UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Research Report

EVALUATING THE CLINICAL UTILITY OF A SNP-BASED MICROARRAY PLATFORM – A COMPARATIVE STUDY

By

Minenhle Peniel Mayisela

1880919

Master of Science in Medicine in Genomic Medicine

Supervised by:

Ms Nkateko Mayevu

Dr Fiona Baine–Savanhu

NATIONAL HEALTH LABORATORY SERVICE

A research report (in the format of a “submissible” paper) submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in partial fulfilment of the requirements for the degree of Master of Science in Medicine (Genomic Medicine)

Johannesburg, 2023

Declaration

I, Minenhle Peniel Mayisela, declare this research report (in the format of a “submissible” paper) is my own, unaided work. It is being submitted for the Degree of Master of Science in Medicine at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

(Signature of candidate)

4th day in June 2023, in Johannesburg

Contribution of the candidate to the paper

Declaration: Student's contribution to article(s) and agreement of co-author

I, Minenhle Peniel Mayisela, student number 1880919, declare that this Research Report is my own work and that I contributed significantly towards research findings presented in the paper intended for publication below.

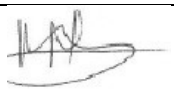
Signature of Student
Date: 4/06/2023
Name of Primary Supervisor: Ms Nkateko Mayevu

Signature of Primary Supervisor
Date: 4/06/2023

Signature of Co-Supervisor
Date: 4/06/2023

Agreement by co-authors: By signing this declaration, the co-authors listed below agree to the use of the article(s) by the student as part of her Research Project.

Article Title: Evaluating the clinical utility of a SNP-based microarray platform – a comparative study

Authors	Name	Signature	Date
First Author	Minenhle Mayisela		4/06/2023
Second Author	Nkateko Mayevu		4/06/2023
Third Author	Fiona Baine - Savanhu		4/06/2023

Dedication

This research report is dedicated to my parents, Friction and Qondile Mayisela.

Abstract

Developmental disorders make up a majority of cases seen in genetic clinics at the National Health Laboratory Service (NHLS). Chromosomal microarray analysis (CMA) is the first line testing for individuals with developmental disorders and congenital anomalies. This study compared the clinical utility of single nucleotide polymorphism (SNP) - based array (CytoScan® Optima array) and a comparative genomic hybridization (CGH) array (SurePrint G3 Unrestricted CGH ISCA v2, 8x60 kb microarray) in diagnosis of developmental disorders. This was done by looking at the differences in copy number variant calls, segment size differences and gene content between the two microarrays. The cost to run each type of microarray test for diagnosis was also compared. The copy number variant calls made by the two platforms were comparable. The SNP-based array did provide additional information that would be useful in molecular diagnosis. The final recommendation was for the CGH array and the SNP-based array to be used interchangeably based on clinical requests at the Division of Human Genetics.

Acknowledgements

I would like to thank my supervisors for all their support and guidance this year. Thank you for your guidance in all my research assessments this year. To Nkateko Mayevu, thank you for all the time you put in towards my learning about microarrays. Thank you for your continued patience as I navigated my way in writing and reporting. Thank you for looking at my work even after work hours. Thank you for your support and help in my presentations for my coursework. Your reassurances throughout the year were one of the reasons I did not give up. To Dr Fiona Baine – Savanhu, thank you for all your guidance with my research this year and thank you for showing me and directing me towards good writing that I hope to achieve as I continue on my research journey - slowly but surely. It has been as honour to receive guidance and support from you, and I pursue my research journey.

Thank you to Ms Mirriam Hlongwane, Mr Daniel Nkwane and Mr Naseem Cassim for all their help in developing a costing tool for SNP-array testing. I appreciated the invaluable time you put in.

Thank you to Dr Robyn Kerr who helped with putting together the research ethics application.

Thank you to Prof Zane Lombard and Dr Khuthala Mnika for their support as course coordinators for this programme.

Thank you to my parents, Qondile and Friction Mayisela who funded my studies, and for their support and love throughout the journey. Thank you to my friends who always encouraged me to keep pressing on. Most importantly, thank you to God, for the strength to keep going.

Table of Contents

Declaration ii
Contribution of the candidate to the paper iii
Dedication v
Abstract vi
Acknowledgements vii
List of Tables 1
List of Figures 2
List of Abbreviations 3
List of Appendices 4
Abstract for paper submission 9
Introduction 10
Materials and methods 13
Sample/Patient selection 13
CytoScan® Optima Array 13
Chromosome Analysis Suite (ChAS) v4.3 14
Copy number variant calling 14
Genotype-phenotype correlation of CytoScan Optima results 15
Costing 16
Results 17
Copy number variants detected by the CytoScan Optima SNP array 17
Comparing CGH array and SNP-based array 18
Genotype-phenotype correlation 21
Patient 3 23
Aneuploidies 24

Mosaicism 24
Mechanisms underlying CNVs 25
Loss of Heterozygosity 26
Costing 29
Discussion 30
Limitations 36
Conclusion 37
References 38
Appendices 46
Appendix A. Approved research protocol 46
Appendix B. Ethics clearance 75
Appendix C. Turn-it-in report 76
Appendix D. Author guidelines for the Journal of Molecular Diagnostics 80
Appendix E. Exchange rates indicators for South Africa 99

List of Tables

Table 1. Genomic breakpoints. Size and associated CNV syndromes detected by CytoScan Optima array and SurePrint ISCA array……20

Table 2. CNV syndromes confirmed in the sample of patients in this study……22

Table 3. List of rearrangements and the inferred mechanisms.....26

Table 4. List of LOH segments greater than 8 Mb identified in the patient sample….26

Table 5. The cost of running a SNP-based array test ……29

Table 6. The cost of running a CGH array test ……29

List of Figures

Figure 1. Screenshot from the chromosome analysis suite v4.3 showing the copy number graph data, the log2 and smooth signal of chromosome 5……18

Figure 2. Screenshot of chromosome analysis suite v4.3 showing a CNV loss at chromosome 18p11.31……25

Figure 3. Screenshot of chromosomal analysis suite v4.3 showing the LOH (segment), allele difference, LOH and BAF tracks of chromosome 6q33.3q37.3……...27

Figure 4 Screenshot of chromosome analysis suite v4.3 showing the LOH (segment), allele difference, LOH and BAF tracks of chromosome 22……...28

Figure 5. Screenshot of chromosome analysis suite v4.3 showing the LOH (segment), allele difference, LOH and BAF tracks of chromosome X in Patient 1……28

List of Abbreviations

CGH - Comparative genomic hybridization

ChAS - Chromosomal analysis suite

ClinGen – The Clinical Genome Resource

CMA - Chromosomal microarray analysis

CNV - Copy number variant

DECIPHER – Database of chromosomal imbalance and Phenotype in Humans Using Ensembl Resources

FISH - Fluorescence in-situ hybridization

HMM – Hidden Markov Model

ISCA - International Standards for Cytogenomic Arrays

LOH - Loss of heterozygosity

MLPA - Multiplex ligation-dependent probe amplification

NHLS - National Health Laboratory Service

OMIM – Online Mendelian Inheritance of Man

SV – Structural variants

UPD - Uniparental disomy

List of Appendices

Appendix A. Approved research protocol……46
Appendix B. Ethics clearance for sub study M220295……75
Appendix C. Turn-it-in-report……..……..76
Appendix D. Author guidelines for The Journal of Molecular Diagnostics……...80
Appendix E. Exchange Rate Indicators for South Africa…….……99

Research report in the format of a “submissible paper”

This work has been written up as a paper for submission to The Journal of Molecular Diagnostics. The author guidelines such as abstract, referencing and word count have been adhered to and can be found in Appendix D. Tables and Figures have been included in the text for ease of reference.

Evaluating the clinical utility of SNP-based microarray platform – a comparative study

Minenhle Mayisela[1] *Nkateko Mayevu*[2] *Fiona Baine-Savanhu*[1, 2]

[1] Division of Human Genetics, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[2] Division of Human Genetics, National Health Laboratory Service, Corner Hospital and De Korte Streets, Johannesburg, South Africa

Minenhle Mayisela
Email: 1880919@students.ac.za
Contact number: +27 78 788 8219

Nkateko Mayevu
Email: nkateko.mayevu@nhls.ac.za
Contact number: +27 11 489 9237

Fiona Baine-Savanhu
Email: fiona.baine-savanhu@nhls.ac.za
Contact number: +27 11 489 9158

Abstract for paper submission

Chromosomal microarray analysis (CMA) is the first line testing for individuals with developmental disorders and congenital anomalies. The evolution of cytogenetic techniques has brought about improvements in resolution and the turnaround time for genetic testing and disease diagnosis. The National Health Laboratory Service in Johannesburg, South Africa offers a testing service for South African patients with developmental disorders and congenital abnormalities and currently uses a Comparative Genomic Hybridization (CGH) array (SurePrint G3 Unrestricted CGH ISCA v2, 8x60 kb microarray). This study aimed to compare the clinical utility of a Single Nucleotide Polymorphism (SNP) array (CytoScan® Optima array) to a CGH array (SurePrint ISCA microarray) in the diagnosis of patients referred for microarray testing at the Division of Human Genetics, National Health Laboratory Service in Johannesburg, South Africa. This was done by looking at the differences in copy number variant calls, segment size differences and gene content between the two microarrays. The cost of running each type of microarray was also compared. The results of the study show that copy number variant calls made by the two platforms were comparable. The different microarrays were able to detect genomic losses and gains, including aneuploidies, in 17 samples. The SNP-based array provided additional SNP genotype information that would be useful in molecular diagnosis. The SNP-based microarray and the CGH array currently in use in the division may be used interchangeably as this comparison yielded similar results.

Introduction

Copy number variants (CNVs) are DNA segments greater than 50 bp in size that show a variable copy number compared to the reference genome[1]. Large CNVs that encompass dosage sensitive genes are likely to be disease causing[2]. Copy number variants play a role in the aetiology of developmental disorders, intellectual disability and congenital anomalies[3]. Cytogenetic techniques have long been used to identify and analyse structural and numerical rearrangements of chromosomes that cause losses and gains in genetic material[4]. The evolution of cytogenetic techniques has brought about improvements in resolution and the turnaround time for genetic testing and disease diagnosis.

Karyotyping allows the viewing of all chromosomes under a microscope and enables the detection of changes in chromosome number (aneuploidy), and large structural chromosome abnormalities by analysis of banding patterns[5]. It is still considered the gold standard for confirming and informing the mechanisms that underlie the gains and losses of genomic material in many chromosomal aberrations. Karyotyping is able to detect balanced and unbalanced chromosomal rearrangements because it uses banding patterns to visualise different regions of the chromosomes[5]. The technique is however limited to detecting CNVs that are greater than 5 Mb making it unsuitable for submicroscopic aberrations[5].

Fluorescence in-situ hybridization (FISH) has a higher resolution of between 100-200 kb compared to karyotyping[6]. This method uses a fluorescently labelled DNA probe to hybridize to its complementary segment on chromosomes and is visualised with a fluorescent microscope. Its limitation, however, is that the FISH probes are designed for specific locus on the chromosome.

This means that a disorder must be suspected for FISH to be used since one probe is used for one chromosome abnormality. Compared to FISH, multiplex ligation-dependent probe amplification (MLPA) has the advantage of simultaneously targeting multiple regions across the genome by quantifying dosage of DNA regions based on directed ligation[7]. Although MLPA probes have multiple targets, they are designed for known submicroscopic abnormalities related to specific diseases, meaning only conditions of known aetiology can be tested using this technique[8].

Chromosomal microarray analysis (CMA) has become a useful approach in the molecular diagnosis of developmental delay and congenital anomalies of unknown aetiology. This has led to a consensus by the International Standards for Cytogenomic Arrays (ISCA) Consortium recommending CMA as a first line test for the diagnosis of patients with developmental disorders and/or congenital anomalies[9]. Alternative methods of detecting chromosomal aberrations such as karyotyping, FISH and MLPA are limited by low resolution, low diagnostic yields and longer turnaround time for testing. Microarrays may have a resolution as high as 16 kb[10] depending on the probe number, spacing and distribution of its design[11]. A study undertaken to diagnose structural aberrations using CMA, gave a diagnostic yield of 15-20% compared to the 3% yield from karyotyping[9].

Microarrays use deoxyribonucleic acid-based methods to detect CNVs as well as the size of the aberrations present caused by unbalanced chromosomal rearrangements (submicroscopic) and aneuploidies[9]. There are two types that are commonly used: comparative genomic hybridization (CGH) array and single nucleotide polymorphism (SNP) microarrays. The CGH array, referred to as a true comparative hybridization array, directly compares the DNA from a test sample to a normal reference sample[11]. The microarrays emit a fluorescent light that is converted to a log2 ratio value. The copy number changes of this array are represented graphically as a log2 ratio which informs the type of copy number change i.e., a gain or a loss.

The SNP microarray on the other hand, only requires the test sample and no reference[12]. This array functions by directly hybridizing the test DNA sample to a microarray slide that contains probes designed to be complementary to target DNA segments flanking SNPs[13]. The results of the array show the gain or loss of DNA segments as well as the different SNP genotypes for each SNP locus[13]. Since SNP arrays provide genotyping information, they can be used to detect loss of heterozygosity (LOH) in genomic regions and can also detect low levels of mosaicism[14].

The CGH array test was implemented in 2020 as the first line testing strategy for patients presenting with developmental disorders and congenital abnormalities at the Division of Human Genetics at the National Health Laboratory Service (NHLS), a state diagnostic laboratory in the Republic of South Africa. The CGH array used is the SurePrint G3 Unrestricted CGH ISCA v2, 8x60K microarray (Agilent technologies, Santa Clara, CA). For the initial testing and validation of the system, the CytoScan® Optima array (Thermo Fisher Scientific, Waltham, MA), a low-resolution microarray designed specifically for prenatal testing was included.

This project aimed to compare the detection capacity, ease of use and cost of the CytoScan® Optima array to the SurePrint CGH ISCA, 8x60K array currently used for diagnostic testing of postnatal samples in the Division. This was done by testing 17 samples with known CNVs detected by the CGH array. The CytoScan® Optima contains both copy number and SNP probes and thus provides SNP genotype which may be used to detect LOH. SNP arrays have also been shown to perform better in the detection of low-level mosaicism.

Materials and methods

Sample/Patient selection

This comparative study used residual DNA samples from patients referred to the Division of Human Genetics diagnostic laboratory for CMA testing. Microarray testing records were reviewed and samples with a range of chromosomal abnormalities previously detected using the CGH array were selected. The patients had clinical features such as intellectual disability, dysmorphic features and a variety of congenital anomalies. The total number of samples investigated was 17 and all samples were tested using the Cytoscan® Optima array. The kit used has the capacity to test 24 samples. The 17 samples analysed here were selected from the 24, to represent a range of chromosomal abnormalities (including structural variants and copy number variants). The data previously generated from the SurePrint G3 Unrestricted CGH ISCA v2, 8x60 kb microarray were compared with data from the Cytoscan® Optima array. The SurePrint ISCA array has a probe-spacing of approximately 60 kb and analyses eight samples on a single slide. The study was granted ethics approval by the Human Research Ethics Committee (Medical) of the University of Witwatersrand, certificate number: M220295.

CytoScan® Optima Array

The Cytoscan® Optima Array (Thermo Fisher Scientific, Waltham, MA) was used to generate copy number data and SNP genotype data. This is an Affymetrix SNP-based array that consists of 18 018 copy number probes to detect any gains or losses as well as 148 450 SNP probes which target SNP sites at different polymorphic locations[15]. For comparison, the SurePrint ISCA array (Agilent technologies, Santa Clara, CA) has 60 000 copy number probes concentrated at genomic regions of clinical relevance as determined by the ISCA consortium. The laboratory workflow for the CytoScan Optima array (including hybridization, washing and scanning) was performed by staff members at the Division according to the manufacturer's protocol for the 17 selected samples. The GeneChip® Scanner 3000 7G (Thermo Fisher Scientific, Waltham, MA) generates CEL files. CEL files are files that summarise probe intensity signals from the scanner. These raw data CEL files were provided to the candidate and analysed using Chromosome Analysis Suite (ChAS) software v4.3 (Thermo Fisher Scientific, Waltham, MA).

Chromosome Analysis Suite (ChAS) v4.3

The 17 CEL files were analysed using ChAS v4.3 software (Thermo Fisher Scientific, Waltham, MA). This a cytogenetics analysis software that analyses chromosomal changes across the genome. It automatically converts CEL files into cychp files which contain information on copy number, LOH, mosaicism and genotype calls. Before further analysis, the cychp files must meet all the quality control metrics: median absolute pairwise >/= 0.25, waviness </= 0.12 and SNP quality control >/= 15. The software uses computational statistical models to determine copy number state and SNP genotype for each CNV call. The algorithms used for copy number state and SNP genotyping are the Hidden Markov Model (HMM) and the Bayesian Robust model respectively. Normalised log2 ratio (weighted log2 ratio) for each probe is plotted and the copy number determined from the graph using the Bayesian HMM algorithm.

Copy number variant calling

The information provided by the copy number probes and the SNP probes is displayed on different tracks in the ChAS browser (Thermo Fisher Scientific, Waltham, MA). The log2 ratio and the copy number state tracks show information obtained from the copy number probes. The B allele frequency, allele difference and smooth signal tracks show the information provided by the SNP probes.

The CNVs are shown as gains or losses based on the log2ratio displayed by the tracks on the graph in the ChAS browser. A normal diploid copy number has a log2 ratio of 0. The ratio used to calculate a gain is 3/2 therefore a single copy gain has a log2 ratio of 0.58 and the ratio used to calculate a loss is 1/2 therefore a single copy loss has a log2 ratio of -1.

SNP genotype information is derived from the allele difference and B allele frequency given by the ChAS software. The alleles are designated A and B, and each assigned a value of 0.5. The allele difference is the difference between the number of two alleles using the formula A – B. The B allele frequency is the number of B alleles divided by the total number of allele combinations of a SNP (AA, BB or AB), and is calculated using the formula B/(A+B) is used to calculate it. The smooth signal is a smoothened calibrated copy number estimate. Genomic imbalances were annotated

against the primary genome assembly, Genome Reference Consortium Human Build 37 (GRCh 37). In this study, the software was set to call CNVs that are greater than 200 kb and LOH regions greater than 8 Mb in size. It is recommended that larger size thresholds provide more sensitivity and reduce false positive rates[16].

Genotype-phenotype correlation of CytoScan Optima results

The patients' clinical features were made available to ensure the molecular finding could be associated with the phenotype i.e., genotype-phenotype correlation. The public databases embedded in the ChAS software provide information on the associated CNV syndrome if any, pathogenicity, deleterious genes contained, dosage sensitivity of the genes, and existing data on patients with the same/similar CNV. The databases used are Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER)[17], Clinical Genome Resource (ClinGen)[18], ClinVar[19,20] and Online Mendelian Inheritance of Man (OMIM)[21].

DECIPHER is a clinical research tool that helps in interpreting data from CNV analysis. DECIPHER facilitates the sharing of information between clinical researchers across the world to enable progress in defining new syndromes and gene function[17] and is accessible here: https://www.deciphergenomics.org/browser). ClinVar is a public archive of human variants and the interpretation of their relationship to disease and other conditions. Many different organisations, including laboratories and patient registries, submit variants to ClinVar for public access. Submissions include a description of CNV syndrome for which variant was interpreted as well as the clinical significance of the variant[19,20]. ClinVar is accessible here: https://www.ncbi.nlm.nih.gov/clinvar/). Clinical Genome Resource (ClinGen) is a curated collection of data on copy number variants and sequence variants derived from publications of case studies looking at CNV syndromes. ClinGen provides an interconnected community for resources to improve understanding of genomic variants and improve its use in clinical care[18]. ClinGen is accessible here: https://clinicalgenome.org/).

Online Mendelian Inheritance of Man (OMIM) is a primary repository of information on genes and genetic phenotypes derived from clinical studies. OMIM synthesizes

and summarises new and important information based on expert review of biomedical literature. OMIM also plays a role in naming and classifying genetic phenotypes[22] (https://www.omim.org/). Altogether, these databases contain genetic and phenotypic information of variation in the genome. They allow for storage, sharing and comparison of genetic data across institutions and between researchers.

Costing

The cost of performing the SNP array test was determined using the NHLS costing tool. This tool is designed to standardise the costing of diagnostic tests within the organisation. The calculation considered the cost of the array reagents, consumables, labour and fixed equipment. This price was then compared to that of the CGH array already in use in the Division.

Results

Copy number variants detected by the CytoScan Optima SNP array

In the sample of 17 patients, 24 genomic copy number variants were observed by the CytoScan Optima SNP array alone. This included 12 losses and 12 gains. The genomic coordinates of the copy number variants for each sample tested can be found in Table 1. Some of the CNVs detected are associated with known syndromes and most are suspected to have occurred de novo.

The copy number state and SNP genotype information for each software call were visualised using the ChAS v3 browser. Figure 1 shows the different tracks that can be seen on the ChAS browser and the information it provides. The copy number state represents the copy number variation in the target genome relative to the reference genome. It is displayed as a log2 ratio based on the intensity of the probe represented by that genomic region. The SNP genotype information is represented by calculation of the allele difference and B allele frequency, as well as smooth signal.

