DETECTION OF COPY NUMBER VARIATIONS IN INTELLECTUALLY DISABLED PATIENTS IN INDONESIA USING HIGH RESOLUTION GENOME WIDE ARRAY ANALYSIS

DETEKSI *COPY NUMBER VARIATIONS* PADA PASIEN DISABILITAS INTELEKTUAL DI INDONESIA MENGGUNAKAN ANALISIS *GENOME WIDE ARRAY* RESOLUSI TINGGI

THESIS

Submitted to fulfill the assignment and fit-out requisite in passing Post Graduate Program

Master of Biomedical Science

Almira Zada
22010110400095

FACULTY OF MEDICINE
DIPONEGORO UNIVERSITY SEMARANG
2013

THESIS

DETECTION OF COPY NUMBER VARIATIONS IN INTELLECTUALLY DISABLED PATIENTS IN INDONESIA USING HIGH RESOLUTION GENOME WIDE ARRAY ANALYSIS

Arranged by
ALMIRA ZADA, MD
22010110400095

Has been defended in front of the defense committee on 23rd December 2013 and has been approved by:

The Netherlands	**Indonesia**
Principal Supervisor,	Principal Supervisor,
Nicole de Leeuw	Prof. Sultana MH Faradz, MD,PhD NIP. 19520202 197901 2 001
Supervisor,	Supervisor,
Helger Ijntema, PhD	Farmaditya EP Mundhofir, MD, PhD NIP. 19810425 200812 1 002

Approved by,
Head of Master Degree Program in Biomedical Science
Medical Faculty of Diponegoro University

Prof. Dr. dr. Tri Nur Kristina, DMM, M.Kes
NIP.1950527 198603 2 001

DECLARATION

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgement is made in the text

Semarang, November 2013

Almira Zada

ACKNOWLEDGEMENT

This work would have been impossible without the contribution of the skill and assistance of many people. I give my deepest gratitude to my supervisor, Nicole de Leeuw PhD, for her patient and encouragement in guiding and teach me, and all of the opportunities that have been given to me during my research in The Departement of Human Genetics Radboud University Medical Centre (RUMC) Nijmegen, the Netherlands for the array analysis. I would also like to thank Rolph Pfundt PhD for giving me opportunities and more practice in array diagnostics. I gratefully thank to Prof. Ben Hamel who gave me a wonderful opportunity to study in RUMC, guiding and teaching me in the clinical examination and of course Helger Ijntema PhD and Willy Nillesen for teaching me basic techniques of molecular genetics.

I would like to give countless thanks to Prof. Dr. Sultana MH Faradz, PhD my supervisor for all her support, concern and attention to me. I also would like to thank Farmaditya Eka Putra Mundhofir, MD, PhD for his guidance and suggestions from the early stage of this research especially in collecting Intellectually Disabled patients and sharing secondary data regarding his previous research.

I wish to thank Nico Leijsten for guiding me in the array laboratory flow and teaching me the array analysis, to Trudi Machielsen, Ronald van Beek, Martijn Bannings, Monique Wunderink, Sabine Kempers van Gessel for their endless guidance. I express my gratitude to Guillaume van de Zande and Liesbeth

van Rossum for guiding and teaching me in the Flourescence in situ hybridization laboratory flow. I would also like to thank Martine van Zweeden for her guiding in the short tandem repeats assay laboratory flow and analysis. To all the staff DNA Diagnostic, RUNMC, Nijmegen, The Netherlands, thank you for your cooperation and friendship you have shared over time. And, my gratitude for Bregje van Bon PhD and Ineke van der Burgt for discussion and suggestion for the clinical knowledge.

I express my gratitude to Dr. Tri Indah Winarni for her assistance with cytogenetics and all the staff of Division of Human Genetics CEBIOR Faculty of Medicine Diponegoro University in Semarang Indonesia for your cooperation. Particularly I would like to thank Wiwik Lestari, Lusi Suwarsi, Dwi Kustiyani and Rita Indiarti for their time and technical guidance in routine cytogenetics analysis.

I wish to thank Drs. Ciptono, Poerwoko BA and Drs.Slamet from SLB Negeri Semarang, drs. Djiwaningsih from SLB Temanggung for organizing the collection of samples and my visit to their special schools. My deepest gratitude to the parents for providing records and clinical data and allowing me to examine their children and to collect blood samples. This research would not be possible without their cooperation.

This work is funded by Beasiswa Unggulan BPKLN Ministry of Education Republic Indonesia. My thanks to the staffs of Beasiswa Unggulan Project especially DR. Abe Susanto and Dina Ardina, S.Sos from the staff of Biomedical Science Post Graduate Program. Thanks to Rens Jongerius and Ineke Zaalmink for their help in organizing my stay in Netherlands.

