Available online at www.sciencedirect.com

ScienceDirect

Journal of Genetics and Genomics 42 (2015) 151−159

JGG

ORIGINAL RESEARCH

The Performance of Whole Genome Amplification Methods and Next-Generation Sequencing for Pre-Implantation Genetic Diagnosis of Chromosomal Abnormalities

Na Li [a,b,1], Li Wang [c,1], Hui Wang [c], Minyue Ma [c], Xiaohong Wang [a], Yi Li [a], Wenke Zhang [c], Jianguang Zhang [d], David S. Cram [d,*], Yuanqing Yao [c,*]

[a] *Department of Gynecology and Obstetrics, Reproductive Medicine Center, Tangdu Hospital, Fourth Military Medical University, Xi'an 710038, China*
[b] *Department of Obstetrics and Gynecology, Affiliated Hospital of Academy of Military Medical Sciences, Beijing 100071, China*
[c] *Department of Obstetrics and Gynecology, Chinese PLA General Hospital, Beijing 100853, China*
[d] *Berry Genomics, C., Limited, Beijing 100015, China*

Received 20 October 2014; revised 6 March 2015; accepted 8 March 2015
Available online 14 March 2015

ABSTRACT

Reliable and accurate pre-implantation genetic diagnosis (PGD) of patient's embryos by next-generation sequencing (NGS) is dependent on efficient whole genome amplification (WGA) of a representative biopsy sample. However, the performance of the current state of the art WGA methods has not been evaluated for sequencing. Using low template DNA (15 pg) and single cells, we showed that the two PCR-based WGA systems SurePlex and MALBAC are superior to the REPLI-g WGA multiple displacement amplification (MDA) system in terms of consistent and reproducible genome coverage and sequence bias across the 24 chromosomes, allowing better normalization of test to reference sequencing data. When copy number variation sequencing (CNV-Seq) was applied to single cell WGA products derived by either SurePlex or MALBAC amplification, we showed that known disease CNVs in the range of 3−15 Mb could be reliably and accurately detected at the correct genomic positions. These findings indicate that our CNV-Seq pipeline incorporating either SurePlex or MALBAC as the key initial WGA step is a powerful methodology for clinical PGD to identify euploid embryos in a patient's cohort for uterine transplantation.

KEYWORDS: Single cells; Whole genome amplification; Next-generation sequencing; Copy number variation; Pre-implantation genetic diagnosis

INTRODUCTION

Whole genome amplification (WGA) is a technology designed to amplify small amounts of DNA down to the single cell level and generate a representative DNA template sufficient for downstream genetic analysis using conventional molecular techniques. A variety of WGA methodologies have evolved but are essentially based on principles of primer extension PCR (PEP) (Zhang et al., 1992), degenerate oligonucleotide primed PCR (DOP-PCR) (Telenius et al., 1992) or multiple displacement amplification (MDA) (Dean et al., 2002). Applications of WGA are widespread in clinical practice, including genetic typing of circulating tumor cells (Czyz et al., 2014), identification of forensic samples (Cai et al., 2010), and chromosome and single gene analysis of embryos from patients undergoing assisted reproduction and pre-implantation genetic diagnosis (PGD) (Zhang et al., 1992).

In the field of PGD, array-based comparative genomic hybridization (CGH) and single-nucleotide polymorphism (SNP) arrays are widely used to discriminate between euploid and

* Corresponding authors. Tel/fax: +86 10 6693 9258 (Y. Yao); Tel: +86 10 5325 9188x5073, fax: +86 10 8430 6824 (D.S. Cram).
E-mail addresses: david.cram@berrygenomics.com (D.S. Cram); yqyao_ghpla@163.com (Y. Yao).
[1] These two authors contributed equally to this work.

http://dx.doi.org/10.1016/j.jgg.2015.03.001

aneuploid embryos in the cohort produced by assisted reproductive technologies (Munne, 2012; Handyside, 2013). More recently, several next-generation sequencing (NGS) protocols have been developed and validated (Hou et al., 2013; Fiorentino et al., 2014; Huang et al., 2014; Wang et al., 2014b, 2014c; Wells et al., 2014). Both array and NGS methodologies rely solely on a pre WGA step to amplify the small amount of DNA in an embryo biopsy sample with high fidelity, in order to generate sufficient embryo DNA for either hybridization or sequencing.

Currently, the WGA kits that are used in PGD are commercially available (de Bourcy et al., 2014), and have been specifically optimized for single cell application. However, no WGA method has been reported with the capacity to replicate the entire single cell genome (Zheng et al., 2011). Further, from the analysis of the sequences that are amplified by WGA, there is a substantial bias due to the cumulative effects of allelic dropout or preferential allelic amplification (Findlay et al., 1995; Zheng et al., 2011; Binder et al., 2014). To further improve NGS-based technologies for comprehensive PGD of the full spectrum of chromosomal abnormalities afflicting human embryos (Vanneste et al., 2009; Voet et al., 2011), the WGA reaction remains the most critical step that determines the overall diagnostic performance of both array and sequencing based methods.