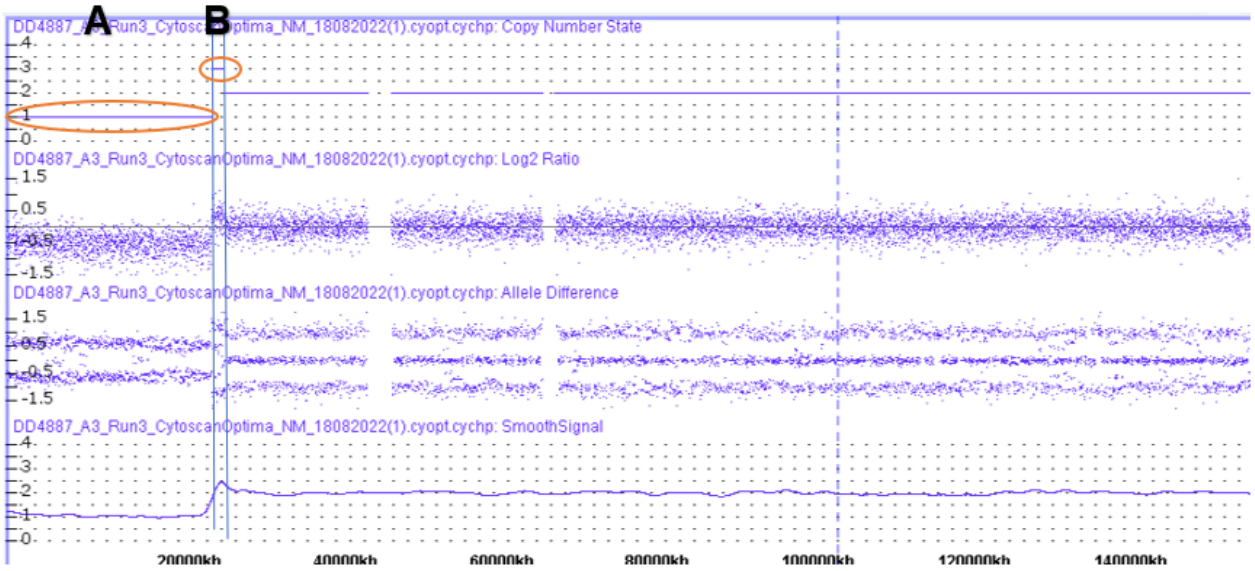


Figure 1. Screenshot from the chromosome analysis suite v4.3 showing the copy number graph data, the log2 and smooth signal of chromosome 5. The graph shows chromosome 5 with a 26 Mb terminal loss adjacent to a 1 kb gain on Patient 14. **A)** A copy number loss. Copy number state is 1.0 which indicates a loss of that genomic region. The loss is indicated by the red circle. The allele difference shows two tracks only indicating that there are two genotypes for that segment indicating a loss. **B)** A copy number gain on Patient 14. Copy number state shows 3.0 which indicates a gain of that genomic region. The gain is indicated by a red circle. The allele difference shows four tracks indicating that there are four genotypes for that segment indicating a gain.

Comparing CGH array and SNP-based array

The SNP array (CytoScan Optima) used in this study detected 24 CNVs. Of these, 23 were previously detected by the SurePrint ISCA CGH array used for diagnostic testing. Table 1 presents the genomic breakpoints for each CNV call. The genomic coordinates of the copy number changes were compared and found to be similar between the two arrays for a majority of the patients. The OMIM database, DECIPHER genome browser and ClinGen were used to look at the gene content of the CNVs detected. The coordinates of each CNV detected were entered into the DECIPHER genome browser and gene content assessed linked were provided directly to the OMIM database for genes known to be associated with mendelian disorders.

The SNP array detected a 981.958 kb gain of 15q26.2 in Patient 10 that the CGH array did not detect. Upon looking up this CNV on the public databases, it was found to not be clinically significant and therefore was not the molecular cause of the patient's phenotype.

Table 1. Genomic breakpoints. Size (kb) and associated CNV syndromes detected by CytoScan Optima array and SurePrint ISCA array.

Patient	CytoScan Optima array results			SurePrint ISCA array results			
	Breakpoint coordinates	Copy number state	Segment size (kbs	Breakpoint coordinates	Segment size	Size difference	Associated CNV syndrome
1	arr[GRCh37] 11p14.3p13(24385583_33146139)x1	1	8760.556	arr[GRCh37] 11p14.3p13(24381118_33162974)x1	8781.856	21.3	WAGR syndrome - 11p12del
2	arr[GRCh37] 3q13.2q22.3(111612191_136676525)x3	3	25064.334	arr[GRCh37] 3q13.2q22.3(111848681_136983572)x3	25134.891	70.557	
3	arr[GRCh37] 13q11q12.11(19436287_31787356)x1	1	12351.069	arr[GRCh37] 13q12.11(20407295_21861250)x1	1453.955	10897.114	
3	arr[GRCh37] 18p11.31(3412357_5191958)x1	1	1779.601	arr[GRCh37] 18p11.31(3415823_5195070)x1	1779.247	0.354	
3	negative			arr[GRCh37] 18p11.32p11.31(236820_3414693)x1	3177.873	3177.873	
4	arr[GRCh37] 2q26q27 (208150499_242783384)x1	1	34632.885	arr[GRCh37] 2q33.3q37.3(208174860_243068396)x3	34893.536	260.651	
4	arr[GRCh37] 6q26q27 (161069941_170919482)x1	1	9849.541	arr[GRCh37] 6q26q27(161176622_170921089)x1	9744.467	105.074	
5	arr[GRCh37] 8q21.11q22.1(75661636_94779028)x1	1	19117.392	arr[GRCh37] 8q21.11q22.1(75624070_94892477)x1	19268.407		8q21.11 microdeletion
6	arr[GRCh37] 22q11.21(18166423_21108369)x1	1	2941.946	arr[GRCh37] 22q11.21(18917031_21443283)x1		2941.946	22q.11.2 microdeletion
7	arr[GRCh37] 7p22.3p11.1(43361_58006205)x3	3	57962.844	arr[GRCh37] 7p22.3p11.2(54185_56174888)x3	56120.703	1842.141	7p trisomy
8	arr[GRCh37] 5p15.33p14.1(113577_27953367)x1	1	27839.79	arr[GRCh37] 5p15.33p14.1(22149_27908206)x1	27886.057	46.267	Cri Du Chat syndrome
9	arr[GRCh37] 13q11q34(19436287_115107733)x3	3	95671.446	arr[GRCh37] 13q12.11q34(20407295_115092648)x3	94685.353	986.093	13q trisomy
10	arr[GRCh37] 15q26.2(95721415_96703373)x3	3	981.958	negative		981.958	
10	arr[GRCh37] 16p12.2p11.2(21620626_29379768)x3	3	7759.142	arr[GRCh37] 16p12.2p11.2(21475060_29331350)x3	7856.29	97.148	16p11.2p12.2 duplication
11	arr[GRCh37] 4q12q35.2(58793349_190957473)x3	3	132164.124	arr[GRCh37] 4q12q35.2(58881742_190896674)x3	132014.932	149.192	
11	arr[GRCh37] Xq28(147973535_155233731)x1	1	7260.196	arr[GRCh37] Xq28 (147991745_155232907)x1	7241.162	19.034	
12	arr[GRCh37] 6q27(167711355_170610394)x1	1	2899.039	arr[GRCh37] 6q27(168364343_170906796)x1	2542.453	356.586	
12	arr[GRCh37] 17p13.3p13.2(150733_4116672)x3	3	3965.939	arr[GRCh37] 17p13.3p13.2(11807_4045078)x3	4033.271	67.332	
13	arr[GRCh37] 17q12(34831963_36350005)x3	3	1518.042	arr[GRCh37] 17q12(34611352_36248918)x3	1637.566	119.524	17q12 duplication syndrome
14	arr[GRCh37] 5p15.33p14.1(113577_26368797)x1	1	26255.22	arr[GRCh37] 5p15.33p14.1(22149_26287349)x1	26265.2	9.98	
14	arr[GRCh37] 5p14.1(26380734_27858577)x3	3	1477.843	arr[GRCh37] 5p14.1(26422464_28052827)x4	1630.363	152.52	Cri Du Chat syndrome
15	arr[GRCh37] Xq13.1q21.1(70902832_83032256)x2	2	12128.424	arr[GRCh37] Xq13.1q21.1 (71030284_83015968)x3	11985.684	142.74	
16	arr[GRCh37] 1q44(246267566_249224684)x3	3	2957.118	arr[GRCh37] 1q44(246081208_249212668)x3	3131.46	174.342	
16	arr[GRCh37] 10q25.1q26.3(109292731_132504534)x3	3	23211.803	arr[GRCh37] 10q25.1q26.3(109312155_132468422)x	22156.267	1055.536	
16	arr[GRCh37] 10q26.3(132516104_135427143)x1	1	2911.039	arr[GRCh37] 10q26.3(132580596_135434178)x1	2853.583	57.456	
17	negative			negative			

Genotype-phenotype correlation

The clinical features of the patients were provided to match the molecular diagnoses to the phenotype. Of the 17 patients, six were found to have CNVs were associated with known microduplication/deletion syndromes. Microdeletion syndromes were confirmed in four patients: Patient 1 – Wilms tumour, aniridia, genitourinary anomalies and mental retardation (WAGR) syndrome (OMIM: 194072)[23,24] from a 8.7 Mb interstitial deletion, Patient 5 - 8q21.11 microdeletion syndrome (OMIM: 614230)[25] from a 19.1 Mb interstitial deletion. Patient 6 - 22q11.2 microdeletion syndrome (OMIM: 611867)[26,27] from a 2.9 Mb interstitial deletion, Patient 8 – Cri Du Chat syndrome (OMIM: 123450)[28–30] from a 27.8 Mb terminal deletion and Patient 14 – Cri Du Chat syndrome (OMIM: 123450)[28–30] from a 1.5 Mb terminal deletion. Microduplication syndromes were confirmed in two patients: Patient 10 - 16p11.2 duplication syndrome (OMIM: 614671)[31] from a 7.8 Mb gain and Patient 12 – 17q12 duplication syndrome (OMIM: 614562)[32] from a 3.9 Mb kb gain. The clinical features for each of these patients matched the phenotype of the confirmed disorder, as shown in Table 2.

For five of the patients: Patient 3, Patient 4, Patient 11, Patient 15 and Patient 16, there were no syndromes associated with the CNVs detected. The CNVs detected in these patients had no clinical significance.

Table 2. CNV syndromes confirmed in the sample of patients in this study.

Patient	CNV syndrome	Clinical features of the patient	Features of syndrome	Reference
1	WAGR 11p13p13 syndrome (OMIM: 194072)	overweight-for-age, developmental delay, dysmorphic features, hypotonia, cryptorchidism,aniridia, behavioural problems	Wilms tumour, Aniridia, genital anomalies, mental retardation, cryptorchidism	23,24
5	8q21.11 microdeletion syndrome (OMIM: 614230)	severe-profound developmental delay, self-injurious behaviour, bilateral eye abnormalities ((R) buphthalmos and (L) microphthalmia), dysmorphic features, craniosynostosis	Intellectual disability, dysmorphic features, camptodactyly, syndactyly, hypotonia, sensoneural hearing loss	25
6	22q11.21 microdeletion syndrome (OMIM: 611867)	premature baby with tracheoesophageal fistula, cyanotic heart with common atrium, pulmonary atresia, abnormally shaped low set ears, long fingers and toes, intermediate fontanelle	Cardiac anomalies, immunodeficiency, speech delay, hypocalcaemia, neurological and developmental behavioural issues	26,27
8	Cri Du Chat syndrome (OMIM: 123450)	global developmental delay, failure to thrive, microcephaly, dysmorphic features (including strabismus, microstomia and marked microretrognathia), chronic upper airway obstruction, high-pitched cry	Microcephaly, round face, hypotonia, psychomotor symptoms, macrostomia, mental retardation, high pitched cry	28–30
10	16p11.2 Microduplication (OMIM: 614671)	postnatal short stature, moderate-to-severe developmental delay, dysmorphic features, limb abnormalities	Speech delay, microcephaly, autism, epilepsy, diaphragmatic hernia, smooth philtrum, small ears, developmental delay	31
13	17p13.3 duplication syndrome (OMIM: 614562)	prenatal short stature, microcephaly, mild-moderate developmental delay, subtle dysmorphic features, CNS abnormalities, limb abnormalities	Hypotonia, dysmorphic features, failure to thrive, poor growth, brain abnormalities	32
14	Cri Du Chat syndrome (OMIM: 123450)	overweight-for-age, microcephaly (postnatal onset), dysmorphic features, cryptorchidism, structural brain abnormalities (hydrocephalus and Dandy-Walker variant), and mild developmental delay	Microcephaly, round face, hypotonia, psychomotor symptoms, macrostomia, mental retardation, high pitched cry	28–30

Patient 3

Several findings from the results from Patient 3 aided in the distinguishing of the SNP array and the CGH array. These were associated with CNV calls, genomic coordinates and size breakpoints, gene content from CNVs, and genotype-phenotype correlation.

The CGH array detected a loss of 18p11.32p11.31 in Patient 3 that the SNP array did not detect. This CNV was considered clinically significant as it included the OMIM morbid gene *SMCHD1* (OMIM :614982). Haploinsufficiency of *SMCHD1* is associated with Bosma arrhinia microphthalmia with coloboma, arhinia and endocrine findings [33]. This patient presented with the following overlapping clinical feature: choanal atresia – a narrowing of the nasal passageway, which is feature of Bosma arrhinia microphthalmia syndrome.

There was a notable difference for Patient 3 who was found to have a loss of 13q12.11 detected by both arrays. The genomic coordinates determined by the SNP array were 13:19436287-31787356; while CGH array coordinates were 13:20407295-21661250. This equated to a CNV of 12351 kb on SNP array compared to 1454 kb on CGH array; a difference of 10897 kb. Table 1 shows the genomic coordinates and the sizes of all CNVs detected in the sample patients by the CGH array and SNP array. The observed differences in genomic coordinates and CNV size are indication that the SNP array and the CGH array differ in probe design and placement across the genomic regions they target.

A difference in gene content can be observed in instances where both arrays detect the CNV but because of probe placement differences and distribution, one array includes a gene or part of a gene that is not shown by the other microarray.

Patient 3 had a CNV loss of 18p11.32p11.31 detected by the CGH array copy number probes that the SNP array copy number probes did not detect. This may be because the SNP arrays copy number probes may not have covered that region. However, the smooth signal and BAF tracks derived from SNP probes from the SNP

array showed evidence of this loss. The gene(s) contained in the CNV included the OMIM morbid gene *SMCHD1.*

Aneuploidies

Two of the 17 patients were confirmed to have trisomies: Patient 9 had a mosaic partial trisomy 13q, shown by a gain of the entire q arm (95 Mb) at approximately 63% mosaicism and Patient 7 had partial trisomy 7p, shown by a gain of the entire p arm (57 Mb). Partial trisomies can result from a gain in relevant large portion of a chromosome[34] due to chromosomal rearrangements. Both the SNP array and the CGH array detected these large chromosomal changes and estimation of mosaicism for Patient 9 was comparable between the two and is addressed below.

Mosaicism

Patient 9 was found to have mosaic trisomy 13 at approximately 63% based on the CytoScan Optima SNP array data on chromosome 13. The ChAS software is able to estimate the level of mosaicism by averaging the smooth signal values of the CNV detected. The SurePrint ISCA array detected the gain 13q; however, the level of mosaicism has to be manually estimated by working backwards from the log2 ratio. Patient 9 was diagnosed with trisomy 13, in approximately 60% of cells. This is comparable to the automated software estimate of 63% made by ChAS.

In order to investigate the ability of the SNP array to detect mosaicism, Patient 17 was included in this study because of the reported clinical phenotype (skin pigmentary abnormalities) which is strongly suggestive of mosaicism. SNP arrays are more reliable at detecting mosaicism because they give SNP genotype information unlike the CGH array. Like the CGH array, the CytoScan® Optima array used in this study was unable to pick up mosaicism in this patient. Patient 17 was therefore reported to have no clinically significant CNVs detected by the microarrays.

Patient 3 had been previously diagnosed with a loss of 18p11.32 using the SurePrint (Agilent Technologies, Santa Clara, CA). This deletion was no detected by the CytoScan® using copy number probes. The SNP probes however, showed this as a potentially mosaic loss. In Figure 2, there is no copy number variant call made by the software (see copy number state track at the top); however, the smooth signal track

is lower than 2 and the BAF track shows a split suggesting the presence of a mosaic loss. The 18p11.32 (1.7 Mb) deletion was estimated to be approximately 59% mosaic.

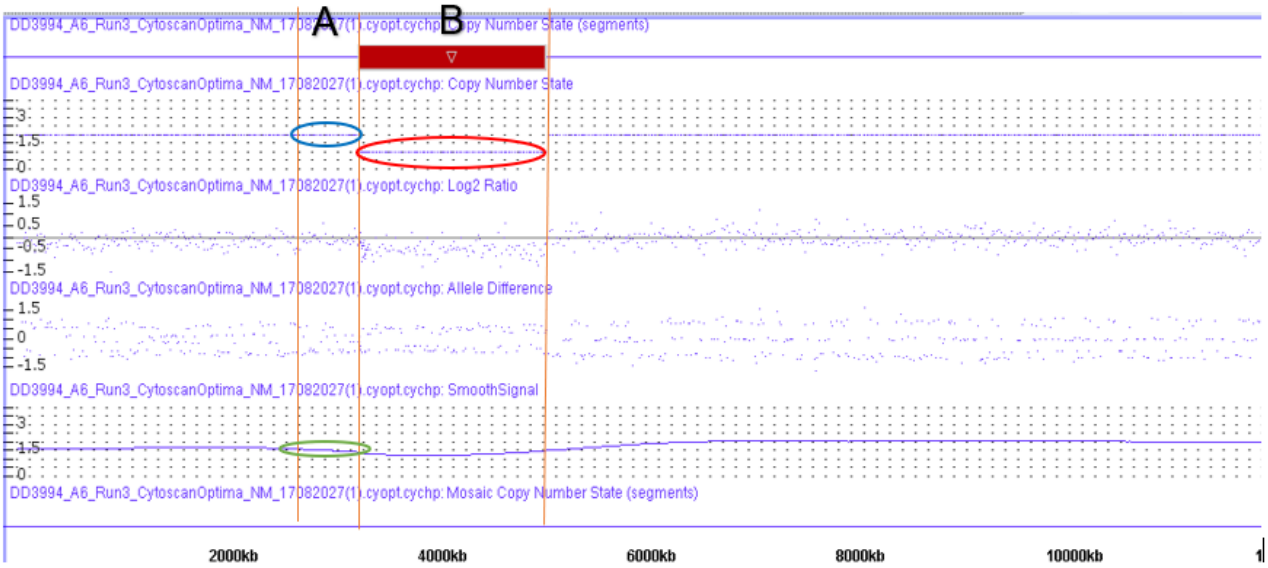


Figure 2. Screenshot of chromosome analysis suite v4.3 showing a CNV loss at chromosome 18p11.31. **A)** The copy number state indicates a normal number of copies for the region (blue circle) however the smooth signal is lower than 2 (green circle) indicating the possibility of a mosaic loss in that region **B)** The copy number state of 1.0 for that segment indicates a loss of that genomic region (red circle). The allele difference track has two tracks instead of three, indicating that there are two possible genotypes for that segment i.e., a loss of genomic material.

Mechanisms underlying CNVs

The mechanisms underlying the CNVs detected by both the SNP array and the CGH array can be inferred from visualising the CNVs in the respective software. This applies to patients with more than one CNV e.g., a terminal loss on one chromosome and a terminal gain on another chromosome is suggestive of an unbalanced translocation. Table 3 presents the list of patients with multiple CNVs and their inferred mechanism.

Table 3. List of rearrangements and the inferred mechanisms.

Patient	CNVs	Inferred mechanism
4	2q26q27 terminal loss 6q26q27 terminal loss	Unbalanced translocation
11	4q12q35.2 terminal gain Xq28 terminal loss	Unbalanced translocation
12	6q27 terminal loss 17p13.3p13.2 terminal gain	Unbalanced translocation
14	5p15.33p14.1 terminal loss 5p14.1 terminal gain	Inversion
16	1q44 terminal gain 10q25.1q26.3 interstitial gain 10q26.3 terminal loss	Complex rearrangement

Loss of Heterozygosity

The SNP array was able to detect loss of heterozygosity greater than 8 Mb based on filter settings. The LOH was visualised using the ChAS browser. The LOH regions are determined by the SNP genotype information: BAF and allele difference displayed as tracks in the ChAS browser which can be seen in Figures 5, 6 and 7. Nine of the 17 patients, had LOH regions larger than 8 Mb as presented on Table 4.

Table 4. List of LOH segments greater than 8 Mb identified in the patient sample.

Patient	Sex chromosomes	LOH segment	Corresponding deletion
1	Male	Xp22.33p11.1 X11.1q28	-
4	Female	6q25.3q27	6q26q27
8	Female	5p15.33p14.1	5p15.33p14.1
10	Male	Xp22.33p11.1 X11.1q28	-
12	Female	Xp11.23p11.1	-
13	Male	Xp22.33p11.1 X11.1q28	-
14	Male	5p15.33	5p14.1
		Xp22.33p11.1 X11.1q28	-
16		Xp22.33p11.1 X11.1q28	-
17	Female	22q11.23q12.3	-

For three of the patients, the LOH regions were associated with a deletion in the same region. The LOH tracks on the browser indicated LOH was present, however the allele difference track did not support LOH, but rather a deletion (Figure 3). The allele difference indicates only indicates the presence of A and B genotypes, which indicates a loss.

