Int J Clin Exp Pathol 2016;9(7):6965-6976
www.ijcep.com /ISSN:1936-2625/IJCEP0028503

Original Article

Whole genome amplification effect on segmental copy-number changes and copy-number neutral loss of heterozygosity analysis by oligonucleotide-based array-comparative genomic hybridization in human myeloma cell line

Aneta Mikulasova[1,2,3,4], Jan Smetana[1,2], Marketa Wayhelova[1,2], Helena Janyskova[1], Samuel A Okubote[3], Roman Hajek[4,5], Petr Kuglik[1,2]

[1]Department of Experimental Biology, Faculty of Science, Masaryk University, Brno, Czech Republic; [2]Department of Medical Genetics, University Hospital Brno, Brno, Czech Republic; [3]Department of Pathological Physiology, Faculty of Medicine, Masaryk University, Brno, Czech Republic; [4]Faculty of Medicine, University of Ostrava, Ostrava, Czech Republic; [5]Department of Hematooncology, University Hospital Ostrava, Ostrava, Czech Republic

Received March 17, 2016; Accepted May 26, 2016; Epub July 1, 2016; Published July 15, 2016

Abstract: Whole genome amplification (WGA) is an approach designed to overcome small amounts of DNA for genome-wide genetic tests used in many clinical applications. Various strategies of WGA have been developed; however, none of them can guarantee the absence of amplification bias. In this study, a total of four multiple displacement amplification (MDA)-based and two PCR-based WGA kits were compared in their effect on segmental copy-number (CN) changes and copy-number neutral loss of heterozygosity (cnnLOH) detection by three microarray platforms: CGH/4×44K (Agilent), CGH+SNP/4×180K (Agilent) and CGH+SNP/4×180K (OGT). Genomic imbalances-rich myeloma cell line U266 was used as material. The main outcomes are as follows: 1) MDA-based WGAs showed higher tendency to generate false positive imbalances in contrast to PCR-based WGAs with higher risk of false negativity; 2) the specific risk of false positivity and/or negativity increased with decreasing CN segments size; 3) single-cell WGAs showed significantly worse effect on results in comparison to WGAs with nanogram level of DNA as input; 4) PCR-based WGAs were incompatible with cnnLOH analysis based on SNP in restriction digestion sites and also showed higher risk of cnnLOH false negativity when combined with analysis based on simple hybridization. In conclusion, the results of this study help to choose WGA according to individual user requirements and options. Moreover, we have shown a strategy to verify and validate segmental CN changes detection by DNA array protocol including any WGA for any purpose to attain the highest efficiency without an unnecessary WGA bias.

Keywords: Whole genome amplification, array-comparative genomic hybridization, copy-number changes, copy-number neutral loss of heterozygosity, genotyping

Introduction

Germline and somatic genetic mutations are causative of serious congenital and acquired human diseases. Early and accurate diagnosis of these disorders plays a significant role in clinical genetics, for instance, it can help to avoid the birth of an affected child, to explain reasons for infertility or determine patients' prognosis and modulate effective therapy. A number of modern methods, from basic targeted PCR-based analysis to genome-wide screening techniques, such as array-comparative genomic hybridization (aCGH) or next generation sequencing (NGS), have been developed to test genetic causes of disorders. However, performing an adequate diagnostics is limited by insufficient amount of DNA in a number of cases. To address this problem, whole genome amplification (WGA) methods have been developed using two different strategies. The first strategy is based on temperature cycling (i.e. PCR-based) WGAs, for instance degenerate oligonucleotide primed-PCR (DOP-PCR) [1, 2], or

primer extension preamplification (PEP) [3, 4]. More recently, Omniplex WGA as a newer PCR-based WGA has been developed as a random fragmentation of genomic DNA followed by isothermal primer extensions to prepare the OmniPlex library amplified by PCR using universal primers [5, 6]. The second strategy involves isothermal WGAs as multiple displacement amplification (MDA) based on phi29 DNA polymerase activity and random primers to amplify the entire genome [7, 8]. In contrast to PCR-based WGAs, MDA provides long DNA products (on average about 10 kbp) gained by isothermal reaction without repeated cycling. Furthermore, multiple annealing and looping based amplification cycles (MALBAC) strategy has been developed as a PCR-based method that includes a limited MDA pre-amplification phase preceding PCR cycles [9]. Despite the advancement of WGA methods, the risk of amplification bias may affect analysis results. Recently, WGAs have been combined with various techniques to evaluate their suitability for different clinical applications; for instance, single nucleotide polymorphisms (SNPs) and/or short tandem repeats (STRs) genotyping [10-15], restriction fragment length polymorphism (RFLP) analysis [8], Southern blotting analysis [8], quantitative real-time PCR [8, 16], high resolution melting analysis [17-19], CGH [8, 20], aCGH [21-24] or recently NGS [25-28].

The use of WGA in combination with aCGH has become very promising, especially for genome-wide imbalances analysis in preimplantation [29, 30] or tumor genetics [16, 31, 32]. More or less, many studies have shown the suitability of both PCR-based WGA [6, 16, 33] and MDA-based WGA [21, 30, 34] for aCGH. Nevertheless, the results of these studies are often difficult to compare to each other due to methodological background variability, such as differences in DNA resource used as template, WGA protocol procedure, aCGH platform or bioinformatics/biostatistics algorithms and tools. Therefore, evaluation of WGA effect on downstream methods with a comparable background, and detailed and systematic analysis of WGA effect on aCGH analysis is challenging.

In the current study, we designed an experiment to assess the effect of WGAs on segmental copy-number (CN) changes and copy-number neutral loss of heterozygosity (cnnLOH) using oligonucleotide DNA arrays. We chose six commercial WGA kits and tested them in combination with three different aCGH platforms using a human myeloma cell line (HMCL) U266 which is rich in genomic imbalances of variable size across the genome.

Materials and methods

DNA samples

DNA was isolated from HMCL U266 using Gentra Puregene Kit (Qiagen, Hilden, Germany) from 10 million cells gained from one passage. As reference DNA, Human Genomic DNA (Promega, Madison, WI, USA) and Human Reference DNA (Agilent Technologies, Santa Clara, CA, USA) were used. Unamplified and WGA DNA samples were quantified by Qubit Fluorometer 1.0 (Thermo Fisher Scientific, Waltham, MA, USA). The length of WGA products was verified by 2% agarose electrophoresis, using 1 Kbp Plus DNA Ladder as control (Thermo Fisher Scientific).

Whole genome amplification

In total, 50 ng of HMCL U266 DNA and reference DNA were amplified by four commercial WGA kits, hereafter referred to as 'standard WGAs' (three MDA-based and one PCR-based) as follows: REPLI-g Mini Kit, REPLI-g Midi Kit (both Qiagen), illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare Life Sciences, Piscataway, NJ, USA) and GenomePlex Complete WGA Kit (Sigma-Aldrich, St. Louis, MO, USA). In total, 0.15 ng of HMCL U266 DNA and reference DNA were amplified by two commercial WGA kits, hereafter referred to as 'single-cell WGAs' (scWGAs) (one MDA-based and one PCR-based) as follows: REPLI-g Single Cell Kit (Qiagen) and PicoPlex WGA Kit (Rubicon Genomics, Ann Arbor, MI, USA). All WGAs protocols were done according to manufacturer recommendation. WGA DNA samples were purified by ethanol precipitation and precipitated DNA samples were diluted in TE buffer.