In this study, we examined the performance of two PCR-based methods SurePlex (Fiorentino et al., 2011; Shen et al., 2013) and MALBAC (Zong et al., 2012) as well as the multiple displacement amplification (MDA)-based method REPLI-g (Treff et al., 2011) for genome coverage and bias, and their potential as a starting template for detection of copy number variation (CNV) using our recently described NGS method called CNV-Seq (Liang et al., 2014). Here, we showed that both PCR-based WGA methods used in conjunction with CNV-Seq, are highly effective for identification, delineation and quantitation of small CNVs in single cells.

RESULTS

Evaluation of WGA methods for genome coverage of low template DNA

Three commonly used WGA methods based on either PCR (SurePlex, MALBAC) or MDA (REPLI-g) were evaluated with the aim of determining which methodology can provide the most uniform and reproducible coverage of the 24 chromosomes in a single cell for CNV analysis. For initial experimentation, we used 15 pg genomic DNA samples, which corresponds to approximately the same chromosomal DNA content in a single cell (6.6 pg of DNA), taking into account the random loss of DNA due to the effects of limiting dilution. Replicate 46,XX genomic DNA samples (15 pg) were prepared and subjected to WGA using the three different methods. Agarose gel electrophoresis was used to determine the yield and size range of the WGA reaction products (Fig. 1A). From five replicates, the mean yield was 4.15 ± 0.40 μg (SurePlex), 3.18 ± 0.39 μg (MALBAC) and 21.91 ± 3.50 μg (REPLI-g). The size range of the different WGA products was 0.2–0.8 kb (SurePlex), 0.2–2.0 kb (MALBAC) and >10 kb (REPLI-g), respectively, and patterns were consistent with previous results from single cells (Treff et al., 2011; Zong et al., 2012; Wang et al., 2014c).

To evaluate the genome coverage of the different WGA systems in more depth, we first performed WGA reactions from 15 pg samples, constructed libraries from 50 ng of the WGA products and 50 ng of unamplified DNA, and then performed massively parallel sequencing and chromosome mapping of the unique reads. Density plots of the distribution sequencing reads (100 kb sequencing bins) for all 24 chromosomes were compared between the three different WGA methods and representative results were shown for Chr. 1, Chr. 12 and Chr. 18 (Fig. 1B; refer to Fig. S1 for remaining chromosomes). In contrast to the uniform density of sequencing reads across Chr1, Chr12 and Chr18 for non-amplified genomic DNA, the patterns for amplified low template genomic DNA were non-uniform with a significant number of regions with high read numbers (peaks) and low read numbers (troughs). Nonetheless, the overall density profiles were relatively similar, regardless of the WGA method and this held true for the other chromosomes (Fig. S1). As judged by the distribution and fluctuation of the peaks and troughs, both PCR-based WGA methods exhibited more uniformity and potentially less regional amplification bias than the MDA-based WGA system.

The two PCR-based WGA systems were further evaluated for reproducing consistent chromosomal density read profiles. This is of particular importance for the application of NGS-based technologies for CNV detection since a high degree of reproducibility would allow the application of normalization algorithms and thus more meaningful comparisons of test to reference samples. Three replicate WGA products derived from 15 pg of genomic DNA were sequenced and density plots for Chr1, Chr12 and Chr18 were analyzed (Fig. 2). Replicate sequencing read density distributions across these three example chromosomes, and the remaining chromosomes, were again very similar for both SurePlex and MALBAC methods. Thus, both PCR-based WGA systems showed reproducible chromosome-specific coverage from a low template DNA sample mimicking the DNA content of a single cell.

Sensitivity, specificity and reproducibility of CNV diagnosis from different WGA templates

The suitability of the three different WGA systems to provide a starting template for reliable and accurate chromosome analysis by CNV-Seq was assessed using 15 pg DNA samples with known clinically-significant CNVs. For these experiments, we selected four genomic DNA disease models with variably sized terminal deletions or duplications (Table 1), including an unbalanced t(1,X) translocation (52.24 Mb Xp del, 92.8 Mb 1q dup), Sotos syndrome (15.4 Mb 5p dup, 6.44 Mb 5q del), Jacobsen syndrome (12.38 Mb 11q del) and Wolf-Hirschhorn syndrome (3.32 Mb 4p del). CNV-Seq analysis of SurePlex and MALBAC products detected all the six types of CNVs with sizes and map intervals similar to those detected in matching unamplified genomic DNA