. There was a linear relationship between gd-

cfDNA and graft fraction ($r = 0.71$, $p < 0.0001$), however gd-cfDNA only accounted for 50% of the variance of graft fraction ($r^2 = 0.5$).



5.6.2.6 Effect of Biopsy Procedure

13/60 (22%) blood samples were obtained after biopsy of the allograft. 11/13 (85%) of these were obtained on the same day as the biopsy, and 2/13 (15%) on the day after biopsy. The median gd-cfDNA concentration for samples collected after biopsy was 13 (IQR 7-25) copies/mL compared to 8 (IQR 5-14) copies/mL for samples collected before biopsy ($W = 232$, $p \approx 0.19$) (Supplemental Figure 5-17, left panel). The median graft fraction for samples collected after biopsy was 0.69% (IQR 0.14-2.4%) compared to 0.52% (IQR 0.26-0.94%) for samples collected before biopsy ($W = 284$, $p \approx 0.71$) (Supplemental Figure 5-17, right panel).

Of the 60 samples with complete collection and biopsy time data, 28 samples were associated with negative biopsies. 8/28 (29%) of these were obtained after biopsy of the allograft, 7/8 (88%) on the same day as the biopsy. The median negative biopsy gd-cfDNA concentration for samples collected after biopsy was 8 (IQR 5-14) copies/mL compared to 7 (IQR 5-10) copies/mL for samples collected before biopsy ($W = 68$, $p \approx 0.56$) (Supplemental Figure 5-18, left panel). The median negative biopsy graft fraction for samples collected after biopsy was 0.29% (IQR 0.07-0.70%) compared to 0.26% (IQR 0.12-0.40%) for samples collected before biopsy ($W = 76$, $p \approx 0.86$) (Supplemental Figure 5-18, right panel).

Figure 5-17: Distribution of cell-free DNA measurements for samples taken before or after their corresponding biopsy. There were no statistically significant differences in the distribution of gd-cfDNA, total cfDNA, and graft fraction measurements before compared to after biopsy.

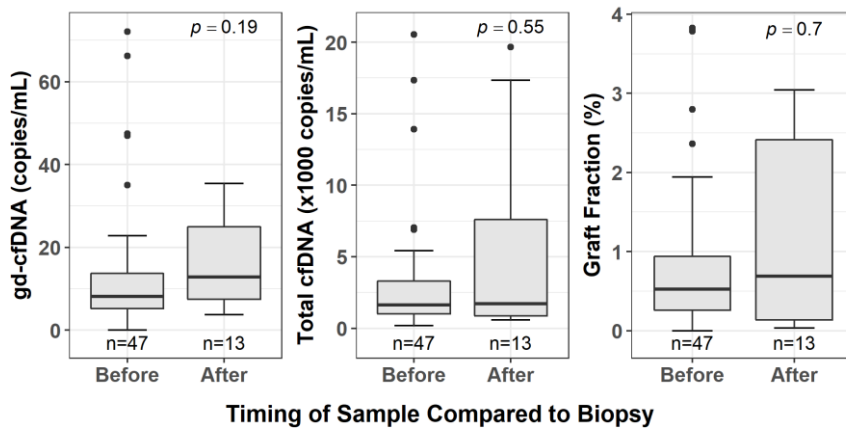
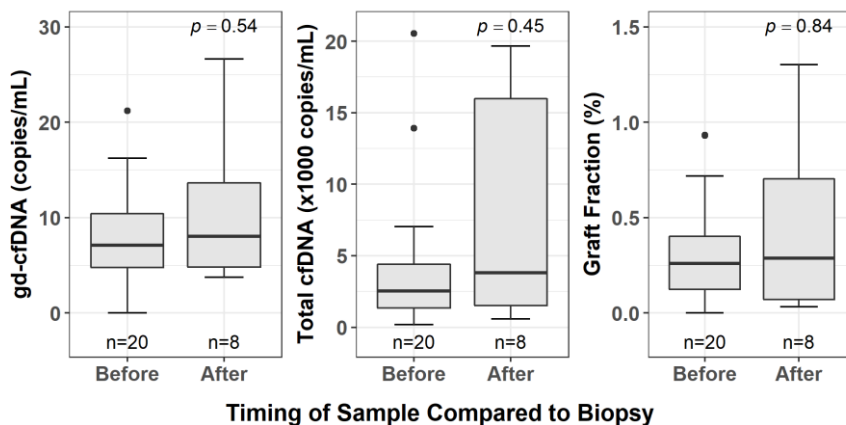


Figure 5-18: Distribution of cell-free DNA measurements for negative biopsy samples taken before or after their corresponding biopsy. There were no statistically significant differences in the distribution of gd-cfDNA, total cfDNA, and graft fraction measurements for negative biopsy bloods samples before compared to after biopsy.



5.6.2.7 Distribution of Cases by Time Since Transplant

Figure 5-19: Distribution of cfDNA measurements by ABMR classification and time since transplantation. The dashed horizontal lines represented the proposed diagnostic cut-offs for each measure (top panel 21 copies/mL, 13 copies/mL; bottom panel 0.75%). As recipients within 3 weeks of transplantation were excluded from the study, the effects of ischaemia-reperfusion injury are not readily apparent. Total cfDNA concentration tended to decline with time following

transplantation ($r = -0.32$, $p = 0.01$), but not gd-cfDNA ($r = 0.02$, $p = 0.85$) or graft fraction ($r = 0.17$, $p = 0.18$). Note Log10 transformation of x-axis.

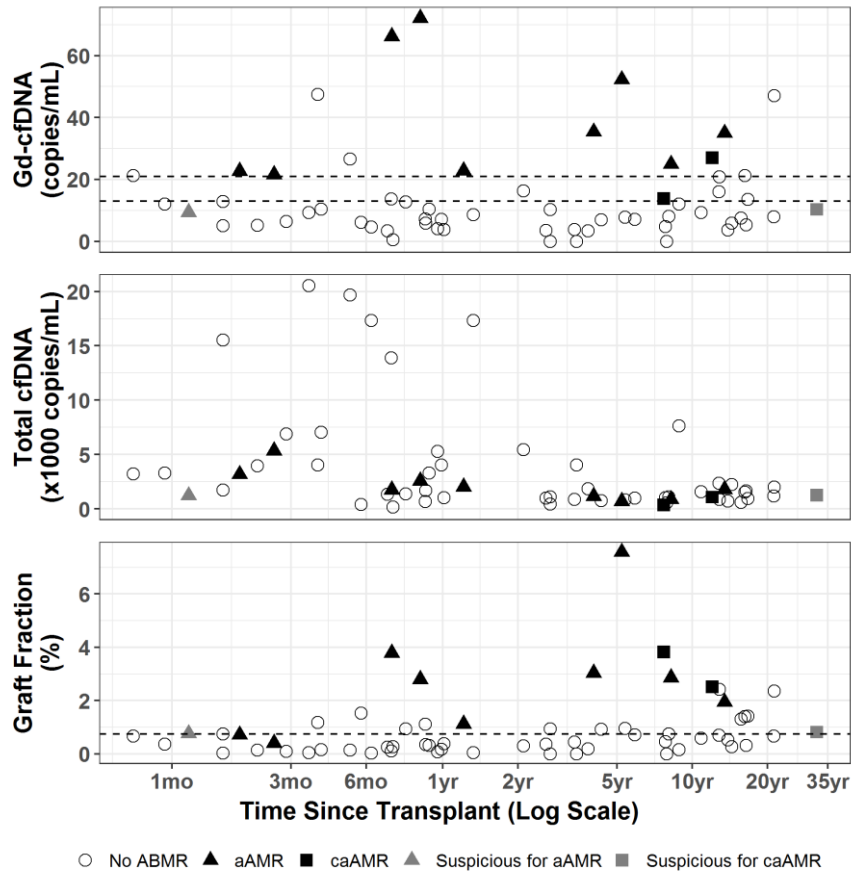
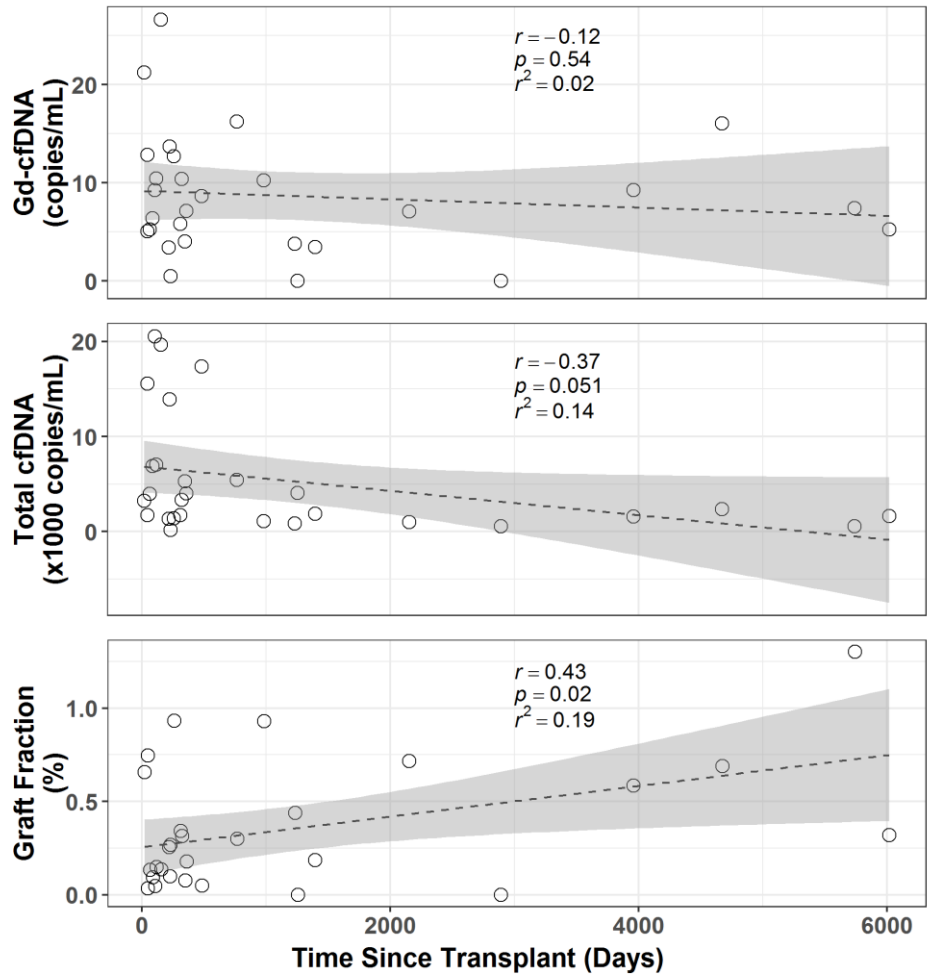


Figure 5-20: Distribution of cfDNA measurements from negative biopsy cases over time since transplantation. Linear models (dashed line) with 95% confidence intervals (grey shading) are plotted. The graft fraction of negative biopsy cases appeared to increase with time after

transplantation ($r = 0.43$, $p = 0.02$), due to a decline in total cfDNA concentration ($r = -0.37$, $p = 0.051$). Gd-cfDNA in negative biopsy cases did not change over time after transplantation.



CHAPTER 6 HEMATOPOIETIC CELL TRANSPLANTATION CHIMERISM

6.1 ABSTRACT

Chimerism analysis has an important role in the management of allogeneic haematopoietic stem cell transplantation. It informs response to disease relapse, graft rejection and graft-versus-host disease. We have developed a method for chimerism analysis using ubiquitous copy number variation (CNV), which has the benefit of a “negative background” against which multiple independent informative markers are absolutely quantified using digital droplet PCR. A panel of up to 38 CNV markers with homozygous deletion frequencies of approximately 0.4-0.6 were used. Sensitivity, precision, reproducibility and informativity were assessed. CNV chimerism results were compared against established fluorescence in situ hybridization (FISH), single nucleotide polymorphism (SNP) and short tandem repeat (STR)-based methods with excellent correlation. Using 30ng of input DNA per well, the limit of detection was 0.05% chimerism and the limit of quantification was 0.5% chimerism. High informativity was seen with a median of four informative markers detectable per individual in 39 recipients and 43 donor genomes studied. The strength of this approach was exemplified in a multi-donor case involving four genomes (three related). The precision, sensitivity and informativity of this approach recommended it for use in clinical practice.

6.2 MAIN BODY TEXT

6.2.1 Introduction

Chimerism analysis provides important information for the clinical management of allogeneic haematopoietic stem cell transplantation (HSCT). Specifically, it is used to detect disease relapse and guide treatment intensity so as to maximize graft-versus-tumour effect and minimize graft rejection, opportunistic infection and graft-versus-host disease risk.

Current approaches for chimerism analysis vary between institutions but broadly include measurement of aetiology-specific pathologic mutations and methods that quantify the degree of cellular chimerism in bone marrow or peripheral blood leukocyte compartments. For the latter group, fluorescence *in situ* hybridization (FISH) of white cell nuclei using X- and Y-specific chromosome probes has been used historically where there is a sex mismatch. More recently, highly heterozygous genetic polymorphisms including short tandem repeats (STR), single nucleotide polymorphisms (SNP)[295] and small insertions and deletions (indels)[296, 297] have been exploited. STR-based chimerism analysis is presently the most commonly adopted approach, but has a sensitivity limited to 1% chimerism[298].

For non-STR-based chimerism analysis, quantification was typically achieved with real-time PCR (rtPCR), which has limitations particularly at lower levels of chimerism[299]. Digital droplet

PCR (ddPCR) overcomes many of the inherent limitations of rtPCR, permitting absolute quantification without requirement for a standard curve, protection from PCR cycling variation and better performance at lower concentrations in part because of simplicity distinguishing positive droplets containing target from those with no template[48, 53]. A number of approaches exploiting this platform have already been published including assays comprising 8 indels and a sex-mismatch marker (amelogenin)[300], 5 SNP markers[301], 38 indel markers[302] and 29 indel polymorphisms with two reference markers (beta-globin and a Y-chromosome marker)[303].

Existing ddPCR-based approaches have demonstrated improved sensitivity with detection to below 0.01% chimerism at the cost of higher input DNA requirements, which may be prohibitive in commonly encountered circumstances such as early count recovery or in flow sorted cell populations[300, 303]. Additionally, two of the proposed methods have small panel sizes which may limit informativity in closely related donor, or multiple serial or coincident donor circumstances. All methods, either by design or limitation, measure chimeric fraction using usually one informative marker with or without requirement of a reference marker. We feel that multiple independent measures of chimeric fraction, ideally incorporating donor and recipient informative markers, are required to act as internal controls, provide data on precision and support confidence in the results. Addressing these limitations, we have developed upon a copy number variation (CNV)-based method previously demonstrated to measure recipient and donor/graft cell-free DNA in solid organ transplantation[145].

CNV describes losses or gains of genetic segments more than 50 base pairs long, sufficiently large enough to accommodate primer and probe binding sites (as compared to SNPs and many indels). CNV accounts for the majority of the variation in the human genome[304]. A panel of ubiquitous CNVs that have homozygous copy number deletion (CND) frequencies in the population of 0.4-0.6 was developed. In post-HSCT samples, where two or more genomes are present, an informative CNV is defined as a detectable heterozygous (one-copy) or homozygous (two-copy) CNV in one genome that is undetectable homozygous CND (zero-copy) in all other genomes. Homozygous CND thus creates a “zero background” against which the absolute concentration of informative markers can be measured and chimeric fraction calculated. Recipient homozygous CND markers can therefore be used to monitor engraftment, graft-versus-host disease risk and graft rejection when donor genomes are informative for that marker. Conversely, donor homozygous CND markers can be used to monitor for relapse and presence of minimal residual disease where the recipient genome is informative for that marker.

6.2.2 **Materials and Methods**

6.2.2.1 *Assay Development*

38 CNV loci were identified using in silico analysis as previously described[145]. DdPCR primers and Zen double-quenched probes (Integrated DNA Technologies, Australia) were designed within each of the CNV loci in the panel (Supplementary Table 6-1). All targets are within agenic or intronic regions and have in themselves no clinical significance. An angiotensin converting enzyme gene target sequence (ACE, NM_000789) was used as a non-polymorphic, two-copy reference marker.

6.2.2.2 *DNA Extraction*

Genomic DNA was extracted from blood and bone marrow aspirates, anticoagulated with EDTA, using the NucleoBond CB100 kit (Macherey-Nagel, Germany) according to manufacturer instructions. The concentration of extracted genomic DNA was quantified using the NanoDrop 2000 (ThermoFisher Scientific, USA).

6.2.2.3 *CNV Chimerism Analysis*

Samples were run against the CNV assay panel using the Bio-Rad QX200 Droplet Digital System (Bio-Rad, USA) and a standardised input of 30ng of unfragmented genomic DNA. PCR reactions were run in duplex using FAM labelled probes for one CNV marker and HEX labelled probes for a second CNV marker. Each reaction consisted of 12.5µL Bio-Rad ddPCR Supermix for Probes (no dUTP), 2.25µL of 10µM forward and reverse primers for both the FAM & HEX markers, 0.65µL each of the 10µM FAM and HEX-labelled probes, DNA and distilled water to make a reaction volume of 25µL. PCR cycling conditions were 95° for 10 minutes then 44 cycles of 94° for 30 seconds and 60° for 1 minute, then 98° for 10 minutes followed by an indefinite hold at 12°. ddPCR data was processed using Bio-Rad QuantaSoft software (version 1.6) according to the manufacturer instructions. Droplet thresholds were set based on clear separation of positive and negative clusters (Supplementary Figure 6-8).

Pre-transplant peripheral blood or buccal swab samples (DNA Genotek, Canada) were used to obtain donor (if required) and recipient genotypes for the assay panel. Participants consented to storage of DNA and research participation. Genotype dosage (zero-copy, one-copy or two-copy) for each marker was determined by comparing its concentration with the two-copy reference marker *ACE* and other panel markers tested. At least three informative markers was sought for both recipient and donor(s). This was usually achieved using a smaller panel of 15 CNV markers. If this was not achieved or the transplant arrangement was complex (multiple donors or related pairs), an expanded panel of between 30-38 markers was used.

Post-transplant peripheral blood samples were used for chimerism analysis. Only informative markers were assayed. For each chimerism study, FAM- and HEX-labelled probes were selected to optimise the duplex assay capability of the method. For example, if three informative markers were available for both the donor and recipient, where possible three FAM and three HEX probes were used to permit chimerism analysis using only three PCR wells. Each well contained one FAM- and one HEX-based assay which could be directed at either an informative donor or recipient marker. No reference marker is required although *ACE* was used as a positive control and provides quality assurance capability as a measure of total DNA concentration. If a two-copy informative marker was used, the concentration of this marker was halved before chimerism fraction was calculated. The chimeric fraction was calculated using the average concentration of informative CNV markers specific for the target individual (recipient or donor) divided by the sum of the averaged concentrations of the recipient and donor(s) informative markers.

6.2.2.4 Sensitivity and Precision Analysis

26 CNV markers from the panel underwent sensitivity analysis. Three pairs consisting of a one-copy individual and a zero-copy individual were identified based upon maximal marker coverage to permit inter-assay comparison of 12, 9 and 5 markers respectively. Genomic DNA from the individual shown to be one-copy for the pair's CNV markers was spiked into DNA from the individual shown to be zero-copy for those markers to create a spiking series covering 100%, 60%, 25%, 10%, 5%, 1%, 0.5%, 0.1%, 0.05%, 0.01% and 0% chimerism. *ACE* was used as a positive control for each pair. A no template control (NTC) was also used.

Spiked samples in triplicate underwent ddPCR as described using a standardised input of 30ng of genomic DNA per well across the fractional range. Wells with less than 10,000 total droplets were excluded from the analysis. Observed chimeric fraction was calculated by dividing the measured concentration of the marker by the mean 100% chimerism concentration. A linear model for each CNV assay was determined. Intra-assay variation was measured by calculating the coefficient of variation (CV) for the triplicate measurement of each CNV marker at each dilution. Inter-assay variation was measured by calculating the CV for the pair's CNV markers at each dilution step. Limit of detection was defined as the lowest detectable chimeric fraction at which the majority of replicates detected the fraction. Limit of quantification was defined as the lowest detectable chimeric fraction at which all replicates were positive and intra- and inter- assay variation were equal to or less than 50%. Limit of blank was defined as the highest apparent chimeric fraction detectable in 0% chimeric samples (i.e. neat zero-copy individual DNA).

Inter-assay variation was also assessed in practice using longitudinal chimerism analysis data from an unrelated HSCT donor-recipient pair for which there were a total of 13 informative

markers. The standard deviation and CV were calculated from all informative markers for the donor and recipient at each time point.

6.2.2.5 *Reproducibility Analysis*

Genomic DNA from two individuals was spiked to create two samples with chimeric fractions of 1% and 5%. These samples were run on three separate days using the CNV chimerism method as described. Five and two informative markers respectively for each individual were measured. The concentration of each CNV marker in the reaction was used to calculate the reproducibility CV for each marker over time. Additionally, the informative markers were used to calculate the chimeric fraction for each sample on each run and this was compared across time points.

6.2.2.6 *Comparison with FISH-Based Method*

Whole blood samples were obtained from one female and one male. White blood cell counts were obtained for each sample. The samples were centrifuged and buffy coat isolated. The cells were washed twice in phosphate buffered saline before being diluted in phosphate buffered saline to the original sample volume. The female and male samples were then mixed to create a spiked series consisting of approximately 100%, 80%, 60%, 40%, 20% and 0% female cells.

Slides were prepared for FISH according to routine procedures. Analysis was performed using CytoCell AneuVysion X centromeric (DXZ1) and Y centromeric (DYZ3) probes simultaneously following standard methods. A total of 300 interphase cells were analysed by two independent analysts (200 by analyst and 100 by checker) at each chimeric fraction.

The spiked series underwent CNV chimerism analysis as described. The concentration of each CNV marker in the reaction was used to calculate the CNV marker CV at each dilution step. Chimeric fractions were then calculated and compared with those measured by the FISH method.

6.2.2.7 *Comparison with SNP-Based Method*

Fifteen longitudinally collected peripheral blood samples from four recipients of single-donor HSCT underwent clinically-indicated SNP-based rtPCR chimerism analysis, using the established method in our institution modified from Harries et al.[295] Donor and pre-transplant recipient blood samples were genotyped. White cells were separated and counted using a Coulter Ac-T diff Analyzer (Beckman Coulter, Australia) then diluted with PBS to a concentration of approximately 5×10^6 white blood cells in 200 μ L. DNA extraction was performed using the QIAamp DNA Blood Mini-Kit (QIAGEN, Germany) according to manufacturer instructions.

SNP-based allelic discrimination was performed using three assays targeting TSC0955234, rs3918344 and rs338773[295]. Genotyping from pre-transplant recipient and donor samples was used to confirm informative combinations. Each sample and positive control (50:50 mix of donor

and recipient cells) was run in triplicate with a negative control. Each reaction consisted of 50ng DNA, 10 μ L TaqMan Master Mix (ThermoFisher Scientific, USA), 1 μ L primer/probe mix and 4 μ L distilled water. rtPCR was performed using the Corbett Rotorgene 3000 (QIAGEN, Germany) running at 95° for 10 minutes followed by 40 cycles of 92° for 15 seconds and 60° for 60 seconds. Threshold definition and analysis was performed using the stock Rotor-Gene 6000 Series Software. Chimerism quantitation was performed using the delta delta cycle threshold method.

Blinded chimerism analysis was also performed on the same samples using the CNV method as described and chimeric fractions compared.

Comparison with STR-Based Method

Whole blood samples obtained from two unrelated individuals underwent genomic DNA extraction from buffy coat cells as previously described. A working solution of 5ng/ μ L was mixed to create a spiking series consisting of 100%, 80%, 60%, 40%, 20%, 10%, 5%, 1%, 0.1% and 0% chimerism.

STR-based chimerism analysis was performed by a collaborator using the PowerPlex 16 HS System (Promega, USA) and a panel of fifteen STR loci including Penta E, D18S51, D21S11, TH01, D3S1358, FGA, TPOX, 8S1179, vWA, Penta D, CSF1PO, D16S539, D7S820, D13S317, D5S818 and Amelogenin. Detection of repeat sequence copy number was performed using the Applied Biosystems 3730 DNA Analyzer (ThermoFisher Scientific, USA). Genotyping and chimerism analysis was done using ChimerMarker software version 3.1.0 (SoftGenetics, USA).

Chimerism analysis was performed using the CNV method as described above. Comparison between the two methods was then performed.

6.2.2.8 Informativity Analysis

Genotyping results from a series of 39 recipients and 43 donor genomes that had undergone CNV chimerism analysis for clinical indications at our institution were reviewed. The number of markers tested and the number of markers that were informative for each individual in the transplant arrangement was determined. The number of individuals meeting the target of three informative markers was calculated. The contribution of relatedness between transplant pairs and presence of multiple donors to reduced marker informativity was assessed.

6.2.2.9 Statistical Analysis

CV was calculated by dividing the standard deviation (SD) by the mean. Pearson's product-moment correlation coefficients and linear regression parameters were calculated for linear correlations. Fisher's Exact Test for count data was used to calculate odds ratios (OR) and 95%

confidence intervals (95% CI). The Wilcoxon rank sum test with continuity correction was used to compare non-parametric distributions.

6.2.3 **Results**

6.2.3.1 *Sensitivity and Precision Analysis*

The mean total droplet number was 18,371 (SD 2,105). 13/1044 results (1%) were excluded due to a total droplet count of <10,000 droplets. These failed droplet generation wells (8/576 wells) were not confined to a particular marker, plate or pair although tended to occur at higher chimeric fractions.

Of the 26 tested CNV markers, 3/76 (4%) NTCs (affecting 2 markers) and 3/78 (4%) 0% chimerism samples (affecting 3 markers) had detectable one-copy target. In these 6 false positive reactions, the number of positive droplets was 1 (corresponding with a chimeric fraction of ~0.01%). In 4/6, the false positive droplet fluorescence amplitude was less than half that of the positive cluster suggesting threshold misattribution as a possible contributor. There was good separation between positive and negative droplet clusters (Supplementary Figure 6-8). Limit of blank was thus 1 positive droplet (~0.01% chimerism), although this was only seen in 3 of the 26 markers tested (the remainder had no positive droplets at 0% chimerism).

All replicates of the 26 CNV markers tested detected chimerism down to 0.5% (Supplementary Table 6-2). At 0.1% chimerism, all replicates were positive in 12/26 markers and two of three replicates were positive in the remainder. At 0.05% chimerism, all replicates were positive in 10/26 markers and two replicates were positive in 9/26. Sensitivity was reduced at 0.01% chimerism, where 11/26 markers failed to detect chimerism and only one replicate was positive in 10/26. Limit of detection was thus 0.05% chimerism.

Performance for all 26 CNV across the dynamic range (from 0% to 100%) was linear ($r > 0.97$, $p < 0.001$) (Figure 6-1, Supplementary Table 6-3 and Supplementary Figure 6-9). CV within the replicates for each marker (intra-assay variation) was <20% (mean 6%, SD of CVs 4.7%) for chimeric fractions from 100% down to 5%, with the exception of two outliers at 60% chimerism (both from the same well) (Figure 6-2, top panel). At 1% and 0.5% chimerism, the intra-assay CV was up to 50% (mean 27%, SD 12%). At 0.1% chimerism and below, the intra-assay CV was higher and subject to more variation (mean of 65%, SD 67%).

CV of the means of the marker replicates within each pair (inter-assay variation) was <8% (mean 5%, SD of CVs 2%) from 100% to 5% chimerism (Figure 6-2, bottom panel). At 1% and 0.5% chimerism, inter-assay CV was <25% (mean 16%, SD 4%). At 0.1% and 0.05%, inter-assay CV was as high as 80% (mean 51%, SD 17%). Limit of quantification was thus set at 0.5% chimerism.

The inter-assay CV results obtained from the spiking series matched those obtained from a longitudinally sampled clinical case in which the donor had 5 informative CNV markers and the recipient had 8 informative CNV markers (Figure 6-3). Inter-assay chimeric fraction CV ranged from 2-11% across the chimeric fractional range 3%-97%.

Figure 6-1: Measurement linearity and precision across the chimeric range of three representative target markers (CNV5, CNV14 and CNV39). DNA from an individual known to be one-copy for the respective marker was spiked into DNA which was zero-copy for the marker to give the chimeric fractions indicated and measured in triplicate. Note Log2(x+1) scale on both axes.

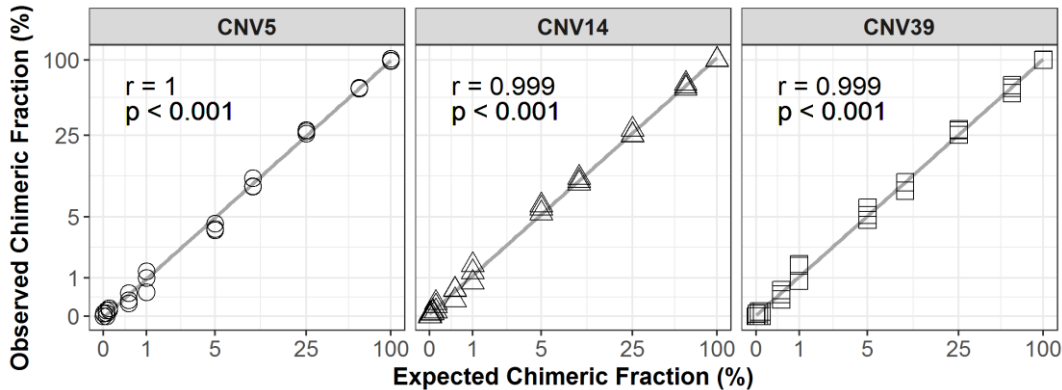


Figure 6-2: Intra- (top panel) and inter- (lower panel) assay variation of 26 CNV markers measured in triplicate across the spiking series indicated. Inter-assay variation was calculated amongst markers covered by three pairs of individuals in which one was one-copy for the markers and the other was zero-copy for the markers. Pairs 1, 2 and 3 included 12, 9 and 5 markers respectively.