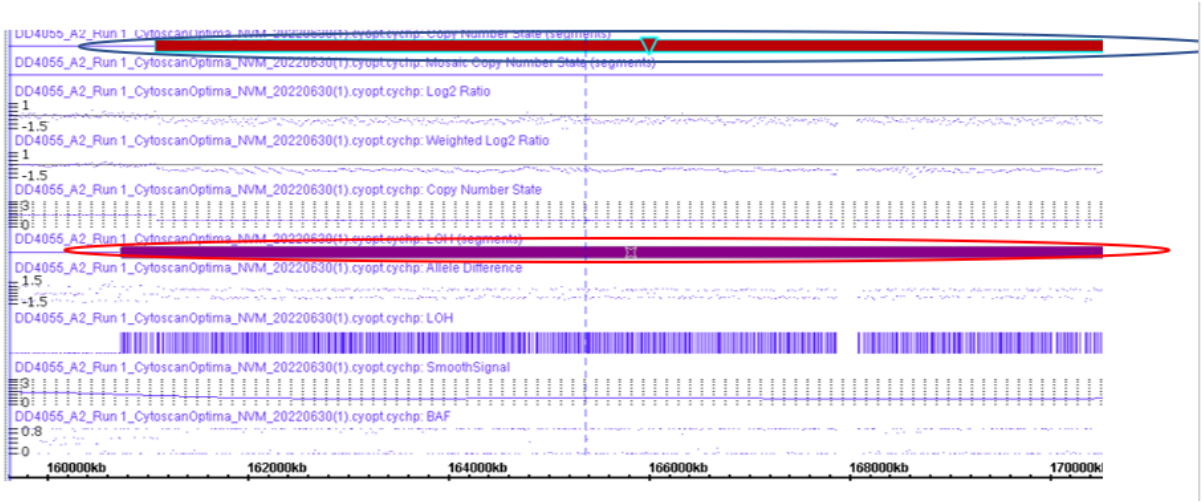


Figure 3. Screenshot of chromosomal analysis suite v4.3 showing the LOH (segment), allele difference, LOH and BAF tracks of chromosome 6q33.3q37.3. The LOH segment (red circle) track show that the LOH segments correspond to a loss (blue circle) in the same region at chromosome 6q33.3q37.3. The LOH tracks indicate a presence of LOH in the region however, the allele difference tracks do not support LOH.

Of the nine patients found to have LOH, only three patients had regions greater than 8 Mb that were associated with a deletion. Figure 4 is a visual representation of a segment of chromosome 22 of Patient 17. On the browser, both the LOH and allele difference track indicate a LOH on chromosome 22q11.23. The allele difference tracks show AA and BB genotypes only, indicating there is neither loss nor gain therefore there is loss of heterozygosity in that region. It is important to note that the CGH array is unable to detect LOH therefore the CGH test result was negative for any abnormalities for this patient. Unfortunately, there is no published information on the effect of LOH on the 22q11.23 region therefore LOH studies on 22q11.23 need to be done. Patient 12 also had an LOH region at chromosome Xp11.23p11.1 that was not associated with a deletion and had AA and BB genotypes.

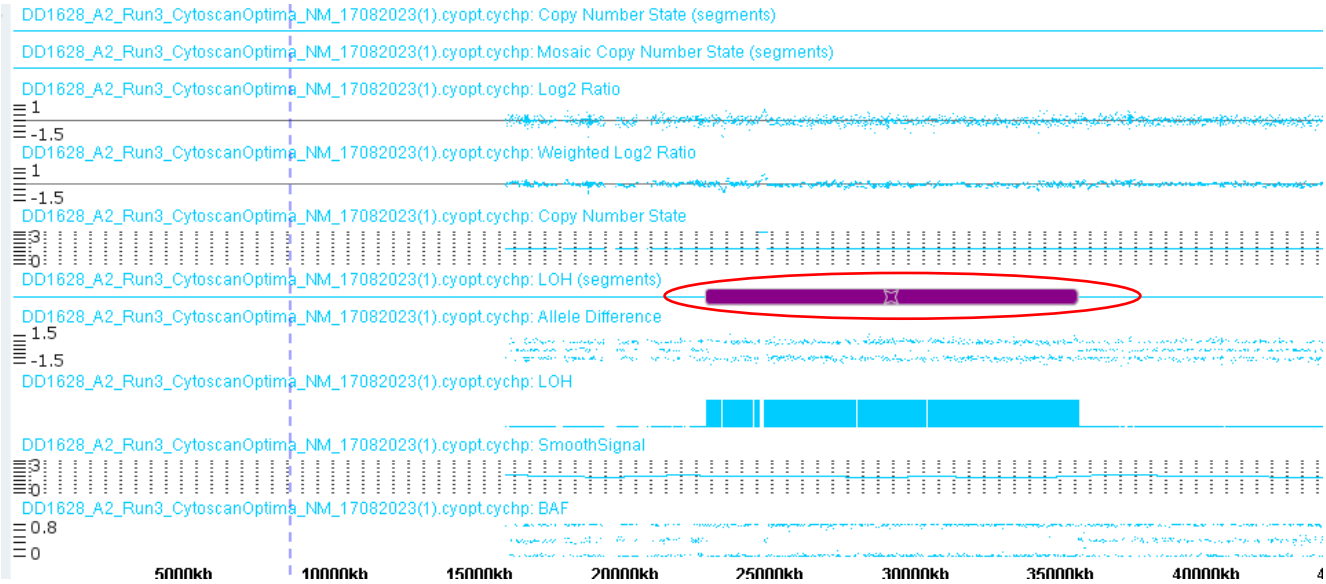


Figure 4. Screenshot of chromosome analysis suite v4.3 showing the LOH (segment), allele difference, LOH and BAF tracks of chromosome 22. The tracks show that the LOH segments (red circle) were associated with a loss in the same region, chromosome 22q11.23q12.3. Both the LOH and allele difference tracks indicate a presence of LOH in the region.

The ChAS browser indicated LOH for the entire X chromosome in all male patients in this sample as seen in Figure 5. This may have been due to the settings of the ChAS browser.

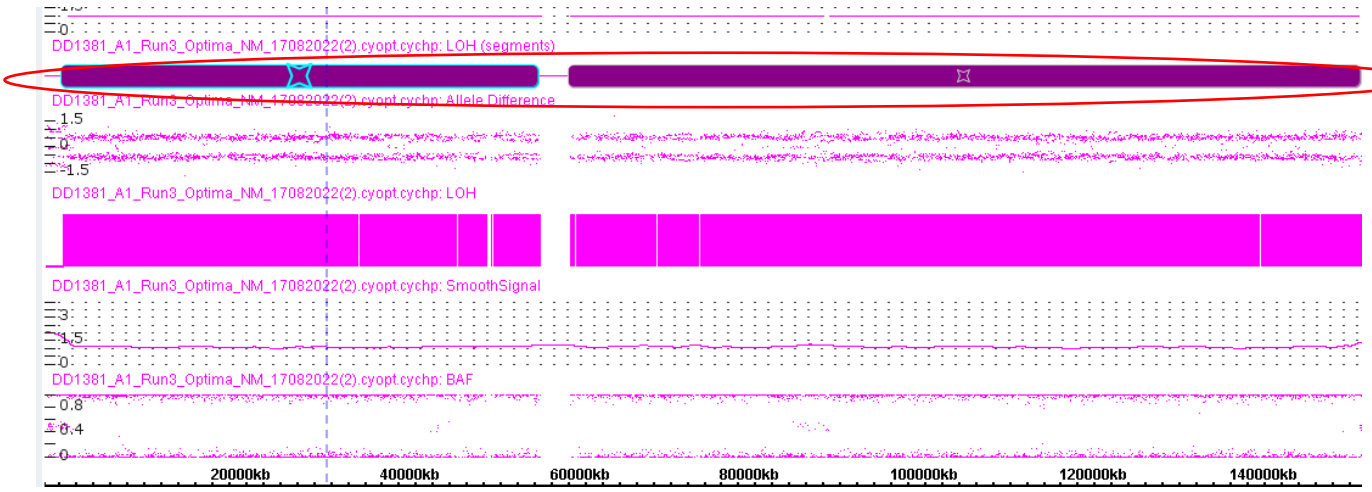


Figure 5. Screenshot of chromosome analysis suite v4.3 showing the LOH (segment), allele difference, LOH and BAF tracks of chromosome X in Patient 1. The tracks show that the LOH segments (red circle) runs on the entire X chromosome.

Costing

The price of a single sample SNP array test was determined to be $352.33 (R5601.57) at exchange rate $1 = R15.91 (9 January 2023) with DNA extraction included. This price was determined using the NHLS costing tool which considers the consumables and reagents, labour (medical scientists and genetic counsellor) and fixed equipment. Table 5 shows the breakdown of the amounts calculated for the different pricing components. The cost of the CGH array already implemented in the Division is $533.94 (8494.98) at exchange rate $1 = R15.91 (9 January 2023) with DNA extraction included. Table 6 shows the breakdown of the costs of the CGH array test. This makes the cost of SNP array $181.61 (R2889.41) at exchange rate $1 = R15.91 (9 January 2023) lower than the CGH array. The price breakdown for each test shows that the difference in total price between the two microarrays is due to a difference in the reagents and consumables used for each test.

Table 5. The cost of running a SNP-based array test

Component	**Cost for each test**
Reagents and Consumables	$ 268.93
Labour	$ 37.29
Fixed equipment	$ 13.59
DNA extraction	$ 32.52
TOTAL	$ 352.33

Table 6. The cost of running a CGH array test

Component	**Cost for each test**
Reagents and Consumables	$ 450.54
Labour	$ 37.29
Fixed equipment	$ 13.59
DNA extraction	$ 32.52
TOTAL	$ 533.94

Discussion

Chromosomal microarray analysis has proved to be useful in the detection of CNVs that cause developmental delay in individuals. In this study, the clinical utility and cost of a SNP-based array, the CytoScan® Optima microarray was compared to that of the SurePrint G3 Unrestricted CGH ISCA v2, 8x60K microarray, which is currently being used at the Division of Human Genetics, NHLS. The CNV calls made by each array type were compared for genomic coordinates, segment size and gene content. The genotype-phenotype correlations were only made for patients with associated syndromes.

The SNP array detected 24 CNVs in total and the CGH array detected 24 CNVs. Between the CGH array and the SNP array, there were 23 common CNVs detected. The SNP array was unable to detect the 18p11.32p11.31 loss that was detected by the SNP array and the CGH array was unable to detect the 15q26.2 gain that was detected by the SNP array. The ability of a microarray to detect genomic imbalances is determined by its design[35]. Although the designs of the two microarrays investigated are different, the CNVs detected were comparable.

The difference in design also suggests that there are likely to be differences in the genomic coordinates of the CNVs detected. This also explains the observed differences in the segment sizes of the CNVs detected by the two microarrays. An example of a difference in the starting breakpoint detected was seen in Patient 3 where the CytoScan® Optima detected a 13q11q12.11 deletion at coordinates 13:19436287-31787356 whereas the SurePrint ISCA array detected a 13q12.11 deletion at coordinates 13:20407295-21661250. The size difference in starting breakpoint was 9.71 Mb.

The sizes of the segments detected by both the CytoScan® Optima and the SurePrint ISCA array were comparable. In this study sample, the CytoScan® Optima array detected CNVs ranging in size from 0.98 Mb to 94.36 Mb. Other studies have reported the CytoScan® arrays detecting CNVs as small as 0.76 Mb[36].

Patient 2 was previously confirmed to have a gain at 3q21 by karyotyping, however karyotyping does not provide the breakpoint of the genomic imbalance. The two microarrays defined this CNV as a gain of 3q13.2q22.3. Microarrays have the advantage of being able define the exact genomic breakpoints of the CNV as shown in Table 1. This capacity to precisely characterize breakpoints is important in accurately determining the specific genes included in the CNV and therefore the genetic disorders that may be caused by that CNV[37].

The probe-spacing and design of the arrays compared here differs. For this study, a comparison between the CytoScan® 750K array (Thermo Fisher Scientific, Waltham, MA) and SurePrint ISCA may have been more suitable because similar to the SurePrint ISCA, the CytoScan® 750K array (Thermo Fisher Scientific, Waltham, MA) is designed for postnatal analysis.

The SNP array is designed with copy number probes alongside SNP probes therefore, unlike the CGH array, this array is able to genotype the SNP locus in that genomic region. This allows the detection of LOH in patients. Loss of heterozygosity is the loss of one of the pair of alleles at a constitutional heterozygous locus; it is assessed by detecting small DNA sequence changes at specific chromosomal loci[38]. LOH has been detected in events of UPD[39]. UPD is when the two homologous chromosomes are inherited from one parent with no contribution from the other parent[40]. LOH studies show that UPD has a prevalence of 1 in 176 in patients with developmental delay[41]. The American College of Medical Genetics (ACMG) recommends reporting LOH greater than 5 Mb[40]. In this study we analysed LOH regions greater than 8 Mb to decrease cases of false positives. This was necessary as false positives may lead to incorrect diagnoses.

Long regions of LOH greater than 8 Mb were detected in 9 patients. Seven of the LOH were false positives as the allele difference track that indicates genotype was not indicative of LOH region but a loss or the hemizygous X chromosome in males. Care must be taken in interpreting the results as shown deletions may be called as LOH due to the fact that in that region there is only one allele represented due to the loss of genomic material on the other chromosome.

Although the ChAS software was able to detect true LOH regions, there no evidence to support how this was the underlying cause of the patient phenotypes. The LOH region found in chromosome 22q11.23q12.3 in Patient 17 does not include the three known imprinting genes – *DGCR6L, DGCR6* and *FLJ204* found on chromosome 22[42]. Furthermore, the 36 genes in the LOH region on chromosome Xp11.23p11.1 in Patient 12 are not reported to be imprinted genes.

Being able to detect LOH is also important during diagnosis of patients because of the important information on consanguinity that may be derived Individuals born of a consanguineous union have segment of their genomes that are homozygous as a result of inheriting identical ancestral genomic segments through both parents[43]. Identifying LOH caused by consanguineous unions can help in diagnosing recessive conditions. This is why it is important to do trio testing whereby the maternal and paternal parent are also tested in order to determine whether a CNV was de novo or inherited. Information on family history and trio studies were not provided in this study and therefore the significance of LOH in the patients was not determined.

The difference in design between the two microarrays is influenced by the purpose for which it is intended. The CytoScan® Optima array was designed as a low-resolution prenatal array. This means it is limited to large and known pathogenic CNVs. The SurePrint ISCA array was designed for postnatal CNV detection and therefore has a much higher resolution as CNVs may be analysed in association with clinical features. This difference informs the probes and markers that manufacturers include in the microarray. For the CytoScan® Optima array, probe coverage needs to be sufficient for the detection of aneuploidies, large chromosomal aberrations and clinically relevant regions underlying known syndromes associated with congenital anomalies and developmental disorders.

The samples selected for this project were predominantly from patients with known CNVs including confirmed microduplications and microdeletion syndromes. Of the 17 patients, four had microdeletion syndromes and two had microduplication syndromes

confirmed diagnoses, and two confirmed trisomy of chromosome 7p in Patient 7 and trisomy 13q of chromosome 13q in Patient 9.

Inferences on the mechanism leading to a loss or gain can be made from the data from both microarrays. A quarter of the CNVs were suggested to be a result of unbalanced translocation which are known to cause developmental disorders[44]. Translocations are structural variants (SV) caused by the incorrect joining of co-existing chromosome breaks[45]. The joining together of chromosome breaks is facilitated by DNA repair mechanisms. Translocations are a result of irregularities in the proteins involved in the cellular mechanisms used for repair of chromosome breaks. Mechanisms include non-homologous end joining and homologous recombination[44]. The analysis software used to display the microarray results allow one to visualise positions of these structural variants. Microarrays can only infer the mechanism underlying unbalanced structural aberrations. Confirming the mechanism may be helpful in determining recurrence rate of the CNV for risk assessment during genetic counselling.

The two arrays compared in this study are able to detect mosaicism. The main difference is that the SNP-based array used SNP genotype data to provide a more reliable estimate of mosaicism. Using ChAS mosaicism can be displayed as a segment on the browser or inferred from the smooth signal, BAF and the allele difference track. Mosaicism on the CGH array can only be inferred from the log2 ratio.

Another difference is how the level of mosaicism detected for a particular CNV is estimated. In the CytoGenomics software of the Agilent platform, the level of mosaicism is manually calculated from the log2 ratios whereas in ChAS, the level of mosaicism is estimated by the software from the average of the smooth signal values. In this study, two samples were found to present with mosaic chromosomal aberrations, Patient 9 with gain of 13q at 63% mosaicism and Patient 3 with loss of 18p11.32 at 59%. SNP-based arrays are reported to be able to detect mosaic levels of as low as 4%[46]. The detection of mosaicism is important in diagnosis, predicting

recurrence rate, as well as in genetic counselling. Failure to identify low levels of mosaicism during diagnostic testing may lead to inaccurate interpretations of genetic results[47]. Mosaicism in copy number change have been reported in some developmental conditions[48,49]. SNP arrays are known to detect lower levels of mosaicism and being able to do that could lead to more accurate reporting genetics results to patients.

The two microarrays did not have 100% concordance for the known CNVs in the sample of patients investigated in this study. For Patient 3, the SurePrint ISCA array detected a deletion that was linked to the patient's clinical features (choanal atresia and subtle dysmorphism). This region included the gene *SMCHD1*; loss of function mutations in this gene causes Bosma arrhinia micropthalmia syndrome (OMIM: 603457), a rare condition of the eyes and nose[50]. Choanal atresia is one of the reported features in affected patients. In addition, there is one case recorded on ClinVar with a loss of the *SMCHD1* gene causing the disease[51].

The ChAS software did not call this deletion as a copy number change. However, the smooth signal and BAF values on the browser were indicative of a change in copy number. This may be explained by the lower resolution of the CytoScan Optima array compared to the SurePrint ISCA array. Although the smooth signal track and BAF track are able to indicate a copy number change, this may be missed by an inexperienced analyst.

The differences in the CNV calls made following analysis by the CGH array and the SNP-based array are important to note because as seen in Patient 3, the CNV that the ChAS software did not call contained the gene that explained the phenotype of the patient. This suggests the difference in the genomic regions that each array has a higher density of copy number probes in. This shows that both microarrays are relevant in detection of CNVs associated with neurodevelopmental and congenital disorders as they have the capacity to detect regions that are relevant to these disorders.

Patients 3, 4, 11, 15 and 16 presented with clinical phenotypes related to congenital disorders and developmental delays, however, the CNVs associated with the phenotypes were not clinically significant therefore a molecular diagnosis was not reached. This may be due to public databases not having enough information on a CNV, patients with that CNV not having a clinical diagnosis or insufficient information on the dosage sensitivity of the genes in the CNV. Cytogeneticists who analyse CNVs in clinical settings use public databases to inform the classification of a CNV and reach a molecular diagnosis for the patient. These databases having insufficient information limits the correlation of certain CNVs with the patients clinical features and/or condition.

The CytoScan® Optima array has benefits in that it is much easier to use. The washing step of the workflow happens in a closed fluidics system as opposed to the SurePrint ISCA array workflow whereby the washing step is done manually. The wash buffer conditions need to be temperature controlled therefore this step in the SurePrint ISCA is prone to error. A limitation of the CytoScan® Optima array is that it is labour intensive and has sensitive enzymes.

The price of detecting CNVs using a SNP array is $181.61 less than using the CGH array. This is the cost that patients will have to pay in order to have the test conducted on them if they present with intellectual disabilities and developmental delays and congenital anomalies. A cheaper price for these patients would be more beneficial as the NHLS services the public sector; a cheaper price may mean improved affordability for lower income patients.

Other studies that have compared the utility of CGH array and SNP array have found that the CGH array is better suitable for detection of CNVs because its probe is more clinically targeted. This means its usage leads to better results and coverage for target sites[52]. Another study however, found that the SNP array was more suitable for detection of copy neutral abnormalities such as UPD, however, the fewer the SNP probes on the array, the more likely LOHs that are due to identity by descent will be called[53].

Limitations

1. The study analysed a small sample size (17) due to:
   a. The CytoScan Optima kit used has the capacity to test only 24 samples.
   b. The enzymes used for the CytoScan Optima test are very sensitive therefore some tests were unsuccessful.
2. The public databases used to interpret and classify CNVs are limited; in some cases, there is little or no information and a classification may not be made.

Conclusion

The utility of the CytoScan Optima array was studied by comparing it to the SurePrint ISCA array that is already in use for diagnostic service at the Division of Human Genetics, National Health Laboratory Service in Johannesburg. The two microarrays investigated here have comparable detection capacity. Compared to CGH arrays, SNP arrays provide information on LOH that could be relevant in the diagnosis of genetic disorders. The SNP array costs less than the CGH array which would make it more advantageous to the lower income group. It is suggested that the SNP-based array and the CGH array be used interchangeably based on clinical requests. Since the CytoScan® Optima was designed specifically for prenatal studies, it can be used for prenatal testing to limit the detection of CNVs you cannot interpret. The Division should consider incorporating SNP probes into the microarray testing strategy for a suspected chromosomal abnormality.