Finally, I would give my special thanks for my parents, M. Noor Kholis and Diah Susana, my sister Amalina Fakhriah, and my partner Akbar Yoga Wibawa, MD who have been always supporting me in any situations, and of course my friends that always giving me endless support during this work.

CURRICULUM VITAE

Personal Data

Name	: Almira Zada
Sex	: Female
Nationality	: Indonesian
Place & Date of Birth	: Kudus/ July 31st, 1987
Latest Degree	: Medical Doctor
Address	: Jl. Permai XVIII No. 09, Kudus, Central Java, Indonesia, 59361
Email	: almira.zd@gmail.com

Education

1993-1999	Elementary School at SD Prambatan Lor 1 Kudus
1999-2002	Junior High School at SMP N 1 Kudus
2002-2005	High School at SMA N 1 Kudus Majoring Natural Science
2005-2009	Diponegoro University, Medical Faculty (Bachelor Degree)
2009-2011	Diponegoro University, Medical Faculty (Medical Doctor)
2011- present	Post Graduate Program Diponegoro University, Master in Biomedical Science Majoring in Genetic Counseling

	(Twinning Program with Radboud University medical Centre Nijmegen, The Netherlands)

Experiences

2006-2008	Department of Pharmacology and Therapeutics , medical Faculty of Diponegoro University. Teaching assistant and practical laboratory assistant for drugs toxicity study
2007-2009	Department of Biochemistry, Medical faculty of Diponegoro University. Teaching assistant and practical laboratory assistant for enzyme and protein study
2007-2009	Department of Chemistry, Medical faculty of Diponegoro University. Teaching assistant and practical laboratory assistant for organic chemistry study
2012-2013	Department of Human Genetics, Radboud University Medical Centre, Nijmegen, the Netherlands: Internship Master Student
2013	Trained and Certified Analyst in Array Diagnostics, Department of Human Genetics, Radboud University Medical Centre, Nijmegen, the Netherlands

Trainings & Courses

2012	Workshop on Neurogenetics, Semarang Indonesia
2012	Advanced Medical Genetics Course from Basic to

	Clinic, Semarang Indonesia
2012	Medical Genetics Course from Basic to Clinic, Semarang Indonesia
2012	Radboud University Medical Centre, Nijmegen, the Netherlands: • Genomics in Health and Disease- Towards Personal Genomics • Neurodevelopmental Disorders from Mechanism to Clinical Trials • The Role of X-Linked Mental Retardation Genes in Synapse Functions • Synapse Evolution and Mechanism of Cognition • Open Lecture in Genetics. Prof. dr. J.R Lupski, Baylor College of Medicine. Houston, Texas USA
2013	The 1st Cognomics Symposium, Radboud University Medical Centre, Nijmegen, the Netherlands
2013	The 8th Goldrain Course in Clinical Cytogenetics, South Tyrol, Italy
2013	Nederlandse Vereniging voor Humane Genetica (NVHG) Symposium (Symposium of The Netherlands Human Genetic Association), Arnhem, the Netherlands
2013	Seminar Neurodevelopmental Disease: From Copy

	Number Variations to Genes, Radboud University Medical Centre, Nijmegen, the Netherlands in Collaboration with University of Washington U.S.A

Scientific Papers

2009	Zada, A. The Effect of *Eucheuma sp*. Seaweed Diet in Erythrocyte Counts in Wistar Rats with Diabetic Alloxan. Undergraduate Thesis. Medical faculty of Diponegoro University, Semarang Indonesia
2013	Zada, A. A rare, recurrent, *de novo* 14q32.2q32.31 microdeletion of 1.1 Mb in a 20-year-old female patient with a maternal UPD(14)-like phenotype. Presented in The Short Lecture Competition The 8th Goldrain Course in Clinical Cytogenetics 2013, South Tyrol, Italy

Awards

2004	International Biology Olympiade 2004 (Indonesian Representative)
2005	National Chemistry Olympiade Gadjah Mada University, Indonesia (Finalist)
2005	*Annual Scientific Fair Faculty of Science Diponegoro University (2nd Winner in Scientific Research*

	Competition)
2013	The Best Short Lecture Presentation in the Short Lecture Competition The 8th Goldrain Course in Clinical Cytogenetics, South Tyrol, Italy

Scholarship/Grant

2011	Excellent Scholarship Program of The Bureau of Planning and International Cooperation, Ministry of National Education, Government of Indonesia
2013	The Grant from The European Cytogeneticist Association for 8th Goldrain Course in Clinical Cytogentics, South Tyrol, Italy