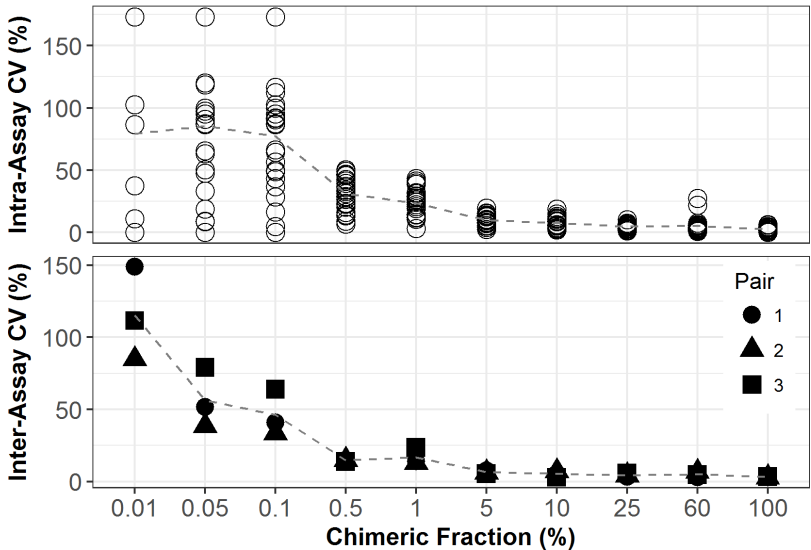
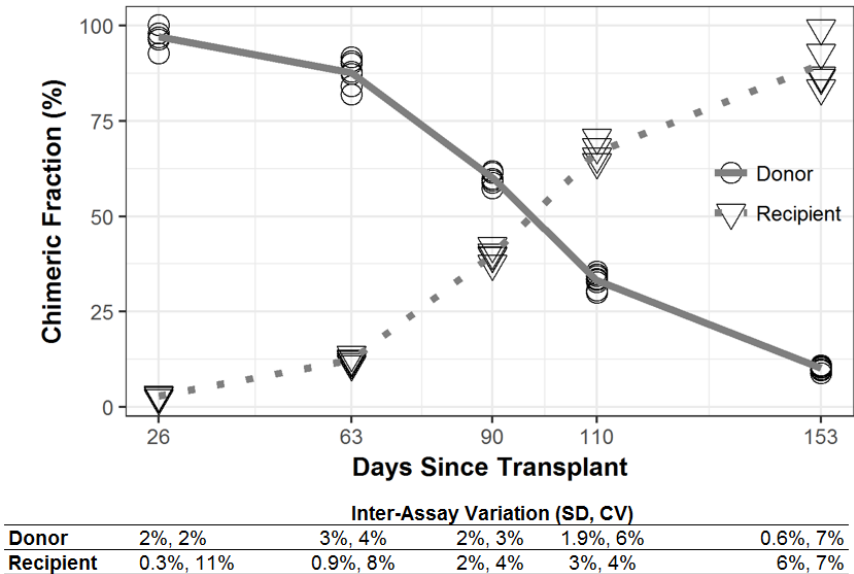


Figure 6-3: Inter-assay variation of 8 informative donor CNV markers and 5 informative recipient CNV markers at each time point during longitudinal sampling from a clinical case. The chimeric fraction SD at each time point and corresponding CV are presented below the graph.



6.2.3.2 Reproducibility Analysis

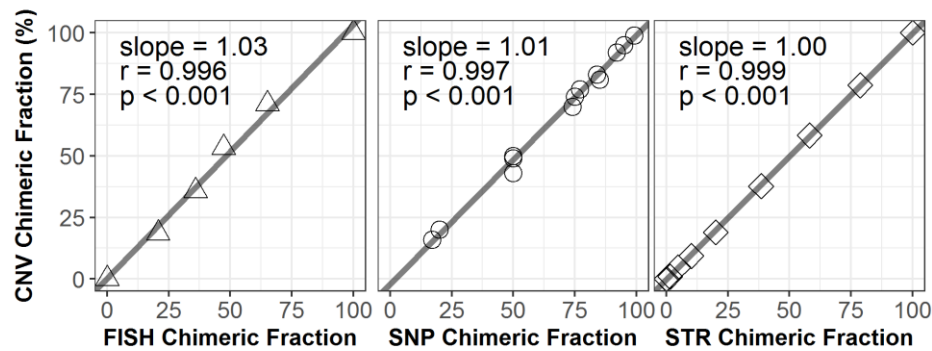
Repeated analysis of two chimeric samples using seven CNV markers on three occasions yielded mean intra-assay reproducibility CVs of 18% at 1% chimerism (SD of CVs 10%) and 5% at 5% chimerism (SD of CVs 2%). When these markers were used to determine the chimeric fraction, chimerism results from the same sample were highly consistent across time with a reproducibility CV of 2% for 1% chimerism (corresponding with a chimerism SD of 0.02%) and 1.7% for 5% chimerism (chimerism SD 0.08%).

6.2.3.3 Comparison with FISH-based Method

Four unique informative markers from the male and female were used to determine chimerism fraction via the CNV method. The inter-assay CV among the markers at each dilution step was less than 8%. There was a strong linear correlation between chimerism fraction measured using the CNV method and chimerism fraction as determined by sex mismatch FISH analysis (r=0.9957, 95% CI 0.9591-0.9995, p<0.001) (Figure 6-4, left panel).

Figure 6-4: Comparison of chimerism measurements using the CNV method with (left panel) an X/Y FISH method using a spiked series of male and female cells; (middle panel) an SNP/rtPCR

method using DNA from 15 longitudinally collected blood samples from 4 transplant recipients; (right panel) an STR-based method using a spiked series.



6.2.3.4 Comparison with SNP-based Method

The chimerism fraction measured by SNP assay ranged from 10-95% (the limits of sensitivity for this method). The CNV method identified multiple informative CNVs: pair 1 (recipient 3, donor 2), pair 2 (recipient 3, donor 4), pair 3 (recipient 1, donor 2) and pair 4 (recipient 5, donor 2). Chimerism fraction measured using the CNV methodology correlated linearly with SNP assay results ($r=0.997$, 95% CI 0.992-0.999, $p<0.001$) (Figure 6-4, middle panel).

6.2.3.5 Comparison with STR-based Method

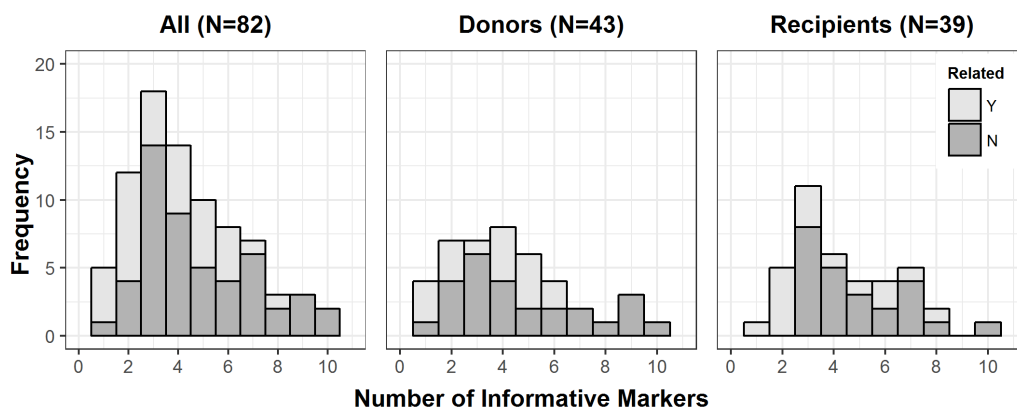
2 and 5 informative markers from the first and second individuals respectively comprising the spiking series were used to detect chimerism fractions from 0.1% to 100% using the CNV method. The STR method was not able to detect chimerism at 0.1%. At 1% expected chimerism, the STR method overestimated expected chimerism by 55% while the CNV method underestimated expected chimerism by 6%. There was a linear correlation between chimerism as measured by the STR approach with chimerism measured using CNV ($r=0.999$, 95% CI 0.9997-0.9999, $p<0.001$) (Figure 6-4, right panel).

6.2.3.6 Informativity Analysis

At least one informative marker was identified in all of the 82 donor and recipient genotyping data reviewed (Figure 6-5, left panel). There was a median of 4 informative markers. 77/82 (94%) had two or more informative markers. 65/82 (79%) had three or more informative markers. The likelihood of obtaining less than three informative markers was no different between donors and recipients (OR 0.53, 95% CI 0.14-1.8, $p=0.29$). 12/82 (15%) individuals (8 donors and 4 recipients) were part of multiple-donor transplant (MDT) arrangements (i.e. the chimerism study had to distinguish the recipient and more than one possible donor genome, either because of concurrent or prior transplantation). The distribution of informative markers was lower in individuals participating in a MDT (median 2 in MDT, 4 in non-MDT; $W=651$, $p=0.002$). 32/82

(39%) individuals were related to another individual in the same transplant arrangement. The median number of informative markers in related and unrelated individuals was similar (3.5 in related, 4 in unrelated), however, the distribution was lower in related individuals ($W=1032$, $p=0.026$) and the likelihood of achieving less than three informative markers was increased (OR 5.3, 95% CI 1.5-22, $p=0.005$).

Figure 6-5: Histograms of method informativity in 39 transplant recipients and 43 donors. In all individuals genotyped, at least one informative marker was present. The distribution of informative markers in related individuals was lower than that of unrelated individuals.



6.2.3.7 Application to a Complex HSCT Case

Chimerism analysis was performed on serial samples from a young man with an X-linked immune deficiency disorder. He had received a “piggy-back” HSCT at age 3 from his similarly affected brother, who had previously received a mismatched HSCT from his aunt. Eleven years later he developed uncontrolled systemic vasculitis that was refractory to immunosuppressant therapy and attributed to inadequate donor engraftment within the myeloid cell line. A second HSCT was performed using an unrelated male donor. In this case, neither FISH nor SNP methods were adequate to distinguish four different genomes.

Informative CNVs were identified for the aunt (3x one-copy), brother (2x one-copy), unrelated donor (2x one-copy) and recipient (1x one-copy) (Figure 6-6). Before the second transplant, most DNA (98.4%) was of recipient origin, consistent with the clinical picture of graft failure. The remaining 1.6% originated from his aunt, consistent with the history of “piggy-back” HSCT in which she was the original donor (Figure 6-7). DNA from his brother was not detected. Serial quantification of informative markers after the second transplant indicated 100% engraftment at day 24 which was sustained to day 180 (Figure 6-7).

Figure 6-6: Concentrations of 15 selected CNVs in blood collected from the recipient immediately before the second transplant (Tx#2) (top panel). Two-copy (CNVs 1, 2), one-copy (CNVs 3-10) and zero-copy (CNVs 11-15) genotypes are of recipient-origin. Trace amounts of

CNV 11-13 (asterisked) from the first transplant are specifically from the donor aunt; the non-polymorphic *ACE* target is a two-copy positive control (recipient plus any other detectable donor DNA). (bottom panel) Concentrations of the same CNVs in DNA from the unrelated donor. CNVs 14 and 15 (arrowed) are one-copy in the donor but zero-copy in the recipient, so can be used to monitor engraftment. Conversely, CNVs 1, 3, 5, 6, 9-10 are informative recipient markers that are absent in the donor (arrowed in top panel) so can be used for monitoring minimal residual disease or relapse.

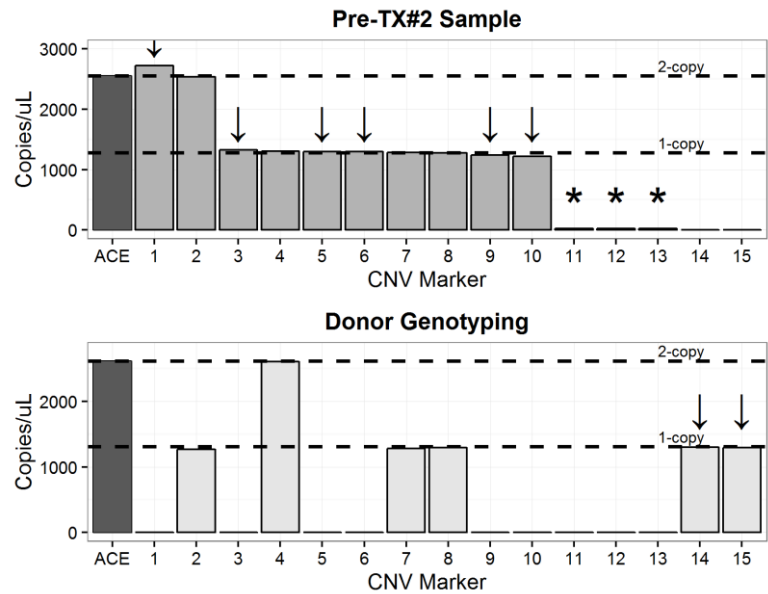
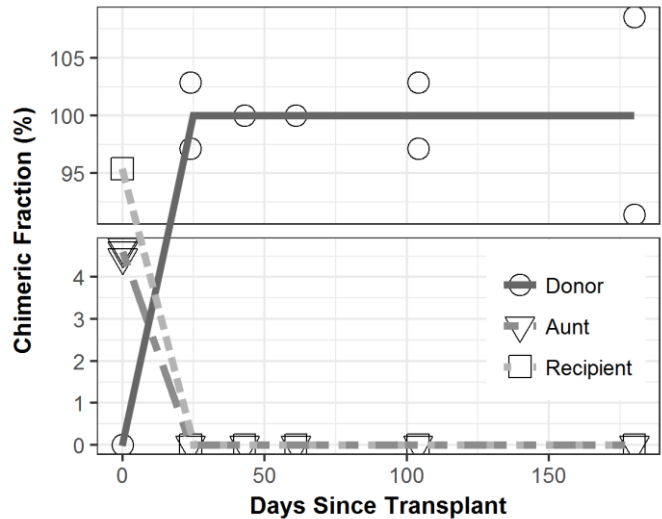


Figure 6-7: Chimeric fraction before and after Tx#2 using unique informative CNVs for donor (two CNVs), the aunt (three CNVs) and the recipient (one CNV graphed). There are an additional six CNVs that are informative if only the recipient and donor genomes are considered (arrowed in Figure 6-6). Some markers appear to indicate >100% chimerism – this reflects inter-assay variation around the mean of the informative markers for each individual that forms the basis for calculation of the chimeric fraction.



6.2.4 **Discussion**

We describe a method for chimerism analysis in allogeneic HSCT that provides incremental improvement over existing methods by balancing methodological simplicity, high informativity, sensitivity and precision against practical limitations such as availability of DNA for analysis and yet to be established significance of variation at extremely low levels of chimerism. The proposed method exploits ubiquitous CNV to distinguish recipient from donor genomic DNA present in clinical samples such as peripheral blood and bone marrow aspirates, before or after flow cytometry cell sorting.

The CNV methodology is uncomplicated, automatable, cost-effective and can be performed within a day. Interpretation of panel genotyping and identification of informative markers can be performed algorithmically. After genotyping to identify informative markers for donor(s) and recipient, subsequent chimerism analysis can be performed in 3-4 wells of a 96 well digital droplet PCR plate, exploiting duplex probe assays directed at donor and recipient informative markers within the same well if required. Unlike other ddPCR-based methods which employ reference markers (some of which require sex-mismatch) as part of the chimerism study, ACE serves only as a positive droplet control and provides a measure of the total DNA concentration, facilitating quality assurance. Additionally, no modification to standard PCR reagents is needed and extracted genomic DNA does not require additional processing such as sonic fragmentation or ddPCR concentration measurement prior to analysis.

The variable size genotype panel-based approach allows a staged strategy to screening for informative markers that at its simplest permits identification of informative markers using a panel of 15 CNV markers in most unrelated single donor HSCTs without need for donor genotyping. In related transplants or those in which more than one donor genome may be present, the panel of screened CNV markers can be increased up to 38, facilitating identification of at least three informative markers in most cases. This number (3 markers for donor(s) and 3 for recipient) permits multiple independent assessments of chimeric fraction that can be used to provide redundancy in the unlikely event of assay failure, determine precision and provide confidence in the results. This is necessary particularly at low chimeric fractions, where intra- and inter-assay variation are higher, and multiple informative markers provide internal assay replication. This is supported by observation of lower inter-assay CVs and improved reproducibility CV of chimerism fraction compared with individual markers at low levels of chimerism.

The benefits of panel informativity were exemplified in a complex HSCT case in which up to four possible genotypes were present. Although close familial relationships in the first transplant and the use of a same sex donor in the second transplant precluded use of institutional SNP/rPCR and sex mismatch FISH assays, unique CNV markers were identified for all four possible

genomes. Consequently, successful engraftment and freedom from relapse on serial follow-up was confirmed. High test informativity is likely to become more necessary with the increasing use of dual cord donors and repeat transplantation.

Compared with other methods such as SNP and STR, sensitivity is enhanced and noise reduced by exploitation of copy number deletion. These variants are large enough to permit selection of forward and reverse PCR primer target sequences within the deleted regions. This produces a truly “negative background” in null alleles against which informative markers undergo quantification using ddPCR. The mean number of positive droplets detected in NTC and 0% chimerism samples for the CNV markers studied was 0.39 (corresponding to a chimeric fraction well below 0.01%). 4/6 assays with one false positive droplet were caused by threshold misattribution which arose because thresholds for droplet positivity were defined based on the droplet clusters visualised en bloc for each assay across the spiking series. In practice, thresholding is applied to samples on a per-well basis and such misattribution is unlikely. Limit of blank was conservatively set at one positive droplet (~0.01% chimerism) but it is almost certainly below this.

Sensitivity studies indicated the majority of 26 tested CNV markers were able to detect chimeric fractions to 0.05%, the limit of detection defined for the assay panel. The sensitivity and accuracy of this approach exceed that of the STR-based comparator at the lower chimeric fractions. For clinical application, the limit of quantification was set at 0.5% chimerism. Inter- and intra-assay performance below this chimeric fraction deteriorates significantly using 30ng of template DNA. At 0.5% chimerism, the maximum intra-assay CV was 50% (mean CV 31%) and maximum inter-assay CV was 15% (mean CV 14.6%). In practical terms, at this chimeric fraction, the intra-assay CV of 50% correlates with a SD of 0.025%. Given present applications of chimerism analysis, it is unlikely that this variation will be of clinical significance. If improved resolution at or below this chimeric fraction is desired, higher input DNA quantities or summation of wells with the same target will permit improved precision and discriminative capacity if clinically desired. Other groups have demonstrated the high discriminative capacity of digital droplet PCR at low levels of chimerism using 200ng[300] and 100ng-600ng[303] of input DNA. Disadvantages of using higher input DNA quantities include need for sonic fragmentation prior to ddPCR (BioRad manufacturer instructions) and limited application to flow sorted cell populations where sufficient DNA may not be available to permit chimerism study using multiple independent informative markers.

6.2.5 Conclusion

The simplicity of the CNV approach, combined with the advantages of high informativity and ddPCR technology, recommend it as an improved approach to monitoring of allogeneic HSCT

chimerism. Newer genomic methods such as this one offer markedly improved sensitivity but little is presently known about the clinical relevance or utility of detecting mixed chimerism at such low levels. Opportunities exist to study possible roles in guiding treatment intensity, early detection of relapse and prognostication.

6.3 FUNDING

This work was supported by the Roche Organ Transplant Research Fund (Grant 529172731), the Murdoch Children's Research Institute and the Victorian Government's Operational Infrastructure Support Program. Stipend support for author JBW from Kidney Health Australia and the National Health and Medical Research Council.

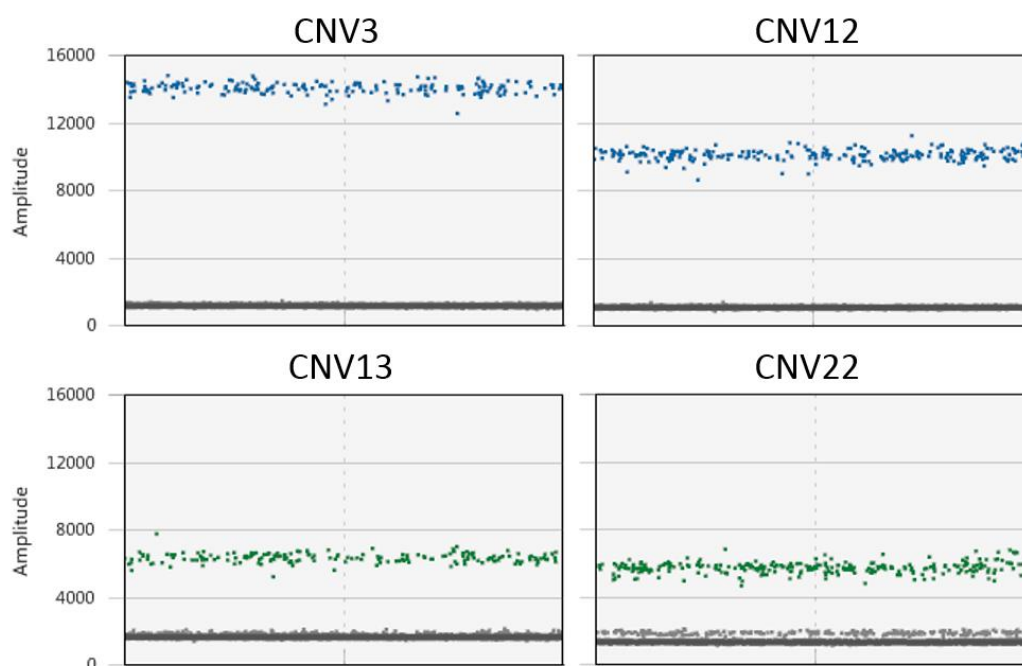
6.4 DISCLOSURE OF CONFLICTS OF INTEREST

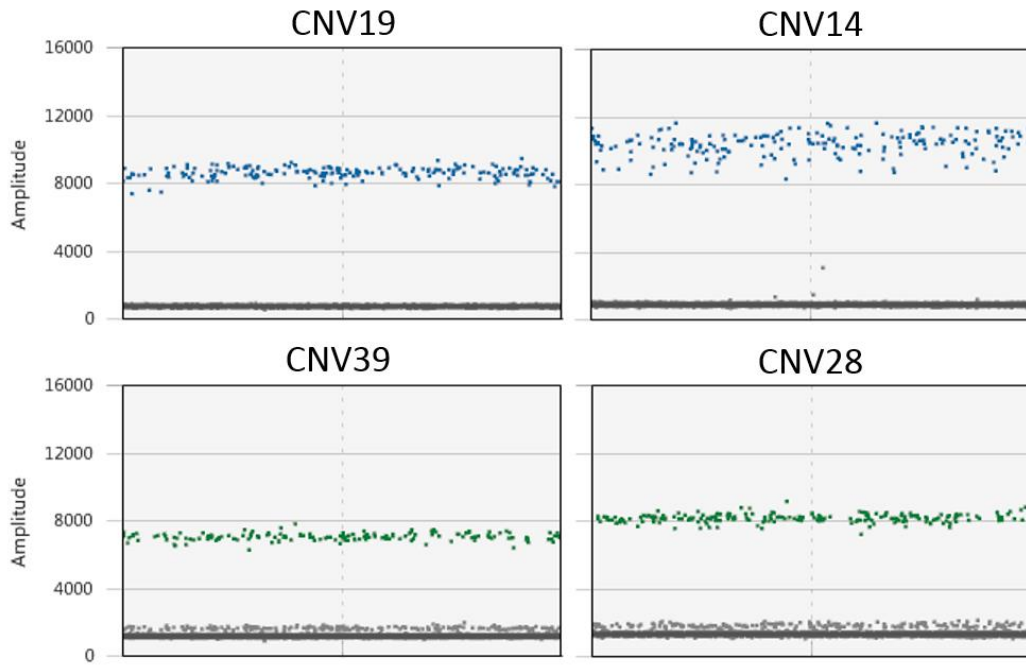
JBW: None; LL: None; MS: None, TH: None, OM: None, IB: None, SC: None, JC: None, VP: None; FM: None; RC: None. DLB & HRS: This work is the subject of a patent application (PCT/AU2016/050499) in which DLB & HRS are listed as inventors.

6.5 SUPPLEMENTARY MATERIAL

6.5.1 Droplet Digital PCR Plots

Figure 6-8: 1D ddPCR plots of one copy markers at 10% chimerism demonstrating separation of positive clusters from negative clusters for selected representative assays under standard conditions. The markers in the odd rows were duplexed with the corresponding markers in the even rows. Y-axis scales and limits have been standardised to permit inter-assay comparison. X-axis plots event number.





6.5.2 CNV Loci

Table 6-1: Genomic locations, sizes and population frequencies of the 38 CNV panel.

CNV	Chromosome	hg18 Start	hg18 End	Size (bp)	0 Copy Freq.	1 Copy Freq.	2 Copy Freq.
01	chr1	157134152	157136674	2522	0.51	0.42	0.07
02	chr18	53097735	53099702	1967	0.52	0.4	0.08
03	chr8	112363260	112365469	2209	0.5	0.41	0.08
04	chr14	21951540	21952070	530	0.46	0.43	0.11
05	chr16	25247611	25250595	2984	0.45	0.47	0.08
06	chr7	24004755	24006584	1829	0.47	0.38	0.15
07	chr13	37955318	37958191	2873	0.43	0.43	0.14
08	chr8	4110308	4112335	2027	0.43	0.43	0.13
09	chr18	33560073	33560645	572	0.42	0.45	0.12
10	chr10	108020303	108022518	2215	0.41	0.42	0.17
12	chr7	73467042	73469172	2130	0.55	0.33	0.12
13	chr10	4698559	4700493	1934	0.61	0.31	0.08
14	chr2	54419132	54420976	1844	0.59	0.37	0.04
15	chr13	38831627	38833387	1760	0.55	0.33	0.12
16	chr2	76627234	76628943	1709	0.6	0.33	0.08
17	chr1	185731458	185733106	1648	0.62	0.33	0.04
18	chr9	17900038	17901633	1595	0.61	0.36	0.03
19	chr2	121513996	121515257	1261	0.53	0.36	0.1
20	chr2	159668040	159669209	1169	0.5	0.43	0.07
21	chr2	219761436	219762572	1136	0.39	0.42	0.18
22	chr13	71743783	71744892	1109	0.41	0.55	0.04
23	chr22	33975445	33976472	1027	0.45	0.43	0.12
24	chr6	95250114	95251113	999	0.4	0.47	0.13

25	chr8	112363455	112364436	981	0.57	0.38	0.05
26	chr1	157915386	157916253	867	0.6	0.35	0.05
27	chr9	88344772	88345596	824	0.56	0.36	0.07
28	chr6	18510099	18510860	761	0.41	0.44	0.15
29	chr15	97392250	97392985	735	0.51	0.43	0.06
30	chr15	37531690	37532136	446	0.52	0.42	0.07
31	chr1	208148236	208150607	2371	0.38	0.45	0.17
32	chr1	150822234	150856715	34481	0.39	0.46	0.15
33	chr1	72538870	72584557	45687	0.39	0.46	0.15
34	chr20	21234267	21236529	2262	0.52	0.42	0.07
36	chr4	9823254	9844366	21112	0.55	0.33	0.12
37	chr5	57359283	57369499	10216	0.61	0.36	0.02
38	chr6	100141275	100141994	719	0.62	0.38	0
39	chr7	147704059	147707263	3204	0.66	0.31	0.03
41	chr5	1977604	1978111	507	0.67	0.33	0

6.5.3 Sensitivity Analysis

Table 6-2: Proportion of tested marker replicates that were positive at each chimeric fraction.

		Chimerism Fraction (%)										
		0	0.01	0.05	0.1	0.5	1	5	10	25	60	100
CNV Marker	3	0.33	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	5	0.00	1.00	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	6	0.00	0.00	0.67	0.33	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	8	0.00	0.33	0.33	0.33	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	10	0.00	0.00	0.67	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	12	0.00	0.33	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	13	0.33	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	14	0.00	0.33	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	15	0.00	0.33	1.00	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	17	0.00	0.00	1.00	0.33	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	19	0.00	0.00	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	22	0.00	0.33	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	23	0.00	0.00	0.33	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	24	0.00	0.67	0.67	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	25	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	27	0.00	0.33	0.33	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	28	0.00	0.67	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	29	0.00	0.33	0.33	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	30	0.00	0.00	0.00	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	31	0.00	0.33	0.33	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	32	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	33	0.00	0.00	1.00	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	34	0.00	0.00	0.33	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

36	0.33	0.00	0.67	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00
37	0.00	0.33	0.67	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
39	0.00	0.33	0.67	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00

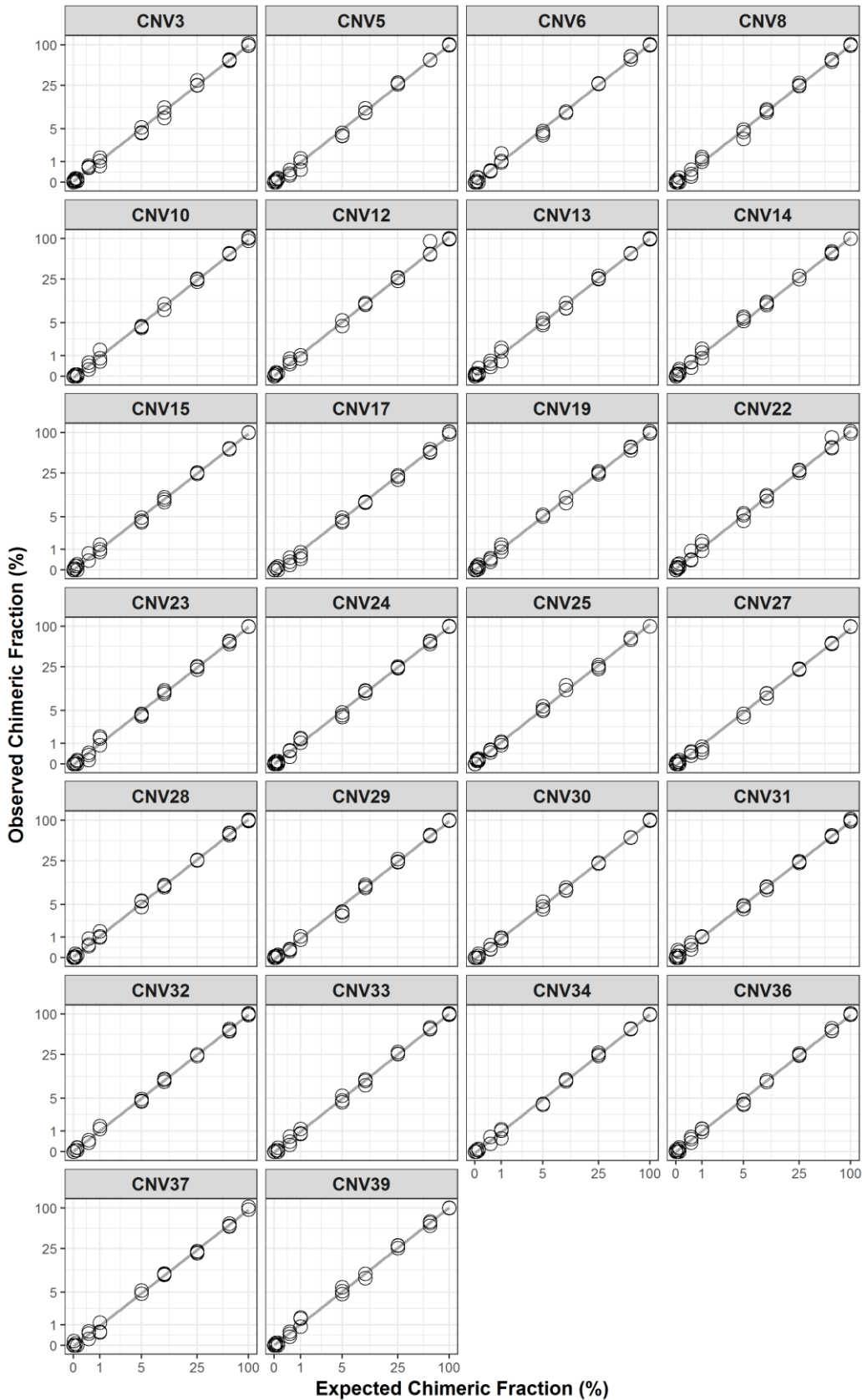
6.5.4 Assay Linear Model Characteristics

Table 6-3: Linear model characteristics and correlation coefficient values from marker performance in the sensitivity analysis spiking series.

