CHAPTER 1

INTRODUCTION

1.1 Introduction

In vitro fertilization (IVF) is a form of assisted reproduction technology where fertilization occurs outside of the body. A technique known as controlled ovarian hyperstimulation is utilized in IVF to induce ovulation or the release of mature eggs from the ovary, so fertilization with sperm can happen *in vitro* (Gindoff, Hall, & Stillman, 1990). After fertilization, the zygote goes through embryo culture for 3 – 5 days before it is returned and implanted to the uterus. However, embryo could have genetic defects which would lead to miscarriage or genetic disease. To prevent such cases, preimplantation genetic testing (PGT) can be performed through the removal of few embryo cells that have developed for about 5 or 6 days for genetic disorders screening (Sermon, Van Steirteghem, & Liebaers, 2004).

One of the common disorders screened through PGT is aneuploidy. Aneuploidy is an abnormal condition where the number of chromosomes in the cell is abnormal; normal or euploid embryos would have 23 pairs of chromosomes (Anthony et al., 2000; Painter, 1923). Aneuploidy arises during cell division because of abnormal separation of the chromosomes (Dierssen, Herault, & Estivill, 2009). An extra or missing chromosome could have a detrimental health effect and causes disorders, such as Down syndrome, Patau syndrome, Turner syndrome, and Edwards syndrome. Preimplantation genetic testing for aneuploidy (PGT-A) is the name of the test employed for aneuploidy detection.

Even when aneuploidy screening is performed for all 46 chromosomes, a successful pregnancy is still not a guarantee. Many conditions could increase the chance of miscarriage during pregnancy, such as embryonic stress, autoimmune factors, anatomic abnormalities, and other conditions that are currently unknown (Brezina, Anchan, & Kearns, 2016). Even so, ensuring embryo is free from aneuploidy prior implantation remains is fundamental to ensure a successful pregnancy.

In the early days of PGT, fluorescence in situ hybridization (FISH) is a way to perform cheap, accurate, and specific for aneuploidy detection (Philip, Bryndorf, & Christensen, 1994). However, FISH cannot be used to evaluate the ploidy status of all 23 chromosomes pairs (Treff & Scott, 2012). Furthermore, a prospective and randomized study published in 2007 failed to show an increase in pregnancy rates with PGT (Mastenbroek et al., 2007). The 2007 study is not the only study which further discourages the use of PGT. The American College of Obstetrics and Gynecology (ACOG), the European Society of Human Reproduction and Embryology (ESHRE), and the American Society for Reproductive Medicine (ASRM) also discourages the use PGT (ACOG, 2009; ASRM, 2008; ESHRE, 2002).

With further discouragements towards the use of PGT with FISH, embryologist and geneticists explored other ways to perform PGT for aneuploidy detection. All currently available methods for PGT-A are able to evaluate the ploidy status for all 23 chromosome pairs at trophectoderm stage that resulted in significantly higher implantation and live birth rates compared with FISH (Schoolcraft et al., 2011). The current methods are microarrays, Single Nucleotide Polymorphism (SNP) microarrays, Comparative Genomic Hybridization (CGH) microarrays, quantitative PCR (qPCR), and Next Generation Sequencing (NGS) (Brezina et al., 2016). From these platforms, NGS technologies can provide more data points compared to the other platforms that allow further analysis such as mosaicism detection and significant deletions and duplications detection (Fiorentino et al., 2014).

Here we propose a new strategy for aneuploidy assessment with the use of targeted Next Generation Sequencing (tNGS) to produce coverage and SNP variant data as the main parameter for aneuploidy detection. We used AmpliSeq for Illumina, a suite of chemistry products that are compatible with Illumina NGS platforms, to allow rapid and efficient sequencing of thousands of SNP amplicon locus that are spaced throughout the human genome. The main outputs were variation data stored in Variant Call Format (VCF) files and genomic reads coverage data that were stored in coverage files. Genotype information in the VCF files was utilized to construct B allele frequency (BAF) plots, and the read count

information in the coverage files was utilized to construct a chromosome copy number plot. Together, the two strategies can be used simultaneously for quick and precise aneuploidy assessment.

In addition to aneuploidy assessment, copy number and BAF analysis can also be used to check for deletion and duplication, including those that span less than 5 Mb, known as microdeletion and microduplication syndrome (MMS) (Weise et al., 2012). MMS can have adverse health effects, such as the 3.7 Mb microdeletion in a specific region of chromosome 17 (17p11.2), which causes Smith-Magenis syndrome (Bi et al., 2002). With targeted sequencing, MMS regions of interest could be sequenced without the need of whole genome sequencing which would allow for a faster and more efficient MMS detection.

Several studies found out that miscarriages in human pregnancy are highly correlated with chromosomal aneuploidy, and aneuploidy is known to be the most common cause for reproductive failure (Brezina & Kutteh, 2015; Morales et al., 2008). As such, technology for aneuploidy assessment is necessary to reduce the rate of reproduction failure. The use of targeted sequencing for aneuploidy assessment not only allow the evaluation of all 46 chromosomes for ploidy status but also the evaluation of microdeletion and microduplication. The scope and focus of this thesis are the development and validation of bioinformatics application to perform aneuploidy detection from targeted sequencing data. The objective of this study is to develop an efficient and intuitive application that is wrapped inside a user-friendly interface program that is used not only for aneuploidy detection but also microdeletion and microduplication evaluation from targeted sequencing data.