Array-comparative genomic hybridization

Unamplified DNA samples and WGA DNA samples were tested in duplicates by three aCGH platforms: Human Genome CGH Microarray, 4×44K and SurePrint G3 CGH+SNP Microarray, 4×180K (both Agilent Technologies); and

Table 1. WGAs false positivity and false negativity copy-number (CN) changes risk analysis

WGA (type)	CN type	False positivity			False negativity		
		Median size (range) [Mbp]	Increase compared to control	*P*	Median size (range) [Mbp]	Increase compared to control	*P*
Control	Losses	0.3 (0.1–2.0)	---	---	1× 0.1, 1× 0.3	---	---
	Gains	0.8 (0.3–6.9)	---	---	1× 0.4	---	---
RgMidi (standard MDA-based)	Losses	0.3 (0.1–2.0)	+5.4% (1.6×)	0.14	0.1 (0.1–0.7)	+3.8% (4.0×)	0.10
	Gains	0.3 (0.1–1.1)	+6.5% (2.3×)	0.06	none	no increase	1.00
RgMini (standard MDA-based)	Losses	0.4 (0.1–4.9)	+14.5% (2.6×)	1.92×10^{-4}	0.4 (0.1–0.7)	+1.3% (2.0×)	0.68
	Gains	0.4 (0.1–0.9)	+0.7% (1.1×)	1.00	none	no increase	1.00
GenomiPhi (standard MDA-based)	Losses	0.2 (0.1–2.8)	+8.1% (1.9×)	3.08×10^{-2}	0.3 (0.2–0.4)	+1.9% (2.5×)	0.45
	Gains	0.8 (0.2–6.8)	+12.2% (3.4×)	1.55×10^{-3}	2× 0.4	+0.8% (2.0×)	1.00
RgSC (single-cell MDA-based)	Losses	3.3 (0.1–121.7)	+31.4% (4.4×)	$< 10^{-9}$	0.2 (0.1–2.5)	+35.4% (29.0×)	$< 10^{-9}$
	Gains	0.4 (0.1–10.7)	+24.5% (5.7×)	1.60×10^{-8}	0.4 (0.4–6.6)	+10.9% (15.0×)	3.54×10^{-4}
GenomePlex (standard PCR-based)	Losses	0.3 (0.1–1.1)	no increase	0.14	0.2 (0.1–0.7)	+4.4% (4.5×)	0.06
	Gains	1× 0.4	no increase	0.07	1× 0.4	no increase	1.00
PicoPlex (single-cell PCR-based)	Losses	0.4 (0.1-3.0)	no increase	0.21	0.2 (0.1–5.3)	+37.3% (30.5×)	$< 10^{-9}$
	Gains	3.4 (0.4–9.4)	no increase	1.00	3.6 (0.4–48.2)	+21.1% (28.0×)	2.40×10^{-8}

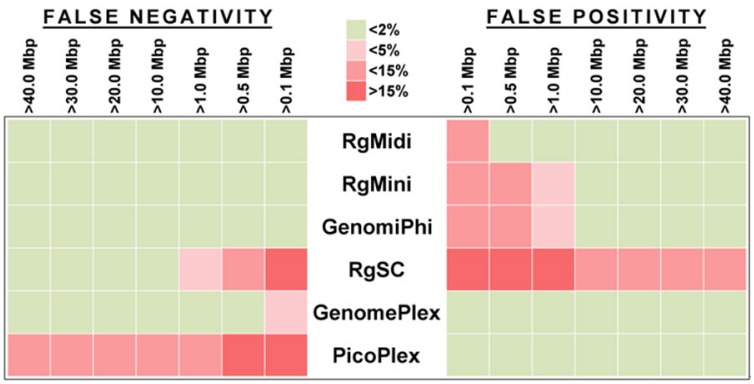


Figure 1. WGAs false positivity and false negativity copy-number changes risk in relation to copy-number changes size. Different colors show false positivity and false negativity percentage increase in each WGA in comparison to control, gradually for segments with the size of > 0.1 Mbp, > 0.5 Mbp, > 1.0 Mbp, > 10.0 Mbp, > 20.0 Mbp, > 30.0 Mbp and > 40.0 Mbp.

CytoSure ISCA UPD Array, 4×180 K (OGT; Oxford Gene Technology, Oxfordshire, UK). In total, 42 hybridizations were included in analysis with an input of 1.5 µg DNA. Tested HMCL U266 DNA samples were co-hybridized with reference DNA samples prepared by the same procedure, meaning pairs of unamplified HMCL U266 DNA versus unamplified reference DNA and WGA HMCL U266 DNA versus WGA reference DNA were prepared. Unamplified and MDA-based WGA DNA samples were fragmented by *AluI* and *RsaI* (both Promega) restriction enzymes. PCR-based WGA DNA samples were digested only in combination with SurePrint G3 CGH+ SNP Microarray, 4×180K (Agilent Technologies) platform. DNA samples were fluorescent labeled by BioPrime Total for Agilent aCGH Kit (Thermo Fisher Scientific); Alexa Fluor 3 for reference DNA and Alexa Fluor 5 for HMCL U266 DNA. After purification of labeled DNA, reference DNA and HMCL U266 DNA samples were together with COT Human DNA (Hoffman-La Roche, Basel, Switzerland) and hybridization mix (Oligo aCGH Hybridization Kit, Agilent Technologies) co-hybridized to the arrays. After 24-hour (Agilent platforms) or 48-hour (OGT platform) hybridization and washing, DNA arrays were scanned by Microarray Scanner (Agilent Technologies). Feature Extraction Software-v12.0.2.2 (Agilent Technologies) was used for data extraction and quality control evaluation. Genomic Workbench v7.0.4.0 (Agilent Technologies) and CytoSure Interpret Software v4.3.2 (OGT) were used for data analysis. For segmentation, ADM-2 algorithm and following setting: ≥ 100 kbp size, ≥ 0.25 fold change of log2 ratio (L2R), ≥ 5 and ≥ 3 (180K and 44K, respectively) consecutive probes were used. Database of Genomic Variants (http://www.openhelix.com) for hg19 was used to differentiate between somatic CN alterations (CNAs) and common germinal CN variations (CNVs). Autosomal CNAs/CNVs were included in the analysis. Manufacture recommended setting was used for LOH calling: confidence level 0.95 with threshold 6.0 in case of SurePrint G3 CGH+SNP Microarray, 4×180K

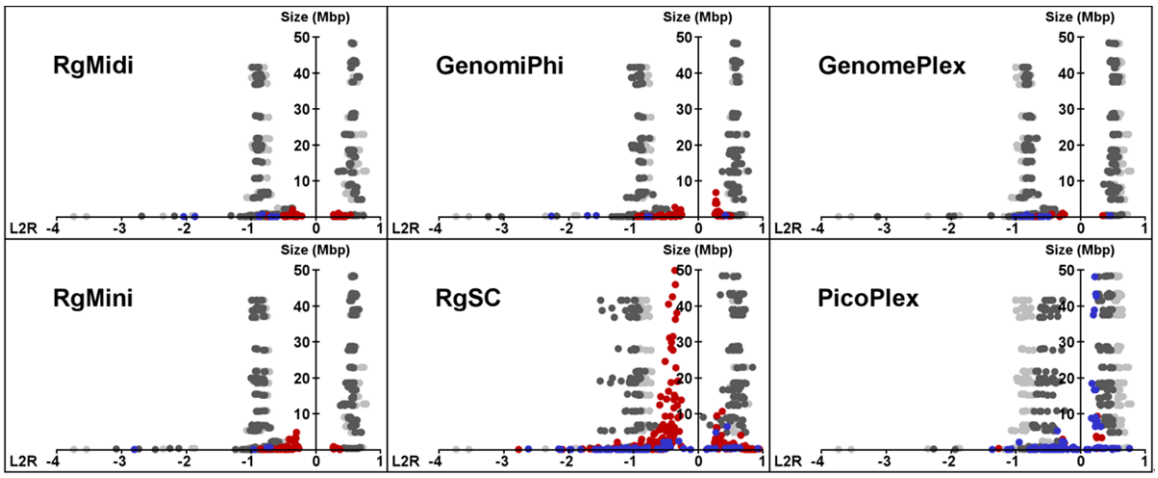


Figure 2. Graphical view of genome-wide copy-number changes in human myeloma cell line U266 in WGA samples. Log2 ratio (L2R) and size of copy-number changes are shown in axis x and y, respectively. Light and dark grey points represent proper copy-number changes in control and WGA samples, respectively. Red and blue points show false positive and false negative copy-number changes in WGA samples, respectively.

Agilent platform; and 80% threshold homozygosity and score above 140 in case of CytoSure ISCA UPD Array, 4×180K OGT platform. Evaluation of unamplified control aCGH profiles are described in Supplemental Methods. The array data supporting the results of this article are available at Gene Expression Omnibus (GEO), National Center for Biotechnology Information (NCBI) under the accession number GSE73513.

Statistical analysis

Basic statistical analysis was done using IBM SPSS Statistics v22 (IBM, Armonk, NY, USA), MedCalc software v14.8.1 (MedCalc Software, Ostend, Belgium) and Statistica software v12 (StatSoft, Praque, Czech Republic). Statistical tests were used as follows: Fisher's exact test for categorical data, Mann-Whitney U test and Wilcoxon test for continuous variables and Pearson correlation. P values ≤ 0.05 were considered statistically significant.