CNV	Slope	y Intercept (%)	Correlation Coefficient (r)	95% Confidence Interval	p-value
3	0.9993	0.05072	0.999	0.998-0.9995	<0.001
5	1.002	-0.003717	1	0.9991-0.9998	<0.001
6	1.023	0.0911	0.998	0.9965-0.9992	<0.001
8	0.9973	-0.08561	0.999	0.9988-0.9997	<0.001
10	1.002	-0.29	0.999	0.9973-0.9994	<0.001
12	1.04	0.2862	0.986	0.9716-0.9931	<0.001
13	1.002	0.1011	1	0.9992-0.9998	<0.001
14	1.007	0.3131	0.999	0.9989-0.9997	<0.001
15	0.9901	-0.1117	1	0.9994-0.9998	<0.001
17	0.9737	-0.6833	0.997	0.9936-0.9985	<0.001
19	0.9953	-0.02334	0.999	0.9976-0.9995	<0.001
22	1.032	0.4212	0.991	0.9812-0.9955	<0.001
23	0.9917	-0.1466	0.999	0.9985-0.9996	<0.001
24	0.9943	-0.09322	0.999	0.9988-0.9997	<0.001
25	1.036	0.2654	0.998	0.9963-0.9992	<0.001
27	0.9856	-0.4574	0.999	0.9985-0.9997	<0.001
28	1.012	0.2115	0.999	0.9983-0.9996	<0.001
29	1.001	-0.2179	1	0.9993-0.9998	<0.001
30	0.9838	-0.4252	0.999	0.9983-0.9996	<0.001
31	0.9947	-0.2169	0.999	0.9981-0.9995	<0.001
32	0.9922	-0.1077	0.999	0.9987-0.9997	<0.001
33	1.007	0.01953	0.999	0.9989-0.9997	<0.001
34	1.004	-0.1212	1	0.9995-0.9999	<0.001
36	0.9956	-0.1378	0.999	0.9988-0.9997	<0.001
37	0.9749	-0.3253	0.998	0.9953-0.9989	<0.001
39	0.9979	0.1546	0.999	0.9981-0.9996	<0.001

6.5.5 Assay Performance

Figure 6-9: Individual marker performance across the spiked chimeric range. Note $\text{Log}_2(x+1)$ scale on both axes.



CHAPTER 7 CONCLUSIONS AND FUTURE DIRECTIONS

A state of genetic chimerism is created by allogenic transplantation as well as natural events such as pregnancy.[4] Such chimerism may exist in a variety of compartments, including a broad range of cell types and tissues, as well as in cfDNA detectable in bodily fluids. The quality and quantity of chimerism is highly variable between compartments, individuals, and over time. The factors affecting this chimerism, as well as the effects it has, are complex, poorly understood, and remain under active investigation. It is highly likely that chimerism has active functional effects, rather than simply being a passive state, and preliminary evidence exists that it plays a role in alloimmunisation and tolerance.[76-78, 83]

Change in chimerism status after allogenic transplantation may provide insights into allograft health. Donor chimerism assessment of circulating blood or bone marrow cells is a fundamental guiding metric in the management of allogeneic HCT recipients, providing an assessment of engraftment, relapse and graft rejection, as well as prognostic information.[57] After solid organ transplantation, proof-of-principle and diagnostic validity studies, including the present work, indicate that elevated levels of cfDNA expressing a graft/donor genotype may be reflective of graft injury.[120, 143, 146, 148, 151, 153, 160, 161] Whilst the diagnostic performance of these approaches has been well studied, the underlying pathophysiological mechanisms affecting chimerism in the transplant context remains poorly understood.

The ability to robustly discriminate and quantify genetic chimerism is an important pre-requisite for the study of chimerism. This relies upon the identification of discriminatory markers for each chimeric genotype, irrespective of donor-recipient sex and relatedness, and the use of a sensitive and precise quantification platform. Traditional approaches have been hampered by lack of broadly applicable discriminative capability, poor informativity, qualitative or cumbersome methods, or quantitative techniques that were inadequately sensitive. Whilst these approaches are responsible for much of the current understanding of chimerism, the emerging diagnostic role of chimerism analysis, as well as the study of lower levels of chimerism in smaller compartments, requires the development of improved methods of chimerism analysis.

The sequencing of the human genome and related projects cataloguing human genetic variation has permitted the development of chimerism analysis approaches targeting this variation. This thesis presents the development of a chimerism analysis method targeting highly polymorphic CNV DNA sequences. To the best of our knowledge, this is the first and only chimerism analysis approach targeting this form of genetic variation. The size and abundance of CNV permitted the development of a relatively large panel of ddPCR TaqMan hydrolysis probe assays sited within, and highly specific to, these CNV. The consequence of this design, and the target of structural variation rather than sequence variation, is that the presence of bi-allelic copy number deletion (a

zero copy genotype) creates a negative background against which to detect and quantify a non-zero copy genotype for chimerism analysis.

The size of the panel, as well as the strategically selected allelic frequencies of the CNV in the population, permit high informativity, which has been demonstrated in this work in related donor-recipients pairs in KT and allogeneic HCT, as well as more complex multi-donor and piggy-back donor allogeneic HCT transplants. In some of these latter scenarios, the full CNV assay panel of 39 assays was required, highlighting that panel size, not just allelic informativity, is important. Whilst published methods targeting larger indel polymorphisms also allow a negative background genotype, these approaches have relatively small panels of less than 10-15 assays that would make chimerism analysis difficult in closely related donor-recipient pairs or multiple donor transplant scenarios.[147, 148]

The chimerism analysis method developed in this work uses ddPCR to perform absolute quantification of the graft and recipient contribution to the compartment under study, from which the chimeric fraction can be calculated. To the best of our knowledge, the diagnostic validity studies presented in this work represent the first time such an approach has been studied in KT and allogeneic HCT. Other methods have quantified chimeric fraction using relative measures, after pre-amplification of DNA, using sequence variation, or using qrtPCR that relies upon quantification against a standard curve.[120, 140, 153-155, 157, 160, 163]

7.1 CELL-FREE DNA CHIMERISM ANALYSIS

Analytical validation of the CNV ddPCR method for cfDNA chimerism analysis indicated adequacy for the purposes of conducting a diagnostic validity study in KT recipients. Conceptual validation of a basic chimeric sample genotyping method confirmed that separate donor and recipient genotyping was not required. Whilst not unique to this method, the genotyping approach developed is simpler than other methods, which require more complex bioinformatic approaches based upon probability modelling and mixture modelling.

Comparison of analytical performance with other methods was not possible due to limited availability of those methods, proprietary or incompletely published methodologies, relative quantification or incomparable quantification platform, or absence of analytical validation. The most directly comparable approach, a qrtPCR method targeting 12 indels, only studied 8 KT recipients for the first 3 days after transplantation and did not include analytical validation.[148]

The analytical validation identified the need to better study performance at lower levels of target cfDNA (between 0-25 copies/mL), the possible development of improved assay performance through optimisation of the quantity of DNA used in the reaction, and the need to assess inter-assay variation using repeated testing, as areas of future work. Standardisation for cfDNA

extraction yield using spiked non-human DNA sequences may additionally improve the robustness of the absolute quantification.

CNV-based ddPCR absolute quantification of graft-derived cfDNA and calculation of graft fraction identified significantly elevated levels of these measurements in ABMR, particularly aAMR. For the diagnosis of ABMR, this test performed well with a ROC AUC of 0.91. These findings are consistent with those of a number of other studies that have been published recently (Table 7-1). Notably, and particularly for the diagnosis of aAMR, we found that absolute measurement of graft-derived cfDNA tended to perform better than the relative measurement of graft fraction. Oellerich et al. have also reported that multiplying graft fraction by a separate ddPCR measurement of total cell-free DNA improved diagnostic performance for acute rejection.[120] These findings argue that relative measures of cfDNA chimerism may be prone to confounding by variation in total cfDNA levels.

Table 7-1: Diagnostic validity studies of cell-free DNA chimerism analysis in kidney transplantation.

Publication	Number of samples studied	Method	ABMR vs no ABMR ROC AUC (95% CI)
Bloom et al. 2017 [153] (CareDx Allosure)	107 graft dysfunction biopsy-matched	Relative quantification using SNPs and sequencing	0.87 (0.75-0.97)
Jordan et al. 2018 [154] (CareDx Allosure)	33 DSA positive biopsy-matched	Relative quantification using SNPs and sequencing	0.86 (0.70-0.98)
Whitlam et al. 2019 [146]	61 graft dysfunction biopsy-matched	Absolute quantification using CNVs and ddPCR	0.91 (0.82-0.98)
Oellerich et al. 2019 [120] (Chronix Biomedical)	22 biopsy proven rejection, 408 “stable phase” samples	Relative quantification using SNPs and ddPCR	0.73 (0.61-0.82) (included CMR)
Gielis et al. 2019 [160]	43 protocol biopsy-matched, 46 graft dysfunction biopsy-matched	Relative quantification using SNPs and sequencing	0.64 (0.50-0.78) (included CMR)
Sigdel et al. 2019 [158] (Natera Prospera)	217 protocol and graft dysfunction biopsy-matched	Relative quantification using SNPs and sequencing	0.87 (0.80-0.95) (included CMR, bCMR)

Whilst the association between graft-derived cfDNA and rejection supports the hypothesis that alloimmune injury to the allograft is associated with increased release of donor-specific cfDNA as a result of apoptosis and necrosis, this mechanism remains unproven. In particular, no studies to date have been performed examining the tissue of origin of graft-derived cfDNA. Furthermore, limited examination of the quantity and quality of cellular microchimerism after KT has been performed, and the relationship between cellular and cfDNA chimerism after KT is unknown.

The ability to measure chimerism in absolute and relative terms, as well as the analytical validation of the same ddPCR CNV method for cfDNA and cellular chimerism analysis, uniquely positions this approach to answer these questions. To the best of our knowledge, no other cfDNA chimerism method has been studied in this way.

The commercial investment in this biomarker, the results of diagnostic validity studies to date, and the United States Medicare and Medicaid subsidization of the CareDx Allosure test, presently makes plasma graft-derived cfDNA the most widely adopted KT rejection biomarker in clinical practice. It should be noted, however, that this has only really occurred in the United States. In countries with a heavily subsidised health service, such as Australia, there will probably be an expectation that clinical utility is first demonstrated prior to funding. The CNV ddPCR approach proposed in this work offers a faster turnaround time than sequencing based techniques (1 day vs 2-3 days) and at lower cost. The assay costs less than AU\$300 to run, which is in contrast to current Medicare reimbursements for the CareDx Allosure test of US\$2841 per test.

The existing clinical data concerning graft-derived cfDNA indicates that this is not a perfect KT rejection biomarker. Limitations with respect to diagnosis of CMR and other allograft pathology means that invasive biopsy is still required for definitive diagnosis. Further studies are also needed to determine the diagnostic performance of this biomarker in the absence of graft dysfunction. Additional opportunities exist to study the role of this biomarker in monitoring the response to treatment, prognostication, and as a marker of immunosuppression, as well as to better define the contribution that urinary cfDNA chimerism analysis may provide.

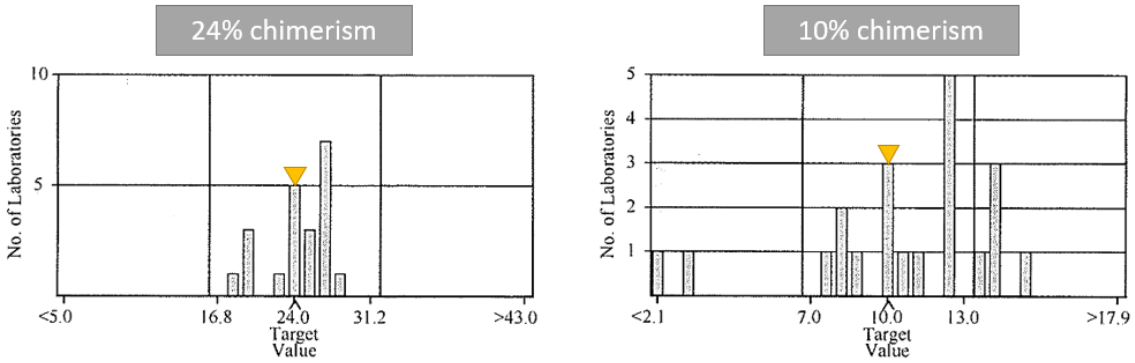
7.2 CELLULAR DNA CHIMERISM ANALYSIS

The validity of the ddPCR CNV approach for analysis of blood cellular chimerism analysis was also demonstrated in this work. When compared to other common methods of chimerism analysis, including FISH, SNP-based, and STR-based chimerism analysis, the ddPCR CNV method performed comparably and demonstrated better sensitivity. Using 30 ng of DNA, chimeric fractions as low as 0.05% were detected. Higher assay sensitivity, and improvements in precision, would be expected using higher concentrations of DNA, although this is not always available when the cell population under study is relatively small, such as during early cell count recovery after bone marrow transplantation or in cellular subpopulations isolated by flow sorting.

Whilst a key finding of this work was that the CNV ddPCR assay permits the measurement of very low level microchimerism, the clinical significance of this in allogeneic HCT is not known. Early detection of relapse via detection of low level microchimerism warrants investigation, and may permit early treatment with lower toxicity. Presently, however, the assay performance characteristics for quantification of microchimerism during routine post-allogeneic HCT chimerism assessment are probably in excess of what is required.

The relative ease of genotyping donor and recipient, the high throughput achievable, prospects of an automated workflow, and the scalability of a highly informative assay panel that permits chimerism analysis in complex and related donor-recipient transplants, recommend this approach for translation to clinical practice. At our institution (Royal Children’s Hospital, Melbourne), this methodology has been implemented for routine post-allogeneic HCT chimerism analysis and has undergone National Association of Testing Authorities (NATA) accreditation. Annual Royal College of Pathologists Australia Quality Assurance Program participation indicates that our laboratory routinely performs at or near that top of the field with respect to cellular chimerism analysis nationally (Figure 7-1).

Figure 7-1: Royal College of Pathologists Australia chimerism analysis Quality Assurance Program results from 2017, reproduced with permission. The arrow head indicates that the CNV-based ddPCR chimerism result was on target in both samples. The histogram indicates the distribution of results generated by 21 other national laboratories.



The established analytical performance characteristics of this approach also recommend it for use in a research setting to study chimerism in other states and compartments. The ddPCR CNV approach could be used to study cellular microchimerism during and following pregnancy, or in the context of solid organ transplantation. Additionally, this approach may be used to study the fate of emerging cellular therapies such as use of *in vivo* infused or implanted stem cells for organ regeneration, or use of cytotoxic T cells to treat infections after transplantation.[13, 15] In this latter application, the informativity of this approach holds promise as multiple donors (e.g. kidney donor, cytotoxic T cell donor) must be distinguished from the recipient chimeric sample.

The concurrent analytical validation of CNV-based ddPCR chimerism analysis also positions it well to study the significance of cfDNA, and its relationship to cellular chimerism, after allogeneic HCT. This analyte has been poorly studied in this context, and early studies indicate potential utility in monitoring graft-versus-host disease and identification of early relapse.[125, 126]

To conclude, this thesis presents the development and validation of a novel method of chimerism analysis based upon CNV polymorphism and ddPCR. In addition to its use in the diagnosis of KT rejection and detection of chimerism after allogeneic HCT, the method holds promise as a means to study chimerism more broadly and in a variety of compartments.

REFERENCES

- [1] Anderson D, Billingham RE, Lampkin GH, Medawar PB. The use of skin grafting to distinguish between monozygotic and dizygotic twins in cattle. *Heredity*. 1951;5:379-397.
- [2] Gartler SM, Waxman SH, Giblett E. An XX/XY human hermaphrodite resulting from double fertilization. *Proceedings of the National Academy of Sciences of the United States of America*. 1962;48:332-335.
- [3] Malan V, Vekemans M, Turleau C. Chimera and other fertilization errors. *Clin Genet*. 2006;70:363-373.
- [4] Nelson JL. The otherness of self: microchimerism in health and disease. *Trends Immunol*. 2012;33:421-427.
- [5] Lo YM, Lau TK, Chan LY, Leung TN, Chang AM. Quantitative analysis of the bidirectional fetomaternal transfer of nucleated cells and plasma DNA. *Clin Chem*. 2000;46:1301-1309.
- [6] Maloney S, Smith A, Furst DE, et al. Microchimerism of maternal origin persists into adult life. *J Clin Invest*. 1999;104:41-47.
- [7] Bianchi DW, Zickwolf GK, Weil GJ, Sylvester S, DeMaria MA. Male fetal progenitor cells persist in maternal blood for as long as 27 years postpartum. *Proceedings of the National Academy of Sciences of the United States of America*. 1996;93:705-708.
- [8] Gammill HS, Harrington WE. Microchimerism: Defining and redefining the prepregnancy context - A review. *Placenta*. 2017;60:130-133.
- [9] Dutta P, Burlingham WJ. Microchimerism: tolerance vs. sensitization. *Curr Opin Organ Transplant*. 2011;16:359-365.
- [10] Nelson JL. Maternal-fetal immunology and autoimmune disease: is some autoimmune disease auto-alloimmune or allo-autoimmune? *Arthritis Rheum*. 1996;39:191-194.
- [11] Stevens AM, Hermes HM, Rutledge JC, Buyon JP, Nelson JL. Myocardial-tissue-specific phenotype of maternal microchimerism in neonatal lupus congenital heart block. *Lancet*. 2003;362:1617-1623.
- [12] Lo YM, Corbetta N, Chamberlain PF, et al. Presence of fetal DNA in maternal plasma and serum. *Lancet*. 1997;350:485-487.

- [13] Leen AM, Christin A, Myers GD, et al. Cytotoxic T lymphocyte therapy with donor T cells prevents and treats adenovirus and Epstein-Barr virus infections after haploidentical and matched unrelated stem cell transplantation. *Blood*. 2009;114:4283-4292.
- [14] Alizadeh M, Bernard M, Danic B, et al. Quantitative assessment of hematopoietic chimerism after bone marrow transplantation by real-time quantitative polymerase chain reaction. *Blood*. 2002;99:4618-4625.
- [15] Gutierrez-Fernandez M, Rodriguez-Frutos B, Ramos-Cejudo J, et al. Effects of intravenous administration of allogenic bone marrow- and adipose tissue-derived mesenchymal stem cells on functional recovery and brain repair markers in experimental ischemic stroke. *Stem Cell Res Ther*. 2013;4:11.
- [16] Lo YM, Tein MS, Pang CC, Yeung CK, Tong KL, Hjelm NM. Presence of donor-specific DNA in plasma of kidney and liver-transplant recipients. *Lancet*. 1998;351:1329-1330.
- [17] Durnam DM, Anders KR, Fisher L, O'Quigley J, Bryant EM, Thomas ED. Analysis of the origin of marrow cells in bone marrow transplant recipients using a Y-chromosome-specific in situ hybridization assay. *Blood*. 1989;74:2220-2226.
- [18] Eikmans M, van Halteren AG, van Besien K, van Rood JJ, Drabbels JJ, Claas FH. Naturally acquired microchimerism: implications for transplantation outcome and novel methodologies for detection. *Chimerism*. 2014;5:24-39.
- [19] Gielis EM, Ledeganck KJ, De Winter BY, et al. Cell-Free DNA: An Upcoming Biomarker in Transplantation. *Am J Transplant*. 2015;15:2541-2551.
- [20] Zhang J, Tong KL, Li PK, et al. Presence of donor- and recipient-derived DNA in cell-free urine samples of renal transplantation recipients: urinary DNA chimerism. *Clin Chem*. 1999;45:1741-1746.
- [21] Lui YY, Woo KS, Wang AY, et al. Origin of plasma cell-free DNA after solid organ transplantation. *Clin Chem*. 2003;49:495-496.
- [22] McDaniel HB, Yang M, Sidner RA, Jindal RM, Sahota A. Prospective study of microchimerism in transplant recipients. *Clin Transplant*. 1999;13:187-192.
- [23] Lawler M, Humphries P, McCann SR. Evaluation of mixed chimerism by in vitro amplification of dinucleotide repeat sequences using the polymerase chain reaction. *Blood*. 1991;77:2504-2514.
- [24] Pujal JM, Grinyo JM, Gil-Vernet S, et al. Early hematopoietic microchimerism predicts clinical outcome after kidney transplantation. *Transplantation*. 2007;84:1103-1111.

- [25] Thiede C, Florek M, Bornhauser M, et al. Rapid quantification of mixed chimerism using multiplex amplification of short tandem repeat markers and fluorescence detection. *Bone Marrow Transplant*. 1999;23:1055-1060.
- [26] Gadi VK, Nelson JL, Boespflug ND, Guthrie KA, Kuhr CS. Soluble donor DNA concentrations in recipient serum correlate with pancreas-kidney rejection. *Clin Chem*. 2006;52:379-382.
- [27] Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409:860-921.
- [28] International HapMap C. The International HapMap Project. *Nature*. 2003;426:789-796.
- [29] Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. The next-generation sequencing revolution and its impact on genomics. *Cell*. 2013;155:27-38.
- [30] Genomes Project C, Abecasis GR, Altshuler D, et al. A map of human genome variation from population-scale sequencing. *Nature*. 2010;467:1061-1073.
- [31] Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. *Nature reviews Genetics*. 2006;7:85-97.
- [32] Genomes Project C, Abecasis GR, Auton A, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491:56-65.
- [33] International HapMap C, Frazer KA, Ballinger DG, et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature*. 2007;449:851-861.
- [34] Hinds DA, Stuve LL, Nilsen GB, et al. Whole-genome patterns of common DNA variation in three human populations. *Science*. 2005;307:1072-1079.
- [35] Sebat J, Lakshmi B, Troge J, et al. Large-scale copy number polymorphism in the human genome. *Science*. 2004;305:525-528.
- [36] Iafrate AJ, Feuk L, Rivera MN, et al. Detection of large-scale variation in the human genome. *Nat Genet*. 2004;36:949-951.
- [37] Tuzun E, Sharp AJ, Bailey JA, et al. Fine-scale structural variation of the human genome. *Nat Genet*. 2005;37:727-732.
- [38] Conrad DF, Pinto D, Redon R, et al. Origins and functional impact of copy number variation in the human genome. *Nature*. 2010;464:704-712.
- [39] Sykes PJ, Neoh SH, Brisco MJ, Hughes E, Condon J, Morley AA. Quantitation of targets for PCR by use of limiting dilution. *Biotechniques*. 1992;13:444-449.

- [40] Vogelstein B, Kinzler KW. Digital PCR. *Proceedings of the National Academy of Sciences of the United States of America*. 1999;96:9236-9241.
- [41] Bhat S, Herrmann J, Armishaw P, Corbisier P, Emslie KR. Single molecule detection in nanofluidic digital array enables accurate measurement of DNA copy number. *Anal Bioanal Chem*. 2009;394:457-467.
- [42] Warren L, Bryder D, Weissman IL, Quake SR. Transcription factor profiling in individual hematopoietic progenitors by digital RT-PCR. *Proceedings of the National Academy of Sciences of the United States of America*. 2006;103:17807-17812.
- [43] Spurgeon SL, Jones RC, Ramakrishnan R. High throughput gene expression measurement with real time PCR in a microfluidic dynamic array. *PloS one*. 2008;3:e1662.
- [44] Morrison T, Hurley J, Garcia J, et al. Nanoliter high throughput quantitative PCR. *Nucleic Acids Res*. 2006;34:e123.
- [45] Hindson BJ, Ness KD, Masquelier DA, et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem*. 2011;83:8604-8610.
- [46] Pinheiro LB, Coleman VA, Hindson CM, et al. Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Analytical chemistry*. 2012;84:1003-1011.
- [47] Bustin SA, Nolan T. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *J Biomol Tech*. 2004;15:155-166.
- [48] Sanders R, Huggett JF, Bushell CA, Cowen S, Scott DJ, Foy CA. Evaluation of digital PCR for absolute DNA quantification. *Analytical chemistry*. 2011;83:6474-6484.
- [49] Dingle TC, Sedlak RH, Cook L, Jerome KR. Tolerance of droplet-digital PCR vs real-time quantitative PCR to inhibitory substances. *Clin Chem*. 2013;59:1670-1672.
- [50] Weisenberger DJ, Trinh BN, Campan M, et al. DNA methylation analysis by digital bisulfite genomic sequencing and digital MethyLight. *Nucleic Acids Res*. 2008;36:4689-4698.
- [51] Whale AS, Huggett JF, Cowen S, et al. Comparison of microfluidic digital PCR and conventional quantitative PCR for measuring copy number variation. *Nucleic Acids Res*. 2012;40:e82.
- [52] Whale AS, Cowen S, Foy CA, Huggett JF. Methods for applying accurate digital PCR analysis on low copy DNA samples. *PloS one*. 2013;8:e58177.

- [53] Hindson CM, Chevillet JR, Briggs HA, et al. Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat Methods*. 2013;10:1003-1005.
- [54] Weaver S, Dube S, Mir A, et al. Taking qPCR to a higher level: Analysis of CNV reveals the power of high throughput qPCR to enhance quantitative resolution. *Methods*. 2010;50:271-276.
- [55] Gratwohl A, Baldomero H, Aljurf M, et al. Hematopoietic stem cell transplantation: a global perspective. *JAMA*. 2010;303:1617-1624.
- [56] Nivison-Smith I, Bardy P, Dodds AJ, et al. A Review of Hematopoietic Cell Transplantation in Australia and New Zealand, 2005 to 2013. *Biol Blood Marrow Transplant*. 2016;22:284-291.
- [57] Bader P, Niethammer D, Willasch A, Kreyenberg H, Klingebiel T. How and when should we monitor chimerism after allogeneic stem cell transplantation? *Bone Marrow Transplant*. 2005;35:107-119.
- [58] Baron F, Sandmaier BM. Chimerism and outcomes after allogeneic hematopoietic cell transplantation following nonmyeloablative conditioning. *Leukemia*. 2006;20:1690-1700.
- [59] Spitzer TR, McAfee S, Sackstein R, et al. Intentional induction of mixed chimerism and achievement of antitumor responses after nonmyeloablative conditioning therapy and HLA-matched donor bone marrow transplantation for refractory hematologic malignancies. *Biol Blood Marrow Transplant*. 2000;6:309-320.
- [60] Mattsson J, Uzunel M, Remberger M, Ringden O. T cell mixed chimerism is significantly correlated to a decreased risk of acute graft-versus-host disease after allogeneic stem cell transplantation. *Transplantation*. 2001;71:433-439.
- [61] Baron F, Baker JE, Storb R, et al. Kinetics of engraftment in patients with hematologic malignancies given allogeneic hematopoietic cell transplantation after nonmyeloablative conditioning. *Blood*. 2004;104:2254-2262.
- [62] Huss R, Deeg HJ, Gooley T, et al. Effect of mixed chimerism on graft-versus-host disease, disease recurrence and survival after HLA-identical marrow transplantation for aplastic anemia or chronic myelogenous leukemia. *Bone Marrow Transplant*. 1996;18:767-776.
- [63] Zetterquist H, Mattsson J, Uzunel M, et al. Mixed chimerism in the B cell lineage is a rapid and sensitive indicator of minimal residual disease in bone marrow transplant recipients with pre-B cell acute lymphoblastic leukemia. *Bone Marrow Transplant*. 2000;25:843-851.

- [64] Orsini E, Alyea EP, Chillemi A, et al. Conversion to full donor chimerism following donor lymphocyte infusion is associated with disease response in patients with multiple myeloma. *Biol Blood Marrow Transplant*. 2000;6:375-386.
- [65] Kolb HJ, Schattenberg A, Goldman JM, et al. Graft-versus-leukemia effect of donor lymphocyte transfusions in marrow grafted patients. *Blood*. 1995;86:2041-2050.
- [66] Dey BR, McAfee S, Colby C, et al. Impact of prophylactic donor leukocyte infusions on mixed chimerism, graft-versus-host disease, and antitumor response in patients with advanced hematologic malignancies treated with nonmyeloablative conditioning and allogeneic bone marrow transplantation. *Biol Blood Marrow Transplant*. 2003;9:320-329.
- [67] Karasu GT, Yesilipek MA, Karauzum SB, et al. The value of donor lymphocyte infusions in thalassemia patients at imminent risk of graft rejection following stem cell transplantation. *Pediatr Blood Cancer*. 2012;58:453-458.
- [68] Owen R. Immunogenetic Consequences of Vascular Anastomoses between Bovine Twins. *Science*. 1945;102:400.
- [69] Scandling JD, Busque S, Dejbakhsh-Jones S, et al. Tolerance and chimerism after renal and hematopoietic-cell transplantation. *The New England journal of medicine*. 2008;358:362-368.
- [70] Kawai T, Cosimi AB, Spitzer TR, et al. HLA-mismatched renal transplantation without maintenance immunosuppression. *The New England journal of medicine*. 2008;358:353-361.
- [71] Sachs DH. Mixed chimerism as an approach to transplantation tolerance. *Clin Immunol*. 2000;95:S63-68.
- [72] Sykes M. Mixed chimerism and transplant tolerance. *Immunity*. 2001;14:417-424.
- [73] Sachs DH, Sykes M, Kawai T, Cosimi AB. Immuno-intervention for the induction of transplantation tolerance through mixed chimerism. *Semin Immunol*. 2011;23:165-173.
- [74] Storb R, Yu C, Sandmaier BM, et al. Mixed hematopoietic chimerism after marrow allografts. Transplantation in the ambulatory care setting. *Annals of the New York Academy of Sciences*. 1999;872:372-375; discussion 375-376.
- [75] Scandling JD, Busque S, Lowsky R, et al. Macrochimerism and clinical transplant tolerance. *Hum Immunol*. 2018;79:266-271.
- [76] Mold JE, Michaelsson J, Burt TD, et al. Maternal alloantigens promote the development of tolerogenic fetal regulatory T cells in utero. *Science*. 2008;322:1562-1565.