CONTENTS

ABBREVIATIONS

AAIDD	American Association on Intellectual and Developmental Disabilities
AGCC	Affymetrix GeneChip Command Console
ASHG	American Society of Human Genetics
BAC	Bacterial artificial Chromosome
BAZ1B	Brodomain Adjacent to the Zinc finger domain 1 B
BEGAIN	Brain Enriched Guanylate Kinase Associated Protein
CLIP2	CAP-Gly domain containing Linker Protein 2
CGH	Comparative Genomic Hybridization
ChAS	Chromosome Analysis Suite
CNV	Copy Number Variation
DAEMON	Disk And Execution Monitor
DECIPHER	Database of Chromosomal Imbalance and Phenotype in Human Using Ensembl Resources
DEGS2	Delta(4)-desaturase,sphingolipid 2
DGV	Database of Genomic Variants
DLK1	Delta Like 1
DNA	Deoxyribo Nucleic Acid
ECARUCA	European Cytogeneticists Association Register of Unbalanced Chromosome Abberations

ELN	Elastin
EML1	Echinoderm Microtubule associated protein like 1
ESHG	European Society of Human Genetics
EVL	Enah/Vasp-like
FBS	Fetal Bovine Serum
FCN3	Ficolin 3
FISH	Fluorescent in situ Hybridization
FMR1	Fragile X Mental Retardation 1
FoSTeS	Fork Stalling and Template Switching
GECCO	Genomic CNV Classification Object
HR	Homologous Recombination
ID	Intellectual Disability
LCR	Low Copy Repeat
LOH	Loss of Heterozigosity
LIMK1	LIM domain Kinase 1
MAPD	Median of the Absolute values of all Pairwise Differences
MEG	Maternally Expressed Gene
MEM	Minimum Essential Medium
MLPA	Multiplex Ligation-dependent Probe Amplification
MMBIR	Microhomology Mediated Break Induce Replication
NAHR	Non-Allelic Homologous Recombination
NHEJ	Non-Homologous End Joining
NROB2	Nuclear Receptor Subfamily O, grup B, member 2

OMIM	Online Mendelian Inheritance in Man
PCR	Polymerase Chain Reaction
PEG	Paternally Expressed Gene
ROH	Region Of Homozigosity
RTL1	Retrotransposon- Like 1
RTL1as	Retrotransposon- Like 1 anti sense
RUMC	Radboud University Medical Centre
SNP	Single Nucleotide Polymorphism
UCSC	University of California Santa Cruz
UPD	Uni Parental Disomy
VOUS	Variation of Unknown Significance
WARS	Tryptophanyl-tRNA synthethase
WBS	Williams- Beuren Syndrome
WDR25	WD repeat domain 25
WSTF	William Syndrome Transcription Factor
WINAC	WSTF Including Nucleosome Assembly Complex
YY1	Yin Yang 1

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

GLOSSARY

Copy number variation	A genomic structural variation ranging from kilobases (kb) to megabases (Mb) that are not identifiable by conventional chromosomal banding
Intellectual disability	Neurodevelopmental disorder that is characterized by an intelligence quotient (IQ) of 70 or below, and deficits in at least two behaviors related to adaptive functioning diagnosed by 18 years of age
Array	A collection of microscopic DNA spots attached to a solid surface. Each spot contains picomoles (10^{-12} moles) of a specific DNA sequence known as probes.
FISH	Fluorescent in situ hybridization. A cytogenetic technique that is used to detect and localize the presence or absence of specific DNA sequence on chromosomes by using fluorescent probes
MLPA	Multiplex Ligation-dependent Probe Amplification. A variation of the multiplex polymerase chain reaction that permits multiple targets to be amplified with only a single primer pair. Only those primers that hybridize to the target sequences are amplified, and the resulting products can be analyzed by capillary electrophoresis
qPCR	Quantitative Polymerase Chain Reaction. A

	quantitative measurement of DNA based polymerase chain reaction
ROH	Region Of Homozigosity. When the genomics copies derived from each parent have the same base for the polymorphic region
SNP	Single Nucleotide Polymorphism. a DNA sequence variation occuring when a single nucleotide (A,T,C or G) in the genome differs between paired chromosomes
Trio	A group consists of child, father, and mother

ABSTRACT

Background: Intellectual disability (ID) is a neurodevelopmental disorder that is characterized by an intelligence quotient (IQ) of 70 or below, and deficits in at least two behaviors related to adaptive functioning diagnosed by 18 years of age. The etiology of ID can be classified as genetic and non-genetic factors. Chromosomal abnormality is a frequent genetic cause of ID. Routine cytogenetic analysis remained the major diagnostic test for ID patients in Indonesia. However, it cannot detect the submicroscopic chromosomal rearrangement termed copy number variation (CNV). The use of genome wide array analysis to detect CNVs might increase the rate of diagnostic yield in ID patients in Indonesia.