Results

General WGA DNA products evaluation

For each WGA, DNA yields of eight HMCL U266 and eight reference DNA independent replicates were fluorometrically assessed. MDA-based WGAs synthesized higher amount of DNA in the median and showed longer-sized product (Supplementary Figure 1) in comparison to WGAs including PCR cycles. RgMidi (35.6 µg, range 27.9–41.0 µg) and RgSC (35.7 µg, range 31.9–38.5 µg) provided the highest amount of DNA, followed by RgMini (5.3 µg, range 4.4–7.2 µg), GenomiPhi (4.7 µg, range 4.0–5.2 µg), GenomePlex (4.6 µg, range 4.1–6.4 µg) and PicoPlex (2.7 µg, range 2.5–2.9 µg).

WGAs false positivity and false negativity of copy-number changes

DNA isolated from HMCL U266 and reference DNA samples were amplified separately by six commercial WGA kits and processed by the same aCGH protocol and data analysis as unamplified control samples. A total of 143 CN changes (median size of 12.7 Mbp, range 0.1–48.4 Mbp) were included in this study. FP and FN of CN changes analysis for each WGA is summarized in **Table 1**.

We observed that MDA-based WGAs showed higher risk of FP than PCR-based WGAs. Significant FP increase was present in RgMini for losses ($P = 1.92 \times 10^{-4}$), in GenomiPhi for both losses and gains ($P = 3.08 \times 10^{-2}$ and $P = 1.55 \times 10^{-3}$, respectively) and the most significant in RgSC for losses as well as gains ($P < 10^{-9}$ and $P = 1.60 \times 10^{-8}$, respectively). Statistically significant FN was identified only in both scWGAs RgSC and PicoPlex for losses (both $P < 10^{-9}$) as well as gains ($P = 3.54 \times 10^{-4}$ and $P = 2.40 \times 10^{-8}$, respectively).

As we expected, the risk of both FP and FN increased with decreasing CN segments size in general (**Figure 1**, Supplementary Figure 2). In

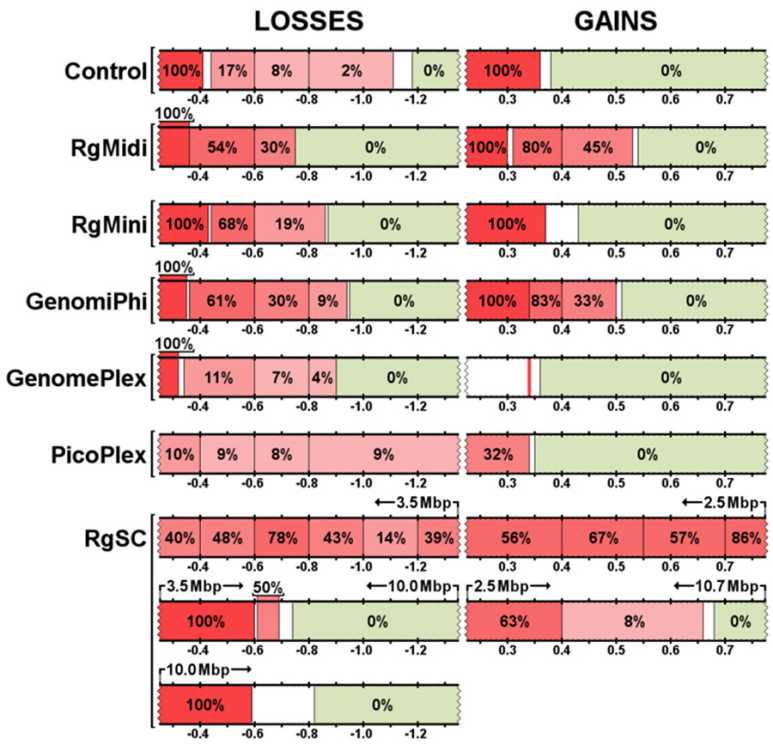


Figure 3. False positivity risk in relation to copy-number change log2 ratio values in control and each WGA. Copy-number change log2 ratio values are displayed in horizontal axis separately for losses (left) and gains (right). Different colors support false positivity risk rate, from green (0%) to red (100%), in specific log2 ratio values intervals. Losses and gains < 10.0 Mbp were included in analysis in case of control, RgMidi, RgMini, GenomiPhi, GenomePlex and PicoPlex; and all losses (divided into three intervals: < 3.5 Mbp, 3.5 Mbp – 10.0 Mbp and > 10.0 Mbp) and gains < 10.7 Mbp (divided into two intervals: < 2.5 Mbp and 2.5 Mbp – 10.7 Mbp) were included in analysis in case of RgSC.

control and in all standard WGAs, FP of losses, FP of gains, FN of losses and FN of gains were < 5.0 Mbp, < 10.0 Mbp, < 1.0 Mbp and < 0.5 Mbp, respectively. MDA-based scWGA RgSC showed FP of losses also in size 5.0–10.0 Mbp (47.1%) and even in size > 10.0 Mbp (30.5%) with a maximum size of 121.7 Mbp. The size of FP of gains slightly overcame 10.0 Mbp with the maximum size of 10.7 Mbp. Interestingly, FP discrepancies in PCR-based scWGA PicoPlex were comparable to standard WGAs. On the other hand, both scWGAs RgSC and PicoPlex showed very high FN of losses as well as gains. Overall, 80.3% ($P < 10^{-9}$) and 83.3% ($P < 10^{-9}$) of losses < 1.0 Mbp were not recognized in RgSC and PicoPlex, respectively. Moreover, 62.5% of losses 1.0–5.0 Mbp were not found in both scWGAs and 5.6% of losses 5.0–10.0 Mbp were not found in PicoPlex. Only 10.0% ($P = 1.09 \times 10^{-3}$) and even none ($P = 1.19 \times 10^{-4}$) of gains < 0.5 Mbp were recognized in RgSC and PicoPlex, respectively, and 66.7% of gains 0.5–1.0 Mbp were not found in both scWGAs. In addition, RgSC and PicoPlex showed 9.1% and 27.3% FN risk of gains 1.0–10.0 Mbp, respectively, and PicoPlex caused even 8.9% FN risk of gains > 10.0 Mbp with the maximum size of 48.2 Mbp.

WGAs false positivity of copy-number changes prediction in relation to log2 ratio values

aCGH allows the detection of unbalanced chromosome changes (losses and gains). Although these changes are counted by advanced statistical algorithm, essentially, segments calling comes from L2R of two different fluorescent signals getting from tested and reference samples, which are counted for each probe and finally for recognized segments. Therefore, each segment can be described by two features – size and L2R value. The overview of HMCL U266 genome-wide segments described by these two features in each WGA is summarized in **Figure 2**.

In this part of analysis, we focused on L2R study of CN segments < 10.0 Mbp as an upper limit of problems with FP in general. Only in case of RgSC, segments > 10.0 Mbp were studied due to FP presence over this size. We found that L2R values of specific segments can be helpful for FP prediction (Supplementary Table 1). In control and standard WGAs, L2R values of FP segments were significantly closer to L2R = 0 than real CN segments. In all these cases, we clearly recognized L2R limits specific for FP and real segments and usually also a region of L2R values as a mixture of both with the FP risk increasing with L2R value approaching to L2R = 0 (**Figure 3**).

In case of PicoPlex, we identified less significant differentiation of FP from real gains and