- [77] Loubiere LS, Lambert NC, Flinn LJ, et al. Maternal microchimerism in healthy adults in lymphocytes, monocyte/macrophages and NK cells. *Lab Invest*. 2006;86:1185-1192.
- [78] Sunku CC, Gadi VK, de Laval de Lacoste B, Guthrie KA, Nelson JL. Maternal and fetal microchimerism in granulocytes. *Chimerism*. 2010;1:11-14.
- [79] Burlingham WJ, Grailer AP, Heisey DM, et al. The effect of tolerance to noninherited maternal HLA antigens on the survival of renal transplants from sibling donors. *The New England journal of medicine*. 1998;339:1657-1664.
- [80] van den Boogaardt DE, van Rood JJ, Roelen DL, Claas FH. The influence of inherited and noninherited parental antigens on outcome after transplantation. *Transpl Int*. 2006;19:360-371.
- [81] Ichinohe T, Uchiyama T, Shimazaki C, et al. Feasibility of HLA-haploidentical hematopoietic stem cell transplantation between noninherited maternal antigen (NIMA)-mismatched family members linked with long-term fetomaternal microchimerism. *Blood*. 2004;104:3821-3828.
- [82] Opiela SJ, Adkins B. The pendulum swings: Tolerance versus priming to NIMA. *Chimerism*. 2010;1:36-38.
- [83] Honger G, Fornaro I, Granado C, Tiercy JM, Hosli I, Schaub S. Frequency and determinants of pregnancy-induced child-specific sensitization. *Am J Transplant*. 2013;13:746-753.
- [84] Tajik N, Singal D, Pourmand G, et al. Prospective study of microchimerism in renal allograft recipients: association between HLA-DR matching, microchimerism and acute rejection. *Clin Transplant*. 2001;15:192-198.
- [85] Elwood ET, Larsen CP, Maurer DH, et al. Microchimerism and rejection in clinical transplantation. *Lancet*. 1997;349:1358-1360.
- [86] Starzl TE, Demetris AJ, Murase N, Ildstad S, Ricordi C, Trucco M. Cell migration, chimerism, and graft acceptance. *Lancet*. 1992;339:1579-1582.
- [87] Suberbielle C, Caillat-Zucman S, Legendre C, et al. Peripheral microchimerism in long-term cadaveric-kidney allograft recipients. *Lancet*. 1994;343:1468-1469.
- [88] Lagaaij EL, Cramer-Knijnenburg GF, van Kemenade FJ, van Es LA, Bruijn JA, van Krieken JH. Endothelial cell chimerism after renal transplantation and vascular rejection. *Lancet*. 2001;357:33-37.

- [89] Mengel M, Jonigk D, Marwedel M, et al. Tubular chimerism occurs regularly in renal allografts and is not correlated to outcome. *Journal of the American Society of Nephrology : JASN*. 2004;15:978-986.
- [90] Peters DL, Pretorius PJ. Origin, translocation and destination of extracellular occurring DNA--a new paradigm in genetic behaviour. *Clinica chimica acta; international journal of clinical chemistry*. 2011;412:806-811.
- [91] Jiang P, Chan CW, Chan KC, et al. Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. *Proceedings of the National Academy of Sciences of the United States of America*. 2015;112:E1317-1325.
- [92] Burnham P, Kim MS, Agbor-Enoh S, et al. Single-stranded DNA library preparation uncovers the origin and diversity of ultrashort cell-free DNA in plasma. *Sci Rep*. 2016;6:27859.
- [93] Burnham P, Khush K, De Vlaminck I. Myriad Applications of Circulating Cell-Free DNA in Precision Organ Transplant Monitoring. *Ann Am Thorac Soc*. 2017;14:S237-S241.
- [94] Jahr S, Hentze H, Englisch S, et al. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer research*. 2001;61:1659-1665.
- [95] van der Vaart M, Pretorius PJ. Circulating DNA. Its origin and fluctuation. *Annals of the New York Academy of Sciences*. 2008;1137:18-26.
- [96] Bronkhorst AJ, Wentzel JF, Aucamp J, van Dyk E, du Plessis L, Pretorius PJ. Characterization of the cell-free DNA released by cultured cancer cells. *Biochim Biophys Acta*. 2016;1863:157-165.
- [97] Stroun M, Lyautey J, Lederrey C, Olson-Sand A, Anker P. About the possible origin and mechanism of circulating DNA apoptosis and active DNA release. *Clinica chimica acta; international journal of clinical chemistry*. 2001;313:139-142.
- [98] Lo YM, Chan KC, Sun H, et al. Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. *Sci Transl Med*. 2010;2:61ra91.
- [99] Nagata S, Nagase H, Kawane K, Mukae N, Fukuyama H. Degradation of chromosomal DNA during apoptosis. *Cell death and differentiation*. 2003;10:108-116.
- [100] Mouliere F, Robert B, Arnau Peyrotte E, et al. High fragmentation characterizes tumour-derived circulating DNA. *PloS one*. 2011;6:e23418.

- [101] Lui YY, Chik KW, Chiu RW, Ho CY, Lam CW, Lo YM. Predominant hematopoietic origin of cell-free DNA in plasma and serum after sex-mismatched bone marrow transplantation. *Clin Chem*. 2002;48:421-427.
- [102] Zheng YW, Chan KC, Sun H, et al. Nonhematopoietically derived DNA is shorter than hematopoietically derived DNA in plasma: a transplantation model. *Clin Chem*. 2012;58:549-558.
- [103] Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. *Cell*. 2016;164:57-68.
- [104] Haghiac M, Vora NL, Basu S, et al. Increased death of adipose cells, a path to release cell-free DNA into systemic circulation of obese women. *Obesity*. 2012;20:2213-2219.
- [105] Lapaire O, Volgmann T, Grill S, et al. Significant correlation between maternal body mass index at delivery and in the second trimester, and second trimester circulating total cell-free DNA levels. *Reproductive sciences*. 2009;16:274-279.
- [106] Yu SC, Lee SW, Jiang P, et al. High-resolution profiling of fetal DNA clearance from maternal plasma by massively parallel sequencing. *Clin Chem*. 2013;59:1228-1237.
- [107] Tsui NB, Jiang P, Chow KC, et al. High resolution size analysis of fetal DNA in the urine of pregnant women by paired-end massively parallel sequencing. *PloS one*. 2012;7:e48319.
- [108] Lo YM, Zhang J, Leung TN, Lau TK, Chang AM, Hjelm NM. Rapid clearance of fetal DNA from maternal plasma. *American journal of human genetics*. 1999;64:218-224.
- [109] Beiter T, Fragasso A, Hudemann J, Niess AM, Simon P. Short-term treadmill running as a model for studying cell-free DNA kinetics in vivo. *Clin Chem*. 2011;57:633-636.
- [110] Chused TM, Steinberg AD, Talal N. The clearance and localization of nucleic acids by New Zealand and normal mice. *Clin Exp Immunol*. 1972;12:465-476.
- [111] Korabecna M, Opatrna S, Wirth J, et al. Cell-free plasma DNA during peritoneal dialysis and hemodialysis and in patients with chronic kidney disease. *Annals of the New York Academy of Sciences*. 2008;1137:296-301.
- [112] Jylhava J, Lyytikainen LP, Kahonen M, et al. A genome-wide association study identifies UGT1A1 as a regulator of serum cell-free DNA in young adults: The Cardiovascular Risk in Young Finns Study. *PloS one*. 2012;7:e35426.

- [113] Nakamura A, Nakajima M, Yamanaka H, Fujiwara R, Yokoi T. Expression of UGT1A and UGT2B mRNA in human normal tissues and various cell lines. *Drug Metab Dispos.* 2008;36:1461-1464.
- [114] Breitbach S, Sterzing B, Magallanes C, Tug S, Simon P. Direct measurement of cell-free DNA from serially collected capillary plasma during incremental exercise. *Journal of applied physiology.* 2014.
- [115] Urato AC, Peter I, Canick J, et al. Smoking in pregnancy is associated with increased total maternal serum cell-free DNA levels. *Prenatal diagnosis.* 2008;28:186-190.
- [116] Suzuki N, Kamataki A, Yamaki J, Homma Y. Characterization of circulating DNA in healthy human plasma. *Clinica chimica acta; international journal of clinical chemistry.* 2008;387:55-58.
- [117] Anker P, Lyautey J, Lederrey C, Stroun M. Circulating nucleic acids in plasma or serum. *Clinica chimica acta; international journal of clinical chemistry.* 2001;313:143-146.
- [118] Ziegler A, Zangemeister-Wittke U, Stahel RA. Circulating DNA: a new diagnostic gold mine? *Cancer treatment reviews.* 2002;28:255-271.
- [119] Rhodes A, Wort SJ, Thomas H, Collinson P, Bennett ED. Plasma DNA concentration as a predictor of mortality and sepsis in critically ill patients. *Critical care (London, England).* 2006;10:R60.
- [120] Oellerich M, Shipkova M, Asendorf T, et al. Absolute Quantification of Donor-Derived Cell-Free DNA as a Marker of Rejection and Graft Injury in Kidney Transplantation - Results From a Prospective Observational Study. *Am J Transplant.* 2019.
- [121] Sparks AB, Wang ET, Struble CA, et al. Selective analysis of cell-free DNA in maternal blood for evaluation of fetal trisomy. *Prenatal diagnosis.* 2012;32:3-9.
- [122] Taglauer ES, Wilkins-Haug L, Bianchi DW. Review: cell-free fetal DNA in the maternal circulation as an indication of placental health and disease. *Placenta.* 2014;35 Suppl:S64-68.
- [123] Anker P, Mulcahy H, Chen XQ, Stroun M. Detection of circulating tumour DNA in the blood (plasma/serum) of cancer patients. *Cancer metastasis reviews.* 1999;18:65-73.
- [124] Diaz LA, Jr., Bardelli A. Liquid biopsies: genotyping circulating tumor DNA. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2014;32:579-586.

- [125] Duque-Afonso J, Waterhouse M, Pfeifer D, et al. Cell-free DNA characteristics and chimerism analysis in patients after allogeneic cell transplantation. *Clin Biochem*. 2017.
- [126] Aljurf M, Abalkhail H, Alseraihy A, et al. Chimerism Analysis of Cell-Free DNA in Patients Treated with Hematopoietic Stem Cell Transplantation May Predict Early Relapse in Patients with Hematologic Malignancies. *Biotechnol Res Int*. 2016;2016:8589270.
- [127] Lo YM, Tein MS, Lau TK, et al. Quantitative analysis of fetal DNA in maternal plasma and serum: implications for noninvasive prenatal diagnosis. *American journal of human genetics*. 1998;62:768-775.
- [128] Martins PN, Mashreghi MF, Reutzel-Selke A, et al. Quantification of donor-derived DNA in serum: a new approach of acute rejection diagnosis in a rat kidney transplantation model. *Transplant Proc*. 2005;37:87-88.
- [129] Fu YW, Wang WG, Zhou HL, Cai L. Presence of donor-and-recipient-derived DNA microchimerism in the cell-free blood samples of renal transplantation recipients associates with the acceptance of transplanted kidneys. *Asian J Androl*. 2006;8:477-482.
- [130] Garcia Moreira V, Prieto Garcia B, Baltar Martin JM, Ortega Suarez F, Alvarez FV. Cell-free DNA as a noninvasive acute rejection marker in renal transplantation. *Clin Chem*. 2009;55:1958-1966.
- [131] Snyder TM, Khush KK, Valantine HA, Quake SR. Universal noninvasive detection of solid organ transplant rejection. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108:6229-6234.
- [132] Macher HC, Suarez-Artacho G, Guerrero JM, et al. Monitoring of transplanted liver health by quantification of organ-specific genomic marker in circulating DNA from receptor. *PloS one*. 2014;9:e113987.
- [133] Zhong XY, Hahn D, Troeger C, et al. Cell-free DNA in urine: a marker for kidney graft rejection, but not for prenatal diagnosis? *Annals of the New York Academy of Sciences*. 2001;945:250-257.
- [134] Zhang Z, Ohkohchi N, Sakurada M, et al. Analysis of urinary donor-derived DNA in renal transplant recipients with acute rejection. *Clin Transplant*. 2002;16 Suppl 8:45-50.
- [135] Li Y, Hahn D, Zhong XY, Thomson PD, Holzgreve W, Hahn S. Detection of donor-specific DNA polymorphisms in the urine of renal transplant recipients. *Clin Chem*. 2003;49:655-658.

- [136] Zhang Z, Ohkohchi N, Okazaki H, Guo Y. Use of PCR and PCR-SSP for detection of urinary donor-origin DNA in renal transplant recipients with acute rejection. *Chin Med J (Engl)*. 2003;116:191-194.
- [137] Sigdel TK, Vitalone MJ, Tran TQ, et al. A rapid noninvasive assay for the detection of renal transplant injury. *Transplantation*. 2013;96:97-101.
- [138] Rutkowska J, Interewicz B, Rydzewski A, et al. Donor DNA is detected in recipient blood for years after kidney transplantation using sensitive forensic medicine methods. *Ann Transplant*. 2007;12:12-14.
- [139] Hidestrand M, Tomita-Mitchell A, Hidestrand PM, et al. Highly Sensitive Non-Invasive Cardiac Transplant Rejection Monitoring Using Targeted Quantification of Donor Specific Cell Free DNA. *Journal of the American College of Cardiology*. 2013.
- [140] Lee H, Park YM, We YM, et al. Evaluation of Digital PCR as a Technique for Monitoring Acute Rejection in Kidney Transplantation. *Genomics Inform*. 2017;15:2-10.
- [141] Beck J, Bierau S, Balzer S, et al. Digital droplet PCR for rapid quantification of donor DNA in the circulation of transplant recipients as a potential universal biomarker of graft injury. *Clin Chem*. 2013;59:1732-1741.
- [142] Oellerich M, Schutz E, Kanzow P, et al. Use of graft-derived cell-free DNA as an organ integrity biomarker to reexamine effective tacrolimus trough concentrations after liver transplantation. *Therapeutic drug monitoring*. 2014;36:136-140.
- [143] Schütz E, Fischer A, Beck J, et al. Graft-derived cell-free DNA, a noninvasive early rejection and graft damage marker in liver transplantation: A prospective, observational, multicenter cohort study. *PLoS medicine*. 2017;14.
- [144] Zou J, Duffy B, Slade M, et al. Rapid detection of donor cell free DNA in lung transplant recipients with rejections using donor-recipient HLA mismatch. *Human Immunology*. 2017;78:342-349.
- [145] Bruno DL, Ganesamoorthy D, Thorne NP, et al. Use of copy number deletion polymorphisms to assess DNA chimerism. *Clin Chem*. 2014;60:1105-1114.
- [146] Whitlam JB, Ling L, Skene A, et al. Diagnostic application of kidney allograft-derived absolute cell-free DNA levels during transplant dysfunction. *Am J Transplant*. 2019;19:1037-1049.

- [147] Goh SK, Muralidharan V, Christophi C, Do H, Dobrovic A. Probe-Free Digital PCR Quantitative Methodology to Measure Donor-Specific Cell-Free DNA after Solid-Organ Transplantation. *Clin Chem*. 2017;63:742-750.
- [148] Macher HC, Garcia-Fernandez N, Adsuar-Gomez A, et al. Donor-specific circulating cell free DNA as a noninvasive biomarker of graft injury in heart transplantation. *Clinica chimica acta; international journal of clinical chemistry*. 2019.
- [149] De Vlaminck I, Valentine HA, Snyder TM, et al. Circulating cell-free DNA enables noninvasive diagnosis of heart transplant rejection. *Sci Transl Med*. 2014;6:241ra277.
- [150] De Vlaminck I, Martin L, Kertesz M, et al. Noninvasive monitoring of infection and rejection after lung transplantation. *Proceedings of the National Academy of Sciences of the United States of America*. 2015.
- [151] Agbor-Enoh S, Wang Y, Tunc I, et al. Donor-derived cell-free DNA predicts allograft failure and mortality after lung transplantation. *EBioMedicine*. 2019;40:541-553.
- [152] Grskovic M, Hiller DJ, Eubank LA, et al. Validation of a Clinical-Grade Assay to Measure Donor-Derived Cell-Free DNA in Solid Organ Transplant Recipients. *J Mol Diagn*. 2016;18:890-902.
- [153] Bloom RD, Bromberg JS, Poggio ED, et al. Cell-Free DNA and Active Rejection in Kidney Allografts. *Journal of the American Society of Nephrology : JASN*. 2017;28:2221-2232.
- [154] Jordan SC, Bunnapradist S, Bromberg JS, et al. Donor-derived Cell-free DNA Identifies Antibody-mediated Rejection in Donor Specific Antibody Positive Kidney Transplant Recipients. *Transplant Direct*. 2018;4:e379.
- [155] Huang E, Sethi S, Peng A, et al. Early clinical experience using donor-derived cell-free DNA to detect rejection in kidney transplant recipients. *Am J Transplant*. 2019;19:1663-1670.
- [156] Khush K, Zarafshar S. Molecular Diagnostic Testing in Cardiac Transplantation. *Curr Cardiol Rep*. 2017;19:118.
- [157] Altug Y, Liang N, Ram R, et al. Analytical Validation of a Single-Nucleotide Polymorphism-Based Donor-Derived Cell-Free DNA Assay for Detecting Rejection in Kidney Transplant Patients. *Transplantation*. 2019.
- [158] Sigdel TK, Archila FA, Constantin T, et al. Optimizing Detection of Kidney Transplant Injury by Assessment of Donor-Derived Cell-Free DNA via Massively Multiplex PCR. *J Clin Med*. 2018;8.

- [159] Gielis EM, Beirnaert C, Dendooven A, et al. Plasma donor-derived cell-free DNA kinetics after kidney transplantation using a single tube multiplex PCR assay. *PloS one*. 2018;13:e0208207.
- [160] Gielis EM, Ledeganck KJ, Dendooven A, et al. The use of plasma donor-derived, cell-free DNA to monitor acute rejection after kidney transplantation. *Nephrol Dial Transplant*. 2019.
- [161] Khush KK, Patel J, Pinney S, et al. Noninvasive detection of graft injury after heart transplant using donor-derived cell-free DNA: A prospective multicenter study. *Am J Transplant*. 2019.
- [162] Yao W, Mei C, Nan X, Hui L. Evaluation and comparison of in vitro degradation kinetics of DNA in serum, urine and saliva: A qualitative study. *Gene*. 2016;590:142-148.
- [163] Adamek M, Opelz G, Klein K, Morath C, Tran TH. A fast and simple method for detecting and quantifying donor-derived cell-free DNA in sera of solid organ transplant recipients as a biomarker for graft function. *Clin Chem Lab Med*. 2016;54:1147-1155.
- [164] Gordon PM, Khan A, Sajid U, et al. An Algorithm Measuring Donor Cell-Free DNA in Plasma of Cellular and Solid Organ Transplant Recipients That Does Not Require Donor or Recipient Genotyping. *Front Cardiovasc Med*. 2016;3:33.
- [165] Gala-Lopez BL, Neiman D, Kin T, et al. Beta Cell Death by Cell-free DNA and Outcome After Clinical Islet Transplantation. *Transplantation*. 2018;102:978-985.
- [166] Shen J, Zhou Y, Chen Y, et al. Dynamics of early post-operative plasma ddcfDNA levels in kidney transplantation: a single-center pilot study. *Transpl Int*. 2019;32:184-192.
- [167] Li Y, Hahn D, Wenzel F, Holzgreve W, Hahn S. Detection of SNPs in the plasma of pregnant women and in the urine of kidney transplant recipients by mass spectrometry. *Annals of the New York Academy of Sciences*. 2006;1075:144-147.
- [168] Wolfe RA, Ashby VB, Milford EL, et al. Comparison of mortality in all patients on dialysis, patients on dialysis awaiting transplantation, and recipients of a first cadaveric transplant. *The New England journal of medicine*. 1999;341:1725-1730.
- [169] McDonald SP. Australia and New Zealand Dialysis and Transplant Registry: Thirty Fifth Annual Report. 2012.
- [170] Meier-Kriesche HU, Schold JD, Kaplan B. Long-term renal allograft survival: have we made significant progress or is it time to rethink our analytic and therapeutic strategies? *Am J Transplant*. 2004;4:1289-1295.

- [171] Morath C, Ritz E, Zeier M. Protocol biopsy: what is the rationale and what is the evidence? *Nephrol Dial Transplant*. 2003;18:644-647.
- [172] Solez K, Colvin RB, Racusen LC, et al. Banff 07 classification of renal allograft pathology: updates and future directions. *Am J Transplant*. 2008;8:753-760.
- [173] Roufosse C, Simmonds N, Clahsen-van Groningen M, et al. A 2018 Reference Guide to the Banff Classification of Renal Allograft Pathology. *Transplantation*. 2018;102:1795-1814.
- [174] Solez K, Racusen LC. The Banff classification revisited. *Kidney international*. 2013;83:201-206.
- [175] Racusen LC, Solez K, Colvin RB, et al. The Banff 97 working classification of renal allograft pathology. *Kidney international*. 1999;55:713-723.
- [176] Bates WD, Davies DR, Welsh K, Gray DW, Fuggle SV, Morris PJ. An evaluation of the Banff classification of early renal allograft biopsies and correlation with outcome. *Nephrol Dial Transplant*. 1999;14:2364-2369.
- [177] Halloran PF. T cell-mediated rejection of kidney transplants: a personal viewpoint. *Am J Transplant*. 2010;10:1126-1134.
- [178] Halloran PF, Miller LW, Urmson J, et al. IFN-gamma alters the pathology of graft rejection: protection from early necrosis. *J Immunol*. 2001;166:7072-7081.
- [179] August C, Schmid KW, Dietl KH, Heidenreich S. Prognostic value of lymphocyte apoptosis in acute rejection of renal allografts. *Transplantation*. 1999;67:581-585.
- [180] Noronha IL, Oliveira SG, Tavares TS, et al. Apoptosis in kidney and pancreas allograft biopsies. *Transplantation*. 2005;79:1231-1235.
- [181] Famulski KS, Kayser D, Einecke G, et al. Alternative macrophage activation-associated transcripts in T-cell-mediated rejection of mouse kidney allografts. *Am J Transplant*. 2010;10:490-497.
- [182] Famulski KS, Broderick G, Einecke G, et al. Transcriptome analysis reveals heterogeneity in the injury response of kidney transplants. *Am J Transplant*. 2007;7:2483-2495.
- [183] Jabs WJ, Sedlmeyer A, Ramassar V, et al. Heterogeneity in the evolution and mechanisms of the lesions of kidney allograft rejection in mice. *Am J Transplant*. 2003;3:1501-1509.

- [184] Einecke G, Fairhead T, Hidalgo LG, et al. Tubulitis and epithelial cell alterations in mouse kidney transplant rejection are independent of CD103, perforin or granzymes A/B. *Am J Transplant.* 2006;6:2109-2120.
- [185] 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. *Am J Transplant.* 2007;7:1121-1130.
- [186] Wiebe C, Gibson IW, Blydt-Hansen TD, et al. Evolution and clinical pathologic correlations of de novo donor-specific HLA antibody post kidney transplant. *Am J Transplant.* 2012;12:1157-1167.
- [187] Racusen LC, Colvin RB, Solez K, et al. Antibody-mediated rejection criteria - an addition to the Banff 97 classification of renal allograft rejection. *Am J Transplant.* 2003;3:708-714.
- [188] Haas M, Sis B, Racusen LC, et al. Banff 2013 meeting report: inclusion of c4d-negative antibody-mediated rejection and antibody-associated arterial lesions. *Am J Transplant.* 2014;14:272-283.
- [189] Halloran PF, Schlaut J, Solez K, Srinivasa NS. The significance of the anti-class I response. II. Clinical and pathologic features of renal transplants with anti-class I-like antibody. *Transplantation.* 1992;53:550-555.
- [190] Erlich HA, Opelz G, Hansen J. HLA DNA typing and transplantation. *Immunity.* 2001;14:347-356.
- [191] Sumitran-Holgersson S, Wilczek HE, Holgersson J, Soderstrom K. Identification of the nonclassical HLA molecules, mica, as targets for humoral immunity associated with irreversible rejection of kidney allografts. *Transplantation.* 2002;74:268-277.
- [192] Austen WG, Jr., Zhang M, Chan R, et al. Murine hindlimb reperfusion injury can be initiated by a self-reactive monoclonal IgM. *Surgery.* 2004;136:401-406.
- [193] Dragun D, Muller DN, Brasen JH, et al. Angiotensin II type 1-receptor activating antibodies in renal-allograft rejection. *The New England journal of medicine.* 2005;352:558-569.
- [194] Colvin RB, Smith RN. Antibody-mediated organ-allograft rejection. *Nat Rev Immunol.* 2005;5:807-817.

- [195] Nakashima S, Qian Z, Rahimi S, Wasowska BA, Baldwin WM, 3rd. Membrane attack complex contributes to destruction of vascular integrity in acute lung allograft rejection. *J Immunol.* 2002;169:4620-4627.
- [196] Biglarnia AR, Nilsson B, Nilsson T, et al. Prompt reversal of a severe complement activation by eculizumab in a patient undergoing intentional ABO-incompatible pancreas and kidney transplantation. *Transpl Int.* 2011;24:e61-66.
- [197] Locke JE, Magro CM, Singer AL, et al. The use of antibody to complement protein C5 for salvage treatment of severe antibody-mediated rejection. *Am J Transplant.* 2009;9:231-235.
- [198] Stegall MD, Diwan T, Raghavaiah S, et al. Terminal complement inhibition decreases antibody-mediated rejection in sensitized renal transplant recipients. *Am J Transplant.* 2011;11:2405-2413.
- [199] Hidalgo LG, Sis B, Sellares J, et al. NK cell transcripts and NK cells in kidney biopsies from patients with donor-specific antibodies: evidence for NK cell involvement in antibody-mediated rejection. *Am J Transplant.* 2010;10:1812-1822.
- [200] Meier-Kriesche HU, Schold JD, Srinivas TR, Kaplan B. Lack of improvement in renal allograft survival despite a marked decrease in acute rejection rates over the most recent era. *Am J Transplant.* 2004;4:378-383.
- [201] Djamali A, Kaufman DB, Ellis TM, Zhong W, Matas A, Samaniego M. Diagnosis and management of antibody-mediated rejection: current status and novel approaches. *Am J Transplant.* 2014;14:255-271.
- [202] Sellares J, de Freitas DG, Mengel M, et al. Understanding the causes of kidney transplant failure: the dominant role of antibody-mediated rejection and nonadherence. *Am J Transplant.* 2012;12:388-399.
- [203] El-Zoghby ZM, Stegall MD, Lager DJ, et al. Identifying specific causes of kidney allograft loss. *Am J Transplant.* 2009;9:527-535.
- [204] Jackson JA, Kim EJ, Begley B, et al. Urinary chemokines CXCL9 and CXCL10 are noninvasive markers of renal allograft rejection and BK viral infection. *Am J Transplant.* 2011;11:2228-2234.
- [205] Buchmann TN, Wolff T, Bachmann A, et al. Repeat true surveillance biopsies in kidney transplantation. *Transplantation.* 2012;93:908-913.

- [206] Dorje C, Mjoen G, Strom EH, et al. One-year protocol biopsies from ABO-incompatible renal allografts compared with a matched cohort of ABO-compatible allografts. *Clin Transplant*. 2015;29:268-276.
- [207] Wilkinson A. Protocol transplant biopsies: are they really needed? *Clin J Am Soc Nephrol*. 2006;1:130-137.
- [208] Thierry A, Thervet E, Vuiblet V, et al. Long-term impact of subclinical inflammation diagnosed by protocol biopsy one year after renal transplantation. *Am J Transplant*. 2011;11:2153-2161.
- [209] Nankivell BJ, Chapman JR. The significance of subclinical rejection and the value of protocol biopsies. *Am J Transplant*. 2006;6:2006-2012.
- [210] Rush D, Nickerson P, Gough J, et al. Beneficial effects of treatment of early subclinical rejection: a randomized study. *Journal of the American Society of Nephrology : JASN*. 1998;9:2129-2134.
- [211] Moreso F, Ibernón M, Goma M, et al. Subclinical rejection associated with chronic allograft nephropathy in protocol biopsies as a risk factor for late graft loss. *Am J Transplant*. 2006;6:747-752.
- [212] Korbett SM, Volpini KC, Whittier WL. Percutaneous renal biopsy of native kidneys: a single-center experience of 1,055 biopsies. *American journal of nephrology*. 2014;39:153-162.
- [213] Nicholson ML, Wheatley TJ, Doughman TM, et al. A prospective randomized trial of three different sizes of core-cutting needle for renal transplant biopsy. *Kidney Int*. 2000;58:390-395.
- [214] Yablon Z, Recupero P, McKenna J, Vella J, Parker MG. Kidney allograft biopsy: timing to complications. *Clin Nephrol*. 2010;74:39-45.
- [215] Furness PN, Philpott CM, Chorbajian MT, et al. Protocol biopsy of the stable renal transplant: a multicenter study of methods and complication rates. *Transplantation*. 2003;76:969-973.
- [216] Furness PN, Taub N, Assmann KJ, et al. International variation in histologic grading is large, and persistent feedback does not improve reproducibility. *The American journal of surgical pathology*. 2003;27:805-810.
- [217] Marcussen N, Olsen TS, Benediktsson H, Racusen L, Solez K. Reproducibility of the Banff classification of renal allograft pathology. Inter- and intraobserver variation. *Transplantation*. 1995;60:1083-1089.

- [218] Gibson IW, Gwinner W, Brocker V, et al. Peritubular capillaritis in renal allografts: prevalence, scoring system, reproducibility and clinicopathological correlates. *Am J Transplant.* 2008;8:819-825.
- [219] Furness PN, Taub N, Project CoERTPAPC. International variation in the interpretation of renal transplant biopsies: report of the CERTPAP Project. *Kidney international.* 2001;60:1998-2012.
- [220] Adam B, Randhawa P, Chan S, et al. Banff Initiative for Quality Assurance in Transplantation (BIFQUIT): reproducibility of polyomavirus immunohistochemistry in kidney allografts. *Am J Transplant.* 2014;14:2137-2147.
- [221] Sis B, Mengel M, Haas M, et al. Banff '09 meeting report: antibody mediated graft deterioration and implementation of Banff working groups. *Am J Transplant.* 2010;10:464-471.
- [222] American Society of Nephrology Renal Research Report. *Journal of the American Society of Nephrology : JASN.* 2005;16:1886-1903.
- [223] Lo DJ, Kaplan B, Kirk AD. Biomarkers for kidney transplant rejection. *Nature reviews Nephrology.* 2014;10:215-225.
- [224] Netto MV, Fonseca BA, Dantas M, Saber LT, Castro MC, Ferraz AS. Granzyme B, FAS-ligand and perforin expression during acute cellular rejection episodes after kidney transplantation: comparison between blood and renal aspirates. *Transplant Proc.* 2002;34:476-478.
- [225] Vasconcellos LM, Schachter AD, Zheng XX, et al. Cytotoxic lymphocyte gene expression in peripheral blood leukocytes correlates with rejecting renal allografts. *Transplantation.* 1998;66:562-566.
- [226] Graziotto R, Del Prete D, Rigotti P, et al. Perforin, Granzyme B, and fas ligand for molecular diagnosis of acute renal-allograft rejection: analyses on serial biopsies suggest methodological issues. *Transplantation.* 2006;81:1125-1132.
- [227] Heidt S, San Segundo D, Shankar S, et al. Peripheral blood sampling for the detection of allograft rejection: biomarker identification and validation. *Transplantation.* 2011;92:1-9.
- [228] Reeve J, Sellares J, Mengel M, et al. Molecular diagnosis of T cell-mediated rejection in human kidney transplant biopsies. *Am J Transplant.* 2013;13:645-655.
- [229] Sellares J, Reeve J, Loupy A, et al. Molecular diagnosis of antibody-mediated rejection in human kidney transplants. *Am J Transplant.* 2013;13:971-983.