References

1. MacDonald JR, Ziman R, Yuen RKC, Feuk L, Scherer SW. The Database of Genomic Variants: a curated collection of structural variation in the human genome. Nucleic Acids Res, 2014, 42

2. Levy B, Wapner R. Prenatal diagnosis by chromosomal microarray analysis. Fertil Steril, 2018, 109:201–12

3. Kaminsky EB, Kaul V, Paschall J, Church DM, Bunke B, Kunig D, Moreno-De-Luca D, Moreno-De-Luca A, Mulle JG, Warren ST, Richard G, Compton JG, Fuller AE, Gliem TJ, Huang S, Collinson MN, Beal SJ, Ackley T, Pickering DL, Golden DM, Aston E, Whitby H, Shetty S, Rossi MR, Rudd MK, South ST, Brothman AR, Sanger WG, Iyer RK, Crolla JA, Thorland EC, Aradhya S, Ledbetter DH, Martin CL. An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. Genet Med, 2011, 13:777–84

4. Liehr T. “Classical cytogenetics” is not equal to “banding cytogenetics.” Mol Cytogenet, 2017, 10:2

5. Shemilt L, Verbanis E, Schwenke J, Estandarte AK, Xiong G, Harder R, Parmar N, Yusuf M, Zhang F, Robinson IK. Karyotyping Human Chromosomes by Optical and X-Ray Ptychography Methods. Biophys J, 2015, 108:706

6. Cui C, Shu W, Li P. Fluorescence in situ hybridization: Cell-based genetic diagnostic and research applications. Front Cell Dev Biol, 2016, 4:89

7. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res, 2002, 30:57

8. Madrigal I, Rodríguez-Revenga L, Badenas C, Sánchez A, Martinez F, Fernandez I, Fernández-Buriel M, Milà M. MLPA as first screening method for the detection of microduplications and microdeletions in

patients with X-linked mental retardation. Genetics in Medicine 2007, 9:117–22

9. Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ, Faucett WA, Feuk L, Friedman JM, Hamosh A, Jackson L, Kaminsky EB, Kok K, Krantz ID, Kuhn RM, Lee C, Ostell JM, Rosenberg C, Scherer SW, Spinner NB, Stavropoulos DJ, Tepperberg JH, Thorland EC, Vermeesch JR, Waggoner DJ, Watson MS, Martin CL, Ledbetter DH. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. Am J Hum Genet, 2010, 86:749–64

10. Zarrei M, Fehlings DL, Mawjee K, Switzer L, Thiruvahindrapuram B, Walker S, Merico D, Casallo G, Uddin M, MacDonald JR, Gazzellone MJ, Higginbotham EJ, Campbell C, Deveber G, Frid P, Gorter JW, Hunt C, Kawamura A, Kim M, Mccormick A, Mesterman R, Samdup D, Marshall CR, Stavropoulos DJ, Wintle RF, Scherer SW. De novo and rare inherited copy-number variations in the hemiplegic form of cerebral palsy. Genet Med, 2018, 20:172–80

11. Bi W, Borgan C, Pursley AN, Hixson P, Shaw CA, Bacino CA, Lalani SR, Patel A, Stankiewicz P, Lupski JR, Beaudet AL, Cheung SW. Comparison of chromosome analysis and chromosomal microarray analysis: What is the value of chromosome analysis in today's genomic array era? Genetics in Medicine, 2013, 15:450–7

12. Peiffer DA, Le JM, Steemers FJ, Chang W, Jenniges T, Garcia F, Haden K, Li J, Shaw CA, Belmont J, Sau WC, Shen RM, Barker DL, Gunderson KL. High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. Genome Res, 2006, 16:1136

13. Bignell GR, Huang J, Greshock J, Watt S, Butler A, West S, Grigorova M, Jones KW, Wei W, Stratton MR, Futreal PA, Weber B, Shapero MH, Wooster R. High-Resolution Analysis of DNA Copy Number Using Oligonucleotide Microarrays. Genome Res, 2004, 14:287–94

14. Markello TC, Carlson-Donohoe H, Sincan M, Adams D, Bodine DM, Farrar JE, Vlachos A, Lipton JM, Auerbach AD, Ostrander EA, Chandrasekharappa SC, Boerkoel CF, Gahl WA. Sensitive Quantification of Mosaicism Using High Density SNP Arrays and the Cumulative Distribution Function. Mol Genet Metab, 2012, 105:665

15. Thermo Fisher Scientific. CytosScan Optima Assay Manual Workflow: User Guide, 2018

16. Hoppman N, Rumilla K, Lauer E, Kearney H, Thorland E. Patterns of homozygosity in patients with uniparental disomy: detection rate and suggested reporting thresholds for SNP microarrays. Genet in Med, 2018, 20:1522–7

17. Firth H v., Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, Vooren S van, Moreau Y, Pettett RM, Carter NP. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J Hum Genet, 2009, 84:524–33

18. Rehm HL, Berg JS, Brooks LD, Bustamante CD, Evans JP, Landrum MJ, Ledbetter DH, Maglott DR, Martin CL, Nussbaum RL, Plon SE, Ramos EM, Sherry ST, Watson MS. ClinGen — The Clinical Genome Resource. N Eng J Med, 2015, 372:2235–42

19. Landrum MJ, Chitipiralla S, Brown GR, Chen C, Gu B, Hart J, Hoffman D, Jang W, Kaur K, Liu C, Lyoshin V, Maddipatla Z, Maiti R, Mitchell J, O'Leary N, Riley GR, Shi W, Zhou G, Schneider V, Maglott D, Holmes JB, Kattman BL. ClinVar: improvements to accessing data. Nucleic Acids Res, 2020, 48:D835–44

20. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Jang W, Karapetyan K, Katz K, Liu C, Maddipatla Z, Malheiro A, McDaniel K, Ovetsky M, Riley G, Zhou G, Holmes JB, Kattman BL, Maglott DR. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res, 2018, 46:D1062–7

21. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res, 2005, 33:D514–7

22. Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. Nucleic Acids Res, 2015, 43:D789–98

23. Crolla JA, Cawdery JE, Oley CA, Young ID, Gray J, Fantes J, van Heyningen V. A FISH approach to defining the extent and possible clinical significance of deletions at the WAGR locus. J Med Genet, 1997, 34:207–12

24. Miller RW, Fraumeni JFJr, Manning MD. Association of Wilms's Tumor with Aniridia, Hemihypertrophy and Other Congenital Malformations. N Engl J Med, 1964, 270:922-7

25. Palomares M, Delicado A, Mansilla E, de Torres ML, Vallespín E, Fernandez L, Martinez-Glez V, García-Miñaur S, Nevado J, Simarro FS, Ruiz-Perez VL, Lynch SA, Sharkey FH, Thuresson AC, Annerén G, Belligni EF, Martínez-Fernández ML, Bermejo E, Nowakowska B, Kutkowska-Kazmierczak A, Bocian E, Obersztyn E, Martínez-Frías ML, Hennekam RCM, Lapunzina P. Characterization of a 8q21.11 microdeletion syndrome associated with intellectual disability and a recognizable phenotype. Am J Hum Genet, 2011, 89:295–301

26. Lindsay EA, Greenberg F, Shaffer LG, Shapira SK, Scambler PJ, Baldini A. Submicroscopic deletions at 22q11.2: Variability of the clinical picture and delineation of a commonly deleted region. Am J Med Genet, 1995, 56:191–7

27. Bassett AS, Chow EWC, Husted J, Weksberg R, Caluseriu O, Webb GD, Gatzoulis MA. Clinical features of 78 adults with 22q11 Deletion Syndrome. Am J Med Genet A, 2005, 138:307–13

28. Overhauser J, Huang X, Gersh M, Wilson W, Mcmahon J, Bengtsson U, Rojas K, Meyer M, Wasmuth JJ. Molecular and phenotypic mapping of

the short arm of chromosome 5: sublocalization of the critical region for the cri-du-chat syndrome. Hum Mol Genet, 1994, 3:247–52

29. van Buggenhout GJCM, Pijkels E, Holvoet M, Schaap C, Hamel BCJ, Fryns JP. Cri du Chat Syndrome: Changing Phenotype in Older Patients. J Med Genet, 2000, 90:203–15

30. Fang JS, Lee KF, Huang CT, Syu CL, Yang KJ, Wang LH, Liao DL, Chen CH. Cytogenetic and molecular characterization of a three-generation family with chromosome 5p terminal deletion. Clin Genet, 2008, 73:585–90

31. Fernandez BA, Roberts W, Chung B, Weksberg R, Meyn S, Szatmari P, Joseph-George AM, MacKay S, Whitten K, Noble B, Vardy C, Crosbie V, Luscombe S, Tucker E, Turner L, Marshall CR, Scherer SW. Phenotypic spectrum associated with de novo and inherited deletions and duplications at 16p11.2 in individuals ascertained for diagnosis of autism spectrum disorder. J Med Genet, 2010, 47:195–203

32. Roos L, Jønch AE, Kjaergaard S, Taudorf K, Simonsen H, Hamborg-Petersen B, Brøndum-Nielsen K, Kirchhoff M. A new microduplication syndrome encompassing the region of the Miller-Dieker (17p13 deletion) syndrome. J Med Genet, 2009, 46:703–10

33. Brasseur B, Martin CM, Cayci Z, Burmeister L, Schimmenti LA. Bosma arhinia microphthalmia syndrome: Clinical report and review of the literature. Am J Med Genet , 2016, 170:1302–7

34. Morrow EM. Genomic Copy Number Variation in Disorders of Cognitive Development. J Am Acad Child Adolesc Psychiatry, 2010, 49:1091–104

35. Rodríguez-Revenga L, Vallespín E, Madrigal I, Palomares M, Mur A, García-Miñaur S, Santos F, Mori MÁ, Lapunzina P, Mila M, Nevado J. A parallel study of different array-CGH platforms in a set of Spanish patients with developmental delay and intellectual disability. Gene, 2013, 521:82–6

36. Bernardini L, Alesi V, Loddo S, Novelli A, Bottillo I, Battaglia A, Digilio MC, Zampino G, Ertel A, Fortina P, Surrey S, Dallapiccola B. High-

resolution SNP arrays in mental retardation diagnostics: how much do we gain? Eur J Hum Gen 2010, 18:178–85

37. Chen W, Ullmann R, Langnick C, Menzel C, Wotschofsky Z, Hu H, Döring A, Hu Y, Kang H, Tzschach A, Hoeltzenbein M, Neitzel H, Markus S, Wiedersberg E, Kistner G, van Ravenswaaij-Arts CM, Kleefstra T, Kalscheuer VM, Ropers HH. Breakpoint analysis of balanced chromosome rearrangements by next-generation paired-end sequencing. Eur J Hum Gen 2010, 18:539–43

38. Thiagalingam S, Laken S, Willson JKV, Markowitz SD, Kinzler KW, Vogelstein B, Lengauer C. Mechanisms underlying losses of heterozygosity in human colorectal cancers. Proc Natl Acad Sci U S A, 2001, 98:2698–702

39. Murthy SK, DiFrancesco LM, Travis Ogilvie R, Demetrick DJ. Loss of Heterozygosity Associated with Uniparental Disomy in Breast Carcinoma. Mod Path 2002, 15:1241–50

40. del Gaudio D, Shinawi M, Astbury C, Tayeh MK, Deak KL, Raca G. Diagnostic testing for uniparental disomy: a points to consider statement from the American College of Medical Genetics and Genomics (ACMG). Genet Med, 2020, 22:1133–41

41. Dong Z, Chau MHK, Zhang Y, Yang Z, Shi M, Wah YM, Kwok YK, Leung TY, Morton CC, Choy KW. Low-pass genome sequencing-based detection of absence of heterozygosity: validation in clinical cytogenetics. Genet Med, 2021, 23:1225

42. Li X, Liu Y, Yue S, Wang L, Zhang T, Guo C, Hu W, Kagan KO, Wu Q. Uniparental disomy and prenatal phenotype: Two case reports and review. Medicine, 2017, 96

43. Woods CG, Cox J, Springell K, Hampshire DJ, Mohamed MD, McKibbin M, Stern R, Raymond FL, Sandford R, Sharif SM, Karbani G, Ahmed M, Bond J, Clayton D, Inglehearn CF. Quantification of Homozygosity in Consanguineous Individuals with Autosomal Recessive Disease. The Am J Hum Genet, 2006, 78:889–96

44. Ramsden DA, Nussenzweig A. Mechanisms driving chromosomal translocations: lost in time and space. Oncogene 2021, 40:4263–70

45. Li Y, Roberts ND, Wala JA, Shapira O, Schumacher SE, Kumar K, Khurana E, Waszak S, Korbel JO, Haber JE, Imielinski M, Akdemir KC, Alvarez EG, Baez-Ortega A, Beroukhim R, Boutros PC, Bowtell DDL, Brors B, Burns KH, Campbell PJ, Chan K, Chen K, Cortés-Ciriano I, Dueso-Barroso A, Dunford AJ, Edwards PA, Estivill X, Etemadmoghadam D, Feuerbach L, Fink JL, Frenkel-Morgenstern M, Garsed DW, Gerstein M, Gordenin DA, Haan D, et al. Patterns of somatic structural variation in human cancer genomes. Nature, 2020, 578:112–21

46. Markello TC, Carlson-Donohoe H, Sincan M, Adams D, Bodine DM, Farrar JE, Vlachos A, Lipton JM, Auerbach AD, Ostrander EA, Chandrasekharappa SC, Boerkoel CF, Gahl WA. Sensitive quantification of mosaicism using high density SNP arrays and the cumulative distribution function. Mol Genet Metab, 2012, 105:665–71

47. Qin L, Wang J, Tian X, Yu H, Truong C, Mitchell JJ, Wierenga KJ, Craigen WJ, Zhang VW, Wong LJC. Detection and Quantification of Mosaic Mutations in Disease Genes by Next-Generation Sequencing. J Mol Diagn, 2016, 18:446–53

48. Yurov YB, Vorsanova SG, Iourov IY, Demidova IA, Beresheva AK, Kravetz VS, Monakhov V v, Kolotii AD, Voinova-Ulas Y. Unexplained autism is frequently associated with low-level mosaic aneuploidy. J Med Genet, 2007, 44:521–5

49. Pham J, Shaw C, Pursley A, Hixson P, Sampath S, Roney E, Gambin T, Kang SHL, Bi W, Lalani S, Bacino C, Lupski JR, Stankiewicz P, Patel A, Cheung SW. Somatic mosaicism detected by exon-targeted, high-resolution aCGH in 10,362 consecutive cases. Eur J Hum Genet, 2014, 22:969–78

50. Graham JM, Lee J. Bosma arhinia microphthalmia syndrome. Am J Med Genet A, 2006, 140:189–93

51. Shaw ND, Brand H, Kupchinsky ZA, Bengani H, Plummer L, Jones TI, Erdin S, Williamson KA, Rainger J, Stortchevoi A, Samocha K, Currall BB, Dunican DS, Collins RL, Willer JR, Lek A, Lek M, Nassan M, Pereira S, Kammin T, Lucente D, Silva A, Seabra CM, Chiang C, An Y, Ansari M, Rainger JK, Joss S, Smith JC, Lippincott MF, Singh SS, Patel N, Jing JW, Law JR, Ferraro N, et al. SMCHD1 mutations associated with a rare muscular dystrophy can also cause isolated arhinia and Bosma arhinia microphthalmia syndrome. Nature Genet 2017, 49:238–48

52. Mason-Suares H, Kim W, Grimmett L, Williams ES, Horner VL, Kunig D, Goldlust IS, Wu BL, Shen Y, Miller DT, Martin CL, Rudd MK. Density matters: Comparison of array platforms for detection of copy-number variation and copy-neutral abnormalities. Genet Med, 2013, 15:706–12

53. Shen Xuan, Goidin Didier, Kim Young, Happe Scott. Technical overview of In silico Probe Coverage Comparision Between Agilent GentiSure Cyto vs CytoScan Microarrays. Santa Clara, Carlifornia, 2021

Appendices

Appendix A. Approved research protocol

CANDIDATE'S SURNAME: MAYISELA	FIRST NAME/S: MINENHLE		STUDENT NUMBER: 1880919
CURRENT QUALIFICATIONS: BSC (HONS) GENETICS AND DEVELOPMENTAL BIOLOGY			
TEL:	CELL: 078 788 8219	E-MAIL: 1880919@STUDENTS.WITS.AC.ZA	FAX:
DEGREE FOR WHICH PROTOCOL IS BEING SUBMITTED: MSC (MED) GENOMIC MEDICINE			
PART-TIME OR FULL-TIME: FULL-TIME			
FIRST REGISTERED FOR THIS DEGREE: 2022	TERM: 1	YEAR: 2022	
DEPARTMENT: HUMAN GENETICS			
TITLE OF PROPOSED RESEARCH: EVALUATING THE CLINICAL UTILITY OF A SNP-BASED ARRAY PLATFORM – A COMPARATIVE STUDY			
CANDIDATE'S SIGNATURE:		DATE: 28 APRIL 2022	
SUPERVISOR 1 (NAME & SURNAME): NKATEKO MAYEVU		% SUPERVISION: 60%	
SUPERVISOR'S QUALIFICATIONS: MSC (MED) MEDICAL BIOCHEMISTRY			
SUPERVISOR'S DEPARTMENT: HUMAN GENETICS			

SUPERVISOR'S ADDRESS / TEL / E-MAIL: DIVISION OF HUMAN GENETICS, NHLS BRAAMFONTEIN / 011 489 9237 / NKATEKO.MAYEVU@NHLS.AC.ZA	
SUPERVISOR 2 (NAME & SURNAME): FIONA BAINE-SAVANHU	% SUPERVISION: 40%
SUPERVISOR'S QUALIFICATIONS: PHD	
SUPERVISOR'S ADDRESS / TEL / E-MAIL: FIONA.BAINE@WITS.AC.ZA	

SYNOPSIS OF RESEARCH:*(Brief summary of proposed research project; between 200-300 words only; with sub-headings: an introduction and justification for study, aim/s, proposed methodology and expected outcome/s)*

Introduction and justification of study:

The proposed study will compare the clinical utility of a SNP-based array (Cytoscan® Optima) to that of the comparative genomic hybridization array (aCGH) currently used by the Division of Human Genetics. The aCGH platform has been used by the Division of Human Genetics, National Health Laboratory Service as a diagnostic method for conditions caused by copy number variants and structural abnormalities. The long-term goal of this project is to make recommendations regarding whether the Division should implement a SNP-based array in diagnostic setting.

Aim:

The aim of the study is to evaluate the clinical utility of a SNP-based array for diagnostic CNV detection compared to an CGH array.

Proposed methodology:

The study will analyse SNP-based array data using the Chromosomal Analysis Suite (ChAS) v4.3, which is an Affymetrix copy number variant (CNV) calling tool. The CNV analysis and interpretation will consider the size of the CNV, the gene content of the CNV and the use of databases such as database of Genomic Variants (DGV). A comparison of the available CNV results obtained from the CGH array and the SNP-based array by analysing the data from 10 samples with known CNVs that have been identified on the CGH array platform. Lastly the cost of offering the SNP-based array test in the Division of Human Genetics will determined using the National Health Laboratory Services (NHLS) costing tool.

Expected outcome:

A direct comparison of a SNP-based array and CGH array for diagnostic CNV detection

This project will explore the use in using SNP microarrays for detecting copy number variants in our human DNA sequences. These copy number variants are cause of disease for many South Africans and most of the time they go undiagnosed because of lack of technology to detect the disease. This project will show why using SNP array will improve diagnostic rate of disease and this is important because diagnosis of disease informs prognosis and will be helpful in offering personalised therapy and even prompt scientist to start clinical trials to treat the genetic cause of disease.

WITS ETHICS NOT REQUIRED: Yes No WITS ETHICS PENDING: Yes No WITS ETHICS APPROVED: Yes No (circle appropriate symbol)* ***Please note the final human ethics clearance certificate or animal ethics certificate must be available prior to starting research**		IF Y SUPPLY ETHICS CLEARANCE CERTIFICATE AS ATTACHMENT AND INCLUDE ETHICS NUMBER HERE: **M220295**
As supervisor/s, I/we confirm that I have read the protocol which has been submitted for assessment.		
SIGNATURE OF SUPERVISOR/S:		
SIGNATURE PG OFFICE STAFF ……..	REGISTERED YES….. NO…..	STAMP

SYNOPSIS OF RESEARCH CONTINUED

Research Proposal

Evaluating the clinical utility of a SNP-based array platform – a comparative study

By

Minenhle Mayisela

1880919

Master of Science in Medicine in Genomic Medicine

Supervised by:

Ms Nkateko Mayevu

Dr Fiona Baine - Savanhu

Table of Contents

1. Introduction 55
1.2. Evolution of Cytogenetics 55
1.2.2. Fluorescence in situ hybridization (FISH) 56
1.2.3. Multiplex Ligation Dependent Probe Amplification (MLPA) 56
1.2.4. Chromosomal Microarray 57
1.2.5. Copy number variant testing at the Division of Human Genetics 58
2. Study Aim and Objectives 61
3. Methods 61
3.1. Objective 1 – Analysis of CNVs from SNP array 61
3.2. Objective 2 – Comparison of CNVs between the two platforms 61
3.3. Objective 3 - Costing analysis of SNP array 62
4. Ethics 62
5. Timeline 62
6. Funding 63
7. Problems/Limitations 63
8. References 63
9. Appendices 70

Table of figures

Figure 1. Cytogenomics image showing a terminal 13 deletion……8

Figure 2. Affymetrix® Chromosome Analysis Suite image showing normal chromosome

4……
……9

List of abbreviations

CGH – comparative genomic hybridization

CMA – Chromosomal microarray analysis

CNV – copy number variants

FISH – Fluorescence in situ hybridization

kb – kilobases

LOH – loss of heterozygosity

Mb – Megabases

MLPA – Multiplex ligation-dependent probe amplification

PCR – Polymerase chain reaction

SNP – single nucleotide polymorphism

UPD – Uniparental disomy

ISCA - International Standards for Cytogenomic Arrays

1. Introduction

1.1. Copy Number Variants

Copy number variants (CNVs) are DNA segments that are greater than 50 bp in length and show variable copy number compared to reference genome [1] This suggests that the genome of individuals not only varies in DNA sequence but also in the number of copies of a particular DNA segment [54] . A greater proportion of CNVs are part of normal variation and do not cause disease.