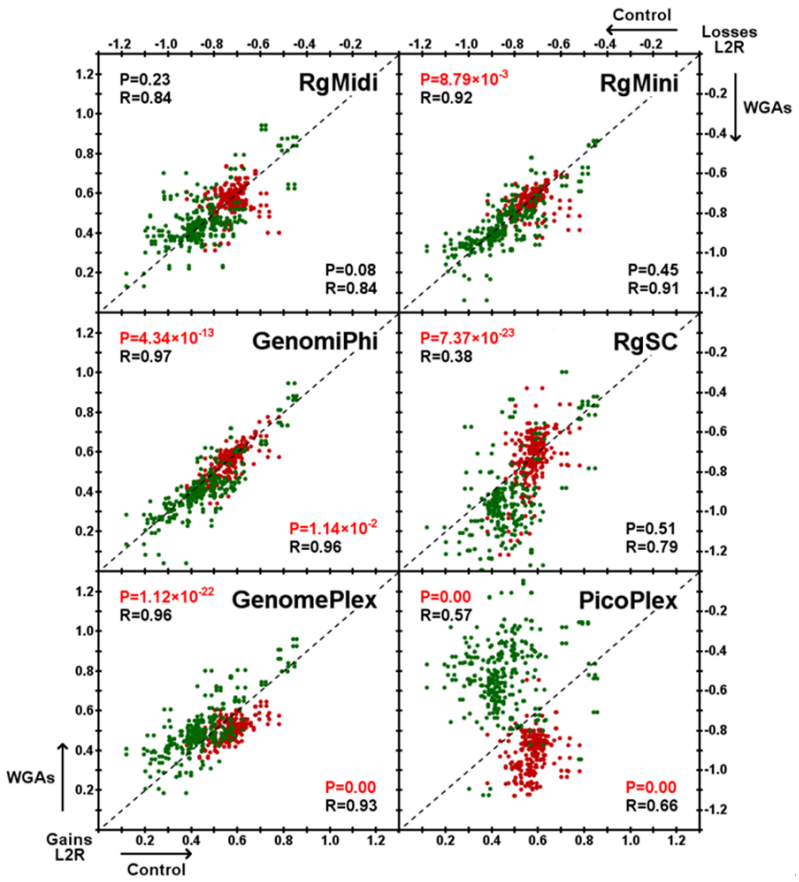


Figure 4. WGA effect on log2 ratio values of real copy-number changes. Log2 ratio (L2R) values of copy-number changes in control and WGA samples are shown in axis x and y, respectively. Red and green points show gains and losses, respectively. *P* value (Wilcoxon test) describes a shifting of log2 ratios after WGAs and *R* value (Pearson correlation) describes a correlation between log2 ratio values of WGAs with unamplified control (for all $P_R < 10^{-4}$)-losses analysis in upper left corners and gains analysis in lower right corners.

no statistic difference between FP and real losses corresponding with low but continuous risk of FP losses independently of L2R values. Interestingly, GenomePlex also showed smaller, but still significant FP and real losses differentiation in comparison to standard MDA-based WGAs ($P < 10^{-2}$ versus $P < 10^{-17}$ – $< 10^{-6}$). On the other hand, overall FP risk in both PicoPlex and GenomePlex was much lower in comparison to MDA-based WGAs as we described above.

In case of RgSC, we found three and two size intervals of losses and gains, respectively, with a different behavior of L2R values. Based on these results, we determined the minimal size limit for CN changes detection as 3.5 Mbp for losses and 2.5 Mbp for gains. Up to these sizes, very unreliable data was present as 42.6% and 72.9% of losses and gains, respectively, were FP with no possibility to recognize FP segments from real segments using L2R values; and 78.4% and 81.3% proper losses and gains, respectively, were not found at all. In the second size interval (3.5–10.0 Mbp for losses and 2.5–10.7 Mbp for gains), the risk of FN was significantly reduced (no risk for losses and 9.1% risk for gains), but the risk of FP remained high (57.1% for losses and 33.3% for gains). On the other hand, FP segments were significantly closer to L2R = 0 than real CN segments in the second intervals as in standard WGAs analysis for CN changes < 10.0 Mbp. Moreover, high FP risk of losses > 10.0 Mbp described above in RgSC was completely eliminated by more strict L2R threshold as all proper losses > 10.0 Mbp showed L2R < -0.8 while FP losses > 10.0 Mbp had L2R > -0.6.

WGAs effect on log2 ratio values

In theory, when one sample is tested by aCGH repeatedly, the same L2R value of each probe as well as each called loss and gain should be the same, in an ideal situation log2 (1/2) = -1.0 for losses, log2 (3/2) = 0.58 for gains and log2 (2/2) = 0 for diploid regions. In reality, L2R values are influenced by random effects within a particular experiment which can lead to natural differences of L2R values. With a regard to this natural background variability, we studied how WGAs changed L2R values of individual probes and segments and we found some specific effects which can partly explain the presence of higher FP or FN in some WGAs.

We analyzed L2R values changes of each proper loss and gain after each WGA and we found that MDA-based WGAs RgMini, GenomiPhi and

Table 2. Analysis of WGA effect on genotyping tested by Agilent SNP platform

WGA	% of failed SNP loci	SNP locus failures increase compared to control	Genotype concordance	AA/BB:AB loci failure ratio	LOH (increase)	GOH (increase)	LOH:GOH
Control	23.4%	---	98.9%	1.6:1.0	0.2%		---
RgMidi	41.7%	1.8× (+18.3%)	97.9%	1.0:1.4	0.6% (3.8×)	0.6% (3.7×)	1.0:1.0
RgMini	36.8%	1.6× (+13.3%)	98.7%	1.0:1.1	0.3% (2.3×)	0.5% (3.3×)	1.0:1.5
GenomiPhi	35.0%	1.5× (+11.6%)	98.6%	2.2:1.0	0.2% (1.4×)	0.6% (4.0×)	1.0:2.8

LOH: loss of heterozygosity (AB to AA/BB). GOH: gain of heterozygosity (AA/BB to AB).

especially RgSC caused a significant shift of losses in direction from L2R = 0 ($P = 8.79 \times 10^{-3}$, $P = 4.34 \times 10^{-13}$ and $P = 7.37 \times 10^{-23}$, respectively) or intensified L2R values of losses (**Figure 4**). Both PCR-based WGAs GenomePlex and especially PicoPlex showed significant shifting of losses ($P = 1.12 \times 10^{-22}$ and $P = 0.00$, respectively) as well as gains (both $P = 0.00$), both in direction to L2R = 0. Interestingly, a tendency of losses L2R value shift was in opposite direction in comparison to MDA-based WGAs. Only RgMidi did not show any significant movement of L2R values. As we expected, scWGAs RgSC and PicoPlex showed decrease of segments L2R values correlation ($R < 0.80$, $P_R < 10^{-4}$) while standard WGAs kept high correlation with the unamplified control ($R > 0.80$, $P_R < 10^{-4}$; **Figure 4**).

Compatibility of WGAs and SNP arrays

Taken together, cnnLOH analysis by restriction digestion based Agilent SNP platform was successful only in combination with standard MDA-based WGAs as RgMidi, RgMini and GenomiPhi, while cnnLOH analysis by restriction digestion independent OGT SNP platform was successful in combination with standard MDA-based WGAs as RgMidi, RgMini and GenomiPhi as well as PCR-based WGAs as GenomePlex and PicoPlex. Only RgSC did not provide analyzable cnnLOH data in combination with both Agilent and OGT SNP platforms presumably due to high DLRS as the main indicator of aCGH data quality.

Despite the fact that both PCR-based WGAs GenomePlex and PicoPlex were digested by *AluI/RsaI* enzymes before hybridization to Agilent SNP platform, they were not compatible with an analysis of cnnLOH by this platform, likely due to unspecific DNA fragmentation during WGA protocols. We assumed there might a connection between incompatibility of specific WGAs with Agilent SNP platform and a high number of SNP loci with failed genotype. In total, 79.1%, 75.5% and 75.3% of SNP loci were unrecognized in GenomePlex, PicoPlex and RgSC, respectively. All standard MDA-based WGAs as RgMidi, RgMini and GenomiPhi showed not so high, but still higher number of SNP loci with failed genotype (41.7%, 36.8% and 35.0%, respectively) in comparison to control (23.4%), but provided analyzable SNP data (**Table 2**).

WGAs effect on copy-number neutral loss of heterozygosity analysis

All 10 cnnLOH areas were found and no FP cnnLOH area was detected by Agilent and OGT SNP platforms in case of all standard MDA-based WGAs RgMidi, RgMini and GenomiPhi. In case of PCR-based WGAs, which were compatible only with OGT SNP platform, GenomePlex showed the same match in cnnLOH areas detection as MDA-based WGAs, but 40.0% (8/20) of cnnLOH areas with the median size of 28.0 Mbp (15.8–68.3 Mbp) were not recognized in PicoPlex. Although the FN risk of cnnLOH was detected only in case of PicoPlex, additional detailed analysis of cnnLOH areas showed that all other WGAs were also associated with the potential risk of FN. This FN risk was manifested as partial failures of homozygosity in detected cnnLOH areas. Overall, 9.4%, 9.4% and 6.0% of a total of 303.8 Mbp of cnnLOH across the tested genome failed in RgMidi, RgMini and GenomiPhi, respectively, thus higher than in control (1.1%; Supplementary Table 2) in case of Agilent SNP platform. Interestingly, using OGT SNP platform, these failures were more significant in PCR-based WGAs GenomePlex and PicoPlex (5.3% and 47.6% of a total 301.9 Mbp of cnnLOH, respectively) than in MDA-based WGAs RgMidi, RgMini and GenomiPhi (1.5%, 0.6% and 0.5% of a total 301.9 Mbp of cnnLOH, respectively; Supplementary Table 2).