- [230] Li L, Khatri P, Sigdel TK, et al. A peripheral blood diagnostic test for acute rejection in renal transplantation. *Am J Transplant*. 2012;12:2710-2718.
- [231] Roedder S, Sigdel T, Salomonis N, et al. The kSORT Assay to Detect Renal Transplant Patients at High Risk for Acute Rejection: Results of the Multicenter AART Study. *PLoS medicine*. 2014;11:e1001759.
- [232] Afaneh C, Muthukumar T, Lubetzky M, et al. Urinary cell levels of mRNA for OX40, OX40L, PD-1, PD-L1, or PD-L2 and acute rejection of human renal allografts. *Transplantation*. 2010;90:1381-1387.
- [233] Li B, Hartono C, Ding R, et al. Noninvasive diagnosis of renal-allograft rejection by measurement of messenger RNA for perforin and granzyme B in urine. *The New England journal of medicine*. 2001;344:947-954.
- [234] Suthanthiran M, Schwartz JE, Ding R, et al. Urinary-cell mRNA profile and acute cellular rejection in kidney allografts. *The New England journal of medicine*. 2013;369:20-31.
- [235] Anglicheau D, Sharma VK, Ding R, et al. MicroRNA expression profiles predictive of human renal allograft status. *Proceedings of the National Academy of Sciences of the United States of America*. 2009;106:5330-5335.
- [236] Ho J, Rush DN, Gibson IW, et al. Early urinary CCL2 is associated with the later development of interstitial fibrosis and tubular atrophy in renal allografts. *Transplantation*. 2010;90:394-400.
- [237] Ho J, Wiebe C, Rush DN, et al. Increased urinary CCL2: Cr ratio at 6 months is associated with late renal allograft loss. *Transplantation*. 2013;95:595-602.
- [238] Hricik DE, Nickerson P, Formica RN, et al. Multicenter validation of urinary CXCL9 as a risk-stratifying biomarker for kidney transplant injury. *Am J Transplant*. 2013;13:2634-2644.
- [239] Lorenzen JM, Volkmann I, Fiedler J, et al. Urinary miR-210 as a mediator of acute T-cell mediated rejection in renal allograft recipients. *Am J Transplant*. 2011;11:2221-2227.
- [240] Nather BJ, Nickel P, Bold G, et al. Modified ELISPOT technique--highly significant inverse correlation of post-Tx donor-reactive IFNgamma-producing cell frequencies with 6 and 12 months graft function in kidney transplant recipients. *Transpl Immunol*. 2006;16:232-237.
- [241] Kim SH, Oh EJ, Kim MJ, et al. Pretransplant donor-specific interferon-gamma ELISPOT assay predicts acute rejection episodes in renal transplant recipients. *Transplant Proc*. 2007;39:3057-3060.

- [242] Sherwood K, Weimer ET. Characteristics, properties, and potential applications of circulating cell-free dna in clinical diagnostics: a focus on transplantation. *J Immunol Methods*. 2018;463:27-38.
- [243] McCarroll SA, Kuruvilla FG, Korn JM, et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet*. 2008;40:1166-1174.
- [244] Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome research*. 2002;12:996-1006.
- [245] Conrad DF, Andrews TD, Carter NP, Hurles ME, Pritchard JK. A high-resolution survey of deletion polymorphism in the human genome. *Nat Genet*. 2006;38:75-81.
- [246] Hsu F, Kent WJ, Clawson H, Kuhn RM, Diekhans M, Haussler D. The UCSC Known Genes. *Bioinformatics*. 2006;22:1036-1046.
- [247] Pruitt KD, Tatusova T, Maglott DR. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res*. 2005;33:D501-504.
- [248] Haeussler M, Gerner M, Bergman CM. Annotating genes and genomes with DNA sequences extracted from biomedical articles. *Bioinformatics*. 2011;27:980-986.
- [249] Sherry ST, Ward MH, Kholodov M, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*. 2001;29:308-311.
- [250] MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res*. 2014;42:D986-992.
- [251] Jurka J. Repbase update: a database and an electronic journal of repetitive elements. *Trends Genet*. 2000;16:418-420.
- [252] Owczarzy R, Tataurov AV, Wu Y, et al. IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Res*. 2008;36:W163-169.
- [253] Bio-Rad. Droplet Digital PCR Applications Guide. Bulletin 6407 Rev A ed; 2014.
- [254] Kent WJ. BLAT—the BLAST-like alignment tool. *Genome research*. 2002;12:656-664.
- [255] Sudmant PH, Huddleston J, Catacchio CR, et al. Evolution and diversity of copy number variation in the great ape lineage. *Genome Research*. 2013;23:1373-1382.

- [256] Coe BP, Witherspoon K, Rosenfeld JA, et al. Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nature genetics*. 2014;46:1063-1071.
- [257] Tired L, Rigat B, Visvikis S, et al. Evidence, from combined segregation and linkage analysis, that a variant of the angiotensin I-converting enzyme (ACE) gene controls plasma ACE levels. *American journal of human genetics*. 1992;51:197-205.
- [258] Mills RE, Walter K, Stewart C, et al. Mapping copy number variation by population-scale genome sequencing. *Nature*. 2011;470:59-65.
- [259] Bhagwat M, Young L, Robison RR. Using BLAT to Find Sequence Similarity in Closely Related Genomes. *Current protocols in bioinformatics / editorial board, Andreas D Baxevanis [et al]*. 2012;0 10:Unit10.18-Unit10.18.
- [260] Cooper DN. Functional intronic polymorphisms: Buried treasure awaiting discovery within our genes. *Hum Genomics*. 2010;4:284-288.
- [261] Beck J, Oellerich M, Schulz U, et al. Donor-Derived Cell-Free DNA Is a Novel Universal Biomarker for Allograft Rejection in Solid Organ Transplantation. *Transplant Proc*. 2015;47:2400-2403.
- [262] Streck. Cell-Free DNA BCT IFU.
[https://www.streck.com/resources/Cell_Stabilization/Cell-Free_DNA_BCT/01_Instructions_\(IFU\)/01_IFU_Cell-Free_DNA_BCT_IFU.pdf](https://www.streck.com/resources/Cell_Stabilization/Cell-Free_DNA_BCT/01_Instructions_(IFU)/01_IFU_Cell-Free_DNA_BCT_IFU.pdf); 2017.
- [263] Streck. Streck Cell-Free DNA BCT® and Cell-Free RNA BCT® Phlebotomy Best Practices. [https://www.streck.com/resources/Cell_Stabilization/Cell-Free_DNA_BCT/01_Instructions_\(IFU\)/05_Streck_Cell-Free_DNA%20BCT_and_Cell-Free_RNA_BCT_Phlebotomy_Best_Practices.pdf](https://www.streck.com/resources/Cell_Stabilization/Cell-Free_DNA_BCT/01_Instructions_(IFU)/05_Streck_Cell-Free_DNA%20BCT_and_Cell-Free_RNA_BCT_Phlebotomy_Best_Practices.pdf).
- [264] QIAGEN. QIAamp Circulating Nucleic Acid Handbook. Third ed.
<https://www.qiagen.com/au/resources/resourcedetail?id=0c4b31ab-f4fb-425f-99bf-10ab9538c061>; 2013.
- [265] Team RC. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
- [266] Botezatu I, Serdyuk O, Potapova G, et al. Genetic analysis of DNA excreted in urine: a new approach for detecting specific genomic DNA sequences from cells dying in an organism. *Clin Chem*. 2000;46:1078-1084.

- [267] Su YH, Wang M, Brenner DE, et al. Human urine contains small, 150 to 250 nucleotide-sized, soluble DNA derived from the circulation and may be useful in the detection of colorectal cancer. *J Mol Diagn*. 2004;6:101-107.
- [268] Olsson M, Kierczak M, Karlsson A, et al. Absolute quantification reveals the stable transmission of a high copy number variant linked to autoinflammatory disease. *BMC Genomics*. 2016;17:299.
- [269] Martins WFS, Subramaniam K, Steen K, et al. Detection and quantitation of copy number variation in the voltage-gated sodium channel gene of the mosquito *Culex quinquefasciatus*. *Sci Rep*. 2017;7:5821.
- [270] Harmala SK, Butcher R, Roberts CH. Copy Number Variation Analysis by Droplet Digital PCR. *Methods in molecular biology*. 2017;1654:135-149.
- [271] Bell AD, Usher CL, McCarroll SA. Analyzing Copy Number Variation with Droplet Digital PCR. *Methods in molecular biology*. 2018;1768:143-160.
- [272] Karlin-Neumann G, Bizouarn F. Entering the Pantheon of 21(st) Century Molecular Biology Tools: A Perspective on Digital PCR. *Methods in molecular biology*. 2018;1768:3-10.
- [273] Bustin SA, Benes V, Garson JA, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem*. 2009;55:611-622.
- [274] Zare F, Dow M, Monteleone N, Hosny A, Nabavi S. An evaluation of copy number variation detection tools for cancer using whole exome sequencing data. *BMC Bioinformatics*. 2017;18:286.
- [275] Zhao M, Wang Q, Wang Q, Jia P, Zhao Z. Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives. *BMC Bioinformatics*. 2013;14 Suppl 11:S1.
- [276] do Nascimento F, Guimaraes KS. Copy Number Variations Detection: Unravelling the Problem in Tangible Aspects. *IEEE/ACM Trans Comput Biol Bioinform*. 2017;14:1237-1250.
- [277] Kim HY, Choi JW, Lee JY, Kong G. Gene-based comparative analysis of tools for estimating copy number alterations using whole-exome sequencing data. *Oncotarget*. 2017;8:27277-27285.
- [278] Huggett JF, Foy CA, Benes V, et al. The digital MIQE guidelines: Minimum Information for Publication of Quantitative Digital PCR Experiments. *Clin Chem*. 2013;59:892-902.

- [279] Hudecova I. Digital PCR analysis of circulating nucleic acids. *Clin Biochem*. 2015.
- [280] Miotke L, Lau BT, Rumma RT, Ji HP. High sensitivity detection and quantitation of DNA copy number and single nucleotide variants with single color droplet digital PCR. *Analytical chemistry*. 2014;86:2618-2624.
- [281] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.
- [282] Armbruster DA, Pry T. Limit of blank, limit of detection and limit of quantitation. *Clin Biochem Rev*. 2008;29 Suppl 1:S49-52.
- [283] Thierry AR, El Messaoudi S, Gahan PB, Anker P, Stroun M. Origins, structures, and functions of circulating DNA in oncology. *Cancer metastasis reviews*. 2016.
- [284] Trypsteen W, Vynck M, De Neve J, et al. ddpcRquant: threshold determination for single channel droplet digital PCR experiments. *Anal Bioanal Chem*. 2015;407:5827-5834.
- [285] Merker JD, Oxnard GR, Compton C, et al. Circulating Tumor DNA Analysis in Patients With Cancer: American Society of Clinical Oncology and College of American Pathologists Joint Review. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2018;36:1631-1641.
- [286] Mauger F, Dulary C, Daviaud C, Deleuze JF, Tost J. Comprehensive evaluation of methods to isolate, quantify, and characterize circulating cell-free DNA from small volumes of plasma. *Anal Bioanal Chem*. 2015.
- [287] Perez-Barrios C, Nieto-Alcolado I, Torrente M, et al. Comparison of methods for circulating cell-free DNA isolation using blood from cancer patients: impact on biomarker testing. *Transl Lung Cancer Res*. 2016;5:665-672.
- [288] Sorber L, Zwaenepoel K, Deschoolmeester V, et al. A Comparison of Cell-Free DNA Isolation Kits: Isolation and Quantification of Cell-Free DNA in Plasma. *J Mol Diagn*. 2017;19:162-168.
- [289] Tug S, Helmig S, Deichmann ER, et al. Exercise-induced increases in cell free DNA in human plasma originate predominantly from cells of the haematopoietic lineage. *Exerc Immunol Rev*. 2015;21:164-173.
- [290] Yaron Y. The implications of non-invasive prenatal testing failures: a review of an under-discussed phenomenon. *Prenatal diagnosis*. 2016;36:391-396.

- [291] Buderer NM. Statistical methodology: I. Incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. *Academic emergency medicine : official journal of the Society for Academic Emergency Medicine*. 1996;3:895-900.
- [292] Haas M. The Revised (2013) Banff Classification for Antibody-Mediated Rejection of Renal Allografts: Update, Difficulties, and Future Considerations. *Am J Transplant*. 2016;16:1352-1357.
- [293] Whitlam JB, Ling L, Swain M, et al. Use of ubiquitous, highly heterozygous copy number variants and digital droplet polymerase chain reaction to monitor chimerism after allogeneic haematopoietic stem cell transplantation. *Exp Hematol*. 2017;49:39-47 e35.
- [294] Perez-Jaume S, Skaltsa K, Pallarès N, Carrasco JL. ThresholdROC: Optimum Threshold Estimation Tools for Continuous Diagnostic Tests in R. *Journal of Statistical Software*. 2017;82:21.
- [295] Harries LW, Wickham CL, Evans JC, Rule SA, Joyner MV, Ellard S. Analysis of haematopoietic chimaerism by quantitative real-time polymerase chain reaction. *Bone Marrow Transplant*. 2005;35:283-290.
- [296] Kroger N, Zagrivnaja M, Schwartz S, et al. Kinetics of plasma-cell chimerism after allogeneic stem cell transplantation by highly sensitive real-time PCR based on sequence polymorphism and its value to quantify minimal residual disease in patients with multiple myeloma. *Exp Hematol*. 2006;34:688-694.
- [297] Jimenez-Velasco A, Barrios M, Roman-Gomez J, et al. Reliable quantification of hematopoietic chimerism after allogeneic transplantation for acute leukemia using amplification by real-time PCR of null alleles and insertion/deletion polymorphisms. *Leukemia*. 2005;19:336-343.
- [298] Odriozola A, Riancho JA, Colorado M, Zarrabeitia MT. Evaluation of the sensitivity of two recently developed STR multiplexes for the analysis of chimerism after haematopoietic stem cell transplantation. *Int J Immunogenet*. 2013;40:88-92.
- [299] Stahl T, Badbaran A, Kroger N, et al. Minimal residual disease diagnostics in patients with acute myeloid leukemia in the post-transplant period: comparison of peripheral blood and bone marrow analysis. *Leuk Lymphoma*. 2010;51:1837-1843.
- [300] George D, Czech J, John B, Yu M, Jennings LJ. Detection and quantification of chimerism by droplet digital PCR. *Chimerism*. 2013;4:102-108.

- [301] Taira C, Matsuda K, Yamaguchi A, et al. Rapid single nucleotide polymorphism based method for hematopoietic chimerism analysis and monitoring using high-speed droplet allele-specific PCR and allele-specific quantitative PCR. *Clinica chimica acta; international journal of clinical chemistry*. 2015;445:101-106.
- [302] Santurtun A, Riancho JA, Arozamena J, Lopez-Duarte M, Zarrabeitia MT. Indel analysis by droplet digital PCR: a sensitive method for DNA mixture detection and chimerism analysis. *International journal of legal medicine*. 2016.
- [303] Stahl T, Rothe C, Bohme MU, Kohl A, Kroger N, Fehse B. Digital PCR Panel for Sensitive Hematopoietic Chimerism Quantification after Allogeneic Stem Cell Transplantation. *International journal of molecular sciences*. 2016;17.
- [304] Zarrei M, MacDonald JR, Merico D, Scherer SW. A copy number variation map of the human genome. *Nature reviews Genetics*. 2015;16:172-183.

APPENDIX A SELECTED UCSC GENOME BROWSER TRACKS

Figure A-1: Numerous common SNPs and repeating elements precluded assay development in the CNV 35 locus.

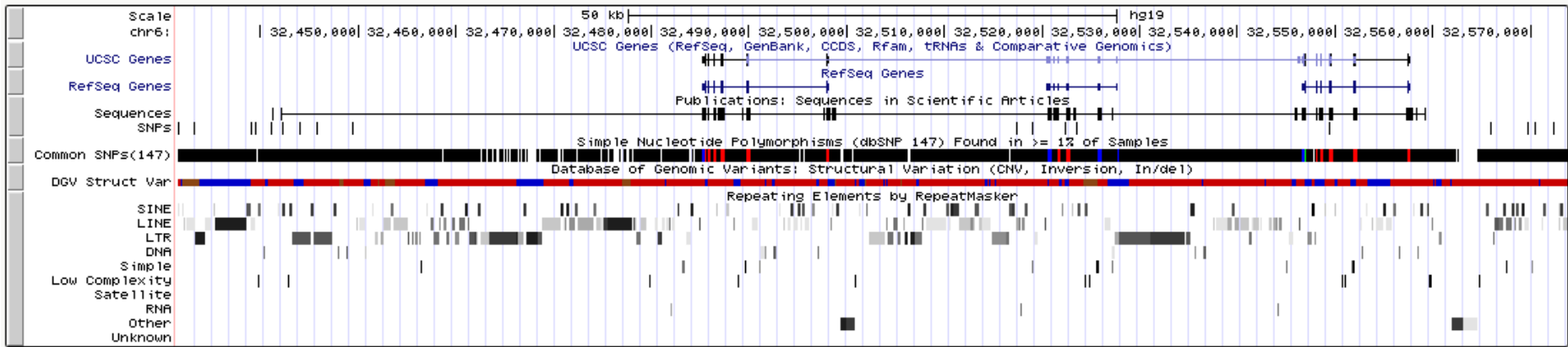


Figure A-2: The CNV 37 locus, demonstrating that the entire CNV 40 locus sits within this locus. No assay could be designed for CNV 40 that wouldn't also sit within CNV 37.

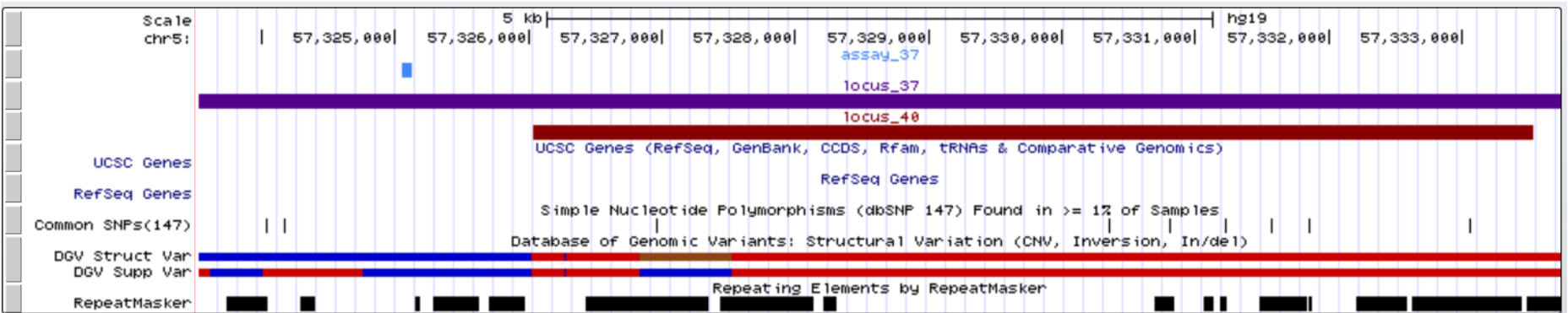


Figure A-3: Amplicon of assay CNV 15A demonstrating the position of SNP rs111271716 within the probe binding region.

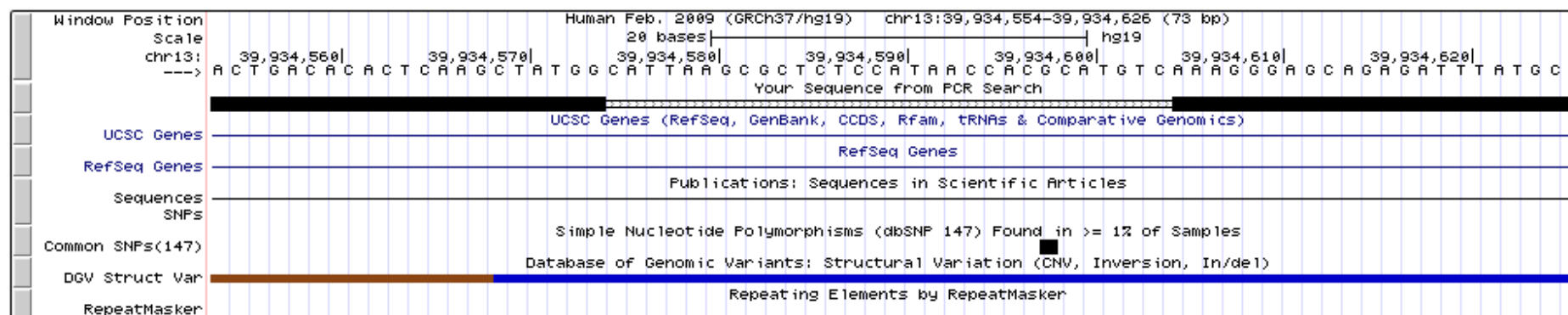


Figure A-4: Overlapping CNV locus 3 (locus_3) and CNV locus 25 (locus_25) demonstrating the position of their assays (assay_3, assay_25) with respect to the overlapping loci. The CNV 25 locus sits entirely within the CNV 3 locus, and as such assay CNV 25A also sits within the CNV 3 locus. The CNV 03B assay does not sit within the CNV 25 locus.

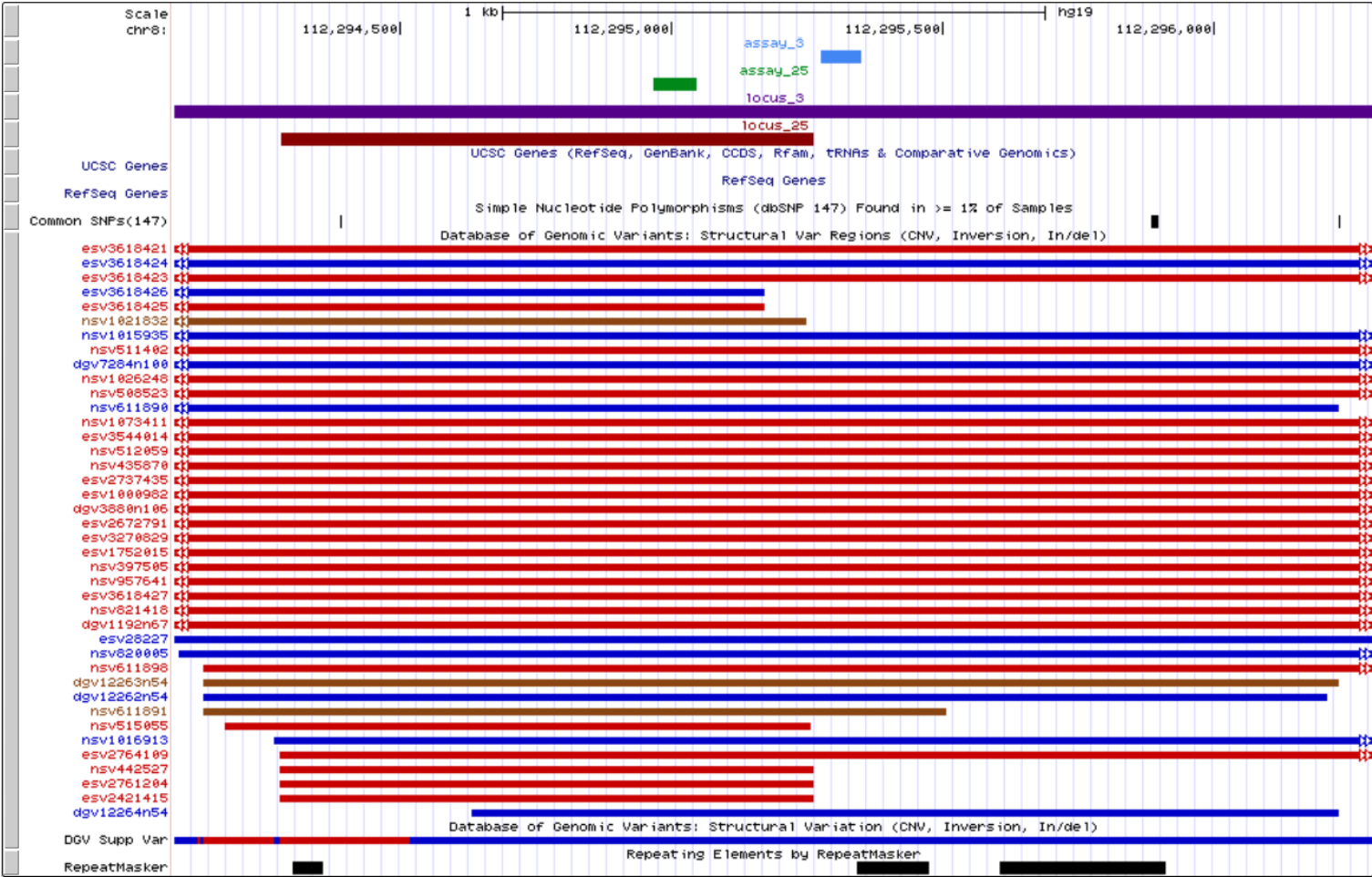


Figure A-5: Overlapping CNV 11 locus (locus_11) and CNV 21 locus (locus_21) demonstrating the position of assay amplicons (assay_11, assay_21) with respect to the overlapping loci. The CNV 21 locus sits entirely within the CNV 11 locus and as such assay CNV 21A also sits within the CNV 11 locus. The

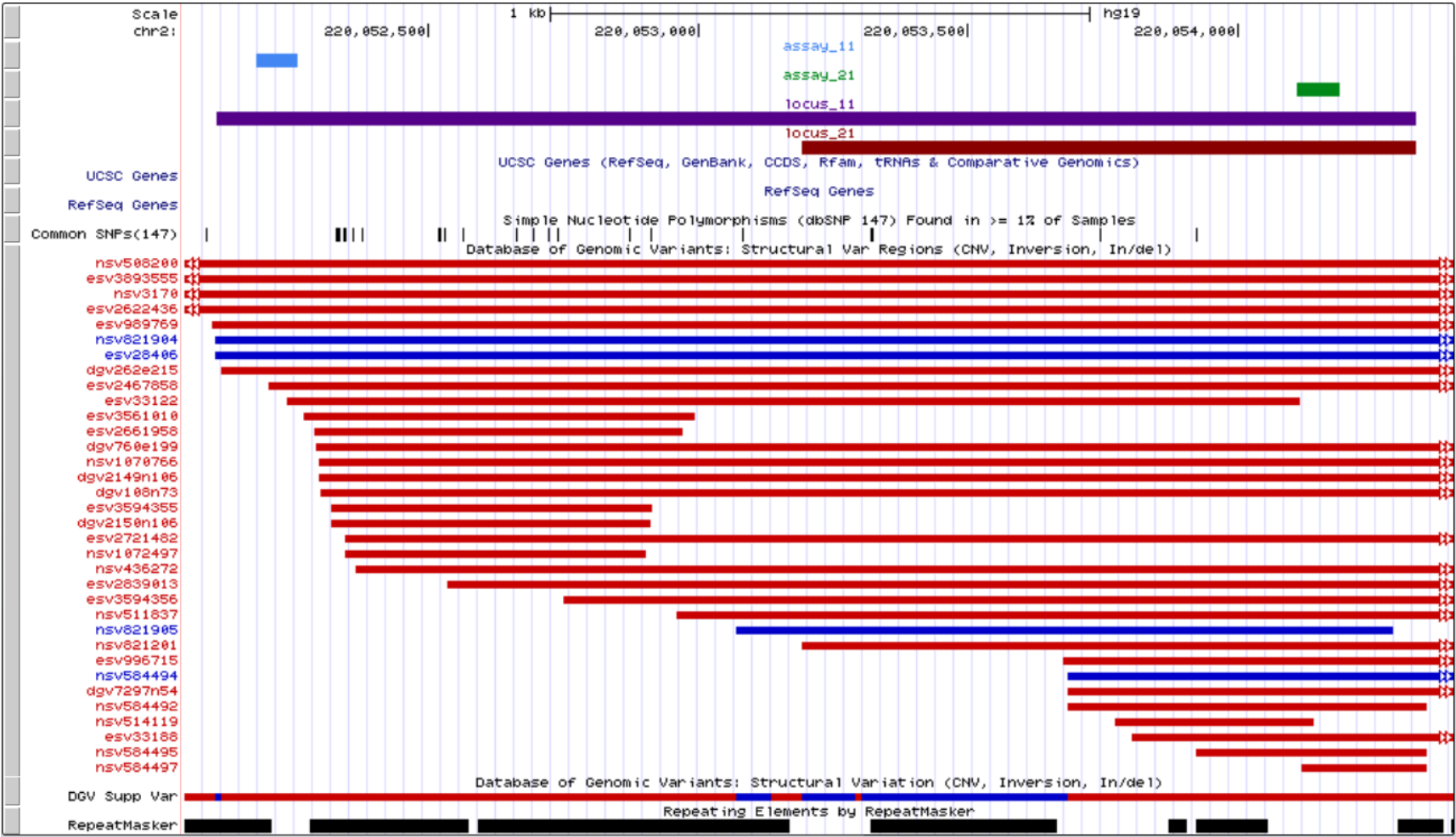


Figure A-6: Overlapping CNV 37 locus (locus_37) and CNV 40 locus (locus_40) demonstrating the position of assay 37A (assay_37) with respect to the overlapping loci. The assay amplicon for CNV 37A does not overlap with the CNV 40 locus.

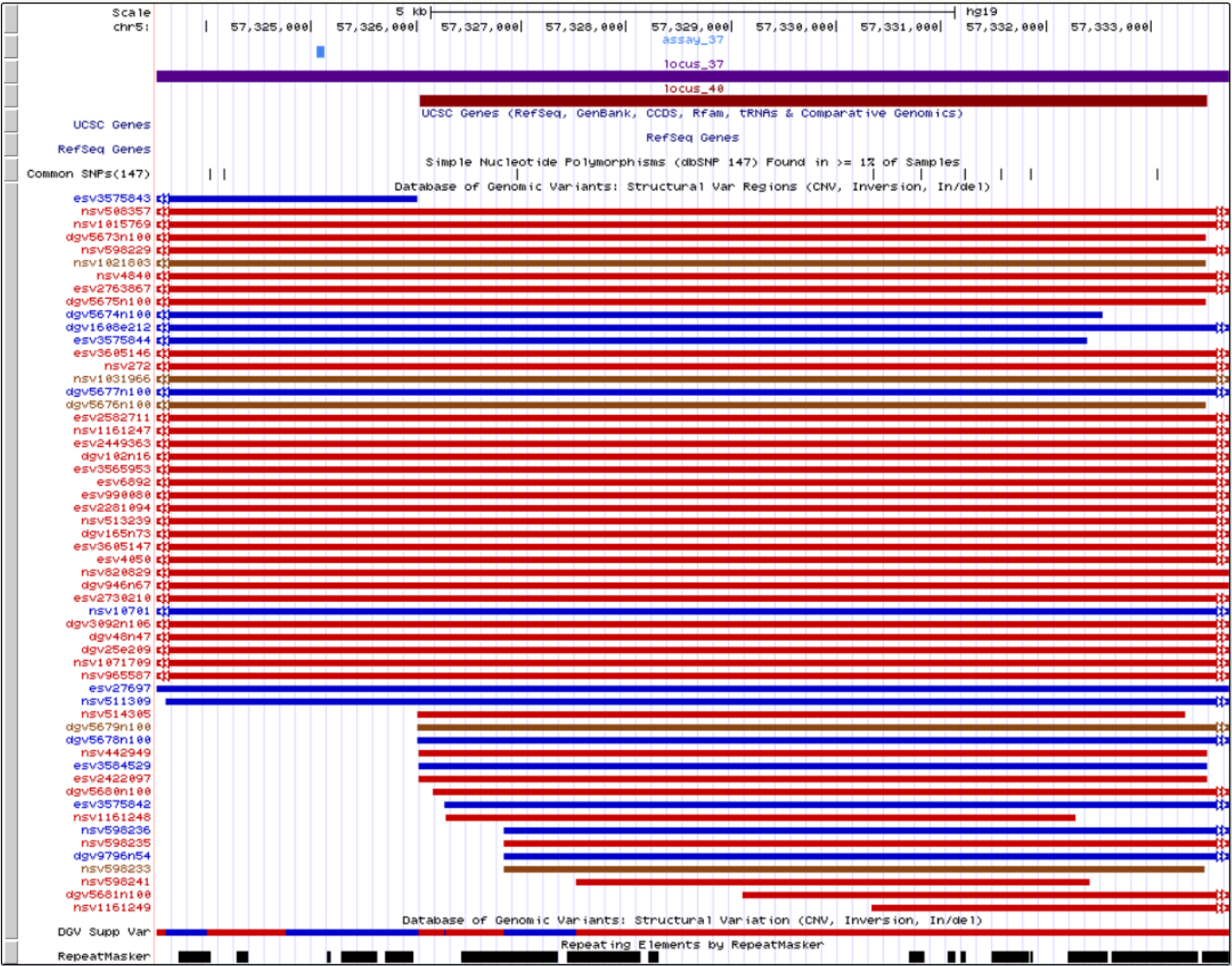


Figure A-7: Position of the *ACE* assay within the *ACE* gene relative to common SNPs, repeating elements and structural variation.