The CNVs that are disease-causing are usually large, encompassing many genes and/or are found in dosage sensitive genes [54]. They play a role in the aetiology of diseases such as developmental disorders [55]. The study of CNVs in relation to disease contributes to determining the copy loss or gains that cause disease [55]. It also involves discovery of rare de novo alterations that cause neurodevelopmental disorders and congenital anomalies [55]. These CNVs are detected using several laboratory techniques.

Cytogenetic methods have been used for a long time to detect CNVs that cause disease in humans. Over time, the methods have evolved from conventional cytogenetic methods such as karyotyping to molecular cytogenetics methods such chromosomal microarray platforms.

2. Evolution of Cytogenetics

Cytogenetic techniques are used to identify and analyse structural rearrangements of chromosomes caused by cellular processes such as incorrect recombination or DNA repair. They detect balanced and unbalanced structural changes to chromosomes such as CNVs, translocations and inversions [4]. The techniques have evolved over the years, with each successive technique improving resolution and turnaround time for genetic testing and disease diagnosis.

2.1.1. Karyotyping

Karyotyping allows scientists to view all chromosomes under a microscope and detect numerical and large structural chromosome abnormalities by analysing banding patterns [5]. Giemsa/G banding is a commonly used technique in karyotyping. The

giemsa stain and trypsin solution create dark and light banding patterns on the chromosomes [5]. This helps visualise the different regions of the chromosomes and to study any structural changes. More advanced molecular cytogenetic techniques have been introduced; however, karyotyping is still considered the gold standard for confirming and informing the mechanism underlying many chromosomal aberrations. The resolution of karyotyping is between 5-10 Mb, meaning it is unable to detect submicroscopic chromosomal abnormalities [5]. Improved cytogenetic techniques have helped detect these abnormalities over the years.

3. Fluorescence in situ hybridization (FISH)

The development of fluorescence in situ hybridization (FISH) made the detection of submicroscopic chromosomal abnormalities possible, addressing some of the limitations of the karyotype. It is a molecular cytogenetic technique that uses fluorescently labelled probes hybridized on chromosomes to detect known structural changes in chromosomes [56] The chromosomes are analysed under a fluorescent microscope. The resolution of FISH is between 100 – 200 kb, depending on the probes used [57]. The FISH probes are designed for specific regions of interest on the chromosome; thus, to employ FISH, there should be a suspected disorder [58]. Another drawback is that FISH can only test for one suspected chromosomal abnormality at a time making it less cost effective. T

4. Multiplex Ligation-Dependent Probe Amplification (MLPA)

Multiplex ligation-dependent probe amplification (MLPA) is a technique that quantifies the dosage of DNA regions based on directed ligation [7]. It has the advantage of targeting multiple regions on the genome, which FISH fails to do so. This makes MLPA more cost effective. The MLPA technique uses probes specific to the region of interest to assess the relative copy number of the DNA sequence [8]. The method uses two probes that are ligated if hybridization takes place [8]. Ligated probes are then amplified by a polymerase chain reaction (PCR) and the products visualised using specialised software called Coffalyser.

The results are generated as peaks on a graph, which are then compared to separate peaks from control probes [7]. The MLPA probes can be used with sequence targets of between 50 – 100 bases long. This means that this technique may be useful in fragmented DNA as well as in detecting single gene changes that cause disease and cannot be detected by FISH [7].

Although MLPA has multiple targets, the probes are designed for known submicroscopic abnormalities related to specific diseases, meaning only those conditions can be tested using this technique. With the field of genetics rapidly changing, testing strategies and techniques have also evolved to provide more information for clinical diagnosis. A broader technique which targets the whole genome - the microarray - makes it possible to detect de novo aberrations.

5. Chromosomal Microarray

In 2010, the International Standards for Cytogenomic Arrays (ISCA) Consortium recommended chromosomal microarray analysis (CMA) as the first line of testing for patients that present with developmental delays, congenital anomalies, intellectual disabilities and autism spectrum disorder [59]. Microarray addressed the resolution limits imposed by conventional cytogenetic techniques such as karyotyping and FISH [11]. The CMA technique is a deoxyribonucleic acid-based method that detects CNVs as well as the size of the aberrations present [60].

Microarrays detect CNVs, unbalanced chromosomal rearrangements and aneuploidies. The resolution depends on probe number, spacing and distribution when designing the microarray [11].

There are two microarray approaches used: array comparative genome hybridization (aCGH) or single nucleotide polymorphism (SNP) array [61]. SNP-based arrays specifically are capable of detecting loss heterozygosity (LOH) and uniparental disomies (UPD) [61]. A study undertaken to diagnose structural aberrations in patients with Down syndrome using CMA, gave a diagnostic yield that was 15-20% higher compared to the 3% yield from conventional methods, specifically karyotyping [59,62].

5.1.1.1. Comparative Genomic Hybridization array

The aCGH is referred to as true comparative hybridization because it demonstrates the amount of DNA at various areas of the genome by comparing a test sample to a normal reference sample [11]. This is done by differentially labelling the reference DNA and the patient's DNA with fluorophores and the two samples are then co-hybridized onto a microarray (a glass slide containing multiple probes from various regions of the genome) [63]. The results indicate the fluorescence ratio of the amount of reference sample to the test sample. A higher presence of reference sample indicates that there is a loss of DNA and a higher presence of the test sample indicates a gain of DNA [63] .

The results are represented graphically as a $\log_2$ ratio. Important to note is that aCGH requires both a test and normal reference in order to determine if there is a CNV at a region [11].

5.1.1.2. Single Nucleotide Polymorphism array (SNP array)

The SNP-based microarray requires only the test sample, in what is called a one – colour test [12]. The SNP array functions by directly hybridizing the test DNA sample to a microarray slide that contains probes designed to bind to a target DNA sequence flanking SNPs [13]. The fluorescent colour intensity is dependent on the amount of target DNA in the sample. The results are displayed as a graph with two tracks. One track shows the exact copy number of the gene region of the sample represented by a probe. Another track shows the B-allele frequency which shows the genotype of each SNP locus. The B allele measures the intensity ratio of the two [54].

The SNP microarrays are a tool that have gone from genotyping a few thousands of SNPs to presently genotyping millions [64]. Genotyping is made possible by SNP arrays being developed to include SNP probes alongside copy number probes. The results of the array show the gain or loss of DNA as well as the different SNP genotypes for each SNP locus [13]. Since SNP arrays are a genotyping tool, they can detect long regions of loss of heterozygosity [65]. They can also detect low level mosaicism [14].

Two chromosomes will give three possible combinations: AA, AB and BB and the graph will show three tracks [66]. The probes are designed to hybridize two alleles (A and B alleles). Once hybridized, a scanner measures the intensity of fluorescence from the hybridized A and B probes on the array. The SNP genotypes are analysed by comparing A and B allele fluorescence intensity. A genotype of A-A is called when A fluorescence intensity is strong and B allele intensity is low; B-B is called when B fluorescence intensity is strong and A allele intensity is low. When A-B is called, that indicates that the two intensities are similar and of medium level of intensity (Levy & Burnside, 2019).

This is possible because the number of homozygous SNPs is consistent with a longer region of loss of heterozygosity in chromosomes [65]

6. Diagnostic copy number variant testing at the Division of Human Genetics

Comparative genomic hybridization (CGH) array was implemented in 2020 as the first line of testing for patients with developmental disorders and congenital abnormalities

at the Division of Human Genetics, National Health Laboratory Service and the University of the Witwatersrand. The Division uses the SurePrint G3 Unrestricted CGH ISCA v2, 8x60 kb microarray (Agilent Technologies). This microarray has 60 000 copy number probes that are concentrated at genomic regions of clinical relevance as determined by the ISCA consortium. The CNV calling is performed on Cytogenomics and visualised in Figure 1. The array has a resolution of 200kb and analyses eight samples on a single slide. It is unable to detect loss of heterozygosity LOH, UPD and low-level mosaicism (<30%).

The Division recently acquired a Affymetrix SNP- based array system to advance genomic research. The data for this study will be generated using the CytoScan Optima array (Thermo Fisher Scientific). This array consists of 18 018 copy number probes and 148 450 SNP probes. The copy number probes detect the CNVs of the test sample at the gene region the probe hybridizes. The SNP probes have a SNP site at different locations. CNV calling is performed using Chromosomal Analysis suite v4.3. The resolution offered by this array is low compared to the Agilent array currently in use in the Division. The resolution is 1 Mb for losses, 2 Mb for gains and 5 Mb for loss of heterozygosity. The resolution is low because this array is designed for prenatal testing therefore the probes are limited to specific regions with clear clinical significance. However, the availability of SNP probes makes it useful for detection of LOHs, UPDs and low-level mosaicism.

Given that CGH array has been implemented in the Division of Human Genetics, the question this project seeks to answer is whether a SNP array may add more value to the diagnostic service.

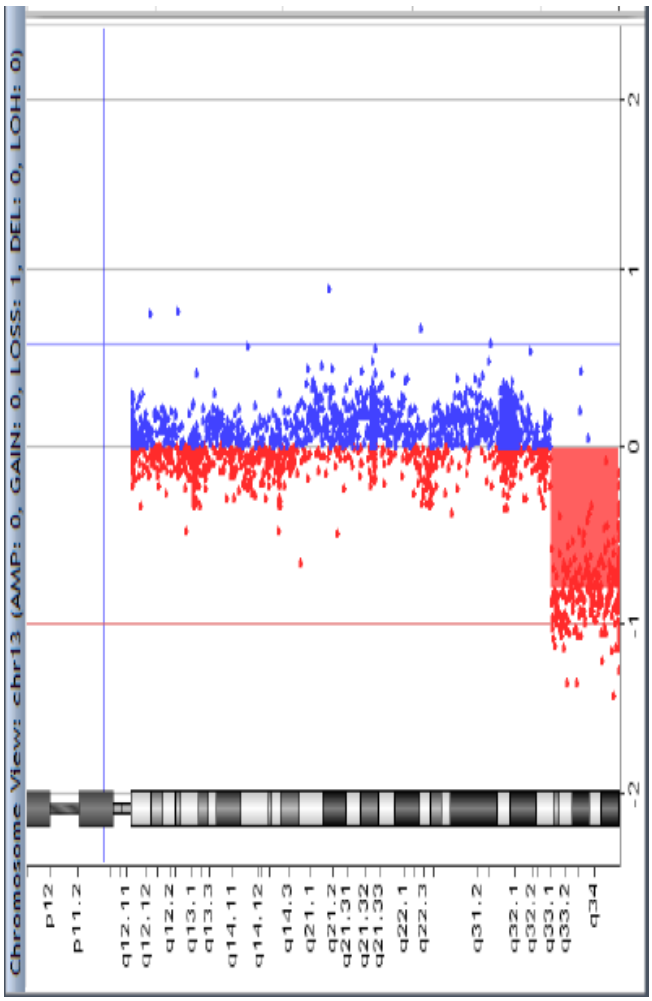


Figure 1. Cytogenomics image showing a terminal 13 deletion. The red line represents $\log_2$ ratio of 1 which indicates a deletion.

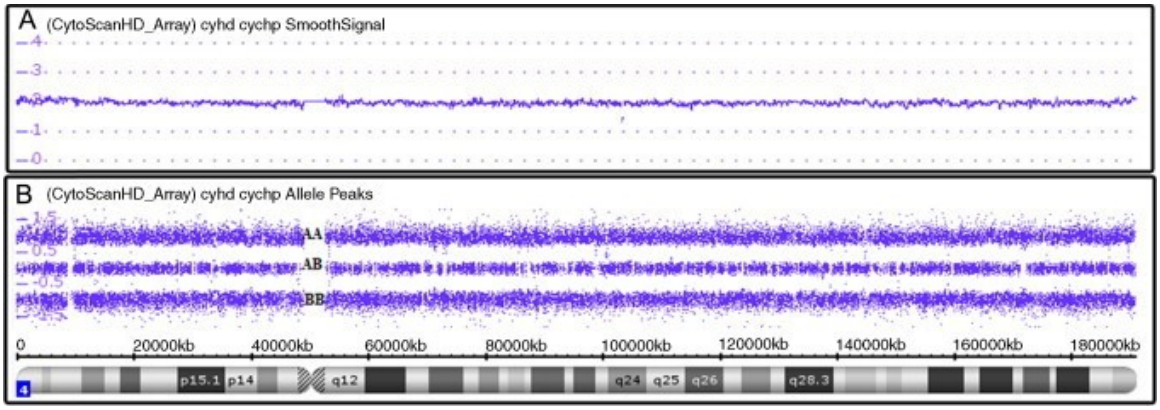


Figure 2. Affymetrix® Chromosome Analysis Suite image showing normal chromosome 4. (A) Smooth signal representing a normal copy number (purple line CN: 2.00). The information on copy number variations is from non-polymorphic probes. (B) Allele peaks representing a normal genotype with AA, AB and BB alleles (three purple lines). Information on the genotype is from polymorphic probes. (Figure adapted from Rev Bras Hematol Hemoter. 2015 Jan-Feb; 37(1): 48–54).

7. Study Aim and Objectives

The aim of this study is to evaluate the clinical utility of a SNP-based microarray in comparison to the CGH array platform for diagnostic CNV detection in the Division of Human Genetics.

The specific objectives of this study are to:

1. To analyse CNVs detected from Affymetrix SNP-based array data for 10 selected samples.
2. To compare CNV calls/results obtained from the Affymetrix SNP-based array to the available diagnostic results from the Agilent platform.
3. To conduct a costing analysis of the Affymetrix SNP array platform

8. Methods

9. Objective 1 – Analysis of CNVs from SNP array

The CNVs detected from the CytoScan Optima array performed on the Affymetrix platform will be analysed. The dataset consists of CNV and SNP data from 10 samples with a variety of known pathogenic chromosomal rearrangements representing gains and losses. The 10 samples were selected from the diagnostic records (results) available to represent a range of structural variants of different kinds including gains, losses and large aberrations. A GeneChip Scanner 3000, using the GeneChip Command Console Software (Thermo Fisher Scientific), will be used to generate the CEL files that summarise the intensity probe signals.

The data will be analysed using a CNV-calling tool called Chromosomal Analysis Suite (ChAS) v4.3. The tool computes and plots the genome and chromosome overviews with tracks incorporated to interpret the aberrations. The ChAS software also contains algorithms for the detection of LOH and UPD. The software will be used to classify and annotate the observed aberrations. Through the databases linked to the software, the observed aberrations can be classified for pathogenicity. The databases that will be used are: Databases of Genomic Variants (DGV), DECIPHER and ClinGen. These databases contain genetic and phenotypic information of variation in the genome. They allow for storage, sharing and comparison of genetic data across research studies.

10. Objective 2 – Comparison of CNVs between the two platforms

The data generated from the CytoScan® Optima array will be compared to data that has already been generated from the Agilent platform.

The following data generated from each platform will be compared:

- The probe density at each DNA region for each platform
- The $\log_2$ ratio figures for each CNV call
- The number of calls for each sample based on a 200 kb threshold
- The segment coordinates of each CNV call that will be detected
- The gene content of the CNVs called made

The results will be validated by correlating whether the CNV with the clinical phenotype.

11. Objective 3 - Costing analysis of SNP array

A costing analysis of the Affymetrix SNP-based array will be performed using the NHLS costing model. This tool includes calculation cost of reagents, the fixed facilities and the labour cost of lab personnel. This will allow for the estimation of the total cost per sample of the SNP array. This cost will be compared to the current cost of the aCGH.

12. Ethics

A conditional ethics clearance (M220295) was obtained from the Human Research Ethics Committee (HREC) to undertake the sub study under M18056 with the following conditions: evidence of protocol approval to be submitted as soon as it is approved. The ethics clearance certificate can be found in Appendix B

13. Timeline

	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Literature Review												
Ethics Application												
Protocol Preparation												
Protocol Assessment												
Introduction to software												
Data Analysis												

Written Report												

14. Funding

No funding is required for this study.

15. Problems/Limitations

- The SNP-based array has a lower resolution and may not pick up some of the variants detected on a CGH array
- The sample size is small due to limited funds available for this analysis
- Neither of the microarray platforms is a gold standard, therefore results that are picked up from only one of the platforms cannot be truly validated

16. References

1. MacDonald JR, Ziman R, Yuen RKC, Feuk L, Scherer SW. The Database of Genomic Variants: a curated collection of structural variation in the human genome. Nucleic Acids Res, 2014, 42

2. Levy B, Wapner R. Prenatal diagnosis by chromosomal microarray analysis. Fertil Steril, 2018, 109:201–12

3. Kaminsky EB, Kaul V, Paschall J, Church DM, Bunke B, Kunig D, Moreno-De-Luca D, Moreno-De-Luca A, Mulle JG, Warren ST, Richard G, Compton JG, Fuller AE, Gliem TJ, Huang S, Collinson MN, Beal SJ, Ackley T, Pickering DL, Golden DM, Aston E, Whitby H, Shetty S, Rossi MR, Rudd MK, South ST, Brothman AR, Sanger WG, Iyer RK, Crolla JA, Thorland EC, Aradhya S, Ledbetter DH, Martin CL. An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. Genet Med, 2011, 13:777–84

4. Liehr T. “Classical cytogenetics” is not equal to “banding cytogenetics.” Mol Cytogenet, 2017, 10:2

5. Shemilt L, Verbanis E, Schwenke J, Estandarte AK, Xiong G, Harder R, Parmar N, Yusuf M, Zhang F, Robinson IK. Karyotyping Human Chromosomes by Optical and X-Ray Ptychography Methods. Biophys J, 2015, 108:706

6. Cui C, Shu W, Li P. Fluorescence in situ hybridization: Cell-based genetic diagnostic and research applications. Front Cell Dev Biol, 2016, 4:89

7. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res, 2002, 30:57

8. Madrigal I, Rodríguez-Revenga L, Badenas C, Sánchez A, Martinez F, Fernandez I, Fernández-Buriel M, Milà M. MLPA as first screening method for the detection of microduplications and microdeletions in patients with X-linked mental retardation. Genet Med, 2007, 9:117–22

9. Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ, Faucett WA, Feuk L, Friedman JM, Hamosh A, Jackson L, Kaminsky EB, Kok K, Krantz ID, Kuhn RM, Lee C, Ostell JM, Rosenberg C, Scherer SW, Spinner NB, Stavropoulos DJ, Tepperberg JH, Thorland EC, Vermeesch JR, Waggoner DJ, Watson MS, Martin CL, Ledbetter DH. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. Am J Hum Genet, 2010, 86:749–64

10. Zarrei M, Fehlings DL, Mawjee K, Switzer L, Thiruvahindrapuram B, Walker S, Merico D, Casallo G, Uddin M, MacDonald JR, Gazzellone MJ, Higginbotham EJ, Campbell C, Deveber G, Frid P, Gorter JW, Hunt C, Kawamura A, Kim M, Mccormick A, Mesterman R, Samdup D, Marshall CR, Stavropoulos DJ, Wintle RF, Scherer SW. De novo and rare inherited copy-number variations in the hemiplegic form of cerebral palsy. Genet Med, 2018, 20:172–80

11. Bi W, Borgan C, Pursley AN, Hixson P, Shaw CA, Bacino CA, Lalani SR, Patel A, Stankiewicz P, Lupski JR, Beaudet AL, Cheung SW. Comparison of chromosome analysis and chromosomal microarray analysis: What is the value of chromosome analysis in today's genomic array era? Genetic in Medicine, 2013, 15:450–7

12. Peiffer DA, Le JM, Steemers FJ, Chang W, Jenniges T, Garcia F, Haden K, Li J, Shaw CA, Belmont J, Sau WC, Shen RM, Barker DL, Gunderson KL. High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. Genome Res, 2006, 16:1136

13. Bignell GR, Huang J, Greshock J, Watt S, Butler A, West S, Grigorova M, Jones KW, Wei W, Stratton MR, Futreal PA, Weber B, Shapero MH, Wooster R. High-Resolution Analysis of DNA Copy Number Using Oligonucleotide Microarrays. Genome Res, 2004, 14:287–94

14. Markello TC, Carlson-Donohoe H, Sincan M, Adams D, Bodine DM, Farrar JE, Vlachos A, Lipton JM, Auerbach AD, Ostrander EA, Chandrasekharappa SC, Boerkoel CF, Gahl WA. Sensitive Quantification of Mosaicism Using High Density SNP Arrays and the Cumulative Distribution Function. Mol Genet Metab, 2012, 105:665