Methods: Eighteen patients with ID who had normal results after karyotyping, CGG repeat analysis in the *FMR1* gene and subtelomeric MLPA testing were included. Detection of CNVs were done by using high resolution genome wide array analysis (Affymetrix CytoScan HD Array platform) with an average test resolution of approximately 20 kb then analyzed by using Chromosome Analysis Suite (ChAS) Software V.2.0. The various CNV detected were classified by comparing in house and international normal and affected individual datasets, gene contents and literature studies. FISH studies were performed for carrier testing and the possible presence of a balanced chromosome rearrangement in the parents and/or mosaic imbalance.

Results: Three out of 18 patient with ID were found to carry pathogenic CNV i.e one patient with 14q32.2q32.31 microdeletion of 1.1 Mb and 2 patients with 7q11.23 microdeletion of 1.4 Mb (Williams-Beuren Syndrome). One patient with 1p36.11p35.3 microduplication of 1.7 Mb were found to carry likely pathogenic CNV. In addition, there is one patient carried large homozygous regions totaling ~7% of the autosomal genome which one of these homozygous stretches harbours a mutated recessive disease gene.

Conclusion: The genome wide array analysis in this study has an additional detection rate of 16.7% (3 out of 18 patients) causative CNVs in selected ID patients after routine karyotyping combined with subtelomeric MLPA and CGG(n) repeat analysis in the *FMR1* gene.

Keywords: Intellectual disability, copy number variation, genome wide array analysis

ABSTRAK

Latar belakang: Disabilitas intelektual adalah kelainan perkembangan syaraf yang ditandai dengan IQ ≤ 70 dan deficit sekurangnya dua perilaku terkait fungsi adadptif yang terdiagnosis sebelum usia 18 tahun. Penyebabnya dapat diklasifikasikan sebagai factor genetic dan non-genetik. Kelainan kromosom adalah penyebab genetik tersering pada disabilitas intelektual. Analisis sitogenetik rutin masih menjadi tes diagnosis utama pada pasien disabilitas intelektual di Indonesia. Meskipun demikian, tes tersebut tidak dapat mendeteksi kelainan kromosom submikroskopik yang disebut *copy number variation* (CNV). Analisis menggunakan *genome wide array* kemungkinan menaikkan tingkat penegakan diagnosis pada pasien disabilitas intelektual di Indonesia.

Metode: Sampel DNA diambil dari delapan belas pasien dengan disabilitas intelektual yang mempunyai hasil normal setelah analisis karyotyping, CGG *repeat* pada gen *FMR1* dan subtelomerik MLPA. Deteksi CNV dilakukan dengan menggunakan teknik *genome wide array* (Affymetrix CytoScan HD Array platform) dengan resolusi rata-rata 20 Kb yang selanjutnya dianalisis menggunakan *Chromosome Analysis Suite (ChAS) Software V.2.0*. Variasi CNV yang terdeteksi diklasifikasikan dengan cara membandingkan dataset *in house* dan internasional dari individu normal dan penderita, isi gen, dan studi literatur. Uji FISH dilakukan dalam rangka validasi rutin untuk mengetahui adanya *balance rearrangement* atau mozaik tingkat rendah pada orang tua.

Hasil: *CNV* patogenik ditemukan pada 3 dari 18 pasien dengan disabilitas intelektual yaitu satu pasien dengan mikrodelesi 1.1 Mb pada kromosom 14q32.2q32.31 dan 2 pasien dengan mikrodelesi 1.4 Mb pada kromosom 7q11.23 (Williams-Beuren Syndrome). *Likely pathogenic CNV* ditemukan pada satu pasien dengan mikroduplikasi 1.7 Mb pada kromosom 1p36.11p35.3. Terdapat pula satu pasien dengan regio homozigot luas (7% dari total autosom) dimana salah satu daerah homozigot ini mengandung gen penyakit resesif yang termutasi.

Kesimpulan: Analisis dengan genome wide array pada studi ini memiliki tambahan tingkat deteksi CNV penyebab disabilitas intelektual sebesar 16.7% (3 dari 18 pasien) setelah analisis karyotype rutin dikombinasi dengan subtelomerik MLPA dan CGG rpeat pada gen *FMR1*.

Kata kunci: Disabilitas intelektual, *copy number variation, genome wide array*