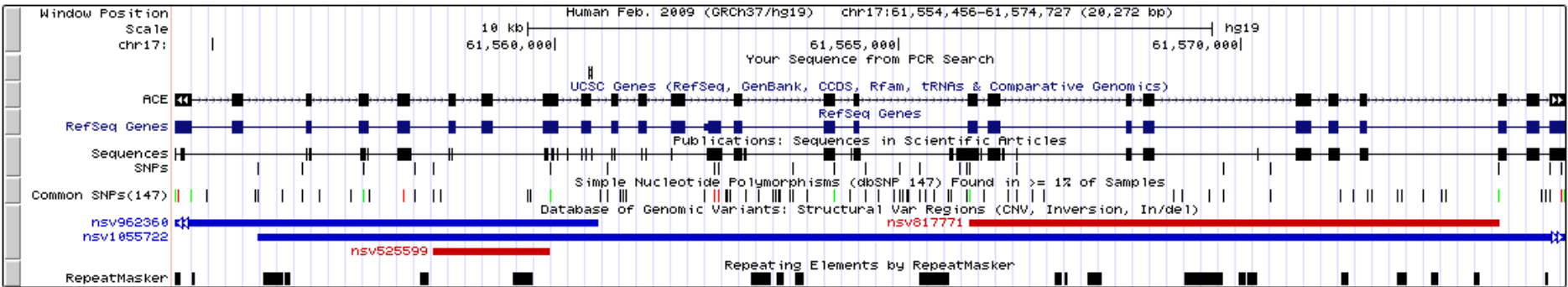


Figure A-8: *ACE* assay spanning the junction of exon 9/25 and intron 9/24 of the *ACE* gene. It also sits within two reported structural variants.

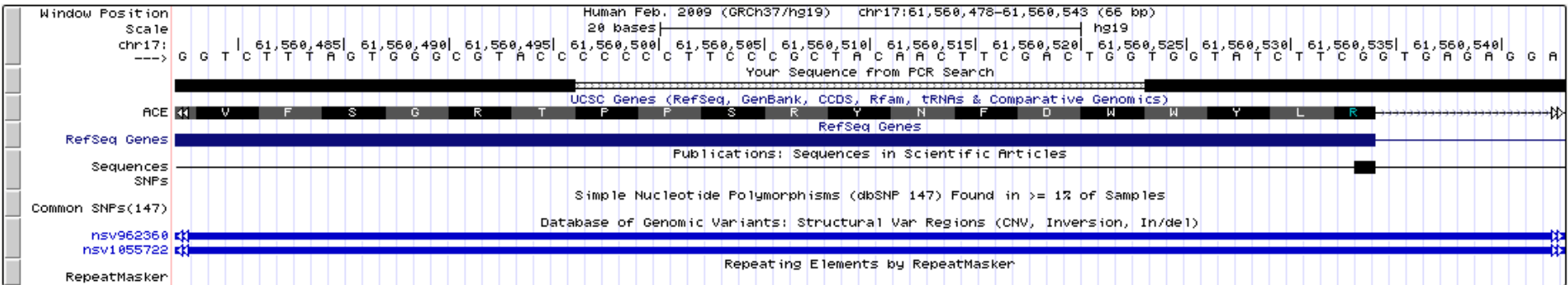
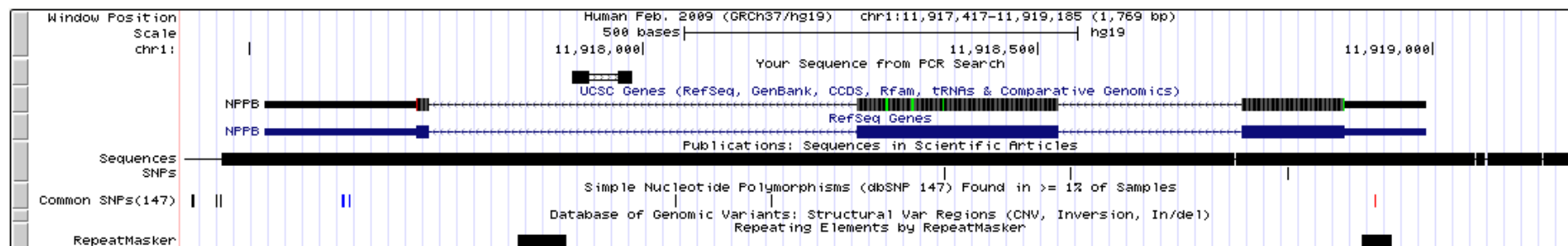


Figure A-9: Position of the *NPPB* assay within intron 2/2 of the *NPPB* gene relative to common SNPs and repeating elements.



APPENDIX B ASSAY UCSC GENOME BROWSER TRACKS

Figure B-1: The CNV 1 locus and position of the CNV 01B assay with respect to genes, common SNPs, structural variation and repeating elements.

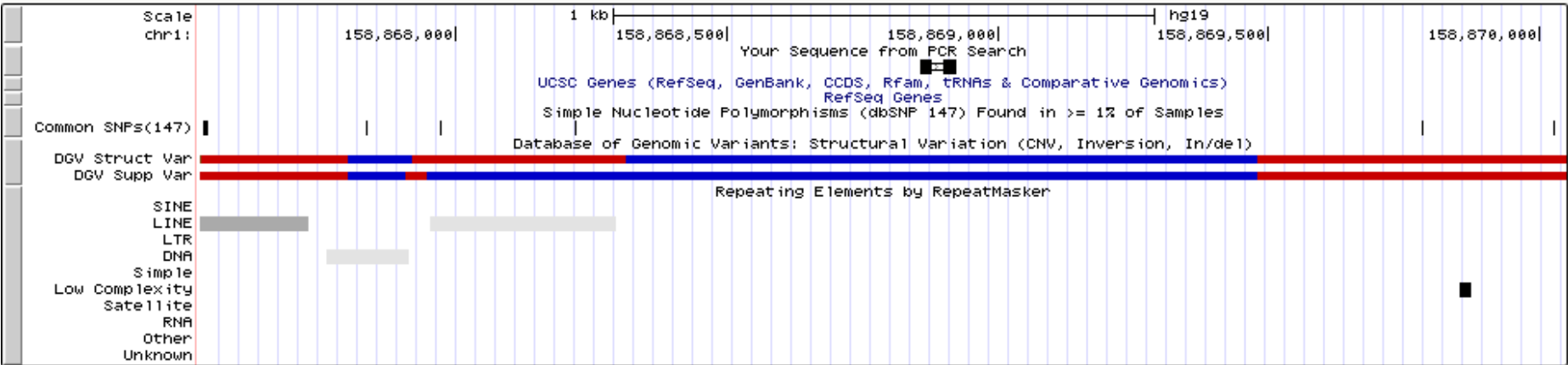


Figure B-2: The CNV 2 locus and position of the CNV 02B assay with respect to genes, common SNPs, structural variation and repeating elements.

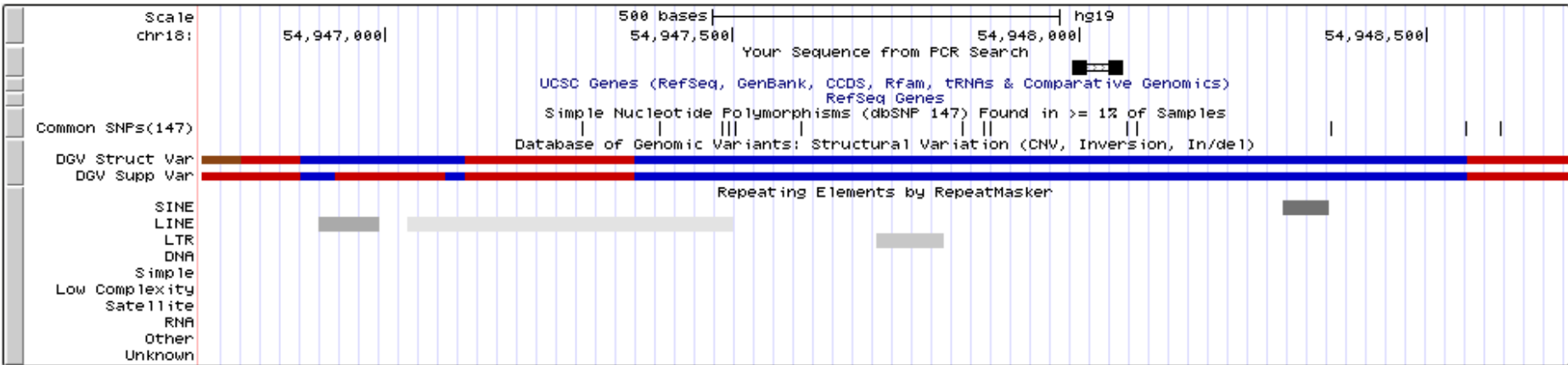


Figure B-3: The CNV 3 locus and position of the CNV 03B assay with respect to genes, common SNPs, structural variation and repeating elements.

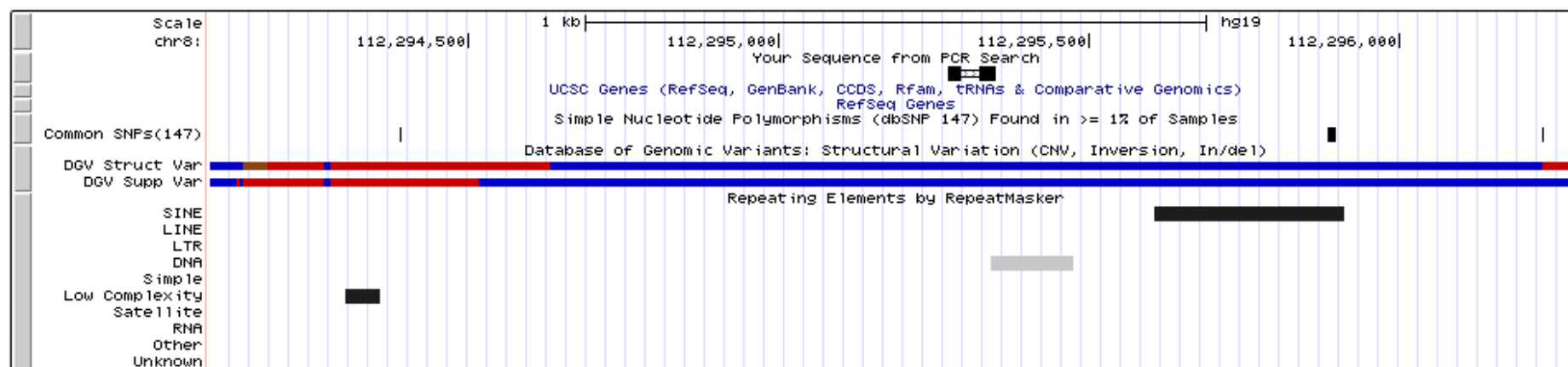


Figure B-4: The CNV 4 locus and position of the CNV 04B assay with respect to genes, common SNPs, structural variation and repeating elements.

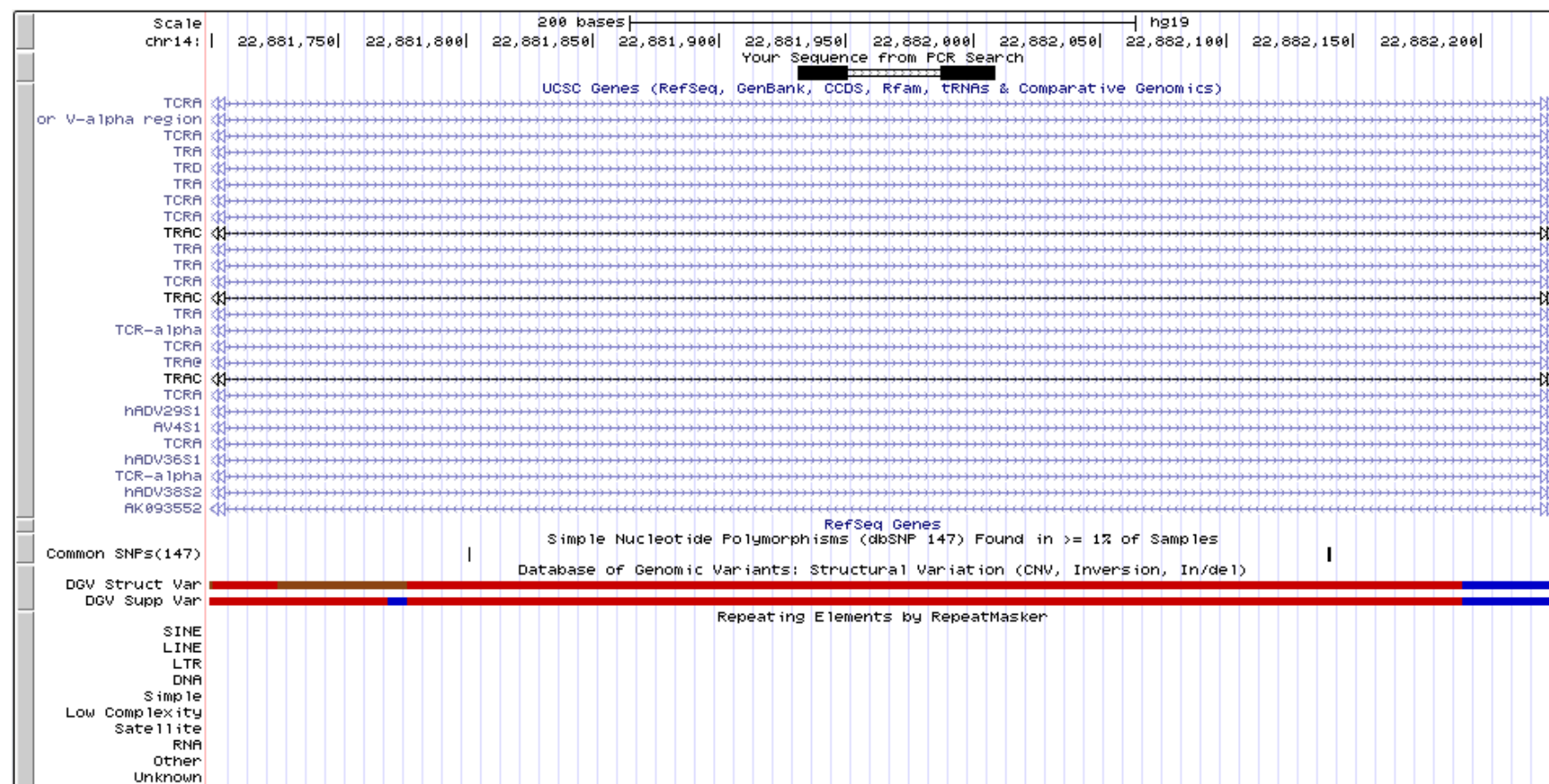


Figure B-5: The CNV 5 locus and position of the CNV 05B assay with respect to genes, common SNPs, structural variation and repeating elements.

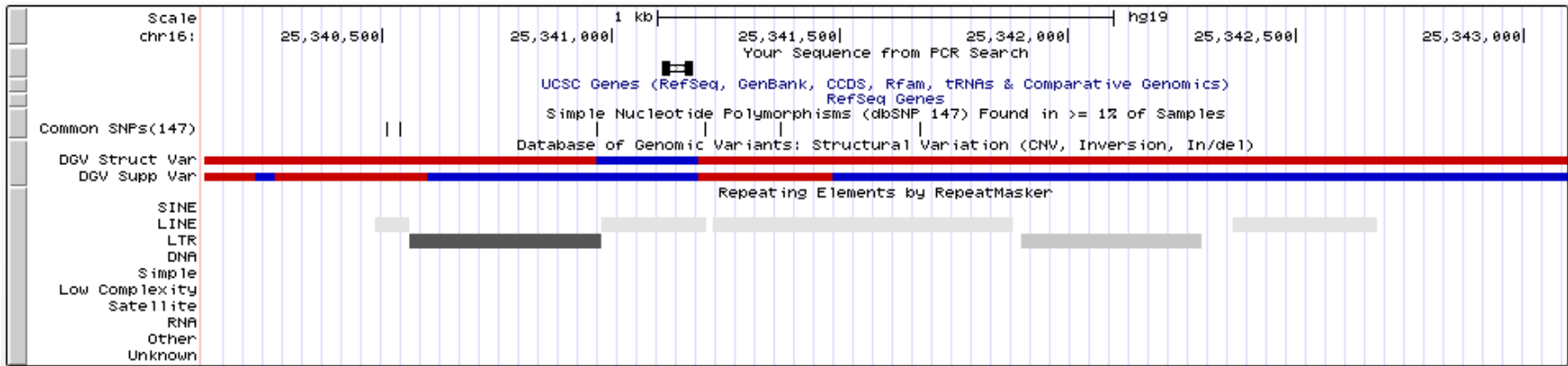


Figure B-6: The CNV 6 locus and position of the CNV 06B assay with respect to genes, common SNPs, structural variation and repeating elements.

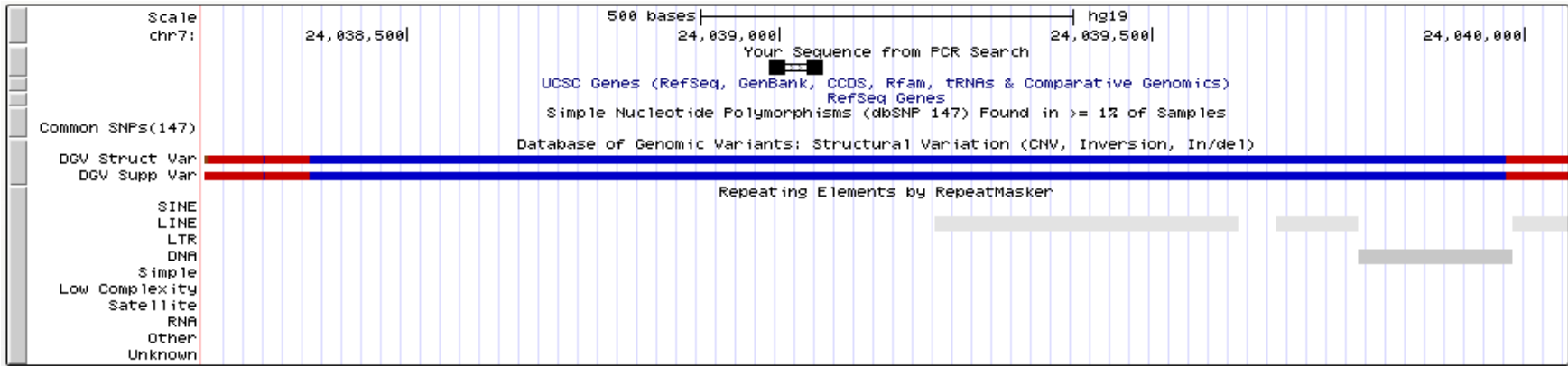


Figure B-7: The CNV 7 locus and position of the CNV 07B assay with respect to genes, common SNPs, structural variation and repeating elements.

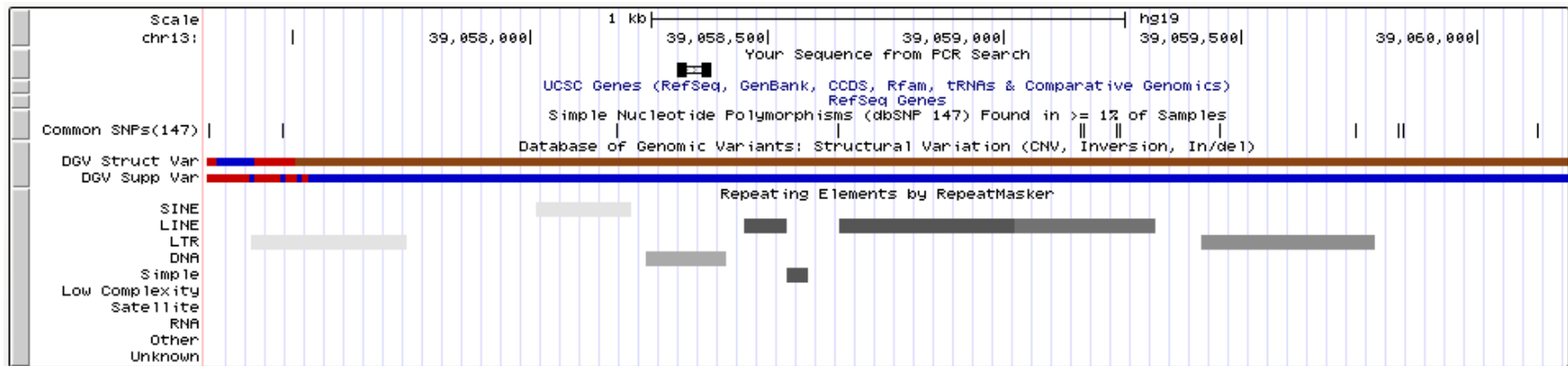


Figure B-8: The CNV 8 locus and position of the CNV 08B assay with respect to genes, common SNPs, structural variation and repeating elements.

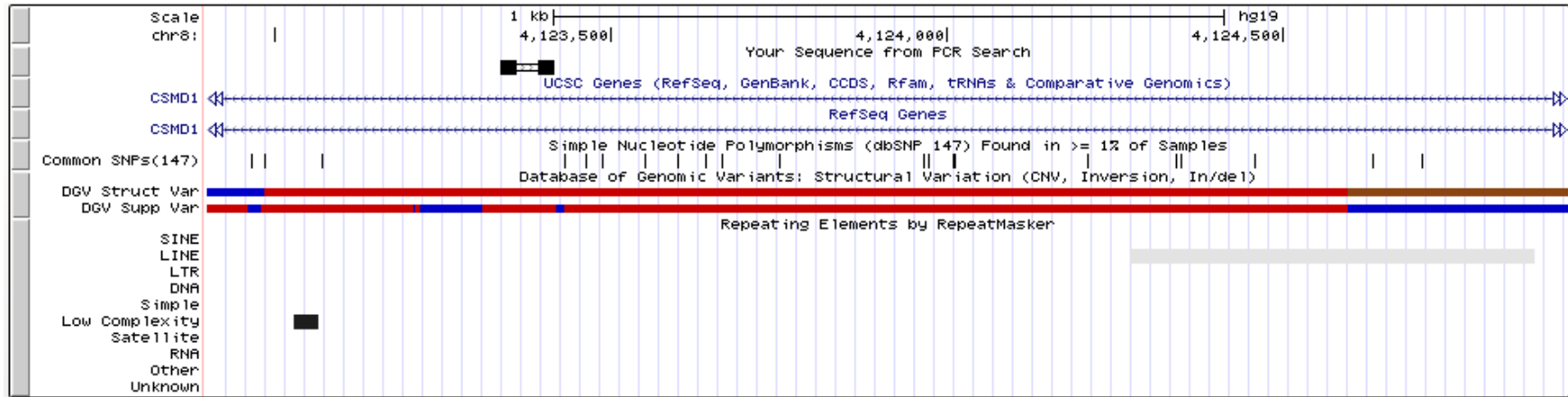


Figure B-9: The CNV 9 locus and position of the CNV 09B assay with respect to genes, common SNPs, structural variation and repeating elements.

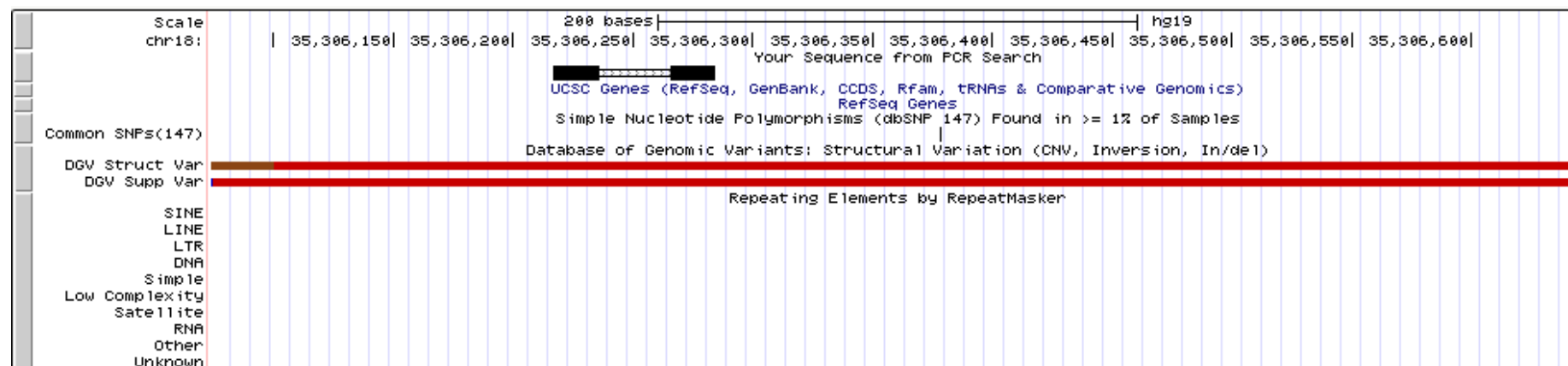


Figure B-10: The CNV 10 locus and position of the CNV 10C assay with respect to genes, common SNPs, structural variation and repeating elements.

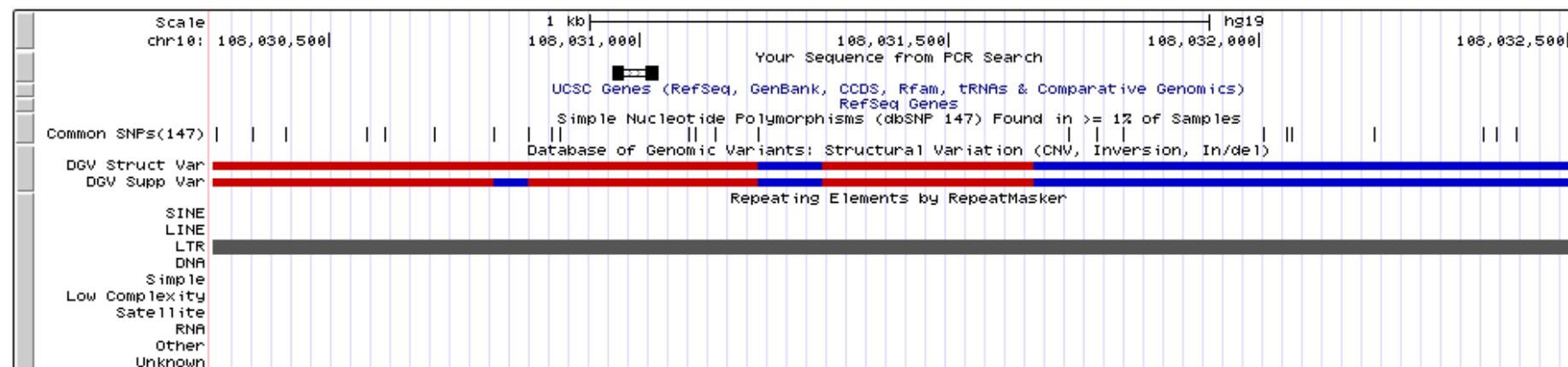


Figure B-11: The CNV 11 locus and position of the CNV 11A assay with respect to genes, common SNPs, structural variation and repeating elements.

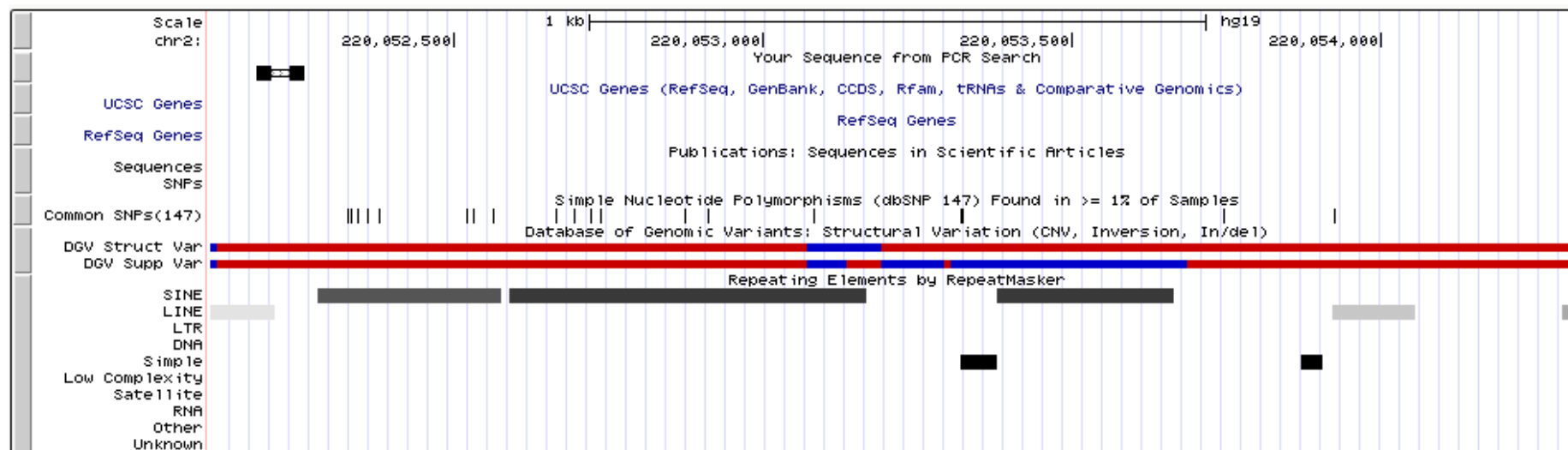


Figure B-12: The CNV 12 locus and position of the CNV 12A assay with respect to genes, common SNPs, structural variation and repeating elements.