15. Thermo Fisher Scientific. CytosScan Optima Assay Manual Workflow: User Guide, 2018

16. Hoppman N, Rumilla K, Lauer E, Kearney H, Thorland E. Patterns of homozygosity in patients with uniparental disomy: detection rate and suggested reporting thresholds for SNP microarrays. Genetics in Medicine, 2018, 20:1522–7

17. Firth H v., Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, Vooren S van, Moreau Y, Pettett RM, Carter NP. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J Hum Genet, 2009, 84:524–33

18. Rehm HL, Berg JS, Brooks LD, Bustamante CD, Evans JP, Landrum MJ, Ledbetter DH, Maglott DR, Martin CL, Nussbaum RL, Plon SE, Ramos EM, Sherry ST, Watson MS. ClinGen — The Clinical Genome Resource. New England Journal of Medicine, 2015, 372:2235–42

19. Landrum MJ, Chitipiralla S, Brown GR, Chen C, Gu B, Hart J, Hoffman D, Jang W, Kaur K, Liu C, Lyoshin V, Maddipatla Z, Maiti R, Mitchell J, O'Leary N, Riley GR, Shi W, Zhou G, Schneider V, Maglott D, Holmes JB, Kattman BL. ClinVar: improvements to accessing data. Nucleic Acids Res, 2020, 48:D835–44

20. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Jang W, Karapetyan K, Katz K, Liu C, Maddipatla Z, Malheiro A, McDaniel K, Ovetsky M, Riley G, Zhou G, Holmes JB, Kattman BL, Maglott DR. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res, 2018, 46:D1062–7

21. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res, 2005, 33:D514–7

22. Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. Nucleic Acids Res, 2015, 43:D789–98

23. Crolla JA, Cawdery JE, Oley CA, Young ID, Gray J, Fantes J, van Heyningen V. A FISH approach to defining the extent and possible clinical significance of deletions at the WAGR locus. J Med Genet, 1997, 34:207–12

24. Miller RW, Fraumeni JFJr, Manning MD. Association of Wilms's Tumor with Aniridia, Hemihypertrophy and Other Congenital Malformations. Http://DxDoiOrg/101056/NEJM196404302701802, 2010, 270:922–7

25. Palomares M, Delicado A, Mansilla E, de Torres ML, Vallespín E, Fernandez L, Martinez-Glez V, García-Miñaur S, Nevado J, Simarro FS, Ruiz-Perez VL, Lynch SA, Sharkey FH, Thuresson AC,

Annerén G, Belligni EF, Martínez-Fernández ML, Bermejo E, Nowakowska B, Kutkowska-Kazmierczak A, Bocian E, Obersztyn E, Martínez-Frías ML, Hennekam RCM, Lapunzina P. Characterization of a 8q21.11 microdeletion syndrome associated with intellectual disability and a recognizable phenotype. Am J Hum Genet, 2011, 89:295–301

26. Lindsay EA, Greenberg F, Shaffer LG, Shapira SK, Scambler PJ, Baldini A. Submicroscopic deletions at 22q11.2: Variability of the clinical picture and delineation of a commonly deleted region. Am J Med Genet, 1995, 56:191–7

27. Bassett AS, Chow EWC, Husted J, Weksberg R, Caluseriu O, Webb GD, Gatzoulis MA. Clinical features of 78 adults with 22q11 Deletion Syndrome. Am J Med Genet A, 2005, 138:307–13

28. Overhauser J, Huang X, Gersh M, Wilson W, Mcmahon J, Bengtsson U, Rojas K, Meyer M, Wasmuth JJ. Molecular and phenotypic mapping of the short arm of chromosome 5: sublocalization of the critical region for the cri-du-chat syndrome. Hum Mol Genet, 1994, 3:247–52

29. van Buggenhout GJCM, Pijkels E, Holvoet M, Schaap C, Hamel BCJ, Fryns JP. Cri du Chat Syndrome: Changing Phenotype in Older Patients. J Med Genet, 2000, 90:203–15

30. Fang JS, Lee KF, Huang CT, Syu CL, Yang KJ, Wang LH, Liao DL, Chen CH. Cytogenetic and molecular characterization of a three-generation family with chromosome 5p terminal deletion. Clin Genet, 2008, 73:585–90

31. Fernandez BA, Roberts W, Chung B, Weksberg R, Meyn S, Szatmari P, Joseph-George AM, MacKay S, Whitten K, Noble B, Vardy C, Crosbie V, Luscombe S, Tucker E, Turner L, Marshall CR, Scherer SW. Phenotypic spectrum associated with de novo and inherited deletions and duplications at 16p11.2 in individuals ascertained for diagnosis of autism spectrum disorder. J Med Genet, 2010, 47:195–203

32. Roos L, Jønch AE, Kjaergaard S, Taudorf K, Simonsen H, Hamborg-Petersen B, Brøndum-Nielsen K, Kirchhoff M. A new microduplication syndrome encompassing the region of the Miller-Dieker (17p13 deletion) syndrome. J Med Genet, 2009, 46:703–10

33. Brasseur B, Martin CM, Cayci Z, Burmeister L, Schimmenti LA. Bosma arhinia microphthalmia syndrome: Clinical report and review of the literature. Am J Med Genet A, 2016, 170:1302–7

34. Morrow EM. Genomic Copy Number Variation in Disorders of Cognitive Development. J Am Acad Child Adolesc Psychiatry, 2010, 49:1091–104

35. Rodríguez-Revenga L, Vallespín E, Madrigal I, Palomares M, Mur A, García-Miñaur S, Santos F, Mori MÁ, Lapunzina P, Mila M, Nevado J. A parallel study of different array-CGH platforms in

a set of Spanish patients with developmental delay and intellectual disability. Gene, 2013, 521:82–6

36. Bernardini L, Alesi V, Loddo S, Novelli A, Bottillo I, Battaglia A, Digilio MC, Zampino G, Ertel A, Fortina P, Surrey S, Dallapiccola B. High-resolution SNP arrays in mental retardation diagnostics: how much do we gain? European Journal of Human Genetics 2010 18:2, 2009, 18:178–85

37. Chen W, Ullmann R, Langnick C, Menzel C, Wotschofsky Z, Hu H, Döring A, Hu Y, Kang H, Tzschach A, Hoeltzenbein M, Neitzel H, Markus S, Wiedersberg E, Kistner G, van Ravenswaaij-Arts CM, Kleefstra T, Kalscheuer VM, Ropers HH. Breakpoint analysis of balanced chromosome rearrangements by next-generation paired-end sequencing. European Journal of Human Genetics 2010 18:5, 2009, 18:539–43

38. Thiagalingam S, Laken S, Willson JKV, Markowitz SD, Kinzler KW, Vogelstein B, Lengauer C. Mechanisms underlying losses of heterozygosity in human colorectal cancers. Proc Natl Acad Sci U S A, 2001, 98:2698–702

39. Murthy SK, DiFrancesco LM, Travis Ogilvie R, Demetrick DJ. Loss of Heterozygosity Associated with Uniparental Disomy in Breast Carcinoma. Modern Pathology 2002 15:12, 2002, 15:1241–50

40. del Gaudio D, Shinawi M, Astbury C, Tayeh MK, Deak KL, Raca G. Diagnostic testing for uniparental disomy: a points to consider statement from the American College of Medical Genetics and Genomics (ACMG). Genetics in Medicine, 2020, 22:1133–41

41. Dong Z, Chau MHK, Zhang Y, Yang Z, Shi M, Wah YM, Kwok YK, Leung TY, Morton CC, Choy KW. Low-pass genome sequencing-based detection of absence of heterozygosity: validation in clinical cytogenetics. Genet Med, 2021, 23:1225

42. Li X, Liu Y, Yue S, Wang L, Zhang T, Guo C, Hu W, Kagan KO, Wu Q. Uniparental disomy and prenatal phenotype: Two case reports and review. Medicine, 2017, 96

43. Woods CG, Cox J, Springell K, Hampshire DJ, Mohamed MD, McKibbin M, Stern R, Raymond FL, Sandford R, Sharif SM, Karbani G, Ahmed M, Bond J, Clayton D, Inglehearn CF. Quantification of Homozygosity in Consanguineous Individuals with Autosomal Recessive Disease. The American Journal of Human Genetics, 2006, 78:889–96

44. Ramsden DA, Nussenzweig A. Mechanisms driving chromosomal translocations: lost in time and space. Oncogene 2021 40:25, 2021, 40:4263–70

45. Li Y, Roberts ND, Wala JA, Shapira O, Schumacher SE, Kumar K, Khurana E, Waszak S, Korbel JO, Haber JE, Imielinski M, Akdemir KC, Alvarez EG, Baez-Ortega A, Beroukhim R, Boutros PC, Bowtell DDL, Brors B, Burns KH, Campbell PJ, Chan K, Chen K, Cortés-Ciriano I, Dueso-Barroso A, Dunford AJ, Edwards PA, Estivill X, Etemadmoghadam D, Feuerbach L, Fink JL, Frenkel-Morgenstern M, Garsed DW, Gerstein M, Gordenin DA, Haan D, et al. Patterns of somatic structural variation in human cancer genomes. Nature 2020 578:7793, 2020, 578:112–21

46. Markello TC, Carlson-Donohoe H, Sincan M, Adams D, Bodine DM, Farrar JE, Vlachos A, Lipton JM, Auerbach AD, Ostrander EA, Chandrasekharappa SC, Boerkoel CF, Gahl WA. Sensitive quantification of mosaicism using high density SNP arrays and the cumulative distribution function. Mol Genet Metab, 2012, 105:665–71

47. Qin L, Wang J, Tian X, Yu H, Truong C, Mitchell JJ, Wierenga KJ, Craigen WJ, Zhang VW, Wong LJC. Detection and Quantification of Mosaic Mutations in Disease Genes by Next-Generation Sequencing. J Mol Diagn, 2016, 18:446–53

48. Yurov YB, Vorsanova SG, Iourov IY, Demidova IA, Beresheva AK, Kravetz VS, Monakhov V v, Kolotii AD, Voinova-Ulas Y. Unexplained autism is frequently associated with low-level mosaic aneuploidy. J Med Genet, 2007, 44:521–5

49. Pham J, Shaw C, Pursley A, Hixson P, Sampath S, Roney E, Gambin T, Kang SHL, Bi W, Lalani S, Bacino C, Lupski JR, Stankiewicz P, Patel A, Cheung SW. Somatic mosaicism detected by exon-targeted, high-resolution aCGH in 10,362 consecutive cases. Eur J Hum Genet, 2014, 22:969–78

50. Graham JM, Lee J. Bosma arhinia microphthalmia syndrome. Am J Med Genet A, 2006, 140:189–93

51. Shaw ND, Brand H, Kupchinsky ZA, Bengani H, Plummer L, Jones TI, Erdin S, Williamson KA, Rainger J, Stortchevoi A, Samocha K, Currall BB, Dunican DS, Collins RL, Willer JR, Lek A, Lek M, Nassan M, Pereira S, Kammin T, Lucente D, Silva A, Seabra CM, Chiang C, An Y, Ansari M, Rainger JK, Joss S, Smith JC, Lippincott MF, Singh SS, Patel N, Jing JW, Law JR, Ferraro N, et al. SMCHD1 mutations associated with a rare muscular dystrophy can also cause isolated arhinia and Bosma arhinia microphthalmia syndrome. Nature Genetics 2017 49:2, 2017, 49:238–48

52. Mason-Suares H, Kim W, Grimmett L, Williams ES, Horner VL, Kunig D, Goldlust IS, Wu BL, Shen Y, Miller DT, Martin CL, Rudd MK. Density matters: Comparison of array platforms for detection of copy-number variation and copy-neutral abnormalities. Genetics in Medicine, 2013, 15:706–12

53. Shen Xuan, Goidin Didier, Kim Young, Happe Scott. Technical overview of In silico Probe Coverage Comparision Between Agilent GentiSure Cyto vs CytoScan Microarrays. Santa Clara, Carlifornia, 2021

54. Levy B, Wapner R. Prenatal diagnosis by chromosomal microarray analysis. Fertil Steril, 2018, 109:201–12

55. Kaminsky EB, Kaul V, Paschall J, Church DM, Bunke B, Kunig D, Moreno-De-Luca D, Moreno-De-Luca A, Mulle JG, Warren ST, Richard G, Compton JG, Fuller AE, Gliem TJ, Huang S, Collinson MN, Beal SJ, Ackley T, Pickering DL, Golden DM, Aston E, Whitby H, Shetty S, Rossi MR, Rudd MK, South ST, Brothman AR, Sanger WG, Iyer RK, Crolla JA, Thorland EC, Aradhya S, Ledbetter DH, Martin CL. An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. Genetics in Medicine, 2011, 13:777–84

56. Benedek K, Chudoba I, Klein G, Wiener F, Mai S. Rearrangements of the telomeric region of mouse chromosome 11 in Pre-B ABL/MYC cells revealed by mBANDing, spectral karyotyping, and fluorescence in-situ hybridization with a subtelomeric probe. Chromosome Res, 2004, 12:777–85

57. Pinkel D, Straume T, Gray JW. Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. Proc Natl Acad Sci U S A, 1986, 83:2934–8

58. Yakut T, Sara AE, Kilic S, Ergun AE, Esra C, Ae Y, Egeli U. FISH investigation of 22q11.2 deletion in patients with immunodeficiency and/or cardiac abnormalities. Pediatr Surg Int, 2006, 22:380–3

59. Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ, Faucett WA, Feuk L, Friedman JM, Hamosh A, Jackson L, Kaminsky EB, Kok K, Krantz ID, Kuhn RM, Lee C, Ostell JM, Rosenberg C, Scherer SW, Spinner NB, Stavropoulos DJ, Tepperberg JH, Thorland EC, Vermeesch JR, Waggoner DJ, Watson MS, Martin CL, Ledbetter DH. Consensus Statement: Chromosomal Microarray Is a First-Tier Clinical Diagnostic Test for Individuals with Developmental Disabilities or Congenital Anomalies. Am J Hum Genet, 2010, 86:749

60. Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ, Faucett WA, Feuk L, Friedman JM, Hamosh A, Jackson L, Kaminsky EB, Kok K, Krantz ID, Kuhn RM, Lee C, Ostell JM, Rosenberg C, Scherer SW, Spinner NB, Stavropoulos DJ, Tepperberg JH, Thorland EC, Vermeesch JR, Waggoner DJ, Watson MS, Martin CL, Ledbetter DH. Consensus statement: chromosomal microarray is a first-tier clinical

diagnostic test for individuals with developmental disabilities or congenital anomalies. Am J Hum Genet, 2010, 86:749–64

61. Wiszniewska J, Bi W, Shaw C, Stankiewicz P, Kang SHL, Pursley AN, Lalani S, Hixson P, Gambin T, Tsai CH, Bock HG, Descartes M, Probst FJ, Scaglia F, Beaudet AL, Lupski JR, Eng C, Wai Cheung S, Bacino C, Patel A. Combined array CGH plus SNP genome analyses in a single assay for optimized clinical testing. European Journal of Human Genetics, 2014, 22:79–87

62. Oğuz S, Arslan UE, Kiper PÖŞ, Alikaşifoğlu M, Boduroğlu K, Utine GE. Diagnostic yield of microarrays in individuals with non-syndromic developmental delay and intellectual disability, 2021. https://doi.org/10.1111/jir.12892

63. Shaw-Smith C, Redon R, Rickman L, Rio M, Willatt L, Fiegler H, Firth H, Sanlaville D, Winter R, Colleaux L, Bobrow M. Microarray based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. J Med Genet, 2004, 41:241–8

64. Kennedy GC, Matsuzaki H, Dong S, Liu WM, Huang J, Liu G, Su X, Cao M, Chen W, Zhang J, Liu W, Yang G, Di X, Ryder T, He Z, Surti U, Phillips MS, Boyce-Jacino MT, Fodor SPA, Jones KW. Large-scale genotyping of complex DNA. Nat Biotechnol, 2003, 21:1233–7

65. Wang JC, Ross L, Mahon LW, Owen R, Hemmat M, Wang BT, el Naggar M, Kopita KA, Randolph LM, Chase JM, Aguilera MJM, Siles JL, Church JA, Hauser N, Shen JJ, Jones MC, Wierenga KJ, Jiang Z, Haddadin M, Boyar FZ, Anguiano A, Strom CM, Sahoo T. Regions of homozygosity identified by oligonucleotide SNP arrays: evaluating the incidence and clinical utility. European Journal of Human Genetics, 2015, 23:663

66. Levy B, Burnside RD. Are all chromosome microarrays the same? What clinicians need to know. Prenat Diagn, 2019, 39:157–64

17. Appendices

Appendix 1

NATIONAL HEALTH LABORATORY SERVICE
School of Pathology, University of the Witwatersrand

DIVISION OF HUMAN GENETICS

Hospital Street, Johannesburg, 2001 I PO Box 1038, Johannesburg, 2000
[T] +27 11 489 9223 I [M] +27 78 080 8841 I [F] +27 11 489 9226 I [E] human.genetics@nhls.ac.za

Date: 25 February 2022

I wish to apply for permission to use data/samples from the Division of Human Genetics covered under the Ethics Clearance Certificate M180506:

Using anonymised residual diagnostic samples and –file data for audit, research and development:

Name of Staff member/Student:	Minenhle Mayisela
Staff/ Student Number:	1880919
Degree/Non-degree purposes	MSc Genomic Medicine
E-mail address:	1880919@students.wits.ac.za
Contact Number:	0787888219

Name of Supervisor (if PI is a student): Nkateko Mayevu

Staff/ Student Number:	NHLS: 6247
E-mail address:	nkateko.mayevu@nhls.ac.za
Contact Number:	(011) 489 9237

My study title is: Evaluating a SNP based array for diagnostic copy number variation detection - a comparative study

I require access to the following samples/patient files (Give numbers and description):

I require 10 samples with complex structural chromosomal rearrangements identified using CGH microarray (Agilent). The data from the samples will be compared with newly generated data from the SNP array platform (Affymetrix). The resolution of the two microarray platforms will be assessed.

I will use samples/data covered under the following Section of the above mentioned Ethics clearance:

Category D/E

Retrospective database of clinical patient records and residual diagnostic samples:

> **Category A**: Existing patient clinical records from 1970s to 1 September 2018, for which no consent has been obtained
>
> **Category B**: Existing residual diagnostic samples from mid 1980s for which no consent has been obtained

Prospective database of clinical patient records and residual diagnostic samples (after 1 September 2018)

> **Category C**: Clinical records of patients who presented to Human Genetics Clinics after 1 September 2018 from whom written consent is obtained.
>
> **Category D**: Residual Diagnostic samples from patients who presented to Human Genetics Clinics after 1 September 2018 from whom written consent is obtained.
>
> **Category E**: Residual Diagnostic samples from patients referred to the Division of Human Genetics from whom consent has not be obtained I attach a summary of my main project objectives.

I am aware that if I may still need to apply to the Human Ethics Research Committee for approval for my research project

Signature: M.M

Name: Minenhle Mavisela

Supervisor signature:

Name: Nkateko Mayevu

Approved:

Amanda Krause, MBBCh, PhD
Medical Geneticist/Associate Professor
Head of Division

2

Appendix 2

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

Office of the Deputy Vice-Chancellor (Research and Innovation)

TO:

Ms M Mayisela
School of Pathology
Division of Human Genetics
Medical School
University

E-mail: 1880919@students.wits.ac.za

CC:

Supervisor: Ms N Mayevu and Dr F Baine-Savanhu
<Nkateko@Mayevu@nhls.ac.za>
and <HREC-Medical Research Office@wits.ac.za>

FROM:

Mr Iain Burns
Human Research Ethics Committee (Medical)
Tel: 011 717 1252

E-mail: Iain.Burns@wits.ac.za

DATE: 2022/03/16

REF: R14/49

PROTOCOL NO: **M220295** (This is your ethics application reference number. Please quote it in all enquiries, oral or written, relating to this study.)

PROJECT TITLE: *Chromosomal re-arrangements investigated on two microarray platforms: comperative study*

Please find attached the Clearance Certificate for the above project. I hope it goes well and that an article in a recognized publication comes out of it. This will reflect well on your professional standing and contribute to Government funding of the University.

MSWorks2000/Iain2007/Clearcon.wps

Appendix B. Ethics clearance

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

R49 Ms M Mayisela

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M220295

NAME: (Principal Investigator)	Ms M Mayisela
DEPARTMENT:	School of Pathology Division of Human Genetics Medical School University
PROJECT TITLE:	*Chromosomal re-arrangements investigated on two microarray platforms: comparative study*
DATE CONSIDERED:	Ad hoc
DECISION:	Approved unconditionally
CONDITIONS:	Sub-study under M180506 Prior condition (Faculty protocol approval) satisfied on 2022/07/13
NOTE:	If contact information regarding student study participants is required, please contact the Registrar's office - <Nicoleen.Potgieter@wits.ac.za>
SUPERVISOR:	Ms N Mayeyu and Dr F Baine-Savanhu
APPROVED BY:	Dr CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	2022/03/16

This Clearance Certificate is valid for 5 years from the date of approval. An extension may be applied for.