When two unamplified controls analyzed by Agilent SNP platform were compared, ratio of cnnLOH areas with decreased and increased

homozygosity was 1.0:1.0 and 100.0% cnnLOH cases showed < 1.0% homozygosity deviation. In case of OGT SNP platform, perhaps due to a greater spacing between tested SNP loci, higher homozygosity deviation of 10 cnnLOH areas between two unamplified controls was found. First control and second control showed 5.0× higher tendency to increase and to decrease homozygosity, respectively. Moreover, only 40.0% of cnnLOH areas showed homozygosity deviation < 1.0%, thus much less than 100.0% in case of Agilent SNP platform. However, in total, 100.0% of control cnnLOH areas showed homozygosity deviation < 4.0%.

Detailed analysis of WGA effect on 10 cnnLOH areas is shown in Supplementary Tables 3 and 4. MDA-based WGAs showed higher homozygosity deviation compared to unamplified controls than the comparison of two unamplified controls. In total, 82.5%/87.5%, 70.0%/90.0% and 67.5%/85.0% of cnnLOH detected by Agilent/OGT platform showed < 1.0%/< 4.0% homozygosity deviation in case of RgMidi, RgMini and GenomiPhi, respectively. Moreover, MDA-based WGAs led to clear tendency to decrease overall homozygosity in combination with Agilent SNP platform. The proportion of cnnLOH areas with decreased homozygosity was 4×, 9× and 7× higher than with increased homozygosity with the maximum homozygosity decrease -2.1%, -3.5% and -3.0% in RgMidi, RgMini and GenomiPhi, respectively. Interestingly, this trend was not found in combination with OGT platform. We found 4.0× (5.6:1.4) higher tendency to decrease than to increase homozygosity in RgMini, 2.0× (3.1:1.6) and 1.1× (2.1:1.9) higher tendency to increase than to decrease homozygosity in RgMidi and GenomiPhi, respectively, which does not reflect any specific trend in comparison to two unamplified controls reciprocal analysis. Importantly, in contrast to MDA-based WGAs analyzed by OGT SNP platform, both PCR-based WGAs GenomePlex and PicoPlex led to significant homozygosity decrease tendency as no cnnLOH area with homozygosity increase was present in both and the maximum of cnnLOH homozygosity decrease was even -19.8% and -36.4%, respectively. Moreover, homozygosity deviation < 4.0% was present only in 20.0% and 2.5% of cnnLOH areas in GenomePlex and PicoPlex, respectively. Additionally, GenomePlex and PicoPlex showed only 47.5% and 7.5% of cnnLOH areas, respectively, with homozygosity deviation < 10.0%, much lower than MDA-based WGAs (all 100.0%).

WGAs effect on genotyping analysis

Effect of WGAs on genotyping analysis was studied by Agilent SNP platform, which allows generating SNP genotype. Unfortunately, only standard MDA-based WGAs were compatible with the platform (RgMidi, RgMini and GenomiPhi). Although standard MDA-based WGAs increased proportion of failed SNP sites (< 2.0×) compared to control, overall genotype concordance with unamplified control was very high (> 97.0%) (**Table 2**).

SNP loci with a diploid character were used for loss of heterozygosity (LOH, AB to AA/BB), gain of heterozygosity (GOH, AA/BB to AB) and change of homozygosity (COH, AA to BB) analysis. MDA-based WGAs increased LOH and GOH (≤ 4.0×) compared to control, however a real proportion of loci with LOH and GOH was very low (≤ 0.6%) (**Table 2**). COH occurred as a rare discrepancy (< 2.0% of non-concordat calls) with a low overall frequency (< 0.1%) in MDA-based WGAs and was not present in control.

Discussion

In this study, we designed an experimental approach to assess the effect of six commercial WGAs on aCGH analysis of segmental CN and cnnLOH using oligonucleotide DNA arrays. We described specific outcomes of each WGA, especially specific risk of CN changes FP and FN, its relation to the CN segments size and L2R values, effect on cnnLOH as well as genotyping analysis. However, more importantly, our results showed specific feature of each WGA strategy, especially when MDA-based and PCR-based or scWGA and standard WGA were compared. For each experiment, reference DNA was amplified using the same WGA procedure and the same amount of starting DNA material as tested DNA as other studies [14, 21, 33, 35], however not all [24, 29], have showed that WGA bias is reduced when reference DNA is amplified.

We confirmed previous observations that MDA-based WGAs has a tendency to produce FP CN changes [34-36] and provide very high LOH concordance (nearly 100%) with unamplified

control [37]. We also showed that scWGAs provided more negative effect on aCGH analysis than standard WGAs. This was not surprising as it was found in other studies that WGA bias depends on the starting amount of DNA [21, 34, 38]. All standard WGAs showed problems in CN detection only < 10.0 Mbp in general, while scWGA showed specific risk of FP and/or FN over 10.0 Mbp. Effect of scWGAs were manifested as high FP/FN risk of CN changes in MDA-based scWGA and high FN risk of CN changes in PCR-based scWGA. Le Caignec *et al.* [29] described that the combination of MDA-based scWGA and BAC-based aCGH allowed to get no FP/FN at the whole chromosome level and recognized segmental deletions with the size of 34 Mbp and 58 Mbp as well as segmental gain with the size of 47 Mbp. Compared to that, we showed that using MDA-based scWGA and oligonucleotide-based aCGH, all segmental CN changes > 10.0 Mbp were recognized, but with 30.5% FP risk of losses. Using detailed study of effect of WGA to L2R values we also specified a minimal size limit for CN changes detection as 3.5 Mbp for losses and 2.5 Mbp for gains. Fiegler *et al.* [6] showed a possibility to detect segmental CN changes, for example a microdeletion with the size of ~10.8 Mbp, when DNA was amplified by PCR-based scWGA and combined with BAC-based aCGH. And, they also found FP CN changes in various experiments. As opposed to these findings, our analysis by oligonucleotide-based aCGH showed higher tendency of FN than of FP. However, comparison with these results can be questionable due to different type of WGA as well as aCGH platform. Effect of scWGA was reflected in the L2R values of CN changes. We hypothesized that L2R correlation decrease and specific and opposite tendency to change L2R values could partly explain the high FP in RgSC and the high FN in PicoPlex. Significant FN was present also in RgSC, but in smaller sizes in comparison to PicoPlex and so it was caused more likely by substantially reduced correlation and thus high background.

Barker *et al.* [39] showed comparable usability of MDA-based and PCR-based WGAs for SNP genotyping, however SNP and STR genotyping analysis done by Bergen *et al.* [10] or Shojaei Saadi *et al.* [38] showed that PCR-based WGAs produce more WGA bias than MDA-based WGAs. Both DNA array platforms used in this study are not primarily intended for genotyping; however, they are based on SNP analysis in principle. Agilent SNP platform allows generating SNP genotype, which was used for MDA-based WGA effect analysis. We showed that MDA-based WGAs are suitable for genotyping analysis with almost 100% genotype concordance and that these WGAs has a tendency to loci failure as has been described in previous studies [12, 15, 37, 40]. Our results are also consistent with the notion that COH represents a small portion of non-concordant calls found in previous studies [35, 41]. Xing *et al.* [41] observed higher tendency of MDA-based WGA to cause GOH than LOH, oppositely to other studies [35, 40]. We found LOH:GOH ratio 1.0:1.0, 1.0:1.5 and 1.0:2.8 for RgMidi, RgMini and GenomiPhi, respectively, corresponding more likely with the results of Xing *et al.* [41]. However, we supported their hypothesis that different results are associated with different array platforms and statistical algorithms. We also hypothesize that increased WGA bias found in this study for genotyping analysis was present more likely due to effect on SNP genotype/cnnLOH calling algorithms and so it did not completely reflect the real effect of WGA on DNA sequence. This hypothesis could explain RgSC incompatibility with cnnLOH analysis more likely due to aCGH quality than due to bias in amplified DNA product. To support that, it has been also shown that DNA amplified from single cell by MDA is suitable for SNP and STR genotyping analysis [13], although lower amount of DNA input is associated with higher WGA bias, especially with allele or locus dropout [11, 14].

