

EXPLORING COPY NUMBER VARIATIONS IN A THAI POPULATION**CHAIWAT NAKTANG 5536368 SCBC/M****M.Sc.(BIOCHEMISTRY)****THESIS ADVISORY COMMITTEE: VARODOM CHAROENSAWAN, Ph.D.
NATINI JINAWATH, M.D., Ph.D. BHOOM SUKTITIPAT, M.D., Ph.D.
SUMALEE TUNGPRADAPKUL, Ph.D.****ABSTRACT**

Copy Number Variation (CNV) is one of the major structural variations in a human genome. CNV has been associated with several human diseases such as neurodevelopmental diseases, including Autism Spectrum Disorder (ASD), and neuropsychiatric diseases. Recently, reference CNVs in normal subjects from certain populations, such as African-American, Caucasian and East Asian, are available from a number of CNV databases. These CNV databases can facilitate clinical interpretation of CNVs, which can be categorized into three main groups: pathogenic (disease-related), unknown clinical significant, or benign. So far there is no normal CNV database available for the Thais, and existing CNV databases of different ethnic groups are by no mean an ideal reference for the CNV interpretation for the Thai population, due to divergent genetic backgrounds. In this study, we combine the genome-wide Single Nucleotide Polymorphism (SNP) genotyping data from previous studies, consisting of 3,017 Thai subjects with no known genetic disorders. We perform CNV discovery from these datasets using PennCNV and CNV Workshop software to ensure the highest possible confident of CNV calls, and using the combining CNV sets to create the largest CNV reference for the Thais to date. Moreover, we perform population analysis by using the program Plink to compare the Thai population with eleven HAPMAP3 populations. Hierarchical clustering analysis (HCA) using frequency of candidate genes is used to assess similarity between the Thai population and other HAPMAP3 populations. The results show that CNVs found in the Thai population cluster with other Asian populations. Having population-specific CNV database will improve the accuracy for the interpretation of clinical significant CNVs in the Thais, and serve as one of the most informative population-specific CNV reference databases for population geneticists.

**KEY WORDS: COPY NUMBER VARIATION (CNV) / SINGLE NUCLEOTIDE
POLYMORPHISM (SNP) / THAI POPULATION / DATABASE****80 pages**