DECLARATION OF INVESTIGATORS

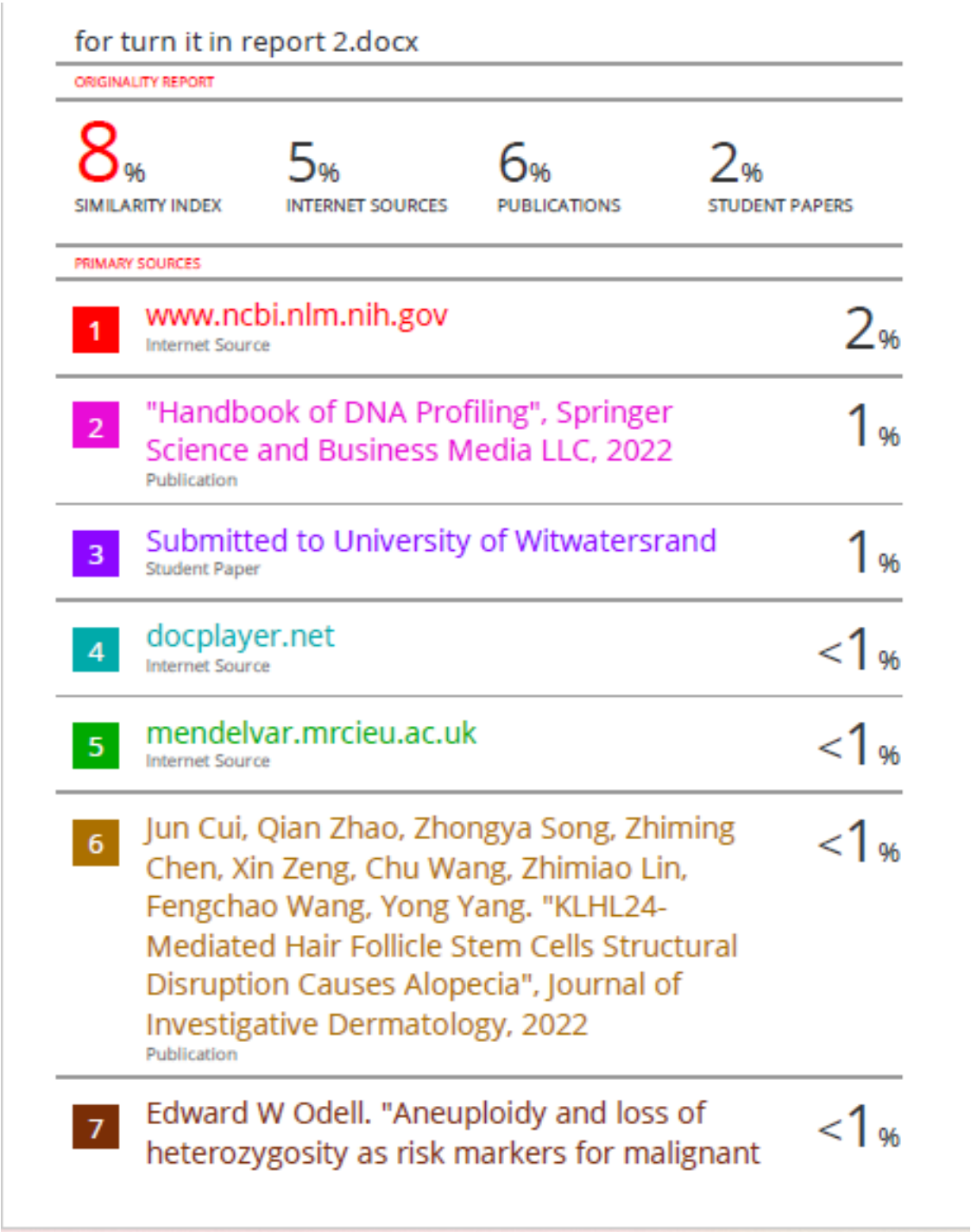
To be completed in duplicate and **ONE COPY** returned to the Research Office secretariat on the 3rd floor, Phillip Tobias Building, Parktown, University of the Witwatersrand, Johannesburg.

I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated from the research protocol as approved, I/we undertake to submit details to the Committee. I agree to submit a yearly progress report. When a funder requires annual re-certification, the application date will be one year after the date when the study was initially reviewed. In this case, the study was initially reviewed in **February** and therefore reports and re-certification will be due in the month of Feb. Unreported changes to the study may invalidate the clearance given by the HREC (Medical).

18 August 2022

Signature of Principal Investigator — Date

Appendix C. Turn-it-in report



for turn it in report 2.docx

ORIGINALITY REPORT

8%	5%	6%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.ncbi.nlm.nih.gov Internet Source	2%
2	"Handbook of DNA Profiling", Springer Science and Business Media LLC, 2022 Publication	1%
3	Submitted to University of Witwatersrand Student Paper	1%
4	docplayer.net Internet Source	<1%
5	mendelvar.mrcieu.ac.uk Internet Source	<1%
6	Jun Cui, Qian Zhao, Zhongya Song, Zhiming Chen, Xin Zeng, Chu Wang, Zhimiao Lin, Fengchao Wang, Yong Yang. "KLHL24-Mediated Hair Follicle Stem Cells Structural Disruption Causes Alopecia", Journal of Investigative Dermatology, 2022 Publication	<1%
7	Edward W Odell. "Aneuploidy and loss of heterozygosity as risk markers for malignant	<1%

transformation in oral mucosa", Oral Diseases, 2021
Publication

8	Gulrukh Begum. "Consanguinity: A Form of Social Capital Among the Muslims of Char Areas in Barpeta District, Assam", Journal of the Anthropological Survey of India, 2021 Publication	<1 %
9	Daniela del Gaudio, Marwan Shinawi, Caroline Astbury, Marwan K. Tayeh, Kristen L. Deak, Gordana Raca. "Diagnostic testing for uniparental disomy: a points to consider statement from the American College of Medical Genetics and Genomics (ACMG)", Genetics in Medicine, 2020 Publication	<1 %
10	Xudong Cai, Bo Fan, San H. Thang, Calum J. Drummond, Nhiem Tran, Jiali Zhai. "Paclitaxel-loaded cubosome lipid nanocarriers stabilised with pH and hydrogen peroxide-responsive steric stabilisers as drug delivery vehicles", Journal of Materials Chemistry B, 2023 Publication	<1 %
11	www.researchgate.net Internet Source	<1 %
12	phcfm.org Internet Source	<1 %

13	wiredspace.wits.ac.za Internet Source	<1 %
14	pure.uva.nl Internet Source	<1 %
15	www.nature.com Internet Source	<1 %
16	journals.lww.com Internet Source	<1 %
17	research-publishing.net Internet Source	<1 %
18	www.linknovate.com Internet Source	<1 %
19	www.mdpi.com Internet Source	<1 %
20	Christopher W. Whelan, Robert E. Handsaker, Giulio Genovese, Seva Kashin et al. "Insights into dispersed duplications and complex structural mutations from whole genome sequencing 706 families", Cold Spring Harbor Laboratory, 2020 Publication	<1 %
21	Ho, Rosettia. "Genetic Determinants Underlying Rare Diseases Identified Using Next-Generation Sequencing Technologies",	<1 %

	The University of Western Ontario (Canada), 2022 Publication	
22	Molecular Diagnostics, 2014. Publication	<1 %
23	eje.bioscientifica.com Internet Source	<1 %
24	spandidos-publications.com Internet Source	<1 %
25	www.eurofinsclinicalgenetics.co.in Internet Source	<1 %
26	Marília Moreira Montenegro, Débora Camilotti, Caio Robledo D'Anglioli Costa Quaio, Yanca Gasparini et al. "Expanding the Phenotype of 8p23.1 Deletion Syndrome: Eight New Cases Resembling the Clinical Spectrum of 22q11.2 Microdeletion", The Journal of Pediatrics, 2023 Publication	<1 %
27	Helen V. Firth, Shola M. Richards, A. Paul Bevan, Stephen Clayton et al. "DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources", The American Journal of Human Genetics, 2009 Publication	<1 %

Appendix D. Author guidelines for the Journal of Molecular Diagnostics

The Journal of Molecular Diagnostics *(JMD)*, the official journal of the Association for Molecular Pathology (AMP), co-owned by the American Society for Investigative Pathology (ASIP) and published by Elsevier, Inc., seeks to publish high-quality original papers on scientific advances in the field of molecular diagnostics, the translation and validation of molecular discoveries in medicine into the clinical diagnostic setting, and the description and application of technical advances in the field of molecular diagnostic medicine. The editors welcome for review manuscripts that contain: novel discoveries with direct application to clinical diagnostics or clinicopathologic correlations including studies in oncology, infectious diseases, inherited diseases, predisposition to disease, clinical informatics, or the description of polymorphisms linked to disease states or normal variations; the application of diagnostic methodologies in clinical trials; or the development of new or improved molecular methods which may be applied to diagnosis or monitoring of disease predisposition.

Manuscript Categories. In addition to **Regular Articles**, the Journal also publishes **Technical Advances**, which are intended to report breakthroughs in methodology or analytical tools applied to molecular diagnostics. Submissions to this category should include a detailed description of the methodological design and discussion of how this technique improves the practice of molecular diagnostics. There are no official limits to word count or number of tables and figures that can be included in for Regular Articles or Technical Advances; authors should consider their budgetary restrictions when preparing their manuscript. The Editors reserve the right to request reduction of text length or removal of data.

The Journal also features **Reviews, Mini-Reviews, Commentaries, Special Articles, Practice Advances, Guest Editorials,** and **Perspectives** solicited by the Editor-in-Chief, as well as AMP meeting abstracts and other articles of interest to the membership of AMP. Authors who wish to publish a **Review** or **Mini-Review** should send their curriculum vitae along with an outline of the proposed article for prior approval by the Editor-in-Chief; proposals should be emailed to jmd@asip.org. Reviews have a maximum of 5,000 words, 75 references, and 4 tables or figures (at

least one of which is mechanistic). Mini-Reviews have a maximum of 3,000 words, 40 references, and 2 tables or figures (one of which is mechanistic). Supplemental data should not be overly utilized to expand beyond these listed parameters. **Practice Advances** are submitted from AMP Education programs to highlight webinar offerings from the same group. This article format is not currently open to external submissions. **Commentaries** are submitted by invitation only and relate to a concurrently published manuscript. **Guest Editorials** and **Perspectives** are published by invitation only.

Review Process. The Editors perform an initial evaluation on all submissions to determine whether they believe the manuscript is within the scope of *JMD* and will also achieve a sufficient priority score to warrant publication. Priority is determined by the Editors' assessment of the manuscript relative to other papers being considered. The aim is that expeditious treatment will enable authors whose manuscripts do not meet these criteria to submit their manuscript elsewhere as soon as possible without unnecessary delays. For manuscripts accepted for external review, the Editor-in-Chief assigns an Associate Editor according to their expertise. The Associate Editor will solicit reviewers (typically, two external reviews are sought). Authors are encouraged to propose potential reviewers during the submission process. Authors may also request that specific reviewers not be used due to prior collaborations, known conflicts of interest, or direct competition. The Editors will make every effort to respect requests that are well-founded; however, the Editors do have the authority to utilize such a reviewer if it is necessary for expert peer review. It is the practice of the Journal to conduct a blinded peer-review process. The peer-review process is kept completely confidential; it is considered a violation of this confidentiality for authors to identify or attempt to communicate directly with peer reviewers or Associate Editors regarding their manuscript. All editorial communications should be directed through the Editorial Office at jmd@asip.org. The reviewer comments and Associate Editor's recommendation are evaluated by the Editor-in-Chief for disposition and transmittal to the authors. Every effort is made to complete the review process within 35 days of the date received.

Review Decisions. Only a portion of manuscripts will be accepted for publication. A number of worthy manuscripts will be rejected based on priority or for being out of scope. The Journal will advise authors whether the manuscript is accepted, eligible

for resubmission after minor or major revisions, or rejected. A manuscript may be returned to the authors without outside review if the Editors find it inappropriate for publication in this Journal. **Appeals** (rebuttals) to rejection decisions must be submitted within 60 days to be considered.

Journal Scientific Integrity Policy. The Journal has developed principles for defining proper conduct and scientific misconduct as well as procedures for handling such matters. General guidelines are described below; detailed information can be viewed at *https://www.jmdjournal.org/content/integrity*. To report suspected misconduct relating to authors, reviewers, or Editors, send written complaint to the Editorial Office at *The Journal of Molecular Diagnostics*, 1801 Rockville Pike, Suite 350, Rockville, Maryland 20852-9975 or jmd@asip.org. Issues relating to staff conduct should be directed to the ASIP Executive Officer at the American Society for Investigative Pathology, 1801 Rockville Pike, Suite 350, Rockville, Maryland 20852-9975 or WBColeman@asip.org.

Authors

Authorship. Authorship is defined as 1) substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data; 2) drafting the manuscript or revising it critically for important intellectual content; 3) final approval of the version to be published; and 4) agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Authors should meet all four conditions. When a group serves as a coauthor, the list of members (individuals or institutions) should be included in the Acknowledgements. When submitting a manuscript to the Journal, the corresponding author takes responsibility on behalf of all authors for the authorship, authenticity, and integrity of the research being reported. A current and functional email address needs to be provided for each author at submission, professional and institutional email addresses are strongly preferred. All coauthors will receive a notification email at the time of submission that requests their confirmation to be affiliated with the manuscript. Authors added at revision also need an email address provided in the system and will be sent the confirmation email request. Any change made to the list of authors (addition, removal, change in order) after manuscript acceptance requires consent of all authors and editorial approval.

Author Conduct. Authors should take special care that manuscripts submitted to the Journal are prepared in accordance with the Uniform Requirements for Manuscripts Submitted to Biomedical Journals (see *http://www.icmje.org*); in particular the ethical considerations regarding authorship, conflicts of interest, redundant publication, and treatment and confidentiality of research subjects should be carefully adhered to. Additionally, the Journal takes great care to secure the confidentiality and integrity of the peer-review process; it is considered a violation of this confidentiality for authors to identify or attempt to communicate directly with peer reviewers or Associate Editors regarding their manuscript. All editorial communications should be directed through the Editorial Office at jmd@asip.org. The Editors will consider any deliberate ethical violation in either the reported research or the manuscript preparation and review to be actionable misconduct, the potential results of which may be manuscript rejection or public article retraction, reporting of conduct to the authors' governing institutions, and/or the denial to consider any future submissions to the Journal. Willful misconduct does not include incidents of honest misjudgment or inadvertent error. Detailed information regarding possible misconduct can be found in the Journal Scientific Integrity Policy at *https://www.jmdjournal.org/content/integrity*.

Article Transfer Service This journal is part of our Article Transfer Service. This means that if the Editor feels your article is more suitable in one of our other participating journals, then you may be asked to consider transferring the article to one of those. If you agree, your article will be transferred automatically on your behalf with no need to reformat. Please note that your article will be reviewed again by the new journal. More information.

Manuscript Submission. Manuscripts should be submitted online via the Editorial Manager system (*https://www.editorialmanager.com/jmdi/default.aspx*). Detailed instructions on preparing and submitting files can be found on the author submission website at the above URL. Authors having difficulty submitting files online should contact the editorial office for assistance. Manuscript text, figures, tables, and supplemental data files can be emailed to the Editorial Office at jmd@asip.org. The cover letter must state any conflicts of interest (both financial and personal), affirm that the manuscript has not been published previously and is not being considered concurrently by another publication, and affirm that all authors and acknowledged contributors have read and approved the manuscript. Submissions will be eligible for

review when the manuscript has not been previously published and is not under copyright at another publication (print, online, or in a non-English journal). Previous presentation of an abstract, poster, webinar, or oral presentation is permissible. Hosting of a manuscript on a pre-print server is not preferred; however, it may not exclude a manuscript from consideration at *JMD*. These may be handled on a case by case basis and authors should disclose in their cover letter if their manuscript is already available as a pre-print.

Manuscript Preparation. Manuscripts should be prepared in the style of the Journal and in accordance with The Uniform Requirements for Manuscripts Submitted to Biomedical Journals (see *http://www.icmje.org*). The required file format for text is Microsoft Word. Standard abbreviations can be found in the CSE Style Manual: *Scientific Style and Format* (8th ed., 2014). All pages of the manuscript should be double-spaced and numbered (including references, tables and figure legends); line numbers should also be included to assist reviewers in making comments. Other formatting specifications (eg, font size and type, margin settings, etc) are left to the authors' discretion, as papers accepted for publication will be reformatted according to the print specifications of the Journal. Manuscripts not prepared in accordance with the submission guidelines detailed below may be returned to the authors. Authors are encouraged to include a list of nonstandard abbreviations to aid reviewers; however, the Journal does not publish abbreviations or keywords as part of the final article. In addition, authors should be ready to comply with Editors' requests for copies of any similar works in preparation, copies of cited manuscripts that are submitted or in press, and/or supporting manuscript data (eg, data not shown, but summarized in the manuscript) that may aid the review process.

Title Page. The title page must include a concise title accurately reflecting the findings of the work. Manuscript titles should not be formatted as a question and single word titles with a descriptive sub-title are not preferred. Full names (not initials) of all authors; department, institution and address where each author was affiliated at the time the research was performed; number of text pages, tables and figures; a short running head (40 characters or less); grant numbers and sources of support (including departmental or institutional funding when no extramural funding was received); name, mailing address, phone, fax, and email address of the corresponding author(s). Departmental affiliations should be in English. The Journal

allows for two co-corresponding authors maximum. If an author changes employment after the study was performed, the new affiliation information for that author should be included as a footnote. Any relationships (eg, employment, consultancies, board membership, stock ownership, funding, honoraria, expert testimony, patents or royalties, travel reimbursements, industry-supplied free reagents, etc) with any organization or entity having a direct financial or personal interest in the subject matter or materials discussed in the article should also be clearly stated. Disclosures not directly related tothe manuscript are also encouraged to be included for transparency.

Abstract. An abstract of 220 words or less should be prepared on a separate page and should be intelligible to the general reader without reference to the text. The abstract should clearly summarize the background, methodology, results, and significance of the study. Structured abstracts are not Journal style and will need to be revised to eliminate section headings. Abbreviations, external website links, and citations should be avoided.

Text. The remaining sections of the text, which should include Introduction, Materials and Methods, Results, and Discussion (in this order), need not begin on new pages. Conclusions, Acknowledgements and Author Contributions are optional sections that appear after the Discussion if authors choose to include them. Commonly-abbreviated terms should be spelled out in their first occurrence and then may be referenced in abbreviation through the remainder of the manuscript. The Journal requires the use of official gene and protein symbols, to facilitate standardization of scientific communication. Consult the Human Genome Organization Gene Nomenclature Committee website (*http://www.genenames.org*) for gene names and symbols and UniProt (*http://www.uniprot.org*) for protein names and symbols. For standard mutation nomenclature, refer to Ogino et al, J Mol Diagn 2007 doi:10.2353/jmoldx.2007.060081. For a complete list of other approved nomenclature organizations (eg, bacteria, viruses, mice), please contact the Editorial Office. Note that genes, loci, alleles, and genotypes are styled as italic; proteins, gene products, phenotypes, mRNAs, and cDNAs are not.

In Regular Articles and Technical Advances first-person personal pronouns (I, me, my, we, us, our) should be kept to a minimum. Replace active verbs with passive verbs, especially with regard to eliminating first-person personal pronouns. Use

present tense for figure legends since the data is static and is viewable now (except for describing what was done in an experiment) but past tense in the Results because this describes results that WERE obtained in the past. Headings in the Results section should be in present tense.

For **Materials and Methods,** authors should describe experimental and statistical methods in enough detail that other researchers can replicate results and evaluate claims. enough detail that other researchers can replicate results and evaluate claims. Authors are encouraged to review the NIH Principles and Guidelines for Reporting Preclinical Research (see *http://www.nih.gov/about/reporting-preclinical-research.htm*). Inclusion nclusion of method or reagent details as supplementary material is not acceptable. In addition, linking to protocols online is not permissible as there is no guarantee that the information will remain unchanged and accessible in perpetuity. The sequences of oligonucleotides, if not previously published, should be provided. Novel DNA or protein sequences should be deposited to an appropriate database (eg, Genbank, EMBL, SWISS-PROT), with the accession numbers included in the manuscript. When providing supplier information for materials sources, company name and location (city and state, or city and country) should be provided. When describing reagents such as antibodies, cell lines, animal strains, bacteria, and viruses, authors should include the source, characteristics, dilutions, strain, species, sex, authentication, etc. as necessary for repetition of the experiments. Website references to company or commercial information are not permitted. All novel materials and the procedures to prepare them should be described in sufficient detail to allow their reproduction (eg, DNA constructs, analytical software). Materials that are approved for investigational-use only should be clearly indicated.

Methods should state whether sample size was determined statistically prior to experimentation, whether samples were randomized (and how), whether data acquisition was blinded (particularly for subjective scoring methodologies), and what criteria were used to include/exclude data points or subjects. Experimental procedures should include the number of replicates performed and the number of samples in each experimental condition. A careful power analysis undertaken at the beginning of work can help assure the significance of study findings. "Discrepant analysis" is regarded as statistically biased and should not be employed for

assessment of assay sensitivity and specificity. Reports of assay sensitivity and specificity should include appropriately calculated confidence intervals. Special care should be taken to assure that statistical methods are appropriate, with clearly defined statements of the statistical test(s) used, sample size, and measures reported (eg, mean, median, SD, SEM, confidence intervals). Statistical methods may be assessed by a professional statistician as part of the peer review process.