In summary, based on our finding, we provide general recommendations for WGA selection as follows: 1) If possible, standard WGAs (nanogram quantities of DNA as input) should be used to avoid unnecessary bias of scWGAs (picogram quantities of DNA as input). 2) If segmental CN changes analysis is necessary at the single-cell level, genome-wide profiles should be interpreted cautiously, considering higher risk of FP and FN in case of MDA-based and PCR-based WGAs, respectively. 3) It should be taken into account that DNA fragmentation during PCR-based WGAs can interfere with some downstream applications. For instance, incompatibility of PCR-based WGAs with cnnLOH analysis based on SNP in restriction sites

was found in this study. 5) If high DNA yield of WGA is essential for some reason (for instance, downstream applications requirement, repeat testing possibility or archiving material for future purposes), MDA-based WGA is a more appropriate option. 6) If DNA array analysis is used only for segmental CN changes detection, PCR-based WGAs are more appropriate as they produce fewer imbalances than MDA-based WGAs. If cnnLOH analysis is included, MDA-based WGAs should be chosen as they provide more accurate results in cnnLOH detection than PCR-based WGAs. This last recommendation is applicable for most cases, but not generally. In this study, for instance, MDA-based scWGA did not show analyzable data for cnnLOH recognition. Analysis of cnnLOH at the single-cell level was possible only when PCR-based scWGA and OGT SNP platform were combined; however, high FN risk was found.

Acknowledgements

This work was supported by Internal Grant Agency of the Ministry of Health of the Czech Republic NT13492.

Disclosure of conflict of interest

None.

Address correspondence to: Petr Kuglik, Department of Experimental Biology, Faculty of Science, Masaryk University, Kotlarska 267/2, 611 37 Brno, Czech Republic. Tel: +420549495446; Fax: +420-549491070; E-mail: kugl@sci.muni.cz

References

[1] Telenius H, Carter NP, Bebb CE, Nordenskjold M, Ponder BA and Tunnacliffe A. Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. Genomics 1992; 13: 718-725.

[2] Cheung VG and Nelson SF. Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. Proc Natl Acad Sci U S A 1996; 93: 14676-14679.

[3] Zhang L, Cui X, Schmitt K, Hubert R, Navidi W and Arnheim N. Whole genome amplification from a single cell: implications for genetic analysis. Proc Natl Acad Sci U S A 1992; 89: 5847-5851.

[4] Sermon K, Lissens W, Joris H, Van Steirteghem A and Liebaers I. Adaptation of the primer extension preamplification (PEP) reaction for preimplantation diagnosis: single blastomere analysis using short PEP protocols. Mol Hum Reprod 1996; 2: 209-212.

[5] Langmore JP. Rubicon Genomics, Inc. Pharmacogenomics 2002; 3: 557-560.

[6] Fiegler H, Geigl JB, Langer S, Rigler D, Porter K, Unger K, Carter NP and Speicher MR. High resolution array-CGH analysis of single cells. Nucleic Acids Res 2007; 35: e15.

[7] Lizardi PM, Huang X, Zhu Z, Bray-Ward P, Thomas DC and Ward DC. Mutation detection and single-molecule counting using isothermal rolling-circle amplification. Nat Genet 1998; 19: 225-232.

[8] Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, Bray-Ward P, Sun Z, Zong Q, Du Y, Du J, Driscoll M, Song W, Kingsmore SF, Egholm M and Lasken RS. Comprehensive human genome amplification using multiple displacement amplification. Proc Natl Acad Sci U S A 2002; 99: 5261-5266.

[9] Zong C, Lu S, Chapman AR and Xie XS. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. Science 2012; 338: 1622-1626.

[10] Bergen AW, Haque KA, Qi Y, Beerman MB, Garcia-Closas M, Rothman N and Chanock SJ. Comparison of yield and genotyping performance of multiple displacement amplification and OmniPlex whole genome amplified DNA generated from multiple DNA sources. Hum Mutat 2005; 26: 262-270.

[11] Bergen AW, Qi Y, Haque KA, Welch RA and Chanock SJ. Effects of DNA mass on multiple displacement whole genome amplification and genotyping performance. BMC Biotechnol 2005; 5: 24.

[12] Berthier-Schaad Y, Kao WH, Coresh J, Zhang L, Ingersoll RG, Stephens R and Smith MW. Reliability of high-throughput genotyping of whole genome amplified DNA in SNP genotyping studies. Electrophoresis 2007; 28: 2812-2817.

[13] Kumar G, Garnova E, Reagin M and Vidali A. Improved multiple displacement amplification with phi29 DNA polymerase for genotyping of single human cells. Biotechniques 2008; 44: 879-890.

[14] Ling J, Zhuang G, Tazon-Vega B, Zhang C, Cao B, Rosenwaks Z and Xu K. Evaluation of genome coverage and fidelity of multiple displacement amplification from single cells by SNP array. Mol Hum Reprod 2009; 15: 739-747.

[15] Philips S, Rae JM, Oesterreich S, Hayes DF, Stearns V, Henry NL, Storniolo AM, Flockhart DA and Skaar TC. Whole genome amplification of DNA for genotyping pharmacogenetics candidate genes. Front Pharmacol 2012; 3: 54.

[16] Wang G, Brennan C, Rook M, Wolfe JL, Leo C, Chin L, Pan H, Liu WH, Price B and Makrigiorgos

GM. Balanced-PCR amplification allows unbiased identification of genomic copy changes in minute cell and tissue samples. Nucleic Acids Res 2004; 32: e76.
[17] Cho MH, Ciulla D, Klanderman BJ, Raby BA and Silverman EK. High-resolution melting curve analysis of genomic and whole-genome amplified DNA. Clin Chem 2008; 54: 2055-2058.
[18] van Eijk R, van Puijenbroek M, Chhatta AR, Gupta N, Vossen RH, Lips EH, Cleton-Jansen AM, Morreau H and van Wezel T. Sensitive and specific KRAS somatic mutation analysis on whole-genome amplified DNA from archival tissues. J Mol Diagn 2010; 12: 27-34.
[19] Winkel BG, Hollegaard MV, Olesen MS, Svendsen JH, Haunso S, Hougaard DM and Tfelt-Hansen J. Whole-genome amplified DNA from stored dried blood spots is reliable in high resolution melting curve and sequencing analysis. BMC Med Genet 2011; 12: 22.
[20] Harada T, Okita K, Shiraishi K, Kusano N, Furuya T, Oga A, Kawauchi S, Kondoh S and Sasaki K. Detection of genetic alterations in pancreatic cancers by comparative genomic hybridization coupled with tissue microdissection and degenerate oligonucleotide primed polymerase chain reaction. Oncology 2002; 62: 251-258.
[21] Lage JM, Leamon JH, Pejovic T, Hamann S, Lacey M, Dillon D, Segraves R, Vossbrinck B, Gonzalez A, Pinkel D, Albertson DG, Costa J and Lizardi PM. Whole genome analysis of genetic alterations in small DNA samples using hyperbranched strand displacement amplification and array-CGH. Genome Res 2003; 13: 294-307.
[22] Hughes S, Lim G, Beheshti B, Bayani J, Marrano P, Huang A and Squire JA. Use of whole genome amplification and comparative genomic hybridisation to detect chromosomal copy number alterations in cell line material and tumour tissue. Cytogenet Genome Res 2004; 105: 18-24.
[23] Iwamoto K, Bundo M, Ueda J, Nakano Y, Ukai W, Hashimoto E, Saito T and Kato T. Detection of chromosomal structural alterations in single cells by SNP arrays: a systematic survey of amplification bias and optimized workflow. PLoS One 2007; 2: e1306.
[24] Talseth-Palmer BA, Bowden NA, Hill A, Meldrum C and Scott RJ. Whole genome amplification and its impact on CGH array profiles. BMC Res Notes 2008; 1: 56.
[25] Pinard R, de Winter A, Sarkis GJ, Gerstein MB, Tartaro KR, Plant RN, Egholm M, Rothberg JM and Leamon JH. Assessment of whole genome amplification-induced bias through high-throughput, massively parallel whole genome sequencing. BMC Genomics 2006; 7: 216.
[26] Zhang C, Zhang C, Chen S, Yin X, Pan X, Lin G, Tan Y, Tan K, Xu Z, Hu P, Li X, Chen F, Xu X, Li Y, Zhang X, Jiang H and Wang W. A single cell level based method for copy number variation analysis by low coverage massively parallel sequencing. PLoS One 2013; 8: e54236.
[27] Rykalina VN, Shadrin AA, Amstislavskiy VS, Rogaev EI, Lehrach H and Borodina TA. Exome sequencing from nanogram amounts of starting DNA: comparing three approaches. PLoS One 2014; 9: e101154.
[28] Huang L, Ma F, Chapman A, Lu S and Xie XS. Single-Cell Whole-Genome Amplification and Sequencing: Methodology and Applications. Annu Rev Genomics Hum Genet 2015; 16: 79-102.
[29] Le Caignec C, Spits C, Sermon K, De Rycke M, Thienpont B, Debrock S, Staessen C, Moreau Y, Fryns JP, Van Steirteghem A, Liebaers I and Vermeesch JR. Single-cell chromosomal imbalances detection by array CGH. Nucleic Acids Res 2006; 34: e68.
[30] Hellani A, Coskun S, Benkhalifa M, Tbakhi A, Sakati N, Al-Odaib A and Ozand P. Multiple displacement amplification on single cell and possible PGD applications. Mol Hum Reprod 2004; 10: 847-852.
[31] Cardoso J, Molenaar L, de Menezes RX, Rosenberg C, Morreau H, Moslein G, Fodde R and Boer JM. Genomic profiling by DNA amplification of laser capture microdissected tissues and array CGH. Nucleic Acids Res 2004; 32: e146.
[32] Hirsch D, Camps J, Varma S, Kemmerling R, Stapleton M, Ried T and Gaiser T. A new whole genome amplification method for studying clonal evolution patterns in malignant colorectal polyps. Genes Chromosomes Cancer 2012; 51: 490-500.
[33] Guillaud-Bataille M, Valent A, Soularue P, Perot C, Inda MM, Receveur A, Smaili S, Roest Crollius H, Benard J, Bernheim A, Gidrol X and Danglot G. Detecting single DNA copy number variations in complex genomes using one nanogram of starting DNA and BAC-array CGH. Nucleic Acids Res 2004; 32: e112.
[34] Arriola E, Lambros MB, Jones C, Dexter T, Mackay A, Tan DS, Tamber N, Fenwick K, Ashworth A, Dowsett M and Reis-Filho JS. Evaluation of Phi29-based whole-genome amplification for microarray-based comparative genomic hybridisation. Lab Invest 2007; 87: 75-83.
[35] Pugh TJ, Delaney AD, Farnoud N, Flibotte S, Griffith M, Li HI, Qian H, Farinha P, Gascoyne RD and Marra MA. Impact of whole genome amplification on analysis of copy number variants. Nucleic Acids Res 2008; 36: e80.
[36] Corneveaux JJ, Kruer MC, Hu-Lince D, Ramsey KE, Zismann VL, Stephan DA, Craig DW and