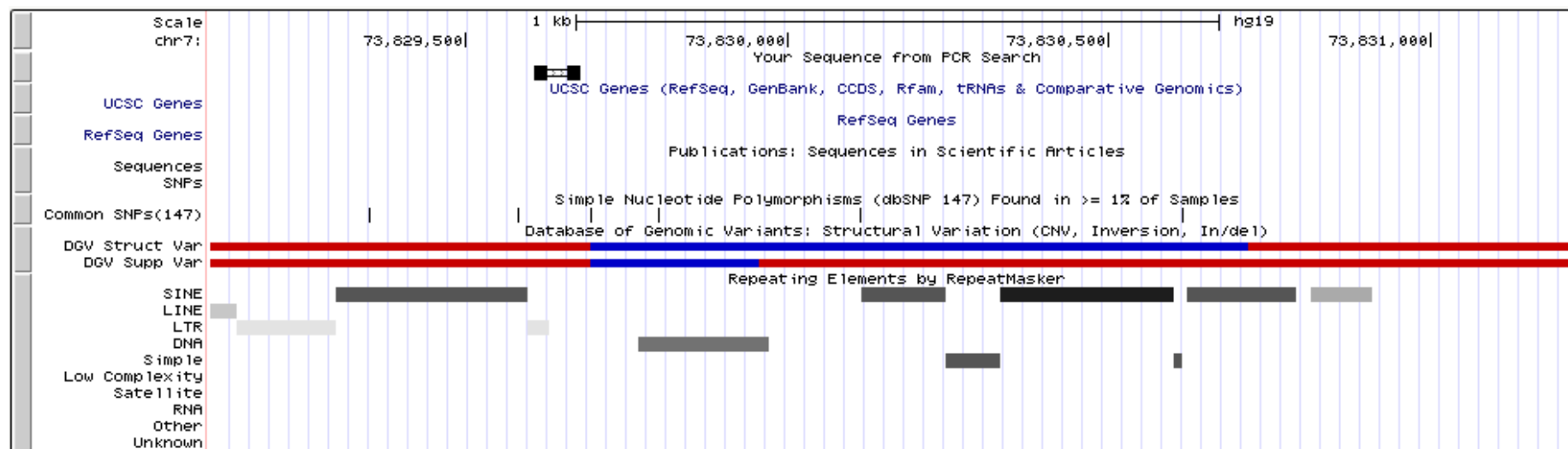


Figure B-13: The CNV 13 locus and position of the CNV 13A assay with respect to genes, common SNPs, structural variation and repeating elements.

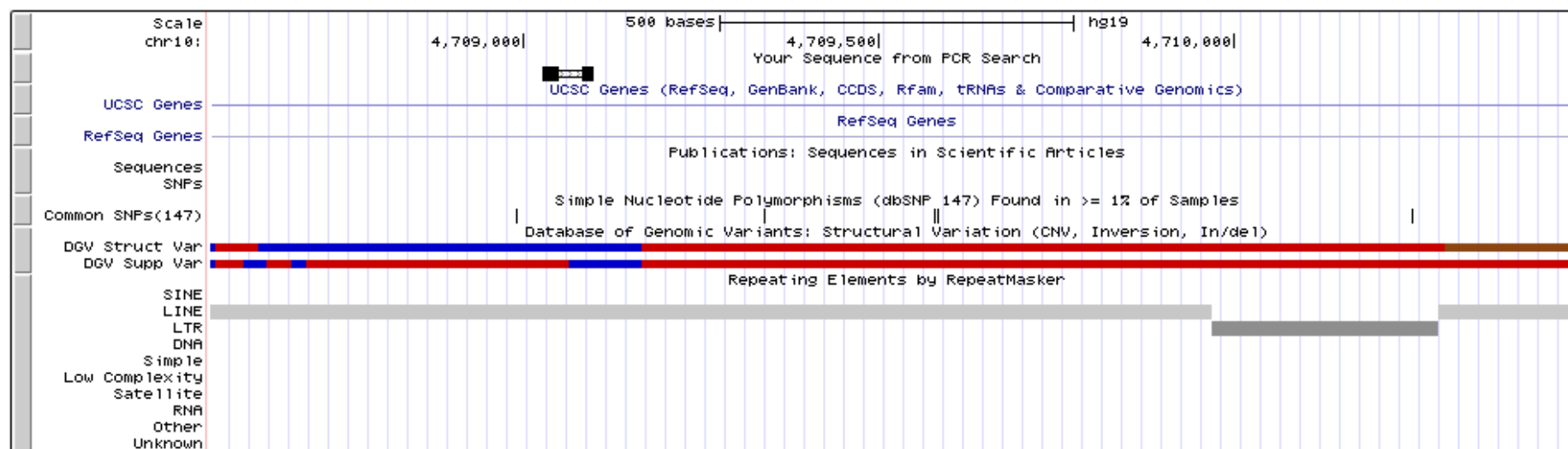


Figure B-14: The CNV 14 locus and position of the CNV 14A assay with respect to genes, common SNPs, structural variation and repeating elements.

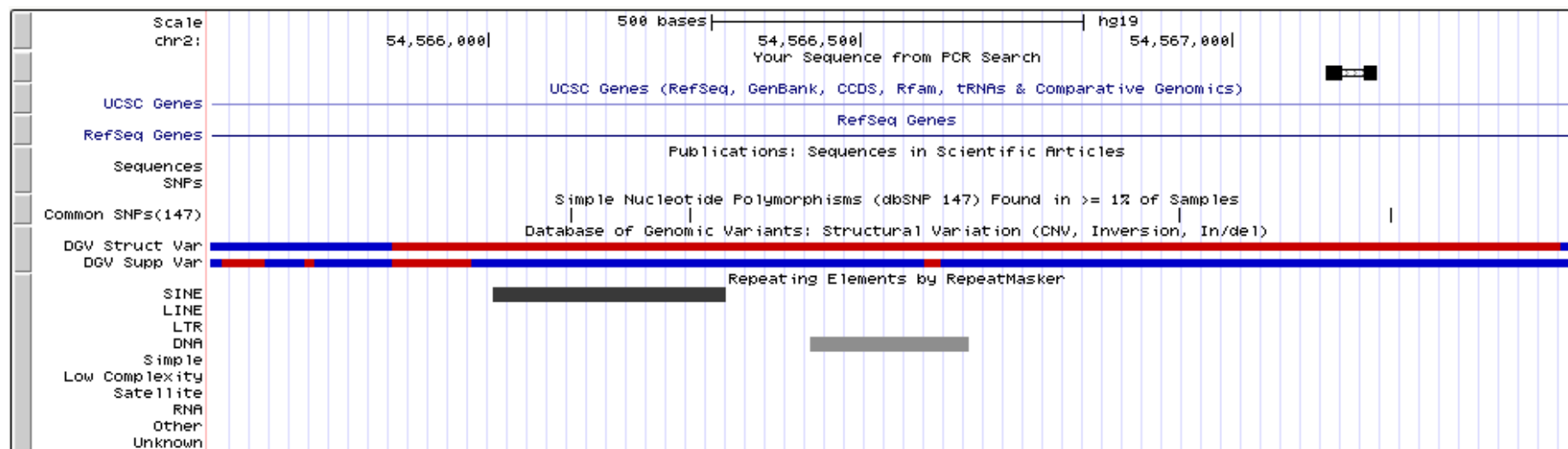


Figure B-15: The CNV 15 locus and position of the CNV 15A assay with respect to genes, common SNPs, structural variation and repeating elements.

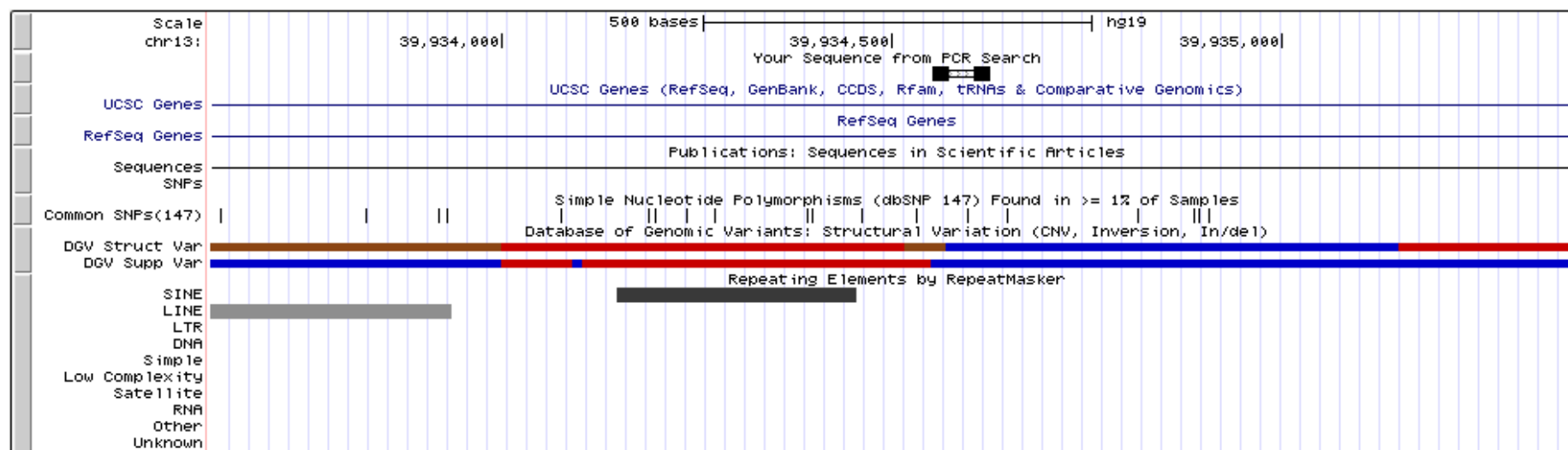


Figure B-16: The CNV 16 locus and position of the CNV 16A assay with respect to genes, common SNPs, structural variation and repeating elements.

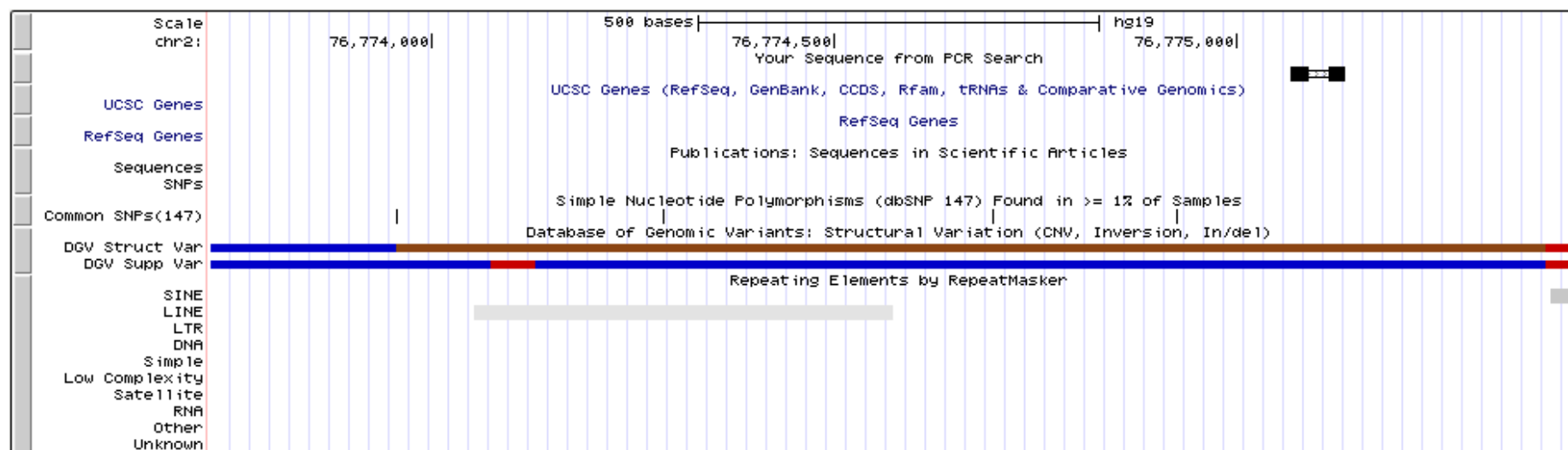


Figure B-17: The CNV 17 locus and position of the CNV 17A assay with respect to genes, common SNPs, structural variation and repeating elements.

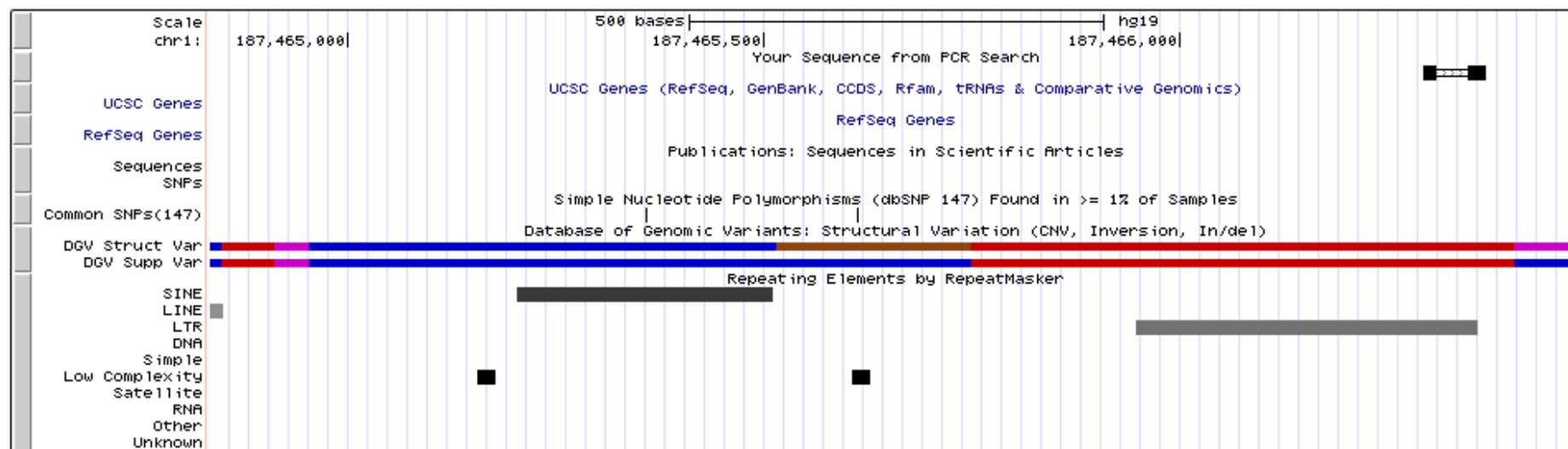


Figure B-18: The CNV 18 locus and position of the CNV 18A assay with respect to genes, common SNPs, structural variation and repeating elements.

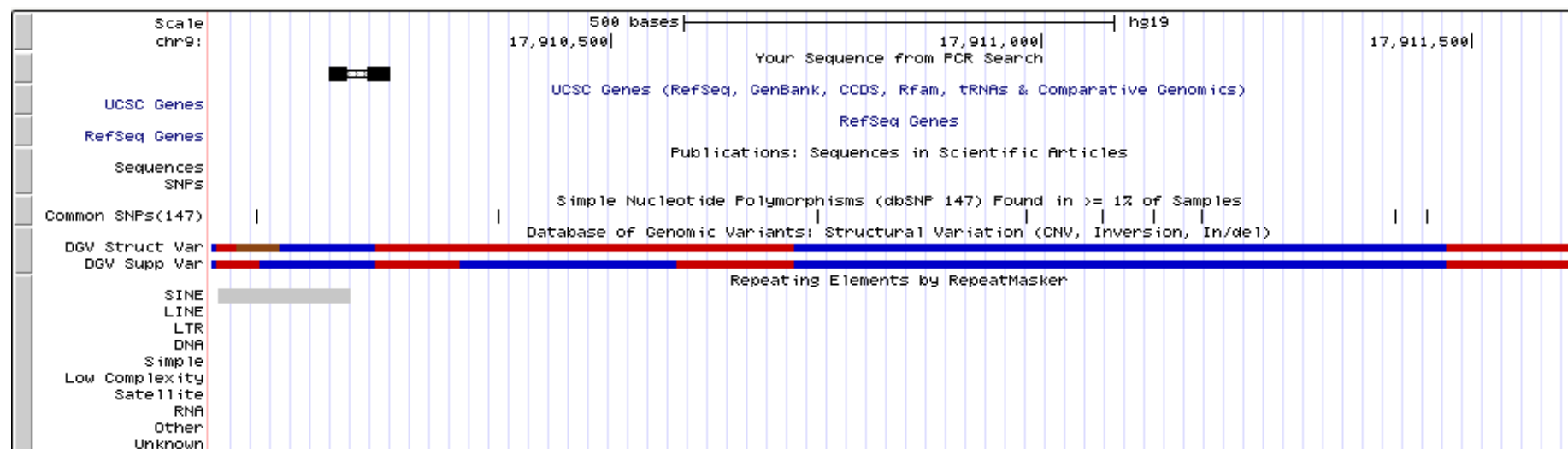


Figure B-19: The CNV 19 locus and position of the CNV 19B assay with respect to genes, common SNPs, structural variation and repeating elements.

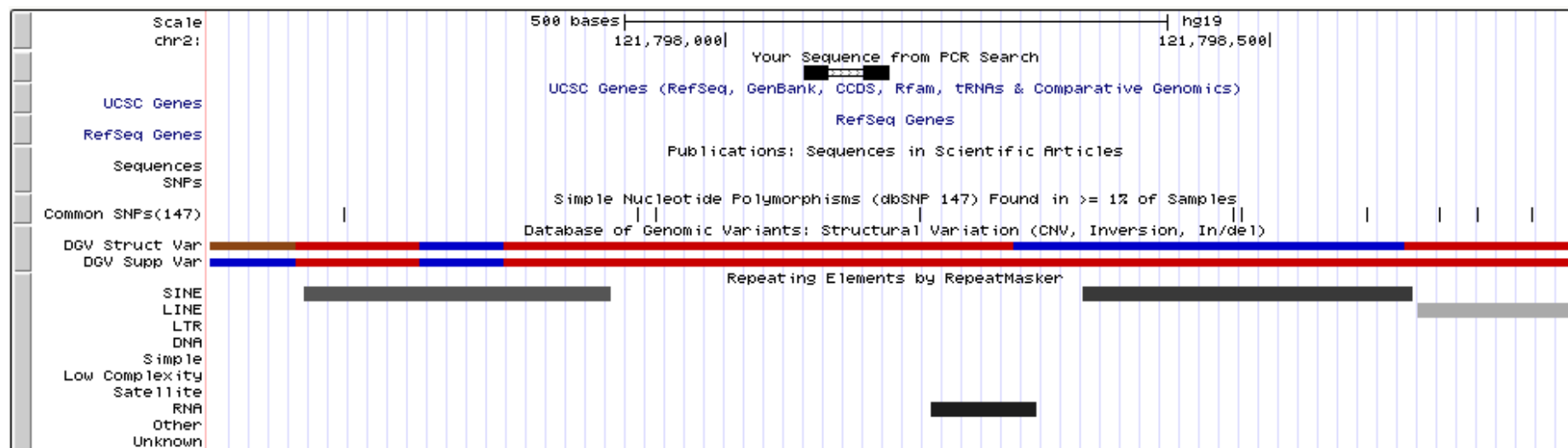


Figure B-20: The CNV 20 locus and position of the CNV 20A assay with respect to genes, common SNPs, structural variation and repeating elements.

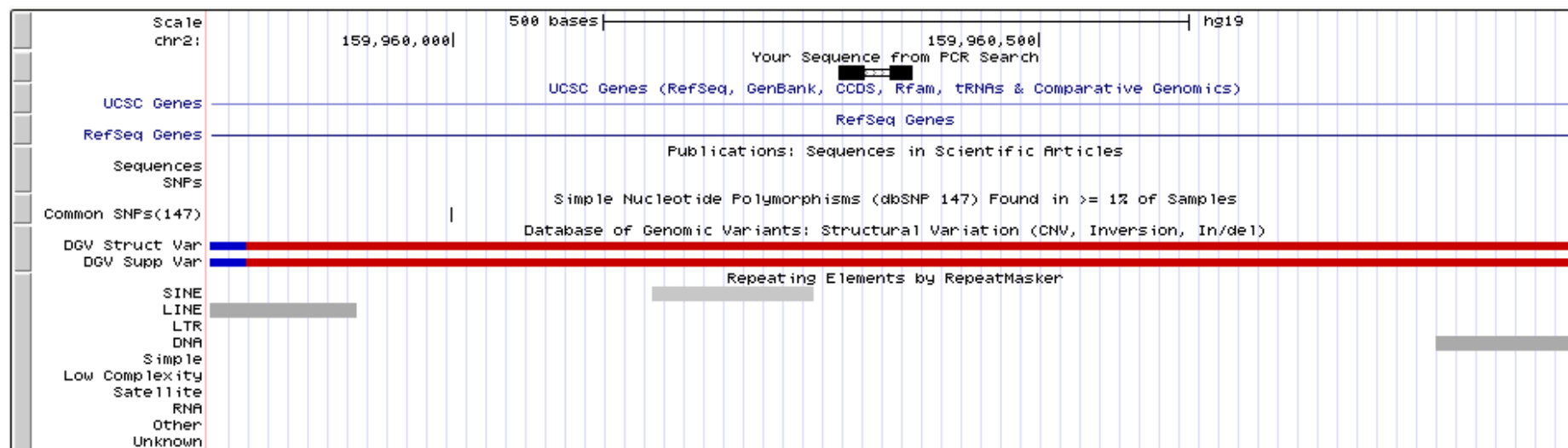


Figure B-21: The CNV 21 locus and position of the CNV 21A assay with respect to genes, common SNPs, structural variation and repeating elements.

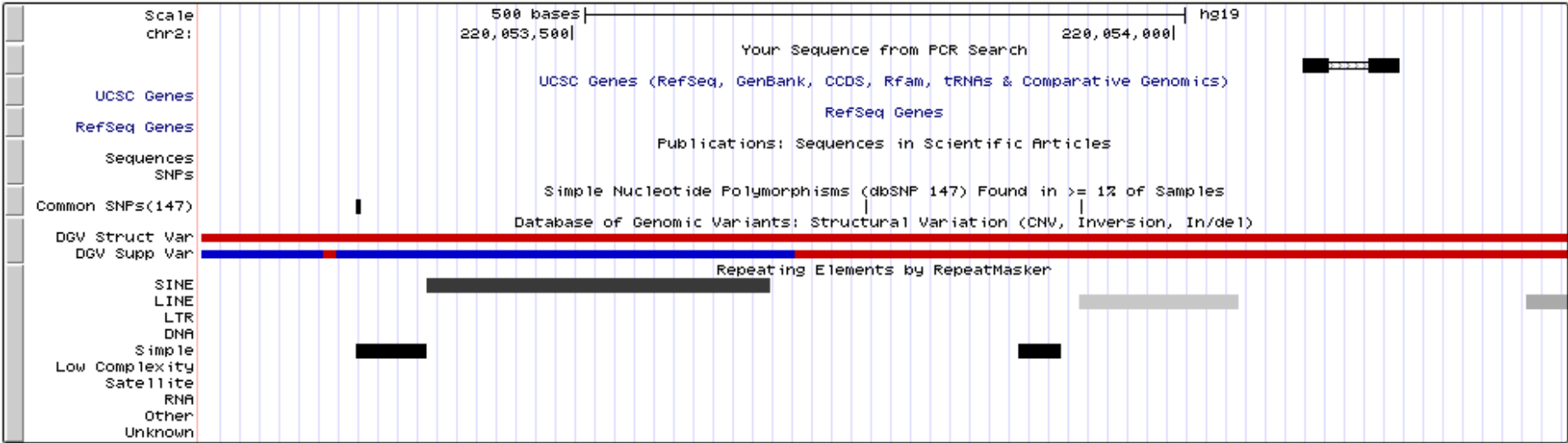


Figure B-22: The CNV 22 locus and position of the CNV 22A assay with respect to genes, common SNPs, structural variation and repeating elements.

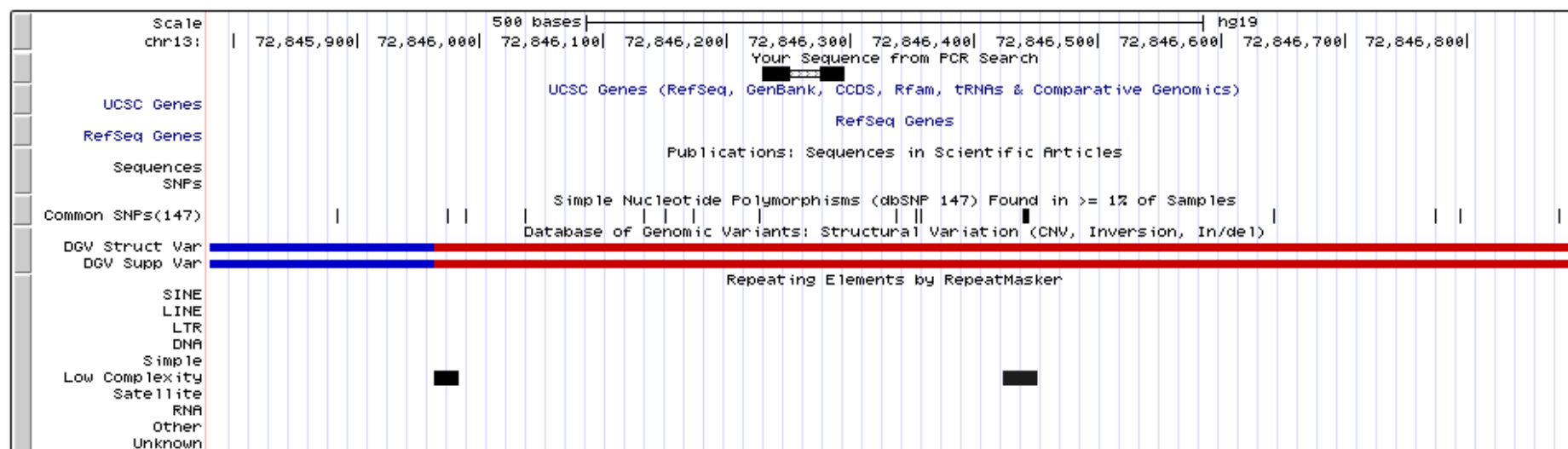


Figure B-23: The CNV 23 locus and position of the CNV 23A assay with respect to genes, common SNPs, structural variation and repeating elements.

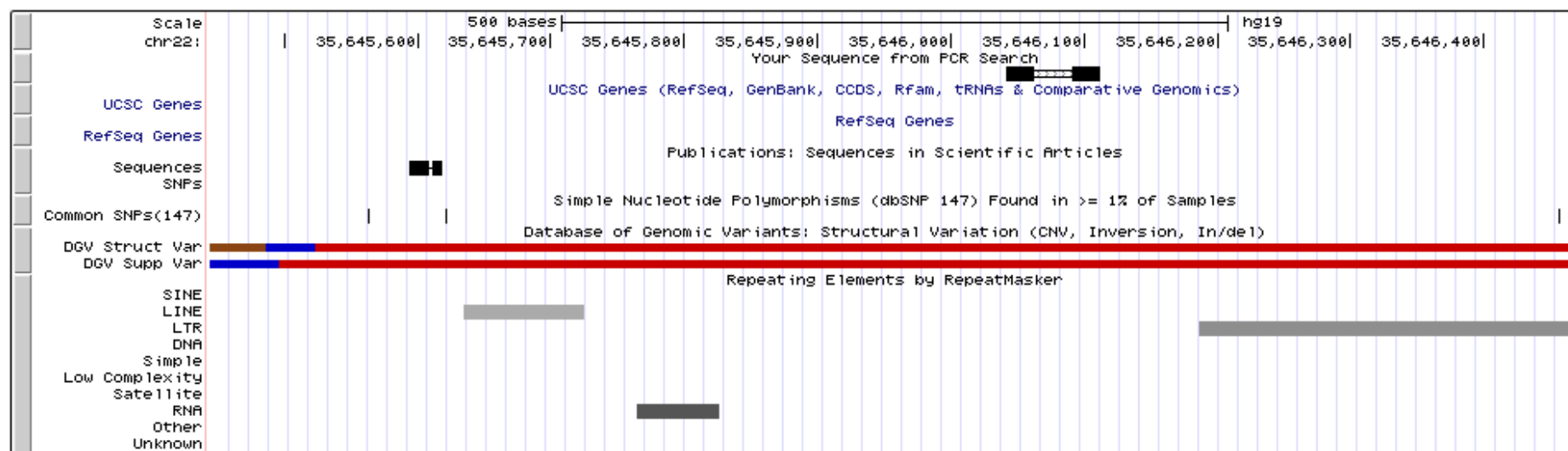


Figure B-24: The CNV 24 locus and position of the CNV 24A assay with respect to genes, common SNPs, structural variation and repeating elements.

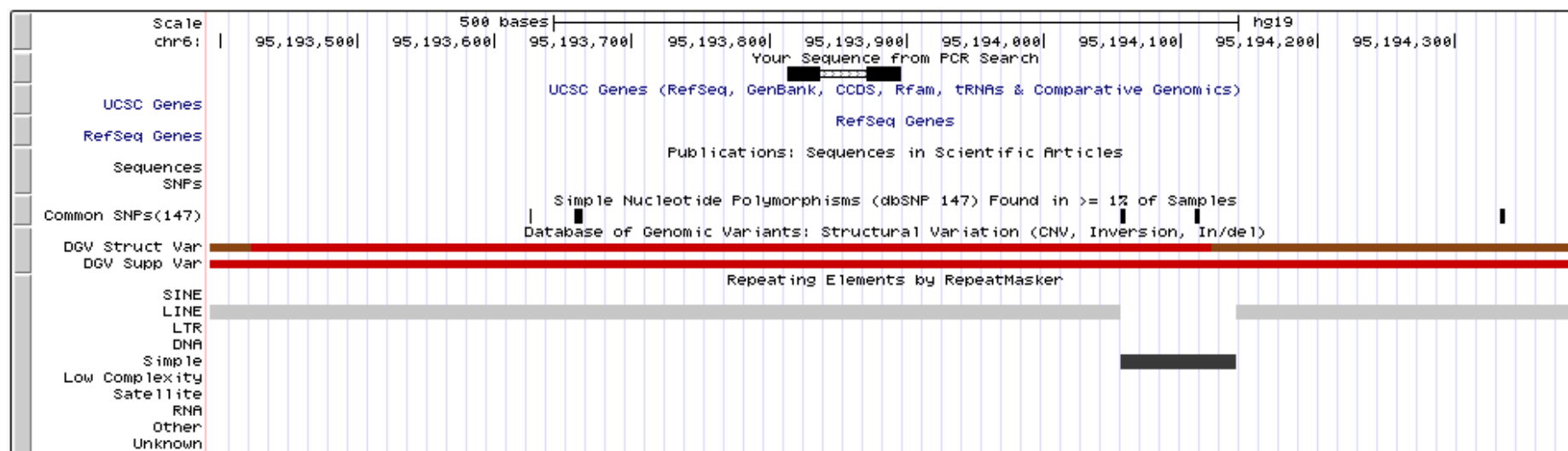


Figure B-25: The CNV 25 locus and position of the CNV 25A assay with respect to genes, common SNPs, structural variation and repeating elements.