Publication in the Journal implies that the authors agree, upon reasonable request, to share any materials or data that are integral to the results presented in the article, including whatever would be necessary for a skilled investigator to verify or replicate the claims. This may include original software code used in the data analysis. Agreement to share reagents or software code does not preclude the authors from implementing a Data Use Agreement. Authors must disclose upon submission any restrictions on the availability of materials or information, such as for patented or dual-purpose materials. Reporting guidelines for specific study designs (eg, randomized controlled trials) can be found online via the Enhancing the QUAlity and Transparency Of health Research (EQUATOR) network (see *http://www.equator-network.org/reporting-guidelines*). Authors must affirm that the research protocol was approved by the appropriate institutional review boards or ethics committees for human (including use of human cells or tissues) or animal experiments and that all human subjects provided appropriate informed consent and/or that regulations concerning the use of animals in research were adhered to. If race/ethnicity is reported, authors should state who determined race/ethnicity, how the options were defined, and why race/ethnicity was important in the study. Authors should be prepared to provide study protocol number(s) if requested.

Manuscripts that meet the criteria set forth by **STARD** (Standards for Reporting of Diagnostic Accuracy) will be annotated with a special footnote on the title page. Authors who wish their manuscripts to receive this designation should submit a completed checklist (available at *http://www.equator-network.org/reportingguidelines/stard*) at the time of submission for evaluation by the Editors at the time of acceptance. Further information about STARD can be found in Bossuyt PM et al, STARD 2015: an updated list of essential items for reporting diagnostics accuracystudies. BMJ 2015; 351: h5527.

Authors should obtain permission from all individuals named (by full name) in the **Acknowledgments** who contributed substantially to the work reported (eg, data collection, analysis, or writing/editing assistance) but did not fulfill the authorship criteria. Likewise, authors should receive permission from all individuals named as sources for personal communication or unpublished data. Such permissions should be affirmed by the corresponding author in the cover letter. The list of members of a group author should also be included in this section. A list of individual **Author Contributions** may be placed at the end of the manuscript. Author initials should be used, and authors should be grouped according to applicable responsibility with each responsibility separated by a semi-colon. "A.B., C.D., and Y.Z. performed experiments and analyzed the data; W.X. and Y.Z. secured funding and designed the experiment." The corresponding author is also encouraged to include a statement of ethical assurance, such as, "X.X. is the guarantor of this work and, as such, had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis."

References. References should begin on a new page, be double-spaced and numbered in order of citation in the text, including citations in tables and figure legends. Citations that first appear in tables, figures, or supplemental data should be numbered according to the item's first call out in the text; a separate reference list should not be prepared for supplemental data. Complete author citation is required, use of "et al" is only acceptable for sources with more than 35 authors. In this instance, list the first 6 author names and then et al. References should conform to the style of the Journal. Examples follow:

Journals: Gong JZ, Cook JR, Greiner TC, Hedvat C, Hill CE, Lim MS, Longtine JA, Sabath D, Wang YL: Laboratory Practice Guidelines for Detecting and Reporting JAK2 and MPL Mutations in Myeloproliferative Neoplasms: A Report of the Association for Molecular Pathology. J Mol Diagn 2013, 15:733-744

Electronic Journals: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group: Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 2009, 6:e1000097. http://dx.doi.org/10.1371/journal.pmed.1000097

Preprints: Metsky H, Freije C, Kosoko-Thoroddsen T-S, Sabeti P, Myhrvold C: CRISPR-based surveillance for COVID-19 using genomically-comprehensive machine learning design. bioRxiv 2020, [Preprint] doi: 10.1101/2020.02.26.967026

Books: Funkhouser WK: Pathology: The Clinical Description of Human Disease. Edited by Coleman WB and Tsongalis GJ. Burlington, MA, Academic Press, 2010, pp. 137-142

In Press: To be used only for papers accepted for publication. Cite according to previous instructions with (in press) in place of volume and page numbers. Include digital object identifier (DOI) when available for online early publications (Epub ahead of print). Accepted manuscripts that are not yet in press (no DOI available) should not be included in references and are considered unpublished data.**Submitted Papers/Unpublished Data:** Cite in text only as (unpublished data).

Product Inserts: Cite in text only: (Affymetrix technical note: Globin Reduction Protocol: A Method for Processing Whole Blood RNA Samples for Improved Array Results. Santa Clara, CA).

Web sites: Cite in text only. See **Data Supplements** section below for proper use of web site references. Use the DOI when available. Include the name of the institution sponsoring the web site, URL address with direct linkage to the referenced information, and date of last access.

Data References: This journal encourages you to cite underlying or relevant datasets in your manuscript by citing them in your text and including a data reference in your Reference List. Data references should include the following elements: author name(s), dataset title, data repository, version (where available), year, and global persistent identifier. Add [dataset] immediately before the reference so we can properly identify it as a data reference. The [dataset] identifier will not appear in your published article.

Tables. Tables should be typed double-spaced and submitted after the main text on separate pages, as part of the manuscript. The preferred file format for Tables is MS Word; figure file formats (including those embedded in the text) are unacceptable. Tables should be black and white text only and should not include figures or other non-typeset images. Color or gray shading is not permitted. Emphasis may be

represented by bold, italics, and underlining which is explained in the table legend. Nucleotide sequences should be capitalized, not lowercase. Table footnotes should use the following sequential symbols: *, †, ‡, §, ‖, ¶ these may be doubled up if needed. All symbols and abbreviations in the table should be defined in the footnotes.

Figures. Authors should give considerable care in preparing figures. Resolution and quality of submitted images is the responsibility of the author; as a routine, the Journal does not provide figure enhancement services. Images (such as graphs and schematics) should have a white background; color, black, or gray is generally unacceptable. Authors are encouraged to carefully consider whether bar or line graphs should be submitted in grayscale or color, paying special attention to the ability of a reader to distinguish between different data points. **It is preferred that graphs be published in grayscale unless color is necessary for reader understanding**. Patterns should not be used to fill in bars in bar graphs. Authors should refrain from using three-dimensional bars for single sets of data. All labels in Figures need to be sized so they will be legible in print. Figures may consist of multiple related panels (labeled A, B, C, etc) described under one figure legend. Each figure (with all of its related panels) should be arranged on a single page, with panels reading from left to right, as it should appear in final publication. Text labels in figures should use a professional looking font such as Arial; Comic Sans should be avoided. Figure labels should appear in the upper left corner as uppercase letters; refrain from using sub-labels (Aa, Ab, Ai, Aii, A1, A2, etc) or titles for each panel label; panels should be labeled chronologically from left to right. Figure panels submitted on separate pages will be arranged at the publisher's discretion. Figures should be sized to fit one column (8 cm) or two columns (17 cm). Maximum page length is 22.5 cm. Figures deviating from these dimensions will be sized at the publisher's discretion. Unwanted background material should be excluded, and edges should be straight. Scale bar measures included in images need to be legible to the reader and be replicated in the figure legend. If a scale bar is used, scale bars must be present on every applicable image. Ensure that numbers >10,000 are set with commas. Add spaces before unit of measure (eg, 2kDa to 2 kDa). Set with space on either side of math symbols (eg, 2 + 2 = 4; not 2+2=4). Text in Figures can be capitalized as title case, or initial cap only, but must be consistent throughout all figures. X and Y axes should be aligned, when possible. Ensure that X and Y axis

labels are phrased consistently: eg, "Time, days"; or "Time (days)"; avoid using simply "Days." Ensure that sample size info, *P* values, scale bars, and original magnification are removed from the image and summarized at the end of the legend.

***Figure Legends*.** Figure legends should be submitted in a list as part of the manuscript (separate from the figure files) and should adequately describe all descriptors: arrows, arrowheads, scale bars, insets, asterisks, boxes/circles/etc. in line graphs, and any other notations. Histology images should have all staining methods described, and either the magnification or the measure of the scale bar in the image should also be included. Descriptive text, rather than graphics inserted into the text, should be used when possible (eg, closed circles, open boxes, etc.); alternatively, a legend key should be included in the figure. Statistical significance (eg, *P* values) should be clearly defined by asterisks (*, **, ***) or by other sequential symbols: *, †, ‡, §, ¶, ||; these may be doubled up if needed. *P* values should be labeled by symbol in the image, with definitions appearing only in the legend (eg, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$). Labeling of data as not significant (NS) is unnecessary. All abbreviations should also be defined.

Preparation of Images (Ethics): Taking photographs of the same source under varied fields of view, light intensity, magnifications, or contrast conditions without disclosing that the data are not unique to the present study constitutes suspect scientific conduct. Further, unless serial sections are used, the publication of identical-appearing images labeled with different staining techniques in different papers raises legitimate questions. Finally, reuse of one's data (ie, "self-plagiarism") is a copyright violation if the authors signed over copyright to the publisher; the reuse of such copyrighted images is at the sole discretion of the publisher, with proper attribution of the original publication a requirement of reuse. No specific feature within an image may be enhanced, obscured, moved, removed, or introduced. The grouping of images from different parts of the same gel or blot, or from different gels or blots, fields, or exposures must be made explicit by the arrangement of the figure (eg, using dividing lines) and in the figure legend. Adjustments of brightness, contrast, or color balance are acceptable only if they are applied to the whole image, whether experimental or control image, and as long as they do not obscure or eliminate any information present in the original (Portions adapted with permission from the JCB). Any evidence of inappropriate manipulation may prompt the Editors to

request an explanation and access to original data, which the authors must make available. Information can be found in the Journal Scientific Integrity Policy at *https://www.jmdjournal.org/content/integrity*.

***Digital Art*:** Digital figure files are required for submission. The Journal of Molecular Diagnostics' requirement for color images is RGB (Red, Green, Blue) color mode. Images submitted in RGB will retain the vivid reds, greens, and blues of the original digital files for online publication. The preferred file formats for digital figures are PDF, TIFF, and JPG. Embedding images in MS Word is strongly discouraged as individual file images will ultimately be required for accepted manuscripts when being sent into production. Use of PowerPoint (PPT) files is also discouraged due to known conversion errors when consolidating to PDF as well as when processed through the production pipeline. For detailed instructions on preparing digital art for submission or production, visit *https://www.elsevier.com/authors/author-schemas/artwork-and-media-instructions* or contact the Editorial Office (jmd@asip.org).

- * ***Data Supplements and Non-Traditional Media*.** Figures and tables that are critical to the evaluation and understanding of the research presented, but which cannot be accommodated via the print medium (eg, video), will be considered part of the manuscript submission. Details of reagents (ie, oligonucleotides, antibodies) should be part of the main manuscript and not Supplemental Data (see Materials and Methods). Such files will be peer reviewed and published on the Journal website if the manuscript is accepted. Inclusion of supplementary material is at the discretion of the Editors, and data may be moved to the main article at their request. Supplemental figures (PDF, TIFF, or JPG), tables (MS Word or Excel), and legends (MS Word) should be prepared as above. Supplemental material published on the Journal website is subject to the same copyright as applies to the printed article. Online data will remain associated with its article and is not subject to any modifications or updates after publication. As of January 1, 2016, authors of accepted manuscripts are no longer charged for supplemental data being published on the Journal website. However, authors are still encouraged to minimize the size of their supplemental data files.

Complex data sets such as microarray data and gene sequences should be deposited in a reliable public archive. Microarray repositories should comply with the

Minimum Information About a Microarray Experiment (MIAME) guidelines (*https://www.ncbi.nlm.nih.gov/geo/info/MIAME.html*). Examples include Gene Expression Omnibus (*http://www.ncbi.nlm.nih.gov/geo*) and Array Express (*http://www.ebi.ac.uk/arrayexpress*). Tissue microarray data exchange specifications can be found at http://www.biomedcentral.com/1472-6947/3/5. Otherwise, the data may be published as supplemental data on the Journal's website (*https://www.jmdjournal.org/*).

Other materials that are not required to understand and evaluate the article may be stored on an institutional website and referenced as a URL within the manuscript text. In this case, the authors must assume responsibility for maintaining a live, unrestricted link to the material from the URL published within the article, in perpetuity. Failure to maintain a live, unrestricted link may result in retraction of the article by the Editors. Material that has been published previously (print or online) is not acceptable for posting as supplementary data. Instead, the appropriate reference(s) to the original publication should be made in the text and references.

The Editors will make the final determination as to whether a given data set is essential to the manuscript and whether they will require publication of the material on the Journal website. Failure to relinquish required materials for online publication can result in the reversal of a manuscript's acceptance.

For more information, please see the Research Data section below.

RESEARCH DATA

This journal encourages and enables you to share data that supports your research publication where appropriate, and enables you to interlink the data with your published articles. Research data refers to the results of observations or experimentation that validate research findings. To facilitate reproducibility and data reuse, this journal also encourages you to share your software, code, models, algorithms, protocols, methods and other useful materials related to the project. Below are a number of ways in which you can associate data with your article or make a statement about the availability of your data when submitting your manuscript. If you are sharing data in one of these ways, you are encouraged to cite the data in your manuscript and reference list. Please refer to the "References"

section for more information about data citation. For more information on depositing, sharing and using research data and other relevant research materials, visit the research data page.

Data linking.If you have made your research data available in a data repository, you can link your article directly to the dataset. Elsevier collaborates with a number of repositories to link articles on ScienceDirect with relevant repositories, giving readers access to underlying data that give them a better understanding of the research described. There are different ways to link your datasets to your article. When available, you can directly link your dataset to your article by providing the relevant information in the submission system. For more information, visit the database linking page .For supported data repositories a repository banner will automatically appear next to your published article on ScienceDirect.In addition, you can link to relevant data or entities through identifiers within the text of your manuscript, using the following format: Database: xxxx (e.g., TAIR: AT1G01020; CCDC: 734053; PDB: 1XFN).

Mendeley Data.This journal supports Mendeley Data, enabling you to deposit any research data (including raw and processed data, video, code, software, algorithms, protocols, and methods) associated with your manuscript in a free-to-use, open access repository. During the submission process, after uploading your manuscript, you will have the opportunity to upload your relevant datasets directly to *Mendeley Data*. The datasets will be listed and directly accessible to readers next to your published article online. For more information, visit the Mendeley Data for journals page.

Data statement.To foster transparency, we encourage you to state the availability of your data in your submission. This may be a requirement of your funding body or institution. If your data is unavailable to access or unsuitable to post, you will have the opportunity to indicate why during the submission process, for example by stating that the research data is confidential. The statement will appear with your published article on ScienceDirect. For more information, visit the Data statement page.

Copyright. Copyright of published manuscripts is held by the Association for Molecular Pathology and the American Society for Investigative Pathology, which

must receive the assignment of copyright from the authors of accepted manuscripts. For US government employees, the above assignment applies only to the extent allowed by law. Details regarding copyright transfer and author rights will be presented by Elsevier, Inc., at the time of article production.

Authors wishing to republish images, tables, or text previously published elsewhere should provide proof of permission with their submission and should include the appropriate attribution in the figure or table legend or in the text. It is the responsibility of the authors, not the Journal, to obtain such permission from the copyright holder.

Publishing in *The Journal of Molecular Diagnostics* automatically places authors in compliance with the NIH Public Access Policy (see http://publicaccess.nih.gov/submit_process.htm, Submission Method A). Any article noted as being funded by NIH, Wellcome Trust, MRC, or other groups for which Elsevier, Inc., has a transfer agreement (https://www.elsevier.com/about/open-science/open-access/agreements; such list is subject to change) will be deposited in PubMed Central (PMC), to be made available to the public twelve or six months after final print publication (as stipulated by the funding agency). Authors therefore should NOT complete a separate deposit of their material but will be contacted by PubMed Central for grant verification once the article has been received by the PMC article system.

Contact healthpermissions@elsevier.com regarding permission to deposit manuscripts in other government-sponsored repositories in cases where *The Journal of Molecular Diagnostics* does not have a system in place to automatically deposit materials on behalf of its authors. Deposit of accepted or published manuscripts in any non-JMD repository without prior permission by the Journal is a violation of copyright.

Elsevier supports responsible sharing.Find out how you can share your research published in Elsevier journals.

Embargo Policy. All information regarding the content of submitted or accepted manuscripts is strictly confidential. Information contained in or about accepted articles cannot appear in print, audio, video, or digital form or be released by the news media until the Journal embargo date has passed, not to exceed the

publication date of the article. For detailed information on embargo release dates or for news media requests for preprint copies of specific articles, contact asipproduction@elsevier.com.

Financial Disclosure and Conflicts of Interest. All authors must disclose any current or former relationships held by the author or an immediate family member (eg, employment, consultancies, board membership, stock ownership, funding, honoraria, expert testimony, patents or royalties, travel reimbursements, industry-supplied free reagents, etc) with any organization or entity having a direct financial or personal interest in the subject matter or materials discussed in the article. Authors should err on the side of full disclosure and should contact the Editorial Office if they have questions or concerns. This information should be provided at the time of submission and reiterated as part of copyright assignment. Failure to do so may result in manuscript rejection or editorial retraction of the article. Further information can be found in the Journal Scientific Integrity Policy at http://jmd.amjpathol.org/content/integrity. Upon manuscript acceptance, all coauthors will be instructed to formally disclose all potential conflicts.

Publication Charges. As of January 1, 2016, there are no longer separate charges for color figures, black & white or grayscale figures, or composed tables. Corresponding authors who are current, dues-paying Regular Members of ASIP or of AMP at the time of submission will be charged US $75 per page. All other Societies members and nonmembers will be charged US $100 per printed page. Manuscripts arising from research in developing countries may be eligible for waiver of publication charges only if all authors are located in a qualifying country (for a list of eligible countries, see Group A and Group B countries at *https://www.research4life.org/access/eligibility/*). Otherwise, page charges will not be waived, except for solicited editorials.

Approval of Fees and Invoicing: Upon submission of an article, the corresponding author must state their agreement to pay the publications costs, as outlined above. A final invoice will be presented

by Elsevier, Inc., immediately after publication; full payment is expected at that time. Failure to pay publication charges may result in the authors' inability to publish future articles in The Journal of Molecular Diagnostics.

Open Access. In addition to publication charges outlined above, authors may elect to participate in the Journal's Open Access program. Articles published under Open Access will utilize the Creative Commons (CC) BY-NC-ND 4.0 license (*http://creativecommons.org/licenses/by-nc-nd/4.0*), allowing users to access and share the article easily. CC BY 4.0 license (*http://creativecommons.org/licenses/by/4.0*) will be made available only when required by the funding agency. **All other copyright restrictions regarding reuse still apply.** Cost to participate is US$2400 for ASIP and AMP Regular Members. All other Society members and non-members will be charged $3000, which must be received before article restriction is lifted. Instructions for obtaining Open Access will be presented by Elsevier, Inc., at the time of article production. Authors wishing to participate may also contact the Editorial Office after manuscript acceptance.

Reprints. Reprints should be ordered when page proofs are returned; an order form is included with the proofs. Late reprint orders may result in additional fees. There is no extra charge for color reprints. Author reprints will be supplied in CMYK color mode because it is a printed medium, but electronic reprints will be supplied in RGB color mode for visualization on screen. Each author is eligible to receive one complimentary electronic (secure PDF) reprint. No hardcopy complimentary reprints are provided.

Proofs. Article proofs are prepared and corrected using Elsevier's Proof Central online interface (*https://www.elsevier.com/authors/journal-authors/proof-central*). Once the proofs are ready, the corresponding author will be contacted by email with a link to access the manuscript in Proof Central. Electronic page proofs will be provided in RGB, thus reflecting the online quality. Image files will be converted to CMYK for the printed journal; all images published online will be in RGB. Full instructions on completing proof corrections will be provided with the downloaded proof. The author should check the proofs carefully, mark any printer's errors, and answer queries as requested. Author changes should be kept to a minimum. Proof corrections and replacement figures (if any) must be returned within 24 hours to avoid any delay in publication.

Editorial Communications. All correspondence concerning editorial matters should be addressed to the Editor-in-Chief at The Journal of Molecular Diagnostics, 1801 Rockville Pike, Suite 350, Rockville, Maryland 20852-9975, or emailed to jmd@asip.org. **Appeals** (rebuttals) to editorial decisions must be submitted within 60 days to be considered. Letters to the Editor will be printed at the Editors' discretion in the **Correspondence** section. The Editors will consider all refutations of previously published work, according to our usual standards of quality, with the original authors given the opportunity to respond to such criticisms.

Corrections are published upon request and after editorial review. **Retractions** are published upon request of authors or their institutions and may also be published by the Journal following a determination of scientific misconduct. **Notes of Concern** are published in response to editorial concerns relating to scientific or publishing misconduct by authors or reviewers or to alert the scientific community of an ongoing investigation. [Revised: January 2022]

Appendix E. Exchange rates indicators for South Africa

Exchange Rates selected indicators

South Africa

Source: International Financial Statistics (IFS)

Metadata by Country (IFS)

Concept	Unit	Code	Scale	2017	2018	2019	2020	2021
National Currency per SDR, End of Period	Rate	ENSE_XDC_XDR_RATE	Units	17.54	19.99	19.39	21.15	22.26
National Currency per SDR, Period Average	Rate	ENSA_XDC_XDR_RATE	Units	18.47	18.74	19.96	22.93	21.05
Domestic Currency per U.S. Dollar, End of Period	Rate	ENDE_XDC_USD_RATE	Units	12.32	14.38	14.03	14.69	15.91
Domestic Currency per U.S. Dollar, Period Average	Rate	ENDA_XDC_USD_RATE	Units	13.32	13.23	14.45	16.46	14.78
Nominal Effective Exchange Rate, Trade Partners by Consumer Price Index	Index	ENEER_IX	Units	65.27	65.31	62.08	55.67	60.03
Real Effective Exchange Rate, based on Consumer Price Index	Index	EREER_IX	Units	79.65	80.83	78.27	71.02	77.62