Huentelman MJ. SNP-based chromosomal copy number ascertainment following multiple displacement whole-genome amplification. Biotechniques 2007; 42: 77-83.

[37] Paez JG, Lin M, Beroukhim R, Lee JC, Zhao X, Richter DJ, Gabriel S, Herman P, Sasaki H, Altshuler D, Li C, Meyerson M and Sellers WR. Genome coverage and sequence fidelity of phi29 polymerase-based multiple strand displacement whole genome amplification. Nucleic Acids Res 2004; 32: e71.

[38] Shojaei Saadi HA, Vigneault C, Sargolzaei M, Gagne D, Fournier E, de Montera B, Chesnais J, Blondin P and Robert C. Impact of whole-genome amplification on the reliability of pre-transfer cattle embryo breeding value estimates. BMC Genomics 2014; 15: 889.

[39] Barker DL, Hansen MS, Faruqi AF, Giannola D, Irsula OR, Lasken RS, Latterich M, Makarov V, Oliphant A, Pinter JH, Shen R, Sleptsova I, Ziehler W and Lai E. Two methods of whole-genome amplification enable accurate genotyping across a 2320-SNP linkage panel. Genome Res 2004; 14: 901-907.

[40] Tzvetkov MV, Becker C, Kulle B, Nurnberg P, Brockmoller J and Wojnowski L. Genome-wide single-nucleotide polymorphism arrays demonstrate high fidelity of multiple displacement-based whole-genome amplification. Electrophoresis 2005; 26: 710-715.

[41] Xing J, Watkins WS, Zhang Y, Witherspoon DJ and Jorde LB. High fidelity of whole-genome amplified DNA on high-density single nucleotide polymorphism arrays. Genomics 2008; 92: 452-456.

Supplemental methods

Human myeloma cell line U266 genome-wide copy-number changes in unamplified control

Control DNA sample of HMCL U266 without any WGA was tested in duplicates by three different aCGH platforms to assess the presence of CN changes. Segmentation by ADM-2 algorithm showed slightly differing genome-wide profiles as a result of three basic reasons: the number and the distribution of DNA probes specific for each platform, and the type of DNA reference which can differ in CN status in CNV loci. Results of duplicates in each platform and then results of three platforms were compared to analyze the number of CN segments, to distinguish CNAs from CNVs and to assess false positive (FP) and false negative (FN) segments used as specific array platform background for following WGA effect evaluation. FP and FN were calculated as follows: FP = [A/(N + A) + B/(N + B)]/2 and FN = (C/N + D/N)/2, where N is overall number of proper segments found by specific platform; A and B is number of FP segments found in first and second duplicate, respectively; C and D is number of FN segments found in first and second duplicate, respectively.

Overall, 39 CN changes (20 losses/19 gains; 38 CNAs/1 CNV) with the median size of 18.4 Mbp (range 0.1–48.2 Mbp) were found using Agilent Human Genome CGH Microarray, 4×44K platform. In comparison to higher-resolution platforms, overall 12 segments (10 losses and two gains) were not recognized due to low number of DNA probes in these loci. On the other hand, all 39 CN segments were recognized by higher-resolution platforms. In addition, one gain was reassessed as two gains separated by small loss using higher-resolution platforms. We also detected six FP losses (two in both duplicates, four in one of duplicates) with the median size of 0.3 Mbp (range 0.1–2.0 Mbp) which were not found by higher-resolution arrays despite the denser DNA probes coverage of these loci. Overall, two segments (one loss with the size of 0.1 Mbp and one gain with the size of 0.4 Mbp) were not found in one of the duplicates and marked as FN segments. Overall risk of FP and FN was 9.3% and 2.6%, respectively.

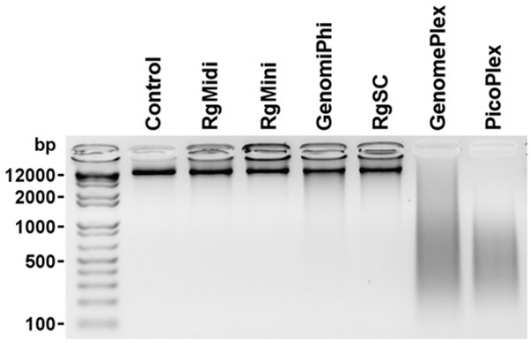
Overall, 52 CN changes (29 losses/23 gains; 45 CNAs/7 CNVs) with the median size of 9.9 Mbp (range 0.1–48.4 Mbp) were detected using Agilent SurePrint G3 CGH+SNP Microarray, 4×180K platform. In comparison to OGT 180K platform, overall five segments (median size of 0.1 Mbp, range 0.1–0.7 Mbp) were not found due to low number of DNA probes in these loci. We also found five segments (median size of 0.6 Mbp, range 0.1–2.1 Mbp), which were recognized only by this platform, as a result of different and specific Agilent DNA reference CN status in CNVs loci. No FP segment was found and only one loss with the size of 0.3 Mbp was not detected in one of the duplicates and marked as FN segment. Overall risk of FP and FN was 0.0% and 1.0%, respectively.

Overall, 52 CN changes (30 losses/22 gains; 48 CNAs/4 CNVs) with the median size of 9.9 Mbp (range 0.1–48.4 Mbp) were identified using OGT CytoSure ISCA UPD Array, 4×180K platform. In this platform, the highest variability was found in control sample duplicates analysis. In total, 15 segments (median size of 0.4 Mbp, range 0.1–6.9 Mbp) were found only in one of the duplicates and marked as FP segments as they were not recognized by other two platforms. On the other hand, all CN changes detected by Agilent platforms were also found using this OGT platform and no FN segment was present. Overall risk of FP and FN was 12.6% and 0.0%, respectively.