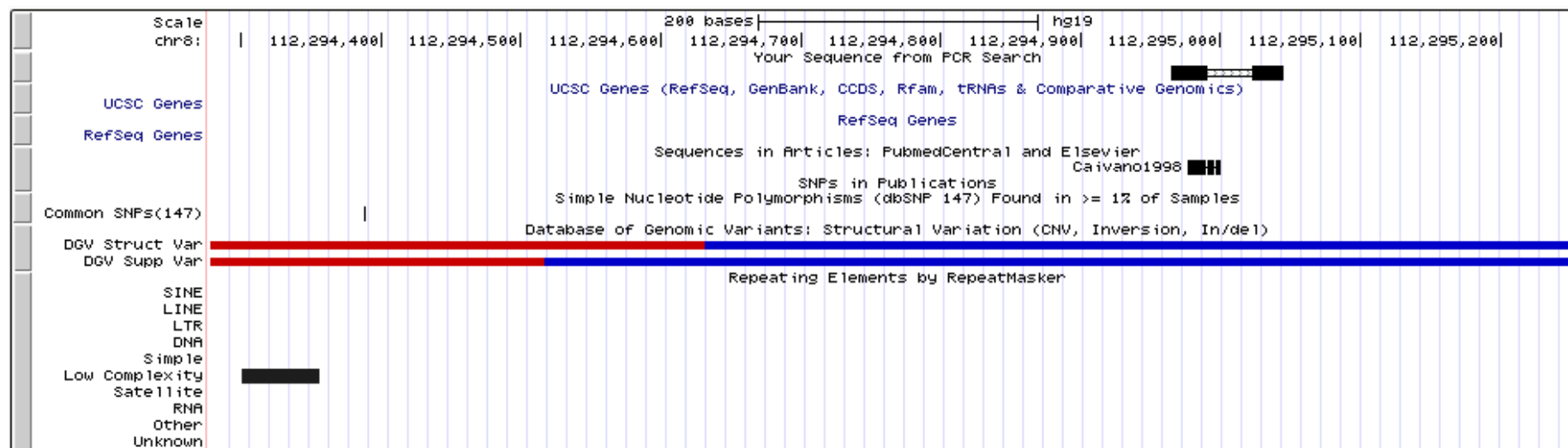


Figure B-26: The CNV 26 locus and position of the CNV 26A assay with respect to genes, common SNPs, structural variation and repeating elements.

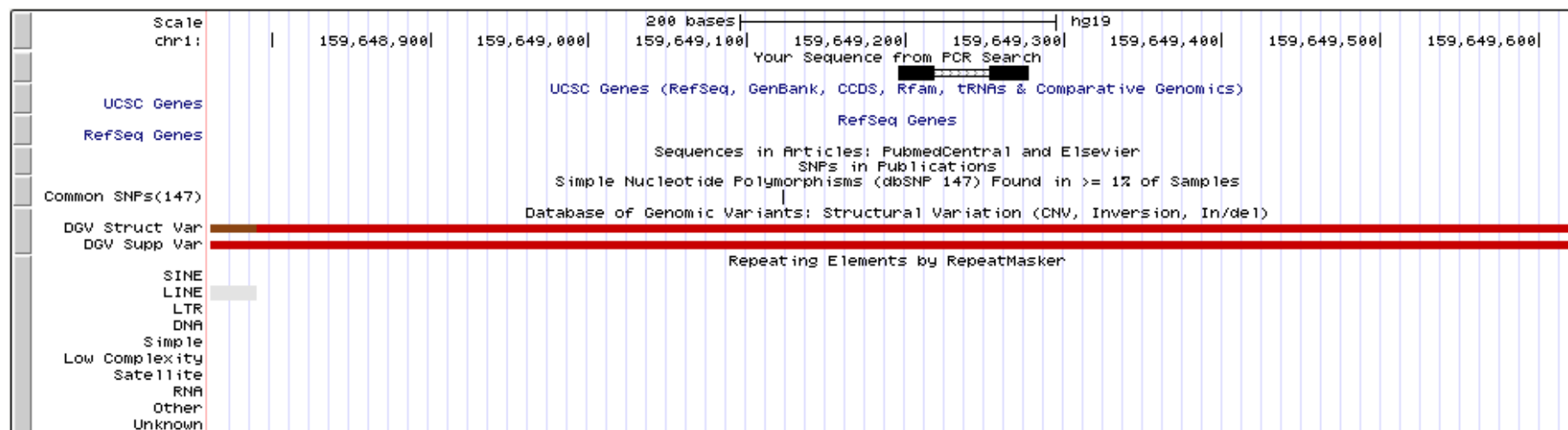


Figure B-27: The CNV 27 locus and position of the CNV 27A assay with respect to genes, common SNPs, structural variation and repeating elements.

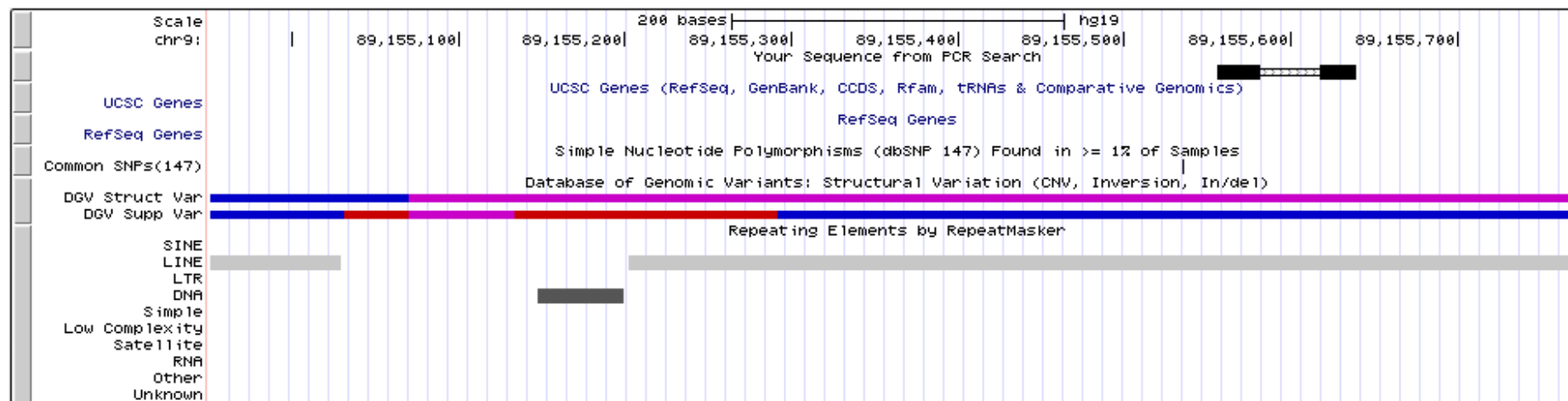


Figure B-28: The CNV 28 locus and position of the CNV 28A assay with respect to genes, common SNPs, structural variation and repeating elements.

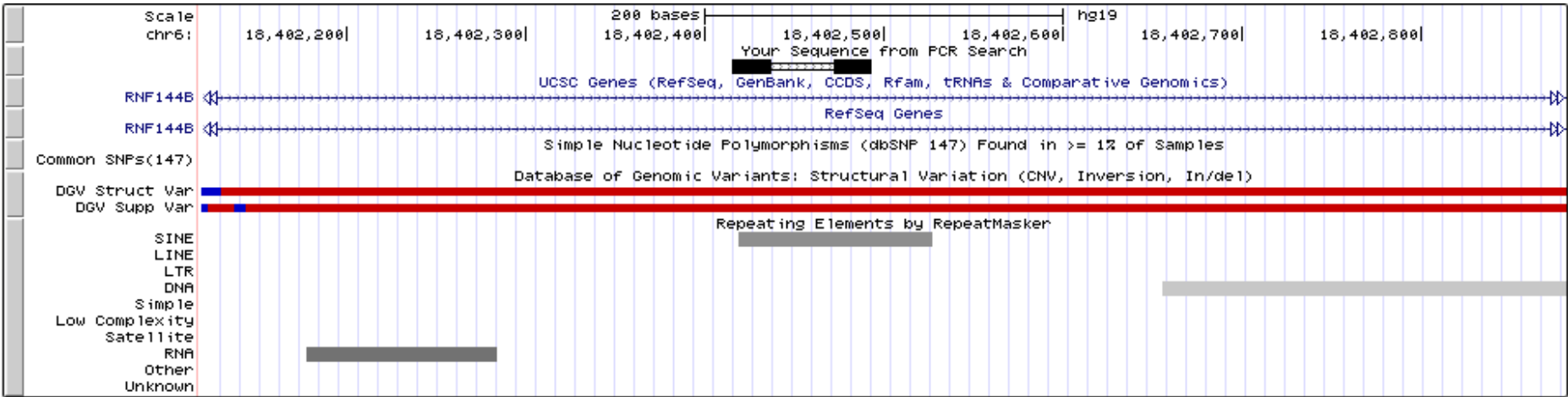


Figure B-29: The CNV 29 locus and position of the CNV 29A assay with respect to genes, common SNPs, structural variation and repeating elements.

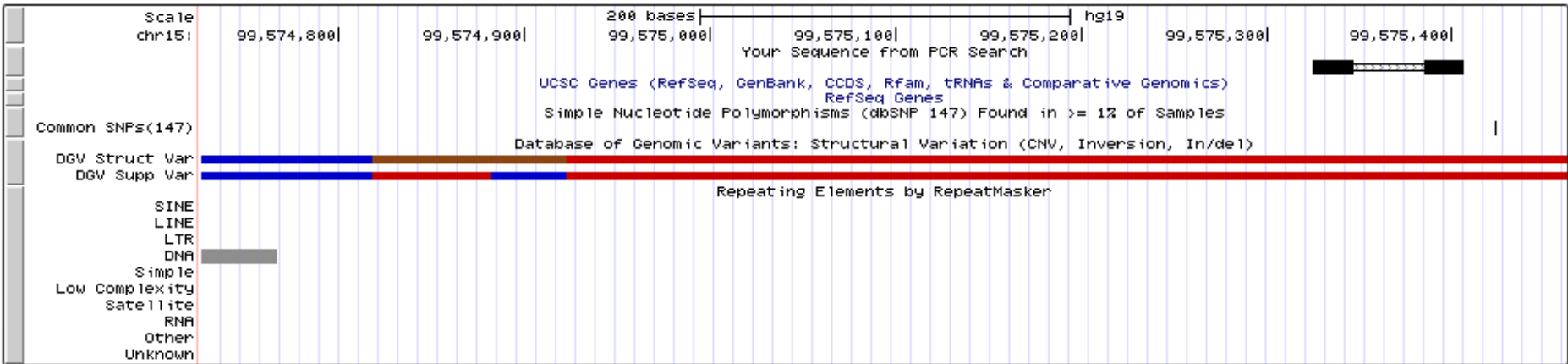


Figure B-30: The CNV 30 locus and position of the CNV 30A assay with respect to genes, common SNPs, structural variation and repeating elements.

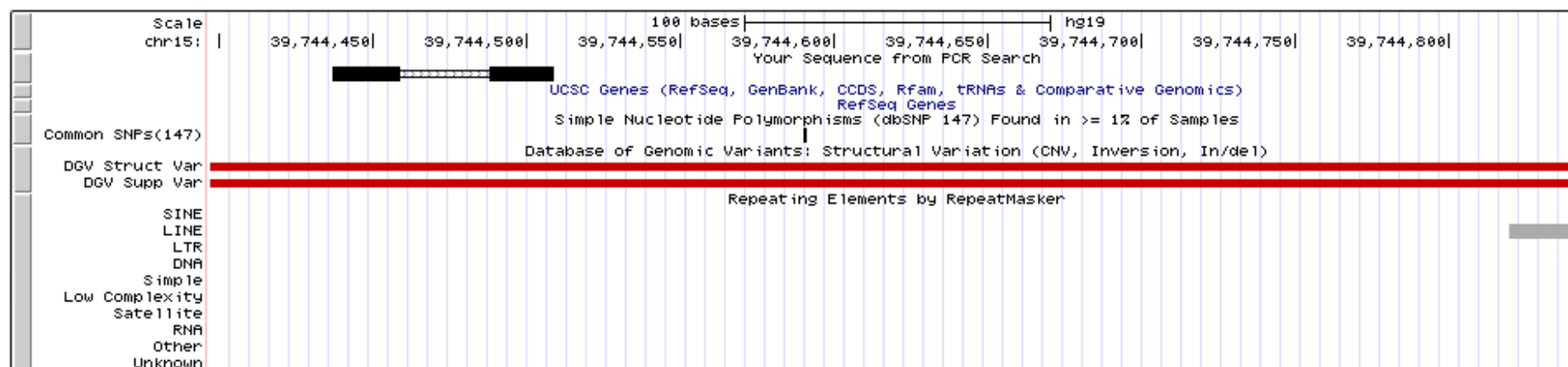


Figure B-31: The CNV 31 locus and position of the CNV 31A assay with respect to genes, common SNPs, structural variation and repeating elements.

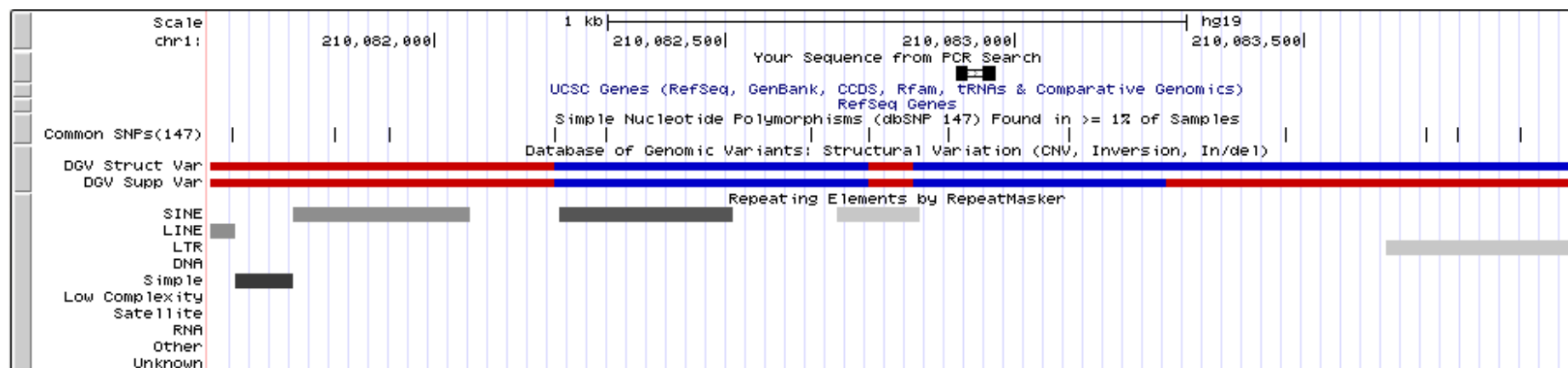


Figure B-32: The CNV 32 locus and position of the CNV 32A assay with respect to genes, common SNPs, structural variation and repeating elements.

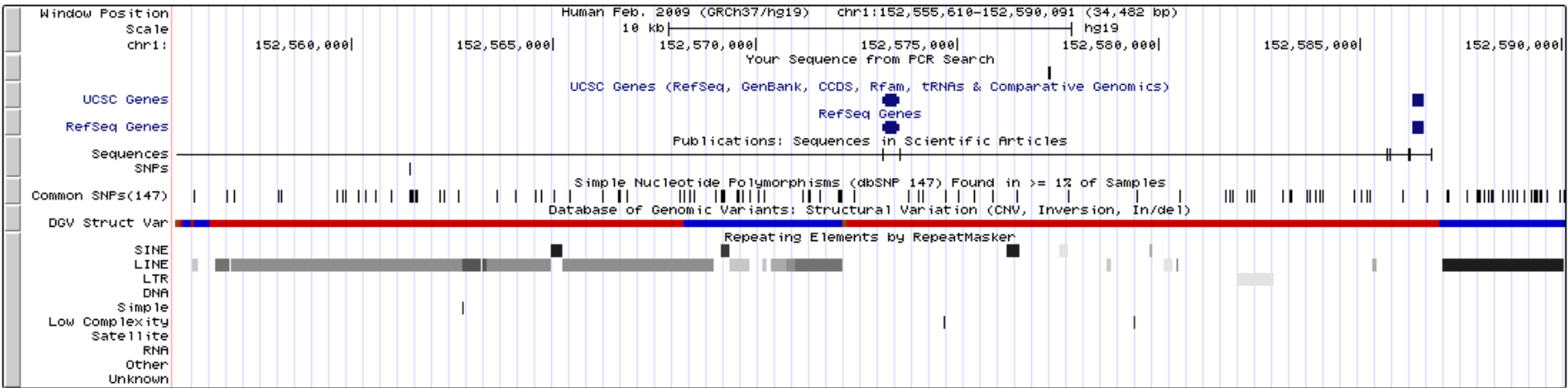


Figure B-33: The CNV 33 locus and position of the CNV 33A assay with respect to genes, common SNPs, structural variation and repeating elements.

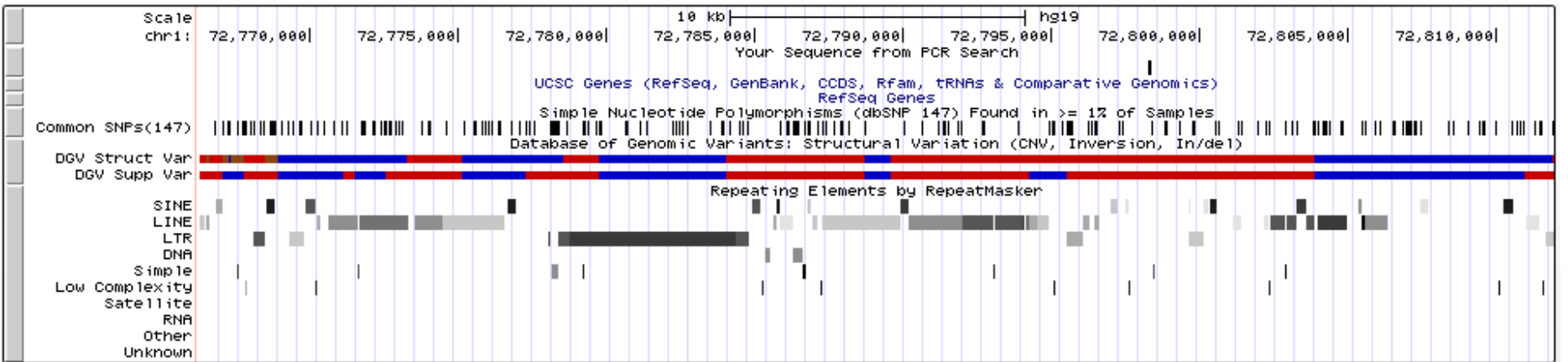


Figure B-34: The CNV 34 locus and position of the CNV 34A assay with respect to genes, common SNPs, structural variation and repeating elements.

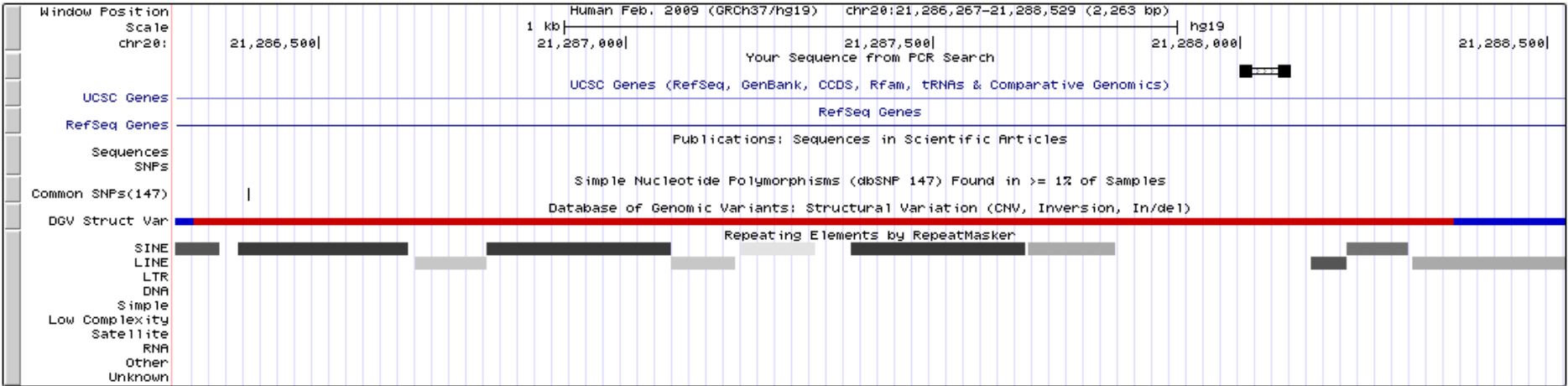


Figure B-35: The CNV 36 locus and position of the CNV 36A assay with respect to genes, common SNPs, structural variation and repeating elements.

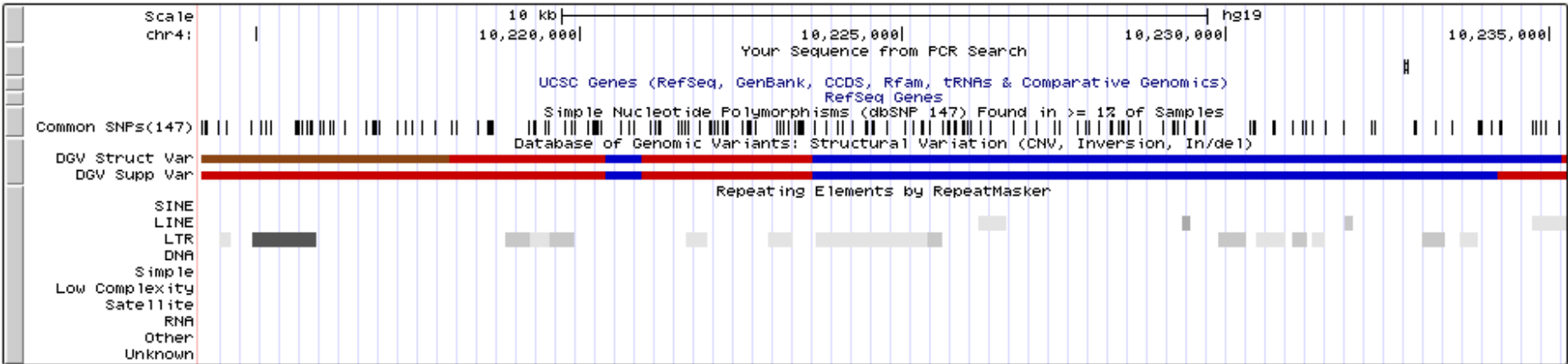


Figure B-36: The CNV 37 locus and position of the CNV 37A assay with respect to genes, common SNPs, structural variation and repeating elements.

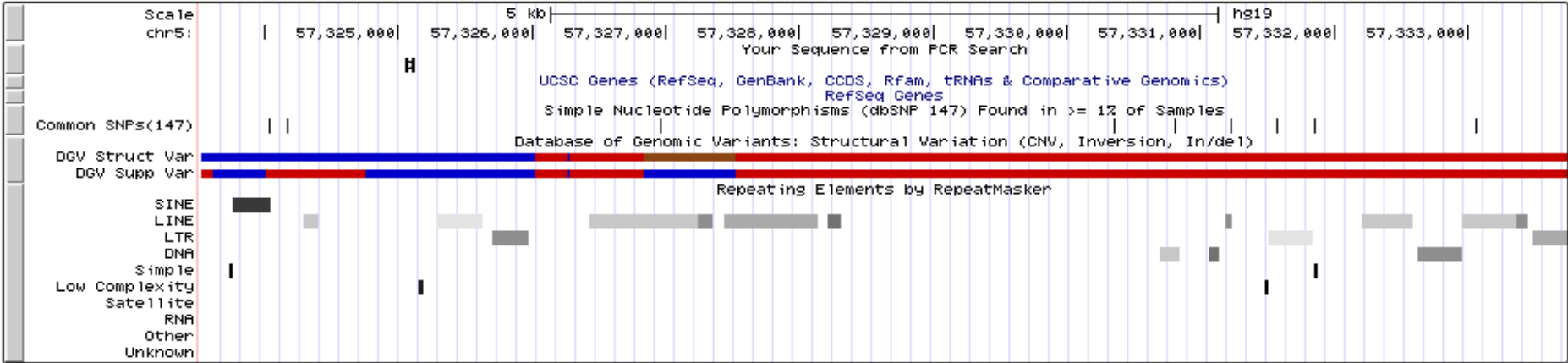


Figure B-37: The CNV 38 locus and position of the CNV 38A assay with respect to genes, common SNPs, structural variation and repeating elements.

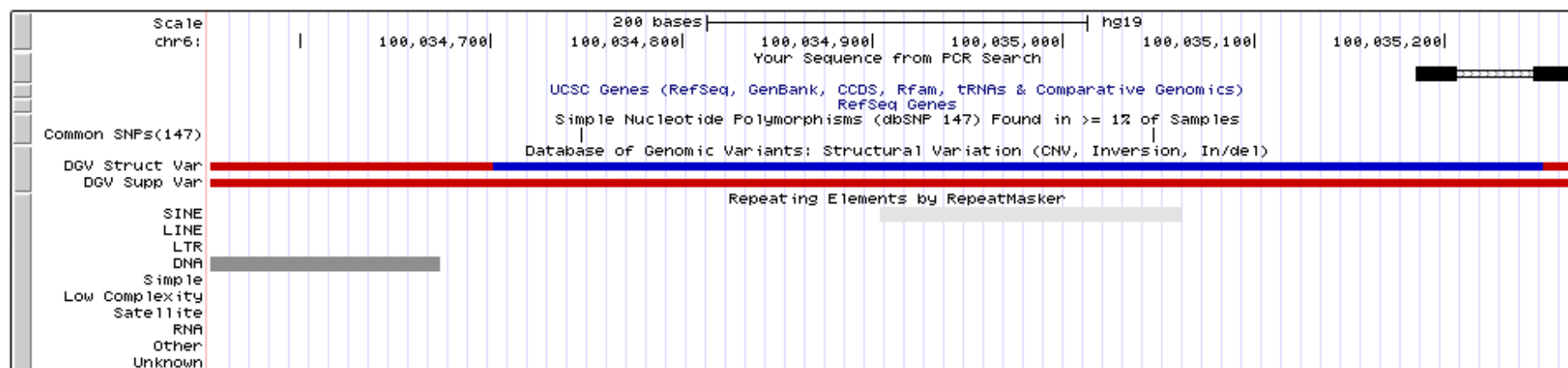


Figure B-38: The CNV 39 locus and position of the CNV 39A assay with respect to genes, common SNPs, structural variation and repeating elements.

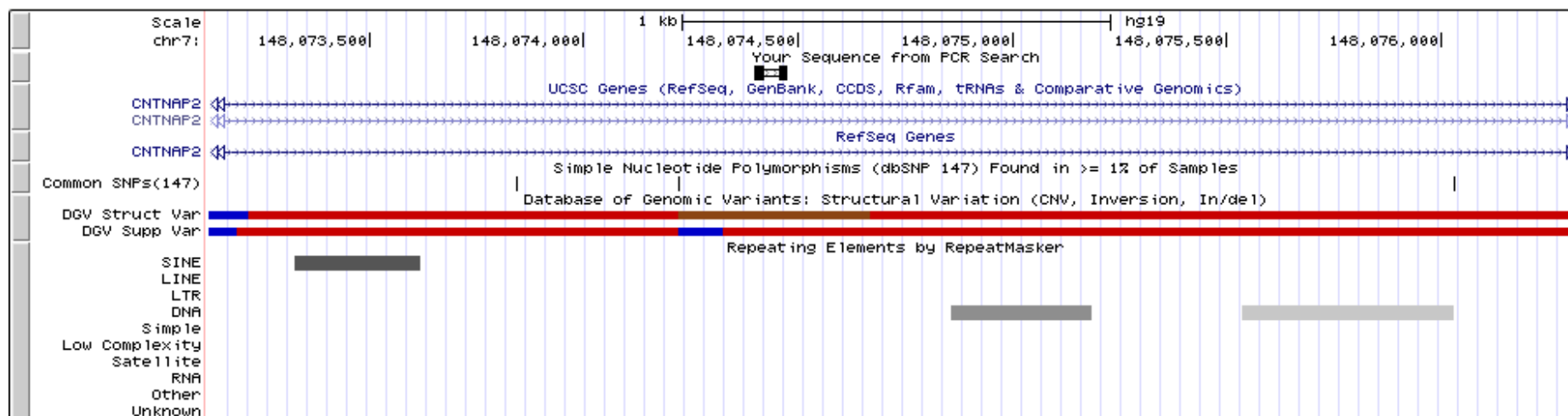
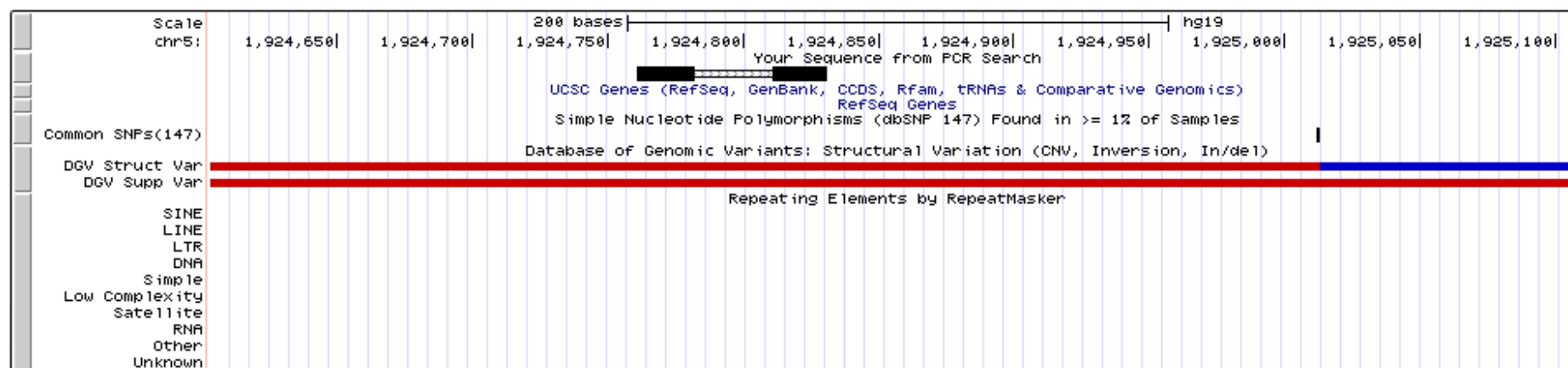


Figure B-39: The CNV 41 locus and position of the CNV 41A assay with respect to genes, common SNPs, structural variation and repeating elements.



This page is deliberately blank.