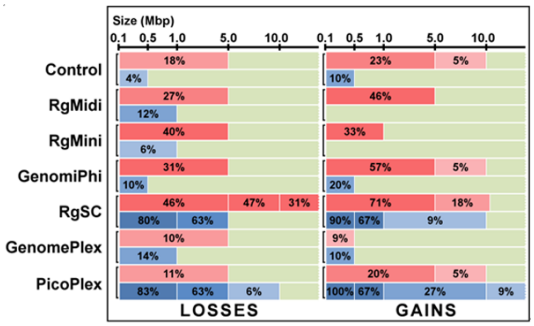
Human myeloma cell line U266 genome-wide copy-number neutral loss of heterozygosity areas in unamplified control

In our experiment, two SNP array platforms were used: 1) Agilent SurePrint G3 CGH+SNP Microarray, 4×180 K, which is based on 59 645 SNPs in restriction digestion sites and requires specific DNA digestion by *AluI* and *RsaI* restriction enzymes; and 2) OGT CytoSure ISCA UPD Array, 4×180K, which is based on a simple preferential hybridization to one of two DNA probes determined for 6 186 SNP positions across the genome. Although these SNP platforms differ in principle, both are used for the detection of genomic regions where only SNP homozygous constitutions are present. There are two cases of LOH: 1) presence of only one copy of this region due to loss of second allele; 2) cnnLOH originated likely as loss followed by replication of remaining allele. In this part of study, we focused on cnnLOH analysis. Although OGT SNP platform contains 9.6× smaller number of SNP probes in analysis than Agilent SNP platform,

the same 10 cnnLOH areas > 5.0 Mbp were recognized by both platforms in unamplified U266 DNA sample. The median size of detected cnnLOH areas was 23.6 Mbp (range 15.5–68.8 Mbp) and 23.0 Mbp (range 15.8–68.3 Mbp); and the homozygosity median was 99.6% (range 98.2–100.0%) and 95.5% (range 87.5–97.7%) in Agilent platform and OGT platform, respectively.



Supplementary Figure 1. Electrophoretic analysis of WGA products. 100 ng of each DNA WGA sample by RgMidi, RgMini, GenomiPhi, RgSC, GenomePlex and PicoPlex WGA kits were analyzed on a 2% agarose gel and compared to unamplified genomic DNA control.



Supplementary Figure 2. Detailed analysis of false positivity/false negativity copy-number segments risk in relation to copy-number segments size. The risk of false positivity (shades of red) and the risk of false negativity (shades of blue) is showed in specific copy-number segments size intervals.

Supplementary Table 1. Log2 ratio (L2R) values of real and false positive copy-number (CN) segments comparison

WGA	CN type	Tested size interval [Mbp]	L2R median (range)		*P*
			FP segments	Real segments	
Control	Losses	0.1–10.0	-0.39 (-0.27 – -1.11)	-0.85 (-0.44 – -3.74)	1.29×10^{-7}
	Gains		0.31 (0.26 – 0.36)	0.54 (0.38 – 1.85)	3.33×10^{-5}
RgMidi	Losses	0.1–10.0	-0.44 (-0.21 – -0.75)	-0.84 (-0.36 – -2.69)	1.06×10^{-11}
	Gains		0.35 (0.27 – 0.53)	0.60 (0.31 – 1.73)	8.16×10^{-8}
RgMini	Losses	0.1–10.0	-0.43 (-0.26 – -0.86)	-0.89 (-0.44 – -3.08)	2.27×10^{-18}
	Gains		0.30 (0.26 – 0.37)	0.59 (0.43 – 1.78)	1.13×10^{-5}
GenomiPhi	Losses	0.1–10.0	-0.51 (-0.25 – -0.94)	-0.91 (-0.36 – -3.24)	6.72×10^{-12}
	Gains		0.31 (0.26 – 0.50)	0.55 (0.34 – 1.86)	1.74×10^{-10}
RgSC	Losses	0.1–3.5	-0.77 (-0.26 – -2.77)	-1.06 (0.13 – -5.22)	0.09
	Gains	3.5–10.0	-0.41 (-0.28 – -0.69)	-0.91 (-0.61 – -1.51)	7.64×10^{-8}
		> 10.0	-0.37 (-0.26 – -0.59)	-0.99 (-0.82 – -1.51)	1.08×10^{-14}
		0.1–2.5	0.76 (0.25 – 1.92)	0.54 (-0.23 – 1.87)	0.09
		2.5–10.7	0.33 (0.26 – 0.66)	0.49 (0.08 – 0.74)	1.24×10^{-2}
GenomePlex	Losses	0.1–10.0	-0.45 (-0.25 – -0.90)	-0.83 (-0.34 – -3.14)	1.52×10^{-3}
	Gains		1× 0.34	0.48 (0.36 – 1.50)	---
PicoPlex	Losses	0.1–10.0	-0.44 (-0.25 – -1.26)	-0.52 (0.49 – -2.27)	0.65
	Gains		0.26 (0.24 – 0.34)	0.39 (0.17 – 1.06)	2.51×10^{-2}

Supplementary Table 2. Analysis of partial cnnLOH areas failures by Agilent and OGT SNP platforms

WGA	Platform	Size of cnnLOH failures Median (range) [Mbp]	CnnLOH failures > 5 Mbp [Mbp]	Overall % failures*
Control	Agilent	1.2 (0.04–4.1)	none	1.1%
	OGT	none	none	0.0%
RgMidi	Agilent	2.0 (0.2–5.5)	n = 3 (5.2, 2× 5.5)	9.4%
	OGT	1.4 (1.2–1.8)	none	1.5%
RgMini	Agilent	1.6 (0.1–7.8)	n = 2 (5.9, 7.8)	9.4%
	OGT	0.8 (0.5–1.2)	none	0.6%
GenomiPhi	Agilent	2.4 (0.6–4.5)	none	6.0%
	OGT	1.4 (1.2–1.7)	none	0.5%
GenomePlex	Agilent	---	---	---
	OGT	3.8 (0.1–6.9)	n = 1 (6.9)	5.3%
PicoPlex	Agilent	---	---	---
	OGT	7.1 (0.2–68.3)	n = 10 (median 19.2 Mbp, range 6.1–68.3 Mbp)	47.6%

*% of cnnLOH failures from overall 303.8 Mbp (Agilent SNP platform) or 301.9 Mbp (OGT SNP platform) of cnnLOH across the genome.

Supplementary Table 3. Detailed analysis of WGAs effect on 10 cnnLOH areas by Agilent SNP platform

WGA	% of failed SNP probes Median (range)	Homozygosity deviation from control Median (range)	Change of homozygosity ratio*	% of cnnLOH cases with a change lower than:		
				+/-1%	+/-2%	+/-3%
Control	18.5% (13.5–24.7%)	+/-0.1% (-0.7–0.7%)	1.0:1.0:0.0	100.0%	100.0%	100.0%
RgMidi	36.3% (30.5–42.2%)	-0.6% (-2.1–0.6%)	4.0:1.0:0.0	82.5%	97.5%	100.0%
RgMini	30.1% (26.2–34.1%)	-0.6% (-3.5–0.6%)	9.0:1.0:0.0	70.0%	92.5%	97.5%
GenomiPhi	30.6% (27.1–33.4%)	-0.8% (-3.0–0.6%)	7.0:1.0:0.0	67.5%	92.5%	97.5%

*Change of homozygosity ratio of cases with homozygosity decreasing : increasing : no change.

Supplementary Table 4. Detailed analysis of WGAs effect on 10 cnnLOH areas by OGT SNP platform

WGA	Homozygosity deviation from control Median (range)	Change of homozygosity ratio*	% of cnnLOH cases with a change lower than:			
			+/-4%	+/-5%	+/-10%	+/-20%
Control	+/-0.6% (-3.6–3.6%)	1.0:5.0:4.0 5.0:1.0:4.0	100.0%	100.0%	100.0%	100.0%
RgMidi	1.3% (-6.3–7.1%)	1.6:3.1:1.0	87.5%	92.5%	100.0%	100.0%
RgMini	-2.0% (-6.3–4.8%)	5.6:1.4:1.0	90.0%	95.0%	100.0%	100.0%
GenomiPhi	0.0% (-4.5–7.1%)	1.9:2.1:1.0	85.0%	95.0%	100.0%	100.0%
GenomePlex	-10.3% (-19.8–0.0%)	12.3:0.0:1.0	20.0%	25.0%	47.5%	100.0%
PicoPlex	-21.2% (-36.4–-1.8%)	1.0:0.0:0.0	2.5%	2.5%	7.5%	42.5%

*Change of homozygosity ratio of cases with homozygosity decreasing : increasing : no change.